Identification of the CEP1 peptide receptor in *Medicago truncatula*

Han Chung Lee

Doctor of Philosophy

The Australian National University

Research School of Biology

07/2020

Table of Contents

Acknowledgements

I would like to dedicate this thesis to the numerous people who have provide support, understanding and professional advice that resulted in the completion this thesis. First, I would like to thank my family to support me through all my academic pursuits and understanding through tough times during my Ph.D. I especially thank my wife, Phoebe Wong for her support.

I sincerely thank Michael for supervising this study, John Rathjan, Nijat Imin, Amani Batareh for all the scientific suggestions and supports. I also thank Adam Carroll, Joseph Boileau from ANU mass spec facility, Ben Crossett, Angela Connolly from Sydney mass spec facility for all the MS support.

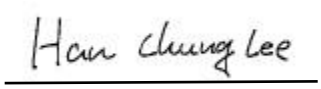
I thank Neha Patel and Simon Williams for providing technical support on protein purification and biological activity assay, Daryl Webb, Rosemary White and The Centre for Advanced Microscopy for providing technical assistance on confocal microscopy. and Steve Binos for providing technical advice on MS and PD search.

Finally, I thank for the support from Taiwan-ANU Scholarship and Australian Research Council grant to MAD (DP150104250) for all this work.

Declaration

This thesis is my own work and does not contain results that have been generated by persons other than myself, except where due the reference and acknowledgement has been made. Results presented in this thesis have not been used for the award of any other degree at any other university.

Han Chung Lee



Research School of Biology

The Australian National University

July 2020

Thesis Abstract

Peptide hormones can mediate the communication between plant cells by binding their corresponding plasma membrane-localized receptor and inducing downstream effects. Among the peptide hormones, CEPs (C-TERMINALLY ENCODED PEPTIDEs), play multiple roles in plant growth and development. In *Medicago truncatula* (Medicago), the receptor that interacts with CEP peptides remains unknown. Thus, this thesis aimed to identify the CEP receptor directly via bottom up proteomics and by using *in vivo* ligand-receptor binding assays in CEP responding and non-responding backgrounds.

First, plant transmembrane proteins (TMPs) are essential for normal cellular homeostasis, nutrient exchange and responses to environmental cues. Commonly used bottom-up proteomic approaches fail to identify a broad coverage of peptide fragments derived from TMPs. Here, we used mass spectrometry to compare the effectiveness of two solubilization and protein cleavage methods to identify shoot-derived TMPs from the legume Medicago, including the putative CEP receptor, COMPACT ROOT ARCHITECTURE 2 (CRA2). We compared a urea solubilisation, trypsin Lys-C (UR-TLC) cleavage method to a formic acid solubilisation, cyanogen bromide and trypsin Lys-C (FA-CTLC) cleavage method. We assessed the effectiveness of these methods by (i) comparing total protein identifications, (ii) determining how many TMPs were identified, and (iii) defining how many peptides incorporate all, or part, of transmembrane domains (TMD) sequences. The results show that the FA-CTLC method identified nine-fold more TMDs, and enriched more hydrophobic TMPs than the UR-TLC method. FA-CTLC identified more TMPs, particularly transporters, whereas, UR-TLC preferentially identified TMPs with one TMD, particularly signaling proteins. We identified 8993 protein groups including 3289 TMPs in young shoot tissues, which accounts for around 36 % of the predicted Medicago TMPs, however, we were unable to identify CRA2 with a high degree of certainty. Nevertheless, the results suggest that combining plant membrane purification techniques with both the FA-CTLC and UR-TLC methods will achieve a more complete identification and coverage of TMPs.

Second, we use formaldehyde or photo-activation to cross-link fluorescently tagged group 1 or group 2 CEPs to receptors in semi-purified Medicago or *Arabidopsis thaliana* (Arabidopsis) leaf vascular tissue. In the beginning, we synthesized the CEP peptides and their modified derivatives and tested their biological activities to confirm that the synthetic peptide still binds its receptor. Formaldehyde cross-linked fluorescently (FITC)-tagged Medicago group 1 CEP (MtCEP1) bound to a specific subset of vascular tissue cells in wild type leaves, but not those from Arabidopsis or Medicago CEP receptor mutants. Binding competition showed that unlabeled MtCEP1 displaced the FITC-MtCEP1 from the CEP receptor *in vivo*. By contrast, the group 2 CEP, FITC-AtCEP12, which has several divergent amino acids to group 1 CEPs, bound to vascular tissue cells independently of the confirmed Arabidopsis CEP receptor, CEPR1, or the putative Medicago CEP receptor, CRA2. The crosslinking of photo-activated FITC-MtCEP1 to the periphery of vascular tissue cells suggested that the putative Medicago CEP receptor, CRA2, localized to the plasma membrane. We visualized the photo-cross-linking of FITC-MtCEP1 to CRA2 using SDS-PAGE and fluorescence-detection. Mass spectrometry analysis identified CRA2-specific peptides in this band. Collectively these results indicate that FITC-MtCEP1 binds to CRA2, and that MtCRA2 and AtCEPR1 are functionally similar. The use of formaldehyde or photo-activation to crosslink biologically-active, fluorescently-tagged ligands to receptors is a simple technique that may find wide utility to demonstrate ligand receptor pairings in specific tissues, augment genetic evidence supporting ligand-receptor pairing, or identify ligand binding sites.

Chapter 1: Introduction

1.1 Plant peptide hormones regulate growth

Peptide hormones are short polypeptides that consist of 5-60 amino acids (Tager and Steiner, 1974). In the past two decades peptide hormones have been recognized as key elements that plants use for short and long-distance cell-cell interactions. About 20 different types of secreted peptide hormones are known to control specific growth and development responses in plants (Fig. 1.1) (Czyzewicz et al., 2013; Tavormina et al., 2015; Breiden and Simon, 2016; Roy et al., 2018). These include Rapid Alkalinisation Factor (RALF) peptides and its receptor Feronia (FER) and Lorelei (LRE)-Like Glycosylphosphatidylinositol (GPI)-anchored protein 1 (LLG1) for immune signalling, Phytosulfokine (PSK) and its receptor PSKR1 for growth regulation, Shafer's Classic Reproductions (SCR/SP22) and its receptor S locus receptor kinase (SRK) for prevent inbreeding, CLAVATA 3 (CLV3) and its receptors CLAVATA 1 (CLV1), CLAVATA 2 (CLV2), and CORYNE for shoot development, as well as C-terminally encoded peptide (CEP) and its receptor CEPR1 (Matsubayashi and Sakagami, 2006; Ohyama et al., 2008; Djordjevic et al., 2011; Stahl and Simon, 2012; Delay et al., 2013; Imin et al., 2013; Roberts et al., 2013; Tabata et al., 2014; Djordjevic et al., 2015; Mohd-Radzman et al., 2015; Mohd-Radzman et al., 2016; Ohkubo et al., 2017; Imin et al., 2018; Kereszt et al., 2018; Chapman et al., 2019; Delay et al., 2019; Xiao et al., 2019; Chapman et al., 2020). Most peptide hormones are secreted to the apoplast where they are thought to interact with plasma membrane located receptors on the surfaces of nearby cells (Franssen and Bisseling, 2001; Murphy et al., 2012) or cells at great distance from the site of ligand production, which has been demonstrated with certain species of CLE and CEP peptides (Huault et al., 2014; Tabata et al., 2014; Mohd-Radzman et al., 2015; Notaguchi and Okamoto, 2015; Okamoto et al., 2015; Mohd-Radzman et al., 2016; Patel et al., 2018; Chapman et al., 2019; Laffont et al., 2019; Chapman et al., 2020; Gautrat et al., 2020). The class XI of leucine rich repeat receptor like kinases (LRR-RLKs) are thought to preferentially interact with small (12 -18 amino acid) secreted peptide hormones (Chakraborty et al., 2019).

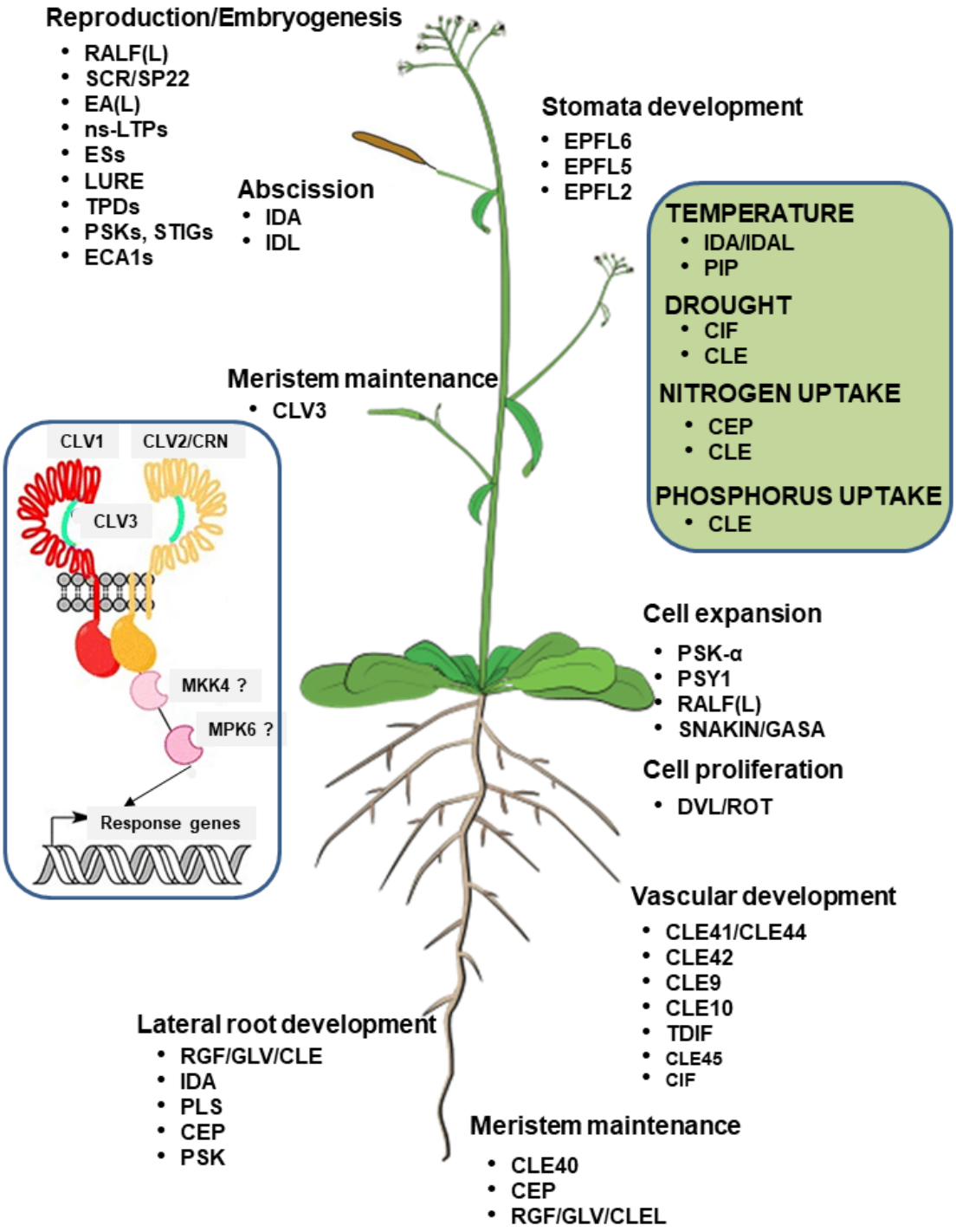


Figure 1.1 Signaling peptides implicated in specific plant developmental processes are indicated (diagram based on Czyzewicz et al., 2013). The inner diagram illustrated how CLV3 binds to its receptor, CLV1 and CLV2 to trigger downstream signal pathway for meristem development potentially through MAP kinases (Breiden and Simon, 2016).

.

1.2 C-TERMINALLY ENCODED PEPTIDEs (CEPs) and the function of CEPs

My thesis will focus on a family of proline-rich signalling peptides, called *C*-TERMINALLY ENCODED PEPTIDEs (CEPs). CEP peptide encoding multigene families occur universally in the genomes of flowering and seed-bearing plants (Ogilvie et al., 2014). Most CEP genes encode a short prepropeptide of about 100 amino acids, but some CEP genes can have 2 or more 15 amino acid CEP domains and encode larger precursors. These precursors encode an *N*-terminal signal peptide, a variable domain, one or more CEP domains and, often but not always, a short *C*-terminal extension (Delay et al., 2013; Imin et al., 2013; Ogilvie et al., 2014) (Fig 1.2 A). Upon *N*-terminal processing as the product transitions into the ER, the precursor is further processed by unknown proteases (most likely located in the Golgi apparatus) to liberate 15 amino acid, secreted peptide(s) corresponding to the CEP domain (Ohyama et al., 2008; Mohd-Radzman et al., 2015; Ogilvie et al., 2014). During processing one or more of the proline residues is/are hydroxylated by prolyl-4-hydroxylases (P4Hs) (Fig. 1.2A), which are encoded by 13 genes in Arabidopsis (Ohyama et al., 2008; Tabata et al., 2014; Mohd-Radzman et al., 2015) or further modified by glycosyl residues catalysed by glycosyl transferases (Mohd-Radzman et al., 2015). P4Hs and glycosyl transferases are located in the Golgi (Ogilvie et al., 2014). In *Medicago truncatula* and *Arabidopsis thaliana* respectively there are 13 and 12 group 1 CEP genes (MtCEP1-13; AtCEP1-12), and 5 and 3 so-called group 2 CEP genes (MtCEP14-17, MtCEP19; AtCEP13-15) (Fig. 1.2 B&C) that encode one or more CEP peptides. Medicago CEP1 (MtCEP1), which encodes two group 1 CEP peptides, controls root growth by reducing the number of lateral roots and increasing root nodule number (Imin et al., 2013; Mohd-Radzman et al., 2015; Mohd-Radzman et al., 2016), whereas Arabidopsis CEP3, a group 1 CEP peptide, decreases main root and lateral root growth in Arabidopsis (Delay et al., 2013; Chapman et al., 2019; Delay et al., 2019; Chapman et al., 2020). In Arabidopsis and legumes, CEPs can enter the xylem stream (Tabata et al.,

2014; Ohkubo et al., 2017; Patel et al., 2018). CEP genes are induced by several environmental conditions, including low nitrogen (N), or N starvation, hypoxia, low sulphate, low phosphate, low iron, high CO_2 and sucrose, and osmotic stress caused by high levels of salt or mannitol (Imin et al., 2013; Chapman et al., 2019) (green box in Fig 1.1). The role of group 2 CEPs remains unknown. This thesis addresses functional aspects of group 1 and group 2 CEPs.

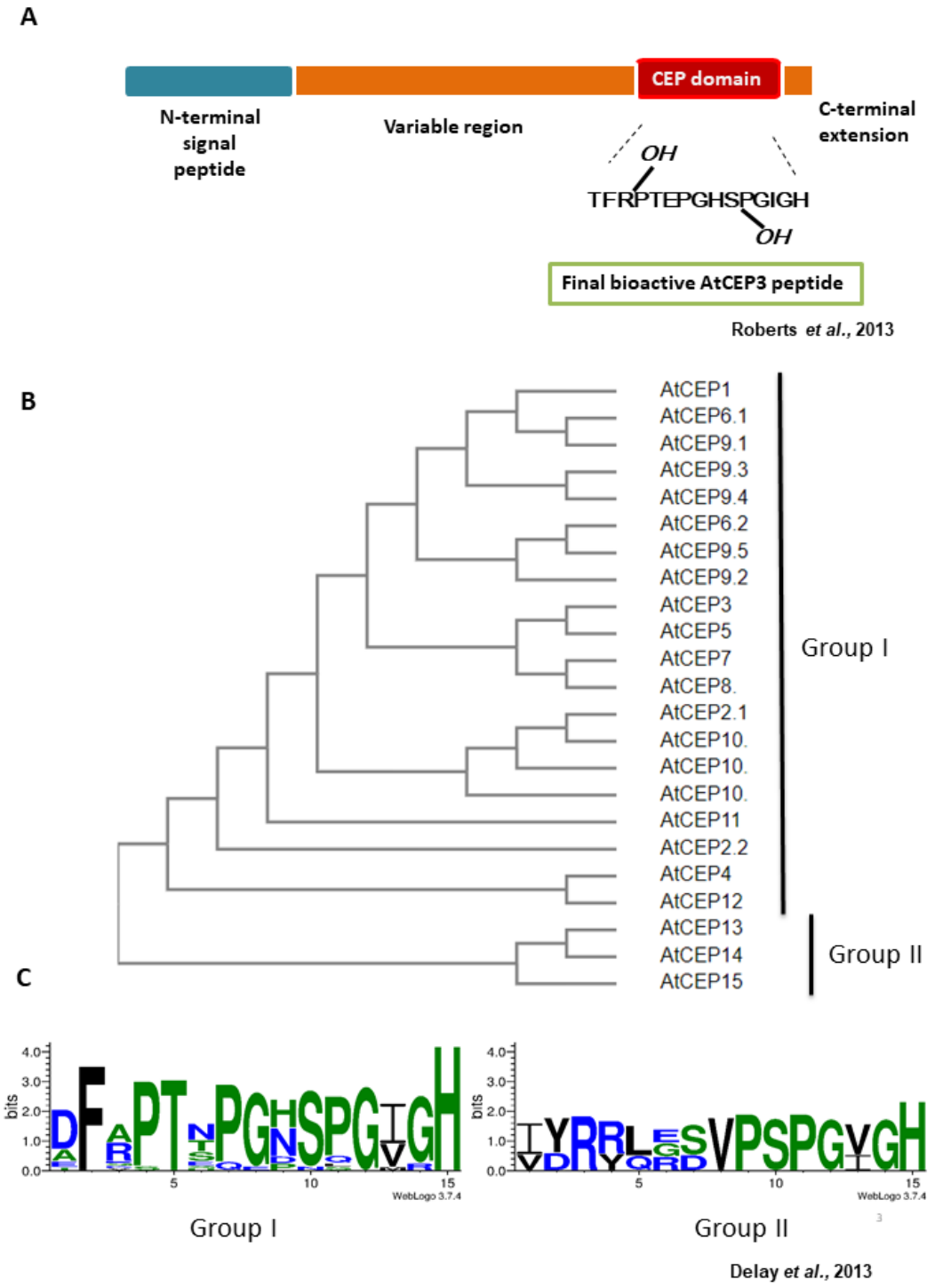


Figure 1.2. The CEP family groups into two groups based on CEP domain. (A) The biological active AtCEP3 peptide is a fifteen amino acid long peptide commonly hydroxylated at proline at position 4 and 11. (B) Phylogenetic tree based on CEP domain sequences (C) Weblogo representation of group I and group II CEP domains. Group 2 CEP domains show *C*-terminal conservation with group 1 CEP domains but differ in *N*-terminal residues and are the most distantly related group of CEPs.

CEPs play a central role in controlling major aspects controlling root development, root system architecture, root nodule formation and nitrogen (N) demand signalling auxin transport and resource allocation supporting yield (Ohyama et al., 2008; Delay et al., 2013; Imin et al., 2013; Roberts et al., 2013; Tabata et al., 2014; Mohd-Radzman et al., 2016; Ohkubo et al., 2017; Taleski et al., 2018; Chapman et al., 2019; Delay et al., 2019; Chapman et al., 2020; Taleski et al., 2020) where they mediate local and systemic interactions (Huault et al., 2014; Tabata et al., 2014; Mohd-Radzman et al., 2016; Chapman et al., 2019; Laffont et al., 2019; Chapman et al., 2020). Differences in the expression pattern of, and phenotypic responses to AtCEP1 and MtCEP1 in Arabidopsis and Medicago species respectively, suggests CEPs contribute to root system plasticity differently in divergent plant species (Delay et al., 2013; Roberts et al., 2013; Tabata et al., 2014; Chapman et al., 2019; Chapman et al., 2020; Ogilvie et al., 2014)ab . AtCEP1 was first described as a negative root growth regulator in Arabidopsis (Ohyama et al., 2008). By contrast, the application of the MtCEP1 peptide to Medicago not only decreased root growth but also increased rhizobia infection rates, nodule number and the production of “fused” nodules. In fused nodules, two developmentally distinct nodule foci in close proximity to one another fuse together (Mohd-Radzman et al., 2016). in Medicago, a low nitrogen condition will trigger the MtCEP1 expression in root as shown in Fig. 1.3. The MtCEP1 will travel to the shoot interact with its putative receptor, MtCRA2 (Compact Root Architecture 2), to trigger micro-RNA miR2111 expression. miR2111 targets the transcripts of downstream genes TML1 (Too Much Love 1) and TML2 (Too Much Love 2), which encode F-box Kelch repeat proteins in the root that negatively regulate nodule number (Gautrat et al., 2020). In Arabidopsis, like Medicago, the low nitrogen conditions will trigger the production of AtCEPs in the root, which travel to the shoot and interact to the CEPR1. As a consequence, this triggers the production of the type III glutaredoxins, CEPD1 and CEPD2, which travel to the root to regulate *NRT 2.1* expression (Ohkubo et al., 2017).

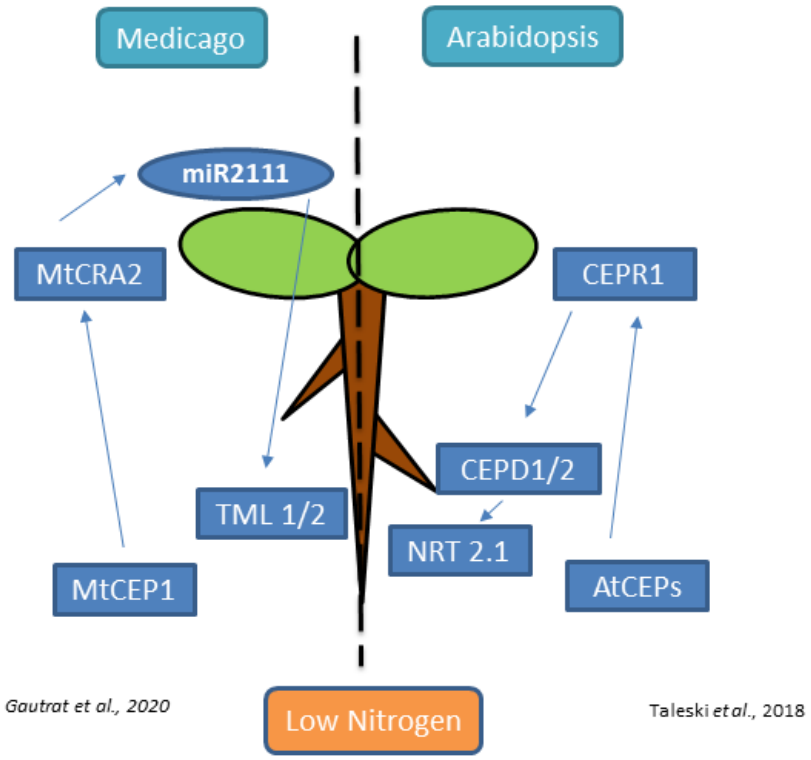


Figure 1.3. CEP-CEPR pathways control plant root development in Arabidopsis and Medicago

1.3 The receptor of CEPs in Arabidopsis and Medicago

CEPs control root growth and development by interacting with membrane-bound receptors on target cells (Tabata et al., 2014; Mohd-Radzman et al., 2016), which are the leucine rich repeat receptor like kinases (LRR-RLKs). LRR-RLKs play critical roles in plant growth, morphogenesis, disease immunity responses and stress regulation (Mitra et al., 2015). LRR-RLKs form a large family of membrane bound receptor consisting of an extracellular domain connected by a single-pass transmembrane sequence to a cytoplasmic kinase domain (Fig 1.4). It is thought that the auto-phosphorylation of specific Ser and/or Thr and/or Tyr residues in the kinase domain or the juxtamembrane domain is often critical for the activation of LRR-RLKs family members. Arabidopsis encodes 220 LRR-RLKs (Shiu and Bleecker, 2001) subdivided into multiple sub-families. It is hypothesised that 100 of these Arabidopsis LRR-RLKs interact with different classes of secreted peptides (Matsubayashi, 2018) to mediate a plethora of physiological processes central to plant growth, development, and defence (Matsubayashi and Sakagami, 2006).

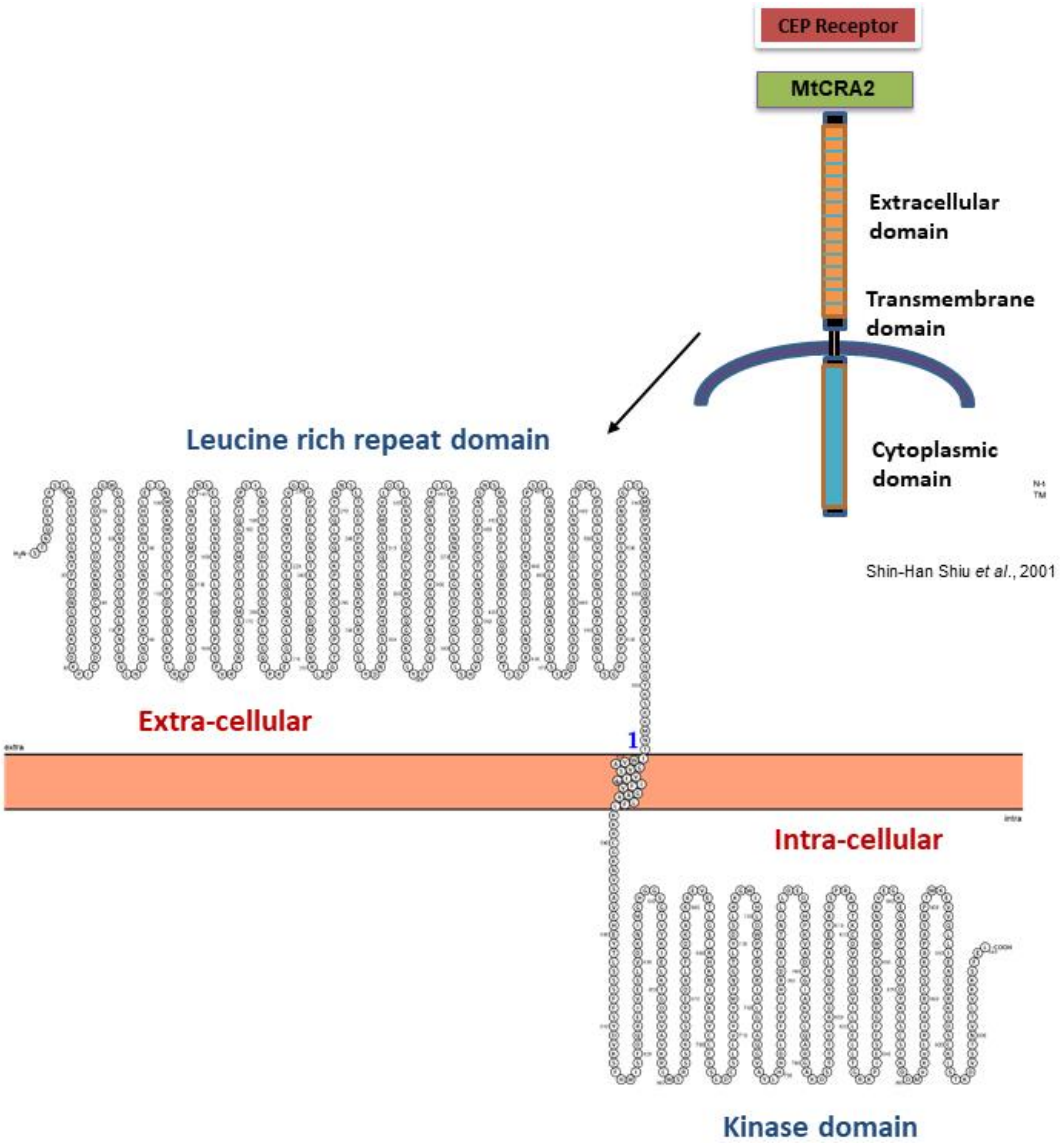


Fig 1.4. The schematic diagrams of MtCRA2

The schematic diagrams were made using the Protter online tool (Omasits et al., 2014) (http://wlab.ethz.ch/protter/start/).

There are 13 different classes of LRR-RLKs in Arabidopsis (Ten Hove et al., 2011). A previous study identified two CEP receptors (CEPRs) that bind CEP peptides in Arabidopsis, XYLEM INTERMIXED WITH PHLOEM 1 (XIP1) /CEP RECEPTOR1 (CEPR1) (At5g49660) and CEPR2 (At1g72180) (Tabata et al., 2014) which both belong to class XI. In this study, the binding to radio-labelled CEP peptide could be competed off the CEPR1 receptor by several non-radiolabelled CEP peptides (Tabata et al., 2014) indicating CEPR1 is the receptor for several CEP peptides. Mutation of AtCEPR1 was also shown to disrupt nitrate-uptake machinery (Tabata et al., 2014) and a more branched

root architecture was observed in Arabidopsis (Fig. 1.5 A) (Chapman et al., 2019) and in Medicago (Fig. 1.5 B) (Huault et al., 2014). Additionally, COMPACT ROOT ARCHITECTURE 2 (CRA2), was proposed to be a XIP1/ CEPR1 homologue which expressed in the vascular tissue of roots and shoots (Laffont et al., 2019). Previous studies showed MtCRA2 being phloem localized as suggested by *in situ* gene expression studies (Fig. 1.6 A) (Huault et al., 2014) and CEPR1 was identified as being expressed mainly in vasculature (Fig. 1.6 B) (Taleski et al., 2020). Since Medicago MtCRA2 is homologous to CEPR1 (Huault et al., 2014) and *MtCRA2 /CEPR1* mutants are insensitive to CEP peptides, MtCRA2 has been proposed to be a CEP receptor in Medicago (Mohd-Radzman et al., 2016). These results suggest that MtCRA2 and CEPR1 are functionally similar, however the receptor-peptide interaction in Medicago is yet to be validated.

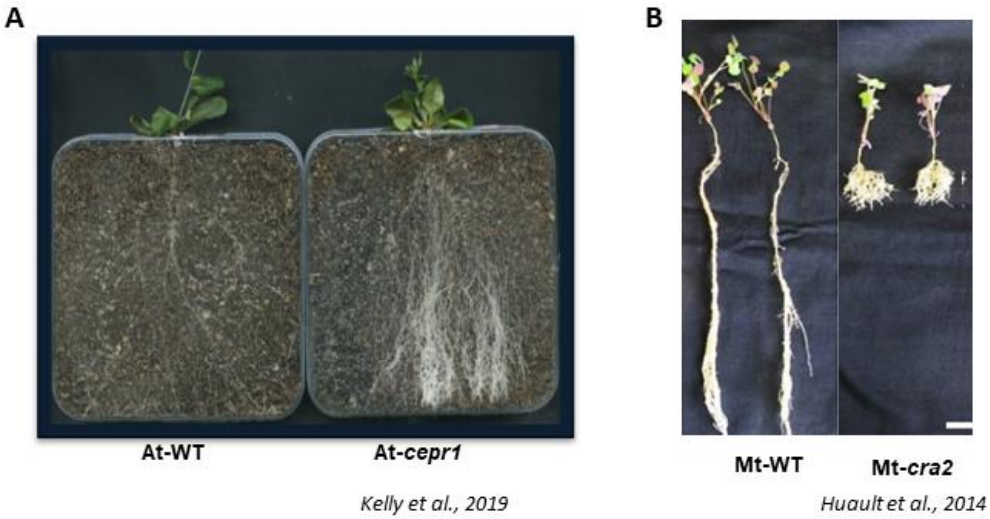


Figure 1.5. Phenotypes of CEP receptor mutants.

(A) Arabidopsis cepr1 mutants have more branches roots than wild-type plants. (B) The Medicago “COMPACT ROOT ARCHITECTURE 2” mutant (*cra2*) has short roots and more lateral roots that are independent of the growth conditions.

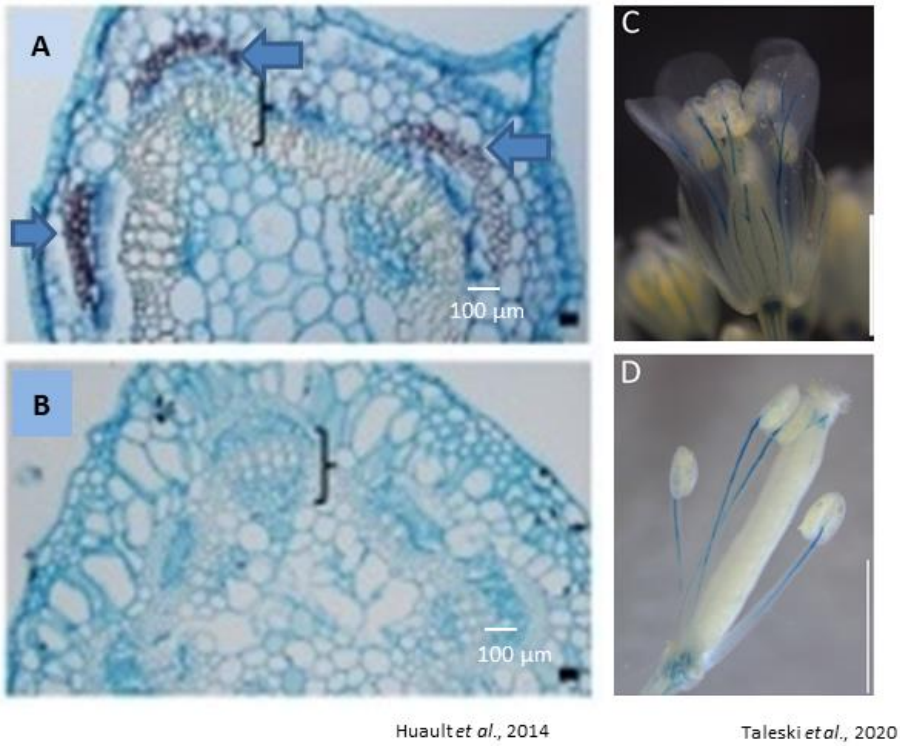


Figure 1.6. CEPR expressed in vasculature in Medicago and Arabidopsis. *In situ* hybridization of the *MtCRA2* transcripts in stem transversal sections (A) With an anti-sense probe (purple signal) (B) with a sense probe used as a negative control. GUS staining of *CEPR1* expression in the Arabidopsis reproductive organs. (C) stage 12 gynoecium (D) mature silique. Scale bars in (C) (D)= 1 mm

Many receptor-peptide ligands pairs remain to be elucidated. For example, several recently identified novel xylem-mobile peptides, termed XAPs, do not yet have receptor partners (Okamoto et al., 2015; Patel et al., 2018). Biochemical binding assays, bioassay guided purification, isothermal titration calorimetry, gel filtration and genetic approaches are effective at identifying peptide ligand-receptor pairings (Tabata et al., 2014; Santiago et al., 2016; Song et al., 2016), however, each technique has limitations. Therefore, there is a need to devise simple, complementary, and rapid methods for identifying or validating peptide hormone-receptor pairings *in vivo*.

1.4 CEP receptor may need a co-receptor to form stable binding complex

The signal transduction through LRR-RLKs is thought to initiate with the peptide ligand binding to the corresponding receptor. After ligand binding the extracellular domain of the LRR-RLKs, the receptors initiate a ligand-driven stable complex, homodimerically or heterodimerically (Li et al., 2002; Sun et al., 2013; Meng et al., 2015; Wang et al., 2015; Meng et al., 2016; Song et al., 2016; Zhang et al., 2016; Li et al., 2017; Hu et al., 2018).

In Arabidopsis, several cases showed the LRR-RKs signalling involves complexes with co-receptor(s). This ligand-receptor-co receptor complex such as BRI1 (brassinosteroid insensitive 1) can work with BAK1/SERK3 (brassinosteroid insensitive 1-associated receptor kinase 1/ somatic embryogenesis receptor-like kinases 3) (Hothorn et al., 2011; Sun et al., 2013) to form receptor complex for the binding of the brassinosteroid ligand. The ligand-receptor-coreceptor complex can initiate the transphosphorylation of the kinase domains to initiate a downstream signal pathway(Sun et al., 2013; Bojar et al., 2014). The peptide ligand IDA binds to its receptor HAESA, also a LRR-RLK, to trigger floral organ abscission, and this ligand-receptor complex requires SERK1 (Somatic embryogenesis receptor kinase 1) as a co-receptor to form stable binding (Santiago et al., 2016). In this study, they demonstrated the ligand-receptor-co receptor complex structure by crystalizing the ectodomain of HAESA, SERK1 with IDA (Fig. 1.7). The co-receptor SERK1 can significantly increase the binding efficiency between HAESA and IDA. Based on the prevalence of SERK as a coreceptor for several peptide-receptor interactions, it is possible that there is a co-receptor to stabilize the MtCRA2-MtCEP1 complex. However, the co-receptor for CEPR1 is still unclear.

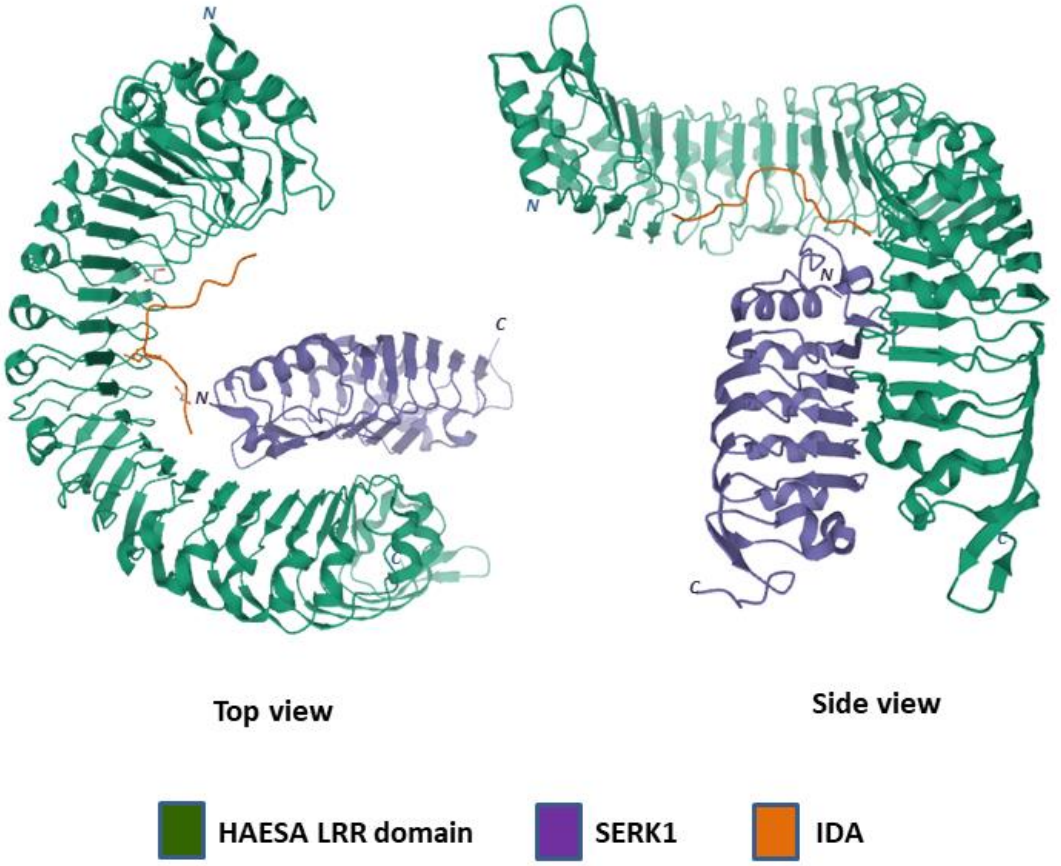


Figure. 1.7. The ligand-receptor-co receptor complex structure of HAESA, SERK1 with IDA (ref)

1.5 Identification of protein-protein interactions (PPIs)

1.5.1 Recent methodology in protein-protein interactions (PPIs) study

To identify the protein-protein interaction, there are several commonly used approaches, such as the isothermal titration calorimetry (ITC) (Duff and Howell, 2015), gel filtration chromatography (Bai, 2015), pull down assay (Miura and Imaki, 2008; Gerace and Moazed, 2015), reflectometric interference spectroscopy (RIFS) (Abdiche et al., 2008; Wilson et al., 2010), surface plasmon resonance spectroscopy (SPR) (Patching, 2014), circular dichroism spectroscopy (CD) (Siligardi et al., 2014), Raman optical activity spectroscopy (ROA) (Zhu et al., 2005), and cryo-electron microscopy (Cryo-EM) (Hauer et al., 2015) (Table 1.1) (Miura, 2018). These candidates can be further confirmed *in vivo* (Xing et al., 2016), by approaches such as yeast two-hybrid (Piya et al., 2014) and the split-ubiquitin system (Johnsson and Varshavsky, 1994) in yeast

model systems, as well as the fluorescence resonance energy transfer (FRET) (Sun et al., 2013), bimolecular fluorescence complementation (BiFC) (Berendzen et al., 2012), split-luciferase (Stynen et al., 2012) and co-immunoprecipitation (Co-IP) (Masters, 2004) in plant model systems (Table 1.2). However, *in vitro* screening potential PPIs candidates for plant is usually costly and cannot completely reproduce the *in vivo* condition, since the artificial binding condition can lead to high false positive screening results which need extra step for further confirmation (Awuni and Mu, 2015; Deng et al., 2015). Although many of these approaches can be used to validate protein-protein interactions, most *in vivo* PPIs experiments are not suitable to identify peptide ligand-receptor interaction since the molecular weight of peptide is usually much smaller than the receptor and need alternative approach (McFedries et al., 2013; Du et al., 2016). For example, stable isotope labelling for ligand-receptor identification combining the genetical mutant plant to verify the interaction were used as alternative approach in plant such as Arabidopsis and tomato to verify specific ligand-receptor pairs (Matsubayashi and Sakagami, 2006). However, these approaches were usually costly, highly time consuming and only can be applied in a case dependent manner.

