Asia-Pacific Journal of Science and Technology: Volume: 27. Issue: 02. Article ID.: APST-27-02-12. Research Article

APST

Asia-Pacific Journal of Science and Technology

https://www.tci-thaijo.org/index.php/APST/index

Published by the Research and Graduate Studies,
Khon Kaen University, Thailand

Release of alcohol oxidase 1 (AOX1) from *Pichia pastoris* during heterologous expression of rice enzymes

Istiftakhun Nikmah[1], Manatchanok Kongdin[1], Yanling Hua[1,2], Sittiruk Roytrakul[3], Fatchiyah Fatchiyah[4] and James R. Ketudat-Cairns[1,5,*]

[1]School of Chemistry, Institute of Science and Center of Biomolecular Structure, Function, and Application, Suranaree University of Technology, Nakhon Ratchasima, Thailand
[2]Center for Scientific and Technological Equipment, Suranaree University of Technology, Nakhon Ratchasima, Thailand
[3]National Center of Genetic Engineering and Biotechnology, Thailand Science Park, Pathum Thani, Thailand
[4]Research Center of Smart Molecule of Natural Genetic Resources and Department of Biology, Faculty of Mathematics and Natural Science, Brawijaya University, Malang, East Java, Indonesia
[5]Laboratory of Biochemistry, Chulabhorn Research, Bangkok, Thailand
[*]Corresponding author: cairns@sut.ac.th

Received 10 July 2021
Revised 15 September 2021
Accepted 3 October 2021

Abstract

Pichia pastoris is a popular host for recombinant expression of proteins due to its ability to secrete proteins, grow at high cell density with methanol as a carbon source, and support methanol-induced protein expression with the *alcohol oxidase 1 (AOX1)* promoter. However, overproduction of native proteins in the medium during induction may decrease its advantage as a host for heterologous protein expression. Here, we sought to identify what native protein is overexpressed in the medium during heterologous expression of rice enzymes in *P. pastoris*. The genes encoding rice enzymes, including Os4BGlu11 β-glucosidase and two putative serine carboxypeptidase-like acyltransferases, OsSCPL2a and OsSCPL7, were inserted in the pPICZαBNH8 expression vector and used to transform *P. pastoris*. Upon induction of expression, a 75 kDa protein was observed in the medium of cells expressing OsSCPL7 and Os4BGlu11 but not in OsSCPL2a. The protein was purified by immobilized metal affinity chromatography from the media, excised from sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) gels, and identified by liquid chromatography and tandem mass spectrometry (LC/MS/MS) as *P. pastoris* alcohol oxidase 1 (AOX1) protein. The purified AOX1 was found to convert cyanidin-3-*O*-glucoside to compounds that no longer exhibit absorbance in the 520 nm range, but absorb relatively strongly at 360 nm, as detected in ultra-high-performance liquid chromatography. These results suggest that although it is normally found in the peroxisome, AOX1 protein may be released into medium during expression of certain proteins, but not other closely related proteins, and that AOX1 has the ability to oxidize complex phenolic alcohols, such as anthocyanins.

Keywords: Alcohol oxidase 1, Anthocyanin, β-glucosidase 11, *Pichia pastoris*, Serine carboxypeptidase-like acyltransferase.

1. Introduction

Pichia pastoris (P. pastoris, also called *Komagataella phaffii*) is a methylotrophic yeast that is very useful for heterologous production of proteins. It has the advantage of being able to grow at high cell densities on simple media and produce high concentrations of expressed proteins from either constitutive or induced promoters [1].

The most popular inducible promoter is derived from the *alcohol oxidase 1* (*AOX1*) gene, which is used for methanol-induced expression of foreign genes [1,2].

The *alcohol oxidase 1 (AOX1)* gene encodes the AOX1 enzyme (EC 1.1.3.13), which belongs to the family of glucose-methanol-choline (GMC) oxidoreductases [3], and allows *P. pastoris* to live on methanol as a sole carbon source. Based on substrate specificity, AOX enzymes are categorized to four groups, consisting of short chain alcohol oxidase (SCAO), long chain alcohol oxidase (LCAO), aromatic alcohol oxidase (AAO) and secondary alcohol oxidase (SAO) [4]. Among those groups, SCAO is found in *P. pastoris*. SCAO are intracellular enzymes located in the peroxisome of yeasts that catalyze the oxidation of alcohol substrates with chain lengths in the range of one to eight carbons [4,5]. The structure and details of the catalytic activity of AOX have been clearly described [2]. Yet, to our knowledge, there is no information about the action of AOX protein toward phenolic alcohols, such as anthocyanins.

Carbohydrate active enzymes (Cazymes) include enzymes that modify carbohydrates, such as glycoside hydrolases (GH), glycosyltransferases (GT), polysaccharide lyases (PL), carbohydrate esterases, lytic polysaccharide mono-oxygenases and associated enzymes (CAZY database: www.cazy.org) [6]. Among these, the glycoside hydrolases are perhaps most abundant and are critical for a number of functions both in living organisms and in industry. For instance, β-glucosidases are needed for glycolipid recycling in most eukaryotes, and play roles in defense, signaling and cell wall recycling in plants, as well as being critical for biomass breakdown by microorganisms, both in nature and industry [7]. As such, it is often desired to characterize such enzymes in recombinant systems to evaluate their substrate specificity as a clue to their role in nature, as well as to evaluate their potential for biotechnological application. Aside from the "official" carbohydrate active enzymes, several other groups of enzymes act on carbohydrates to modify them or utilize carbohydrates for leaving groups or other metabolic functions. For instance, acyl glucose-dependent acyl transferases, which belong to the serine carboxypeptidase-like (SCPL) enzyme subfamily, transfer an acyl group from a 1-O-β-glucosyl ester onto another molecule, such as a saponin or the sugar on a flavonoid glycoside [8].

