Cloning of the Transketolase Gene from Erythritol-Producing Yeast Candida magnoliae

Boung-Hyuk Yoo†, Eun-Hee Park†, Jin-Ho Seo2, and Myoung-Dong Kim1*

1Department of Food Science and Biotechnology, Kangwon National University, Chuncheon 200-701, Republic of Korea
2Department of Agricultural Biotechnology and Center for Food and Bioconvergence, Seoul National University, Seoul 151-921, Republic of Korea

Introduction

Erythritol (1,2,3,4-butanetetrol; molecular weight (MW) 122.12), a four-carbon sugar alcohol, occurs naturally in a variety of foods such as fruits, mushrooms, and fermented foods, including cheese, wine, and soy sauce [8]. Because of its low energy value, non-insulin stimulant properties, and excellent taste, erythritol is used as a pharmaceutical excipient and noncaloric sweetener [34]. Most ingested erythritol is not metabolized by the human body and is excreted unchanged in the urine without altering blood glucose and insulin levels. It also prevents dental caries because the bacteria that cause dental caries are not able to utilize erythritol as a carbon source [24].

Erythritol is produced routinely by microbial fermentation, not by chemical synthesis processes such as hydrogenation at high temperature and pressure under a nickel catalyst [29]. It is known that osmophilic bacteria, fungi, and yeast can produce erythritol [9, 16, 17, 27, 48], and a number of studies have been carried out to obtain high-production organisms and to improve fermentation methods for high-efficiency production [4, 33]. Erythritol has been produced commercially using the Aureobasidium sp. SN-C42 that produces erythritol in high yield [13]. A high erythritol-producing yeast was isolated from honeycombs and identified as Candida magnoliae JH110 [17]. Various biological processes using C. magnoliae JH110 and its mutant have been developed to improve erythritol production [20, 33, 49].

Although it has been generally thought that erythrose-4-phosphate is biosynthesized through the pentose phosphate pathway, there have been few studies regarding the biosynthetic pathway of erythritol [43]. Tokuoka et al. [47] reported that erythrose reductase (ER), the second group of aldose reductase (E.C. 1.1.1.21), in Aureobasidium sp. converts erythrose to erythritol, and that there are three isozymes. Ookura et al. [28] determined the gene sequences of the three isozymes of erythrose reductase in Trichosporonoides megachiliensis SNG-42.

Recently, a complete gene sequence of the C. magnoliae ER was determined and its enzymatic properties were examined after heterologous expression in recombinant Escherichia coli [22]. Although ER catalyzing the final

The entire nucleotide sequence of the TKL1 gene encoding transketolase (TKL) in an erythritol-producing yeast of Candida magnoliae was determined by degenerate polymerase chain reaction and genome walking. Sequence analysis revealed an open reading frame of C. magnoliae TKL1 (CmTKL1) that spans 2,088 bp and encodes 696 amino acids, sharing 61.7% amino acid identity to Kluyveromyces lactis TKL. Functional analysis showed that CmTKL1 complemented a Saccharomyces cerevisiae tkl1 tkl2 double mutant for growth in the absence of aromatic amino acids and restored transketolase activity in this mutant. An enzyme activity assay and RT-PCR revealed that the expression of CmTKL1 is induced by fructose, H2O2, and KCl. The GenBank accession number for C. magnoliae TKL1 is KF751756.

Keywords: Candida magnoliae, transketolase, degenerate PCR, genome walking
reaction might play a crucial role in synthesizing erythritol, the correlation between erythritol productivity and enzyme activities in the biosynthetic pathways have not been studied sufficiently.

Transketolase (TKL, E.C. 2.2.1.1) is a ubiquitously occurring enzyme that catalyzes the reversible transfer of a glyceraldehyde group from an activated ketose (donor) to an activated aldose (acceptor) [37]. The enzyme functions as a homodimer and requires thiamine pyrophosphate as a co-organ to generate reducing power by transferring a ketol to an aldose in the pentose phosphate pathway to glycolysis to generate reducing power in the form of NADPH. TKL catalyzes glyceraldehyde-3-phosphate and fructose-6-phosphate or sedoheptulose-7-phosphate to xylulose-5-phosphate and erythrose-4-phosphate, which might be a prominent intermediate for erythritol biosynthesis or ribose-5-phosphate [19]. Either the immediate substrates or the products of the reaction are the precursors for the synthesis of nucleic acids, aromatic amino acids, and vitamins [2, 19]. Sawada et al. [36] evaluated the enzyme activities of the pentose phosphate pathway in *T. megachiliensis* SN-G42 to reveal the crucial role of transketolase in erythritol production. The activity of TKL, which is dependent upon thiamine, correlated well with erythritol productivity under different cultivation conditions.

