Experimental validation of in silico model-predicted isocitrate dehydrogenase and phosphomannose isomerase from Dehalococcoides mccartyi

The MIT Faculty has made this article openly available. Please share how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1111/1751-7915.12315">http://dx.doi.org/10.1111/1751-7915.12315</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>Wiley Blackwell</td>
</tr>
<tr>
<td>Version</td>
<td>Final published version</td>
</tr>
<tr>
<td>Accessed</td>
<td>Wed Feb 06 19:48:44 EST 2019</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/99704">http://hdl.handle.net/1721.1/99704</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Creative Commons Attribution</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td><a href="http://creativecommons.org/licenses/by/4.0/">http://creativecommons.org/licenses/by/4.0/</a></td>
</tr>
</tbody>
</table>
Experimental validation of in silico model-predicted isocitrate dehydrogenase and phosphomannannose isomerase from Dehalococcoides mccartyi

M. Ahsanul Islam, Anatoli Tchigvintsev, Veronica Yim, Alexei Savchenko, Alexander F. Yakunin, Radhakrishnan Mahadevan and Elizabeth A. Edwards

Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON M5S 3E5, Canada

Summary

Gene sequences annotated as proteins of unknown or non-specific function and hypothetical proteins account for a large fraction of most genomes. In the strictly anaerobic and organohalide respiring Dehalococcoides mccartyi, this lack of annotation plagues almost half the genome. Using a combination of bioinformatics analyses and genome-wide metabolic modelling, new or more specific annotations were proposed for about 80 of these poorly annotated genes in previous investigations of D. mccartyi metabolism. Herein, we report the experimental validation of the proposed reannotations for two such genes (KB1_0495 and KB1_0553) from D. mccartyi strains in the KB-1 community. KB1_0495 or DmDH was originally annotated as an NAD-dependent isocitrate dehydrogenase, but biochemical assays revealed its activity primarily with NADP+ as a cofactor. KB1_0553, also denoted as DmpPMI, was originally annotated as a hypothetical protein/sugar isomerase domain protein. We previously proposed that it was a bifunctional phosphoglucose isomerase/phosphomannose isomerase, but only phosphomannose isomerase activity was identified and confirmed experimentally. Further bioinformatics analyses of these two protein sequences suggest their affiliation to potentially novel enzyme families within their respective larger enzyme super families.

Introduction

As one of the smallest free-living organisms, Dehalococcoides mccartyi are important for their ability to detoxify ubiquitous and stable groundwater pollutants such as chlorinated ethenes and benzenes into benign or less toxic compounds (Maymó-Gatell et al., 1997; Adrian et al., 2000; 2007a; He et al., 2003; Löffler et al., 2012). Only these strictly anaerobic niche specialists are capable of harnessing energy for growth from the complete detoxification of known human carcinogens, trichloroethylene (TCE) and vinyl chloride (VC) (EPA, 2011, Guha et al., 2012) to benign ethene. This energy-conserving metabolic process, termed reductive dechlorination, is catalysed by reductive dehalogenases, the respiratory enzyme system of D. mccartyi. Reductive dechlorination, also known as organohalide respiration, is useful not only for the bioremediation of toxic chlorinated solvents, but also a key component to comprehend the recycling of naturally occurring organohalides (Gribble, 2010; 2012). As such, the fundamental understanding of D. mccartyi metabolism, including the genes and enzymes involved in metabolic processes, has wide-ranging implications for the remediation of toxic compounds and global element cycling (Gribble, 2012). So far, numerous systems-level studies on D. mccartyi metabolism, including the construction of a pan-genome-scale metabolic model (Ahsanul Islam et al., 2010), and various transcriptomic and proteomic analyses (Morris et al., 2006; 2007; Johnson et al., 2008; 2009; Lee et al., 2012), have begun to shed light on key metabolic processes and associated genes.

