CHARACTERIZATION OF β-GLUCOSIDASE FROM THE RICE *SFR2* GENE

Massalin Nakphaichit

A Thesis Submitted in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Biotechnology

Suranaree University of Technology

Academic Year 2006

การศึกษาคุณสมบัติของเอนไซม์บีต้ากลูโคซิเดสจากยีน *SFR2* ในข้าว

นางสาวมัสลิน นาคไพจิตร

วิทยานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิทยาศาสตรมหาบัณฑิต
สาขาวิชาเทคโนโลยีชีวภาพ
มหาวิทยาลัยเทคโนโลยีสุรนารี
ปีการศึกษา 2549

CHARACTERIZATION OF β-GLUCOSIDASE FROM THE RICE *SFR2* GENE

Suranaree University of Technology has approved this thesis submitted in partial fulfillment of the requirements for a Master's Degree.

Thesis Examining Committee

_____

(Asst. Prof. Dr. Chokchai Wanapu)

Chairperson

_____

(Asst. Prof. Dr. Mariena Ketudat-Cairns)

Member (Thesis Advisor)

_____

(Assoc. Prof. Dr. James Ketudat-Cairns)

Member

_____

(Dr. Rodjana Opassiri)

Member

_____ _____

(Assoc. Prof. Dr. Saowanee Rattanaphani) (Asst. Prof. Dr. Suwayd Ningsanond)

Vice Rector for Academic Affairs Dean of Institute of Agricultural Technology

MASSALIN NAKPHAICHIT : CHARACTERIZATION OF β-GLUCOSIDASE FROM THE RICE *SFR2* GENE. THESIS ADVISOR : ASST. PROF. MARIENA KETUDAT-CAIRNS, Ph.D. 100 PP.

RICE/FREEZING SENSITIVE/PROTEIN LOCALIZATION

The rice *sensitive to freezing2* or rice *SFR2* gene was cloned and expressed. The SFR2 protein encoding an enzyme in the group of glycosyl hydrolase family 1. It's amino acid sequence was more similar to β-glycosidases from thermophilic and halophilic archea and bacteria than any other non-SFR2 enzymes in plants. Several strains of *Escherichia coli*, vectors (pET32a, pSY5 and pGEX-4T-3), induction times and induction temperatures (4, 20, 37 and 45°C) were investigated. No condition tried was able to produce SFR2 protein in *E. coli*. *Pichia pastoris* was also used to express *SFR2*. The mRNA of *SFR2* was detected in methanol induction but no protein was detected by activity nor SDS-PAGE technique. Therefore, plant cells were used to investigate the expression and protein localization. The detection of GFP and GUS protein indicated that the *SFR2* gene could express in plant cells under control of the 35S promoter. The constructs containing *SFR2* transit peptide at the N-terminal showed GFP fluorescence or GUS staining in the chloroplast.

School of Biotechnology Student's Signature_____

Academic Year 2006 Advisor's Signature_____

Co-advisor's Signature_____

มัสลิน นาคไพจิตร : การศึกษาคุณสมบัติของเอนไซม์บีต้ากลูโคซิเดสจากยีน *SFR2* ในข้าว (CHARACTERIZATION OF β-GLUCOSIDASE FROM THE RICE *SFR2* GENE) อาจารย์ที่ปรึกษา : ผู้ช่วยศาสตราจารย์ ดร.มารินา เกตุทัต-คาร์นส์, 100 หน้า.

งานวิจัยนี้เป็นการศึกษาคุณลักษณะของเอนไซม์บีต้ากลูโคซิเดสจากยีน *SFR2* ในพืชตระกูลข้าว โดยทำการศึกษาผ่านระบบรีคอมบิแนนท์โปรตีน ผลการเปรียบเทียบลำดับเบสของยีน *SFR2* กับโปรตีนในฐานข้อมูล พบว่าโปรตีน SFR2 อยู่ในกลุ่มของเอนไซม์ไกลโคซิลไฮโดรเลสที่ 1 (GH1) และลำดับเบสคล้ายคลึงกับ (GH1) จากแบคทีเรีย Thermophilic และ Halophilic มากกว่ากลุ่มของเอนไซม์บีต้ากลูโคซิเดสจากพืชต่างๆ จากผลการศึกษาพบว่ารีคอมบิแนนท์โปรตีน SFR2 ไม่สามารถผลิตได้ใน *Escherichia coli* ทั้งในสายพันธุ์และระบบเวคเตอร์ต่างๆ (pET32a, pSY5 และ pGEX-4T-3) ซึ่งการศึกษานี้ยังได้มีการทดลองหาสภาวะอุณหภูมิที่เหมาะสม (4, 20, 37 และ 45°C) ต่อการผลิตและระยะเวลาที่เหมาะสมต่อการชักนำให้มีการผลิตโปรตีนด้วย แต่ทุกสภาวะที่ศึกษาไม่สามารถผลิตโปรตีน SFR2 ใน *E. coli* ได้ จึงได้นำรีคอมบิแนนท์โปรตีน SFR2 มาผลิตในยีสต์ *Pichia pastoris* ซึ่งจากการทดลองพบ mRNA ของยีน *SFR2* แต่ไม่สามารถตรวจสอบกิจกรรมของเอนไซม์และโปรตีนได้ ดังนั้นรีคอมบิแนนท์โปรตีน SFR2 จึงถูกนำมาผลิตในเซลล์พืช โดยโปรตีน SFR2 ถูกเชื่อมกับโปรตีนติดตาม GFP และ GUS ทางด้านปลาย N หรือ C เพื่อถ่ายเข้าเซลล์สาหร่ายหางกระรอกและเซลล์หัวหอม จากผลการทดลองพบว่าสามารถผลิตรีคอมบิแนนท์โปรตีน SFR2 ภายใต้การควบคุมการแสดงออกของโปรโมเตอร์ 35S ในเซลล์พืช และพบโปรตีน SFR2 ในคลอโรพลาสต์ของเซลล์สาหร่ายหางกระรอก

สาขาวิชาเทคโนโลยีชีวภาพ ลายมือชื่อนักศึกษา_____

ปีการศึกษา 2549 ลายมือชื่ออาจารย์ที่ปรึกษา_____

ลายมือชื่ออาจารย์ที่ปรึกษาร่วม_____

ACKNOWLEDGEMENTS

I am deeply grateful to my advisor, Asst. Prof. Dr. Mariena Ketudat-Cairns for her kind supervision, valuable advice, and encouragement. She always spent a lot of her time to help me solve my problems. I would like to thank Assoc. Prof. James R. Ketudat-Cairns, my coadvisor, for idea to start this project and greatly providing a lot of vectors and *E. coli* strains.

I am very grateful to Dr. Rodjana Opassiri for guidance and supporting data. I would like to thanks the Department of Biotechnology, Thammasat University for providing the use of the gene gun machine. I also thank Dr. Theppanya Charoenrat and Ajarn Saifon Sherbonmee for advice with the gene gun. Thanks to all members of Biotechnology and Biochemistry, SUT for their help. Special thanks Miss Ormruthai Kampong, who is the best sister, best friend and best assistant.

This research was supported by Thailand Research Fund (TRF).

Finally, I would like to express my deepest gratitude to my family for their love, understanding and inspiration.

Massalin Nakphaichit

CONTENTS

Page

CONTENTS (Continued)

CONTENTS (Continued)

Page

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF FIGURES (Continued)

Figure **Page**

LIST OF FIGURES (Continued)

Figure **Page**

LIST OF ABBREVIATIONS

bp	Base pair
°C	Degree Celsius
cDNA	Complementary deoxynucleic acid
DNA	Deoxyribonucleic
dNTP	Deoxynucleotide triphosphate
kDa	Kilodalton
kb	Kilobase
ml	Milliliter
mM	Millimolar
mRNA	Messenger ribonucleic acid
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase chain reaction
p-NP	*p*-Nitrophenol
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Revolution per minute
SDS	Sodium dodecyl sulfate
SFR	Sensitive to freezing
µl	Microliter
µM	Micromolar

CHAPTER I

INTRODUCTION

Rice is one of the world's most important food crops and Thailand is a major producer of rice. Rice is also a model system for the study of monocots. The rice genome databases are available with up to 99% coverage of the rice genes. Therefore, it is an ideal system to study factors affecting growth and response to stress of grain crops, since genes of interest can be identified and cloned directly from their sequences.

The *sensitive to freezing2* or *SFR2* gene is a member of enzyme glycosyl hydrolase family 1. It plays a role in freezing resistance in *Arabidopsis thaliana* (Thorlby, 2004). The amino acid sequence of *SFR2* is more similar to glycosyl hydrolase family 1 enzymes from thermophilic archea and bacteria than any other glycosyl hydrolase family 1 from plants. A search of the National Center for Biotechnology Information Web site shows that other plants species such as rice, maize, soybean, cherry tomato and loblolly pine also contain *SFR2* genes. The rice *SFR2* cDNA was found in accession number AK119461. It is closely related with maize with 82% identity, whereas in *Arabidopsis* only 59% identity was shown. The rice *SFR2* is one of the enzymes in 40 glycosyl hydrolase family 1 gene from rice and was designated as *Os11bglu36*, since it is mapped to chromosome 11 (Opassiri, 2006). The *SFR2* mRNA has been found in shoot and leaf in stress conditions.

In *Arabidopsis*, mutation of *SFR2* gene confers the sensitive to freeze phenotype. However, the SFR2 protein's function is yet to be determined. Therefore, molecular techniques were used to clone the rice (*Oryza sativa*) *SFR2* gene to try to elucidate the SFR2 function. Two clones of *SFR2* were produced, *st-SFR*2 and *co-SFR2*. The *st-SFR2* contained a transit peptide sequence and conserved region of β-glucosidase, whereas *co-SFR2* started with the conserved region of β-glucosidase. Production of the recombinant SFR2 proteins was tried in prokaryote and eukaryote cells in order to characterize the mechanism of this enzyme. Reporter genes were also used to tag the *SFR2* to define the localization of SFR2 in plant cells. The localization was used to predict the function of *SFR2* in stress adaptation. All of these studies are basic knowledge to try to understand at least one rice gene function to develop more productive rice for the benefit of Thailand in the future.

CHAPTER II

LITERATURE REVIEWS

2.1 Rice

Rice is one of the most important food crops in the world. It feeds more people than any other crops. Rice belongs to the genus *Oryza*. There are more than 20 species of rice, of which only two species, *Oryza sativa* (which originated in the humid tropics of South and Southeast Asia) and *O. glaberrima* (which originated in the Niger Basin in Africa), are cultivated (Duffy, 2001).

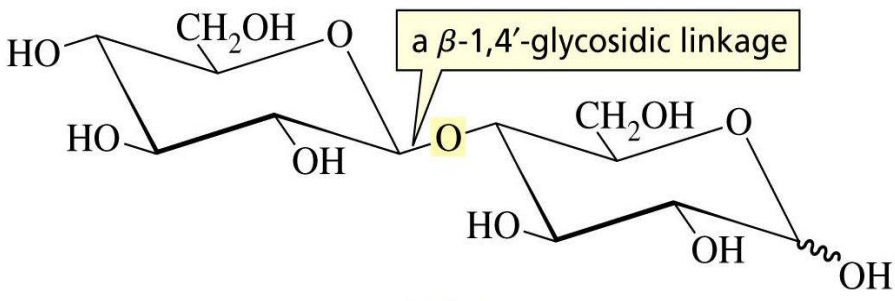
O. sativa has a small diploid genome (2n=24) of about 450 Mb. Based on its economic, social value and the small genome, the international community decided to use it as the monocot plant model for genomic sequencing (the sequence is available at http://www.ncbi.nih.gov). Based on comparative mapping with other Gramineae, the rice genome has a large number of regions exhibiting conserved synteny with other members of the family, such as wheat, maize, and barley. Thus, study of rice genes will aid research on other crops as well.

There are currently three publicly funded rice genome sequencing projects. First, the International Rice Genome Sequencing Project (IRGSP), which is a collaboration of publicly funded projects to sequence the genome of *O. sativa* ssp. *japonica*. The other two projects are funded by the Chinese Academy of Sciences,

japonica by a whole genome shotgun method. The size of the genome has been report as 466 Mb by the Chinese group for *O. sativa* ssp. *indica* and 430 Mb for *O. sativa* ssp. *japonica* by the IRGSP.

There are many studies about rice genes. For example, based on its identification from genomic data, *OsRAA1* (a novel conserved protein in plants) was characterized and found to play an important role in rice root development (Han, 2004). Gu *et al.* (2005) investigated *OsDPKs* (*Oryza sativa* dual-specificity protein kinase) genes and showed that these genes may belong to a novel rice gene family encoding dual-specificity protein kinases involved in the plant responses to abiotic and biotic stresses. These research projects provide are important basic knowledge for develop productive rice or crops in the near future.

2.2 Overview of β-glucosidase



β-glucosidase (β-D-glucoside glucohydrolase, EC.3.2.1.21) constitutes a major group among glucoside hydrolase. They catalyze the hydrolysis of β-O-glucosidic linkages formed between D-glucose and an aglycone or another sugar, releasing glucose and an aglycone. These enzymes occur ubiquitously in all three domains of living organisms, archea, eubacteria and eukarya (Amouri, 2006).

In plants, β-glucosidases have been implicated in a variety of processes, for example, β-glucosidases *BGLU45* and *BGLU46* gene products in *Arabidopsis thaliana* play a role in lignification by hydrolyzing monolignol glucosides (Escamilla-Trevino, 2006). Akiyama *et al.* (1997) suggested that one possible role for the rice β-glucosidase enzyme might be hydrolysis of oligosaccharides released from cell wall β-glucans during rice germination. Moreover, jasmonate-regulated wheat protein that cross-reacts with maize β-glucosidase has the functional role of the plant defense response. Higher plants adapt to changes in abiotic or biotic factors in the environment by altering morphological, physiological and developmental processes (Wang, 2005). Another important function of plant β-glucosidases is activation of plant hormones, *Arabidopsis* β-glucosidases (ATBG1) hydrolyses inactive ABA glucose ester to active ABA under dehydration stress (Lee, 2006). The ABA conjugates such as ABA glucose ester accumulates in plant tissues with age and during stress treatments. The conjugates are mainly compartmenralized into vacuole. In stress conditions they can be transported long distances to other cells. The inactive form of ABA can be cleaved by the extracellular β-glucosidases to release the free active ABA (Dietz, 2000).

2.3 *Sensitive to freezing2* gene (*SFR2* Gene)

Cold-induced injury occurs over a wide range of temperatures, depending on both genotype and environment. This has led to classification of plant species into three groups. Chilling-sensitive plants suffer injury at positive temperatures, usually below 10-15 °C freezing-sensitive plants are damaged by exposure to any temperature

below 0 °C and freezing-resistant plants are able to survive subzero temperature down to a limit which is the characteristic of each genotype.

The *Sensitive to freezing2* or *SFR2* gene was previously studied in *Arabidopsis thaliana* by Thorlby *et al.* (2004). They characterized the *SFR2* gene by comparing mutant and wild type plants and found that these regions mapped on chromosome III between the 28L1SP1 and GT10 markers. The amino acid sequence encoded by the mutant gene (*sfr2-1*) is similar to the amino acid sequence of the SFR2 protein, except for one amino acid change, a glycine into an asparagine at codon 234. The mutant plants show freezing sensitivity when grow under freeze testing conditions about -6 °C.

The localization of arabidopsis *SFR2* expression was determined using β-glucuronidase (GUS) as a reporter. GUS activity was found in hypocotyls, cotyledons, stems, leaves, pedicels, sepals, anthers and pistils, but expression in roots was very limited (Thorlby, 2004).

When the public database was searched with the SFR2 protein sequence, they found that SFR2 is a member of glycosyl hydrolase family 1. The amino acid sequence of SFR2 show higher similarity with β-glycosidases from thermophilic and halophilic archea and bacteria than that of every other family 1 β-glycosidase enzyme in *Arabidopsis* (Thorlby, 2004). The amino acid sequence of SFR2 from rice (*Os11bglu36*) also show similar result that it is more divergent from every other glycosyl hydrolase family 1 enzyme from rice and other plants except for SFR2 orthologes (Fig. 2.1 copied from Opassiri, 2006).

The property and function of *SFR2* gene that support freezing tolerance is still undiscovered. It has only been suggested that perhaps *SFR2* is involved in the turnover of polysaccharides in the cell wall, or perhaps it is involved in protecting chloroplasts from freezing damage (Martin, 2004).

Figure 2.1 Relationship between rice and other plant GH1 protein sequences described by a phylogenetic tree rooted by *Os11bglu36* (copied from Opassiri, 2006). The sequences were aligned with ClustalX, then manually adjusted, followed by removal of N-terminal, C-terminal and large gap regions to build the data model. The tree was produced by the neighbor joining method and analyzed with 1,000 bootstrap

replicates. The internal branches supported by a maximum parsimony tree made from the same sequences are shown as bold lines. The sequences other than rice include: ME AAB71381, *Manihot esculenta* linamarase; RSMyr BAB17226, *Raphanus sativus* myrosinase; BJMyr AAG54074, *Brassica juncea myrosinase*; BN CAA57913, *Brassica napus* zeatin-O-glucoside-degrading β-glucosidase; HB AAO49267, *Hevea brasiliensis* rubber tree β-glucosidase; CS BAA11831, *Costus speciosus* furostanol glycoside 26-O-β-gluco sidase (F26G); PS AAL39079 *Prunus serotina* prunasin hydrolase isoform PHB precursor; PA AAA91166, *Prunus avium* ripening fruit β-glucosidase; TR CAA40057, *Trifolium repens* white clover linamarase; CA CAC08209, *Cicer arietinum* epicotyl β-glucosidase with expression modified by osmotic stress; DC AAF04007, *Dalbergia cochinchinensis* dalcochinin 8'-O-β-glucoside β-glucosidase; PT BAA78708, *Polygonum tinctorium* β-glucosidase; DL CAB38854, *Digitalis lanata* cardenolide 16-O-glucohydrolase; OE AAL93619, *Olea europaea* subsp. europaea β-glucosidase; CR AAF28800, *Catharanthus roseus* strictosidine β-gluco sidase; RS AAF03675, *Rauvolfia serpentina* raucaffricine-O-β-D-glucosidase; CP AAG25897, *Cucurbita pepo* silverleaf whitefly-induced protein 3; AS CAA55196, *Avena sativa* β-glucosidase; SC AAG00614, *Secale cereale* β-glucosidase; ZM AAB03266, *Zea mays* cytokinin β-glucosidase; ZM AAD09850, *Zea mays* β-gluco sidase; SB AAC49177, *Sorghum bicolor* dhurrinase; LE AAL37714, *Lycopersicon esculentum* β-mannosidase; HV AAA87339, barley BGQ60 β-glucosidase; HB AAP51059, *Hevea brasiliensis* latex cyanogenic β-glucosidase; PC AAC69619 *Pinus contorta* coniferin β-glucosidase; GM AAL92115, *Glycine max* hydroxyisourate hydrolase; CS BAC78656, *Camellia sinensis* β-primeverosidase.

2.4 Freezing stress

Cold temperature can cause one type of water deficit stress. Freezing greatly affects cellular water relations. At a given temperature, the chemical potential of ice is less than that of liquid water. The ice formation is initiated in the intercellular spaces, cellular water moves down the water potential gradient, across the plasma membrane, and toward the extracellular ice (Buchanan, 2000). Some freezing-tolerant plants promote the formation of extracellular ice and thereby prevent the formation of damaging ice crystals within the cytoplasm.

Freezing tolerance develops in a process know as cold acclimation, a response to low but nonfreezing temperatures that occur before freezing. The details of the mechanism that permit freezing tolerance are not well understood. Several processes that can occur in the development of freezing are (a) stabilization of membranes (b) accumulation of sugars, other osmolytes, and antifreeze proteins and (c) multiple changes in gene expression (Buchanan, 2000).

Two type of alfalfa cultivars, chilling-resistant and chilling-sensitive, were used to investigate the relationship of chloroplast membrane fatty acid composition at a chilling temperature. Chilling-resistant cultivars showed a high degree unsaturated of the membrane fatty acids, while low concentrations were found in chilling-sensitive cultivars (Peoples, 1997). This phenomenon suggested that the degree of unsaturation of membrane fatty acids influences the sensitivity of the membrane to chilling temperature (Friedman, 1997 and Lyons, 1965). During periods of cold acclimation, some plants accumulate sugar and apoplastic PR proteins that can retard the growth of ice crystals. The antifreezing proteins such as endochitinase, β-endoglucanase and

osmotin-like protein have been shown to form oligomeric complexes. The large surfaces of the complexes interact with ice, inhibiting its growth and recrystallization more effectively than an individual polypeptide can (Buchanan, 2000). The genes responsible for low temperature resistance eg. wheat *Wcor15* has been characterized. The *Wcor15*-GFP expressed under control of the 35S promoter located to the stromal compartment of the chloroplasts. The 5' upstream region of *Wcor15* were used as promoter to drive the *Wcor15* expression. The results showed transgenic *Wor15* tobacco plants had improved freezing tolerance compared with wild type (Shimamura, 2005).

2.5 Gateway® system (Invitrogen)

The Gateway® system is one of the techniques used for cloning and gene expression in this research. The system facilitates the transfer of DNA segments from an entry vector to multiple expression vectors. Gateway Technology is based on the well-characterized lambda phage site-specific recombination system (*att*B x *att*P and *att*L x *att*R). Two reactions, BP and LR, constitute the Gateway Technology (Table 2.1, Fig. 2.2).

Table 2.1 Summary of Gateway reactions and nomenclature. (modified from Invitrogen)

Reaction	**Reacting site**	**Product**	**Structure of product**
BP reaction	*att*B × *att*P	Entry clone	*att*L1-gene-*att*L2
LR reaction	*att*L × *att*R	Expression clone	*att*B1-gene-*att*B2

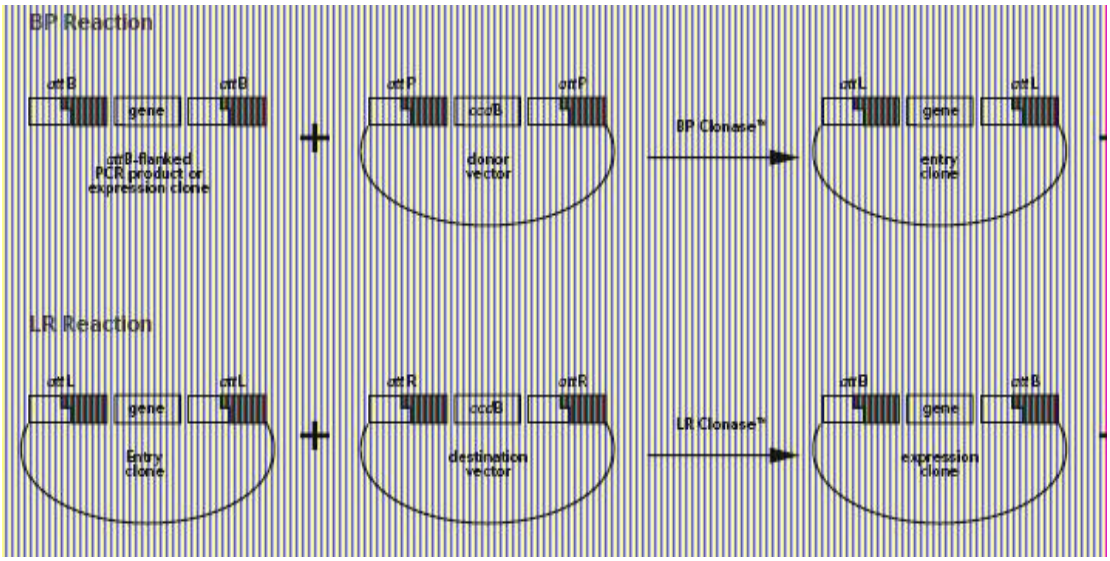


Figure 2.2 Gateway Technology Overview [copied from Invitrogen].

The BP reaction uses a recombination reaction between *att*B sites in a DNA segment or expression clone and *att*P site in a donor vector to create an entry clone. The LR reaction is a recombination between an *att*L entry clone and *att*R site in a destination vector. The LR reaction is used to move the sequence of interest from an entry clone to one or more destination vectors in parallel reactions.

Here is a sample of using the Gateway® system for gene expression and localization study. Jung *et al.* (2005) characterized ROP (Rho-related GTPase from plant) from rice. The full ORF of *OsRacB* was amplified from the rice *Rops* genes with PCR primers containing adapted *att*B sequences in two step PCR amplifications. Amplified PCR products were recombined with BP clonase into the pDONR vector (Invitrogen) yielding an entry clone and a *ccd*B fragment by-product. The entry clone containing *OsRacB* recombined with a destination vector with LR clonase (Invitrogen) to generate GFP:OsRacB chimeric construct according to the procedures described by the company (Invitrogen). The pK7WGF2 (VIB, Belgium) vector with *OsRacB* ORF was used for localization study in onion cells and *Arabidopsis*.

2.6 Protein expression

During the past two decades, technical advances in DNA manipulation in vitro, together with the development of efficient transformation systems for a wide variety of prokaryotic and eukaryotic cells have allowed extensive study of the recombinant protein expression. Such studies have contributed widely to the understanding of the regulation of gene expression via transcription, translation and post translation mechanisms. The increase in protein yield obtained in many recombinant expression systems has facilitated studies of protein structure and biological activity that were formerly limited by the ability to purify sufficient quantities of many biologicaly active proteins from their native systems (Shewry, 1992).

Escherichia coli has obvious advantages as a host for gene expression. It can be grown to very high cell density in simple growth media (Shewry, 1992). All *E. coli* strains used in this study and their descriptions are shown in table 2.2.

