CHARACTERIZATION OF THE ARABIDOPSIS SIEVE

ELEMENT-SPECIFIC EARLY NODULIN PROTEIN

By

NING SONG

Bachelor of Science in Biology

Shandong Normal University

Jinan, People's Republic of China

2007

CHARACTERIZATION OF THE ARABIDOPSIS SIEVE

ELEMENT-SPECIFIC EARLY NODULIN PROTEIN

Thesis Approved:

Dr. Gary A. Thompson

Thesis Adviser

Dr. Junpeng Deng

Dr. Jeanmarie Verchot-Lubicz

Dr. Mark E. Payton

Dean of the Graduate College

ACKNOWLEDGMENTS

First and foremost I wish to express my sincere gratitude to Dr. Gary A. Thompson for giving me the opportunity to work on this project. I would like to thank him for his patience, inspiration, enthusiasm, invaluable guidance and immense knowledge that he shared with me during my research. His constant support, guidance and encouragement kept me going while doing research as well as while writing this thesis.

I am also grateful and owe special thanks to my committee members Dr. Junpeng Deng and Dr. Jeanmarie Verchot-Lubicz for their valuable time, suggestions and insightful comments throughout the degree program. I also wish to extend my sincere thanks to Dr. Rita K. Miller for sharing with me the yeast two hybrid system and constructs for positive control.

I would like to thank our lab members as well as graduate students and Core Facilities staffs, James Anstead, Sampurna Sattar, Cherie Ognibene, Brian Krumm, Annabel Alonso, Maliha Rahman, Jacob Manjarrez, Alisha Dawn, Janet Rogers and Lisa Whitworth for their invaluable help and suggestions.

I would also like to thank Oklahoma Center for the Advancement of Science and Technology (OCAST) for funding my project during my studies.

Finally, I would like to thank my parents and fiancé, Si Zhang, for their concerns, encouragements and continuous supports to help me accomplish all the frustrations and hardships.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

SIEVE ELEMENT STRUCTURE AND FUNCTION

1.1 PHLOEM STRUCTURE AND FUNCTION

The vascular tissues of higher plants, xylem and phloem, are responsible for the long distance transport of water as well as organic and inorganic compounds throughout the plant body. Xylem is composed of continuous elongated conducting cells, called tracheary elements, which primarily transport water and inorganic nutrients. Transport in xylem is driven by a water potential gradient generated by the transpiration of water from surfaces of leaves and maintained through the cohesion-tension mechanism (Holbrook et al., 1995).

In contrast to the xylem, the phloem tissue transports photoassimilates, amino acids, organic acids, hormones, macromolecules, and some mineral nutrients, such as potassium, chloride, phosphate, and magnesium, in water from sites of synthesis to sites of utilization or storage. The phloem is a complex vascular tissue minimally composed of parenchyma cells, companion cells, and sieve elements. Sieve tubes are formed by files of interconnected sieve elements creating the highly specialized, living conduit. Sieve

elements are functionally associated and symplasmically connected via pore-plasmodesmata units with adjacent companion cells. In angiosperms, the sieve element-companion cell complex is derived from a common meristematic mother cell. Structural proteins including integral and membrane-associated proteins are likely synthesized within the immature sieve element. During differentiation, however, immature sieve elements undergo selective autophagy of specific organelles resulting in terminally differentiated cells that are incompetent for subsequent transcription and translation. At maturity, sieve elements are generally considered to rely on the metabolically active companion cells to synthesize macromolecules and other compounds, which are transported between the two cell types through pore-plasmodesmata.

The unique process of selective autophagy involves the degeneration of nuclei, disappearance of golgi apparatus, microtubules, ribosomes, reorganization of endoplasmic reticulum, and formation of sieve plates that join individual sieve elements (Evert R.F., 1990). Nuclear degeneration appears to be caused by chromatolysis. During differentiation, the chromatin becomes denser, showing prominent aggregates until the stainable contents of the nucleus disappear and the nuclear envelope is ruptured. The endoplasmic reticulum (ER) undergoes a series of changes from cisternal, rough-surfaced and randomly dispersed throughout the cytoplasm to stacks of smooth ER forming convoluted, tubular, or latticelike shapes that are in a parietal position attached to the sieve element plasma membrane (Evert et al., 1969; Esau and Hoefert, 1980; Torsch and Esau, 1981). Thus, the mature sieve element possesses a unique structure containing only a few mitochondria, plastids, smooth endoplasmic reticulum (ER) and functional plasma membrane. Ultrastructural studies of well preserved sieve elements suggest that the

parietal P-proteins, plastids, mitochondria and parietal smooth ER are associated together and firmly attached to the longitudinal plasma membrane through clamp-like structures, resulting in a low resistance lumen free of occlusions (Ehlers et al., 2000).

Münch's pressure-flow theory describes the fundamental mechanism that drives the rapid translocation of phloem sap in the unobstructed phloem lumen (reviewed by Schulz and Thompson, 2009). The translocation of phloem sap in the longitudinal sieve tube is driven by an osmotically-generated pressure gradient that forms between source (high solute concentration) and sink (low solute concentration) tissues. This gradient can be generated by several mechanisms that rely on differential concentrations of osmolytes, primarily sugars, being loaded into the sieve elements at the source and/or removed from sieve elements at the sink. Establishing the osmotic gradient is generally thought to be a function of phloem loading in source tissues.

Two generalized mechanisms of phloem loading, apoplastic loading and symplastic loading, are well accepted, although, there is considerable evidence for an alternative mechanism that does not rely on phloem loading to establish the pressure gradient. Apoplastic phloem loading is the typical phloem loading mechanism in plants that primarily or exclusively translocate sucrose. Sucrose, synthesized in photosynthically active mesophyll cells, accumulates in the apoplast of the sieve element-companion cell complex (SE-CC). Apoplastic sucrose is actively loaded via sucrose/proton symporters into the SE-CC utilizing energy from adenosine triphosphate hydrolysis (Lalonde et al., 2003). Symplastic loading involves the movement of sugars through plasmodesmata from the mesophyll cells into the SE-CC. A thermodynamically feasible mechanism of symplastic loading specific to the translocation of high concentration oligosaccharides

3

from raffinose-family or higher polymers of sucrose-galactosyl sugars is termed the polymer trap. In this mechanism, the sucrose symplasmically diffuses from bundle sheath cells to a specific type of companion cells, known as intermediary cells, which are extremely rich in pore-plasmodesmata units that are linked to the sieve element. In the intermediary cells, the sucrose is polymerized into sucrose-galactosyl polymers such as raffinose or stachyose preventing their diffusion back into the bundle sheath cells. Polysaccharide oligomers can, however, diffuse through wider pore-plasmodesmata units into sieve elements resulting in the high concentration of sugars required to establish the osmotic gradient (Turgeon and Medville, 2004; Schulz and Thompson, 2009). In some woody plants, such as *Salix*, the sucrose concentration in the mesophyll cytosol is much higher than that in sieve elements. This is explained by an "alternative" symplasmic loading model in which sucrose passively diffuses from mesophyll cytosol to the sieve element and the gradient appears to be established primarily by phloem unloading at the sink tissues (Schulz and Thompson, 2009).

1.2 SIEVE ELEMENT PROTEINS

1.2.1 P-PROTEINS

In angiosperms, structural phloem proteins (P-proteins) were detected as distinct proteinaceous inclusion bodies under light microscopy as early as 1854 by Hartig (Cronshaw and Esau, 1967). After the sieve element autophagy, the P-proteins are present in sieve elements as various morphological shapes, such as filaments in *Cucurbita maxima* and *Ricinus communis*, tubules in *Nicotiana tabacum* and *Coleus*

blumeii and crystalline in *Phaseolus vulgare* and *Glycine max* (Wergin and Newcomb, 1970; Palevitz and Newcomb, 1971; Cronshaw et al., 1973). The phloem filament protein (PP1) and phloem lectin (PP2) of cucurbits are two well-characterized P-proteins. In mature sieve elements, these proteins are synthesized in companion cells, transported into sieve elements through pore-plasmodesmata units, and assembled in sieve element parietal layer (Clark et al., 1997). It appears that P-proteins are maintained in a dynamic equilibrium between unpolymerized subunits that are translocated in the phloem sap and large, ultrastructurally distinct polymers (Golecki et al., 1999). Although a variety of alternative functions are being attributed to PP2 (Bencharki et al., 2010), P-protein filaments are generally considered to maintain turgor pressure within damaged sieve tubes by accumulating at sieve plates, thereby, occluding the sieve elements that are adjacent to the damaged cells (Clark et al., 1997; Golecki et al., 1999).

Elegant studies on the forisomes (crystalline P-proteins) of the Fabaceae have revealed the regulatory events that occur during the occlusion of sieve elements in response to vascular damage (Knoblauch and Peters, 2004). Forisomes are calcium-dependent contractile proteins located adjacent to the sieve plates, but not fixed to the parietal endomembrane system (Peters et al., 2007). In *Vicia faba* leaflets, the diameter of sieve element has a direct correlation with the size of the forisome (Peters et al., 2006). Interestingly, forisomes structurally and reversibly transform from "crystalline" to "amorphous" forms in response to changing calcium concentrations and pH, thereby providing a mechanism for rapid, reversible blocking of individual sieve tubes. Evidence is accumulating that calcium concentrations in wounded sieve elements can be

manipulated by phloem-feeding insects to prevent sieve element occlusion by forisomes (Will et al., 2007).

1.2.2 SOLUBLE SIEVE ELEMENT PROTEINS

In addition to the structural P-proteins, a large number of soluble proteins have been detected in phloem sap that are significantly different from protein profiles found in adjacent cells (Turgeon and Wolf, 2009). Complete sets of enzymes that function in several different biosynthetic reactions have been detected in sieve elements. For example, many enzymes needed for ascorbic acid, jasmonic acid and alkaloid biosynthesis have been identified in phloem sap from *squash, tomato,* and *opium poppy.* The products of these enzymatic pathways participate in stress and defense responses (Bird et al., 2003; Hancock et al., 2003; Hause et al., 2003). Several kinases, including calmodulin-like domain protein kinase and calcium-independent protein kinase, have been detected in phloem sap. These kinases are thought to be responsive to abiotic and biotic stresses as well as influence plant development (Yoo et al., 2002). Another group of important phloem sap proteins are antioxidants, such as superoxide dismutases, reductases, and peroxidases. These enzymes ensure the maintenance of stable physiological conditions in sieve elements and could also serve in defense mechanisms (Ishiwatari et al., 1995; Szederkenyi et al., 1997; Walz et al., 2002).

Recently, considerable work has been done in identifying and characterizing RNA binding proteins in phloem sap. Several types of RNAs, such as mRNAs, microRNAs, and small interfering RNAs have been detected in phloem sap (Turgeon and Wolf, 2009)

that appear to be mobile regulatory signals in the plant body (Deeken et al., 2008). The first RNA binding protein, CmPP16 in pumpkin was detected as a homolog of viral movement protein and functions to mediate RNA transport (Xoconostle-Cazares et al., 1999). Interestingly, the P-protein CmPP2 in cucumber was shown to play a role in systemic infection by binding viral RNA in addition to phloem mRNA (Gomez and Pallas, 2004). RNA binding proteins, such as the 50 kDa polypyrimidine tract binding protein RBP50 in pumpkin, form a functional core in the ribonucleoprotein complex (Ham et al., 2009). It is now clear that both protein- and RNA-based communication networks function within the phloem to transmit regulatory information throughout the plant body (Lough and Lucas, 2006).

1.2.3 SIEVE ELEMENT MEMBRANE PROTEINS

The physiology of the phloem is highly dependent upon proteins that are integrated or associated with the sieve element membranes, especially the sieve element plasma membrane. Integral membrane proteins play essential roles in regulating the movement of metabolites, water, and ions across the sieve element plasma membrane. The best studied integral membrane proteins are the sucrose transporters that facilitate the uptake of sucrose into sieve elements (Stadler et al., 1995; Kuhn et al., 1997).

Sucrose transporters (SUTs) can be categorized into three subclasses, SUT1/SUT3, SUT2 and SUT4, based on their protein sequence homology, substrate affinity and functions. SUT1 subfamily members have the highest affinity for the substrate. In *Solanum tuberosum*, the activity of StSUT1 can be regulated by redox reagents and both

the monomer and dimer form proteins are detected *in planta*. SUT3 was only found in tobacco plants and shares high sequence homology with the SUT1 subfamily (Krugel et al., 2008). In *Arabidopsis*, AtSUC2 is the predominant sucrose proton symporter with dual functions of efficiently transporting sucrose and as a sucrose sensor in sieve elements (Barker et al., 2000). SUT4 subfamily members have the lowest substrate affinity. All the sucrose transporter proteins have been detected in sieve elements, whereas SUT1 mRNA was only detected in companion cells (Hackel et al., 2006).

Other membrane proteins also play important roles in sieve element physiology. A phloem-specific ATPase has been identified that is thought to be responsible for establishing the proton gradient across the plasma membrane that is required for active phloem loading as well as other physiological processes in the phloem (DeWitt and Sussman, 1995). Maintaining proper water relations within the sieve element is another important function that affects long-distance transport, and aquaporins, a group of integral membrane proteins with roles in regulating water transport, have been identified in the phloem (Barrieu et al., 1998). In higher plants, amino acids that are produced in roots or shoots are transported throughout the plant organs through the phloem and xylem (Hirner et al., 2006). In *Arabidopsis*, the amino acids transporters are classified into six family members *AAP1-AAP6* (Okumoto et al., 2004). Two of the family members, *AAP4* and *AAP5,* are primarily expressed in mature leaves and play major roles in amino acid translocation through phloem loading (Fischer et al., 1995). Ion channels, including calcium ion channels, a photosynthate-induced potassium channel, and a sulfate transporter are less well characterized, but have been detected in the sieve element plasma membrane (Lacombe et al., 2000; Lytle et al., 2000; Yoshimoto et al., 2003).

1.3 SIEVE ELEMENT EARLY NODULIN-LIKE PROTEIN (SE-ENOD)

The work described in this thesis focuses on a novel sieve element plasma membrane protein that was initially identified and characterized by Dr. Junaid Khan and Dr. Qi Wang in a collaboration between Dr. Richard Sjolund and Dr. Gary Thompson (Khan et al., 2007). A collection of monoclonal antibodies (mabs) was generated from partially purified sieve elements isolated from *Streptanthus tortuosus* (Brassicaceae) callus cultures. One of the mab lines, RS6, immunolocalized an antigen to the sieve element plasma membrane within the phloem of *Arabidopsis* and *Brassica* species. High resolution immunolocalization experiments showed that the antigen accumulated in the plasma membrane at the earliest stages of sieve element differentiation and persisted in mature sieve elements. Affinity purification of the protein antigen from *S. tortuosus* phloem (+) tissue cultures followed by N-terminal peptide sequencing of the entire peptide and a tryptic fragment provided two peptide sequences (amino-terminus REFAVGGAKGWTIPS and internal TSFSLTHSGPYYFISGNK) that corresponded to blue copper binding protein (cupredoxin) in BLASTP searches.

Cupredoxins are a superfamily of type 1 blue copper binding proteins that facilitate single electron transport both in plants and bacteria (Adman, 1991; Ryden and Hunt, 1993). These proteins are typically composed of a single polypeptide chain of 100 to 140 amino acids that share a conserved copper binding center core domain. X-ray crystallographic structure analysis revealed that many family members share a highly similar core structure consisting of a four to seven β-strands folded into a Greek key β-barrel structure (Chothia and Lesk, 1982). In order to stabilize the binding of copper, the binding center consists of three or four copper binding ligands and two highly conserved

cysteines forming a disulfide bond. Consequentially, the geometry of the copper binding center is a distorted tetrahedral architecture with a single copper atom (Ainscough et al., 1987). Plastocyanin is the type member of cupredoxins in higher plants (phytocyanins) that transfers single electrons from cytb6/f complex in photosynthesis system II to P700* of photosynthesis system I (Ubbink et al., 1998). Based on the spectroscopic properties, primary peptide sequence, glycosylation, and copper axial ligands, the phytocyanins can be classified into four subgroups: mavicyanins, stellacyanins, uclacyanins and early nodulins (ENODs) (Nersissian et al., 1998).

While the experimentally-derived peptide sequences could be used to initially identify the RS6 sieve element antigen as a phytocyanin, the specific protein could not be determined. This was accomplished by designing two sets of degenerate oligonucleotide primers based on the peptide sequences for nested reverse transcriptase polymerase chain reactions (nested RT-PCR) of total RNA isolated from *Arabidopsis* leaves. BLASTN searches of the TAIR (The *Arabidopsis* Information Resource) database revealed that the cDNA sequence obtained from nested RT-PCR corresponded to the *At3g20570* gene. The deduced peptide encoded by *At3g20570* revealed similar domain architecture with *Medicago truncatula* ENOD16 and 20 proteins and the related pea PsENOD5 protein.

The ENODs are encoded by nodule-specific genes that are expressed at the earliest stages of nodule development during rhizobia interactions with legumes (Mylona et al., 1995). Sequence alignment revealed that ENODs are structurally related to cupredoxins, belonging to a subfamily of phytocyanin-like proteins. The typical ENOD precursor protein domain structure is composed of four domains: signal peptide; plastocyanin-like copper binding domain; a variable-length proline/serine rich domain; and a hydrophobic

10

domain. Significantly, the ENODs contain single amino acid substitutions for the copper binding ligands (Greene et al., 1998). Three of the four copper binding ligands (histidine, cysteine, histidine and methionine or glutamine) are not conserved in Pea PsENOD5, vetch VsEOND5, soybean ENOD55 and *Medicago truncatula* MtENOD16 and 20 (Scheres et al., 1990; Vijn et al., 1995; Greene et al., 1998). These observations have let to the general assumption that ENODs have lost the ability to bind copper and have functions other than electron transport. The tertiary structures of the plastocyanin-like copper-binding domain in several ENODs have been predicted by homology modeling, and the core structures were predicted to be a Greek key β-barrel.

Multiple sequence alignment of twenty-nine *Arabidopsis* phytocyanin-like proteins and construction of the phylogenetic tree indicated that the RS6 antigen detected in sieve element plasma membrane belongs to the ENOD-like protein subfamily and contains the identical domain architecture of signal peptide (domain I), plastocyanin-like copper binding domain (domain II), proline/serine rich domain (domain III) and hydrophobic domain (domain IV) (Figure 1.1). Other computer algorithms predicted that the sieve element-specific ENOD-like protein (SE-ENOD) contains phosphate-binding loop (P-loop), two arabinogalactan glycosylation recognition sites, and alternative ω sites for glycosylphosphatidylinositol (GPI) anchor. The signal peptide and the hydrophobic domain are posttranslationally cleaved from the precursor protein and generate a mature protein with plastocyanin-like copper binding domain and proline/serine rich domain, which is proposed to be attached to the plasma membrane by a GPI anchor (Khan et al., 2007). To determine whether the SE-ENOD is GPI anchored in the sieve element plasma membrane, young cauliflower stem sections were treated with phosphatidylinositol-

specific phospholipase C (PI-PLC), which specifically cleaves the phosphatidylinositol bond in GPI-anchored proteins (Pfaffmann et al., 1987). In immunolocalization experiments, the RS6 mab failed to detect the presence of the SE-ENOD in the sieve elements of fresh stem sections pre-treated with PI-PLC, suggesting that the enzymatic treatment was sufficient to release the antigen and that the SE-ENOD is attached to the sieve element plasma membrane by a GPI anchor.

Immunological and molecular data obtained using the RS6 mab were consistent with the conclusion that the *At3g20570* gene encoded the SE-ENOD. Furthermore, approximately 1000 bp of 5' flanking sequences encompassing the *At3g20570* promoter directed β-glucuronidase activity specifically in the phloem of the vascular of leaves, stems, roots, and inflorescences. The analysis of T-DNA insertion mutants provided the final confirmation of the linkages among the mab, protein, and gene. Immunolocalization of *Arabidopsis Columbia-0* and two independent, homozygous T-DNA insertion lines of *At3g20570*, SALK 105873 and GABI-371E08, with RS6 antibody showed that the SE-ENOD is present in the wild type plant and absent in both mutants. The phenotypes between the *Col-0* and SALK 105873 were compared, and while a significant difference was detected in the number of siliques, the mutants failed to produce a convincing phenotype.

Plastocyanin-like copper binding domain

Signal Peptide

MARNLKSMMLCGFGLLCFLMIVDRAYAREFTVGGATGWTVPSGSQVYSQW
AEQSRFQIGDSLLFVYQSNQDSVLQVTRDAYDSCNTDSPTAKFADGKTSVTL
NHSGPYYFISGNKDNCKKNEKLVVIMADRSGNKNTASSPPSPAPAPSGESA
PSPPVSGTFEMIPAPTPTTSEDTPNSAASSLSFVAALLGAALASTLFLH

ω

Proline/Serine rich domain

Hydrophobic domain
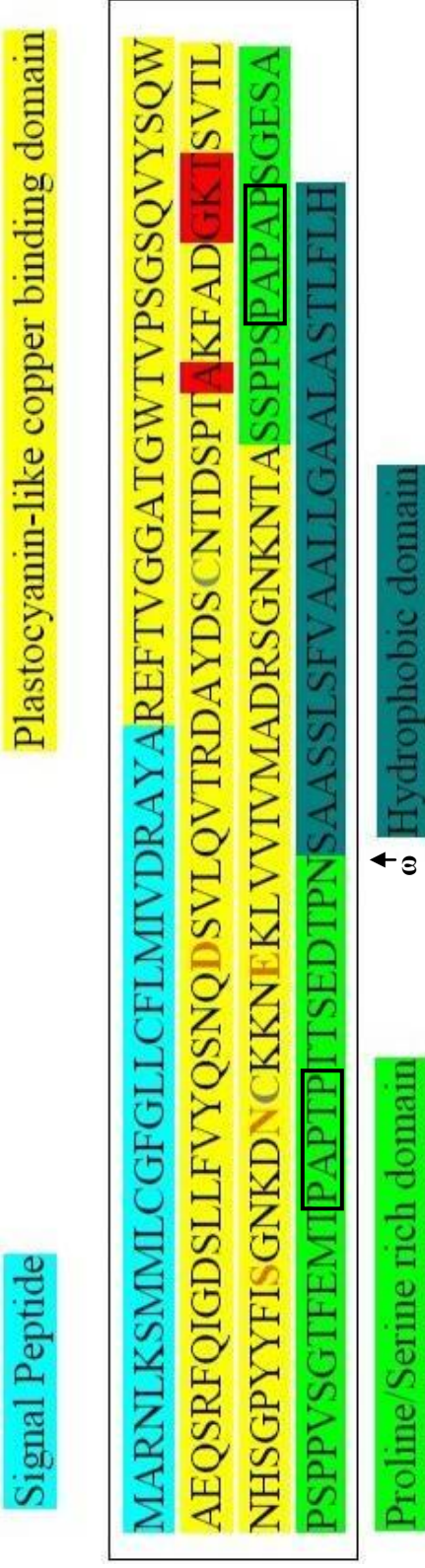
Figure 1.1, Domains in the SE-ENOD.

