**Escherichia coli** D-Malate Dehydrogenase, a Generalist Enzyme Active in the Leucine Biosynthesis Pathway**

Received for publication, July 9, 2014, and in revised form, August 22, 2014 Published, JBC Papers in Press, August 26, 2014, DOI 10.1074/jbc.M114.595363

Anastassia A. Vorobieva†, Mohammad Shahneawz Khan‡, and Patrice Soumillion†

From the†Laboratoire de Biochimie, Biophysique et Génétique des Microorganismes (BBGM), Institut des Sciences de la Vie, Université catholique de Louvain, B-1348 Louvain-la-Neuve, Belgium and the‡Department of Fisheries, University of Dhaka, Dhaka 1000, Bangladesh

**Background:** D-Malate dehydrogenase (DmlA) is involved in D-malate catabolism in *Escherichia coli*.

**Results:** When expressed in the presence of D-malate, DmlA allows growth of a ΔleuB strain without leucine.

**Conclusion:** DmlA physiologically contributes to two core metabolic reactions.

**Significance:** Several high level activities may coexist in one enzyme and be maintained during the course of evolution.

The enzymes of the β-decarboxylating dehydrogenase superfamily catalyze the oxidative decarboxylation of D-malate-based substrates with various specificities. Here, we show that, in addition to its natural function affording bacterial growth on D-malate as a carbon source, the D-malate dehydrogenase of *Escherichia coli* (EcDmlA) naturally expressed from its chromosomal gene is capable of complementing leucine auxotrophy in a leuB− strain lacking the paralogous isopropylmalate dehydrogenase enzyme. To our knowledge, this is the first example of an enzyme that contributes with a physiologically relevant level of activity to two distinct pathways of the core metabolism while expressed from its chromosomal locus. EcDmlA features relatively high catalytic activity on at least three different substrates ([(+)-tartrate, D-malate, and 3-isopropylmalate]). Because of these properties both in vivo and in vitro, EcDmlA may be defined as a generalist enzyme. Phylogenetic analysis highlights an ancient origin of DmlA, indicating that the enzyme has maintained its generalist character throughout evolution. We discuss the implication of these findings for protein evolution.

It is regularly assumed that superfamilies of enzymes evolved from ancestors featuring broad substrate specificity but low catalytic efficiency. After gene duplication, the redundant copies of the ancestor diverged and specialized to effectively accommodate one substrate and occupy a defined position in a metabolic pathway (1–5). Supporting this evolutionary scenario, modern specialized enzymes frequently retained weak activities on secondary substrates (“substrate promiscuity”) or the ability to catalyze secondary reactions (“catalytic promiscuity”), and the secondary activity of one member of the family is often the principal activity of another member of the same family (6–8). Moreover, many enzymes feature secondary activities that are unrelated to their evolutionary history (9–11). In the course of evolution, “promiscuous” activities may be improved to a physiologically relevant level by just a few mutations in the protein and are therefore considered as potential starting points for neofunctionalization (see reviews in Refs. 12–14). For example, one point mutation in the glutamyl phosphate reductase (ProA) enables the enzyme to serve both its original metabolic function and a new one in arginine biosynthesis (15). The mutation increases a promiscuous activity on N-acetylglutamyl phosphate by an order of magnitude and results in overexpression of the protein. Overexpression is necessary to simultaneously support both the original and new functions because the mutation had a significant negative effect on the original activity. In some cases, however, the original and promiscuous activities may coexist at a relatively high level in the wild-type enzyme. Genetic complementation of a functional knock-out then only requires protein overexpression (9–11, 16–18). In most of these studies, knock-out complementations were artificially achieved by overexpressing the proteins from multicopy vector systems. Recently, Nam et al. (19) estimated that 37% of *Escherichia coli* enzymes are generalists, i.e. they promiscuously accept more than one substrate with a relatively high efficiency. By using experimental data to model metabolic networks and fluxes in *E. coli* cells, they found that these generalists can contribute to the catalysis of 65% of the known metabolic reactions while expressed under the control of their natural promoter. These results are significant because enzymes with promiscuous activities on natural metabolites are believed to play an important role in the metabolic plasticity of organisms (20). However, there is no experimental evidence showing that a naturally expressed promiscuous enzyme may simultaneously contribute to more than one metabolic flux in vivo.

The superfamily of β-decarboxylating dehydrogenases gathers metabolically important enzymes catalyzing the successive oxidation and decarboxylation of substrates with a common R-malate moiety and different γ-substituents on C-3, with various sizes, polarity, or configurations (see Fig. 1). Up to now, five families have been identified: the NAD-dependent isopropylmalate dehydrogenases (IPMDHs),2 the NAD-dependent isoc-
An Enzyme Active in Two Metabolic Pathways

FIGURE 1. Members of the β-decarboxylating dehydrogenase family catalyze several reactions on d-malate-based substrates. a, c, and d, oxidative decarboxylation catalyzed by Dml/TDHs, IPMDHs, and NAD(P)-IDHs, respectively. b, oxidation described for PpTDH.

trate dehydrogenases (IDHs), the NADP-dependent IDHs, the NAD-dependent d-malate/tartrate dehydrogenases (Dml or TDHs), and the NAD-dependent homoisocitrate dehydrogenases (HDHs). Crystal structures are available for several IDHs (21–23) and IPMDHs (24, 25), for the Thermus thermophilus HDH (26), and for the Pseudomonas putida TDH (PpTDH) (27). All of them present a conserved dimeric structure. The enzymes of the superfamily are Mg2+- or Mn2+-dependent and share conserved active site residues. They also have very different substrate specificities (see Fig. 1). IPMDHs are specific toward hydrophobic alkyd malate substrates (28), and IDHs exhibit high specificity toward the negatively charged isocitrate (29).