Although D. mccartyi metabolism is well studied, the activities of only a handful of metabolic genes have been experimentally validated. These include an Re-citrate synthase (Marco-Urrea et al., 2011) involved in the tricarboxylic acid (TCA)-cycle, a bifunctional mannosylglycerate (MG) synthase/phosphatase with a potential role in osmotic stress adaptation (Empadinhas et al., 2004), 12 reductive dehalogenases involved in res-
piration and energy conservation (Magnuson et al., 1998; 2000; Krajmalkin-Brown et al., 2004; Müller et al., 2004; Adrian et al., 2007b; Tang et al., 2013; Wang et al., 2014), and two hydrogenases (Jayachandran et al., 2004; Nijenhuis and Zinder, 2005) potentially involved in the electron transport chain of *D. mccartyi*. Genome sequences of several strains of these bacteria (Kube et al., 2005; Seshadri et al., 2005) revealed the presence of ~50% hypothetical proteins and proteins with unknown or non-specific functions in their genomes. The primary gene annotations, including the annotations for more than 80 metabolic genes, were reviewed, and in some instances corrected during the construction and manual curation of the genome-wide *D. mccartyi* metabolic model (Ahsanul Islam et al., 2010). Also, a recent systems-level study (Ahsanul Islam et al., 2014) on *D. mccartyi* transcriptions provided additional confidence to some of the proposed gene reannotations used in the metabolic model and helped predict putative functions for five hypothetical proteins. One such hypothetical protein was proposed to be a putative bifunctional phosphoglucose isomerase (PGI; EC 5.3.1.8)/ phosphomannose isomerase (PMI; EC 5.3.1.9) from the gene-expression analysis (Ahsanul Islam et al., 2014) and metabolic modelling studies on *D. mccartyi* (Ahsanul Islam et al., 2010). Another metabolic gene that was primarily annotated as a putative NADP+-isocitrate dehydrogenase (IDH) (EC 1.1.1.41) was reannotated as a putative NADP+-dependent IDH (EC 1.1.1.42) during the modelling study. However, these proposed reannotations were not supported by any experimental data.

In this study, we report the heterologous expression and biochemical characterization of the aforementioned putative IDH (KB1_0495) and PGI/PMI (KB1_0553) from *D. mccartyi* in KB-1 (Hug, 2012; Ahsanul Islam et al., 2014). KB-1 is a mixed enrichment culture containing several strains of *D. mccartyi* (Duhamel and Edwards, 2006; Hug et al., 2012). These two genes were selected because they produced soluble proteins when expressed in *Escherichia coli*. The orthologs of these genes in pure strains of *D. mccartyi* (DETO450, cbdbA408, DehaBAV_0427, DhcVS_392, DehalGT_0391, btf_415, dcmb_461, GY50_0375 and DET0509, cbdbA472, DehaBAV1_0485, DhcVS_0450, DehalGT_0448, btf_472, dcmb_518, GY50_0435) are 98–100% similar at the amino acid level (Markowitz et al., 2012). Although IDH is an important TCA-cycle enzyme catalysing the formation of 2-oxoglutarate and CO2 with NAD+ or NADP+ as a cofactor (Nelson and Cox, 2006; Madigan et al., 2010; Kanehisa et al., 2011), the physiological role of a bifunctional PGI/PMI in *D. mccartyi* is unclear. PGI, in general, plays a central role in sugar metabolism via glycolysis and gluconeogenesis in all life forms (Hansen et al., 2004a,b; Nelson and Cox, 2006), whereas PMI helps to produce precursors for cell wall components, glycoproteins, glycolipids and storage polysaccharides (Hansen et al., 2004a; Quevillon et al., 2005; Rajesh et al., 2012). However, glycolysis is inactive in *D. mccartyi* (Kube et al., 2005; Seshadri et al., 2005; Ahsanul Islam et al., 2010), and cells of these bacteria also lack a typical bacterial cell wall (Löffler et al., 2012). The genes (KB1_0495 and KB1_0553) were, thus, heterologously expressed in *E. coli*, and the purified proteins were experimentally tested for activities with enzymatic assays to confirm or refute the proposed annotations. Further bioinformatics analyses of their sequences suggested their affiliation to potentially new enzyme families within their respective larger enzyme super families.