Table 1.1 *In vivo* protein-protein interaction test (Miura, 2018)

Methods	Protein Binding	Interactor Protein	Protein Complex	Affinity Analysis	Kinetic Analysis	Higher Structure Analysis	Reporter
Pull down	**+**	**+**	**+**	**-**	**-**	**-**	Band/spot by SDS-PAGE
Two-hybrid	**+**	**+**	**-**	**-**	**-**	**-**	Colony formation
Gel Filtration	**+**	**-**	**+**	**-**	**-**	**-**	Retention time
ITC	**+**	**-**	**-**	**+**	**-**	**-**	Temperature
FRET	**+**	**-**	**-**	**+**	**-**	**-**	Fluorescence
LOCI	**+**	**-**	**-**	**+**	**-**	**-**	Luminescence

RIFS	**+**	**-**	**-**	**+**	**+**	**-**	Interference spectrum
SPR	**+**	**-**	**-**	**+**	**+**	**-**	Resonance angle
CD	**±**	**-**	**-**	**-**	**-**	**±**	Absorbance of circularly polarized light
ROA	**±**	**-**	**-**	**-**	**-**	**±**	Raman scattering by circularly polarized light
SAXS	**±**	**-**	**-**	**-**	**-**	**±**	Scattered X rays
NMR	**±**	**-**	**-**	**-**	**-**	**±**	Chemical shift
Cryo-EM	**+**	**-**	**-**	**-**	**-**	**+**	Images of proteins

+: possible and suitable to verify PPI;

±: possible to verify PPI depending on conditions;

-: rather unsuitable or impossible to verify PPI

Table 1.2 *In vivo* protein-protein interaction test (Xing et al., 2016)

Method	Organism/ System	Reporter	Detection of Protein Binding	Quantification	Reference
Yeast two-hybrid	Yeast	Survival on depleted medium or enzymatic	+	++	Piya et al., 2014
Split-ubiquitin system	Yeast	Survival on depleted medium or enzymatic	+	++	Johnsson and Varshavsky, 1994
Fluorescence resonance energy transfer	Plant	Fluorescence	+	++	Sun et al., 2013
Bimolecular fluorescence complementation	Plant	Fluorescence	+	+	Berendzen et al., 2012
Split-luciferase	Plant	Bioluminescence, enzymatic	+	++	Stynen et al., 2012
Co-immunoprecipitation	Plant	Immunostaining	+	+	Masters, 2004

Table 1.2 *In vivo* protein-protein interaction test (continued) (Xing et al., 2016)

Method	Pro	Contra
Yeast two-hybrid	Simple and cheap technique	Truncation and mis-localization of proteins necessary, limited to nucleus

Split-ubiquitin system	Simple and cheap technique, potentially can be used on all proteins	Possible low protein expression, probes need to face cytosol
Fluorescence resonance energy transfer	Can visualized interaction site, Simple technique	Fluorescence spectral might self-interfering or bleaching, concentration dependence
Bimolecular fluorescence complementation	Simple technique and can predicted in topology	Many effects overlay PPI
Split-luciferase	Kinetic and dynamic analysis can be revealed	Exogenous substrate application
Co-immunoprecipitation	Simple technique, can be performed without tags	Not suitable for membrane proteins

1.5.2 Formaldehyde and pBpa mediated crosslinking for PPIs identification

To identify the protein-protein interaction, one commonly used approach is using cross-linking reagents. A number of formaldehyde based experimental approaches are very promising in their ability to accurately and efficiently identify novel protein-protein interaction complexes (Sutherland et al., 2008). Formaldehyde is routinely used to cross-link proteins, which can inactivate, fix or stabilise nearby macromolecules (Toth and Biggin, 2000; Casadonte and Caprioli, 2011; Hoffman et al., 2015; Sinz et al., 2015). A key advantage of formaldehyde as a cross-linking reagent is its small size. This enables the molecule to easily permeate cells to facilitate very specific cross-linking

reactions between amino acids on two proteins in close proximity (2.3-2.7 Å) (Metz et al., 2004). Additionally, formaldehyde cross-linking reactions are reversible and compatible with mass spectrometry (MS) (Sutherland et al., 2008). Although several approaches exist to demonstrate PPIs studies, formaldehyde assisted crosslinking has been rarely used to demonstrate peptide-receptor interactions. Formaldehyde can, however, mediate peptide-peptide crosslinking, which requires suitable amino acid side groups (e.g. Cysteine, Histidine, Tryptophan and Arginine groups and Lysine groups) to be present in the interacting molecules (Metz et al., 2004). Therefore, it is possible that confocal microscopy could be used to visualize the formaldehyde assisted crosslinking of a fluorescently tagged CEP ligand to it receptor *in vivo* by exploiting the natural affinity of the ligand for its receptor. This has the advantage of obviating the need to manipulate the putative receptor (e.g. by adding an *N*- or *C*-terminal tag), however, success is contingent on the fluorescent tag not destroying the biological activity of the peptide ligand by negating binding. It should also be noted that formaldehyde is indiscriminate and will cross link all proximal proteins in the cell.

In contrast to formaldehyde-mediated crosslinking, we also used a more selective binding approach using photo-crosslinking. The photo-active amino acid analog p-benzoyl-L-phenylalanine (pBpa) was also selected in this study as a selective crosslinking reagent. pBpa was introduced as a cross-linker in the research of peptide-protein interactions since 1986 (Kauer et al., 1986). Upon UV irradiation at a wavelength around 365 nm, the photo-generated triplet state of the benzophenone group abstracts a hydrogen atom from geometrically accessible carbon-hydrogen bond of the polypeptide backbone. The side chain methylene groups then combines with the resulting radical to generate a covalent link between the ketone carbonyl carbon and the attacked methylene carbon. Unlike formaldehyde-mediated crosslinking, only carbon-hydrogen bonds within 3 Å of the carbonyl oxygen are targets for cross-linking (Farrell et al., 2005). Due to the specificity of photoactivation, the cross-linked protein complex can be purified and identified by using tandem mass spectrometry. However, few publications have reported using this technique in plant tissues.

In this thesis, we aim to gain better understanding of the peptide hormone-receptor pairings and to identify unknown receptor from known ligand by using CEP/receptor interactions in the model legume Medicago.

1.6 Thesis research aims

We used three strategies to investigate the interactions between MtCEP1 and MtCRA2. First, we used Medicago shoot to purify microsomal membrane protein followed by bottom up proteomics to directly identify possible LRR-RLKs and MtCRA2. Second, we used formaldehyde mediated crosslinking and semi-purified Medicago leaf vasculature to identify the ligand receptor interaction between MtCEP1 and MtCRA2. Finally, we used site directed photo-activated crosslinker to purify and identify the MtCEP1-MtCRA2 complex from the Medicago leaf vasculature. The over-view flow chart of the thesis is presented in Fig. 1.8. An outline of the experimental approach follows:

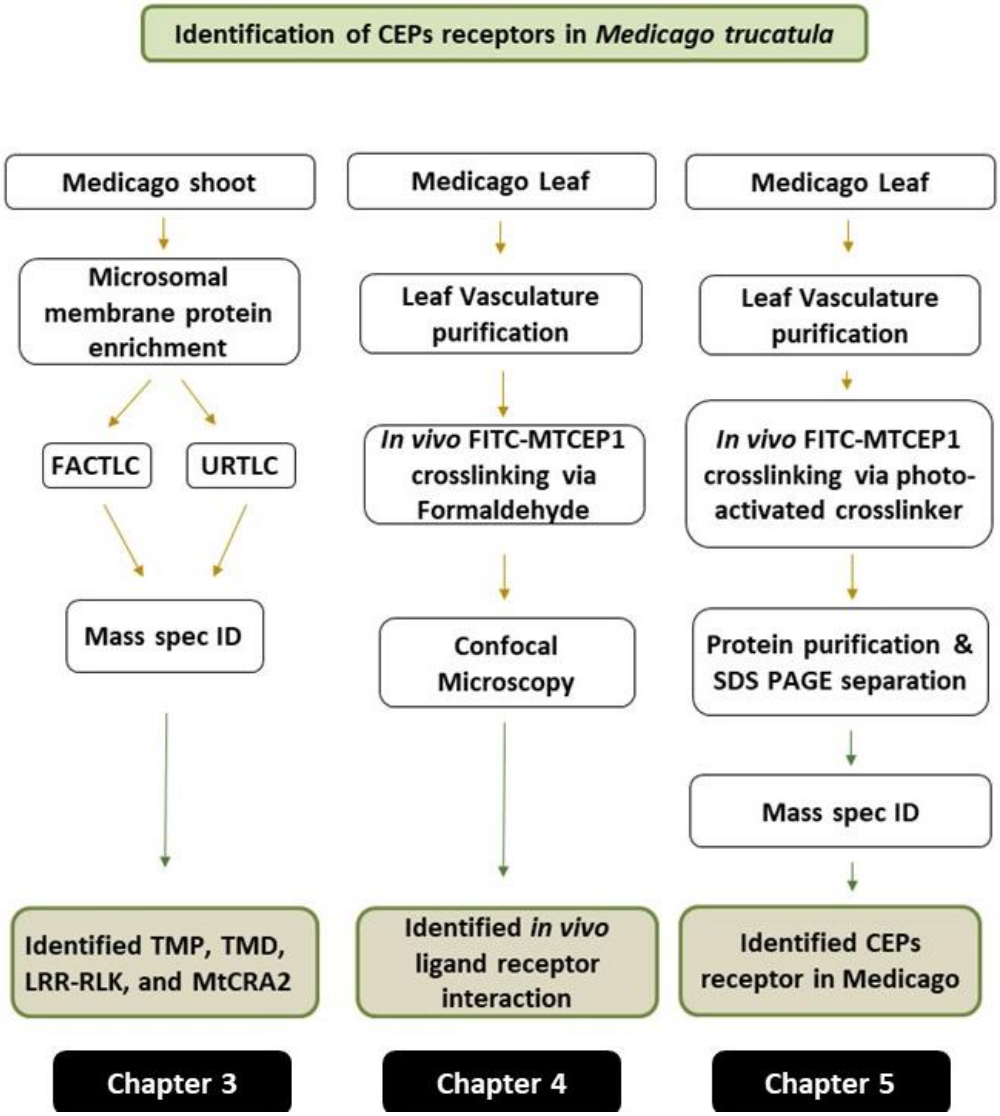


Figure 1.8. Overview of the three strategies taken to study the CEP1-CRA2 interaction. The generic methods underpinning each of the results chapters appear in Chapter 2.

1.7 Chapter conclusion

By using the Medicago model legume, this research aims to unravel the currently unknown CEP/receptor interactions and provides actual proof of the recognition of CEP receptors by CEPs. All the materials and method were described in chapter 2. In chapter 3 we started with leaf microsomal membrane proteins extraction with bottom up proteomics. We establish the non-detergent based and non-urea-based purification procedure and compared the established method to the commonly used urea-based method to see which

one is superior in transmembrane protein (TMP) and transmembrane domain (TMD) enrichment. In chapter4, we used formaldehyde mediated crosslinking to demonstrate the MtCEP1 localisation via semi-purified Medicago leaf vasculature tissue. We show the semi-purified Medicago leaf vasculature tissue was truly enriched in vasculature related protein, and MtCEP1 bound to Medicago leaf vasculature tissue in a receptor dependent manner. In chapter 5, we provided evidence that MtCRA2 is the receptor of MtCEP1 by purifying the ligand-receptor complex and confirming it by mass spectrometry.

Literature cited

Abdiche Y, Malashock D, Pinkerton A, Pons J (2008) Determining kinetics and affinities of protein interactions using a parallel real-time label-free biosensor, the Octet. Anal Biochem **377:** 209-217

Awuni Y, Mu Y (2015) Reduction of false positives in structure-based virtual screening when receptor plasticity is considered. Molecules **20:** 5152-5164

Bai Y (2015) Detecting protein-protein interactions by gel filtration chromatography. Methods Mol Biol **1278:** 223-232

Berendzen KW, Bohmer M, Wallmeroth N, Peter S, Vesic M, Zhou Y, Tiesler FK, Schleifenbaum F, Harter K (2012) Screening for in planta protein-protein interactions combining bimolecular fluorescence complementation with flow cytometry. Plant Methods **8:** 25

Bojar D, Martinez J, Santiago J, Rybin V, Bayliss R, Hothorn M (2014) Crystal structures of the phosphorylated BRI1 kinase domain and implications for brassinosteroid signal initiation. The Plant Journal **78:** 31-43

Breiden M, Simon R (2016) Q&A: How does peptide signaling direct plant development? BMC Biology **14:** 58

Casadonte R, Caprioli RM (2011) Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry. Nature protocols **6:** 1695-1709

Chakraborty S, Nguyen B, Wasti SD, Xu G (2019) Plant Leucine-Rich Repeat Receptor Kinase (LRR-RK): Structure, Ligand Perception, and Activation Mechanism. Molecules (Basel, Switzerland) **24:** 3081

Chapman K, Ivanovici A, Taleski M, Sturrock CJ, Ng JLP, Mohd-Radzman NA, Frugier F, Bennett MJ, Mathesius U, Djordjevic MA (2020) CEP receptor signalling controls root system architecture in Arabidopsis and Medicago. New Phytol

Chapman K, Taleski M, Ogilvie HA, Imin N, Djordjevic MA (2019) CEP-CEPR1 signalling inhibits the sucrose-dependent enhancement of lateral root growth. Journal of experimental botany **70:** 3955-3967

Czyzewicz N, Yue K, Beeckman T, De Smet I (2013) Message in a bottle:

small signalling peptide outputs during growth and development. J Exp Bot **64:** 5281-5296

Delay C, Chapman K, Taleski M, Wang Y, Tyagi S, Xiong Y, Imin N, Djordjevic MA (2019) CEP3 levels affect starvation-related growth responses of the primary root. J Exp Bot **70:** 4763-4774

Delay C, Imin N, Djordjevic MA (2013) CEP genes regulate root and shoot development in response to environmental cues and are specific to seed plants. J Exp Bot **64:** 5383-5394

Deng N, Forli S, He P, Perryman A, Wickstrom L, Vijayan RS, Tiefenbrunn T, Stout D, Gallicchio E, Olson AJ, Levy RM (2015) Distinguishing binders from false positives by free energy calculations: fragment screening against the flap site of HIV protease. J Phys Chem B **119:** 976-988

Djordjevic MA, Mohd-Radzman NA, Imin N (2015) Small-peptide signals that control root nodule number, development, and symbiosis. J Exp Bot **66:** 5171-5181

Djordjevic MA, Oakes M, Wong CE, Singh M, Bhalla P, Kusumawati L, Imin N (2011) Border sequences of *Medicago truncatula* CLE36 are specifically cleaved by endoproteases common to the extracellular fluids of Medicago and soybean. J Exp Bot **62:** 4649-4659

Du X, Li Y, Xia YL, Ai SM, Liang J, Sang P, Ji XL, Liu SQ (2016) Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. Int J Mol Sci **17**

Duff MR, Jr., Howell EE (2015) Thermodynamics and solvent linkage of macromolecule-ligand interactions. Methods **76:** 51-60

Farrell IS, Toroney R, Hazen JL, Mehl RA, Chin JW (2005) Photo-cross-linking interacting proteins with a genetically encoded benzophenone. Nat Methods **2:** 377-384

Franssen HJ, Bisseling T (2001) Peptide signaling in plants. Proceedings of the National Academy of Sciences of the United States of America **98:** 12855-12856

Gautrat P, Laffont C, Frugier F (2020) Compact Root Architecture 2 Promotes Root Competence for Nodulation through the miR2111 Systemic Effector. Current Biology

Gerace E, Moazed D (2015) Affinity Pull-Down of Proteins Using Anti-FLAG M2 Agarose Beads. Methods Enzymol **559:** 99-110

Hauer F, Gerle C, Fischer N, Oshima A, Shinzawa-Itoh K, Shimada S, Yokoyama K, Fujiyoshi Y, Stark H (2015) GraDeR: Membrane Protein Complex Preparation for Single-Particle Cryo-EM. Structure **23:** 1769-1775

Hoffman EA, Frey BL, Smith LM, Auble DT (2015) Formaldehyde crosslinking: a tool for the study of chromatin complexes. The Journal of biological chemistry **290:** 26404-26411

Hothorn M, Belkhadir Y, Dreux M, Dabi T, Noel JP, Wilson IA, Chory J (2011) Structural basis of steroid hormone perception by the receptor kinase BRI1. Nature **474:** 467

Hu C, Zhu Y, Cui Y, Cheng K, Liang W, Wei Z, Zhu M, Yin H, Zeng L, Xiao Y, Lv M, Yi J, Hou S, He K, Li J, Gou X (2018) A group of receptor kinases are essential for CLAVATA signalling to maintain stem cell homeostasis. Nature Plants **4:** 205-211

Huault E, Laffont C, Wen J, Mysore KS, Ratet P, Duc G, Frugier F (2014) Local and systemic regulation of plant root system architecture and symbiotic nodulation by a receptor-like kinase. PLoS Genet **10:** e1004891

Imin N, Mohd-Radzman NA, Ogilvie HA, Djordjevic MA (2013) The peptide-encoding CEP1 gene modulates lateral root and nodule numbers in *Medicago truncatula*. J Exp Bot **64:** 5395-5409

Imin N, Patel N, Corcilius L, Payne RJ, Djordjevic MA (2018) CLE peptide tri-arabinosylation and peptide domain sequence composition are essential for SUNN-dependent autoregulation of nodulation in Medicago truncatula. New Phytol **218:** 73-80

Johnsson N, Varshavsky A (1994) Split ubiquitin as a sensor of protein interactions in vivo. Proc Natl Acad Sci U S A **91:** 10340-10344

Kauer JC, Erickson-Viitanen S, Wolfe HR, Jr., DeGrado WF (1986) p-Benzoyl-L-phenylalanine, a new photoreactive amino acid. Photolabeling of calmodulin with a synthetic calmodulin-binding peptide. J Biol Chem **261:** 10695-10700

Kereszt A, Mergaert P, Montiel J, Endre G, Kondorosi É (2018) Impact of

Plant Peptides on Symbiotic Nodule Development and Functioning. Frontiers in plant science **9:** 1026-1026

Laffont C, Huault E, Gautrat P, Endre G, Kalo P, Bourion V, Duc G, Frugier F (2019) Independent Regulation of Symbiotic Nodulation by the SUNN Negative and CRA2 Positive Systemic Pathways. Plant Physiol **180:** 559-570

Li J, Wen J, Lease KA, Doke JT, Tax FE, Walker JC (2002) BAK1, an Arabidopsis LRR Receptor-like Protein Kinase, Interacts with BRI1 and Modulates Brassinosteroid Signaling. Cell **110:** 213-222

Li Z, Wang Y, Huang J, Ahsan N, Biener G, Paprocki J, Thelen JJ, Raicu V, Zhao D (2017) Two SERK Receptor-Like Kinases Interact with EMS1 to Control Anther Cell Fate Determination. Plant Physiology **173:** 326

Masters SC (2004) Co-immunoprecipitation from transfected cells. Methods Mol Biol **261:** 337-350

Matsubayashi Y (2018) Exploring peptide hormones in plants: identification of four peptide hormone-receptor pairs and two post-translational modification enzymes. Proceedings of the Japan Academy. Series B, Physical and biological sciences **94:** 59-74

Matsubayashi Y, Sakagami Y (2006) Peptide hormones in plants. Annu Rev Plant Biol **57:** 649-674

McFedries A, Schwaid A, Saghatelian A (2013) Methods for the elucidation of protein-small molecule interactions. Chem Biol **20:** 667-673

Meng X, Chen X, Mang H, Liu C, Yu X, Gao X, Torii Keiko U, He P, Shan L (2015) Differential Function of Arabidopsis SERK Family Receptor-like Kinases in Stomatal Patterning. Current Biology **25:** 2361-2372

Meng X, Zhou J, Tang J, Li B, de Oliveira MVV, Chai J, He P, Shan L (2016) Ligand-Induced Receptor-like Kinase Complex Regulates Floral Organ Abscission in Arabidopsis. Cell Rep **14:** 1330-1338

Metz B, Kersten GF, Hoogerhout P, Brugghe HF, Timmermans HA, de Jong A, Meiring H, ten Hove J, Hennink WE, Crommelin DJ, Jiskoot W (2004) Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. J Biol Chem **279:** 6235-6243

Mitra SK, Chen R, Dhandaydham M, Wang X, Blackburn RK, Kota U, Goshe MB, Schwartz D, Huber SC, Clouse SD (2015) An autophosphorylation site database for leucine-rich repeat receptor-like kinases in *Arabidopsis thaliana*. Plant J **82:** 1042-1060

Miura K (2018) An Overview of Current Methods to Confirm Protein-Protein Interactions. Protein and peptide letters **25:** 728-733

Miura K, Imaki J (2008) Molecular cloning of Ebitein1: a novel extracellular signal-regulated kinase 2-binding protein in testis. Biochem Biophys Res Commun **368:** 336-342

Mohd-Radzman NA, Binos S, Truong TT, Imin N, Mariani M, Djordjevic MA (2015) Novel MtCEP1 peptides produced *in vivo* differentially regulate root development in *Medicago truncatula*. Journal of Experimental Botany **66:** 5289-5300

Mohd-Radzman NA, Laffont C, Ivanovici A, Patel N, Reid D, Stougaard J, Frugier F, Imin N, Djordjevic MA (2016) Different Pathways Act Downstream of the CEP Peptide Receptor CRA2 to Regulate Lateral Root and Nodule Development. Plant physiology **171:** 2536-2548

Murphy E, Smith S, De Smet I (2012) Small signaling peptides in Arabidopsis development: how cells communicate over a short distance. Plant Cell **24:** 3198-3217

Notaguchi M, Okamoto S (2015) Dynamics of long-distance signaling via plant vascular tissues. Frontiers in Plant Science **6:** 161

Ogilvie HA, Imin N, Djordjevic MA (2014) Diversification of the C-TERMINALLY ENCODED PEPTIDE (CEP) gene family in angiosperms, and evolution of plant-family specific CEP genes. BMC Genomics **15:** 870

Ohkubo Y, Tanaka M, Tabata R, Ogawa-Ohnishi M, Matsubayashi Y (2017) Shoot-to-root mobile polypeptides involved in systemic regulation of nitrogen acquisition. Nat Plants **3:** 17029

Ohyama K, Ogawa M, Matsubayashi Y (2008) Identification of a biologically active, small, secreted peptide in Arabidopsis by in silico gene screening, followed by LC-MS-based structure analysis. The Plant Journal **55:** 152-160

Okamoto S, Suzuki T, Kawaguchi M, Higashiyama T, Matsubayashi Y

(2015) A comprehensive strategy for identifying long-distance mobile peptides in xylem sap. Plant J **84:** 611-620

Patching SG (2014) Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery. Biochim Biophys Acta **1838:** 43-55

Patel N, Mohd-Radzman NA, Corcilius L, Crossett B, Connolly A, Cordwell SJ, Ivanovici A, Taylor K, Williams J, Binos S, Mariani M, Payne RJ, Djordjevic MA (2018) Diverse Peptide Hormones Affecting Root Growth Identified in the *Medicago truncatula* Secreted Peptidome. Mol Cell Proteomics **17:** 160-174

Piya S, Shrestha SK, Binder B, Stewart CN, Jr., Hewezi T (2014) Protein-protein interaction and gene co-expression maps of ARFs and Aux/IAAs in Arabidopsis. Front Plant Sci **5:** 744

Roberts I, Smith S, De Rybel B, Van Den Broeke J, Smet W, De Cokere S, Mispelaere M, De Smet I, Beeckman T (2013) The CEP family in land plants: evolutionary analyses, expression studies, and role in Arabidopsis shoot development. Journal of Experimental Botany **64:** 5371-5381

Roy S, Lundquist P, Udvardi M, Scheible W-R (2018) Small and Mighty: Peptide hormones in plant biology Plant Cell**:** - tpc.118.tt0718

Santiago J, Brandt B, Wildhagen M, Hohmann U, Hothorn LA, Butenko MA, Hothorn M (2016) Mechanistic insight into a peptide hormone signaling complex mediating floral organ abscission. eLife **5:** e15075

Shiu S-H, Bleecker AB (2001) Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. Proceedings of the National Academy of Sciences **98:** 10763-10768

Siligardi G, Hussain R, Patching SG, Phillips-Jones MK (2014) Ligand- and drug-binding studies of membrane proteins revealed through circular dichroism spectroscopy. Biochim Biophys Acta **1838:** 34-42

Sinz A, Arlt C, Chorev D, Sharon M (2015) Chemical cross-linking and native mass spectrometry: A fruitful combination for structural biology. Protein science : a publication of the Protein Society **24:** 1193-1209

Song W, Liu L, Wang J, Wu Z, Zhang H, Tang J, Lin G, Wang Y, Wen X, Li

W, Han Z, Guo H, Chai J (2016) Signature motif-guided identification of receptors for peptide hormones essential for root meristem growth. Cell Res **26:** 674-685

Stahl Y, Simon R (2012) Peptides and receptors controlling root development. Philosophical transactions of the Royal Society of London. Series B, Biological sciences **367:** 1453-1460

Stynen B, Tournu H, Tavernier J, Van Dijck P (2012) Diversity in genetic in vivo methods for protein-protein interaction studies: from the yeast two-hybrid system to the mammalian split-luciferase system. Microbiol Mol Biol Rev **76:** 331-382

Sun Y, Han Z, Tang J, Hu Z, Chai C, Zhou B, Chai J (2013) Structure reveals that BAK1 as a co-receptor recognizes the BRI1-bound brassinolide. Cell Research **23:** 1326

Sun Y, Li L, Macho AP, Han Z, Hu Z, Zipfel C, Zhou J-M, Chai J (2013) Structural Basis for flg22-Induced Activation of the Arabidopsis FLS2-BAK1 Immune Complex. Science **342:** 624

Sun Y, Rombola C, Jyothikumar V, Periasamy A (2013) Forster resonance energy transfer microscopy and spectroscopy for localizing protein-protein interactions in living cells. Cytometry A **83:** 780-793

Sutherland BW, Toews J, Kast J (2008) Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. J Mass Spectrom **43:** 699-715

Tabata R, Sumida K, Yoshii T, Ohyama K, Shinohara H, Matsubayashi Y (2014) Perception of root-derived peptides by shoot LRR-RKs mediates systemic N-demand signaling. Science **346:** 343-346

Taleski M, Chapman K, Imin N, Djordjevic MA, Groszmann M (2020) The peptide hormone receptor CEPR1 functions in the reproductive tissue to control seed size and yield. Plant Physiology**:** pp.00172.02020

Taleski M, Imin N, Djordjevic MA (2018) CEP peptide hormones: key players in orchestrating nitrogen-demand signalling, root nodulation, and lateral root development. J Exp Bot **69:** 1829-1836

Tavormina P, De Coninck B, Nikonorova N, De Smet I, Cammue BPA (2015) The Plant Peptidome: An Expanding Repertoire of Structural Features and Biological Functions. The Plant Cell **27:** 2095-2118

Ten Hove CA, Bochdanovits Z, Jansweijer VMA, Koning FG, Berke L, Sanchez-Perez GF, Scheres B, Heidstra R (2011) Probing the roles of LRR RLK genes in Arabidopsis thaliana roots using a custom T-DNA insertion set. Plant Mol Biol **76:** 69-83

Toth J, Biggin MD (2000) The specificity of protein-DNA crosslinking by formaldehyde: in vitro and in drosophila embryos. Nucleic acids research **28:** e4-e4

Wang J, Li H, Han Z, Zhang H, Wang T, Lin G, Chang J, Yang W, Chai J (2015) Allosteric receptor activation by the plant peptide hormone phytosulfokine. Nature **525:** 265

Wilson JL, Scott IM, McMurry JL (2010) Optical biosensing: Kinetics of protein A-IGG binding using biolayer interferometry. Biochem Mol Biol Educ **38:** 400-407

Xiao Y, Stegmann M, Han Z, DeFalco TA, Parys K, Xu L, Belkhadir Y, Zipfel C, Chai J (2019) Mechanisms of RALF peptide perception by a heterotypic receptor complex. Nature **572:** 270-274

Xing S, Wallmeroth N, Berendzen KW, Grefen C (2016) Techniques for the Analysis of Protein-Protein Interactions in Vivo. Plant physiology **171:** 727-758

Zhang H, Lin X, Han Z, Qu L-J, Chai J (2016) Crystal structure of PXY-TDIF complex reveals a conserved recognition mechanism among CLE peptide-receptor pairs. Cell Research **26:** 543

Zhu F, Isaacs NW, Hecht L, Barron LD (2005) Raman optical activity: a tool for protein structure analysis. Structure **13:** 1409-1419

Chapter 2: Methods

2.1 Plant growth

Surface-sterilized *M. truncatula* cv. Jemalong A17 and *cra2-11* seed (Imin et al., 2013) were germinated and grown on Fåhraeus medium plates (Mohd-Radzman et al., 2015). Eight seedlings per plate were grown for 14 days in a Conviron growth chamber at 25 °C with a 16 h photoperiod and a photon flux density of 100 µm mol m^{-2} s^{-1} (Imin et al., 2013). Shoots were harvested separately and frozen in liquid nitrogen for immediate extraction or stored at -80°C before use.

2.2 Microsomal membrane (MM) preparation

Proteins were extracted from homogenized and ground tissue based on published methods (Marx et al., 2016) with slight modifications. In brief, 12 gram of shoot tissues (leaves and cotyledons) from 14 day old seedlings were ground into a fine powder in liquid nitrogen using a mortar and pestle. After grinding, five volumes of ice-cold extraction buffer (290 mM sucrose, 250 mM Tris (pH 7.6), 25 mM EDTA (pH 8.0), 10 mM KCl, 25 mM NaF, 50 mM sodium pyrophosphate, 1 mM ammonium molybdate, 1 mM PMSF, mini EDTA-free protease inhibitor (Roche) was added to the ground plant tissue samples. The ground tissue was further homogenized by repeated probe sonication (MSE: Imgen technologies) (10 cycles of 1-minute sonication on ice followed by a 30 second rest period on ice). The homogenized plant tissue was filtered through *100 µm* filter (BD Falcon, Bedford, MA, USA) and subsequently centrifuged for 10 min (4,000 *g*, 4 °C) to remove the remaining tissue debris. MMs were prepared by ultracentrifugation for 30 min at 100,000 *g* (4 °C) to remove cytoplasmic proteins. After ultracentrifugation, the MM pellet was re-suspended in 1M Na_2CO_3 (pH 11) and incubated on ice for 5 min to remove weakly associated proteins (Huang et al., 2013). After incubation, the MMs were subject to another ultracentrifugation (100,000 g, 4 °C) for 30 min (Abas and Luschnig, 2010).

2.3 TCA precipitation and protein solubilisation

Medicago MM proteins were purified by TCA precipitation (Link and LaBaer, 2011) with a slight modification to remove the non-protein contamination. In brief, 11 µL ice-cold 100% TCA solution was added into 89 µL of the MM sample to achieved final TCA concentration as 11% and incubated on ice for 20 min. Another 500 µL of ice-cold 10% TCA solution was added, and the sample was incubated at -20 °C overnight. The solution was centrifuged at 20,000 *g* for 30 min to recover the precipitated protein, and the supernatant discarded. The protein pellet was rinsed 3 times with 80% acetone (Marx et al., 2016), centrifuged at 20,000 *g* for 10 min and dried using a vacuum evaporator (VirTis, bench TopK). The protein sample was divided into two to assess the effectiveness of the two protein solubilization and cleavage protocols.

2.4 Urea solubilisation followed by trypsin -Lys-C digestion (UR-TLC)

The dried protein pellet was re-solubilized in dissolving buffer (8 M urea, 50 mM Tris-HCl (pH 8.0), 30 mM NaCl, 1 mM $CaCl_2$, 20 mM sodium butyrate, 10 mM nicotinamide, mini EDTA-free protease inhibitor (Roche). To improve protein solubility, protein samples were subjected to repeated probe sonication (10 times for 10 sec of pulse and 10 sec of rest on ice). Protein concentration was estimated using a Bradford assay (Bio-Rad). Proteins were reduced with 5 mM dithiothreitol at 60 °C for 40 min. The reduced proteins were alkylated with 15 mM iodoacetamide in the dark at room temperature for 40 min. Alkylation was quenched by adding 5 mM dithiothreitol and incubated at room temperature for 15 min. The protein solution was enzymatically digested in a two-step process using a Trypsin-Lys-C mix (Promega) as the previous study (Saveliev et al., 2013). A 25:1 molar ratio of the enzyme was added to the protein solution and digested for 3 hours at 37 °C. After 3 hours the urea concentration was adjusted to 2 M by dilution with 50 mM Tris, pH 8.0 and the reaction kept at 37 °C overnight. Trypsin can maintain the enzyme activity when the urea concentration was 2 M (Harris, 1956). After overnight digestion, the sample was run over a Sep-Pak C18 classic cartridge (Waters, Milford, MA, US) to remove the salts and the peptides were eluted using 100% acetonitrile.

2.5 Formic Acid solubilisation followed by initial CNBr cleavage and trypsin -Lys-C digestion (FA-CTLC)

The dried protein pellet was dissolved in 70% FA using probe sonication and chemically-digested with CNBr (Sigma) with a 100-fold molar ratio excess to the amount of starting dried protein (Wong and King, 2015). The CNBr solution was prepared as described (Crimmins et al., 2005). Essentially, CNBr crystal was dissolved in acetonitrile to make a 5 M final concentration. After 24 hours of incubation at room temperature in the dark, the supernatant was collected by centrifugation at 20,000 *g* for 10 min, and the FA and CNBr were removed by lyophilization. The dried peptide was dissolved in 10% acetonitrile and 25 mM ammonium bicarbonate. The sample was by reduced, alkylated, and then digested with Trypsin Lys-C, before being lyophilized (see above). The effectiveness of each procedure was assessed by centrifuging (10 min, 10,000 *g*) after UR-TLC or FA-CTLC treatments and examining the residual pellet. A residual pellet that remains after the UR-TLC was subjected to a further FA-CTLC procedure to validate that the pellet contained poorly solubilized undigested protein. The remaining CNBr solution was destroyed by adding 5 volume of 1 M Sodium hydroxide (Lunn and Sansone, 1985) before disposed into chemical waste containers.

2.6 Reverse phase high pH fractionation

The peptide digests were separated into eight fractions using the high pH reversed-phase peptide fractionation Kit (Thermo Fisher Scientific, Waltham, MA, USA) (Kulak et al., 2017; Baldan-Martin et al., 2018). In brief, the C-18 spin column was equilibrated with 0.1% trifluoroacetic acid (TFA) before the digested peptides were loaded. Peptide samples (100 µg) dissolved in 0.1% TFA were loaded onto the spin column and washed with MilliQ H_2O. The peptides were eluted into 16 fractions of increasing concentrations of acetonitrile (ACN) in 0.1% Triethylamine: 5% ACN (fraction 1), 7.5% ACN (fraction 2), 10% ACN (fraction 3), 12.5% ACN (fraction 4), 15% ACN(fraction

5), 17.5% ACN (fraction 6), 20% ACN (fraction 7), 25% ACN (fraction 8), 30% ACN (fraction 9), 35% ACN (fraction 10), 40% ACN (fraction 11), 45% ACN (fraction 12), 50% ACN (fraction 13), 60% ACN (fraction 14), 70% ACN (fraction 15), 95% ACN (faction 16). The 16 fractions were collected and then recombined into eight fractions. Final fraction 1 comprised RP fractions 1 and 16; final fraction 2 comprised RP fractions 2 and 15; and so on until the final fraction 8 comprised RP fractions 8 and 9. The fractionated peptides were lyophilised to remove the solvent, re-dissolved into 100 µl 0.1% TFA, and cleaned up by C18 Ziptip (Merck Millipore, Burlington, MA, US) and subjected to MS analysis.

2.7 Mass spectrometry

The liquid chromatography (LC) was performed by using Thermo Scientific UltiMate™ 3000 RSLCnano system with the setting at 60° C with customized columns. The columns were packed in-house using a laser puller and a pressure bomb, and the length of the columns were generally 35-40 cm with a 75 µm ID fused silica housing. The packing material used was Reprosil-Pur 120 C18-AQ, 1.9 µm particle size. The digested peptides were initially loaded onto the LC system with the mobile phases as 95% buffer A (0.1% formic acid/water) and 5% buffer B (0.1% formic acid/80% ACN/water). Samples were reconstituted in 10 µl loading buffer (as above) and 3 µL directly injected for each run. Peptides were eluted with a 5–40% buffer B gradient for 90 min. The total acquisition time was 140 min, including a 95% acetonitrile wash and re-equilibration.

The LC was either coupled to a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA), or a to a Q Exactive HF-X Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher Scientific, Waltham, MA, USA).

MS scans acquired from Q-Exactive Plus was 70,000 at m/z 200 (mass resolution); and mass analyser range was m/z 350–2000. The 20 most intense

ions with a charge state ≥1 were fragmented in the high energy C-trap dissociation collision (HCD) cell, and subsequently, tandem mass spectra were acquired in the Orbitrap mass analyser with a resolution of 35,000 at m/z =200. The Q Exactive HF-X can be used to achieve almost twice the scan speed while remaining at nearly the same resolution compared to Orbitrap elite (Scheltema et al., 2014), a strategy we utilize here for Q Exactive HF-X MS scan setting was acquired with 60,000 resolution, mass analyser range was 300–2000 m/z. followed by MS/MS data-dependent acquisition of the top 10 ions with HCD (15,000 resolution, 1.6 m/z isolation width).

2.8 Data analysis

All raw files generated by LC-MS/MS were processed by Proteome Discoverer 2.1 (Thermo Fisher Scientific) using the Sequest HT data analysis program to search against the *Medicago* protein sequence databases (UniProt, 2014.12.18.)(Bairoch et al., 2005). Database searching against the corresponding reversed database was also performed to evaluate the false discovery rate of peptide identification. The search parameters of Sequest HT were set as follows: precursor ion mass tolerance +/-10 ppm and product ion mass tolerance of 0.05 m/z units. Standard peptide modification was as follows: carbamidomethylation (CAM) at cysteine residues was set as a fixed modification, while oxidation at methionine, lysine and proline residues, *C*-terminal amidation and deamidation at asparagine and glutamine, as well as *N*-terminal glutamine to pyroglutamic acid were set as variable modifications. When CNBr was chosen as the cleavage agent, methionine was set to be homoserine (Met->Hse) or homoserine lactone (Met->Hsl) as a variable modification.

For all experiments, peptides and their corresponding proteins were both filtered to a 1% FDR. Due to the potential sequence redundancy, the proteins which had shared the same set of identified peptides were grouped into protein groups.

2.9 TMD prediction

The TMD prediction was done by using the TMHMM Server, v. 2.0 (http://www.cbs.dtu.dk/services/TMHMM/). TMHMM is a membrane protein topology prediction method based on a hidden Markov model (Sonnhammer et al., 1998). The web server-based search engine correctly predicts 97-98 % of the transmembrane helices and can distinguish between soluble and membrane proteins with both specificity and sensitivity better than 99 % (Krogh et al., 2001).

2.10 TMD (TRAMDOMI, TRAns Membrane Domain Motif Identification) and GRAVY (grand average of hydropathy) analysis

The TMD analysis was annotated by a customized python script. In brief, two peptide sequence sets were prepared for TMD mapping. One was the detected peptide set, which was derived from the MS analysis of the samples. The second one was the complete *Medicago* TMD motif set, which was predicted by TMHMM server using the *M. truncatula* sequence databases (UniProt, 2014.12.18). TMD identification was done by mapping the detected peptide set to the *Medicago* TMD motif set. The mapping rules were defined as followed. Any detected peptide that satisfied one of the following rules was considered a hit:

1. The length of the detected peptide derived from MS encompassed the complete predicted *Medicago* TMD sequence or laid within the predicted TMD.
2. The detected peptide derived from MS extends from a position outside the TMD to within the TMD with a minimum of a 2 amino acids overlap, or started within the TMD and extended to a position outside of the TMD with a minimum of 2 TMD amino acid overlap.

The python script is presented as supplemental material.

The grand average of hydropathy (GRAVY) algorithm (Kyte and Doolittle, 1982) was used to access the hydrophobicity of the identified proteins from both method. Proteins with a GRAVY scores above 0 are more likely to be

hydrophobic proteins (Magdeldin et al., 2012).

2.11 Subcellular location prediction

The subcellular protein location prediction was done by using LOCALIZER, a machine learning method for predicting subcellular protein localization in plant cells and is available at http://localizer.csiro.au/. It identifies proteins localised to chloroplasts and mitochondria by identifying the presence of transit peptides, and nucleus by using a collection of nuclear localization signals. It can achieve a prediction accuracy of over 90% for chloroplast and mitochondria, and 73% for nuclear proteins (Sperschneider et al., 2017). The queries of protein sequence were submitted directly to the server, and full plant sequences were chosen to perform the prediction.

2.12 Protein functional annotation

The functional annotation was done using Mercator. Mercator is a web-based annotation application (http://mapman.gabipd.org/web/guest/app/Mercator) that achieves accuracies above 90% in predicted functional annotations when compared to manual annotation (Lohse et al., 2014). The queries of protein sequence were submitted directly to the server, searched against the database including TAIR Release 10 and SwissProt/UniProt plant proteins database, and classified into functional plant categories according to MapMan BINs (Thimm et al., 2004).

2.13 Calculate the grand average of hydropathy (GRAVY) value for protein sequences

The GRAVY value is calculated by the sum of hydropathy values of all amino acids divided by the protein length(Kyte and Doolittle, 1982). Hydrophobicity score (arbitrary unit) below 0 are more likely cytoplasmic protein (hydrophilic protein), while scores above 0 are more likely TMPs (hydrophobic).

2.14 Synthetic peptides

The MtCEP1 (Domain1 Hyp 4, 11) (Mohd-Radzman et al., 2015), FITC-MtCEP1, FITC-MtCEP1F* and FITC-MtCEP14 (Fig. 1) were synthesized by GL BioChem Ltd., Shanghai, China with a greater than 95% purity, and validated by mass spectrometry (Djordjevic et al., 2011). The peptides were re-suspended in sterile water to prepare the 1 mM stock solution before being used at final concentration of 1 µM. Aliquots of the stock solutions were stored at -20°C and used within 2 months. The photo-activated crosslinking of FITC-MtCEP1F* was confirmed by MS after directly UV 365 nm exposure.