In addition to homoisocitrate, the HDH of Deinococcus radiodurans also accepts isocitrate and isopropylmalate (30), whereas the HDHs of T. thermophilus and Pyrococcus horikoshii are active on isocitrate and homoisocitrate only (31, 32). Finally, although PpTDH was first identified as an enzyme supporting growth on l(+)-tartrate, it was later shown to be also active on d-malate and 3-isopropylmalate in vitro (33, 34). Although the enzyme transforms both d-malate and 3-isopropylmalate by oxidative decarboxylation, a non-decarboxylating oxidation is observed with l(+)-tartrate (Fig. 1).

In E. coli, the d-malate dehydrogenase EcDmlA (dmlA gene) with high similarity to PpTDH (76% of identity) is responsible for aerobic growth on d-malate as the sole carbon source (35, 36). It coexists with a paralogous EcIPMDH (leuB gene) involved in leucine biosynthesis. It has long been known that E. coli strains harboring the leuB6 genotype produce revertants under leucine starvation conditions (37). However, the fact that most of these revertants are suppressors (i.e., no mutation could be detected in the leuB locus) indicates that the loss of EcIPMDH can be complemented by another protein (37, 38). In this work, we show that chromosomal expression of EcDmlA affords complementation of the leuB+ phenotype when induced by d-malate. The EcDmlA-dependent leuB+ phenotype is not observed with glucose. These results are further supported by the biochemical characterization of the purified enzyme, and we discuss its evolutionary implications with the help of a phylogenetic analysis of the superfamily.

EXPERIMENTAL PROCEDURES

Strains and Chemicals—The leuB strain HBLB1 was constructed from the HB101 strain using a standard phage P1 transduction protocol (39). The leuB6 gene was replaced with a kanamycin resistance cassette using the JW5807-2 strain (ΔleuB::kan) from the Keio Collection (40) as a donor. The cassette insertion was validated by PCR and sequencing.

Cloning—The dmlA gene was amplified from the E. coli DH12S genome by PCR (Phusion polymerase, Fisher Scientific) and cloned between the NdeI and BamHI sites of the pMal-pIII vector (New England Biolabs) in C-terminal fusion with a Strep-tag II. In the final construct called phDmlA, the dmlA gene replaced the entire malE-containing ORF and was under the control of the P_tac promoter. The genetic construction was verified by sequencing.

Phenotypic Auxotrophy Tests—All phenotypic tests were performed on amino acid-free EZ Rich defined medium. The medium was prepared as described by Neidhardt et al. (41), and the 10× EZ solution used to prepare the medium was purchased from Teknova (Hollister, California). For solid medium culture, the medium was solidified with 1.5% agar. The carbon sources were added up to 0.1% (d-glucose) or 0.15% (d-malate), and leucine was used at 1 mM if specified. Proline (1 mM) was added to the culture medium to support growth of the HBLB1 strain. Liquid cultures (100 ml) were inoculated from an overnight preculture of HBLB1 cells preliminary washed with M9 buffer. CaCO₃ (10 mM) was added to the medium to reduce the lag time as described (41). Despite this step, a lag time of ~16 h was observed. The cells were inoculated at an initial A₅₆₀ of 0.05 and incubated overnight at 37 °C and 180 rpm. Growth was monitored the next day. For growth tests on solid medium, precultures grown overnight at 37 °C and 180 rpm in LB/kanamycin (50 µg/ml) were harvested by centrifugation and resuspended in 3 ml of M9 buffer. The cells were then sequentially diluted in M9 buffer up to 10⁻³-fold. 10 µl of each dilution was spotted on the agar. The plates were incubated at 37 °C for 24 h.

Activity Assay in Cell Extracts—50 ml of exponentially grown cell cultures (A₅₆₀ = 0.2 for the culture with d-malate (0.15%) and leucine (1 mM)), A₅₆₀ = 0.1 for the culture with d-malate (0.15%), and A₅₆₀ = 0.45 for the culture with d-glucose (0.1%) and leucine (1 mM)) were centrifuged at 4000 rpm for 10 min, and the cells were resuspended in 1 ml of KAC buffer (25 mM MOPS and 100 mM KCl, pH 7.5) (29, 42) supplemented with one tablet of Complete mini EDTA-free protease inhibitor mixture (Roche Applied Science). The samples were flash-frozen in liquid nitrogen and kept at −80 °C. For activity tests, the cells were crushed in a FastPrep-24 instrument (MP Biomedicals, Santa Ana, California) for 60 s (250 µl of glass bead suspension for 750 µl of cell suspension). The extracts were cleared by centrifugation and diluted 10 times for the activity measurements in KAC buffer. To start the reaction, 25 µl of diluted extracts was added to 125 µl of reaction mixture (500 µM d-malate or 3-isopropylmalate, 250 µM NAD⁺, 5 mM MnCl₂, 1...
was purified by DTT buffer (25 mM MOPS, 100 mM KCl, and 1 mM DTT, pH 8.0) supplemented with one tablet of Complete mini EDTA-free protease inhibitor mixture and flash-frozen in liquid nitrogen. Thawed cells were treated with lysozyme (1 mg/ml) and resuspended in 20 ml of buffer (100 mM Tris and 150 mM NaCl, pH 8.0). The concentration of protein in the extracts, measured by BCA protein assay (Pierce).

