Isolation and expression of the gene which encodes a novel enzyme with polymethoxygalacturonate-degrading activity in Trichosporon penicillatum

Takuo Sakai^{a,b,*}, Naoki Sirasaka^a, Haruka Hirano^a, Masao Kishida^a, Haruhiko Kawasaki^a

^aDepartment of Applied Biochemistry, College of Agriculture, Osaka Prefecture University, Gakuen-cho 1-1, Sakai, Osaka 593. Japan ^bDepartment of Food science, Faculty of Agriculture, Kinki University, 3327-204, Nakamachi, Nara 631. Japan

Received 25 June 1997

Abstract The novel gene named PSX1, encoding a new protopectinase with the polymethoxygalacturonase activity, was isolated from Trichosporon penicillatum. Nucleotide sequencing revealed that the PSX1 gene is composed of 1080 bases (360 amino acids, 38747 Da). The N-terminal amino acid sequences of the open reading frame correspond to a signal peptide and propeptide processed by a Kex2-like proteinase. Mature PPase SX1 was composed of 334 amino acids (36121 Da). PPase SX1 produced by a S. cerevisiae transformant harboring the PSX1 gene degraded methoxylated polygalacturonic acid as a substrate, but not degraded unmethoxylated polygalacturonic acid. © 1997 Federation of European Biochemical Societies.

Key words: Gene cloning; Protopectinase; Polygalacturonase; Polymethoxygalacturonase;

Trichosporon penicillatum; Saccharomyces cerevisiae

1. Introduction

Protopectin is the original water-insoluble parent pectic substance which occurs in plants and which yields pectin or pectic acids upon restricted hydrolysis [1]. Many microorganisms produce the enzyme, named protopectinase (PPase) by Sakai et al. [2], that releases a water-soluble pectin from protopectin. Two types of PPases ahve been reported: type A PPases, polygalacturonase (PGase), degrade the polygalacturonic acid (PGA) chain, which is composed of the unsubstituted region (smooth region) of the protopectin structure, and type B PPases degrade the neutral sugar chain, which connect the smooth region to the other cell wall constituents [2,3]. Type A PPases have been isolated from Trichosporon penicillatum, Kluyveromyces wickerhamii, K. fragilis, and Aspergillus awamori [4-6]. Three PPases, PPase SE1, SE2, and SE3 were isolated from T. penicillatum B2 strain, biochemically characterized, and found to have the different ratios of PPase activities to PGase activities [7]. We suggested that the three PPases have the different protein conformations which affect their enzymatic stabilities and immunological reactivities. However, it may be difficult to analyze the conformation of the two minor PPases, SE2 and SE3, since they are produced in only small amounts (0.15 and 0.007% of total PPases, respectively) by T. penicillatum B2 strain [7].

To determine the entire primary structure of the PPases and

*Corresponding author. Fax: (81) (722) 52-0341.

Abbreviations: PPase, protopectinase; PGA, polygalacturonic acid; PGase, polygalturonase; PMGase, polymethoxygalturonase; bp, base pairs; kb. kilobases; ORF, open reading frame

prepare a large amount of the minor PPases from Saccharomyces cerevisiae transformants, we have tried to isolate the PPase genes from the genomic library of the T. penicillatum B2 strain. In our another study, we isolate the PSE3 gene, which encodes the PPase SE3, and analyzed its nucleotide and amino acid sequences [8]. A Southern hybridization analysis using the genomic DNA of T. penicillatum indicated that at least three genes are homologous to the PSE3 gene. We tried to isolate these genes, which may be the genes of PPase SE1, PPase SE2, and a novel PPase-related enzyme. This paper deals with the isolation of the gene encoding a novel PPase and the analysis of its product.

2. Materials and methods

2.1. Strains and culture condition

Strains used were Trichosporon penicillatum B2 [7], Saccharomyces cerevisiae DKD-5DH (MATa trp1 leu2 his3), and Escherichia coli DH5α (supE44 DlacU169(f80 lacZDM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1). Culture media used were GYP (2% glucose, 0.5% yeast-extract and 0.5% peptone) and leucine-free SC (2% glucose, 0.7% yeast nitrogen base, and 50 µl/ml of each amino acids) for yeasts and LB for bacteria, added 50 µg/ml ampicillin when needed. Culture temperatures were 30°C for yeasts and 37°C for bacteria.