### Results and discussion

**Biochemical activities of KB1_0495 and KB1_0553**

The heterologously expressed and purified protein from KB1_0495 was tested for the IDH activity using a standard assay (Experimental procedures) to measure the conversion of D-isocitric acid to 2-oxoglutarate and CO2 using NADP+ or NAD+ as a cofactor. The pH range and activity of the enzyme (Table 1) were measured using D-isocitric acid as substrate. The enzyme showed IDH activity using both NADP+ and NAD+ as the cofactor, but the activity with NAD+ was 65 times lower than with NADP+ (Table 1). This finding confirmed the annotation of KB1_0495 as an NADP+-dependent isocitrate dehydrogenase (*DmlIDH*) proposed by the previous metabolic modelling study (Ahsanul Islam et al., 2010). The kinetic parameters for *DmlIDH* were also estimated using both NADP+ and NAD+ as cofactors (Table 1), and the values obtained for isocitrate and both cofactors are quite

### Table 1. Kinetic Parameters of *DmlIDH* and *DmPmi* from *D. mccartyi*†

<table>
<thead>
<tr>
<th>Enzyme tested (varying substrate)</th>
<th>V_{max} (µmoles·min^{-1}·mg^{-1})</th>
<th>K_{m} (mM)</th>
<th>k_{cat} (s^{-1})</th>
<th>k_{cat}/K_{m} (mM^{-1}·s^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>DmlIDH</em> (isocitrate)</td>
<td>21.3 ± 2</td>
<td>0.11 ± 0.01</td>
<td>15 ± 1</td>
<td>139.6 ± 1</td>
</tr>
<tr>
<td><em>DmlIDH</em> (NADP+)</td>
<td>20.8 ± 2</td>
<td>0.027 ± 0.02</td>
<td>14.6 ± 1</td>
<td>540 ± 1</td>
</tr>
<tr>
<td><em>DmlIDH</em> (NAD+)</td>
<td>0.32 ± 0.05</td>
<td>0.12 ± 0.04</td>
<td>0.22 ± 0.03</td>
<td>2 ± 1</td>
</tr>
<tr>
<td><em>DmPmi</em> (MSP)</td>
<td>1.19 ± 0.2</td>
<td>0.09 ± 0.09</td>
<td>0.84 ± 0.1</td>
<td>0.96 ± 0.3</td>
</tr>
</tbody>
</table>

a. Reaction conditions were as described in Experimental procedures. b. Kinetic parameters were determined using NADP+ as cofactor (0.3 mM).
The higher activity and efficiency of DmIDH with NADP$^+$ are probably linked to its physiological roles in D. mccartyi. The incomplete TCA-cycle of these bacteria is mainly used for anabolism (Tang et al., 2009; Marco-Urrea et al., 2011). Hence, DmIDH plays a crucial role in D. mccartyi metabolism by producing the critical biosynthetic precursor 2-oxoglutarate and recycling essential cellular redox currency, NADPH. In fact, most bacteria possess NADP$^+$-dependent IDHs (Chen and Gadal, 1990; Muro-Pastor and Florencio, 1994), and microbes have an incomplete TCA-cycle use IDH to generate precursors (2-oxoglutarate) and reducing power (NADPH) for anaerobic biosynthetic pathways (Dean and Golding, 1997; Steen et al., 1998). However, NAD$^+$ may also work as a cofactor for DmIDH in vivo because the cellular concentration of NAD$^+$ is ~9 times higher than that of NADP$^+$ in a bacterial cell (Wimpenny and Firth, 1972; Andersen and von Meyenburg, 1977).

