

Chimeric Structure of the NAD(P)⁺- and NADP⁺-dependent Malic Enzymes of *Rhizobium (Sinorhizobium) meliloti**

(Received for publication, September 2, 1997, and in revised form, February 3, 1998)

Michael J. Mitsch, Ralf T. Voegelé, Alison Cowie, Magne Osteras, and Turlough M. Finan‡

From the Department of Biology, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1, Canada

Malic enzymes catalyze the oxidative decarboxylation of malate to pyruvate in conjunction with the reduction of a nicotinamide cofactor. We determined the DNA sequence and transcriptional start sites of the genes encoding the diphosphopyridine nucleotide-dependent malic enzyme (DME, EC 1.1.1.39) and the triphosphopyridine nucleotide-dependent malic enzyme (TME, EC 1.1.1.40) of *Rhizobium (Sinorhizobium) meliloti*. The predicted DME and TME proteins contain 770 and 764 amino acids, respectively, and are approximately 320 amino acids larger than previously characterized prokaryotic malic enzymes. The increased size of DME and TME resides in the C-terminal extensions which are similar in sequence to phosphotransacetylase enzymes (EC 2.3.1.8). Modified DME and TME proteins which lack this C-terminal region retain malic enzyme activity but are unable to oligomerize into the native state. Data base searches have revealed that similar chimeric malic enzymes were uniquely present in Gram-negative bacteria. Thus DME and TME appear to be members of a new class of malic enzyme characterized by the presence of a phosphotransacetylase-like domain at the C terminus of the protein.

The reversible oxidative decarboxylation of malate to pyruvate and CO₂ is catalyzed by malic enzymes with the associated reduction of either NAD⁺ (EC 1.1.1.38 and EC 1.1.1.39) or NADP⁺ (EC 1.1.1.39 and EC 1.1.1.40). Eukaryotic malic enzymes have been characterized from a wide variety of sources and most are homotetramers with subunits varying in length from 565 to 584 amino acid residues (65.4–71 kDa) (e.g. Refs. 1–3). The prokaryotic malic enzymes described so far are fewer in number and more varied in structure than their eukaryotic counterparts. In two separate *Pseudomonas* species, a 208-kDa NADP⁺-dependent tetramer (4) and a 680 kDa-decamer (5) have been found. Malic enzymes have also been purified from three thermophilic organisms; a NADP⁺-dependent tetramer from *Clostridium thermocellum* (6), a NADP⁺-dependent dimer from *Sulfolobus solfataricus* (7), and a tetramer utilizing both cofactors from *Bacillus stearothermophilus* (8). Most recently a malic enzyme has been cloned from *Streptococcus bovis* and was found to be a dimer of 40-kDa subunits (9). Early studies with *Escherichia coli* identified two malic enzymes, the NADP⁺-dependent isozyme appears to be an octamer made up of 67-kDa subunits (10), while the NAD⁺-dependent enzyme is made up of four 55-kDa subunits (11).

* This work was supported by grants from the Natural Sciences and Engineering Council of Canada (to T. M. F.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 905-525-9140 (ext. 22932); Fax: 905-522-6066; E-mail: FINAN@MCMASTER.CA.

Our interest in malic enzymes arose through studies directed toward elucidating the pathways of succinate metabolism in *Rhizobium meliloti*. This bacterium forms root nodules on alfalfa, and C₄ dicarboxylic acids, such as malate and succinate, appear to be the major energy source donated by the plant host to the N₂-fixing bacterium (12, 13). *R. meliloti* contains both a diphosphopyridine nucleotide (NAD⁺)-dependent malic enzyme (DME, EC 1.1.1.39)¹ and a triphosphopyridine nucleotide (NADP⁺)-dependent malic enzyme (TME, EC 1.1.1.40) and studies with gene fusions have shown that N₂-fixing bacteroids express *dme* at levels similar to those observed for free living cells, whereas *tme* is poorly expressed in bacteroids relative to free living cells. Mutant analysis has shown that DME, but not TME, is required for symbiotic N₂-fixation (14, 15). These results suggest that DME, together with pyruvate dehydrogenase plays a key role in the conversion of malate to acetyl-CoA for efficient tricarboxylic acid cycle function in N₂-fixing bacteroids. Both NAD⁺- and NADP⁺-dependent malic enzymes have also been identified in the soybean symbiont *Bradyrhizobium japonicum*, where the NADP⁺-dependent enzyme with a subunit molecular mass of 78 kDa, is reported to exist as a dimer or tetramer depending on pH conditions (16–18).

Here we report the nucleotide sequence and determination of the transcriptional start sites for the two malic enzyme genes from *R. meliloti*. The amino acid sequences reveal that the proteins have an unusual chimeric structure consisting of a 400-amino acid long N-terminal region similar to previously characterized malic enzymes, plus an additional C-terminal 340-amino acid extension similar in sequence to previously characterized phosphotransacetylase (PTA, EC 2.3.1.8) proteins. We show that deletion of the C-terminal 300 amino acids results in the loss of oligomerization ability. Comparative sequence analysis revealed that this chimeric protein structure is present in several other putative malic enzymes identified from genome sequencing projects. The *R. meliloti* malic enzymes are therefore the first representatives of a novel group of malic enzymes to be cloned, sequenced, and characterized.