2.2. Manipulation of DNA

Plasmids used were pUC118, YEp13 (yeast-E, coli shuttle vector), and pMA91 (containing the promoter and terminator of the PGK gene). Restriction enzymes, T4 DNA ligase, and Exonuclease III (Takara Shuzo, Kyoto) were used as recommended by the manufacturer. DNA sequences were determined using a DNA Sequencer Core Kit (Toyobo, Osaka) and A.L.F. DNA sequencer (Pharmacia Biotech, Tokyo) and analyzed by GENETYX software (Software Development, Tokyo).

Plasmids were prepared from E. coli according as described by Sambrook et al. [9]. Genomic DNA of T. penicillatum B2 was prepared as described by Wach et al. [10] and Iguchi et al. [8].

E. coli competent cells were prepared as described by Calvin et al. [11]. S. cerevisiae recipient cells were prepared by Becker and Guarente [12]. Transformations of E. coli and S. cerevisiae were done as described by Iguchi et al. [8].

2.4. Cloning strategy

Three gene libraries were prepared as follows. The HindIII-digests of the total DNA of T. penicillatum B2 were analyzed by Southern hybridization using the PSE3 gene (encoding PPase SE3) [8] as a probe. Three DNA fractions (about 3.0, 5.5, and 6.2 kb) without the PSE3 gene, hybridized with the probe, were prepared as described by Sambrook et al. [9]. Three gene libraries, GL1, GL2, and GL3, were constructed by insertion of DNA fragments of about 3.0, 5.5, and 6.2 kb, respectively, into pUC118. Probes were labeled with α-³²P-dCTP (Amersham Japan, Tokyo) with a Random Primer DNA Labeling KIT. Southern hybridization and colony hybridization were done as described by Sambrook et al. [9].

TTCAAGTTTTAAATTAAAAAAAAAATTTGAAGCAAAAGTGCGSCCAAATCATGTACAGTTGAATCAGGTACTGCGTTCAATTGACATAATA TGTTCCCCTTTTAAATGTTAAAACCTGGTTATCTATCCCATTTTTTGGAATTCTATCAAGTTAATTTGGAAAAGCGTTGAGTTTCAACCC TT GTAT GATAACT CT GTT GGAAGACA GT CAT CAAAT GGACTTA CCT CCT GCTT TAATT GTAAAAAAAAT AGGT ACT CTTT CATTATT GCAA GAACCTCCACATAATTTATAGGTCATGTTAATAATTTTATAAAAGCTCAATGTCTGACAACTGACATT<u>ATATAAA</u>GATCTTGGAATAGCC MHLSNIVSAAS L A A L A A A P A E L E R R Q V N C V F T N Y E Q I A S H T A N C D T I T L N N I N V P A G K E L D L T N L K P G A N GTCGTCTTTGAGGGTAGAACCACTTTTGGTTATGCTGAGTGGGCTGGTCCTTTGATCATGGTTTCTGGTGATGACATTACTGTTAGTCAG V V F E G R T T F G Y A E W A G P L I M V S G D D I T V S Q T P G S V I D G E G A R W W D N K G A N G G K V K P R L F Y $\mathsf{GCCCACAATTTGGACAATTCTCATATTAATGGTCTTCATATCAAGAACACCCCTGTCTTTGGTTTCAGTATCGACTCTAAAAACTTGATC$ A H N L D N S H I N G L H I K N T P V F G F S I D S K N L I ATCGACGGCGTTAGAATTGACAACTCCGACGGTGACACCCCAGGGTGCTTTCAATACTGATGCCTTTGACGTTTCCCAATCTTACAACGTTI D G V R I D N S D G D T Q G A F N T D A F D V S Q S Y N V ACCATTCAAAACGCTTGGGTTCATAACCAAGACGATTGTCTGGCCATTAACCAAGGTGAACTTATTCATTTCCTCAACGGTTACTGCTAT T I Q N A W V H N Q D D C L A I N Q G E L I H F L N G Y C Y GGTGGACATGGTTTGTCTATCGGCTCTGTTGGCGGCGGTAATGTGGTCTCAGATGTTGTAATTGCCGATTCCCAAATTATCAACTCTCAA G G H G L S I G S V G G G N V V S D V V I A D S Q I I N S Q NGVRIKTKSGQTGEVRGITYRNIFLSGITD TACGGCCTCATCGTTCAGCAAGATTACAACAACCCCGGTCATGCCACTAACAGCATTAAGATCCATGACATCACTTTCGACAACGTTCAC Y G L I V Q Q D Y N N P G H A T N S I K I H D I T F D N V H GGTACTGCGACACAGCATGGTTTTAACATTGCTATTTTCTGTGGTGATGGTAGCTGTTATGACTGGACATGGAACGAAGTTAAGATTCAT G T A T Q H G F N I A I F C G D G S C Y D W T W N E V K I H GARDYKCQIVPSSASCQAS AGAGAATAGA GCTTAATATTTTTATATA CAAAATTCTATCTTCAATTTTCTAAATGTTTCACA GCAAGTACGTTTTTAAAAATTGTGAT TTGCGCAGACTCCGTCATGCGGACTAATATTAGTTAAATAGATGTTTTTCTTGTTACGACTAATACGTAGTTTGCGTTTCCTCGCTATCT AGAGTAGAAAGTCAAAAAAAAAAAAAAAAGTTGAAAGAACTGAAATTTTTCTCTCTGAAAAAGTTTTTCCTTAACAATGACTTCCTAAA AGT|GACGAATTTTGTGAAGAGATCACTAAATTATCTACAGACAAGGQTTT|AGAGTGTTTTTACACGCTACAATTATAAGAAGAACAGCAG GCAGAAGGAAATCAACTCTTTCTAAATTCTTTTCTACTCCATAAGCACGGTTTGTTAAACAGAAACCCCCAAAACTAGCATTACATTTAT ATATGATTCCTTTTTGTAATAAAGCATGTAATAAGCTGCCGTGTTGTTGTAAGATCAAGCAAACTGACTTATTTTTTTCACAAATGTCAAA TACCTTGTTTTCTGCAATAACATTGAATAGCTAATCTCGTTGCAGAT

