Cloning of a protopectinase gene of *Trichosporon penicillatum* and its expression in *Saccharomyces cerevisiae*

Ken-ichi Iguchi,† Haruka Hirano,† Masao Kishida,† Haruhiko Kawasaki† and Takuo Sakai†,2

Author for correspondence: Takuo Sakai. Tel: +81 722 52 1161 ext. 2470. Fax: +81 722 52 0341.

e-mail: masyksd@biochem.osakafu-u.ac.jp

1 Department of Applied Biochemistry, College of Agriculture, Osaka Prefecture University, Gakuen-cho 1-1, Sakai, Osaka 593, Japan
2 Department of Food Science, Faculty of Agriculture, Kinki University, 3327-204, Nakamachi, Nara 631, Japan

A protopectinase (PPase)-encoding gene, PSE3, from *Trichosporon penicillatum* was cloned by colony hybridization using two oligonucleotide probes synthesized from the N-terminal amino acid sequences of native PPase SE1 and one peptide from a lysyl endopeptidase digest. Nucleotide sequencing revealed that PSE3 contains an ORF encoding a 367 amino acid protein. Mature PPase SE3 is composed of 340 amino acids and the N-terminus of the ORF appeared to correspond to a signal peptide and a propeptide processed by a KEX2-like proteinase. The deduced amino acid sequence of PSE3 was 65.4, 56.7, 58.1, 61.8 and 48.9% homologous to the polygalacturonases of *Aspergillus oryzae*, *Aspergillus niger*, *Aspergillus tubigenes*, *Cochliobolus carbonum* and *Fusarium moniliforme*, respectively. One domain, which might interact with polygalacturonic acid, is highly conserved not only in fungal polygalacturonases but also in bacterial and plant polygalacturonases. PSE3 was expressed in *Saccharomyces cerevisiae*, but three forms (the mature form, a glycosylated form and an uncharacterized processed form) of PPase SE3 were present among the PSE3 products.

Keywords: *Trichosporon penicillatum*, protopectinase, gene cloning, yeast polygalacturonase, *Saccharomyces cerevisiae*

INTRODUCTION

Many micro-organisms produce the pectin-solubilizing enzyme protopectinase (PPase), which releases soluble pectin from the protopectin in plant tissues. Two types of PPases have been reported: type A PPases, which are polygalacturonases (PGases), degrade the polygalacturonic acid chains that compose the unsubstituted region (smooth region) of the protopectin structure, and type B PPases degrade the neutral sugar chain which connects the smooth region to the other cell-wall constituents (Sakai et al., 1993). Type A PPases have been purified from yeast strains belonging to *Kluyveromyces* and *Trichosporon* species (Sakai & Okushima, 1978; Sakai et al., 1984). In another study by our group (Iguchi et al., 1996), three PPases (SE1, SE2 and SE3) were purified from *Trichosporon penicillatum* B2 strain, and the ratios of their activity to PGase activity ([PPase/PGase]) were shown to be different. We are interested in the activities of these three PPases and their similarities to PGases produced by bacteria, fungi and plants. However, analysis of the two minor PPases (SE2 and SE3) is difficult, since they are produced only in small amounts (0.15 and 0.007% of total PPase, respectively). Isolation of the PPase-encoding genes from a genomic library of *T. penicillatum* B2 is important in order to determine the complete primary structure of the PPases and to prepare large amounts of the minor PPases from transformants of *Saccharomyces cerevisiae*. The finding that the N-terminal amino acid sequences of PPases SE1, SE2 and SE3 have high homology (Iguchi et al., 1996) suggests that the genes encoding these PPases are highly conserved. Using DNA probes based on sequences common to all three PPase genes but derived from the amino acid sequences of segments of the PPase SE1, which is easily purified from *T. penicillatum* B2 (Iguchi et al., 1996), we succeeded in cloning the gene encoding PPase SE3.
Here, we describe this cloning, the analysis of the PPase SE3 gene (PSE3), the homology between PPase SE3 and the other proteins, and the production of PPase SE3 by S. cerevisiae.

