Int. J. Indust. Entomol. Vol. 14, No. 2, 2007, pp. 69~74

International Journal of *Industrial Entomology*

Cloning and Characterization of 6-Phosphogluconolactonase Gene in Silkworm *Bombyx mori*

HuaJun Yang, KePing Chen*, Qin Yao and ZhongJian Guo

Institute of Life Sciences, Jiangsu University, 301 Xuefu Road, Zhenjiang 212013, P. R. China

(Received 11 January 2007; Accepted 2 March 2007)

As the genome of *B. mori* is available in GenBank and the EST database of *B. mori* is expanding, identification of novel genes of *B. mori* was conceivable by data-mining techniques and bioinformatics tools. In this study, we used the *in silico* cloning method to get the 6-Phosphogluconolactonase (6PGL) gene of *B. mori* and analysed with bioinformatics tools. The result was confirmed by RT-PCR and prokaryotic expression. The 6PGL cDNA comtains a 702 bp ORF. The deduced protein has 233 amino acid residues, with the predicted molecular weight of 25946. 72 Da , isoelectric point of 5. 41, and contains conserved NagB domains. This gene has been registered in GenBank under the accession number EF198104.

Key words: 6PGL, *Bombyx mori*, Bioinformatics, RT-PCR, Prokaryotic expression

Introduction

The pentose phosphate pathway (PPP) is an important part of the central metabolism in many organisms. The oxidative branch of PPP providing reducing power, and phosphorylated carbohydrates obtained in the non-oxidative branch of PPP, are precursors for nucleotide biosynthesis, aromatic amino acids and vitamins[1]. The second step of the pentose phosphate pathway is the hydrolysis of 6-phosphogluconolactone, a spontaneous reaction that is greatly accelerated by a specific 6-phosphogluconolactonase widely distributed in organism. The enzyme appearing to be a monomer of 26 to 30 kDa, in contrast to other lactonases, does not depend on Mg^{2+} for its activity[2]. There are two kinds of 6-phosphogluconolactone, δ and γ form. The δ form is the only substrate for 6-phosphogluconolactonase. Therefore, 6-phosphogluconolactonase activity accelerates hydrolysis of the δ form, thus preventing it convert into the γ form. Furthermore, 6-phosphogluconolactonase guards against the accumulation of γ-6-phosphogluconolactone, which may be toxic through its reaction with endogenous cellular nucleophiles[3]. 6PGL is an essential enzyme of the cyclic Entner-Doudoroff pathway in the *Pseudomonas aeruginosa*[4].

In the paper, we cloned the 6-phosphogluconolactonase of *B. mori* using data-mining techniques and analysed with bioinformatics tools, including its genomic organization and the deduced amino acid sequence. The gene was confirmd by RT-PCR and prokaryotic expression. The deduced protein sequence was compared with that of some homologous protein from other species.

*To whom the correspondence addressed
Institute of Life Sciences, Jiangsu University, 301# Xuefu Road, Zhenjiang 212013, P. R. China. Tel: 86-511-8791923; Fax: 86-511-8791923; E-mail: kpchen@ujs.edu.cn

Materials and methods

Materials

The silkworm *Bombyx mori* and *E. coli* (strain TG1 and BL21) were inbred in our lab. Silkworm strain 306 were used for this study. Restrictases, T4 DNA ligase, PCR reagents and pMD18-T were obtained from TaKaRa Company (Dalian); primer and other reagents were obtaind from Shanghai Sangon Bio-technology Corpotation.

Data extraction of cDNA sequence of *B. mori* 6PGL gene

The NCBI's (http://www.ncbi.nlm.nih.gov/) EST database is a popular starting point for identifying expressed sequence tags (ESTs) of different species, and more than 110, 000 *B. mori* EST sequences are currently available in GenBank. We also used another silkworm cDNA database BGI (http://silkworm.genomics.org.cn/) in this study.

Data-mining techniques and bioinformatics tools were applied to search for cDNA sequence of the *B. mori* 6PGL gene by repeated cycles of assembling and extending EST sequences.

Genome analysis

In order to establish the genomic organization, the cDNA sequence was BLASTed to the contigs of *B. mori* genome in GenBank. SIM4 (http://pbil.univ-lyon1.fr/sim4.php) was used to align the cDNA sequence with the genomic sequences to search potential introns. Splice Site Prediction by Neural Network (NNSSP) (http://www.fruitfly.org/seqtools/splice.html) was used to predict the potential splice sites.

