Mutation of cysteine residues increases heterologous expression of peach expansin in the methylotrophic yeast *Pichia pastoris*

Kaori Matsuyama¹, Naoki Sunagawa¹, Kiyohiko Igarashi^{1,2,*}

¹Department of Biomaterial Sciences, Graduate School of Agricultural and Life Sciences, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113-8657, Japan; ²VTT Technical Research Centre of Finland, PO Box 1000, Tietotie 2, Espoo FI-02044 VTT, Finland

*E-mail: aquarius@mail.ecc.u-tokyo.ac.jp Tel: +81-3-5841-5258 Fax: +81-3-5841-5273

Received June 4, 2020; accepted July 13, 2020 (Edited by T. Demura)

Abstract The study of Carbohydrate-Active enZymes (CAZymes) associated with plant cell wall metabolism is important for elucidating the developmental mechanisms of plants and also for the utilization of plants as a biomass resource. The use of recombinant proteins is common in this context, but heterologous expression of plant proteins is particularly difficult, in part because the presence of many cysteine residues promotes denaturation, aggregation and/or protein misfolding. In this study, we evaluated two phenotypes of methylotrophic yeast *Pichia pastoris* as expression hosts for expansin from peach (*Prunus persica* (L.) Batsch, PpEXP1), which is one of the most challenging targets for heterologous expression. cDNAs encoding wild-type expansin (PpEXP1_WT) and a mutant in which all cysteine residues were replaced with serine (PpEXP1_CS) were each inserted into expression vectors, and the protein expression levels were compared. The total amount of secreted protein in PpEXP1_WT culture was approximately twice that of PpEXP1_CS. However, the amounts of recombinant expansin were 0.58 and 4.3 mg l⁻¹, corresponding to 0.18% and 2.37% of total expressed protein, respectively. This 13-fold increase in production of the mutant in *P. pastoris* indicates that the replacement of cysteine residues stabilizes recombinant PpEXP1.

Key words: CAZymes, cysteine, expansin, heterologous expression, *Pichia pastoris*.

Introduction

The cell walls of land plants have a complex structure, consisting of a primary cell wall built around the plasma membrane and subsequently reinforced by a secondary cell wall. The primary cell wall is associated with cell formation, and consists mainly of polysaccharides and proteoglycans (Burton et al. 2010; Somerville et al. 2004). Polysaccharide regions are formed from cellulose, hemicellulose, and pectin. The cellulose is mainly crystalline cellulose consisting of β -1,4-glucan, whereas hemicellulose contains many kinds of structural polysaccharides, including mannans, xyloglucans, xylans, and galactans. Hemicellulose serves to crosslink cellulose to provide greater wall strength and flexibility. Pectin is composed of homogalacturonan (HG), rhamnogalacturonan-I (RG-I), and rhamnogalacturonan-II (RG-II). The main chain of HG is α -1,4-linked GalA, and most of the GalA residues are methylated. RG-I is composed of a repeating [α -1,4-

GalA- α -1,2-Rha-] main chain and side chains such as galactans and arabinans. RG-II has very complex structure, with a main chain of α -1,4-bound GalA, and many kinds of side chains (Burton et al. 2010; Somerville et al. 2004). The secondary cell wall is formed to thicken the inner primary cell wall after cell elongation, and is composed of cellulose, hemicellulose, and lignin. Lignin is a phenolic high polymer, poorly soluble in water and resistant to microbial enzymes, thereby contributing to chemical resistance and physiological strength (Cosgrove and Jarvis 2012).

Plants construct and reconstruct these walls using many kinds of enzymes, known as carbohydrate active enzymes (CAZymes, <http://www.cazy.org>, Lombard et al. 2014), to form and/or degrade covalent links or interactions between the various polymers. Thus, in order to understand the structure, role, and biosynthesis or remodeling of the cell wall, it is necessary to characterize the CAZymes. Furthermore, plant cell walls represent the most abundant carbon source on earth, and

Abbreviations: EGF, extra glycerol feed; GF, glycerol feed; MF, methanol feed; PpEXP1_CS, mutant in which all cysteine residues of peach expansin are replaced with serine; PpEXP1_WT, wild-type peach expansin.