METHODS

Strains, culture conditions and chemicals. The T. penicillatum and S. cerevisiae strains used were B2 (Iguchi et al., 1996) and DKD-SDH (MATa trp1 leu2 his3), respectively. Escherichia coli strains used were DH5α and MV1184 (obtained from Toyobo). Culture media used were GYP [2% (w/v) glucose, 0.5% yeast extract and 0.5% peptone] for yeasts and LB for bacteria. For some experiments, 50 μg ampicillin ml⁻¹ was added to the LB medium. Culture temperatures were 30 °C for yeasts and 37 °C for bacteria.

Manipulation of DNA. The plasmid pUC118 (purchased from Takara Shuzo) was used for constructing the gene library and DNA sequencing, and YEpl3 (Rose et al., 1991) was used for the transformation of S. cerevisiae. All reagents used for DNA manipulations were obtained from Takara Shuzo. Cleavage of DNA with restriction enzymes, ligation of DNA with T4 DNA ligase and exonuclease III treatment of DNA were done as recommended by the suppliers. DNA sequencing was carried out with a BcaBEST Dideoxy Sequencing kit and an ALF DNA sequencer (Pharmacia Biotech), after nested deletion with exonuclease III. DNA sequences were analysed by GENETYX software (Software Development).

Preparation of DNA. Plasmid DNAs were prepared from E. coli according to the method of Sambrook et al. (1992). Genomic DNA of T. penicillatum B2 was prepared by the method of Wach et al. (1994) with minor modifications as follows. T. penicillatum B2 cells cultured in the late-exponential phase in GYP were harvested and used for the preparation of protoplasts. About 10⁸ protoplasts prepared by addition of 1000 U Zymolyase 20T (Kirin Brewery) and 1000 U NOVOzyme (NOVO Nordics Japan) in 50 mM EDTA were collected, resuspended in 10 mM Tris/HCl (pH 8.0), 10 mM EDTA with 1% (w/v) SDS and 1 μg proteinase K ml⁻¹ (Sigma), and lysed by incubation at 60 °C for 1 h with occasional shaking. The lysate was mixed with an equal volume of 5 M potassium acetate (pH 5.4), kept on ice for 1 h and then centrifuged at 10,000 g. Total genomic DNAs were prepared by mixing the supernatant with an equal volume of 2-propanol and by collecting the precipitate with a glass rod. Preparation of DNA was done at 43 °C.

Transformation. Transformation of E. coli and S. cerevisiae was done by electroporation. Competent E. coli cells were prepared as described by Calvin & Hanawalt (1988) with minor modifications. E. coli cells cultured until the late-exponential stage for 12 h at 37 °C were harvested at 4 °C, washed twice with water at 4 °C, washed twice with 10% (v/v) glycerol at 4 °C and resuspended in 10% glycerol. Recipient S. cerevisiae cells were prepared by the method of Becker & Guarente (1991) with minor modifications as follows. Cells were harvested after culturing in GYP at 30 °C for 12 h, incubated for 30 min in TEMEA buffer [10 mM Tris/HCl (pH 8.5), 1 mM EDTA, 100 mM LiCl and 10 mM mercaptoethanol], washed twice in 1 M sorbitol and then suspended in the same solution.

Electroporation was performed with an Electric Cell Boler 1001 apparatus (Rikokagaku Kenkyusho) as follows. Cell suspensions of E. coli (40 μl) and S. cerevisiae (100 μl) were mixed with about 1 μg of DNA, transferred to a 0.2 cm cuvette, and exposed to a single electric pulse (22 mF capacitance, 4 ms pulse time, and 25 kV and 15 kV peak voltages, respectively). These voltages generated field strengths in the cuvette up to 12 kV cm⁻¹ and 7 kV cm⁻¹, respectively.