**DmlA Expression and Purification**—The protein was expressed from BL21 cells transformed with phDmlA. Typically, 500 ml of Terrific Broth medium was inoculated and grown at 37 °C and 180 rpm until A<sub>600</sub> reached 0.6. Isopropyl β-D-thiogalactopyranoside was added (1 mM final concentration), and the protein was produced at 25 °C and 180 rpm for 4 h. The cells were centrifuged and resuspended in 20 ml of buffer (100 mM Tris and 150 mM NaCl, pH 8.0) supplemented with one tablet of Complete mini EDTA-free protease inhibitor mixture and flash-frozen in liquid nitrogen. Thawed cells were treated with lysozyme (1 mg/ml, 1 h), DNase (5 μg/ml) was added for an additional 20 min, and lysis was completed by 30 min of sonication (cycles of 30 s of ultrasound and 30 s of resting on ice). The protein was purified from the cleared lysate by Strep-Tactin affinity chromatography (IBA, Göttingen, Germany). The fractions containing the protein were pooled, desalted (final buffer of 10 mM Tris, pH 8.0), and concentrated by ultracentrifugation in Amicon tubes (10-kDa cutoff; Millipore). The concentration of the protein was determined by measuring the absorbance at 280 nm using a predicted extinction coefficient of 64,400 M<sup>-1</sup>·cm<sup>-1</sup> (ProtParam tool on the ExPASy server (43)). The final concentration was typically in the range of 6 mg/ml. 25-[<sup>34</sup>P]orthophosphate (ProtParam tool on the ExPASy server (43)). The final concentration of protein in the extracts, measured by BCA protein assay (Pierce).

**Determination of Catalytic Parameters for Different Substrates**—All catalytic measurements were performed in KAC/DTT buffer (25 mM MOPS, 100 mM KCl, and 1 mM DTT, pH 7.5) at 25 °C. The pH of the substrates was adjusted to pH 7.5 with KOH. MnCl<sub>2</sub> or MgCl<sub>2</sub> (100 mM stock solutions in water) was added up to 5 mM to each reaction mixture. NAD<sup>+</sup> was added up to 500 μM (for d-malate and 3-isopropylmalate assays) or 2 mM (for L(+)tartrate assays). For the other substrates, the NAD concentration was arbitrarily fixed at 1 mM to obtain apparent kinetic parameters. The reaction was started by the addition of 25 μl of the enzyme diluted in KAC/DTT buffer. The concentration of d-malate and 3-isopropylmalate ranged from 10 to 500 μM. The concentration of L(+)tartrate ranged from 0.1 to 4 mM. DmlA was loaded on the gel. The size of the DmlA monomer was 42 kDa.

![FIGURE 2. SDS-PAGE analysis of the purified DmlA protein.](image-url)
ing mutations of conserved active site residues were discarded. As most of the sequences obtained were from /H9251-,
/H9252- and /H9253-Proteobacteria, we kept only the sequences from these classes of organisms in the final data set for a more clear analysis. Sequences from fungi, archaea, and other Proteobacteria were discarded, as well as two HDH sequences. Finally, the data set was screened for highly similar sequences. We tried to keep Dml/TDHs and IPMDHs from the same organisms in the data set when it was possible. The resulting final data set contains 95 different IDH, Dml/TDH, IPMDH, and TtuC sequences (see supplemental material). The annotation attributed to some proteins did not correspond to the conserved motif in the substrate recognition structure (45). These sequences were properly re-annotated. Alignments were done with the Clustal algorithm. The best fit model for amino acid replacement was determined using the ProtTest 3 server (46). The LG model was selected. The phylogenetic tree was calculated with the PhyML program, and 100 rounds of bootstrapping analysis were performed (47). The trees were generated using the TreeDyn algorithm, available online at the Phylogeny.fr web site (48). Finally, to evaluate the robustness of the tree, similar analyses from initial Muscle and T-Coffee alignments and from the Clustal alignment using the WAG substitution model were performed. Similar conclusions could be drawn from all constructed trees.

