Identification and Mutation of a Gene Required for Glyceraldehyde Kinase Activity from a Facultative Methylotroph, *Methylobacterium extorquens* AM1

LUDMILA CHISTOSERDOVA1 AND MARY E. LIDSTROM1,2*

Department of Chemical Engineering2 and Department of Microbiology,1 University of Washington, Seattle, Washington 98195

Received 12 March 1997/Accepted 29 May 1997

A gene (gckA) responsible for the activity of glyceraldehyde kinase has been identified within a chromosomal fragment of the serine cycle methylotroph *Methylobacterium extorquens* AM1. A mutation in gckA leads to a specific C₃-negative phenotype. The polypeptide sequence derived from gckA showed high similarity to a product of *tteD* essential for tartrate metabolism in *Agrobacterium vitis*. Our data suggest that gckA and *tteD* might be structural genes for glyceraldehyde kinase and that the serine cycle and the tartrate utilization pathway share a series of reactions.

The serine cycle is one of the assimilatory pathways found in methylotrophic bacteria, and it is known to be present in α- proteobacterial methylotrophs of the genera *Methylobacterium*, *Hyphomicrobium*, *Methylosinus*, and *Methylocystis* (16). This pathway effects the net conversion of formaldehyde to C₃ compounds (15) as follows. Formaldehyde in the form of methylentetrahydrofolate is condensed with glycine by serine hydroxymethyltransferase to produce serine. Serine is transaminated with glyoxylate to produce hydroxypyruvate in the serine glyoxylate aminotransferase (HPR); the latter is phosphorylated by glyceraldehyde kinase (GK) to produce phosphoglyceraldehyde. Phosphoglyceraldehyde is converted into phosphoenolpyruvate (PEP) by enolase, PEP is carboxylated to oxaloacetate by PEP carboxylase (PEPC), and the latter is converted to malate, which is transformed to malyl-CoA by malate thiokinase, PEPC, and malyl-CoA lyase. Glyoxylate is then used to regenerate glycine. The second molecule of glycine is regenerated from acetyl-CoA via an unknown pathway. Although the reactions of the serine cycle were identified over 20 years ago, it was not until recently that genetic studies of the serine cycle began to catch up with the biochemistry. At present, four unlinked chromosomal regions of *Methylobacte- rium extorquens* AM1 have been shown to contain genes for serine cycle enzymes. One of these loci contains six genes, all closely linked and transcribed in the same direction, encoding serine glyoxylate aminotransferase, HPR, the two subunits of malate thiokinase, PEPC, and malyl-CoA lyase (1, 6). The second locus contains the gene for serine hydroxymethyltransferase (3). The third locus contains three genes encoding enzymes essential for the conversion of acetyl-CoA into glyoxylate (5). The fourth locus has been shown to complement a number of chemically induced mutants defective in the activity of GK (22). Except for the genes for PEPC and GK, all of the genes noted above have been characterized at the molecular level, insertion mutations in these genes have been obtained, and the mutants have been characterized. The gene for PEPC has been localized on the chromosome and partially sequenced (4, 6), while the gene for GK has remained uncharacterized. This paper deals with identification and characterization of this gene in *M. extorquens* AM1.

In this study, *Escherichia coli* strains were grown in Luria-Bertani medium in the presence of appropriate antibiotics as described by Maniatis et al. (17). *M. extorquens* AM1 was grown in the minimal medium described previously (12). Suc- cinate (20 mM), methanol (100 mM), methylamine (20 mM), ethanol (40 mM), or ethylamine (20 mM) was used as substrate. Methanol induction of mutants was carried out as described by Dunstan et al. (9). The following antibiotics were used for *M. extorquens* AM1: tetracycline, 10 mg/liter; kanamycin, 100 mg/liter; and rifampin, 50 mg/liter. DNA-DNA hybridizations were carried out with dried agarose gels as described by Meinkoth and Wahl (18) at 68°C. For hybridizations, 6× SSC (1× SSC is 0.15 M NaCl and 0.015 M sodium citrate) was used, and 0.5× SSC was used for washes. Plasmid isolation, *E. coli* transformation, restriction enzyme digestion, ligation, blunting of ends with T4 DNA polymerase, or filling in of ends with Klenow enzyme was carried out as described by Maniatis et al. (17). The chromosomal DNA of *M. extorquens* AM1 was isolated by the procedure of Saito and Miura (21). DNA sequencing was carried out with an Applied Biosystems automated sequencer by the Caltech Sequencing Facility, from both strands. Translation and analyses of DNA and DNA-derived polypeptide sequences were carried out with Genetic Computer Group (Madison, Wis.) programs. The activity of GK was determined as described by Goodwin (10). Spectrophotometric methods (14, 23) were used for protein determination. Triparental or biparental matings between *E. coli* and *M. extorquens* AM1 were performed overnight on nutrient agar. Cells were then washed with sterile medium and plated on selective medium at appropriate dilutions. In triparental matings, pRK2013 (8) was used as a helper plasmid. Rifampin was used for *E. coli* counterselection.

