A Gene from the Mesophilic Bacterium *Dehalococcoides ethenogenes* Encodes a Novel Mannosylglycerate Synthase

Nuno Empadinhas,1 Luciana Albuquerque,1 Joana Costa,1 Stephen H. Zinder,2 Manuel A. S. Santos,3 Helena Santos,4 and Milton S. da Costa1*

Department of Bioquímica, Centro de Neurociências e Biologia Celular, and Departamento de Zootecnia, Universidade de Coimbra, 3004-517 Coimbra,1 Centro de Biologia Celular, Departamento de Biologia, Universidade de Aveiro, 3810-193 Aveiro,3 and Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, 2780-156 Oeiras,4 Portugal, and Department of Microbiology, Cornell University, Ithaca, New York 148532

Received 27 January 2004/Accepted 22 March 2004

Mannosylglycerate (MG) is a common compatible solute found in thermophilic and hyperthermophilic prokaryotes. In this study we characterized a mesophilic and bifunctional mannosylglycerate synthase (MGSD) encoded in the genome of the bacterium *Dehalococcoides ethenogenes*. *mgsD* encodes two domains with extensive homology to mannosyl-3-phosphoglycerate synthase (MPGS, EC 2.4.1.217) and to mannol-3-phosphoglycerate phosphatase (MPGP, EC 3.1.3.70), which catalyze the consecutive synthesis and dephosphorylation of mannosyl-3-phosphoglycerate to yield MG in *Pyrococcus horikoshii*, *Thermus thermophilus*, and *Rhodothermus marinus*. The bifunctional MGSD was overproduced in *Escherichia coli*, and we confirmed the combined MPGS and MPGP activities of the recombinant enzyme. The optimum activity of the enzyme was at 50°C. To examine the properties of each catalytic domain of MGSD, we expressed them separately in *E. coli*. The monofunctional MPGS was unstable, while the MPGP was stable and was characterized. *Dehalococcoides ethenogenes* cannot be grown sufficiently to identify intracellular compatible solutes, and *E. coli* harboring MGSD did not accumulate MG. However, *Saccharomyces cerevisiae* expressing *mgsD* accumulated MG, confirming that this gene product can synthesize this compatible solute and arguing for a role in osmotic adjustment in the natural host. We did not detect MGSD activity in cell extracts of *S. cerevisiae*. Here we describe the first gene and enzyme for the synthesis of MG from a mesophilic microorganism and discuss the possible evolution of this bifunctional MGSD by lateral gene transfer from thermophilic and hyperthermophilic organisms.

Several compatible solutes such as ectoine, glycine betaine, and trehalose are widespread in mesophilic bacteria, where they accumulate in response to stress imposed by salt or heat. Thermophilic and hyperthermophilic bacteria and archaea generally accumulate unusual compatible solutes, such as di-*myo*-inositol-phosphate, di-*mannosyl*-di-*myo*-inositol-phosphate, diglycerol-phosphate, and mannosylglycerate, that have not yet been identified in mesophilic prokaryotes (40). Mannosylglycerate (MG) has been encountered in several thermophilic bacteria, such as *Thermus thermophilus*, *Rhodothermus marinus*, and *Rubrobacter xylanophilus*, and in hyperthermophilic archaea such as *Arbopyrum pernix*, *Pyrococcus* spp., and *Thermococcus* spp. and some strains of the species *Archaeoglobus* (40). Despite its scattered distribution, MG has been primarily associated with prokaryotes that grow at high temperatures (13, 40). While the role of MG during osmotic stress is controversial in some of these organisms (42), in vitro evidence indicates a role for this compatible solute on the protection of proteins against thermal denaturation (5, 14, 36). However, MG was also encountered in marine red algae and is not therefore restricted to thermophilic and hyperthermophilic prokaryotes (25).

Identification of the genes and biosynthetic pathways involved in the synthesis of MG is essential to elucidate the physiological role of this compatible solute in the osmotic stress responses of thermophiles and hyperthermophiles and to understand its puzzling distribution among prokaryotes. Two independent biosynthetic pathways for the synthesis of MG have been elucidated; one involves the direct condensation of GDP-*mannose* and *d*-glycerate into MG by mannosylglycerate synthase (MGS). This pathway has only been encountered in *R. marinus* (29). The alternative pathway channels GDP-*mannose* and *d*-3-phosphoglycerate (3-PGA) through the intermediate mannosyl-3-phosphoglycerate (MPG) to MG by mannosyl-3-phosphoglycerate synthase (MPGS), encoded by *mpgs*, and by mannosyl-3-phosphoglycerate phosphatase (MPGP), encoded by *mpgp*. This pathway has been found in *Pyrococcus* spp., *T. thermophilus*, and, recently, in *R. marinus* as well. The genes of the two-step pathway are always contiguous in the thermophilic and hyperthermophilic archaea and bacteria that have been examined (6, 12, 13).

The bacterium tentatively named *Dehalococcoides ethenogenes* is an unusual microorganism whose energy metabolism relies on the reductive dechlorination of chlorinated ethene solvents, which it uses as electron acceptors (30). This strictly anaerobic organism, with an optimum growth temperature near 35°C, represents a deep line of descent in the bacterial phylum *Chloroflexi* (16). *D. ethenogenes* and a closely related isolate designated CBDB1 were recovered from sewage sludge.
(1, 30). Preliminary experiments have shown that D. ethenogenes is capable of growth at NaCl concentrations as high as 0.5 M, near the salinity of seawater (L. Hsu and S. Zinder, unpublished data). Moreover, environmental 16S rRNA gene sequences of organisms that are closely related to D. ethenogenes have also been detected in marine environments, indicating that they are halotolerant or slightly halophilic and may therefore require compatible solutes for osmotic adjustment (22, 23, 37) (GenBank accession nos. AJ431246 and AJ431247). However, the extremely poor growth of D. ethenogenes does not allow the determination of low-molecular-weight intracellular solutes.