This article can be found at <http://www.jspcmb.jp/>

Published online November 17, 2020

an understanding of CAZymes is essential for utilization of this sustainable resource. Therefore, substantial amounts of plant enzymes are required for research purposes. However, purification of the native enzymes from plants is usually inefficient, and consequently heterologous protein expression has attracted great interest. There are many types of expression host, such as *Escherichia coli*, *Saccharomyces cerevisiae*, *Pichia pastoris*, and insect cells. In this study, we focused on the methylotrophic yeast *P. pastoris*, which has similar molecular, genetic, and biochemical features to higher eukaryotes (Cereghino and Cregg 2000). The production level of recombinant proteins in *P. pastoris* is generally high, often reaching g l^{-1} quantities in the culture medium (Higgins and Cregg 2007), although plant proteins represent challenging targets, because their high contents of cysteine residues can result in denaturation, aggregation and/or protein misfolding (Nielsen et al. 2018). Some strategies are available to overcome this problem. For example, in *E. coli*, co-expression of disulfide isomerase increased soluble protein production (Liu et al. 2005; Yuan et al. 2004). Also, replacement of cysteine residues with serine increased productivity in an *S. cerevisiae* expression system (Ito et al. 2011).

Many kinds of *P. pastoris* host cells are available, with different genotypes, nutritional requirements, levels

of resistance to methanol, and deletion of protease to prevent the degradation of recombinant protein. Strain SMD1168H is one of major protease-deficient strains (Ahmad et al. 2014; Higgins and Cregg 2007). *P. pastoris* is methylotrophic yeast, but a methanol utilization positive (Mut^+) strain can produce excessive amounts of toxic formaldehyde, and therefore a methanol utilization slow (Mut^s) strain, such as KM71H, whose alcohol oxidase gene (*aox1*) is knocked out (Higgins and Cregg 2007), is often preferred.

Several vectors are commercially available for *Pichia* expression. For example, the pPIC and pGAP series (Thermo Fischer) contain methanol-inducible and constitutive promoters, respectively. pPIC vector carries the alcohol oxidase promoter *AOX1*, while pGAP contains the glyceraldehyde-3-phosphate dehydrogenase (GAP) promoter, which is constitutively expressed in the growth phase (Higgins and Cregg 2007). Extra- or intra-cellular production can be achieved through the incorporation or omission of the *Saccharomyces cerevisiae* α -factor sequence, respectively.

Expansin is cell wall localized protein which plays a role in plant cell wall loosening (Cosgrove 2000, 2016). They induce pH-dependent “acid growth” and cell expansion (Li et al. 1993; McQueen-Mason et al. 1992; Shcherban et al. 1995). In plantae, all land plants and

Figure 1. Amino acid sequence of *PpEXP1* and the positions of mutated cysteine residues. A: Amino acid sequence alignment of *PpEXP1_WT* and *PpEXP1_CS* was done with ENBOSS Needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/, Madeira et al. 2019). *PpEXP1_WT* has eight cysteines (GenBank accession number BAB19676.1), which are shown in red in this figure. In this study, we mutated all of them to serine. B: Three-dimensional structure model of *PpEXP1_WT*, built by SWISS-MODEL (<https://swissmodel.expasy.org/>; Benkert et al. 2011; Bertoni et al. 2017; Bienert et al. 2017; Guex et al. 2009; Waterhouse et al. 2018). Cyan, red, blue, and yellow indicate carbon, oxygen, nitrogen, and sulfur atoms of cysteine residues, respectively. The structure of β -expansin from *Zea mays* (*ZmEXPB1*, PDB ID: 2HCZ) was used as a template. Sequence similarity of *PpEXP1_WT* and *ZmEXPB1* was 34.12%, and coverage was 84.06%. All structural figures were drawn by using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