**Sequence analysis.** The 8.5-kb HindIII fragment responsible for complementation of chemically induced mutants defective in GK was found in an *M. extorquens* AM1 gene library (plasmid pSS13-1) and is described by Stone and Goodwin (22). Plasmid pSS13-1, as well as the GK mutants PG2 and PT1001, was kindly provided by P. Goodwin. In this work, we have sequenced a 5.6-kb BamHI-HindIII fragment within this 8.5-kb...
fragment. An overlapping PstI fragment was also isolated and characterized after a partial open reading frame (ORF) similar to those of pyruvate kinase (PK) genes was discovered at one end of the insert in pSS13-1. The physical map of the region sequenced in this study is shown in Fig. 1. Seven complete ORFs and one partial putative ORF have been found in the region (Fig. 1). These ORFs were identified as having a codon preference typical of M. extorquens AM1. The 1,535-bp stretch separating orf5 and orf6 most probably represents a noncoding region, since none of the potential ORFs in this region showed appropriate codon usage. It is interesting to point out that the fourth ORF shows perfect M. extorquens AM1 codon usage over its 5′-terminal part (nucleotides 1823 to 2800), but the codon usage in its 3′ part (nucleotides 2800 to 3136) is rather poor.

Complementation analysis. The 5-kb BamHI and the 6-kb PstI fragments (Fig. 1) were subcloned into the conjugative plasmid pRK310 (8) for complementation analysis. The resulting plasmids were transferred into two GK-negative mutants, PG21 and PT1001, and transconjugants were selected by their resistance to tetracycline. The transconjugants containing the plasmid with the 5-kb BamHI fragment regained their ability to grow on C1 compounds. The transconjugants containing the 6-kb PstI fragment did not (Fig. 1).

Amino acid sequence comparisons. The amino acid sequences deduced from complete and partial ORFs found in the region under study were compared against the protein data bank, including the translated products of the available microbial genomes. The first ORF showed high identity with a nonidentified partial ORF translated from the trpE(D) gene region of Azospirillum brasilense (GenBank accession no. U44127; 71% identity in 28-amino-acid overlap), and this gene was designated orf1. The polypeptide deduced from the fourth ORF showed high identities with the products of two plasmid-borne ttuD genes participating in tartrate utilization in Agrobacterium vitis (GenBank accession no. U32375 and U25634; 50 and 44% identity over the entire length, respectively [7]). This gene became a candidate for encoding GK, and it was designated gckA. The polypeptide derived from the sixth ORF showed considerable identity with PKs from various sources (33 to 39% identity), and the gene was tentatively designated pykA. The polypeptide derived from the seventh ORF showed low identity with proline aminopeptidase from Lactobacillus delbrueckii, poly-3-hydroxyalkanoate depolymerase from Pseudomonas oleovorans, and soluble epoxide hydrolase from mammalian tissues (GenBank accession no. P46544, P26495, and P80299, respectively; 23 to 33% identity). This gene was designated orf5. The polypeptide translated from the eighth, partial ORF showed identity with an unidentified ORF in the mgoS region of the Rhodobacter sphaeroides chromosome (39% identity in 64-amino-acid overlap [20]). This gene was designated orf6. Amino acid sequences deduced from the second, third, and fifth ORFs did not reveal strong identity to any known proteins, and those were designated orf2, orf3, and orf4, respectively.

Insertion mutagenesis of gckA and pykA. Insertion mutations were generated in two ORFs in the region, gckA and pykA, as described earlier for other M. extorquens AM1 genes (2, 5). The sites used for insertion of a Km gene were the SmaI site in the first half of gckA and the HindIII site in the beginning of pykA. Mutants were selected on succinate-containing medium in the presence of kanamycin and checked for their resistance to tetracycline and for growth on C1 compounds. About 4% of the gckA mutants were Tc° (putative double-crossover recombinants), and all of these were unable to grow on C1 compounds. The mutants were still able to grow on C2 compounds, indicating the specific involvement of gckA in C1 metabolism. About 20% of the pykA mutants were putative double crossovers, and these were able to grow on both C1 and C2 compounds, indicating that pykA is not involved in C1 metabolism. In a separate mutagenesis experiment in which insertion mutants were selected on methanol plates, putative double-crossover mutations in pykA appeared with the same frequency as on succinate plates, and these C1-positive insertion mutants were able to grow on succinate. Examples of all putative double-crossover mutants were used in Southern hybridization experiments to confirm the expected recombination event (data not shown).