Table 2.2 All *E. coli* strains used in this study (adapted from Novagen pET system manual).

***E. coli* strain**	**Application description**	**Antibiotic Resistance**
BL21 (DE3)	General purpose expression host	None
Origami (DE3)	Enhances disulfide bond formation in the cytoplasm by mutations in redox genes.	Tetracycline, Kanamycin
Origami B (DE3)	Enhances disulfide bond formation in the cytoplasm by mutations in redox genes and *lacZY* mutant enhance the solubility and activity of difficult target proteins.	Tetracycline, Kanamycin
Rosetta (DE3)	Enhances the expression of eukaryotic proteins that contain codons rarely used in *E. coli*.	Chloramphenicol
Rosetta-gami (DE3)	Combines the properties of Origami and Rosetta	Kanamycin, Tetracycline, Streptomycin, and Chloramphenicol
BL21 pLySs (DE3)	Toxic protein expression	Chloramphenicol

The cDNA of two β-glucosidase isozymes (*bglu1* and *bglu2*) from rice were express in *E. coli* (Opassiri, 2003). They found the protein produced from *bglu1* in active form and can hydrolyze various substrates, but bglu2 protein expression resulted in mainly insoluble and inactive protein. A mannose-specific jacalin-related lectin from rice has been cloned (Jiang, 2005). The *OsJAC1* contain two subunits, dirigent (amino acids 26 – 139) and jacalin (amino acid 175 – 305). Jacalin domain at the C-terminal was ligated into the pGEX4T-1 (GE healthcare, Amershambiosciences) and expressed in *E. coli* BL21 (DE3). The expressed recombinant protein showed agglutination activity towards rabbit red blood cells, which was inhibited by mannose, a result that helped reveal the functional role of this protein. Efficient translation in *E. coli* requires the presence of a ribosome binding site. This consists of the initiation codon, AUG, together with a sequence (lying 3 – 12 bases upstream) know as the Shine-Dalgano (SD) sequence. The SD sequence (3 – 9 bases) is complementary to the 3' end of the 16S rRNA in the 30S ribosomal subunit. It is believed that hybridization of the 5' end of the mRNA to the 3' end of 16S rRNA help generate the translation initiation complex (Steitz, 1979).

However, attempts to express eukaryotic proteins in *E. coli* have met with a number of problems, the most important of which is the lack of post-translational modification of proteins in bacterial cells. Processes such as glycosylation and phosphorylation are necessary for the biological activity and production in *E. coli* may not result in a product with all the required properties. Yeast has several advantages over *E. coli* as an expression systems. In particular, yeast can secrete proteins into the medium, which aids protein purification and also allows disulphide bond formation

and glycosylation of proteins to occur (Shewry, 1992).

Many foreign proteins have been successfully mass produced in the yeast *Saccharomyces cerevisiae*. This particular species is popular largely because it is familiar to molecular biologists. There are large amounts of knowledge that have been accumulated about its genetics and physiology. While this species has been used for the production of some eukaryotic foreign proteins, it has several limitations. Generally, the product yields are low, instability of plasmid may occur and many of the proteins secreted by *S. cerevisiae* are not found free in the medium, but rather in the periplasmic space. This leads to problems with purification and further decreases product yield (Available: http://wwwchem.csustan.edu/chem4400/SJBR/amarjit.htm).

One of the alternative species that has been used is *Pichia pastoris*. There are several reasons that this particular species is appealing. The protocols for its growth did not have to be worked out because it can be grown under conditions that are similar to *S. cerevisiae*. *P. pastoris* has a strong, inducible promoter that can be used for protein production. It is capable of generating post-translational modifications of products better than *S. cerevisiae*. Isolation of foreign protein is facilitated by the fact that *P. pastoris* does not secrete a lot of its own proteins (Available: http://wwwchem. csustan.edu/chem4400/SJBR/amarjit.htm). Park *et al.* (2002) cloned acidic (OsChib1a) and basic (OsChib1b) class III chitinases from rice (*Oryza sativa* L.). Both were expressed from the corresponding cDNAs in *P. pastoris* and purified to homogenously. OsChib1b had lower specific chitinase activity than OsChib1a, but it showed a strong lytic activity and significant antifungal activity. This expression system also makes it possible to produce genetically engineered enzymes for the study

of the structure-function relationship of class III chitinases and other enzymes of interest.

The *P. pastoris* strain Y-11430 was used for SFR2 expression in yeast. It is a wild-type methylotrophic yeast, which can utilize methanol as its sole carbon source (Ilgen, 2005). No antibiotic resistance gene or metabolic mutation is present which allows for selection of expression vectors containing zeocin resistance or another selectable marker gene upon transformation.

2.7 Protein localization

A typical plant cell contains 5,000 to 10,000 different polypeptide sequences and billions of individual protein molecules. If such a cell is to function properly, it must direct these proteins to specific metabolic compartments, cytoplasmic structures, and membrane systems. Cells therefore require the necessary machinery to sort each protein and direct it to its proper destination (Buchanan, 2000).

All proteins, except those that remain in the compartment where they are translated, include one or more targeting signals that act as an address label. Targeting signals are usually short peptides or amino acid motifs and are often located at the amino-(N-) terminal end of a protein but may be present in the carboxyl (C) terminus or elsewhere in the amino acid sequence. Different names apply to targeting signals, depending on the organelle to which a protein is being targeted (table 2.3) (Buchanan, 2000).

Table 2.3 Peptide targeting signals for transport to different organelles (modified from Buchanan, 2000)

Organelle	**Address label (targeting domain)**	**Characteristic of sorting signal**
ER	Signal peptide (SP)	Consists of one or more positively charged amino acids followed by a stretch of 6 to 12 hydrophobic amino acids.
Chloroplast	Transit peptide (TP)	Contains 40 - 50 amino acids at the N-terminus of the sequence. Rich in basic, hydroxylate and hydrophobic amino acids and no acidic residues present. Contain amphipathic β-stands secondary structure.
Mitrochondrion	Presequence	Similar to TP but contain amphipathic α-helics secondary structure.
Nucleus	Nuclear localization signal (NLS)	Consists of one or more short, internal sequences with basic amino acids. Generally contain Arg and Lys and also have residues disrupt helical domain. Position in the amino acid sequence and not cleaved from nuclear protein.
Peroxisome	Peroxisome targeting signal (PTS)	Protein synthesize on free ribosome and contain conserved C-terminal sequence as (Ser, Ala or Cys) – (Lys, Arg or His) – Leu.
Vacuole	Vacuolar sorting signal (VSS)	Contain signal peptide to ER and the VSS at the N, C-terminal or internal sequence will sort them to vacuole.

Although the targeting sequence is essential for protein transport, it may not be part of the active protein, and proteases in the target location often remove the targeting sequence to create a functional, mature polypeptide (Buchanan, 2000).

Today large numbers of proteins are being tracked in living cell using a fluorescent marker called green fluorescent protein (GFP). Tagging proteins with GFP is as simple as attaching the gene for GFP to one end of the gene that encodes a protein of interest. In most cases, the resulting GFP fusion protein behaves in the same way as the original protein, and its movement can be monitored by following its fluorescent inside the cell by fluorescent microscopy. In order to learn more about the function of tomato ripening-associated membrane protein (TRAMP), the cDNA was fused to the GFP coding region (Kim, 2005). Then onion cells were bombarded to allow visualization of the TRAMP protein within the plant cell. The GFP-tagged TRAMP constructs was observed predominantly in the plasma membrane, but it was also found in the nucleus or possibly the ER surrounding it. In plasmolysed onion epidermal cells, in which the cell wall was separated from plasma membrane, GFP tagged TRAMP expression was observed in the plasma membrane but not the cell. The clear location of TRAMP in the plasma membrane suggested that the TRAMP were not involved in the passage of solutes between the vacuole and the cytosol like Chen *et al.* (2001) had modeled.

2.8 Chloroplast proteins

Although chloroplasts contain their own DNA and ribosomes, most of their proteins are encoded in the nuclear DNA and imported from the cytoplasm.

Chloroplast proteins encoded in nuclear DNA are synthesized in the cytosol by free ribosomes. These proteins are translated as precursors with an N-terminal transit peptide of 40-50 amino acids. The transit peptides are both necessary and sufficient for chloroplast import. The import of chloroplast precursors occurs at contact sites (proteinaceous channels) between the outer and inner envelope membranes, and precursor proteins that are in transit span both envelop membranes. The heat-inducible filamentation temperature-sensitive H (*ftsH*) like gene from tomato was targeted to the chloroplast. Under heat stress, FtsH is involved in the degradation of the unassembled Rieske FeS protein and 23 kDa fragment of oxidatively-damaged D1 proteins. These mechanisms help to protect chloroplast proteins from denaturation and repair the damaged proteins cell in heat stress (Sun, 2005). The protein localization assay was used to study the function of *WCor15* gene (cold response gene) from wheat in the development of freezing tolerance in plant cells. The two different plasmid constructs, 35S::*Wcor15*-GFP and *Wcor15*::*Wcor15*-GFP, were introduced into the tobacco genome by *Agrobacterium*-infection method. The results showed transport and abundant accumulation of the WCOR15 protein in the stromal compartment of the chloroplasts and induced expression of the wheat *Wcor15* gene positively correlates to the development of freezing tolerance in the heterologous tobacco plants (Shimamura, 2006).

2.9 Research objectives

Since the function of Arabidopsis *SFR2* gene that support freezing tolerance is still undiscovered. This study rice *SFR2* was cloned and expressed in recombinant

protein expression systems. Reporter genes were also used to tag the SFR2 to define the expression and localization of SFR2 in plant cell. The localization was used to predict the SFR2 function in the plant cell to stress adaptation.

CHAPTER III

MATERIALS AND METHODS

3.1 Cloning

3.1.1 Materials

3.1.1.1 Rice *SFR2* cDNA

The rice *SFR2* cDNA (DDBJ Accession AK119461) in the vector pME18SFL3 from clone 001-133-E10 provided by the Rice Genome Resource Center (National Institute for Agrobiological Sciences, Japan) was used as template.

3.1.1.2 Chemicals

Reagents for PCR amplification contained 10X *Pfu* buffer ($MgSO_4$), 10 mM dNTPs mix, 10 μM primers and *Pfu* polymerase. Reagents for agarose gel electrophoresis contained 0.5% TAE buffer, Agarose low EEO, 1 kb Ladder marker, 100 bp Ladder marker and Loading dye. PCR Purification and gel purification kits were purchased from Qiagen. The *Pst*I , *Hin*dIII, *Eco*RI, *Not*I enzymes and T4 DNA ligase were purchased from Promega. The LR clonase and BP clonase were purchased from Invitrogen. All chemicals used were molecular grade or analytical grade.

3.1.1.3 Primers

Gene specific primers derived from sequence of rice *SFR2* cDNA were obtained from Proligo/Geneset (Singapore)

Table 3.1 The oligonucleotide primers to prepare *SFR2* constructs and sequences in *E. coli* and *P. pastoris*

Name	**Sequence**	**Primer Size (bp)**	**Tm (°C)**	**GC Content (%)**
Start_f	CACCATGCCACTACCGGCGTTC	22	60	64
Costart_f	CACCGCTGACGACTCAGAGGAG	22	60	64
Stop_r	CTATGATGACAGGGCATCAG	20	52	50
co-pSY5 _f	CTATGCGGCCGCAAGCCGACTTC CGCGCC	29	71	72
stop-pSY5_r	CCGAATTCCTATAGTAGCAGGGC	23	57	52
co- pGEX-4T-3_f	CCGAATTCCGCCGACTTCCGCGC CTTTCCC	30	70	67
stop-pGEX-4T-3_r	CTATGCGGCCGCTATAGTAGCAG GGCATCAG	31	67	58
Fai4_f	TATGGCGGGTGGAAGATGG	19	53	58
Fai8_f	CCTGAGGAACCAACCGAAGAATT G	24	57	50
M13 F (-20)	GTAAAACGACGGCCAG	16	46	56
M13 R	CAGGAAACAGCTATGAC	17	45	47
T7 promoter	TAATACGACTCACTATAGGG	20	48	40
T7 terminator	TAGTTATTGCTCAGCGGTGG	20	52	50
pGEX5	GGCAAGCCACGTTTGGTG	18	53	61
pGEX3	GGAGCTGCATGTGTCAGAGG	20	56	60

3.1.1.4 Plasmids

1. Gateway® system (Invitrogen)

- Entry clone

The pENTR Directional TOPO® Cloning Kit was used to directionally clone a blunt-end PCR product, which includes the four base pair sequences (CACC) necessary for directional cloning at the 5′ end of the forward primer.

- Destination vectors

pET32a/DEST (provided by J. R. Ketudat Cairns) is a plasmid constructed from pET32a (Novagen) by inserting Gateway cloning cassette reading frame A into the *Eco*RV site (206).

The pPICZ∝BNH8/DEST (provided by J. R. Ketudat Cairns) contains an 8 histidine tag at the N-terminus of the expressed proteins after the yeast ∝-factor proper peptide sequence which directs the protein to secretion (Ketudat Cairns, personal communication). It also contains the *att*R1 and *att*R2 sites to allow recombination with Gateway entry clones by the LR cloning technique, so it serves as a destination vector for expression of the cDNA transferred from the entry clone (Invitrogen).

2. Conventional ligation

pSY5 His-tagged vector (derived from pET21d (+), R. Robinson personal communication) contains 8 histidine and prescission site at the N-terminus of the fusion protein.

pGEX-4T-3 (GE healthcare, Amershambiosciences) contains glutathione S-transferase and a thrombin recognition site at the N-terminus of the target protein.

3.1.2 Methods

3.1.2.1 Amplification of *SFR2* gene from AK119461 clone 001-133-E10

1. Prokaryotic vectors

- pET32a/DEST

The PCR technique was used to amplify the *SFR2* gene by designing two pairs of primers from the full-length rice *SFR2* cDNA sequence AK119461. The Start_f and Stop_r primers were used to amplify the *SFR2* gene with its N-terminal transit peptide sequence and β-glucosidase encoding sequence. The Costart_f and Stop_r primers were used to amplify the *SFR2* gene from the conserved region of β-glucosidases. Briefly, each 150 µl PCR reaction consisted of 1X *Pfu* buffer, 2.5 mM dNTPs, 0.2 µM primers, 2 µl of 1 ng/µl template and 1.5 U *Pfu* polymerase. The amplification conditions are presented in table 3.2. The PCR products were run on a 1% agarose gel.

Table 3.2 The PCR conditions used to amplify *SFR2* for pENTR/D-TOPO and pSY5 systems.

Condition	**Temperature(°C)**	**Time**	**Number of cycles**
Initial denaturation	95	2 min	1
Denaturation	95	30 sec	35
Annealing	55	45 sec	
Extension	72	3 min	
Final extension	72	5 min	1
Hold	4	∞	1

The bands of PCR products were excised from agarose gels and purified by gel purification kit. The pure PCR products were cloned into the pENTR/D-TOPO cloning vector (Invitrogen). The reaction consisted of 1 µl of 100 ng/µl PCR product, 1 µl of salt solution and 1 µl of pENTR/D-TOPO vector. The solutions were incubated at 25 °C for 30 minutes and transformed into *E. coli* (strain TOP10) by the heat shock technique and the *E. coli* cells were plated on LB agar containing 50 µg/ml kanamycin. The positive clones (TOPO-st and TOPO-co vectors) were extraction for DNA sequencing analysis with M13F (-20), Fai4_f and M13R primers.

The enzyme LR clonase from Gateway® technology (Invitrogen) was used to exchange two regions of the *SFR2* cDNA from the TOPO-st and TOPO-co vectors to pET32a/DEST (Invitrogen). The LR reaction consisted of 1 µl of 300 ng entry clone (TOPO-st or TOPO-co), 2 µl of 150 ng/µl destination vector (pET32a/DEST), 1X LR

buffer and 1 µl of LR clonase. The reactions were incubated at 25°C overnight and then 1 µl of proteinase K was added to stop the reaction. Transformation of plasmids containing rice *SFR2* into *E. coli* strain TOP10 was done by the heat shock technique and the cells were plated on LB agar containing 100 µg/ml ampicillin. The plasmids were extracted from positive clones for DNA sequencing analysis. The T7 promoter and T7 terminator primers were used to sequence pET32a/DEST-st and pET32a/DEST-co.

- pSY5 vector

The conserved region of β-glucosidases the rice *SFR2* cDNA sequence AK119461 was amplified with co-pSY5_f and stop-pSY5_r primers containing sites for *Not*I and *Eco*RI, respectively. Briefly, each 150 µl PCR reaction consisted of 1X *Pfu* buffer, 2.5 mM dNTPs, 0.2 µM primers, 2 µl of 1 ng/µl template and 1.5 U *Pfu* polymerase. The amplification conditions are presented in table 3.2. The PCR product was run on a 1% agarose gel.

The band of PCR product was excised from the agarose gel and purified by gel purification kit. The pure PCR product and pSY5 vector were cut with *Eco*RI at 37 °C for 6 hours and then the digested cDNA and vector were purified by the PCR purification kit. The products were then cut again with *Not*I at 37 °C overnight and the cDNA was purified again by the PCR purification kit. The linear pSY5 vector and PCR product were ligated by T4 DNA ligase and incubated at 4 °C overnight. The reaction was transformed into *E. coli* strain TOP10 by heat shock and the cells were plated on LB agar containing 100 µg/ul ampicillin. The plasmids were extracted from positive clones for DNA sequencing with T7 promoter, Fai4_f and T7 terminator primers.

- pGEX-4T-3

The full-length rice AK119461 *SFR2* cDNA was used as template to amplify the *SFR2* gene starting from the conserved region of β-glucosidases. The co-pGEX-4T-3_f and stop-pGEX-4T-3_r primers, which contain recognition sites for *Eco*RI and *Not*I, respectively, were used. Briefly, each 150 µl PCR reaction consisted of 1X *Pfu* buffer, 2.5 mM dNTPs, 0.2 µM primers, 2 µl of 1 ng/µl template and 1.5 U *Pfu* polymerase. The amplification conditions are presented in table 3.3. The PCR product was run on a 1% agarose gel.

Table 3.3 The PCR conditions used to amplify *SFR2* for pGEX-4T-3 system.

Condition	**Temperature(°C)**	**Time**	**Number of cycles**
Initial denaturation	95	2 min	1
Denaturation	95	30 sec	5
Annealing	45	45 sec	
Extension	72	2 min	
Denaturation	95	30 sec	30
Annealing	65	45 sec	
Extension	72	2 min	
Final extension	72	5 min	1
Hold	4	∞	1

The band of PCR product was excised from the agarose gel and purified by the gel purification kit. The pure PCR product and pGEX-4T-3 vector were cut with *Eco*RI at 37°C for 6 hours and the products were purified by the PCR purification kit.

The purified products were cut with *Not*I at 37 °C for overnight and then reactions were purified again by PCR purification kit. The linear pGEX-4T-3 vector and PCR product were ligated by T4 DNA ligase and incubated at 4 °C overnight. The reaction was used to transform to *E. coli* strain TOP10 by heat shock and the cells were plated on LB agar containing 100 μg/ul ampicillin. The plasmids were extracted from positive clones for DNA sequencing analysis with pGEX5, Fai4_f and pGEX3 primers.

2. Eukaryotic vector

- pPICZαBNH8/DEST

The LR clonase (Invitrogen) was used to exchange two regions of the *SFR2* cDNA from the TOPO-st and TOPO-co vectors to pPICZαBNH8/DEST (Invitrogen). The LR reaction consisted of 1 μl of 300 ng entry clone (TOPO-st or TOPO-co), 2 μl of 150 ng/ul destination vector (pPICZαBNH8/DEST), 1X LR buffer and 1 μl of LR clonase. The reactions were incubated at 25 °C overnight and then 1 μl of proteinase K was added to stop the reaction. Transformation of plasmids containing rice *SFR2* into *E. coli* strain TOP10 was done by heat shock technique and plate on LB agar containing 100 μg/ml ampicillin. The plasmids were extracted from positive clones for DNA sequencing analysis. The 5'AOX and 3'AOX primers were used to sequence the plasmid pPICZαBNH8/DEST-st and pPICZαBNH8/DEST-co.

3.1.2.2 Bioinformatics

The National Center for Biotechnology Information Web site (http://www.ncbi.nlm.nih.gov/) was used to search the pubic DNA sequence database for AK119461. The Expasy translation tool program was used to translate the rice *SFR2* gene to amino acid sequences (http://www.expasy.org/tools/dna.html). The Blastp program (http://www.ncbi. nih.gov) was used to aligned rice SFR2 protein with other proteins in the database. The Clustal multiple alignment program (version 1.82) was used to make alignment between amino acid sequences of rice SFR2 and SFR2 orthologs.

3.2 Protein expression

3.2.1 Materials

The bacteria strains were used to be host contained *E. coli.* BL21 (DE3), Origami (DE3), Origamin B (DE3), Rosetta (DE3), Rosetta-gami (DE3) and BL21 pLysS (DE3). The yeast strain was used to be host as *P. pastoris* Y-11430.

Reagents for *E. coli* competent cell preparation contained 10% ice-cold glycerol and ice-cold GYT media. Reagents for *P. pastoris* competent cell preparation contained ice-cold deionize water and 1 M sorbitol. Reagents for *E. coli* breaking buffer contained 1 mg/ml lyzozyme, 20 mM Tris-HCl pH 8 and 1 mM PMSF. Reagents for *P. pastoris* breaking buffer contained 50 mM sodium phosphate buffer pH 7.5, 1 mM EDTA, 5% glycerol and 1 mM PMSF. Reagents for RNA assay in *P. pastoris* contained Trizol reagent, Chloroform, Isopropyl alcohol, 75% ethanol,

Superscrip III RT, RNase inhibitor, DNase and buffer RQ1, Q_t primer, 100 pmol dNTPs and DTT. Substrate for activity assay contained *p*NP-β-D-glucoside, *p*NP-β-D-galactoside, *p*NP-β-D-fucoside, *o*NP-β-D- glucoside, *p*NP-1-thio-β-D-glucoside and *p*NP-β-D-xylotoside dissolved in 50 mM sodium acetate buffer pH 5. The Ni-column purification was purchased from Qiagen. All chemicals used were molecular grade or analytical grade.

3.2.2 Methods

3.2.2.1 Protein expression in *E. coli*

1. Survey for appropriate strains

The pET32a/DEST containing *st-SFR2* and *co-SFR2* were transformed into the expression hosts (table 2.2) by electroporation technique. The selected clones were grown in LB broth containing appropriate antibiotics (table 2.2) at 37 °C to an OD_{600} of 0.5 - 0.6. Then, IPTG was added to a final concentration of 1 mM and the cultures were incubated at 20 °C for 12-16 h. Induced cultures were harvested by centrifugation at 4,000 rpm 4 °C for 20 minutes. The cell pellets were resuspended in 5 ml of freshly prepared extraction buffer (20 mM Tris-HCl pH 8.0, 1 mg/ml lysozyme, 1 mM PMSF) per gram cells, then incubated at room temperature for 30 minutes, and disrupted by sonication. Centrifugation at 10,000 rpm 4 °C for 20 minutes were use to separate soluble and insoluble fractions. The soluble and insoluble fractions were collected for activity and protein assay.

The empty pET32a vector was used as control. It contains thioredoxin protein and associated tags.

2. Survey for appropriated vectors

The pSY5 and pGEX-4T-3 containing *co-SFR2* were transformed into Rosetta-gami (DE3) and DH5α, respectively. The selected clones were grown in LB broth containing 100 μg/ml ampicillin and appropriate antibiotics for the host cells (table 2.2). The expression processes and cell harvest followed method 1.1.

The empty vector, pSY5 or pGEX-4T-3 that contained glutathione-s-transferase, was used as control.

3. Induction time assay

The pET-st, pET-co and pSY5-co were transformed into Rosetta-gami (DE3). The selected clones were grown in LB broth containing 100 μg/ml ampicillin, 50 μg/ml kanamycin and 10 μg/ml tetracycline. The expression process followed method 1.1, except for the cell suspension were collected every hour before induction and every 30 minutes for 3 hours after induction.

4. Induction temperature assay

The pET-st and pET-co were transformed into Rosetta-gami (DE3). The selected clones were grown in LB broth containing 100 μg/ml ampicillin, 50 μg/ml kanamycin and 10 μg/ml tetracycline at 37 °C to an OD_{600} of 0.5-0.6. Then, IPTG was added to 1 mM final concentration and the cultures were incubated at three different conditions, 4, 37 and 45 °C, for 12-16 hours. Then the cells were harvest following method 1.1.

3.2.2.2 Protein expression in *P. pastoris*

1. RNA assay

The pPICZ∝BNH8/DEST containing *st-SFR2* and *co-SFR2* were transformed to the expression host; *P. pastoris* strain Y-11430. The clones were collected by colony PCR technique. The selected clones were inoculated in BMGY medium (1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer, pH 7.0, and 1% glycerol) and grown at 30 °C, 250 rpm 48 hours. Then, methanol was add every 12 hours to final concentrations of 0.5, 1 and 3 % until 72 hours. The 1 ml of cell cultures were collected by centrifugation at 14,000 rpm for 10 seconds in micro centrifuge tubes. The cell pellets were resuspended in 100 µl trizol reagent and glass beads were added for disruption of the cells by vortexing. The supernatant were transfer to new micro centrifuge tubes and 400 µl trizol reagent were added and incubated 5 minutes at room temperature. One hundred microliters of chloroform were added and the tubes were shake vigorously by hand for 15 seconds. The aqueous phase that contained RNA was transferred to new microcentrifuge tubes and the RNA was precipitated by adding 250 µl of isopropyl alcohol. The solutions were incubated 15 °C for 10 minutes and then centrifuged at 12,000 rpm 4 °C for 10 minutes. The RNA were washed with 75% ethanol and collected by centrifugation at 12,000 rpm 4 °C for 5 minutes. The ethanol was removed by pipetting and dried for short time. The RNA was dissolved in RNAse free water.