some algae express many expansin genes in various organs, for instance seeds, leaves, stems, flowers, fruits, pollens, and so on. A number of previous studies have shown that expansins are important for plant growth, development, and maturation, but the molecular mechanism of their function remains unclear (Cosgrove 2015, 2016; Marowa *et al.* 2016; Nikolaidis *et al.* 2014; Sampedro and Cosgrove 2005; Sharova 2007). Although the amino acid sequences of expansins are similar to those of endoglucanases, which degrade cellulose, their hydrolytic activity and molecular mechanism have not been established (Cosgrove 2000; Yennawar *et al.* 2006). Therefore, there is a need for recombinant expansins. However, expansins are among the most difficult proteins to produce heterologously (Cosgrove 2015; Kerff *et al.* 2008; Yennawar *et al.* 2006), and additionally, it is said to precipitate at concentrations above $10 \mu\text{g ml}^{-1}$, especially for α -expansin (Cosgrove 2015). Only one α -expansin, from tomato (*Solanum lycopersicon*, LeEXP2), has so far been successfully expressed in *P. pastoris*. Expansins contain many cysteine residues and some of them are predicted not to form disulfide bonds. For instance, fruit-ripening-related α -expansin from peach (*Prunus persica*, (L.) Batsch, PpEXP1) has eight cysteine residues among the 227 residues of the mature protein (Hayama *et al.* 2000, 2003, 2006), and based on structure modeling, they localize on the surface (Figure 1). In this study, we aimed to examine whether mutation of all the cysteine residues in PpEXP1 to serine would stabilize the recombinant protein and lead to efficient methanol-induced production of soluble mutant PpEXP1 in *P. pastoris*.

Materials and methods

Construction of expression systems

The cDNA of wild-type α -expansin from peach fruit (PpEXP1, GenBank accession number is AB029083) was a kind gift from Dr. Hiroko Hayama of the National Agriculture and Food Research Organization and Prof. Megumi Ishimaru of Kindai University. The cDNA fragment encoding PpEXP1 excluding the signal peptide was amplified by means of polymerase chain reaction (PCR) with specific primers (Table 1), and *EcoRI* and *XbaI* restriction sites were created at the 5' and 3' ends of the open reading frame, respectively. The amplified cDNA fragments were cloned into restriction-digested pGAPZ α A

vector. cDNA encoding a mutant in which all the cysteine residues of PpEXP1 were changed to serine (PpEXP1_CS), and with codons optimized for *P. pastoris*, was synthesized based on the cDNA sequence of PpEXP1_WT using a commercial service (Genscript, Inc., New Jersey, USA). The synthesized DNA was amplified and cloned into pPICZ α A vector (Invitrogen Corporation, Massachusetts, USA) by the in-fusion method, using the primers shown in Table 1.

pGAPZ α A_PpEXP1_WT and pPICZ α A_PpEXP1_CS were transformed into *E. coli* strain TOP10 (Invitrogen), then extracted, and linearized using *BglII* or *Bpu1102I*, respectively (TaKaRa Bio Inc., Shiga, Japan). The linearized vectors were transformed into *P. pastoris* SMD1168H and KM71H, respectively (Invitrogen) by electroporation. After transformation, the cells were plated on YPDS agar plates containing zeocin ($100 \mu\text{g ml}^{-1}$) and incubated at 30°C for 5 days.

Expression check

The protein expression of colonies on YPDS-zeocin plates was confirmed on a small scale. We selected eight SMD1168H_pGAPZ α A_PpEXP1_WT colonies and eight KM71H_pPICZ α A_PpEXP1_CS colonies with zeocin and pre-cultured them overnight at 30°C in 3 ml of YPD medium containing zeocin (final concentration: $50 \mu\text{g ml}^{-1}$). One milliliter aliquots of the pre-cultures of SMD1168H_pGAPZ α A_PpEXP1_WT were inoculated in 200 ml of basal salt medium (described in the Supplementary information) and cultured at 30°C for 4 days with shaking at 150 rpm. During incubation, 50% (w/v) glycerol was added every 24 h so that the final concentration was 2% (w/v). In contrast, 1 ml of KM71H_pPICZ α A_PpEXP1_CS pre-culture medium was inoculated into 200 ml of YPD medium (described in the Supplementary information) and incubated overnight at 30°C with shaking at 150 rpm. The cultures were centrifuged at $2,000 \times g$ for 10 min. Each supernatant was collected, and the cell pellet was suspended in sterile ultrapure water and centrifuged again. The resulting pellet was resuspended in YP medium (described in the Supplementary information), and transgene expression was induced with methanol (final concentration: 1% v/v). Incubation was carried out at 26.5°C for 4 days with shaking at 150 rpm. During the induction period, methanol was added every 24 h to a final concentration of 1% (w/v).

Each supernatant was collected by centrifugation at $7,000 \times g$ for 10 min after the end of the culture. The total amount of

Table 1. Macromolecule production information.