Anthocyanins are colored water-soluble pigments and belong to the flavonoid family [9]. Several modifications occur on anthocyanins, such as hydroxylation and methoxylation at various positions, followed by attachment of sugars or organic acid, or a combination of these by glycosylation via glycosyltransferases (GTs), and acylation catalyzed by acyltransferases (ATs) [10]. As noted above, SCPL ATs transfer acyl groups from 1-*O*-β-glucose esters, mainly in the vacuole, while BAHD ATs transfer from acyl-coenzyme A, primarily in the cytoplasm. In the last decade, it has been reported that many ATs play roles in modification of anthocyanins to form complex anthocyanins. For instance, CtAT1 in *Clitoria ternatea*, AtSAT in *Arabidopsis thaliana*, DgSCPL1 and DgSCPL2 in *Delphinium grandiflorum* [11-13]. These play important roles in the chemical stabilization of anthocyanins. However, it is important to look into the potential of other enzymes to modify anthocyanin structure. OsSCPL2a and OsSCPL7 are two putative rice SCPL AT enzymes that are expressed in rice flowers and grains, where anthocyanins may occur in black rice. Therefore, we set out to express these enzymes in *P. pastoris* to see whether they are capable of acylating anthocyanins. In the process, we found that the wrong protein accumulated in the medium during OsSCPL7 expression and was bound to immobilized metal affinity resin, and this also occurred during the expression of a putative rice β-glucosidase, so we endeavored to identify this protein.

2. Materials and methods

2.1 Strains and plasmid

Two optimized genes encoding *Oryza sativa* SCPL2a (OsSCPL2a, Genbank accession XP_015614311.1) and SCPL7 (OsSCPL7, Genbank accession XP_015617342.1) without their predicted signal sequences (predicted by SignalP: http://www.cbs.dtu.dk/services/SignalP/) were synthesized and inserted into yeast expression vector pPICZαBNH8 [14] at the *Pst*I and *Xba*I sites, while a cDNA optimized for rice β-glucosidase Os4BGlu11 was synthesized and inserted into pUC57 vector (GenScript, Piscataway, NJ, USA). The optimized Os4BGlu11 cDNA was cut at *PstI* and *XbaI* and ligated into the corresponding sites in pPICZαBNH8 [14]. The ligation reaction was transformed into *Escherichia coli* strain DH5α. The sequences of the expression cassettes of all plasmids were verified by automated DNA sequencing (Genscript or Macrogen Corp. Seoul, South Korea). The corresponding recombinant plasmids, pPICZαBNH8/Os4BGlu11, pPICZαBNH8/OsSCPL2a and pPICZαBNH8/OsSCPL7 were propagated in *E. coli* strain XL1-Blue to generate a sufficient amount of recombinant plasmid for *Pichia* transformation. Competent cells of *P. pastoris* strain SMD1168H were used for protein expression and purification.

2.2 Plasmid preparation, linearization and transformation.

The recombinant plasmids pPICZαBNH8/Os4BGlu11, pPICZαBNH8/OsSCPL2a and pPICZαBNH8/ OsSCPL7 were extracted by alkaline lysis [15]. The quantity of plasmid DNA was determined by measuring the absorbance at 260 nm with a nanodrop spectrophotometer. In addition, 1% (w/v) gel agarose electrophoresis was performed to check the purity and quality of DNA [16]. The recombinant plasmids were linearized by digestion with *Sac*I for integration into the AOX1 promoter in the *P. pastoris* genome and analyzed by 1% (w/v) agarose gel electrophoresis. Approximately, 5-10 µg of DNA plasmids were transformed into 50 µL of SMD1168H competent cells by electroporation. In each case, the contents of the tube were mixed properly and transferred into prelabelled and prechilled electroporation cuvettes. Electric pulses were generated with the following settings (voltage: 1500 VH, capacitor: 25 µF using using a BTX, ECM electroporator (Biorad Corp**.**) 20 and time: 5 s) [17]. Immediately, 1 mL of ice-cold 1 M sorbitol (as recovery medium) was added into the cuvettes and mixed. The transformed *Pichia* was transferred into a new tube and incubated at 30℃ for 1 h. The mixture was divided into three plates following these volumes (200 µL, 300 µL and 500 µL), spread on YPGS agar (1% yeast extract, 2% peptone, 2% filtered glucose, 2% agar and 18.2% sorbitol, w/v) containing 100 µg/mL zeocin and incubated at 30℃ for 2-3 days. The colonies were selected and screened again on YPG agar (1% yeast extract, 2% peptone, 2% filtered glucose, 2% agar, w/v) containing 250 µg/mL zeocin. Subsequently, colonies from each recombinant plasmid were randomly chosen and grown in 10 mL Buffered Glycerol Complex Medium for Yeast (BMGY) medium. The cultures were incubated for 24 h at 30℃ with constant shaking at 250 rpm. The medium was replaced to Buffered Methanol Complex Medium for Yeast (BMMY) and the culture was incubated at 20℃ with constant shaking at 250 rpm. The protein was induced with methanol every 24 h to a final concentration of 1% (v/v) for 5 days. Furthermore, the media was collected by centrifugation at 3,500 rpm, 20 min at 4℃. The media was precipitated with 50% saturation ammonium sulphate by adding it and stirring gently overnight at 4℃. After centrifugation at 12,000 rpm, 20 min at 4℃, the precipitated protein was collected and washed with buffer (20 mM Tris-HCl, pH 8.0) thrice. The final pellet was resuspended with buffer and checked by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE).