EXPERIMENTAL PROCEDURES

Bacterial Strains, Plasmids, and Media—Media, growth conditions, and antibiotic concentrations have been described previously (15, 19). M9 minimal media plates (5.8 g of Na₂HPO₄, 3 g of KH₂PO₄, 0.5 g of NaCl, and 1 g of NH₄Cl per liter) containing 15 mM succinate or glucose, 1 mM MgSO₄, 0.25 mM CaCl₂, 1 μg/ml D-biotin, and 0.1 mg/ml ampicillin for plasmid selection were used for screening.

Genetic Techniques—The localization of the *R. meliloti dme* gene to the 3.2-kb *Hind*III fragment in plasmid pTH139 has been previously described (20). A 7.1-kb *Bam*HI fragment carrying the *R. meliloti tme* gene was subcloned from plasmid pRmT100 (15) into pUC118 to yield pTH251. A 2.7-kb *Bsp*HI fragment carrying the complete *tme* gene was subcloned into the compatible *Nco*I site of pAB2001 (21) and this 2.7-kb region was then re-isolated as a *Hind*III fragment and ligated into

¹ The abbreviations used are: DME, diphosphopyridine-dependent malic enzyme; TME, triphosphopyridine-dependent malic enzyme; PTA, phosphotransacetylase; kb, kilobase(s).

pUC119 to produce pTH392. The *PstI* deletions of both the *dme* and *tme* genes were produced by digesting pTH139 and pTH392, respectively, with *PstI* followed by religation. The resulting deletions extended from the *PstI* site within the malic enzyme gene to the *PstI* site of the pUC vector polylinker. Analysis of the resulting fusion between the gene and vector sequence showed that the truncated *dme* construct would produce a 452-amino acid polypeptide (3 amino acids derived from pUC119 vector) with a predicted molecular mass of 48.4 kDa while the truncated *tme* construct would yield a 462-amino acid polypeptide (21 amino acids derived from the pUC119 vector) with a predicted molecular mass of 49.4 kDa.

Both DNA strands of the *dme* and *tme* genes were sequenced from a combination of subcloned fragments and deletion derivatives employing the M13 universal primer. Remaining gaps were filled in using sequence obtained from synthetic primers. Single stranded DNA was obtained from transformed XL-1 Blue cells (Stratagene) following infection with M13K07 phage by the method of Vieira and Messing (22). Double stranded DNA was isolated by alkaline lysis according to Sambrook *et al.* (23). The DNA sequence of the region flanking Tn5 insertions (15, 20) was determined using an IS50 primer (5'-TCACAT-GGAAGTCAGATCCT-3') kindly provided by Dr. R. J. Watson. DNA sequencing was performed with an ABI 373 Stretch DNA automated sequencer using dye terminator chemistry and cycle sequencing (MOBIX facility, McMaster University). The completed sequences were deposited in GenBank under accession numbers AF017443 and AF017444 for *dme* and *tme*, respectively.

Determination of Transcriptional Start Sites—Total RNA was isolated from Lbmc grown wild type *R. meliloti* Rm1021 cells using the procedure outlined in the RNeasy Midiprep kit (Qiagen). Total nodule RNA was isolated from 20 alfalfa plants *Medicago sativa* var. *Iroquois* inoculated 28 days previously with Rm1021 (15), using the method of Corbin *et al.* (24). The percentage of bacterial RNA was estimated to be approximately 50% via rRNA comparison by agarose gel electrophoresis. Primers employed to determine the *dme* and *tme* transcriptional start sites were 5'-ATCGATGTCTCCGCTGGCGG-3' (nucleotides 38–57 relative to the translational start site) and 5'-GACCTT-GAGAATGAAAATCG-3' (nucleotides 60–79 relative to the translational start site), respectively. Primers were labeled with [γ -³³P]ATP (NEN Life Science Products Inc.) and polynucleotide kinase (New England Biolabs). 30 μ g of bacterial RNA and 60 μ g of nodule RNA were annealed to 10⁶ cpm of labeled primer and extended according to the protocol of Ausubel *et al.* (25). To align the transcriptional start site, the same primers were used in conjunction with template DNA (pTH139 and pTH251) in a standard sequencing reaction using the Sequenase Kit (Amersham).

Sequence Analysis and Alignments—Data base searches were conducted using the BLAST search engine (26) and alignments were carried out with the ClustalW multiple alignment program (27). Generation of the phylogenetic data was done using the PHYLIP package and the protein sequence parsimony (protpars) and distance (protdist) programs (28).

Protein Isolation and Quantitative Malic Enzyme Detection—Overproduced *R. meliloti* malic enzymes were purified to homogeneity from *E. coli* strains carrying either pTH139 (DME) or pTH392 (TME).² Malic enzyme activities were quantitated using the pyruvate formation assay or following the formation of NAD(P)H at 340 nm (15).² Partial purification of the truncated proteins and SDS-polyacrylamide gel electrophoresis analysis verified the production of a 48–50-kDa polypeptide corresponding to the truncated malic enzymes in accordance with our predicted molecular weights. Protein concentrations were determined by the Bradford method (29) using the Bio-Rad protein assay dye.