Activity of GK in gckA mutants. GK activity was measured in gckA insertion mutants and in mutants complemented for growth on C1 compounds by the 5-kb BamHI fragment containing gckA (Fig. 1) cloned into pRK310. GK activity was not detectable in the mutants (data not shown) and was present in both wild-type M. extorquens AM1 and complemented mutants (20 to 40 nmol min⁻¹ mg of protein⁻¹), confirming the necessity of gckA for GK activity.

Conclusions. This work describes a new gene, gckA, involved in the serine cycle for formaldehyde assimilation in M. extorquens AM1. gckA is located on a DNA fragment not linked to known fragments containing other methylotrophic genes and is required for GK activity. GK enzymes have been characterized for a number of organisms, including a serine cycle methylotroph, Hyphomicrobium methylotrorum GM2 (24), but their amino acid sequences remain unknown. GCK has a predicted molecular mass of 44.5 kDa, which is in good agreement with the molecular mass determined for H. methylotrorum GM2 purified GK (41 to 48 kDa [24]). Thus, gckA probably encodes the structural gene for GK. The product of gckA showed similarity to one other protein in the protein database, TtuD, which is required for tartrate utilization in a plant parasite, A. vitis (7). The proposed tartrate utilization pathway has some reactions in common with a part of the serine cycle (Fig. 2). The reaction carried out by TtuD is unknown, but it has been suggested that TtuD might be an HPr, carrying out the second enzymatic reaction in tartrate utilization (7). However, TtuD shows no similarity to HprA or other HPRs but does show similarity to GckA. Therefore, our data suggest that TtuD is probably a GK and that tartrate metabolism enters glycolysis via 2-phosphoglycerate (Fig. 2). GKS from M. extorquens AM1 and A. vitis may represent a separate class of kinases, since they show no significant similarity to other known kinases.

No similarity was found between GckA and any protein translated from the E. coli chromosome, although E. coli is known to contain a GK. However, the E. coli GK is known to produce 3-phosphoglycerate from glyceral (19), while the
Serine cycle (portion involving GK)

\[
\text{SGAT} \rightarrow \text{glyoxylate} \rightarrow \text{glycerate} \rightarrow \text{2-phosphoglycerate} \rightarrow \text{phosphoglycerate} \rightarrow \text{pyruvate}
\]

Tartrate utilization in \(A. \textit{vitis}\)

\[
\text{TDDH} \rightarrow \text{glyoxylate} \rightarrow \text{glycerate} \rightarrow \text{2-phosphoglycerate} \rightarrow \text{pyruvate}
\]

FIG. 2. Common reactions of the serine cycle and proposed tartrate utilization pathway. SGAT, serine glyoxylate aminotransferase; PGM, phosphoglyceromutase; TDDH, tartrate dehydrogenase. Corresponding genes (if known) are shown in italics.

It is not required for growth on C1 compounds. For putative PK enzyme in \(M. \text{extorquens}\) is also found near regulated by tartrate (7). A gene that apparently encodes a PK \(ttuE\). Although it is not required for tartrate utilization, it is been suggested that TtuE might serve to stimulate the rate of glycolysis, allowing more rapid turnover of tartrate (7), and a been suggested that TtuE might serve to stimulate the rate of glycolysis, allowing more rapid turnover of tartrate (7), and a

\[\text{GK-encoding gene in } M. \text{extorquens AM1}\]

is different from \(M. \text{extorquens}\). In \(A. \textit{vitis}\), a gene whose product shows similarity to a number of PKs is located in the \(ttu\) gene cluster and was designated \(ttuE\). Although it is not required for tartrate utilization, it is regulated by tartrate (7). A gene that apparently encodes a PK is also found near \(gckA\) in \(M. \text{extorquens AM1}\). The role of this putative PK enzyme in \(M. \text{extorquens AM1}\) is unknown, but it is not required for growth on C1 compounds. For \(A. \textit{vitis}\), it has been suggested that TtuE might serve to stimulate the rate of glycolysis, allowing more rapid turnover of tartrate (7), and a PK could play a similar role in \(M. \text{extorquens AM1}\) during growth on C1 compounds, shunting a portion of the carbon to pyruvate for biosynthetic purposes (Fig. 2). It is interesting to note that \(Methyllobacterium\) strains are commonly found on plant surfaces, and some strains can grow poorly on tartrate (11). Since \(A. \textit{vitis}\) \(ttu\) genes are carried on conjugative plasmids (7), it is possible that these genes had a common origin via interspecies gene transfer.

Nucleotide sequence accession number.
The sequence of 7,883 nucleotides has been deposited with GenBank under accession no. U87316.

We thank P. Goodwin for providing plasmid pSS13-1 and mutants PG2 and PT1001.

This work was supported by a grant from the NIH (GM36294).

REFERENCES


7. Ditta, G., T. Schmidhauser, F. Jacobson, P. Lu, X. Liang, D. Finlay, D. Guiney, and D. Helinsky. 1982. Plasmid pSS13-1 and mutants PG2 and PT1001. This work was supported by a grant from the NIH (GM36294).