The SE-ENOD domain arrangement is identical to the ENOD subclass proteins, which are domain I, signal peptide (highlighted in cyan), domain II, plastocyanin-like domain/copper binding domain (highlighted in yellow), domain III, proline-serine rich domain (highlighted in green) and domain IV, hydrophobic domain (highlight in blue). The two bold grey cysteine residues represent the two conserved cysteines predicted to form a disulfide bond. The bold orange amino acid residues D, S, N and E represent the four substituted copper binding ligand sites. The $AX_{(4)}GKT$ residues highlighted in red represent the P-loop. Two sequences in the proline/serine rich domain boxed are arabinogalactan glycosylation recognition sites. The black arrowhead shows the alternative ω site for glycosylphosphatidylinositol (GPI) anchor.

13

1.4 OVERVIEW OF THIS STUDY

The goal of this research project was to further characterize structural components of the sieve element-specific early nodulin-like (SE-ENOD) protein to provide insights into the role this plasma membrane-associated protein plays in plant vascular biology. Three general approaches were taken to achieve this goal. In the first approach, structural predications and models were generated based on known structures of proteins within the same superfamily. In the second approach, two microbial expression systems were used to validate some of the computer-generated models created in the structural studies. In the third approach, studies were conducted to examine the protein *in vivo* using phloem-enriched tissues.

Bioinformatic tools provide powerful approaches to study protein structure-function relationships (Teufel et al., 2006). The second chapter of the thesis presents the strategies and results of *in silico* analyses to determine the secondary and tertiary structure of the SE-ENOD. The predictions suggested that the overall structure of SE-ENOD is four polypeptide β-strands arranged as a Greek key topology structure located between another two β-strands and organized on one side of two polypeptide α-helices. Models were constructed to determine whether the protein with copper-binding pocket that no longer binds the metal can alternatively interact with nucleotide phosphates via well-characterized ATP/GTP-binding motifs. Additional models suggest that the SE-ENOD has the potential of forming homodimers or interacting with other molecules.

Microbial expression systems are well-established molecular tools that contribute to the functional studies of proteins from plant and animals. The third chapter of the

thesis describes the studies that were conducted with recombinant peptide domains expressed and purified from *E. coli* to determine if ATP/GTP-binding motifs in domain II of the SE-ENOD are functional. This chapter of the thesis also describes experiments that were conducted using a modified yeast two-hybrid system to determine whether the SE-ENOD can form homodimers in *Sacchromyces cerevisiae*.

Previous studies of the SE-ENOD in *Arabidopsis* centered on its identification and initial molecular characterization, however, the conformational states of this plasma membrane associated protein *in planta* were not addressed. This was primarily due to the difficulties encountered in isolating significant amounts of enriched vascular tissue. The fourth chapter of this thesis describes the development of a new technique to isolate large amounts of highly enriched vascular tissues from broccoli plants and the use of these samples to initially characterize the SE-ENOD *in planta*.

CHAPTER II

BIOINFORMATICS APPROACHES TO STUDY SE-ENOD STRUCTURE

2.1 INTRODUCTION

The SE-ENOD is a novel protein with unknown structure belonging to the cupredoxin superfamily of copper-binding proteins. The overall structure of cupredoxins is well resolved. Members in this family consist of a single polypeptide chain of approximately 100-140 amino acid residues arranged into a "single copper binding center" core structure that functions in electron transfer reactions by accepting or donating a single electron to their redox partner proteins (Adman, 1991; Ryden and Hunt, 1993). The copper binding center in cupredoxins share high similarity in the copper binding ligands Cys-$X_{(n)}$-His-$X_{(m)}$-Met/Gln located in a single loop region. The fourth ligand, histidine, is normally located near the carboxyl terminus at the peptide. This active copper binding center appears to have a geometry of a distorted tetrahedron and binds a single copper atom (Adman, 1991).

Phytocyanins are the predominant cupredoxins in higher plants. Based on the spectroscopic properties, primary peptide sequence, glycosylation, and copper axial ligands, the phytocyanins can be classified into four subgroups: mavicyanins,

stellacyanins, uclacyanins and early nodulins (ENODs) (Nersissian et al., 1998). The protein domain architecture in this family is conserved and consists of four domains: signal peptide (domain I), plastocyanin-like copper-binding domain (domain II), proline/serine rich domain (domain III), and hydrophobic domain (domain IV). The three-dimensional crystal structures of several phytocyanins, such as mavicyanin from *Cucurbita pepo medullosa* (zucchini), have been resolved at 1.6 Å resolution (Xie et al., 2005).

Cucurbita mavicyanin is a glycosylated 109 amino acid blue copper protein containing four copper binding ligands His[45], Cys[86], His[91], and Gln[96] (Guckert et al., 1995). The protein has a core structure of seven β-strands arranged as a Greek key β-barrel in the middle of two β-sheets and three α-helixes (Xie et al., 2005). A highly conserved disulfide bond between Cys[58] and Cys[92] is significant in maintaining the overall structure of the protein. The X-ray crystallographic analysis of *Cucurbita* mavicyanin revealed that four 109 amino acid subunits are arranged into a single tetrameric asymmetric lattice (Xie et al., 2003).

The overall approach taken to model the structure of the SE-ENOD initially used mavicyanin as the predominant reference molecule for predictions. Secondary structure predictions use algorithms that assign one of three states (H: helix, E: strand and C: the rest) to each amino acid residue. These secondary structure codes are obtained by exploiting evolutionary relationships determined from multiple homologous sequence alignments. Homology modeling requires an accurate alignment between the target protein and its template to identify similar regions. To date, millions of proteins have been sequenced, however, only a limited number of domains and topologies have been

detected (Vitkup et al., 2001). Homology modeling takes advantage of the fact that protein primary sequence changes more rapidly than its structure and the protein amino acid sequence determines its tertiary structure. Therefore, a group of proteins that share a significant level of amino acid identity and similarity should adopt a similar tertiary structure (Chothia and Lesk, 1986; Sander and Schneider, 1991).

Khan and co-workers (2007) identified a phosphate-binding region within domain II of the SE-ENOD that was first described by Walker et al. (1982) as Walker A and B motifs. Several consensus amino acid arrangements of Walker A motifs, which are also known as P-loops, have been reported in the literature (Walker et al., 1982; Satishchandran et al., 1992). These include $(G/A)X_{(4)}GK(T/S)$, $GX_{(4)}GX_{(2)}K$, or $GX_{(2)}GXGKS$ with X representing any amino acid residue. The consensus sequence of Walker B has been defined as $(R/K)X_{(7)}h_{(4)}D$, with X representing any amino acid and h representing hydrophobic residues (Walker et al., 1982). These motifs are commonly found in adenosine triphosphate nucleotide (ATP)-binding proteins that often have ATPase activity. The residues in Walker A, especially Lys, interact with the phosphate group of the ATP, while the residues in Walker B, especially Asp, interact with the magnesium ion (Mg^{2+}) and contribute to the stability of catalytic status (Reinstein et al., 1990; Krell et al., 2001). The presence of both motifs suggests that the SE-ENOD binds Mg^{2+}-ATP. Pocket-Finder is a well-suited program for predicting ATP binding sites. It is an online-based algorithm that employs the LIGSITE program to predict potential pockets. The LIGSITE program identifies small molecule binding sites based on several probe scanning command on a cubic grid. Predicted binding sites are only retained if the probe can detect it as a pocket for at least five times (Hendlich et al., 1997).

The SE-ENOD is a sieve element plasma membrane-anchored protein that is expressed in nucleated sieve elements and persists in mature sieve elements. When this protein is partially purified from enriched vascular tissues, it appears to be in the form of complexed oligomers (See Chapter IV). Several other membrane-associated proteins in sieve elements regulate molecular transport in an oligomeric form (Krugel et al., 2008). Protein-protein docking analysis is a powerful bioinformatic approach to predict protein oligomers or complexes. The prediction employs several algorithms to evaluate parameters such as geometric complementarities, electrostatic energy, and collision possibility to obtain the most accurate model (Fischer et al., 1993; Stites, 1997; Mandell et al., 2001; Comeau et al., 2007). ClusPro is a fully automated, web-based program that can perform rigid body docking prediction. Rigid body docking is a filter method that can refine the docking predictions by screen energy properties and retain the models that with minimized energy potential (Comeau et al., 2007).

In this chapter, the SE-ENOD secondary structure was validated by four different programs, PSIpred, SSpro, HNN and Porter. The synthetic analysis for the results from these programs predicted that the secondary structure of SE-ENOD is composed of six strands, two or three helixes and a few coil structures. Further tertiary structure was predicted by SWISS-MODEL homology modeling and the core structure of the molecule is predicted to be four β-strands arranged into a Greek key β-barrel topology. This Greek key β-barrel is organized in the middle of two antiparellel β-strands, surrounded by two α-helixes and a few coils. Based on the sequence characteristic of the SE-ENOD P-loop, a potential ATP binding pocket was predicted by Pocker-Finder. Additionally, ClusPro

protein-protein docking server predicted that two single SE-ENOD molecules have the

potential to form a homodimer.

2.2 MATERIALS AND METHODS

2.2.1 SE-ENOD domain II (plastocyanin-like copper binding domain) secondary structure prediction by four different web-based programs, PSIpred, SSpro, HNN and Porter.

The primary sequence of the SE-ENOD domain II was input into different secondary structure prediction servers: PSIpred (http://bioinf.cs.ucl.ac.uk/psipred/), SSpro (http://scratch.proteomics.ics.uci.edu/explanation.html#SSpro) , HNN (http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_nn.html) and Porter (http://distill.ucd.ie/porter/). The predictions were run with server default parameters. The outputs of these four well-known programs adopt the identical three secondary structure states for each amino acid residues in the query sequence. The outputs were viewed and analyzed in MEGA 4.0, a software used for molecular evolutionary genetics analysis. It is commonly used to analyze multiple sequences, calculate evolutionary distance, estimate evolution relationships among the sequences, and generate phylogenetic map with multiple choices of distance method and models (Tamura et al., 2007; Kumar et al., 2008). In MEGA 4.0, the secondary structure prediction results of SE-ENOD was labeled by different color and output as a picture. As a control and comparison for the secondary structure prediction, the secondary structure of *Cucurbita* mavicyanin was also predicted by the same programs, and the resolved mavicyanin secondary structure was directly output from DSSPcont (http://cubic.bioc.columbia.edu/services/DSSPcont/).

2.2.2 Homology modeling for the SE-ENOD tertiary structure by SWISS-MODEL

The template for SE-ENOD homology modeling was initially searched by FASTA @ EMBL-EBI Protein Similarity Search (http://www.ebi.ac.uk/Tools/fasta33/). The database Protein Structure Sequences in RCSB protein data bank (PDB http://www.pdb.org/pdb/home/home.do) and all the other defaulted parameters were used. The SE-ENOD amino acid sequence was uploaded in a text file format. The molecules having available crystal structures that share >50% identity and similarity with the SE-ENOD were listed, and chain C of *Cucurbita* mavicyanin was ranked as the top candidate. The PDB (Text) file of *Cucurbita* mavicyanin crystal structure (PDB ID 1WS8) was downloaded from PDB and analyzed by PyMOL Molecular Viewer (http://www.pymol.org/). The four chains of mavicyanin were displayed by checking the "Sequence mode" under "Display" in PyMOL program, and chain C was selected and saved as a separate molecule file. The mavicyanin chain C molecule was uploaded on the SWISS-MODEL server as the template for homology modeling for the SE-ENOD domain II. SWISS-MODEL (Peitsch et al., 1995; Arnold et al., 2006; Kiefer et al., 2009) is an automated protein tertiary structure prediction server based on target-template homologous alignment (Schwede et al., 2003).

2.2.3 The SE-ENOD functional region and complex prediction

Potential pockets were predicted by uploading the predicted SE-ENOD tertiary structure .pdb file onto the website Pocket-Finder (http://www.modelling.leeds.ac.uk/pocketfinder/). Protein-protein docking was predicted by ClusPro (http://cluspro.bu.edu/home.php). The predicted SE-ENOD tertiary structure .pdb file was uploaded to the server. Dimers were predicted by input "2" for the "number

of multimers" under "Advanced Options→Multimer Docking" option, then run the

DOCK. The outputs were viewed and analyzed in PyMOL.

2.3 RESULTS

2.3.1 Predicted secondary structure of the SE-ENOD plastocyanin-like copper binding domain

The *Arabidopsis* SE-ENOD and *Cucurbita* mavicyanin belong to the plastocyanin-like copper binding protein family and share 39.4% amino acid sequence identity and 72.7% amino acid sequence similarity. The two highly conserved cysteine residues, Cys^{58} and Cys^{92}, in mavicyanin are present within the SE-ENOD sequence, however, the four copper binding ligands in mavicyanin, (His^{45}, Cys^{86}, His^{91} and Gln^{96}) are substituted in the SE-ENOD (Gln^{43}, Ser^{85}, Asn^{90} and Glu^{95}). The amino acid sequence alignment between the mature proteins (domains II and III) shown in Figure 2.1 can shed light on the accuracy of the analysis of the SE-ENOD secondary structure.

The SE-ENOD secondary structure predictions output from PSIpred, SSpro, HNN and Porter programs are shown in Figure 2.2. Each residue is assigned by a specific color by MEGA 4.0. The four prediction programs adopt the algorithm that assigns each amino acid residue to one of three states (H: helix, E: strand and C: the rest). The secondary structure of mavicyanin was also predicted by the same programs, and they predicted similar secondary structure (Figure 2.2, SSpro, HNN, and Porter data not shown). The experimentally determined secondary structure of mavicyanin was directly output from DSSPcont, however, the DSSPcont program adopts an alternative algorithm that assigns every residue to one of the eight states (H: α-helix, G: 3-10-helix, I: pi-helix, E: extended strand, B: β-bridge, T: turn, S: bend and C: the rest). Therefore, the secondary structure of H, G, I are grouped into Helix (H); E, B are grouped into Strand (E); S, T and the rest are

grouped into Coil (C) in accordance with the three state algorithm. Comparison between the resolved secondary structure of mavicyanin from DSSPcont and the PSIpred prediction result confirmed that the predictions primarily reflect the actual structure.

Analysis based on the total prediction output suggests that the SE-ENOD secondary structure is composed of six strands, two or three helixes and several coil structures.

```
>>PDB:1WS8_C mol:protein length:109 mavicyanin         (109 aa)
 initn: 240 init1: 208 opt: 267  Z-score: 340.4  bits: 68.6 E(): 5.7e-12
Smith-Waterman score: 267; 39.4% identity (72.7% similar) in 99 aa overlap (5-101:7-102)


              10        20        30        40        50
Sequen    REFTVGGATGWT--VPSGSQVYSQWAEQSRFQIGDSLLFVYQSNQDSVLQVTRDAYDS
          ::  .::::   ::      :..:: ...:..:::::: :...  .::::  ..  . :
PDB:1W MATVHKVGDSTGWTTLVPYD---YAKWASSNKFHVGDSLLFNYNNKFHNVLQVDQEQFKS
          10        20        30        40        50

       60        70        80        90        100       110
Sequen CNTDSPTAKFADGKTSVTLNHSGPYYFISGNKDNCKKNEKLVVIVMADRSGNKNTASSPP
       ::..::.:....:  :. :.. : .::. :    .:. ..:. . :
PDB:1W CNSSSPAASYTSGADSIPLKRPGTFYFLCGIPGHCQLGQKVEIKVDPGSSSA
       60        70        80        90        100

        120       130       140       150
Sequen SPAPAPSGESAPSPPVSGTFEMTPAPTPTTSEDTPN
```

Figure 2.1 Amino acid sequence alignment of the *Arabidopsis* SE-ENOD and *Cucurbita* mavicyanin mature proteins (domains II and III). The *Arabidopsis* SE-ENOD (152 amino acids) is labeled as Sequen and *Cucurbita* mavicyanin (109 amino acids) is labeled as PDB: 1W. The residues marked by the "**:**" indicates identical residues and residues marked by the "**.**" indicates similar residues. The cysteine residues that appear to form a highly conserved disulfide bond are connected by brackets. The copper binding ligands His^{45}, Cys^{86}, His^{91} and Gln^{96} in mavicyanin are indicated by " $\triangle$ " and the corresponding substituted residues Gln^{43}, Ser^{85}, Asn^{90} and Glu^{95} in SE-ENOD are indicated by " ▼ ".

26
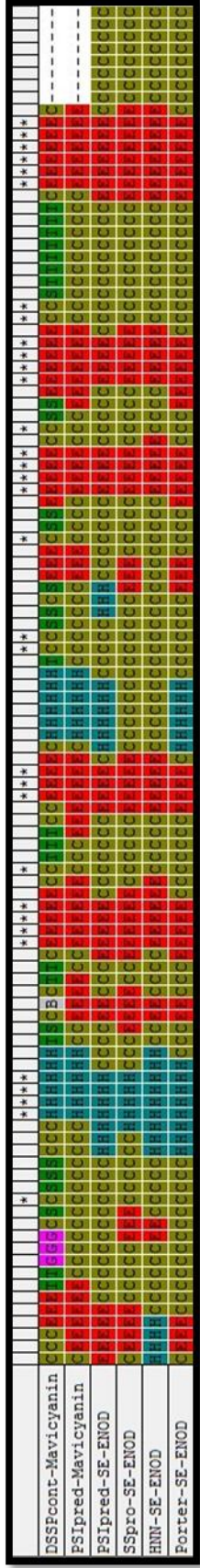
Figure 2.2 Diagrammatic representation of the predicted SE-ENOD domain II (plastocyanin-like copper binding domain) secondary structure comparisons. The predictions from four different web-based programs, PSIpred, SSpro, HNN, and Porter, were viewed and analyzed by MEGA 4.0. The predicted SE-ENOD structures are color coded according to the type of secondary-structure: α-helixes (H) are labeled cyan, β-strands (E) are labeled red, coils (C) are labeled dark yellow, 3-10 helixes (G) are labeled hot pink, bends (S) and turns (T) are labeled green and β-bridges are labeled grey. There are 105 amino acid residues in mavicyanin domain II, 113 amino acid residues in the SE-ENOD domain II, therefore, eight dashes are added to the end of mavicyanin secondary structure prediction. The asterisks " * " indicate consensus structure elements.

27

2.3.2 SE-ENOD domain II tertiary structure homology modeling

The tertiary structure of SE-ENOD was predicted by SWISS-MODEL based on the crystal structure of *Cucurbita* mavicyanin chain C as the template. This template was determined by a Similarity and Homology BLAST on FASTA @ EMBL-EBI. As shown in Figure 2.1, *Cucurbita* mavicyanin shared the greatest identity and similarity with the SE-ENOD. The significantly low E value of 5.7e-12 is a good indicator that the mavicyanin can serve as an appropriate template for the SE-ENOD homology modeling.

The SE-ENOD tertiary structure homology modeling was analyzed by PyMOL and shown in Figure 2.3. The predicted structure of SE-ENOD contains three pairs of antiparallel β-strands, three α helices and several loops. The core structure is composed of the four β-strands E2-E3-E4-E5 that are arranged into a Greek key β-barrel topology. The Greek key β-barrel topology is composed of three up-down-up β-strands (E2, E3, and E4) labeled as hot pink, cyan, and blue, respectively, connected to the fourth down β-strand (E5) labeled red. This fourth down E5 β-strand is located adjacent to the first up E2 β-strand through a long connection composed of a loop-helix-loop topology, labeled wheat. All the other connection loops are also labeled wheat.

It is important to point out that the SE-ENOD tertiary structure prediction is based on a copper binding protein as a homologous template even though the four copper binding ligands shown in Figure 2.4 are all substituted. Residues Gln[43], Ser[85], and Glu[95] are located on predicted loop regions, while Asn[90] is located on the predicted third α-helix (H3) (Figure 2.4). Structural modeling studies of *Medicago truncatula* ENOD proteins, MtENOD16/20, indicated that the members in ENOD-like protein family do not

bind copper (Greene et al., 1998). This raises some uncertainty about the accuracy of the SE-ENOD tertiary structure based on the copper binding mavicyanin as the template.

The common features of proteins in the ENOD-like subfamily of phytocyanins are a highly conserved disulfide bond and the substitution of copper binding ligands by other amino acid residues. Multiple sequence alignment of fourteen ENOD-like proteins generated by Clustal X 2.0.10 exhibited the two conserved Cys^{57} and Cys^{91} residues that form a putative disulfide bond (Figure 2.5). The SE-ENOD plastocyanin-like copper binding domain contains predicted Walker A (P-loop) and Walker B ATP binding motifs. The Walker A motif extends from residue Ala^{64} to Thr^{71} and Walker B motif extends from residue Lys^{92} to Asp^{104} (Figure 2.5). Based on this alignment, it is significant to point out that among the fourteen proteins only the SE-ENOD contains Walker A and Walker B motifs.

