Identification of the *Dekkera bruxellensis* phenolic acid decarboxylase (PAD) gene responsible for wine spoilage

Liliana Godoy, Verónica García, Rubén Peña, Claudio Martínez, Ma Angélica Ganga

*Laboratorio de Biotecnología y Microbiología Aplicada, Departamento en Ciencia y Tecnología de los Alimentos, Universidad de Santiago de Chile, Obispo Manuel Umaña 050, Estación Central, Santiago, Chile*

**Abstract**

The species of the genus *Dekkera* have been described as the main spoilage yeast in wine and are associated with the occurrence of phenolic off-flavor compounds.

In this work, we describe the gene encoding phenolic acid decarboxylase enzyme of *Dekkera bruxellensis* (*DbPAD* gene). The gene has 43% identity with *PAD* nucleotide sequences from different ascomycetes species, and 34% of identity with bacterial sequences. The open reading frame shows an average of 8% identity with the *Pad1* protein of *Giberella zeae*, *Candida albicans* and *Saccharomyces cerevisiae*, and 34.9% with *Pad1* from *Candida guilliermondii*. The heterologous expression of the *DbPAD* gene in *S. cerevisiae* resulted in high 4-vinylphenol production, reaching levels similar to those obtained in *D. bruxellensis* cultures. The enzymatic reaction showed that the *PAD* activity in the transformed yeast was higher than in the wild type and similar to that detected in *D. bruxellensis*. These results demonstrate that the *DbPAD* gene encodes a phenolic acid decarboxylase.

© 2014 Elsevier Ltd. All rights reserved.
Fig. 1. Nucleotide sequence alignment of the Pad proteins of 4 ascomycetes: \textit{S. cerevisiae}, \textit{C. albicans}, \textit{D. hansenii} and \textit{G. zeae}. The conserved sequences on which the primers used were designed are indicated in parentheses [ ].
expression of these paralogous genes of DbARO10 showed that both respond to the presence of p-coumaric acid in the culture medium. This would indicate that there are alternative decarboxylases in the proteome of *D. bruxellensis* which have not been identified.

In this paper we describe the *DbPAD* gene which codes for the enzyme phenolic acid decarboxylase (Pad) and when transformed into *S. cerevisiae* it elevates production of 4-vinylphenol (4-VF).

2. **Materials and methods**

2.1. **Microorganisms**

*D. bruxellensis* L2480, *S. cerevisiae* BY4722 (*MATa leu2delta0 ura3delta0*) and *Escherichia coli* DH5α were obtained from the culture collection of the Biotechnology and Applied Microbiology Laboratory of the Universidad de Santiago, Chile. Strains *E. coli* JM109, LE392 and KW251 were purchased from Promega (WI, USA).

2.2. **Culture conditions**

*E. coli* JM109 was used as the recipient strain for cloning experiments and plasmid amplification. It was grown following standard procedures (*Sambrook & Russell, 2001*). *E. coli* strains KW251 and LE392 were used for the construction of the gene library and were grown in LB medium (0.1 g/L Tryptone, 0.1 g/L NaCl, 0.05 g/L yeast extract) supplemented with 0.02 g/L maltose and 10 mM MgSO4. *E. coli* strain KW251 was grown and maintained in the presence of tetracycline (15 mg/L). *S. cerevisiae* BY4722 was grown in SC medium (6.7 g/L yeast nitrogen base without amino acids and 20 g/L glucose) supplemented with 0.02 g/L leucine and 0.02 g/L uracil. *D. bruxellensis* L2480 was grown in SC medium. To induce the enzymatic activity, 100 mg/L of p-coumaric acid was added to the culture medium.

2.3. **Construction of a genomic library**

A genomic library was constructed of *D. bruxellensis* L2480 using the LambdaGEM®-11 BamHI Arms/Packagene® System (Promega, WI, USA) according to the manufacturer’s instructions.

2.4. **Search for the *DbPAD* gene in the genome library**

The lysis plates obtained from the gene library were transferred to nitrocellulose membranes (immobilon-Ny + charger nylon transfer membrane) according to the LambdaGEM®-11 BamHI Arms/Packagene® System (Promega, WI, USA). The lysis plates containing the gene of interest were analyzed. For this, a 220 bp fragment obtained from the amplification of the *D. bruxellensis* L2480 genome with the degenerate primers 1F and 3R (*Table 1*) was used as a probe. The amplicon obtained and used as a probe was subcloned using pGEM-T Easy Vector Systems kit (Promega, WI, USA). The DIG-High Prime DNA Labeling and Detection Starter Kit II (Roche) was used for labeling of the DNA probe, hybridization and immunological detection reactions. Once the desired clone was identified, the recombinant DNA was isolated from the lysis plate according to the LambdaGEM®-11 BamHI Arms/Packagene® System (Promega, WI, USA). The fragment obtained was then sequenced (Macrogen® GmbH, Seoul, Korea).