Determination of the Native Molecular Weights of Intact and Truncated Malic Enzymes—Standards and crude cell extract samples were electrophoresed using nondenaturing polyacrylamide gels (Sigma Technical Bulletin MKR-137; gel concentrations used: 4.5, 5, 5.5, 6, 7, 8, 9, and 10%) at 4 °C with a constant current of 4 mA for 1 h and 7 mA thereafter until the dye front was 1 cm from the end of the gel. Gels containing protein standards α -lactalbumin (14.2 kDa), carbonic anhydrase (29 kDa), egg albumin (45 kDa), bovine serum albumin (monomer 66 kDa, dimer 132 kDa), and jack bean urease (trimer 272 kDa, hexamer 545 kDa) (Sigma protein molecular weight kit MW-ND-500) were stained with Coomassie Brilliant Blue R-250. To stain for malic enzyme activity, gels were washed once with 100 mM Tris/HCl, pH 7.8, 30 mM K-L-malate, pH 7.8, 3 mM MnCl₂, 50 mM KCl and then immersed in the

same buffer containing 0.3 mg/ml 3-(4,5-dimethyl thiazol-2-yl)-2,5-diphenyl tetrazolium bromide and 0.04 mg/ml phenazine methosulfate. NAD⁺ was added for DME and Δ DME, and NADP⁺ was added for TME and Δ TME to a final concentration of 0.3 mg/ml. Gels were then incubated for approximately 30 min at 37 °C in the dark without shaking. R_F values of the protein standards and the malic enzymes (using all available data points) were plotted as a function of gel concentration in a Ferguson plot ($100 \times \log(100 \times R_F)$ versus % gel concentration). The retardation coefficients (KR) of the protein standards obtained from the slopes of the Ferguson plot graphs were then plotted as a function of the molecular weights of the standards ($\log KR$ versus $\log MW$) (30). The molecular weights for the malic enzymes were determined from their KR values by nonlinear regression. Oligomerization states were estimated by comparison to the predicted subunit molecular weights deduced from the sequences.

RESULTS

Nucleotide Sequence of the *R. meliloti* *tme* and *dme* Genes—The *R. meliloti* *dme* gene was previously delineated to a 3.2-kb *HindIII* fragment (20). We determined the nucleotide sequence of this fragment and analysis of the sequence revealed the presence of a single open reading frame (*dme*) of 2313 nucleotides which encoded a 770-amino acid protein and contained a potential ribosome-binding site (5'-GGAA-3') 10 base pairs upstream of the ATG start codon. The *R. meliloti* *tme* gene was subcloned as a 2.7-kb *BspHI* fragment in plasmid pTH392 which was sufficient to allow *E. coli* EJ1321 (*dme tme pck*) (31) to grow on minimal media with succinate as sole carbon source. The nucleotide sequence of this fragment revealed an open reading frame (*tme*) of 2295 nucleotides encoding a protein of 764 amino acids. A potential ribosome-binding site (5'-GGAA-3') was located 2 base pairs upstream of the ATG start codon. In addition, the deduced molecular masses of the proteins were 83 kDa for DME and 82 kDa for TME and these values are in agreement with the molecular mass estimates of the *R. meliloti* DME and TME proteins which were recently purified from *E. coli* strains carrying plasmids pTH139 and pTH392, respectively.² Furthermore, both the *tme* and *dme* genes showed a clear G + C bias of 83% at the third nucleotide position of the codons compared with the total G + C content of 65%.

Identification of the *dme* and *tme* Promoter Regions—The deduced directions of transcription of the malic enzyme genes were consistent with the directions predicted from previous analyses of *dme* and *tme lacZ* gene fusions (15, 20). In those experiments, *dme* but not *tme* was shown to be expressed in bacteroids from alfalfa root nodules, whereas both genes appeared to be constitutively expressed in free living cells. For comparative purposes it was therefore of interest to identify the promoter regions of these two genes and in addition we wished to establish whether *dme* was expressed from the same promoter in both free living cells and bacteroids. To identify the transcriptional start sites of these genes we employed primer extension experiments using primers from within the *dme* and *tme* structural genes. In the case of *dme*, RNA from free living wild type cells and wild type cells carrying additional copies of *dme* on the cosmid pTH69 were examined. In both cases, the start site of transcription for *dme* mapped to two locations 78 and 80 base pairs upstream from the ATG translational start codon (Fig. 1A, lanes 1 and 2). To map the *dme* transcription start site employed in *N*₂-fixing bacteroids, total RNA from alfalfa nodules was used in extension experiments. The result (Fig. 1A, lane 3) clearly showed that transcription of *dme* occurs from the same start site in both free living cells and nitrogen fixing bacteroids. Employing RNA from wild type cells, the major transcriptional start site for *tme* was located 47 base pairs upstream from the ATG start codon (Fig. 1B, lane 1). A second minor extension product was located 65 base pairs upstream from the ATG, the significance of this product is currently unclear.

² R. T. Voegelé, M. J. Mitsch, and T. M. Finan, manuscript in preparation.

Comparative Analysis of the *R. meliloti* Malic Enzyme Gene Sequences—Comparison of the *dme* and *tme* structural genes revealed they were 58% identical whereas the predicted proteins were 42% identical and 62% similar in amino acid sequence (Table I). Dot-matrix plots comparing the nucleotide and amino acid sequences of the two malic enzymes showed that while the first two-thirds of the genes shared 58% identity and 77% similarity, this degree of homology was reduced in the region constituting the last third of the genes to 25% identity and 47% similarity (data not shown). Searches of the DNA and protein data bases revealed considerable variation between the *R. meliloti* malic enzymes and other prokaryotic malic enzymes (Table I). Among the latter enzymes, only those from the Gram-positive bacteria *S. bovis* (9) and *B. stearothermophilus* (8) have been cloned and biochemically characterized. These two enzymes together with the putative malic enzyme from the blue-green alga *Synechocystis* sp. PCC6803 (32) are considerably shorter than the *R. meliloti* malic enzymes. A multiple alignment of these shorter enzymes with DME and TME showed that the similarity between all the enzymes extended across the entire 415-amino acid N-terminal portion of the DME and

TME proteins (Fig. 2A, and data not shown). On the other hand, the predicted malic enzymes of *Haemophilus influenzae* (33), *E. coli* (accession number AE000333), and *Rickettsia prowazekii*³ (all of which were sequenced as part of genome sequencing projects) were similar in size to the *R. meliloti* malic enzymes, and all these proteins showed homology over their entire length (alignments not shown).