Two step RT-PCR was done. Briefly, the DNase treated solution contained 8 µl RNA, 1 µl 10x RQ_1 buffer and 1 µl DNAse was incubate at 37 °C for 50 minutes. Then, 1 µl of stop solution was added and incubated at 70 °C for 20 minutes. The RT-

PCR solution contained 11 µl DNase treated RNA solution, 1 µl of 100 pmol Q_t primer, 1 µl of 10 mM dNTPs was incubated at 65 °C for 5 minutes. Then, the solution was put on ice for 3 minutes. Four microliters 5X first strand buffer, 1 µl DTT, 1 µl RNAse out and 1 µl superscript II transcriptase was added and incubated at 55 °C for 60 minutes. The solution was inactivated by heat at 70 °C for 15 minutes.

The cDNA were amplified with 5'AOX and 3'AOX primers. Then, nested PCR with a second pair of primers, Costart_f + Stop_r and Fai8_f + Stop_r was done.

2. Expression

The positive clones were inoculated in BMGY medium (1% yeast extract, 2% peptone, 100 mM potassium phosphate buffer, pH 7.0, and 1% glycerol) and grown at 30 °C, 250 rpm 48 hours, then methanol was added every 12 hours to final concentrations of 0.5, 1 and 3% until 72 hours. Induced cultures were harvested by centrifugation at 12,000 rpm at 4 °C for 5 minutes. The crude enzymes were concentrated by ultrafiltration and dissolved in 50 mM sodium acetate pH 5. The pellets were resuspended in breaking buffer and disrupted by glass beads and vortexing. The soluble and insoluble fractions were collected by centrifugation at 14,000 rpm 5 minutes. The crude enzyme, soluble and insoluble fractions were subjected to activity and protein assays.

3.2.2.3 Activity assay

The 20 µl samples were incubated with 70 µl of 2 mM of substrate in 50 mM sodium acetate buffer, pH 5 in microtiter plate wells at 30 °C for 30 minutes. The reactions were stopped by adding 70 µl of 0.4 M Na_2CO. The absorbance of the *p*-nitrophenol liberated were measured at 405 nm.

3.3 Protein localization

3.3.1 Materials

3.3.1.1 Plants tissues and reagents

Plant cells contained Onion (*Allium cepa L.*), Hydrilla (*Hydrilla verticillata (Lf) Royle.*). The PDA-1000/He machine (Biolistic particle delivery system, Biorad) was used to bombardment**.**

Reagents for plasmid construction contained 10X *Pfu* buffer ($MgSO_4$), 10 mM dNTPs mix, 10 μM primers, *Pfu* polymerase, enzyme BP clonase (Invitrogen), 5X BP buffer (Invitrogen), enzyme LR clonase (Invitrogen), 5X LR buffer (Invitrogen) and Gel purification kit (Qiagen). Reagents for preparation of particle contained tungsten M-10 microcarriers (Biorad), 70% ethanol, 0.1 M spermidine and 2.5 M $CaCl_2$. Reagents for GUS assay contained 10 mM sodium phosphate buffer pH 8, 0.5 mM potassium ferricyanide, 0.5 mM potassium ferrocyanide, 0.1% Triton X-100 and 0.5 mg/ml X-gluc (5-bromo-4-chloro-3-β-D-glucuronic acid sodium salt). All chemicals used were molecular grade or analytical grade.

3.3.1.2 Plasmids

- Entry clone

The pDONR/Zeo (Invitrogen) contains *att*P sites used to clone genes of interest PCR products flanked by *att*B and to generate entry clones.

- Destination vectors

The pMDC43 and pMDC83 (Curtis, 2003) are derived from pCAMBIA and contain a GFP reporter gene at the N-terminal and C-terminal ends of the Gateway

cloning cassette, respectively (Fig. 3.1). The pMDC140 (Curtis, 2003) also derived from pCAMBIA contains a GUS reporter gene at the C-terminal end of the Gateway cloning cassette. All of these plasmids contain *att*R sites for recombination with the entry clone in an LR reaction (Invitrogen) to generate an expression clone.

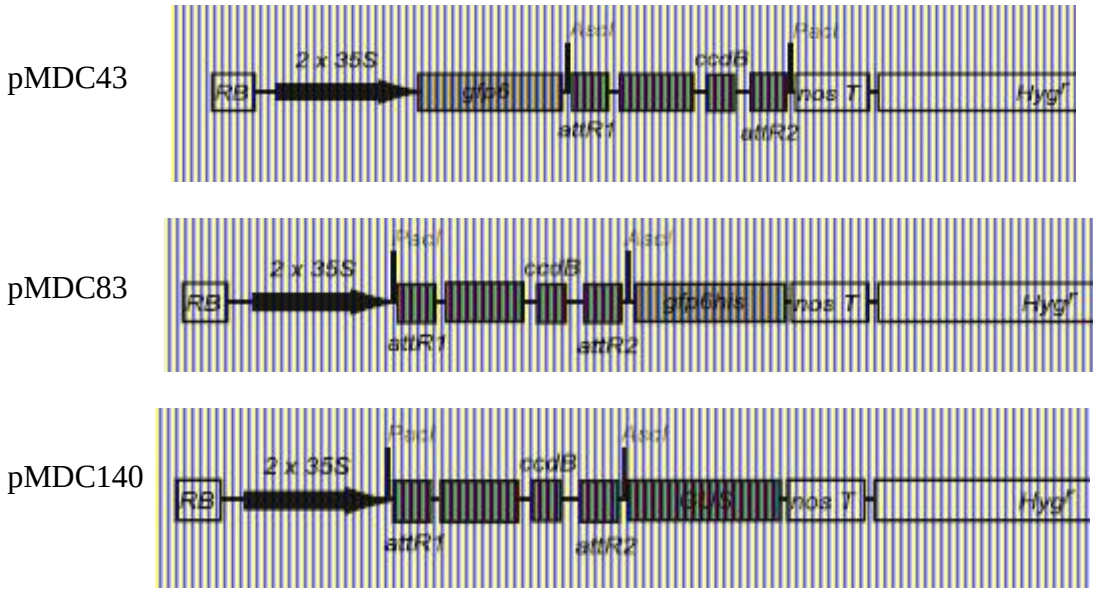


Figure 3.1 Destination vectors used in this report (copied from Curtis, 2003).

3.3.1.3 Primers

Gene specific primers derived from the sequence of rice *SFR2* cDNA were obtained from Proligo/Geneset (Singapore). The underlined *SFR2* sequences and italic att*B* sites are show in table 3.4.

Table 3.4 The oligonucleotide primers for preparing *SFR2* constructs for expression and sequencing in plant cells.

Name	**Sequence**	**Primer Size (bp)**	**Tm (°C)**	**GC Content (%)**
SFR2B1_f	GGGG*ACAAGTTTGTACAAAAAAG* *CAGGCTTC*ATGCCACTACCGGC GTTCGTGGGGG	56	76	55
SFR2B2_r	GGGG*ACCACTTTGTACAAGAAAG* *CTGGGT*TAATAGAATAAGAAAT AACATCATC	54	68	37
pMDC43_GFP	GAAGACGGCGGCGTGCAACTC	21	60	67
pMDC43_nosT	CCGGCAACAGGATTCAATC	19	51	53
pMDC85_139_35S	CATTTGGAGAGGACCTCGAC	20	54	55
pMDC85_GFP	CCTTCACCCTCTCCACTGAC	20	56	60
pMDC139_GUS	CTGGCACAGCAATTGCCCGG	20	58	65

3.3.2 Methods

3.3.2.1 Plasmid construction

The full-length rice *SFR2* cDNA sequence AK119461 was amplified with the SFR2B1_f and SFR2B2_r primers, which contain *att*B sites, allowing for recombination reactions with the pDONR/Zeo (Invitrogen). Briefly, 150 µl PCR reaction consisted of 1X *Pfu* buffer, 2.5 mM dNTPs, 0.2 µM primers, 2 µl of 1 ng/µl template and 1.5 U *Pfu* polymerase. The amplification conditions present in table 3.5.

Table 3.5 The PCR conditions used to amplify *SFR2* gene for pDONR/Zeo system.

Condition	**Temperature(°C)**	**Time**	**Number of cycles**
Initial denaturation	95	2 min	1
Denaturation	95	30 sec	10
Annealing	45	45 sec	
Extension	72	2 min	
Denaturation	95	30 sec	25
Annealing	59	45 sec	
Extension	72	2 min	
Final extension	72	5 min	1
Hold	4	∞	1

The band of PCR product was excised from agarose gel and purified by gel purification kit. The pure PCR product was cloned into pDONR/Zeo cloning vector (Invitrogen). The reaction consisted of 100 fmol PCR product, 150 ng/µl pDONR/Zeo, 1X BP buffer and 2 µl BP clonase. The reaction was incubated at 25 °C overnight and the clonase was inactivated by digestion with 1 µl proteinase K at 37 °C for 10 minutes. The reaction was used to transform TOP10 cells by heat shock and the cells were plated on LB Low-salt containing 100 µg/ml Zeocin.

The colony PCR technique was used to select positive clones with Fai4_f and M13R primers. Selected positive clones were used for plasmid extraction and sequencing analysis. The SFR2-pDONR/Zeo plasmid was sequenced with M13F (-20), costart_f, Fai4_f and M13R primers. Then, the LR reaction was done. Briefly, the

reaction consisted of 300 ng SFR2-pDONR/Zeo, 300 ng destination vector (pMDC43, pMDC83 or pMDC140), 1X LR buffer and 2 µl LR clonase. The reactions were incubated at 25 °C overnight and the clonase was inactivated by digestion with 1 µl proteinase K at 37 °C for 10 minutes. The reaction products were used to transformed TOP10 cells by heat shock, and the cells were plated on LB agar containing 100 µg/ml kanamycin. The recombinant clones from pMDC43, pMDC83 and pMDC140 were sequenced with SM43-GFP_f and SM43-NOS_r, SM83-35S_f and SM83-GFP_r, and SM140-35S_f and SM140-GUS_r primers, respectively, to check only the junction of the SFR2 gene and the plasmids.

3.3.2.2 Localization

1. Preparation of tissue

Onions were washed three times in water and soaked in 30% of sodium hyperchlorite for 15 minutes. Then, they were washed two times with water and soaked in 70% ethanol for 5 minutes. The onions skin were washed again with sterilized water two times before bombardment. After all the washing steps the onions were cut in quarters and sliced parallel to the longitudinal axis of the onion leaf. The peel epidermis was taken off in strip and laid upper side down on MS agar.

Hydrilla young leaves 0.2×2 cm^2 were cut and washed five times in water. The leaves were put with the upper side face down on 2% agar.

2. Preparation of particles

Three milligrams of tungsten M-10 particles (Biorad) were washed with absolute ethanol by vortexing. Then, the particles were centrifuged at 10,000 rpm for 10 seconds and supernatant were removed. Particles were washed 3X with 50 µl of

sterile water and centrifugation was done to remove the water. Fifty microliters of sterile water and 10 µl of 1 µg/ µl plasmid construct were added. The particles were gently mixed by inverting the tube two or three times. Then, 20 µl of 0.1 M spermidine and 50 µl of 2.5 M $CaCl_2$ were added and incubated on ice for 20 minutes with mixing every 2-3 minutes. The supernatant was removed by centrifugation at 10,000 rpm for 10 seconds. Then, 60 µl of absolute ethanol were added and 9 µl of suspension were used per shot.

3. Bombardment

The plasmid constructs were introduced into onion and hydrilla cells by particle bombardment (PDA-1000/He, BioRad) at 1,100 psi per shot at a distance of 6 cm. After that, the tissues were incubated at 25 °C for 48 hours in the dark. The tissues were subjected to fluorescent microscopy analysis or GUS assay to detect GFP protein and GUS protein, respectively.

4. Fluorescent microscopy analysis

The transformed cells on the slide were visualized with an IX71 BP70 microscope at fluorescent filter's wavelength, 520–590 nm, with the 40X objectives (Olympus).

5. GUS assay

The transformed cells were incubated with X-gluc solution (10 mM Sodium phosphate buffer pH8, 0.5 mM Potassium ferricyanide, 0.5 mM Potassium ferrocyanide, 0.1 % Triton X-100 and 0.5 mg/ml 5-bromo-4-chloro-3-β-D-glucuronic acid sodium salt, X-gluc) at 37 °C for 12 hours in the dark and stopped reaction with

70% ethanol. Then the cells were visualized with IX71 BP70 microscope, the objective 40X was used (Olympus).

CHAPTER IV

RESULTS AND DISCUSSION

4.1 Cloning

4.1.1 Amplification of the rice *SFR2* gene from AK119461 clone 001-133-E10

4.1.1.1 Prokaryotic vectors

- pET32a/DEST

The Start_f and Stop_r primers were used to amplify the rice *SFR2* gene starting with transit peptide sequence whereas the Costart_f and Stop_r primers were used to amplify the *SFR2* gene starting with the conserved region of β-glucosidase from clone 001-133-E10. The PCR product of the Start_f and Stop_r primers was designated as *st-SFR2* and was about 1.9 kb, while the product of Costart_f and Stop_r primers, designated as *co-SFR2*, was about 1.7 kb. The DNA band of the two products were excised from an agarose gel and purified for cloning (Fig. 4.1).

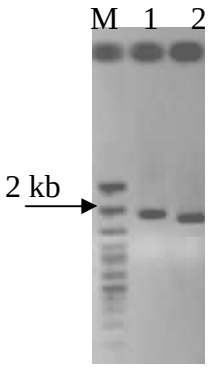


Figure 4.1 Agarose gel electrophoresis of purified *st-SFR2* and *co-SFR2* products. Lane M: 100 bp marker, Lane 1: *st-SFR2*, Lane: *co-SFR2*.

The *st-SFR2* and *co-SFR2* were cloned into the pENTR/D-TOPO vector (Invitrogen) and the product vectors were designated as TOPO-st and TOPO-co, respectively. The TOPO reactions were used to transformed TOP10 competent cells, and five clones of each constructs were chosen for plasmid preparation.

Five clones of TOPO-st were cut with *Pst*I, which gave rise to two DNA fragments of 1.2 and 3.3 kb, whereas *Hin*dIII cut only once to generate a linear plasmid of 4.5 kb (Fig. 4.2).

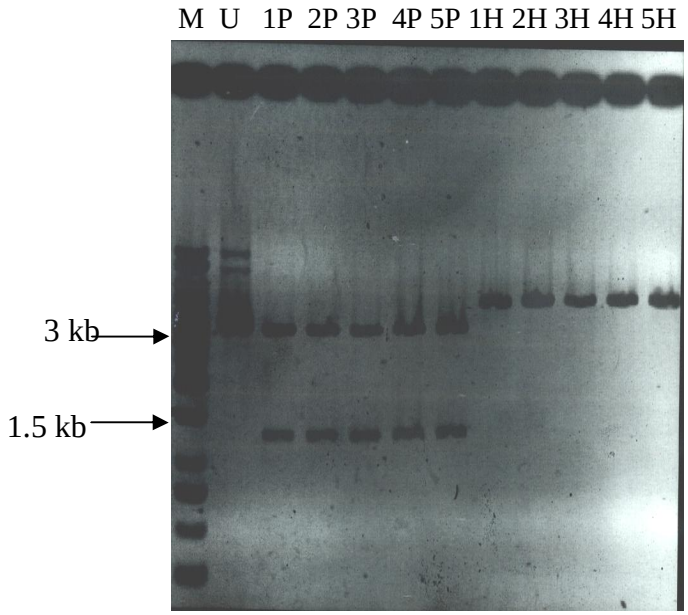


Figure 4.2 Agarose gel electrophoresis of TOPO-st digestion with *Pst*I and *Hin*dIII. Lane M: 1 kb marker, Lane U: TOPO-st, Lanes 1P – 5P: Colonies 1 – 5 of TOPO-st cut with *Pst*I, Lanes 1H – 5H: Colonies 1 – 5 of TOPO-st cut with *Hin*dIII.

Five clones of TOPO-co were cut with *Pst*I, which give rise to two DNA fragments of 1.2 and 3.1 kb, whereas *Hin*dIII cut only once to generate a linear plasmid of 4.3 kb (Fig. 4.3).

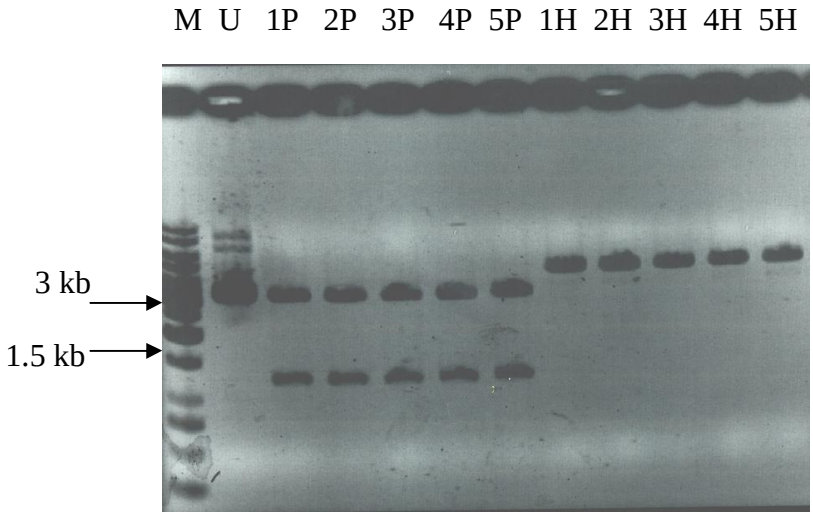


Figure 4.3 Agarose gel electrophoresis of TOPO-co digestion with *Pst*I and *Hin*dIII. Lane M: 1 kb marker, Lane U: TOPO-co, Lanes 1P – 5P: Colonies 1 – 5 of TOPO-co cut with *Pst*I, Lanes 1H - 5H: Colonies 1 – 5 of TOPO-co cut with *Hin*dIII.

The results indicated that all selected colonies of the TOPO-st and TOPO-co vectors contain *SFR2* gene. The TOPO-st colony 5 and TOPO-co colony 3 were used for DNA sequencing analysis. The DNA sequencing results of TOPO-st and TOPO-co were aligned with AK119461 as shown in Fig. 4.4.

```
AK119461        AAGAAAATCACGGAAATGCCACTACCGGCGTTCGTGGCGGCGGCGGCGAGGCTCGCCGTC 120
co-sfr2         ------------------------------------------------------------
st-sfr2         ---------------ATGCCACTACCGGCGTTCGTGGCGGCGGCGGCGAGGCTCGCCGTC 45

AK119461        CTCGTGGCCGCGGCGGCGACGGCGGCCAATGCCGCCTCCTACGCGCGGTACCGGCGGCGC 180
co-sfr2         ------------------------------------------------------------
st-sfr2         CTCGTGGCCGCGGCGGCGACGGCGGCCAATGCCGCCTCCTACGCGCGGTACCGGCGGCGC 105

AK119461        CACCTCCGCCGCATCCCCAGCCCCATCGACGAGTCCGCCGATCCCCTCGCCGACTTCCGC 240
co-sfr2         ------------------------------------------------------------
st-sfr2         CACCTCCGCCGCATCCCCAGCCCCATCGACGAGTCCGCCGATCCCCTCGCCGACTTCCGC 165

AK119461        GCCTTTCCCTCCTCCGACGCTGACGACTCAGAGGAGGATAATTTCTTCTTTGGGCTAGCA 300
co-sfr2         ------------------GCTGACGACTCAGAGGAGGATAATTTCTTCTTTGGGCTAGCA 42
st-sfr2         GCCTTTCCCTCCTCCGACGCTGACGACTCAGAGGAGGATAATTTCTTCTTTGGGCTAGCA 225
                            ****************************************
```

Figure 4.4 Alignment of two *SFR2* regions in pENTR/D-TOPO and AK119461 using ClustalW.

```
AK119461        ACAGCGCCTGCGCATGTTGAGGACAGGCTAGAAGATGCTTGGCTCCAGTTTGCAACTGAA 360
co-sfr2         ACAGCGCCTGCGCATGTTGAGGACAGGCTAGAAGATGCTTGGCTCCAGTTTGCAACTGAA 102
st-sfr2         ACAGCGCCTGCGCATGTTGAGGACAGGCTAGAAGATGCTTGGCTCCAGTTTGCAACTGAA 285
                ************************************************************

AK119461        ACTTCCTGCGATGACAATGGAAACGTGCGCGACCAGAGGCCAGTAGATGCACTGATGGCA 420
co-sfr2         ACTTCCTGCGATGACAATGGAAACGTGCGCGACCAGAGGCCAGTAGATGCACTGATGGCA 162
st-sfr2         ACTTCCTGCGATGACAATGGAAACGTGCGCGACCAGAGGCCAGTAGATGCACTGATGGCA 345
                ************************************************************

AK119461        TCTGCTGCTGGTGATGGAGGCTCCCAGCAATCTTGGAGGTCAACAGGTGGGGAAAATATT 480
co-sfr2         TCTGCTGCTGGTGATGGAGGCTCCCAGCAATCTTGGAGGTCAACAGGTGGGGAAAATATT 222
st-sfr2         TCTGCTGCTGGTGATGGAGGCTCCCAGCAATCTTGGAGGTCAACAGGTGGGGAAAATATT 405
                ************************************************************

AK119461        GGTGATAGAGAGCAGAGGAAACCACTTAGGGTAGCCATGGAGGCTATGCTCAGGGGGTTT 540
co-sfr2         GGTGATAGAGAGCAGAGGAAACCACTTAGGGTAGCCATGGAGGCTATGCTCAGGGGGTTT 282
st-sfr2         GGTGATAGAGAGCAGAGGAAACCACTTAGGGTAGCCATGGAGGCTATGCTCAGGGGGTTT 465
                ************************************************************

AK119461        GAAATTCTTGCTGAGAGTGGAGAATCTGCTGGCGGCGATAACTGCAGCCACAATGTTGCT 600
co-sfr2         GAAATTCTTGCTGAGAGTGGAGAATCTGCTGGCGGCGATAACTGCAGCCACAATGTTGCT 342
st-sfr2         GAAATTCTTGCTGAGAGTGGAGAATCTGCTGGCGGCGATAACTGCAGCCACAATGTTGCT 525
                ************************************************************

AK119461        GCTTGGCATAATGTTCCATGCCCGCAAGAAAGGCTTAGATTTTGGTCTGATCCTGATGCT 660
co-sfr2         GCTTGGCATAATGTTCCATGCCCGCAAGAAAGGCTTAGATTTTGGTCTGATCCTGATGCT 402
st-sfr2         GCTTGGCATAATGTTCCATGCCCGCAAGAAAGGCTTAGATTTTGGTCTGATCCTGATGCT 585
                ************************************************************

AK119461        GAGTTGAAACTTGCTAAGGAGACCGGCATCAGCGTTTTCCGCATGGGGGTAGATTGGGCA 720
co-sfr2         GAGTTGAAACTTGCTAAGGAGACCGGCATCAGCGTTTTCCGCATGGGGGTAGATTGGGCA 462
st-sfr2         GAGTTGAAACTTGCTAAGGAGACCGGCATCAGCGTTTTCCGCATGGGGGTAGATTGGGCA 645
                ************************************************************

AK119461        AGGTTAATGCCTGAGGAACCAACCGAAGAATTGAAGAGCTCAGTTAATTTTGCAGCACTT 780
co-sfr2         AGGTTAATGCCTGAGGAACCAACCGAAGAATTGAAGAGCTCAGTTAATTTTGCAGCACTT 522
st-sfr2         AGGTTAATGCCTGAGGAACCAACCGAAGAATTGAAGAGCTCAGTTAATTTTGCAGCACTT 705
                ************************************************************

AK119461        GAGCGGTATAGATGGATCATTCAAAGGGTTCGCGAATATGGAATGAAAGTAATGCTTACT 840
co-sfr2         GAGCGGTATAGATGGATCATTCAAAGGGTTCGCGAATATGGAATGAAAGTAATGCTTACT 582
st-sfr2         GAGCGGTATAGATGGATCATTCAAAGGGTTCGCGAATATGGAATGAAAGTAATGCTTACT 765
                ************************************************************

AK119461        CTGTTTCATCACTCACTTCCACCCTGGGCTGGAAAATATGGCGGGTGGAAGATGGAAAAA 900
co-sfr2         CTGTTTCATCACTCACTTCCACCCTGGGCTGGAAAATATGGCGGGTGGAAGATGGAAAAA 642
st-sfr2         CTGTTTCATCACTCACTTCCACCCTGGGCTGGAAAATATGGCGGGTGGAAGATGGAAAAA 825
                ************************************************************

AK119461        ACTGTCACCTATTTCATGGATTTTGTGAGGCTTGTTGTTGATTGTGTATCCAATTTAGTG 960
co-sfr2         ACTGTCACCTATTTCATGGATTTTGTGAGGCTTGTTGTTGATTGTGTATCCAATTTAGTG 702
st-sfr2         ACTGTCACCTATTTCATGGATTTTGTGAGGCTTGTTGTTGATTGTGTATCCAATTTAGTG 885
                ************************************************************

AK119461        GACTACTGGGTGATTTTCAATGAACCTCACGTTTTTGTGATGCTAACTTATTGTGCCGGT 1020
co-sfr2         GACTACTGGGTGATTTTCAATGAACCTCACGTTTTTGTGATGCTAACTTATTGTGCCGGT 762
st-sfr2         GACTACTGGGTGATTTTCAATGAACCTCACGTTTTTGTGATGCTAACTTATTGTGCCGGT 945
                ************************************************************

AK119461        GCTTGGCCTGGTGGAGACCCTAATGCAATTGAAGTAGCAACATCTACTCTGCCAACTGGT 1080
co-sfr2         GCTTGGCCTGGTGGAGACCCTAATGCAATTGAAGTAGCAACATCTACTCTGCCAACTGGT 822
st-sfr2         GCTTGGCCTGGTGGAGACCCTAATGCAATTGAAGTAGCAACATCTACTCTGCCAACTGGT 1005
                ************************************************************
```