Fig. 1. Nucleotide and amino acid sequences of the *PSX1* gene. A 2390-bp region that included an ORF was sequenced. Single and double underlined sequences are the putative TATA-like sequence and polyadenylation signals, respectively. Dotted and boxed sequences are the CAAT sequence and transcription termination signals, respectively.

2.5. Protein analysis

PPase SE1, SE2, and SE3 were provided by Iguchi et al. [7]. PPase SX1 was prepared from the culture filtrate of the *S. cerevisiae* transformant cultured with shaking on leucine-free SC for 3 days and then on GYP for 1 day, and partially purified by CM-Toyopearl 650M (Tosoh, Tokyo) column chromatography according to Iguchi et al. [7]. PPase and PGase activities were assayed as described by Iguchi et al. [7]. Pectinase and PMGase activities were assayed by the method used for PGase activity using pectin and methoxylated PGA as substrates, respectively. Methoxylation of PGA was done by the method described by Jansen and Jeng [13], and methoxylation frequency was determined by the method of Wood and Sidduqui [14].

3. Results

3.1. Isolation and sequencing of the PPase-related gene

To clone a new gene which encodes a new PPase, a colony hybridization analysis was done on three gene libraries, GL1, GL2, and GL3, using the *PSE3* gene as a probe. As the result of colony hybridization, one positive clone in the colonies harboring the gene library GL3 was hybridized with the *PSE3* gene. The plasmid recovered from this clone contained a 6.2-kb *HindIII* fragment.

The nucleotide sequence (2,390 bp) of the probe-hybridizing region in the cloned DNA was determined and was found to contain a 1080-bp ORF that encoded 360 amino acid residues with a calculated mass of 38 737 Da (Fig. 1). The G+C ratio of the ORF was 46.6%. The frequencies of A, T, G, and C of the ORF were 24.2, 29.3, 22.1, and 24.4%, respectively. The

third codon position in the ORF was relatively biased to T. These results are as same as those of PPase SE3. Putative promoter regions (TATA- and CAAT-box like sequence) were found 80 and 100 bases upstream from the ORF, respectively, and the transcriptional terminator sequences and the polyadenylation signal (AATAAA) were found 248 (and 415) and 570 bases downstream, respectively (Fig. 1).