The ability to completely dechlorinate polychlorinated compounds prompted the sequencing of the D. ethenogenes genome. A BLAST search of genes involved in the synthesis of MG at the TIGR unfinished-genomes database (http://tigrblast.tigr.org/ufmg/) showed that D. ethenogenes has a gene sequence potentially coding for two functional domains, one with high homology to MPGS (EC 2.4.1.217) and the other to MPGP (EC 3.1.3.70).

The identification of a putative bifunctional gene for the synthesis of MG in D. ethenogenes raises several important questions about the function of the enzyme, the synthesis of MG by the natural host, and the role of this compatible solute in D. ethenogenes. To answer some of these questions, we expressed the gene in Escherichia coli and ultimately in Saccharomyces cerevisiae. Our results show that this gene, designated mgsD, encodes a bifunctional mannosylglycerate synthase (MGSD) that synthesizes MG at low temperatures. Furthermore, expression of the gene in S. cerevisiae led to in vivo accumulation of MG, strongly suggesting that the function of the gene product is to synthesize MG in D. ethenogenes. This is the first report of a bifunctional gene and enzyme for the synthesis of MG in a mesophilic bacterium and is a novelty of evolutionary relevance.

### MATERIALS AND METHODS

#### Strains, plasmids, and culture conditions.

\textit{E. coli} strains XL1-Blue and BL21-Rosetta (Novagen) were used as hosts for cloning and gene expression. The BL21-Rosetta strain carries a plasmid with the TRNA genes for the codons rarely used by \textit{E. coli}. Expression vector pKK223-3 (Amersham Pharmacia Biotech, Uppsala, Sweden) was used for gene expression in \textit{E. coli}. The medium for growth was Luria-Bertani (LB) at 37°C, unless otherwise stated. \textit{Escherichia coli} strains were transformed by the calcium chloride method (38) and selected on LB agar plates with the appropriate antibiotics. Ampicillin and chloramphenicol were added where necessary at concentrations of 100 and 40 µg/ml, respectively. Positive clones were identified by restriction analysis, and plasmids were purified with the High-Pure plasmid purification kit (Roche).

\textit{Saccharomyces cerevisiae} strain CEN.PK2 (MATa/MATb dua1-52/ura3-52 trp1-289/trp1-289 leu2-3,112/leu2-3,112 his3/ his3 matα/A α ura3-52/can1-100) and a yeast shuttle vector derived from pRS425 (ATCC 77106) obtained from LGC Promoclean (American Type Culture Collection, Barcelona, Spain; http://www.lgcpromoclean.com /atcc/) were used for gene expression. \textit{S. cerevisiae} cells were grown in YPD medium (10 g of yeast extract, 20 g of peptone, and 20 g of glucose per liter), transformed by the lithium acetate protocol (17), and recombinants selected on MM-LEU medium (6.7 g of yeast nitrogen base without amino acids [Difco] per liter, 20 g of glucose per liter, and a leucine drop-out mixture of amino acids). The cultures were grown in 1-liter baffled metal-capped Erlenmeyer flasks containing 250 ml of medium at 30°C in an orbital shaker at 160 rpm.

#### Isolation of DNA from \textit{D. ethenogenes} and \textit{PCR}.

DNA was extracted from \textit{D. ethenogenes} with phenol-chloroform-isooamyl alcohol and purified with a Bio 101 (Vista) GeneClean spin kit as previously described (15). All PCR amplifications were carried out with 2.5 U of \textit{Pfu} DNA polymerase (Stratagene) and the conditions described elsewhere (13).

#### Identification and analysis of mgsD and flanking genes in the \textit{D. ethenogenes} genome.

BLAST searches were carried out at the Institute for Genomic Research unfinished genomes database (http://tigrblast.tigr.org/ufmg/) with amino acid sequences from \textit{Pyrococcus horikoshii} MPGS and MPGP (PfMPGS, GenPept BAA30023; PfMPGP, BAA30022). Preliminary sequence data were obtained from TIGR at http://www.tigr.org. All open reading frames (ORFs) surrounding \textit{mgsD} were identified by BLAST with the translated sequences at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov /BLAST).

Cloning the bifunctional \textit{mgsD} and overexpression in \textit{E. coli}.

\textit{D. ethenogenes} genomic DNA was used as the template to amplify \textit{mgsD} with primers DET1 and DET2 (Table 1), constructed based on the nucleotide sequence obtained from TIGR. The PCR product was purified after digestion with EcoRI and HindIII and ligated into the corresponding sites of pMMB66H (Novagen) to yield pDET. The insert was sequence-matched on both strands (AGOWA, Berlin, Germany) for confirmation purposes.

\textit{E. coli} BL21-Rosetta was transformed with pDET, and transformants were selected on LB agar plates with appropriate antibiotics. Eight clones were separately grown in LB medium containing both antibiotics with continuous shaking (180 rpm). One clone carrying an empty pKK223-3 was grown under identical conditions. Growth was allowed to proceed to the mid-exponential phase (optical density at 600 nm [OD\textsubscript{600} = 0.8]) at 37°C, after which the temperature was lowered to 30°C, followed by induction with 0.5 mM isopropylthio-galactopyranoside (IPTG), and the clones were grown at this temperature for an additional 8 h. Aliquots of the cell suspensions were harvested by centrifugation at 10,000 x g for 30 s; the pellet was suspended in Laemmli buffer, boiled for 5 min, and loaded on a denaturing polyacrylamide gel (27). The remaining cultures were allowed to grow overnight. The clone showing greater expression band of the recombinant MGSD on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels was selected for enzyme production. LB medium (1 liter) containing both antibiotics was inoculated with 1 ml of the selected overnight culture, grown to the mid-exponential phase (OD\textsubscript{600} = 0.8), induced with