The Walker A motif residues are also shown on the predicted tertiary structure model. The full Walker A motif extends from the third β-strand (E3) to the fourth β-strand (E4) with the catalytic residue Lys^{70} side chain facing to the surface of the predicted model (Figure 2.4). In addition, the sequence of *Cucurbita* mavicyanin domain II (105 amino acids) is eight residues shorter than that of *Arabidopsis* SE-ENOD (113 amino acids), thus the homology modeling for the SE-ENOD ends at residue Ile^{100}. This eliminated the full sequence of Walker B motif (from Lys^{92} to Asp^{104}) from showing in the predicted tertiary structure of the SE-ENOD.
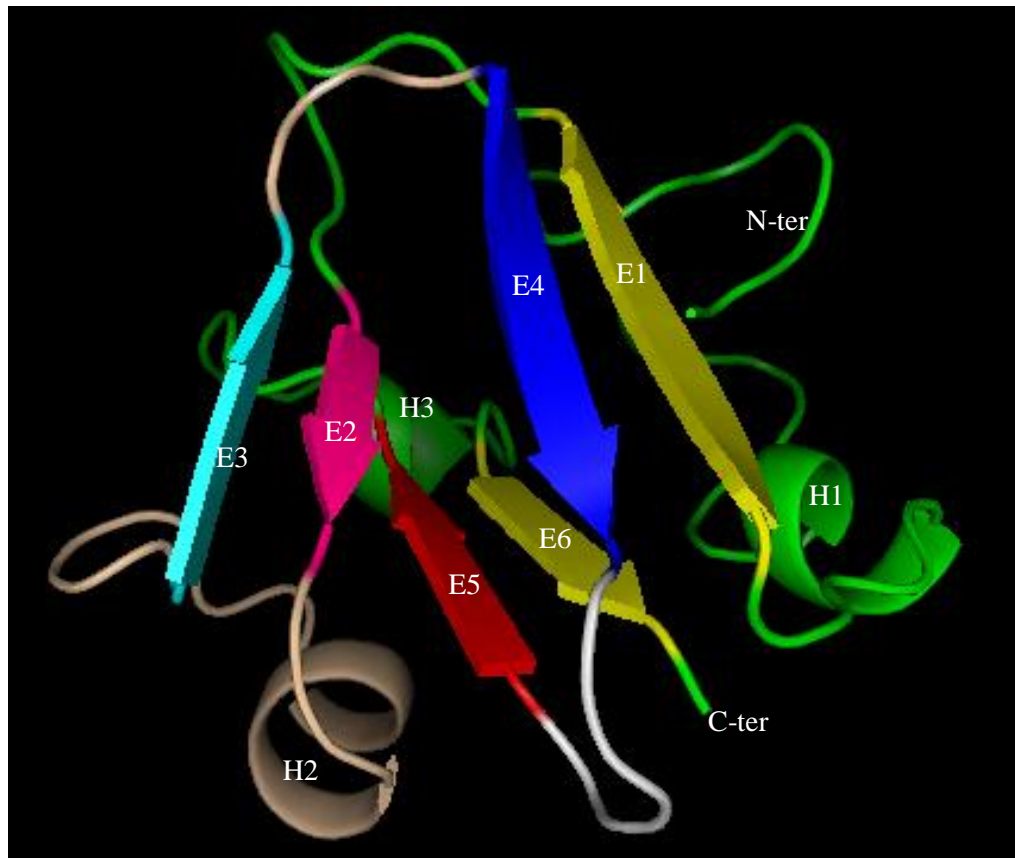
Figure 2.3. Ribbon diagram of the predicted *Arabidopsis* SE-ENOD domain II tertiary structure with *Cucurbita* mavicyanin chain C as a template viewed from the side of the molecule. Based on the SWISS-MODEL homology modeling with mavicyanin chain C as the template, the SE-ENOD tertiary structure is predicted to be consisting of six β-strands, three α-helixes and a few loops. The core structure is composed of a Greek key β barrel topology structure, containing E2-hot pink, E3-cyan, E4-blue, E5-red, arranged into two antiparallel β-strands (E1and E6, both labeled yellow), organized on one side of two α-helixes which are H1and H3 (both labeled green). The β-strands in the Greek key β barrel topology is normally connected by loops, however, in this prediction the connections between E2 and E3 are connected by loop-helix-loop structure, while other β-stands connections are made by loops labeled wheat in this diagram.
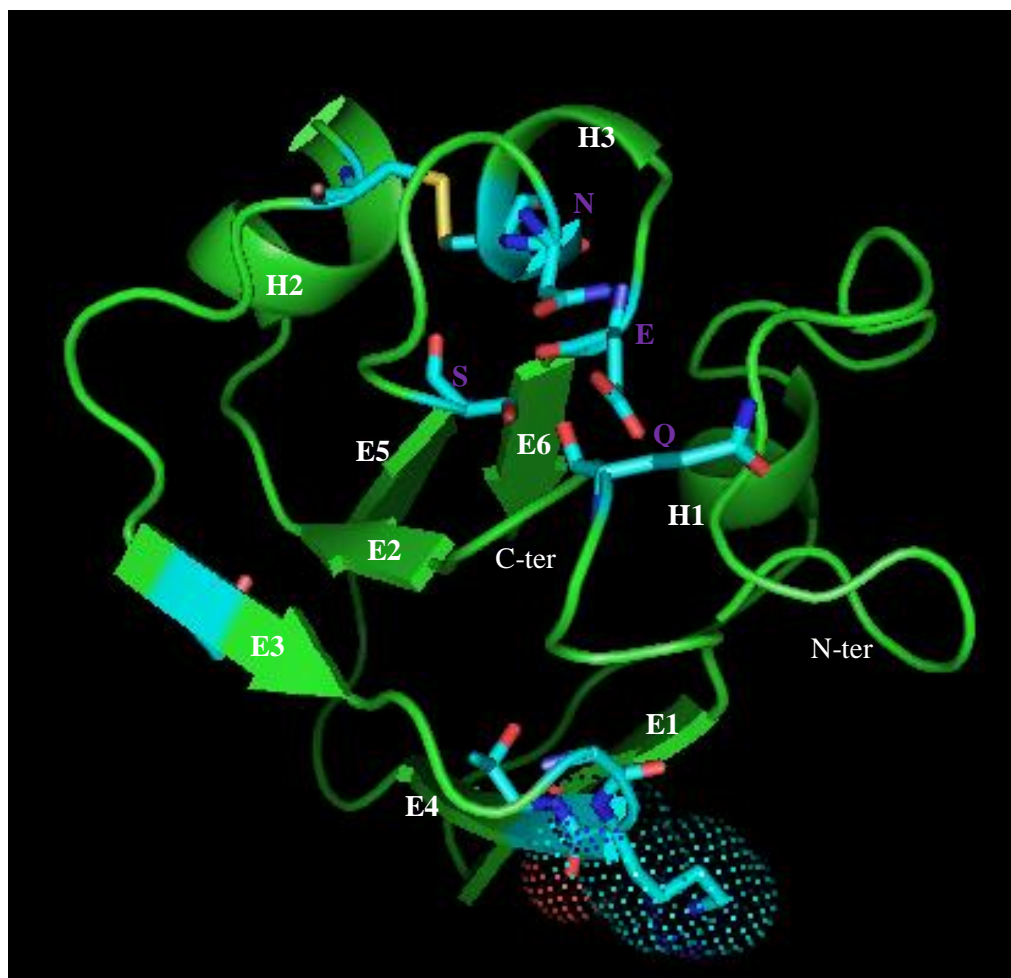
Figure 2.4. Ribbon diagram of the predicted functional regions and motifs of SE-ENOD domain II viewed above the theoretical copper binding center. The atoms of functional residues, two highly conserved cysteines, four substituted copper binding ligands and Walker A motif, are displayed in stick model with atoms labeled as carbon-cyan, hydrogen-wheat, nitrogen-blue, oxygen-red and sulfur-orange. The disulfide bond formed by Cys^{57} and Cys^{91} is labeled yellow. The four substituted copper binding ligands, Gln^{43}, Ser^{85}, Asn^{90} and Glu^{95} are indicated by their single letter code in purple, Q, S, N and E, respectively. The residues of Walker A motif located on the third β-strand (E3) and fourth β-strand (E4). The important catalytic residue Lys^{70} of the Walker A motif on E4 is indicated by dots.
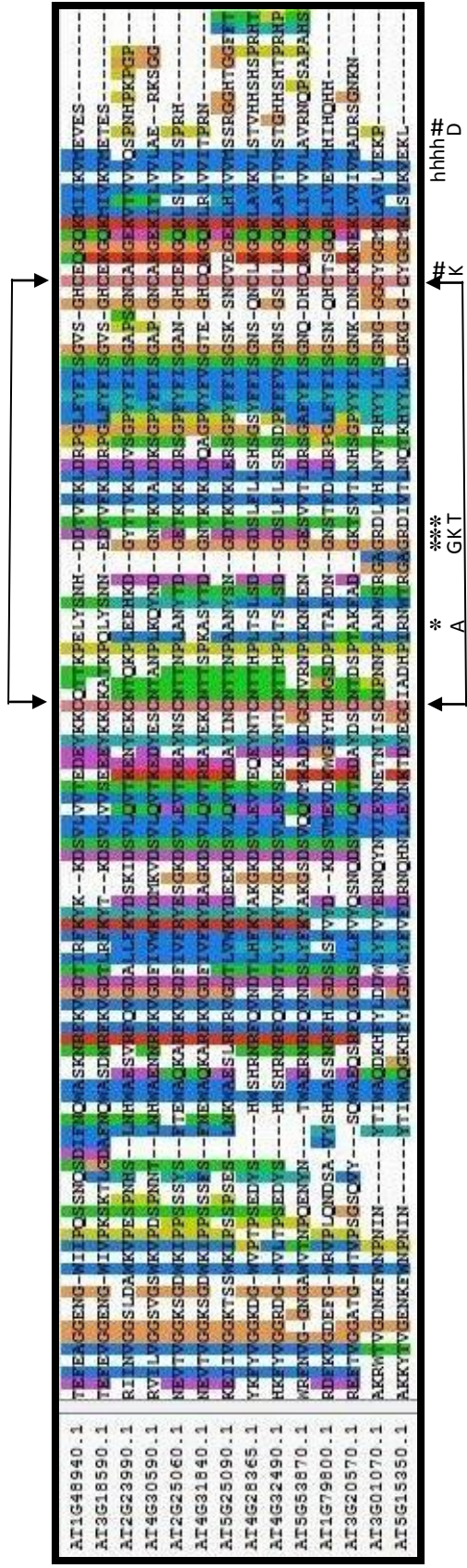
Figure 2.5 Multiple sequence alignment of domain II for fourteen *Arabidopsis* ENOD-like proteins using ClustalX. The residues have been assigned to different color according to default parameter by Clustal X. The sequence of SE-ENOD is indicated by an arrow. The highly conserved cysteines are connected by bracket. The Walker A motif residues, Ala^{64}, Gly^{69}, Lys^{70} and Thr^{71} are indicated by " * " and their single letter code, the Walker B motif residues, Lys^{92} and Asp^{104}, are indicated by " # " and their single letter code, the four hydrophobic residues, Ile^{100}, Val^{101}, Met^{102} and Ala^{103} are indicated by " h ".

32

2.3.3 SE-ENOD functional regions and complex prediction

As shown in Figures 2.4 and 2.5, the SE-ENOD contains a Walker A (P-loop) ATP-binding motif and Walker B magnesium-binding motif (Walker et al., 1982). Potential ATP binding pockets were predicted by an online based program Pocket-Finder (Hendlich et al., 1997), and the result was viewed and analyzed by PyMOL. This analysis predicted a potential ATP binding pocket in the SE-ENOD (Figure 2.6) established by three criteria. Firstly, the distance between the potential pocket and the Walker A motif residues should be suitable for binding an ATP molecule. Secondly, the diameter of the pocket needs to fit the size of ATP molecule. Thirdly, the prior potential pocket should be located on or near the surface of the SE-ENOD to allow accessibility of ATP to the binding site. Due to the length limitation of the mavicyanin sequence used for homology modeling, the full residues in Walker B motif were not included in the predicted tertiary structure. Therefore, the pocket for the magnesium ion is not predicted in this structure.

The SE-ENOD and mavicyanin are phytocyanins belonging to a highly conserved protein family. The dimer complex form of the SE-ENOD is predicted by a protein-protein docking server with suggestions from the tetrameric mavicyanin complex structure. An online-based docking prediction was performed by ClusPro. Most of the twelve docking complexes were filtered and one potential complex form was targeted as the most promising prediction (Figure 2.7). The criteria used for this filtering is mainly the potential ionic or hydrogen bond interaction formed by the interface residues and the hydrophobicity of the interface surface (Figures 2.7 and 2.8). Because relatively strong interactions need to be formed at the adjacent interface to guarantee a stable complex formation, the best prediction showed the interactions at the β-strand or α-helix interface

and the potential interacting residues located on their interface are labeled (Figure 2.7). It is possible that Lys^{65} from one SE-ENOD and Tyr^{81} from the other SE-ENOD, Thr^{64} from one SE-ENOD and Gly^{79} from the other SE-ENOD, Arg^{51} from one SE-ENOD and Ser^{61} from the other SE-ENOD, form hydrogen bonds at the interface of the predicted complex. An ionic interaction possibly forms between Arg^{51} from one SE-ENOD and Asp^{52} from the other SE-ENOD (Figure 2.7). More importantly, hydrophobic interactions are very critical for protein complexes, therefore, the electrostatic potential map on the overall dimer complex and their interface were analyzed and revealed that the major interaction between the two SE-ENOD single molecules is hydrophobic (Figure 2.8).
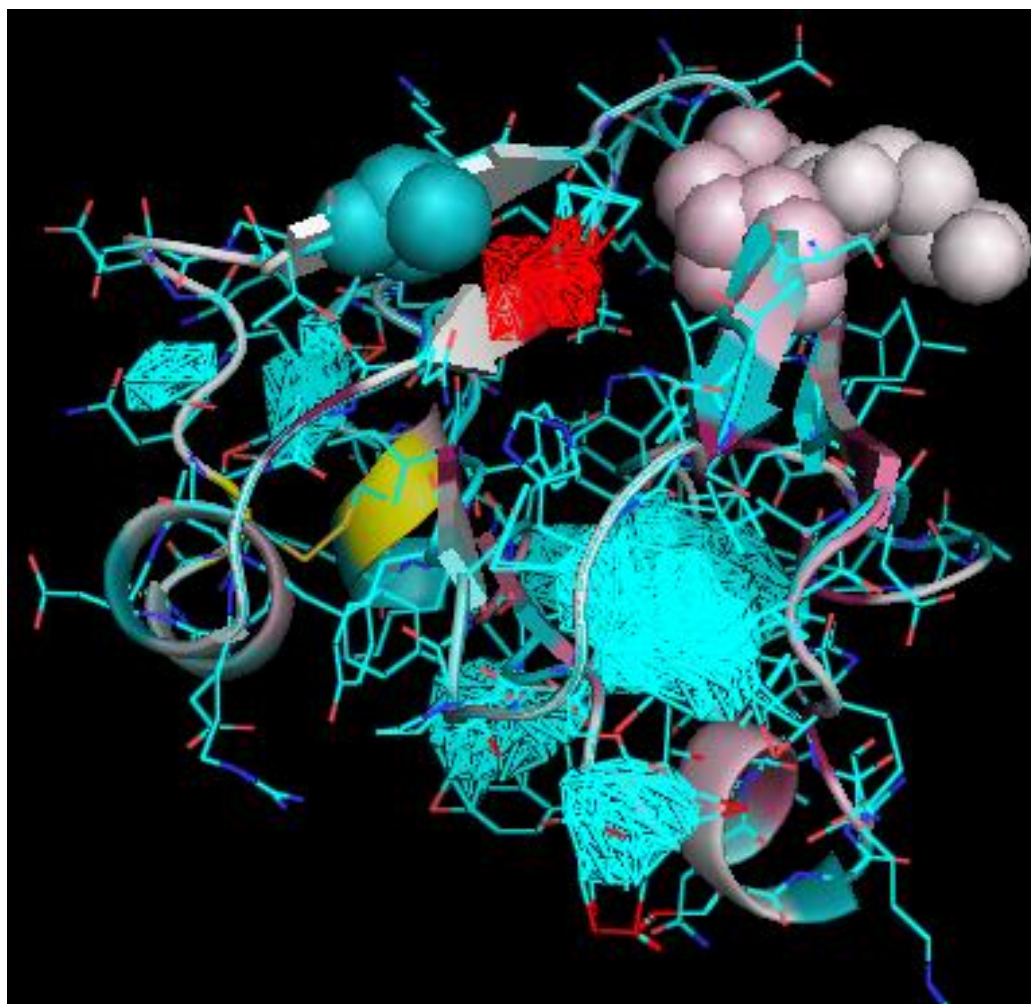
Figure 2.6. Ribbon diagram of the SE-ENOD predicted tertiary structure and potential pocket for ATP binding viewed from the side of the molecule. The molecule is displayed by both line and cartoon models and atoms are labeled as carbon-cyan, hydrogen-wheat, nitrogen-blue, oxygen-red and sulfur-orange. Each residue, except Cys[57] and Cys[91] labeled in yellow, is assigned to one color according to its conserved property (variable-cyan→neutral-white→conserved-dark purple). The residues of Walker A motif, $AX_{(4)}GKT$ are displayed by a spheres model. All the predicted pockets in the SE-ENOD are labeled cyan, except the best candidate for ATP binding labeled in red.
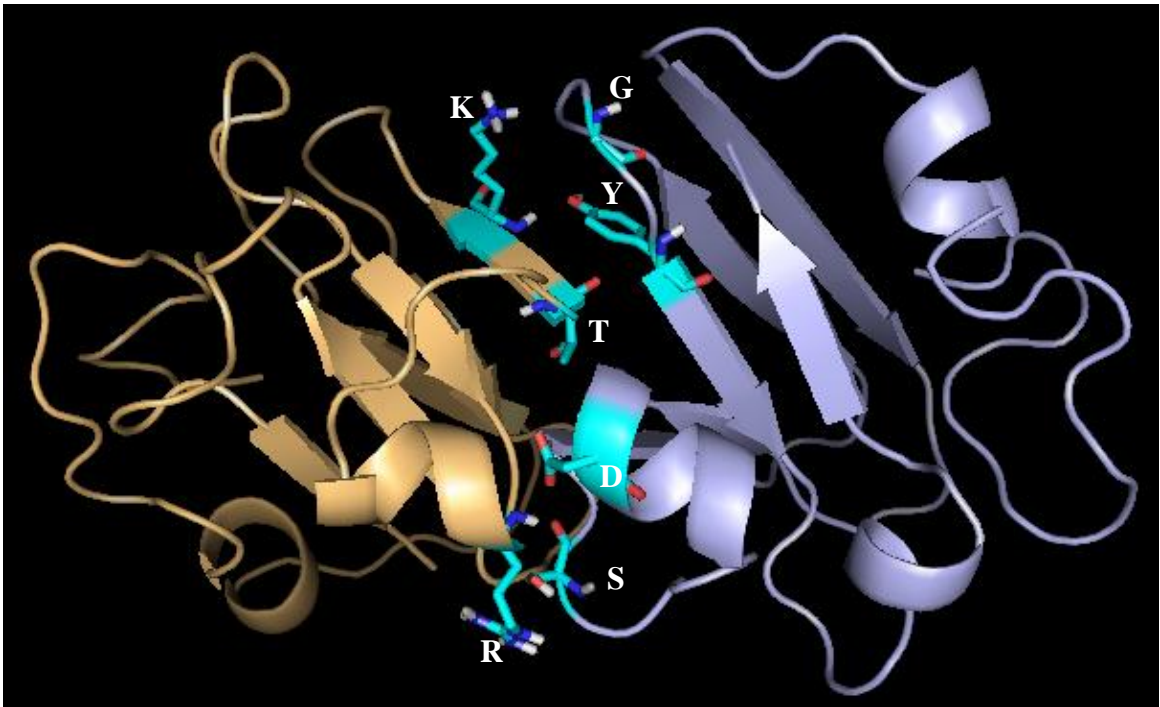
Figure 2.7, Ribbon diagram of the predicted SE-ENOD dimer complex viewed from one side of the molecules. The two single SE-ENOD molecules were labeled light orange on the left and light purple on the right. The residues, Lys[65], Tyr[81], Gly[79], Thr[64], Asp[52], Ser[61] and Arg[51], from different molecules having potentials to interaction with each other are showed in stick model with atom labeled as carbon-cyan, hydrogen-wheat, nitrogen-blue, oxygen-red and sulfur-orange.
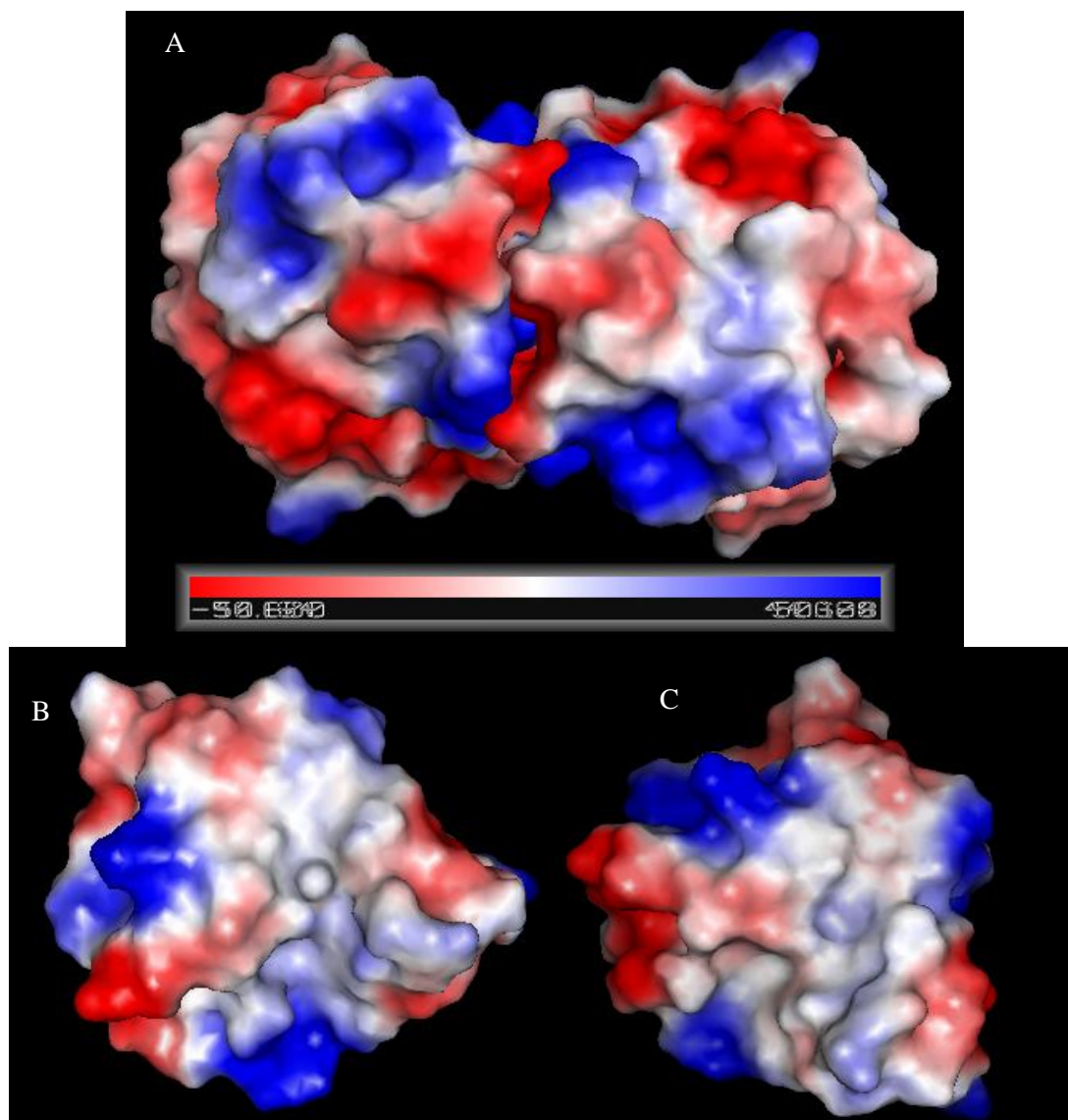
Figure 2.8. Electrostatic potential surface of the predicted docking model on the SE-ENOD viewed from the side of the complex. The figure was generated by PyMOL using the relative vacuum electrostatics option contoured from ~ -50 (red) to ~ +50 (blue). Panel A shows the electrostatic surface of the overall dimer complex. Panel B shows the electrostatic interface of the left molecule. Panel C shows the electrostatic interface of the right molecule.