Protein prediction and analysis

We used the ExPASy Translate tool (http://au. expasy. org/tools/dna. html) to deduce the cDNA's amino acid sequence, and similarity analysis was performed using the BLAST tool in GenBank. We used PLOC (http://www.genome.jp/SIT/plocdir/) to predict the subcellular location of *B. mori* 6PGL. Using InterPro Scan (http://www.ebi.ac.uk/InterProScan/) to analyse the deduced amino acid sequence.

RT-PCR

A pair of specific primers were designed based on the sequence we obtained. The forward primer (5'GGAATTCATGACGATCATCAAAG 3') contained a EcoR I restriction site (underlined), and the reverse primer (5' GAAGCTTACCCACAGAATAGCAC 3') contained a Hind III restriction site (underlined). cDNA was prepared from midgut RNA with M-MLV reverse transcriptase and an oligodT primer. PCR reaction was carried out with Taq polymerase for 35 amplification cycles (94°C/45 sec, 53°C/45 sec, 72°C/1 min). PCR product was examined by electrphoresis in 1% agarose gel with the ethidium bromide staining.

Cloning and Sequencing

The PCR product was ligated into pMD18-T vector using T4 DNA ligase and then transformed into *E. coli* (TG1 strain). Plasmid was purified with MiniBEST Plasmid Purification Kit (Takara). The sequencing was performed using an automatic sequence: CEQ8000(Beckman company).

Construction of expression plasmid

The plasmid pMD18-T/6PGL was digested with EcoRI and Hind III, and then purified. The purified fragment was ligated with the EcoRI-Hind III digested pET30a vector and transformed into *E. coli* (BL21 strain). The transformants harboring the recombinant plasmid were confirmed by restriction enzyme analysis.

Expression of fusion protein in *E. coli* and SDS-PAGE

For expression of recombinant protein, a positive clone was cultured in LB medium supplement with Kanamycin (50 μg/ml) overnight at 37°C with shaking. This culture was added into fresh LB medium and cultured at 37°C with vigorous shaking to A_{600} about 0.6. The culture was then induced with IPTG (final concentration of 0.2 mmol/L) and further cultured for another 10 hours. 15% SDS polyacrylamide gel was used to analyze the recombinant protein. SDS-PAGE was performed in the Mini-Protein system (Bio-Rad, USA). After electrophoresis, the gel was stained with Coomassie Brilliant Blue R250 to visualize the protein bands.

Result

Nucleotide sequence analysis

A 1294bp of full-length cDNA was obtained by data-mining techniques. The cDNA sequence included an open reading frame, beginning with the initiation codon ATG at position 123 and ending with a termination codon TAA at position 824. The cDNA also included a TATA£|box and PolyA signal. The nucleotide sequence of the full-length 6PGL cDNA is shown in Fig. 1.

BLASTing the cDNA to contigs of *B. mori* genome in GenBank revealed that contig21251, contig186799 and contig44358 having a high similarity. After analyzing these three contigs, we found that contig186799 and contig44358 were a part of contig21251, and that the *B. mori* 6PGL gene was a single copy gene in the whole genome. Using SIM4, four extrons were found in the relevant DNA sequence. The length of the exons were 118 bp, 201 bp, 281 bp and 102 bp. Splicing signals (exon /GU-intron-AG/exon) were also identified by NNSSP and were shown in Fig. 2.