RESULTS

Expression of DmlA under the Control of Its Native Promoter Complements the ΔleuB Phenotype—Given the high similarity of EcDmlA to PpTDH and the activity of PpTDH detected on 3-isopropylmalate, we hypothesized that EcDmlA could be responsible for suppressive leuB knock-out complementation. To study the IPMDH function of EcDmlA in vivo and to avoid direct leuB+ reversion, we constructed the HBLB1 strain, in which the leuB gene of the HB101 strain was deleted (ΔleuB strain). Leucine auxotrophy was tested in chemically defined medium with different carbon sources. As shown in Fig. 3a, with D-glucose as the carbon source, leucine supplementation was necessary to support HBLB1 growth. With D-malate, however, HBLB1 cells grew without leucine, and growth was inhibited when glucose was added as an alternative carbon source to D-malate. The same result was obtained on solid medium (data not shown). As EcDmlA expression is induced by D-malate and repressed by glucose (49), this experiment supports our initial
hypothesis. In liquid medium, $A_{600}$ reached $\sim 0.4$ on D-malate and leucine and 0.2 on D-malate at stationary phase. The lower plateau reached by the cells in the absence of leucine indicates that the fitness of the strain is limited by either the D-malate catabolism or the synthesis of leucine, as both pathways are supported by the same enzyme and compete with each other. We collected samples of exponentially growing cells and tested the activities on D-malate (Dml) and 3-isopropylmalate (IPMDH) in the extracts. Both activities were present in cells grown on D-glucose medium supplemented with leucine and absent from cells grown on D-glucose medium supplemented with leucine (Fig. 3b). No difference in the level of Dml or IPMDH activity was detected between cells grown on D-malate in the presence or absence of leucine. Altogether, these results indicate that normal induction of EcDmlA by D-malate is sufficient for leuB+ complementation and that no overexpression mechanism is involved. To confirm the role of EcDmlA in leucine prototrophy, its gene was cloned in an expression plasmid under the control of the P_lac promoter. Competent HBLB1 cells were transformed with the resulting phDmlA vector. With glucose as the carbon source, leucine prototrophy of the phDmlA-HBLB1 strain depended on DmlA induction by isopropyl β-D-thiogalactopyranoside (IPTG) (Fig. 3c). Besides the selection of a few revertants, no growth was observed for the control strain transformed with the empty vector.

EcDmlA is a Generalist Enzyme—Strep-tag II-fused EcDmlA was expressed, purified, and biochemically characterized. Similar to other β-decarboxylating dehydrogenases, EcDmlA activity depended on the presence of divalent metal ion (Mn$^{2+}$ or Mg$^{2+}$) (data not shown). As several TDHs and HDHs are activated by K$^+$ or NH$_4^+$ and are less active in the presence of Na$^+$ (33, 50, 51), we tested the effect of these monovalent cations on the activity of EcDmlA and found a similar behavior (Fig. 4). Enzyme kinetics (50) and structural data (27) suggest that the K$^+$ or NH$_4^+$ ion interacts with the NAD and helps in cofactor positioning. The activity-pH profile has a bell shape with an optimum at pH 7.0–7.5 (data not shown). As predicted by amino acid sequence, EcDmlA preferred NAD to NADP by $\sim 500$-fold (data not shown). The activity of the enzyme on D-malate ($k_{cat}/K_m$(malate) = $10^5$ M$^{-1}$ s$^{-1}$) (Table 1) corresponded to that reported for a moderately efficient enzyme in energy metabolism (52). Strikingly, the enzyme exhibited high activity on 3-isopropylmalate, with $k_{cat}/K_m$(IPMDH) = $4 \times 10^3$ M$^{-1}$ s$^{-1}$ (Fig. 4). This is only 50 times lower than the constant for EcIPMDH at the same temperature ($k_{cat}/K_m$(IPMDH) = $2 \times 10^4$ M$^{-1}$ s$^{-1}$) (28), supporting leuB+ complementation. Finally, significant activity on L(+)-tartrate was also detected ($k_{cat}/K_m$(L(+)-tartrate) = $5 \times 10^5$ M$^{-1}$ s$^{-1}$). The impact of divalent metal ion on substrate specificity was investigated by replacing Mn$^{2+}$ with Mg$^{2+}$ for activity measurements. With the Mg$^{2+}$ metal ion, all three activities decreased by 2–5-fold, but their relative levels remained roughly unchanged (data not shown).

The $k_{cat}/K_m$ parameters obtained for EcDmlA are very similar to the data reported for PpTDH at the same temperature for D-malate, 3-isopropylmalate, and L(+)-tartrate (Table 1; data from Ref. 33). However, overexpressed EcDmlA did not afford any growth of the HBLB1 strain on L(+)-tartrate (Table 1; data from Ref. 33). In a similar experiment, Lukas et al. (49) reported a very slow growth but only when the medium was enriched with amino acids. The turnover rate constant ($k_{cat}$) of DmlA with L(+)-tartrate was twice as high as that with 3-isopropylmalate. However, the $K_m$ for L(+)-tartrate and

![FIGURE 4. Monovalent cations modulate the catalytic activity of DmlA.](image)

The activity of DmlA on D-malate was tested in 25 mM MOPS buffer, pH 7.5, supplemented with 150 mM K$^+$, NH$_4^+$, or Na$^+$ chloride salts. The activity of the enzyme with KCl was set as the reference.

### TABLE 1

| Substrate     | $k_{cat}$ (μmol/min/mg) | $K_m$ (μmol/L) | $k_{cat}/K_m$(substrate) | $K_m$(NAD) (μmol/L) | $k_{cat}/K_m$(NAD) | $k_{cat}/K_m$(substrate) |
|---------------|-------------------------|----------------|--------------------------|---------------------|------------------|--------------------------|
| D-Malate$^a$ | 11.7 ± 0.3              | 96 ± 7         | $1 \times 10^5$          | $1 \times 10^5$     | $2 \times 10^6$   | $1 \times 10^5$          |
| 3-Isopropylmalate$^a$ | 0.102 ± 0.003          | 24 ± 3         | $4 \times 10^3$          | $70 \pm 3$          | $2 \times 10^4$   | $4 \times 10^3$          |
| L(+)-Tartrate$^a$ | 0.26 ± 0.01            | 567 ± 84       | $5 \times 10^4$          | $600 \pm 200$       | $4 \times 10^5$   | $5 \times 10^4$          |
| D-Isocitrate$^a$ | (5.4 ± 0.2) × 10$^{-4}$ | 1300 ± 125     | $4 \times 10^3$          | ND                  | ND               | ND                       |
| L(-)-Tartrate$^a$ | ND                      | $\geq 8000$    | $3 \times 10^3$          | ND                  | ND               | ND                       |
| D-Lactate$^a$ | ND                      | $\geq 14,000$  | $1 \times 10^4$          | ND                  | ND               | ND                       |

$^a$ Apparent $K_m$ for the specified substrate with a fixed concentration of NAD.