2.15 Biological activity of MtCEP1 peptide derivatives on *Medicago truncatula*

Surface-sterilized A17 seed were germinated and transferred to 150 mm Petri plates containing Fåhraeus medium with 5 mm KNO_3, with or without synthetic peptides (1 µM). LR number was scored at day 7 and 14 (Mohd-Radzman et al., 2015). Four day old seedlings were inoculated with *Sinorhizobium meliloti* WSM1022 (Imin et al Mohd Radzman et al 2016) and grown on nitrogen-free Fåhraeus medium with or without peptide added. Nodule number were determined at day 7 and 14 post-inoculation (Imin et al., 2013). Plants were grown in a Thermoline growth chamber at 25 °C with a 16-hour photoperiod and a photon flux density of 100–120 µmol m^{-2} s^{-1} (Imin et al., 2013).

2.16 Isolation and enrichment of viable Medicago and Arabidopsis leaf vascular tissue (VT)

Medicago leaves of 14-day old A17 or *cra2* seedlings were harvested and processed (Endo et al., 2016) with the following modifications. An individual leaf was immersed in 1 mL of the cell wall degrading enzyme solution (0.75% Cellulase (Sigma), 0.25% Macerozyme (Sigma), 0.4 M mannitol, 5 mM MES-KOH (pH 5.6), 8 mM $CaCl_2$) with the cOmplete™, EDTA-free Protease Inhibitor Cocktail (Roche) in a 1.5 mL tube. For Arabidopsis, cauline leaves from 5-week-old flowering plants were cut into ~10 mm^2 pieces and immersed in enzyme

solution in a 1.5 mL tube. The tube was sonicated on ice 5 times for a maximum of 10 seconds with 1-minute rest intervals on ice using a probe sonicator (MSE: Imgen technologies) until most (~80%) of the mesophyll tissue dissociated from the VT. The remaining VT was then washed in 1x MES buffer (pH 5.6) to remove the enzyme solution so that it was ready for *in vivo* peptide binding. The viability of the isolated Medicago and Arabidopsis VT was confirmed by propidium iodide staining (Chapman et al., 2019) using a Leica DM5500 microscope with a 560 nm excitation filter.

2.17 Propidium iodide staining

The viability of the isolated Medicago VT was confirmed by propidium iodide staining following the published protocol (Chapman et al., 2019). In brief, the purified VT was stained with 100 µM propidium iodide for 2 minutes, washed with milliQ water, and mounted on glass slides. The VT viability was confirmed by Leica DM5500 microscope with a 560 nm excitation filter.

2.18 Crosslinking of FITC-MtCEP1 to Medicago VT *in vivo*

The interaction of FITC-MtCEP1 with A17 or *cra2-11* vascular tissue *in vivo* was assessed using a modification on the published method (Ogawa et al., 2008) with the following modifications. After removing the enzyme solution by aspiration, the purified Medicago VT and was incubated with 500 µL of 1 µM of FITC-MtCEP1 in 5 mM MES buffer (pH=5.6) for 30 minutes on ice. After incubation, the VT was washed with 5 mM PBS (pH=7) buffer to remove unbound peptide and subjected to crosslinking with 4% formaldehyde in PBS (pH=7), or with 200 mM Dimethyl-pimelimidate-2-HCl (DMP) in 0.2M triethanolamine (pH 8.0) by incubating for 30 minutes at room temperature with consistent agitation using a rotating wheel.

2.19 Defining the minimum concentration of formaldehyde required for crosslinking of FITC-MtCEP1 to VT

The purified Medicago VT was incubated with 500 µL of 1 µM FITC-MtCEP1 in 5 mM MES buffer (pH 5.6) for 30 minutes on ice. The VT was then washed with 5 mM PBS buffer (pH 7) to remove unbound peptide. The complex was cross-linked with 4% formaldehyde suspended in PBS for 5 minutes at room temperature with consistent agitation and washed with PBS to remove the formaldehyde and unbound CEP peptide. A series of lower formaldehyde concentrations were used for crosslinking as specified. FITC-MtCEP1 crosslinking to VT was observed using confocal microscopy (Zeiss Confocal LSM 780, see the confocal microscopy section)

2.20 Photo-activated crosslinking of FITC-MtCEP1F* to VT *in vivo*

After removing the enzyme solution, the purified Medicago VT was incubated with 1 µM of FITC-MtCEP1F* in 5 mM MES buffer (pH=5.6) for 30 minutes on ice. After incubation, the VT was washed with PBS (pH=7) buffer to remove unbound peptide and exposed to UV 365 nm for 3 hours on ice with consistent agitation. MES (pH=5.6) was not used since it was not sufficient to remove the unbound peptide. After 3 hours the VT was washed with PBS (pH=7) before FITC-MtCEP1F* crosslinking was observed using confocal microscopy.

2.21 Binding site competition assays between FITC- MtCEP1/FITC-MtCEP1F* and MtCEP1

After removed the enzyme solution, the purified Medicago VT was co-incubated with 250 µL of 1 µM of FITC-MtCEP1 and (1) 250 µL of 1 µM of MtCEP1 or (2) 250 µL of 0.1 µM of MtCEP1 or (3) 250 µL of 0.01 µM of MtCEP1 in 5 mM MES buffer (pH=5.6) for 30 minutes on ice. After incubation, the VT was washed with 5 mM Phosphate Buffered Saline (PBS) (pH=7) buffer to remove unbound peptide and subjected to crosslinking with 4% formaldehyde unless otherwise specified in PBS (pH=7) by incubating for 30 minutes at room temperature with consistent agitation using a rotating wheel. The peptide crosslinked VT was

ready for localization of FITC-MtCEP1 using confocal microscopy.

For the photo-activated FiTC-MtCEP1F*, the purified Medicago VT was co-incubated with 250 µL of 1 µM of FITC-MtCEP1F* and (1) 250 µL of 1 µM of MtCEP1 or (2) 250 µL of 0.1 µM of MtCEP1 or (3) 250 µL of 0.01 µM of MtCEP1 or (4) 250 µL of 0.001 µM of MtCEP1 in 5 mM MES buffer (pH=5.6) for 30 minutes on ice. After incubation, the VT was washed with 5 mM Phosphate Buffered Saline (PBS) (pH=7) buffer to remove unbound peptide and exposed to UV 365 nm for 3 hours on ice with consistent agitation. After 3 hours the VT was washed with PBS (pH=7). The peptide crosslinked VT was ready for localization of FITC-MtCEP1F* using confocal microscopy.

2.22 Confocal microscopy

We used confocal microscopy (Zeiss Confocal LSM 780) with argon laser excitation at 485 nm to image the FITC-MtCEP1F* binding to VT localization. The fluorescent spectra of FITC and VT auto fluorescence were determined by spectral scanning module of Zeiss LSM 780 (Ohad et al., 2007), in which the FITC emission peak was at 525 nm and VT auto fluorescence was at around 600 nm (Ohad et al., 2007). All pictures were representative from three independent experiments. Panels in each figure were imaged on the same day with the same confocal microscope settings and exposure time to enable a direct comparison of the results.

2.23 Extraction and separation protein from Medicago VT by SDS-PAGE

Proteins were extracted from the semi-purified VT (Saur et al., 2016) after crosslinking with the following modifications. VT from 40 post-cross-linked preparations was frozen in liquid nitrogen, ground into a fine powder with a mortar and pestle. The ground material was transferred to a 1.5 ml tube containing 1ml extraction buffer (150mM Tris-HCl pH 7.5, 150mM NaCl, 10mM EDTA, 10% glycerol, 15mM DTT, 1mM NaF, 1mM Na_2MoO_4, 0.55mM $NaVO_3$, 1% (v/v) IGEPAL, Protease Inhibitor cocktail (Roach)) and subjected to 3 probe

sonication cycles of 10 seconds with an intervening 1 minute rest incubation on ice. VT proteins were TCA precipitated by adding 125 µL ice-cold 100% TCA solution added into 1000 µL of the sample to achieve a final TCA concentration of 11% followed by 20 min incubation on ice (Link and LaBaer, 2011). The solution was centrifuged at 20,000 *g* for 30 min to recover the precipitated protein, and the supernatant discarded. The protein pellet was rinsed three times with 80% acetone, centrifuged at 20,000 *g* for 10 min and followed by several 100% acetone washes until there was no visible green pigment in the protein pellet. The clean protein pellet was dried using a vacuum evaporator (VirTis, bench TopK) for 3 minutes. The TCA precipitated pellet was re-suspended in 500 µL SDS sample buffer (1M Tris-HCL (pH 6.8), 8M urea, 10% SDS, 0.1% bromophenol blue, 0.1M DTT) and incubated in 65 °C for 10 minutes. The heated sample solution was centrifuged at 20,000 *g* for 1 minute to remove any non-dissolving material. An aliquot (60 µL) of the protein sample was separated using a 10% SDS-PAGE with 8M urea/MOBS running buffer (Dutta et al., 2010).

2.24 Visualization FITC- MtCEP1F* labeled protein separated by SDS-PAGE

FITC fluorescence in the SDS-PAGE was monitored by using a ChemiDoc Imaging Systems ChemiDoc™ (Bio-rad) with a built-in FITC/Alexfluor 488 fluorophore labelling setting. The SDS-PAGE separated VT protein was exposed for 120 seconds to image the FITC fluorescence signal. Gel bands containing the FITC signal were excised from the SDS-PAGE gel and subjected to in-gel trypsin digestion and cysteine carbamidomethylation. The tryptic peptide sample was examined using MS to identify the proteins in target band.

Literature cited

Abas L, Luschnig C (2010) Maximum yields of microsomal-type membranes from small amounts of plant material without requiring ultracentrifugation. Analytical Biochemistry **401:** 217-227

Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh L-SL (2005) The Universal Protein Resource (UniProt). Nucleic Acids Research **33:** D154-D159

Baldan-Martin M, Lopez JA, Corbacho-Alonso N, Martinez PJ, Rodriguez-Sanchez E, Mourino-Alvarez L, Sastre-Oliva T, Martin-Rojas T, Rincon R, Calvo E, Vazquez J, Vivanco F, Padial LR, Alvarez-Llamas G, Ruiz-Hurtado G, Ruilope LM, Barderas MG (2018) Potential role of new molecular plasma signatures on cardiovascular risk stratification in asymptomatic individuals. Sci Rep **8:** 4802

Chapman K, Taleski M, Ogilvie HA, Imin N, Djordjevic MA (2019) CEP–CEPR1 signalling inhibits the sucrose-dependent enhancement of lateral root growth. Journal of Experimental Botany **70:** 3955-3967

Crimmins DL, Mische SM, Denslow ND (2005) Chemical cleavage of proteins in solution. Curr Protoc Protein Sci **Chapter 11:** Unit 11.14

Djordjevic MA, Oakes M, Wong CE, Singh M, Bhalla P, Kusumawati L, Imin N (2011) Border sequences of *Medicago truncatula* CLE36 are specifically cleaved by endoproteases common to the extracellular fluids of Medicago and soybean. J Exp Bot **62:** 4649-4659

Harris JI (1956) Effect of Urea on Trypsin and Alpha-Chymotrypsin. Nature **177:** 471-473

Huang T-K, Han C-L, Lin S-I, Chen Y-J, Tsai Y-C, Chen Y-R, Chen J-W, Lin W-Y, Chen P-M, Liu T-Y, Chen Y-S, Sun C-M, Chiou T-J (2013) Identification of Downstream Components of Ubiquitin-Conjugating Enzyme PHOSPHATE2 by Quantitative Membrane Proteomics in Arabidopsis Roots. The Plant Cell **25:** 4044-4060

Imin N, Mohd-Radzman NA, Ogilvie HA, Djordjevic MA (2013) The peptide-encoding CEP1 gene modulates lateral root and nodule

numbers in *Medicago truncatula*. J Exp Bot **64:** 5395-5409

Kulak NA, Geyer PE, Mann M (2017) Loss-less Nano-fractionator for High Sensitivity, High Coverage Proteomics. Mol Cell Proteomics **16:** 694-705

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol **157:** 105-132

Link AJ, LaBaer J (2011) Trichloroacetic acid (TCA) precipitation of proteins. Cold Spring Harb Protoc **2011:** 993-994

Lohse M, Nagel A, Herter T, May P, Schroda M, Zrenner R, Tohge T, Fernie AR, Stitt M, Usadel B (2014) Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. Plant Cell Environ **37:** 1250-1258

Lunn G, Sansone EB (1985) Destruction of cyanogen bromide and inorganic cyanides. Anal Biochem **147:** 245-250

Magdeldin S, Yoshida Y, Li H, Maeda Y, Yokoyama M, Enany S, Zhang Y, Xu B, Fujinaka H, Yaoita E, Sasaki S, Yamamoto T (2012) Murine colon proteome and characterization of the protein pathways. BioData mining **5:** 11

Marx H, Minogue CE, Jayaraman D, Richards AL, Kwiecien NW, Siahpirani AF, Rajasekar S, Maeda J, Garcia K, Del Valle-Echevarria AR, Volkening JD, Westphall MS, Roy S, Sussman MR, Ane JM, Coon JJ (2016) A proteomic atlas of the legume *Medicago truncatula* and its nitrogen-fixing endosymbiont *Sinorhizobium meliloti*. Nat Biotechnol **34:** 1198-1205

Mohd-Radzman NA, Binos S, Truong TT, Imin N, Mariani M, Djordjevic MA (2015) Novel MtCEP1 peptides produced in vivo differentially regulate root development in *Medicago truncatula*. Journal of Experimental Botany **66:** 5289-5300

Ogawa M, Shinohara H, Sakagami Y, Matsubayashi Y (2008) Arabidopsis CLV3 peptide directly binds CLV1 ectodomain. Science **319:** 294

Ohad N, Shichrur K, Yalovsky S (2007) The Analysis of Protein-Protein Interactions in Plants by Bimolecular Fluorescence Complementation. Plant Physiology **145:** 1090-1099

Saveliev S, Bratz M, Zubarev R, Szapacs M, Budamgunta H, Urh M (2013)

Trypsin/Lys-C protease mix for enhanced protein mass spectrometry analysis. Nature Methods **10:** i-ii

Scheltema RA, Hauschild J-P, Lange O, Hornburg D, Denisov E, Damoc E, Kuehn A, Makarov A, Mann M (2014) The Q Exactive HF, a Benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer. Molecular & cellular proteomics : MCP **13:** 3698-3708

Sonnhammer EL, von Heijne G, Krogh A (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol **6:** 175-182

Sperschneider J, Catanzariti AM, DeBoer K, Petre B, Gardiner DM, Singh KB, Dodds PN, Taylor JM (2017) LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. Sci Rep **7:** 44598

Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J **37:** 914-939

Wong SH, King CY (2015) Amino Acid Proximities in Two Sup35 Prion Strains Revealed by Chemical Cross-linking. J Biol Chem **290:** 25062-25071

Chapter 3: Improving the identification and coverage of plant transmembrane proteins in Medicago using bottom-up proteomics

3.1 Introduction

The aim of this chapter was to assess the feasibility for using mass spectrometry to directly identify LRR-RLKs and specifically MtCRA2 in Medicago leaf tissue using advanced bottom-up proteomics.

3.1.1 TMPs are crucial for regulating growth

TMPs play critical roles in the function of all living organisms. Approximately 20% to 30% of the sequenced genomes from microbial, or eukaryotic organisms encode TMPs (Komatsu et al., 2007; Schey et al., 2013; Almeida et al., 2017). In humans, roughly 50% of all known drug targets are TMPs (Hopkins and Groom, 2002; Lappano and Maggiolini, 2011; Tautermann, 2014), and many human diseases are caused by malfunctioning TMPs (Hardy, 2017; Hasegawa et al., 2017; Hattori et al., 2017; Szablewski, 2017). In plants, TMPs play important roles in solute transport (Pellizzaro et al., 2015), signal recognition and transduction (Kemmerling et al., 2011; Mohd-Radzman et al., 2015; Song et al., 2015; Imin et al., 2018), growth and development (Clark et al., 1997; van der Knaap et al., 1999; Osakabe et al., 2005; ten Hove et al., 2011), and photosynthesis (Liu and Last, 2015; Okumura et al., 2016). In this thesis, the potential Medicago CEP receptor of interested, MtCRA2, is a TMP with one TMD. The aim of the chapter is to establish a bottom up proteomics method for enriching TMP identification, especially for the CEP receptors.

3.1.2 Difficulty in TMPs purification and identification

TMPs can be difficult to identify by proteomic strategies due to their high hydrophobicity (Seddon et al., 2004; Carpenter et al., 2008; Rawlings, 2016), and low abundance (Vit and Petrak, 2017). All proteomic strategies aimed at identifying TMPs begin with membrane enrichment that involves the purification of microsomal, organelle, or plasma membranes (Mirza et al., 2007; Ogawa et al., 2008; Huang et al., 2013; Abdallah et al., 2014; Guillier et al., 2014; Avila, 2015; Aloui et al., 2018). The effectiveness of the proteomic identification of TMPs from these membrane-enriched fractions, however, is compromised by

contaminating cytoplasmic or membrane-associated proteins without transmembrane domains (TMDs). High ionic strength buffers can remove some of these contaminants with varied degrees of success (Marx et al., 2016; Vit and Petrak, 2017). Typically, TMPs represent roughly 20% of proteomic identifications (Chen et al., 2008) and the standard trypsin based purification method typically provides sequence coverage limited to the soluble loops and terminal tails of TMPs (Schey et al., 2013). In addition, TMPs are difficult to solubilize and digest using standard urea solubilization and trypsin Lys-C based procedures. Various detergents, chaotropic agents, organic solvents as well as proteinases and chemical cleavage reagents have been used to help solubilize and/or digest TMPs (Bennett et al., 1992; Newby et al., 2009; Guillier et al., 2014). There are few reports, however, comparing the efficiency of proteomic procedures that aim to identify plant TMPs using mass spectrometry (MS).

3.1.2 Improved bottom up proteomics for TMPs identification

To determine the protein make-up of a given sample using bottom up proteomics requires maximal peptide coverage of the sample. TMPs with a high content of TMDs are under-represented in MS identification since the most commonly used protein purification method for bottom-up proteomics uses 8M urea to solubilize the sample. The poor solubilization and thus poor trypsin digestion (UR-TLC), reduces the identification rate for proteins with a large number of TMPs (Long et al., 2018). The inability of urea to dissolve the membrane most likely contributes to trypsin Lys-C failing to cleave sites including TMDs. Almost all studies of TMPs use certain kinds of detergents to provide solubility (Kar et al., 2017). Common detergents, however, are not MS-friendly and an additional clean-up step is required (Yeung and Stanley, 2010). In addition, detergents are selective in their ability to solubilize TMPs (Arachea et al., 2012; Laganowsky et al., 2013). Formic acid has been used as a substitute for SDS in the solubilization of membrane samples (Zhao et al., 2013).

3.1.3 Improving the identification and coverage of plant TMP in *Medicago* using bottom-up proteomics

In this thesis, we tested a non-detergent based purification and identification strategy designed for analysing TMPs using MS. We tested this method using the model legume, Medicago truncatula, which is an important nitrogen-fixing agricultural crop to augment the prior proteomic analyses of this plant (Natera et al., 2000; Djordjevic et al., 2003; Djordjevic, 2004; Zhang et al., 2006; Djordjevic et al., 2007; Kusumawati et al., 2008; Lee et al., 2013; Marx et al., 2016). In order to directly identify LRR-RLKs and specifically MtCRA2, the microsomal membrane preparations were used to increase the relative abundance of LRR-RLKs in the starting material. All the proteins from microsomal membrane preparation were precipitated by Trichloroacetic Acid (TCA). The protein samples were divided and then subjected to two solubilization extraction and protein cleavage methods to determine if effective solubilization of membrane fractions was responsible for the relatively poor coverage and analysis of Medicago LRR-RLKs in the literature (Marx et al., 2016). The first method utilized the most commonly used method for bottom-up proteomics, which uses a standard 8M urea solubilization step followed by a trypsin Lys-C double digestion (termed UR-TLC). The disadvantage of this technique is that a significant amount of insoluble material remains after solubilization and it is not known how effectively membrane proteins are solubilized for subsequent analysis. It is well recognized that urea solubilization followed by enzymatic cleavage has a low efficiency in solubilizing and cleaving TMPs (Seddon et al., 2004; Carpenter et al., 2008; Rawlings, 2016). Therefore, the relatively poor analysis of LRR-RLKs could simply be due to ineffective solubilization. The second method utilized the superior solubilization and chemical cleavage ability of formic acid (FA) combined with cyanogen bromide (CNBr). FA solubilizes membranes effectively and CB breaks proteins into larger fragments since CNBr chemically cleaves each protein at relatively rare Met residues, which occur on average every 50 amino acids. The large peptide fragments generated by the FACN solubilization/cleavage step were then rendered more compatible for MS analysis by using trypsin/Lys-C double enzymatic digestion, thus generating fragments of MS-suitable average size

with terminal 2+ charged amino acids (i.e. Arginine or Lysine). The second method was abbreviated to FAC-TLC. To increase the resolving power of the mass spectrometer, we separated the tryptic peptides generated by both methods using reverse phase-high pH fractionation. Additionally, we used 2D LC MS-MS combined with the use of in-house columns to run each fraction separately on the mass spectrometer. The in-house columns were packed using a laser puller and a pressure bomb to a length of about 400 mm with a 75 µm ID fused silica housing. These columns gave sharper chromatography than proprietary columns primarily due to their narrower internal diameter, and this resulted in better presentation and fragmentation of the peptides to the mass spectrometer and therefore high quality results. Finally, we used a Q-Exactive plus orbitrap mass spectrometer and analyzed the 20 most abundant ions with a charge state >1, which were fragmented in the high energy C-trap dissociation collision cell to acquire tandem mass spectra, before analyzing the results using Proteome Discoverer.

(i) the popular urea solubilization and trypsin Lys-C digestion based method or (ii) a method for improving the cleavage of TMPs, which utilizes formic acid (FA) solubilization followed by cyanogen bromide (CNBr) cleavage and then trypsin Lys-C digestion (Quach et al., 2003; Girolamo et al., 2010; Vit and Petrak, 2017). The cleaved peptides derived from the two methods were separated by high pH reversed-phase peptide fractionation, and identified by orbitrap-based MS. The effectiveness of the two methods was assessed by (i) comparing total protein identifications, (ii) determining how many TMPs were identified, and (iii) defining how many peptides incorporate all, or part, of the TMD sequences using a new algorithm. In addition, we determined the subcellular location and predicted the function of the proteins identified. The objective of this study was to establish a detergent-free and effective strategy to identify plant TMPs and improve the overall identification and coverage of these proteins using MS. We examined the effectiveness of the two methodological approaches by (i) comparing the number of proteins identified, (ii) determining how many transmembrane proteins (TMPs) were identified, and (iii) defining how many peptides incorporate all, or part, of transmembrane domains (TMD) sequences and (iv) determining the capability of these methods to identify members of the

LRR-RLK family, and specifically MtCRA2, which is a receptor of interest in Chapters 5 and 6.

3.2 Results

3.2.1 Establishment of a FA-CTLC method for improving bottom-up proteomics and membrane protein identification

As a preliminary assessment step, we determined the effectiveness of FA solubilization combined with CNBr treatment to validate that CNBr cleaves to the *C*-terminal side of the comparatively rare methionine (Met) residues to generate large peptide fragments, since methionine occurs, on average, every 50 amino acids. The MS results (Table 3. 1) confirmed that FA solubilization followed by CNBr cleavage alone resulted in the production of large peptides with *C*-terminal Met residues, as expected (Appendix Table 3.1).

Table 3. 1. Improvement of FA-CTLC method in protein identification

Treatment	Identified protein groups number
FA-CNBr cleavage method	2566
FA-CTLC cleavage method	7118
FA-CTLC re-solubilized debris material from URTLC	5644

Peptides of large molecular mass with an uncharged, *C*-terminal Met residue do not ionize well and give poor quality MS/MS spectra. As expected, this resulted in poor coverage of the proteome; only 2566 protein groups were observed using the CNBr cleavage method only (at a 1%FDR; Table 3. 1). This can be remedied by following the CNBr cleavage with trypsin Lys-C digestions.

3.2.2 The FA-CTLC method more effectively solubilized & digest proteins from MM preparations

A summary of the key steps of the two procedures is shown in Fig. 3. 1. To compare the effectiveness of the solubilization and digestion after applying the UR-TLC or FA-CTLC methods, we centrifuged the samples after applying the two methods to identical starting material (MM preparations). The results (Fig. 3. 1, red box) showed that a considerable pellet of insoluble material remained after UR-TLC, but not after FA-CTLC. A common practice in MS is to apply the same extraction procedure to the insoluble material so that more hydrophobic components can be liberated and identified by MS. A second round of the UR-TLC was applied to the pellet, but after overnight digestion and subsequent re-centrifugation, the pellet remained. By contrast, an application of the FA-CTLC method to the insoluble pellet that remained after using the UR-TLC method resulted in no observable pellet after centrifugation (Fig. 3. 1, red box). An MS analysis of this FA-CTLC re-solubilized pellet material identified 5644 protein groups (at a 1% FDR) in the insoluble material that remained after UR-TLC (Appendix Table 3. 2), in which the second round of FA-CTLC identified 979 protein groups (Appendix Table 3. 3), that UR-TLC failed to identify. This demonstrated that considerable protein material remained in the pellet, which may or may not be represented in the proteome. These results also show that the FA-CTLC is more effective in solubilizing and digesting the proteins from MM preparations as shown in the method flowchart (Fig. 3. 1, Table. 3. 1).

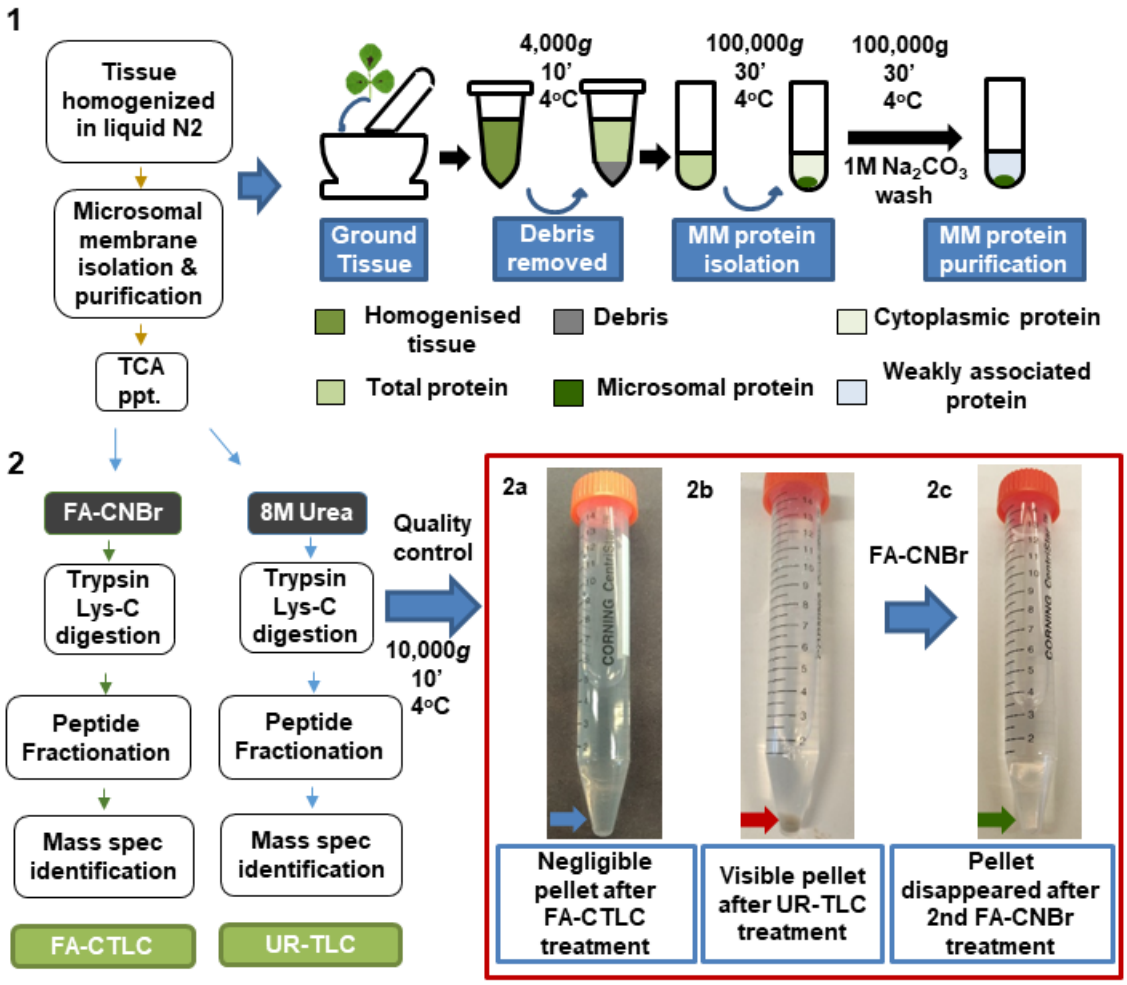


Figure 3.1. Flow chart summarizing the two solubilization and protein cleavage methodologies used in this study

Red box: Negligible insoluble material remained after using the FA-CTLC method (2a, blue arrow), whereas significant insoluble material remained after using the UR-TLC method (2b, red arrow). The remaining insoluble material after UR-TLC was solubilized and digested using the FA-CTLC method. MS analysis showed the presence of a wide range of proteins in the pellet (Table. S2). Subsequent centrifugation of this re-solubilized and FA-CTLC treated material showed negligible insoluble material remained (2c, green arrow).

3.2.3 Assessment of inter-sample reproducibility and the effectiveness protein identification by MS after using the FA-CTLC or the UR-TLC method

We assessed the reproducibility of the two methods by examining the proteins

identified using three biological repeats from each treatment (Fig. 3. 2). When considering proteins with a FDR of <1%, we identified 4171 proteins common to all three biological repeats after UR-TLC, and 3609 proteins groups common to all three biological repeats after FA-CTLC. The reproducibility of the proteins identified was 51.1% and 48.8% for the UR-TLC and FA-CTLC methods, respectively (Fig. 3. 2A and 3. 2B). We further compared the 4171 proteins groups common to all three biological repeats after UR-TLC (Fig. 3. 2A) to the 3609 proteins groups common to all three biological repeats after FA-CTLC (Fig. 3. 2B) and found 2981 groups common proteins were identified by both methods (Fig. 3. 2C). Based on the results, 629 protein groups are unique to FA-CLC identification method while 1191 protein groups are unique to UR-TLC. Appendix Table 3. 4 shows the complete list of proteins groups identified in three biological repeats from both methods. The 1191 UR-TLC-specific protein groups are shown in Appendix Table 3. 5 and the 629 FA-CTLC-specific protein groups identified after using the FA-CTLC are shown in Appendix Table 3. 6.

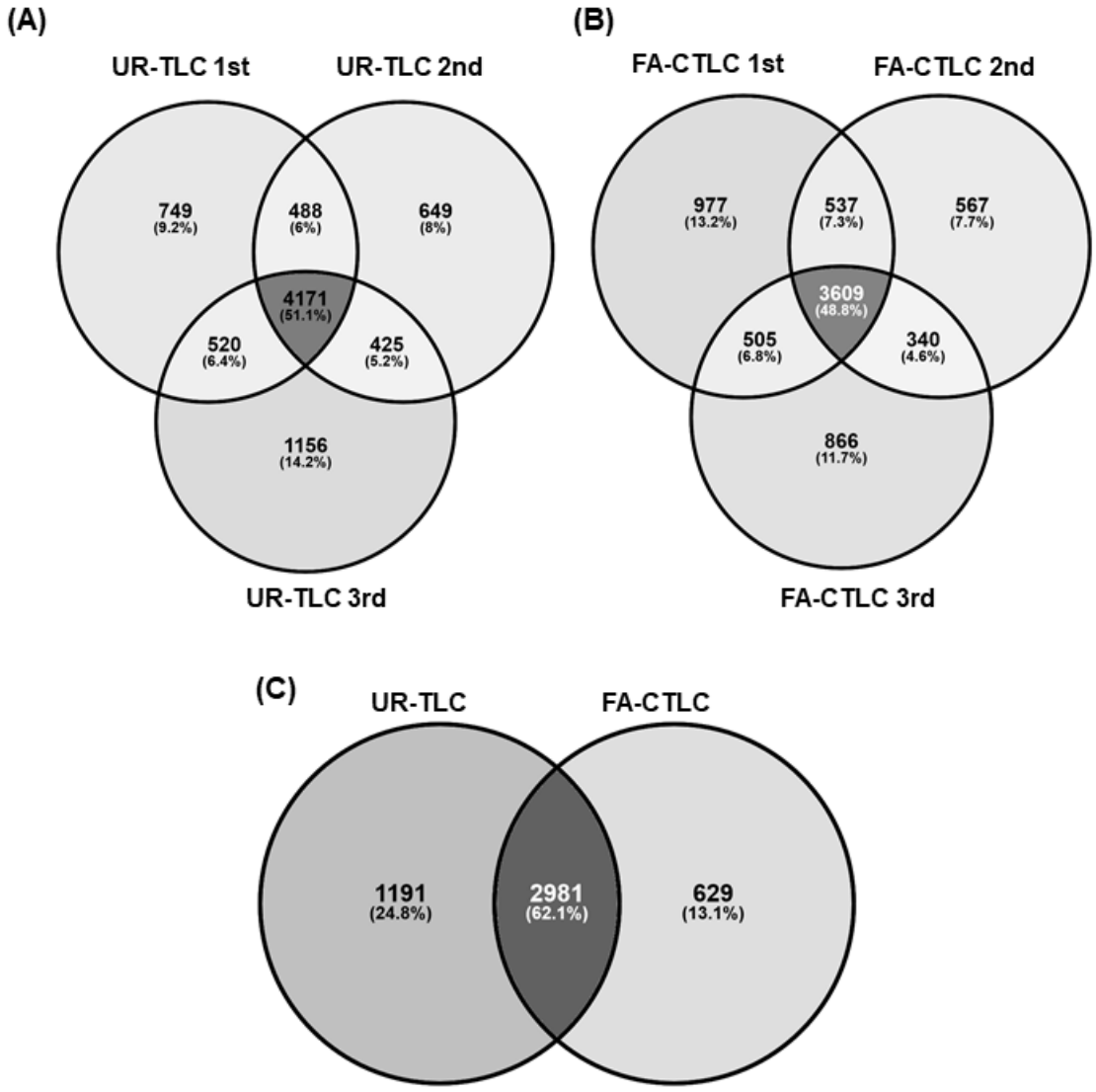


Figure 3.2. Analysis of the proteins identified by MS-MS after UR-TLC or FA-CTLC

(A) The reproducibility of protein identifications between the three biological repeats following UR-TLC treatment (B) Reproducibility of protein identifications between the three biological repeats following FA-CTLC treatment (C) A comparison of proteins identified by each treatment.

(http://bioinfogp.cnb.csic.es/tools/ venny/)

When considering proteins identified from three repeats, 7946 protein groups were identified following UR-TLC. 7118 protein groups were identified following FA-CTLC (Appendix Table 3. 7). After combining the results derived from the two methods, we identified 8993 protein groups in total (Appendix Table 3. 8). To determine the effectiveness of the two methods at identifying TMPs, the identified proteins were analyzed for the presence of TMDs.

3.2.4 The FA-CTLC method preferentially identifies TMPs with a higher number of TMDs

By using the TMHMM algorithm, 23.26% of the proteins in the *Medicago* Uniprot database were identified as TMPs. From the 7946 protein groups identified after UR-TLC treatment, 2817 protein groups (35.45%) contained at least one TMD, and of the 7118 protein groups identified from the FA-CTLC treatment, 2784 protein groups (39.11%) contained at least one TMD. Therefore, a higher percentage of TMPs can be identified by using the FA-CTLC method compared to using the UR-TLC method. Additional analysis showed that there were 5129 protein groups identified from UR-TLC treatment and 4334 protein groups identified from FA-CTLC treatment with 0 TMDs. This result suggests that the published procedures for removing non-membrane proteins (e.g. using Na carbonate washes at pH 11) have poor efficacy. We further analyzed the TMPs distribution in different biological repeats from each purification method using the TMHMM algorithm. The TMPs results are shown in Fig. 3. About half of the TMD containing proteins identified using either method had only one TMD and both purification methods gave no significance difference in distribution of TMDs to that predicted by analyzing the theoretical distribution of TMDs in all *Medicago* TMPs (inset of Fig. 3. 3, confirmed by Chi-Square Test, $p=0.32$). This suggests that there is no major bias of either method in identifying TMPs, and that the most abundant TMPs are likely to populate the lists of proteins identified. However, the significant difference between the proteins identified was that UR-TLC method preferentially identified proteins with only one TMD, whereas the FA-CTLC method preferentially identified significantly more proteins with greater than 4 TMDs (Fig. 3. 3; $p < 0.05$). By combining the TMPs identified by

UR-TLC and FA-CTLC, 3289 TMP groups were identified, or 36.57% of all TMPs.

The proteins identified after using UR-TLC or FA-CTLC methods were also examined using the grand average of hydropathy (GRAVY) algorithm (Kyte and Doolittle, 1982). The GRAVY index indicates the hydrophobicity of the proteins, calculated by adding the hydropathy value for each residue and dividing by the length of the sequence. Proteins with a GRAVY scores above 0 are more likely to be hydrophobic proteins (Magdeldin et al., 2012). The GRAVY results (Fig. 3. 3B) showed that there were 1420 (19.95%) protein groups identified by FA-CTLC, and 1313 (16.52%) proteins identified by UR-TLC, which displayed a GRAVY score greater than zero. These results indicated that the FA-CTLC can preferably purified proteins which are hydrophobic.

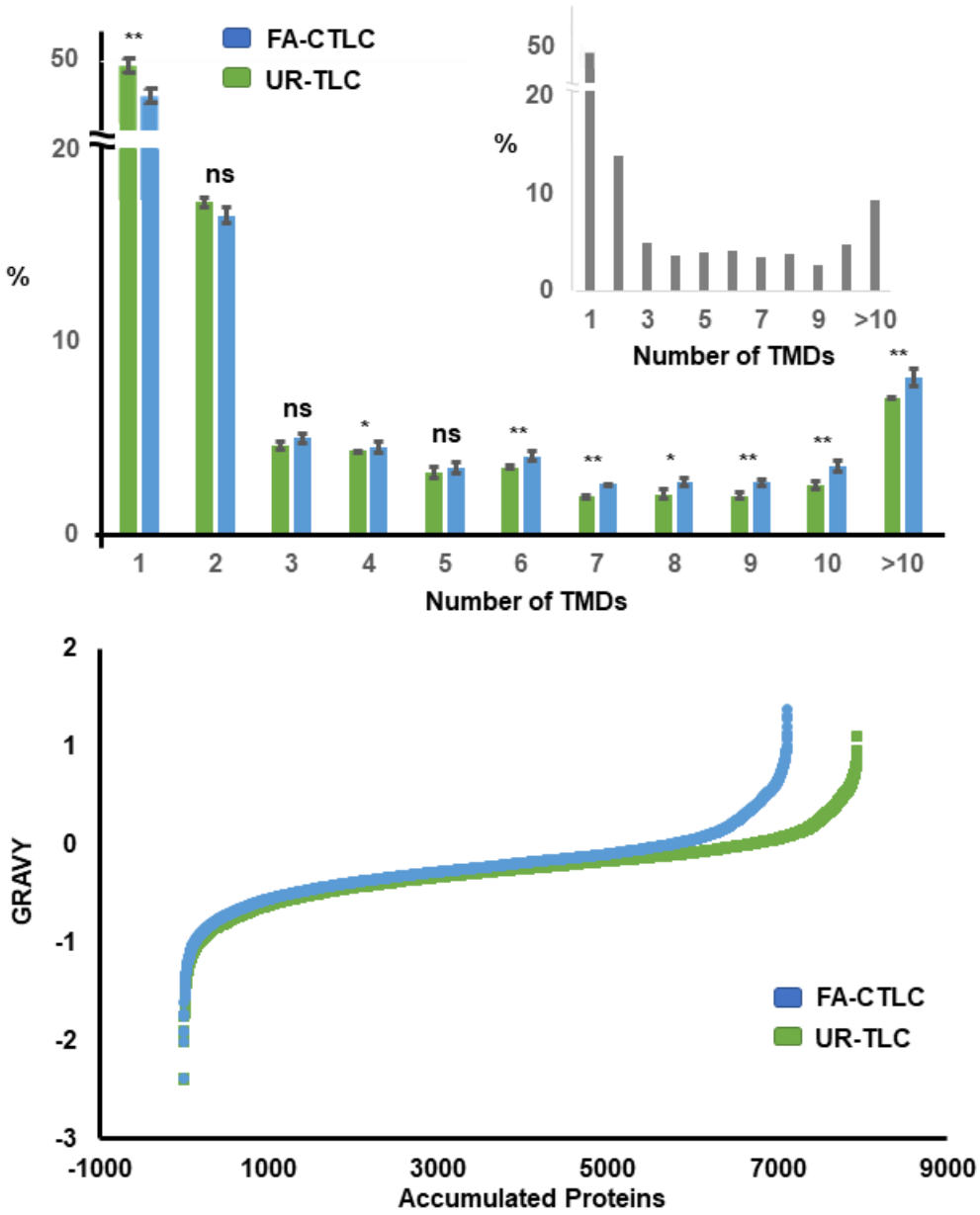


Figure 3.3. Distribution of proteins identified after applying the FA-CTLC or UR-TLC methods to *Medicago* MM preparations from three biological repeats based on TMD number and GRAVY score.