2.5. **DbPAD gene sequence**

Using the DNA walking method (BD Biosciences Clontech) primers were designed that allowed the 5’ and 3’ amplification of the sequence. For sequencing, the PCR products were subcloned in the vector pGEM-T easy (Promega, WI, USA). The nucleotide sequence of the *DbPAD* gene was submitted to GenBank under the accession number JX 997332.

2.6. **Cloning of the *D. bruxellensis* *DbPAD* gene**

The ORF of the *DbPAD* gene was obtained from the amplification of chromosomal DNA of *D. bruxellensis* L2480 using the primers PADDbF and PADDbR (*Table 1*). These primers have the Ncol and HindIII restriction sites incorporated, respectively. The amplified fragment was subcloned in the pGEM-T easy vector (Promega, WI, USA) and digested with Ncol and HindIII. The fragment obtained was purified using the AxyPrep DNA gel extraction kit (Axygen, CA, USA) and subcloned in the vector yEPACT4 (*Sánchez-Torres, González-Candelas, & Ramón, 1998*) in the above mentioned restriction sites. Subsequently, this construct was used to transform strain *S. cerevisiae* BY4722 by electrotransformation (*Becker & Guarente, 1991*). The transformed yeast colonies were selected in SD media plates supplemented with 0.02 g/L of uracil.

2.7. **Quantification of p-coumaric acid and 4-VF**

The concentration of the phenolic compounds was quantified using the method described by *Ross, Beta, and Armitfield* (2009). For this, an analytical HPLC unit (Shimadzu Corporation, Japan) using a reversed-phase column C18 Shimadzu (150 × 4.6 mm) was used. The solvent system used was a gradient of water/formic acid (90:10% v/v) and methanol. Standard curves were prepared for p-coumaric and 4-VF (Sigma–Aldrich, USA) in the range of 0.5–100 mg/L.

In culture medium: 50 mL of the culture media of strains which reached an OD600 of 0.8 were centrifuged at 800 x g for 10 min. The obtained supernatant was analyzed.

In enzymatic reaction “in vitro”: The protein extracts were prepared according to *Godoy, Martinez, Carrasco, and Ganga* (2008). The protein concentration was quantified using the method described by *Bradford* (1976) using bovine serum albumin as standard. The enzymatic reaction was carried out in a total volume of 1 mL. The reaction mixture contained 50 mM phosphate buffer pH 6.0, 1980 mg/L p-coumaric acid and 25 μg of proteins. The reaction mixture was incubated at 40 °C for 40 min. A 1 mL volume of culture media or enzymatic reaction mixture was filtered through a 0.2 μm Millipore syringe to quantify the phenol compounds. All assays were carried out in duplicate.

One unit (U) of enzymatic activity was defined as the amount of enzyme that consumes 1 μmol of p-coumaric acid per minute.

2.8. **Statistical analysis**

Data were subjected to analysis of variance (ANOVA) and the mean values of the experiments were statistically analyzed using the LSD test. Differences were considered significant when the probability was <0.05.
3. Results

3.1. DbPAD gene of D. bruxellensis

Our project started before the first whole D. bruxellensis genome sequence (Curtin, Borneman, Chambers, & Pretorius, 2012; Piskur et al., 2012) were available. Thus, we used a PCR with degenerated primers approach. The initial approach used to determine the sequence of the DbPAD gene of D. bruxellensis was a multiple alignment of the ORFs that code for phenolic acid decarboxylases described for four ascomycetes. For this, the sequences of S. cerevisiae YJM789, Dekkera hansenii CBS767, Candida albicans WO-1 and Giberella zeae PH-1 available in the database http://www.ncbi.nlm.nih.gov were analyzed. The results of the alignment showed highly conserved zones from which the degenerate primers 1F and 3R were designed (Table 1). A phage containing the required sequence was found and DNA isolated from this lysis plate was used as a template for the subsequent amplification reactions.