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5′-3′)</th>
<th>Direction</th>
<th>Target sequence</th>
<th>Restriction enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>DET1</td>
<td>GCCGAATTCTAGCGCATGCCATGAAGCGCTGCGTCCC</td>
<td>Forward</td>
<td>mgsD</td>
<td>EcoRI</td>
</tr>
<tr>
<td>DET2</td>
<td>GCCGAACCGTTTACACGTCCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>mgsD</td>
<td>HindIII</td>
</tr>
<tr>
<td>DETS</td>
<td>GCCGAACGGTTTACAGCCTCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>Dempgs</td>
<td>HindIII</td>
</tr>
<tr>
<td>DETP</td>
<td>CGGAAATCTGATATTAACTTACTTACTTACTTATTGTG</td>
<td>Reverse</td>
<td>Dempgs</td>
<td>EcoRI</td>
</tr>
<tr>
<td>DEF</td>
<td>GCCGAACCTATGCGCATTGAAGCGCTGCGTCCC</td>
<td>Forward</td>
<td>Dempgs</td>
<td>HindIII</td>
</tr>
<tr>
<td>DER</td>
<td>GCGTCGACGATCACAGCTCTCCGTCAGGCAGAAC</td>
<td>Reverse</td>
<td>mgsD</td>
<td>NruI</td>
</tr>
<tr>
<td>ENO2F</td>
<td>AGCTAGGGCCCAATCCTACTTGGCGTGTCG</td>
<td>Forward</td>
<td>ENO2 promoter</td>
<td>Apal</td>
</tr>
<tr>
<td>ENO2R</td>
<td>GATTCCAGCGTTTACAGCCTCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>ENO2 promoter</td>
<td>HindIII</td>
</tr>
<tr>
<td>DETS-L</td>
<td>GCCGAACGGTTTACAGCCTCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>Dempgs + linker</td>
<td>HindIII</td>
</tr>
<tr>
<td>DETS-450</td>
<td>GCCGAACGGTTTACAGCCTCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>Dempgs + linker + 33 aa</td>
<td>HindIII</td>
</tr>
<tr>
<td>DETS-549</td>
<td>GCCGAACGGTTTACAGCCTCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>Dempgs + linker + 132 aa</td>
<td>HindIII</td>
</tr>
</tbody>
</table>

\[a\] The restriction enzyme recognition sequences are underlined.
[\[b\]] aa, amino acids.

### Table 1. Primers

<table>
<thead>
<tr>
<th>Primer</th>
<th>Sequence (5′-3′)</th>
<th>Direction</th>
<th>Target sequence</th>
<th>Restriction enzyme</th>
</tr>
</thead>
<tbody>
<tr>
<td>DET1</td>
<td>GCCGAATTCTAGCGCATGCCATGAAGCGCTGCGTCCC</td>
<td>Forward</td>
<td>mgsD</td>
<td>EcoRI</td>
</tr>
<tr>
<td>DET2</td>
<td>GCCGAACCGTTTACACGTCCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>mgsD</td>
<td>HindIII</td>
</tr>
<tr>
<td>DETS</td>
<td>GCCGAACGGTTTACAGCCTCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>Dempgs</td>
<td>HindIII</td>
</tr>
<tr>
<td>DETP</td>
<td>CGGAAATCTGATATTAACTTACTTACTTACTTATTGTG</td>
<td>Reverse</td>
<td>Dempgs</td>
<td>EcoRI</td>
</tr>
<tr>
<td>DEF</td>
<td>GCCGAACCTATGCGCATTGAAGCGCTGCGTCCC</td>
<td>Forward</td>
<td>Dempgs</td>
<td>HindIII</td>
</tr>
<tr>
<td>DER</td>
<td>GCGTCGACGATCACAGCTCTCCGTCAGGCAGAAC</td>
<td>Reverse</td>
<td>mgsD</td>
<td>NruI</td>
</tr>
<tr>
<td>ENO2F</td>
<td>AGCTAGGGCCCAATCCTACTTGGCGTGTCG</td>
<td>Forward</td>
<td>ENO2 promoter</td>
<td>Apal</td>
</tr>
<tr>
<td>ENO2R</td>
<td>GATTCCAGCGTTTACAGCCTCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>ENO2 promoter</td>
<td>HindIII</td>
</tr>
<tr>
<td>DETS-L</td>
<td>GCCGAACGGTTTACAGCCTCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>Dempgs + linker</td>
<td>HindIII</td>
</tr>
<tr>
<td>DETS-450</td>
<td>GCCGAACGGTTTACAGCCTCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>Dempgs + linker + 33 aa</td>
<td>HindIII</td>
</tr>
<tr>
<td>DETS-549</td>
<td>GCCGAACGGTTTACAGCCTCGTCAGCGCAAGAAC</td>
<td>Reverse</td>
<td>Dempgs + linker + 132 aa</td>
<td>HindIII</td>
</tr>
</tbody>
</table>
IPTG, and grown for another 8 h, as described above. Gene expression was monitored by SDS-PAGE.

**Construction of monofunctional enzymes and expression in E. coli.** The 694-amino-acid sequence of the bifunctional enzyme was aligned with the MGPS and MPGP sequences from the *P. horikoshii* (PhMPGP) and *T. thermophila* (TtMPGP). Prime DET1 and DET2 (Table 1) was designed to introduce an ATG start codon and an additional ATT codon 3′ of the linker region. This resulted in the creation of a DeMPGP with an N-terminal Met followed by Ile to reproduce the N termini of both PhMPGP and TtMPGP. Primers DET1 and DET2 were used to amplify the sequence encoding the bifunctional enzyme was aligned with the MPGP and TtMPGP. The sequence encoding the 549-amino-acid polypeptides (DeMPGS-450 and DeMPGS-549, respectively). All engineered mpg genes were cloned into pKK223-3 and transformed into *E. coli* BL21-Rosetta as described above.