Figure 4.4 Alignment of two *SFR2* regions in pENTR/D-TOPO and AK119461 using ClustalW. (Continue)

```
AK119461        GTATACAATCAGGCTTTGCATTGGATGGCTATTGCACATTCTGAAGCCTATGACTACATA 1140
co-sfr2         GTATACAATCAGGCTTTGCATTGGATGGCTATTGCACATTCTGAAGCCTATGACTACATA 882
st-sfr2         GTATACAATCAGGCTTTGCATTGGATGGCTATTGCACATTCTGAAGCCTATGACTACATA 1065
                ************************************************************

AK119461        CATTCGAAAAGCAAGAACGAAAGGAAGCCAATAGTTGGTGTTGCACACCATGTATCGTTT 1200
co-sfr2         CATTCGAAAAGCAAGAACGAAAGGAAGCCAATAGTTGGTGTTGCACACCATGTATCGTTT 942
st-sfr2         CATTCGAAAAGCAAGAACGAAAGGAAGCCAATAGTTGGTGTTGCACACCATGTATCGTTT 1125
                ************************************************************

AK119461        ACAAGGCCATATGGGCTATTTGATGTTGCTGCTGTCGCGCTAGCTAACTCATTGACCCTT 1260
co-sfr2         ACAAGGCCATATGGGCTATTTGATGTTGCTGCTGTCGCGCTAGCTAACTCATTGACCCTT 1002
st-sfr2         ACAAGGCCATATGGGCTATTTGATGTTGCTGCTGTCGCGCTAGCTAACTCATTGACCCTT 1185
                ************************************************************

AK119461        TTTCCTTACGTGGATAGCATATGTGATAAATTGGACTTCATTGGAATCAACTACTATGGA 1320
co-sfr2         TTTCCTTACGTGGATAGCATATGTGATAAATTGGACTTCATTGGAATCAACTACTATGGA 1062
st-sfr2         TTTCCTTACGTGGATAGCATATGTGATAAATTGGACTTCATTGGAATCAACTACTATGGA 1245
                ************************************************************

AK119461        CAGGAGGTTATATCAGGACCTGGTCTAAAGCTTGTGGATAATGATGAGTATAGTGAATCT 1380
co-sfr2         CAGGAGGTTATATCAGGACCTGGTCTAAAGCTTGTGGATAATGATGAGTATAGTGAATCT 1122
st-sfr2         CAGGAGGTTATATCAGGACCTGGTCTAAAGCTTGTGGATAATGATGAGTATAGTGAATCT 1305
                ************************************************************

AK119461        GGTCGTGGGGTTTATCCTGATGGGCTGTTCCGCATCCTGATTCAATTCAATGAACGATAT 1440
co-sfr2         GGTCGTGGGGTTTATCCTGATGGGCTGTTCCGCATCCTGATTCAATTCAATGAACGATAT 1182
st-sfr2         GGTCGTGGGGTTTATCCTGATGGGCTGTTCCGCATCCTGATTCAATTCAATGAACGATAT 1365
                ************************************************************

AK119461        AAGAGATTAAATATACCTTTTGTCATTACTGAAAATGGAGTTTCTGATGAGACTGATCTG 1500
co-sfr2         AAGAGATTAAATATACCTTTTGTCATTACTGAAAATGGAGTTTCTGATGAGACTGATCTG 1242
st-sfr2         AAGAGATTAAATATACCTTTTGTCATTACTGAAAATGGAGTTTCTGATGAGACTGATCTG 1425
                ************************************************************

AK119461        ATACGGAAACCATATATACTGGAACACTTGTTAGCCACATACGCTGCCATCATTATGGGT 1560
co-sfr2         ATACGGAAACCATATATACTGGAACACTTGTTAGCCACATACGCTGCCATCATTATGGGT 1302
st-sfr2         ATACGGAAACCATATATACTGGAACACTTGTTAGCCACATACGCTGCCATCATTATGGGT 1485
                ************************************************************

AK119461        GTCCGTGTACTTGGTTATTTGTTTTGGACAACATCAGATAATTGGGAATGGGCGGATGGC 1620
co-sfr2         GTCCGTGTACTTGGTTATTTGTTTTGGACAACATCAGATAATTGGGAATGGGCGGATGGC 1362
st-sfr2         GTCCGTGTACTTGGTTATTTGTTTTGGACAACATCAGATAATTGGGAATGGGCGGATGGC 1545
                ************************************************************

AK119461        TATGGTCCCAAGTTTGGGCTTGTTGCTGTTGACCGTGCTAATAACCTAGCACGGAAACCT 1680
co-sfr2         TATGGTCCCAAGTTTGGGCTTGTTGCTGTTGACCGTGCTAATAACCTAGCACGGAAACCT 1422
st-sfr2         TATGGTCCCAAGTTTGGGCTTGTTGCTGTTGACCGTGCTAATAACCTAGCACGGAAACCT 1605
                ************************************************************

AK119461        CGGCCTTCATACTTTTTATTCTCCAGGGTTGTTACAACTGGAAAAATTACAAGACAGGAC 1740
co-sfr2         CGGCCTTCATACTTTTTATTCTCCAGGGTTGTTACAACTGGAAAAATTACAAGACAGGAC 1482
st-sfr2         CGGCCTTCATACTTTTTATTCTCCAGGGTTGTTACAACTGGAAAAATTACAAGACAGGAC 1665
                ************************************************************

AK119461        AGAATGTCTGCTTGGAGGGAGCTGCAACAAGCTGCAGTTCAAAAGAAAACACGTCCATTT 1800
co-sfr2         AGAATGTCTGCTTGGAGGGAGCTGCAACAAGCTGCAGTTCAAAAGAAAACACGTCCATTT 1542
st-sfr2         AGAATGTCTGCTTGGAGGGAGCTGCAACAAGCTGCAGTTCAAAAGAAAACACGTCCATTT 1725
                ************************************************************

AK119461        TTCAGGGCAGTGGATAAGCATGGTCGGATGTATGCAGGTGGTCTAGATCGGCCTATTCAG 1860
co-sfr2         TTCAGGGCAGTGGATAAGCATGGTCGGATGTATGCAGGTGGTCTAGATCGGCCTATTCAG 1602
st-sfr2         TTCAGGGCAGTGGATAAGCATGGTCGGATGTATGCAGGTGGTCTAGATCGGCCTATTCAG 1785
                ************************************************************
```

Figure 4.4 Alignment of two *SFR2* regions in pENTR/D-TOPO and AK119461 using ClustalW. (Continue)

```
AK119461        AGGCCCTTCATATTGCGGGATTGGAGATTTGGTCACTATAAAATGGAAGGCTTGCAGGAT 1920
co-sfr2         AGGCCCTTCATATTGCGGGATTGGAGATTTGGTCACTATAAAATGGAAGGCTTGCAGGAT 1662
st-sfr2         AGGCCCTTCATATTGCGGGATTGGAGATTTGGTCACTATAAAATGGAAGGCTTGCAGGAT 1845
                ************************************************************

AK119461        CCTTTGAGTTGCTTTATAAGATGCATTTTTGCACCATTTTCACGCCAAAAGAAGATTCAC 1980
co-sfr2         CCTTTGAGTTGCTTTATAAGATGCATTTTTGCACCATTTTCACGCCAAAAGAAGATTCAC 1722
st-sfr2         CCTTTGAGTTGCTTTATAAGATGCATTTTTGCACCATTTTCACGCCAAAAGAAGATTCAC 1905
                ************************************************************

AK119461        TACATAGAAGATGATGTTATTTCTTATTCTATTAACTGATGCCCTGCTACTATAGTTGCT 2040
co-sfr2         TACATAGAAGATGATGTTATTTCTTATTCTATTAACTGATGCCCTGCTACTATAG----- 1777
st-sfr2         TACATAGAAGATGATGTTATTTCTTATTCTATTAACTGATGCCCTGCTACTATAG----- 1960
                ********************************************************
```

Figure 4.4 Alignment of two *SFR2* regions in pENTR/D-TOPO and AK119461 using ClustalW. (Continue)

Based on the result, the sequences of TOPO-st and TOPO-co were correct with no mutations. The *st-SFR2* is 1,960 bp long and the *co-SFR2* is 1,777 bp in length. The LR clonase (Invitrogen) was used to transfer *st-SFR2* and *co-SFR2* from pENTR/D-TOPO vector to expression vector, pET32a/DEST. The reactions were transformed to *E. coli* TOP10 by heat shock. The *st-SFR2* and *co-SFR2* in pET32a/DEST were designated as pET-st and pET-co, respectively.

The PCR technique was used to screen the recombinant plasmids of pET-st and pET-co with Start_f and Stop_r, and Costart_f and Stop_r primers, respectively. Some positive clones showed the expected bands of pET-st and pET-co about 1.9 and 1.7 kb. (data not show). The DNA sequence analysis were used to identify *st-SFR2* and *co-SFR2* in pET32a/DEST with T7 promoter and T7 terminator primers. The result indicated that *st-SFR2* and *co-SFR2* were inserted in the correct reading frame following the thioredoxin protein (Trx) in pET32a/DEST (Fig. 4.5 and 4.6).

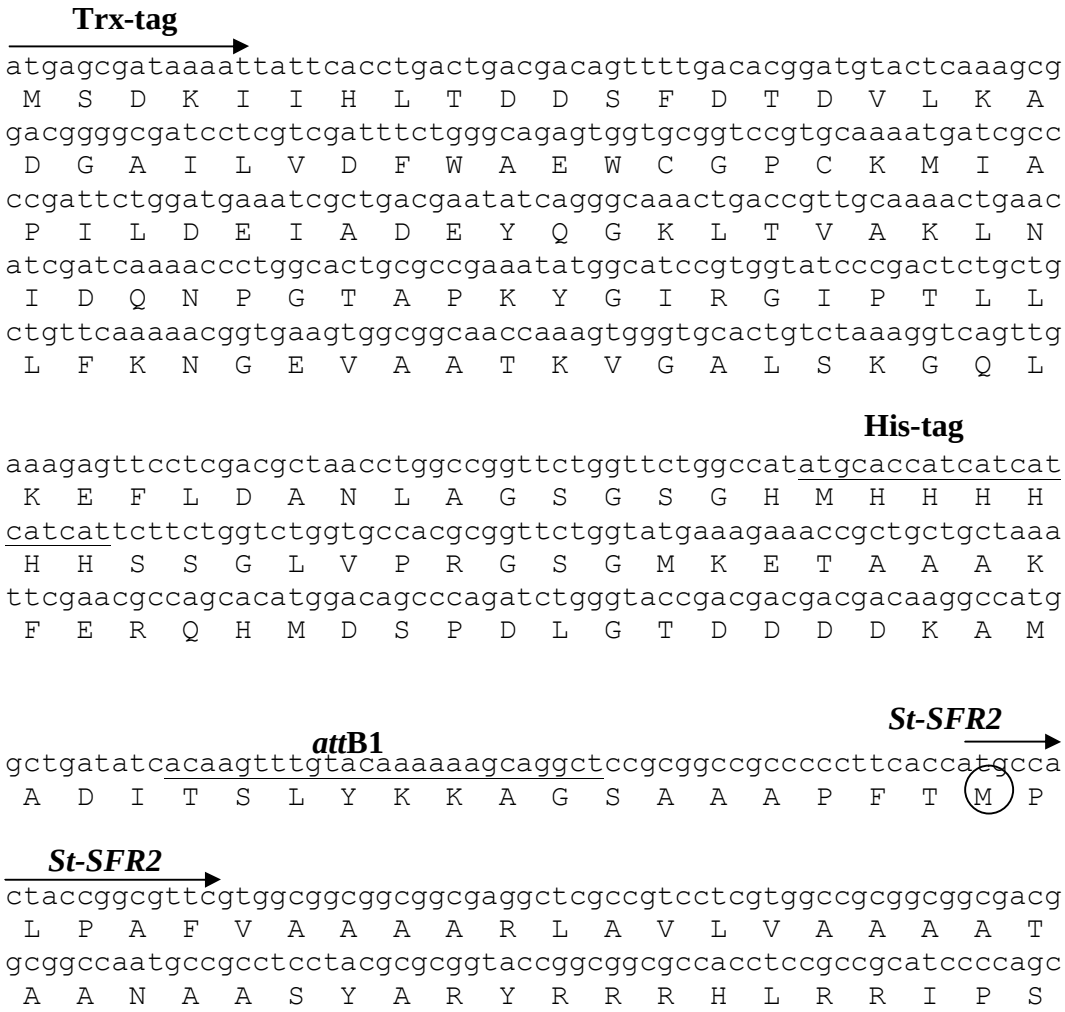


Figure 4.5 Translated DNA sequence of pET-st. The DNA sequence was translated by the Expasy Translation Tool (http://www.expasy.org/tools/dna.html).

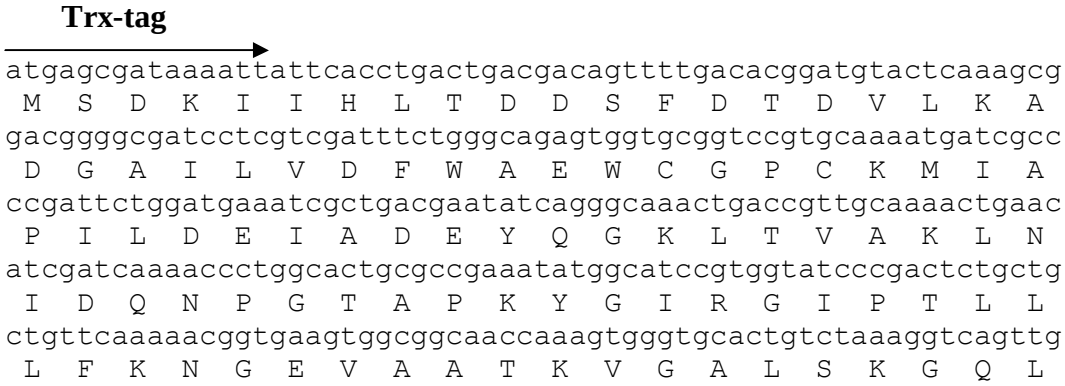


Figure 4.6 Translated DNA sequence of pET-co. The DNA sequence was translated by the Expasy Translation Tool (http://www.expasy.org/tools/dna.html).

His-tag

aaagagttcctcgacgctaacctggccggttctggttctggccatatgcaccatcatcat
K E F L D A N L A G S G S G H M H H H H
catcattcttctggtctggtgccacgcggttctggtatgaaagaaaccgctgctgctaaa
H H S S G L V P R G S G M K E T A A A K
ttcgaacgccagcacatggacagcccagatctgggtaccgacgacgacgacaaggccatg
F E R Q H M D S P D L G T D D D D K A M

att*B1** **Co-*SFR2

gctgatatcacaagtttgtacaaaaaagcaggctccgcggccgccccttcaccgctgac
A D I T S L Y K K A G S A A A P F T A D

gactcagaggaggataatttcttctttgggctagccacagcgcctgcgcatgttgaggac
D S E E D N F F F G L A T A P A H V E D

Figure 4.6 Translated DNA sequence of pET-co. The DNA sequence was translated by the Expasy Translation Tool (http://www.expasy.org/tools/dna.html). (Continue)

- pSY5

The co-pSY5 _f and stop-pSY5_r primers were used to amplify the rice *SFR2* gene from AK119461 starting with the conserved β-glucosidase region. The PCR product was then run on a 1% agarose gel. The DNA band was excised from the gel and purified. The PCR product and pSY5 were cut with *Not*I and *Eco*RI, to produce the expected size bands, as shown in Fig. 4.7.

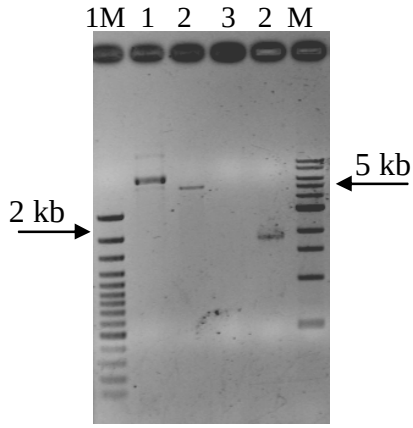


Figure 4.7 Agarose gel electrophoresis of the pSY5 and *co-SFR2* digestion with *Eco*RI and *Not*I. Lane 1M: 100 bp marker, Lane 2M: 1 kb marker, Lane 1: uncut

pSY5, Lane 2: pSY5 cut with *Eco*RI and *Not*I, Lane 3: PCR product of co-pSY5 _f and stop-pSY5_r cut with *Eco*RI and *Not*I.

T4 DNA ligase was used to ligate the linear plasmid and *co-SFR2* PCR product. Then ligation reaction was transformed to TOP10 competent cell. The colony PCR technique was used to screen for positive clones, designated as pSY5-co.

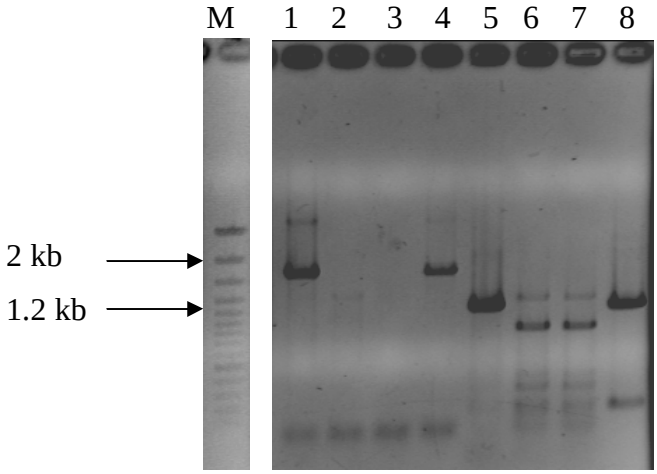


Figure 4.8 Agarose gel electrophoresis of pSY5-co colony screening with colony PCR. Lane M: 100 bp marker, Lanes 1 - 4: Colonies 1 – 3 of pSY5-co clones and plasmid AK119461 amplified with co-pSY5_f and stop-pSY5_r, Lanes 5 – 8: Colonies 1 – 3 of co-pSY5 clones and plasmid AK119461 amplified with Fai4_f and stop-pSY5_r.

The PCR product of pSY5-co colony 1 and AK119461 plasmid gave 1.8 kb and 1.2 kb bands when amplify with co-PSY5_f and stop-PSY5_r primers (Fig. 4.8 lanes 1 and 4) and Fai4_f and stop-PSY5_r primers (Fig. 4.8 lanes 5 and 8), respectively. The colony 1 was then used for plasmid preparation and DNA sequencing was done using T7 promoter, Fai4_f and T7 terminator primers. The DNA sequence of *co-SFR2* in pSY5 was in the correct reading frame containing 8 His-tag at the N-terminal side of the SFR2 protein.

His-tag

atggcagaagaacaccaccaccaccaccaccaccacctggaagttctgttccaggggccc
M A E E H H H H H H H H L E V L F Q G P

Co-SFR2

gggcggccgcaagccgacttccgcgcctttccctcctccgacgctgacgactcagaggag
G R P Q A D F R A F P S S D A D D S E E
gataatttcttctttgggctagcaaccgcgcctgcgcatgttgaggacaggctagaagat
D N F F F G L A T A P A H V E D R L E D
gcttggctccagtttgcaactgaaatttcctgcgatgacaatggaaacgtgcgcgaccag
A W L Q F A T E I S C D D N G N V R D Q

Figure 4.9 Translated DNA sequence of pSY5-co. The DNA sequence was translated by the Expasy Translation Tool (http://www.expasy.org/tools/dna.html).

- pGEX-4T-3

The co- pGEX-4T-3_f and stop- pGEX-4T-3_r primers were used to amplify *SFR2* gene from AK119461 starting from the region conserved in β-glucosidases. The PCR product was then ran on a 1% agarose gel. The DNA band was excised from the gel and purified (Fig. 4.10).

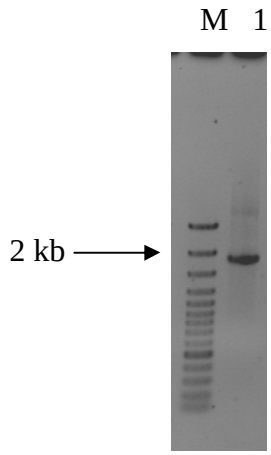


Figure 4.10 Agarose gel electrophoresis of purified *co-SFR2* PCR product for the pGEX-4T-3 system. Lane M: 100 bp marker, Lane 1: pure PCR product of co- pGEX-4T-3_f and stop- pGEX-4T-3_r primers.

The pGEX-4T-3 and PCR product were cut with *Eco*RI and *Not*I. Agarose gel electrophoresis was used to detect the result (Fig. 4.11).

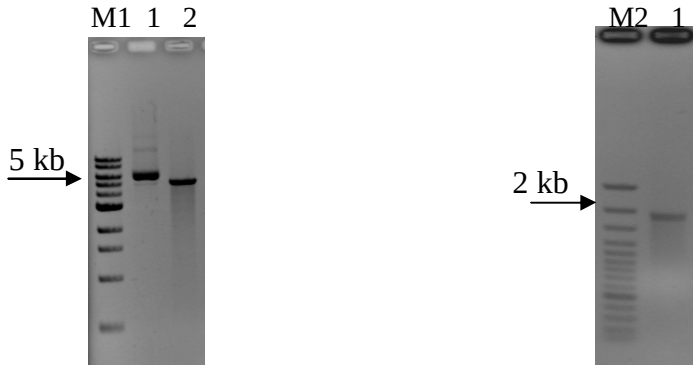


Figure 4.11 Agarose gel electrophoresis of pGEX4T-3 and *co-SFR2* digestion with *Eco*RI and *Not*I. A: pGEX-4T-3, Lane M1: 1 kb marker, Lane 1: pGEX-4T-3, Lane 2: pGEX-4T-3 cut with *Eco*RI and *Not*I. B: PCR product, Lane M2: 100 bp marker, Lane 1: PCR product cut with *Eco*RI and *Not*I.

The 4.9 kb linear pGEX-4T-3 plasmid and *co-SFR2* 1.8 kb PCR product were ligated with T4 DNA ligase. Then, the ligation reaction was transformed into TOP10 competent cell. The colony PCR technique was used to check for positive clones, designate as pGEX-4T-3-co, with the Fai4_f and Stop_r primers (Fig. 4.12).

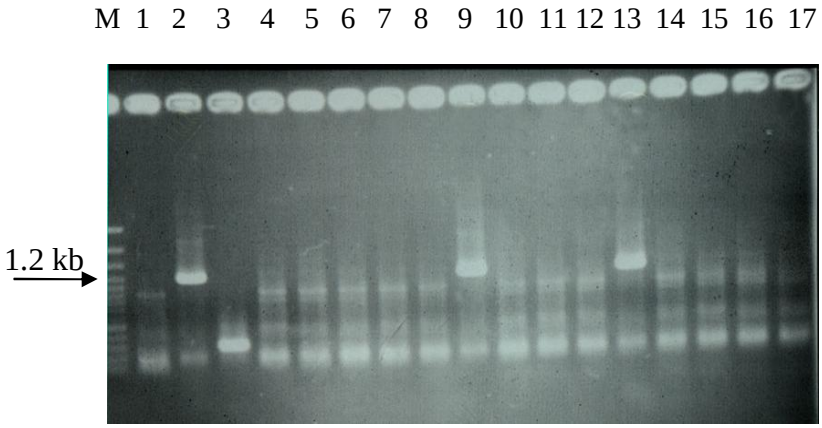


Figure 4.12 Agarose gel electrophoresis of pGEX-4T-3-co colony screening by colony PCR. Lane M: 100 bp marker, Lanes 1 – 17: Colonies 1- 17 of pGEX-4T-3-co amplified with Fai4_f and Stop_r.

PCR products of colonies 2, 9 and 13 showed the expected size of about 1.2 kb. The pGEX-4T-3-co colony 2 was extracted and used for DNA sequencing with pGEX5, Fai4_f and pGEX3 primers. The DNA sequence of *co-SFR2* in pGEX-4T-3 was in the correct reading frame fused to the glutathione-S-transferase (GST) at the *SFR2* N-terminal end (Fig. 4.13).

GST-tag ***co-SFR2***

```
cctccaaaatcggatctggttccgcgtggatccccgaattccgccgacttccgcgccttt
 P  P  K  S  D  L  V  P  R  G  S  P  N  S  A  D  F  R  A  F
ccctcctccgacgctgacgactcagaggaggataatttcttctttgggctagcaacagcg
 P  S  S  D  A  D  D  S  E  E  D  N  F  F  F  G  L  A  T  A
cctgcgcatgttgaggacaggctagaagatgcttggctccagtttgcaactgaaacttcc
 P  A  H  V  E  D  R  L  E  D  A  W  L  Q  F  A  T  E  T  S
tgcgatgacaatggaaacgtgcgcgaccagaggccagtagatgcactgatggcatctgct
 C  D  D  N  G  N  V  R  D  Q  R  P  V  D  A  L  M  A  S  A
gctggtgatggaggctcccagcaatcttggaggtcaacaggtggggaaaatattggtgat
 A  G  D  G  G  S  Q  Q  S  W  R  S  T  G  G  E  N  I  G  D
agagagcagaggaaaccacttagggtagccatggaggctatgctcaggggtttgaaatt
 R  E  Q  R  K  P  L  R  V  A  M  E  A  M  L  R  G  F  E  I
```

Figure 4.13 Translated DNA sequence of pGEX-4T-3-co. The DNA sequence was translated by the Expasy Translation Tool (http://www.expasy.org/tools/dna.html).