Cloning and identification of 6PGL

The cDNA of midgut as the tamplate, PCR amplification of the cDNA was performed using the twe specific primers. After electrphoresis in 1% agarose gel,we obtained a band about 700 bp, which was consistent with the expected molecular mass (Fig. 3). The PCR product was ligated into pMD18-T vector and confirmed by restriction endonuclease digestion and DNA sequencing. The 6PGL fragment could be isolated from the pMD18-T vector after the recombinant plasmid was digested with EcoRI and Hind III (data not shown). Sequencing result indicated that the sequence was the 6PGL sequence.

```
-122                                                              CG
     AATTCATGATAAGACTCGGCAGTGATAATATTATCCTATAATTACTATCAAGAAAAACAA
     AAGTTTTTGTATATTTCTATGAATTATACTATTTTCTGAATTATACTAGTTTAAATTACA
   1 ATGACGATCATCAAAGTTATTGATGAAGAAGAAATCATCAGAAAACTATCGACATACATT
     M  T  I  I  K  V  I  D  E  E  E  I  I  R  K  L  S  T  Y  I
  61 CAAAAGATTTCAAACGACGCGATCTTGAATAGAAATAAATTCGTAGTTGGACTTTCTGGC
     Q  K  I  S  N  D  A  I  L  N  R  N  K  F  V  V  G  L  S  G
 121 GGCTCTGTCGTCAAATATTTATGTGAAGGTTTACCACAAGTTGAAACAGACTGGTCTAAA
     G  S  V  V  K  Y  L  C  E  G  L  P  Q  V  E  T  D  W  S  K
 181 TGGACTCTAGCCTTCTGTGATGAAAGAGTAGTACCTGAAGACAGCAGCGATTCTACATTC
     W  T  L  A  F  C  D  E  R  V  V  P  E  D  S  S  D  S  T  F
 241 GGAATATACAAGAAAGATTTAATCCCTAAGACTGAATTAAAGGAGAGCCAGTTTATAACA
     G  I  Y  K  K  D  L  I  P  K  T  E  L  K  E  S  Q  F  I  T
 301 ATCAAACAAGGAGCTACAGCTCAAGAAACAGCAAAGGATTACATTGAAAAGTTAAGAAAA
     I  K  Q  G  A  T  A  Q  E  T  A  K  D  Y  I  E  K  L  R  K
 361 GTCTTTGGTGGAGATGATTTCAAATTTGACTTACTTCTACTTGGCATGGGACCAGATGGA
     V  F  G  G  D  D  F  K  F  D  L  L  L  L  G  M  G  P  D  G
 421 CACACATGCTCCCTGTTCCCAGGTCACAAATGGCTTGAAGAAACAGAAGATAAGGTAGCA
     H  T  C  S  L  F  P  G  H  K  W  L  E  E  T  E  D  K  V  A
 481 GCTATAACAGATTCCCCAAAACCTCCACCGGAGAGAATCACATTAACTTATCCAGTTATA
     A  I  T  D  S  P  K  P  P  P  E  R  I  T  L  T  Y  P  V  I
 541 AATGGAGCAAGGAACTGTATTTTGGCTATTTCTGGGGCAGGGAAGTCAGAAAGGGCTAAA
     N  G  A  R  N  C  I  L  A  I  S  G  A  G  K  S  E  R  A  K
 601 CGTATTCTAAAAGACAAAGAAGATTTACCAGCTGGTCGAGTGAAGCCGACAGACGGCTCT
     R  I  L  K  D  K  E  D  L  P  A  G  R  V  K  P  T  D  G  S
 661 CTATATGGGATCCTGGACGAAGAAGCAGCTGTACATTGGTAAAATATGACTAGGTCATGA
     L  Y  G  I  L  D  E  E  A  A  V  H  W  .
 721 CTGCAAGGAAAAAACTTGTGCTATTCTGTGGGTATCACTCATCATGTAAAAACAGGGCAG
 781 ATCATGCAAATTCATTAAAAAAAATCAACTTCTTGAGGTTTTAAATATTATCTGGAAAAA
 841 AAACAAGAAACATAATTACTTCAGTTGCTGTAATCTCCGATGATGAACACATCCTACCAG
 901 ATGGGCCCACAGGGTGTGGTATGTATGCCGGTCTGTCTACAAAGATAACAAAAACGAATA
 961 CGAGATTGTAGTTAAAAACATGACTGTCTCATAGCAAAATCACACAGGATAGTCTTGGAT
1021 ACTTTTTTTGTCCTAAAAACTGCTTTAGAAAATAAATGAAAAAATATAAAAATAATGATA
1081 TAATATTGTGCAGCTTTTTGTCAATGTTGTGTATTGATCAGTTCATTTAAATAAGGAAGG
1141 TTCAATTATCACCTACATTTCAAAAATAAAAA
```

Fig. 1. Nucleotide sequence and deduced amino acid sequence of the 6PGL gene. The predicted amino acid is represented by the one letter code designation below the nucleotide sequence. The initiate codes and the stop codes are framed.The TATA-box is italic and underlined. PolyA signal are underlined.

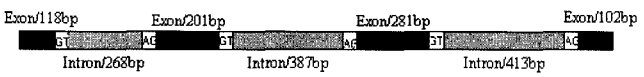

Fig. 2. DNA sequence frame of the 6PGL gene. Extrons are black framed, and introns are gray framed. The splicing signals (exon/GT-intron-AG/exon) are also indicated.

Analysis of the deduced amino acid sequence

The open reading frame encoded a protein of 233 amino acids, as deduced from the nucleotide sequence (Fig. 1). The molecular weight of the encoded protein was predicted to be 25946.72 Da with an isoelectric point of 5.41.