Design of oligonucleotide probes and Southern hybridization. PPase SE1 was digested with Achromobacter lysyl endopeptidase in 50 mM Tris/HCl buffer (pH 8.0) with 4 M urea. Peptides were prepared by hydrophobic interaction chromatography using an HPLC system with a Capcell C18-300 column (Shiseido), and N-terminal amino acid sequences determined using a PSQ1 protein sequencer (Shimadzu). Oligonucleotides corresponding to the predicted DNA sequences from the N-terminal sequences of native PPase SE1 were labelled with [³²P]ATP (Amersham Japan) with a Megalabel 5' end-labelling kit. Southern hybridization to gel-fractionated total DNA of T. penicillatum B2 digested with HindIII and transferred to a cellulose nitrate membrane (Toyo Roshi) was done as described by Sambrook et al. (1992).
Preparation and screening of gene libraries. The HindIII-digested genomic DNA was fractionated by electrophoresis in a 0.8% agarose gel, and fragments of approximately 4-4 kb which included the probe-hybridizing fragments were purified and inserted into pUC118. Colonies transformed with the gene library were transferred to a cellulose nitrate membrane and hybridized with the oligonucleotide probes according to Sambrook et al. (1992).

Protein analysis. Original PPase SE3 was obtained from our previous study (Iguchi et al., 1996). Partially purified PPase SE3 was prepared from the S. cerevisiae transformant as follows. Cells were cultured at 30 °C for 3 d with shaking and PGase fractions of the culture filtrate obtained by CM-Toyopearl 650M (Tosoh) column chromatography were then dialysed in 20 mM acetate buffer (pH 5.0). Approximately 20 mg protein was deglycosylated by digestion with 10 mU endoglycosidase H (Takara Shuzo) at 37 °C for 1 h. For Western blotting, about 50 μg protein was separated on SDS-PAGE gels by the method of Laemmli (1970), transferred to a cellulose nitrate membrane and stained with 0.1% Coomassie Brilliant Blue R-250. After destaining, the membrane was incubated in 5% non-fat milk PBS for 1 h and washed with PBS containing 0.1% Tween-20. The anti-PPase S antibody, which reacts against PPases SE1, SE2 and SE3, was prepared by the method described by Sakai al. (1989). The anti-PPase S antibody, which reacts against PPases SE1, SE2 and SE3, was prepared by the method described by Sakai et al. (1989).

RESULTS AND DISCUSSION

Cloning of the PPase gene from T. penicillatum

Two oligonucleotide probes, P-1 and P-2, were synthesized according to the DNA sequences predicted from the N-terminal amino acid sequences of the original PPase SE1 and a fragment of a lysyl endo-

![Fig. 2. Restriction map of pPSE3-1 plasmid insert. The single arrow indicates the predicted ORF and the broken double arrow the sequenced region of DNA. The narrow bars show the hybridization with the P-1 and P-2 probes. H, HindIII; K, KpnI; S, SacI; T2, EcoT22I; Ba, BamHI.](image)

![Fig. 3. Nucleotide and predicted amino acid sequences of PSE3. Single-underlined, double-underlined and dotted sequences are a putative TATA-like sequence, polyadenylation signals and termination signals, respectively. Boxed sequences correspond to the complementary sequences of the P-1 and P-2 probes. The amino acid sequence in italics is identical to the determined N-terminal amino acid sequence of mature PPase SE3. The outlined amino acid (Ala87) differs from the corresponding sequence in the P-1 probe.](image)
Fig. 4. For legend see facing page.
peptidase digest (Fig. 1a). Southern hybridization of both sequences to a HindIII digest of the genomic DNA of T. penicillatum B2 detected a fragment of about 44 kb (Fig. 1b). This suggested that the gene encoding Pase SE1 might be present on this 44 kb DNA fragment. A gene library was constructed containing HindIII fragments of 40–45 kb in the plasmid pUC118. About 5 x 10^8 colonies were screened for hybridization to both probes. One such colony was found. A plasmid, pPSE3-1, recovered from this colony hybridized to both the P-1 and P-2 probes. The restriction map of this DNA insert is shown in Fig. 2. The regions that hybridized to the P-1 and P-2 probes were identified using several deletion plasmids (Fig. 2).