The present study describes the isolation and sequence analysis of the *TKL* gene from an erythritol-producing yeast of *C. magnoliae*. The functionality of the cloned gene was confirmed by complementation study and TKL enzyme activity assay. In addition, the effects of a variety of stress conditions on the expression and enzyme activity of TKL in *C. magnoliae* were examined.

**Materials and Methods**

**Strains and Culture Conditions**

*C. magnoliae* KFCC 01900 [21] and a *tkl1 tkl2* mutant of *Saccharomyces cerevisiae* BY4742 [Mata his3Δ1 leu2Δ0 lys2Δ0 ura3Δ0 tkl1Δ:kanMX4 tkl2Δ:Ura3Δ3] (lab stock) were used in this study. *Escherichia coli* TOP10 (Invitrogen, USA) was used for plasmid DNA preparation and was routinely grown at 37°C. LB medium (0.5% bacto-yeast extract, 1% bacto-tryptone, and 1% NaCl) supplemented with 100 µg/ml ampicillin was used for plasmid DNA preparation. *C. magnoliae* was grown in YEPD medium at 30°C to an OD (optical density) of 0.6 harvested by filtration, transferred to YEPD medium at 30°C and then harvested by filtration [41] for TKL enzyme assay.

**Nucleic Acid Isolation and Plasmids**

The genomic DNA of *C. magnoliae* was prepared from yeast cells in the exponential growth phase, as described previously [30]. Plasmid DNA was isolated and purified using the AccuPrep Plasmid Mini Extraction Kit (Bioneer, Korea). Plasmids pGEM-T (Promega, USA) and p415GPD (ATCC, USA) were used for general cloning and the complementation study, respectively.

**Degenerate PCR and Genome Walking**

The two degenerate oligonucleotide primers, TK-For and TK-Rev (Table 1), were designed based on the core consensus conserved regions, ANSGHPG and VTTGPLGQ, of TKLs reported in the literature [14, 15, 46]. PCRs were performed using 2.5 units of nTaq polymerase (Enzymomics, Korea), 1 µg of template DNA, and 100 µM of each degenerate primer. The reaction conditions were programmed as 10 min at 94°C followed by 35 cycles of denaturation for 1 min at 94°C, annealing for 1 min at 50~61°C, and extension for 1 min 30 sec at 72°C. The amplified fragment was cloned into the pGEM-T vector by TA cloning and sequenced. The complete open reading frame (ORF), and promoter and terminator regions were identified by genomic walking, which was performed with the DNA Walking SpeedUp Kit (Seegene, Korea) according to the manufacturer’s protocol.

**Sequence Analysis**

The nucleotide and deduced amino acid sequences were analyzed by BLAST (The National Center for Biotechnology Information, NCBI; http://www.ncbi.nlm.nih.gov/BLAST/). Alignment of multiple sequences was carried out by the SDSC (San Diego Supercomputer Center) Biology Workbench (http://workbench.sdsc.edu/). The theoretical molecular mass of the protein was calculated using ExPaSy (http://au.expasy.org/ compute_pi/). The conserved domain of the TKL was identified by InterProScan (http://www.ebi.ac.uk/ Tools/InterProScan/).