4.1.1.2 Eukaryotic vector

-pPICZαBNH8/DEST

The PCR technique was also used to screen the recombinant plasmid of st-pPIC and co-pPIC with the Start_f and Stop_r and Costart_f and Stop_r primers, respectively. The positive result showed the expected band of pPIC-st and pPIC-co at about 1.9 and 1.7 kb. (data not show). DNA sequence analysis with the 5'AOX and 3'AOX primers was used to identify *st-SFR2* and *co-SFR2* in pPICZαBNH8/DEST (Invitrogen). The results indicated that *st-SFR2* and *co-SFR2* were inserted in the correct reading frame in pPICZαBNH8/DEST, which contains the *S. cerevisiae* α-factor prepropeptide and an eight histidine tag protein at the N-terminal end of the SFR2 protein (Fig. 4.14 and 4.15).

α-factor

```
atgagatttccttcaattttactgctgttttattcgcagcatcctccgcattagctgct
 M  R  F  P  S  I  F  T  A  V  L  F  A  A  S  S  A  L  A  A
ccagtcaacactacaacagaagatgaaacggcacaaattccggctgaagctgtcatcggt
 P  V  N  T  T  T  E  D  E  T  A  Q  I  P  A  E  A  V  I  G
tactcagatttagaaggggatttcgatgttgctgttttgccattttccaacagcacaaat
 Y  S  D  L  E  G  D  F  D  V  A  V  L  P  F  S  N  S  T  N
aacgggttattgtttataaatactactattgccagcattgctgctaaagaagaaggggta
 N  G  L  L  F  I  N  T  T  I  A  S  I  A  A  K  E  E  G  V

                                          His-tag
tctctcgagaaaagagaggctgaagctgcacatcaccatcaccatcatcaccatgctgca
 S  L  E  K  R  E  A  E  A  A  H  H  H  H  H  H  H  H  A  A

                    attB1                                 St-SFR2
gtacatcaaacaagtttgtacaaaaaagcaggctccgcggccgccccttcaccatgcca
 V  H  Q  T  S  L  Y  K  K  A  G  S  A  A  A  P  F  T  M  P

ctaccggcgttcgtggcggcggcggcgaggctcgccgtcctcgtggccgcggcggcgacg
 L  P  A  F  V  A  A  A  A  R  L  A  V  L  V  A  A  A  A  T
```

Figure 4.14 Translated DNA sequence of pPIC-st. The DNA sequence was translated by the Expasy Translation Tool (http://www.expasy.org/tools/dna.html).

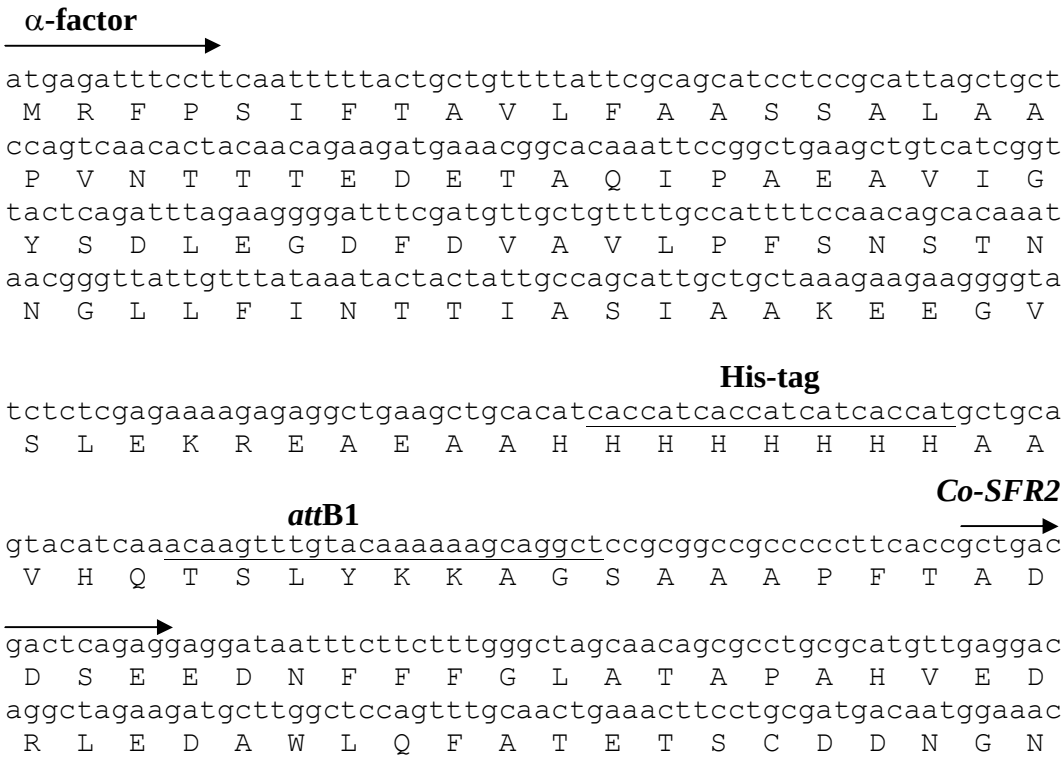


Figure 4.15 Translated DNA sequence of pPIC-co. The DNA sequence was translated by the Expasy Translation Tool (http://www.expasy.org/tools/dna.html).

4.1.2 Bioinformatics

Public database was searched for protein sequences homologous with the conceptual translation product of rice SFR2. We found that other plants species such as maize, soybean, cherry tomato and loblolly pine also each contained an *SFR2* ortholog. The SFR2 from rice was most closely related to the maize ortholog with 82% identity. Whereas for Arabidopsis (Thorlby, 2004) only 59% identity was seen (table 4.1). These enzymes were members of glycosyl hydrolase family 1 (Henrissat, 1991). The characteristic of glycosyl hydrolase family 1 motif was detected in rice SFR2 protein at residues 300 to 304 and 463 to 468. (Fig. 4.16). Moreover, Blastp

program (NCBI) showed that rice SFR2 and other SFR2 orthologs were more closely related to enzyme glycosyl hydrolase family 1 from thermophilic bacteria than other family 1 β-glucosidases in plants (table 4.1). The result was similar with Opassiri *et al.* that rice SFR2 was divergent from other plants β-glucosidases.

```
oryza           --MPLPAFVAAAARLAVLVAAAATAANAASYARYRRRHLRRIPSPIDESADPLADFRAFP 58
Zea             --MPLPAFLAAAARLAVLLAAAVTAANAASFARYRRRHLRRIPNPIDESADPVADFRALR 58
Glycine         --MTVVGLFLTATKLAGALFTLTVAANAFSYSRFRKKNLRGFRSPIDESQDTLAHFNV-- 56
Solanum         --MSVIALFTAATKLAGVLVTITVAANAFSFSVYRKKNLKRFRSPIDDSADVLAHFNLNP 58
Arabidopsis     -----MELFALLIKVAGLLATVTVGANVVSYSRFRRQNLAKFRSPIDESKEVLADFNSIE 55
Pinus           MGIHFVPILLNASKLVGVLATISVTANTIAFIRFRKKFLEPLDNPLDESEEVLATFDTDK 60

oryza           SSDADDSEED-NFFFGLATAPAHVEDRLEDAWLQFATETSCDDNGNVRDQ--RPVDALMA 115
Zea             PASAEDAAEDGSFFFGLATAPAHVEDRLEDAWLQFAVEHSCDDKEAMRDP--TTADAVMA 116
Glycine         AEGED------GFFFGLATAPAHVEDRLDDAWIQFAEKKSGEGGGDSEGK--QRVDAVIG 108
Solanum         SEGEK------GFFFGLATAPAHVEDRLDDAWLQFAKNTES---HEIQQP--QTADAIMG 107
Arabidopsis     HEEGK-------FFFGLATAPAHAEDDLDDAWLQFAKETPCSAEEAEAAD--KKA----- 101
Pinus           NEDDG------GFFFGLATAPAHVEDNLHDAWLEFAENDTADSNPQPSDQPPK------- 107

oryza           SAAGDGGSQQS-WRSTGGEN--IGDREQR-KPLRVAMEAMLRGFEILAESGES----AGG 167
Zea             SAAGDGGAQLASCRSSRGDDDRAGDGELRTKPLKIAMEAMLRGFEMFAEGGESGSAAAAG 176
Glycine         SASGDGGSQPAITSPR-----------SRKPLKVAMEAMIRGIEKYMEVEGK----ERE 152
Solanum         SATGDGGSQQALLPQREATK-----TIKRKKSLKIAIEAQIRGFEKYIEVEEL----TPT 158
Arabidopsis     ----------------------------RRKKVKLAVGAITKGLAKNTHGKEDKN--AAD 131
Pinus           ITRSKGGIQEPTLQYT----------VTKKGKKHFGPKWLWRPLPEVSKDSLK----KKK 153

oryza           DNCSHNVAAWHNVPCPQERLRFWSDPDAELKLAKETGISVFRMGVDWARLMPEEPTEELK 227
Zea             DSCSHNVAAWHNVPCPQERLKFWSDPDTELKLAKETGISVFRMGIDWTRVMPKEPIEDLK 236
Glycine         EEARPNVTAWHNVPCPEERLRFWSDPETEIKLAKDTGVTVFRMGIDWTRIMPVEPVSSLN 212
Solanum         EQCPHNVAAWHNVPHPEERLRFWSDPDIELKLAKNTGVQVFRMGVDWSRIMPEEPLGGLK 218
Arabidopsis     KPPSKNVAAWHNAPHAEDRLKFWSDPDKEVKLAKDTGVTVFRMGVDWSRIMPVEPTKGIK 191
Pinus           KDECHNVAAWKNAYHPEERLRFWSDPDTELRLAKDTGVTVFRMGVDWTRIMPVEPIDGIP 213

oryza           SSVNFAALERYRWIIQRVREYGMKVMLTLFHHSLPPWAGKYGGWKMEKTVTYFMDFVRLV 287
Zea             SSVNFAALERYRWIVQRVRDYGMKVMLTLFHHSLPPWAGEYGGWRMEKTVKYFLDFVRLV 296
Glycine         QSVNYAALERYKWIINRVRSYGMKVMLTLFHHSLPPWAGEYGGWKLEKTVDYFMDFTRLV 272
Solanum         ETVNFAALERYKWIINRVRSYGMKVMLTLFHHSLPPWAGEYGGWKLEKTVDYFMEFTRLI 278
Arabidopsis     EAVNYEAVEHYKWILKKVRSNGMKVMLTLFHHSLPPWAADYGGWKMEKTVDYFMDFTRIV 251
Pinus           NSVNQAALEHYRWIIERVHAYGMRVMLTLFHHSLPPWAAAYGGWKVEKTVNYFLEFTKIA 273

                         *****
oryza           VDCVSNLVDYWVIFNEPHVFVMLTYCAGAWPGGDPNAIEVATSTLPTGVYNQALHWMAIA 347
Zea             VDRVSDLVDYWVVFNEPHVFVMLTYCAGAWPGGDPNAIEVATSALPTGVYNQALHWMAIA 356
Glycine         VDSVSDLVDYWVTFNEPHVFCMLTYCAGAWPGGHPDMLEAATSALPTGVFQQAMHWMSIA 332
Solanum         VDSVADIVDYWVTFNEPHVFCMLTYCAGAWPGGNPDMLEVATSALPTGVFNQTMNWIAIA 338
Arabidopsis     VDSMYDLVDSWVTFNEPHIFTMLTYMCGSWPGNNPDFLEIATSTLPMGVFHRALHWMAVA 311
Pinus           VENFAQLVDYWVTFNEPHVFTMLTYCAGAWPGGHPDLLEVATAAMPQGVFNHVMHWMAVA 333
```

Figure 4.16 Clustalw alignment of *SFR2* orthologs from *Oryza sativa*, *Zea may*, *Glycine max*, *Solanum lycopersicum*, *Pinus taeda* and *Arabidopsis thaliana*. The family 1 glycosyl hydrolase motif (TFNEP, I/VTENG) is shown with ∗ above.

```
oryza           HSEAYDYIHSKSKNE-RKPIVGVAHHVSFTRPYGLFDVAAVALANSLTLFPYVDSICDKL 406
Zea             HAEAYDYIHSKGESK-RKPVVGVAHHVSFTRPYGLFDVAAVTLANSLTLFPYIDSICDKL 415
Glycine         HSKAYDYIHGLSNP--LNSIVGVAHHVSFMRPYGLFDIAAVSLANSLTLFPYIDEISEKL 390
Solanum         HTKAYDYIHEKSKP--ASAIVGVAHHVSFMRPYGLFDVAAVSVANSMTLFPFLDCISDKM 396
Arabidopsis     HSKAYDYIHG-KISL-KKPLVGVAHHVSFMRPYGLFDIGAVTISNSLTIFPYIDSICEKL 369
Pinus           HSKAFDLIHEFSKNSSLNARVGISHHVSFMRPYGLFDVPGVVLSNSMTLFPYIDSISEKL 393

                                                                    ****
oryza           DFIGINYYGQEVISGPGLKLVDNDEYSESGRGVYPDGLFRILIQFNERYKRLNIPFVITE 466
Zea             DFIGINYYGQEVISGPGLKLVDNDEYSESGRGVYPDGLFRILIRFNERYKSLNIPFIVTE 475
Glycine         DYIGINYYGQEVVSGAGLKLVENDEYSESGRGVYPDGLYRMLLQYHERYKHLNIPFIITE 450
Solanum         DYIGINYYGQEVICGAGLKLVETDEYSESGRGVYPDGLFRVLLQFDERYKHLNLPFIITE 456
Arabidopsis     DFIGINYYGQEAVCGAGLKLVETDEYSESGRGVYPDGLYRVLLMFHERYKHLKVPFIVTE 429
Pinus           DFLGLNYYGQEVLSAPGLKLVYNDEYSESGRAVYPDGLYRMLLKFHERYKHLNLPFIITE 453

                **
oryza           NGVSDETDLIRKPYILEHLLATYAAIIMGVRVLGYLFWTTSDNWEWADGYGPKFGLVAVD 526
Zea             NGVSDETDLIRKPYILEHLLAIYAAILMGVPVLGYLFWTTSDNWEWADGYGPKFGLVAVD 535
Glycine         NGVSDETDLIRRPYLLEHLLAIYAAMIMGVRVLGYLFWTISDNWEWADGYGPKFGLVAVD 510
Solanum         NGVSDGTDLIRQPYLLEHLLATYAAMMMGVRVLGYLFWTISDNWEWADGYGPKFGLVAVD 516
Arabidopsis     NGVSDETDVIRRPYLIEHLLALYAAMLKGVPVLGYIFWTISDNWEWADGYGPKFGLVAVD 489
Pinus           NGVSDATDLIRRPYILEHLLAVRAAMNKGVQVLGYLFWTTSDNWEWADGYGPKFGLVAVD 513

oryza           RANNLARKPRPSYFLFSRVVTTGKITRQDRMSAWRELQQAAVQKKTRPFFRAVDKHGRMY 586
Zea             RANNLARKPRPSYYLFSKIVTTGKITRHDRLAAWRELQQAAFQKKTRPFFRAVDKHGRMY 595
Glycine         RANNLARTPRPSYHLFSKIVNTGKVTHEDRERAWDELQRVAKEKKTRPFYRAVDKHRLMY 570
Solanum         RANDLARIPRPSYNLFSKVAESGKITREDREQVWGELQTAAKEGKRRPFYRSVNKYGLMY 576
Arabidopsis     RSHDLARTLRQSYHLFSKIVKSGKVTRKDRSLAWNELQKAAKAGKLRPFYRGVDNHNLMY 549
Pinus           RANNLARIPRPSYFLFSEVVKTGKVTRQQREIAWNELQIAAAEGKTRPFYRAVNKLGLMY 573

oryza           AGGLDRFIQRPFILRDWRFGHYKMEGLQDPLSCFIRCIFAP---FSRQKK-----IHYIE 638
Zea             AGGLDRPIQRSFILRDWRFGHYEMEGLQDPFSRFITSIIST---ILWKKKK----IRYIE 648
Glycine         AGGLDKPEQRPYIERDWRFGHYQMDGLQDPLSRFSRSIFRPFSLFSLKRKP----KSQKK 626
Solanum         AGGLDEPIWRPYIKRDWRFGHYEMEGLQDPLSRLARYLLHP---LSFKQKA----QTQRE 629
Arabidopsis     ADGLDKPQWRPFVDRDWRFGHYQMDGLQDPLSRVARTLLIWP--LIMKKRIRKVKIKHTD 607
Pinus           SGGLDIPIQRPLIQRDWRFNHYQFDGLKDPLSCTIRFVFAVV--FFWKRGGNGKLKKSST 631

oryza           DDVISYS-IN------------------------- 647
Zea             DEDISYS-LSC------------------------ 658
Glycine         NAKLILQPLET------------------------ 637
Solanum         SDQLTLEPLSANI----------------------- 642
Arabidopsis     DAGLVLHPALASPFD--------------------- 622
Pinus           EGYLDHNYSSITTKDKFITSIGAEVNEENLQPAAV 666
```

Figure 4.16 Clustalw alignment of *SFR2* orthologs from *Oryza sativa*, *Zea may*, *Glycine max*, *Solanum lycopersicum*, *Pinus taeda* and *Arabidopsis thaliana*. The family 1 glycosyl hydrolase motif (TFNEP, I/VTENG) is shown with ∗ above. (Continue)

Table 4.1 Blastp results SFR2 protein search against protein database (NCBI).

Accession number	Source	% Identity
CAJ87638	*Zea mays* (*SFR2*)	82
CAJ87636	*Glycine max* (*SFR2*)	69
CAJ87637	*Solanum lycopersicum* (*SFR2*)	67
CAJ87639	*Pinus taeda* (*SFR2*)	61
NP_566285	*Arabidopsis thaliana* (*SFR2*)	59
ZP_01188186	*Halothermothrix orenii* H 168	30
ZP_01462797	*Stigmatella aurantiaca* DW4/3-1	30
ZP_01533349	*Roseiflexus castenholzii* DSM 13941	30
YP_184240	*Thermococcus kodakarensis* KOD1	29
NP_577802	*Pyrococcus furiosus* DSM 3638	29
YP_001041390	*Caldivirga maquilingensis* IC-167	29
NP_142340	*Roseiflexus sp.* RS-1	29
CAA94187	*Fervidobacterium nodosum* Rt17-B1	29
ZP_01710738	*Staphylothermus marinus* F1	28
ZP_01358522	*Pyrococcus horikoshii* OT3	28
ZP_01523917	*Thermococcus sp.*	28

4.2 Protein Expression

4.2.1 Protein expression in *E. coli*

4.2.1.1 Survey for Appropriated strains assay

The pET32a/DEST plasmid was used to express *st-SFR2* and *co-SFR2* as N-terminal thioredoxin fusion proteins. The proteins were induced with 1 mM IPTG at 20 °C in *E. coli* strains BL21 (DE3), Origami (DE3), Origami B (DE3), Rosetta (DE3), Rosetta-gami (DE3) for 16 hours. The cells were collected and broken by sonication in breaking buffer containing lysozyme. The cell debris was collected by centrifugation and the supernatant was called the soluble fraction whereas pellets were called insoluble fraction. The soluble fractions were analyzed by *p*NP-β-D-glucoside hydrolysis. No hydrolysis of *p*NP-β-D-glucoside was detected. The SDS-PAGE technique was used to visualize the proteins. Fig. 4.17 shows the results when Rosetta (DE3) were used as host strain.

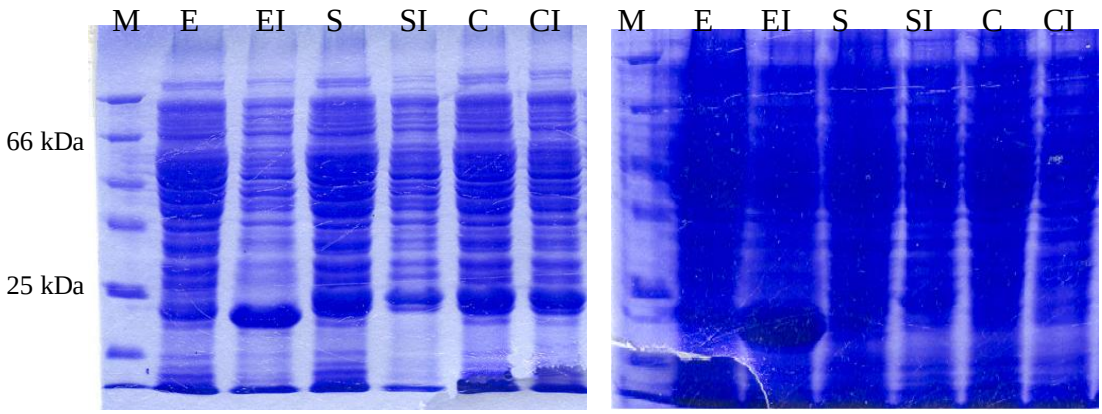


Figure 4.17 SDS-PAGE analysis of soluble protein (A) and insoluble protein (B) of rice SFR2 expression in Rosetta (DE3). M: protein marker, E: empty vector (control), S: st-SFR2 protein, C: co-SFR2 protein and I: induced culture

The protein profiles in SDS-PAGE of controls and samples were similar, except that an intense band of 20 kDa was detected in induced control cells in both soluble and insoluble fractions. This intense band represented the expressed thioredoxin and fusion partners in this system. The intensity of the 24 kDa bands were a bit higher in st-SFR2 and co-SFR2 protein expression samples induced and non-induced cells. These results suggested that maybe the proteins were expressed from the N-terminal thioredoxin fusion protein to a small region of SFR2 in *E. coli*. The pET32a system contains His-tag at the N-terminal of fusion protein. Therefore, the purification with Ni-column was investigated.

Purification of soluble proteins in induce culture was done with Ni-NTA superflow resin (Qiagen) according to the manufacturer's instructions.

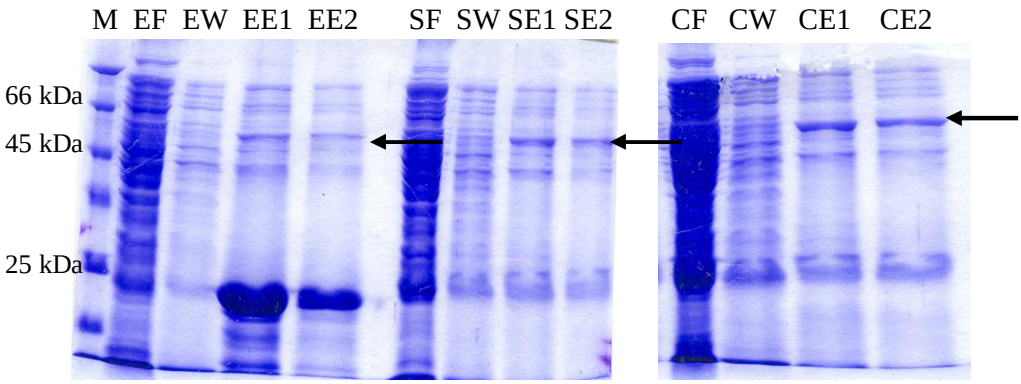


Figure 4.18 SDS-PAGE analysis of purified recombinant pET32a, st-SFR2 and co-SFR2 proteins from Rosetta (DE3). M: protein marker, E: empty vector (control), S: st-SFR2 protein in induced culture, C: co-SFR2 protein in induce culture, F: flow through, W: wash, E1: first of elution, E2: second of elution. The arrow lines show purified protein.

The intense band of thioredoxin was found in elution fraction of the control. However, in the two SFR2 strains, 24 kDa proteins were also detected in the flow through and elution fractions (Fig. 4.18). The result supported that the protein band of 24 kDa in samples were not N-terminal His-tag thioredoxin protein and SFR2 region. The arrows show proteins purified by Ni-column which were detected only in elution fractions. These bands were host protein that were detected in both control and samples.

This result was similar in all *E. coli* strains used. Other strains of *E. coli* that is better support production of toxic protein, BL21 (DE3) pLysS (Novagen) was further investigated. The soluble proteins at 3 hours of induction and overnight induction were analyzed by *p*NP-β-D-glucoside hydrolysis. No activity was detected. The SDS-PAGE technique was used to visualize proteins (Fig. 4.19).