	Wild type (WT)	Cysteine to Serine mutant (CS)
Source organism	<i>Prunus persica</i> (L.) Batsch fruit of "Akatsuki"	<i>Prunus persica</i> (L.) Batsch fruit of "Akatsuki"
DNA source	cDNA	Gene synthesis and codon optimization
Forward primer	ATGAATTCGACTATGGAGGAGGATGGGAAG	AGGGGTATCTCTCGAGAAAAGAGATTACGGTGGTGGT
Reverse primer	GCTCTAGAGCGAATTGACCGCCCGAAAAATG	AGAAAGCTGGCGCGCGCTTAAATTTGACCACCTGAAA
Ligation method	Restriction enzyme and ligation	In-fusion
Expression vector	pGAPZ α A	pPICZ α A
Expression host	<i>P. pastoris</i> SMD1168H	<i>P. pastoris</i> KM71H

expressed protein was determined by Bradford's method using Bio-Rad protein assay reagent (Bio-Rad Laboratories, Inc., Hercules, California, USA) and bovine serum albumin as a standard, and expression of the induced protein was confirmed by SDS-PAGE on 15% acrylamide gel.

Production of recombinant protein in a jar fermenter

Recombinant protein was produced in a 5 l scale mini-jar fermenter (TSC-M5L; Takasugi Seisakusho, Tokyo, Japan). A single colony of SMD1168H_pGAPZαA_PpEXP1_WT or KM71H_pPICZαA_PpEXP1_CS was inoculated in 5 ml of YPG medium and cultured overnight at 30°C with stirring at 300 rpm. Three ml of the preculture was added to 2 l of basal salt medium containing PTM1 salt adjusted to pH 5.0 and incubated at 27°C. The pH of the medium was fixed at pH 5.0 with 10% ammonium solution. After one and a half days of incubation, 50% (w/v) glycerol was added (glycerol feed; GF) at approximately 20 g per hour to make a total of 180 g. At the end of GF, the supply of 100% glycerol (extra glycerol feed; EGF) or 100% methanol (methanol feed; MF) was started for SMD1168H_pGAPZαA_PpEXP1_WT or KM71H_pPICZαA_PpEXP1_CS, respectively, under regulation by the oxygen controller. The supernatant was collected by centrifuging the culture medium 96 h after the start of EGF for SMD1168H_pPICZαA_PpEXP1_WT and 72 h after the start of MF for KM71H_pGAPZαA_PpEXP1_CS. Each supernatant was applied to a 100 kDa cut-off ultrafilter (Kvick Lab Cassette, GE Healthcare, Chicago, USA, for PpEXP1_WT) or a 300 kDa cut-off ultrafilter (Pellicon XL Cassette, Merck Millipore KgaA, Darmstadt, Germany, for PpEXP1_CS). The filtrate was concentrated using a 5 kDa cut-off ultrafilter (Pellicon XL Cassette, Merck Millipore KgaA). Ultrafiltration was performed on ice, and the crude protein solution was pre-cooled.

Evaluation of expressed protein

The integrity of the constructed vectors was confirmed by gene sequencing. The total amount of expressed protein was determined by Bradford's method. All secreted proteins were visualized by SDS-PAGE on 15% acrylamide gel. Two micrograms of protein was run in each lane. The bands of each recombinant expansin were estimated from the molecular weight, and the percentage of it to total protein was determined by ImageJ (Schneider et al. 2012) based on band intensity.

Results and discussion

Vector construction and expression check

The workflow of this study is illustrated in Figure 2. The two expression vectors were constructed as shown in Figure 3 and their sequences were confirmed. The vectors were linearized with restriction enzymes and transformed into *P. pastoris*. Stable transformants were selected based on zeocin resistance and their expression was confirmed. All colonies expressed the target protein,

Figure 2. Schematic illustration of workflow in this study. In this study, the expansin gene was inserted into two kinds of *P. pastoris* expression vectors and transformed into host cells. Transformants were fermented in a 5 l jar fermenter. After fermentation, cultures were centrifuged twice. The expressed protein amount in the supernatant was assayed and visualized by means of SDS-PAGE.