2.3 Protein expression and purification

Protein induction and expression was conducted by inoculating a single colony into BMGY medium (YPG broth plus 100 mM potassium phosphate buffer, pH 6.0, 1.34% (w/v) yeast nitrogen base (YNB) without amino acids, 1% (v/v) glycerol and 4 x 10^{-5}% (w/v) biotin). The culture was grown for 24 h at 30℃ with shaking at 250 rpm [18]. The cells were centrifuged at 3,500 rpm, 20 min at 4℃. The BMGY medium was replaced with 200 mL BMMY (YPG broth + 100 mM potassium phosphate buffer pH 6.0, 1.34% (w/v) YNB w/o amino acids, 1% (v/v) methanol for OsSCPL proteins or 0.5% (v/v) for Os4BGlu11 and 4 x 10^{-5}% (v/v) biotin). At this step, the culture was incubated at 20℃ with constant shaking 250 rpm. Thereafter, protein was induced by adding methanol every 24 h to a final concentration of 1% (v/v) for SCPL proteins and 0.5% (v/v) for Os4BGlu11 for a total period of 5 days. For screening, 30 colonies were picked and screened every day for hydrolysis of 1 mM *p*-nitrophenyl-β-D-glucoside (pNPGlc) [14] for Os4BGlu11 or after completion of the culture for bands on an SDS-PAGE gel for the OsSCPL proteins. After large scale culture, the culture supernatant was collected by centrifugation at 3,500 rpm, 20 min at 4℃ and added with 1 mM phenylmethylsulfonyl fluoride (PMSF). The pH of supernatant was adjusted to 7.5 by adding 2 M dibasic potassium phosphate and filtered through a 0.45 µm membrane (Millipore).

For protein purification, immobilized metal affinity chromatography (IMAC) was employed, as described [19] with some modifications. The IMAC sepharose resin (GE Healthcare) bound to Ni^{+2} was equilibrated with equilibrium (Eq) buffer 1 (20 mM Tris-HCl pH 7.5, 150 mM NaCl) for OsSCPLs protein, and IMAC bound to Co^{2+} equilibrating with Eq buffer 2 (50 mM sodium phosphate buffer pH 7.5) for Os4BGlu11. The supernatant was incubated with IMAC resin at 4℃ for 1 h with moderate shaking and the suspension was transferred to a column. Eq buffer was used to wash the column twice, followed by W_1 (5 mM imidazole in the appropriate Eq buffer) and W_2 (10 mM imidazole in Eq buffer). The protein was eluted from the column by elution solution (100 mM and 250 mM imidazole in Eq buffer). The fractions were collected and analysed by 12% (w/v) SDS-PAGE gel with Coomassie brilliant blue staining. To assess glycoproteins, the protein was digested with endoglycosidase-H at 4℃ overnight.

2.4 Tryptic digestion and Liquid Chromatography with tandem mass spectrometry (LC/MS/MS)

The major purified protein bands from yeast system expression and purification were excised from SDS-PAGE gel. In-gel digestion with trypsin was employed to prepare the proteins for LC/MS/MS analysis, following the previously described protocol [20,21]. At the final step, a total solution in the master tube was dried. The purified OsSCPL proteins were analyzed with LC/MS/MS in the National Center for Genetic Engineering and

Biotechnology, Pathum Thani, Thailand, on an Impact II UHR-TOF MS System (Bruker Daltonics Ltd., Germany), as previously described [22]. The analysis for Os4BGlu11 was conducted on a micrOTOF-Q II ESI-Qq-TOF mass spectrometer at Suranaree University of Technology. The Mascot search engine (http://matrixscience.com), specifically the MS/MS ion search, was used to identify the protein hits. The libraries contaminant (AA), SwissProt (AA) and UP59680_O_sativa (AA) were searched. Trypsin enzyme was chosen since it was used for in-gel digestion and carbamidomethyl (C) was selected as a fixed modification.

2.5 Enzyme assay and Ultra-high-performance liquid chromatography (UHPLC)

The enzyme activity was determined by setting up 100 µL reaction consist of 1 µg/µL of protein, 100 mM citrate buffer, pH 5.0, and 300 µM substrate cyanidin-3-*O*-glucoside (Cy3G) [10,13]. The mixture was incubated at 30℃ overnight and the reaction was terminated with 1x volume of 0.2% (v/v) formic acid in methanol. The mixture could be kept at -20℃ before analysing with UHPLC. The ultra-filtrated product from the enzyme reaction and standard anthocyanin were detected by UHPLC with 2.1x150 mm SC-C18 column (Agilent, USA). The column temperature was maintained at 40℃. The flow rate was 0.2 mL min^{-1} with injection volume of 10 µL. The signal was captured by a diode array detector (DAD) at wavelengths of 260, 280, 325, 360, 520, and 530 nm. For gradient elution, 0.2% (v/v) formic acid in high-performance liquid chromatography (HPLC) water was used as solvent A and 100% (v/v) acetonitrile as solvent B. The gradient elution program was as follows: 0-2 min, 95% A; 2-13 min, 95 to 50% A; 13-14 min 50 to 30% A, 14-16 min, 30 to 0% A; 16-20 min, 0 to 95% A; and 20-25 min, 95% A.

3. Results and discussion

3.1 Production and purification of serine carboxypeptidase-like (SCPL) protein and releasing of alcohol oxidase 1 (AOX1)

The optimized genes of rice SCPLs were inserted to the pPICZαBNH8, resulting two constructs of pPICZαBNH8/OsSCPL2a and pPICZαBNH8/OsSCPL7, which encode a yeast alpha-factor secretion signal, followed by KEX2 and STE13 cleavage sites for removal of the secretory signal sequence, then an 8-histidine affinity tag in frame with the predicted N-terminus of the SCPL protein. The plasmids were transformed into the protease-deficient SMD1168H strain of *Pichia* for protein production. The heterologous gene expression was induced from the *AOX1* promoter with 1% (v/v) methanol for 5 days and the medium used to purify protein, after adjusting the pH of the medium to 7.5 to allow binding to the IMAC column.