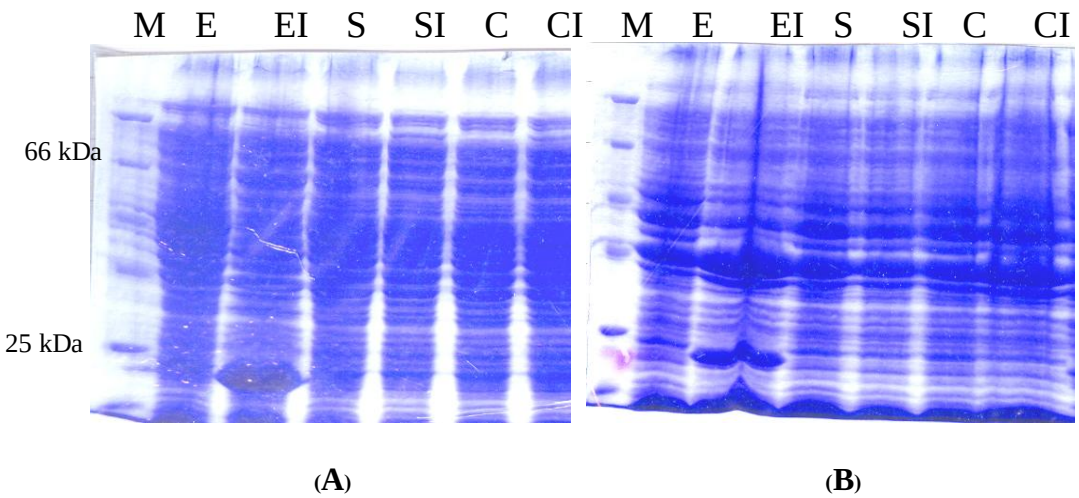


Figure 4.19 SDS-PAGE analysis of soluble protein (A) and insoluble protein (B) of rice SFR2 expression in BL21 (DE3) pLysS. M: protein marker, E: empty vector (control), S: st-SFR2 protein, C: co-SFR2 protein and I: induced culture.

The SDS-PAGE analysis of BL21 (DE3) pLysS showed negative results similar to other strains. These indicated that all strains of *E. coli* used are not appropriate for SFR2 protein production with pET32a system.

4.2.1.2 Screening for appropriate vectors

Other vector systems were used to express SFR2 protein. Conventional ligation techniques was use to ligate *co-SFR2* in pSY5 and pGEX-4T-3, which contains a glutathio-s-tranferase (GST) fusion protein.

The pSY5-co was transformed into Rosetta-gami (DE3) and selected clones were expressed in LB broth containing antibiotics. The co-SFR2 protein production was induced with 1 mM IPTG at 20 °C for 16 hours. The cells were broken by sonication and cell debris was collected by centrifugation. The supernatant was called the soluble fraction, whereas the pellets were called the insoluble fraction. The soluble fractions were analyzed by *p*NP-β-D-glucoside hydrolysis. No activity was detected. The SDS-PAGE technique was used to visualized protein bands (Fig. 4.20).

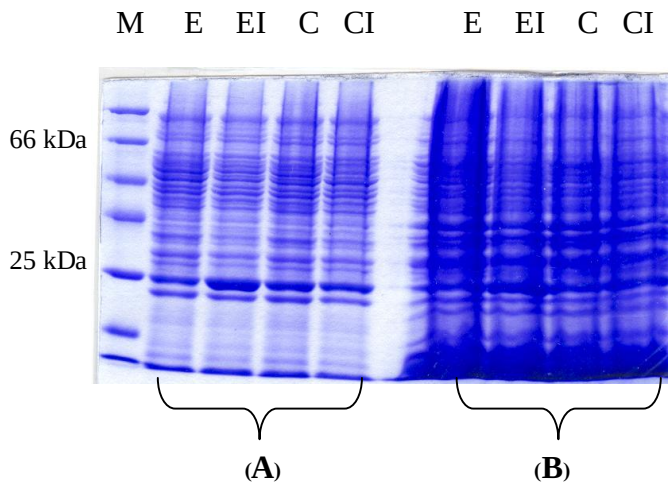


Figure 4.20 SDS-PAGE analysis of soluble protein (A) and insoluble protein (B) of rice SFR2 expression in pSY5 system. M: protein marker, E: empty vector (control), C: co-SFR2 protein and I: induced culture.

The SDS-PAGE protein profiles of the pSY5 system expression with control and sample vectors were similar. No intense band was detected in the pSY5 control due to the lake of a fusion partner in the pSY5 vector. No intense band was detected in co-pSY5 construct indicated that no recombinant SFR2 protein was produced in the pSY5 system.

The pGEX-4T-3-co was used to transformed into DH5α and selected clones were expressed in LB broth containing antibiotics. The co-SFR2 protein was induced with 1 mM IPTG at 20 °C for 16 hours. The cells were collected and broken by sonication in breaking buffer containing lysozyme. The proteins were collected by centrifugation, and the supernatant was called the soluble fraction, whereas the pellets were called the insoluble fraction. The soluble fractions were analyzed by *p*NP-β-D-

glucoside hydrolysis. The activity also was not detected similar to the pSY5 system. The SDS-PAGE technique was used to visualize protein bands (Fig. 4.21).

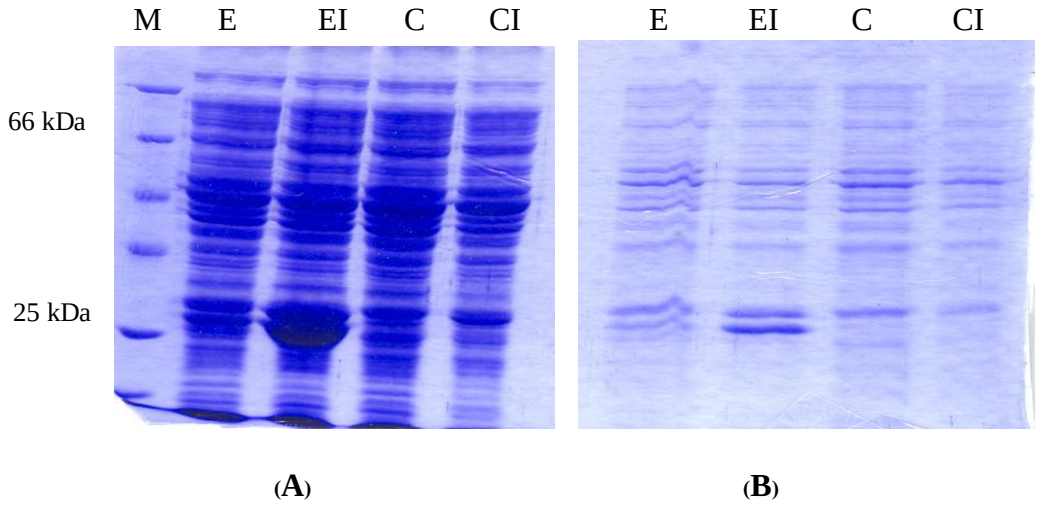


Figure 4.21 SDS-PAGE analysis of soluble protein (A) and cell suspension, contained media and cells (B) of rice SFR2 expression in pGEX-4T-3 system. M: protein marker, E: empty vector (control), C: co-SFR2 protein and I: induced culture.

The SDS-PAGE protein profiles of the control and sample were similar except for the intense 26 kDa band of glutathione-S-transferase in the soluble and insoluble fractions (Fig. 4.21). This indicated that no recombinant SFR2 protein was produced in the pGEX-4T-3 system.

The result concluded that all three vector systems used cannot produce SFR2 protein after 16 hours of induction.

4.2.1.3 Induction time assay

According to the last results, the SFR2 protein was not detect in *E. coli* after 16 hours of induction. More detailed investigation was checked at time before and after induction.

The pET-st, pET-co and pSY5-co were transformed into Rosetta-gami (DE3) and selected clones were expressed in LB broth containing antibiotics. The cell suspension that contained media and cells were collected at 1, 2 and 3 hours before induction and every 30 minutes for 3 hours after induction. The cells suspension were then mixed with dye, boiled and loaded in SDS-PAGE. Fig. 4.22 show the result from the pET-st system.

The protein concentration increased whereas the band patterns were similar before and after induction. This indicated that this *E. coli* system cannot produce SFR2 protein prior to or a short time after induction.

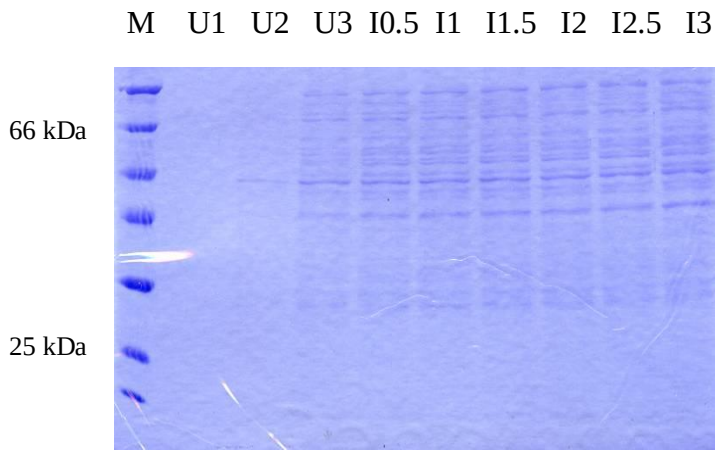


Figure 4.22 SDS-PAGE analysis of pET-st cell suspension in induction time assay. M: protein marker, U: uninduction phase, I: induction phase and number show point of time (hour).

4.2.1.4 Test of induction temperature

Since, *SFR2* was reported to be responsible for freezing resistance in *Arabidopsis* (Thorlby, 2004) and table 4.1 indicated that SFR2 is more closely related to the Thermophilic β-glucosidase, additional induction temperatures of 4, 37 and 45 °C were tested.

The selected clones from pET-st and pET-co in Rosetta-gami (DE3) were used to express at three different temperatures of induction. The soluble fractions of all conditions were analyzed by *p*NP-β-D-glucoside hydrolysis. No activity was detected from any of the conditions. The SDS-PAGE technique was used to visualize protein bands from all three conditions (Fig. 4.23, 4.24 and 4.25).

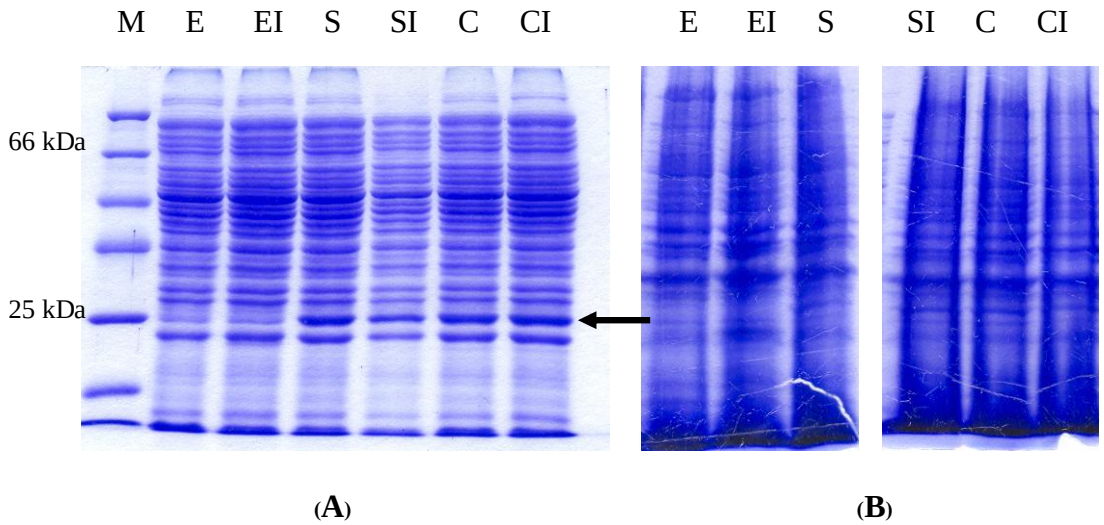
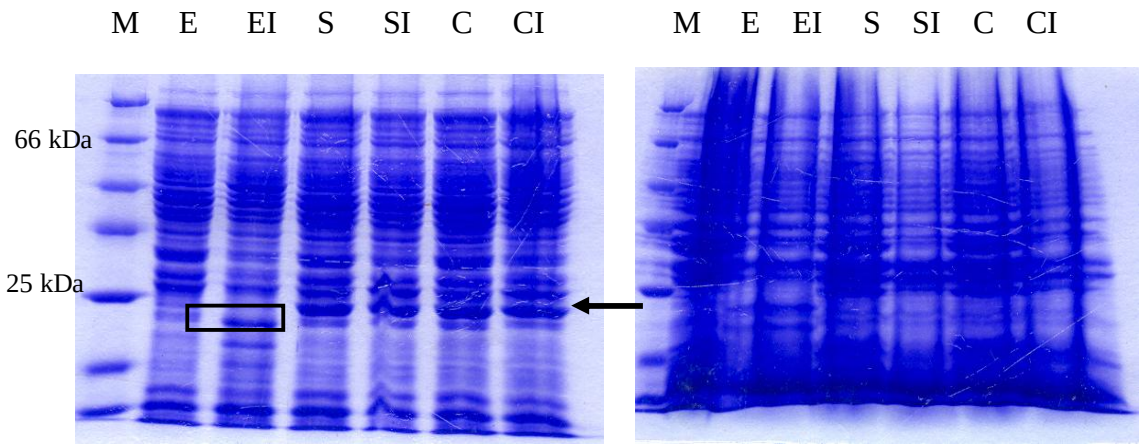
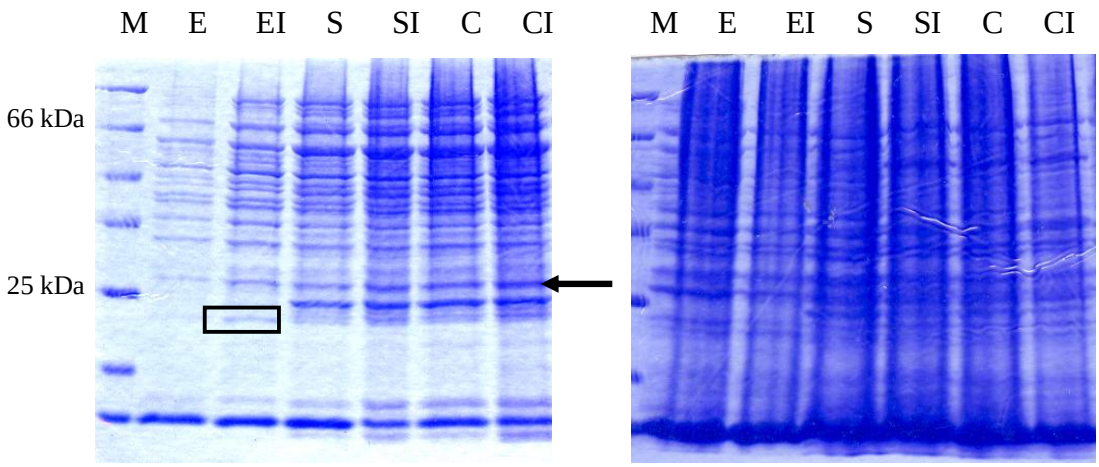


Figure 4.23 SDS-PAGE analysis of soluble protein (A) and insoluble protein (B) from pET32 expression in Roseseta-gami (DE3) at 4 °C induction temperature. M: protein marker, E: control, S: st-SFR2 protein, C: co-SFR2 protein and I: induced culture. The arrow line shows protein bands were found only in samples.



(A) **(B)**

Figure 4.24 SDS-PAGE analysis of soluble protein (A) and insoluble protein (B) from pET32 expression in Rosetta-gami (DE3) at 37 °C induction temperature. M: protein marker, E: control, S: st-SFR2 protein, C: co-SFR2 protein and I: induced culture. The arrow line shows protein bands were found only in samples. The box indicates thioredoxin and fusion tag.



(A) **(B)**

Figure 4.25 SDS-PAGE analysis of soluble protein (A) and insoluble protein (B) from pET32 expression in Rosetta-gami (DE3) at 45 °C induction temperature. M: protein

marker, E: control, S: st-SFR2 protein, C: co-SFR2 protein and I: induced culture. The arrow line shows protein bands were found only in samples. The box indicates thioredoxin and fusion tag.

The protein profiles at each induction temperature were nearly the same. The 20 kDa band of thioredoxin and fusion tags was detected in control induced culture at 37 °C and 45 °C induction temperatures whereas at 4 °C it was not detected. This suggested that the temperature may affect expression in *E. coli*. The arrows (Fig. 4.23 – 4.25) point 25 kDa protein bands which were bigger than the fusion protein in control and found only in sample. This band might be the N-terminal His-tag and a small region of SFR2. The Ni-column (Qiagen) purification technique was used to investigate the 25 kDa band. The result found this band was in the flow through and elution fractions (data not show) similar to result of Fig. 4.19. Therefore, the 25 kDa of protein band is likely not the N-terminal His-tag and small region of SFR2. It might be the host protein which expressed from pET32a/DEST system. The results indicted that the *SFR2* gene was not induced at higher and lower temperatures, since no bands at 70 kDa were detected.

In conclusion, our results suggested that *E. coli* cannot express SFR2 in all conditions tested. Two hypotheses were used to answer this problem. First, RNA is usually thought of as a single stranded linear molecule, however, in a biological system this is not the case. Frequently, different regions of the same RNA strand will fold together via base pair interactions to make intricate secondary and tertiary structures that are essential for correct biological function. Common secondary structure motifs include hairpin loops, stems, and bulges have been found in RNA

molecules (Bialy, 1987). Unpredicted secondary structure in the mRNA, might also be the case for SFR2, blocking expression in *E. coli*.

Second, bacteria have a complicated immune system that helps them recognize and isolate foreign DNA trying to invade their cell membrane. The H-NS protein, like human body immune system, recognizes foreign DNA and prevents it from becoming active. These are a problem for the biotech industry which used bacterial cell for recombinant protein production. The H-NS may be purified out from the bacteria before their use as hosts in recombinant protein industry (Science Express, 2006).

4.2.2 Protein expression in *P. pastoris*

To investigate the expression of SFR2 in a eukaryotic system, wildtype Y11430 *P. pastoris* was selected as the host strain. The pPIC-st and pPIC-co were transformed to *P. pastoris*. The selected clones were grown in BMGY media for 48 hours and methanol induction was done for 72 hours. The trizol extraction method was used to extract RNA from cell pellets. The cDNA were amplify with 5' AOX and 3'AOX and then nested PCR were done with internal primers, Fai8_f and Stop_r (Fig. 4.26).

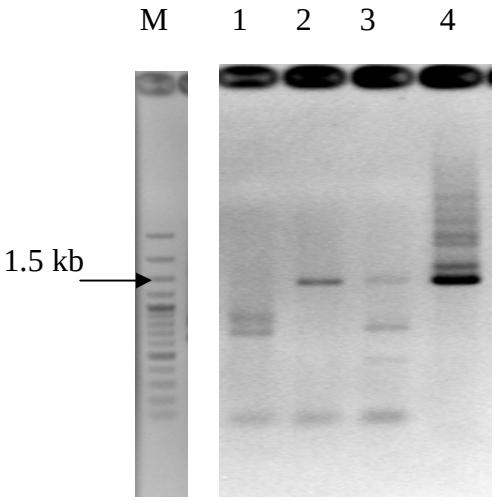


Figure 4.26 Agarose gel electrophoresis of nested PCR products encoding recombinant proteins in pPICZαB system. Lane M: 100 bp marker, Lane 1: cDNA of pPICZ∞BNH8 clone (negative control) , Lane 2: cDNA of pPIC-st clone, Lane 3: cDNA of pPIC-co clone, Lane 4: positive control, plasmid pPIC-co.

The RT-PCR products (Fig 4.26) of st-pPIC (lane 2) and co-pPIC (lane 3) show intense bands at 1.4 kb, the same position as the positive control (lane 4), whereas the empty vector (lane 1) was not detected. The intensity of 1.4 kb PCR product in pPIC-co was less than pPIC-st and the positive control. The error might be that amount of RNA templates were not equal in the RT-PCR step. Based on this result, it was concluded that the selected clones contained RNA of SFR2.

The media from the positive clones that contained RNA of *SFR2* were concentration by ultrafiltration to about 60 fold. The activity assay was assayed with several substrates, *p*NP-β-D-glucoside, *p*NP-β-D-galactoside, *p*NP-β-D-fucoside, *o*NP-β-D-glucoside and *p*NP-β-D-xyloside. The samples and control show no significant differences in activity values. The control that contained empty vector

showed activity with some substrates because high backgrounds were obtained from intracellular β-glucosiase in *P. pastoris* (Turan, 2005). The protein from pPIC-st and pPIC-co should contain a His-tag at the N-terminus. Therefore, the crude enzymes were purified by Ni-NTA superflow resin (Qiagen) and the protein were analyzed with SDS-PAGE (Fig. 4.27).

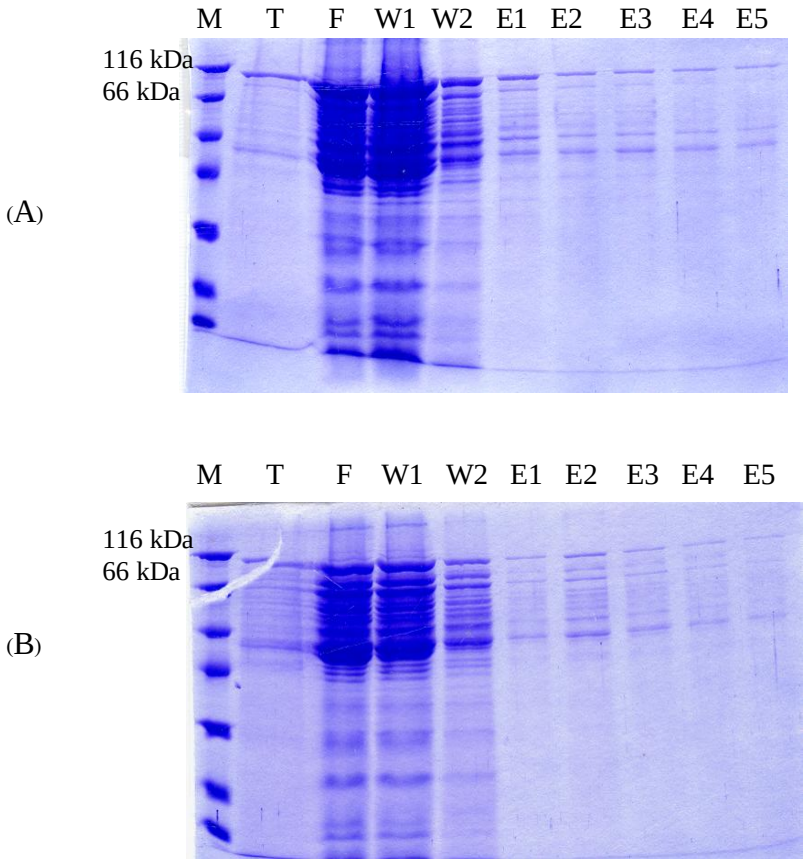


Figure 4.27 SDS-PAGE analysis of recombinant pPIC-st (A) and pPIC-co (B) purified with a Ni-column. M: protein marker, T: total cell contained media and cell, F: flow through, W: wash, E: elution.

The protein profiles of wash and elution fractions were similar and no major protein product was detected in the elution fractions. This indicated that no

recombinant proteins were purified by Ni-NTA. But the results could not verify whether SFR2 was produced from *P. pastoris* or not. Since the secondary or tertiary structure of the protein and subunit interaction can give negative results with Ni-NTA chromatography (Hang, 2004). The denatured proteins were future investigate by western blot analysis. The anti-6His antibody was used to detect the 8His-SFR2 protein by western blot. No intense band was detected (data not show). These result may come from two circumstances, the His-tag may have been cleaved from the SFR2 protein by the host cell proteinase or the translation process may have been blocked.

In summary, when *P. pastoris* was used as host for SFR2 expression. The RNA of *SFR2* was detected, but no detectable SFR2 was found by activity nor protein assay, SDS-PAGE or western blot.

4.3 Protein localization

4.3.1 Plasmids construction

The SFR2B1_f and SFR2B2_r primers were used to amplify *SFR2* from the AK119461 cDNA clone. The PCR product was loaded onto a 1% agarose gel and the 1.9 kb band of the *SFR2* cDNA was excised from the gel. The gel purification kit (Qiagen) was used to purify the PCR product from the gel. The product was checked again on 1% agarose gel (Fig. 4.28).

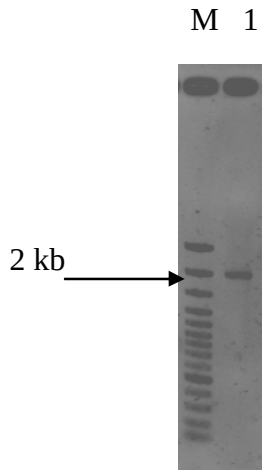


Figure 4.28 Agarose gel electrophoresis of purified PCR product amplification with SFR2B1_f and SFR2B2_r primers. Lane M: 100 bp marker, 1: *SFR2*.