The deduced amino acid sequence was analyzed using InterPro Scan software in the ExPASy website. The results indicated that the deduced amino acid sequence had a domain named 6-Phosphogluconolactonase (IPR 005900). The domain analysis indicated that the encoded protein was probably 6-Phosphogluconolactonase. The subcellular location of *B. mori* 6PGL was predicted in the cytoplasm. The deduced protein was BLASTed in the NCBI, we got a Conserved NagB domain (Fig. 4).

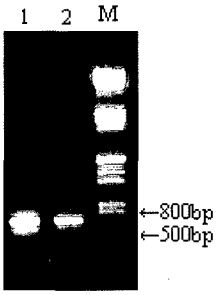

Fig. 3. The result of RT-PCR for 6PGL gene. 1, 2 lane. RT-PCR product M. DNA moleculor mass maker.

Fig. 4. Conserved NagB domains.

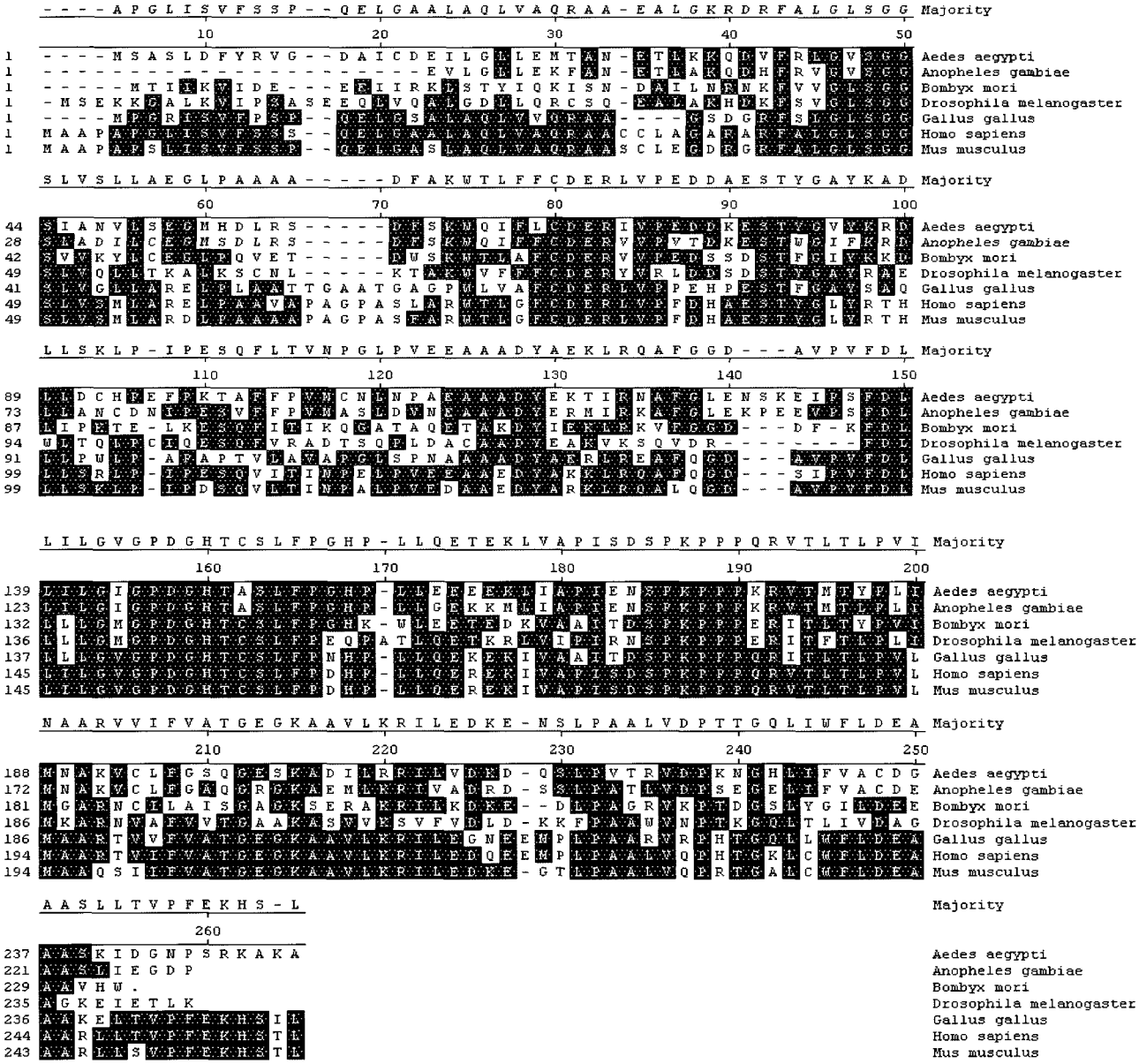

Fig. 5. Alignment of the sequence of B. mori 6PGL with corresponding protein from Aedes aegypti (EAT33522), Anopheles gambiae (EAA05555), Drosophila melanogaster (AAF47963), Gallus gallus (NP_001026759), Homo sapiens (NP_036220) and Mus musculus (NP_079672). The sequences were aligned using DNAstar CLUSTAL W program.

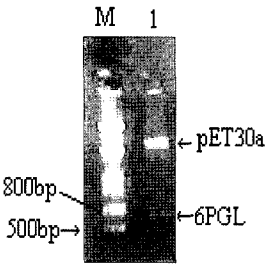

Fig. 6. Identification of the recombinant plasmid pET30a/6PGL. 1. pET30a/6PGL; M. DNA moleculor mass maker.

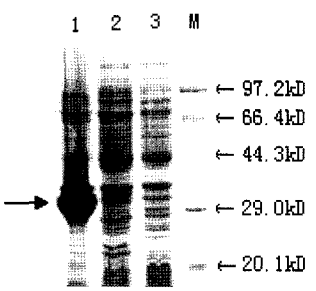

Fig. 7. SDS-PAGE of fusion protein. 1. Protein of E. coli BL21 contained pET30a/6PGL induced by IPTG; 2.Protein of E. coli BL21 induced by IPTG; 3.Protein of E. coli BL21 contained pET30a induced by IPTG; M.Protein marker.

Using BLAST software of NCBI to search for homology in the GenBank database, the deduced amino acid sequence showed an identity of 43%, 44%, 46%, 47%, 50% and 49% to the corresponding genes of *Aedes aegypti* (EAT33522), *Anopheles gambiae* (EAA05555), *Drosophila melanogaster* (AAF47963), *Gallus gallus* (NP_001026759), *Homo sapiens* (NP_036220) and *Mus musculus* (NP_079672) respectively. Comparison