Protein purification was performed by IMAC. After collecting all fractions, the protein samples were observed with SDS-PAGE, in which only a light band at 45 kDa and smear was seen for OsSCPL2a, but discrete bands of 75 and 45 kDa were seen in crude and flow-through fractions from IMAC purification of OsSCPL7 (Figure 1B and C). However, after washing with Eq buffer and low concentrations of imidazole (10 and 20 mM, respectively), the band of approximately 75 kDa was obtained during expression of OsSCPL7, but not in OsSCPL2a. After elution with 100 mM and 250 mM imidazole in Eq buffer, a smaller band around 50 kDa appeared, which was presumed to be the expected protein size. The protein band was clearly observed in OsSCPL7. Meanwhile, a protein band at 45 kDa was consistently yet poorly observed in OsSCPL2a expression.

For Os4BGlu11 expression from the same vector, the methanol concentration used in induction was reduced to 0.5% to reduce the stress that might lead to release of extra proteins. In addition, the metal on the IMAC resin was changed from nickel to cobalt, which in our experience binds less tightly but more specifically to His-tagged proteins. Nevertheless, bands similar to those found with OsSCPL7 were observed, although the main bands were at approximately 70 and 75 kDa and the 45 kDa protein band was more prominent in the IMAC elution fractions (Figure 1D). These bands eluted at lower imidazole concentration from the cobalt resin compared to the similar bands in the nickel resin purification of OsSCPL7.

OsSCPL7 fractions were divided into two parts, the washed fraction and elution fraction, before concentrating and digesting with endoglycosidase-H (EC 3.2.1.96) overnight to remove any carbohydrate and see whether it is a glycoprotein, which may occur by glycosylation process. Glycosylation is a common post-translational modification in eukaryotic cells and closely related to the folding, stability and activity of proteins. The abnormal modification by glycosylation may affect protein function and yeasts are known to add larger high-mannose glycans, rather than the smaller but more complex N-linked glycans found on glycoproteins in higher eukaryotes [23]. After deglycosylation, one major band was observed at 75 kDa, while another at around 50 kDa in the undigested protein disappeared (Figure 1C). As explained above, the expected OsSCPL protein size was around 50 kDa. The endoglycosidase H protein also appears in the SDS-PAGE (29 kDa). The loss of the 50 kDa band upon deglycosylation may have been due to protease contamination that digested the protein during the Endo-H

digest, since Endo-H digest is expected to decrease the size of the protein and no prominent lower molecular weight bands were observed.