(A) The proteins identified after analyzing the UR-TLC or FA-CTLC treated samples from three biological repeats were submitted to the TMHMM sever, respectively. The total predicted number of TMP groups in the UR-TLC and FA-CTLC samples was 2817 (35.45%) and 2784 (39.11%), respectively. The predicted TMD distribution of the Medicago proteins in the UniProt database is shown in the inset panel as a comparison. There were 23.26% proteins were predicted to be TMPs. * $P \leq 0.05$, ** $P \leq 0.01$, *** $P \leq 0.001$ (Two tail Student's t - test). Error bars =standard error, n=3. (B) The GRAVY scores were calculated

(Kyte and Doolittle, 1982) from the proteins identified after analyzing the UR-TLC or FA-CTLC base on the previous published literature (Kyte and Doolittle, 1982). Approximately 20% of proteins identified by FA-CTLC displayed a GRAVY score greater than zero, and 17% of proteins identified by UR-TLC. Proteins had a hydrophobicity scores above 0 are more likely to be TMPs (hydrophobic).

We developed the TRAMDOMI algorithm to identify the peptides that contain all or part of TMDs motifs within the TMPs. This algorithm enabled us to quantify the relative ability of each method to identify peptides with TMD motifs. The search results showed that the FA-CTLC method can identify 9.36 times more TMD-containing peptides than the UR-TLC method (811 compared to 87; Fig. 3. 4A). Therefore, the results indicate that the FA-CTLC method is more effective at detecting peptides within TMPs that have TMD motifs, which boosts the number of TMPs identified. A list of identified TMPs and the TMDs identified using both purification methods is shown in Supplemental Appendix Table 3.9. To further illustrate the difference between the two methods, we compared the peptides identified for the MFS/sugar transporter (MTR_7g005910), which has 12 predicted TMDs (Fig. 3. 4B and 3. 4C). Clearly, the FA-CTLC method identified more peptides within MTR_7g005910 with TMD motifs, whereas the UR-TLC method only identified MTR_7g005910 peptides predicted to loop into the cytoplasm and none that contained TMD motifs.

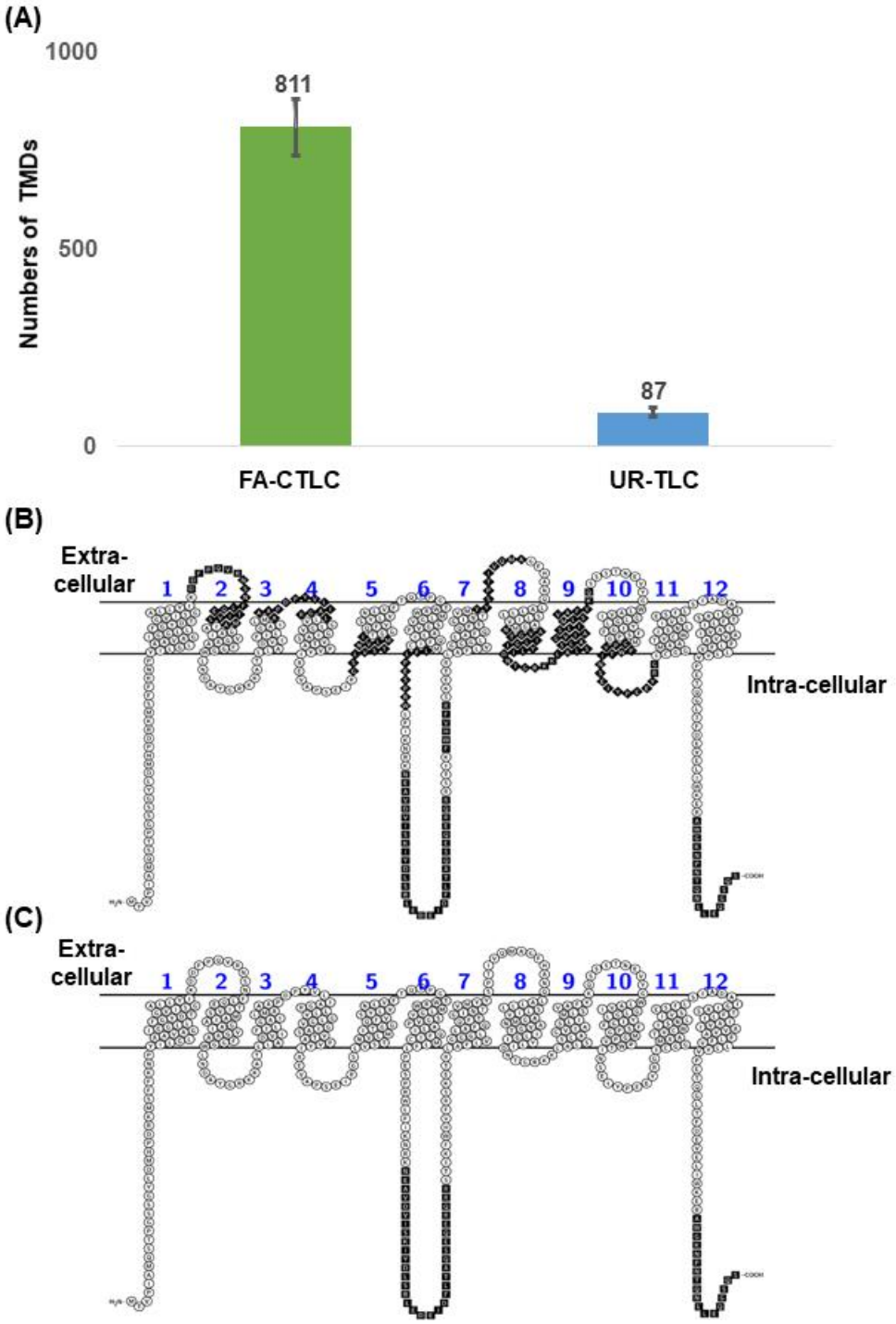


Figure 3.4. FA-CTLC method preferentially identifies peptides with TMD motifs.

(A) A comparison of the total TMDs identified numbers from both purification methods. The FA-CTLC methods identified 9.3-fold more TMDs than the UR-TLC method. (B, C) The peptides identified in the MTR_7g005910 transporter of *Medicago* after using the FA-CTLC (B) or UR-TLC method (C). The peptides which were identified by mass spectrometry (MS) were colored in black. The

schematic diagrams were made using the Protter online tool (Omasits et al., 2014)(http://wlab.ethz.ch/protter/start/).

3.2.5 Transporter proteins are preferentially identified by using the FA-CTLC method

Given that each solubilization and cleavage method identified distinct classes of peptides, a Mercator analysis was done to determine if the two methods resulted in the enrichment of the identification of proteins with different functions. The results (Fig. 3. 5) show that the TMPs containing one TMD, which were preferentially identified by UR-TLC, were mostly functionally assigned as being signaling proteins (20.81%), and the proteins preferentially identified by FA-CTLC, which had either greater than 4 TMDs were predominantly functionally assigned as being transporters (58.01%). We further examined the difference between two data sets by a Binomial test. The results showed the proteins identified by the FA-CLTC method were significant different in 18 categories when compared to the UR-TLC method (Fig. 3. 5). The complete protein functional analysis list is shown in Appendix Table 3.10 and the binomial test results is shown in Appendix Table 3.11.

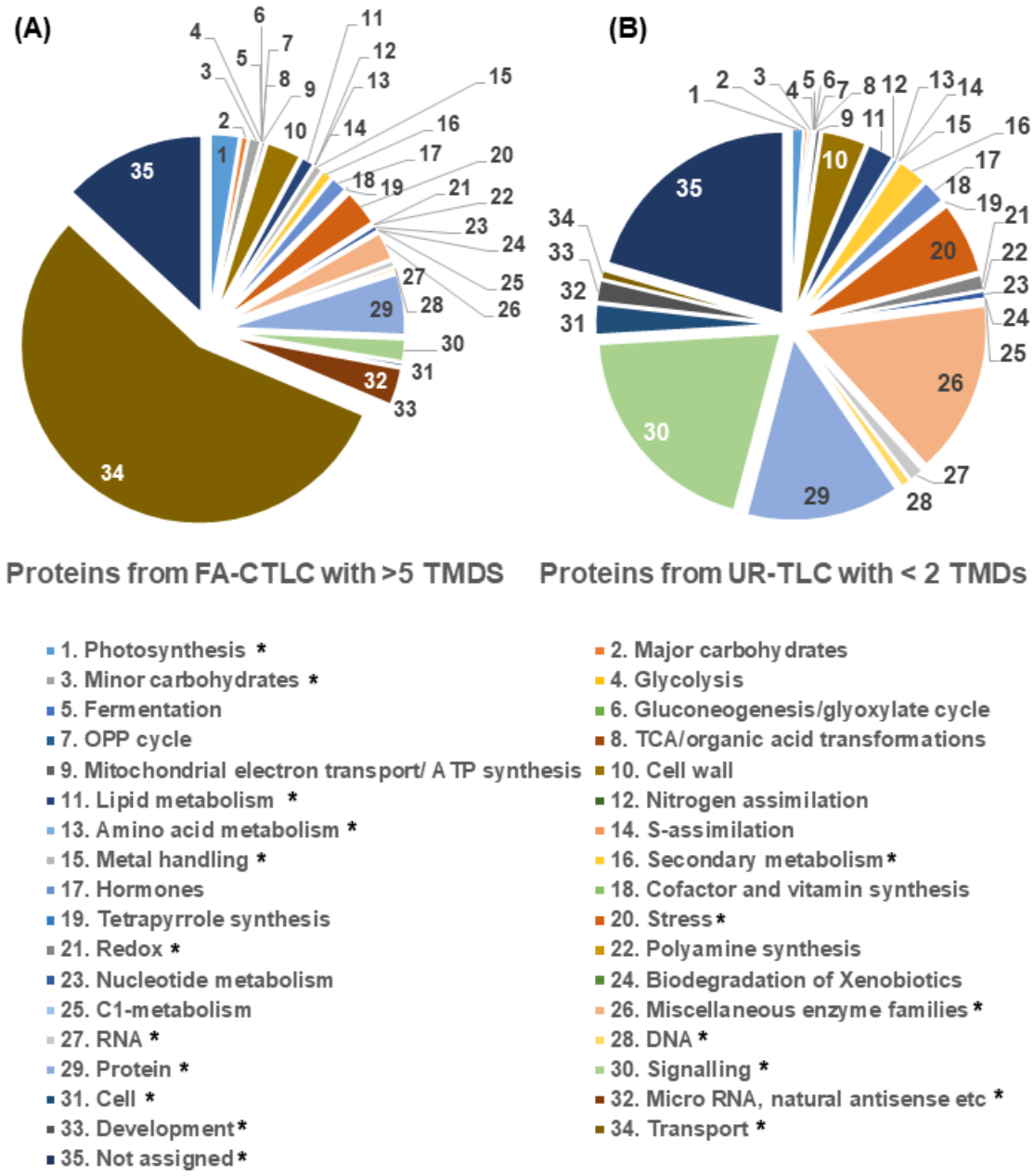


Figure 3.5. Functional analysis of the proteins preferentially identified using either the FA-CTLC or UR-TLC methods

(A) The proteins preferentially identified after using FA-CTLC (i.e. with > 5 TMDs) were predominantly transporters. (B) The proteins preferentially identified after using UR-TLC (i.e. with one TMD) were predominantly signaling proteins. The category which FA-CTLC method had significant difference with p-value < 0.05 were labelled with *.

To determine the likely membrane where the TMPs identified reside, all proteins were analyzed for their sub-cellular location using LOCALIZER (Figure 3. 6). Irrespective of the method used, an analysis of the TMPs identified showed that

was a similar distribution of proteins predicted to reside in the membranes of the nucleus, chloroplast, or mitochondria. For both methods, the FA-CTLC method identified significantly more TMPs predicted where the subcellular location could not be assigned to an organelle (p=0.044, n=3). The complete subcellular location prediction list is shown in Appendix Table 3.12.

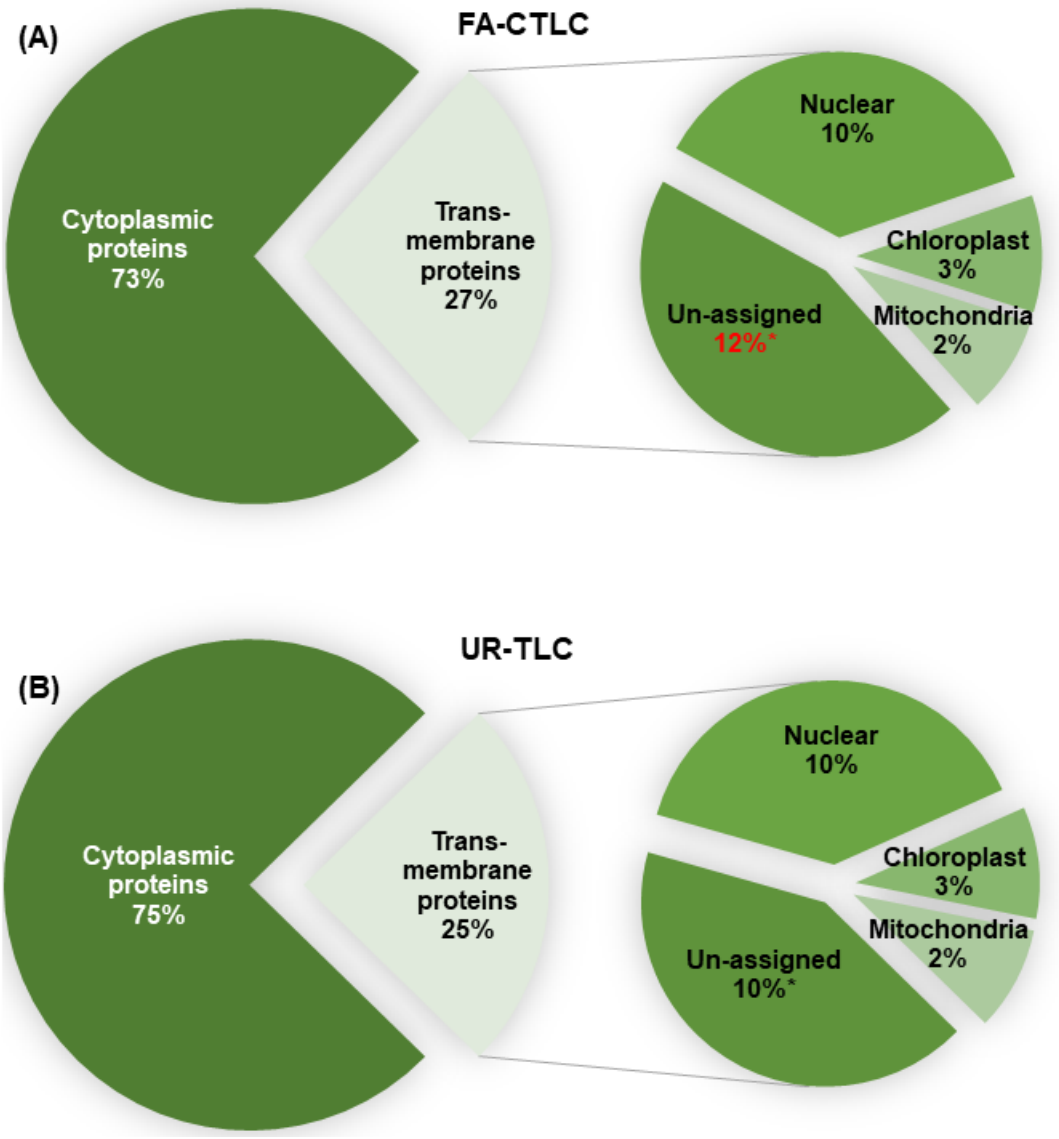


Figure 3.6. The predicted subcellular location of the proteins identified after using FA-CTLC or UR-TLC.

(A, B) Between 73-75% of the proteins identified in the MM preparations were cytoplasmic proteins. Of the proteins identified to be TMPs there was no significant difference in the identity of the proteins predicted to reside in the nuclear, chloroplast or mitochondrial membranes. A t-test confirmed a significance difference ($p < 0.05$, n=3) between two methods in identifying

proteins where the subcellular location could not be assigned (the “unassigned” category).

3.2.6 There was insufficient enrichment of LRR-RLKs after using the FA-CTLC or the UR-TLC method

In this study, the original purpose is to establish a methodology to identify the ligand receptor interaction between CEP peptide and its receptor, potentially MtCRA2. Based on the uniport database, there were 538 LRR-RLKs family proteins in Medicago (Appendix Table 3. 13) including MtCRA2. LRR-RLKs have a single pass TMD and are located in the plasma membrane where they serve roles in interacting with extracellular signals (such as metabolite and peptide hormones) (Liu et al., 2017; Campos et al., 2020; Man et al., 2020). As the identification and analysis of this class of proteins was fundamental for later steps in this thesis, the ability of both methods to identify these proteins was assessed. From the MS results we identified 137 LRR-RLKs protein groups (25.46 %) of from UR-TLC methods, 120 LRR-RLKs protein groups (22.30%) from FA-CTLC methods, which had at least one unique peptide. The average coverage of identified LRR-RLKs were 10.53% and 10.31%, with average of 6 and 5 unique peptides identified for UR-TLC and FA-CTLC, respectively. The UR-TLC identified a slightly more LRR-RLKs then FA-CTLC. The results matched previous theory that UR-TLC can preferable enrich the TMP with single TMD, such as membrane signal proteins like LRR-RLKs family. Besides, both purifications only identified 1% of MtCRA2 with one unique peptide. Thus, both purifications methods were not suitable for extensive LRR-RLKs analysis due to low sequence coverage. Considering previous studies showed using in situ hybridization that MtCRA2 was expressed vascular tissue of leaves (Huault et al., 2014) we established an alternative approach in Chapters 4 and 5, which combined Medicago leaf vascular tissue purification and ligand-receptor crosslinking to examine MtCRA2 function and localization.

3.3 Conclusion

In this chapter, we compared the capability of FA-CTLC to identify TMPs in MM preparation from young Medicago shoot tissue with the commonly used UR-TLC method. The GRAVY score and the number of TMDs are the two main factors used in this chapter to determine the hydrophobicity of proteins. The results both indicated that FA-CTLC can significantly identify more TMPs with a higher proportion of TMDs, and the more hydrophobic proteins which had GRAVY scores greater than 0. The FA-CTLC methods resulted in the identification the proteins with a higher proportion of TMDs with a positive bias towards transporter proteins, whereas UR-TLC preferentially enriched signaling proteins, which contain one TMD. We further compared the identified peptide with the TMDs by TRANDOMI algorithm. The FA-CTLC method identified 9 times more TMDs and enriched more TMPs than UR-TLC method. The results suggested FA-CTLC is a more effective approach for TMDs identification, and better coverage for TMPs.

However, there were no significant enrichment of proteins from the LRR-RLK family (for example MtCRA2) utilizing either the UR-TLC or the FA-CTLC method. In order to identify the potential receptor for MtCEP1, a more direct method was employed in Chapter 4 and 5, which involved the enrichment vascular tissue, combined with crosslinking labelling technique for investigate the interaction of CEP peptide ligand, CEP, with the putative CEP receptor CRA2.

Literature cited

Campos R, Goff J, Rodriguez-Furlan C, Van Norman JM (2020) The Arabidopsis Receptor Kinase IRK Is Polarized and Represses Specific Cell Divisions in Roots. Dev Cell **52:** 183-195.e184

He Y, Zhou J, Shan L, Meng X (2018) Plant cell surface receptor-mediated signaling - a common theme amid diversity. Journal of cell science **131:** jcs209353

Huault E, Laffont C, Wen J, Mysore KS, Ratet P, Duc G, Frugier F (2014) Local and systemic regulation of plant root system architecture and symbiotic nodulation by a receptor-like kinase. PLoS Genet **10:** e1004891

Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol **157:** 105-132

Liu PL, Du L, Huang Y, Gao SM, Yu M (2017) Origin and diversification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants. BMC Evol Biol **17:** 47

Magdeldin S, Yoshida Y, Li H, Maeda Y, Yokoyama M, Enany S, Zhang Y, Xu B, Fujinaka H, Yaoita E, Sasaki S, Yamamoto T (2012) Murine colon proteome and characterization of the protein pathways. BioData mining **5:** 11

Man J, Gallagher JP, Bartlett M (2020) Structural evolution drives diversification of the large LRR-RLK gene family. New Phytol

Marx H, Minogue CE, Jayaraman D, Richards AL, Kwiecien NW, Siahpirani AF, Rajasekar S, Maeda J, Garcia K, Del Valle-Echevarria AR, Volkening JD, Westphall MS, Roy S, Sussman MR, Ane JM, Coon JJ (2016) A proteomic atlas of the legume *Medicago truncatula* and its nitrogen-fixing endosymbiont *Sinorhizobium meliloti*. Nat Biotechnol **34:** 1198-1205

Omasits U, Ahrens CH, Muller S, Wollscheid B (2014) Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics **30:** 884-886

Tang D, Wang G, Zhou J-M (2017) Receptor Kinases in Plant-Pathogen Interactions: More Than Pattern Recognition. The Plant cell **29:** 618-

637

Chapter 4: Medicago vasculature purification and formaldehyde mediated crosslinking ligand- receptor interaction identification

4.1 Introduction

In Chapter 3, an approach was developed to enrich and identify membrane proteins isolated from shoot tissue using MS. Although this method was advantageous, MS lacked the ability to comprehensively and reliably identify low abundance membrane proteins such as LRR-RLKs. One of the aims of this thesis was to develop ways to demonstrate that the LRR-RLK, MtCRA2, was a cognate receptor partner for the peptide ligands of the CEP family.

Therefore, this chapter developed a new approach to validate that members of the CEP family interacted with the MtCRA2 receptor and that CRA2 is a functionally similar to the verified CEP receptor, CEPR1. Since *in situ* localization, grafting and genetic studies demonstrated that MtCRA2 and its likely functional orthologue, AtCEPR1 were expressed in the vascular tissue of leaves in the Fabaceae species (Huault et al., 2014) Medicago (Huault et al., 2014; Mohd-Radzman et al., 2016; Laffont et al., 2019) and the Brassicaceae species Arabidopsis (Tabata et al., 2014), respectively, a chemical crosslinking strategy was devised to demonstrate an interaction between the CEP ligand and Medicago CRA2 *in vivo*. To achieve this, live Medicago vascular tissue (VT) tissues were semi-purified from leaves and biologically active and fluorescently tagged CEP ligands were synthesized. The purification of the VT afforded unhindered accessibility of this tissue to the synthesized ligand as well as allowed the visualisation of any interaction of the fluorescently CEP ligand to specific sites on these tissues to be assessed by confocal microscopy. We also assessed if a natural affinity of the CEP ligand towards its receptor could be cemented by using crosslinking strategies. Formaldehyde and dimethyl pimelimidate (DMP) were assessed for their effectiveness as crosslinking reagents.

4.1.1 Vasculature tissue (VT) isolation for VT specific protein purification

The vascular system of plants is an assemblage of conducting tissues and associated supportive fibres. In leaf tissue the vasculature is extended from stem, which contain vascular bundles composed of xylem and phloem. The xylem consists of tracheids and xylem vessels, which bring water and minerals

into the leaves along with root to shoot signal peptides, such as CEPs, CLEs (Shabala et al., 2016). Phloem is composed of various specialized cells called sieve tubes, companion cells, phloem fibres, and phloem parenchyma cells. The phloem transports the photosynthetic products from the leaf to the other parts of the plant (Fig. 4.1) (Smith A et al., 2009). Several receptors are localized in VT, such as the TDIF (tracheary element differentiation inhibitory factor) and its receptor TDR/PXY (TDIF Receptor/Phloem Intercalated With Xylem), which belongs to the class XI LRR-RLK family (Hirakawa et al., 2008) and is crucial for vascular stem cell maintenance (Lucas et al., 2013). Previous studies showed MtCRA2 is phloem localized as suggested by *in situ* gene expression studies (Fig. 1.6A) (Huault et al., 2014). However, if we used the whole leaf for ligand-receptor localization experiment, the mesophyll cells provide too much auto fluorescence which makes the localisation of ligand-receptor interaction hard to distinguish by confocal microscopy.

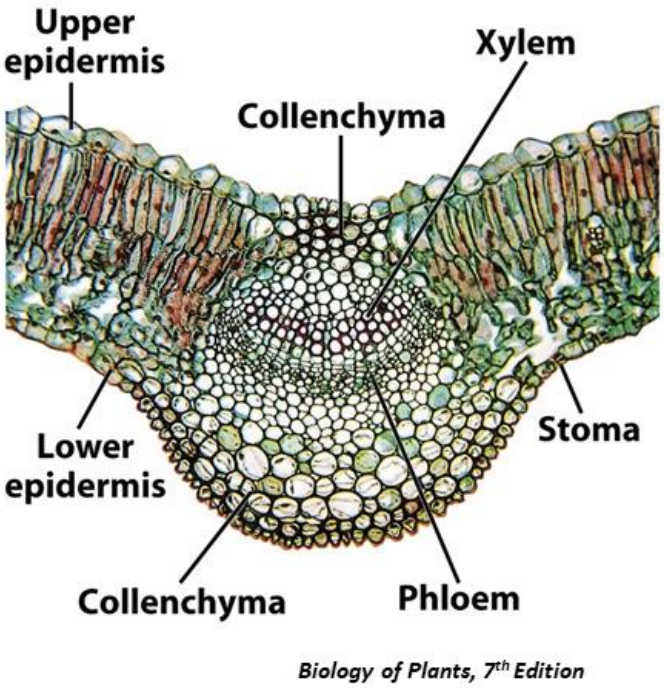


Figure 4.1. The cross-section anatomy of leaf

4.1.2 Formaldehyde mediated crosslinking for PPIs identification

In recent studies, cross-linking reagents are more commonly used to identify the protein-protein interaction. A number of formaldehyde based experimental approaches are very promising in their ability to accurately and efficiently identify novel protein-protein interaction complexes (Sutherland et al., 2008). Formaldehyde is routinely used to cross-link proteins, inactivating, immobilising or stabilising nearby macromolecules (Toth and Biggin, 2000; Casadonte and Caprioli, 2011; Hoffman et al., 2015; Sinz et al., 2015). A key advantage of formaldehyde as a cross-linking reagent is its small size. This enables the molecule to easily permeate cells and facilitates very specific cross-linking reactions between amino acids on two proteins in close proximity (2.3-2.7 Å) (Metz et al., 2004). Additionally, formaldehyde cross-linking reactions are reversible and compatible with mass spectrometry (MS) (Sutherland et al., 2008). Although several approaches exist to demonstrate PPIs studies, formaldehyde assisted crosslinking has been rarely used to demonstrate peptide-receptor interactions. It mainly because formaldehyde mediate peptide-peptide crosslinking requires suitable amino acid side groups (e.g. Cysteine, Histidine, Tryptophan and Arginine groups and Lysine groups) to be present in the interacting molecules (Metz et al., 2004). Therefore, it is possible that confocal microscopy could be used to visualize the formaldehyde assisted crosslinking of a fluorescently tagged CEP ligand to it receptor *in vivo* by exploiting the natural affinity of the ligand for its receptor. This has the advantage of obviating the need to manipulate the putative receptor (e.g. by adding an *N*- or *C*-terminal tag), however, success is contingent on the fluorescent tag not destroying the biological activity of the peptide ligand by negating binding.

Using formaldehyde, we examined the interaction between fluorescein isothiocyanate (FITC) labelled MtCEP1 peptide (FITC-MtCEP1) and MtCRA2 using semi-purified Medicago leaf vascular tissue. Next, the binding efficacy of FITC-MtCEP1 to MtCRA2 was defined and the minimum formaldehyde

concentration for effective ligand receptor crosslinking determined. Finally, we tested the utility of this methodology to see whether it can be employed to validate ligand receptor interaction in different plant species.

4.2 Results

4.2.1 Synthesis of biologically active FITC-MtCEP1

CEP peptides have a 15 amino acid conserved domain, which is conserved across species (Delay et al., 2013; Imin et al., 2013; Ogilvie et al., 2014) and native CEP peptides predominantly secreted from plant roots or isolated from xylem fluids correspond to this 15 amino acid domain (Ohyama et al., 2008; Tabata et al., 2014; Mohd-Radzman et al., 2015; Okamoto et al., 2015; Patel et al., 2018). Some naturally occurring CEPs, however, have *N*- or *C*-terminal extensions that include amino acids flanking the 15 amino acid conserved domain(Patel et al., 2018). When the 15 amino acid MtCEP1 derivative is added to Medicago roots, it decreases the number of lateral roots and increases the number of root nodules induced by *Sinorhizobium meliloti* (Imin et al., 2013; Mohd-Radzman et al., 2015). By contrast, when the naturally occurring MtCEP1 peptides with *N*- or *C*- terminal extensions were synthesized and tested for biological activity, a diminished ability to enhance root nodule number was often observed although the ability to reduce lateral root number was retained. These observations collectively suggest that the interaction of CEP peptide with the native Medicago CEP receptor is sensitive to peptide structure and that loss of the ability to induce root nodules is more sensitive to peptide structure than the ability to reduce lateral root number. Therefore, care needed to be exercised to design MtCEP1 derivatives that retained the full spectrum of biological activity.

We synthesized an MtCEP1 peptide derivative with an N-terminally labelled fluorescent FITC tag (Fig. 4.2). We added the FITC tag to the *N*-terminus based on prior results where a carboxytetramethylrhodamine (TAMRA) tag was added to either the *N*- or the *C*-terminus of MtCEP1. Since the *N*-terminally tagged TAMRA peptides only retained full biologically activity, and TAMRA and FITC

have similar molecular masses (386 and 389, respectively), we tested the biological activity of FITC-MtCEP1 relative to MtCEP1. FITC-MtCEP1 retained the expected biological activity (Fig. 4.3) which is and inhibition of lateral root number and a promotion of root nodule number. This result validated that the *N*-terminal FITC modification did not inactivate the biological activity of the CEP peptide. Thus, we used FITC-MtCEP1 for the following experiments.

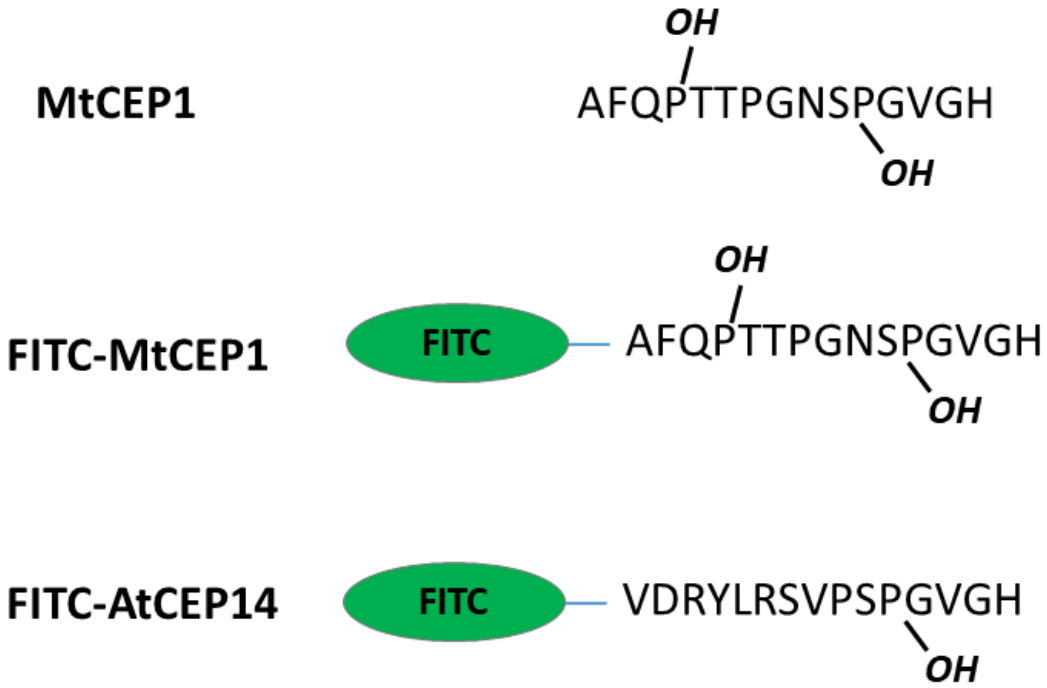


Figure 4.2. The CEP peptide derivatives used in this study

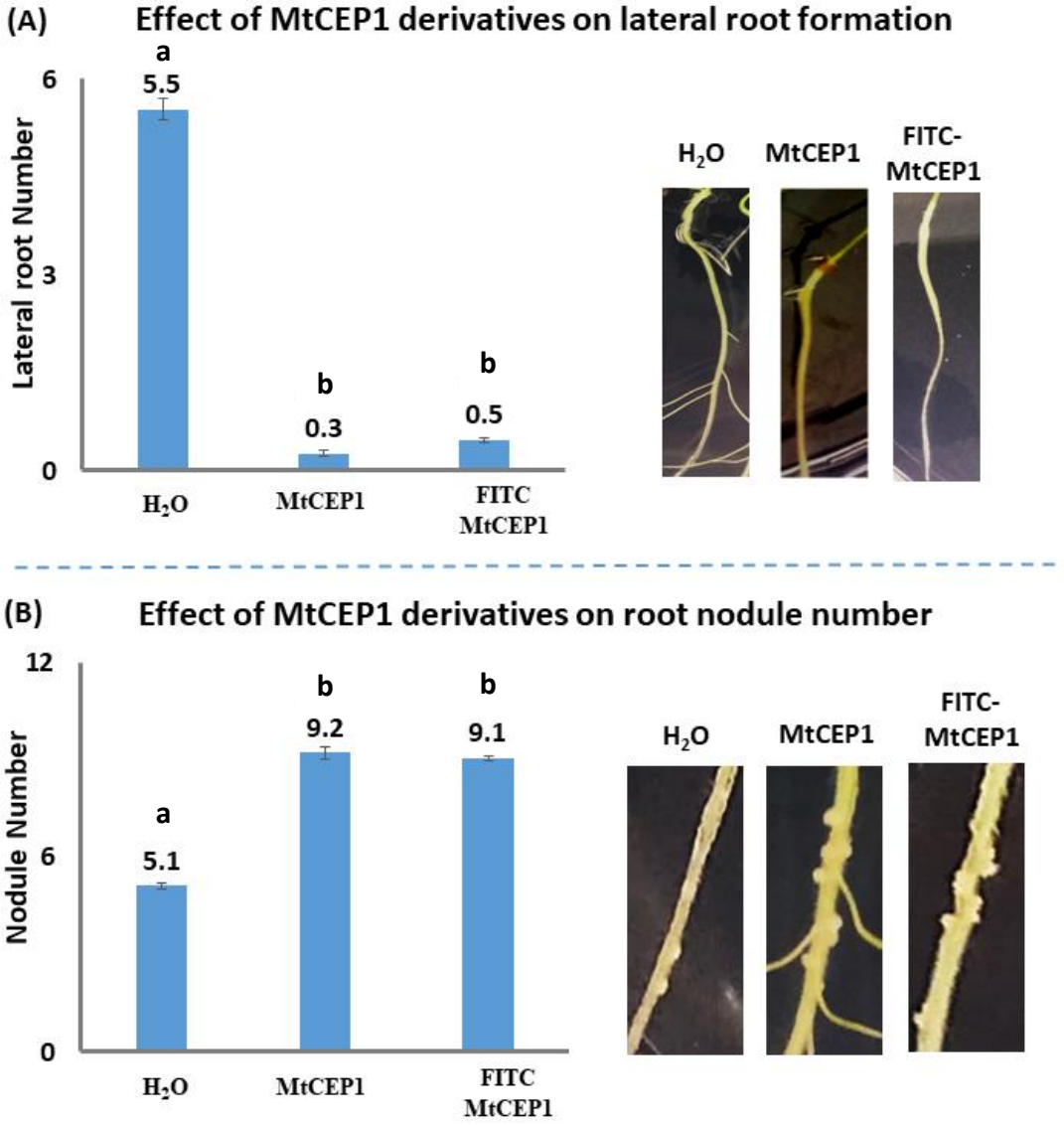


Figure 4.3. The MtCEP1 derivative with a FITC tag is biologically active in Medicago. The synthetic peptide was added into the medium for Medicago seeding growth. The Medicago seedings were tested the lateral root and nodule numbers for peptide biological activity (see section 2.15) **(A)(B)** The biological activity of FITC-MtCEP1 derivative was accessed on A17 roots. **(A)** The effect of FITC-MtCEP1 (1 x 10^{-6}M) derivative on lateral root and **(B)** root nodule number. Lateral organ formation in A and B was scored at 14 dpi. Different letters indicate a statistically significant difference ($P \leq 0.05$, one-way analysis of variance (ANOVA) followed by Tukey multiple comparisons test) (error bars, ±SE, $n \geq 15$)..

4.2.2 Medicago leaf VT isolation

Next, we developed strategies for the isolation Medicago VT and confirmed that the cells in the VT were viable. The technique devised involved using Medicago leaves soaked in buffer containing cell wall degrading enzymes. These leaves were subjected to probe sonication to remove most (~80%) of the mesophyll cells and expose the VT. The isolated VT was subjected to light microscopy to make sure the sonication procedure is sufficient to remove the mesophyll cells but not to cause the dissociation of the VT (Fig. 4.4A). The viability of semi-purified VT cells was verified by using propidium iodide (PI) staining (Fig. 4.4B), which discriminates between viable and non-viable cells. PI binds to DNA through the intercalation of nucleobases. Because PI is not permeable to live cells, it is also widely used to detect dead cells in a population. The results showed that sonication of leaves in a pH buffered solution containing cell wall degrading enzymes was an effective method to (a) remove mesophyll cells, which do not express CRA2, (b) retain VT cell viability, (c) provide an enrichment of tissue that could interact with CEP peptides, and, (d) provide easy access for a crosslinking chemical. This procedure would also enable direct visualization of potential CEP-CEP receptor interactions, and perhaps other puttive ligand-receptor pairsvia confocal microscopy.

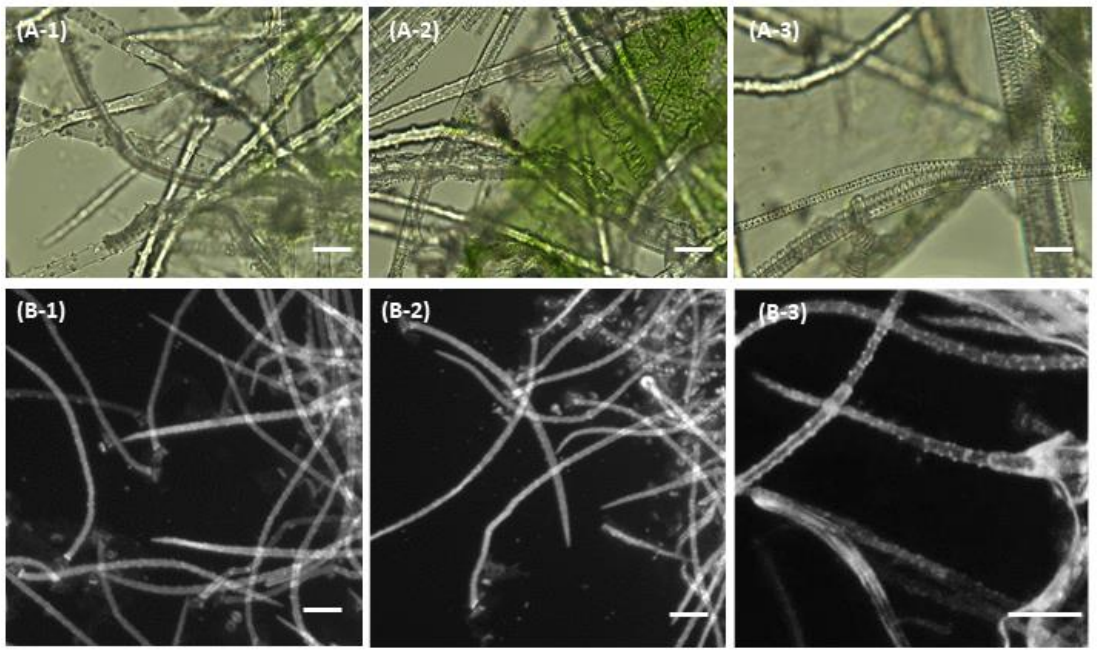


Figure 4.4. Light microscopy images of Medicago vasculature tissue viability staining. The Medicago leaves were soaked in the enzyme solution

and subjected to probe sonication to remove the attached mesophyll cell (see section 2.16) The isolated Medicago VT was confirmed that **(A)** most of the mesophyll cells (green) were removed from VT **(B)** the viability of the vascular tissue cells using propidium iodide staining. Representative pictures from 3 independent experiments are shown. Scale bar = 100 µm

4.2.3 VT related protein purification and identification

To further validate VT enrichment, we employed a detergent based approach to extract the proteins from the semi purified VT preparations and we examined the protein content using mass spectrometry to determine if the isolated VT really enriched in VT related proteins. Purified VT protein (10 mg) was separated by SDS-PAGE, cut into 10 gel slices based on the molecular weight and each gel slice subjected to in-gel trypsin digestion (Fig. 4.5A). Based on the MS identification, 4950 proteins (2380 protein groups) were identified in the Medicago VT. Among these 4950 proteins, there were 931 membrane proteins, (522 protein group) identified from isolated Medicago VT (Fig. 4.5B).