Figure 3. Vector map of constructed pGAPZαA_PpEXP1_WT (A) and pPICZαA_PpEXP1_CS (B). A: PpEXP1_WT gene was fused to a c-myc epitope and 6×histidine tag coding sequence in the C-terminal and inserted under the control of GAP promoter. Recombinant PpEXP1_WT was designed to be constantly expressed. B: PpEXP1_CS gene was inserted under the control of AOX1 promoter so that the recombinant protein was expressed in response to the addition of methanol. PpEXP1_CS did not contain either c-myc epitope or histidine tag. Both vectors have a zeocin restriction site, and transformants were selected using antibiotic resistance.

and the one with the highest expression level was used for cultivation in the 5 l jar fermenter.

Comparison of production of the two proteins in the jar fermenter

In the culture of PpEXP1_WT, the 50% glycerol feed batch (GF) phase and the 100% extra glycerol feed batch (EGF) phase were initiated at 38 and 45 h after the start of the culture, respectively. The protein concentration increased significantly from the start of the culture until approximately 100 h later. The supernatant was collected after approximately 160 h of incubation (Figure 4A).

In the culture of PpEXP1_CS, GF was initiated at 38 h after the start of the culture, as in the case of WT. The methanol feed (MF) was started at approximately 45 h. Between 50 and 62 h after the start of the cultivation, a sharp drop and then increase of DO value and an increase in pH value occurred. These changes suggest that some of the cells could not tolerate the added methanol and may have lysed. After about 120 h of incubation, the protein concentration stopped increasing, and the DO and pH values increased rapidly. The culture supernatant was collected about 133 h after the start of

the culture (Figure 4B).

The concentration of secreted total protein in each construct was assayed by the Bradford method (plotted with black dots in Figure 4), and the result of SDS-PAGE is shown in Figure 5. The total protein amounts were 916 ± 148 mg in the 2.85 l culture supernatant of *PpEXP1_WT* and 438 ± 43 mg in the 2.40 l culture supernatant of *PpEXP1_CS*. Thus, the total amount of secreted protein in *PpEXP1_WT* culture was approximately twice that of *PpEXP1_CS*. However, the production levels of *PpEXP1_WT* and *PpEXP1_CS* were 0.58 and 4.3 mg l^{-1} , respectively, and the percentages of recombinant protein with respect to total expressed protein were 0.18% and 2.37% after concentration with the 5 kDa ultrafiltration membrane. Thus, the ratio of target protein *PpEXP1_CS* to total protein was about 13 times higher than in the case of *PpEXP1_WT*. Absolutely, this result should be paid attention because of the strains of *P. pastoris* and expression vectors of *PpEXP1_WT* and *PpEXP1_CS* were different. Therefore, the

increased expression level in the current mutant could not be said simply an effect of the introducing mutation of cysteine to serine. However, our method reported in this report can become one of choices to heterologously express of hardly expressed protein such as expansin.

Other studies of protein expression in *P. pastoris*

There have been many studies on heterologous protein production using *P. pastoris*. In general, fungal proteins are expressed at a high level in *P. pastoris* without any mutations. In our laboratory, production of endoglucanases (*PcCel45A*) and cellobiohydrolase (*PcCel6A*) from the basidiomycete *Phanerochaete chrysosporium*, amounted to 1.6 and 4.6 g l^{-1} , respectively, using *P. pastoris* strain KM71H in a jar fermenter (Igarashi *et al.* 2012; Nakamura *et al.* 2013). However, expression of plant enzymes in *P. pastoris* has resulted in large yield differences, depending on the type of enzyme. For example, α -amylase from rice (*Oryza sativa*) and α -expansin from tomato were relatively

Figure 4. Production of recombinant *PpEXP1_WT*(A) and *PpEXP1_CS* (B) in a 51 jar fermenter. GF, EGF, and MF show the timings of the start of 50% glycerol feed, extra glycerol feed, and methanol feed, respectively. Each color and symbols show parameter values. Red open circle is DO (ppm), blue cross is pH, green open triangle is temperature ($^{\circ}\text{C}$), yellow green cross is rotation speed (per minutes), and black circle is total protein concentration ($\mu\text{g ml}^{-1}$).