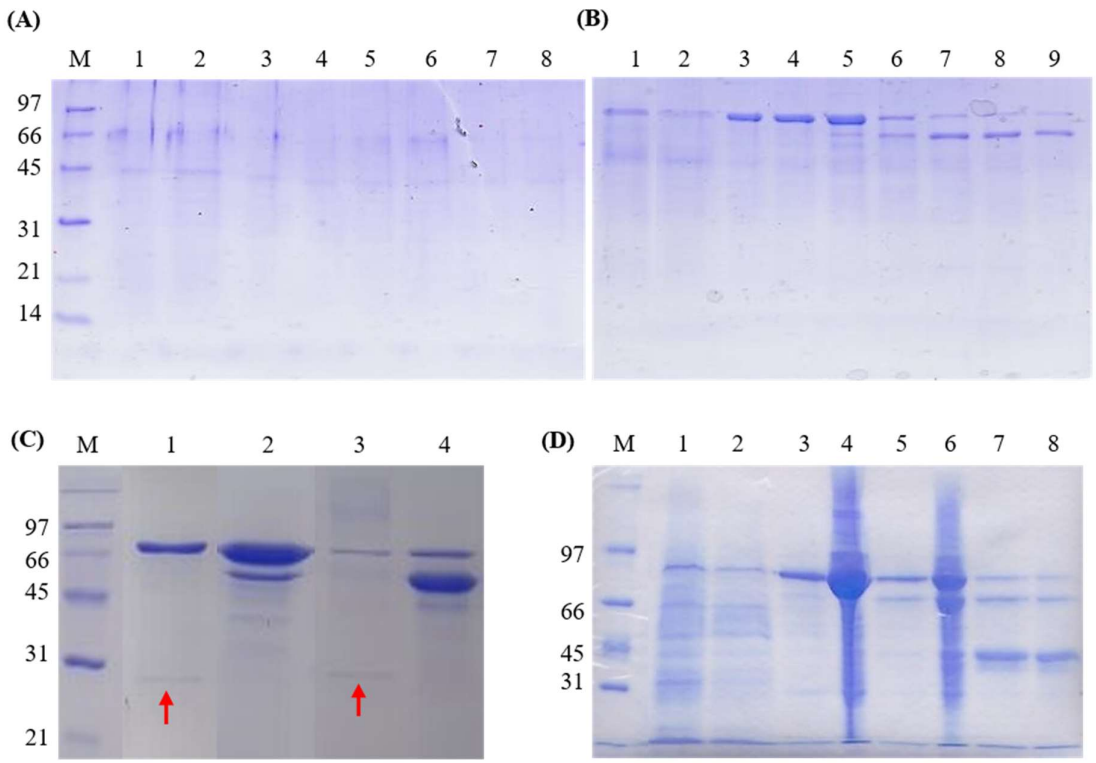

Figure 1 SDS-PAGE results from protein purification. Protein bands are seen in OsSCPL2a (A) and OsSCPL7 (B) expression. Only in OsSCPL7 expression, the 75 kDa band consistently appeared in elution fractions. Lane M, marker; 1, crude; 2, flow-through; 3, wash with Eq buffer; 4 & 5, washes with 10- and 20 mM imidazole; 6 & 7, elution with 100 mM imidazole; 8 & 9, elution with 250 mM imidazole. (C) The SDS-PAGE gel of OsSCPL7 from IMAC purification and digestion with endoglycosidase-H. Lane M, marker; 1 & 2, digested and undigested wash fraction pool, 3 & 4, digested and undigested elution fraction pool. The endoglycosidase-H band is seen at 29 kDa (red arrow). (D) Expression and purification of Os4BGlu11. Lane M, marker; 1, crude medium; 2, flow through; 3 & 4, pool of washes with 10 mM imidazole and its concentrated fraction; 5 & 6, pool of washes with 20 mM imidazole and its concentrated fraction; 7 & 8, fractions eluted with 250 mM imidazole.

3.2 Identification of expressed protein by LC/MS/MS

After purification by IMAC, the protein was separated by SDS-PAGE and the gel stained to visualize the protein band. The in-gel trypsin-digests of OsSCPL7 protein and concentrated fraction of the Os4BGlu11 wash with 10 mM imidazole were analyzed with LC/MS/MS, and peptide masses were used to search databases with Mascot (http://matrixscience.com). The best hit, with a score of 457, was identified as alcohol oxidase 1 (AOX1) protein from *K. phaffii* instead of our expected target protein OsSCPL7 (Figure 2). *K. phaffii* is another name for the *Pichia* strain that is commonly used in gene expression studies [24]. Fourteen peptides and nine sequences were matched with this protein, covering 23% of the total of 663 protein residues (Table 1). In addition, tryptic peptide LC/MS/MS data from Os4BGlu11 matched the same *Pichia* AOX1 protein with a hit score of 1176, including 27 matched peptides (Table 2).

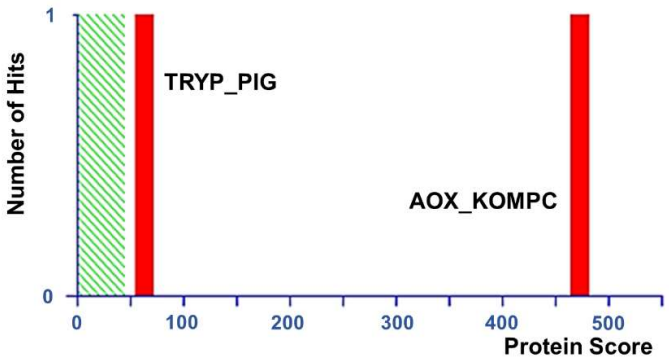

Figure 2 Protein score distribution based on LC/MS/MS result from in-gel-trypsin-digest of OsSCPL7. The protein hit with the score at 457 was identified as alcohol oxidase 1 from *K. phaffii* (strain ATCC 76273/ CBS 7435/ CECT 11047/ NRRL Y-114, accession ID: XP_002494271.1). In addition, the trypsin used to digest the protein was detected as trypsin *Sus scrofa* (ID: XP_020934119.1), with a hit score of 87 for this contaminant.

P. pastoris AOX1 is normally found in the peroxisome, where it helps to metabolize methanol and other short chain alcohols [5]. The appearance of AOX1 in the medium is likely an indication of stress. The two OsSCPLs were expressed under identical conditions, but release of AOX1 in the medium was only observed in clones expressing OsSCPL7, while expression of Os4BGlu11 also resulted in AOX1 in the medium. Interestingly, AOX1 protein in the medium bound to the IMAC column and was found in every elution fraction, possibly reflecting the fact that it is a metalloenzyme. In fact, IMAC resin bound to different metal ions has been used to fractionate metalloproteins from cyanobacterium cell extracts, with some proteins showing binding to IMAC resin loaded with either Ni^{2+} or Co^{2+} ions [25]. Although AOX1 in *P. pastoris* is known as an SCAO involved in methanol metabolism and is naturally an intracellular enzyme [2], AAO and SAO are mostly secreted to the medium [4]. AAO (EC 1.1.3.7) has molecular weights around 63.5 to 78 kDa, suggesting that it could be included in the band from SDS-PAGE gel (around 75 kDa), although the tryptic peptide MS analysis only identified AOX1. In fact, we could find no study that has yet reported that AAO can be found in *P. pastoris* nor a study reporting the trouble shooting to avoid the releasing of AOX1 or any other protein background from the host cell.

Table 1 LC/MS/MS result of OsSCPL7 purification showed fourteen matched peptides to AOX1 protein from *K. phaffii*. Individual ion score > 45 indicate identity or extensive homology.

Observed	Mr (expt)[1]	Mr (calc)[1]	Score	Before [2]	Peptide	After[2]
900.2	2697.7	2697.4	27	K	VGLIEAGENNLNNPWVYLPGIYPR	N
583.6	1747.8	1747.8	46	K	TASFYTSNPSPHLNGR	R
875.0	1748.0	1747.8	87	K	TASFYTSNPSPHLNGR	R
838.4	1674.8	1674.7	73	R	GSASDYDDFQAEGWK	T
838.5	1674.9	1674.7	78	R	GSASDYDDFQAEGWK	T
834.9	1667.8	1667.8	73	R	ACNNPDIHGFEGPIK	V
545.3	1088.6	1088.6	40	R	SGFGDPIKLR	A
724.4	1446.9	1446.9	54	R	AAGVKPLVNLPGVGR	N
483.3	1446.9	1446.9	42	R	AAGVKPLVNLPGVGR	N
594.3	1779.8	1779.7	50	R	NFQDHYCFFSPYR	I
890.9	1779.8	1779.7	56	R	NFQDHYCFFSPYR	I
803.5	2407.6	2407.2	72	R	VFDQWYANGTGPLATNGIEAGVK	I
1204.7	2407.4	2407.2	60	R	VFDQWYANGTGPLATNGIEAGVK	I
899.5	2695.3	2695.3	55	K	VGDLSVCPDNVGCNTYTTALLIGEK	T

[1]Note: Mr (expt), expected mass; Mr (calc), calculated mass. [2]"Before" means the amino acid that occurs before the peptide in the protein sequence (K, lysine; R, Arginine); "After" means the amino acid that occurs after the peptide in the peptide sequence (N, Asparagine; T, Threonine; V, Valine; A, Alanine; I, Isoleucine). The amino acid residues before and after the peptides and peptide sequences are indicated in standard one-letter codes for amino acids.

