Nucleotide sequence and expression analysis of the *Acetobacter xylinum* phosphoglucomutase gene

Trygve Brautaset, Rune Standal, Espen Fjærvik and Svein Valla

Author for correspondence: Svein Valla. Tel: +47 73598694. Fax: +47 73598705.

The *Acetobacter xylinum* gene (*celB*) encoding phosphoglucomutase (EC 5.4.2.2) has previously been cloned by complementation of cellulose-negative mutants. In the present report the nucleotide sequence of a 2.0 kb DNA fragment containing *celB* is described. Expression analysis using the bacteriophage T7 RNA polymerase promoter 410 resulted in identification of a probable translational start codon of *celB*, and this conclusion was confirmed by N-terminal amino acid sequencing of the recombinant protein. From the nucleotide sequence data it was deduced that *celB* encodes a protein with a calculated molecular mass of 59.6 kDa. A protein of similar size was visualized after in vitro transcription and translation, using the cloned 2.0 kb fragment as template. The results of an amino acid sequence comparison and a biochemical analysis indicated that the Cel6 protein is structurally and functionally related to the previously characterized human and rabbit phosphoglucomutases.

**Keywords**: *Acetobacter xylinum*, phosphoglucomutase, *celB*, cellulose

**INTRODUCTION**

Phosphoglucomutase catalyses the interconversion of D-glucose 6-phosphate and D-glucose 1-phosphate, which represents a branch point in carbohydrate metabolism. Biochemical studies indicate that phosphoglucomutases from a variety of organisms convert glucose 1-phosphate much more efficiently than the corresponding 6-phosphate isomer (Lowry & Passonneau, 1969; Ray & Peck, 1972). Glucose 6-phosphate enters catabolic processes to yield energy and reducing power, whereas glucose 1-phosphate is the precursor of sugar nucleotides that are used by the cells in the synthesis of various glucose-containing polysaccharides. In a previous report we showed that phosphoglucomutase is essential for the formation of extracellular cellulose in *Acetobacter xylinum*, as mutants deficient in the corresponding gene are unable to produce cellulose (Fjærvik *et al.*, 1991). A similar observation has also recently been reported for xanthan production in *Xanthomonas campestris* (Köpín *et al.*, 1992). In *Escherichia coli* it has been shown that mutants deficient in phosphoglucomutase accumulate intracellular amylase when grown in the presence of maltose (Adhya & Schwartz, 1971). The probable reason for this is that maltose metabolism results in the formation of glucose 1-phosphate which cannot be channelled into catabolism due to the phosphoglucomutase deficiency.

In the active form of phosphoglucomutase, a divalent metal ion is bound to the enzyme, and a serine residue at the catalytic site is phosphorylated (Ray & Peck, 1972). This phosphate group is initially transferred to the substrate, and the serine residue is rephosphorylated concomitant with the release of either of the two phosphoglucose isomers (Rhyu *et al.*, 1984). Recently the crystal structure of rabbit muscle phosphoglucomutase was described, and it was found that three loops are closely spaced in the active-site cleft (Dai *et al.*, 1992). One loop contains the active-site serine, one is a metal-ion-binding loop, and the third is suggested to be involved in substrate-binding specificity.

Phosphomannomutase catalyses a reaction similar to the phosphoglucomutase reaction, and it has been reported that the phosphomannomutase (AlgC) of *Pseudomonas aeruginosa* contains both the active-site and the metal-binding-loop sequence motifs typical of phosphoglucomutases (Zielinsky *et al.*, 1991). Interestingly, it has been shown that an enzyme from *X. campestris* can efficiently convert both phosphorylated glucose and mannose (Köpín *et al.*, 1992).