PTA-like Domain at the C Terminus of the *R. meliloti* Malic Enzymes—BLAST searches employing the C-terminal 340 amino acids from the DME and TME proteins revealed that these regions resembled PTA enzymes (EC 2.3.1.8) from several bacterial species. Alignment of the C-terminal 340 amino acids from both DME and TME with the 333-amino acid phosphotransacetylase enzyme from *Methanosarcina thermophila*, which has been sequenced and characterized biochemically (34), revealed 26% identity (46% similarity) between DME and the PTA enzyme while TME shared 23% amino acid identity and 49% similarity with the *M. thermophila* PTA (Fig. 2A). Also of interest was the fact that the similarity between the DME, TME, and PTA sequences was scattered along the entire length of the PTA enzyme sequence (Fig. 2A) with no large clusters of identical or conserved residues.

Conserved Regions within the Malic Enzyme Proteins—Eukaryotic malic enzymes unlike their prokaryotic counterparts have been reasonably well characterized with respect to conserved motifs and amino acids essential for activity (2, 3, 35). Motifs containing the putative cofactor and Mn²⁺-binding sites have been identified within the malic enzymes from *B. stearothermophilus* and *S. bovis* (8, 9). An alignment of eight prokaryotic malic enzymes allowed us to identify eight regions which are highly conserved within these proteins (Fig. 2B, regions A-H). Some of these regions correspond to those implicated in the functioning of the eukaryotic malic enzymes.

Region A contains a tyrosine residue conserved in all of the prokaryotic malic enzymes examined. Chemical modification experiments with eukaryotic malic enzymes have implicated a corresponding tyrosine residue in L-malate binding (35, 36). Regions B and E have the required glycine, alanine, and lysine residues indicative of cofactor binding sites (37–39). The two lysine residues in region B are believed to be required for NAD(P)⁺ binding (36). Region C contains a conserved cysteine residue thought to be located at or near the active site of the eukaryotic enzymes (40, 41) and may be responsible for coordinating the binding of the substrate L-malate or the divalent metal ion (2). Region D represents a large block of 13 amino acid residues that are absolutely conserved in the prokaryotic malic enzymes. This domain contains two histidine residues which, through chemical disruption studies of eukaryotic malic

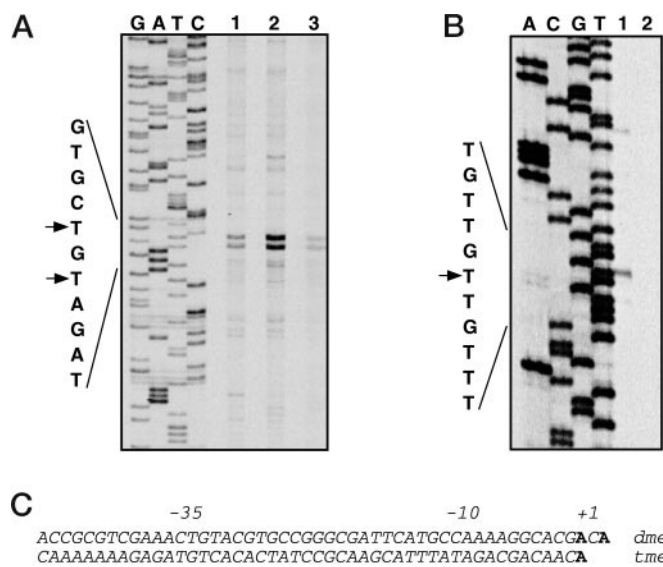


FIG. 1. Primer extensions of the two malic enzymes of *R. meliloti*. A, primer extension of *dme* with associated nucleotide sequence. Arrows denote transcriptional start sites. Lane 1, mRNA isolated from free living Rm1021; lane 2, mRNA isolated from H240 a *Rhizobium* strain carrying pTH69 overexpressing *dme*; lane 3, mRNA isolated from Rm1021 induced nodules. B, primer extension of *tme* with associated nucleotide sequence. Arrow indicates transcriptional start site. Lane 1, mRNA isolated from free living Rm1021; lane 2, tRNA control. C, comparison of the promoter regions of *dme* and *tme*. The -35, -10, regions and transcriptional start sites (in bold type) are indicated.

³ S. Andersson, unpublished data.

TABLE I
Comparison of different prokaryotic malic enzymes to *Rhizobium meliloti* DME and TME

Source	EC designation ^a	% Identity ^b		Length ^c	Accession No. ^d	Ref.
		DME	TME			
<i>H. influenzae</i>	ND ^e	57	43	756	P43837	36
<i>E. coli</i>	ND	58	43	759	AE000333	51
<i>R. prowazekii</i>	ND	42	52	766		Andersson (Footnote 3)
<i>B. stearothermophilus</i>	1.1.1.38	40	40	478	M19485	8
<i>S. bovis</i>	1.1.1.39	45	47	389	U35659	9
<i>S. sp.</i> PCC6803	ND	40	38	463	D90899	35
<i>R. meliloti</i> DME	1.1.1.39		42	770	AF017443	This study
<i>R. meliloti</i> TME	1.1.1.40	42		764	AF017444	This study

^a Classified according to cofactor specificity and oxaloacetate decarboxylation activity.