**Analysis of nucleotide and amino acid sequences of the PPase gene**

About 2 kb of the cloned DNA sequence (Fig. 2) was determined and is shown in Fig. 3. It contained a predicted ORF of 1101 bp encoding a 367 amino acid protein with a calculated molecular mass of 37967 Da. However, the cloned gene did not encode PPase SE1 since the amino acid sequence on which the P-2 probe was based did not completely correspond to that in the predicted ORF. When the N-terminal amino acid sequences of PPases SE1, SE2 and SE3 were compared to the sequenced ORF, amino acids 28–45 of this ORF were found to completely match the N-terminal sequence of PPase SE3 (Fig. 3). These results indicate that the cloned gene encodes PPase SE3, and this gene was named PSE3.

PSE3 contains no introns, although fungal PGase genes usually contain a few introns. The G + C ratio of the PSE3 ORF is 48.9 mol %. The percentages of A, T, G and C in the third position are 21.6, 29.5, 23.0 and 25.9 %, respectively, showing that T is used in the third position of codons at a relatively high rate. Since the codons in highly expressed genes in fungi show a bias to C in the third position (Gurr et al., 1987), this may indicate that PPase SE3 is produced at a low level by T. penicillatum.

The first 27 amino acid residues of the PSE3 ORF are similar to a processing peptide of the mating pheromones in S. cerevisiae (Julius et al., 1984). This suggests that PPase SE3 is a pre-pro type protein, as are the fungal PGases (Bussink et al., 1991; Kitamoto et al., 1993). Upstream of the mature N-terminal amino acid sequence, the first 17 of the 27 amino acid residues are hydrophobic and are probably a signal peptide, involved in the secretion of PPase SE3. The dipeptide Ala26-Arg27 is similar to the processing site of the KEX2 proteinase in S. cerevisiae, suggesting that the 10 amino acid residues from Ala18 to Arg27 are processed by a KEX2-like proteinase in T. penicillatum. Mature PPase SE3 is composed of 340 amino acids with a calculated molecular mass of 35204 Da. The molecular mass of the original PPase SE3 was estimated to be about 40 kDa by SDS-PAGE (Iguchi et al., 1996). The cause of this size difference is not understood.

**Analysis of amino acid sequence of PPase SE3**

The deduced amino acid sequence of PPase SE3 was used to search for homologous sequences in the SWISS-PROT database. This sequence showed significant homology to a PGase of Aspergillus oryzae (Kitamoto et al., 1993), PGase II of Aspergillus niger (Bussink et al., 1991), PGase II of Aspergillus tubingensis (Bussink et al., 1991), a PGase of Cochliobolus carbonum (Scott-Craig et al., 1990) and a PGase of Fusarium moniliforme (Caprari et al., 1993). The alignments of these homologous sequences are given in Fig. 4. Sequence similarities were 65-4 % for the PGase of A. oryzae, 56-7 % for PGase II of A. niger, 58-1 % for PGase II of A. tubingensis, 61-8 % for the PGase of C. carbonum and 48-9 % for the PGase of F. moniliforme. Except for the PGase of F. moniliforme, whole regions of the mature enzymes are conserved (Fig. 4a). In the PGase of F. moniliforme, partial regions of the N-terminal sequence, including VTFKGTTF at positions 69–77, WD5KGSNS at positions 113–120, and the C-terminal region from Ala186 are conserved. The amino acid sequence of PPase SE3 was compared to that of the endo-PGase II of Erwinia carotovora (Saarialhti et al., 1990), the PGase of Actinidae chinensis (kiwifruit) (Atkinson & Gardner, 1993) and an exo-PGase of Zea mays (Niogret et al., 1991). The similarities with the entire sequences of the bacterial and plant PGases (24.4, 25.8 and 23.4 %, respectively) were lower than with those of the fungi indicating that PPase SE3 is more closely related to the fungal PGases than to the bacterial and plant PGases. However the sequence from Asp204 to Thr262 in PPase SE3 is highly conserved among these enzymes. Fig. 4(b) shows this sequence in PPase SE3, the PGase of A. oryzae, the endo-PGase II of E. carotovora, the PGase of Act. chinensis and the exo-PGase of Z. mays. These domains, conserved in fungi, bacteria and higher plants, might correspond to the substrate (i.e. polygalacturonic acid)-binding site and the active centre of the enzyme.