Table 2 LC/MS/MS result showed the purification of Os4Bglu11 matched with 27 peptides to AOX1 *K. phaffii*.

Observed	Mr (expt)[1]	Mr (calc)[1]	Score	Before [2]	Peptide	After[2]
401.2	800.4	800.4	41	K	IIVEDGR	A
410.7	819.4	819.4	23	R	SGFGDPIK	L
439.7	877.5	877.5	22	K	AIENYIR	E
470.2	938.5	938.5	37	K	WGGVLDHR	S
487.2	972.6	972.6	25	K	DLLPLMKK	T
505.8	1009.6	1009.6	21	R	LANLDHSLK	V
527.8	1053.6	1053.6	63	R	TVPSKPLNAK	K
572.3	1142.6	1142.6	64	K	VDKIIVEDGR	A
482.6	1444.7	1444.7	79	R	SDSAHAFVHSTMR	N
723.3	1444.9	1444.7	47	R	SDSAHAFVHSTMR	N
724.4	1446.9	1446.9	72	R	AAGVKPLVNLPGVGR	N
731.3	1460.7	1460.7	64	R	SDSAHAFVHSTMR	N
487.9	1460.7	1460.7	66	R	SDSAHAFVHSTMR	N
534.6	1600.8	1600.8	55	R	RSDSAHAFVHSTMR	N
548.6	1642.8	1642.8	93	R	IKPQYESFDDFVR	G
874.9	1747.8	1747.8	92	K	TASFYTSNPSPHLNGR	R
583.6	1747.8	1747.8	81	K	TASFYTSNPSPHLNGR	R
634.4	1899.8	1899.8	58	K	YMTMFHFLEYPFSR	G
950.9	1899.8	1899.8	56	K	YMTMFHFLEYPFSR	G
1204.6	2407.2	2407.2	89	K	VFDQWYANGTGPLATNGIEAGVK	I
803.4	2407.2	2407.2	35	K	VFDQWYANGTGPLATNGIEAGVK	I
804.7	2411.1	2411.1	74	K	IRPTPEELSQMDESFQEGYR	E
1214.6	2427.1	2427.1	27	K	IRPTPEELSQMDESFQEGYR	E
810.0	2427.1	2427.1	92	K	IRPTPEELSQMDESFQEGYR	E
839.1	2514.1	2514.1	96	K	MDHFAGEVTSHHPLFPYSSEAR	A
938.1	2811.2	2811.2	127	R	NEGHVTSNQVELHPDIEYDEEDDK	A
772.6	3086.5	3086.4	26	R	EYFEDKPDKPVMHYSIIAGFFGDHTK	I

[1]Note: Mr (expt), expected mass; Mr (calc), calculated mass. [2]"Before" means the amino acid that occurs before the peptide in the protein sequence (K, lysine; R, Arginine); "After" means the amino acid that occurs after the peptide in the peptide sequence (A, Alanine; L, Leucine; E, Glutamic acid; S, Serine; T, Threonine; V, Valine; N, Asparagine; G, Glycine; I, Isoleucine; E, Glutamic acid). The amino acid residues before and after the peptides and peptide sequences are indicated in standard one-letter codes for amino acids.

3.3 The AOX activity towards cyanidin-3-O-glucoside (Cy3G)

Reactions of the purified protein with Cy3G were conducted before the protein was identified as AOX1. Cy3G was detected by its absorbance at 280 and 520 nm. Anthocyanins, including Cy3G, have UV-visible absorption spectra with λ_{max}= 280 nm in the UV range and λ_{max}= 520 nm in the visible range [26]. The standard Cy3G was detected at a retention time of 7.6 min in absorbance 520 nm, while two peaks were observed at 7.6 and 8.0 min in 360 nm (Figure 3). In the reaction containing the enzyme, extra peaks were observed at the retention times around 9.5 min with absorbance at 360 nm (Figure 3A). This result indicates that there was a product from AOX1 activity, which presumably oxidized C3G. The additional peaks were not detected in 520 nm absorbance, at which wavelength a peak would be expected for acylated anthocyanin, although the peak area for C3G was significantly decreased (Figure 3B). The additional peaks were observed in absorbance 360 nm suggesting that AOX protein may catalyze the oxidation reaction that changes the C3G structure so that it is no longer absorbs at 520 nm. Because only a small amount of this compound was produced, we could not determine the structure to verify the nature of the reaction.