2.4 DISCUSSION

In this chapter, bioinformatic approaches were applied to predict and characterize the structure of the *Arabidopsis* sieve element-specific early nodulin-like protein (SE-ENOD). The secondary structure predictions suggested that the protein structure was composed of six strands, two or three helixes and a few random coils. Further SE-ENOD tertiary structure prediction indicated four β-strands arranged into a core Greek key β-barrel organizing into another two antiparallel β-strands, termed as a cupredoxin fold that is a typical core structure in phytocyanins (Roberts et al., 2002). Phytocyanins in higher plants such as plastocyanins, mavicyanins and stellacyanins bind copper and participate in electron transport reactions (Gewirth and Solomon, 1988; Hart et al., 1996; Xie et al., 2005). The ENOD5-type subfamily of the phytocyanins, however, adopt the cupredoxin fold structure, but have substituted the copper binding ligands and have theoretically lost the electron transport function (Greene et al., 1998).

ENODs are a heterogeneous group of plant proteins that are normally expressed in response to lipo-chitooligosaccharide Nod factors synthesized by rhizobia at the initial stages of nodulation in legumes (Combier et al., 2007). Multiple sequence alignment with the *Arabidopsis* ENODs reveals that the domain II is highly conserved among the family members (Figure 2.5). It is remarkable that 95% of the extremely conserved residues such as Ser[45], Val[46], Leu[47], Val[49], Tyr[82], Phe[83], Ile[84], Ser[85], Gly[86], Val[99], and Val[101] in the SE-ENOD are hydrophobic residues within the cupredoxin fold and are probably involved in forming elements of the secondary structure. Typical of the ENOD subfamily of phytocyanins, the two conserved cysteines formed a putative disulfide bond in the predicted tertiary structure. This disulfide bond has been proposed to be essential in

stabilizing the twisted adjacent β-strands (Guss et al., 1996). In Figure 2.4, this disulfide bond connects the second helix (H2) and third helix (H3) closing the distance between the second β-strand (E2) and the fifth β-strand (E5), that is important to form the Greek key β-barrel topology. This topology requires the fourth β-strand (E5 in the SE-ENOD predicted structure) to be located adjacent to the first β-strand (E2 in the SE-ENOD) in a reverse orientation (Figure 2.3 and 2.4).

Arabidopsis ENOD-like genes are expressed in highly-specialized cell types including phloem sieve elements as well as the male and female gametophytes in reproductive tissues (Kouchi and Hata, 1993; Scutt et al., 2003). The presence of Walker A (P-loop) and Walker B motifs (Figure 2.5) only in the SE-ENOD suggests that the protein could have evolved a unique function among the *Arabidopsis* ENODs. The Walker A motif interacts with the phosphate groups of the nucleotide in the presence of magnesium ions. The functional residues on Walker A motifs (Figure 2.4) are predicted to be located on the surface of the protein with Lys^{70} side chain exposed to the solvent area and the candidate pocket was located near the ATP binding center (Figure 2.6). The combined results suggest the SE-ENOD could perform an alternative function that involves the binding of ATP in the vascular tissue.

Sieve elements are terminally differentiated cells with a well-developed system of endomembranes composed primarily of the sieve element plasma membrane and sieve element reticulum. In light of the highly modified anatomy of sieve elements, the relationships of membrane proteins could serve critical roles in regulating physiological interactions of these cells. It has been proposed that microdomains, also called lipid rafts, play an important role in organizing proteins within the plasma membrane of sieve

elements. The targeting of proteins to lipid raft microdomains can increase the local protein concentration and consequently facilitate their interaction, including homo- or hetero-oligomeric interactions (Grennan, 2007). There is experimental evidence indicating that the SE-ENOD is a glycosylphosphatidylinositol-anchored protein (GPI-anchored protein) (Khan et al., 2007). The association of GPI-anchored proteins with lipid rafts and the importance of these interactions has been the topic of considerable research (Sangiorgio et al., 2004). GPI-anchored proteins could play an important role in stabilizing the position of lipid rafts by maintaining interactions with components of the plant cell wall. On the other hand, the lipid environment could enhance protein activities by influencing their conformation (Garcia et al., 2003). It has yet to be demonstrated whether the SE-ENOD co-localizes with integral plasma membrane proteins to affect their function. Recent evidence has shown that the dimer form of the SUT1 sucrose transporter in *Solanum tuberosum* was present in lipid rafts and the activity of sucrose uptake was enhanced by this form. It was proposed that the significant increase in activity was a result of conformational changes induced by oxidants with inter- or intramolecular disulfide bridges as the major contributors. This oxidative stimulation can positively regulate the flexibility of the SUT1, and stimulation can be abolished by thiol modifying agents (Krugel et al., 2008). A similar regulatory mechanism could affect the SE-ENOD because it also contains structurally important disulfide bridges.

Recent studies revealed that a complex protein profile of 270 proteins had been identified by proteomic approach in lipid rafts of *Medicago truncatula* root plasma membrane, including proteins such as H^+-ATP synthase and voltage-dependent anion channels (Lefebvre et al., 2007). Significantly, a complete redox protein system was

shown to exist in the *Medicago truncatula* root plasma membrane lipid rafts. Several putative cytosolic NADH-ubiquinone oxidoreductases, apoplasm-facing L-ascorbate oxidase and peroxidases had been detected (Lefebvre et al., 2007). Based on the electrostatic potential distribution of the predicted SE-ENOD dimer, several charged hydrophilic residues were regularly distributed on the surface (Figure 2.8). It is possible that the solvent-exposed side chains of these residues are involved in interactions with other lipid raft proteins. These interactions could provide an alternative function of SE-ENOD in sieve element plasma membrane and further build a cohesive picture of membrane-associated protein-protein interactions in combination with signal transduction pathways involved in phloem long distance transport.

Based on the structural predictions, it is reasonable to conduct further investigations to determine whether the SE-ENOD is involved in protein-protein interactions. Empirical studies to resolve the structure by X-ray crystallography or nuclear magnetic resonance are required to validate the predicted structures and docking models for the SE-ENOD.

CHAPTER III

FUNCTIONAL STUDIES OF SE-ENOD

3.1 INTRODUCTION

Sieve elements, with their highly modified cellular anatomy, rely on both the intimately associated companion cells as well as a highly organized endomembrane system to transduce information with functional consequences. The SE-ENOD is a plasma membrane associated protein with several structural characteristics that are consistent with a function as an information transducing molecule. The protein is post-translationally modified by the covalent attachment of a glycosylphosphatidylinositol-anchor (GPI-anchor), which is thought to attach the protein to the extracellular leaflet of the plasma membrane (Khan et al., 2007). Previous studies indicate that GPI-anchored proteins play important roles in connecting the communication of cell exterior and interior through their function as enzymes or receptors in the extracellular matrix (Borner et al., 2002; Elortza et al., 2003; Sardar and Showalter, 2007). Recently, a GPI-anchored arabinogalactan-protein was shown to structurally interact with transmembrane callose synthase and produce downstream effects on class VIII myosin in *Tobacco* BY-2 cells (Sardar and Showalter, 2007).

The mature SE-ENOD is composed of two domains, the plastocyanin-like copper binding domain II and the serine/proline-rich domain III. The proline/serine-rich domain III, which is GPI anchored to the plasma membrane, could serve as a flexible hinge for the globular domain II. As is the case for other members of the cupredoxin superfamily, domain II is predicted to adopt an overall cupredoxin fold structure. Amino acid substitutions of the four copper binding ligands in domain II, however, suggest that like other members of the ENOD subfamily, the protein does not bind copper and has undefined physiological functions other than electron transport. The presence of Walker A (P-loop) and B motifs in this domain suggests that the protein could serve an alternative function as an ATP-binding molecule within the extracellular matrix. Structural modeling experiments of domain II described in Chapter II predicted that the functional residues of Walker A motif are solvent exposed within a suitable ATP sitting pocket. Interestingly, recent studies revealed that a putative ATP-binding protein located in the extracellular leaflet of *Arabidopsis* root plasma membrane performed the functions of regulating extracellular ATP and initiating plant wound responses (Demidchik et al., 2009).

The regulatory roles of extracellular nucleotides in plant signaling responses are poorly studied, despite the observations that these signaling mechanisms exist in very primitive plants. In contrast, the regulatory function of extracellular nucleotides in animals is well established, and animal studies provide good examples to guide investigations of extracellular nucleotide signaling in plants. For example, purinoceptors in green algae are structurally similar to mammalian receptors and were shown to regulate the eATPs that function as wound signals (Burnstock, 1996). It is now believed

that plant responses to extracellular ATP, ADP, and analogues, such as ATPγS and ADPβS, occur primarily through fluxes in cytoplasmic calcium (Jeter and Roux, 2006). There are several examples demonstrating the importance of calcium fluxes in regulating physiological events within mature sieve elements (Eckardt, 2001).

The structural modeling studies described in Chapter II also predicted that SE-ENOD subunits interact either as a homodimer or with other proteins primarily through hydrophobic interactions on the surface of domain II. The GPI anchored proteins, like the SE-ENOD, are often located in lipid raft microdomains placing them in proximity to other proteins that are associated with the sieve element plasma membrane. Conformational changes induced by this specialized lipid environment could influence such intra- or inter-molecular interactions. For example, the SUT1 sucrose transporter in *Solanum tuberosum* showed elevated sucrose transport activity as a dimer when targeted to lipid rafts (Krugel et al., 2008). SE-ENOD interactions could be a function of its conformation in the highly specialized environment within sieve element plasma membranes.

Microbial expression systems, including *Escherichia coli* (*E. coli*) and *Sacchromyces cerevisiae,* are well-established molecular tools to generate recombinant proteins for functional and mutagenesis studies (Frommer and Ninnemann, 1995; Rajagopala and Uetz, 2009). In this chapter, wild-type and mutated versions of the SE-ENOD domain II were expressed in *E. coli* to test whether the ATP-binding motifs are functional. The ability of domain II to form homodimers as predicted *in silico* was analyzed in yeast two hybrid experiments.

3.2 MATERIALS AND METHODS

3.2.1 Amplifying wild type and mutated versions of SE-ENOD domain II

The 336 bp DNA sequence encoding the SE-ENOD domain II extending from amino acid Arg^1 to Ala^{112} was PCR amplified from the RS6-ORF full-length cDNA (NCBI accession no. AAP88350) in pCDNA3.0 vector with primers No.1 and No.2 (Table 3.1). The template and primers were added to PCR Master Mix (Promega, catalog no. M7502), and PCR amplification was performed following a standard program (94°C 2 min: 35 cycles of 94°C, 30 sec - 57°C, 30 sec - 72°C, 30 sec: 72°C, 10 min).

The Walker A motif $A^{64}X_{(4)}G^{69}K^{70}T^{71}$ in domain II was mutated to $T^{64}X_{(4)}A^{69}A^{70}A^{71}$ through a series of PCR reactions. The 336 bp wild-type domain II amplicon generated in the initial experiment was used as a template for two PCR reactions using the conditions as described above. Segment I was amplified using primers No. 1 and 3 (Table 3.1) resulting in a 237 bp DNA fragment that encoded the equivalent peptide extending from Arg^1 to Gly^{79}. Segment II was amplified using primers No. 4 and 5 (Table 3.1) resulting in a 171 bp DNA fragment that encoded the equivalent peptide extending from Cys^{56} to Ala^{112}. The mutated domain II sequence was obtained by a third PCR reaction that amplified a 336 bp DNA sequence from segments I and II as templates and primers No. 1 and 5 (Table 3.1). PCR reactions were conducted using the conditions previously described.

3.2.2 Amplifying ATP binding domain from *E. coli* DnaA

A 717 bp DNA sequence encoding the ATP binding domain, with both Walker A and B motifs, from the chromosomal replication initiator DnaA protein of *E. coli*

BL21(DE3) was PCR amplified using High Fidelity Platinum Taq DNA Polymerase kit (InVitrogen, catalog no. 11304-011). Genomic DNA from an *E. coli* BL21(DE3) single colony culture served as the template for the PCR reaction using the standard program (94°C 2 min: 35 cycles of 94°C, 30 sec - 57°C, 30 sec - 72°C, 60sec: 72°C, 10 min) with primers No. 6 and 7 (Table 3.1).

3.2.3 Constructs, bacterial strains and culture conditions for *E. coli*

The amplicons encoding the SE-ENOD wild type and mutated domain II as well as the DnaA ATP binding domain were individually inserted into the 5' *Nde*I and 3' *Xho*I restriction sites of the modified pET28 vector (pSKB$_3$) containing an N-terminal 6xHis tag and a *Tobacco etch virus* (TEV) cleavage site (ENLYFQG). Plasmids containing the recombinants were inserted into competent *E. coli* BL21(DE3) cells using the Electroporator 2510 (1800 V).

Recombinant proteins were expressed in 1 liter of *E. coli* BL21(DE3) cells grown in Luria Bertani (LB) medium containing 50µg/ml kanamycin (Sigma, EC no. 246-933-9) at 18°C on a rotary shaker (250 rpm) for 22 hours. Growth of the cultures was monitored by measuring absorbance at A_{600} and were induced by the addition to a final concentration of 1 mM IPTG (GOLDBIO, catalog no. I2481C25) when cultures had reached an absorbance of A_{600}=0.8 to 1.0. Initial induction experiments were conducted under similar growth conditions at 37°C for a total of 4 hours. The induced cells were harvested by centrifugation at 5,000 rpm for 25 minutes with a JLA 9.1000 rotor in Avanti J-E centrifuge (Beckman Coulter) and the pellets were stored at -80°C (Krumm et al., 2008).

3.2.4 Recombinant protein purification

The induced *E. coli* cell pellets were resuspended with 30 ml of 1% sarkosyl (Catalog No.61207-5000; Acros, Morris Plains, NJ, USA) in buffer A (50 mM Tris, pH 7.9, 500 mM NaCl, 25 mM imdizole and 10% glycerol) by rocking at room temperature for 30 minutes. The cells were lysed by sonication (550 Sonic Dismembrator, Fisher Scientific) with two bursts of 3 seconds on, 3 seconds off for a total time of 1 minute. After sonication, the cell lysates were centrifuged at 16,000 rpm for 30 minutes using a JA-20 rotor in Avanti J-E centrifuge (Beckman Coulter) at 4°C. Supernatants (30.25 ml) were transferred into a new 50 ml tube. CHAPS (J.T. Baker Inc, catalog no. 4145-00) and Triton X 100 (Fisher Biotech, catalog no. CAS9002-93-1) were added to the final concentration of 20mM CHAPS and 2% Triton X 100. The 35.25 ml mixtures were rocked at room temperature for 30 minutes and loaded onto 500 µl of resuspended Ni-nitrilotriacetic acid (NTA) (Qiagen, catalog no.1018611) affinity column matrix. Non-specific background proteins were washed off with 800 ml of buffer A. Finally, 6xHis-tagged recombinant proteins were eluted by buffer B (20 mM Tris, pH 7.9, 500 mM NaCl, 250 mM imidizole and 10% glycerol) at 4°C (Krumm et al., 2008; Tao et al., 2010).

The proteins were further purified by high-performance liquid chromatography (HPLC). The Ni-NTA purified recombinant proteins were centrifuged at 14,000 rpm for 15 minutes at 4°C in a microfuge to clarify the samples of insoluble aggregates. The Superdex 200 (GE Healthcare, catalog no. 17-5175-01) size exclusion column was equilibrated by loading pre-chilled equilibrium buffer (20 mM Tris, pH 7.9, 500 mM NaCl and 10% glycerol) onto the AKTA Explorer HPLC (Pharmacia Biotech). Cleared

samples were injected into the HPLC with a 1ml syringe and separated using the program "S200 without equilibrium" with a flow rate of 0.3 ml/min. Samples were collected in 42 fractions representing different molecular weights. Proteins (10 µl) from even numbered fractions from 16-40 were visualized by 4-20% SDS-PAGE, and the recombinant proteins were present in fractions 30-32. Fractions 30, 31, and 32 (500 µl each) were combined and concentrated with a 10 kDa molecular weight cut off Amicon Ultra centrifugal filter device (MILLIPORE, catalog no. UFC901008) to a final concentration of 0.5 mg/mL (Krumm et al., 2008).

3.2.5 Nitrocellulose filter ATP-binding assay

A nitrocellulose filter binding assay suitable for detecting ATP-protein interactions was adapted from Sekimizu and co-workers (1987). Briefly, the negatively charged nitrocellulose filter binds and immobilizes proteins with a net positive charge. Proteins binding [α-^{32}P]ATP will be retained on the filter and bound radioactivity can be measured in a liquid scintillation counter.

Individual reactions containing 3 picomoles or 30 picomoles of recombinant wild-type domain II, mutated domain II, DnaA ATP binding domain, and BSA were incubated with 1 mM [α-^{32}P]ATP (25 Ci/mmol) (MP BIOMEDICALS, LLC, serial no. 427121) at 0°C for 15 minutes in 25µl of buffer C (50 mM Tricine-KOH pH 8.25, 0.5 mM magnesium acetate, 0.3 mM EDTA, 20% glycerol (v/v), 0.007% Triton X-100, and 7 mM dithiothreitol). Individual reactions were filtered through a glass filter apparatus containing a 0.22 µM nitrocellulose membrane (MILLPORE, catalog no. GSWP02500) that had been presoaked in buffer D (50 mM Tricine-KOH pH 8.25, 10 mM (NH$_4$)$_2$SO$_4$,

0.5 mM magnesium acetate, 0.3 mM EDTA, 5 mM dithiothreitol, 17% glycerol (v/v), 0.005% Triton X-100). Filters were washed with 30 ml of buffer D to remove any unbound $[\alpha\text{-}^{32}P]$ATP. Radioactivity retained on the filters was measured in a Multi Purpose Scintillation Counter (Beckman Coulter) in 4 ml Ecolume (MP BIOMEDICALS, LLC, catalog no. 882470) (Sekimizu et al., 1987). One replicate was conducted for each of three experiments. The data were analyzed with the ^{32}P radioactivity counts recorded by liquid scintillation counter in Microsoft Excel. The change in ATP binding was calculated by dividing the difference in the ^{32}P radioactivity counts between the high (30 picomole) and low (3 picomole) protein concentration by the difference in the protein concentration (27). The average counts per picomole protein and the standard errors of the three replicates were calculated and shown in Figure 3.3.

3.2.6 Yeast two hybrid analysis

A 336 bp DNA sequence encoding the SE-ENOD domain II was amplified by PCR using the previously described wild type domain II PCR amplicon as the template. The template DNA and primers No. 8 and 9 (Table 3.1) were added to the PCR Master Mix (Promega, catalog no.M7502) and amplification was conducted using the standard program (94°C, 2 min: 35 cycles of 94°C, 30 sec - 57°C, 30 sec - 72°C, 30 sec: 72°C, 10 min). A 456 bp DNA sequence encoding the SE-ENOD domain II/III extending from amino acid Arg[1] to Asn[152] was PCR amplified with the RS6-ORF full-length cDNA (NCBI accession no. AAP88350) in pCDNA3.0 vector as the template using the same conditions with primers No.10 and No.11. The template and primers were added to PCR Master Mix (Promega, catalog no. M7502), PCR amplification was performed following

a standard program (94°C 2 min: 35 cycles of 94°C, 30 sec - 57°C, 30 sec - 72°C, 30 sec: 72°C, 10 min).

The domain II and domain II/III amplicons were inserted into the 5' *Eco*RI and 3'*Bgl*II restriction sites of the plasmid pRM1151 *GAD-C1(LEU)* and pRM1154 *GBDU-C1(URA)*. The recombinant pRM1151 plasmid and pRM1154 plasmid were mixed with the *S. cerevisiave MATa* PJ69-4A *(*yRM1757*)* and *MATα* PJ69-4α (yRM1756) strains, respectively, in 40% PEG transformation buffer and vortexed for 1 minute followed by incubation for 1 hour at 30°C on a roller drum. The cells were then heat shocked for 10 minutes at 42°C and pelleted by centrifugation at 3,000 rpm for 3 minutes in a microfuge. Cells were resuspended in sterile ddH$_2$O and plated on selective plates.

The yRM1756 containing pRM1154 plasmids and yRM1757 strains containing pRM1151 plasmids were mated on yeast-peptone-dextrose (YPD) plates over night at 30°C. Diploids were selected on synthetic complete (SC) plates lacking both uracil and leucine over night at 30°C. Diploids were transferred with a multiprong transfer device to SC plates lacking both uracil and leucine (-ura/-leu), histidine (-his) and adenine (-ade). Interactions were detected by scoring the growth of diploid colonies on the respective plates after incubation at 30°C for 2-3 days.

Table 3.1. DNA sequences of primers used in PCR reactions for microbial expression experiments[1].

No.	Primers	Oligonucleotides
1	II-wt-forward	5'-GGGTTTCATATGAGAGAGTTTACGGTCGGTG-3'
2	II-wt-reverse	5'-CCGCTCGAGTTATGATGCGGTGTTTTTG- 3'
3	II-mut-1-reserve	5'-GCCCAGAGTGGTTCAGCGTGACAGAAGCTGCAGC ATCGGCGAATTTAGTGGTTGGTGAGTCCGTGTTGCAG-3'
4	II-mut-2-forward	5'-CTGCAACACGGACTCACCAACCACTAAATTCGCC GATGCTGCAGCTTCTGTCACGCTGAACCACTCTGGGC-3'
5	II-mut-reverse	5'-CCGCTCGAGTCATGATGCGGTGTTTTTG-3'
6	DnaA^{ATP}-forward	5'-GGGTTTCATATGAAACACACGTTTGATAAC-3'
7	DnaA^{ATP}-reverse	5'-CCGCTCGAGTCACAGTTTTTCCTGCAATGCC-3'
8	Y2H-II-forward	5'-CCGGAATTCAGAGAGTTTACGGTCGGTG-3'
9	Y2H-II-reverse	5'-GGAAGATCTTCATGATGCGGTGTTTTTG-3'
10	Y2H-II/III-forward	5'-CCGGAATTCAGAGAGTTTACGGTCGGTG-3'
11	Y2H-II/III-reverse	5'-GGAAGATCTTCAGTTTGGAGTGTCTT-3'

[1]Primers were diluted to a final concentration of 0.5μM for individual PCR reactions.

3.3 RESULTS

3.3.1 Domain II ATP binding assays

The ENOD subfamily of cupredoxins represents a significant functional divergence from the single electron transport reactions that typify other members of the superfamily. While the function of members of the ENOD5 subfamily is unknown, it is of interest to note that this class of proteins is found in highly specialized tissues such as nodules, gametophytes, and sieve elements of the phloem. ENODs appear to have lost the ability to bind copper, essential for electron transport, but retain the structural copper-binding pocket (see Chapter II). Interestingly, the *Arabidopsis* SE-ENOD has evolved signature motifs of the well characterized Walker A (P-loop) and B nucleotide binding sites located near the surface of domain II. Significantly, the SE-ENOD is the only member of the *Arabidopsis* ENOD-like protein family to contain complete Walker A and B domains, although, other members contain partial motifs (see Chapter II).