The second purified enzyme (KB1_0553) was tested for both PGI/PMI activities using four standard assays (Experimental procedures). Although the PGI activity (Glucose-6-phosphate, G6P ↔ Fructose-6-phosphate, F6P) was tested in both directions with two assays, the enzyme showed no activity with either G6P or F6P as substrates (data not shown). The presence of PMI activity (Mannose-6-phosphate, M6P ↔ Fructose-6-phosphate, F6P) was confirmed in the direction of F6P formation using two assays (Experimental procedures). The pH range and kinetic parameters (Table 1) were also estimated using M6P as substrate. Thus, only the PMI activity was detected and confirmed for KB1_0553 (DmPMI). This activity is typically associated with peptidoglycan and teichoic acid biosynthesis (Kanehisa et al., 2011). Therefore, the physiological role of DmPMI in D. mccartyi is unclear because they lack a typical bacterial cell wall in favour of an archaean S-layer like protein and cell membrane (Löffler et al., 2012). DmPMI may be involved in cell membrane biogenesis in these bacteria as suggested by the previous transcriptomic study (Ahsanul Islam et al., 2014). It can also be involved in the mannosylglycerate (MG) biosynthesis pathway (Fig. 1) because a bifunctional mannosyl-3-phosphoglycerate synthase (EC 2.4.1.217)/phosphatase (EC 3.1.3.70) enzyme encoded from D. mccartyi strain 195 (formerly D. ethenogenes strain 195) (Empadinhas et al., 2004) showed no activity with either G6P or F6P as substrates (data not shown). The presence of PMI activity from DmPMI was confirmed previously from D. mccartyi strain 195 (formerly D. ethenogenes strain 195) (Empadinhas et al., 2004). MG is an unusual compatible solute that helps thermophilic or hyperthermophilic bacteria and archaea to adjust osmotic stress and heat (Santos and da Costa, 2002; Empadinhas and da Costa, 2011). Because D. mccartyi strains are halotolerant or slightly halophilic (Empadinhas et al., 2004), MG was hypothesized to be involved in cellular osmotic stress adaptation in these bacteria (Hendrickson et al., 2002; Empadinhas et al., 2004). However, this suggested physiological role is a non-essential or specialized function rather than a major biological function, which probably explains why DmPMI has relatively lower catalytic activity and efficiency (Table 1) than similar enzymes in BREnda (Chang et al., 2009).

Sequence homology and phylogenetic analyses of DmIDH and DmPMI sequences

The sequence homology analysis of DmIDH and DmPMI protein sequences (Fig. 2) revealed the remarkably conserved nature of DmIDH across the domains of life (Fig. 2A) as compared with DmPMI (Fig. 2B). Being a...
A TCA-cycle enzyme, DmIDH was found to share >40% amino acid sequence identity with other homologous IDHs from eukarya, archaea, and bacteria (Fig. 2A), whereas DmPMI showed <30% amino acid sequence identity with the majority of its homologs (Fig. 2B). This difference in sequence conservation was also observed from the higher bootstrap values in the DmIDH maximum likelihood (ML) tree (Fig. 3) as compared with those in the ML tree of DmPMI (Fig. 4). The DmIDH tree further showed its separate clustering (Fig. 3) from the previously described (Steen et al., 1997) subfamilies I and II of bioc hemically characterized IDHs, as well as from subfamily III of NAD+-dependent eukaryotic IDHs (Steen et al., 1997). Thus, DmIDH belongs to a potentially new subfamily that also includes bacterial IDHs from Planctomycetes, Firmicutes and Cyanobacteria (Fig. 3). Interestingly, no homolog of DmPMI was identified within the three types of PMIs characterized and described previously (Schmidt et al., 1992; Proudfoot et al., 1994). Among the biochemically characterized enzymes, the closest homologs of DmPMI were identified to be archaeal bifunctional PGI/PMIs (Fig. 4), which represent a novel enzyme family.
Fig. 3. Phylogenetic analysis of DmIDH protein sequence. Maximum likelihood (ML) tree for DmtIDH and its homologous protein sequences was constructed by PHYML (Guindon et al., 2010) plugin in GENEIOUS (Biomatters, 2011). Protein sequences were mined from UniProt (Apweiler et al., 2012) and aligned with MUSCLE (Edgar, 2004) plugin in GENEIOUS. Then, the ML tree was constructed under WAG (Whelan and Goldman, 2001) model of amino acid substitution with 100 bootstrap resampling trees were conducted. Bootstrap values are shown as branch labels, and the biochemically characterized genes are marked by asterisks. Organism names are coloured according to different kingdoms (Orange = Archaea, Green = Bacteria, and Purple = Eukarya).