**Cloning of *CmTKL1***

A *LEU2*-marked plasmid, p415GPD (6.92 kb), was used to express *CmTKL1*. The entire ORF of *CmTKL1* was amplified by PCR using the primers CmTKL-F and CmTKL-R, shown in Table 1. The PCR product of the expected size was digested with the *XhoI* and *HindIII* restriction enzymes and then inserted into plasmid p415GPD to construct plasmid pME1330 (8.9 kb).
Table 1. Oligonucleotide primers used in this study.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5'-3')</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>TK-For</td>
<td>GCIAAYAGYGGICAYCCIGG</td>
<td>Degenerate PCR</td>
</tr>
<tr>
<td>TK-Rev</td>
<td>YTICCIARIGGICGCTG14TAC</td>
<td>Degenerate PCR</td>
</tr>
<tr>
<td>CmTK-up-TSP1</td>
<td>ATGTCGTAGCCGTAAGCTTGC</td>
<td>Genome walking</td>
</tr>
<tr>
<td>CmTK-up-TSP2</td>
<td>CTGGGACACGAGAAGGC</td>
<td>Genome walking</td>
</tr>
<tr>
<td>CmTK-up-TSP3</td>
<td>TCGGCTGAACCTTCATGCAGA</td>
<td>Genome walking</td>
</tr>
<tr>
<td>CmTK-down1-TSP1</td>
<td>TCAACGCGGACGCCTG</td>
<td>Genome walking</td>
</tr>
<tr>
<td>CmTK-down1-TSP2</td>
<td>GGCTACGACATCAGG</td>
<td>Genome walking</td>
</tr>
<tr>
<td>CmTK-down1-TSP3</td>
<td>CAGCGCGGTCCACCG</td>
<td>Genome walking</td>
</tr>
<tr>
<td>CmTK-down2-TSP1</td>
<td>GAGTACAGCTGTCG</td>
<td>Genome walking</td>
</tr>
<tr>
<td>CmTK-down2-TSP2</td>
<td>TGGAGCGAAGTACGCTG</td>
<td>Genome walking</td>
</tr>
<tr>
<td>CmTK-down2-TSP3</td>
<td>ACAAGGCGTCGACTTCG</td>
<td>Genome walking</td>
</tr>
<tr>
<td>CmTKL-RT-F</td>
<td>ATGTCTTCTCAGAAAGTTGC</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>CmTKL-RT-R</td>
<td>GTGACCTCTGATGACCGG</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>CmURA3-RT-F</td>
<td>GGATTTGCGTGTCG</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>CmURA3-RT-R</td>
<td>GGCCAACGATGAGTC</td>
<td>RT-PCR</td>
</tr>
<tr>
<td>CmTKL-F</td>
<td>GGCTCTTCTAAGCAGTCTTGM</td>
<td>pME1330 construction</td>
</tr>
<tr>
<td>CmTKL-R</td>
<td>GGCAAGCTTATTATGTGGT</td>
<td>pME1330 construction</td>
</tr>
</tbody>
</table>

Transformation

*S. cerevisiae* BY4742 was transformed by the lithium acetate method [7]. *E. coli* cells were transformed as described previously [35].

Complementation

*S. cerevisiae* BY4742 transformants harboring plasmid p415GPD (control) and pME 1330 (p415GPD-CmTKL1) were grown in SC-Leu− and their culture broth was diluted with sterile water to achieve an OD$_{600}$ of 0.2. The cell suspension taken by 5-fold serial dilutions was spotted onto a SC-Phe− Tyr− Trp− plate, which lacks phenylalanine, tyrosine, and tryptophan. The plates were incubated at 30°C and cell growth was assessed after 3 days.

Preparation of Cell Extracts

Harvested yeast cells were broken by vortexing with acid-washed glass beads (Sigma-Aldrich) in 1 ml of glycyglycine buffer (50 mM, pH 7.6) for 2 min [30]. Cell homogenates were centrifuged at 15,000 ×g for 5 min at 4°C, and the supernatants were used as cell extracts. Cell extracts were prepared from three independent cultures. Protein concentrations were determined by a Bio-Rad assay using bovine serum albumin as the standard.

TKL Enzyme Activity Assay

TKL enzyme activity was assayed as described in the literature [18]. The reaction mixture containing 50 mM glycyglycine buffer (pH 7.6), 2 mM MgCl$_2$, 0.1 mM thiamine pyrophosphate, 1.4 mM K$_2$Fe(CN)$_6$, and 3 mM fructose-6-phosphate was incubated for 5 min at 30°C and then protein extract was added to start the reaction for 10 min. A molecular extinction coefficient (1,000 M$^{-1}$ cm$^{-1}$) for ferricyanide was used. One unit of TKL activity was defined as the amount that catalyzes the conversion of 1 µmole of fructose-6-phosphate per minute at 30°C and pH 7.6.