Experiments were conducted to determine whether the domain II of the SE-ENOD was capable of binding ATP. As a true negative control, the Walker A motif $A^{64}X_{(4)}G^{69}K^{70}T^{71}$ in domain II was mutated to $T^{64}X_{(4)}A^{69}A^{70}A^{71}$. Both SE-ENOD wild-type domain II and mutated domain II are 112 amino acids in length with a calculated molecular mass of 12 kDa. The recombinant proteins included a 6xHis-tag and a TEV cleavage site extending the length to 138 amino acids with a calculated molecular mass of 15 kDa. In an induction test conducted with cultures grown at 37°C, a fusion protein with an apparent molecular weight of 16 kDa was induced within one hour after the addition of IPTG that accumulated to a maximum amount within three hours (Figure 3.1A).

Similar results were obtained for both the wild type and mutated domain II recombinant proteins. The induced protein was confirmed as SE-ENOD domain II by immunoblotting with the RS6 monoclonal antibody (Figure 3.1B).

The recombinant proteins primarily accumulated within insoluble inclusion bodies in *E. coli* cells grown at 37°C. Initial attempts to isolate soluble domain II peptides using native Ni-NTA purification resulted in co-purification of multiple *E. coli* proteins. Because the mutated domain II protein serves as a true negative control, colorimetric ATPase activity assays (QuantiChrom^{TM} ATPase/GTPase Assay Kit, BioAssay Systems, catalog no. DATG-200) were conducted using the partially purified wild type and mutant domain II (data not shown). ATPase activity assays showed similar ATPase activity (16 micromoles ATP hydrolyzed by 1 microgram protein) for both wild type and mutated fusion domain II samples. The ATPase activity was attributed to endogenous *E. coli* proteins that co-purified with the recombinant domain II proteins.

An increased amount of soluble recombinant proteins was obtained by growing the *E. coli* cultures overnight at 18°C. Furthermore, recombinant proteins were isolated from the *E. coli* cultures following the procedures for difficult to solublize proteins described by Tao et al. (2010). This proceedure utilizes sarkosyl in the isolation buffer to gently enhance the solublization and recovery of recombinant proteins that readily form inclusion bodies. The initial native Ni-NTA affinity purification step resulted in considerable enrichment of the domain II recombinant proteins (Figure 3.2, lanes 2 and 5). The proteins were further purified by HPLC size exclusion chromatography (Figure 3.2, lanes 3 and 6).

The initial colorimetric assays to assess ATPase activity were replaced with a radioactive-ATP binding assay as a more sensitive and directed test to determine whether domain II binds ATP. The ATP binding domain from the chromosomal replication initiator DnaA protein of *E. coli* BL21(DE3) contains both Walker A ($GX_{(4)}GKT$) and B motifs ($KX_{(7)}hhhhD$) that are very similar to the Walker A ($AX_{(4)}GKT$) and B ($KX_{(7)}hhhhD$) motifs in the SE-ENOD domain II. The 239 amino acids domain is of similar size to domain II and thus, served as a good positive control for the [α-^{32}P]ATP-binding assays. The mutated domain II recombinant protein served as the true negative control, and bovine serum albumin (BSA) served as a non-specific protein negative control. ATP binding was assessed as the difference in the amount of radioactivity in binding reactions with 3 picomoles and 30 picomoles of test protein in three replicated experiments. As shown in Figure 3.3, the binding of ATP by the wild type SE-ENOD domain II (162 counts/picomole) was about half of the binding recorded for the DnaA positive control (312 counts/picomole). Both negative controls, the mutated SE-ENOD domain II (-7.6 counts/picomole) and BSA (3.58 counts/picomole) showed minimal differences in ATP binding between the two protein concentrations. These results indicate that domain II of the SE-ENOD is capable of binding ATP, however, these experiments are preliminary and additional experiments must be conducted to fully attribute this function to the SE-ENOD.
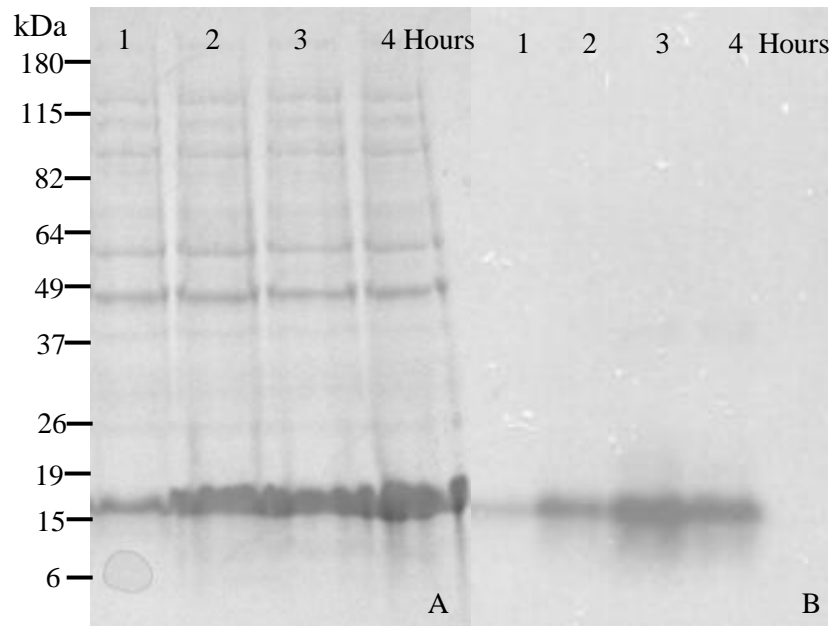
Figure 3.1. Time course analysis of the recombinant SE-ENOD domain II expressed in *E. coli* and immunoblotting with the RS6 monoclonal antibody. Expression of the recombinant SE-ENOD domain II was induced by addition of 1mM IPTG in *E. coli* BL21(DE3) cells cultured at 37°C. Aliquots of 1 ml were collected at 1, 2, 3, and 4 hours following the addition of IPTG. The cells were collected by centrifugation and lysed in a buffer containing 20 mM Tris pH 8.0, 8 M Urea and 2% SDS. Molecular weight standards in kDa are shown at the left. Panel A. Proteins in the *E. coli* cell lysate (20 µL) were separated on a 4-20% gradient SDS-polyacrylamide gel and stained with Coomassie Blue. Panel B. Immunoblot of duplicate SDS-PAGE of proteins in *E. coli* cell lysate (10 µl) detected with the RS6 monoclonal antibody (1:5000 dilution). The time points for each sample are indicated in both panels A and B.
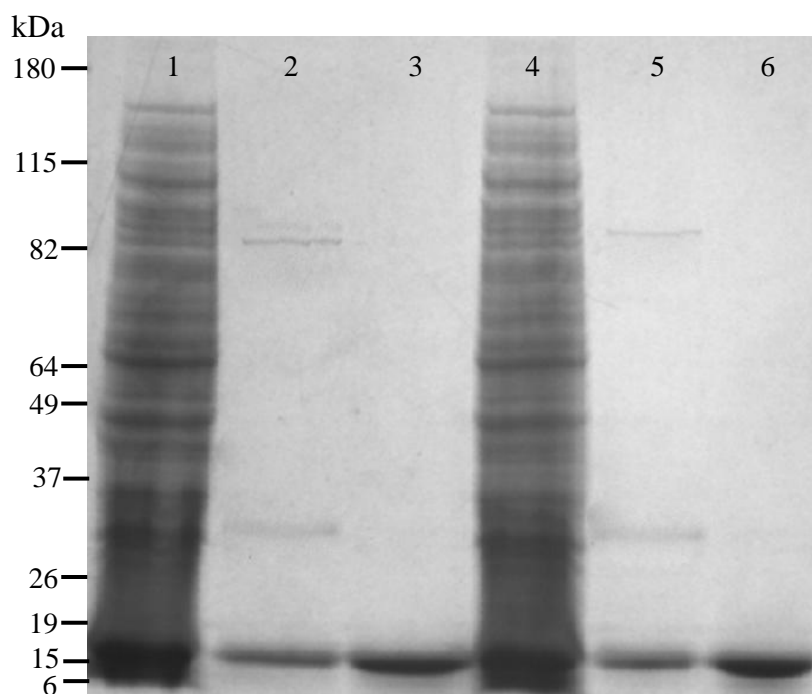
Figure 3.2. Purification of recombinant SE-ENOD wild type (lanes 1-3) and mutated (lanes 4-6) domain II proteins. Recombinant SE-ENOD peptides in *E. coli* lysates (lanes 1 and 4) were sequentially purified on Ni-NTA affinity chromatography (lanes 2 and 5) followed by HPLC size exclusion chromatography (lanes 3 and 6). Protein fractions (10 µl lanes 1 and 4; 10 µg lanes 2, 3, 5, 6) were loaded on a 4-20% gradient SDS-polyacrylamide gel and visualized by Coomassie blue stain. The position of molecular mass standards in kDa is shown on the left.
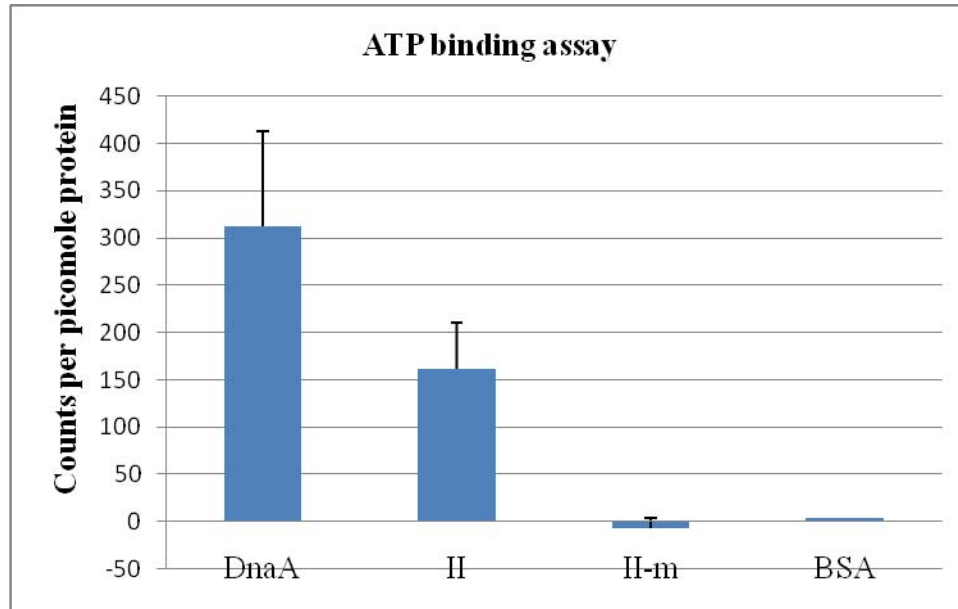
Figure 3.3. Nitrocellulose filter ATP-binding assay. The ATP binding activity ([α-
^{32}P]ATP radioactivity counts per picomole protein) for DnaA ATP binding domain
(DnaA), wild type domain II (II), mutated domain II (II-m) and BSA (BSA) are
represented as blue bars. The blue bars are the mean $\pm$SD of three replicates.

3.3.2 Dimerization analysis by yeast-two-hybrid system

The structural predictions described in Chapter II indicated that an exposed hydrophobic face in domain II would allow homodimers to form between two SE-ENOD subunits. Yeast two hybrid experiments were conducted to provide an empirical test to determine whether SE-ENOD proteins (domains II/III) or the isolated domain II peptides dimerize. The yeast two hybrid system takes advantage of the properties of the budding yeast *Sacchromyces cerevisiae* (*S. cerevisiae*) GAL4 protein that functions as a transcriptional activator of the genes encoding enzymes in galactose utilization process. The GAL4 protein has an N-terminal DNA binding domain (BD) and a C-terminal transcriptional activation domain (AD). The two domains can be physically separated and maintain their specific individual functions. When the two domains are brought together into close physical proximity by intermolecular protein interactions between proteins fused to the domains, the GAL4 protein is functionally reconstituted as a transcriptional activator. This is accomplished experimentally by mating two haploid yeast strains; one containing a BD plasmid and the other an AD plasmid. The expression of reporter genes, such as *HIS* or *ADE*, under the regulation of promoters that contain the activation sequences recognized by the DNA binding domain of the system serve as indicators for positive protein-protein interactions that will allow the mated diploid yeast strains to grow on histidine or adenine deficient media (McAlister-Henn et al., 1999).

The budding yeast *S. cerevisiae* spindle positioning protein Kar9 has a well-defined interaction with another spindle positioning protein Bim1 (Meednu et al., 2008). To serve as a positive control for the analysis, pRM1154-<u>BD</u> plasmid containing Kar9 protein sequence and pRM1151-<u>AD</u> plasmid containing Bim1 protein sequence were

transformed into *S. cerevisiae* yRM1757 and yRM1756 strains, respectively. As negative controls, strains containing empty AD plasmid (E-AD) were mated with strains containing empty BD plasmid (E-BD). The pRM1151-AD plasmid contains the gene for synthesizing leucine, therefore, the yeast strains containing AD plasmid can be selected by growing on leucine deficient plates (-leu). The pRM1154-BD plasmid contains the gene for synthesizing uracil, therefore, the yeast strains containing BD plasmid can be selected by growing on uracil deficient plates (-ura). The diploid yeast containing both of the plasmids can be selected on leucine and uracil double deficient plates (-leu/-ura).

Growth of the diploid colonies on the -leu/-ura media (top) of panels A-E in Figure 3.4 demonstrated successful matings between all of the constructs made in the BD-plasmids and AD-plasmids, including the empty vectors (E-BD and E-AD). Colony growth on both the –his (middle) and –ade (bottom) media of the Kar9 and Bim1 positive control and the lack of growth of the empty vector negative controls on the respective selective media shown in Panel A demonstrated that the experimental system was functioning correctly.

The SE-ENOD domain II and the mature SE-ENOD protein (domains II/III) were each cloned into pRM1151-AD and transformed into yRM1757 strains. Likewise, domain II or domains II/III were each cloned into pRM1154-BD and transformed into yRM1756 strains. Haploid yeast cells were mated by crossing the two strains containing corresponding plasmids to generate multiple combinations that were replicate-plated on either –ura/-leu media, -his media, or –ura media. Figure 3.4 shows the results of the various combinations: domain II-AD x domain II-BD (Panel B); domain II/III-AD x domain II/III-BD (Panel C); domain II-AD x domain II/III-BD (Panel D); and domain

II/III-AD x domain II-BD (Panel E). All of the yeast containing SE-ENOD combinations failed to grow on both –his and –ade plates. These results failed to confirm the predicted structural model that the SE-ENOD forms homodimers.
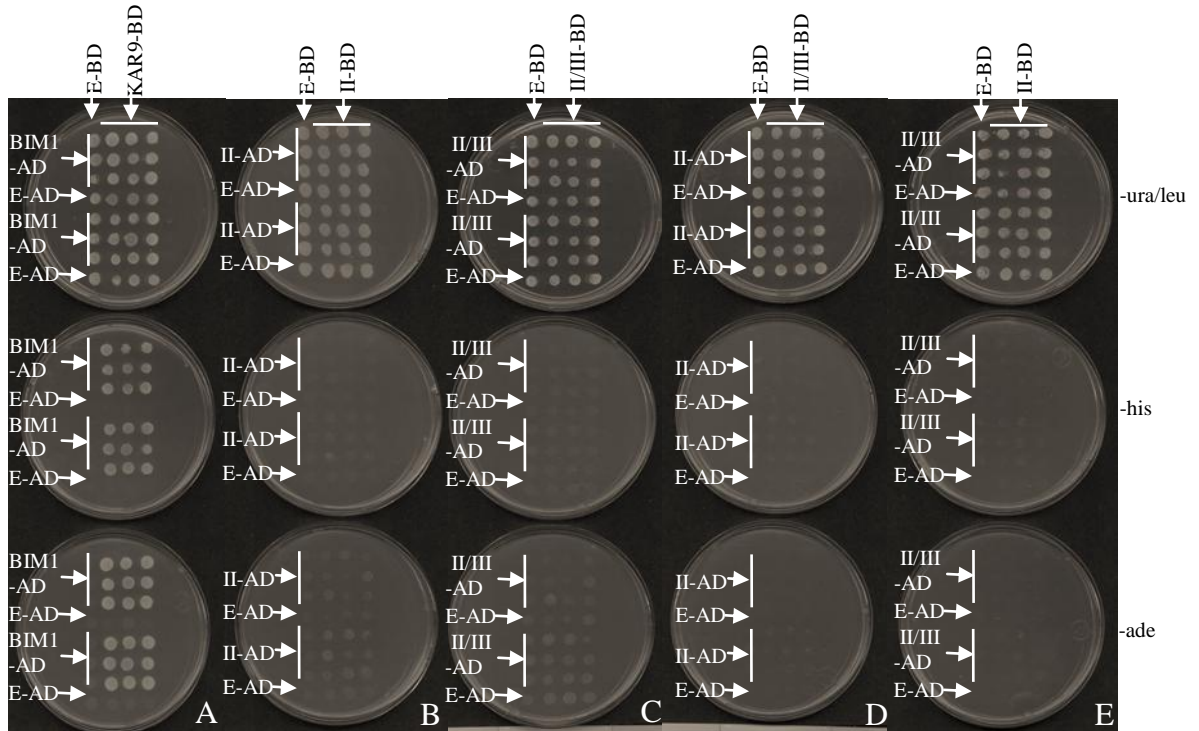
Figure 3.4. Domain II and domain II/III dimerization analysis in the yeast two hybrid system. Diploid two-hybrid reporter strains were generated by crossing yRM1757 containing recombinant II-AD, II/III-AD in pRM1151or empty-AD (pRM1151) plasmid with yRM1756 containing II-BD, II/III-BD in pRM1154 or empty-BD (pRM1154). Diploid colonies were replicate plated on SC-ura/-leu, SC-his and SC-ade media (from top to bottom, mating types are indicated on each plate). Panel A shows the Kar9-BD and Bim1-AD positive interaction. The dimerization tests for SE-ENOD domain II or II/III are shown as follows: panel B is II-BD x II-AD, panel C is II/III-BD x II/III-AD, panel D is II/III-BD x II-AD, and panel E is II-BD x II/III-AD. The negative control strains containing the empty-BD (E-BD) plasmid x empty-AD (E-AD) plasmid were present multiple times on each plate. Two replicates were blotted on each plate.

3.4 DISCUSSION

Extracellular ATP signaling in phloem: a functional model for the SE-ENOD

Plant cells respond to wide range of intrinsic and extrinsic signals, such as hormones, elicitors, stimuli, and some intermediate metabolites. Besides these traditional signals, studies focusing on nucleotide effects on plant biological processes proposed extracellular adenosine triphosphates (eATP) to be an additional extrinsic signal in plants (Jeter and Roux, 2006). Evidence is accumulating that eATP influences physiological development in *Arabidopsis*, such as plasma membrane depolarization in root hairs (Lew and Dearnaley, 2000), inhibition of auxin transport (Tang et al., 2003), and obstruction of pollen germination (Roux and Steinebrunner, 2007). ATP is the energy currency for all living organisms and mostly exists within organelles, such as mitochondria or chloroplasts, and the cytoplasm. Increasingly, studies revealed that plant cells efflux ATP from the cytoplasm into the extracellular matrix mainly through two processes. One is that ATP passively diffuses from the intracellular matrix to extracellular matrix when the plasma membrane is disrupted by wounding. The other alternative process is that efflux occurs by active transport via three mechanisms: ATP exocytosis by secretory vesicles, anion channels, and multidrug resistance (MDR) transporters (Sauer et al., 2000; Bowler et al., 2001; Song et al., 2006). The function of eATPs by active transport into the extracellular matrix is poorly defined, however, evidence of eATP influencing plant cell viability has been reported (Tang et al., 2003; Jeter et al., 2004; Chivasa et al., 2005).

Studies suggest that eATP released by disruptions in the plasma membrane serves as an important signal for plant stress responses. The concentration of eATP at plant

wound sites was recorded as high as millimolar levels at the moment of the wound, gradually decreasing to micromolar levels around the wound site. However, the micromolar levels of eATP appear to be sufficient to trigger wound responses (Song et al., 2006). It was suggested that plants respond to eATP by rapidly increasing cytoplasmic calcium as an early step in multiple signal transduction pathways (Demidchik et al., 2003). Notably, callose deposition and forisome plugging are two critical calcium-dependent injury responses in the sieve elements of the phloem (Zabotin et al., 2002; Furch et al., 2007). In plants, the cytosolic free calcium concentration is maintained by membrane associated Ca^{2+}-release channels (Dodd et al., 2010). A recent study revealed that a putative ATP binding protein in the plasma membrane of *Arabidopsis* roots senses eATP to activate a plasma membrane NADPH oxidase resulting in the production of reactive oxygen species (ROS) (Demidchik et al., 2009). This signal further activated plasma membrane hyperpolarization-activated Ca^{2+} influx channels and promoted Ca^{2+} release from endoplasmic reticulum or vacuole. Eventually, the transcription of stress-signaling mitogen-activated protein kinase MPK3 was activated.

Efficient metabolism of eATP within the extracellular matrix appears to be critical to maintain normal physiological conditions. Therefore, multiple mechanisms responding to eATP could exist in plants. The SE-ENOD has several characteristics that suggest a role in the eATP regulatory pathway as an ATP sensor. The SE-ENOD is anchored to the extracellular leaflet of sieve element plasma membrane where it could sense increasing eATP concentrations as cytoplasmic ATP is released to extracellular matrix in response to wounding. In order to amplify the signal and initiate wounding response reactions, the SE-ENOD could indirectly interact with a low-affinity Ca^{2+}-H^+ antiporter to induce

calcium release from sieve element endoplasmic reticulum by a high-affinity Ca^{2+}-pumping ATPase. The increase of cytoplasmic Ca^{2+} concentration can induce calcium-dependent callose deposition and P-protein dispersal to occlude damaged or disrupted sieve tubes (Furch et al., 2007). Furthermore, it is possible that the SE-ENOD functions as an ATP sensor and interacts with other proteins that are present in the sieve element plasma membrane.

Can local membrane environments affect the structure and function of the SE-ENOD?