© 2015 The Authors. Microbial Biotechnology published by John Wiley & Sons Ltd and Society for Applied Microbiology
Fig. 4. Phylogenetic analysis of DmPMI protein sequence. Maximum likelihood (ML) tree for DmPMI and its homologous protein sequences was constructed by PhyML (Guindon et al., 2010) plugin in GENEIOUS (Biomatters, 2011). Protein sequences were mined from UniProt (Apweiler et al., 2012) and aligned with MUSCLE (Edgar, 2004) plugin in GENEIOUS. Then, the ML tree was constructed under WAG (Whelan and Goldman, 2001) model of amino acid substitution with 100 bootstrap resampling trees were conducted. Bootstrap values are shown as branch labels, and the biochemically characterized genes are marked by asterisks. Organism names are coloured according to different kingdoms (Orange = Archaea and Green = Bacteria).
within the PGI superfamily (Hansen et al., 2004a,b). Thus, DmPMI likely constitutes a novel class of bacterial PMIs, including the homologs from Actinobacteria, Firmicutes and Bacteroidetes (Fig. 4).

**Structure-based analysis of DmIDH and DmPMI sequences**

To verify the aforementioned differences of DmIDH and DmPMI from the previously characterized similar enzymes, we conducted further bioinformatics analyses of their sequences. Multiple sequence alignment (MSA) of DmIDH sequence and its biochemically characterized homologs from SWISSPROT identified 66 completely conserved residues (Fig. 5). These included all residues (indicated by blue and red boxes in Fig. 5) involved in substrate and coenzyme binding in Archaeoglobus fulgidus and E. coli IDHs (Hurley et al., 1991; Steen et al., 1997; Stokke et al., 2007), suggesting a similar reaction mechanism for DmIDH. Although the signature motif of both isocitrate and isopropylmalate dehydrogenases (IDH and IPMDH) was identified in DmIDH (indicated by a green box in Fig. 5), only the IDH activity was tested and confirmed because of its location in an operon containing other putative TCA-cycle genes (Fig. S2). Also, the D. mccartyi genomes harbor a putative IPMDH gene (KB1_0839, cddbA804, DET0826, DhcVS_730, DehBAV1_0745 and DehalGT_0706) located in the L-leucine biosynthesis operon (Fig. S4). Comparison of the predicted secondary structure of DmIDH (Fig. 5) with the crystal structures of A. fulgidus and E. coli IDHs (Hurley et al., 1991; Stokke et al., 2007) showed that most helix and strand regions were conserved among them in spite of the presence of some subtle structural differences. For instance, DmIDH has 12 α-helices and 11 β-strands (Fig. 5), whereas A. fulgidus has 18 α-helices, 16 β-strands (Stokke et al., 2007) and E. coli has 13 α-helices, 12 β-strands (Hurley et al., 1991).

MSA of DmPMI sequence with its biochemically characterized or manually reviewed homologs identified only 20 completely conserved residues (Fig. 6). In addition to two signature motifs (marked by blue boxes in Fig. 6) of archaeal PGI/PMIs (Hansen et al., 2004b), a SIS (sugar isomerase) domain (marked by a green box in Fig. 6) was identified in DmPMI. All residues proposed to be important for substrate binding and catalysis in archaeal PGI/PMIs (Hansen et al., 2004a; Swan et al., 2004) are also found to be conserved in DmPMI (marked by red boxes in Fig. 6), except a few notable ones: Thr60, Ser103, Arg152, Ser154, Pro341, and Ile342. The most important change was detected at residue position 154, where a Serine (S) substituted an Arginine (R) found in other archaeal PGI/PMIs (marked by a yellow S in Fig. 6). This residue is pivotal for the catalytic activities of archaeal PGI/PMIs because it stabilizes the formation of crucial enediol intermediates during their PGI activity (Seeohlzer, 1993; Hansen et al., 2004b; Swan et al., 2004). Hence, these changes in the DmPMI sequence are likely responsible for its inability to function as PGI. A similar notion was also obtained from the comparison of the predicted secondary structure of DmPMI (12 α-helices and 9 β-strands) (Fig. 6) with the crystal structure of PaPGI/PMI from P. aerophilum (Swan et al., 2004). However, the mechanism for PMI activity of DmPMI is likely similar to PaPGI/PMI because the presence of Thr291 in the PaPGI/PMI sequence is key to its PMI activity (Swan et al., 2004), and the equivalent residue in DmPMI is an Isoleucine (Ile342) (Fig. 6).