RNA Isolation and RT-PCR

Total RNA was prepared using TRIzol reagent (Invitrogen). RT-PCR was performed using the cDNA TOPscript RT DryMIX (Enzymomics) according to the manufacturer’s instructions. Total RNA (1.0 µg) and full-length primers of CmURA3 [32] and CmTKL1 (Table 1) were used for 22 cycles of RT-PCR amplification. Aliquots of the PCR product were separated on a 1.0% agarose gel.

Statistical Analysis

Statistical analysis was performed using Student’s t-test [1] and based on at least three independent experiments. The results were considered to be statistically significant when $p < 0.05$.

Results and Discussion

Cloning of CmTKL1

Degenerate PCR primers (Table 1) targeting the conserved regions of the TKL1 gene were designed based on an alignment of published TKL amino acid sequences. A PCR product of approximately 300 bp in size was amplified at an annealing temperature of 51°C, TA-cloned, and then
sequenced (data not shown). Alignment of the deduced primary structure of the cloned gene product with other yeast TKL homologs indicated that the cloned gene contains a conserved sequence of the yeast TKLs. Based on the cloned gene sequence, nine gene-specific primers (listed in Table 1) were designed to further isolate the 5’ and 3’ regions of the CmTKL1 gene. The entire nucleotide sequence of the CmTKL1 gene, including approximately 350 bp of the 5’ and 3’ untranslated regions, was obtained after genome walking and sequence assembly and was deposited in GenBank (Accession No. KF751756).

**Sequence Analysis**

Analysis of the DNA sequence revealed that the CmTKL1 gene consisted of an uninterrupted open reading frame of 2,088 bp in size. The predicted amino acid sequence is 696 amino acids long, with an estimated molecular mass of 76.2 kDa. The putative transcriptional regulatory element, TGATTT [10], was found at 285 nucleotides upstream of the initiation codon in the 5’-flanking region. These elements are very similar to the conserved binding site, TGANTN, for the transcription factor Gcn4 in *S. cerevisiae* [11, 50].

Based on the sequence analysis of the 5’-flanking region, a putative stress response element (STRE, 5’-AGGGG-3’) found in HSP104, HSP26, DDR2, and GLO1 in *S. cerevisiae* was found at 46 nucleotides upstream of the initiation codon [39]. At 51 bp downstream of the TAA stop codon, a putative polyadenylation signal sequence, ATAAATTT (consensus AATAAAA) [5], was observed.

Alignment of the amino acid sequences revealed that the *C. magnoliae* TKL shared significant identity with other homologs (Fig. 1), with the highest homology being shared with the TKL of *Kluyveromyces lactis* (61.7%), followed by those of *Komagataella pastoris* (61.5%), *Candida tropicalis* (61.3%), and *Aspergillus niger* (60.6%). Sequence comparison also showed that *C. magnoliae* TKL contains a highly conserved thiamine diphosphate-binding domain at the N-terminus, which is similar to the structure of other TKLs [6].

The predicted open reading frame of *C. magnoliae* TKL contains a specific sequence motif of THDSIGLGEDGPTHQPIE that has been identified previously in other TKL proteins [3, 14, 38] and corresponds to amino acids 469–486 in this enzyme. Interestingly, a second sequence motif, which is common to thiamine cofactor-binding enzymes, is a GDG consensus motif (amino acids 159–161) followed by 21 amino acids varying in sequence identity [6]. Amino acid residues involved in substrate or cofactor binding (Arg359, His69, and His263) and active site (His30) were also conserved [23, 45].

TKLs have been identified in several organisms, including humans, *S. cerevisiae*, *E. coli*, maize, and spinach [14, 44, 46]. These are typically cytosolic enzymes that have a molecular mass of 70–75 kDa, with the homodimer being the active entity [23]. This class of enzyme uses the cofactor thiamine pyrophosphate and a divalent metal cation to catalyze the cleavage of carbon-carbon bonds to transfer two ketol carbon units from donor ketose sugars, such as ribose-5-phosphate or erythrose-4-phosphate, resulting in the production of sedoheptulose-7-phosphate or fructose-6-phosphate, respectively [25, 42].