In this paper the nucleotide sequencing and expression analysis of the *A. xylinum* phosphoglucomutase gene *celB*...
Table 1. Bacterial strains and plasmids used in this study

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Characteristics* and references</th>
</tr>
</thead>
<tbody>
<tr>
<td>E. coli Hfr3000</td>
<td>pgm+ (Adhya &amp; Schwartz, 1971)</td>
</tr>
<tr>
<td>PGM1</td>
<td>pgm derivative of Hfr3000 (Adhya &amp; Schwartz, 1971)</td>
</tr>
<tr>
<td>PUC18.8.6B.dE</td>
<td>pUC18 with a 3.8 kb EcoRI-BamHI insert; celB+ Ap’ (Fjærvik et al., 1991)</td>
</tr>
<tr>
<td>pUC128</td>
<td>Inducible lac promoter upstream of a polylinker; Ap’ (Keen et al., 1988)</td>
</tr>
<tr>
<td>pT7.4</td>
<td>Bacteriophage T7 φ10 promoter upstream of a polylinker; Ap’ (Tabor &amp; Richardson, 1985)</td>
</tr>
<tr>
<td>pGP1-2</td>
<td>Thermo-inducible T7 RNA polymerase gene; Km’ (Tabor &amp; Richardson, 1985)</td>
</tr>
<tr>
<td>pT7-4S</td>
<td>pT7-4 with an Sphl linker ligated into the polylinker SmaI site (this study)</td>
</tr>
<tr>
<td>pTB2A/B</td>
<td>pUC128 with a 2.0 kb Sphl insert from pUC18.8.6B.dE in orientation A and B, respectively; celB+ (this study)</td>
</tr>
<tr>
<td>pTB3A/B</td>
<td>pT7-4S with a 2.0 kb Sphl insert from pUC18.8.6B.dE in orientation A and B, respectively; celB+ (this study)</td>
</tr>
<tr>
<td>pTB1</td>
<td>pTB3A carrying nt 99 through 2058 from celB (this study)</td>
</tr>
<tr>
<td>pTB8</td>
<td>pTB3A carrying nt 110 through 2058 from celB (this study)</td>
</tr>
<tr>
<td>pTB16</td>
<td>pTB3A carrying nt 161 through 2058 from celB (this study)</td>
</tr>
</tbody>
</table>

* Ap’, ampicillin resistance; Km’, kanamycin resistance.

is reported, and some biological properties of the gene product are described.

METHODS

Bacterial strains, plasmids and growth media. Strains and plasmids used in this study are listed in Table 1. E. coli strains were, unless otherwise stated, grown at 37 °C in L-broth or on L-agar (Valla et al., 1989) supplemented when appropriate with ampicillin (100 μg ml⁻¹) or kanamycin (50 μg ml⁻¹).

Preparation of DNA, molecular cloning, construction of deletion mutants, DNA sequencing, and computer analysis. Plasmid isolations, agarose gel electrophoresis, restriction endonuclease digestions, ligations and other routine DNA manipulations were performed according to standard protocols (Sambrook et al., 1989). Transformations were performed according to Chung et al. (1989). Unidirectional exonuclease degradations (Henikoff, 1987) of plasmids pTB3A and pTB3B were done with the Erase-a-base system (Promega). The resulting deletion derivatives were used for sequencing of both DNA strands by the dideoxy chain-termination method of Sanger et al. (1977). Homology searches with the deduced amino acid sequence of CelB were done using the BLAST network service at NCBI and the non-redundant DNA and protein sequence databases (Altschul et al., 1990).

In vitro expression of celB. The 2.0 kb Sphl fragment (2 μg) containing celB was electrophoresed from an agarose gel and expressed in vitro using an E. coli S30 coupled transcription-translation system (Promega). The translation product was labelled with ³⁵S)methionine (1200 Ci/nmol⁻¹; 44-4 TBq mmol⁻¹) according to the manufacturer’s instructions. Aliquots (5–15 μl), together with ³⁴C-labelled protein molecular mass markers, were subjected to SDS-PAGE (8%, w/v, acrylamide). The gel was fixed and soaked in Amplify (Amersham), dried, and exposed to X-ray film.