3.2. Analysis of the deduced protein

The homologies of the deduced protein from the ORF to PPase SE3, PPase SE2 were 52.5% and 52.0%. The deduced protein contained a highly conserved domain (Asp200 to Thr258) of PGases (Fig. 2). These results suggests that the cloned gene encodes an enzyme related to the PGase of *T. penicillatum*. The cloned gene and the coding protein were named *PSX1* and PPase SX1, respectively.

The dipeptide residue, Arg25 and Arg26, of PPase SX1 was identical to the processing site of the Kex2 proteinase in *S. cerevisiae* [15]. This result suggests that the first 25 amino acid residues are probably processed by the signal peptidase and the KEX2-like proteinase in *T. penicillatum*. Mature PPase SX1 is probably composed of 334 amino acids with a calculated mass of 36,121 Da.

3.3. Analysis of PPase SX1 produced by Saccharomyces cerevisiae

The cloned PSX1 gene was inserted in a YEp13 vector and

```
1: mhl-s-nivsaaslaalaaaapaelerrQVNCVFTNYEQIASHTANCDTITLNNINVPAG
                                                                                        58
 PPase SX1
                     fsksaifamaa<u>i</u>avaap<u>tegdlaa</u>rG<mark>G</mark>ACVF<u>R</u>DA<u>H</u>SAI<u>A</u>GKKSCSSITLENIAVPAG
                                                                                        59
 PPase SE2
                    lfsksaifamaalavaaptegdlqarG<mark>S</mark>ACVFKDAKSAIAGKKSCSSITLENIAVPAG
                                                                                        59
 PPase SE3
              59: KELDLTNLKPGANVVFEGRTTFGYAEWAGPLIMVSGDDITVSQTPGSVIDGEGARWWDNK 118
 PPase SX1
              60: OTLDLTGLAKGTVVTFAGTTTFGYKEWEGPLISVSGDSITVKQASGGKIDCGGSRWWDGK
 PPase SE2
              60: OTLDLTGLAKGTVVTFAGTTTFGYKEWAGPLISVSGDSITVKQASGGKIDCGGSRWWDGK 119
 PPase SE3
            119: GAN-GGKVKPRLFYAHNLDNSHINGLHIKNTPVFGFSIDSKNLIIDGVRIDNSDGDTQGA 177
            120: GSNS-GKTKPKFFAAHKLQNSNIQGLQVYNTPVQAFSILSDHLTLSNILIDNRAGDKAGG
 PPase SE2
            120: GSNSGGKQKPKFFYAHKLQNSNIQGLQVYNTPVQAFSILSDHLTLSNILVDNRAGDKAGG 179
 PPase SE3
            178: FNTDAFDVSQSYNVTIQNAWVHNQDDCLAINQGELIHFLNGYCYGGHGLSIGSVGG-G-N 236
 PPase SX1
            179: HNTDAFDVGSSTFITIDHATVYNQDDCLAINSGDHI<mark>I</mark>FQNGFCSGGHGLSIGSVGGRSDN
 PPase SE2
            180: HNTDAFDVG<u>TSTY</u>ITIDHATVYNQDDCLAINSG<u>D</u>HI<mark>T</mark>FQNGFCSGGHGLSIGSVGGRS<mark>L</mark>N 239
 PPase SE3
            237: VVSDVVIADSQIINSQNGVRIKTKSGQTGEVRGITYRNIFLSGITDYGLIVQQDY-N-NP
             239: <u>SVTNVQIINNQ</u>VVNSDNGVRIK<u>SV</u>SG<u>I</u>TG<u>II</u>SGVKFQDITLSNIAKYGIDVQQDYRNGGP
                                                                                       298
 PPase SE2
            240: TVSNVNILNSQVVNSDNGVRIKTISGATGSVSGVKFQDITLSNIAKYGIDVQQDYRNGGP 299
 PPase SE3
             295: -GHATNSIKIHDITEDNVHGTATOHGENIAIECGDGSCYDWTWNEVKIHGARDY-KCQIV
 PPase SX1
                                                                                       352
             299: TGNPTNGVKITGIEFINVHGTVKSSGTNAYILCGSGSCSNWTWSQINVKGGKDSGACKNV
 PPase SE2
            300: TGNPTNGVKITGIEFINIHGSVKSSGTNAYLLCGSGSCSNWTWSKINVKGGKDSGACKNV
                                                                                       359
• PPase SE3
                                                                                       361
            353:
 PPase SX1
                                                                                       366
            359: PAGATCKI
 PPase SE2
                                                                                       367
 PPase SE3
             360: PSGATCKI
```