Plasma membrane studies revealed that the biological membranes are usually not homogeneous, instead, contain organized microdomains that are rich in sterols and sphingolipids (Martin et al., 2005). The microdomains, often called lipid rafts, have been proposed to enhance oligomeric interactions as well as alter the function of membrane associated proteins by shortening the distance between proteins that are located within this specialized lipid environment (Grossmann et al., 2006; Krugel et al., 2008). For example, dimers of the sucrose transporter SUT1, a sieve element plasma membrane protein located in lipid rafts of potato phloem sieve elements, were the dominant structural and most active forms of the protein (Krugel et al., 2008). While the SE-ENOD did not dimerize in the yeast-two-hybrid analysis (Figure 3.4), detergent extractions of the protein from phloem-enriched tissues showed immunoreactive aggregates that are consistent with dimers, tetramers or oligomeric complexes with other membrane proteins (see Chapter IV). The SE-ENOD domain II and domain II/III were expressed as soluble proteins in the yeast system, which could have significantly influenced their structure and subsequent interactions. There is good evidence that the SE-ENOD is embedded in the sieve element plasma membrane via the post-translational addition of a GPI anchor to the

carboxy-terminus of domain III (Khan et al., 2007). A number of studies have shown that GPI-anchored proteins are components of lipid rafts (Sangiorgio et al., 2004). Therefore, the unique environment of these plasma membrane microdomains could be required to promote protein-protein interactions between the SE-ENOD subunits or with other proteins.

Previous studies have confirmed the presence of NADPH oxidoreductases in *Medicago truncatula* root plasma membrane lipid rafts (Lefebvre et al., 2007). The model of the SE-ENOD indirectly interacting with Ca^{2+}-H^+ antiporter could be further extended through activation of NADPH oxidase and the production of reactive oxygen species as a signal to trigger Ca^{2+} release from sieve element endoplasmic reticulum. Whether the lipid raft environment plays a critical role in the formation of SE-ENOD dimers or complex interactions with other proteins could be analyzed by expressing the unprocessed SE-ENOD protein in yeast. With the presence of the signal peptide and GPI anchor ω site, the SE-ENOD should be post-translationally modified by adding a GPI anchor and inserted into the plasma membrane. Lipid raft microdomains can be isolated by non-ionic detergents and the presence of the SE-ENOD can be analyzed by blue-native PAGE gel. Additionally, a similar analysis can be conducted *in vivo* utilizing an *Arabidopsis* mutant deficient in ergosterol biosynthesis. The lipid rafts microdomains in this mutant line will be very rare because of the low ergosterol level in the plant (Pierson et al., 2004). Therefore, the SE-ENOD should not be targeted into the lipid raft region in the mutant. The conformation of SE-ENOD in plasma membrane can be analyzed by isolating plasma membrane vesicles and visualized on a blue-native PAGE gel. To further confirm the role of lipid raft, the mutant can be complemented with the functional

gene effectively reconstituting lipid raft regions in the mutant plants. Such studies could provide an insightful perspective to better understand the proposed eATP regulatory function and the native conformation of the SE-ENOD *in planta*.

CHAPTER IV

IN VIVO CHARACTERIZATION OF THE SE-ENOD

4.1 INTRODUCTION

The sieve element-specific early nodulin like protein (SE-ENOD) is specifically expressed in the phloem sieve elements in members of the Brassicaceae family, including species in the genera *Streptanthus*, *Arabidopsis,* and *Brassica* (Khan et al., 2007). Immunolocalization studies using the SE-ENOD-specific RS6 monoclonal antibody showed that the protein accumulates in the sieve element plasma membrane of *Arabidopsis* at the earliest developmental stages and persists in mature sieve elements. The pre-protein appears to be synthesized and extensively processed in the endomembrane system of immature sieve elements prior to selective autophagy. The N-terminal signal peptide (domain I) and C-terminal hydrophobic domain (domain IV) of the pre-protein are post-translationally cleaved, and the mature protein, composed of the plastocyanin-like copper binding domain (domain II) and proline/serine rich domain (domain III), is attached to the outer leaflet of the plasma membrane via a covalently bound glycosylphosphatidylinositol (GPI)-anchor. Experiments conducted with live tissue sections of cauliflower (*Brassica oleracea* var. *botrytis*) stems demonstrated that treatment with phosphotidylinositol-specific phospholipase C would cleave the protein from the GPI anchor. ATP/GTP-binding motifs in domain II and arabinogalactan

glycosylation motifs in domain III have been identified in the primary amino acid sequence. While *in vitro* data presented in Chapter III of this thesis indicated that the SE-ENOD can bind ATP, there is no *in planta* data in support of these potentially important functional features.

Structure-function studies of the SE-ENOD have been hampered by the difficulty in obtaining large amounts of the protein from vascular tissues of intact plants with functional, translocating sieve elements. Khan and coworkers (2007) purified the SE-ENOD from vascularized callus cultures of *Streptanthus tortuosus* by immuno-affinity chromatography. The parenchyma cells of the vascularized callus cultures were digested with cell wall degrading enzymes and the thicker-walled sieve elements could be obtained by sequential filtering of the cells, resulting in an enriched sample of sieve element membranes. The technique was tedious, and damage to the cells most often resulted in isolating membrane ghosts rather than intact sieve elements. Furthermore, the sieve elements differentiating in callus cultures, while appearing morphologically and developmentally normal, do not function in long-distance transport. This limits the utility of the system to conduct functional studies. Reverse genetics is another approach that often provides insights into the function of specific proteins. Analysis of two independent T-DNA tagged insertion mutants of the gene encoding the *Arabidopsis* SE-ENOD resulted in knock-out mutants without an obvious and consistent phenotype (Khan et al. 2007). The gene encoding the SE-ENOD is a member of a small multigene family in *Arabidopsis*, and while there is no evidence for a second sieve element-specific gene, functional complementation of the mutants by another family member is possible.

GPI-anchored proteins appear to be abundant in plant plasma membranes, where they are thought to function in cell-to-cell signaling as well as interactions within the extracellular matrix that impact reproduction, cellular growth, and differentiation (Schultz et al., 1998). One model for the function of the SE-ENOD presented in Chapter III focuses on its role as a regulatory molecule for sensing extracellular ATP (eATP) to initiate plant wound responses. GPI-anchored proteins are often located within lipid raft microdomains of the plasma membrane. These specialized membrane domains are thought to enhance oligomeric interactions and influence the interactions of membrane associated proteins (Grossmann et al., 2006; Krugel et al., 2008). Structural modeling studies in Chapter II predicteded that the SE-ENOD can adopt a dimer conformation. Although the dimerization analysis in the budding yeast *Sacchromyces cerevisiae* was negative, this result does not preclude the possibility of SE-ENOD dimer conformation exists *in planta.*

This chapter describes the development of a simple technique to isolate tissues from broccoli plants that are highly enriched in phloem sieve elements. Extraction of the SE-ENOD from the phloem enriched tissues required the addition of detergents to the extraction buffers. Immunoblot analysis using the SE-ENOD-specific RS6 mab, however, showed extensive aggregation that was insensitive to reducing agents, boiling, and denaturants.

4.2 MATERIALS AND METHODS

4.2.1. Isolation of phloem-enriched tissues from broccoli

Stems from commercially grown broccoli crowns were scored with a double-edged razor blade near the base into cylinder-like sections ~3-5 cm wide at a depth of ~1-2 mm. A vertical slice was made to expose the cambium, and the exterior layer composed mostly of the epidermis was peeled off using fine forceps under a binocular microscope. The majority of the phloem tissue was removed with the epidermal peel, leaving behind the xylem tissues. Strands of phloem-enriched tissue were prepared by peeling phloem fibers from the epidermal peel with a probe under the binocular microscope, which were flash frozen in liquid nitrogen. The phloem strands were weighed and immediately stored at -80°C. Phloem enrichment was confirmed by visualization by light microscopy and immunolocalization with phloem specific antibodies.

4.2.2 Phloem membrane protein extraction

Two grams of phloem-enriched tissue were ground in liquid nitrogen with a mortar and pestle and extracted with 4 ml of soluble protein extraction buffer (10 mM Tris pH 7.2, 10 mM EGTA, 150 mM NaCl, 10 mM KCl, 1% Sigma plant protease inhibitor cocktail, 20 mM dithiothreitol). The tissue was incubated in the soluble extraction buffer for one hour on a rocking platform at 4°C. Soluble proteins in 4.5 ml of the supernatant were removed following centrifugation at 17,000 rpm for 25 minutes in JA 20 rotor in Avanti J-E centrifuge (Beckman Coulter). Tissues were resuspended in 4 ml of buffers containing either CHAPS or SDS detergents and incubated at room temperature for 1 hour on a rocking platform. CHAPS buffer was composed of 1%

CHAPS, 10 mM Tris pH 7.2, 10 mM EGTA, 150 mM NaCl, 10 mM KCl, 1% Sigma plant protease inhibitor cocktail, 20 mM dithiothreitol. SDS buffer was composed of 4% SDS, 125 mM Tris-HCl pH 7.2, 150 mM NaCl, 10 mM KCl, 50 mM dithiothreitol, 1% Sigma plant protease inhibitor cocktail. Four ml of the supernatant containing total membrane proteins were collected following centrifugation as described above. Protein concentrations were determined with the RCDC protein assay kit (Bio-Rad, catalog no.500-0119), which is compatible with CHAPS and SDS. Serial dilutions of BSA in the respective detergent buffers served as concentration standards. Aliquots of the aqueous and detergent soluble protein fractions were flash frozen in liquid nitrogen and stored at -80°C.

4.2.3 Immunolocalization of the SE-ENOD in phloem-enriched tissues

Several newly isolated phloem-enriched strands were blocked with 2.5% dry milk in PBS (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4 and 1.47 mM KH_2PO_4, pH 7.4) for 20 minutes at room temperature on a rocking platform. Tissues were incubated with the RS6 (1:100 dilution in PBS) primary antibody for 45 minutes to 1 hour at room temperature with rocking. Tissues were then washed three times for 10 minutes each with PBS followed by incubation with Alexa Fluor 488-labeled goat anti-mouse IgG (InVitrogen, A11001) secondary antibody (1:1000 dilution in PBS) in the dark at room temperature with rocking. The tissues were then washed three times for 10 minutes each with PBS and visualized by light and fluorescence microscopy.

4.2.4 Immunoblotting proteins extracted from phloem-enriched tissues

Different concentrations of membrane proteins, as specified by the experiment, were extracted from 2 grams phloem-enriched tissues and separated by SDS-PAGE. Proteins were transferred for 90 minutes from the gel to a PVDF membrane (GE Healthcare, pack no. RPN303F) in a western blot wet transfer apparatus in 1 liter Tris-glycine transfer buffer (28 mM Tris base pH 8.3, 190 mM glycine) at 90 volts. Following transfer, the PVDF membrane was blocked with 5% dry milk in TBS (10 mM Tris, pH 7.5, 150 mM NaCl) for 1 hour at room temperature. The membrane was then washed five times with for five minutes each with TTBS (20 mM Tris, pH 7.5, 500 mM NaCl, 0.05% Tween 20). The membrane was incubated in RS6 mab (1:5000 dilution in TTBS) for 1 hour at room temperature followed by five washes for five minutes each with TTBS. The membrane was then incubated with anti-mouse IgG horseradish peroxidase (HRP) (Amersham, NXA931) (1:4000 dilution in TTBS) for 1 hour at room temperature followed by five washes for five minutes each with TTBS. Immunoreactive proteins were visualized following treatment of the membranes with five ml of western blotting luminal reagent buffer A mixed with five ml of buffer B (Santa Cruz Biotechnology, Inc. catalog no. sc-2048) on X ray films and AFP Imaging (AFP Imaging Corp).

4.3 RESULTS

4.3.1 Isolation and immunolocalization of broccoli phloem-enriched tissues

The SE-ENOD was initially immunopurified from dispersed islands of phloem that were induced to differentiate in *Streptanthus tortuosus* callus cultures. While the protein accumulated in the plasma membrane of phloem tissues generated *in vitro*, the sieve elements are discontinuous and non-functional. Furthermore, the maintenance of the tissue cultures combined with low yields of enriched sieve elements prompted the development of an alternative technique to easily isolate significant amounts of phloem tissue from whole plants. Previous work had shown that the SE-ENOD-specific RS6 mab readily labeled sieve elements in fresh tissue sections of commercially available *Brassica* species including both cauliflower and broccoli (Khan et al, 2007). The large stems of broccoli crowns proved to be a useful source to isolate strands of phloem-enriched tissues. The outer layer composed mostly of epidermis and adjacent cells was easily peeled from the stem. These sections also contained vertical files of phloem tissue that had separated at the cambium from the xylem tissue. The phloem fibers allowed strands of phloem-enriched tissue to be removed with a probe and forceps.

Sieve tubes with their connecting sieve plates in the isolated strands could be clearly observed by light microscopy (Figure 4.1B). Immunolocalization experiments with the RS6 mab confirmed the presence of SE-ENOD in sieve elements within the strands (Figure 4.1C). Interestingly, the SE-ENOD was abundant in these phloem strands, but its distribution showed a clustered pattern rather than homogeneous distribution along the length of the contiguous sieve elements.
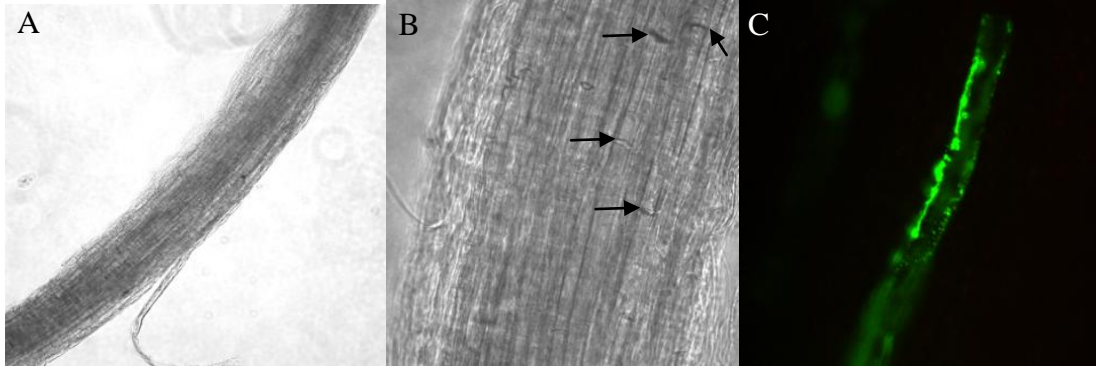
Figure 4.1 Broccoli phloem-enriched tissue strands and SE-ENOD immunolocalization visualized by light and fluorescence microscopy. Panel A. Brightfield image shows the tissue strands under low magnification (100x). Panel B. Brightfield image shows the tissue strands under higher magnification (400x) where sieve tubes can be readily seen. Sieve plates are indicated by black arrows. Panel C. Fluorescence immunolocalization of the SE-ENOD with the RS6 monoclonal antibody in contiguous sieve elements present in the tissue strand shown in Panel A.

4.3.2 Protein extractions from phloem-enriched tissues and immunoblot detection of the SE-ENOD

Proteins were extracted in a two-step process from the phloem-enriched tissue strands. Because the SE-ENOD is attached to sieve element plasma membrane by a GPI anchor, this membrane-bound protein can be further enriched by initially extracting the soluble proteins within the tissue sample with an aqueous buffer. Aqueous insoluble proteins, including membrane bound proteins, were extracted by adding detergents to the extraction buffers to disrupt membrane structure and solubilize hydrophobic proteins.

In the first set of experiments, the total soluble proteins were extracted in the aqueous buffer followed by detergent buffer containing 1% CHAPS. CHAPS is a zwitterionic detergent that is used to solubilize membrane proteins in a native conformation (Khan et al., 2007). Panel A in Figure 4.2 shows the protein profiles separated by SDS-PAGE of the sequential aqueous (lanes 2, 4, and 6) and CHAPS (lanes 1, 3, and 5) extractions. It should be noted that the two profiles are significantly different from one another and that the proteins appear to be well resolved. Immunoblot analysis with the RS6 mab shown in Panel B of Figure 4.2 failed to detect immunoreactive proteins in the two lowest concentrations of soluble protein extractions (lanes 7 and 9) and appeared to cross-react with abundant soluble protein species in the highest concentration of soluble proteins resolved on the gel (lane 11). Unexpectedly, the SE-ENOD was detected as high molecular weight smears ranging from 40 kDa to 120 kDa in all concentrations of CHAPS extracted proteins that were loaded on the gel (Figure 4.2, lanes 8, 10, and 12). Another phloem-specific monoclonal antibody, RS32, consistently detected a well-resolved, detergent soluble protein with apparent molecular weight of 35 kDa. The presence of the 35 kDa protein indicated that the high molecular weight smear

detected by RS6 mab is not simply the result of non-specific interactions of detergent soluble proteins with the antibodies. Rather the results suggest that the SE-ENOD forms complexes that are not resolved after boiling in loading buffer containing SDS and reducing agents and separation in SDS-PAGE. These complexes were thought to be either an artifact of extraction with CHAPS or hydrophobic interactions with other cellular components. In contrast to these results, previous investigators reported extracting and purifying the 15 kDa SE-ENOD in a 1% CHAPS buffer from vascularized *Streptanthus* callus cultures (Khan et al., 2007).

In a second set of experiments, total proteins were extracted from the phloem-enriched tissues with one step protocol using 4% sodium dodecyl sulfate (SDS). SDS is an anionic surfactant that has a high solubility of proteins. The total proteins showed a similar smeared immunoblot pattern to what had been previously observed with the CHAPS extraction (Figure 4.3, lanes 1, 3, 5, and 7). Pre-extraction with aqueous buffer containing moderate concentrations of salt and a reducing agent was performed before the sequential extraction with 4% SDS added to an identical buffer. The aqueous buffer pre-extraction step followed by the addition of 4% SDS detergent resolved the high molecular weight smear into three predominant species that are consistent with SE-ENOD monomer (15 kDa), dimer (30 kDa), and tetramer (60 kDa) as well as less resolved species up to 120 kDa (Figure 4.3, lanes 4, 6, and 8).

It appeared that the proteins could be reasonably resolved by the sequential extraction with aqueous and SDS-containing buffers, however, the results were not consistent for every extraction. Questions still persisted as to whether some steps, such as boiling before loading on a SDS-PAGE gel, contributed to the high molecular smear and

whether higher levels of thiol modifying agents might aid in fully resolving the proteins. Figure 4.4 shows the immunoblot results of a single tissue sample that was sequentially extracted with an aqueous buffer (discarded) and SDS-containing buffer (analyzed) diluted with SDS loading buffer containing either 20 mM (lanes 1-4), 100 mM (lanes 5 and 7), or 200 mM (lanes 6 and 8) dithiothreitol (DTT). The samples were either boiled (lanes 1, 2, 5, and 6) or denatured on ice (lanes 3, 4, 7, and 8) prior to separating the proteins by SDS-PAGE. The method of denaturing the proteins or the concentration of reducing agent in the SDS-PAGE loading buffer did not improve the resolution of the SE-ENOD complexes. Additionally, increasing the SDS concentration to 8% in the detergent buffer did not have effect in resolving the complexes (data not shown). Hydrophobic aggregation caused by long incubation periods with extraction buffers or long-term storage was also eliminated by a rapid extraction followed by immediate separation by SDS-PAGE. Finally, phloem-enriched tissues taken from commercially available broccoli crowns were compared with tissues isolated from living plants with no difference in the results.
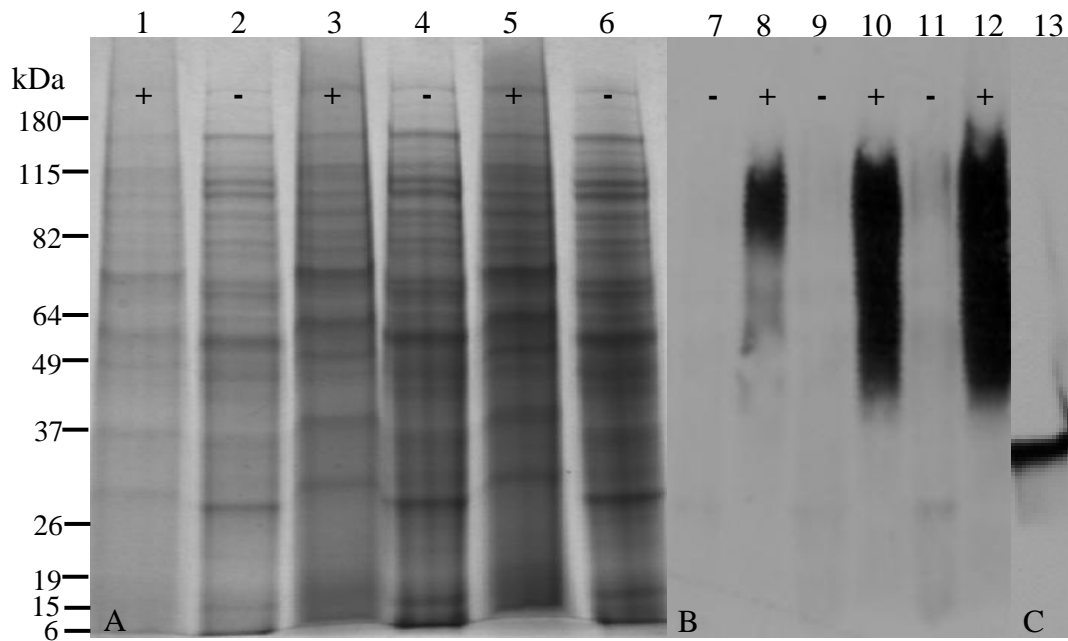
Figure 4.2 SDS-PAGE separation (Panel A) and immunoblot analysis (Panels B and C) of proteins sequentially extracted from broccoli phloem-enriched tissues with aqueous (-) and 1% CHAPS (+) buffers. Panel A. Proteins (25µg in lanes 1 and 2; 50µg in lanes 3 and 4; 100µg in lanes 5 and 6) were separated in a 12% SDS-PAGE gel and visualized by Coomassie blue stain. Panel B. Immunoblot of the protein fractions (10µg in lanes 7 and 8; 25µg in lanes 9 and 10; 50µg in lanes 11 and 12) separated by 12% SDS-PAGE, blotted onto PVDF membrane, and detected with the RS6 mab (1:5000 dilution). Panel C. Immunoblot of a 1% CHAPS-extracted protein sample (10 µg in lane 13; duplicate of lane 8 in Panel B) reacted with the RS32 mab (1:5000 dilution). Molecular weight standards are shown in kDa on the left.
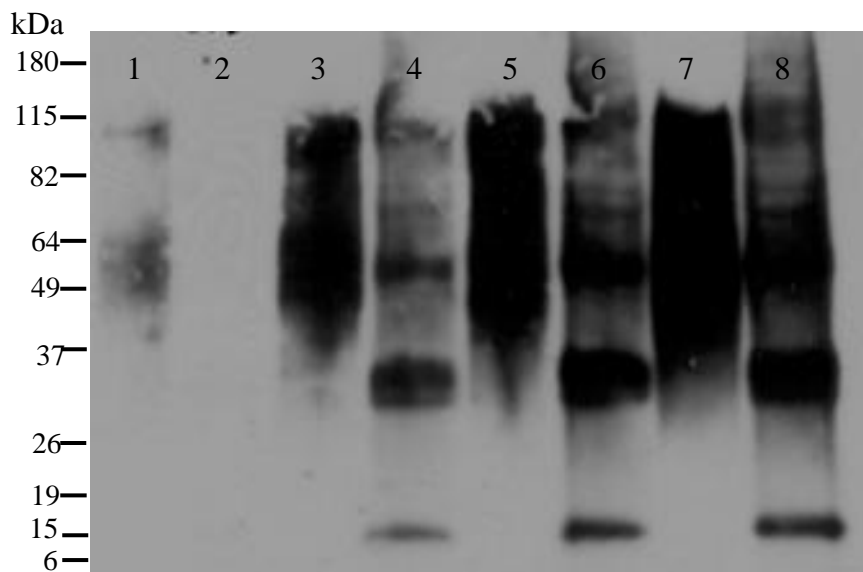
Figure 4.3 Immunoblot analysis of the SE-ENOD extracted from broccoli phloem-enriched tissues with 4% SDS with or without aqueous buffer pre-extraction. Proteins extracted by a one step protocol using a 4% SDS buffer (lanes 1, 3, 5, and 7) or sequentially extracted by a two step protocol using an aqueous buffer followed by a 4% SDS buffer (lanes 2, 4, 6, and 8) were separated in a 4-20% SDS-PAGE gradient gel, blotted onto PVDF membrane, and detected with the RS6 mab (1:5000 dilution). Increasing volumes of the extractions were loaded onto the gel: 5µl (lanes 1 and 2), 10µl (lanes 3 and 4), 15µl (lanes 5 and 6), and 25µl (lanes 7 and 8). Molecular weight standards are shown in kDa on the left.
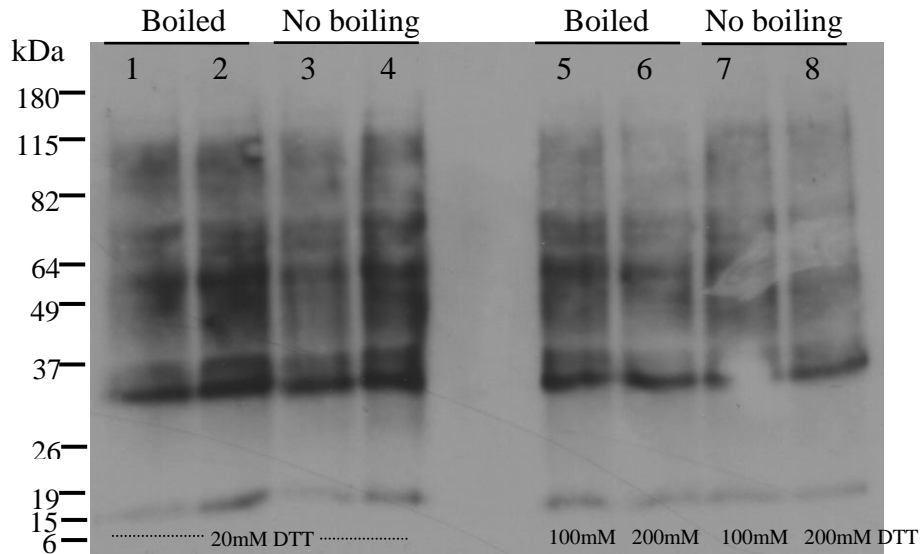
Figure 4.4 Immunoblot analysis of the SE-ENOD extracted from broccoli phloem-enriched tissues: effects of boiling and reducing agents. Detergent soluble proteins were sequentially extracted with aqueous buffer followed by 4% SDS buffer. Two volumes of proteins: 30µl (lanes 1, 3, 5, 6, 7, and 8) and 50µl (lanes 2 and 4) were separated by 4-20% SDS-PAGE. Samples were denatured by boiling at 100°C (lanes 1, 2, 5, and 6) or on ice (lanes 3, 4, 7, and 8) prior to loading onto the gel. Different dithiotreitol (DTT) concentrations (20mM, lanes 1 to 4; 100mM, lane 5 and 7; 200mM, lanes 6 and 8) in SDS loading buffer are shown. The SE-ENOD was detected with the RS6 mab (1:5000 dilution). Molecular weight standards are shown in kDa on the left.