Figure 5. SDS-PAGE of *PpEXP1_WT* (A) and *PpEXP1_CS*(B). M indicates protein molecular weight markers (pre-stained dual color protein markers, Bio-Rad). Lane 1 is the supernatant of fermentation, lane 2 is the fraction concentrated through a 100 or 300 kDa ultrafiltration membrane, lanes 3 and 4 are flow-through fractions of the 100 or 300 kDa ultrafiltration membrane, lane 5 is concentrated fraction of 5 kDa ultra filtration membrane, and lanes 6 and 7 (only B) are flow-through fractions of the 5 kDa ultrafiltration membrane. Two micrograms of protein was applied each lane. The arrows show the band of *PpEXP1_WT* or *PpEXP1_CS*, respectively.

highly expressed at 340 and 70.9 mg l⁻¹ (Lee et al. 2003; Liu et al. 2014), respectively, while rhamnogalacturonan endolyase from Chilean strawberry (*Fragaria chiloensis*) and β -galactosidase from radish (*Raphanus sativus*) were expressed at levels of only 1.6 and 7 mg l⁻¹, respectively (Kotake et al. 2005; Méndez-Yañez et al. 2020). One of the reasons for the relatively high expression of *LeEXP2* (Liu et al. 2014) is probably because that six of the eight cysteine residues correctly formed a disulfide bond and the other two free cysteine residues did not cause abnormal bonds. Kaewthai et al. (2010) cloned nine xyloglucan endo-transglycosylase/hydrolase (XTH) genes from barley (*Hordeum vulgare*) and expressed them in *P. pastoris* strain SMD1168H. Only five isoforms were secreted as soluble proteins, and only four of them showed enzymatic activity. They did not find any correlation between successful protein production and gene/protein sequence characteristics (Kaewthai et al. 2010).

Effect of removing cysteine residues

In the expression of the taste-modifying protein miraculin from miracle fruit (*Synsepalum dulcificum*) in *S. cerevisiae*, the expression level of the mutant containing serine instead of cysteine residues was increased compared to that of the wild-type protein (Ito et al. 2011). Our results are consistent with this, and therefore it appears that the presence of cysteine residues adversely affects the amount of *PpEXP1* secreted by *P. pastoris*. These findings support the idea that replacement of cysteine residues in plant proteins is a promising strategy to increase heterologous expression of eukaryotic proteins. In addition, if the mutation of cysteine to serine results in abnormal protein folding, the protein is expected to be degraded by the quality control system of *P. pastoris*, which suggests that the protein secreted into the culture supernatant is not abnormal (Inan et al. 2006; Ito et al. 2011; Saito et al. 2002).

XTHs have approximately the same proportion of cysteine residues as expansin. It is noteworthy that *PtEG16*, which is homologous to XTH from *Populus trichocarpa*, expressed in *E. coli* had no effect on the molecular weight of barley mixed linkage glucan, lichenan, from Icelandic moss, or xyloglucan from tamarind seeds, but these polysaccharides were efficiently degraded by *PtEG16* in the presence of a reducing agent, 0.25 mM dithiothreitol (DTT, Eklöf et al. 2013). Furthermore, since the “acid growth” activity of expansins extracted from plants has been reported to be promoted in the presence of a reducing agent (Li et al. 1993; McQueen-Mason et al. 1992), *PpEXP1_CS* is expected to be as good or better than *PpEXP1_WT* as a function of recombinant expansins. These results imply that the function of a mutated protein would be retained, although further study would be needed to confirm this.

In conclusion, the widely used *P. pastoris* heterologous expression system might be applicable even to challenging proteins containing large amounts of cysteine residues if the cysteine residues are removed, e.g., by mutation to serine. If the mutated recombinant proteins retain the functional activity of the wild-type proteins as expected, this system could be useful to produce large amounts of plant proteins for studies of the developmental mechanisms of plants and also for the utilization of plants as a biomass resource.

Acknowledgements

We thank for Dr. Hiroko Hayama and Prof. Megumi Ishimaru for kindly providing *PpEXP1* cDNA. This study was supported by Grant-in-Aid for Scientific Research on Innovative Areas No. 18H05484 to KI.

Author contributions

All experiments were done by KM. NS advised and supported the fermentation experiments and data analysis. KI supervised the experiments and wrote the manuscript.