Fig. 2. Alignments of homologous PGase sequences with the *T. penicillatum* PPase SX1 sequence. Sequences of PPase SX1, PPase SE2 and PPase SE3 are predicted from our cloning genes, *PSX1*, *PSE2* (our unpublished data), and *PSE3* [8], respectively. Black-boxed letters are identical amino acids and black-boxed and underlined letters are similar amino acids. Small capitals are the processed peptide residues.

Table 1 Enzyme activities on various substrates

Substrate	Enzyme activity (U/mg protein)			
	PPase SX1	PPase SE1	PPase SE2	PPase SE3
PGA ^a	<1	160000	26000	130000
Protopectin ^b	230000	190000	25000	110000

aPGase activity was assayed.

the new plasmid (pYEPSX1) was used to transform the S. cerevisiae DKD-5DH strain. However, PPase activity was not detected in the culture supernatant of the transformant. Next, the plasmid pMKPSX1 was constructed by the inserting the ORF-region of PSX1 into pMA91. PPase activity was detected in the culture supernatant of the transformant harboring pMKPSX1. PGase activity was assayed using the culture supernatant of the pMKPPSX1 transformant as a crude enzyme preparation. However, PPase SX1 showed less activity than PPase SE1, SE2, and SE3 using pure PGA as a substrate (Table 1). When the enzyme activities on protopectin prepared from a lemon peel were assayed, the PPase activity of PPase SX1 was the same as that of PPase SE1, SE2, and SE3 (Table 1). These results suggest that PPase SX1 more strongly act on PGA-related compounds (probably modified PGA) than on pure PGA in the pectin and protopectin.

Since a part of the PGA molecule is methoxylated in the pectin and protopectin, the effect of methoxylation of PGA on the activity of PPase SX1 was assayed. Fig. 3 shows the enzyme activities of PPase SX1, SE1, SE2, and SE3 on the substrates which contain a various frequency of the methoxylated PGA. The activity of PPase SX1 was larger but the activities of PPase SE1, SE2, and SE3 were smaller using the more methoxylated PGA as the substrate. These results suggest that PPase SX1 is a PMGase which degrades only the methoxylated PGA. Since a PMGase has not yet been reported, PPase SX1 is a novel PPase.

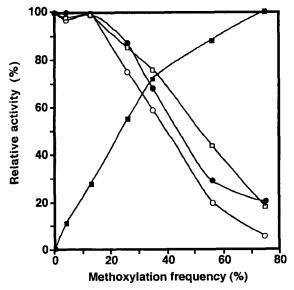


Fig. 3. Effect of methoxylation of PGA on enzyme activity. Enzymes were reacted with 10 mg substrate at 40°C for 30 min. Maximum activities of each enzymes are 100%. PPase SE1, ○; PPase SE2, •; PPase SE3, □; PPase SX1, ■.

4. Discussion

We have succeeded in isolating an interesting gene that encodes a novel PPase that is more active on methoxylated PGA than on pure PGA. The DNA sequence indicates that the PSX1 gene contains a TATA-sequence and other units necessary for the initiation of transcription. However, PSX1 was not expressed in transformed S. cerevisiae. These results are different from those of PSE3, which encodes PPase SE3, since PSE3 was expressed in S. cerevisiae [8]. As the promoter region of PSE3 functions in S. cerevisiae and T. penicillatum, the inability of the promoter region of PSX1 to function in S. cerevisiae suggests that PSX1 is a gene that is not expressed in T. penicillatum. These results suggest that the third position of the codons of the PSXI ORF is biased to T. As the codons of highly expressed genes in fungi are biased to C at third position [16], PSX1 is probably expressed at a low level in T. penicillatum. The reason why PSE3 can express and PSX1 can not express while the both genes have the same units of the transcriptional initiation has not yet been elucidated. But the results that the 5'-upstream sequence of PSX1 is different from PSE3 except for the transcriptional initiation units may suggest that the different cis-elements which regulate the expression of PSX1 and PSE3 is present in the 5'-upstream sequences of the both genes.