**Preparation of cell extracts.** All cultures of recombinant *E. coli* were centrifuged (7,000 × g, 10 min, 4°C) and suspended in 25 mM Tris-HCl (pH 7.5) containing 5 mM MgCl2, DNase I (10 μg/ml), and protease inhibitors (13). Cells were disrupted twice by French press (120 MPa), followed by centrifugation (20,000 × g, 30 min, 4°C). The supernatants were filtered through a 0.22-μm filter and used for the enzymatic assays described above. Clones carrying DeMPGP were grown, and gene overexpression was determined as described above.

**Preparation of cell extracts.** All cultures of recombinant *E. coli* were centrifuged (7,000 × g, 10 min, 4°C) and suspended in 25 mM Tris-HCl (pH 7.5) containing 5 mM MgCl2, DNase I (10 μg/ml), and protease inhibitors (13). Cells were disrupted twice by French press (120 MPa), followed by centrifugation (20,000 × g, 30 min, 4°C). The supernatants were filtered through a 0.22-μm filter and used for the enzymatic assays described above. Clones carrying DeMPGP were grown, and gene overexpression was determined as described above.

**Enzyme assays.** The reaction mixture (100 μl) used to detect enzyme activity of the recombinant MGSD in *E. coli* cell extracts and to monitor activity during purification contained 25 μl of sample, 2.5 mM GDP-mannose, and 2.5 mM 3-phosphoglycerate (3-PGA, sodium salt) in 25 mM Tris-HCl (pH 7.5) with 20 mM CoCl2. The mixture was incubated at 30°C for 30 min and cooled on ice. The assay for detection of the activity of recombinant DeMPGP was as described above, followed by incubation at 90°C for 5 min with 2 μg of pure recombinant MPGP from *T. thermophila* (IspH) (13).

**Protein purification.** The purification of MGSD was carried out by fast protein liquid chromatography in three steps: the extract was loaded onto a Q-Sepharose column (Hi-Load 16/10), and the active fractions were collected and concentrated in a 50-kDa-cutoff Centricron (Amicon), dialyzed threefold with 25 mM Tris-HCl (pH 7.5), and loaded onto a second Q-Sepharose column. The active fractions were treated as described above and loaded onto a Mono-Q column. All columns were equilibrated with 25 mM Tris-HCl (pH 7.5). Elution was carried out with linear NaCl gradients (0.0 to 1.0 M), and the active protein fractions were collected by TLC with the assay described above.

**Overspecification of mpdI in *S. cerevisiae*.** To express the bifunctional gene from *D. ethenogenes* in *S. cerevisiae* via a linear DNA fragment containing the constitutive strong promoter of the *S. cerevisiae eno2* gene (GenBank accession no. NC_001140) was amplified with primers ENO2F and ENO2R (Table 1). *S. cerevisiae* chromosomal DNA for PCR was obtained from the supernatant of a cell lysate after boiling a loopful of cells for 5 min in 200 μl of distilled water, cooling on ice, and high-speed centrifugation for 2 min in a microcentrifuge. The PCR product was cloned between the ApaI and HindIII sites of the pRS425 derivative, yielding pRSpENO.

**mgpD from *D. ethenogenes* was amplified from *D. ethenogenes* DNA with primers DEF and DER (Table 1). The PCR product was ligated into the corresponding sites of pRSPENO, yielding pDES. All reaction reactions were transformed into *E. coli* XL1-Blue, transformants were selected on LB-ampicillin agar plates, and positive clones were identified by restriction analysis. Purified pDES was transformed into *S. cerevisiae* cells, and transformants were selected for uracil auxotrophy on MM-LEU plates. Three clones were maintained on MM-LEU plates. These were inoculated and cultured to the mid-exponential phase (OD600 = 2.0) with orbital shaking at 30°C in baffled flasks containing 250 ml of MM-LEU liquid medium. Cells were collected and disrupted as described above.

**Overspecification of mpdI in *S. cerevisiae*.** To express the bifunctional gene from *D. ethenogenes* in *S. cerevisiae* via a linear DNA fragment containing the constitutive strong promoter of the *S. cerevisiae eno2* gene (GenBank accession no. NC_001140) was amplified with primers ENO2F and ENO2R (Table 1). *S. cerevisiae* chromosomal DNA for PCR was obtained from the supernatant of a cell lysate after boiling a loopful of cells for 5 min in 200 μl of distilled water, cooling on ice, and high-speed centrifugation for 2 min in a microcentrifuge. The PCR product was cloned between the ApaI and HindIII sites of the pRS425 derivative, yielding pRSpENO.

**mgpD from *D. ethenogenes* was amplified from *D. ethenogenes* DNA with primers DEF and DER (Table 1). The PCR product was ligated into the corresponding sites of pRSPENO, yielding pDES. All reaction reactions were transformed into *E. coli* XL1-Blue, transformants were selected on LB-ampicillin agar plates, and positive clones were identified by restriction analysis. Purified pDES was transformed into *S. cerevisiae* cells, and transformants were selected for uracil auxotrophy on MM-LEU plates. Three clones were maintained on MM-LEU plates. These were inoculated and cultured to the mid-exponential phase (OD600 = 2.0) with orbital shaking at 30°C in baffled flasks containing 250 ml of MM-LEU liquid medium. Cells were collected and disrupted as described above.
Phosphate guanylyl transferase/phosphomannose isomerase; \textit{pmm}

\textit{Rhodothermus marinus} \textit{(Tt)} (13), \textit{Rhodothermus marinus} \textit{(Rmar)} (6), and \textit{D. ethenogenes} \textit{(Deth)} (www.tigr.org). Arrows represent genes and their orientations. \textit{mpgs}, mannosyl-3-phosphoglycerate synthase; \textit{mpgp}, mannansyl-3-phosphoglycerate phosphatase; \textit{m1p-gt/pmi}, mannose-1-phosphate guanylyl transferase/phosphomannose isomerase; \textit{pmm}, phosphomannose mutase.

developed (13), and autoradiographed for 3 days. The TLC plate was sprayed with \( \alpha \)-naphthol and visualized by charring at 120°C (13).