The in silico analysis of amino acid sequence showed that the obtained sequence has a 38.9% identity with the protein Pad of Phytophthora nicotianae, 13.8% with PadC of Bacillus subtilis and 15.1% with the FDC of Bacillus pumilus. For the phenylacrylic acid decarboxylase protein described for ascomycetes, the percentage of identity with S. cerevisiae was 7.9%, 8.3% with G. zeae, 9.3% with C. albicans and 34.9% with Candida guillermondii. Sequence analysis of internal peptides obtained from the purified protein (Godoy et al., 2008) showed four peptide sequences with homology to domains present in the DNA fragment obtained in this work, with a similarity that varied between 25% and 40% (data not shown). Additionally, the amino terminal region of the purified protein (Godoy et al., 2008) shares an 80% identity with the amino acid sequence obtained in our study. The segment was then confirmed with the analysis of the genome data (Curtin et al., 2012; Woolfitt, Rozpedowska, Piskur, & Wolfe, 2007).

3.2. Expression of the D. bruxellensis DbPAD gene in S. cerevisiae BY4722

Expression vector was constructed and the D. bruxellensis gene was under the S. cerevisiae actin (ACT1) promoter control. The transformation of this yeast with the vector yEPACT4 without the insert was used as the control. The vector yEPACT4 has the ACT1 promoter of S. cerevisiae allowing the heterologous expression of proteins in this yeast (Sánchez-Torres et al., 1998). The yeasts were grown in culture media supplemented with p-coumaric acid. Once an OD600 of 0.8 was reached, the culture medium was centrifuged and the concentration of 4-VF was quantified in the supernatant. Results shown in Table 2 indicate that strains S. cerevisiae BY4722yEPACT4DbPAD and D. bruxellensis L2480 produced similar detectable quantities of 4-VF in the culture media but not in the control assays (wild S. cerevisiae yeast and S. cerevisiae transformed with the vector without the insert). However, it was observed that the sum of the p-coumaric acid and 4-VF concentrations in the culture media of these yeast was less than expected (approximately 100 mg/L).

Subsequently, the cultures were centrifuged (S. cerevisiae BY4722 transformed with yEPACT4DbPAD, S. cerevisiae BY4722 transformed with yEPACT4 S. cerevisiae BY4722 and D. bruxellensis L2480) and a protein extraction was carried out on the pellet (Godoy et al., 2008). Using these extracts, the Pad activity was determined by quantifying the production of 4-VF (Table 2). Results show that strain S. cerevisiae yEPACT4DbPAD had a greater phenol production than the control microorganism (S. cerevisiae BY4722 transformed with yEPACT4) and the wild yeast S. cerevisiae BY4722.

4. Discussion and conclusion

D. bruxellensis is globally considered as the main microbiological contaminant of wines because it is capable of producing volatile phenols in high concentrations which have a negative effect on its organoleptic qualities. The production of these is due to the metabolization of hydroxycinnamic acids, naturally present in the grape, through two enzymatic reactions: a phenylacrylic acid decarboxylase (Pad) and a vinylphenol reductase (Vr). Decarboxylase activity has been described in a number of species; Lactobacillus plantarum (Cavin et al., 1997), Bacillus pumilus (Degrassi et al., 1995) and in yeast such as S. cerevisiae (Clausen et al., 1994); Pichia guilliermondii (Barata, Nobre, Correia, Malfeito-Ferreira, & Loureiro, 2006) and D. bruxellensis (Edlin, Narbad, Gasson, Dickinson, & Lloyd, 1998; Godoy et al., 2008). In this work, the gene coding for the enzyme Pad was identified in D. bruxellensis. To identify the DbPAD gene of D. bruxellensis, an alignment of the nucleotide sequences of the yeast Pad genes was carried out and degenerate primers were designed in the conserved zones. The alignment obtained showed a conserved zone corresponding to the flavo-protein domain (pfam PF02441), which is found in enzymes characterized for having the capacity to bind and utilize flavin as cofactor. Some flavoproteins such as MrSD, a protein involved in the synthesis of antibiotics and EpiDa peptidil cystein decarboxylase have decarboxylase activity (Blaesse, Kupke, Huber, & Steinbacher, 2000). Furthermore, the Pad1 protein of S. cerevisiae is also included in this family (Rangarajan et al., 2004). The Pad enzymes of lactic bacteria have no homology with the Pad1 protein of S. cerevisiae (Cavin et al., 1998), but they do have a high homology with the Pad enzymes of other bacteria. Cavin et al. (1998) showed that the amino acid sequence of the enzyme phenolic decarboxylase (PadC) of Bacillus subtilis shows an 84% identity with fermentable decarboxylase (Fdc) of B. pumilus and a 71% identity with p-coumaric decarboxylase (Fdc) of L. plantarum. However, the Pad1 protein of E. coli O157:H7 has a 54% identity with the Pad1 protein of S. cerevisiae, indicating that there are highly conserved areas in these types of enzymes (Rangarajan et al., 2004).