The PPase SX1 produced by the S. cerevisiae transformant harboring the PSX1 gene contains PPase activity. However, PPase SX1 is more active against methoxylated PGA than pure PGA, while the other PPases (SE1, SE2, and SE3) of T. penicillatum are more active against pure PGA than methoxylated PGA. The primary structure PPase SX1 has a high similarity to that of PPase SE2 and SE3 (Fig. 2), and especially the conserved domains of the PGases are almost identical. To our knowledge, there are no other reports of fungal PGases with PMGase activity. The results of this study indicate that the substrate specificity of PPase SX1 is much different from those of the other PGases, although the primary structures of these enzymes are very similar. This suggests that the specificity of PPase SX1 for methoxylated PGA is due to a few differences in the amino acid sequence. In fact, the structure of the enzyme surface of PPase SX1 appears to be different from those of the other PPases (SE1, SE2, and SE3) of T. penicillatum, since PPase SX1 does not react with the anti-PPase S antibody that interacts with PPase SE1, SE2, and SE3 (our unpublished data). The study of PPase SX1 will provide important data on the mechanism of protopectin solubilization by the PPases, since PPase SX1 is a novel PPase that has a substrate specificity for methoxylated PGA.

Acknowledgements: We thank Dr. A. Kingsman of Oxford University, UK, for his kind gift of the plasmid used in this study, and Dr.

^bPPase activity was assayed using protopectin prepared from lemon peel as a substrate.

Roque A. Hours of La Plata National University, Argentina, for critically reading the manuscript.

References

- [1] van Buren, J.P. (1991) in: The chemistry and Technology of Pectin, Academic Press, San Diego, pp. 1–22.
- [2] Sakai, T., Sakamoto, T., Hallaert, J. and Vandamme, E.J. (1993) Adv. Appl. Microbiol. 39, 213-294.
- [3] Hours, R.A. and Sakai, T. (1994) Biotech. Lett. 16, 721-726.
- [4] Sakai, T., Okushima, M. and Yoshitake, S. (1984) Agric. Biol. Chem. 48, 1951–1961.
- [5] Yoshitake, S., Numata, T., Katsuragi, T., Hours, R.A. and Sakai, T. (1994) J. Ferment. Bioeng. 77, 370–375.
- [6] Yoshikawa, M., Hayashi, M., Sakamoto, T., Hours, R.A., Katsuragi, T. and Sakai, T. (1995) Appl. Biol. Sci. 1, 71–82.
- [7] Iguchi, K., Kishida, M. and Sakai, T. (1996) Biosci. Biotech. Biochem. 60, 603–607.
- [8] Iguchi, K., Hirano, H., Kishida, M., Kawasaki, H. and Sakai, T. (1997) Microbiology 143, 1657–1664.

- [9] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1992) in: Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, New York.
- [10] Wach, A., Pick, H. and Philippsen, P. (1994) in: Molecular Genetics of Yeast: A Practical Approach. IPL Press, Oxford, pp. 1-16.
- [11] Calvin, N.M. and Hanawalt, P.C. (1988) J. Bacteriol. 170, 2797-2801.
- [12] Becker, D.M. and Guarente, L. (1991) in: Methods in Enzymology, vol. 194: Guide to Yeast Genetics and Molecular Biology, Academic Press, San Diego, pp. 182–187.
- [13] Jansen, E.F. and Jeng, R. (1946) J. Am. Chem. Soc. 68, 1475-1477
- [14] Wood, P.J. and Sidduqui, J.R. (1971) Anal. Bioche. 39, 418-428.
- [15] Julius, D., Brake, A., Blair, L., Kunisawa, R. and Thorner, J. (1984) Cell 37, 1075–1089.
- [16] Bussink, H.J.D., Brouwer, K.B., de Graaff, L.H., Kester, H.C.M. and Visser, J. (1991) Curr. Genet. 20, 301–307.