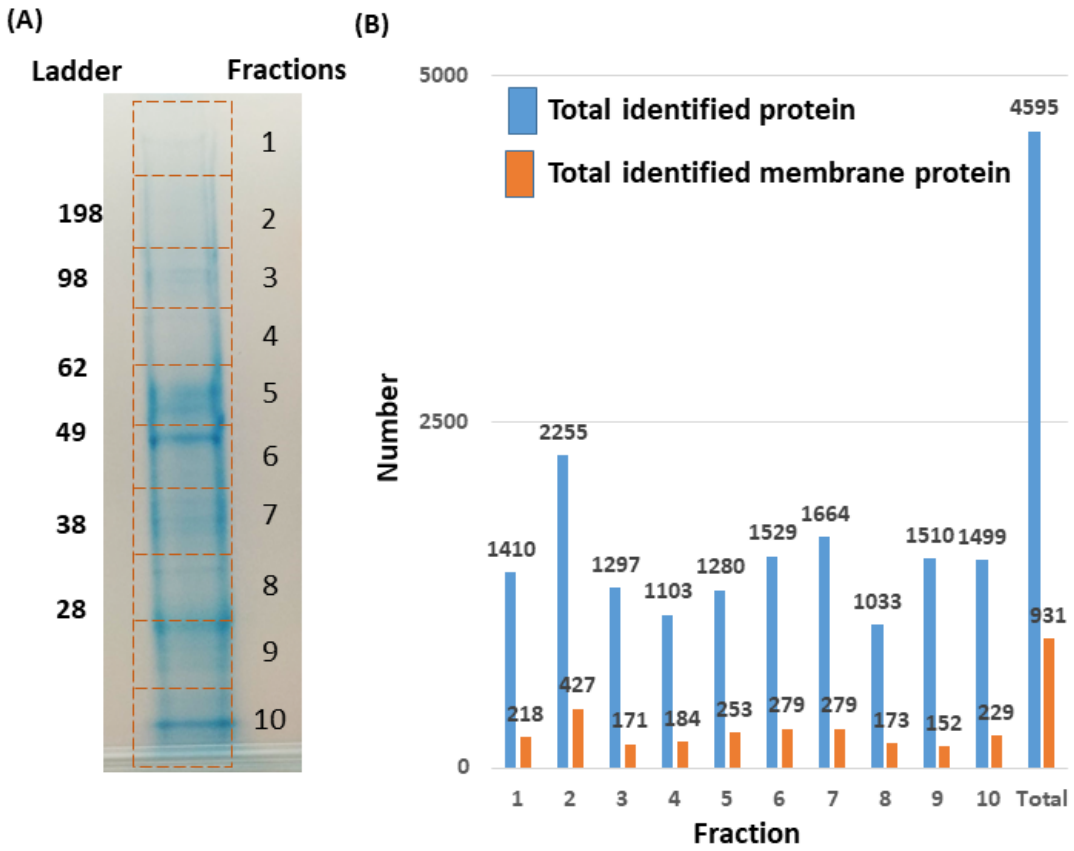


Figure 4.5. SDS-PAGE separation of the Medicago vasculature protein The Medicago VT protein was purified, size separated by SDS-PAGE, and identified by MS (see section 2.23). **(A)** The isolated Medicago VT proteins were separated by SDS-PAGE and cut into 10 pieces according to the MW. **(B)** The identified protein numbers in each fraction

To assess the effectiveness of the Medicago VT isolation, we accessed the vasculature specific gene database (Melnyk et al., 2018) and compared the protein identifications to the gene expression profile identified in this study. In brief, we used the VT related gene lists to find the homologues in Medicago proteins based on the Uniprot database (Fig. 4.6A). There are 939 genes in Arabidopsis which are VT related (Appendix Table 4.1) (Melnyk et al., 2018). Among the 939 genes, there are 596 genes which had homologues and can be found protein data in Medicago Uniprot database (Appendix Table 4.2). We compared the proteins purified from Medicago shoot tissue (Appendix Table 4.3) to the proteins purified from Medicago VT (Appendix Table 3.6) to determine whether vasculature related proteins were enriched. The result (Fig. 4.6) 2.32% (of the 4950 proteins) were VT related (Fig. 4.6B). This compared favorably to

the results in Chapter 3 where of 15656 Medicago leaf proteins identified only 1.84% were VT related. Therefore, the VT purification procedure developed successfully enriched proteins related to VT, especially the proteins related to phloem, stele, and vascular development (Fig. 4.6C). The significance of difference of VT related protein enrichment between leaf tissue and VT was confirmed by a binomial test ($P<0.01$)

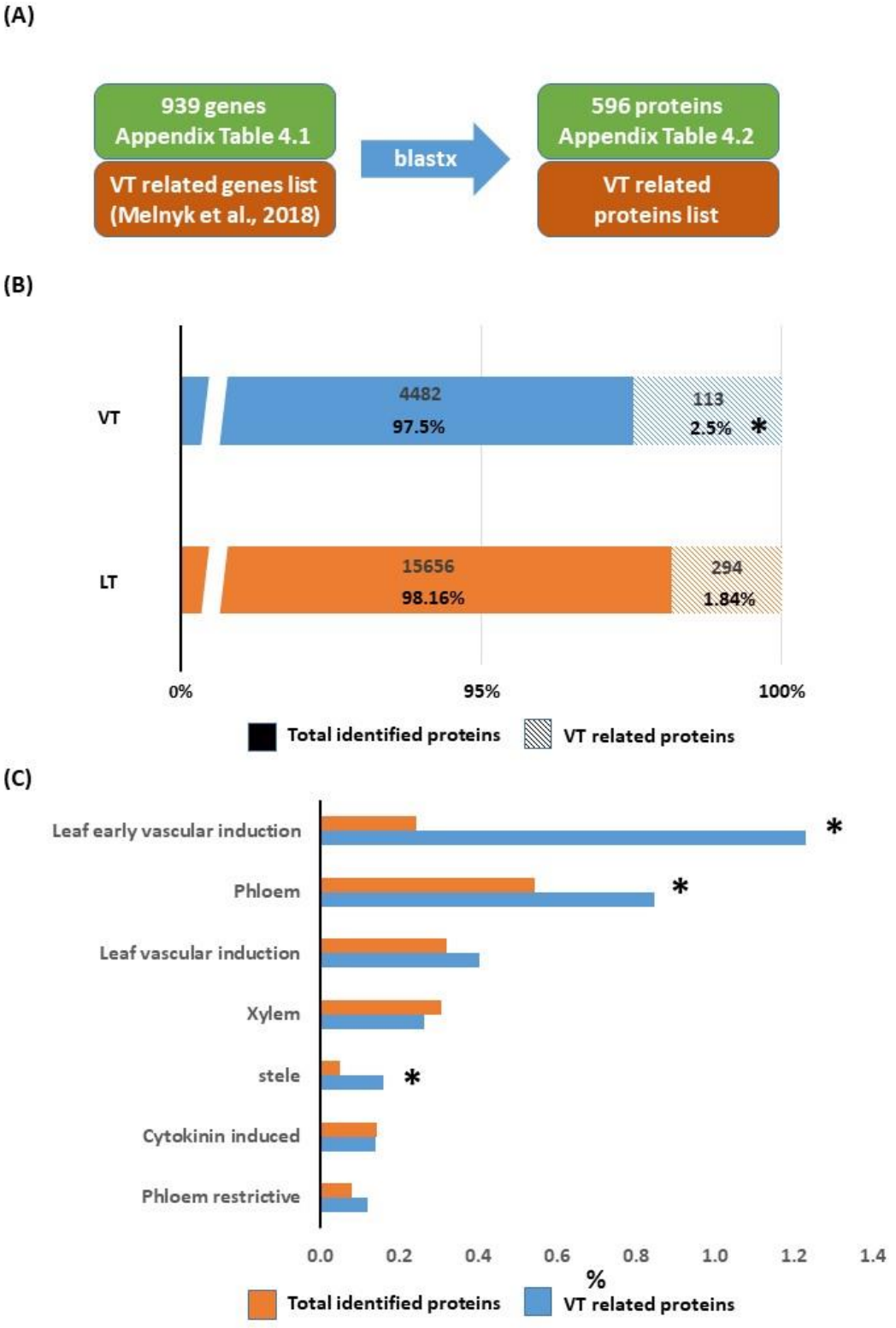


Fig 4.6. Comparison between proteins identified from Medicago Leaf tissue (LT) and vasculature tissue (VT) based on published literature (Melnyk et. al., 2018). (A) Flow chart of generating Medicago VT related proteins list (B) The VT related proteins were significantly enriched after VT

isolation (C) The VT isolation significantly enriched VT related proteins such as phloem, stele, leaf vascular induction. Significant difference was exanimated by Binomial test, $P < 0.01$

4.2.4 Formaldehyde crosslinking of FITC-MtCEP1 to A17 leaf VT is MtCRA2 dependent

To investigate the suitable chemical crosslinking reagents, we assessed the ability of two crosslinking reagents (formaldehyde and DMP) to specifically crosslink FITC-MtCEP1 to Medicago leaf VT (summarized in Fig. 4.7).

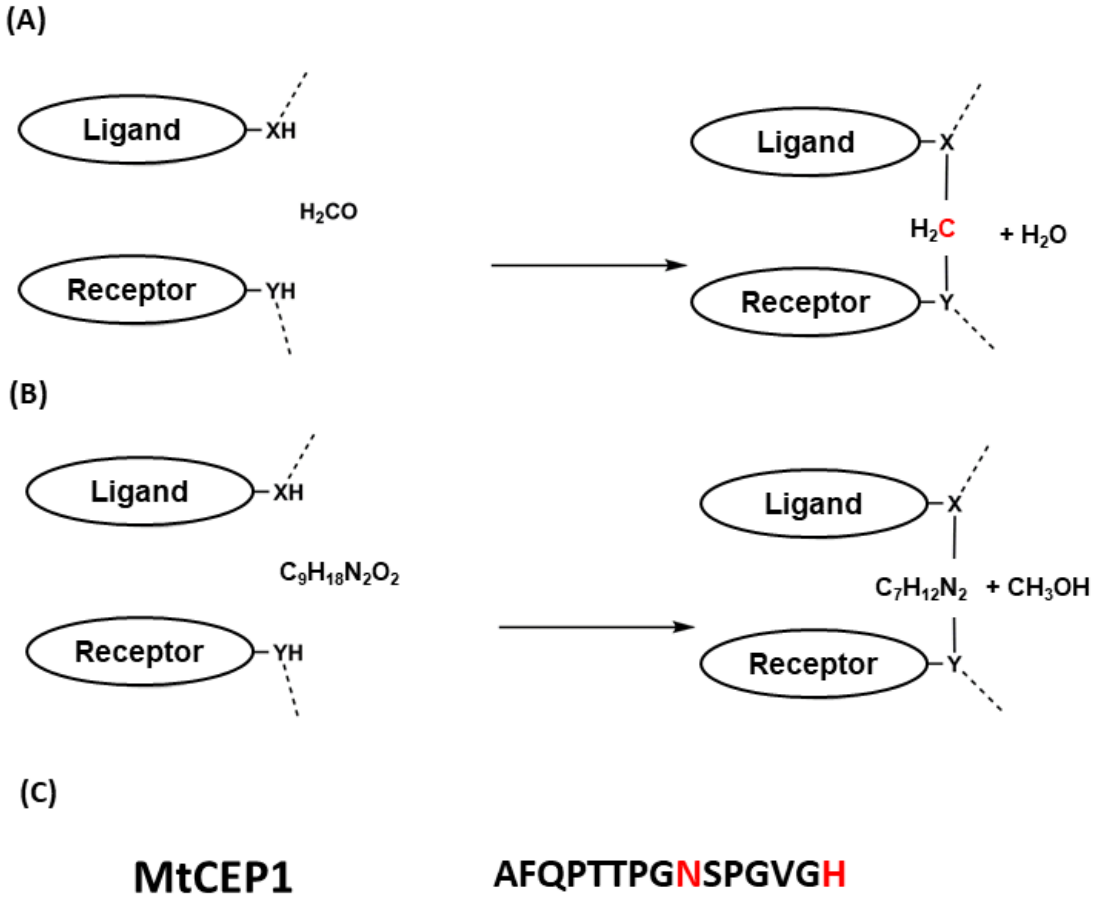


Figure 4.7. An illustration of the potential crosslinking of MtCEP1 peptides to its receptor using formaldehyde or DMP General (A) Formaldehyde (B) DMP assisted crosslinking reaction; X, Y =S or N of Cys, Lys, Trp, Asn, His, or Arg side chain (C) Possible crosslinking sites on MtCEP1 for formaldehyde and DMP were labelled in red. The space arm length is 2.3–2.7 Å for formaldehyde

and 9.2 Å for DMP

During formaldehyde-dependent crosslinking, formaldehyde first reacts with a potential nucleophile on either the ligand or the receptor to form methylol adducts. Suitable nucleophiles include amino acids such as Lys, Cys, Trp, Asn, His, or Arg. The reaction of formaldehyde with an appropriate amino acid results in the loss of water during the formation of a Schiffs base. Second, the Schiff base reacts with a second nucleophile on a closely associating molecule, in this case the receptor or also the ligand, to form covalent crosslinking (Hoffman et al., 2015). Formaldehyde crosslinking reactions are reversible at the temperature above 37°C, and the crosslinking is stable at 4 °C (Kennedy-Darling and Smith, 2014). By contrast to formaldehyde, DMP contains an imidoester at each end of a 7-carbon spacer arm and forms an amidine bond with the primary amines on the amino groups such as Lys, Cys, Trp, Asn, His, or Arg. The cross-linking reaction is more efficient when performed at pH >8 (DeCaprio and Kohl, 2019). The major differences between formaldehyde and DMP are the arm length of the cross-linkers and pH dependency of DMP. The space arm length is 2.3–2.7 Å for formaldehyde (Klockenbusch and Kast, 2010) and 9.2 Å for DMP (McKee et al., 2000). DMP mediated crosslinking reaction is pH dependent and maximal at pH>8, whereas formaldehyde mediated crosslinking reaction is pH independent. This later point is important as CEP-CEP receptor interactions might be expected to occur at pH 5.5, which is the pH of the extracellular space of plants and where the extracellular ligand, MtCEP1, interacts with the ecto-domain of the potential receptor, CRA2, which faces the extracellular space. As a control for these experiments, we used the CRA2 mutant strain, cra2-11, which would be expected to not bind MtCEP1 as this strain is a transposon-induced CEP-receptor null mutant.

We incubated FITC-MtCEP1 with WT or *cra2-11* VT and cross-linked using formaldehyde or DMP. The results (Figs. 4.6&4.7) show that both crosslinking reagents can bind FITC-MtCEP1 to WT VT (Fig. 7A; Fig. 8A), but not the VT from *cra2-11* (Fig. 4.8B, Fig. 4.9C). Therefore, result implies that the binding of FITC-MtCEP1 to Medicago leaf VT is CRA2-dependent, and the DMP mediated crosslinking is both pH and CRA2-dependent (Fig. 4.9B). Since DMP mediated

crosslinking is pH-dependent and formaldehyde crosslinking is independent of pH, we used formaldehyde as the preferred crosslinking reagent in subsequent experiments. In addition to the green fluorescence expected of a FITC-MtCEP1 interaction with its potential receptor, we also observed a purple signal in *cra2-11* vascular tissue (Fig. 4.8B). This purple auto-florescence also occurs in WT VT not cross-linked to FITC-MtCEP1 (Fig. 4.8B) and we conclude that this is due to the weak cell wall auto-fluorescence of the xylem tissue (Willemse, 1989; Donaldson and Williams, 2018), which is known to be infused with phenolic materials which auto-fluoresce. This cell wall auto-fluorescence served as a useful internal control in all samples as well as a tissue-specific marker. In Fig. 4.9A, the cell wall auto-fluorescence was masked by the greater intensity of the FITC-MtCEP1 signal.

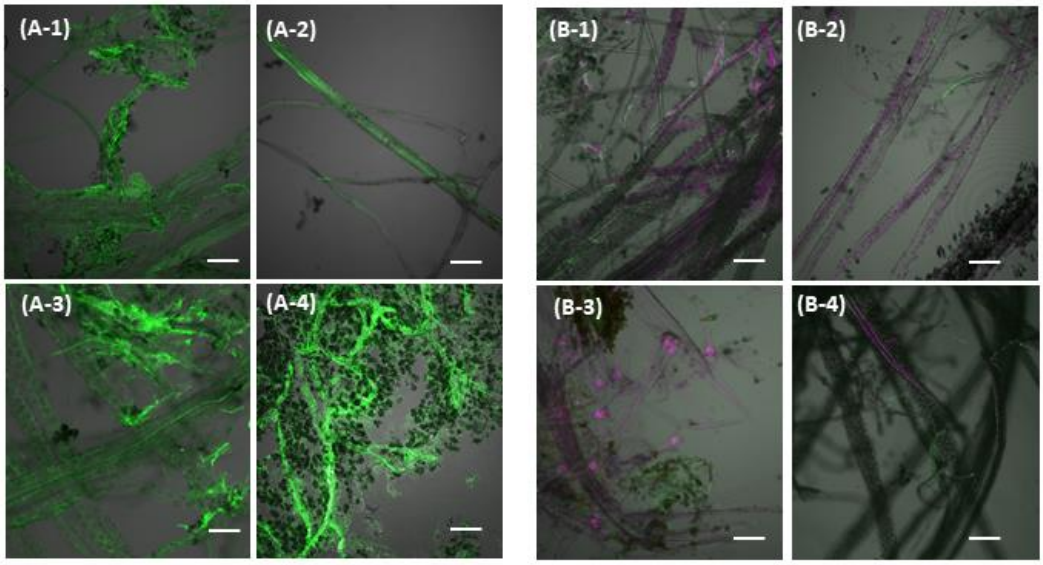


Figure 4.8. Formaldehyde crosslinking of FITC-MtCEP1 to A17 leaf vascular tissue is MtCRA2 dependent. The isolated Medicago leaf VT was incubated with FITC-MtCEP1 and crosslinked by formaldehyde before subjected to confocal microscopy (see section 2.18 & 2.22). **(A, B)** Confocal microscopy images of Medicago vasculature tissue of **(A)** A17 and **(B)** *cra2-11*. Vasculature tissue was incubated with FITC-MtCEP1 for 30 min before formaldehyde crosslinking and the subsequent washing step. FITC fluorescence is green and auto–fluorescence due to cell wall related phenolics is purple. Representative pictures from 3 independent experiments are shown. The VT of A17 + *cra2-11* were prepared and imaged using identical confocal

microscopy setting. Scale bar = 100 µm

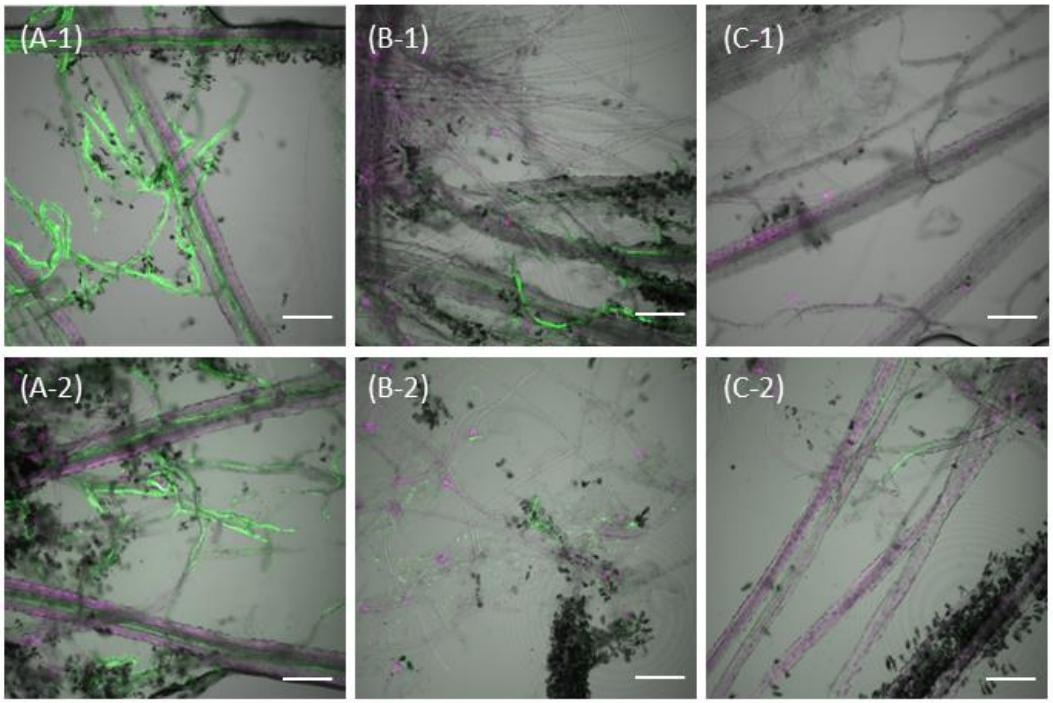


Figure 4.9. Dimethyl pimelimidate crosslinking of FITC-MtCEP1 to A17 leaf vascular tissue is pH dependent. The isolated Medicago leaf VT was incubated with FITC-MtCEP1 and crosslinked by DMP before subjected to confocal microscopy (see section 2.18 & 2.22). **(A,B)** Confocal microscopy images of Medicago A17 vasculature tissue at **(A)** pH=9, which is a preferable pH condition for DMP crosslinking, and **(B)** pH=7. **(C)** Images of Medicago *cra2-11* vasculature tissue at pH=9. Vasculature tissue was incubated with FITC-MtCEP1 for 30 min before DMP crosslinking and the subsequent washing step. FITC fluorescence is green and auto–fluorescence due to cell wall related phenolics is purple. Representative pictures from 2 independent experiments are shown. The VT of A17 + *cra2-11* were prepared and imaged using identical confocal microscopy setting. Scale bar = 100 µm

4.2.5 MtCEP1 can out compete FITC-MtCEP1 for binding sites on WT VT

If the binding of FITC-MtCEP1 to CRA2 is specific, then it should be outcompeted by unlabeled MtCEP1. To determine this, the A17 VT was used in a binding competition experiment, which also enabled us to estimate the relative binding efficacy of MtCEP1 compared to FITC-MtCEP1 (Fig 4.10).

FITC-MtCEP1 bound to WT VT as expected (Fig. 4.10A) however co-incubation of 1 μM MtCEP1 with 1 μM FITC-MtCEP1 resulted in no detectable binding of FITC-MtCEP1 (Fig. 4.10B). This result suggested the MtCEP1 and FITC-MtCEP1 bound to the same receptor and completely displaced FITC-MtCEP1 from its binding site.

We further tested the relative binding efficacy of MtCEP1 compared to FITC-MtCEP1 by co-incubating a fixed amount of FITC-MtCEP1 (1 μM) with MtCEP1 concentrations ranging from 1 μM down to 1nM. Detectable binding of FITC-MtCEP1 started to be restored when the MtCEP1 concentration was reduced to 10 nM. Therefore, these results suggest that MtCEP1 is at least 10-100 times more efficient at binding than FITC-MtCEP1 (Fig. 4.10C-E) and that MtCEP1 peptide can bind to WT VT at 10 nM.

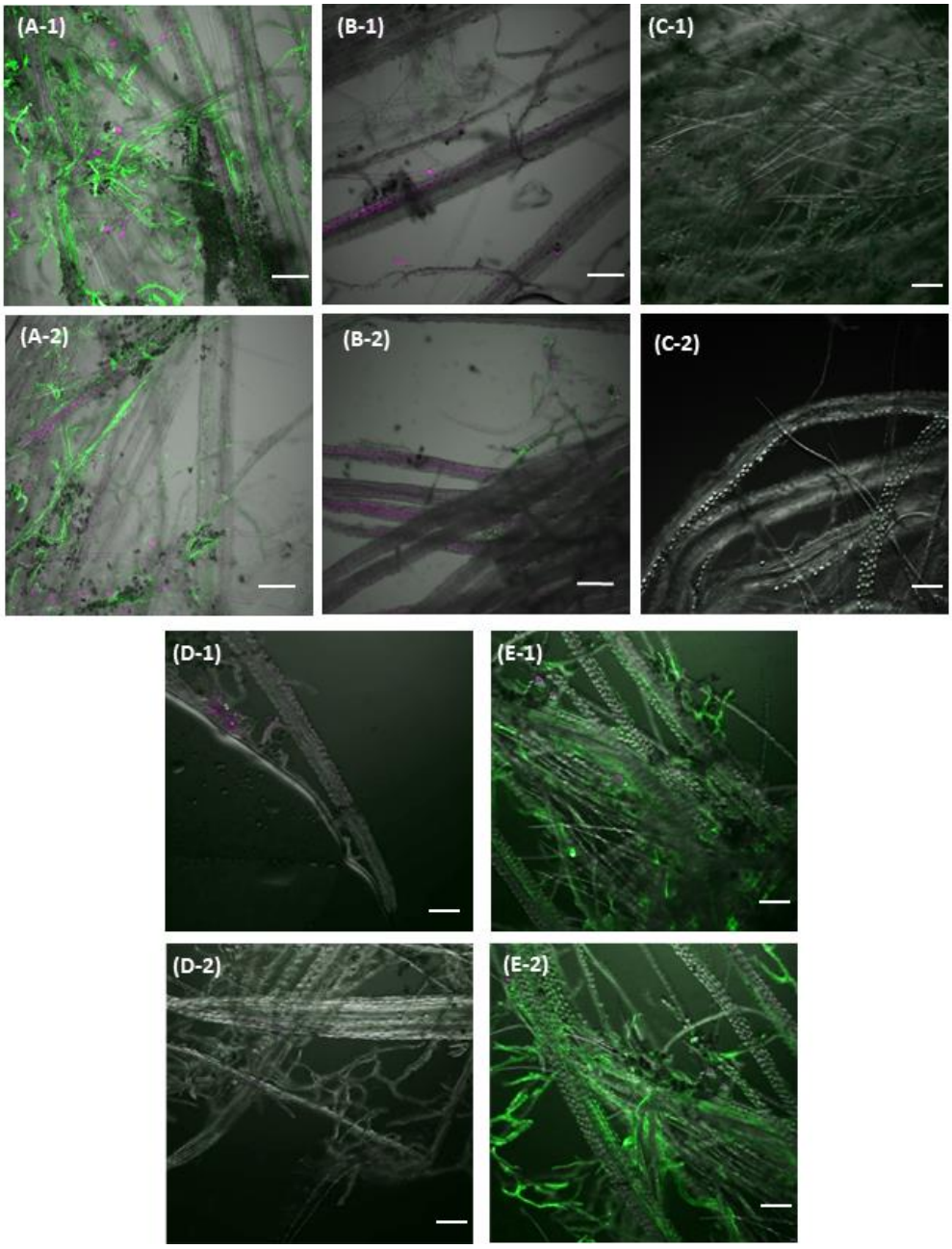


Figure 4.10. Binding site competition assays between FITC- MtCEP1 and MtCEP1. The isolated Medicago leaf VT was co-incubated with 1 µM FITC-MtCEP1 and different concentrations of MtCEP1 and crosslinked by formaldehyde before subjected to confocal microscopy (see section 2.21 & 2.22). Confocal microscopy pictures of WT leaf VT incubated with: **(A)** 1 µM of FITC- MtCEP1; **(B)(C)**1 µM of FITC-MtCEP1 and 1 µM of MtCEP1; **(D)**1 µM of FITC- MtCEP1 and 0.1 µM of MtCEP1; **(E)** 1 µM of FITC- MtCEP1 and 0.01 µM of MtCEP1. FITC fluorescence is green, and auto–fluorescence is purple. The pictures are representative of 3 independent experiments. Scale bar = 100µm

4.2.6 Defining the minimum concentration of formaldehyde required for specific crosslinking of FITC-MtCEP1 to A17 *M. truncatula* VT

To further ensure that the formaldehyde mediated crosslinking was specific we determined the minimum concentration of formaldehyde required for FITC-MtCEP1 crosslinking to semi-purified Medicago VT (Fig. 4.11). This was done because the published concentrations of formaldehyde used for interactome studies (Thavarajah et al., 2012) recommend using 4 % formaldehyde to permeate into cells, whereas the crosslinking reaction in these studies occurs on the cell surface. Since it was unnecessary for formaldehyde to penetrate the VT cells, we expected that the minimum formaldehyde concentration required for MtCEP1-CRA2 crosslinking to be lower than 4%. Figure 10A shows the confocal images of peptide treated VT after crosslinking with a range of formaldehyde concentrations using 4% as a baseline. There is clear FITC fluorescence in the 4% positive control, indicating the FITC-labelled peptide has been successfully cross-linked to proteins in the vasculature. The fluorescence level in samples cross-linked with the $2x10^{-3}$% solution is comparable to the control (Fig. 4.11D). There is a noticeable drop in fluorescence at lower concentrations. Crosslinking with $2x10^{-4}$% (Fig. 4.11E) significantly reduces the intensity of fluorescence but FITC is present in throughout the vascular tissue samples as a green vein. In the samples exposed to the $2x10^{-5}$% solution, much less FITC is evident fluorescence cannot be seen throughout the VT, instead it is localized to a few intense clusters (Fig. 4.11F). There was little to no fluorescence observable after incubation with $2x10^{-6}$% formaldehyde, suggesting this was not enough to induce covalent bond formation (Fig. 4.11G). Therefore, these studies confirm that formaldehyde can crosslink FITC-MtCEP1 to CRA2 at 1000-fold lower concentrations that are required for cellular interactome studies and that the incubation time can be reduced from 30 min to 5 min.

An interesting outcome from the titration experiment is that as the FITC fluorescence intensity is reduced with decreasing formaldehyde, a more

specific MtCEP1 localization within the VT is revealed (Fig. 4.11 H, I). Since the fluorescence of the xylem vessels emitted a strong purple auto-fluorescence signal, which caused by the phenolics in their cell walls, the green vein which caused by FITC labelling is most likely be the phloem tissue, which corresponds to the location where MtCRA2 expresses in prior studies (Tabata et al., 2014; Ohkubo et al., 2017) Taleski et al 2020)

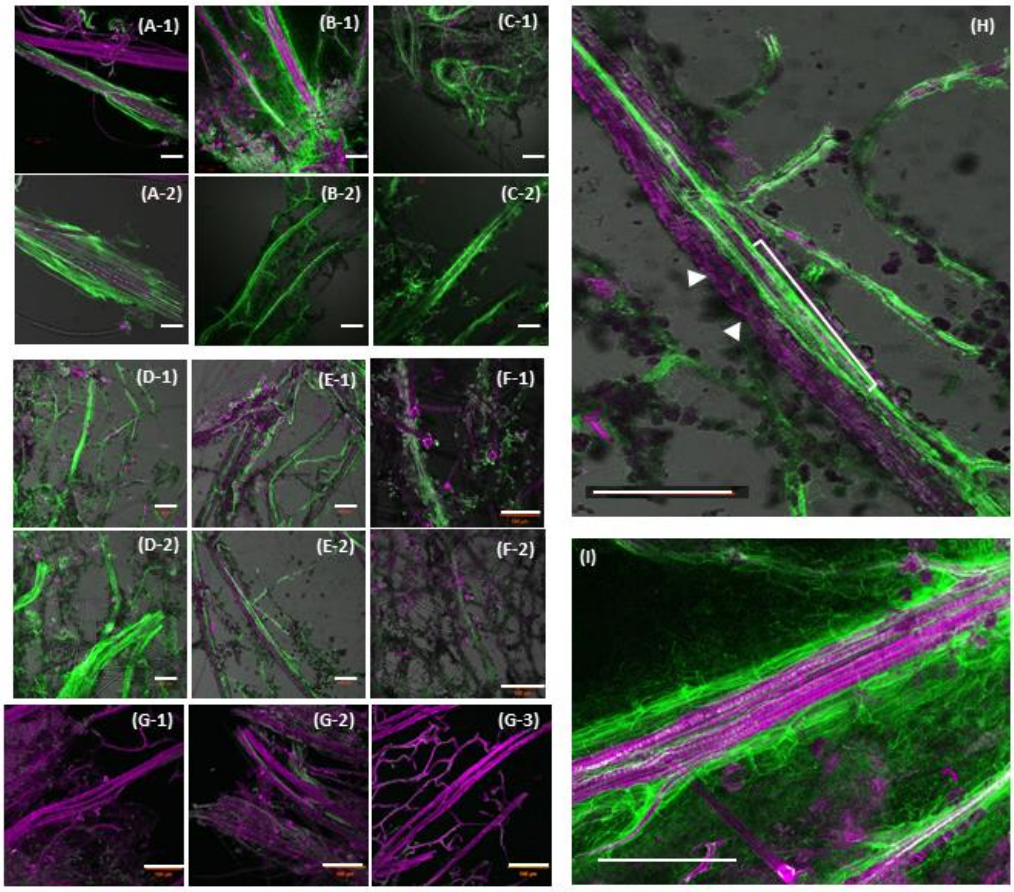


Figure 4.11. Defining the minimum concentration of formaldehyde required for specific crosslinking of FITC-MtCEP1 to A17 *M. truncatula* VT. The isolated Medicago leaf VT was incubated with FITC-MtCEP1 and crosslinked by different concentration of formaldehyde before subjected to confocal microscopy (see section 2.19 & 2.22). The samples in the images were crosslinked for 5 minutes with the following formaldehyde concentrations: **(A)** 4%, **(B)** 0.2%, **(C)** 0.02%, **(D)** 0.002%, **(E)** $2x10^{-4}$%, **(F)** $2x10^{-5}$%, **(G)** $2x10^{-6}$%. The FITC fluorescence is green whereas cell wall auto-fluorescence is purple. The scale bars are 100 µm. **(H-I)** Magnifications of image **(E)** lower panel and **(B)** upper panel. **(H)**The fluorescence due to MtCEP1 binding is discretely localized within the vascular tissue (as see white bracket) whereas white arrows

indicate regions of purple auto-fluorescence that likely correspond to xylem vessels.

4.2.7 Cross-species ligand receptor identification revealed by formaldehyde cross-linking

We investigated if the Medicago CEP1 derivative can interact specifically with the Arabidopsis CEPR1 receptor. First, we tested the biological activity of FITC-MtCEP1 in Arabidopsis to see if it can bind to the AtCEPR1, which is the putative functional orthologue of MtCRA2 (Fig. 4.12A; Mohd-Radzman et al 2016). Second, we incubated the FITC-MTCEP1 with the purified Arabidopsis leaf VT to confirm the ligand receptor interaction *in vivo*. The results (Fig. 4.12) show that FITC-MtCEP1 can bind to Arabidopsis WT VT (Fig. 4.13A), but not the VT from cepr1-1(Fig. 4.13B). Therefore, the binding of FITC-MtCEP1 to Arabidopsis leaf VT is AtCEPR1-dependent. The results also reaffirm that MtCRA2 is functionally orthologous to AtCEP1.

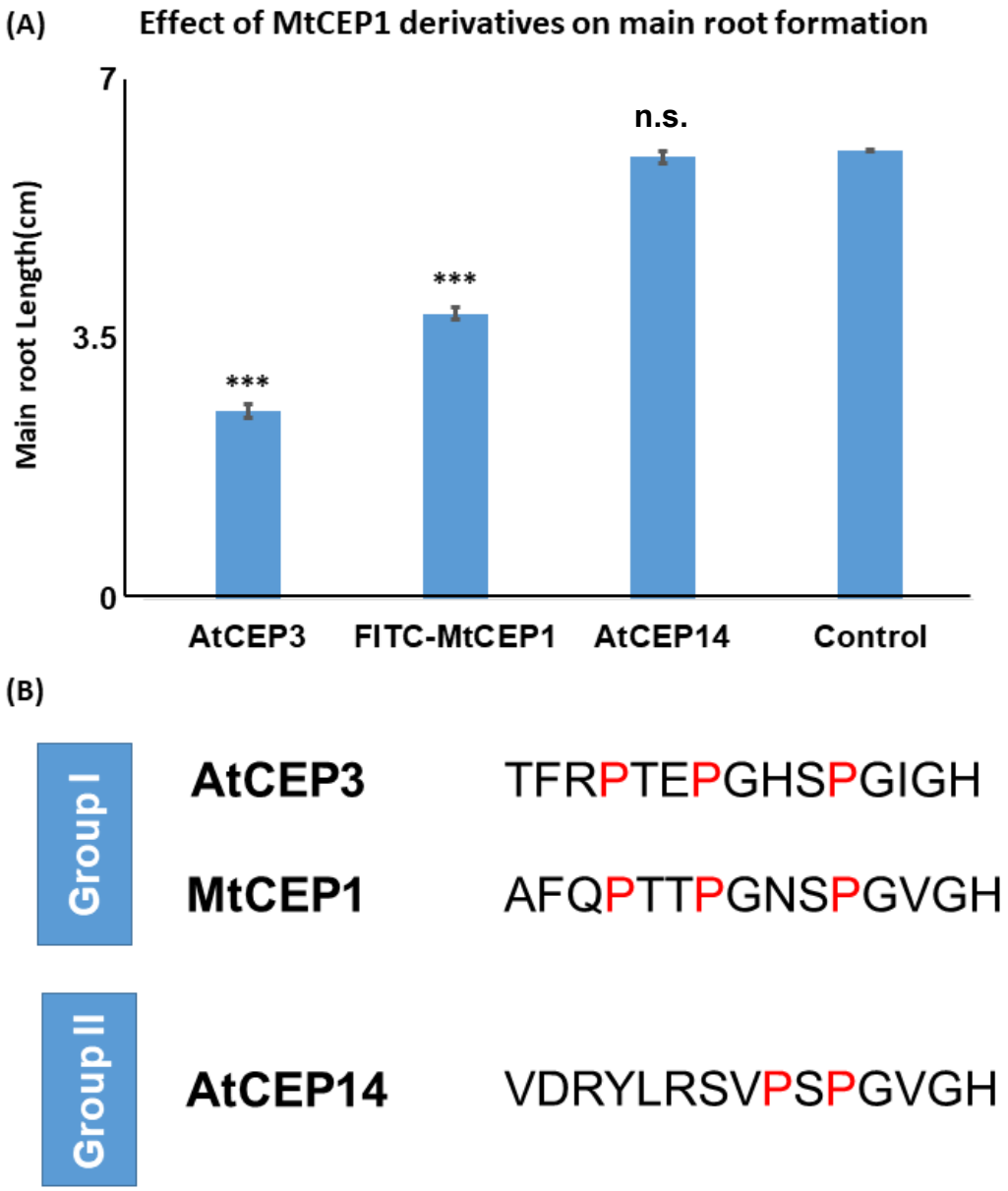


Figure 4.12. The MtCEP1 derivative with a FITC tag is biologically active in Arabidopsis. The synthetic peptide was added into the medium for Arabidopsis seeding growth. The Arabidopsis seedings were tested the main root for peptide biological activity (see section 2.15) **(A)** The effect of FITC-MtCEP1 (1 x 10^{-6}M) derivative on main root. Main root formation was scored at 10 dpi. Statistically significant differences were determined using a Student's *t*-test: ***, $P \leq 0.001$ (error bars, ± standard error, $n \geq 7$). **(B)** The sequence aliment of group I CEP, which were AtCEP3 and MtCEP1, and group II CEP, which was AtCEP14.

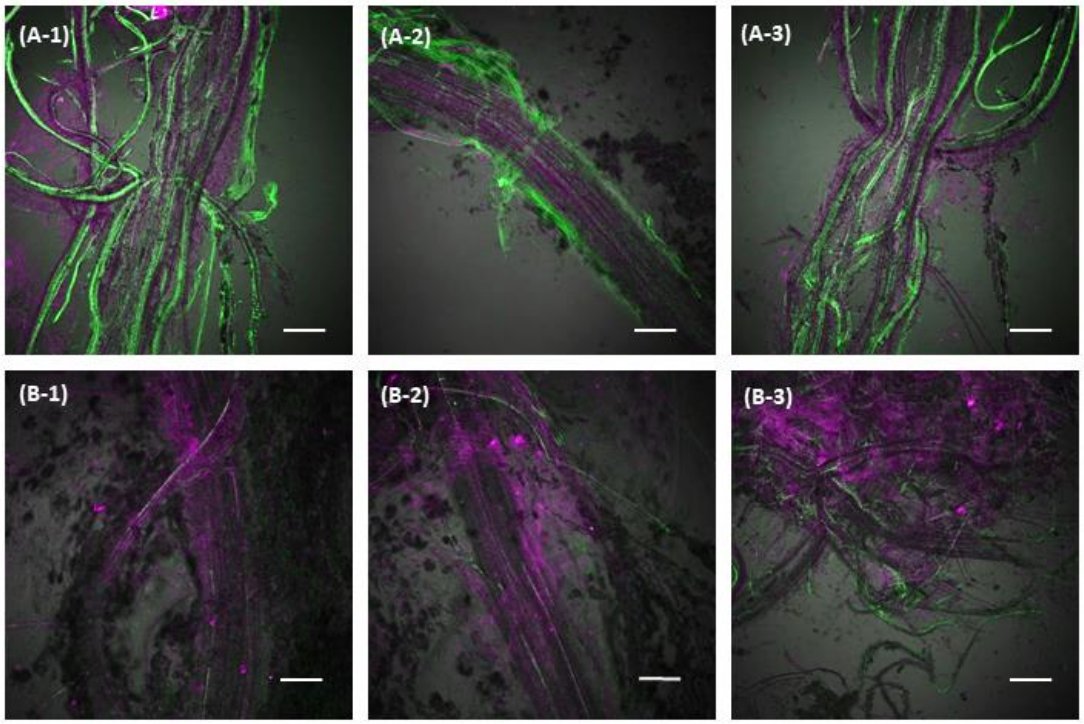


Figure 4.13. Formaldehyde crosslinking of FITC-MtCEP1 to WT Arabidopsis leaf vascular tissue is AtCEPR1 dependent. The isolated Arabidopsis leaf VT was incubated with FITC-MtCEP1 and crosslinked by formaldehyde before subjected to confocal microscopy (see section 2.18 & 2.22). **(A,B)** Confocal microscopy images of mechanically-isolated *Arabidopsis* vasculature tissue of **(A)** WT and **(B)** *cepr1-1*. Vasculature tissue was incubated with FITC-MtCEP1 for 30 min before formaldehyde crosslinking and the subsequent washing step. FITC fluorescence is green and auto–fluorescence due to cell wall related phenolics is purple. Representative pictures from 3 independent experiments are shown. Scale bar = 100μm

To examine the specificity of CEP species to CEPR1/CRA2 we made an FITC-AtCEP14 peptide and assessed its ability to bind to the VT of WT and a CEP receptor mutant. AtCEP14 is a so-called group 2 CEP based on rather weak homology of its CEP domain to the group 1 CEPs (Delay et al., 2013; Roberts et al., 2013). Whilst the C terminal residues of the group 2 CEPs is conserved with the group 1 CEPs, there is considerable divergence in the N-terminal amino acids of group 1 and group 2 CEP family peptides. In addition, AtCEP14 contains two prolines at position 9 and 11 only one of which is conserved in position with the group I CEP family. By contrast most group 1 CEPs contain

conserved prolines at positions 4, 7 and 11, which have been shown to be required for CEP peptide activity (Mohd-Radzman et al., 2015). The biological activity of group 2 CEPs is untested.