The *SFR2* was transfered to pDONR/Zeo by BP clonase reaction (Invitrogen). The reaction was then transformed to TOP10 competent cell. The colony PCR technique was used to select positive clones (data not show). Recombinant plasmid from selected clones were extracted and sequenced.

```
AK119461        AAGAAAATCACGGAAATGCCACTACCGGCGTTCGTGGCGGCGGCGGCGAGGCTCGCCGTC 120
sfr2-pDONR      ---------------ATGCCACTACCGGCGTTCGTGGCGGCGGCGGCGAGGCTCGCCGTC 45
                              ****************************************

AK119461        CTCGTGGCCGCGGCGGCGACGGCGGCCAATGCCGCCTCCTACGCGCGGTACCGGCGGCGC 180
sfr2-pDONR      CTCGTGGCCGCGGCGGCGACGGCGGCCAATGCCGCCTCCTACGCGCGGTACCGGCGGCGC 105
                ************************************************************

AK119461        CACCTCCGCCGCATCCCCAGCCCCATCGACGAGTCCGCCGATCCCCTCGCCGACTTCCGC 240
sfr2-pDONR      CACCTCCGCCGCATCCCCAGCCCCATCGACGAGTCCGCCGATCCCCTCGCCGACTTCCGC 165
                ************************************************************

AK119461        GCCTTTCCCTCCTCCGACGCTGACGACTCAGAGGAGGATAATTTCTTCTTTGGGCTAGCA 300
sfr2-pDONR      GCCTTTCCCTCCTCCGACGCTGACGACTCAGAGGAGGATAATTTCTTCTTTGGGCTAGCA 225
                ************************************************************

AK119461        ACAGCGCCTGCGCATGTTGAGGACAGGCTAGAAGATGCTTGGCTCCAGTTTGCAACTGAA 360
sfr2-pDONR      ACAGCGCCTGCGCATGTTGAGGACAGGCTAGAAGATGCTTGGCTCCAGTTTGCAACTGAA 285
                ************************************************************

AK119461        ACTTCCTGCGATGACAATGGAAACGTGCGCGACCAGAGGCCAGTAGATGCACTGATGGCA 420
sfr2-pDONR      ACTTCCTGCGATGACAATGGAAACGTGCGCGACCAGAGGCCAGTAGATGCACTGATGGCA 345
                ************************************************************

AK119461        TCTGCTGCTGGTGATGGAGGCTCCCAGCAATCTTGGAGGTCAACAGGTGGGGAAAATATT 480
sfr2-pDONR      TCTGCTGCTGGTGATGGAGGCTCCCAGCAATCTTGGAGGTCAACAGGTGGGGAAAATATT 405
                ************************************************************
```

Figure 4.29 Alignment of the *SFR2* region in pDONR/Zeo with AK119461.

```
AK119461        GGTGATAGAGAGCAGAGGAAACCACTTAGGGTAGCCATGGAGGCTATGCTCAGGGGGTTT 540
sfr2-pDONR      GGTGATAGAGAGCAGAGGAAACCACTTAGGGTAGCCATGGAGGCTATGCTCAGGGGGTTT 465
                ************************************************************

AK119461        GAAATTCTTGCTGAGAGTGGAGAATCTGCTGGCGGCGATAACTGCAGCCACAATGTTGCT 600
sfr2-pDONR      GAAATTCTTGCTGAGAGTGGAGAATCTGCTGGCGGCGATAACTGCAGCCACAATGTTGCT 525
                ************************************************************

AK119461        GCTTGGCATAATGTTCCATGCCCGCAAGAAAGGCTTAGATTTTGGTCTGATCCTGATGCT 660
sfr2-pDONR      GCTTGGCATAATGTTCCATGCCCGCAAGAAAGGCTTAGATTTTGGTCTGATCCTGATGCT 585
                ************************************************************

AK119461        GAGTTGAAACTTGCTAAGGAGACCGGCATCAGCGTTTTCCGCATGGGGGTAGATTGGGCA 720
sfr2-pDONR      GAGTTGAAACTTGCTAAGGAGACCGGCATCAGCGTTTTCCGCATGGGGGTAGATTGGGCA 645
                ************************************************************

AK119461        AGGTTAATGCCTGAGGAACCAACCGAAGAATTGAAGAGCTCAGTTAATTTTGCAGCACTT 780
sfr2-pDONR      AGGTTAATGCCTGAGGAACCAACCGAAGAATTGAAGAGCTCAGTTAATTTTGCAGCACTT 705
                ************************************************************

AK119461        GAGCGGTATAGATGGATCATTCAAAGGGTTCGCGAATATGGAATGAAAGTAATGCTTACT 840
sfr2-pDONR      GAGCGGTATAGATGGATCATTCAAAGGGTTCGCGAATATGGAATGAAAGTAATGCTTACT 765
                ************************************************************

AK119461        CTGTTTCATCACTCACTTCCACCCTGGGCTGGAAAATATGGCGGGTGGAAGATGGAAAAA 900
sfr2-pDONR      CTGTTTCATCACTCACTTCCACCCTGGGCTGGAAAATATGGCGGGTGGAAGATGGAAAAA 825
                ************************************************************

AK119461        ACTGTCACCTATTTCATGGATTTTGTGAGGCTTGTTGTTGATTGTGTATCCAATTTAGTG 960
sfr2-pDONR      ACTGTCACCTATTTCATGGATTTTGTGAGGCTTGTTGTTGATTGTGTATCCAATTTAGTG 885
                ************************************************************

AK119461        GACTACTGGGTGATTTTCAATGAACCTCACGTTTTTGTGATGCTAACTTATTGTGCCGGT 1020
sfr2-pDONR      GACTACTGGGTGATTTTCAATGAACCTCACGTTTTTGTGATGCTAACTTATTGTGCCGGT 945
                ************************************************************

AK119461        GCTTGGCCTGGTGGAGACCCTAATGCAATTGAAGTAGCAACATCTACTCTGCCAACTGGT 1080
sfr2-pDONR      GCTTGGCCTGGTGGAGACCCTAATGCAATTGAAGTAGCAACATCTACTCTGCCAACTGGT 1005
                ************************************************************

AK119461        GTATACAATCAGGCTTTGCATTGGATGGCTATTGCACATTCTGAAGCCTATGACTACATA 1140
sfr2-pDONR      GTATACAATCAGGCTTTGCATTGGATGGCTATTGCACATTCTGAAGCCTATGACTACATA 1065
                ************************************************************

AK119461        CATTCGAAAAGCAAGAACGAAAGGAAGCCAATAGTTGGTGTTGCACACCATGTATCGTTT 1200
sfr2-pDONR      CATTCGAAAAGCAAGAACGAAAGGAAGCCAATAGTTGGTGTTGCACACCATGTATCGTTT 1125
                ************************************************************

AK119461        ACAAGGCCATATGGGCTATTTGATGTTGCTGCTGTCGCGCTAGCTAACTCATTGACCCTT 1260
sfr2-pDONR      ACAAGGCCATATGGGCTATTTGATGTTGCTGCTGTCGCGCTAGCTAACTCATTGACCCTT 1185
                ************************************************************

AK119461        TTTCCTTACGTGGATAGCATATGTGATAAATTGGACTTCATTGGAATCAACTACTATGGA 1320
sfr2-pDONR      TTTCCTTACGTGGATAGCATATGTGATAAATTGGACTTCATTGGAATCAACTACTATGGA 1245
                ************************************************************

AK119461        CAGGAGGTTATATCAGGACCTGGTCTAAAGCTTGTGGATAATGATGAGTATAGTGAATCT 1380
sfr2-pDONR      CAGGAGGTTATATCAGGACCTGGTCTAAAGCTTGTGGATAATGATGAGTATAGTGAATCT 1305
                ************************************************************

AK119461        GGTCGTGGGGTTTATCCTGATGGGCTGTTCCGCATCCTGATTCAATTCAATGAACGATAT 1440
sfr2-pDONR      GGTCGTGGGGTTTATCCTGATGGGCTGTTCCGCATCCTGATTCAATTCAATGAACGATAT 1365
                ************************************************************

AK119461        AAGAGATTAAATATACCTTTTGTCATTACTGAAAATGGAGTTTCTGATGAGACTGATCTG 1500
sfr2-pDONR      AAGAGATTAAATATACCTTTTGTCATTACTGAAAATGGAGTTTCTGATGAGACTGATCTG 1425
                ************************************************************
```

Figure 4.29 Alignment of the *SFR2* region in pDONR/Zeo with AK119461.

(Continue)

```
AK119461        ATACGGAAACCATATATACTGGAACACTTGTTAGCCACATACGCTGCCATCATTATGGGT 1560
sfr2-pDONR      ATACGGAAACCATATATACTGGAACACTTGTTAGCCACATACGCTGCCATCATTATGGGT 1485
                ************************************************************

AK119461        GTCCGTGTACTTGGTTATTTGTTTTGGACAACATCAGATAATTGGGAATGGGCGGATGGC 1620
sfr2-pDONR      GTCCGTGTACTTGGTTATTTGTTTTGGACAACATCAGATAATTGGGAATGGGCGGATGGC 1545
                ************************************************************

AK119461        TATGGTCCCAAGTTTGGGCTTGTTGCTGTTGACCGTGCTAATAACCTAGCACGGAAACCT 1680
sfr2-pDONR      TATGGTCCCAAGTTTGGGCTTGTTGCTGTTGACCGTGCTAATAACCTAGCACGGAAACCT 1605
                ************************************************************

AK119461        CGGCCTTCATACTTTTTATTCTCCAGGGTTGTTACAACTGGAAAAATTACAAGACAGGAC 1740
sfr2-pDONR      CGGCCTTCATACTTTTTATTCTCCAGGGTTGTTACAACTGGAAAAATTACAAGACAGGAC 1665
                ************************************************************

AK119461        AGAATGTCTGCTTGGAGGGAGCTGCAACAAGCTGCAGTTCAAAAGAAAACACGTCCATTT 1800
sfr2-pDONR      AGAATGTCTGCTTGGAGGGAGCTGCAACAAGCTGCAGTTCAAAAGAAAACACGTCCATTT 1725
                ************************************************************

AK119461        TTCAGGGCAGTGGATAAGCATGGTCGGATGTATGCAGGTGGTCTAGATCGGCCTATTCAG 1860
sfr2-pDONR      TTCAGGGCAGTGGATAAGCATGGTCGGATGTATGCAGGTGGTCTAGATCGGCCTATTCAG 1785
                ************************************************************

AK119461        AGGCCCTTCATATTGCGGGATTGGAGATTTGGTCACTATAAAATGGAAGGCTTGCAGGAT 1920
sfr2-pDONR      AGGCCCTTCATATTGCGGGATTGGAGATTTGGTCACTATAAAATGGAAGGCTTGCAGGAT 1845
                ************************************************************

AK119461        CCTTTGAGTTGCTTTATAAGATGCATTTTTGCACCATTTTCACGCCAAAAGAAGATTCAC 1980
sfr2-pDONR      CCTTTGAGTTGCTTTATAAGATGCATTTTTGCACCATTTTCACGCCAAAAGAAGATTCAC 1905
                ************************************************************

AK119461        TACATAGAAGATGATGTTATTTCTTATTCTATTAACTGATGCCCTGCTACTATAGTTGCT 2040
sfr2-pDONR      TACATAGAAGATGATGTTATTTCTTATTCTATTAAC---------------------- 1941
                ***********************************
```

Figure 4.29 Alignment of the *SFR2* region in pDONR/Zeo with AK119461. (Continue)

Sequencing of the *SFR2* in pDONR/Zeo indicated that all sequences were correct.

The LR clonase (Invitrogen) was used to transfer *SFR2* from pDONR/Zeo to the pMDC43, pMDC83 and pMDC140 vectors. The colony PCR technique was used to select positive clones (data not show). Some selected clones were used for plasmid extraction and sequencing. The SFR2 in pMDC43, pMDC83 and pMDC140 were designated as pSM43, pSM83 and pSM140, respectively.

DNA sequence analysis was performed to check that the reading frame of the GFP and SFR2 fusion was correct in pMDC43 with the SM43-GFP_f primer. The

result found that *SFR2* was inserted in the correct reading frame in pMDC43, which contains GFP protein and *att*B1 at the N-terminal of the SFR2 protein (Fig. 4.30).

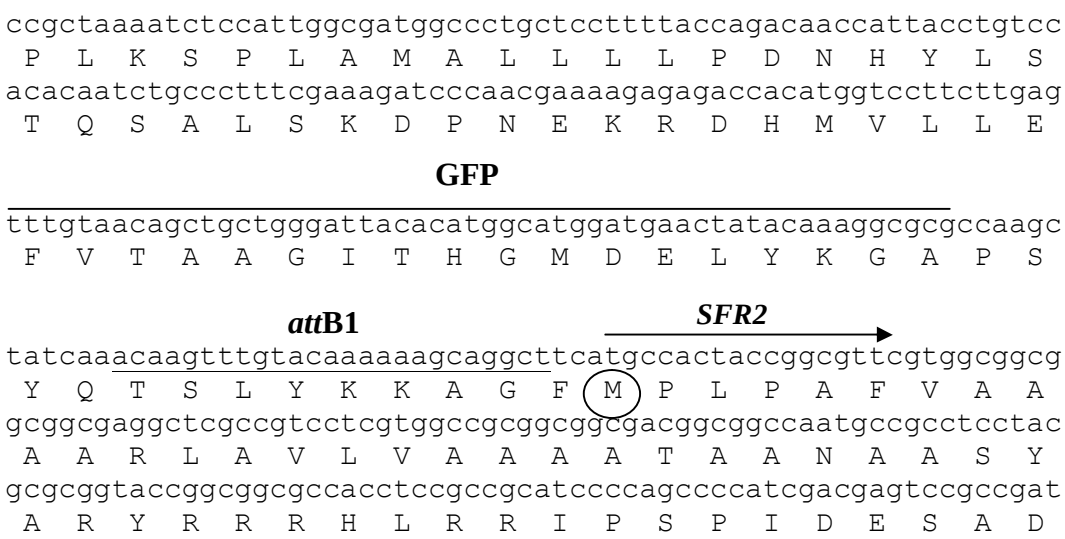


Figure 4.30 Translated DNA sequence of pSM43 (GFP-SFR2). The DNA sequence was translated by the Expasy Translation Tool (http://www.expasy.org/tools /dna. html).

DNA sequence analysis was used to check the fusion junction of SFR2 and GFP in pMDC83 with the SM83-GFP_r primer. *SFR2* was inserted in the correct reading frame in pMDC83 and the GFP protein followed the SFR2 and *att*B2 sequence in the correct frame. If translated, it will give rise to SFR2 fused to the N-terminus of the GFP protein (Fig. 4.31).

Figure 4.31 Translated DNA sequence of pSM83 (SFR2-GFP). The DNA sequence was translated by the Expasy Translation Tool (http://www.expasy.org/tools /dna. html).

DNA sequence analysis with SM140-GUS_r primer was used to confirm the fusion junction of SFR2 in pMDC140. *SFR2* was inserted in the correct reading frame in pMDC140 and the GUS protein followed the SFR2 and *att*B2 sequence in the correct frame. If translated, it will give rise to SFR2 fused to the N-terminus of the GUS protein (Fig. 4.32).

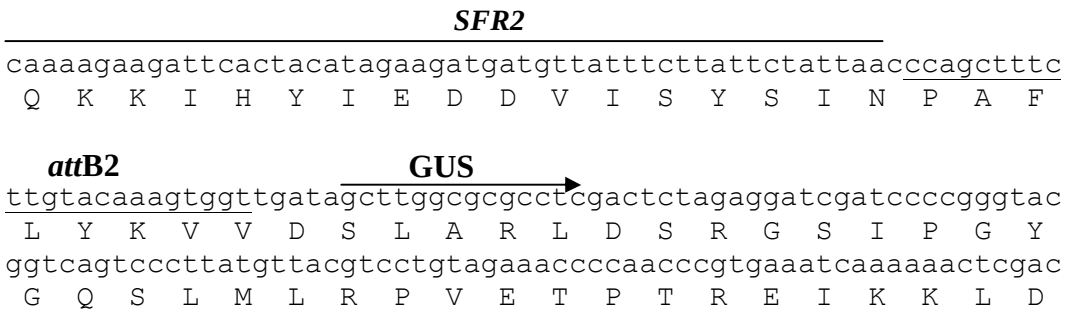


Figure 4.32 Translated DNA sequence of pSM140 (SFR2-GUS). The DNA sequence was translated by the Expasy Translation Tool (http://www.expasy.org/tools /dna. html).

4.3.2 Bombardment

The Gene gun machine (PDA – 1000/ He, Biorad) was used to transform plant cells with the plasmid constructs. The plasmids were coated with 0.7 μM tungsten particles (Biorad) before bombardment. All the plasmid constructs used in the localization assay are shown in table 4.2.

Table 4.2 Plasmid constructs used for protein localization assay.

Name	**Descriptions**	**Transformed tissue**
pMDC43	2x 35S – **GFP – *ccd*B** – *nos T*	Hydrilla cells
pSM43	2x 35S – **GFP – *SFR2*** – *nos T*	Hydrilla cells
pRM83	2x 35s – **AAM08614– GFP** – *nos T*	Hydrilla cells
pSM83	2x 35s – ***SFR2* – GFP** – *nos T*	Hydrilla cells
pMDC140	2x 35s –***ccd*B– GUS** – *nos T*	Hydrilla, onion cells
pSM140	2x 35s – ***SFR2* – GUS** – *nos T*	Hydrilla, onion cells
pRM140	2x 35s – **AAM08614– GUS** – *nos T*	Hydrilla, onion cells

The pSM43 and pSM83 that contained GFP reporter gene were used to transform hydrilla cells. The hydrilla leaves were put on 2% agar and bombarded at a pressure of 1,100 psi per shot at distance 6 cm. After bombardment, the transformed hydrilla leaves were incubated at 25 °C for 48 hours under continuous darkness conditions (Noji, 1998). Under light conditions, the expressed GFP chimeras rapidly undergo proteolytic degradation in the cell (Tamura, 2003). Transformed cells were visualized with a fluorescent microscope at filter wavelength 520 – 590 nm 40X objectives.

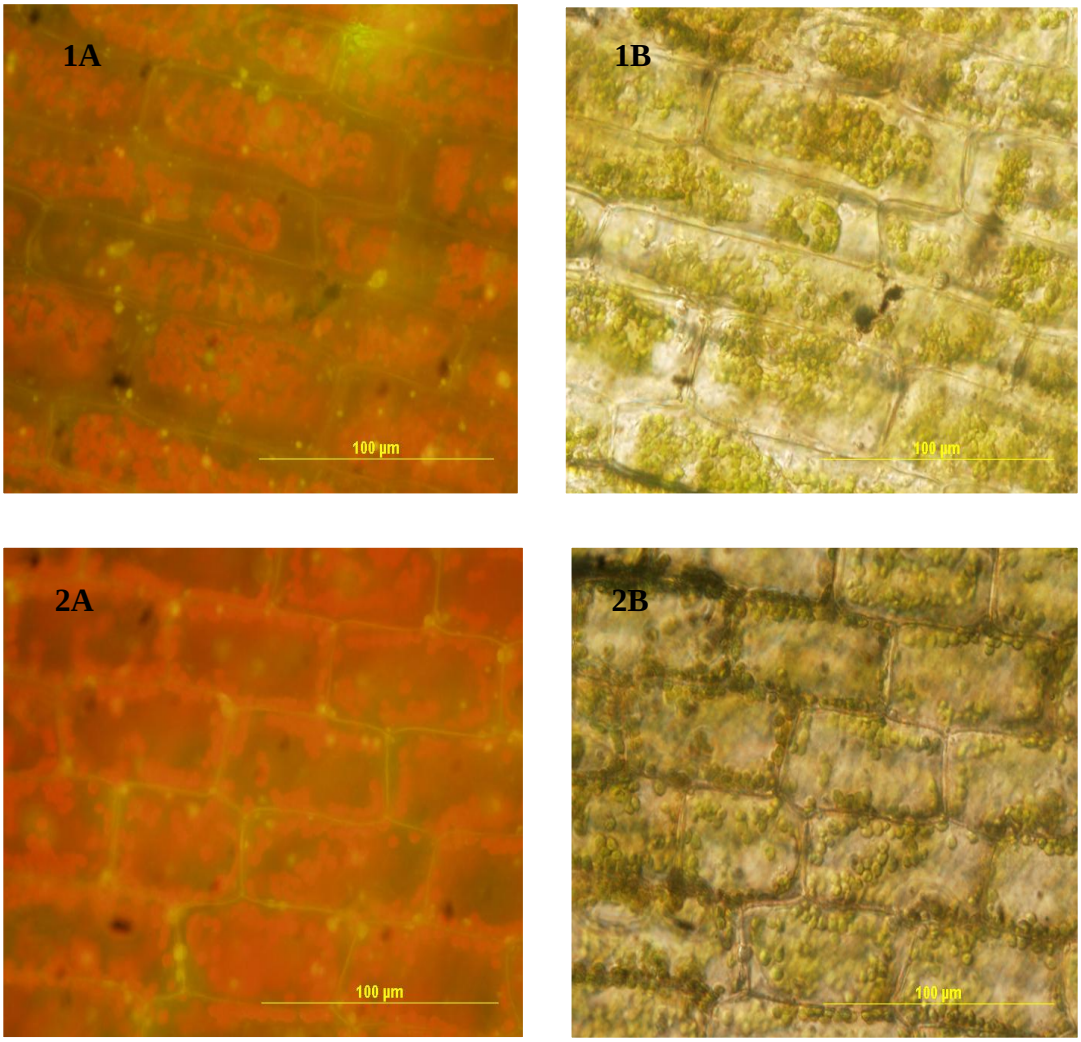


Figure 4.33 Transcient assay of pMDC43 (1) and pSM43 (2) in hydrilla cells.

A: Fluorescent microscope image, B: Light microscope image. The red fluorescent spots in (A) and green spots in (B) show the locations of the chloroplasts.

The fluorescent microscope images of pMDC43 (control) and pSM43 show similar results of GFP protein dispersed all over these cells. Both constructs used the signal peptide of the GFP protein.

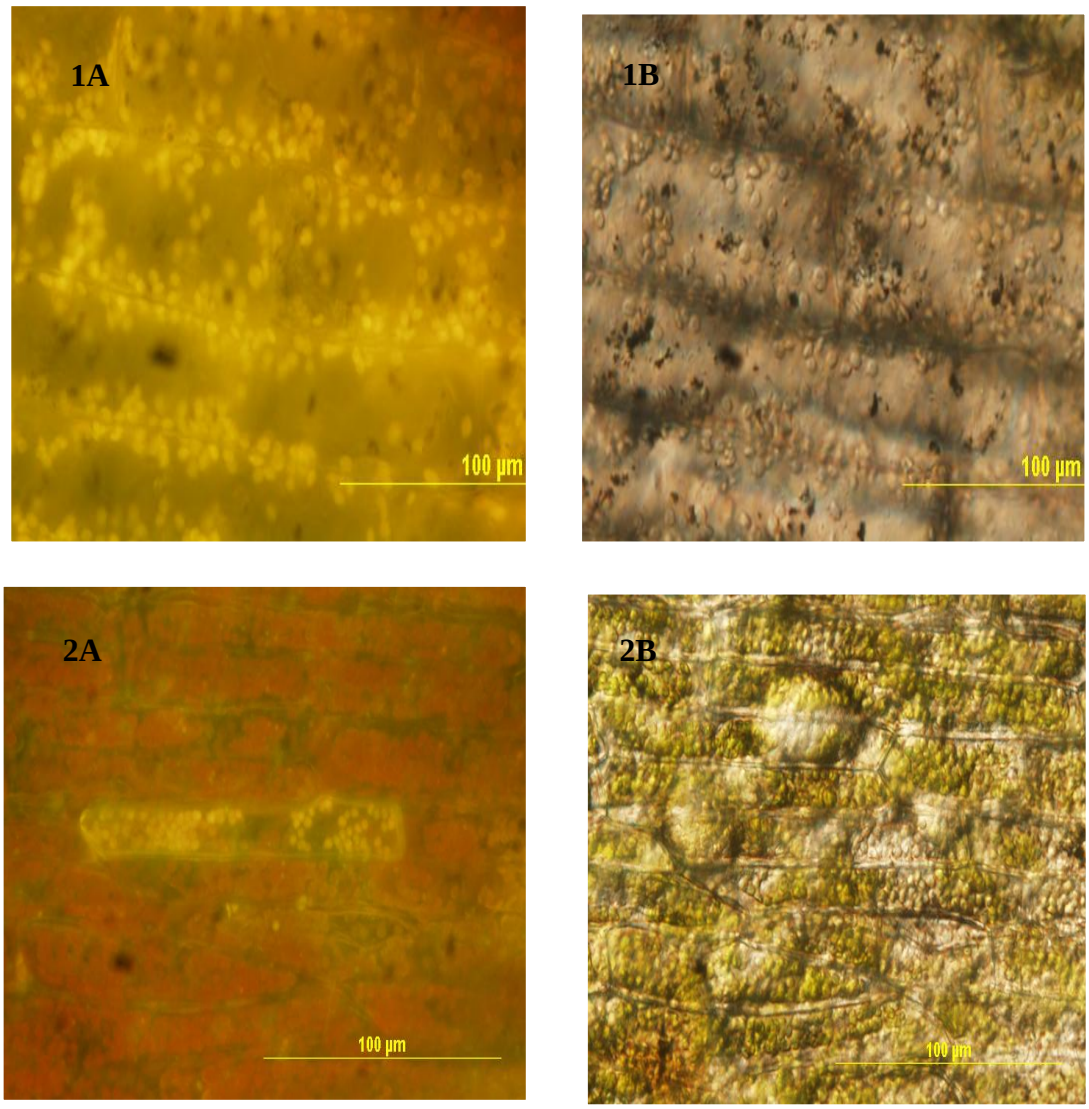


Figure 4.34 Transcient assay of pSM83 (1) and pRM83 (2) in hydrilla cells. A: Fluorescent microscope image, B: Light microscope image. The red fluorescent spots in (A) and green spots in (B) show the locations of chloroplasts.

The fluorescent microscope images of pSM83 shown GFP protein were in chloroplasts. Whereas images AAMO8614, RM83 which contained secreted signal to membrane proteins (PSORT-program and signalP) shown GFP protein all over the cell

except location of chloroplasts. The result suggested that AAM08614 protein (NCBI) were in vacuoles which contain this protein traveling to the membrane.

Thus, both constructs used signal sorting from insertion protein at N –terminal of GFP protein.

The pSM140 that contained GUS reporter gene was transformed to onion and hydrilla cells by biolistic transformation. The 1,100 psi rupture discs were employed at distance 6 cm. After bombardment, the transformed cells were incubated at 25 °C for 48 hours in the dark. The GUS assay reagents were used to stain cells and incubated overnight at 37 °C in dark conditions. Then, the cells were washed briefly with 70% ethanol and were inspected by light microscope (Fig. 4.35).