$^b$ Apparent $K_m$ for NAD with a fixed concentration of the specified substrate.

$^c$ Data are taken from Ref. 33.

$^d$ For these substrates, the concentration of NAD was fixed at 500 μM for the $K_m$(substrate) measurements, and the concentration of substrates was fixed at 500 μM for the $K_m$(NAD) measurements.

$^e$ For the L(+)-tartrate, the concentration of NAD was fixed at 2 mM for the $K_m$(substrate) measurements, and the concentration of substrates was fixed at 5 mM for the $K_m$(NAD) measurements.

$^f$ For these substrates, the concentration of NAD was fixed at 1 mM to obtain apparent $k_{cat}/K_m$(substrate).

$^g$ Because of the high $K_m$(substrate), the kinetic parameters could not be obtained for these substrates, and the $k_{cat}/K_m$(substrate) values were estimated from the linear part of the Michaelis-Menten function.
NAD was 5–20-fold higher than the $K_m$ for all previously tested substrates. Additionally, a low affinity of EcDmlA for L(-)-tartrate was suggested by thermal unfolding of the protein in a differential scanning fluorometry experiment (Fig. 5). In a differential scanning fluorometry assay, the temperature at which a protein unfolds in the presence of a ligand is compared with its unfolding temperature in the apo form. The stabilizing effect increases with binding affinity (53, 54). In the presence of 1 mM D-malate or 3-isopropylmalate, the protein unfolded at higher temperature compared with its apo form (1.5–2 °C higher $T_m$). In contrast, no significant $T_m$ shift was observed in the presence of the same concentration of L(+)-tartrate, indicating that the interaction between EcDmlA and L(+)-tartrate is weaker than the interaction between EcDmlA and D-malate or 3-isopropylmalate. Notably, L(+)-tartrate is the only substrate in Fig. 1 with C-3 in the preferred R-configuration, whereas all other substrates feature the S-configuration. The increase in $K_m$ for both L(+)-tartrate and NAD may be explained by steric hindrance between the substrate and cofactor or between the substituent on C-3 of L(+)-tartrate and a structural part of the active site.

In the results presented above, we showed that EcDmlA was active on a substrate with a large hydrophobic $\gamma$-substituent (3-isopropylmalate), as well as a substrate with hydrogen in that position (D-malate) and a substrate with a polar hydroxyl as a $\gamma$-substituent in an inverted configuration (L(+)-tartrate). We further observed a weak activity of the enzyme on isocitrate ($k_{cat}/K_{m(isocitrate)} = 4 \times 10^{-1} \text{M}^{-1} \text{s}^{-1}$), a molecule with a negatively charged $\gamma$-substituent. D-Lactate, a minimalistic substrate without the C-4 carboxyl, was also a poor substrate for EcDmlA ($k_{cat}/K_{m(lactate)} = 1 \times 10^{-1} \text{M}^{-1} \text{s}^{-1}$). Finally, the enzyme was also active on D(-)-tartrate ($k_{cat}/K_{m(D(-)-tartrate)} = 3 \times 10^{-1} \text{M}^{-1} \text{s}^{-1}$), which has C-3 in the preferred S-configuration but features an inverted configuration of C-2 involved in hydride transfer with the cofactor. Hence, EcDmlA is a highly versatile enzyme and accepts substrates featuring different sizes and polarities of the $\gamma$-substituent, a carboxyl or a hydrogen as the other substituent of C-3, and both configurations of C-2 and C-3. The rate versus substrate concentration data for all tested substrates are shown in Fig. 6.

**Phylogenetic Analysis Suggests an Ancient Origin for Dml/TDH Enzymes—**D-Malate/tartrate dehydrogenases are metabolic enzymes supporting aerobic growth either on D-malate (Dml activity) or on L(+)-tartrate (TDH and TtuC activities). It is important to emphasize that the current data on the Dml/TDH family are limited, especially at the metabolic level. Hence, the enzymes hereafter are annotated according to their initial characterization (Dml or TDH), which may not precisely reflect their actual metabolic role(s). So far, biochemical data are available for PpTDH, *Pseudomonas fluorescens* TDH (55), and *Rhodobacter sphaeroides* TDH (51), as well as EcDmlA (this study). Additionally, the metabolic role of two other proteins was experimentally confirmed: D-malate growth for *Enterobacter aerogenes* Dml (55) and L(+)-tartrate growth for *Agrobacterium vitis* TtuC (56).

The phylogenetic analysis of a data set of sequences homologous to PpTDH and EcDmlA (minimum 35% identity) showed that the sequences cluster into three groups: the IPMDHs, NAD-dependent IDHs, and Dml/TDHs (Fig. 7). In the data set, the function of the proteins was assigned using the conserved motifs in the substrate recognition region identified by Chen and Jeong (45). The motif SXX is conserved in IDHs, whereas the motifs EX$_2$-$a$(L/L)L and LX$_2$L are proper to 3-isopropylmalate and D-malate/tartrate dehydrogenases, respectively (45, 57).