Therefore, we investigated if AtCEP14 is a functional ligand for MtCRA2. (Fig. 4.12B). First, we noted that AtCEP14 shares no obvious phenotypic effect on plants in Arabidopsis and Medicago (Fig. 4.12A), in contrast to the group 1 CEPs. We evaluated a possible interaction between AtCEP14 and MtCRA2 using formaldehyde mediated crosslinking and FITC labelled AtCEP14. The results (Fig. 4.13) show that FITC-AtCEP14 bound to both Medicago WT VT (Fig. 4.14A), and the VT from *cra2-11* (Fig. 4.14B). Therefore, the binding of FITC-AtCEP14 to Medicago leaf VT appears not to be MtCRA2-related. The results suggested the receptor of AtCEP14 is not MtCRA2, and AtCEP14 and other "class 2 CEPs" should be excluded from CEP family as they bind to (an)other vascular receptor(s).

204

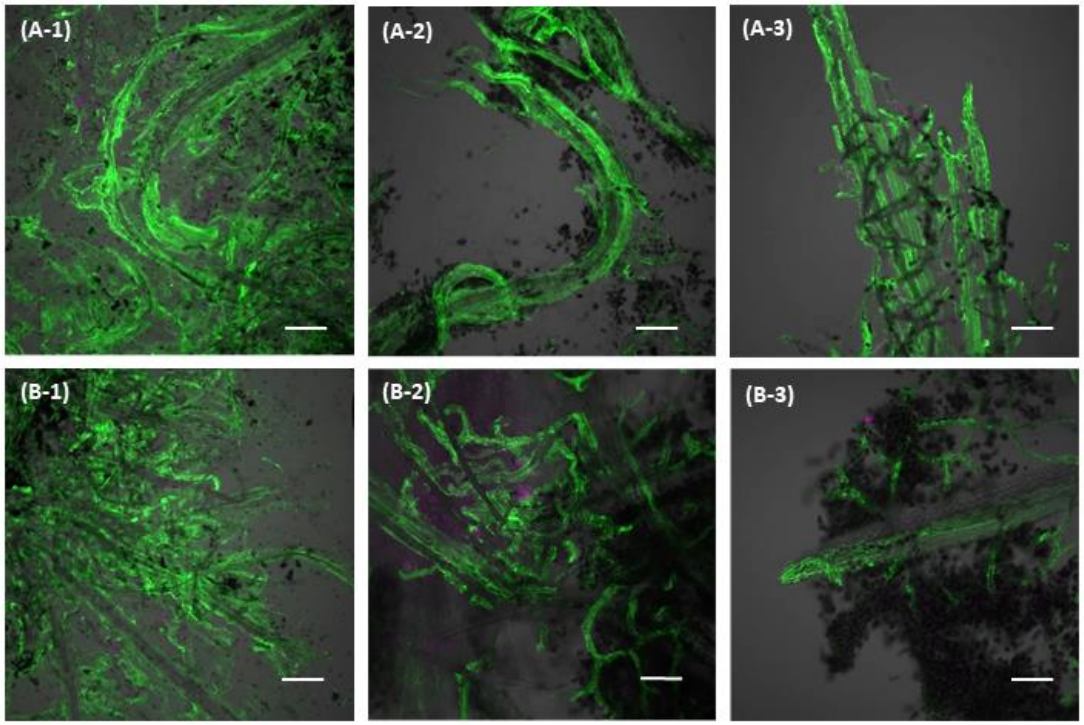


Figure 4.14. FITC-AtCEP14 binds to both A17 and CRA2 MT leaf VT The isolated Medicago leaf VT was incubated with FITC-AtCEP14 and crosslinked by formaldehyde before subjected to confocal microscopy (see section 2.18 & 2.22). **(A, B)** Confocal microscopy images of Medicago vasculature tissue of **(A)** A17 and **(B)** *cra2-11*. Vasculature tissue was incubated with FITC-AtCEP14

for 30 min before formaldehyde crosslinking and the subsequent washing step. FITC fluorescence is green and auto–fluorescence due to cell wall related phenolics is purple. Representative pictures from 3 independent experiments are shown. The VT of A17 + *cra2-11* were prepared and imaged using identical confocal microscopy setting. Scale bar = 100 μm

4.3 Conclusion

In this chapter, we established a methodology to identify ligand receptor interaction *in vivo* by using formaldehyde mediated crosslinking and purified Medicago VT.

First, since the MtCRA2 was mainly expressed in VT, we isolated VT from Medicago leaf as the starting material and verified the isolated Medicago VT retained viability and was enriched for VT related proteins. The viability of the isolated Medicago VT was confirmed by propidium iodide staining. Proteomic analysis indicated that the proteins purified from Medicago VT were enriched with proteins related to the VT, especially the proteins linked to phloem, stele, and vascular development.

Second, we accessed the capability of in vivo ligand receptor interaction using semi-purified Medicago VT and chemical crosslinking. Specifically, we examined formaldehyde and DMP as crosslinking reagents for *in vivo* crosslinking of class 1 CEPs to isolated Medicago VT. Although both reagents can cross-link MtCEP1 to VT in a receptor dependent manner, DMP failed to stabilize the FITC-MtCEP1 to WT Medicago VT at pH<8 condition. Thus, we choose formaldehyde as the preferred crosslinking reagent since the formaldehyde crosslinking was independent of pH.

Third, using formaldehyde-assisted cross-linking, we validated the binding of FITC-MtCEP1 to Medicago leaf VT was MtCRA2 dependent, and this binding can be out competed by unlabeled MtCEP1. This result suggests that MtCRA2 specifically binds FITC-MtCEP1, and MtCEP1 binds CRA2 at least 10 times

more efficiently.

Fourth, to ensure that the formaldehyde mediated crosslinking was specific, we determined minimum concentration of formaldehyde required to cross link MtCEP1 to the ecto domain of CRA2. The crosslinking can be achieved with the formaldehyde concentration as low as $2x\ 10^{-3}$%. As a result of lowering the formaldehyde concentration confocal examination indicated a more tissue-specific binding of FITC-MtCEP1 in the VT. It is clear that FITC-MtCEP1 does not co-localize with the tissues containing the purple auto fluorescence.

Fifth, we expanded the methodology of formaldehyde mediated ligand receptor identification to cross-species ligand receptor identification and examined the interaction between ligand and receptor. The results showed FITC-MtCEP1 bound to the Arabidopsis leaf VT in AtCEPR1 dependent manner, which confirmed the methodology can be applied to identify ligand receptor interaction in different species. This result is consistent with MtCRA2 being a functional orthologue of AtCEPR1.

Sixth, the approach devised in this chapter showed that AtCEP14, which is one of group II CEP peptide, bounds to the Medicago VT independently of MtCRA2. This result suggests that group 2 CEPs may bind to (a)nother receptor(s) in Medicago VT.

Since the formaldehyde is a non-specific crosslinking reagent, it would be difficult to isolate the receptor complex and identify it by MS identification. One possibility might be to extract the cross-linked proteins use blue native gels and use the FITC fluorescence to identify a receptor complex in these gels that could be examined by MS. However, since formaldehyde mediated crosslinking is a random and non-specific crosslinking, it basically crosslinks every protein which is in the suitable range. This makes the MtCEP1-*CRA2* complex hard to identify from the crosslinked protein complex. To address this issue the next chapter uses site-directed photo-activation to cross-link and purify the ligand-receptor complex for MS identification.

Literature cited

Casadonte R, Caprioli RM (2011) Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry. Nature protocols **6:** 1695-1709

DeCaprio J, Kohl TO (2019) Cross-Linking Antibodies to Beads Using Dimethyl Pimelimidate (DMP). Cold Spring Harb Protoc **2019**

Delay C, Imin N, Djordjevic MA (2013) CEP genes regulate root and shoot development in response to environmental cues and are specific to seed plants. J Exp Bot **64:** 5383-5394

Donaldson L, Williams N (2018) Imaging and Spectroscopy of Natural Fluorophores in Pine Needles. Plants (Basel, Switzerland) **7:** 10

Hirakawa Y, Shinohara H, Kondo Y, Inoue A, Nakanomyo I, Ogawa M, Sawa S, Ohashi-Ito K, Matsubayashi Y, Fukuda H (2008) Non-cell-autonomous control of vascular stem cell fate by a CLE peptide/receptor system. Proc Natl Acad Sci U S A **105:** 15208-15213

Hoffman EA, Frey BL, Smith LM, Auble DT (2015) Formaldehyde Crosslinking: A Tool for the Study of Chromatin Complexes. Journal of Biological Chemistry **290:** 26404-26411

Huault E, Laffont C, Wen J, Mysore KS, Ratet P, Duc G, Frugier F (2014) Local and systemic regulation of plant root system architecture and symbiotic nodulation by a receptor-like kinase. PLoS Genet **10:** e1004891

Imin N, Mohd-Radzman NA, Ogilvie HA, Djordjevic MA (2013) The peptide-encoding CEP1 gene modulates lateral root and nodule numbers in *Medicago truncatula*. J Exp Bot **64:** 5395-5409

Kennedy-Darling J, Smith LM (2014) Measuring the formaldehyde Protein-DNA cross-link reversal rate. Anal Chem **86:** 5678-5681

Klockenbusch C, Kast J (2010) Optimization of formaldehyde cross-linking for protein interaction analysis of non-tagged integrin beta1. J Biomed Biotechnol **2010:** 927585

Laffont C, Huault E, Gautrat P, Endre G, Kalo P, Bourion V, Duc G, Frugier F (2019) Independent Regulation of Symbiotic Nodulation by the SUNN Negative and CRA2 Positive Systemic Pathways. Plant

Physiol **180:** 559-570

Lucas WJ, Groover A, Lichtenberger R, Furuta K, Yadav SR, Helariutta Y, He XQ, Fukuda H, Kang J, Brady SM, Patrick JW, Sperry J, Yoshida A, Lopez-Millan AF, Grusak MA, Kachroo P (2013) The plant vascular system: evolution, development and functions. J Integr Plant Biol **55:** 294-388

McKee JA, Kumar S, Ecelbarger CA, Fernandez-Llama P, Terris J, Knepper MA (2000) Detection of Na(+) transporter proteins in urine. J Am Soc Nephrol **11:** 2128-2132

Melnyk CW, Gabel A, Hardcastle TJ, Robinson S, Miyashima S, Grosse I, Meyerowitz EM (2018) Transcriptome dynamics at Arabidopsis graft junctions reveal an intertissue recognition mechanism that activates vascular regeneration. Proceedings of the National Academy of Sciences **115:** E2447-E2456

Metz B, Kersten GF, Hoogerhout P, Brugghe HF, Timmermans HA, de Jong A, Meiring H, ten Hove J, Hennink WE, Crommelin DJ, Jiskoot W (2004) Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. J Biol Chem **279:** 6235-6243

Mohd-Radzman NA, Binos S, Truong TT, Imin N, Mariani M, Djordjevic MA (2015) Novel MtCEP1 peptides produced in vivo differentially regulate root development in *Medicago truncatula*. Journal of Experimental Botany **66:** 5289-5300

Mohd-Radzman NA, Laffont C, Ivanovici A, Patel N, Reid D, Stougaard J, Frugier F, Imin N, Djordjevic MA (2016) Different Pathways Act Downstream of the CEP Peptide Receptor CRA2 to Regulate Lateral Root and Nodule Development. Plant Physiol **171:** 2536-2548

Ogilvie HA, Imin N, Djordjevic MA (2014) Diversification of the C-TERMINALLY ENCODED PEPTIDE (CEP) gene family in angiosperms, and evolution of plant-family specific CEP genes. BMC Genomics **15:** 870

Ohkubo Y, Tanaka M, Tabata R, Ogawa-Ohnishi M, Matsubayashi Y (2017) Shoot-to-root mobile polypeptides involved in systemic regulation of nitrogen acquisition. Nat Plants **3:** 17029

Ohyama K, Ogawa M, Matsubayashi Y (2008) Identification of a biologically active, small, secreted peptide in Arabidopsis by in silico gene screening, followed by LC-MS-based structure analysis. The Plant Journal **55:** 152-160

Okamoto S, Suzuki T, Kawaguchi M, Higashiyama T, Matsubayashi Y (2015) A comprehensive strategy for identifying long-distance mobile peptides in xylem sap. Plant J **84:** 611-620

Patel N, Mohd-Radzman NA, Corcilius L, Crossett B, Connolly A, Cordwell SJ, Ivanovici A, Taylor K, Williams J, Binos S, Mariani M, Payne RJ, Djordjevic MA (2018) Diverse Peptide Hormones Affecting Root Growth Identified in the *Medicago truncatula* Secreted Peptidome. Mol Cell Proteomics **17:** 160-174

Roberts I, Smith S, De Rybel B, Van Den Broeke J, Smet W, De Cokere S, Mispelaere M, De Smet I, Beeckman T (2013) The CEP family in land plants: evolutionary analyses, expression studies, and role in Arabidopsis shoot development. J Exp Bot **64:** 5371-5381

Shabala S, White RG, Djordjevic MA, Ruan Y-L, Mathesius U (2016) Root-to-shoot signalling: integration of diverse molecules, pathways and functions. Functional Plant Biology **43:** 87-104

Sinz A, Arlt C, Chorev D, Sharon M (2015) Chemical cross-linking and native mass spectrometry: A fruitful combination for structural biology. Protein science : a publication of the Protein Society **24:** 1193-1209

Smith A, George Coupland, Liam Dolan, Nicholas Harberd, Jonathan Jones, Cathie Martin, Robert Sablowski, Amey A (2009) Plant Biology. Garland Science

Sutherland BW, Toews J, Kast J (2008) Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. J Mass Spectrom **43:** 699-715

Tabata R, Sumida K, Yoshii T, Ohyama K, Shinohara H, Matsubayashi Y (2014) Perception of root-derived peptides by shoot LRR-RKs mediates systemic N-demand signaling. Science **346:** 343-346

Thavarajah R, Mudimbaimannar VK, Elizabeth J, Rao UK, Ranganathan K (2012) Chemical and physical basics of routine formaldehyde fixation. Journal of oral and maxillofacial pathology : JOMFP **16:** 400-

405

Toth J, Biggin MD (2000) The specificity of protein-DNA crosslinking by formaldehyde: in vitro and in drosophila embryos. Nucleic acids research **28:** e4-e4

Willemse MTM (1989) Cell Wall Autofluorescence. *In* A Chesson, ER Ørskov, eds, Physico-Chemical Characterisation of Plant Residues for Industrial and Feed Use. Springer Netherlands, Dordrecht, pp 50-57

Chapter 5: Validation of CEP peptide binding to MtCRA2

5.1 Introduction

The previous chapter provided evidence for the binding of MtCEP1 to the Medicago leaf VT in a MtCRA2 dependent manner. The results indicated that MtCEP1 and MtCRA2 may form a ligand receptor complex. However, as formaldehyde is a non-specific crosslinking reagent, the formaldehyde mediated crosslinking made the ligand receptor complex purification extremely difficult since crosslinking was not limited to MtCEP1 and MtCRA2, but also all the other proteins complexes within. Although the formaldehyde mediated FITC-MtCEP1 localization to the MtCEP1 was localized majorly on the living cells of the Medicago VT, which might be the parenchyma from xylem or phloem, the non-specific formaldehyde crosslinking still limited the resolution of FITC localization. Therefore, this chapter aimed to develop a methodology to purify the receptor which interacted with MtCEP1 by using a site-directed photo-activatable cross-linker and confirmed the ligand-receptor complex with Mass spectrometry.

5.1.1 Photo-active site directed crosslinking for PPIs identification

Another approach used to examine specific PPIs is to incorporate a photo-active amino acid on one of the interacting partners. There are several photo-active amino acids that can be used and one example is the photo-active amino acid analogue of phenylalanine, *ρ*-benzoyl-L-phenylalanine (*ρBpa*), (Kauer et al., 1986). Upon UV irradiation at a wavelength around 365 nm, the photo-generated triplet state of the benzophenone group abstracts a hydrogen atom from geometrically accessible carbon-hydrogen bond of the polypeptide backbone. The side chain methylene groups then combine with the resulting radical to generate a covalent link between the ketone carbonyl carbon and the attacked methylene carbon on an adjacent protein (Wong and King, 2015). Carbon-hydrogen bonds within 3 Å of the carbonyl oxygen are targets for cross-linking (Farrell et al., 2005). The cross-linked protein complex can be purified and identified using tandem mass spectrometry. However, few publications have reported using this technique in plant tissues to examine putative peptide ligand-receptor interactions.

5.1.2 *In vivo* CEPs receptor identification

The need for a deeper understanding of how intercellular signalling mediated by peptide hormone-receptor interactions control growth and development requires the development of rapid and simple methods for identifying or validating interactions of peptide hormones with their respective receptors *in vivo*. In Chapter 5, we utilized a technique to semi-purify VT from the leaves of Medicago and Arabidopsis to investigate if native CEP receptor complexes known to express in these tissues can bind to biologically active fluorescently tagged group 1 peptides *in vivo*. We then used confocal microscopy to assess if formaldehyde-mediated crosslinking of these peptides to wild type VT of Medicago or Arabidopsis depends on CEP receptors by exploiting the natural expected affinity of group 1 CEP peptides to putative, or known, receptors (e.g. MtCRA2 or AtCEPR1, respectively).

In this chapter, the aim is to discover the potential receptor for MtCEP1 by photo-activated crosslinking and MS identification. First, we tested the biologically activity of varies *N*-terminal extension for affinity and fluorescence tag. Second, we tested if the photo-cross-linker (*p*Bpa) modified FITC-MtCEP1 peptide (FITC-MtCEP1F*) can still maintain the biological activity. Third, we examined the interaction between FITC-MtCEP1F* and MtCRA2 using purified Medicago leaf VT, as well as the binding efficacy of FITC-MtCEP1F* to MtCRA2. Finally, we purified the proteins from Medicago Leaf VT which is cross-linked specifically to FITC-MtCEP1F* and identified the binding partner as MtCRA2 by MS detection.

5.2 Results

5.2.1 Different N-terminal extension can affect the biological activity of MtCEP1

First, in order to purified the MtCEP1-MtCRA2 complex, we sythesized several

MtCEP1 derivatives with different tag on the *N*-terminus, which were FITC tag (named FITC-MtCEP1), HA tag (named HA-MtCEP1), StrepII tag with spacer (named Strep- MtCEP1) (Fig. 5.1). We further compared the biological activities of these MtCEP1 variants to ensure the synthetic peptide still bound to its receptor. The results (Fig. 5.2) showed that the MtCEP1 N-terminal extension variants were biologically activie to different level. All of the MtCEP1 variants can inhibit the laterial roots formation, but only FITC-MtCEP1 can increase the nodule number similarly to MtCEP1. The HA-MtCEP1 and Strep MtCEP1 displayed abolished activity on the enhanced nudulation. Thus, we used FITC-MtCEP1 as the N-teminal tagged ligand for ligand-receptor complex purificaiton.

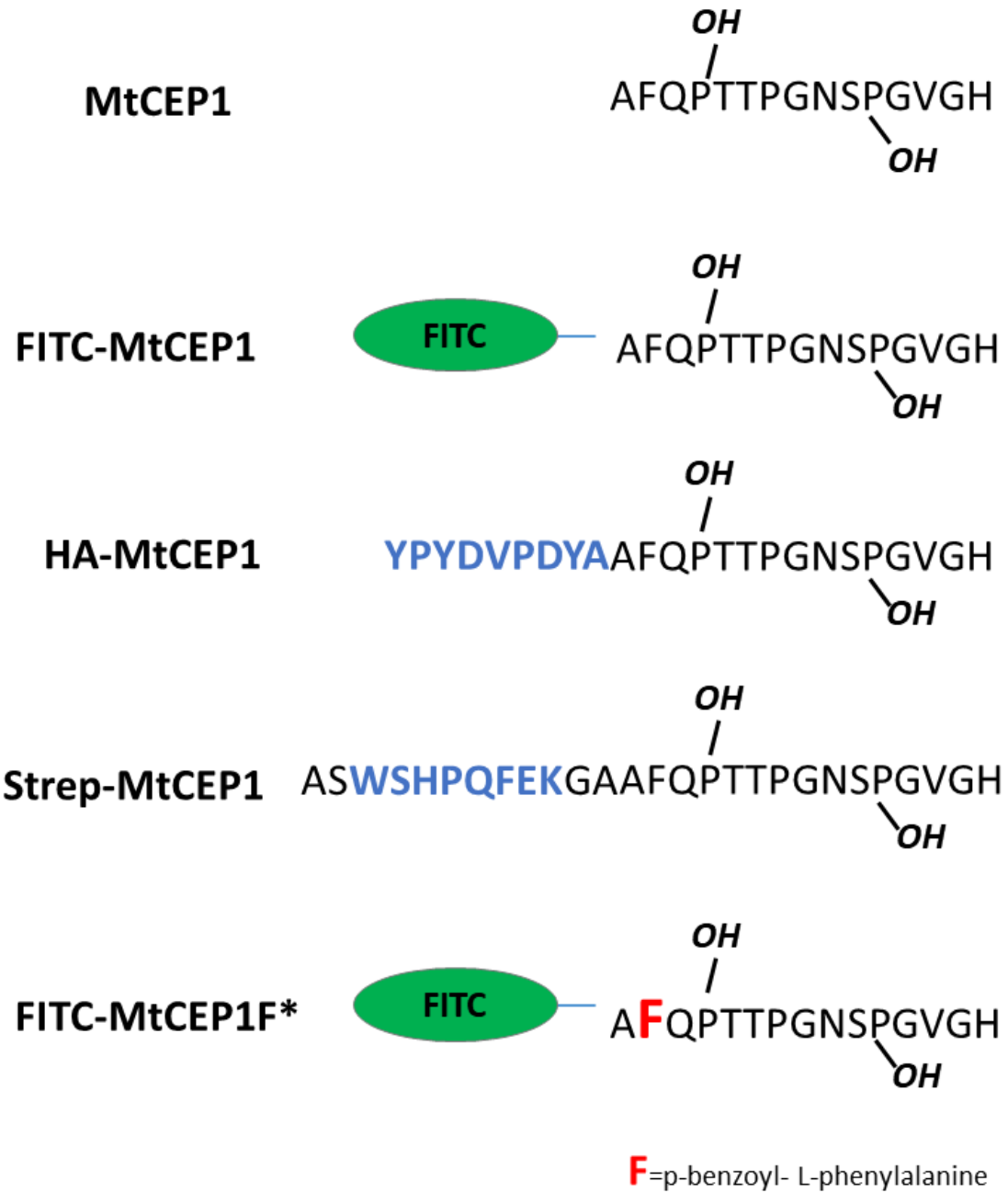


Figure 5.1. The MtCEP1 derivatives used in this study. The HA tag and Strep tag were labeled in blue. The photo-crosslinker was labeled in red.

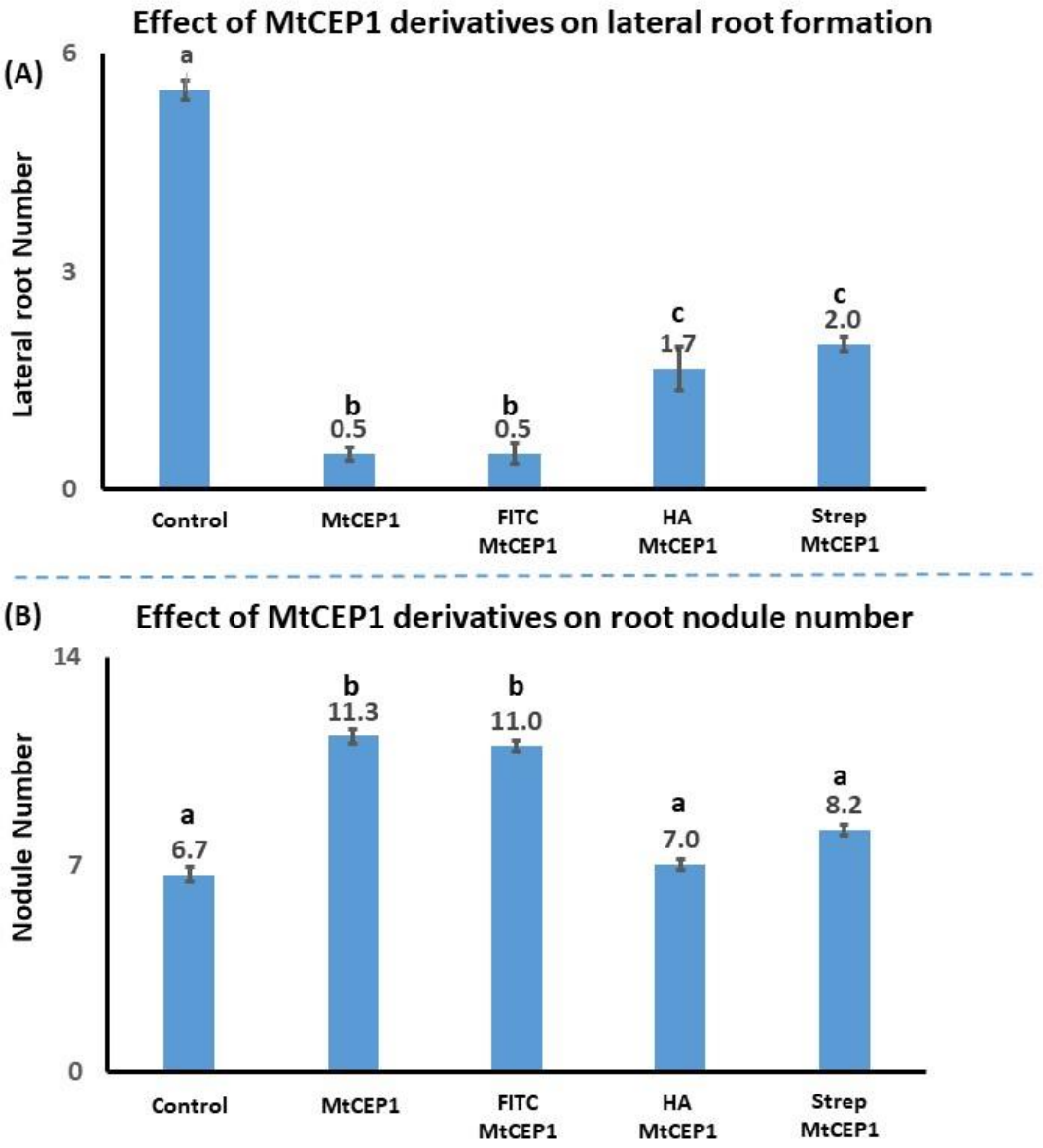


Figure 5.2. N-terminal extension difference of MtCEP1 will affect the biological activity

(A)(B) The biological activity of MtCEP1 derivatives was accessed on A17 roots. **(A)** The effect of root 1 x 10^{-6}M MtCEP1 derivatives on lateral root and **(B)** root nodule number at 14 dpi. Water was used as a control. Different letters indicate a statistically significant difference ($P \leq 0.05$, one-way analysis of variance (ANOVA) followed by Tukey multiple comparisons test) (error bars, ±SE, $n \geq 15$).

5.2.2 MtCEP1 derivatives with FITC tags and photo-cross-linker are biologically active

We further compared MtCEP1 to its variant, FITC-MtCEP1F*, to ensure that FITC-MTCEP1F* retained biological activity. The two peptides were identical, except FITC-MTCEP1F* had a *p*-benzoyl-L-phenylalanine (F*) at its 2nd residue instead of phenylalanine as well as the *N*-terminal FITC tag (Fig. 5.1). Both peptides were biologically active. FITC-MTCEP1F* inhibited Medicago lateral root formation and increased the number of root nodules similarly to that reported for MtCEP1 (Imin et al., 2013; Mohd-Radzman et al., 2015) (Fig. 5.3) and to FITC-MTCEP1 (Fig. 4.2; Chapter 4). This suggested that the modified peptides were biologically active and can bind to the receptor *in vivo* to alter the number of lateral organs produced on Medicago roots.

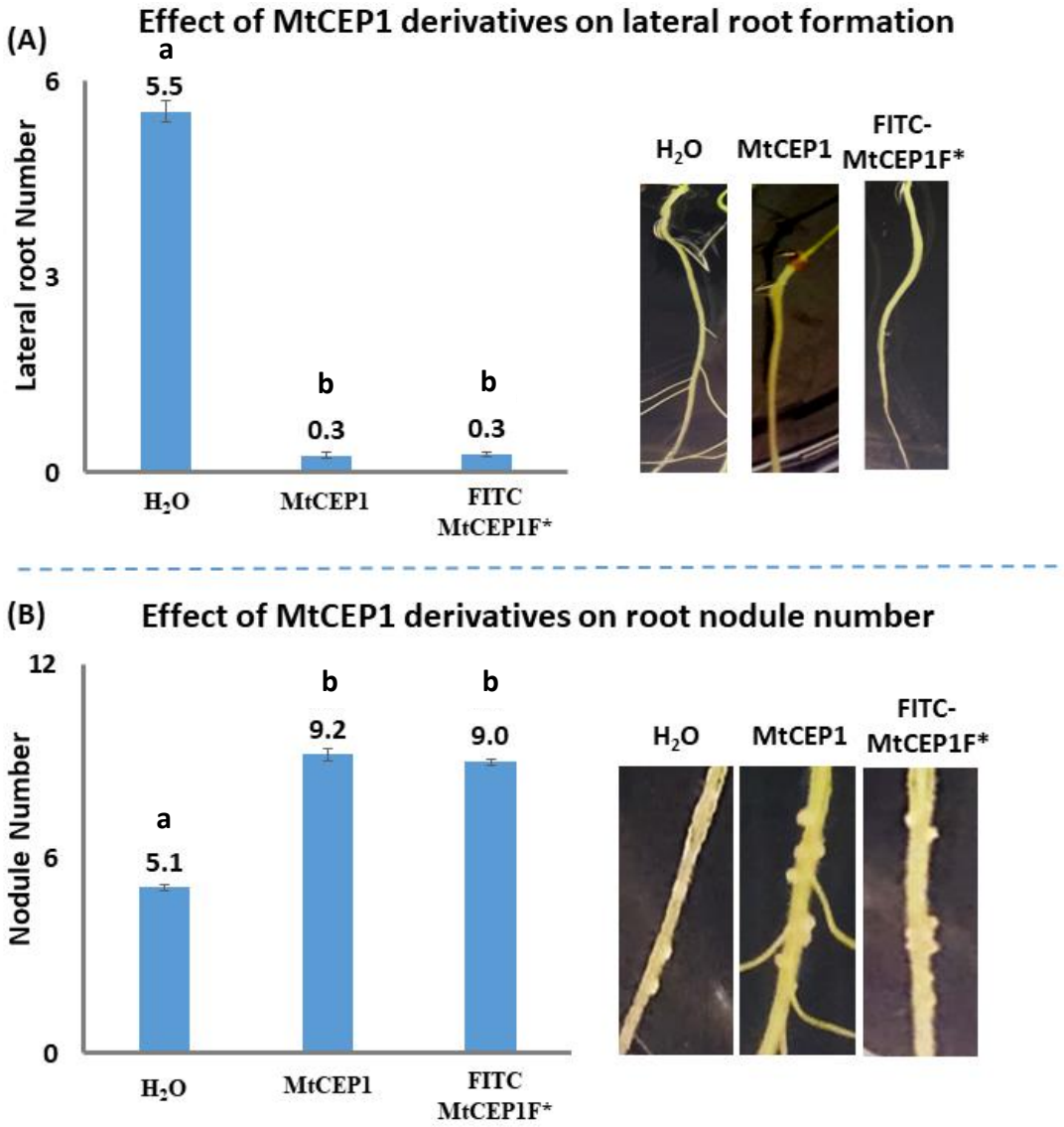


Figure 5.3. The MtCEP1 derivative with a ρBpa residue and FITC tag is biologically active

(A)(B) The biological activity of FITC-MtCEP1F* derivative was accessed on A17 roots. **(A)** The effect of FITC-MtCEP1F* (1 x 10^{-6}M) derivative on lateral root and **(B)** root nodule number. Lateral organ formation in A and B was scored at 14 dpi. Different letters indicate a statistically significant difference ($P \leq 0.05$, one-way analysis of variance (ANOVA) followed by Tukey multiple comparisons test) (error bars, ±SE, $n \geq 15$)..

5.2.3 Photo-activation of FITC-MtCEP1F* resulted in specific crosslinking to WT VT

To test the specificity of MtCEP1 binding to its putative receptor, CRA2, we covalently cross-linked FITC-MtCEP1F* to WT VT *in vivo* using UV at 365 nm. If *p*-benzoyl-L-phenylalanine is sufficiently proximal to a suitable amino acid in CRA2 this should result in a covalent linkage of the CEP peptide to CRA2. Specifically, any CRA2 amino acid side chain in CRA2's ligand binding domain proximal to the position of the phenylalanine residue at position 2 in the ligand will be cross-linked (Fig. 5.4). Confocal microscopy observations showed that FITC-MtCEP1F* bound to the WT VT (Fig. 5.5A). Moreover, a magnification of the tissue showed that FITC-MtCEP1F* bound at the periphery of individual VT cells (Fig. 5.5B) but not any internal membrane, which is consistent with the CRA2 being in the plasma membrane. The binding of FITC-MtCEP1F* was UV and CRA2 dependent (Fig. 5.5 C, D) since there was no FITC-MtCEP1F* binding in the absence of UV exposure or in UV-treated *cra2-11* VT exposed to FITC-MtCEP1F*. Overall, these results indicate that the binding of FITC-MtCEP1F* to the WT VT is CRA2, UV, and ligand dependent.

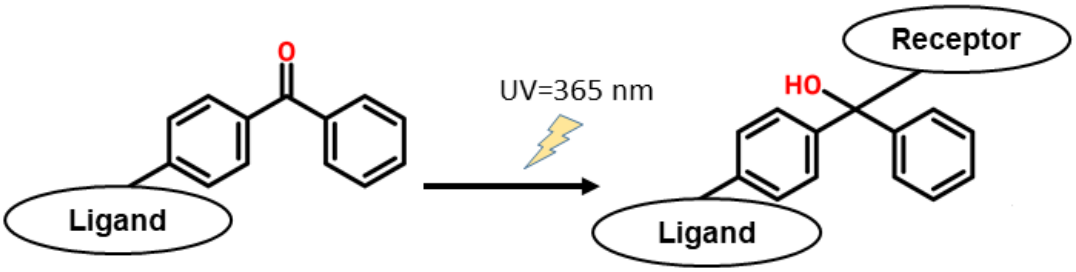


Figure 5.4 Putative photo-activated crosslinking of CEP ligand to the CEP receptor. MtCEP1 with *ρ*-benzoyl-L-phenylalanine (*ρ*bap) as second residue is crosslinked to an adjacent amino acid residue of the receptor.

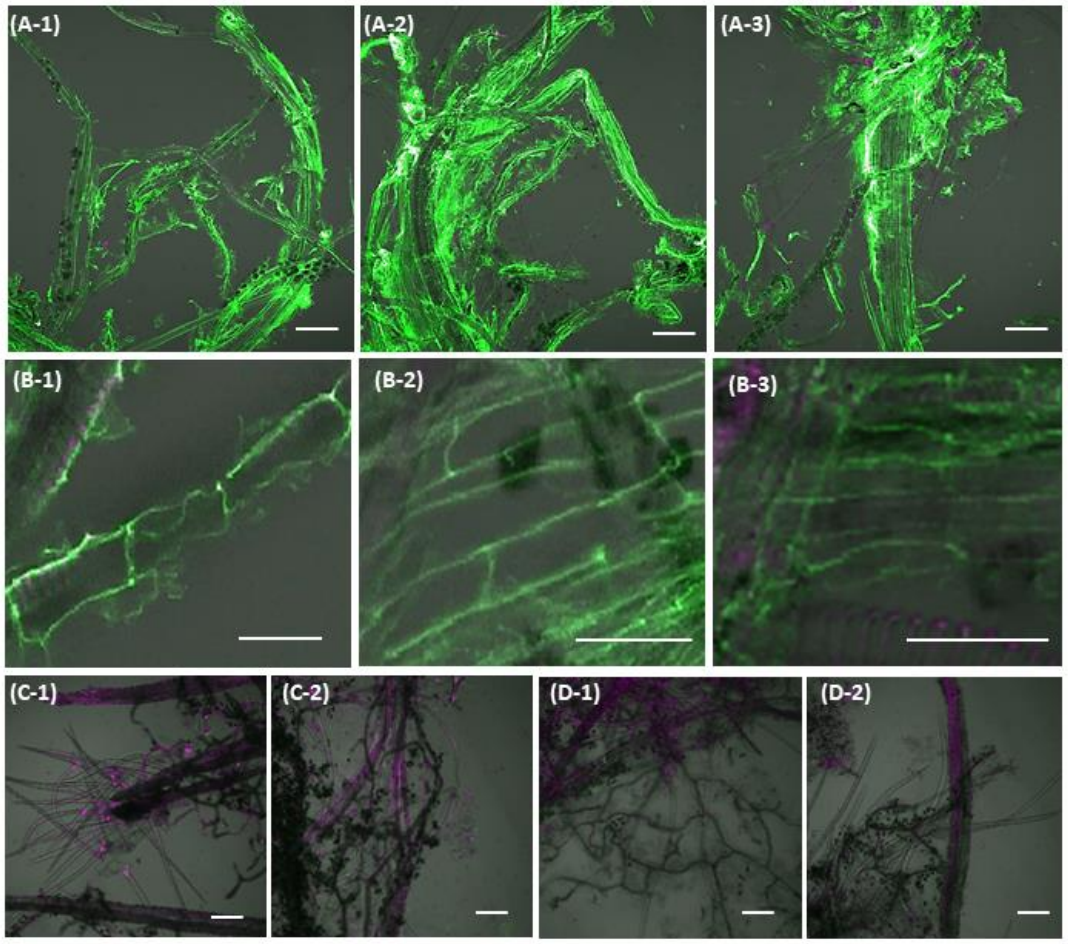


Figure 5.5. Photo-activated crosslinking of FITC-MtCEP1F* to VT is UV- and CRA2-dependent

(A) Confocal microscopy images of Medicago VT incubated with FITC-MtCEP1F* after UV exposure. (B) High magnification confocal microscopy images of Medicago VT incubated with FITC-MtCEP1F* after UV exposure. FITC-MtCEP1F* dependent fluorescence accumulates in the periphery of individual VT cells. (C) Confocal microscopy images of Medicago VT incubated

with FITC-MtCEP1F* without UV exposure. (D) Confocal microscopy images of VT from cra2-11 VT incubated with FITC-MtCEP1F* after UV exposure. The FITC fluorescence is green, and auto–fluorescence is purple. The pictures are representative of three independent experiments. Scale bar in (A)(C)(D) panels are 100 μm, and in (B) panel is 25 μm.

5.2.4 MtCEP1 can out compete FITC-MtCEP1F* for binding sites on WT VT

A17 VT was used for a binding competition assay to estimate the relative binding efficacy of MtCEP1 compared to FITC-MtCEP1F*(Fig 5.6). The co-incubation of 1 μM MtCEP1 with 1 μM FITC-MtCEP1F* resulted in no detectable binding of FITC-MtCEP1F* (Fig. 5.6A). This result suggested the MtCEP1 and FITC-MtCEP1F* bound to the same receptor, but MtCEP1 can outcompete FITC-MtCEP1F* for binding. We further tested the relative binding efficacy of MtCEP1 compared to FITC-MtCEP1F* by co-incubating a fixed concentration and dose of FITC-MtCEP1F* (50 μl at 1 μM) with diminishing concentrations and dosages of MtCEP1. Detectable binding of FITC-MtCEP1F* was partially restored when the MtCEP1 concentration was reduced to 10 nM, and appeared completely restored when the MtCEP1 concentration was reduced to 1 nM (Fig. 5.6 C-E). These results suggested that MtCEP1 is at least 100 to 1000 times more efficient at binding than FITC-MtCEP1F (Fig. 5.6 C-E) and implies that MtCEP1 peptide can bind to WT VT between 1 and 10 nM.