4.4 DISSCUSSION

One of the challenges in studying the function of sieve element membrane proteins is the efficient isolation of significant amounts of phloem-enriched tissue. Historically, this has been resolved by several approaches depending upon the protein and the species being investigated. For example, sugar beet leaves have high concentrations of sucrose transporters that can be efficiently isolated in plasma membrane-enriched vesicles prepared from the leaves. Similar approaches have been used to study sucrose transporters in other species, but the choice of the plant is directed by the ability to isolate the protein. Sieve element-specific proteins, although abundant within the sieve element, are often difficult to detect in whole organ extracts even with sensitive antibody probes. To solve this problem, Sjolund and coworkers developed techniques to induce vascular differentiation in callus cultures by optimizing growth regulators in the culture media (Toth et al., 1994; Khan et al., 2007). The sieve elements formed in callus cultures were shown to be developmentally and structurally similar to the sieve elements in intact plants, however, these sieve elements are not connected and do not function in long distance transport (Evert R.F., 1990).

Another approach has been to isolate sieve element membrane proteins from plants that have easily accessible vasculature. The unusually large petioles of celery contain thick-walled parenchyma associated with the vascular tissues allowing the phloem to be easily stripped from the organ at the vascular cambium. This technique was modified for broccoli as a companion approach to complement molecular and genetic studies that can be conducted in *Arabidopsis*. The technique of collecting phloem-enriched strands from broccoli stems enables the enrichment of phloem tissues directly

81

from living plants. Broccoli stems are large and easy to peel off the epidermis along with adjacent cell layers that contain the phloem tissues. Phloem-enriched tissue strands can be readily dissected from the epidermal peels.

Immunoblot analyses using a number of different sieve element-specific antibodies detected discrete protein antigens as well-resolved, single species from both aqueous- and detergent-soluble extractions of the broccoli phloem-enriched tissues. As expected, the initial experiments to identify the SE-ENOD found that the protein was present only when the phloem-enriched tissues were extracted in the presence of a detergent. This agrees with the previous data that the SE-ENOD is a plasma membrane anchored protein. Surprisingly, the SE-ENOD-specific RS6 monoclonal antibody detected an unresolved smear of proteins ranging from 40 kDa to 120 kDa on the immunoblot when the tissue was extracted with CHAPS. This is in contrast to previous results when a CHAPS buffer of identical composition was used to extract the SE-ENOD from vascularized *Streptanthus* callus cultures. In those experiments, the RS6 mab detected a single 15 kDa protein (Khan et al., 2007). Different methods, such as increasing the concentration of reducing agent, loading proteins on gels without denaturing the sample by boiling, and changing extraction detergent to SDS, were used in attempts to resolve the SE-ENOD to the 15 kDa monomer.

The best results were obtained by first extracting the phloem-enriched tissues with an aqueous buffer containing moderate salt concentrations and reducing agent followed by a second extraction with the 4% SDS added to the aqueous buffer. Following this protocol, the RS6 antibody detected protein species resolved at 15 kDa, 30 kDa, and 60 kDa that could correspond to the SE-ENOD monomer, dimer, and tetramer,

respectively. High molecular weight smears were still present as background. These results suggested that the SE-ENOD forms stable oligomeric conformations in the sieve element plasma membrane of whole plant tissues.

The plasma membrane is primarily composed of phospholipids that are rich in unsaturated aliphatic chains loosely packed into a fluid bilayer. Rapid lateral movement and constant fluidity are characteristic of this disorganized structure. Lipid rafts are specialized regions of the apical plasma membrane that are rich in sphingolipids, ergosterol, and cholesterol that play important roles in sorting GPI-anchored proteins, transmembrane receptors, and palmitoylated signaling proteins (Grossmann et al., 2006). The association of different proteins in lipid rafts appear to alter their functions of these proteins, sometimes inducing oligomeric confirmations, in order to regulate various biological processes such as signaling, cell adhesion, and pathogen recognition (Brown and London, 2000; Edidin, 2003; Zurzolo et al., 2003). Therefore, it is proposed that the different conformations of the SE-ENOD are associated with lipid rafts.

As proposed in the Chapter III, the SE-ENOD could bind eATP in order to regulate the eATP concentration. The regulatory function could be enhanced if the SE-ENOD was located in lipid rafts. The size of lipid rafts is critical in regulating the function of their associated proteins. It is thought that the small size of an individual raft usually keeps the proteins inactive. In order to initiate protein functions in response to a stimulus, several individual rafts cluster with one another to form a platform. Interestingly, the immunolocalization pattern in sieve elements shows the presence of the SE-ENOD in large, but discrete, areas on the surface of the cell. The platform will induce protein-protein interactions that further activate their functions (Garcia et al., 2003;

Mongrand et al., 2004). This could explain why multiple conformations of the SE-ENOD were detected from phloem-enriched tissues.

Other sieve element membrane proteins appear to be localized in lipid rafts as activated multimers. Experimentally, the SUT1 sucrose transporter in *Solanum tuberosum* was extracted from plasma membrane vesicles as dimers and/or monomers depending upon the protein-to-detergent ratio of the extraction buffer. A high protein-to-detergent ratio (200 µg total protein to 0.2% detergent) yielded two distinct proteins corresponding to SUT1 dimers and monomers on blue native-PAGE gel. In contrast, a low protein-to-detergent ratio (20 µg total protein to 2% detergent) yielded a single protein corresponding to the SUT1 monomer (Krugel et al., 2008). Detergent extractions of the broccoli phloem-enriched tissues were performed with a single weight to volume ratio of tissue to extraction buffer. Unlike plasma membrane vesicles, the total amount of protein present in whole tissues is difficult to estimate. Thus, inappropriate protein-to-detergent ratios during the extractions could, in part, explain why oligomeric complexes were only partially resolved in the different experiments. Another explanation for the oligomeric complex pattern observed following SDS-PAGE and immunoblotting could be due to covalent interactions with the SE-ENOD and other proteins. However, the lack of resolution following harsh denaturing conditions and separation in SDS-PAGE is still unexplained.

The approach of bulk extracting proteins from the phloem-enriched tissues with various detergents could also be generating SE-ENOD-specific artifacts. As an alternative approach, phloem-enriched tissues could also be used to isolate microsomes and if necessary, plasma membrane vesicles. Extracting proteins from partially or highly

purified membrane preparations could provide an entirely different result that was obtained from bulk extraction of whole tissues. Furthermore, microsomes or plasma membrane vesicles could be used to conduct the appropriate experiments to determine whether the SE-ENOD is localized in lipid rafts. Lipid raft microdomains contain saturated aliphatic chains that are insoluble in 1% Triton X-114 at 4°C providing a convenient way to isolate lipid raft-associated proteins (Radeva and Sharom, 2004). Future work could include a proteomics approach to analyze lipid raft protein profiles in phloem-enriched tissues. Further functional characterization of the SE-ENOD in the sieve element plasma membrane promises to provide interesting insights into our understanding of how membrane-associated proteins regulate events that influence the physiology of the phloem.

REFERENCES

Adman, E.T. (1991). Copper protein structures. Advances in Protein Chemistry **42,** 145-197.

Ainscough, E.W., Bingham, A.G., Brodie, A.M., Ellis, W.R., Gray, H.B., Loehr, T.M., Plowman, J.E., Norris, G.E., and Baker, E.N. (1987). Spectrochemical studies on the blue copper protein azurin from Alcaligenes denitrificans. Biochemistry **26,** 71-82.

Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics **22,** 195-201.

Barker, L., Kuhn, C., Weise, A., Schulz, A., Gebhardt, C., Hirner, B., Hellmann, H., Schulze, W., Ward, J.M., and Frommer, W.B. (2000). SUT2, a putative sucrose sensor in sieve elements. Plant Cell **12,** 1153-1164.

Barrieu, F., Chaumont, F., and Chrispeels, M.J. (1998). High expression of the tonoplast aquaporin ZmTIP1 in epidermal and conducting tissues of maize. Plant Physiology **117,** 1153-1163.

Bencharki, B., Boissinot, S., Revollon, S., Ziegler-Graff, V., Erdinger, M., Wiss, L., Dinant, S., Renard, D., Beuve, M., Lemaitre-Guillier, C., and Brault, V. (2010). Phloem protein partners of cucurbit aphid borne yellows virus: possible involvement of phloem proteins in virus transmission by aphids. Molecular Plant Microbe Interactions **23,** 799-810.

Bird, D.A., Franceschi, V.R., and Facchini, P.J. (2003). A tale of three cell types: alkaloid biosynthesis is localized to sieve elements in opium poppy. Plant Cell **15,** 2626-2635.

Borner, G.H., Lilley, K.S., Stevens, T.J., and Dupree, P. (2003). Identification of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A proteomic and genomic analysis. Plant Physiology **132,** 568-577.

Borner, G.H., Sherrier, D.J., Stevens, T.J., Arkin, I.T., and Dupree, P. (2002). Prediction of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A genomic analysis. Plant Physiology **129,** 486-499.

Bowler, W.B., Buckley, K.A., Gartland, A., Hipskind, R.A., Bilbe, G., and

Gallagher, J.A. (2001). Extracellular nucleotide signaling: a mechanism for integrating local and systemic responses in the activation of bone remodeling. Bone **28,** 507-512.

Bowles, D. (1998). Signal transduction in the wound response of tomato plants. Philosophical Transactions of the Royal Society Lond B Biological Sciences **353,** 1495-1510.

Brown, D.A. and London, E. (2000). Structure and function of sphingolipid- and cholesterol-rich membrane rafts. Journal of Biological Chemistry **275,** 17221-17224.

Burnstock, G. (1996). P2 purinoceptors: historical perspective and classification. Ciba Foundation Symposium **198,** 1-28; discussion 29-34.

Burnstock, G. (2007). Purine and pyrimidine receptors. Cellular and Molecular Life Sciences **64,** 1471-1483.

Carrington, J.C., Kasschau, K.D., Mahajan, S.K., and Schaad, M.C. (1996). Cell-to-cell and long-distance transport of viruses in plants. Plant Cell **8,** 1669-1681.

Chivasa, S., Ndimba, B.K., Simon, W.J., Lindsey, K., and Slabas, A.R. (2005). Extracellular ATP functions as an endogenous external metabolite regulating plant cell viability. Plant Cell **17,** 3019-3034.

Chothia, C. and Lesk, A.M. (1982). Evolution of proteins formed by beta-sheets. I. Plastocyanin and azurin. Journal of Molecular Biology **160,** 309-323.

Chothia, C. and Lesk, A.M. (1986). The relation between the divergence of sequence and structure in proteins. EMBO Journal **5,** 823-826.

Clark, A.M., Jacobsen, K.R., Bostwick, D.E., Dannenhoffer, J.M., Skaggs, M.I., and Thompson, G.A. (1997). Molecular characterization of a phloem-specific gene encoding the filament protein, phloem protein 1 (PP1), from Cucurbita maxima. The Plant Journal **12,** 49-61.

Combier, J.P., Vernie, T., de Billy, F., El Yahyaoui, F., Mathis, R., and Gamas, P. (2007). The MtMMPL1 early nodulin is a novel member of the matrix metalloendoproteinase family with a role in Medicago truncatula infection by Sinorhizobium meliloti. Plant Physiology **144,** 703-716.

Comeau, S.R., Gatchell, D.W., Vajda, S., and Camacho, C.J. (2004). ClusPro: an automated docking and discrimination method for the prediction of protein complexes. Bioinformatics **20,** 45-50.

Comeau, S.R., Gatchell, D.W., Vajda, S., and Camacho, C.J. (2004). ClusPro: a fully automated algorithm for protein-protein docking. Nucleic Acids Research **32,** W96-99.

Comeau, S.R., Kozakov, D., Brenke, R., Shen, Y., Beglov, D., and Vajda, S. (2007).

ClusPro: performance in CAPRI rounds 6-11 and the new server. Proteins **69,** 781-785.

Crawford, K.M. and Zambryski, P.C. (1999). Phloem transport: Are you chaperoned? Current Biology **9,** R281-285.

Cronshaw, J. and Esau, K. (1967). Tubular and fibrillar components of mature and differentiating sieve elements. The Journal of Cell Biology **34,** 801-815.

Cronshaw, J., Gilder, J., and Stone, D. (1973). Fine structural studies of P-proteins in Cucurbita, Cucumis, and Nicotiana. Journal of Ultrastructure Research **45,** 192-205.

Crooke, E., Castuma, C.E., and Kornberg, A. (1992). The chromosome origin of Escherichia coli stabilizes DnaA protein during rejuvenation by phospholipids. The Journal of Biology Chemistry **267,** 16779-16782.

Demidchik, V., Shang, Z., Shin, R., Thompson, E., Rubio, L., Laohavisit, A., Mortimer, J.C., Chivasa, S., Slabas, A.R., Glover, B.J., Schachtman, D.P., Shabala, S.N., and Davies, J.M. (2009). Plant extracellular ATP signalling by plasma membrane NADPH oxidase and Ca^{2+} channels. The Plant Journal **58,** 903-913.

DeWitt, N.D. and Sussman, M.R. (1995). Immunocytological localization of an epitope-tagged plasma membrane proton pump (H(+)-ATPase) in phloem companion cells. Plant Cell **7,** 2053-2067.

Dodd, A.N., Kudla, J., and Sanders, D. (2010). The language of calcium signaling. Annual Review of Plant Biology **61,** 593-620.

Eckardt, N.A. (2001). A calcium-regulated gatekeeper in phloem sieve tubes. Plant Cell **13,** 989-992.

Edidin, M. (2003). The state of lipid rafts: from model membranes to cells. Annual Review of Biophysics and Biomolecular Structure **32,** 257-283.

Elortza, F., Nuhse, T.S., Foster, L.J., Stensballe, A., Peck, S.C., and Jensen, O.N. (2003). Proteomic analysis of glycosylphosphatidylinositol-anchored membrane proteins. Molecular and Cellular Proteomics **2,** 1261-1270.

Esau, K., and Hoefert, L.L. (1980). Endoplasmic-reticulum and its relation to microtubules in sieve elements of sugarbeet and spinach. Journal of Ultrastructure Research **71,** 249-257.

Evert, R.F., Tucker, C.M., Davis, J.D., and Deshpand.Bp. (1969). Light microscope investigation of sieve-element ontogeny and structure in Ulmus americana. American Journal of Botany **56,** 999-1017.

Evert, R.F. (1990). Dicotyledons *in*: Sieve elements, comparative structure, induction and development. H. D. Behnke and R. D. Sjolund, eds. Springer-Verlag Berlin

Heidelberg. pp 103-137.

Fields, S. and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature **340,** 245-246.

Fischer, D., Norel, R., Wolfson, H., and Nussinov, R. (1993). Surface motifs by a computer vision technique: searches, detection, and implications for protein-ligand recognition. Proteins **16,** 278-292.

Fischer, W.N., Kwart, M., Hummel, S., and Frommer, W.B. (1995). Substrate specificity and expression profile of amino acid transporters (AAPs) in Arabidopsis. The Journal of Biological Chemistry **270,** 16315-16320.

Frommer, W.B. and Ninnemann, O. (1995). Heterologous expression of genes in bacterial, fungal, animal and plant cells. Annual Review of Plant Physiology and Plant Molecular Biology **46,** 419-444.

Fukuda, A., Okada, Y., Suzui, N., Fujiwara, T., Yoneyama, T., and Hayashi, H. (2004). Cloning and characterization of the gene for a phloem-specific glutathione S-transferase from rice leaves. Physiologia Plantarum **120,** 595-602.

Furch, A.C., Hafke, J.B., Schulz, A., and van Bel, A.J. (2007). Ca^{2+}-mediated remote control of reversible sieve tube occlusion in Vicia faba. Journal of Experimental Botany **58,** 2827-2838.

Garcia, A., Cayla, X., Fleischer, A., Guergnon, J., Alvarez-Franco Canas, F., Rebollo, M.P., Roncal, F., and Rebollo, A. (2003). Rafts: a simple way to control apoptosis by subcellular redistribution. Biochimie **85,** 727-731.

Gewirth, A.A., and Solomon, E.I. (1988). Electronic structure of plastocyanin: excited state spectral features. Journal of the American Chemical Society **110,** 3811–3819.

Golecki, B., Schulz, A., and Thompson, G.A. (1999). Translocation of structural P proteins in the phloem. Plant Cell **11,** 127-140.

Greene, E.A., Erard, M., Dedieu, A., and Barker, D.G. (1998). MtENOD16 and 20 are members of a family of phytocyanin-related early nodulins. Plant Molecular Biology **36,** 775-783.

Grossmann, G., Opekarova, M., Novakova, L., Stolz, J., and Tanner, W. (2006). Lipid raft-based membrane compartmentation of a plant transport protein expressed in Saccharomyces cerevisiae. Eukaryotic Cell **5,** 945-953.

Guckert, J.A., Lowery, M.D., and Solomin, E.I. (1995). Electronic structure of the reduced blue copper active site: contributions to reduction potentials and geometry. Journal of the American Chemical Society **117,** 2817–2844.

Guss, J.M., Merritt, E.A., Phizackerley, R.P., and Freeman, H.C. (1996). The structure of a phytocyanin, the basic blue protein from cucumber, refined at 1.8

angstrom resolution. Journal of Molecular Biology **262,** 686-705.

Hackel, A., Schauer, N., Carrari, F., Fernie, A.R., Grimm, B., and Kuhn, C. (2006). Sucrose transporter LeSUT1 and LeSUT2 inhibition affects tomato fruit development in different ways. The Plant Journal **45,** 180-192.

Ham, B.K., Brandom, J.L., Xoconostle-Cazares, B., Ringgold, V., Lough, T.J., and Lucas, W.J. (2009). A polypyrimidine tract binding protein, pumpkin RBP50, forms the basis of a phloem-mobile ribonucleoprotein complex. Plant Cell **21,** 197-215.

Hancock, R.D., McRae, D., Haupt, S., and Viola, R. (2003). Synthesis of L-ascorbic acid in the phloem. BMC Plant Biology **3,** 7-21.

Hart, P.J., Nersissian, A.M., Herrmann, R.G., Nalbandyan, R.M., Valentine, J.S., and Eisenberg, D. (1996). A missing link in cupredoxins: crystal structure of cucumber stellacyanin at 1.6 A resolution. Protein Science **5,** 2175-2183.

Hause, B., Hause, G., Kutter, C., Miersch, O., and Wasternack, C. (2003). Enzymes of jasmonate biosynthesis occur in tomato sieve elements. Plant and Cell Physiology **44,** 643-648.