References

- Ahmad M, Hirz M, Pichler H, Schwab H (2014) Protein expression in *Pichia pastoris*: Recent achievements and perspectives for heterologous protein production. *Appl Microbiol Biotechnol* 98: 5301–5317
- Benkert P, Biasini M, Schwede T (2011) Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* 27: 343–350
- Bertoni M, Kiefer F, Biasini M, Bordoli L, Schwede T (2017) Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci Rep* 7: 1–15
- Bienert S, Waterhouse A, De Beer TAP, Tauriello G, Studer G, Bordoli L, Schwede T (2017) The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Res* 45(D1): D313–D319
- Burton RA, Gidley MJ, Fincher GB (2010) Heterogeneity in the chemistry, structure and function of plant cell walls. *Nat Chem Biol* 6: 724–732
- Cereghino JL, Cregg JM (2000) Heterologous protein expression in the methylotrophic yeast *Pichia pastoris*. *FEMS Microbiol Rev* 24: 45–66
- Cosgrove DJ (2000) Loosening of plant cell walls by expansins. *Nature* 407: 321–326
- Cosgrove DJ (2015) Plant expansins: Diversity and interactions with plant cell walls. *Curr Opin Plant Biol* 25: 162–172
- Cosgrove DJ (2016) Catalysts of plant cell wall loosening. *F1000 Res* 5: 119
- Cosgrove DJ, Jarvis MC (2012) Comparative structure and biomechanics of plant primary and secondary cell walls. *Front Plant Sci* 3: 1–6
- Eklöf JM, Shojania S, Okon M, McIntosh LP, Brumer H (2013) Structure-function analysis of a broad specificity *Populus trichocarpa* endo- β -glucanase reveals an evolutionary link between bacterial licheninases and plant XTH gene products. *J Biol Chem* 288: 15786–15799