The hydropathy profile of PPase SE3 was compared to those of the five above-mentioned fungal PGases (Fig. 5). The complete profile of PPase SE3 is different from those of the PGases of C. carbonum and F. moniliforme, while the profile of the N-terminal half of PPase SE3 is similar to those of the Aspergillus PGases. The profile of the

**Fig. 4.** Alignments of T. penicillatum PPase SE3 with homologous PGase sequences. (a) Alignment with fungal PGases. PPase SE3_Tp, PPase SE3 of T. penicillatum; PGase_Ao, PGase of A. oryzae; PGase_Am, PGase I of A. niger; PGase_At, PGase II of A. tubingensis; PGase_Cc, PGase of C. carbonum; PGase_Fm, PGase of F. moniliforme. (b) Alignments of highly conserved domains with fungal, bacterial and plant PGases. PPase SE3_Tp, PPase SE3 of T. penicillatum; PGase_Ao, PGase of A. oryzae; endo-PGase_Ec, endo-PGase II of E. carotovora; PGase_Ac, PGase of Act. chinensis; exo-PGase_Zm, exo-PGase of Z. mays. Black-boxed letters, with and without underlining, indicate similar and identical amino acids, respectively. Lower case letters indicate the putative signal peptide and propeptide.
Fig. 5. Hydropathy profiles of PPase SE3 of *T. penicillatum* and fungal PGases. (a) PPase SE3, (b) PGase of *A. oryzae*, (c) PGase II of *A. niger*, (d) PGase II of *A. tubigensis*, (e) PGase of *C. carbonum*, (f) PGase of *F. moniliforme*. 
Protopectinase gene of *Trichosporon penicillatum*

Fig. 6. PPases produced by *S. cerevisiae* transformants. Protein samples (50 µg) were partially purified from culture filtrates by CM-Toyopearl column chromatography. Western analysis was done using anti-PPase S antibody. (a) PPases produced by transformants. Lanes: 1, original PPase SE3; 2, YEpl3 transformant; 3, pYEPSE3 transformant. (b) Effect of endoglycosidase H on PPases. Lanes: 1, original PPase SE3; 2, undigested sample; 3, sample digested with endoglycosidase H (70 kDa band). Molecular size markers were phosphorylase b (94 kDa), bovine serum albumin (67 kDa), ovalbumin (43 kDa), carbonic anhydrase (30 kDa) and soybean trypsin inhibitor (20 kDa).

region (from Gly158 to Ser266) including the conserved domain of PPase SE3 is similar to the profile of the same region of *A. oryzae* PGase, but slightly different from the profiles of the same regions of the PGase IIIs of *A. niger* and *A. tubingensis*. These results suggest that the conformation of PPase SE3 is similar to that of the Aspergillus PGases but is slightly different from that of PGase IIIs. The q_{PPase/PGase} ratio of PGase II of *A. niger* was relatively low (0.06) using lemon peel protopectin as a substrate (unpublished data). PPase SE3 shows relatively high PPase activity (q_{PPase/PGase} 0.93) using the same substrate (Iguchi *et al.*, 1996). The finding that the PPase activity of PPase SE3 is about 15-fold greater than that of PGase II of *A. niger* may also indicate that the conserved domain (Asp204-Thr262) of PPase SE3 has a conformation that is different from that of PGase II. These results suggest that this enzyme would be useful for the study of the mechanism of action of PPases.