^b Based on comparison of overlapping amino acid residues of alignment.

^c Number of amino acid residues of malic enzyme protein.

^d GenBank accession code.

^e ND, not determined, information based on sequence data alone.

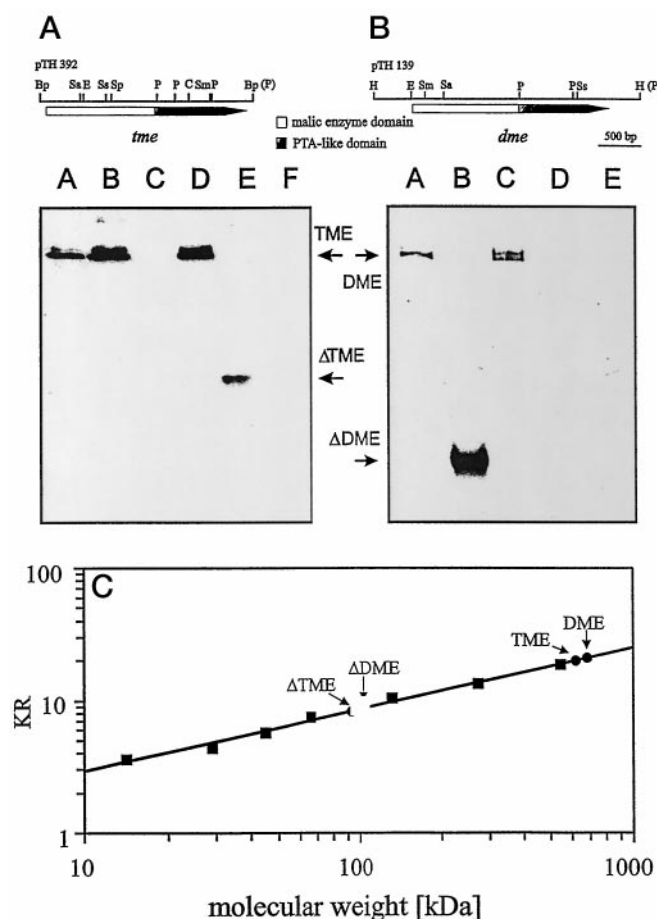


FIG. 3. Restriction maps of the malic enzyme gene regions with direction of transcription indicated and nondenaturing polyacrylamide gels of crude cell extracts stained for malic enzyme activity in several *R. meliloti* and *E. coli* strains. Locations of DME, TME, Δ DME, and Δ TME are shown. **A**, restriction map of *tme* and NAD⁺-dependent malic enzyme-stained gel: lane A, Rm1021; lane B, RmG456 (*dme*⁻); lane C, RmH215 (*tme*⁻); lane D, malic enzyme-deficient *E. coli* strain EJ1321 carrying pTH392; lane E, malic enzyme-deficient *E. coli* strain EJ1321 carrying pTH392 Δ PstI; lane F, malic enzyme-deficient *E. coli* strain EJ1321. **B**, restriction map of *dme* and NAD⁺-dependent malic enzyme stained gel: lane A, Rm1021; lane B, PTA enzyme-deficient *E. coli* strain BW16462 carrying pTH139 Δ PstI; lane C, PTA enzyme-deficient *E. coli* strain BW16462 carrying pTH139; lane D, PTA enzyme-deficient *E. coli* strain BW16462; lane E, malic enzyme-deficient *E. coli* strain EJ1321. **C**, plot of retardation coefficients derived from Ferguson plots versus molecular weights of standard proteins. Position of the *R. meliloti* malic enzymes (DME and TME) and truncated forms (Δ DME and Δ TME) lacking the C-terminal region are indicated based upon migration in nondenaturing polyacrylamide gels. Restriction enzyme sites shown in A and B are: *Bsp*HI, *Bp*; *Cla*I, *C*; *Eco*RI, *E*; *Hind*III, *H*; *Pst*I, *P*; *Sall*, *Sa*; *Sma*I, *Sm*; *Sph*I, *Sp*; *Sst*I, *S*. (*P*) denotes *Pst*I of multiple cloning site used for gene truncation (see "Experimental Procedures").

3-(4,5-dimethyl thiazol-2-yl)-2,5-diphenyl tetrazolium bromide (see "Experimental Procedures"). No activity band was detected in extracts of plasmid-free *E. coli* strains, whereas as expected, *E. coli* strains carrying the complete malic enzyme genes on plasmids pTH139 and pTH392 showed activity bands which co-migrated with the native DME and TME proteins of wild type *R. meliloti* (Fig. 3A, lanes A, B, and D; and B, lanes A and C). The native sizes of the malic enzymes were determined by comparing the migration of these proteins to known marker proteins in different percentage nondenaturing polyacrylamide gels according to the procedure of Hedrick and Smith (30) (Fig. 3C). DME and TME were found to have molecular masses of approximately 680 and 620 kDa, respectively. Based upon the subunit molecular masses of DME and TME (83 and 82 kDa,

respectively) in SDS-polyacrylamide electrophoresis gels,² DME and TME both appear to form homo-octamers.

The truncated *dme* and *tme* genes produced malic enzyme activity bands which migrated much faster than the native enzymes indicating a smaller protein size (Fig. 3, A, lane E, and B, lane B). The native molecular masses of the truncated Δ DME and Δ TME proteins were estimated to be 104 and 93 kDa, respectively (Fig. 3C), which is very close to the size of a homodimer for both truncated forms. This indicates that removal of the C terminus of the proteins impedes oligomerization but that the enzymes are still capable of functioning at a reduced level. We note that in preliminary purification experiments involving gel filtration of truncated DME and TME proteins, Δ TME eluted as a 121-kDa protein and Δ DME as a 120-kDa protein, also suggesting a dimeric state (data not shown).