The evolutionary trajectory that was hypothesized for explaining the low activity of PpTDH on L(+)-tartrate and its activity on isopropylmalate starts by a recent duplication of an IPMDH gene, followed by the evolution of the enzyme into a generalist, and ends with the partial re-specialization for the new catalytic function (33). However, the topology of our phylogenetic tree suggests that the Dml/IDH activity is more ancient than previously thought. Instead of branching to the IPMDH cluster, the Dml/IDHs are monophyletic and diverge more or less from the same point as the NAD-dependent IDHs and IPMDHs. Hence, Dml/IDH proteins likely evolved directly from the common ancestor of the $\beta$-carboxylating dehydrogenases that was proposed to be a generalist enzyme (58) rather than from a duplicated IPMDH gene. Another interesting observation is that the generalist character seems to be a common feature of the Dml/IDH enzymes. Activity on D-malate, L(+)-tartrate, and 3-isopropylmalate was reported for PpTDH and EcDmlA, sharing 76% identity along the entire sequence length. D-Malate and tartrate activities were also reported for R. sphaeroides TDH, sharing 60 and 65% identities, respectively, with the two others and located in a different phylogenetic cluster. The highest activity of the purified *R. sphaeroides* TDH, the Dml activity, is only 2 orders of magnitude higher than its TDH activity (in terms of $k_{cat}/K_{m}$) (51). Moreover, *R. sphaeroides* TDH supports the growth of *R. sphaeroides* on L(+)-tartrate. The activity of this enzyme toward 3-isopropylmalate was not tested. The promiscuous properties of these three divergent enzymes suggest that the generalist trait was maintained through the evolution of Dml/IDHs.

**An Enzyme Active in Two Metabolic Pathways**

![FIGURE 5. DmlA unfolds at a higher temperature in the presence of D-malate and 3-isopropylmalate, but not L(+)-tartrate. $T_m$ was measured by differential scanning fluorometry of the apo form and in the presence of 1 mM D-malate, 3-isopropylmalate, and L(+)-tartrate. The $\Delta T_m$ values represent the difference in $T_m$ between the protein in the presence of the specified substrate and the apoprotein.](image-url)
FIGURE 6. Kinetic characterization of DmA on various substrates. The $k_{obs}$ values were measured for six substrates at different concentrations with fixed concentrations of NAD$^+$ (see “Experimental Procedures” for concentrations) and Mn$^{2+}$ (5 mM). The kinetic parameters were calculated from the data by fitting to the Michaelis-Menten equation (Equation 2) (A–C and F) and by estimating the $k_{cat}/K_m$ parameter from the linear part of the activity-substrate curve (D and E). The calculated parameters are reported in Table 1.
The phylogenetic tree also shows that TtuC-like proteins form a subcluster among the \( \delta \)-malate/tartrate dehydrogenases (Fig. 7). Surprisingly, some active site residues that are usually conserved in the enzymes of the \( \beta \)-decarboxylating dehydrogenase superfamily are modified in TtuC proteins (Table 2). First, the last leucine of the \( \text{LXXXL} \) motif specific to the substrate recognition loop of the Dml/TDH enzymes is replaced with arginine. Second, Arg-97 (\( Ec \)DmlA numbering), which is conserved in all \( \alpha \)-decarboxylating dehydrogenases and interacts with both carboxylates of the common \( \delta \)-malate moiety of the substrates, is mutated to cysteine in TtuC proteins. Despite these mutations, \( A. \ vitis \) TtuC is essential for growth on \( \delta \)-(\( H \))-tartrate (56). This suggests a unique substrate-binding mechanism among the \( \beta \)-decarboxylating dehydrogenases. As many organisms, such as \( R. \ sapheoides \), \( K. \ pneumoniae \), and \( A. \ tumefaciens \), contain both TtuC and TDH, the enzymes may have different metabolic roles or functions.

**DISCUSSION**

In this work, we have shown that the promiscuous activity of wild-type and naturally expressed \( Ec \)DmlA on 3-isopropylmalate is high enough to complement the knock-out of the paralogous LeuB enzyme from the leucine biosynthesis pathway. Moreover, \( Ec \)DmlA is simultaneously active in its original \( \delta \)-malate catabolism pathway and in the leucine biosynthesis pathway. The bifunctionality is significant because relatively high fluxes through these two pathways are essential to the bacteria under the applied growth conditions. To our knowledge, this result provides the first experimental example of a wild-type generalist enzyme supporting two core metabolic fluxes in vivo in the absence of protein overexpression. The capacity of \( Ec \)DmlA to function in both metabolic pathways is explained by its high catalytic activity on \( \delta \)-malate and isopropylmalate. The enzyme also has a relatively high activity on \( \delta \)-(\( H \))-tartrate, although our results indicate a weak affinity of the protein for this alternative substrate compared with \( \delta \)-malate and 3-isopropylmalate. This can be linked to the work on \( Pp \)TDH done by Serfozo and Tipton (59, 60), who suggested a common binding mode for substrates featuring the \( S \)-configuration of C-3 (\( \delta \)-malate, 3-isopropylmalate) and an alternative binding mode for substrates featuring the \( R \)-configuration of C-3 (\( \delta \)-(\( H \))-tartrate). The alternative \( R \)-configuration binding
mode has not been characterized yet and was proposed to explain the release of the oxidized intermediate without further decarboxylation. As the kinetic values obtained for EcDmlA are very similar to the kinetic values reported for PpTDH on all three substrates (33), the catalytic mechanisms of both enzymes should be similar, and the inability of EcDmlA overexpression to support growth of the bacteria on L(+)-tartrate is more likely due to a lack of membrane transporters for L(+)-tartrate or to metabolic differences between E. coli and P. putida.