Determination of the N-terminal amino acid sequence of CelB. Expression of celB was induced by heat in PGM1 cells (10¹ culture) containing plasmids pTB1 and pGP1-2. Cells were harvested by centrifugation, washed twice in 60 ml buffer 1 (20 mM Tris/HCl, 1 mM Na₄EDTA, 1 mM dithiothreitol, 10%, v/v, glycerol, and 100 mM KCl, pH 7.4) and resuspended...
in 8 ml buffer 2 (buffer supplemented with 50 mg phenylmethylsulfonyl fluoride ml⁻¹). All remaining operations were done at 0-4 °C. Cells were disrupted by sonication and the cell debris was removed by centrifugation (12000 g, 10 min), the protein extract was concentrated 20-fold by freeze-drying, and then subjected to SDS-PAGE, as described above. The proteins were transferred to an Immobilon polyvinylidene difluoride membrane (Millipore) by Western blotting, using a buffer containing Tris-glycine. After staining the membrane with Coomassie Brilliant Blue R-250, a band with the expected molecular mass was excised and used directly for amino acid sequence analysis in a model 477A protein sequencer from Applied Biosystems.

**Fig. 1.** Nucleotide sequence of the cloned 2.0 kb Sphl fragment and the deduced amino acid sequence of CellB. The start and stop codons and a putative Shine-Dalgarno sequence are underlined and double-underlined, respectively. Residues identified by N-terminal amino acid sequencing are marked with asterisks.
sonication, and cell debris was removed by centrifugation (10000 g, 1 min). Aliquots (1 ml) of the supernatants were used directly for amylose analysis by adding 0.1 ml staining solution (0.1% I₂, 10% KI). The presence of amylose was detected as appearance of a deep blue colour. The experiments were done three times. Quantitative measurements were done in one of these experiments by measuring the absorbance at 610 nm. Solutions containing pure amylose and the staining mixture were used as standards.

RESULTS AND DISCUSSION

Localization of celB within the cloned fragment

Plasmid pUC18.8.6B.dE was previously shown to express the A. xylinum celB gene in E. coli (Fjervik et al., 1991). Deletion derivatives of the 3.8 kb DNA insert in this plasmid were constructed, and enzyme assays of PGM1 cells containing such plasmid derivatives indicated that a 20 kb Spφl fragment encoded the phosphoglucomutase activity. This was confirmed by enzyme assays of PGM1 cells containing pTB2A and pTB2B, which demonstrated that celB was expressed from both plasmids (data not shown). pTB2A expressed a 10-fold higher phosphoglucomutase activity than pTB2B, indicating that the orientation of the Spφl fragment in pTB2A allowed celB to be transcribed from the lac promoter in the vector.

Identification of the celB coding sequence

The molecular masses of phosphoglucomutases from several organisms have been determined; they range from 60 to 65 kDa (Ray & Peck, 1972). This indicates that the coding sequences for the corresponding genes are between 16 and 18 kb, and we therefore decided to sequence the entire 20 kb Spφl fragment. Prior to sequencing, the fragment was subcloned in both orientations into the corresponding site in the plasmid expression vector pT7-4S, generating plasmids pTB3A and pTB3B. The orientation of the Spφl fragment relative to the lac/φ10 promoters was found to be the same in pTB2A and pTB3A. This indicates that the 5' end of celB in pTB3A is localized proximal to the φ10 promoter. The nucleotide sequence of the Spφl fragment was determined, and the sequence data revealed an open reading frame (ORF1) starting with ATG at nucleotides (nt) 145–147 and terminating with TGA at nt 1810–1812 (Fig. 1).

Inserts cloned in the polylinker of pT7-4S can be transcribed from the φ10 promoter, which is specifically recognized by the bacteriophage T7 RNA polymerase (encoded by pGP1-2). In order to identify the biologically active start codon we performed an expression analysis of three deletion derivatives of pTB3A (pTB1, pTB8 and pTB16). The results of this analysis (Table 2) demonstrated that pTB1 and pTB16 expressed celB, and that the expression was strongly stimulated by T7 RNA polymerase. Very little or no activity was detected in cells containing pTB8. There is no Shine–Dalgarno (SD) sequence in the φ10-polylinker region in pT7-4S, and the results thus indicated the presence of a biologically active SD sequence in the inserts of pTB1 and pTB6. A putative SD sequence (GGAG) was identified upstream of ORF1 at nt 135–138 (Fig. 1). These experiments thus indicated that the ATG at nt 145–147 in ORF1 represents the biologically active start codon of celB. ORF1 corresponds to a polypeptide consisting of 555 amino acids and the calculated molecular mass of the protein is 59.6 kDa. This is consistent with the results of an in vitro expression analysis of celB, which demonstrated that the 20 kb Spφl fragment directed the synthesis of a protein with a molecular mass of approximately 60 kDa (Fig. 2).