Extraction and quantification of \textit{S. cerevisiae} intracellular organic solutes. Solutes accumulated in the recombinant \textit{S. cerevisiae} cells during growth in MM-LEU medium were estimated from cells harvested during the mid-exponential phase of growth (OD\(_{600} \approx 2.0\)) by centrifugation (5,000 \( \times \) g, 10 min, 4°C). Cells were washed once with ice-cold phosphate-buffered saline buffer (140 mM NaCl, 2.7 mM KCl, 10 mM Na\(_2\)HPO\(_4\), 1.8 mM KH\(_2\)PO\(_4\), pH 7.3), suspended in 10 ml of ice-cold water, and frozen at \(-80°C\) before analysis of solutes. Cells were also harvested during the early (OD\(_{600} \approx 4.7\)) and late (overnight growth, OD\(_{600} \approx 6.5\)) stationary phases of growth to evaluate the relationship between MG accumulation and growth phase. The dry mass of the cells was determined under negative pressure (42). For solute quantification by \( ^1H \)-nuclear magnetic resonance (NMR), the final pH of the extracts was adjusted to approximately 8 by the addition of NaOH. Spectra were acquired on a Bruker AMX300 spectrometer as previously described (12).

RESULTS

Identification of bifunctional \textit{mgsD} and analysis of the surrounding ORFs. BLAST searches with \textit{P. horikoshii} MPGS (PhMPGS) and MPGP (PhMPGP) and with \textit{T. thermophilus} MPGS (TtMPGS) and MPGP (TtMPGP) sequences resulted in the identification of a single ORF from \textit{D. ethenogenes} with 2,085 nucleotides encoding a 694-amino-acid polypeptide with two catalytic domains (Fig. 1). The N-terminal 399 residues in the bifunctional enzyme exhibited high sequence identity with TtMPGS (48%, GenPept AA043097), PhMPGS (44%, GenPept BAA30023), the putative MPGS found in a 34-kb DNA fragment from an uncultured crenarchaeote (44%, GenPept CAD42692), and the MPGS from \textit{R. marinus} (43%, GenPept AAP74552), and from \textit{A. pernix} (37%, GenPept BAA79872). The C-terminal 280 residues showed high sequence identity with the MPGP from \textit{R. marinus} (40%, GenPept AAP74553), PhMPGP (37%, GenPept BAA30022), and TtMPGP (34%, GenPept AAO43098) and lower identity with the putative MPGP from \textit{A. pernix} (23%, GenPept BAA79870). An MPGP homologue was absent from the DNA fragment from the uncultured crenarchaeote (GenBank AJ496176). In the \textit{Pyrococcus} sp., \textit{T. thermophilus}, and \textit{R. marinus}, mpgs and mpgp are contiguous genes, without intergenic nucleotides (the mpgs stop codon overlaps the mpgp start codon), while in \textit{A. pernix} they are separated by an intergenic region of 20 nucleotides (Fig. 1). The two functional domains of MGSD were connected by a peptide linker of 15 amino acids (Fig. 2) without homology to the C terminus of PhMPGS and TtMPGS or the N terminus of PhMPGP and TtMPGP.

The region surrounding \textit{mgsD}, as in \textit{T. thermophilus}, \textit{R. marinus}, and \textit{A. pernix}, did not contain \textit{pmm} or \textit{m1p-gt/pmi} (Fig. 2), which are found immediately downstream of mpgs and mpgp, which encode phosphomannomutase and mannose-1-phosphate guanylyltransferase, respectively, the enzymes that convert fructose-6-phosphate into GDP-mannose in the \textit{Pyrococcus} sp. Upstream of the \textit{mgsD} region we found typical bacterial −10 (TAATAA) and −35 (TTGATA) promoter elements separated by a 17-nucleotide spacer and a ribosome-binding site (AGGA) located 10 nucleotides upstream of the \textit{mgsD} start codon (Fig. 2). We also found several genes encoding elements of a putative type IV pilus structure. Downstream of \textit{mgsD} we recognized a putative glutamyl tRNA synthase gene with numerous bacterial homologues, as well as two putative “archaeal-type” genes with homology to a pyruvate-formate lyase-activating enzyme and an iron-sulfur flavoprotein (Fig. 2).

Functional overexpression in \textit{E. coli} and purification of enzymes. The frequency of the isoleucine ATA codon in \textit{mgsD} is high, accounting for nearly 50% of all isoleucine codons, with an overall frequency of 3.9%. In \textit{E. coli}, ATA is a rare codon, occurring at a frequency of about 0.4% (10). To prevent translational constraints, overexpression of \textit{mgsD} and of the truncated constructs Dempgps and Dempgp was carried out in \textit{E. coli} BL21-Rosetta, which carries a plasmid containing the tRNA genes for all rare \textit{E. coli} codons. Clones containing \textit{mgsD}, Dempgps, and Dempgp yielded purified recombinant enzymes, as visualized by SDS-PAGE (Fig. 3). We were, however, unable to produce a recombinant \textit{DeMPGS} stable enough for characterization. We also attempted to stabilize the functional MPGS domain by constructing \textit{DeMPGSs} with the peptide linker and the peptide linker plus 33 and 132 amino acids of the phosphatase domain, designated \textit{DeMPGS-L}, \textit{DeMPGS-450}, and \textit{DeMPGS-549}, respectively.

\textit{E. coli} cell extracts carrying each of the created MPGS domains except \textit{DeMPGS-549} exhibited MPG synthesis from GDP-mannose and 3-PGA, while \textit{DeMPGS} dephosphorylated MPG, as verified by TLC (results not shown). Among the MPGS polypeptides, only \textit{DeMPGS} was purified to near homogeneity but could not be characterized because it lost activity. On the other hand \textit{DeMPGS} was stable, purified to near homogeneity (Fig. 3), and characterized.