In general, our results indicate that the DbPad protein is more similar to proteins with phenylacrylic acid activity from bacteria, however, DbPad sequence share a high degree of identity with CgPad protein of C. guilliermondii identified by Huang et al. (2012),
formulating part of the same clade. These results suggest that more analysis is needed to verify the evolutionary origin of this gene.

Likewise, some of the internal peptides as well as the amino terminal obtained from the protein purified by Godoy et al. (2008) were identified in this sequence. The amino terminal had an 80% identity which supports that the sequence described corresponds to the gene coding for the Pad enzyme. On the other hand, the ORF was 666 bp long corresponding to 221 amino acids with a molecular mass of 25,567 Da approximately, which is close to that described for the purified protein (Edlin et al., 1998; Godoy et al., 2008). To define the expression of the gene, the yeast S. cerevisiae BY4722 was transformed. The culture media of the yeast transformed with the DbIPAD segment showed a concentration of 45.0 ± 0.01 mg/L of 4-VF, whilst in the control cultures (wild-type S. cerevisiae and S. cerevisiae BY4722epact-4) this compound was not detected (Table 2). These results show that the DNA segment isolated in this work contains the sequence coding for the Pad enzyme of D. bruxellensis. This sequence would contain the active site of the protein and therefore shows for the first time that S. cerevisiae is capable of expressing a gene from D. bruxellensis. Previous studies carried out by our group (data not published) show that S. cerevisiae is able to recognize promoter sequences as well as reading frames of D. bruxellensis, and process them correctly.

Furthermore, the transformed strain produced similar amounts of 4-VF as D. bruxellensis L2480. S. cerevisiae has been described as having a basal decarboxylase activity (Claussen et al., 1994; Larsson, Nilvebrant, & Jonsson, 2001; Mukai, Masaki, Fujii, Kawamukai, & Lefuji, 2010; Suárez et al., 2007) as shown in the control assays of this study. This coincides with the experiments carried out by Stratford, Plumridge, and Archer (2007) who showed that S. cerevisiae BY4741 is able to decarboxylate only a small amount of weak acids.

While there are currently two genomes of D. bruxellensis sequenced, one of them of Australian origin (AWR1499 strain) (Curtin et al., 2012) and another of French origin (Y879 strain) (Piskur et al., 2012; Woolfit et al., 2007), they do not describe the sequence of the gene. However, Curtin et al. (2012) suggests that a putative Pad enzyme of D. bruxellensis has greater similarity to bacterial Pad proteins than to S. cerevisiae.

Various studies carried out to characterize the DbIPAD gene of Dekkera sp. (Edlin et al., 1998; Godoy et al., 2008; Harris, Ford, Jiranek, & Grbin, 2008) have shown that there could be more than one isoform. From the amino terminal analysis of the purified Pad protein, Harris et al. (2008) obtained a homologous sequence to Pst2p of S. cerevisiae, a protein of unknown function. However, the expression of the S. cerevisiae Pst2 protein in E. coli showed no detectable decarboxylase activity (Edlin et al., 1998; Harris et al., 2008). Pst2p contains a flavodoxin domain (pfam PF00258) and although both proteins belong to the flavoprotein superfamily, they are classified differently. It is important to highlight that both studies used the N terminal sequence described by Edlin et al. (1998), which explains the results obtained by both groups. Mutagenesis studies could be carried out using the DbIPAD gene to illustrate the pathway responsible for the production of volatile phenols and determine if there are other genes involved in the production of vinyl phenol derivatives. These studies may explain the differences found between isolates as well as the development of technological tools to control the formation of these volatile phenols in wine.

Acknowledgments

This work was supported by Fondecyt—Chile 1080376 and 1110700 grants, Doctoral Fellowship Comisión Nacional de Investigación Científica y Tecnológica and Thesis Support Grant Conicyt 24120969.

References