Hendlich, M., Rippmann, F., and Barnickel, G. (1997). LIGSITE: automatic and efficient detection of potential small molecule-binding sites in proteins. Journal of Molecular Graphics and Modelling **15,** 359-363.

Hirner, A., Ladwig, F., Stransky, H., Okumoto, S., Keinath, M., Harms, A., Frommer, W.B., and Koch, W. (2006). Arabidopsis LHT1 is a high-affinity transporter for cellular amino acid uptake in both root epidermis and leaf mesophyll. Plant Cell **18,** 1931-1946.

Holbrook, N.M., Burns, M.J., and Field, C.B. (1995). Negative xylem pressures in plants: a test of the balancing pressure technique. Science **270,** 1193-1194.

Ishiwatari, Y., Honda, C., Kawashima, I., Nakamura, S., Hirano, H., Mori, S., Fujiwara, T., Hayashi, H., and Chino, M. (1995). Thioredoxin h is one of the major proteins in rice phloem sap. Planta **195,** 456-463.

Jeter, C.R., and Roux, S.J. (2006). Plant responses to extracellular nucleotides: Cellular processes and biological effects. Purinergic Signal **2,** 443-449.

Jeter, C.R., Tang, W., Henaff, E., Butterfield, T., and Roux, S.J. (2004). Evidence of a novel cell signaling role for extracellular adenosine triphosphates and diphosphates in Arabidopsis. Plant Cell **16,** 2652-2664.

Khan, J.A., Wang, Q., Sjolund, R.D., Schulz, A., and Thompson, G.A. (2007). An early nodulin-like protein accumulates in the sieve element plasma membrane of Arabidopsis. Plant Physiology **143,** 1576-1589.

Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L., and Schwede, T. (2009). The SWISS-

MODEL Repository and associated resources. Nucleic Acids Research **37,** D387-392.

Kouchi, H. and Hata, S. (1993). Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. Molecular and General Genetics **238,** 106-119.

Kozakov, D., Brenke, R., Comeau, S.R., and Vajda, S. (2006). PIPER: an FFT-based protein docking program with pairwise potentials. Proteins **65,** 392-406.

Krell, T., Maclean, J., Boam, D.J., Cooper, A., Resmini, M., Brocklehurst, K., Kelly, S.M., Price, N.C., Lapthorn, A.J., and Coggins, J.R. (2001). Biochemical and X-ray crystallographic studies on shikimate kinase: the important structural role of the P-loop lysine. Protein Science **10,** 1137-1149.

Krugel, U., Veenhoff, L.M., Langbein, J., Wiederhold, E., Liesche, J., Friedrich, T., Grimm, B., Martinoia, E., Poolman, B., and Kuhn, C. (2008). Transport and sorting of the solanum tuberosum sucrose transporter SUT1 is affected by posttranslational modification. Plant Cell **20,** 2497-2513.

Krumm, B., Meng, X., Li, Y., Xiang, Y., and Deng, J. (2008). Structural basis for antagonism of human interleukin 18 by poxvirus interleukin 18-binding protein. Proceeding of the National Academy of Sciences of the United States of America **105,** 20711-20715.

Kuhn, C., Franceschi, V.R., Schulz, A., Lemoine, R., and Frommer, W.B. (1997). Macromolecular trafficking indicated by localization and turnover of sucrose transporters in enucleate sieve elements. Science **275,** 1298-1300.

Kumar, S., Tamura, K., Jakobsen, I.B., and Nei, M. (2001). MEGA2: molecular evolutionary genetics analysis software. Bioinformatics **17,** 1244-1245.

Kumar, S., Nei, M., Dudley, J., and Tamura, K. (2008). MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. Brief Bioinformatics **9,** 299-306.

Lacombe, B., Pilot, G., Michard, E., Gaymard, F., Sentenac, H., and Thibaud, J.B. (2000). A shaker-like K(+) channel with weak rectification is expressed in both source and sink phloem tissues of Arabidopsis. Plant Cell **12,** 837-851.

Lalonde, S., Tegeder, M., Throne-Holst, M., Frommer, W.B., and Patrick, J.W. (2003). Phloem loading and unloading of sugars and amino acids. Plant, Cell and Environment **26,** 37 - 56.

Lalonde, S., Boles, E., Hellmann, H., Barker, L., Patrick, J.W., Frommer, W.B., and Ward, J.M. (1999). The dual function of sugar carriers. Transport and sugar sensing. Plant Cell **11,** 707-726.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson,

J.D., Gibson, T.J., and Higgins, D.G. (2007). Clustal W and Clustal X version 2.0. Bioinformatics **23,** 2947-2948.

Lew, R.R., and Dearnaley, J.D.W. (2000). Extracellular nucleotide effects on the electrical properties of growing Arabidopsis thaliana root hairs. Plant Science **153,** 1-6.

Lytle, B.L., Volkman, B.F., Westler, W.M., and Wu, J.H. (2000). Secondary structure and calcium-induced folding of the Clostridium thermocellum dockerin domain determined by NMR spectroscopy. Archives of Biochemistry and Biophysics **379,** 237-244.

Mandell, J.G., Roberts, V.A., Pique, M.E., Kotlovyi, V., Mitchell, J.C., Nelson, E., I., T., and Ten Eyck, L.F. (2001). Protein docking using continuum electrostatics and geometric fit. Protein Engineering **14,** 105-113.

Martin, S.W., Glover, B.J., and Davies, J.M. (2005). Lipid microdomains--plant membranes get organized. Trends in Plant Science **10,** 263-265.

Mayor, S. and Riezman, H. (2004). Sorting GPI-anchored proteins. Nature Reviews Molecular Cell Biology **5,** 110-120.

McAlister-Henn, L., Gibson, N., and Panisko, E. (1999). Applications of the yeast two-hybrid system. Methods **19,** 330-337.

Meednu, N., Hoops, H., D'Silva, S., Pogorzala, L., Wood, S., Farkas, D., Sorrentino, M., Sia, E., Meluh, P., and Miller, R.K. (2008). The spindle positioning protein Kar9p interacts with the sumoylation machinery in Saccharomyces cerevisiae. Genetics **180,** 2033-2055.

Mongrand, S., Morel, J., Laroche, J., Claverol, S., Carde, J.P., Hartmann, M.A., Bonneu, M., Simon-Plas, F., Lessire, R., and Bessoule, J.J. (2004). Lipid rafts in higher plant cells: purification and characterization of Triton X-100-insoluble microdomains from tobacco plasma membrane. The Journal of Biological Chemistry **279,** 36277-36286.

Nakamura, S., Hayashi, H., Mori, S., and Chino, M. (1993). Protein phosphorylation in the sieve tubes of rice plants. Plant and Cell Physiology **34,** 927-933.

Nersissian, A.M., Immoos, C., Hill, M.G., Hart, P.J., Williams, G., Herrmann, R.G., and Valentine, J.S. (1998). Uclacyanins, stellacyanins, and plantacyanins are distinct subfamilies of phytocyanins: plant-specific mononuclear blue copper proteins. Protein Science **7,** 1915-1929.

Norel, R., Fischer, D., Wolfson, H.J., and Nussinov, R. (1994). Molecular surface recognition by a computer vision-based technique. Protein Engineering **7,** 39-46.

Okumoto, S., Koch, W., Tegeder, M., Fischer, W.N., Biehl, A., Leister, D., Stierhof, Y.D., and Frommer, W.B. (2004). Root phloem-specific expression of the plasma membrane amino acid proton co-transporter AAP3. Journal of

Experimental Botany **55,** 2155-2168.

Palevitz, B.A and Newcomb, E.H. (1971). Ultrastructure and Development of tubular and crystalline P-protein in sieve elements of certain papilionaceous legumes. Protoplasma **72,** 399-426.

Peitsch, M.C. (1995). Protein modeling by e-mail. Nature Biotechnology **13,** 658-660.

Peitsch, M.C., Wells, T.N., Stampf, D.R., and Sussman, J.L. (1995). The Swiss-3DImage collection and PDB-Browser on the World-Wide Web. Trends in Biochemical Sciences **20,** 82-84.

Peters, W.S., van Bel, A.J., and Knoblauch, M. (2006). The geometry of the forisome-sieve element-sieve plate complex in the phloem of Vicia faba L. leaflets. Journal of Experimental Botany **57,** 3091-3098.

Peters, W.S., Knoblauch, M., Warmann, S.A., Pickard, W.F., and Shen, A.Q. (2008). Anisotropic contraction in forisomes: simple models won't fit. Cell Motility and the Cytoskeleton **65,** 368-378.

Pfaffmann, H., Hartmann, E., Brightman, A.O., and Morre, D.J. (1987). Phosphatidylinositol specific phospholipase C of plant stems: membrane associated activity concentrated in plasma membranes. Plant Physiology **85,** 1151-1155.

Plesner, L. (1995). Ecto-ATPases: identities and functions. International Review of Cytology **158,** 141-214.

Radeva, G., and Sharom, F.J. (2004). Isolation and characterization of lipid rafts with different properties from RBL-2H3 (rat basophilic leukaemia) cells. Biochemical Journal **380,** 219-230.

Reinstein, J., Schlichting, I., and Wittinghofer, A. (1990). Structurally and catalytically important residues in the phosphate binding loop of adenylate kinase of Escherichia coli. Biochemistry **29,** 7451-7459.

Roberts, S.A., Weichsel, A., Grass, G., Thakali, K., Hazzard, J.T., Tollin, G., Rensing, C., and Montfort, W.R. (2002). Crystal structure and electron transfer kinetics of CueO, a multicopper oxidase required for copper homeostasis in Escherichia coli. Proceedings of the National Academy of Sciences of the United States of America **99,** 2766-2771.

Roux, S.J. and Steinebrunner, I. (2007). Extracellular ATP: an unexpected role as a signaler in plants. Trends in Plant Science **12,** 522-527.

Ryden, L.G. and Hunt, L.T. (1993). Evolution of protein complexity: the blue copper-containing oxidases and related proteins. Journal of Molecular Evolution **36,** 41-66.

Sander, C. and Schneider, R. (1991). Database of homology-derived protein structures

and the structural meaning of sequence alignment. Proteins **9,** 56-68.

Sardar, H.S. and Showalter, A.M. (2007). A cellular networking model involving interactions among glycosylphosphatidylinositol (GPI)-anchored plasma membrane arabinogalactan proteins (AGPs), microtubules and F-actin in tobacco BY-2 cells. Plant Signaling and Behavior **2,** 8-9.

Satishchandran, C., Hickman, Y.N., and Markham, G.D. (1992). Characterization of the phosphorylated enzyme intermediate formed in the adenosine 5'-phosphosulfate kinase reaction. Biochemistry **31,** 11684-11688.

Sauer, H., Hescheler, J., and Wartenberg, M. (2000). Mechanical strain-induced Ca^{2+} waves are propagated via ATP release and purinergic receptor activation. American Journal of Physiology-Cell Physiology **279,** C295-307.

Scheres, B., van Engelen, F., van der Knaap, E., van de Wiel, C., van Kammen, A., and Bisseling, T. (1990). Sequential induction of nodulin gene expression in the developing pea nodule. Plant Cell **2,** 687-700.

Schultz, C., Gilson, P., Oxley, D., Youl, J., and Bacic, A. (1998). GPI-anchors on arabinogalactan-proteins: implications for signalling in plants. Trends in Plant Science **3,** 426-431.

Schulz, A. (1998). Phloem, structure related to function. Progress in Botany **59,** 429-475.

Schulz, A. and Thompson, G.A. (2009) Phloem Structure and Function. *In*: Encyclopedia of Life Sciences (ELS). John Wiley & Sons, Ltd: Chichester.

Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. (2003). SWISS-MODEL: An automated protein homology-modeling server. Nucleic Acids Research **31,** 3381-3385.

Scutt, C.P., Vinauger-Douard, M., Fourquin, C., Ailhas, J., Kuno, N., Uchida, K., Gaude, T., Furuya, M., and Dumas, C. (2003). The identification of candidate genes for a reverse genetic analysis of development and function in the Arabidopsis gynoecium. Plant Physiology **132,** 653-665.

Sekimizu, K., Bramhill, D., and Kornberg, A. (1987). ATP activates dnaA protein in initiating replication of plasmids bearing the origin of the E. coli chromosome. Cell **50,** 259-265.

Song, C.J., Steinebrunner, I., Wang, X., Stout, S.C., and Roux, S.J. (2006). Extracellular ATP induces the accumulation of superoxide via NADPH oxidases in Arabidopsis. Plant Physiology **140,** 1222-1232.

Stadler, R., Brandner, J., Schulz, A., Gahrtz, M., and Sauer, N. (1995). Phloem Loading by the PmSUC2 Sucrose Carrier from Plantago major Occurs into Companion Cells. Plant Cell **7,** 1545-1554.

Stites, W.E. (1997). Proteinminus signProtein Interactions: Interface Structure, Binding

Thermodynamics, and Mutational Analysis. Chemical Reviews **97,** 1233-1250.

Stoddart, A., Dykstra, M.L., Brown, B.K., Song, W., Pierce, S.K., and Brodsky, F.M. (2002). Lipid rafts unite signaling cascades with clathrin to regulate BCR internalization. Immunity **17,** 451-462.

Szederkenyi, J., Komor, E., and Schobert, C. (1997). Cloning of the cDNA for glutaredoxin, an abundant sieve-tube exudate protein from Ricinus communis L and characterisation of the glutathione-dependent thiol-reduction system in sieve tubes. Planta **202,** 349-356.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular Biology and Evolution **24,** 1596-1599.

Tang, W., Brady, S.R., Sun, Y., Muday, G.K., and Roux, S.J. (2003). Extracellular ATP inhibits root gravitropism at concentrations that inhibit polar auxin transport. Plant Physiology **131,** 147-154.

Tao, H., Liu, W., Simmons, B.N., Harris, H.K., Cox, T.C., and Massiah, M.A. (2010). Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS. Biotechniques **48,** 61-64.

Teufel, A., Krupp, M., Weinmann, A., and Galle, P.R. (2006). Current bioinformatics tools in genomic biomedical research. International Journal of Molecular Medicine **17,** 967-973.

Thompson, G.A. and Schulz, A. (1999). Macromolecular trafficking in the phloem. Trends in Plant Science **4,** 354-360.

Torsch, J. and Esau, K. (1981). Changes in the endoplasmic reticulum during differentiation of a sieve element in *Gossypium hirsutum*. Journal of Ultrastructure Research **74,** 183-194.

Toth, K.F., Wang, Q., and Sjolund, R.D. (1994). Monoclonal-antibodies against phloem P-protein from plant-tissue cultures. I. microscopy and biochemical-analysis. American Journal of Botany **81,** 1370-1377.

Turgeon, R. and Medville, R. (2004). Phloem loading. A reevaluation of the relationship between plasmodesmatal frequencies and loading strategies. Plant Physiology **136,** 3795–3803.

Turgeon, R., Beebe, D.U., and Gowan , E. (1993). The intermediary cell: minor-vein anatomy and raffinose oligosaccharide synthesis in the Scrophulariaceae. Planta **191,** 446-456.

Ubbink, M., Ejdeback, M., Karlsson, B.G., and Bendall, D.S. (1998). The structure of the complex of plastocyanin and cytochrome f, determined by paramagnetic NMR and restrained rigid-body molecular dynamics. Structure **6,** 323-335.

Van Bel, A.J.E. (1996). Interaction between sieve element and companion cell and the consequences for photoassimilate distribution. Two structural hardware frames with associated physiological software packages in dicotyledons? Journal of Experimental Botany **47,** 1129-1140.

Varma, R. and Mayor, S. (1998). GPI-anchored proteins are organized in submicron domains at the cell surface. Nature **394,** 798-801.

Vijn, I., Martinez-Abarca, F., Yang, W.C., das Neves, L., van Brussel, A., van Kammen, A., and Bisseling, T. (1995). Early nodulin gene expression during Nod factor-induced processes in Vicia sativa. Plant Journal **8,** 111-119.

Vitkup, D., Melamud, E., Moult, J., and Sander, C. (2001). Completeness in structural genomics. Nature Structural Biology **8,** 559-566.

Walker, J.E., Saraste, M., Runswick, M.J., and Gay, N.J. (1982). Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. EMBO Journal **1,** 945-951.

Walz, C., Juenger, M., Schad, M., and Kehr, J. (2002). Evidence for the presence and activity of a complete antioxidant defence system in mature sieve tubes. The Plant Journal **31,** 189-197.

Walz, C., Giavalisco, P., Schad, M., Juenger, M., Klose, J., and Kehr, J. (2004). Proteomics of curcurbit phloem exudate reveals a network of defence proteins. Phytochemistry **65,** 1795-1804.

Wergin, W.P. and Newcomb, E.H. (1970). Formation and dispersal of crystalline P-protein in sieve elements of soybean (*Glycine max* L.). Protoplasma **71,** 365-388.

Will, T., Tjallingii, W.F., Thonnessen, A., and van Bel, A.J. (2007). Molecular sabotage of plant defense by aphid saliva. Proceedings of the National Academy of Sciences of the United States of America **104,** 10536-10541.

Xie, Y., Inoue, T., Miyamoto, Y., Matsumura, H., Kataoka, K., Yamaguchi, K., Suzuki, S., and Kai, Y. (2003). Crystallization and preliminary X-ray crystallographic studies of mavicyanin from Cucurbita pepo medullosa. Acta Crystallographica. Section D, Biological Crystallography **59,** 1474-1476.

Xie, Y., Inoue, T., Miyamoto, Y., Matsumura, H., Kataoka, K., Yamaguchi, K., Nojini, M., Suzuki, S., and Kai, Y. (2005). Structural reorganization of the copper binding site involving Thr15 of mavicyanin from Cucurbita pepo medullosa (zucchini) upon reduction. The Journal of Biochemistry **137,** 455-461.

Xoconostle-Cazares, B., Xiang, Y., Ruiz-Medrano, R., Wang, H.L., Monzer, J., Yoo, B.C., McFarland, K.C., Franceschi, V.R., and Lucas, W.J. (1999). Plant paralog to viral movement protein that potentiates transport of mRNA into the phloem. Science **283,** 94-98.

Yoo, B.C., Lee, J.Y., and Lucas, W.J. (2002). Analysis of the complexity of protein kinases within the phloem sieve tube system. Characterization of Cucurbita maxima calmodulin-like domain protein kinase 1. The Journal of the Biological Chemistry **277,** 15325-15332.

Yoshimoto, N., Inoue, E., Saito, K., Yamaya, T., and Takahashi, H. (2003). Phloem-localizing sulfate transporter, Sultr1;3, mediates re-distribution of sulfur from source to sink organs in Arabidopsis. Plant Physiology **131,** 1511-1517.

Zabotin, A.I., Barysheva, T.S., Trofimova, O.I., Lozovaya, V.V., and Widholm, J. (2002). Regulation of callose metabolism in higher plant cells in vitro. Russian Journal of Plant Physiology **49,** 792-798.

Zurzolo, C., van Meer, G., and Mayor, S. (2003). The order of rafts. Conference on microdomains, lipid rafts and caveolae. EMBO Reports **4,** 1117-1121.

VITA

NING SONG

Candidate for the Degree of

Master of Science

Thesis: CHARACTERIZATION OF THE ARABIDOPSIS SIEVE ELEMENT-
SPECIFIC EARLY NODULIN PROTEIN

Major Field: Biochemistry and Molecular Biology

Biographical Information:

   Education:

   Completed the requirements for the Master of Science Degree in Biochemistry
   and Molecular Biology, Oklahoma State University, Stillwater, Oklahoma. July,
   2010.

   Bachelor of Science Degree in Biology. Shandong Normal University, Jinan,
   People's Republic of China. June, 2007.

   Experience:

   Graduate Research Assistant. Department of Biochemistry and Molecular
   Biology, Oklahoma State University. August, 2007- August, 2010.

   Undergraduate Research Assistant. College of Life Science, Shandong Normal
   University. April, 2007-June, 2007.

   Honors and Awards:

   The Graduate Research Fellowship. College of Agricultural Sciences and
   Natural Resources (CASNR) at OSU. August, 2007-April, 2008

   Professional Memberships:

   Member, OSU BMB Graduate Student Association from August, 2007-present

   Publication:
   Anstead J., Samuel P., Song N., Wu C., Thompson G. and Goggin F. (2009).
   Activation of ethylene-related genes in response to aphid feeding on resistant
   and susceptible melon and tomato plants. Entomologia experimentalis et
   applicata. **134,** 170-181

Name: NING SONG   Date of Degree: July, 2010

Institution: Oklahoma State University   Location: Stillwater, Oklahoma

Title of Study: CHARACTERIZATION OF THE ARABIDOPSIS SIEVE ELEMENT-
                 SPECIFIC EARLY NODULIN PROTEIN

Pages in Study: 97   Candidate for the Degree of Master of Science

Major Field: Biochemistry and Molecular Biology

Scope and Method of Study: Sieve elements are the terminally differentiated conducting cells in the phloem tissue of higher plants. These highly modified cells have a well-developed endomembrane system that is important in regulating physiological events associated with cellular responses and long-distance transport. To provide insights into these regulatory events, studies were designed to understand functional aspects of the *Arabidopsis* sieve element-specific early nodulin-like protein (SE-ENOD). The SE-ENOD is glycosylphosphatidylinositol (GPI)-anchored in the sieve element plasma membrane and is thought to extend into the extracellular matrix. Three approaches were taken to further characterize the structure and function of this protein: *in silico* structural prediction studies, *in vitro* microbial expression studies, and *in vivo* protein isolation studies.

Findings and Conclusions: *In silico* structural predictions suggested that the overall tertiary structure of the SE-ENOD is a Greek key topology arranged to one side of two β-strands and two α-helixes. The Pocket-Finder program predicted a potential ATP binding pocket near the site of a P-loop sequence in domain II of the SE-ENOD. The ClusPro program predicted that the SE-ENOD has the structural features that would allow homodimerization to occur at the exposed surfaces of domain II. Two microbial expression systems were used to test the *in silico* structural predictions. Purified recombinant SE-ENOD peptides expressed in *E. coli* demonstrated ATP binding activity that through mutant analysis was shown to correspond to ATP binding motifs in domain II. In contrast, yeast two hybrid experiments failed to demonstrate intermolecular interactions between SE-ENOD peptides. A novel technique to isolate phloem-enriched tissues from broccoli plants was developed to provide large amounts of phloem tissues for protein extractions. Immunoblot analyses with a SE-ENOD-specific monoclonal antibody detected the protein in detergent extractions of the phloem-enriched tissues as several species with molecular weights corresponding to the SE-ENOD monomer, dimer, and tetramer conformations. The results of these studies are consistent with a functional model for the SE-ENOD as an extracellular ATP-binding protein attached to the sieve element plasma membrane by a GPI anchor. Membrane microdomains within the plasma membrane could influence the conformation and function of the SE-ENOD.

ADVISER'S APPROVAL: Dr. Gary A. Thompson