Properties of recombinant enzymes. The 694-amino-acid recombinant bifunctional MGSD with a predicted molecular mass of 78.0 kDa was consistent with the expression band visible on the SDS-PAGE gel of MGSD-containing cell extracts. Only GDP-mannose and 3-PGA served as substrates for the synthesis of MG. Mannosyl-3-phosphoglycerate (MPG) was the only sugar phosphate to be dephosphorylated by
FIG. 2. Nucleotide and amino acid sequences of the bifunctional mgsD gene and the corresponding enzyme and flanking regions in *Dehalococcoides ethenogenes*. Arrows represent genes and their directions. a, competence PilM protein homolog; b, putative ATPase involved in pili biogenesis, PilT; c, putative prepilin signal peptidase; d, putative glutamyl-tRNA synthetase; e, putative pyruvate-formate lyase-activating enzyme; f, putative iron-sulfur flavoprotein. Shaded arrows represent genes coding for proteins with generally higher homology to archaeal than to bacterial proteins. The enlargement shows the promoter region of the mgsD gene with the consensus promoter elements (boxes), the probable ribosome-binding site (RBS, underlined), and the ATG start codon (bold). The shaded area represents the peptide linker between the MPGS and MPGP domains of the bifunctional MGSD.
MGSD, yielding MG. The α-configuration of MG was confirmed from the NMR measurement of the coupling constant between the anomeric carbon and the directly bound proton (J = 171.8 Hz) (results not shown).

The enzyme was active between 10 and 60°C, with maximum activity at around 50°C (Fig. 4). The half-life for thermal inactivation was 10.5 min at 50°C. Within the pH range examined (5.5 to 8.5), the activity of the enzyme at 50°C was maximal near pH 7.0 (Fig. 4). The V_max of the enzyme was 3.2 μmol/min/mg of protein with GDP-mannose and 3-phosphoglycerate as the substrate and 6.1 μmol/min/mg of protein with MPG as the substrate. The rate constants (k_cat), substrate K_m, and catalytic efficiencies (k_cat/K_m) are represented in Table 2. The bifunctional MGSD had an absolute requirement for cations; maximal activity was obtained with Co^{2+} ions (chloride or sulfate salts), which could also be partially replaced by Mg^{2+} or Mn^{2+} ions (Fig. 5). Other cations tested did not stimulate MGSD activity at any concentration tested.

Purified DeMPGP showed an apparent 30-kDa band that was in agreement with the genetic construction. The recombinant enzyme had an optimal temperature for activity of about 50°C and a pH optimum of 7.0 (results not shown). The V_max was 16.3 μmol/min/mg of protein, about fivefold higher than the V_max calculated for the bifunctional enzyme and 2.7-fold higher than the dephosphorylating activity of MGSD. The rate constant (k_cat), K_m for MPG, and catalytic efficiency (k_cat/K_m) are represented in Table 2. Recombinant DeMPGP also had an absolute requirement for divalent cations; the preferred metal ion was Ni^{2+} at a concentration of 20 mM, but it could be partially replaced by Co^{2+}, Mg^{2+}, or Mn^{2+} ions (results not shown). Since the activity of DeMPGP was enhanced by Ni^{2+} and the activity of MGSD was enhanced by Co^{2+}, we examined various combinations of these cations on the rate of MGSD activity but found that these combinations did not stimulate the activity of the bifunctional enzyme.

Synthesis and accumulation of compatible solutes by recombinant S. cerevisiae. Recombinant S. cerevisiae carrying pDES was grown in MM-LEU to a cell density of 2.0 (OD_{600}), corresponding to 800 mg of cell dry mass per liter of culture. The recombinant yeast cells accumulated MG to an intracellular concentration of 49 μmol/g of cell dry mass during growth in MM-LEU medium (Table 3), while MG was not detected in the parental strain (data not shown). Glycerol was detected at a concentration of 17 μmol/g of cell dry mass.

The accumulation of MG in recombinant S. cerevisiae cells was higher during the mid-exponential phase of growth (OD_{600} = 2.0), decreasing with culture age to 46 μmol/g of cell dry mass in the early stationary phase (OD_{600} = 4.7) and to 28 μmol/g of cell dry mass in the late stationary phase (OD_{600} = 6.5) (Table 3). Trehalose and glycerol were also detected in small amounts. Trehalose was detected in the early and late stationary phases, while glycerol was present at all phases of growth (Table 3).

The recombinant bifunctional MGSD had in vivo activity under all conditions examined, as verified from the intracellular accumulation of MG. However, incubation of recombinant S. cerevisiae cell extracts with labeled GDP-[U-14C]mannose and unlabeled 3-PGA, followed by autoradiography of the TLC plate, revealed no in vitro activity. Under the same conditions, a pure preparation of the E. coli recombinant MGSD showed synthesis of radiolabeled MG (Fig. 6). An MGSD expression band was not detected in SDS-PAGE of recombinant S. cerevisiae, and we could not purify the recombinant MGSD because we could not detect activity in the cell extracts.

DISCUSSION

Accumulation of mannosylglycerate is known to occur in several species of red algae and in thermophilic and hyperthermophilic bacteria and archaea (25, 40). We now show that a mesophilic bacterium, unrelated to any known (hyper)thermophile, has a gene that is involved in the synthesis of MG. The synthesis and accumulation of MG could not be proven to occur in D. ethenogenes because of insufficient biomass production for analysis of intracellular compatible solutes. However, the role of the bifunctional mgsD was confirmed, to our surprise, in S. cerevisiae, which lacks the genetic machinery to produce MG. The results strongly indicate that the bifunctional enzyme MGSD synthesizes MG in D. ethenogenes and that this compound may serve as a compatible solute in the organism that harbors mgsD.