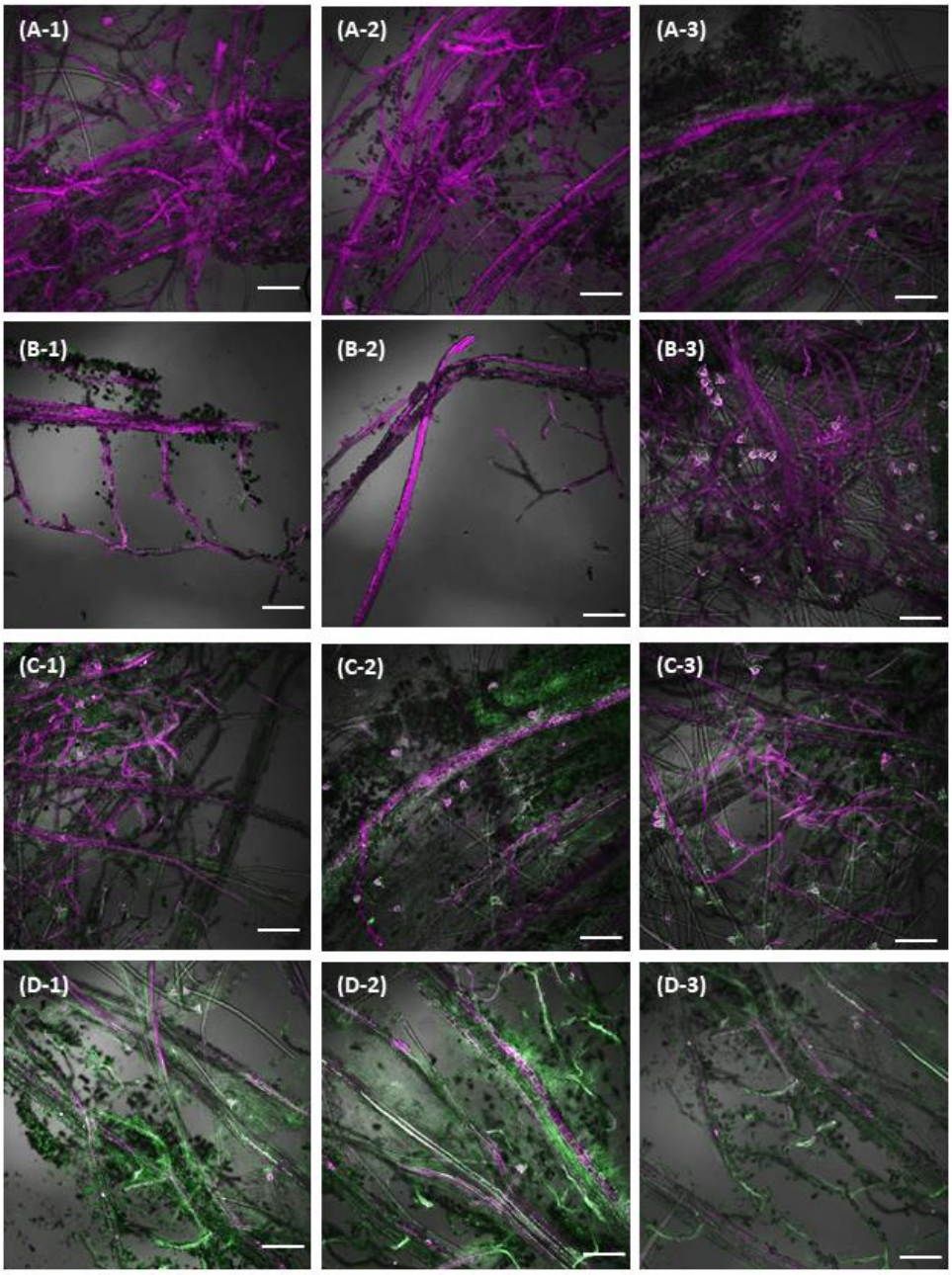


Figure 5.6. Binding site competition assays between FITC-MtCEP1F* and MtCEP1

Confocal microscopy pictures of WT leaf VT incubated with: (A) 1 µM of FITC-MtCEP1F* and 1 µM of MtCEP1; (B)1 µM of FITC- MtCEP1F* and 0.1 µM of MtCEP1; (C) 1 µM of FITC- MtCEP1F* and 0.01 µM of MtCEP1. (D) 1 µM of FITC- MtCEP1F* and 0.001 µM of MtCEP1. FITC fluorescence is green, and auto–fluorescence is purple. The pictures are representative of 3 independent experiments. Scale bar = 100 µm.

5.2.5 Visualization FITC- MtCEP1F* labeled protein separated by SDS-PAGE

To independently validate that CRA2 is the receptor for MtCEP1, we first incubated and UV irradiated WT VT with or without FITC-MtCEP1F*, and as a control, to cra2 VT with FITC-MtCEP1F*. We then extracted and separated the WT and cra2 VT proteins using SDS-PAGE before visualizing the separated proteins using a fluorescence detector set to measure the emissions at 530± 30 nm (which includes the 525 nm emission maximum for FITC) (Fig. 5.7). The results showed a fluorescent band common ~105 kDa (Lane 1 red arrow) but not when FITC-MtCEP1F* was photo-cross-linked to cra2 VT or in the absence of FITC-MtCEP1F* (Lanes 2-4). The band (red arrow) at ~105 kDa corresponds to the expected molecular mass of CRA2 (107.23 kDa). These results suggested that the band indicated with a red arrow (Fig. 5.7) corresponds to CRA2.

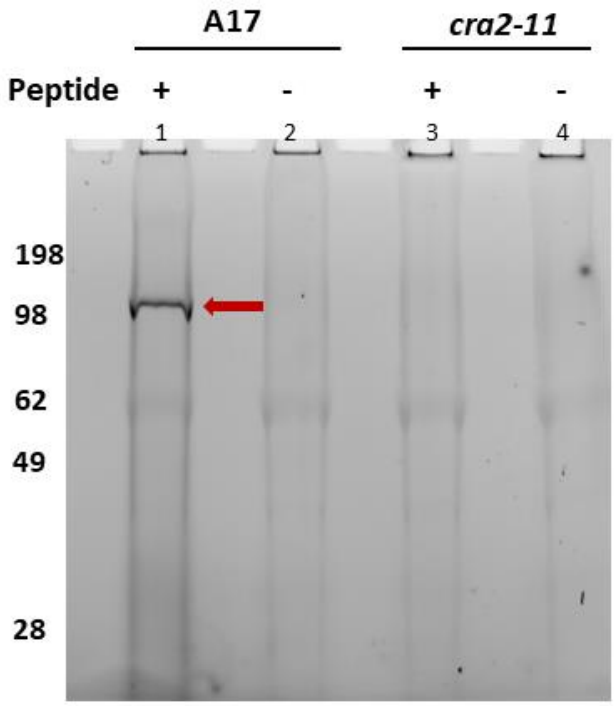


Fig. 5.7 Visualization FITC- MtCEP1F* labeled protein separated by SDS-PAGE Lane 1: 10 mg protein extracted from A17 Medicago vascular tissue labeled with FITC-MtCEP1F*. Lane 2: 10 mg protein extracted from A17 Medicago vascular tissue, no peptide control. Lane 3: 10 mg protein extracted

from *cra2* mutant labeled by FITC-MtCEP1F*. Lane 4: 10 mg protein extracted from *cra2*, no peptide control. The band above 100 kDa (red arrow) corresponds to the expected size of MtCRA2 (MW= 107.23 kDa).

5.2.6 Identification of MtCRA2 by mass spectrometry in protein bands extracted from SDS-PAGE

To validate that the unique bands in lane 2 of the SDS gel were MtCRA2 derivatives we examined the excised band at ~105 kDa for the presence of CRA2 peptides using MS. The MS analysis (Fig. 5. 8) identified peptides corresponding to MtCRA2 in the gel piece. The peptides represented 7% of total protein sequence coverage, and the identified peptides were all unique fragments to MtCRA2 (Fig. 5.8A). Fig. S4B shows a representative spectrum of an MtCRA2 peptide. Thus, we confirmed the binding of FITC-MtCEP1F* to Medicago VT is peptide and MtCRA2 dependent.

Fig. 5.8 Identification of MtCRA2 in protein extracted from SDS-PAGE.
A17 and cra2VT was incubated with FITC- MtCEP1F* and cross-linked. Total protein was extracted and separated by SDS-PAGE, and the FITC labelled protein was visualized. The fluorescent band unique to A17 cross-linked with FITC- MtCEP1F* was excised and subjected to mass spectrometry. (A) shows the protein sequence of MtCRA2. The peptides identified by mass spectrometry (shown in different colors) account for 7% coverage of CRA2. (B) A representative spectrum of a unique MtCRA2 peptide identified from CRA2.

5.3 Conclusion

In this chapter, we established a methodology to identify ligand receptor interaction *in vivo* by using site-direct photo activated crosslinking, purified ligand-receptor complex from Medicago VT and identified by Mass spectrometry.

First, we synthesized various MtCEP1 with different *N*-terminal affinity tag extensions and screened the biological activities of the synthetic peptides. We examined the effects of synthetic peptides on A17 lateral root and nodulation formation to verify the MtCEP1 derivatives can cause same phenotype as MtCEP1. This result implied that the synthetic ligands still bind the same receptor. The results showed all MtCEP1 derivative with different *N*-terminal modifications can reduce the number of lateral roots, but only FITC-MtCEP1 can induce nodulation as well as MtCEP1. Based on the results, we used FITC as the *N*-terminal extension fluorescence tag.

Second, we replaced the second amino acid on FITC-MtCEP1 from phenylalanine to p-benzoyl-L-phenylalanine (ρBpa), which is a site directed, photo activated cross-linker we termed FITC-MtCEP1F*. The photo-active amino acid analog *p*-benzoyl-L-phenylalanine (*pBpa*) was selected in this study as a crosslinking reagent. *p*Bpa has been used as a cross-linker to study peptide-protein interactions since 1986 (Kauer et al., 1986). *p*BPA contains a photo-reactive and chemically stable benzophenone group in its side chain.

Upon UV irradiation at a wavelength around 365 nm, the photo-generated triplet state of the benzophenone group abstracts a hydrogen atom from a geometrically accessible carbon-hydrogen bond of the polypeptide backbone. The UV-activated benzophenone can attack adjacent CH bonds to form covalent cross-linking(Miyazaki et al., 2020). The carbon-hydrogen bonds within 3 Å of the carbonyl oxygen are targets for cross-linking (Farrell et al., 2005). The cross-linked protein complex was capable for following purification and identification with tandem mass spectrometry. A structure study of plant plasma membrane P-type H+-ATPase was published via *pBpa* and tandem MS in yeast model (Nguyen et al., 2018). However, few publications have reported using this technique directly in plant tissues.

Third, we further test the biological activities of FITC-MtCEP1F* to check if the modified peptide still bound to the receptor and cause the same phenotype as MtCEP1. The results showed the FITC-MtCEP1F* can reduce the lateral roots formation and increase root nodule number as well as MtCEP1, which indicated the modified peptide still bounds the same receptor and can be used for the ligand receptor complex purification.

Fourth, with the assistance of ρBpa, we validated the binding of FITC-MtCEP1F* to Medicago leaf VT was MtCRA2 dependent, and this binding can be out competed by unlabeled MtCEP1 by at least 100-fold ratio. The results suggested the MtCRA2 is crucial for the Medicago VT binding of FITC-MtCEP1F*, and MtCEP1 bounds the same receptor as FITC-MtCEP1F* at least 100 times more efficiently.

Fifth, we purified the protein from semi purified VT via a detergent based method. The purified proteins were separated on the SDS-PAGE and a protein band in WT VT was shown to specifically bind to FITC-MtCEP1F* as visualized using a fluorescence detector set to measure FITC. The results showed a unique band was detected in WT VT plus FITC-MtCEP1F* peptide with corresponding to the expected molecular mass of CRA2, which were not detected in the protein sample purified from WT VT alone and receptor mutant VT plus FITC-MtCEP1F*. This unique fluorescence band was cut out, trypsin

digested, and the purified peptide subjected to MS. As the result, unique peptides corresponding to MtCRA2 were identified that corresponded to MtCRA2. The results strongly indicate that FITC-MtCEP1F* directly binds to MtCRA2.

Literature cited

Farrell IS, Toroney R, Hazen JL, Mehl RA, Chin JW (2005) Photo-cross-linking interacting proteins with a genetically encoded benzophenone. Nat Methods **2:** 377-384

Imin N, Mohd-Radzman NA, Ogilvie HA, Djordjevic MA (2013) The peptide-encoding CEP1 gene modulates lateral root and nodule numbers in *Medicago truncatula*. J Exp Bot **64:** 5395-5409

Kauer JC, Erickson-Viitanen S, Wolfe HR, Jr., DeGrado WF (1986) p-Benzoyl-L-phenylalanine, a new photoreactive amino acid. Photolabeling of calmodulin with a synthetic calmodulin-binding peptide. J Biol Chem **261:** 10695-10700

Miyazaki R, Akiyama Y, Mori H (2020) A photo-cross-linking approach to monitor protein dynamics in living cells. Biochimica et Biophysica Acta (BBA) - General Subjects **1864**

Mohd-Radzman NA, Binos S, Truong TT, Imin N, Mariani M, Djordjevic MA (2015) Novel MtCEP1 peptides produced in vivo differentially regulate root development in *Medicago truncatula*. Journal of Experimental Botany **66:** 5289-5300

Nguyen TT, Sabat G, Sussman MR (2018) In vivo cross-linking supports a head-to-tail mechanism for regulation of the plant plasma membrane P-type H(+)-ATPase. The Journal of biological chemistry **293:** 17095-17106

Wong S-H, King C-Y (2015) Amino Acid Proximities in Two Sup35 Prion Strains Revealed by Chemical Cross-linking. The Journal of biological chemistry **290:** 25062-25071

Chapter 6: Discussion

6.1 Thesis Overview and Key Results and Findings:

This thesis developed two broad approaches to investigate the interaction of the LRR-RLK, MtCRA2, with its putative cognate ligands of the CEP family. Chapters 3 4 and 5 utilized direct proteomic approaches and mass spectrometry to analyze the results and Chapters 4 and 5 developed *in situ* methods to visualize the direct binding of CEP ligands to native receptor complexes in the vascular tissues of leaves using confocal microscopy. The overall results show that peptides of the CEP family specifically bind to MtCRA2 localized to a subset of vascular tissue cells in Medicago.

Using direct proteomic approaches chapter 3 identified LRR-RLKs and in particular MtCRA2 using enhanced solubilization and fractionation techniques combined with the latest mass spectrometry instrumentation. Whilst these approaches identified 3289 TMP groups or 36.57% of all transmembrane proteins (TMPs), including trans membrane proteins with a significantly higher number of transmembrane domains (TMDs), and enriched highly hydrophobic proteins such as membrane transporters, they proved unsuitable for an in depth analysis of LRR-RLKs. Neither methodological approach gave particularly high coverage of LRR-RLKs even though 41% of the 538 annotated LRR-RLKs were identified. In addition, only one MtCRA2 peptide was identified using either or the solubilization, fractionation and MS analysis strategies. A proteomic analysis of vascular tissue material in Chapter 4, which should have enriched the relative proportion of MtCRA2 in the sample, also showed poor coverage of LRR-RLKs and MtCRA2

In Chapter's 4 and 5 new approaches were developed to optimize the purification of live vascular tissue and to use fluorescently tagged group 1 CEP ligands to investigate if MtCRA2 was the primary receptor for these ligands. The addition of the fluorescent tag did not unduly affect the biological activity of the ligand. We then exploited the natural affinity of the ligand for the receptor complex to demonstrate specific binding by using two crosslinking and washing strategies combined with confocal microscopy. Visualization of the ligand-receptor complex in an SDS gel finally enabled MS verification of MtCRA2 as

the receptor for group 1 but not group 2 CEPs

6.2 Proteomic approaches

6.2.1 Rationale for the approach in Chapter 3

Chapter 3 utilized mass spectrometry in combination with detergent-free biochemical approaches to identify MtCRA2 and other members of the LRR RLK family directly. Detergent based approaches were avoided because they are difficult to remove prior to MS analysis and detergent contamination of mass spectrometers leads to poor results and instrument fouling and downtime. Chapter 3 used two protein solubilization techniques combined with proteolytic digestion and extensive fractionation before MS analysis. We tested the hypothesis that the poor representation of LRR-RLKs in the Medicago proteomic literature was due to inadequate solubilization of membrane proteins, which leads to ineffective enzymatic digestion of the material. Microsomal membranes (MM) were utilized to enrich the membrane protein component in the starting material, and we applied the reported superior capacity of formic acid (FA) to solubilize the membrane proteins followed by a chemical cleavage using CnBr, which cleaves at rare methionine residue, to break the proteins into large peptide fragments. The cleavage of the protein material was validated by MS analysis (Appendix Table 1). To ensure that the average size of the peptides generated was more suitable for mass spectrometry, we used trypsin and Lys-C to cleave the larger peptide fragments to generate peptides with Arg or Lys termini, which promotes the production of 0.6–4 kDa peptides that are ideal for MS ionization and analysis (Fischer and Poetsch, 2006). As a reference point, we compared this strategy to the universally used bottom up proteomic strategy using urea solubilization followed by trypsin and Lys-C digestion. Following digestion, the peptides were fractionated using high pH reverse phase (section 2.6) and subjected to 2-D LC MS-MS analysis using a high precision orbitrap MS combined with a more effective chromatographic separation using in house columns (section 2.7). The results were analyzed by using the TMHMM, GRAVY, subcellular location prediction, protein functional annotation, TRAMDOMI algorithms (section 2.9 & 2.10 TMD analysis by TMHMM and TRAMDOMI; 2.11

subcellular protein location prediction by LOCALIZER; 2.12 functional annotation by Mercator and 2.13 Calculate the grand average of hydropathy (GRAVY) value for protein sequences), the latter of which was a new development of this thesis. The TRAMDOMI algorithm developed in this thesis is the first customized python script to map the actual detected peptide set to the *Medicago* TMD motif set to reveal the true extent of the TMDs identified (Section 2.10). With the help of TRAMDOMI, one can determine how well the purification method used enriches peptides containing all or part of the TMDs within transmembrane proteins in the samples.

6.2.2 Major findings using proteomic approaches

As Fig. 6.1 shows, the FA-CTLC method in Chapter 3 was superior at solubilizing and digesting proteins from MM (microsomal membrane) preparations (Fig. 3.1). Of the 57065 proteins in the MT data base, 13274 (23.26%) are predicted to be TMPs. The combined output of the two methods identified 3289 TMP groups and 36.57% of all TMPs. This is approximately 1.5-fold more protein identifications achieved in a recent quantitative proteomic analysis of young Medicago seedlings (Long et al., 2018), and more comparable to the number of Medicago proteins identified using similar sampling and bioinformatics procedures and similar instrumentation (Marx et al., 2016). It should be noted that Marx et al (2016) analyzed more individual fractions (31 fractions) from high pH reverse phase HPLC than used in this study (16 fractions). Therefore, this thesis acceived a comparable number of identifications by a cost-effective method with less fractions and MS runs. About 50% of the trans membrane proteins identified using either method had only one TMD but, reassuringly, both purification methods gave no significance difference in distribution of TMDs to that predicted by analyzing the theoretical distribution of TMDs in all *Medicago* trans membrane proteins using the THMMM algorithm (Fig. 3.3). Therefore, this result suggests that there was no major bias of either method in identifying TMPs.

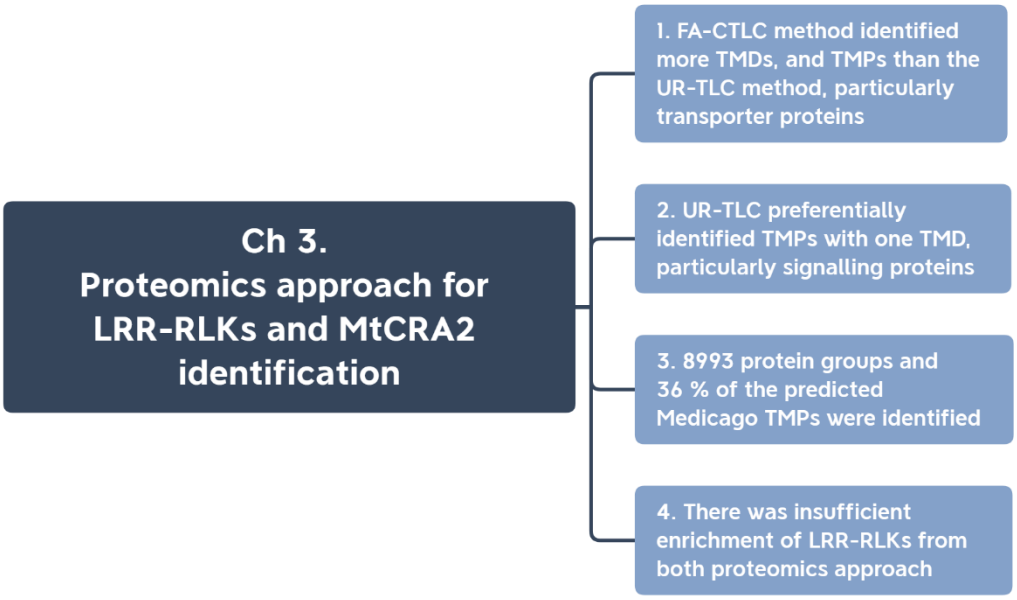


Fig. 6.1. The key results and findings from chapter 3: proteomics approach for LRR-RLKs and MtCRA2 identification

Given the high accuracy and resolution of modern day MS, why was the total % of trans membrane proteins so low (Fig. 3.3) First, an analysis of the results suggest it is likely that only the most abundant trans membrane proteins populate the lists of proteins identified and that trans membrane proteins of lower relative abundance were not detected by the MS. Secondly, subsequent analysis (Fig. 3.3) also showed that a substantial number of protein groups (5209 for UR-TLC and 6342 FA-CTLC) identified using either method had 0 TMDs. The high number of proteins with no TMDs, despite the enrichment of membrane proteins in the initial MM starting material, suggests that the published procedures for removing non-membrane proteins (e.g. using Na carbonate washes at pH 11) have poor efficacy or these cytoplasmic proteins have a strong affinity to membranes or the components within. The presence of these non-membrane proteins in the proteomic output would almost certainly hamper the identification of more trans membrane proteins and further fractionation of the membrane components in the starting material using ultra centrifugation would be recommended to remedy this issue. Finally, we identified a significant benefit of using the FAC-TLC method: the FA-CTLC method preferentially identifies trans membrane proteins with a significantly

higher number (9.36 fold) of TMDs than the UR-TLC method and each method identified partially non-overlapping TMP cohorts, as demonstrated in Fig. 3.4. Each purification method still had its bias, since 629 protein groups are unique to FA-CLC identification method whereas 1191 protein groups are unique to UR-TLC. This result was validated by the FAC-TLC method identifying more transporter proteins, which have multiple > 8 TMDs (Fig. 3.3). FA-CTLC identified more trans membrane proteins with a higher proportion of TMDs with a positive bias towards transporter proteins, whereas UR-TLC preferentially enriched signaling proteins, which contain one TMD. The results implied that the more hydrophobic trans membrane proteins were difficult to denature or solubilize using 8M urea. Therefore, UR-TLC method most likely shaves the exposed extra- and intra- cellular domains that loop away from the regions of trans membrane proteins imbedded inside the membrane. This deficiency leads to lower protein sequence coverage for proteins with a higher number of TMDs. By contrast, the UR-TLC method gave better identification of trans membrane proteins with one TMD. The FA-CTLC method identified 9 times more TMDs and enriched more trans membrane proteins than UR-TLC method. Since increasing the protein sequence coverage of trans membrane proteins can be benefit to improve the quantitative proteomics (Ishihama et al., 2008; Millioni et al., 2011; Koziol et al., 2013) and FA-CTLC can provide higher sequence coverage in TMDs, this method had the potential to provide superior data for quantitative proteomics. Therefore, we recommend that a combination of the two methods be used to achieve better TMP identification and better coverage of the peptides within TMPs. After combining the results derived from the two methods, we identified 8993 protein groups and 3289 trans membrane proteins in young shoot tissues. If more tissues were examined and different, or more extensive, membrane fractionation techniques applied where non-membrane proteins were minimized, the number of trans membrane proteins and their coverage would be expected to increase.

6.2.3 Both proteomic approaches failed to identify LRR-RLKs and MtCRA2

LRR-RLKs have a single pass TMD and given that both methods identified a high proportion of single pass TMD proteins it might be expected that LRR-RLKs would yield high quality results. Disappointingly, after applying MS analysis of the proteins identified using either the FA-CTLC or the UR-TLC method, there was insufficient enrichment and poor coverage of LRR-RLKs as a cohort and this resulted in an equally poor identification of MtCRA2. Of the 538 LRR-RLKs family proteins in Medicago (Appendix Table 3.13, Sheet 1), 224 LRR-RLKs were identified in total by applying both methods (Appendix Table 3.13, Sheet 2). The UR-TLC method identified 137 LRR-RLKs protein groups (25.46 % of the total), and the FA-CTLC method identified 120 LRR-RLKs protein groups (22.30%) (Appendix Table 3.13, Sheet 3&4) with at least one unique peptide. The average peptide coverage for the LRR-RLKs was approximately 5% for most of the identified species but only 1% of MtCRA2 (one unique peptide) was identified (Appendix Table 3.13, Sheet 3&4). Thus, both purifications methods were not suitable to provide an extensive analysis of LRR-RLKs due to the low sequence coverage. Since the results indicated that better enrichment of the starting material might remedy this situation, alternative proteomic strategies were used in Chapters 4 and 5.

6.2.4 Alternative proteomic procedures used in this thesis to identify LRR-RLKs and MtCRA2

In Chapters 4 and 5 alternative approaches were devised that combined Medicago leaf vascular tissue purification and ligand-receptor crosslinking to specifically examine MtCRA2 function and localization. As part of the validation process for the enrichment of vascular tissue, we employed a gel based detergent- and size fractionation-based approach (SDS-PAGE) to extract the proteins from the semi purified VT preparations and subsequently examined the protein content using mass spectrometry. One expectation of this process is that the proteins identified would be TMP proteins known to be localized to VT cells and that this might also include MtCRA2. Among the 4950 proteins

identified 931 were membrane proteins (Fig. 4.4) and there was an enrichment of VT related proteins (Fig. 4.5), however few (only 25) were LRR-RLKs and MtCRA2 was not represented (Appendix Table 4.3). Therefore. despite the enrichment of VT in the starting material, this approach also was not effective at identifying a huge diversity of LRR-RLKs or MtCRA2 and this was most likely due to their low relative abundance. Therefore, alternative approaches were used to analyze MtCRA2 and to demonstrate that it interacted with ligands of the CEP family. These approaches included using formaldehyde and photo-activated crosslinking of fluorescently tagged CEP peptides to Medicago leaf VT, as summarized below.

6.3 Rationale for the approach in Chapter 4

In situ hybridization (Huault et al., 2014) and reporter gene constructs (Bryan et al., 2012; Tabata et al., 2013) indicated that the Arabidopsis CEP receptor, CEPR1 and the putative Medicago CEP receptor, MtCRA2, localize primarily to the root and shoot vascular tissue. Tabata et al (2014) provided the only biochemical proof that CEP peptides bind to CEPR1. They demonstrated that radiolabeled and photoaffinity tagged Arabidopsis CEP1 peptide can bind to a recombinant CEPR1 protein expressed in tobacco BY2 cells. Tabata et al (2014), however, did not demonstrate that AtCEP1 bound to CEPR1 complexes in Arabidopsis vascular tissue. Although pharmacological and genetic evidence suggested that Medicago MtCRA2 may be the functional orthologue of Arabidopsis CEPR1 (Mohd-Radzman et al., 2016), there was no biochemical proof that members of the CEP family bind to MtCRA2, and this remains an important gap in the current knowledge.

Another consideration is that the CEP family is divided into group 1 and 2 CEPs (Delay et al., 2013; Imin et al., 2013; Roberts et al., 2013; de Bang et al., 2017). Tabata et al (2014) showed that the group 1 CEPs physically bind to recombinant CEPR1 in tobacco cells using photoactivation tagging. The biological activity and binding specificity of group 2 CEPs is unknown. Although the group 2 CEPs share *C*-terminal homology to the group 1 CEPs (Fig. 1.5),

they vary considerably in their *N*-terminal amino acid composition to group 1 CEPs, and also lack two conserved proline residues, which are critical for the biological activity of group 1 CEP peptides (Mohd-Radzman et al., 2015). Given the lack of functional analysis of group 2 CEPs it is not known if they bind to CEPR1/CRA2 or other/another receptor/s. In addition, group 2 CEP peptides have not been detected in xylem fluid, whereas several group 1 CEP peptides are readily identified in the xylem fluid (Patel et al., 2018). Therefore, group 2 CEPs do not appear to have a long distance signaling capacity like group 1 CEPs.

6.3.1 Key findings from Chapter 4

As the Fig. 6.2 shows, chapter 4 aimed to determine if (a) group 1 and/or group 2 CEPs bound to MtCRA2/CEPR1 (b) “foreign” group 1 CEPs from Medicago are biologically active on Arabidopsis and (c) Medicago CEP receptors localize to particular cells in the vascular tissue. Chapter 4 utilized confocal microscopy to visualize the direct binding of fluorescently tagged CEP family ligands to native CEP receptor complexes in the vascular tissues of Medicago and Arabidopsis leaves. This involved, firstly, designing MtCEP1 peptide derivatives with a fluorescent tag that did not destroy biological activity (Fig. 4.1; Fig. 4.2). Second, we developed a procedure to isolate intact vascular tissue cells by optimizing approaches to semi-purify live vascular tissue from Medicago and Arabidopsis leaves (Fig. 4.3). We used sonication and cell wall-degrading enzymes to remove leaf mesophyll cells without disrupting or killing vascular tissue cells (Fig. 4.3). The semi-purified vascular tissue preparations then allowed unimpeded access to the vascular tissue of fluorescently tagged CEP peptides and any added crosslinking agent(s). The procedure then determined if the biologically active, fluorescently tagged Medicago CEP1 has a specific affinity for the putative Medicago CEP receptor, MtCRA2. To demonstrate specificity of binding to the putative receptor, vascular tissue was prepared from the MtCRA2 mutant, which is null for MtCRA2. Subsequently, we tested formaldehyde and DMP (Dimethyl pimelimidate) for their ability cross-link these peptides specifically to wild type vascular tissue cells but not those of a MtCRA2

mutant (Fig. 4.6). Finally, we determined if group 1 and group2 CEPs bound to MtCRA2.

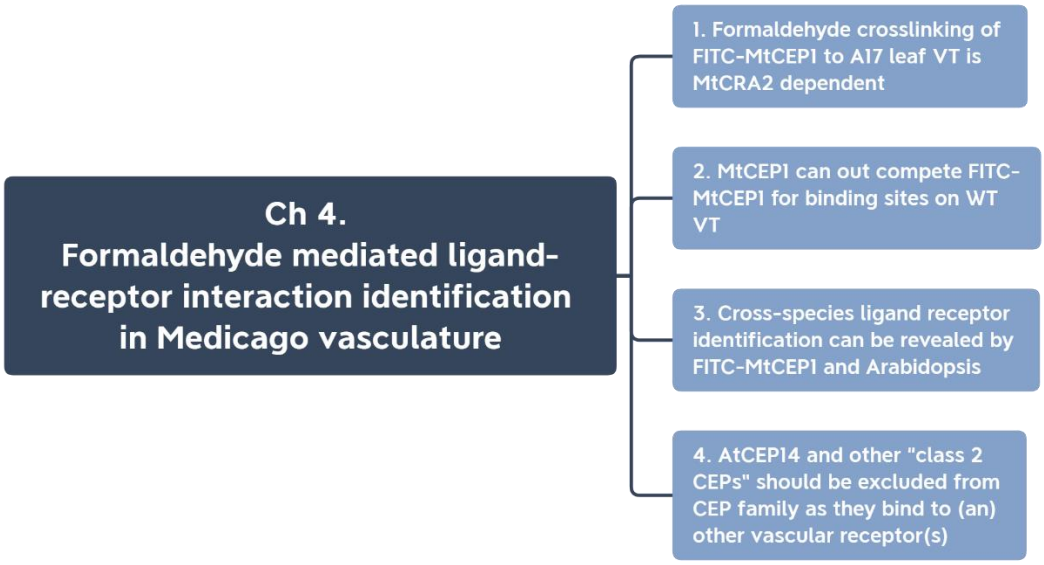


Fig. 6.2. The key results and findings from chapter 4: formaldehyde mediated crosslinking ligand- receptor interaction identification

Using confocal microscopy, we found that formaldehyde and DMP effectively cross-linked the fluorescently tagged group 1 CEP peptide, MtCEP1, to either wild type Medicago or Arabidopsis VT but not to VT of their respective CEP-receptor mutants (i.e. *cra2* and *cepr1*, respectively; Fig. 4.6). This result strongly implies that CEPR1 and MtCRA2 are receptors for group1 CEPs and MtCRA2 is a functional orthologue of AtCEPR1. The results also suggest that the MtCEP1 peptide tightly associates with native CEP receptor complexes in a specific subset of VT cells before crosslinking. FITC-MtCEP1 did not co-localize with the purple auto-fluorescing cells in the vascular tissues, which are most likely to be xylem tracheids and vessel cells (Fig. 4.7; Fig. 4.8), which incorporate phenolic compounds into their cell walls (Willemse, 1989; Donaldson and Williams, 2018). These results are consistent with MtCRA2 being localized in discrete cell types in the vascular tissue distinct from the purple auto-fluorescing cells. It is likely that MtCRA2 localizes to the phloem as in situ studies using promoter:GUS fusions indicate that CEPR1 is phloem localized in Arabidopsis (Bryan et al., 2012; Tabata et al., 2013; Taleski et al., 2020) and Medicago (Huault et al., 2014). However, these studies did not

provide specific knowledge of the tissue type(s) where MtCRA2 is localized. Therefore, a parallel technique that marks specific tissues in the sample after crosslinking would be of particular use. Such as carboxyfluorescein (CF dye) was widely used as a symplastic mobile phloem sap tracer (Jensen et al., 2011; Savage et al., 2013), Calcofluor White for general cell wall stain (Snow, 2016), Basic Fuchsin for lignin stain (Dharmawardhana et al., 1992) and Nile Red for suberin stain (Waduwara et al., 2008).

In addition, unlabeled MtCEP1 peptide effectively competed with FITC-MtCEP1 for binding sites on wild type vascular tissue, and these competition experiments showed that FITC-MtCEP1 was at least 100-fold less efficient at binding. The result also imply that 10 nM MtCEP1 is almost sufficient to completely displace a 100-fold higher concentration of FITC-MtCEP1 molecules from the MtCRA2 binding sites (Fig. 4.9). The reduced activity of FITC-MtCEP1 is consistent with previous results from our laboratory. These published results (Patel et al., 2018) showed that the addition of three or five amino acids located at the *N*-terminal flank of the MtCEP1 peptide domain partially reduced the biological activity of the peptide. Specifically, Patel et al (2018) showed that three to five amino acid extensions abolished the ability of this peptide to enhance root nodule formation but not the ability to reduce lateral root formation. The result in Fig. 4.9 is also consistent with prior unpublished observations using *N*- or *C*-terminal modifications of the MtCEP1 peptide using the TAMRA fluorescent probe; only the *N*-terminal TAMRA derivative retained biological activity whereas the *C*-terminal derivative lacked any biological activity. Overall, these results suggest that the mature size of the CEP peptide affects its biological activity and that modifications or extensions to the *N*- or *C*-terminus can affect biological activity moderately or severely. Therefore, there is a need to screen carefully the effect of modifications to CEP peptides on CEP peptide binding efficiency and biological activity.

The results also showed that much lower formaldehyde concentrations than used in interactome studies (Thavarajah et al., 2012) are sufficient to efficiently crosslink CEP peptides to MtCRA2 (Fig. 4.10). This is significant for future studies as minimizing the formaldehyde concentration can prevent the

excessive formation of higher molecular weight complexes. Therefore, it might be possible to use formaldehyde to cross-link MtCEP1 peptides to MtCRA2 and other closely associating molecules (e.g. membrane localized co receptors, protein scaffolds like Feronia (Li et al., 2016) or other receptors or alternatively to intracellular signal transduction proteins that might be involved in phosphor relays). It is conceivable that the formation of these hypothetical macromolecular complexes could be fractionated, for example, by size exclusion chromatography and specifically detected using the FITC fluorescence signature. A conceptually similar procedure was used recently to macromolecular complexes in plants using a large-scale proteomic analysis (McWhite et al., 2020). Mass spectrometry could determine the nature of proteins forming these complexes.

We found also that formaldehyde can effectively crosslink the Medicago FITC-MtCEP1 to Arabidopsis CEPR1 in vascular tissue preparations (Fig. 4.11). This result is consistent with group 1 peptides having similar biological activity across species and binding to the corresponding orthologous receptors specifically. In addition, this result would explain why group 1 CEPs are strongly conserved across a wide range of seed bearing plants (Delay et al., 2013; Roberts et al., 2013; Ogilvie et al., 2014). These results suggest a conservation of CEP-CEP receptor signaling across a wide range of species.
Finally, we also found that group 2 CEP peptides bound to vascular tissue cells in leaves independently of MtCRA2 or CEPR1, and did not share the same biological activity as group 1 CEPs. This result indicates that group 2 CEPs may bind to other/another receptor and therefore may form a separate class of peptides altogether.

6.4 Rationale for approach in Chapter 5

In chapter 5 the strategy developed in Chapter 4 was extended to address the specific issue: do group 1 CEPs specifically bind to MtCRA2? Here, we used a derivate of FITC-MtCEP1 that incorporated a photoaffinity tag and determined

if it bound specifically to MtCRA2. Again, the relative binding activity of biologically active FITC-MtCEP1F* to vascular tissue derived from wild type and MtCRA2 was assessed. Finally, mass spectrometry was used to verify that peptides corresponding to MtCRA2 were present in a protein band excised from a SDS-PAGE gel that was fluorescently labelled with FITC-MtCEP1F*.

6.4.1 Key results in Chapter 5

As the Fig. 6.3 shows, the approache used to investigate MtCEP1-MtCRA2 interactions involved the use of MtCEP1 derivatives with different tags on the *N*-terminus to facilitate visualization or purification (Fig. 5.1). The addition of HA or Strep tags to the *N*-terminus negatively affected biological activity so these were abandoned (Fig. 5.2). This result, however, further emphasized that longer *N*-terminal extensions negatively impacted biological activity. FITC-MtCEP1F*, however, retained the full range of biological activity (Fig. 5.2 and 5.3) and therefore it was a reasonable expectation that UV crosslinking may bind FITC-MtCEP1F* to the ligand binding site of MtCRA2 (Fig. 5.4). Confocal microscopy examination of UV cross-linked WT and MtCRA2 vascular tissue showed that FITC-MtCEP1F* bound only to wild type (Fig. 5.5) and was competitively inhibited by unlabeled MtCEP1 (Fig. 5.6). The results (Fig. 5.6) also imply that MtCEP1 binds in the low nM range. Using SDS PAGE, we extracted size separated the proteins from wild type and *cra2* vascular tissue that was UV exposed in the presence and absence of FITC-MtCEP1F*. Examination of this gel using fluorescence detection revealed a fluorescent band at the approximate molecular mass of MtCRA2 (~105 kDa) only in the lane with wild type vascular tissue and FITC-MtCEP1F* (Fig. 5.7). The extraction of this band and examination by mass spectrometry showed that peptides corresponding to MtCRA2 were present (Fig. 5.8). Overall, these results demonstrate that MtCEP1 specifically binds to MtCRA2.

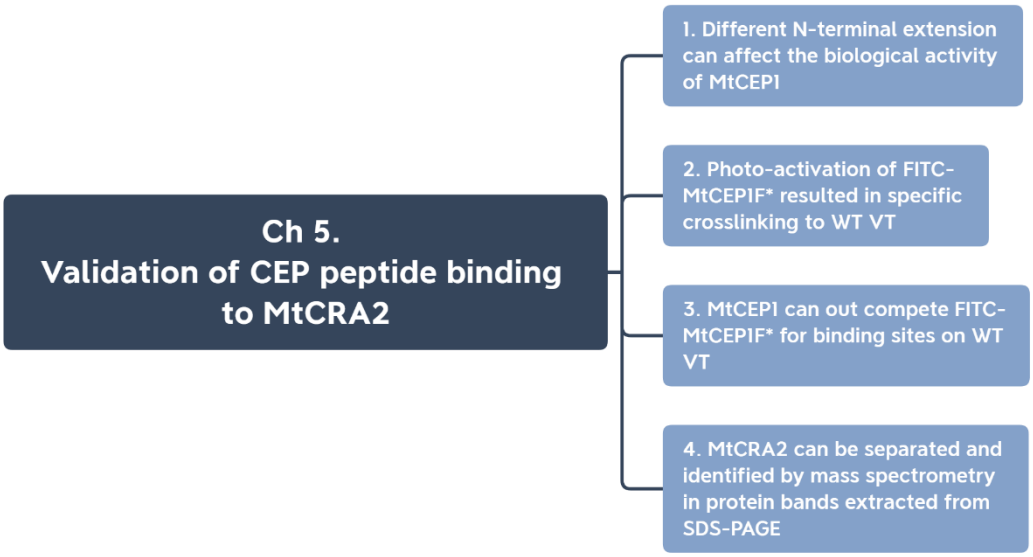


Fig. 6.3. The key results and findings from chapter 5: validation of CEP peptide binding to MtCRA2

6.5 Possible future development

Based on the proteomics approach results in Chapter 3, the FA-CTLC can provide better coverage of more abundant TMDs but failed to provide sufficient coverage of LRR-RLKs and MtCRA2. It is not ideal that we used microsomal membrane proteins in Chapter 3 as starting material as it contains the plasma membrane and membranes derived from the endoplasmic reticulum, Golgi apparatus, transGolgi network, tonoplast and plastids (LaMontagne et al., 2016). Therefore, it is more ideal to enrich the plasma membrane fractions to more effectively undertake these types of direct proteomic studies. The plant plasma membrane protein can be enriched by two-phase partitioning (Yoshida et al., 1983) with purity of about 95% (Alexandersson et al., 2008), or affinity purification by biotinylated the plasma membrane protein (deBlaquiere and Burgess, 1999). Plasma membrane enriched protein fractions would be expected to less contaminated with cytoplasmic proteins and other membrane proteins from organelles. This should result better coverage and more identifications in LRR-RLKs and MtCRA2. One major disadvantage of the FAC-TLC method in formic acid may cause excessive *N*- and *O*-formylation (Zheng and Doucette, 2016) and the strongly acidic conditions may lead to unexpected

cleavages at acid labile residues (Smith, 1988) . It may be good to use Trifluoroacetic acid (TFA) to replace the FA as protein denature reagent to avoid the possible protein modification.

In Chapter 4 we demonstrated the formaldehyde- and DMP-mediated crosslinking of CEPs to the MtCRA2 receptor. It may be possible to use other crosslinking agents instead of formaldehyde such as DSSO. DSSO is advantageous in that it be cleaved in the gas phase during tandem MS (MS/MS) using collision-induced dissociation (CID), whereas formaldehyde crosslinking is reversible at higher temperature (Nilsen, 2014). A major disadvantage of DSSO and DMP is that they require a higher pH for crosslinking due to the reliance of primary amines. Shifting the pH from 5.5 (the pH of the apoplast) to higher pH for crosslinking may cause the ligand to dissociate from the receptor before crosslinking. The success of DMP may suggest this may not be too problematic. It is noteworthy that the fluorescently-tagged derivative of the ligand used does not have a primary amine and if formaldehyde-mediated crosslinking is initiated from the ligand formaldehyde must activate the alternative suitable amino acids (e.g. Asn or His) since Cys, Lys, Trp, and Arg residues are absent in CEP1 ligand sequence. The ability of formaldehyde to activate residues in the ligand or the receptor as well as its pH independence also increases its versatility. It is also very cost-effective compared to other crosslinking agents.