The purified CelB protein obtained by in vitro expression in E. coli was sequenced at the amino-terminal end; in eight of the eleven first sequencing steps the amino acids could be definitely identified. All these amino acids are consistent with the deduced sequence of ORF1 (Fig. 1), with the exception that they were all assigned a position one residue number lower than predicted from the nucleotide sequence. The results thus indicated that the initiating
formylmethionine is removed from CelB when expressed in *E. coli*.

**Effect of CelB on intracellular amylose synthesis in *E. coli***

The *in vivo* effect of CelB on intracellular amylose synthesis in *E. coli* was analysed by growing cells in the presence of glucose or maltose as carbon sources. In agreement with previous observations (Adhya & Schwartz, 1971), it was found that PGM1 accumulated amylose [0.14 mg (mg dry wt$^{-1}$)] when grown on maltose, while no amylose was detected when glucose was used as carbon source. PGM1 did not accumulate amylose when grown under these conditions, and the same was true for the isogenic wild-type strain of PGM1 (Hfr3000). These experiments thus indicated that CelB could substitute for the *E. coli* enzyme in transforming the glucose 1-phosphate formed in the cells during growth on maltose to glucose 6-phosphate. In the biosynthesis of cellulose in *A. xylinum*, CelB is essential for the formation of glucose 1-phosphate. The ability of CelB to catalyse both directions of the glucose phosphate isomerization *in vivo* is thus demonstrated.

**Structural and biochemical comparisons between CelB and other related enzymes***

The amino acid sequence of CelB was compared with the primary structures of previously reported phosphoglucomutases and phosphomannomutases (Fig. 3). The polypeptide chains of CelB and the human and rabbit phosphoglucomutases are about 100 amino acids longer than the phosphomannomutases. Although there was no overall sequence similarity, the analysis revealed two highly conserved regions (Box I and II). The Box I motif contains the catalytic site of the enzymes, which was originally identified in rabbit muscle phosphoglucomutase (Ray *et al.*, 1983); AlgC, *P. aeruginosa* phosphomannomutase (Zielinski *et al.*, 1991); XanA, *X. campestris* phosphogluco-mannomutase (Koplin *et al.*, 1992); CpsG (Se), *Salmonella enterica* phosphomannomutase (Stevenson *et al.*, 1991); RfbK (St), *Salmonella typhimurium* phosphomannomutase (Jiang *et al.*, 1991); RfbK (Ec), *E. coli* phosphomannomutase (Marolda & Valvano, 1993).
The third region of homology (Box III, not reported previously) was found only in CelB and the two eukaryotic (human and rabbit) phosphoglucomutases. This motif consists of 39 amino acids, and 38 of these are the same in the rabbit and human enzymes, while there are 16 matches (41%) between CelB and the two eukaryotic sequences. In the rabbit muscle enzyme three amino acids (ESF) at the end of this motif are exposed in the active-site cavity, and it is suggested that they interact with the substrate glucose ring (Dai et al., 1992).

Among all the phosphoglucomutases previously reported, the X. campestris enzyme (XanA) is unique in that it can efficiently use both phosphorylated mannose and glucose as substrates. We have analysed CelB with respect to this property, and the results of these experiments showed that the activity with mannose 1-phosphate was less than 0.2% of the activity with glucose 1-phosphate. CelB is in this respect similar to the rabbit and human phosphoglucomutases, which have been reported to be specific for glucose phosphates (Lowry & Passonneau, 1969). The presence of the Box III motif only in these three enzymes suggests the possibility that this motif is responsible for their phosphoglucose specificity, and that the enzymes lacking this motif may have a broader substrate specificity. Except for XanA, the ability to convert phosphoglucose has to our knowledge not been determined for the phosphomannomutases included in this comparison.

Acknowledgements

We thank Dr Knut Sletten for his help with the amino acid sequencing.

References


Received 27 September 1993; revised 11 November 1993; accepted 29 November 1993.