- Guex N, Peitsch MC, Schwede T (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis* 30(Suppl 1): 162–173
- Hayama H, Ito A, Moriguchi T, Kashimura Y (2003) Identification of a new expansin gene closely associated with peach fruit softening. *Postharvest Biol Technol* 29: 1–10
- Hayama H, Shimada T, Fujii H, Ito A, Kashimura Y (2006) Ethylene-regulation of fruit softening and softening-related genes in peach. *J Exp Bot* 57: 4071–4077
- Hayama H, Shimada T, Haji T, Ito A, Kashimura Y, Yoshioka H (2000) Molecular cloning of a ripening-related expansin cDNA in peach: Evidence for no relationship between expansin accumulation and change in fruit firmness during storage. *J Plant Physiol* 157: 567–573
- Higgins DR, Cregg JM (2007) *Pichia Protocols Second Edition: Methods in Molecular Biology*. Human Press, Totowa, New Jersey
- Igarashi K, Maruyama M, Nakamura A, Ishida T, Wada M, Samejima M (2012) Degradation of crystalline celluloses by *Phanerochaete chrysosporium* Cellobiohydrolase II (Cel6A) heterologously expressed in methylotrophic yeast *Pichia pastoris*. *J Appl Glycosci* 59: 105–110
- Inan M, Aryasomayajula D, Sinha J, Meagher MM (2006) Enhancement of protein secretion in *Pichia pastoris* by overexpression of protein disulfide isomerase. *Biotechnol Bioeng* 93: 771–778
- Ito K, Sugawara T, Koizumi A, Nakajima KI, Shimizu-Ibuka A, Shiroishi M, Asada H, Yurugi-Kobayashi T, Shimamura T, Asakura T, et al. (2011) Cysteine-to-serine shuffling using a *Saccharomyces cerevisiae* expression system improves protein secretion: Case of a nonglycosylated mutant of miraculin, a taste-modifying protein. *Biotechnol Lett* 33: 103–107
- Kaewthai N, Harvey AJ, Hrmova M, Brumer H, Ezcurra I, Teeri TT, Fincher GB (2010) Heterologous expression of diverse barley XTH genes in the yeast *Pichia pastoris*. *Plant Biotechnol* 27: 251–258
- Kerff F, Amoroso A, Herman R, Sauvage E, Petrella S, Filée P, Charlier P, Joris B, Tabuchi A, Nikolaidis N, et al. (2008) Crystal structure and activity of *Bacillus subtilis* YoaJ (EXLX1), a bacterial expansin that promotes root colonization. *Proc Natl Acad Sci USA* 105: 16876–16881
- Kotake T, Dina S, Konishi T, Kaneko S, Igarashi K, Samejima M, Watanabe Y, Kimura K, Tsumuraya Y (2005) Molecular cloning of a β -Galactosidase from radish that specifically hydrolyzes β -(1 $\rightarrow$ 3)- and β -(1 $\rightarrow$ 6)-galactosyl residues of arabinogalactan protein. *Plant Physiol* 138: 1563–1576
- Lee CY, Nakano A, Shiomi N, Lee EK, Katoh S (2003) Effects of substrate feed rates on heterologous protein expression by *Pichia pastoris* in DO-stat fed-batch fermentation. *Enzyme Microb Technol* 33: 358–365
- Li ZC, Durachko DM, Cosgrove DJ (1993) An oat coleoptile wall protein that induces wall extension in vitro and that is antigenically related to a similar protein from cucumber hypocotyls. *Planta* 191: 349–356
- Liu X, Liu C, Ma Y, Hong J, Zhang M (2014) Heterologous expression and functional characterization of a novel cellulose-disruptive protein LeEXP2 from *Lycopersicon esculentum*. *J Biotechnol* 186: 148–155
- Liu Y, Zhao TJ, Yan Y, Zhou HM (2005) Increase of soluble expression in *Escherichia coli* cytoplasm by a protein disulfide isomerase gene fusion system. *Protein Expr Purif* 44: 155–161
- Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res* 42(D1): 490–495
- Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, et al. (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res* 47(W1): W636–W641
- Marowa P, Ding A, Kong Y (2016) Expansins: Roles in plant growth and potential applications in crop improvement. *Plant Cell Rep* 35: 949–965
- McQueen-Mason S, Durachko DM, Cosgrove DJ (1992) Two endogenous proteins that induce cell wall extension in plants. *Plant Cell* 4: 1425–1433
- Méndez-Yañez A, González M, Carrasco-Orellana C, Herrera R, Moya-León MA (2020) Isolation of a rhamnogalacturonan lyase expressed during ripening of the Chilean strawberry fruit and its biochemical characterization. *Plant Physiol Biochem* 146: 411–419
- Nakamura A, Ishida T, Fushinobu S, Kusaka K, Tanaka I, Inaka K, Higuchi Y, Masaki M, Ohta K, Kaneko S, et al. (2013) Phase-diagram-guided method for growth of a large crystal of glycoside hydrolase family 45 inverting cellulase suitable for neutron structural analysis. *J Synchrotron Radiat* 20: 859–863
- Nielsen LR, Lund MN, Davies MJ, Nielsen JH, Nielsen SB (2018) Effect of free cysteine on the denaturation and aggregation of holo α -lactalbumin. *Int Dairy J* 79: 52–61
- Nikolaidis N, Doran N, Cosgrove DJ (2014) Plant expansins in bacteria and fungi: Evolution by horizontal gene transfer and independent domain fusion. *Mol Biol Evol* 31: 376–386
- Saito A, Usui M, Song Y, Azakami H, Kato A (2002) Secretion of glycosylated α -lactalbumin in yeast *Pichia pastoris*. *J Biochem* 132: 77–82
- Sampedro J, Cosgrove DJ (2005) The expansin superfamily. *Genome Biol* 6: 1–11
- Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9: 671–675
- Sharova EI (2007) Expansins: Proteins involved in cell wall softening during plant growth and morphogenesis. *Russ J Plant Physiol* 54: 713–727
- Shcherban TY, Shi J, Durachko DM, Guiltinan MJ, McQueen-Mason SJ, Shieh M, Cosgrove DJ (1995) Molecular cloning and sequence analysis of expansins: A highly conserved, multigene family of proteins that mediate cell wall extension in plants. *Proc Natl Acad Sci USA* 92: 9245–9249
- Somerville C, Bauer S, Brininstool G, Facette M, Hamann T, Milne J, Osborne E, Paredes A, Persson S, Raab T, et al. (2004) Toward a systems approach to understanding plant cell walls. *Science* 306: 2206–2211
- Waterhouse A, Bertoni M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, De Beer TAP, Rempfer C, Bordoli L, et al. (2018) SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Res* 46(W1): W296–W303
- Yennawar NH, Li L-C, Dudzinski DM, Tabuchi A, Cosgrove DJ (2006) Crystal structure and activities of EXPB1 (Zea m 1), a beta-expansin and group-1 pollen allergen from maize. *Proc Natl Acad Sci USA* 103: 14664–14671
- Yuan S, Duan H, Liu C, Liu X, Liu T, Tao H, Zhang Z (2004) The role of thioredoxin and disulfide isomerase in the expression of the snake venom thrombin-like enzyme calobin in *Escherichia coli* BL21 (DE3). *Protein Expr Purif* 38: 51–60