The V_max values determined for the bifunctional MGSD were significantly lower than those of thermophilic and hyperthermophilic monofunctional MPGSs and MPGP s (6, 12, 13). These results are not unprecedented when comparing mesophilic and thermophilic homofunctional enzymes (44).
TABLE 2. Kinetics of recombinant enzymes

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$K_m$ (mM)</th>
<th>$k_{cat}/K_m$ (M$^{-1}$ s$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Both substrates</td>
<td>GDP-mannose</td>
<td>3-PGA</td>
</tr>
<tr>
<td>MGSD</td>
<td>3.9</td>
<td>7.9</td>
<td>0.70</td>
</tr>
<tr>
<td>DeMPGP</td>
<td>7.9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The effect of the protein fusion on the catalytic efficiency ($k_{cat}/K_m$) of MGSD was partially examined by the kinetic analysis of DeMPGP, since we lacked a stable monofunctional DeMPGS. DeMPGP and MGSD have similar turnover ($k_{cat}$) values at 50°C. Nevertheless, a decrease in the $K_m$ for MPG resulted in a twofold-higher catalytic efficiency of the monofunctional DeMPGP over the bifunctional MGSD. Similar findings were reported in a comparison between fusion enzymes from the aspartate pathway in E. coli and its independent domains (24). Conversely, a higher $k_{cat}/K_m$ for fusion enzymes over their natural independent monofunctional components has also been reported and attributed to more efficient substrate channeling between enzyme domains and tight co-regulation of expression (3, 41, 46).

The evolutionary success of gene fusions depends on factors such as the linker extension between domains or the ability of each of the polypeptides to fold correctly into a bifunctional enzyme (24, 41). The adaptation to mesophily, if we assume horizontal transfer of mpgs and mppg from a thermophilic organism, may have provided the monofunctional enzymes with increased flexibility for catalysis at lower temperatures. However, increased flexibility leads to decreased stability (9). The fusion between mpgs and mppg could have provided stability to the enzyme. The triosephosphate isomerase-phosphoglycerate kinase fusion protein of Thermotoga maritima is, for example, more stable than the artificial triosephosphate isomerase domain (3). The MPGS domains constructed by us were unstable and lost activity during purification. Only the full-length, bifunctional enzyme was stable enough to characterize, indicating that the whole C-terminal domain (MPGP) is required to stabilize the MPGS domain in vitro.

MGSD shares most of its in vitro biochemical properties with its monofunctional thermophilic and hyperthermophilic counterparts. A notable exception is the thermal profile, because MGSD is active at temperatures as low as 10°C, has optimum activity at 50°C, and has no detectable activity at 60°C. MGSD and DeMPGP as well as the MPGSs from T. thermophilus and R. marinus but not P. horikoshii have absolute requirements for divalent cations (6, 12, 13). Curiously, other enzymes involved in the synthesis of other sugar or sugar derivative compatible solutes such as glucosylglycerol-phosphate and sucrose-phosphate synthases from Synechocystis sp. strain PCC6803 and trehalose-phosphate synthase from Mycobacterium smegmatis are active in the absence of cations (8, 21, 34).

The identification of a bifunctional mgsD gene in D. ethenogenes was also unexpected in the sense that all prokaryotes known to produce MG via the two-step pathway have separate mpgs and mppg genes (6, 12, 13). In Pyrococcus spp., mpgs and mppg are contiguous, and two extra genes, pmm and mlp-glp/pmi, encoding enzymes of the MG pathway, are located downstream, forming an operon-like structure (12). The gene organization in T. thermophilus and A. pernix is different, because pmm and mlp-glp/pmi are located elsewhere in the genome (13, 26). Another genetic structure was encountered in a crenarchaeal DNA fragment isolated from soil that contained a putative mpgs gene alone (35). This sequence lacks the adjacent mppg gene that codes for the enzyme that ultimately yields MG from MPG. The mppg gene may be located elsewhere in the genome. A similar organization occurs in Synechocystis sp. strain PCC6803 for the synthesis of the compatible solute glucosylglycerol, where the genes ggpS and stpA, encoding the glucosylglycerol-phosphate synthase and phosphatase, respectively, are controlled by individual promoters (20, 28). It is also possible that, in this uncultured crenarchaeote, the mpgs homologue is involved in the synthesis of an unknown compound.

The assays for determination of activity were performed in the presence of two different concentrations of cobalt (solid bars), magnesium (dark shaded bars), and manganese ions (light shaded bars).

**FIG. 5.** Metal ion dependence of the in vitro activity of MGSD. The assays for determination of activity were performed in the presence of two different concentrations of cobalt (solid bars), magnesium (dark shaded bars), and manganese ions (light shaded bars).

**TABLE 3.** Accumulation of organic solutes by recombinant S. cerevisiae

<table>
<thead>
<tr>
<th>Growth phase (OD$_{600}$)</th>
<th>Solute accumulation (μmol/g of cell dry mass)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MG</td>
</tr>
<tr>
<td>Mid-exponential (2.0)</td>
<td>49</td>
</tr>
<tr>
<td>Early stationary (4.7)</td>
<td>46</td>
</tr>
<tr>
<td>Late stationary (6.5)</td>
<td>28</td>
</tr>
</tbody>
</table>

*Cells carried the bifunctional mgsD from D. ethenogenes. Solute concentrations are the mean values for at least two experiments. ND, not detected.
or synthesizes MPG as an intermediate for other macromolecules.

The presence of a gene for the synthesis of MG in the mesophilic bacterium *D. ethenogenes* may reflect a lateral gene transfer event. The thermophiles and hyperthermophiles *A. pernix*, *Pyrococcus* spp., *Rubrobacter* *xylanophilus*, *T. thermophilus*, and *R. marinus* have an insignificant phylogenetic relationship but share the ability to synthesize MG in response to salt and/or heat stress (6, 13, 40, 42). Transfer of a specific characteristic may occur across vast phylogenetic distances as an adaptation mechanism for improved colonization of specific ecological niches (32). Some genes present in *D. ethenogenes* and *Thermotoga* spp. seem to have been acquired from archaea by lateral gene transfer (33). The archaeal-type *ghiB* gene, encoding a subunit of glutamate synthase, for example, may have been acquired by lateral gene transfer from the *Archaeoglobus* lineage (33). The *mgsD* gene has a higher frequency of purines, especially of adenine, over pyrimidines and a preferred usage of ATA for isoleucine over ATC or ATT codons. This nucleotide usage could reflect a thermophilic origin of *mgsD* because it is more common to genome sequences of thermophilic than mesophilic prokaryotes (43). The successful transfer of a specific gene to a new host is not only a multistep process but naturally depends on transformation, transfection, and conjugation events (11). Type IV pilus-related genes found upstream of *mgsD*, involved in the natural transformation machinery of numerous organisms, indicate natural competence of *D. ethenogenes* and hence ability for DNA acquisition (2).