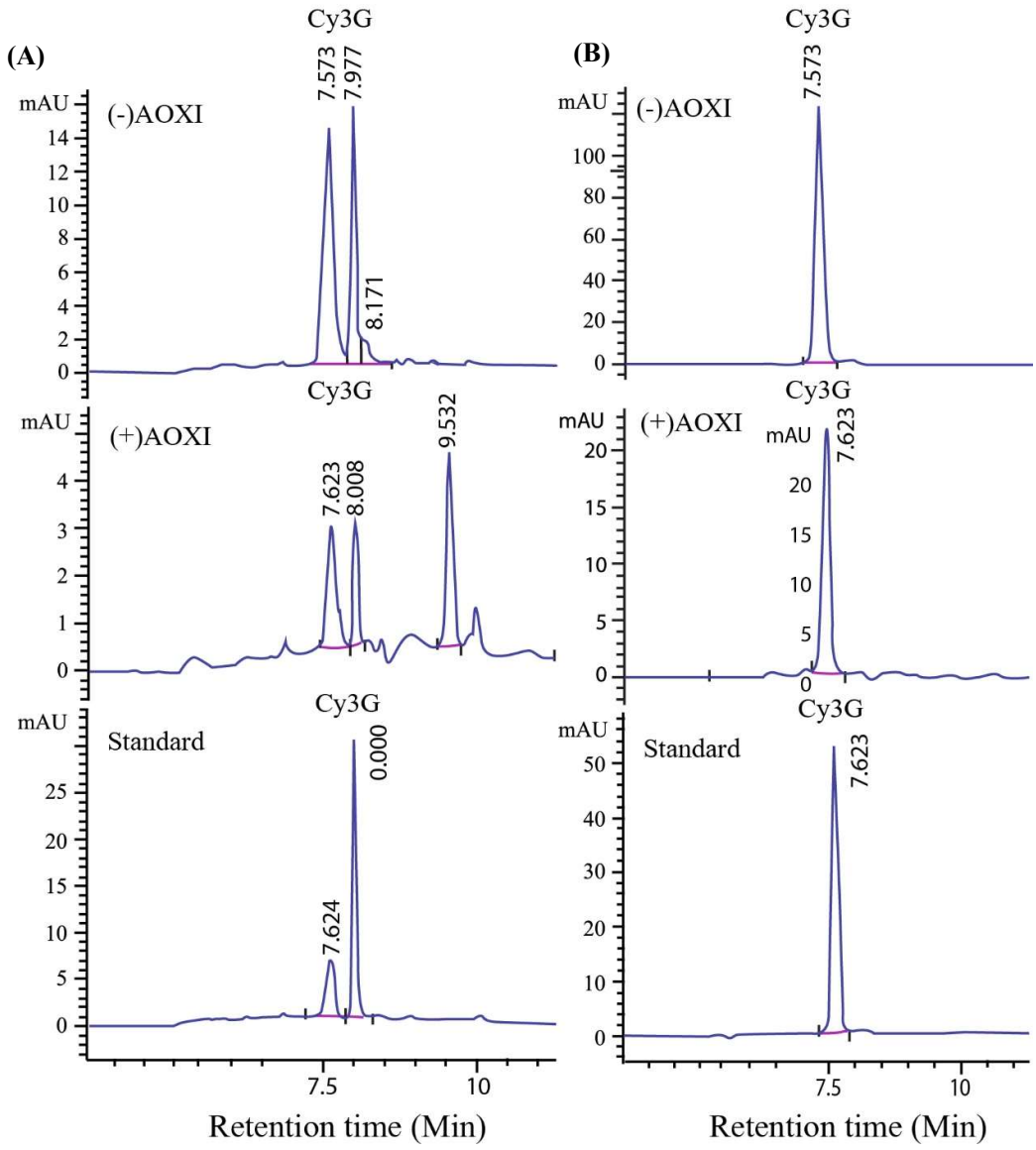

Figure 3 Ultra-high-performance liquid chromatography (UHPLC) chromatograms of AOX1 activity upon cyanidin-3-*O*-glucoside (Cy3G) with detection by absorbance at 360 nm (A) and 520 nm (B), respectively. In each case, the top chromatogram is the control reaction without enzyme, the middle is the reaction with AOX1 enzyme and the bottom is the Cy3G standard.

It is not clear why expression of certain enzymes cause release of peroxisomal AOX1, while that of similar enzymes do not. Yeast cell walls are composed of about 50-60% glucan, which is mainly β-1,3-linked glucan synthesized at the cell surface [27], so in principle Os4BGlu11 β-glucosidase, if it acts on such β-glucans or their precursors, could compromise the cell wall. However, several β-glucosidases and exoglucanases have been produced in *Pichia pastoris*, for example [14], and *Pichia* produces its own exoglucanases that can act on 1,3-β-glucan [28]. It might also cause general cell stress in some other way. It is not clear how an acyltransferase would disrupt the cell membrane, cell wall or cause secretion of peroxisomal contents, but it may be that OsSCPL7 has other unknown activities that are not found in SCPL2a or causes greater stress during production. Since the proteins are closely related and produced under the same conditions, it suggests this property is somehow intrinsic to the expression of specific genes to produce their proteins.

4. Conclusion

AOX1 protein from *P. pastoris* was found to be released into the medium upon induction of expression to produce rice protein SCPL7 and Os4BGlu11, but not OsSCPL2a, and was purified with IMAC resin bound with either Ni^{2+} or Co^{2+}. The activity of AOX1 protein may have modified Cy3G, which may affect the stability of this anthocyanin. Further study is needed to characterize the AOX1 protein substrate specificity and the structure of the new product generated from Cy3G. Researchers expressing proteins in *P. pastoris* should be aware of the potential of AOX1 to be released into the medium and purified by IMAC, as well as its potential for catalysis of unexpected oxidation reactions.

5. Acknowledgements

This work was supported by the Thailand Research Fund International Research Network IRN62W0004 and SUT-PhD scholarship SUT-PhD/06/2556 (National Research University Project of the Committee on Higher Education to Suranaree University of Technology).

6. Conflict of interest

The authors declare that they have no conflict of interest.

7. References

[1] Cregg JM, Cereghino JL, Shi J, Higgins DR. Recombinant protein expression in *Pichia pastoris*. Mol Biotechnol. 2000;16(1):23-52.