Enzymes are generally described in textbooks as remarkable specialists. However, the number of promiscuous enzymes reported in the literature is rapidly growing, suggesting that most enzymes promiscuously accept alternative substrates (13). EcDmlA is a remarkable example because it features relatively high activities on three different substrates and may therefore be defined as a generalist. Moreover, the topology of our phylogenetic tree suggests that Dml/TDHs are ancient and therefore be defined as a generalist. Moreover, the whole family most probably shares the generalist trait, although a more systematic study would be necessary to confirm this hypothesis.

Why was EcDmlA maintained as a generalist in the course of evolution? Was it selected for several activities, or is it just the result of fortuitous enzyme infidelity (i.e. promiscuity (61))? The later hypothesis is supported by the regulation mechanism of the dmlA gene. The expression of the protein is determined by the carbon source used, independent of the presence or absence of leucine, because no overexpression could be detected under leucine starvation conditions. This indicates that its IPMDH activity is a consequence of enzyme infidelity rather than a selected alternative route for leucine biosynthesis. We propose two hypotheses to explain the evolutionary retention of the generalist trait.

(i) The selection pressure for specialization may be low. In other words, promiscuity may be maintained as long as it does not interfere with any metabolic function. Although EcDmlA will have to support a high metabolic flux when L-malate is used as the carbon source, these conditions may be very occasional (49). Moreover, when DmlA is expressed in a wild-type strain with an active leucine biosynthesis pathway, the two pathways will probably not interfere significantly. Indeed, as an intermediate metabolite, 3-isopropylmalate will probably be maintained at a low concentration. In contrast, if a source of exogeneous leucine is present in the medium, the expression of the genes encoded in the leu operon will be down-regulated (62), and the intracellular concentration of the intermediate metabolites will be even lower.

(ii) The full specialization may be difficult to achieve in this particular active site scaffold. Indeed, the positioning of the cofactor relative to the substrate is a critical aspect for catalysis by the β-decarboxylating dehydrogenases (63, 64). The cofactor may be constrained to its optimal conformation by active site residues (as for IPMDHs (63)) or by the γ-substituent of the substrate (as for IDHs (64)). For PpTDH, structural data suggest that cofactor positioning probably involves an interaction with the monovalent cation (K⁺ or NH₄⁺) that is necessary for catalysis (27, 33). However, the interactions between K⁺ ions and their ligands are dominantly electrostatic, and a wider range of distances and coordination numbers are then acceptable (65). An active site geometry shaped by a K⁺ ion would probably allow more conformational changes to accommodate different substrates. This hypothesis is supported by the biochemical data obtained for other β-decarboxylating dehydrogenases. Indeed, among the biochemically characterized HDHs, the cation-dependent HDH of D. radiodurans features broader substrate specificity than the HDHs of T. thermophilus and P. horikoshii, which do not require any monovalent cation (30–32). The IDHs and IPMDHs, which are cation-independent, are more specific to their respective natural substrates.

How can the generalist traits of an enzyme impact the metabolic plasticity of the organism? Enzyme promiscuous activities are believed to be good starting points for evolution by providing possibilities of innovation before or after a gene duplication event. As discussed in the Introduction, McLoughlin and Copley (15) reported an example of an enzyme featuring one unique mutation that affords bifunctionality but at the expense of a significant negative trade-off between the new and old functions. Hence, at the early stage of evolution, a compromise between the newly evolved and original activities has to be found to allow the survival of the organism. The particular example of EcDmlA shows that a secondary high level activity may coexist with the original one without any mutation-associated compromise. Despite the fact that the activity of the enzyme on 3-isopropylmalate is high (kcat/Km ~ 10³ M⁻¹ s⁻¹) and sufficient to affect the phenotype of the organism, its principal activity on L-malate remains compatible with an efficient enzyme from secondary metabolism (52). There is, however, a compromise imposed by gene sharing but at the metabolic level when both activities are simultaneously required. Indeed, the
growth of the ΔleuB strain on D-malate is lower when exogenous leucine is absent, indicating a competition between the two metabolic pathways sustained by the same enzyme. If the selection pressure for one or both activities increases because higher or more constant metabolic fluxes are required, the fitness will probably improve upon a simple overexpression of the protein by mutation in the promoter region or by gene duplication. Interestingly, sequence alignments suggest that TtuC enzymes form a distinct cluster among Dml/TDHs and feature a different substrate-binding motif. The gene coding for TtuC of A. vitis (a grapevine pathogen) belongs to the tartrate operon (ttu), suggesting that TtuC-like enzymes could have arisen from the specialization of a tdh gene for tartrate utilization (56). In A. tumefaciens, the tttuC gene is encoded in a similar tartrate ttu operon and coexists with a tdh gene (see supplemental material), suggesting a gene duplication scenario. If TtuC enzymes evolved to efficiently catabolize tartrate, then this specialization required consequent mutations of otherwise conserved residues. Hence, a deeper study of substrate specificities among Dml/TDHs regarding the potential existence of structural features promoting enzyme promiscuity or specialization and the trade-offs between the different activities in the course of evolution would be very informative.