the deduced amino acid sequence with that of the corresponding genes of the above six species is shown in Fig. 5. Alignment of several such members revealed important regions of conservation (motifs) that may include the active-site residues of 6-phosphogluconolactonase. Some conserved regions include SGGS, CDER, FDLL, LG-GPDGHT-SLFP, SPKPPP.

Construction of expression plasmid

The plasmid pMD18-T/6PGL was digested with EcoRI and Hind III, and then the fragment was ligated with the EcoRI-Hind III digested pET30a vector. The 6PGL fragment could be isolated from the pET30a vector after the recombinant plasmid was digested with EcoR I and Hind III. The expression plasmid pET30a/6PGL was successfully constructed (Fig. 6).

Expression of fusion protein in *E. coli* and SDS-PAGE

IPTG induced the *E. coli* BL21 transformed with the pET30a/6PGL plasmid to express the His-6PGL recombinant fusion protein of about 30 kDa (Fig. 7), which was consistent with the expected molecular mass of the fusion protein of pET30a/6PGL.

Discussion

Silkworm (Bombyx mori) is an important economic insect and is regarded as a model insect of Lepidoptera. Studies on structures and functions of certain related genes in *B. mori*, has attracted more and more attention. 6-Phosphogluconolactonase is the second enzyme in the pentose-P pathway and converts 6-phosphogluconolactone to 6-phosphogluconate. 6-phosphogluconate is used as a substrate for both the pentose-P pathway and the inducible pathway for catabolism of gluconate, the Entner-Doudoroff pathway. Thus, the 6PGL enzyme has the potential to contribute a metabolite to two pathways[1]. Interestingly, in some eukaryote 6PGL and Glucose-6-phosphate dehydrogenases (G6PD) activities are both present in a single bifunctional enzyme[10,11].