[2] Koch C, Neumann P, Valerius O, Feussner I, Ficner R. Crystal structure of alcohol oxidase from *Pichia pastoris*. PLoS One. 2016;11(2):e0149846.

[3] Kiess M, Hecht HJ, Kalisz HM. Glucose oxidase from *Penicillium amagasakiense*: primary structure and comparison with other glucose-methanol-choline (GMC) oxidoreductases. Eur J Biochem. 1998;252(1): 90-99.

[4] Goswami P, Chinnadayyala SS, Chakraborty M, Kumar AK, Kakoti A. An overview on alcohol oxidases and their potential applications. Appl Microbiol Biotechnol. 2013;97(10):4259-4275.

[5] Ozimek P, Veenhuis M, van der Klei IJ. Alcohol oxidase: a complex peroxisomal, oligomeric flavoprotein. FEMS Yeast Res. 2005;5(11):975-983.

[6] Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat, B, The carbohydrate-active enzymes database (CAZy) in 2013, Nucleic Acids Res. 2014; 42(D1):D490-D495.

[7] Ketudat Cairns JR, Esen A. β-Glucosidases. Cell Mol Life Sci. 2010;67:3389-3405.

[8] Li AX, Steffens JC. An acyltransferase catalyzing the formation of diacylglucose is a serine carboxypeptidase-like protein. Proc Natl Acad Sci USA. 2000;97(12):6902-6907.

[9] Khoo HE, Azlan A, Tang ST, Lim SM. Anthocyanidins and anthocyanins: colored pigments as food, pharmaceutical ingredients, and the potential health benefits. Food Nutr Res. 2017;61(1):1361779.

[10] Yoshida K, Mori M, Kondo T. Blue flower color development by anthocyanins: from chemical structure to cell physiology. Nat Prod Rep. 2009;26(7):884-915.

[11] Fraser CM, Thompson MG, Shirley AM, Ralph J, Schoenherr JA, Sinlapadech T, et al. Related arabidopsis serine carboxypeptidase-like sinapoylglucose acyltransferases display distinct but overlapping substrate specificities. Plant Physiol. 2007;144(4):1986-1999.

[12] Noda N, Higaki T, Sano T, Kazuma K, Sasaki T, Hasegawa S, et al. The structural localization and substrate specificity of 1-O-acylglucose dependent anthocyanin acyltransferase from blue petals of butterfly pea (*Clitoria ternatea*). Plant Cell Physiol. 2007;48:S222.

[13] Nishizaki Y, Yasunaga M, Okamoto E, Okamoto M, Hirose Y, Yamaguchi M, et al. p-Hydroxybenzoyl-glucose is a Zwitter donor for the biosynthesis of 7-polyacylated anthocyanin in delphinium. Plant Cell. 2013;25(10):4150-4165.

[14] Toonkool P, Metheenukul P, Sujiwattanarat P, Paiboon P, Tongtubtim N, Ketudat-Cairns M, et al. Expression and purification of dalcochinase, a β-glucosidase from *Dalbergia cochinchinensis* Pierre, in yeast and bacterial hosts. Protein Expr Purif. 2006;48:195-204.

[15] Zhou A, Jiang X, Xu X. Improved alkaline lysis method for rapid isolation of plasmid DNA. Biotechniques. 1997;23(4):592-594.

[16] Kotchoni SO, Gachomo EW, Betiku E, Shonukan OO. A home made kit for plasmid DNA mini-preparation. Afr J Biotechnol. 2003;2(4):88-90.

[17] Kumar R. Simplified protocol for faster transformation of (a large number of) *Pichia pastoris* strains. Yeast. 2019;36(6):399-410.
[18] Inan M, Aryasomayajula D, Sinha J, Meagher MM. Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. Biotechnol Bioeng. 2006;93(4):771-778.
[19] Hays FA, Roe-Zurz Z, Stroud RM. Overexpression and purification of integral membrane proteins in yeast. Methods Enzymol. 2010;470:695-707.
[20] Shevchenko A, Wilm M, Vorm O, Mann M. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal Chem. 1996;68(5):850-858.
[21] Wall MJ, Crowell AM, Simms GA, Liu F, Doucette AA. Implications of partial tryptic digestion in organic-aqueous solvent systems for bottom-up proteome analysis. Anal Chim Acta. 2011;703(2):194-203.
[22] Sari DRT, Paemanee A, Roytrakul S, Ketudat CJR, Safitri A, Fatchiyah F. Black rice cultivar from Java Island of Indonesia revealed genomic, proteomic, and anthocyanin nutritional value. Acta Biochim Pol. 2021;68:5386.
[23] Cereghino JL, Cregg JM. Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. FEMS Microbiol Rev. 2000;24(1):45-66.
[24] Kurtzman CP. Biotechnological strains of *Komagataella (Pichia) pastoris* are *Komagataella phaffii* as determined from multigene sequence analysis. J Ind Microbiol Biotechnol. 2009;36(11):1435.
[25] Barnett JP, Scanlan DJ, Blindauer CA. Fractionation and identification of metalloproteins from a marine cyanobacterium. Anal Bioanal Chem. 2012;402:3371-3377.
[26] Qin C, Li Y, Niu W, Ding Y, Zhang R, Shang X. Analysis and characterization of anthocyanins in mulberry fruit. Czech J Food Sci. 2010;28(2):117-126.
[27] Bowman SM, Free SJ. The structure and synthesis of the fungal cell wall. BioEssays. 2006;28:799-808.
[28] Xu Z, Shih MC, Poulton JE. An extracellular exo-β-(1,3)-glucanase from *Pichia pastoris*: purification, characterization, molecular cloning, and functional expression. Protein Expr Purif. 2006;47:118-127.