In summary, the two characterized enzymes, DmIDH and DmPMI, from D. mccartyi strains in KB-1 likely have reaction mechanisms similar to the previously characterized enzymes although they appear to belong to novel IDH and PMI enzyme families. DmIDH showed a higher catalytic activity with NADP+ as a cofactor than with NAD+, implicating its involvement in anabolic biosynthetic pathways in D. mccartyi, whereas the lower activity of DmPMI suggests that the physiological substrate is possibly a different sugar than M6P, or that it is possibly involved in non-essential physiological roles such as osmotic stress adjustment in these bacteria.

**Conclusions**

Automated and non-curated primary annotations of genes are necessary but highly problematic if they are wrong. Because initial annotations are often non-specific or incorrect, correcting these annotations is very important and challenging. In this study, corrected or reviewed initial annotations that were proposed from previously published modelling and transcriptomic studies were experimentally tested and shown to be essentially correct, illustrating the utility of metabolic modelling and bioinformatics approaches as important tools of hypothesis generation for gene annotations. Each new biochemical confirmation of gene functions adds to the confidence of propagated annotations, contributing to a stronger database for sequence analysis in general. In the specific case of the unusual organohalide respiring anaerobes D. mccartyi, fundamental understanding of the physiology and biochemistry of these organisms will contribute to better deployment of bioremediation approaches for environmental stewardship.

**Experimental procedures**

**Bacterial culture, reagents, and chemicals**

Genomic DNA (gDNA) was collected from KB-1, a D. mccartyi-containing anaerobic mixed culture growing...
Fig. 5. Structure-based multiple sequence alignment (MSA) of DmIDH. MSA of DmIDH protein sequence and its homologous, biochemically characterized sequences are shown. Protein sequences were mined from the SWISSPROT curated database (Boeckmann et al., 2003), and MSA was performed by CLUSTALX (Larkin et al., 2007). Completely conserved residues are marked by asterisks, whereas residues reported to be important for substrate and coenzyme binding (Hurley et al., 1991; Stokke et al., 2007) are marked by blue and red boxes respectively. Green box indicates signature motif for isocitrate/isopropylmalate dehydrogenases as identified by SCANPROSITE (de Castro et al., 2006). Protein secondary structure was predicted by PREDICTPROTEIN (Rost and Sander, 1994), and indicated by bars (α-helices) and arrows (β-strands). Protein accession numbers (UniProt) are: A. fulgidus (O29610), C. noboribetus (P96318), S. aureus (P99167), B. subtilis (P39126), E. coli (P08200), C. maris (P41560), B. taurus (P41563) and H. sapiens (P50213).
Fig. 6. Structure-based multiple sequence alignment (MSA) of DmPMI.
MSA of DmPMI protein sequence and its homologous, manually reviewed/biochemically characterized sequences are shown. Protein sequences were mined from the SWISSPROT curated database (Boeckmann et al., 2003), and MSA was performed by CLUSTALX (Larkin et al., 2007). Completely conserved residues are marked by asterisks, whereas residues reported to be important for substrate binding and catalysis (Hansen et al., 2004a; Swan et al., 2004) are marked by red boxes. Blue boxes indicate two signature motifs for bifunctional PGI/PMI protein family (Hansen et al., 2004b), and green box indicates Pfam (Punta et al., 2012) SIS (sugar isomerase) domain identified by SCANPROSITE (de Castro et al., 2006). Protein secondary structure was predicted by PREDICTPROTEIN (Rost and Sander, 1994), and indicated by bars (α-helices) and arrows (β-strands). Protein accession numbers (UniProt) are: C. bescii (Q44407), T. volcanium (Q978F3), T. acidophilum (Q9HIC2), A. aeolicus (O66954), S. tokodaij (Q96YC2), S. acidocaldarius (Q4JCA7), S. solfataricus (Q82945), P. aerophilum (Q82945), and A. pernix (Q9YE01).
on TCE and methanol following the procedure described previously (Duhamel and Edwards, 2006). The PCR primers for amplifying D. mccartyi gDNA were synthesized by Integrated DNA Technologies (Coralville, IA, USA). Luria broth and terrific broth powder were purchased from EMD Chemicals (Gibbstown, NJ, USA), and the Bradford assay reagent from Bio-Rad (Hercules, CA, USA). Lysozyme, protease K, agarose, glycerol, ampicillin, kanamycin, SDS and IPTG were obtained from BioShop (Burlington, ON, Canada), and all other chemicals were purchased from Sigma-Aldrich (St. Louis, MO, USA) with greater than 98% in purity. Nickel-Nitrilotriacetic acid (NiNTA) resin and the QIAquick PCR Purification Kit were purchased from Qiagen (Mississauga, ON, Canada), whereas the In-Fusion PCR Cloning Kit was purchased from Clontech (Palo Alto, CA, USA). The commercially available kits were used according to the manufacturers’ instructions.