**Complementation**

Since erythrose-4-phosphate, which is produced by TKL, is used as substrate for biosynthesis of aromatic amino acids [2, 46], the ability of the cloned CmTKL1 gene to complement the transketolase-negative (TKL-) phenotype, the growth defect of the *tkl1* *tkl2* mutant in the absence of aromatic amino acids, was tested by transformation of the *S. cerevisiae* *tkl1* *tkl2* mutant strain. As shown in Fig. 2A, the *S. cerevisiae* *tkl1* *tkl2* strain harboring p415GPD (vector) did not grow in the SC-Trp Phe Tyr plate. However, the *S. cerevisiae* transformant harboring the pME1330 plasmid successfully grew in the SC- Trp Phe Tyr medium. TKL enzyme activity assays for the *S. cerevisiae* *tkl1* *tkl2* transformants also indicated that the *TKL1* gene cloned from *C. magnoliae* could recover the TKL- phenotype of the *S. cerevisiae* *tkl1* *tkl2* strain (Fig. 2B).

**Expression Analysis**

Expression of the CmTKL1 gene was examined in the presence of different carbon sources, and the effects of stress inducers, including NaCl, KCl, sorbitol, and H₂O₂, on gene transcription and TKL enzyme activity were investigated. As fructose concentration increased from 2% to 30% (Fig. 3), the transcript level of TKL was notably enhanced with concomitant more than 2-fold enhancement in TKL enzyme activity, which is in part compatible with the enhanced erythritol production rate for fructose compared with glucose or sucrose [17]. However, increase in the concentrations of sucrose and glucose did not promote TKL enzyme activity. In addition, TKL activity in the presence of 2% sucrose was notably higher than that observed for 2% glucose and 2% fructose.

A significant increase in both TKL transcript level and enzyme activity was also observed when the *C. magnoliae* cells were grown under stress conditions induced by KCl,
Fig. 1. Alignment of multiple amino acid sequences from *C. magnoliae* TKL and other homologous fungal TKLs.

Fully conserved residues are marked in black. The conserved regions on which degenerate primers were designed are marked with asterisks, substrate binding sites are marked with inverted open triangles, and binding sites for Mg\(^{2+}\) ion and thiamine pyrophosphate are marked with inverted filled triangles and filled triangles, respectively. The numbers on the left and right indicate the positions of the amino acids. *Candida magnoliae* (GenBank Accession No. KF751756); *Candida tropicalis* (EER32747); *Sheffersomyces stipitis* (CAA81260); *Komagataella pastoris* (ACN76561); *Saccharomyces cerevisiae* (CAA85074); *Kluyveromyces lactis* (AAC05935); *Aspergillus oryzae* (XM 001727478); *Aspergillus niger* (XM 001392735); and *Mucor circinelloides* (EPB92998).
indicating that salt stress, as well as sugar stress observed for fructose, may be involved in the expression of \textit{CmTKL1}, which may support the finding that showed the erythritol production rate for \textit{C. magnoliae} was significantly enhanced by supplementation of KCl into the cultivation medium [4]. However, treatment with 0.5 M NaCl did not exert a profound effect on either the transcript level or TKL enzyme activity, indicating that regulation of \textit{CmTKL1} expression may rather be salt-specific. Neither transcript level nor enzyme activity was enhanced by osmotic stress caused by sorbitol. Since the TKL enzyme activity was enhanced in the presence of KCl and H$_2$O$_2$, the upregulation of the \textit{CmTKL1} gene was thought to be caused by salt stress and oxidative stress rather than by osmotic stress by sorbitol or high concentrations of glucose and sucrose. Accordingly, it was concluded that \textit{CmTKL1} expression was regulated in a sugar- and stress-dependent manner.

Results in the present study also suggest that the \textit{TKL1} gene deletion mutant must be constructed in order to elucidate the properties of TKL, and thus, this remains the subject of future studies. In summary, the gene encoding a transketolase in an erythritol-producing yeast, \textit{C. magnoliae}, was cloned and its function was verified in a \textit{S. cerevisiae} mutant.

\section*{Acknowledgments}

This work was carried out with the support of

\textbf{References}


\textit{J. Microbiol. Biotechnol.}