Di-*myo*-inositol phosphate, like MG, is a common compatible solute of hyperthermophilic archaea but also accumulates in some *Thermotoga* species (40). The *ino1* gene found in these *Thermotoga* spp. encodes *myo*-inositol-1-phosphate synthase, which catalyzes a step in *di-myo*-inositol phosphate biosynthesis (7). The adaptation of some *Thermotoga* spp. to saline environments may have relied, in part, on the acquisition of *ino1* from hyperthermophilic archaea for the synthesis of di-*myo*-inositol phosphate (33). The *ino1* gene has a scattered and restricted distribution in bacteria, being found only in some *Thermotoga* strains and *Aquifex alicicus* but also in the mesophilic bacteria *D. ethenogenes*, *Mycobacterium* spp., and *Streptomyces* spp. (33). The function of *ino1* in *D. ethenogenes* is unknown, but it codes for enzymes involved in the synthesis of cell wall components in *Mycobacterium* spp. and antibiotics in *Streptomyces* spp. (4, 45). The putative acquisition of *mgsD* by lateral gene transfer may have provided *D. ethenogenes* with the ability to colonize estuarine and deep-sea hydrothermal vent environments (19, 23).

The small number of sequences from genes involved in the synthesis of MG that are currently available are not sufficient to confirm lateral gene transfer, but the overall absence of these genes from mesophilic bacteria, the scattered distribution of MG in prokaryotes and eukaryotes, and its relevance to osmotic adaptation in saline environments lead us to envision lateral gene transfer for acquisition of *mgsD* by *D. ethenogenes*.

The biosynthesis and role of MG in *D. ethenogenes* remain elusive. We did not obtain synthesis of MG in the MGSD-containing *E. coli*, perhaps because the required substrates were not available. Similar findings were observed in *E. coli* strains expressing the *mpgs* gene from *T. thermophilus*, *R. marinus*, and *P. horikoshii* (6, 12, 13). Trace amounts of MG were synthesized by a recombinant *E. coli* strain expressing the *R. marinus* MGS gene only after genetic manipulation to increase the flux to the synthesis of GDP-mannose (39). Moreover, *mgsD* was expressed in this metabolically engineered *E. coli* strain, but MG was not detected (M. M. Sampaio and H. Santos, unpublished data). However, after expression of *mgsD* in *S. cerevisiae* in an attempt to produce an enzyme with higher specific activity, we found that the recombinant yeast cells accumulated MG. This remarkable result showed that MGSD synthesizes MG in vivo and argues for a role in the osmotic adaptation of the *D. ethenogenes* to salt stress. However, we were unable to measure MGSD activity in the *S. cerevisiae* cell extracts. Similar results have been obtained previously for heterologous gene expression in this yeast (31). The in vivo half-life of a protein is extremely variable and depends on numerous factors, namely, the N-terminal residue, specific sequences or domains, and posttranslational modifications that are recognized by components of the ubiquitin proteolytic pathway that target proteins for destruction (18). A ubiquitin-mediated high turnover of MGSD could explain the absence of activity in the yeast cell extracts.

*S. cerevisiae* strains overproducing MG provide a good in vivo model and a valuable tool for unraveling the specific role of MG in osmotic and thermal protection of biomolecules and cells. Ongoing studies with double-knockout mutants on the genes leading to the synthesis of glycerol, the canonical compatible solute of *S. cerevisiae*, will also elucidate the role of MG in the osmotic adjustment of this yeast. Whatever the results, the fact remains that a mannosylglycerate synthase with low optimal temperature for activity, encoded by a gene from a mesophilic bacterium, produces this rare compatible solute in *S. cerevisiae*.

ACKNOWLEDGMENTS

This research was funded by the European Commission 5th Framework Program, project QLK3-CT-2000-00640, and FCT/FEDER project POCTI/35715/BIO/2000. Preliminary *D. ethenogenes* sequence data were obtained from the Institute for Genomic Research website. Se-
quencing of \textit{D. ethenogenes} was accomplished with support from the U.S. Department of Energy. N. Empadinhas acknowledges a Ph.D. grant from FCT (PRA4XXI/BID/21665/99). S. Zinder acknowledges support from the NSF Metabolic Biochemistry program.

We thank Ana Isabel Mingote (ITQB, Oeiras, Portugal) for quantification of intracellular solutes by NMR.

**REFERENCES**

1. Adrian, L., U. Swenyk, J. Wecke, and H. Gorisch. 2000. Bacterial dehalo-

2. Averhoff, B., and A. Friedrich. 2003. Type IV pili-related natural transfor-

mophilic bacterium \textit{Thermotoga maritima}: design and characterization of the separate triosephosphate isomerase. Protein Sci. 6:2159–2165.


glycerate in osmoadaptation and thermoadaptation of \textit{Rhodothermus marti-
na}. J. Biol. Chem. 279:8992–8998.


8. Curatti, L., E. Folco, P. Desplats, G. Abrati, V. Limones, L. Herrera-


17. Gietz, R. D., R. H. Schiestl, A. W. Willems, and R. A. Woods. 1995. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG pro-

18. Glickman, M. H., and A. Ciechanover. 2002. The ubiquitin-proteasome pro-


**BIFUNCTIONAL MG SYNTHASE FROM D. ETHENOGENES**

**4083**