Gene cloning, protein overexpression, and purification

The selected genes (KB1_0495 and KB1_0553) were PCR-amplified using D. mccartyi gDNA and the PCR primers containing the restriction sites for BamHI and NdeI, and were cloned into the modified pET-15b vector (Novagen, Madison, WI, USA) containing a 5’ N-terminal hexahistidine tag (6xHis-tag) and an ampicillin resistance gene as described previously (Zhang et al., 2001). In the modified vector, the tobacco etch virus protease cleavage site replaced the thrombin cleavage site, and a double stop codon was introduced downstream from the BamHI site (Zhang et al., 2001). These vectors were subsequently transformed into E. coli strain BL21 (DE3) (Stratagene, La Jolla, CA, USA) for overexpression of the targeted fused genes. The cells were grown aerobically in 1 L flasks containing tryptone-phosphate medium at 37°C and ~220 rpm until the OD<sub>600</sub> reached around 1.0 (approximately in 3 h). Expression of the cloned genes was induced by adding 100 mg IPTG, and cells were harvested the following day by centrifugation (Zhang et al., 2001). The overexpressed, fused 6xHis-tagged proteins were purified to more than 95% homogeneity (Fig. S1) using metal-chelate affinity chromatography on nickel affinity resin and gel filtration on a Superdex 200 26/60 column (Amersham Biosciences, Piscataway, NJ, USA) as described before (Zhang et al., 2001; Proudfoot et al., 2008).

Isocitrate dehydrogenase (IDH) assay

IDH activity in KB1_0495 (DmiDH) was confirmed by an enzymatic assay (Steen et al., 1997; 1998) for the reaction described by the following equation:

\[
\text{D - isocitric acid + NAD(P)^+} \leftrightarrow 2 - \text{oxoglutarate} + \text{CO}_2 + \text{NAD(P)H}
\]

The standard 1 ml assay contained 50 mM tris-hydrochloride (Tris-HCl) buffer (pH 7.5), 0.3 mM NADP<sup>+</sup>, 1 mM D-isocitric acid and 10 mM MgCl<sub>2</sub>. The reaction was started by adding 1 µg of purified protein into the reaction mixture, and the product formation was inferred from the measurement of NAD(P)H formation as indicated by the increase in absorbance at 340 nm and 30°C with a spectrophotometer. The IDH activity in KB1_0495 was also measured using NAD<sup>+</sup> as a cofactor.

Phosphoglucone isomerase (PGI) assay

PGI activity in KB1_0553 (DmPMI) was tested using two standard assays (Hansen et al., 2004b) for the reaction described by the following equations:

Assay 1:

\[
\text{F6P} \overset{\text{PGI}}{\rightarrow} \text{G6P}
\]

where F6P refers to fructose-6-phosphate, G6P refers to glucose-6-phosphate, G6P15L refers to 6-phospho-D-glucono-1,5-lactone, and G6PDH is glucose-6-phosphate dehydrogenase (EC. 1.1.1.49).

Assay 2:

\[
\text{G6P} \overset{\text{PGI}}{\rightarrow} \text{F6P}
\]

where F6P refers to fructose-6-phosphate, G6P refers to glucose-6-phosphate, G6P15L refers to 6-phospho-D-glucono-1,5-lactone, and G6PDH is glucose-6-phosphate dehydrogenase (EC. 1.1.1.49).

Phosphomannose isomerase (PMI) assay

PMI activity in KB1_0553 (DmPMI) was tested with two standard assays (Hansen et al., 2004a) for the reactions described by the following equations:

© 2015 The Authors. Microbial Biotechnology published by John Wiley & Sons Ltd and Society for Applied Microbiology
Assay 1:

\[
\begin{align*}
\text{PMI} & : 
M6P \leftrightarrow F6P \\
\text{M1PDH} & : 
F6P + NADH \leftrightarrow M1P + NAD^+ 
\end{align*}
\]

where M6P refers to mannose-6-phosphate.