**Characterization of genomic DNA**

When genomic DNA was digested with various restriction enzymes and analysed by Southern hybridization using the PSE3-containing DNA as the probe, three or four bands were obtained from each digest (for example, 4.4, 5.5, 6.5 and 9.0 kb from HindIII-digested DNA, data not shown). Since one of the bands on all digests is derived from PSE3, the *T. penicillatum* genome possesses two or three copies of the PSE3 homologues. We are trying to isolate these PSE3 homologues, which probably include the genes encoding PPases SE1 and SE2.

**Production of PPase SE3 from the PSE3 gene in *Saccharomyces cerevisiae***

The cloned DNA fragment containing PSE3 was inserted in the YEpl3 vector and the new plasmid, pYEPSE3, was used to transform *S. cerevisiae* DKD-5DH. The transformants were grown in leucine-free medium. PPase SE3 in the culture filtrate of the pYEPSE3 transformant was partially purified and assayed by Western blotting using the anti-PPase S antibody (Fig. 6a). The pYEPSE3 transformant produced three proteins that reacted with this antibody. The smallest of these proteins had the same molecular mass (approx. 40 kDa) as the original PPase SE3 purified from *T. penicillatum* B2. This result suggests that the promoter of the PSE3 gene was functioning and that PPase SE3 was produced in *S. cerevisiae*. The total PPase production in the culture filtrate of the transformant was about 10-fold greater than the PPase SE3 production in the culture filtrate of *T. penicillatum* B2. This may indicate that PSE3 is repressed in B2. The estimated molecular mass of approximately 40 kDa of PPase SE3 produced by the *S. cerevisiae* transformant was also larger than that calculated from the amino acid sequence (35204 Da). The reason for this has not yet been elucidated.

The other two proteins that reacted with the anti-PPase S antibody had larger molecular masses (about 41 and 45 kDa). When the partially purified enzyme was digested with endoglycosidase H, the largest protein (45 kDa) was not detected (Fig. 6b). This result indicates that the 45 kDa protein is a glycosylated protein derived from PPase SE3 or from the 41 kDa protein. The original PPase SE3 produced by *T. penicillatum* B2 shows a single band on SDS-PAGE gels (Iguchi *et al.*, 1996). This suggests the glycosylation mechanism in the secretion pathway of *S. cerevisiae* is different from that in *T. penicillatum*. Another protein (41 kDa) that was 1 kDa
larger than PPase SE3 could not be identified. One possibility is that it is a premature enzyme (pro-PPase SE3) that is not processed by the KEX2 proteasine. In S. cerevisiae, the KEX2 proteasine recognizes certain dipeptides, such as Arg-Arg, Lys-Arg and Pro-Arg, and cleaves at the C-terminal side of the propeptide (Julius et al., 1984). The C-terminal dipeptide of the propeptide of pro-PPase SE3 (Ala26-Arg27) differs from that at the cleavage site recognized by the KEX2 proteasine. The cleavage efficiency of the KEX2 proteasine is markedly decreased (by more than 90%) when the first amino acid residue of the dipeptide is exchanged for another, such as His (Brennan et al., 1990). We attempted to determine the N-terminal amino acid sequence of the 41 kDa protein, but this band appeared to be contaminated with the N-terminal amino acid sequence of the 41 kDa protein, but this band appeared to be contaminated with other proteins (data not shown). If pro-PPase SE3 was incompletely cleaved by the KEX2 proteasine, pro-PPase SE3 (Fig 6b), suggesting that the KEX2 proteasine partially cleaves the Ala-Arg dipeptide residue of pro-PPase SE3.

ACKNOWLEDGEMENTS

We thank Dr Roque A. Hours (La Plata National University, Argentina) for helpful suggestions and for reading the manuscript.

REFERENCES