In this study, we identified a novel *B. mori* 6PGL gene through bioinformatics approaches. We found that this gene locates within contig21251 of *B. mori* genome, and encodes 233 amino acids. Through amino acid analysis, it was found that the sequence of *B. mori* 6PGL has a high homology with the 6PGL sequences of other insect species. Some conserved regions include SGGS, CDER, FDLL, LG-GPDGHT-SLFP, SPKPPP that may include the active-site residues of 6PGL. Curiously the LG-GPDGHT-SLFP region was very similar to a region of the rabbit microsomal glucose-6-phosphate dehydrogenase (residues 399 to 411 of the rabbit microsomal glucose-6-phosphate dehydrogenase are GMGtDGHTASLFP). Also this regiong contains a consensus coenzyme-binding motif GxGxxG/A[12]. However some glucose-6-phosphate dehydrogenase also contains a 6PGL domain,for example: the unusual glucose- 6-phosphate dehydrogenase of the *Plasmodium falciparum* contains a 6PGL domain.

Some animals have evolved a simple solution for efficient metabolic flux through glucose-6-phosphate dehydrogenase and 6-phosphogluconolactonase by combining both activities into a single protein, which can catalyses the first two steps of the PPP. Whether the *B. mori* has a single bifunctional enzyme, which has 6PGL and G6PD activities still needs further investigation.

Acknowledgments

This work was supported by a grant from Academy Nat-

ural Science Foundation (BK2006074) and National Basic Research Program of China under grant (2005CB 121000)

Reference

Bauer, H. P., Srihari, T., Jochims, J. C. and Hofer, H. W. (1983) 6-Phosphogluconolactonase: purification, properties and activities in various tissues. Eur. J. Biochem., **133**, 163-168.

Clarke, J. L., D. A. Scopes, O. Sodeinde and P. J. Mason. (2001) Glucose-6-phosphate dehydrogenase-6-phosphogluconolactonase. A novel bifunctional enzyme in *malaria parasites*. Eur. J. Biochem., **268**, 2013-2019.

Clarke, J. L. and P. J. Mason. (2003) Murine hexose-6-phosphate dehydrogenase: a bifunctional enzyme with broad substrate specificity and 6-phosphogluconolactonase activity. Arch. Biochem. Biophys., **415**, 229-234.

Collard, F., J. Collet, I. Gerin, M. Veiga-da-Cunha and E. Van Schaftingen. (1999) Identification of the cDNA encoding human 6-phosphogluconolactonase,the enzyme catalyzing the second step of the pentose phosphate pathway.FEBS Lett., **459**, 223-226.

Duffieux, F., Van Roy, J., Michels, P. A. and Opperdoes, F. R. (2000) Molecular characterization of the first two enzymes of the pentose-phosphate pathway of *Trypanosoma brucei*. Glucose-6-phosphate dehydrogenase and 6-phosphogluconolactonase. J. Biol. Chem., **275**, 27559-27565.

Hager, P. W., M. W. Calfee, and P. V. Phibbs. (2000) The *Pseudomonas aeruginosa devB/SOL* homolog, *pgl*, is a member of the *hex* regulon and encodes 6-phosphogluconolactonase. J.Bacteriol., **182**, 3934–3941.

Kupor, S. R. and D. G. Fraenkel. (1972) Glucose metabolism in 6 phosphogluconolactonase mutants of *Escherichia coli*. J. Biol. Chem., **247**, 1904-1910.

Kupor, S. R. and D. G. Fraenkel. (1969) 6-Phosphogluconolactonase mutants of *Escherichia coli* and a maltose blue gene. J. Bacteriol., **100**, 1296-1301

Miclet, E., Stoven, V., Michels, P. A., Opperdoes, F. R., Lallemand, J. Y. and Duffieux, F. (2001) NMR spectroscopic analysis of the first two steps of the pentose-phosphate pathway elucidates the role of 6-phosphogluconolactonase. J. Biol. Chem. **276**, 34840-34846.

Ozols, J. (1993) Isolation and the complete amino acid sequence of lumenal endoplasmic reticulum glucose-6-phosphate dehydrogenase. Proc. Natl. Acad. Sci. USA, **90**, 5302-5306.

Thomason, L. C., Court, D. L., Datta, A. R., Khanna, R. and Rosner, J. L. (2004) Identification of the *Escherichia coli* K-12 *ybhE* gene as *pgl*, encoding 6-phosphogluconolactonase. J. Bacteriol., **186**, 8248-8253.

Zimenkov, D., Gulevich, A., Skorokhodova, A., Biriukova, I., Kozlov, Y. and Mashko, S. (2005) *Escherichia coli* ORF *ybhE* is *pgl* gene encoding 6-phosphogluconolactonase (EC 3.1.1.31) that has no homology with known 6PGLs from other organisms. *FEMS Microbiol Lett*, **244**, 275-280.