DISCUSSION

We have sequenced and analyzed the two malic enzyme genes of *R. meliloti* and characterized the enzymes with respect to native size and domain structure. We suggest that the two malic enzymes from *R. meliloti*, DME and TME, are representatives of a new class of large subunit malic enzymes characterized by the presence of a C-terminal region of approximately 320 amino acids which shows homology to PTA enzymes (EC 2.3.1.8). All attempts to detect PTA activity using purified DME and TME proteins have been unsuccessful and thus the function of the C-terminal PTA-like domain is currently unclear.

The presence of the PTA-like domain in both DME and TME, which have different cofactor specificities and effectors,² suggests that it is related to neither cofactor specificity nor effector binding. This view is further strengthened by the fact that many "shorter" eukaryotic malic enzymes show similar properties to DME or TME (46, 47). The observation that both DME and TME deletion derivatives lacking the C-terminal region retain (albeit reduced) malic enzyme activity clearly suggests that these two proteins have at least two major domains. We found that while both DME and TME deletion derivatives retained malic enzyme activity, these proteins now appeared to form dimers in contrast to the octameric structure found for the native enzymes. While this result suggests that the PTA-like domain may be involved in octamerization, it is also possible that the C-terminal deletion indirectly interferes with formation of the native oligomeric state. On the other hand, the formation of dimers by these deletion derivatives suggests that this may be the first step in the formation of the octamer.

In view of the unusual structure of the DME and TME proteins as discussed above, we wish to summarize the evidence that DME and TME are in fact malic enzymes. In previous studies, *dme* and *tme* mutants were shown to lack the respective NAD⁺- and NADP⁺-dependent malic enzyme activities. In appropriate genetic backgrounds, these mutations resulted in growth phenotypes which were entirely consistent with an *in vivo* role for these enzymes in the conversion of malate to pyruvate (14, 20). Furthermore, expression of either the *dme* or *tme* gene in an *E. coli* malic enzyme double mutant restored its ability to grow on succinate as carbon source (Ref. 20, this work). We have also determined the kinetic characteristics of the two purified proteins obtained from overexpression of the *tme* and *dme* genes and these proteins catalyze the oxidative decarboxylation of malate and the simultaneous reduction of NAD⁺ or NADP⁺ in the case of DME, and NADP⁺ in the case of TME.² Moreover, the N-terminal sequence of the purified proteins is the same as that predicted by the nucleic acid sequence of the *dme* and *tme* genes. We therefore conclude that DME and TME are indeed malic enzymes.

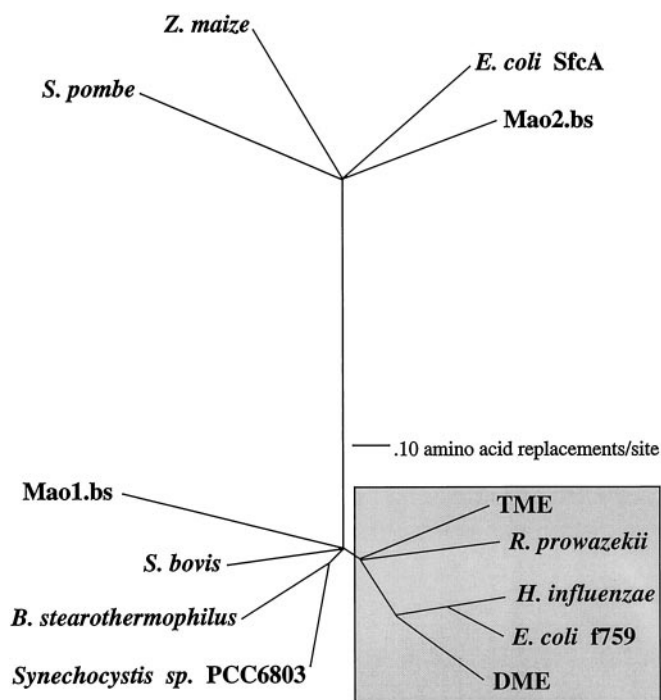


FIG. 4. Unrooted tree of the various prokaryotic malic enzymes available from the genomic data base using *fitch* analysis with the PHYLIP software package. Mao1.bs and Mao2.bs are putative malic enzymes from *B. subtilis* while SfcA is the NAD⁺-dependent malic enzyme of *E. coli*. The lengths of the branches reflect the distances of divergence and the highlighted box represents the elongated malic enzyme subdivision with a confidence level of 100% generated from 100 bootstrap replicates.