Based on the results of Fig. 4.35, the image of transformed onion with pMDC140 showed blue color distributed throughout the cells, including the nucleus, cell wall and cytoplasm, whereas onion transformed with pSM140 were different. In onion cells transformed with pSM140, blue color was found only in the cell, but the location was not clear. It was hard to detect location of chloroplasts in onion cells, therefore hydrilla cells which present clear chloroplasts's location were investigated. In hydrilla cells the results were similar that control show blue color dispersed all over the cells, but pSM140 shown blue color in only chloroplasts. The GUS assay found high background of concentrated blue color distributed all over the cells, therefore it is hard to identify the location of the blue GUS staining.

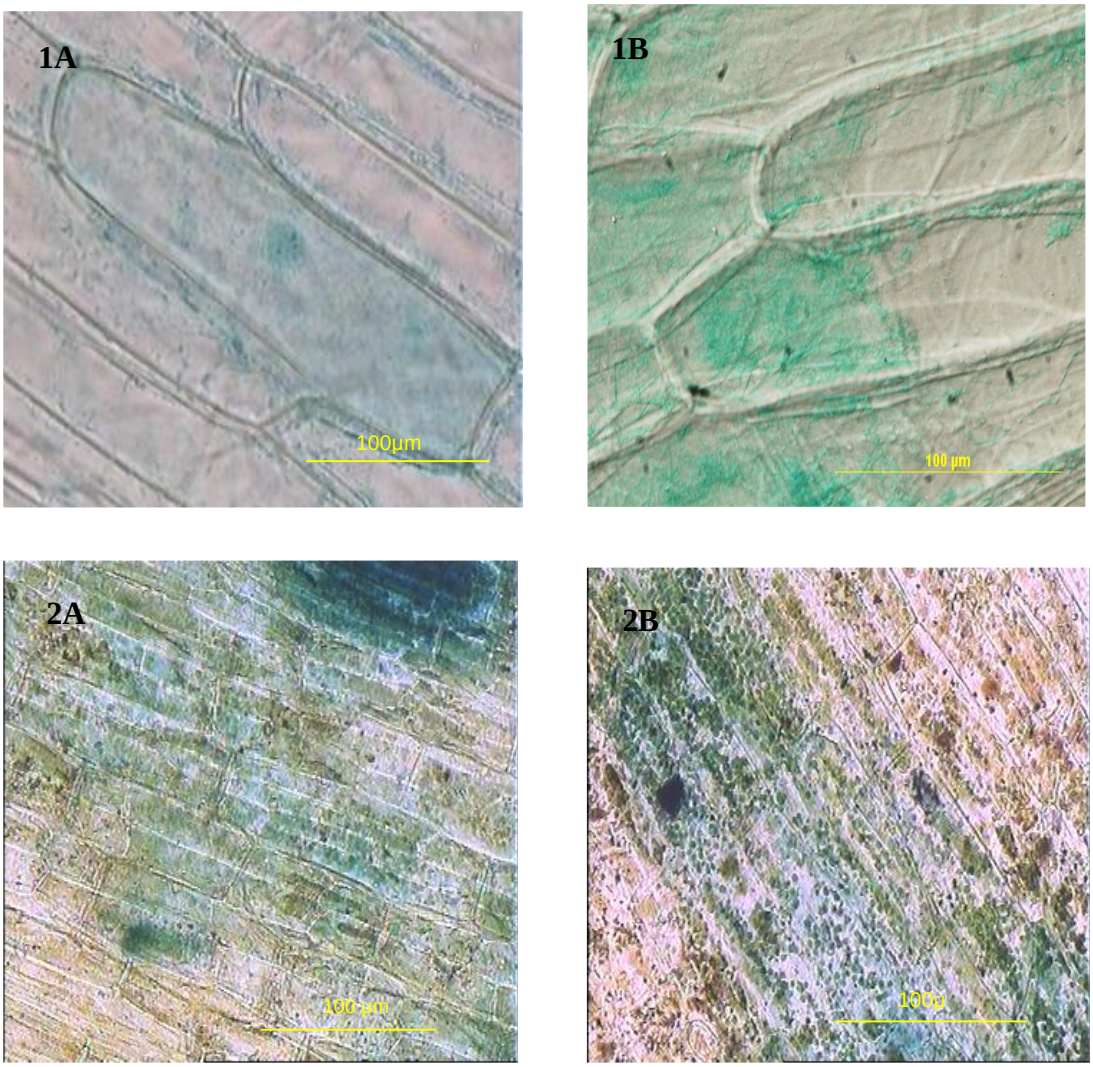


Figure 4.35 Transcient assay of pMDC140 (A) and pSM140 (B) in onion (1) and hydrilla (2) cells. The blue color shows positive results with the GUS assay.

In summary, the results of protein localization showed that *SFR2* gene could be expressed in plant cells under control of the 35S promoter. The SFR2 protein contained a transit peptide to chloroplast as predicted by PSORT and ChloroP programs. The result was similar to Thorlby (2004) *et al.* that arabidopsis SFR2 expression was only found in green tissues.

According to mutant arabidopsis *SFR2* show freezing damage in low temperature (Thorlby, 2004). And Northern blot analysis of rice *SFR2* with 3'UTR in stress conditions found that the transcript of *SFR2* was induced and maintained a similarly high level from 6 to 48 hours when treated at 5 °C whereas signal were not detected in 150 mM salt treatment (Fig. 4.36 and Opassiri, personal communication). The *SFR2* RNA was not detected in control plant that grown at 28 °C (Data not shown). Base on this result, the *SFR2* transcript level was upregulated in response to cold stress at 5 °C in rice seeding.

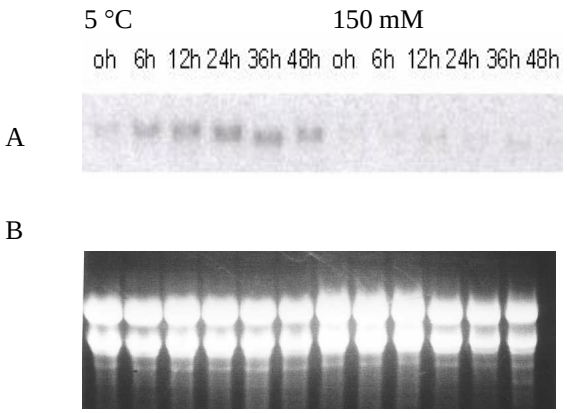


Figure 4.36 Northern bolt analysis of SFR2 transcript levels at 5 °C and 150 mM salt treatment (A), rRNA loading control (B) (Opassiri, personal communication).

The data supports and current result can used to help predict the SFR2 function in freezing stress.

Stabilization of membranes is one of mechanisms that help to protect cell damage in freezing condition. The membrane biogenesis in plants involves a massive movement of fatty acids or lipid from one organelle to another. The major component of chloroplast membrane is glycolipid. However, the mechanism of lipids movement between ER to chloroplast or chloroplast to chloroplast's membrane are not know

(Buchanan, 2000). The lipid transfer in glycoside form with SFR2 function might help to cleave glycolipid from ER to release lipid into chloroplast membrane or with in the cell.

Accumulated sugar is one of mechanism that occurs in the development of freezing tolerance. Although the sugars are proposed to protect the membrane, the accumulation of sugar alone is not sufficient for the development of freezing tolerance (Buchanan, 2000). The enzyme like SFR2 may activate the accumulation by conjugation of sugar to aglycones to help with the freezing tolerance. However, the role of the accumulated sugar promoting tolerance remain poorly understood (Buchnan, 2000).

Therefore, more work is needed on these hypotheses. They may be investigated by producing recombinant protein from plant system or extract crude enzyme from induced plant, can be used for analysis of possible substrates, which may offer the possibility for functional identification.

CHAPTER V

CONCLUSION

The rice *SFR2* was amplified from AK119461 (Rice Genome Resource Center). The 1.9 kb *st-SFR2* contained a transit peptide and conserved region of β-glucosidases, whereas 1.7 kb *co-SFR2* contained only conserved region of β-glucosidases but no transit peptide. The amino acid sequence of rice SFR2 was more closely related to thermophilic bacteria than any other β-glucosidases in other plants.

Two constructs of *SFR2* gene were cloned and expressed in *E. coli* strains, BL21 (DE3), Origami (DE3), Origami B (DE3), Rosetta (DE3), Rosetta-gami (DE3) and BL21 pLysS (DE3). Three vector systems, pET32a/DEST, pSY5 and pGEX-4T-3 were tested. Several induction time and induction temperatures (4, 20, 37 and 45 °C) were also investigated. However, no condition tested was able to produce SFR2 protein in *E. coli*.

The pPICZαBNH8/DEST system was used to express SFR2 protein in *P. pastoris* Y11430 wildtype strain. The RNA of *SFR2* was detected when amplified with specific primers but no activity nor protein were detected by nitrophenol β-glucosidase hydrolysis assay and SDS-PAGE.

The protein localization techniques were used to determine the localization of SFR2 protein in plant cells. The *SFR2* gene was fused with GFP and GUS reporter

gene at the N or C-terminal and transformed to hydrilla and onion cells by particle gun bombardment. The *SFR2*-GFP and *SFR2*-GUS constructs showed SFR2 protein in the chloroplasts.

REFERENCES

REFERENCES

Akiyama, T., Kaku, H., and Shibuya, N. (1997). A cell wall-bound β-glucosidase from germinated rice: purification and properties. **Phytochemistry.** 48: 49-54.

Amouri, B., and Gargouri, A. (2006). Characterization of a novel β-glucosidase from a *Stachybotrys* strain. **Biochemical Engineering Journal.** 32: 191-197.

Bialy, H. (1987). Recombinant proteins: virtual Authenticity. **Biotechnology.** 5: 883-90.

Buchnan, B. B., Gruissem, W., and Jones, R. L. (2000). **Biochemistry and molecular biology of plants.** American Socity of Plant Physiologists Rockville, Maryland.

Chen, G., Wilson, I. D., Kim, S. H., and Grierson, D. (2001). Inhibiting expression of a tomato ripening-associated membrane protein increase organic acid and reduced sugar levels of fruit, **Planta.** 212: 799-807.

Curtis, M. D., and Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. **Plant Physiology.** 133: 462-469.

Dietz, K-J., Sauter, A. Wichert, K., Messaghi, D., and Hartung, W. (2000). Extracellular β-glucosidase activity in barley involved in the hydrolysis of ABA glucose conjugate in leaves. **Journal of Experimental Botany.** 51: 937-944.

Duffy, R. (2001). **Speciality rices of the world breeding, production and marketing**. Food and Agricultural Organization of the United Nations Rome.

Escamilla-Trevino, L. L., Chen, W., Card, M. L., Shih, M. C., Cheng, C. L., and Poulton, J. E. (2006). *Arabidopsis thaliana* β-glucosidase BGLU45 and BGLU46 hydrolase monolignol glucosidase. **Phytochemistry.** 67: 165-1660.

Friedman, K. J. (1977). Role of lipids in the *Neurospora crassa* membrane and influence of fatty acid composition on membrane lipid phase transition. **Journal of Membrane Biology.** 32: 33-47.

Gu, Z., Wang, J., Huang, J., and Zhang, H. (2005). Cloning and characterization of a rice gene family encoding putative dual-specificity protein kinases, involved in responses to abiotic and biotic stresses. **Plant Science.** 169: 470-477.

Han, Y., Wang, X., Jiang, J., Xu, Y., Xu, Z., and Chong K. (2004). Biochemical character of the purified OsRAA1, a novel rice protein with GTP-binding activity, and its expression pattern in *Oryza sativa*. **Plant Biology.** 162: 1057-1063.

Hang, S-K., Salamone, P. R., Kavakli, H., Salttery, C. J., and Okita, T. W. (2004). Rapid purification of the potato ADP – glucose pyrophosphorylase by polyhistidine – mediated chromatography. **Protein Expression and Purification.** 38: 99-107.

Ilgen, C., Lin-Cereghino, J. and Cregg, J. M. (2005). *Pichia pastoris* in: production of recombinant proteins – novel microbial and eukaryotic expression systems (Gellissen, G., Ed.), pp. 143-162. Wiley-VCH, Weinheim.

Jiang, J. F., Han, Y., Xing, L. J., XU, Y. Y., XU, Z. H., and Chong, K. (2005). Cloning and expression of a novel cDNA encoding a mannose-specific jacalin-related lectin from *Oryza sativa*. **Toxicon.** 47: 133-139.

Jung, Y. H., Agrawal, G. K., Rakwal, R., Kim, J. A., Lee, M. O., Choi, P. G., Kim, Y. J., Kim, M. J., Shibato, J., Kim, S. H., Iwahashi, H., and Jwa, N. S. (2005). Functional characterization of *OSRacB* GTPase- a potentially negative regulator of basal disease resistance in rice. **Plant Physiology and Biochemistry.** 44: 68-77.

Kim, S. H., and Grierson, D. (2005). Subcellular localization and silencing of ripening-associated membrane protein (TRAMP) in tomato *Lycopersicon esculentum* Mill. **Plant Science.** 169: 1022-1029.

Lee, K. H., Piao, H. L., Kim, H. Y., Choi, S. M., Jiang, F., Hartung, W., Hwang, L., Kwak, J. M., Lee, I. J., and Hwang, I. (2006). Activation of glucosidase via stress-induce polymerization rapidly incresses active pools of abscisic acid. **Cell**. 126: 1109-1120.

Lyons, J. M., and Asmundson, C. M. (1965). Solidification of unsaturated/saturated fatty acid mixtures and its relationship to chilling sensitivity in plants. **Journal of the American Oil Chemists' Society.** 42: 1056-1058.

Martin, C., and Jones, J. (2004). Cold tolerance, *SFR2*, and the legacy of gary warren. **The Plant Cell**. 16: 1955-1957.

Noji, M., Inoue, K., Kimura, N., Gouda, A., and Saito, K. (1998). Isoform- dependent differences in feedback regulation and subcellular localization of serine acetyltransferase involved in cysteine biosynthesis from Arabidopsis thaliana. **Journal of Biological Chemistry.** 273: 32739-32745.

Opassiri, R., Ketudat Cairns, J. R., Akiyama, T., Wara-Aswapati, O., Svasti, J. and Esen A. (2003). Characterization of a rice β-glucosidase highly expressed in

flower and germinating shoot. **Plant Science**. 165: 627-638.

Opassiri, R., Pomthong, B., Onkoksoong, T., Akiyama, T., Esen, A., and Ketudat Cairns, J. R. (2006). Analysis of rice glycosyl hydrolase family 1 and expression of Os4bglu12 β-glucosidase. **BMC Plant Biology: 33** doi:10.1186/1471-2229-6-33.

Park, S-M., Kim, D-H., Truong, N. H., and Itoh, Y. (2002). Heterologous expression and characterization of class III chitinase from rice (*Oryza sativa L.*). **Enzyme and Microbial Technology.** 30: 697-702.

Peoples, T. R., Koch, D. W., and Smith, S. C. (1977). Relationship between chloroplast membrane fatty acid composition and photosynthetic response to a chilling temperature in four alfalfa cultivars. **Plant Physiol.** 61: 472-473

Science express. (2006). **Bacteria have their own immune system protecting against outside DNA.**

Shewry, P. R., and Gutterride, S. (1992). **Plant protein engineering**. Cambridge University Press 3-4.

Shimamura, C., Ohno, R., Nakamura, C., and Takumi, S. (2006). Improvement of freezing tolerance in tobacco plants expressing a cold-responsive and chloroplast-targeting protein WCOR15 of wheat. **Journal of Plant Physiology.** 163: 213-219.

Steitz, J.A. (1979). Genetic signals and nucleotide sequences in messenger RNA. **In Biological regulation and development 1: Gene expression.** Plenum press, New York.

Sun, A-Q., Yi, S-Y., Yang, J-Y., Zhao, C-M., and Liu, J. (2005). Identification and characterization of a heat-inducible *ftsH* gene from tomato (*Lycopersicon esculentum Mill.*). **Plant Science.** 170: 551-562.

Tamura, K., Shimada, T., Ono, E., Tanaka, Y., Nagatani, A., Higashi, S., Watanabe, M., Nishimura, M., and Hara-Nishimura, I. (2003). Why green fluorescent fusion proteins have not been observed in the vacuoles of higher plants. **Journal of Plant.** 35: 545- 555.

Thorlby, G., Fourrier N., and Warren, G. (2004). The *SENSITIVE TO FREEZING2* gene, required for freezing tolerance in *Arabidopsis thaliana*, encodes a β-glucosidase. **The Plant Cell.** 16: 2192-2203.

Wang, X. M., and Ma, Q. H. (2005). Characterization of a jasmonate-regulated wheat protein related to a beta-glucosidase-aggregating factor. **Plant Physiology and Biochemistry.** 43: 185-192.

APPENDIX

APPENDIX

1. Media

1.1 LB Medium (1 L)

Dissolve 10 g Tryptone, 5 g Yeast Extract and 5 g NaCl in 1,000 ml distilled water. Autoclave the solution at 121 °C for 20 min. Allow the medium to cool before adding antibiotics with concentration recommended. To prepare plates, add 1.5% agar before autoclaving.

1.2 LB Low salt medium with 100 µg/ml Zeocin (1 L)

Dissolve 10 g Tryptone, 5 g Yeast Extract and 2.5 g NaCl in 1,000 ml distilled water. Autoclave the solution at 121 °C for 20 min. Allow the medium to cool then, add 1 ml of 100 mg/ml Zeocin. To prepare plates, add 1.5% agar before autoclaving.

1.3 YPD medium with 100 µg/ml Zeocin (1 L)

Dissolve 20 g Peptone and 10 g Yeast Extract in 900 ml distilled water. Autoclave the solution at 121 °C for 20 min. Allow the medium to cool then, add 100 ml of 10x D-glucose sterile (0.2 g/ml) and 100 ml of 1 mg/ml Zeocin. To prepare plates, add 1.5% agar before autoclaving.

1.4 BMGY medium (1 L)

Dissolve 10 g Yeast Extract, 20 g Peptone and 5 g Glycerol in 200 mM potassium phosphate buffer pH7 and adjust volume to 1 L. Sterilize by autoclaving for 20 minutes.

1.5 GYT medium (100 ml) for *E. coli* completent cell preparation.

Dissolve 0.25 g Tpytone, 0.125 g Yeast Extract and 10 g glycerol and adjust the volume to 100 ml with distilled water. Sterilize by autoclaving for 20 minutes. Store at 4 °C.

1.6 MS medium for tissue culture

Mix 50 ml Stock I (Macro/MS), 5 ml Stock II(Micro/MS), 5 ml StockIII (Fe-EDTA), 1 ml Vitamin MS, 1 ml 2,4-D 2000 ppm, 1 ml myo-inosital and 30 g sucrose . Adjust pH to 5.8. The volume adjust to 1 L with distilled water. Add 1.5% agar and sterilize by autoclaving for 20 minutes.

2. Antibiotics

Table 1 The antibiotics used in this research.

Name	**Stock concentration**	**Working concentration**
Ampicillin	100 mg/ml in H_2O	100 μg/ml
Kanamycin	50 mg/ml in H_2O	50 μg/ml
Chloramphenicol	25 mg/ml in ethanol	25 μg/ml
Tetracycline	10 mg/ml in ethanol	10 μg/ml
Zeocin (Invitrogen)	100 mg/ml in H_2O	100 μg/ml

Filter sterile all antibiotic solution (except chlorampenical and Tetracyclin) and keep at - 20°.

3. Reagents for *E. coli* breaking buffer

3.1 100 mM Tris-HCl pH 8.0 buffer (100 ml)

Dissolve 1.2 g Tris base in 80 ml distilled water. Adjust pH to 8.0 with HCl add distilled water up to 100 ml.

3.2 10 mg/ml Lyzozyme stock solution (10 ml)

Dissolve 10 mg lyzozyme in 10 ml distilled water. Keep at -20°C.

3.3 5 mM PMSF stock solution (100 ml)

Dissolve 0.871 g PMSF in 100 ml isopropanol. Keep at -20°C.

4. Reagents for *P. pastoris* breaking buffer

4.1 1 M sodium phosphate buffer pH 7.5 (100 ml)

Mix 77.4 ml of 1 M Na_2HPO_4 and 22.6 ml of 1 M NaH_2PO_4.

4.2 5 mM EDTA pH 8.0 stock solution (100 ml)

Dissolve 0.186 g EDTA (disodium ethylene tetraacetrate $\cdot 2H_2O$) in 70 ml distilled water. Adjust pH to 8.0 with NaOH and add distilled water up to 100 ml.

4.3 10 % glycerol stock solution (100 ml)

Dissolve 10 g glycerol in distilled water and adjust the volume to 100 ml.

4.4 5 mM PMSF stock solution (100 ml)

Follow by 7.3.

5. Reagents for activity assay

5.1 50 mM NaOAc pH 5.0 buffer (100 ml)

Dissolve 0.41 g NaOAc in 80 ml distilled water and adjust pH to 5.0 with glacial acetic acid and add distilled water up to 100 ml.

5.2 0.4 M Na_2CO_3 (10 ml)

Dissolve 0.42 g Na_2CO_3 in distilled water and adjust the volume to 100 ml.

6. Reagents for SDS-PAGE

6.1 SDS-gel loading buffer (5X stcck, 8 ml)

Dissolve 0.30 g Tris Base, 1 g SDS, 0.05 g Bromophenol blue, 5 ml glycerol and adjust pH to 6.8 with HCL and the volume to 8 ml with distilled water. Before used add 20 µl of 2-mercapthoethanol to 80 µl of solution mixture. Store at room temperature.

6.2 1.5 M Tris pH 8.8 (100 ml)

Dissolve 18.17 g Tris Base in 80 ml distilled water. Adjust pH to 8.0 with HCl and adjust the volume to 100 ml with distilled water. Store 4 °C.

6.3 0.5 M Tris pH 6.8 (100 ml)

Dissolve 6.06 g Tris Base in 80 ml distilled water. Adjust pH to 6.8 with HCl and adjust the volume to 100 ml with distilled water. Store 4 °C.

6.4 30 % Acrylamide solution (100 ml)

Dissolve 29 g acrylamide and 1 g N, N'-methyene-bis-acrylamide in distilled water and adjust the volume to 100 ml. Mix the solution by stirring in dark bottle for 1

hour or till homogenious and filter through Whatman membrane No. 1. Store in the dark bottle at

4 °C.

6.5 Tris – Glycine electrode buffer (5x stock, 1 L)

Dissolve 15.1 g Tris base, 94 g glycine and 5 g SDS in distilled water. Adjust pH to 8.3 with HCl and adjust the volume to 1 L with distilled water.

6.6 Staining solution with coomassie brilliant blue for protein (1 L)

Mix 1 g Coomassie brilliant blue R-250, 400 ml methanol, 500 distilled water and 100 ml glacial acetic acid. Filter through Whatman No. 1.

6.7 Destaining solution for Coomassie Stain (1 L)

Mix 400 ml methanol, 100 ml glacial acetic acid and add distilled water to a final volume of 1 L.

6.8 10% (w/v) Ammonium persulfate (1 ml)

Dissolve 100 mg ammonium persulfate in 1 ml distilled water. Store at -20° C.

6.9 15% Separating gel SDS-PAGE (10 ml)

Mix the solution as follow:

1.5 M Tris pH 8.8	2.5 ml
Distilled water	2.3 ml
10% SDS	0.1 ml
30% acrylamide solution	5.0 ml
10% ammonium persulfate	0.1 ml
TEMED	0.008 ml

6.10 5% Stacking gel SDS-PAGE (5 ml)

Mix the solution as follow:

0.5 M Tris pH 6.8	1.26 ml
Distilled water	2.77 ml
10% SDS	0.05 ml
30% acrylamide solution	0.83 ml
10% ammonium persulfate	0.05 ml
TEMED	0.005 ml

7. Reagents for Western Blot

7.1 Transfer buffer (Towbin transfer buffer, 1 L)

Dissolve 3.03 g Tris Base, 14.41 g Glycine and 1 g SDS in 500 ml distilled water. Adjust pH to 8.2 – 8.4 and add 200 ml methanol. Adjust volume to 1 L with distilled water.

7.2 TBST buffer (2 L)

Mix 500 ml of 100 mM Tris-HCl, 300 ml of 1 M NaCl and 20 ml of 5 % NaCl. Adjust pH to 7.6 with conc. HCl.

7.3 Blocking buffer (1 ml)

Dissolve 25 mg BSA in 1 ml TBST. Freshly prepare.

8. Reagents for GUS assay

8.1 500 mM Sodium phosphate buffer pH8 (100 ml)

Mix 93.2 ml of 1 M Na_2HPO_4 and 6.8 ml of 1 M NaH_2PO_4. Take 50 ml mixture and adjust the volume to 100 ml with distilled water.

8.2 500 mM EDTA (100 ml)

Dissolve 18.61 g EDTA in 100 ml distilled water.

8.3 500 mM Potassium fericyanide (100 ml)

Dissolve 16.4 g Potassium fericyanide in 100 ml distilled water. Keep in dark bottle.

8.4 500 mM Potassium ferrocyanide (100 ml)

Dissolve 21.12 g Potassium ferocyanide in 100 ml distilled water. Keep in dark bottle.

8.5 5% Triton X-100 (100 ml)

Dissolve 5 ml triton X-100 in 70 ml distiied water. Adjust the volume to 100 ml with distilled water.

BIBLIOGRAPHY

Massalin Nakphaichit was born on May 17th, 1982 in Uttaradit, Thailand. She graduated with a bachelor degree from Department of Biotechnology, Faculty of Agro-Industry, Kasetsart University, in 2004. After graduation, she has opportunity to study master degree in School of Biotechnology, Institute of Agricultural at Suranaree University of Technology. She conducted the research in the topic of characterization of β-glucosidase from the rice *SFR2* gene as her thesis work. The research was support by the Thailand Research Fund (TRF).