Are these results relevant in the context of natural evolution? The experimental design applied in this study is artificial, as a metabolic limitation was created by the knock-out of a metabolically important gene. However, it mimics a new requirement for neo-functionalization that could arise naturally due to environmental changes. It is yet unclear if other proteins have secondary activities close to a physiologically relevant level. In particular, it is important to emphasize that all of the systematic screenings for enzyme generalists were realized under overexpression conditions using multicopy vectors (9, 10). However, some of the enzymes selected in these studies may be sufficiently active on the alternative substrates to afford a phenotype without overexpression. Previous studies (11, 15) showed that a single mutation in a chromosomal gene might increase a promiscuous activity at a level sufficient to complement a metabolic knock-out. Hence, the development of new techniques for directed evolution of chromosome-encoded proteins is a key challenge for gaining further insight into the mechanisms governing natural evolution of enzymes.

Acknowledgments—We acknowledge Prof. Jean-François Collet (Université catholique de Louvain) for providing the strain from the Keio Collection. We also thank Dr. Florian Holfelder (University of Cambridge, Cambridge, United Kingdom) for comments on the manuscript.

REFERENCES

1. Ohno, S. (1970) Evolution by Gene Duplication, Springer-Verlag, New York.
2. Voordekers, K., Brown, C. A., Vanneste, K., van der Zande, E., Voet, A., Maere, S., and Verstrepen, K. J. (December 11, 2012) Reconstitution of ancestral metabolic enzymes reveals molecular mechanisms underlying evolutionary innovation through gene duplication. PLoS Biol. 10, e1000446
3. Huang, R., Hippauf, F., Rohrbeck, D., Haustein, M., Wenke, K., Feike, J., Sorrelle, N., Piechulla, B., and Barkman, T. J. (2012) Enzyme functional evolution through improved catalysis of ancestrally nonpreferred substrates. Proc. Natl. Acad. Sci. U.S.A. 109, 2966–2971
4. Jensen, R. A. (1976) Enzyme recruitment in evolution of new function. Annu. Rev. Microbiol. 30, 409–425
5. Copley, S. D. (2012) Toward a systems biology perspective on enzyme evolution. J. Biol. Chem. 287, 3–10
6. Glasner, M. E., Gerlt, J. A., and Babbitt, P. C. (2006) Evolution of enzyme superfamilies. Curr. Opin. Chem. Biol. 10, 492–497
7. Afriat, L., Roordveldt, C., Manco, G., and Tawfik, D. S. (2006) The latent promiscuity of newly identified microbial lactonases is linked to a recently diverged phosphotriesterase. Biochemistry 45, 13677–13686
8. van Loo, B., Jonas, S., Babbie, A. C., Benjdia, A., Berteau, O., Hyvönen, M., and Holfelder, F. (2010) An efficient, multiply promiscuous hydrolyase in the alka-line phosphatase superfamily. Proc. Natl. Acad. Sci. U.S.A. 107, 2740–2745
9. Wallon, G., Kryger, G., Lovett, S. T., Oshima, T., Ringe, D., and Petsko, G. A. (2005) Crystal structure of the tetrameric homoisocitrate dehydrogenase. Biochemistry 44, 7819–7829
10. Yip, S. H., and Matsumura, I. (2013) Substrate ambiguous enzymes within the E. coli proteome offer different evolutionary solutions to the same problem. Mol. Biol. Evol. 30, 2001–2012
11. Khersonsky, O., Roodveldt, C., and Tawfik, D. S. (2006) Enzyme promiscuity: evolutionary and mechanistic aspects. Curr. Opin. Chem. Biol. 10, 498–508
12. Khersonsky, O., and Tawfik, D. S. (2010) Enzyme promiscuity: a mechanistic and evolutionary perspective. Annu. Rev. Biochem. 79, 471–505
13. Miller, B. G., and Raines, R. T. (2004) Identifying latent enzyme activities: substrate ambiguity within modern bacteria sugar kinases. Biochemistry 43, 6387–6392
14. Miller, B. G., and Raines, R. T. (2005) Reconstitution of a defunct glycolytic pathway via recruitment of ambiguous sugar kinases. Biochemistry 44, 10776–10783
15. Fedey, A. E., Yang, N., Martinez, A., Leiros, H. K., and Steen, I. H. (2004) Structural and functional properties of isocitrate dehydrogenase from the psychrophilic bacterium Desulfitobacterium autotrophicum, a phosphite-dependent dehydrogenase. Proc. Natl. Acad. Sci. U.S.A. 101, 7919–7924
16. Hall, W. L., Xu, Y., and Metcalf, W. W. (2003) A new activity for an older enzyme: Escherichia coli bacterial alkaline phosphatase is a phosphate-dependent dehydrogenase. J. Mol. Biol. 330, 8635–8639
17. Wallen, G., Kryger, G., Lovett, S. T., Oshima, T., Ringe, D., and Petsko, G. A. (1997) Crystal structure of E. coli and Salmonella typhimurium 3-isopropylmalate dehydrogenase. J. Mol. Biol. 266, 1016–1031
18. Miyazaki, J., Asada, K., Fushinobu, S., Kuzuyama, T., and Nishiyama, M. (2005) Crystal structure of the tetrameric homoisocitrate dehydrogenase from an extreme thermophile, Thermus thermophilus: involvement of hydrophobic dimer-dimer interaction in extremely high thermostolerance. J. Bacteriol. 187, 6779–6788