Recent genome sequencing projects have revealed putative malic enzyme genes in *E. coli*, *H. influenzae*, and *R. prowazekii* which are predicted to encode proteins similar in length and sequence to the *R. meliloti* malic enzymes. The conservation of the PTA-like region in all these proteins represents strong evidence that this region plays some functional role, be it structural or catalytic. These results prompted us to further examine the phylogenetic relationships of all available prokaryotic malic enzymes. A maximum parsimony analysis and distance measurements using the eukaryotic malic enzymes (*Zea maize*, accession number J05130, and *Schizosaccharomyces pombe*, accession number U00621) as the outgroups was carried out. The resulting tree (Fig. 4) showed that the long form enzymes cluster as a distinct group, and two other groups, one containing malic enzymes from *Synechocystis* sp. PCC6803 (32) and *B. stearothermophilus*, and the other made up of SfcA (accession number D90788) from *E. coli* and Mao2.bs (gene designation YWKA, accession number Z49782) from *Bacillus subtilis* can be distinguished. Interestingly, the other malic enzyme identified in *B. subtilis*, Mao1.bs (gene designation YQKJ, accession number D84432), appears to be more closely related to the other prokaryotic malic enzymes than it is to Mao2.bs. We also note that the same tree was obtained when the PTA-like C-terminal domain was removed from the "long" malic enzyme sequences, which shows that the apparent phylogenetic relationship between the large malic enzymes is not solely a result of the PTA-like domain. The different placements of the two malic enzymes from *E. coli*, SfcA and the chimeric malic enzyme f759, clearly shows that these are paralogous proteins (proteins arising via gene duplication), and the linkage of SfcA with the Mao2.bs from the Gram-positive organism *B. subtilis* suggests that an ancestral malic enzyme gene duplication occurred prior to the divergence generating the Gram-positive and Gram-negative bacteria.

Two malic enzymes have previously been purified from *E. coli*, however, neither of the reported molecular weights of these proteins corresponds to that of the larger enzyme f759 (accession number AE000333) as determined from the sequence of the *E. coli* genome (48). The amino acid composition of the *E. coli* NAD⁺-dependent malic enzyme (EC 1.1.1.38) does match that predicted for the *sfcA* gene product (11, 48, 49). The *sfcA* encoded protein was recently overexpressed and was shown to exhibit NAD⁺-dependent malic enzyme activity (50). We thus assume that the other malic enzyme gene (f759) encodes the previously characterized NADP⁺-dependent malic enzyme (EC 1.1.1.40) (51, 52). However, this inference clearly needs to be verified as the subunit molecular weight and the amino acid composition of the NADP⁺-malic enzyme as reported by Spina *et al.* (10), fails to match the sizes and amino acid composition of the two malic enzymes deduced from the *E. coli* genome sequence.

The mapping of the transcription start sites of the *dme* and *tme* genes allowed us to identify and analyze their promoters (Fig. 1C). No sequences resembling the -24/-12 motif of σ^{54} -dependent promoters employed by many *Rhizobium* genes involved in N₂-fixation could be found in either promoter. Although the constitutive *dme* and *tme* expression in free living cells is likely dependent on the housekeeping σ -factor, no clear homology to the -35/-10 motifs of σ^{70} -dependent promoters could be observed. The *tme* promoter, but not the *dme*, has an AT-rich -10 sequence (5'-TATAGA-3') resembling the *E. coli* consensus (5'-TATAAT-3'). The putative -35 sequences of *dme* (5'-CTGTAC-3') and *tme* (5'-ATGTCA-3') have little resemblance to the putative *R. meliloti* consensus (5'-TTGACC-3') (Ref. 53 and references therein). Divergence from the consensus promoter does not necessarily modify the regulation of expression, but usually affects the rate of transcription (54).

Previous studies have shown that *dme* was expressed in *R. meliloti* bacteroids to levels comparable to free living cells, while *tme* expression in bacteroids appeared to be 20% or less, the level in free living cells (15, 20). In view of these differences in expression, it is perhaps not surprising to find that the sequences of the *dme* and *tme* promoters appear to be quite different (Fig. 1C). As DME activity is absolutely required for N₂-fixation, we reasoned that a symbiotic specific activation of *dme* transcription was possible. The finding that the same *dme* transcriptional start sites are employed in both free living and bacteroids appears to reduce the probability that such symbiotic regulation occurs.

Acknowledgments—We thank Siv Andersson and Raf Podowski for supplying us with the *R. prowazekii* malic enzyme sequence, Brian Golding for generous assistance in analyzing the malic enzyme sequence and phylogeny, and Patrick Chain for molecular weight determinations of DME.

REFERENCES

- Cushman, J. C. (1992) *Eur. J. Biochem.* **208**, 259–266
- Hsu, R. Y., Glynias, M. J., Satterlee, J., Feeney, R., Clarke, A. R., Emery, D. C., Roe, B. A., Wilson, R. K., Goodridge, A. G., and Holbrooks, J. J. (1992) *Biochem. J.* **284**, 869–876
- Walter, M. H., Grima-Pettenati, J., and Feuillet, C. (1994) *FEBS Lett.* **227**, 999–1009
- Garrido-Pertierra, A., Martinez Marcos, C., Martin Fernandez, M., and Ruiz-Amil, M. (1983) *Biochimie (Paris)* **65**, 629–635
- Suye, S. I., Okada, Y., Funada, A., Kawaoe, M., and Inuta, S. (1992) *J. Ferment. Bioeng.* **73**, 343–347
- Lamed, R., and Zeikus, J. G. (1981) *Biochim. Biophys. Acta.* **660**, 251–255
- Bartolucci, S., Rella, R., Guagliardi, A., Raia, C. A., Gambacorta, A., De Rosa, M., and Rossi, M. (1987) *J. Biol. Chem.* **262**, 7725–7731
- Kobayashi, K., Doi, S., Negoro, S., Urabe, I., and Okada, H. (1989) *J. Biol. Chem.* **264**, 3200–3205
- Kawai, S., Suzuki, H., Yamamoto, K., Inui, M., Yukawa, H., and Kumagai, H. (1996) *Appl. Environ. Microbiol.* **62**, 2692–2700
- Spina, J., Jr., Bright, H. J., and Rosenbloom, J. (1968) *Biochemistry* **9**, 3794–3801
- Yamaguchi, M., Tokushige, M., and Katsuki, H. (1973) *J. Biochem. (Tokyo)* **73**, 169–180