Assay 2:

\[
\begin{align*}
\text{PMI} & : 
M6P \rightarrow F6P \\
\text{PGI} & : 
F6P \rightarrow G6P \\
\text{G6PDH} & : 
G6P + NADP^+ \leftrightarrow 6PG15L + NADPH 
\end{align*}
\]

The reaction mixture for assay 1 contained 100 mM Tris-HCl (pH 7.5), 0.5 mM NADH, 10 mM M6P, and 10 μl of M1PDH. The M1PDH was purified from E. coli (Novotny et al., 1984) with >95% purity (see Fig. S1). In assay 2, the reaction mixture contained 100 mM Tris-HCl (pH 7.5), 0.5 mM NADP+, 10 mM M6P and 1.1 U of G6PDH (Sigma-Aldrich, St. Louis, MO) and 1 U of PGI (Sigma-Aldrich, St. Louis, MO). The reactions were started by adding 1 μg of purified protein into the reaction mixture, and the product, F6P formation, was measured by coupling it to the oxidation of NADH in assay 1 and to the formation of NADPH in assay 2. In both instances, the product formation was inferred from the measurement of absorbance with a spectrophotometer at 340 nm and 30°C.

**Determination of pH range and kinetic parameters**

The pH range of purified KB1_0495 (DmIDH) and KB1_0553 (DmPMI) (Table 1) was determined by measuring their activities at pH values from 6 to 8.5 at 30°C using D-isocitric acid and M6P as substrates, respectively. The enzymatic activity of both purified proteins was also measured at selected pHs with varying concentrations of substrates, D-isocitric acid, M6P and cofactors, NADPH and NAD+. These data were, then, fitted to the Michaelis–Menten enzyme kinetics model with the non-linear regression analysis to estimate the maximum enzyme velocity (Vmax) and the Michaelis constant (Km) using GraphPad Prism v 5.0 (GraphPad Software, La Jolla, CA, USA). Then, the turnover number (kcat) and the efficiency (kcat/Km) of both proteins were calculated from Vmax and Km data.

**Bioinformatics analyses of DmIDH and DmPMI sequences**

Homologous protein sequences of DmIDH and DmPMI in eukarya, archaea, and bacteria were identified by blastp in BLAST (Altschul et al., 1997) from the UniProt database (Apweiler et al., 2012). The sequence homology networks of both DmIDH and DmPMI were constructed and visualized by CYTOSCAPE (Smoot et al., 2011) – an open source bioinformatics software platform. The phylogenetic analysis was conducted by the PhyML (Guindon et al., 2010) plugin in the GENEIOUS software platform (Biomatters, 2011) to construct the ML trees for both DmIDH and DmPMI sequences and their homologous protein sequences. First, the homologous sequences were mined from UniProt (Apweiler et al., 2012) with blastp in BLAST and aligned with the MUSCLE (Edgar, 2004) plugin in GENEIOUS. Then, the ML tree was constructed using the WAG (Whelan and Goldman, 2001) model of amino acid substitution with 100 bootstrap resampling trees were conducted. The multiple sequence alignment of DmIDH and DmPMI protein sequences was performed by CLUSTALX (Larkin et al., 2007), and the biochemically characterized homologous protein sequences were mined from the SWISSPROT curated database (Boeckmann et al., 2003). The protein secondary structure of DmIDH and DmPMI was predicted by the PREDICTPROTEIN (Rost and Sander, 1994) software.

**Conflict of interest**

None declared.

**References**


**Supporting information**

Additional Supporting Information may be found in the online version of this article at the publisher’s web-site:

**Fig. S1.** SDS-PAGE of (A) DmIDH, (B) DmPMI, and (C) M1PDH.

**Fig. S2.** Orthologous gene neighborhood analysis of isocitrate dehydrogenase (IDH) from *D. mccartyi*.

**Fig. S3.** Orthologous gene neighborhood analysis of hypothetical protein/SIS domain protein from *D. mccartyi*.

**Fig. S4.** Orthologous gene neighborhood analysis of 3-isopropylmalate dehydrogenase (IPMDH) from *D. mccartyi*.