12. Ronson, C. W., Lyttleton, P., and Robertson, J. G. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 4284–4288
13. Finan, T. M., Wood, J. M., and Jordan, D. C. (1983) *J. Bacteriol.* **154**, 1403–1413
14. Driscoll, B. T., and Finan, T. M. (1993) *Mol. Microbiol.* **7**, 865–873
15. Driscoll, B. T., and Finan, T. M. (1996) *J. Bacteriol.* **178**, 2224–2231
16. Chen, F., Okabe, Y., Osano, K., and Tajima, S. (1997) *Biosci. Biotech. Biochem.* **61**, 384–386
17. Copeland, L., Quinnell, R. G., and Day, D. A. (1989) *J. Gen. Microbiol.* **135**, 2005–2011
18. Kimura, I., and Tajima, S. (1989) *Soil Sci. Plant Nutr.* **35**, 271–279
19. Yarosh, O. K., Charles, T. C., and Finan, T. M. (1989) *Mol. Microbiol.* **3**, 813–823
20. Driscoll, B. T., and Finan, T. M. (1997) *Microbiology* **143**, 489–498
21. Becker, A., Schmidt, M., Jager, W., and Puhler, A. (1995) *Gene (Amst.)* **162**, 37–39
22. Vieira, J., and Messing, J. (1987) *Methods Enzymol.* **153**, 3–34
23. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
24. Corbin, D., Barran, L., and Ditta, G. (1983) *Proc. Natl. Acad. Sci. U. S. A.* **80**, 3005–3009
25. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1987) *Current Protocols in Molecular Biology*, John Wiley and Sons Inc., New York
26. Gish, W., and States, D. J. (1993) *Nat. Genet.* **3**, 266–272
27. Higgins, D. G., and Sharp, P. M. (1988) *Gene (Amst.)* **73**, 237–244
28. Felsenstein, J. (1993) *PHYLIP (Phylogeny inference package)*, Version 3.5c, Department of Genetics, University of Washington, Seattle, WA
29. Bradford, M. M. (1976) *Anal. Biochem.* **72**, 248–254
30. Hedrick, J. L., and Smith, A. J. (1968) *Arch. Biochem. Biophys.* **126**, 155–164
31. Hansen, E. J., and Juni, E. (1975) *Biochem. Biophys. Res. Commun.* **28**, 559–566
32. Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirose, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., and Tabata, S. (1996) *DNA Res.* **3**, 109–136
33. Fleishman, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J.-F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. F., Scott, J. D., Shirley, J., Liu, R., Glodek, L.-L., Kelley, A., Weidman, J. M., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O., and Venter, J. C. (1995) *Science* **269**, 496–512
34. Latimer, M. T., and Ferry, J. G. (1993) *J. Bacteriol.* **175**, 6822–6829
35. Viljoen, M., Subden, R. E., Krizus, A., and Van Vuuren, H. J. J. (1994) *Yeast* **10**, 613–624
36. Hsu, R. Y. (1982) *Mol. Cell. Biochem.* **43**, 3–26
37. Hanukoglu, I., and Gutfinger, T. (1989) *Eur. J. Biochem.* **180**, 479–484
38. Wierenga, R. K., Terpstra, P., and Hol, W. G. J. (1986) *J. Mol. Biol.* **187**, 101–107
39. Scrutton, N. S., Berry, A., and Perham, R. N. (1990) *Nature* **343**, 38–43
40. Tang, C. L., and Hsu, R. Y. (1974) *J. Biol. Chem.* **249**, 3916–3922
41. Drincovich, M. F., Spampinato, C. P., and Andreo, C. S. (1992) *Plant Physiol.* **100**, 2035–2040
42. Chang, G.-G., and Hsu, R. Y. (1977) *Biochim. Biophys. Acta.* **483**, 228–235
43. Jawali, N., and Bhagwat, A. S. (1986) *Phytochemistry* **26**, 1859–1862
44. Rao, S. R., Kamath, B. G., and Bhagwat, A. S. (1991) *Phytochemistry* **30**, 431–435
45. Vernon, C. M., and Hsu, R. Y. (1983) *Arch. Biochem. Biophys.* **225**, 296–305
46. Bukato, G., Kochan, Z., and Swierczynski, J. (1995) *Int. J. Biochem. Cell. Biol.* **27**, 1003–1008
47. Swierczynski, J., Stankiewicz, A., Scislowski, P., and Aleksandrowicz, A. (1980) *Biochim. Biophys. Acta* **612**, 1–10
48. Blattner, F. R., Plunkett, G., Mayhew, G. F., Perna, N. T., and Glassner, F. D. (1997) *NCBI database*
49. Mahajan, S. K., Chu, C. C., Willis, D. K., Templin, A., and Clark, A. J. (1990) *Genetics* **125**, 261–273
50. Stols, L., and Donnelly, M. I. (1997) *Appl. Environ. Microbiol.* **63**, 2695–2701
51. Sanwal, B. D., and Smando, R. (1969) *J. Biol. Chem.* **244**, 1817–1823
52. Sanwal, B. D., Wright, J. A., and Smando, R. (1968) *Biochem. Biophys. Res. Commun.* **31**, 623–627
53. Osters, M., Driscoll, B. T., and Finan, T. M. (1995) *J. Bacteriol.* **177**, 1452–1460
54. Kobayashi, M., Nagata, K., and Ishihama, A. (1990) *Nucleic Acids Res.* **18**, 7367–7372