This study did not identify MtCRA2's co-receptor. Based on numerous studies of LRR-RLKs that recognize endogenous or foreign peptide ligands, SERK (somatic embryogenesis receptor kinase) is commonly a co-receptor (Roux et al., 2011; Brandt and Hothorn, 2016; Chakraborty et al., 2018). These ligands are triggers for a multitude of processes involved in defense and development (as summarized in Fig. 1.3).

Based on this, it is highly possible that MtSERK is a co-receptor involved in forming a functional MtCRA2-MtCEP1 complex to trigger the downstream phosphorylation as a protein kinase. (Fig. 6.4) but this was not addressed in this thesis. However, if short term applications (~5 min) of low formaldehyde

concentrations are used this may minimize the irrelevant crosslinking of intracellular protein-protein complexes and enrich surface receptor-ligand crosslinking. By applying this strategy, a more specific FITC-MtCEP1 localization within the VT was revealed (Figure 4.10). It may be then possible to purify such ligand-receptor-co receptor complex by following the fate of FITC tag during separation such as size exclusion using fluorescence detection. By analyzing the constitution of such complexes with MS, it may be possible to identify the co-receptor(s) for MtCRA2 and possibly any interacting proteins that complex with MtCRA2 on the cytoplasmic (kinase) side of the receptor. Besides, the recently reported TurboID (Branon et al., 2018; Zhang et al., 2019) which based proximity labeling approach could be the possible approach to identify the coreceptor(s) protein(s) of CEPs.

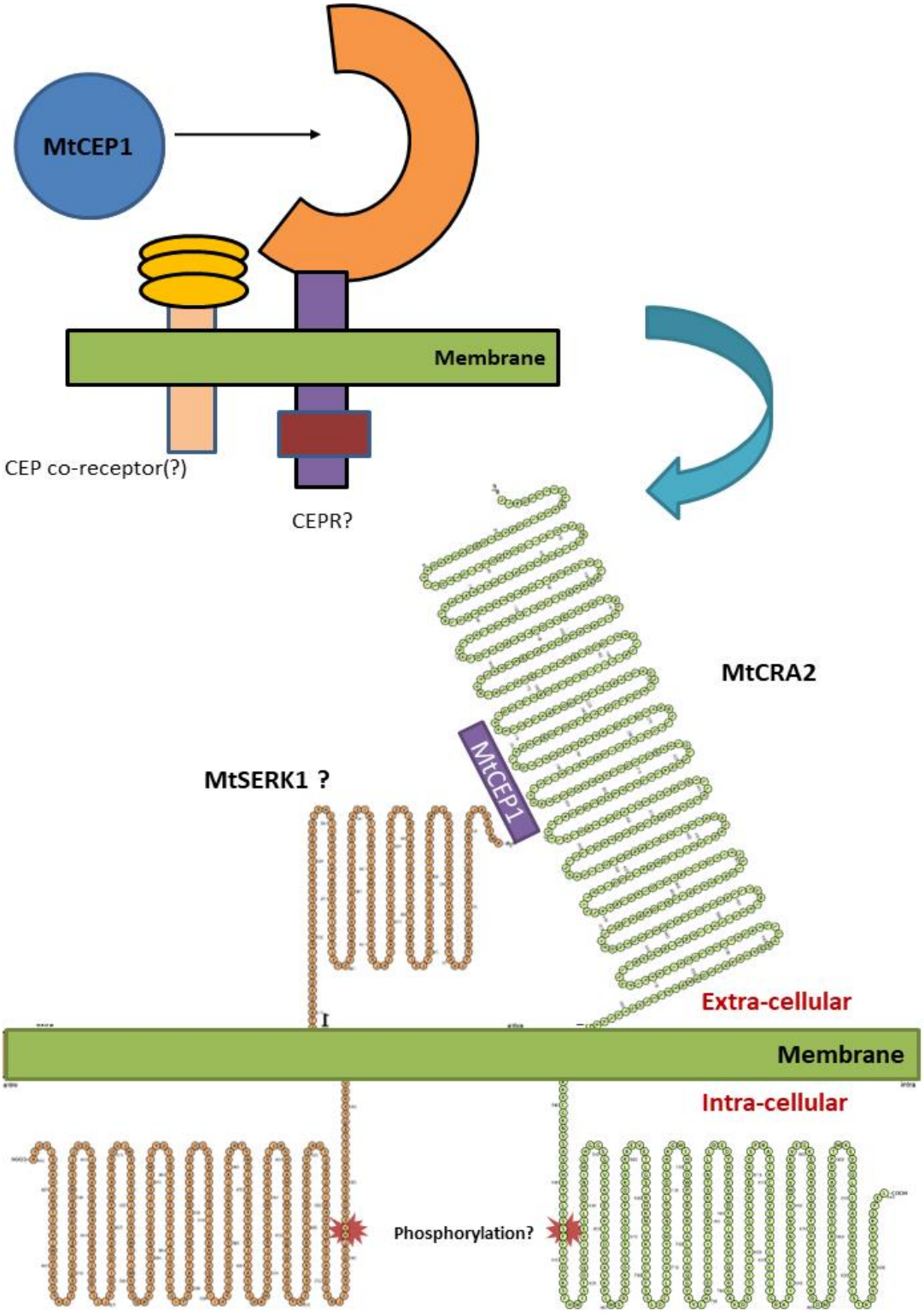


Fig. 6.4. The illustration diagrams of MtCRA2-MtCEP1-MtSERK1 complex

The schematic diagrams were made using the Protter online tool (Omasits et al., 2014) (http://wlab.ethz.ch/protter/start/).

In chapter5, we used the site-directed photo-activated cross-linker, *p*-benzoyl-

L-phenylalanine to more definitively identify MtCRA2 as receptor for MtCEP1. This was supported by MS identification. This strategy specifically and exclusively links the ligand with the binding site of the receptor. With further enrichment of tissues containing MtCRA2 such as using a large scale *in vivo* crosslinking initiative, such as using Bioruptor to handle multiple leaves in one time (Endo et al., 2016), it may be possible using MS to identify the binding site. *p*Bpa can reveal the interaction site in protein-protein interaction by MS identification (Wong and King, 2015). As the Fig. 6.4, the MtCEP1 should sit on the extracellular domain of MtCRA2, the *p*bpa will provide the covalent binding to MtCRA2. Then one could use novel database search engines such as XlinkX to find the crosslinking site between MtCEP1 and MtCRA2 to reveal how MtCEP1 binds to MtCRA2. XlinkX is a cross-link search engine in Proteome Discover for analyzing a multipronged data set (Liu et al., 2015; Liu et al., 2017). The software can identify cross-links obtained with non-cleavable cross-linkers by collision-induced dissociation for protein-protein interactions (Haupt et al., 2017; Liu et al., 2018; Gotze et al., 2019). In addition, size exclusion chromatography by FPLC could be combined with fluorescence detection to purify MtCRA2. This may allow the phosphorylation state of MtCRA2 to be examined more closely before and after ligand addition to determine phosphor sites in the kinase domain of MtCRA2.

Literature cited

Abas L, Luschnig C (2010) Maximum yields of microsomal-type membranes from small amounts of plant material without requiring ultracentrifugation. Analytical Biochemistry **401:** 217-227

Abdallah C, Valot B, Guillier C, Mounier A, Balliau T, Zivy M, van Tuinen D, Renaut J, Wipf D, Dumas-Gaudot E, Recorbet G (2014) The membrane proteome of *Medicago truncatula* roots displays qualitative and quantitative changes in response to arbuscular mycorrhizal symbiosis. J Proteomics **108:** 354-368

Abdiche Y, Malashock D, Pinkerton A, Pons J (2008) Determining kinetics and affinities of protein interactions using a parallel real-time label-free biosensor, the Octet. Anal Biochem **377:** 209-217

Alexandersson E, Gustavsson N, Bernfur K, Karlsson A, Kjellbom P, Larsson C (2008) Purification and Proteomic Analysis of Plant Plasma Membranes. *In* D Pflieger, J Rossier, eds, Organelle Proteomics. Humana Press, Totowa, NJ, pp 161-173

Almeida JG, Preto AJ, Koukos PI, Bonvin A, Moreira IS (2017) Membrane proteins structures: A review on computational modeling tools. Biochim Biophys Acta Biomembr **1859:** 2021-2039

Aloui A, Recorbet G, Lemaitre-Guillier C, Mounier A, Balliau T, Zivy M, Wipf D, Dumas-Gaudot E (2018) The plasma membrane proteome of *Medicago truncatula* roots as modified by arbuscular mycorrhizal symbiosis. Mycorrhiza **28:** 1-16

Arachea BT, Sun Z, Potente N, Malik R, Isailovic D, Viola RE (2012) Detergent selection for enhanced extraction of membrane proteins. Protein Expr Purif **86:** 12-20

Avila JR, Lee, J. S., Torii, K. U. (2015) Co-Immunoprecipitation of Membrane-Bound Receptors. Arabidopsis Book **13**

Awuni Y, Mu Y (2015) Reduction of false positives in structure-based virtual screening when receptor plasticity is considered. Molecules **20:** 5152-5164

Bai Y (2015) Detecting protein-protein interactions by gel filtration chromatography. Methods Mol Biol **1278:** 223-232

Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, Martin MJ, Natale DA, O'Donovan C, Redaschi N, Yeh L-SL (2005) The Universal Protein Resource (UniProt). Nucleic Acids Research **33:** D154-D159

Baldan-Martin M, Lopez JA, Corbacho-Alonso N, Martinez PJ, Rodriguez-Sanchez E, Mourino-Alvarez L, Sastre-Oliva T, Martin-Rojas T, Rincon R, Calvo E, Vazquez J, Vivanco F, Padial LR, Alvarez-Llamas G, Ruiz-Hurtado G, Ruilope LM, Barderas MG (2018) Potential role of new molecular plasma signatures on cardiovascular risk stratification in asymptomatic individuals. Sci Rep **8:** 4802

Bennett MK, Calakos N, Kreiner T, Scheller RH (1992) Synaptic vesicle membrane proteins interact to form a multimeric complex. J Cell Biol **116:** 761-775

Berendzen KW, Bohmer M, Wallmeroth N, Peter S, Vesic M, Zhou Y, Tiesler FK, Schleifenbaum F, Harter K (2012) Screening for in planta protein-protein interactions combining bimolecular fluorescence complementation with flow cytometry. Plant Methods **8:** 25

Bojar D, Martinez J, Santiago J, Rybin V, Bayliss R, Hothorn M (2014) Crystal structures of the phosphorylated BRI1 kinase domain and implications for brassinosteroid signal initiation. The Plant Journal **78:** 31-43

Brandt B, Hothorn M (2016) SERK co-receptor kinases. Curr Biol **26:** R225-226

Branon TC, Bosch JA, Sanchez AD, Udeshi ND, Svinkina T, Carr SA, Feldman JL, Perrimon N, Ting AY (2018) Efficient proximity labeling in living cells and organisms with TurboID. Nature Biotechnology **36:** 880-887

Breiden M, Simon R (2016) Q&A: How does peptide signaling direct plant development? BMC Biology **14:** 58

Bryan AC, Obaidi A, Wierzba M, Tax FE (2012) XYLEM INTERMIXED WITH PHLOEM1, a leucine-rich repeat receptor-like kinase required for stem growth and vascular development in Arabidopsis thaliana. Planta **235:** 111-122

Campos R, Goff J, Rodriguez-Furlan C, Van Norman JM (2020) The Arabidopsis Receptor Kinase IRK Is Polarized and Represses Specific Cell Divisions in Roots. Dev Cell **52:** 183-195.e184

Carpenter EP, Beis K, Cameron AD, Iwata S (2008) Overcoming the challenges of membrane protein crystallography. Current Opinion In Structural Biology **18:** 581-586

Casadonte R, Caprioli RM (2011) Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry. Nature protocols **6:** 1695-1709

Chakraborty S, Nguyen B, Wasti SD, Xu G (2019) Plant Leucine-Rich Repeat Receptor Kinase (LRR-RK): Structure, Ligand Perception, and Activation Mechanism. Molecules (Basel, Switzerland) **24:** 3081

Chakraborty S, Pan H, Tang Q, Woolard C, Xu G (2018) The Extracellular Domain of Pollen Receptor Kinase 3 is structurally similar to the SERK family of co-receptors. Scientific Reports **8:** 2796

Chapman K, Ivanovici A, Taleski M, Sturrock CJ, Ng JLP, Mohd-Radzman NA, Frugier F, Bennett MJ, Mathesius U, Djordjevic MA (2020) CEP receptor signalling controls root system architecture in Arabidopsis and Medicago. New Phytol

Chapman K, Taleski M, Ogilvie HA, Imin N, Djordjevic MA (2019) CEP-CEPR1 signalling inhibits the sucrose-dependent enhancement of lateral root growth. Journal of experimental botany **70:** 3955-3967

Chapman K, Taleski M, Ogilvie HA, Imin N, Djordjevic MA (2019) CEP–CEPR1 signalling inhibits the sucrose-dependent enhancement of lateral root growth. Journal of Experimental Botany **70:** 3955-3967

Chen EI, McClatchy D, Park SK, Yates JR, 3rd (2008) Comparisons of mass spectrometry compatible surfactants for global analysis of the mammalian brain proteome. Analytical Chemistry **80:** 8694-8701

Clark SE, Williams RW, Meyerowitz EM (1997) The CLAVATA1 gene encodes a putative receptor kinase that controls shoot and floral meristem size in Arabidopsis. Cell **89:** 575-585

Crimmins DL, Mische SM, Denslow ND (2005) Chemical cleavage of proteins in solution. Curr Protoc Protein Sci **Chapter 11:** Unit 11.14

Czyzewicz N, Yue K, Beeckman T, De Smet I (2013) Message in a bottle: small signalling peptide outputs during growth and development. J Exp Bot **64:** 5281-5296

de Bang TC, Lundquist PK, Dai X, Boschiero C, Zhuang Z, Pant P, Torres-Jerez I, Roy S, Nogales J, Veerappan V, Dickstein R, Udvardi MK, Zhao PX, Scheible WR (2017) Genome-Wide Identification of Medicago Peptides Involved in Macronutrient Responses and Nodulation. Plant Physiol **175:** 1669-1689

deBlaquiere J, Burgess AW (1999) Affinity purification of plasma membranes. J Biomol Tech **10:** 64-71

DeCaprio J, Kohl TO (2019) Cross-Linking Antibodies to Beads Using Dimethyl Pimelimidate (DMP). Cold Spring Harb Protoc **2019**

Delay C, Chapman K, Taleski M, Wang Y, Tyagi S, Xiong Y, Imin N, Djordjevic MA (2019) CEP3 levels affect starvation-related growth responses of the primary root. J Exp Bot **70:** 4763-4774

Delay C, Imin N, Djordjevic MA (2013) CEP genes regulate root and shoot development in response to environmental cues and are specific to seed plants. J Exp Bot **64:** 5383-5394

Deng N, Forli S, He P, Perryman A, Wickstrom L, Vijayan RS, Tiefenbrunn T, Stout D, Gallicchio E, Olson AJ, Levy RM (2015) Distinguishing binders from false positives by free energy calculations: fragment screening against the flap site of HIV protease. J Phys Chem B **119:** 976-988

Dharmawardhana DP, Ellis BE, Carlson JE (1992) Characterization of vascular lignification in Arabidopsis thaliana. Canadian Journal of Botany **70:** 2238-2244

Djordjevic MA (2004) *Sinorhizobium meliloti* metabolism in the root nodule: a proteomic perspective. Proteomics **4:** 1859-1872

Djordjevic MA, Chen HC, Natera S, Van Noorden G, Menzel C, Taylor S, Renard C, Geiger O, Weiller GF (2003) A global analysis of protein expression profiles in *Sinorhizobium meliloti*: discovery of new genes for nodule occupancy and stress adaptation. Mol Plant Microbe Interact **16:** 508-524

Djordjevic MA, Mohd-Radzman NA, Imin N (2015) Small-peptide signals that control root nodule number, development, and symbiosis. J Exp Bot **66:** 5171-5181

Djordjevic MA, Oakes M, Li DX, Hwang CH, Hocart CH, Gresshoff PM (2007) The *glycine max* xylem sap and apoplast proteome. J Proteome Res **6:** 3771-3779

Djordjevic MA, Oakes M, Wong CE, Singh M, Bhalla P, Kusumawati L, Imin N (2011) Border sequences of *Medicago truncatula* CLE36 are specifically cleaved by endoproteases common to the extracellular fluids of Medicago and soybean. J Exp Bot **62:** 4649-4659

Donaldson L, Williams N (2018) Imaging and Spectroscopy of Natural Fluorophores in Pine Needles. Plants (Basel, Switzerland) **7:** 10

Du X, Li Y, Xia YL, Ai SM, Liang J, Sang P, Ji XL, Liu SQ (2016) Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. Int J Mol Sci **17**

Duff MR, Jr., Howell EE (2015) Thermodynamics and solvent linkage of macromolecule-ligand interactions. Methods **76:** 51-60

Endo M, Shimizu H, Araki T (2016) Rapid and simple isolation of vascular, epidermal and mesophyll cells from plant leaf tissue. Nature Protocols **11:** 1388-1395

Farrell IS, Toroney R, Hazen JL, Mehl RA, Chin JW (2005) Photo-cross-linking interacting proteins with a genetically encoded benzophenone. Nat Methods **2:** 377-384

Fischer F, Poetsch A (2006) Protein cleavage strategies for an improved analysis of

the membrane proteome. Proteome Science **4:** 2

Franssen HJ, Bisseling T (2001) Peptide signaling in plants. Proceedings of the National Academy of Sciences of the United States of America **98:** 12855-12856

Gautrat P, Laffont C, Frugier F (2020) Compact Root Architecture 2 Promotes Root Competence for Nodulation through the miR2111 Systemic Effector. Current Biology

Gerace E, Moazed D (2015) Affinity Pull-Down of Proteins Using Anti-FLAG M2 Agarose Beads. Methods Enzymol **559:** 99-110

Girolamo FD, Ponzi M, Crescenzi M, Alessandroni J, Guadagni F (2010) A simple and effective method to analyze membrane proteins by SDS-PAGE and MALDI mass spectrometry. Anticancer Res **30:** 1121-1129

Gotze M, Iacobucci C, Ihling CH, Sinz A (2019) A Simple Cross-Linking/Mass Spectrometry Workflow for Studying System-wide Protein Interactions. Anal Chem **91:** 10236-10244

Guillier C, Cacas JL, Recorbet G, Depretre N, Mounier A, Mongrand S, Simon-Plas F, Wipf D, Dumas-Gaudot E (2014) Direct purification of detergent-insoluble membranes from *Medicago truncatula* root microsomes: comparison between floatation and sedimentation. BMC Plant Biol **14:** 255

Hardy J (2017) Membrane damage is at the core of Alzheimer's disease. Lancet Neurol **16:** 342

Harris JI (1956) Effect of Urea on Trypsin and Alpha-Chymotrypsin. Nature **177:** 471-473

Hasegawa T, Sugeno N, Kikuchi A, Baba T, Aoki M (2017) Membrane Trafficking Illuminates a Path to Parkinson's Disease. Tohoku J Exp Med **242:** 63-76

Hattori N, Arano T, Hatano T, Mori A, Imai Y (2017) Mitochondrial-Associated Membranes in Parkinson's Disease. Adv Exp Med Biol **997:** 157-169

Hauer F, Gerle C, Fischer N, Oshima A, Shinzawa-Itoh K, Shimada S, Yokoyama K, Fujiyoshi Y, Stark H (2015) GraDeR: Membrane Protein Complex Preparation for Single-Particle Cryo-EM. Structure **23:** 1769-1775

Haupt C, Hofmann T, Wittig S, Kostmann S, Politis A, Schmidt C (2017) Combining Chemical Cross-linking and Mass Spectrometry of Intact Protein Complexes to Study the Architecture of Multi-subunit Protein Assemblies. J Vis Exp

Hirakawa Y, Shinohara H, Kondo Y, Inoue A, Nakanomyo I, Ogawa M, Sawa S, Ohashi-Ito K, Matsubayashi Y, Fukuda H (2008) Non-cell-autonomous control of vascular stem cell fate by a CLE peptide/receptor system. Proc Natl Acad Sci U S A **105:** 15208-15213

Hoffman EA, Frey BL, Smith LM, Auble DT (2015) Formaldehyde crosslinking: a tool for the study of chromatin complexes. The Journal of biological chemistry **290:** 26404-26411

Hoffman EA, Frey BL, Smith LM, Auble DT (2015) Formaldehyde Crosslinking: A Tool for the Study of Chromatin Complexes. Journal of Biological Chemistry **290:** 26404-26411

Hopkins AL, Groom CR (2002) The druggable genome. Nat Rev Drug Discov **1:** 727-730

Hothorn M, Belkhadir Y, Dreux M, Dabi T, Noel JP, Wilson IA, Chory J (2011) Structural basis of steroid hormone perception by the receptor kinase BRI1.

Nature **474:** 467

Hu C, Zhu Y, Cui Y, Cheng K, Liang W, Wei Z, Zhu M, Yin H, Zeng L, Xiao Y, Lv M, Yi J, Hou S, He K, Li J, Gou X (2018) A group of receptor kinases are essential for CLAVATA signalling to maintain stem cell homeostasis. Nature Plants **4:** 205-211

Huang T-K, Han C-L, Lin S-I, Chen Y-J, Tsai Y-C, Chen Y-R, Chen J-W, Lin W-Y, Chen P-M, Liu T-Y, Chen Y-S, Sun C-M, Chiou T-J (2013) Identification of Downstream Components of Ubiquitin-Conjugating Enzyme PHOSPHATE2 by Quantitative Membrane Proteomics in Arabidopsis Roots. The Plant Cell **25:** 4044-4060

Huault E, Laffont C, Wen J, Mysore KS, Ratet P, Duc G, Frugier F (2014) Local and systemic regulation of plant root system architecture and symbiotic nodulation by a receptor-like kinase. PLoS Genet **10:** e1004891

Imin N, Mohd-Radzman NA, Ogilvie HA, Djordjevic MA (2013) The peptide-encoding CEP1 gene modulates lateral root and nodule numbers in *Medicago truncatula*. J Exp Bot **64:** 5395-5409

Imin N, Patel N, Corcilius L, Payne RJ, Djordjevic MA (2018) CLE peptide tri-arabinosylation and peptide domain sequence composition are essential for SUNN-dependent autoregulation of nodulation in Medicago truncatula. New Phytol **218:** 73-80

Ishihama Y, Schmidt T, Rappsilber J, Mann M, Hartl FU, Kerner MJ, Frishman D (2008) Protein abundance profiling of the *Escherichia coli* cytosol. BMC Genomics **9:** 102

Jensen KH, Lee J, Bohr T, Bruus H, Holbrook NM, Zwieniecki MA (2011) Optimality of the Münch mechanism for translocation of sugars in plants. Journal of The Royal Society Interface **8:** 1155-1165

Johnsson N, Varshavsky A (1994) Split ubiquitin as a sensor of protein interactions in vivo. Proc Natl Acad Sci U S A **91:** 10340-10344

Kar UK, Simonian M, Whitelegge JP (2017) Integral membrane proteins: bottom-up, top-down and structural proteomics. Expert Rev Proteomics **14:** 715-723

Kauer JC, Erickson-Viitanen S, Wolfe HR, Jr., DeGrado WF (1986) p-Benzoyl-L-phenylalanine, a new photoreactive amino acid. Photolabeling of calmodulin with a synthetic calmodulin-binding peptide. J Biol Chem **261:** 10695-10700

Kemmerling B, Halter T, Mazzotta S, Mosher S, Nurnberger T (2011) A genome-wide survey for Arabidopsis leucine-rich repeat receptor kinases implicated in plant immunity. Front Plant Sci **2:** 88

Kennedy-Darling J, Smith LM (2014) Measuring the formaldehyde Protein-DNA cross-link reversal rate. Anal Chem **86:** 5678-5681

Kereszt A, Mergaert P, Montiel J, Endre G, Kondorosi É (2018) Impact of Plant Peptides on Symbiotic Nodule Development and Functioning. Frontiers in plant science **9:** 1026-1026

Klockenbusch C, Kast J (2010) Optimization of formaldehyde cross-linking for protein interaction analysis of non-tagged integrin beta1. J Biomed Biotechnol **2010:** 927585

Komatsu S, Konishi H, Hashimoto M (2007) The proteomics of plant cell membranes. J Exp Bot **58:** 103-112

Koziol J, Griffin N, Long F, Li Y, Latterich M, Schnitzer J (2013) On protein abundance distributions in complex mixtures. Proteome Science **11:** 5-5

Kulak NA, Geyer PE, Mann M (2017) Loss-less Nano-fractionator for High Sensitivity, High Coverage Proteomics. Mol Cell Proteomics **16:** 694-705
Kusumawati L, Imin N, Djordjevic MA (2008) Characterization of the secretome of suspension cultures of Medicago species reveals proteins important for defense and development. J Proteome Res **7:** 4508-4520
Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol **157:** 105-132
Laffont C, Huault E, Gautrat P, Endre G, Kalo P, Bourion V, Duc G, Frugier F (2019) Independent Regulation of Symbiotic Nodulation by the SUNN Negative and CRA2 Positive Systemic Pathways. Plant Physiol **180:** 559-570
Laganowsky A, Reading E, Hopper JTS, Robinson CV (2013) Mass spectrometry of intact membrane protein complexes. Nature Protocols **8:** 639-651
LaMontagne ED, Collins CA, Peck SC, Heese A (2016) Isolation of Microsomal Membrane Proteins from Arabidopsis thaliana. Curr Protoc Plant Biol **1:** 217-234
Lappano R, Maggiolini M (2011) G protein-coupled receptors: novel targets for drug discovery in cancer. Nat Rev Drug Discov **10:** 47-60
Lee J, Lei Z, Watson BS, Sumner LW (2013) Sub-cellular proteomics of *Medicago truncatula*. Front Plant Sci **4:** 112
Li C, Wu HM, Cheung AY (2016) FERONIA and Her Pals: Functions and Mechanisms. Plant physiology **171:** 2379-2392
Li J, Wen J, Lease KA, Doke JT, Tax FE, Walker JC (2002) BAK1, an Arabidopsis LRR Receptor-like Protein Kinase, Interacts with BRI1 and Modulates Brassinosteroid Signaling. Cell **110:** 213-222
Li Z, Wang Y, Huang J, Ahsan N, Biener G, Paprocki J, Thelen JJ, Raicu V, Zhao D (2017) Two SERK Receptor-Like Kinases Interact with EMS1 to Control Anther Cell Fate Determination. Plant Physiology **173:** 326
Link AJ, LaBaer J (2011) Trichloroacetic acid (TCA) precipitation of proteins. Cold Spring Harb Protoc **2011:** 993-994
Liu F, Lossl P, Rabbitts BM, Balaban RS, Heck AJR (2018) The interactome of intact mitochondria by cross-linking mass spectrometry provides evidence for coexisting respiratory supercomplexes. Mol Cell Proteomics **17:** 216-232
Liu F, Lössl P, Scheltema R, Viner R, Heck AJR (2017) Optimized fragmentation schemes and data analysis strategies for proteome-wide cross-link identification. Nature communications **8:** 15473-15473
Liu F, Rijkers DT, Post H, Heck AJ (2015) Proteome-wide profiling of protein assemblies by cross-linking mass spectrometry. Nat Methods **12:** 1179-1184
Liu J, Last RL (2015) A land plant-specific thylakoid membrane protein contributes to photosystem II maintenance in *Arabidopsis thaliana*. Plant J **82:** 731-743
Liu PL, Du L, Huang Y, Gao SM, Yu M (2017) Origin and diversification of leucine-rich repeat receptor-like protein kinase (LRR-RLK) genes in plants. BMC Evol Biol **17:** 47
Lohse M, Nagel A, Herter T, May P, Schroda M, Zrenner R, Tohge T, Fernie AR, Stitt M, Usadel B (2014) Mercator: a fast and simple web server for genome scale functional annotation of plant sequence data. Plant Cell Environ **37:** 1250-1258
Long R, Gao Y, Sun H, Zhang T, Li X, Li M, Sun Y, Kang J, Wang Z, Ding W, Yang Q

(2018) Quantitative proteomic analysis using iTRAQ to identify salt-responsive proteins during the germination stage of two Medicago species. Sci Rep **8:** 9553

Lucas WJ, Groover A, Lichtenberger R, Furuta K, Yadav SR, Helariutta Y, He XQ, Fukuda H, Kang J, Brady SM, Patrick JW, Sperry J, Yoshida A, Lopez-Millan AF, Grusak MA, Kachroo P (2013) The plant vascular system: evolution, development and functions. J Integr Plant Biol **55:** 294-388

Lunn G, Sansone EB (1985) Destruction of cyanogen bromide and inorganic cyanides. Anal Biochem **147:** 245-250

Magdeldin S, Yoshida Y, Li H, Maeda Y, Yokoyama M, Enany S, Zhang Y, Xu B, Fujinaka H, Yaoita E, Sasaki S, Yamamoto T (2012) Murine colon proteome and characterization of the protein pathways. BioData mining **5:** 11

Man J, Gallagher JP, Bartlett M (2020) Structural evolution drives diversification of the large LRR-RLK gene family. New Phytol

Marx H, Minogue CE, Jayaraman D, Richards AL, Kwiecien NW, Siahpirani AF, Rajasekar S, Maeda J, Garcia K, Del Valle-Echevarria AR, Volkening JD, Westphall MS, Roy S, Sussman MR, Ane JM, Coon JJ (2016) A proteomic atlas of the legume *Medicago truncatula* and its nitrogen-fixing endosymbiont *Sinorhizobium meliloti*. Nat Biotechnol **34:** 1198-1205

Masters SC (2004) Co-immunoprecipitation from transfected cells. Methods Mol Biol **261:** 337-350

Matsubayashi Y (2018) Exploring peptide hormones in plants: identification of four peptide hormone-receptor pairs and two post-translational modification enzymes. Proceedings of the Japan Academy. Series B, Physical and biological sciences **94:** 59-74

Matsubayashi Y, Sakagami Y (2006) Peptide hormones in plants. Annu Rev Plant Biol **57:** 649-674

McFedries A, Schwaid A, Saghatelian A (2013) Methods for the elucidation of protein-small molecule interactions. Chem Biol **20:** 667-673

McKee JA, Kumar S, Ecelbarger CA, Fernandez-Llama P, Terris J, Knepper MA (2000) Detection of Na(+) transporter proteins in urine. J Am Soc Nephrol **11:** 2128-2132

McWhite CD, Papoulas O, Drew K, Cox RM, June V, Dong OX, Kwon T, Wan C, Salmi ML, Roux SJ, Browning KS, Chen ZJ, Ronald PC, Marcotte EM (2020) A Pan-plant Protein Complex Map Reveals Deep Conservation and Novel Assemblies. Cell

Melnyk CW, Gabel A, Hardcastle TJ, Robinson S, Miyashima S, Grosse I, Meyerowitz EM (2018) Transcriptome dynamics at Arabidopsis graft junctions reveal an intertissue recognition mechanism that activates vascular regeneration. Proceedings of the National Academy of Sciences **115:** E2447-E2456

Meng X, Chen X, Mang H, Liu C, Yu X, Gao X, Torii Keiko U, He P, Shan L (2015) Differential Function of Arabidopsis SERK Family Receptor-like Kinases in Stomatal Patterning. Current Biology **25:** 2361-2372

Meng X, Zhou J, Tang J, Li B, de Oliveira MVV, Chai J, He P, Shan L (2016) Ligand-Induced Receptor-like Kinase Complex Regulates Floral Organ Abscission in Arabidopsis. Cell Rep **14:** 1330-1338

Metz B, Kersten GF, Hoogerhout P, Brugghe HF, Timmermans HA, de Jong A, Meiring H, ten Hove J, Hennink WE, Crommelin DJ, Jiskoot W (2004) Identification of formaldehyde-induced modifications in proteins: reactions with model peptides. J Biol Chem **279:** 6235-6243

Millioni R, Tolin S, Puricelli L, Sbrignadello S, Fadini GP, Tessari P, Arrigoni G (2011) High abundance proteins depletion vs low abundance proteins enrichment: comparison of methods to reduce the plasma proteome complexity. PloS One **6:** e19603-e19603

Mirza SP, Halligan BD, Greene AS, Olivier M (2007) Improved method for the analysis of membrane proteins by mass spectrometry. Physiol Genomics **30:** 89-94

Mitra SK, Chen R, Dhandaydham M, Wang X, Blackburn RK, Kota U, Goshe MB, Schwartz D, Huber SC, Clouse SD (2015) An autophosphorylation site database for leucine-rich repeat receptor-like kinases in *Arabidopsis thaliana*. Plant J **82:** 1042-1060

Miura K (2018) An Overview of Current Methods to Confirm Protein-Protein Interactions. Protein and peptide letters **25:** 728-733

Miura K, Imaki J (2008) Molecular cloning of Ebitein1: a novel extracellular signal-regulated kinase 2-binding protein in testis. Biochem Biophys Res Commun **368:** 336-342

Miyazaki R, Akiyama Y, Mori H (2020) A photo-cross-linking approach to monitor protein dynamics in living cells. Biochimica et Biophysica Acta (BBA) - General Subjects **1864**

Mohd-Radzman NA, Binos S, Truong TT, Imin N, Mariani M, Djordjevic MA (2015) Novel MtCEP1 peptides produced in vivo differentially regulate root development in *Medicago truncatula*. Journal of Experimental Botany **66:** 5289-5300

Mohd-Radzman NA, Binos S, Truong TT, Imin N, Mariani M, Djordjevic MA (2015) Novel MtCEP1 peptides produced *in vivo* differentially regulate root development in *Medicago truncatula*. Journal of Experimental Botany **66:** 5289-5300

Mohd-Radzman NA, Laffont C, Ivanovici A, Patel N, Reid D, Stougaard J, Frugier F, Imin N, Djordjevic MA (2016) Different Pathways Act Downstream of the CEP Peptide Receptor CRA2 to Regulate Lateral Root and Nodule Development. Plant Physiol **171:** 2536-2548

Mohd-Radzman NA, Laffont C, Ivanovici A, Patel N, Reid D, Stougaard J, Frugier F, Imin N, Djordjevic MA (2016) Different Pathways Act Downstream of the CEP Peptide Receptor CRA2 to Regulate Lateral Root and Nodule Development. Plant physiology **171:** 2536-2548

Murphy E, Smith S, De Smet I (2012) Small signaling peptides in Arabidopsis development: how cells communicate over a short distance. Plant Cell **24:** 3198-3217

Natera SH, Guerreiro N, Djordjevic MA (2000) Proteome analysis of differentially displayed proteins as a tool for the investigation of symbiosis. Mol Plant Microbe Interact **13:** 995-1009

Newby ZER, O'Connel JD, Gruswitz F, Hays FA, Harries WEC, Harwood IM, Ho JD, Lee JK, Savage DF, Miercke LJW, Stroud R (2009) A general protocol for the

crystallization of membrane proteins for X-ray structural investigation. Nature Protocols **4:** 619-637

Nguyen TT, Sabat G, Sussman MR (2018) In vivo cross-linking supports a head-to-tail mechanism for regulation of the plant plasma membrane P-type H(+)-ATPase. The Journal of biological chemistry **293:** 17095-17106

Nilsen TW (2014) Preparation of cross-linked cellular extracts with formaldehyde. Cold Spring Harb Protoc **2014:** 1001-1003

Notaguchi M, Okamoto S (2015) Dynamics of long-distance signaling via plant vascular tissues. Frontiers in Plant Science **6:** 161

Ogawa M, Shinohara H, Sakagami Y, Matsubayashi Y (2008) Arabidopsis CLV3 peptide directly binds CLV1 ectodomain. Science **319:** 294

Ogilvie HA, Imin N, Djordjevic MA (2014) Diversification of the C-TERMINALLY ENCODED PEPTIDE (CEP) gene family in angiosperms, and evolution of plant-family specific CEP genes. BMC Genomics **15:** 870

Ohad N, Shichrur K, Yalovsky S (2007) The Analysis of Protein-Protein Interactions in Plants by Bimolecular Fluorescence Complementation. Plant Physiology **145:** 1090-1099

Ohkubo Y, Tanaka M, Tabata R, Ogawa-Ohnishi M, Matsubayashi Y (2017) Shoot-to-root mobile polypeptides involved in systemic regulation of nitrogen acquisition. Nat Plants **3:** 17029

Ohyama K, Ogawa M, Matsubayashi Y (2008) Identification of a biologically active, small, secreted peptide in Arabidopsis by in silico gene screening, followed by LC-MS-based structure analysis. The Plant Journal **55:** 152-160

Okamoto S, Suzuki T, Kawaguchi M, Higashiyama T, Matsubayashi Y (2015) A comprehensive strategy for identifying long-distance mobile peptides in xylem sap. Plant J **84:** 611-620

Okumura M, Inoue S, Kuwata K, Kinoshita T (2016) Photosynthesis Activates Plasma Membrane H+-ATPase via Sugar Accumulation. Plant Physiol **171:** 580-589

Omasits U, Ahrens CH, Muller S, Wollscheid B (2014) Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics **30:** 884-886

Osakabe Y, Maruyama K, Seki M, Satou M, Shinozaki K, Yamaguchi-Shinozaki K (2005) Leucine-Rich Repeat Receptor-Like Kinase1 Is a Key Membrane-Bound Regulator of Abscisic Acid Early Signaling in Arabidopsis. The Plant Cell **17:** 1105-1119

Patching SG (2014) Surface plasmon resonance spectroscopy for characterisation of membrane protein-ligand interactions and its potential for drug discovery. Biochim Biophys Acta **1838:** 43-55

Patel N, Mohd-Radzman NA, Corcilius L, Crossett B, Connolly A, Cordwell SJ, Ivanovici A, Taylor K, Williams J, Binos S, Mariani M, Payne RJ, Djordjevic MA (2018) Diverse Peptide Hormones Affecting Root Growth Identified in the *Medicago truncatula* Secreted Peptidome. Mol Cell Proteomics **17:** 160-174

Pellizzaro A, Clochard T, Planchet E, Limami AM, Morere-Le Paven MC (2015) Identification and molecular characterization of *Medicago truncatula* NRT2 and NAR2 families. Physiol Plant **154:** 256-269

Piya S, Shrestha SK, Binder B, Stewart CN, Jr., Hewezi T (2014) Protein-protein interaction and gene co-expression maps of ARFs and Aux/IAAs in

Arabidopsis. Front Plant Sci **5:** 744

Quach TT, Li N, Richards DP, Zheng J, Keller BO, Li L (2003) Development and applications of in-gel CNBr/tryptic digestion combined with mass spectrometry for the analysis of membrane proteins. J Proteome Res **2:** 543-552

Rawlings AE (2016) Membrane proteins: always an insoluble problem? Biochemical Society Transactions **44:** 790-795

Roberts I, Smith S, De Rybel B, Van Den Broeke J, Smet W, De Cokere S, Mispelaere M, De Smet I, Beeckman T (2013) The CEP family in land plants: evolutionary analyses, expression studies, and role in Arabidopsis shoot development. Journal of Experimental Botany **64:** 5371-5381

Roberts I, Smith S, De Rybel B, Van Den Broeke J, Smet W, De Cokere S, Mispelaere M, De Smet I, Beeckman T (2013) The CEP family in land plants: evolutionary analyses, expression studies, and role in Arabidopsis shoot development. J Exp Bot **64:** 5371-5381

Roux M, Schwessinger B, Albrecht C, Chinchilla D, Jones A, Holton N, Malinovsky FG, Tor M, de Vries S, Zipfel C (2011) The Arabidopsis leucine-rich repeat receptor-like kinases BAK1/SERK3 and BKK1/SERK4 are required for innate immunity to hemibiotrophic and biotrophic pathogens. Plant Cell **23:** 2440-2455

Roy S, Lundquist P, Udvardi M, Scheible W-R (2018) Small and Mighty: Peptide hormones in plant biology Plant Cell**:** - tpc.118.tt0718

Santiago J, Brandt B, Wildhagen M, Hohmann U, Hothorn LA, Butenko MA, Hothorn M (2016) Mechanistic insight into a peptide hormone signaling complex mediating floral organ abscission. eLife **5:** e15075

Savage JA, Zwieniecki MA, Holbrook NM (2013) Phloem Transport Velocity Varies over Time and among Vascular Bundles during Early Cucumber Seedling Development. Plant Physiology **163:** 1409-1418

Saveliev S, Bratz M, Zubarev R, Szapacs M, Budamgunta H, Urh M (2013) Trypsin/Lys-C protease mix for enhanced protein mass spectrometry analysis. Nature Methods **10:** i-ii

Scheltema RA, Hauschild J-P, Lange O, Hornburg D, Denisov E, Damoc E, Kuehn A, Makarov A, Mann M (2014) The Q Exactive HF, a Benchtop mass spectrometer with a pre-filter, high-performance quadrupole and an ultra-high-field Orbitrap analyzer. Molecular & cellular proteomics : MCP **13:** 3698-3708

Schey KL, Grey AC, Nicklay JJ (2013) Mass spectrometry of membrane proteins: a focus on aquaporins. Biochemistry **52:** 3807-3817

Seddon AM, Curnow P, Booth PJ (2004) Membrane proteins, lipids and detergents: not just a soap opera. Biochim Biophys Acta **1666:** 105-117

Shabala S, White RG, Djordjevic MA, Ruan Y-L, Mathesius U (2016) Root-to-shoot signalling: integration of diverse molecules, pathways and functions. Functional Plant Biology **43:** 87-104

Shiu S-H, Bleecker AB (2001) Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. Proceedings of the National Academy of Sciences **98:** 10763-10768

Siligardi G, Hussain R, Patching SG, Phillips-Jones MK (2014) Ligand- and drug-

binding studies of membrane proteins revealed through circular dichroism spectroscopy. Biochim Biophys Acta **1838:** 34-42

Sinz A, Arlt C, Chorev D, Sharon M (2015) Chemical cross-linking and native mass spectrometry: A fruitful combination for structural biology. Protein science : a publication of the Protein Society **24:** 1193-1209

Smith A, George Coupland, Liam Dolan, Nicholas Harberd, Jonathan Jones, Cathie Martin, Robert Sablowski, Amey A (2009) Plant Biology. Garland Science

Smith BJ (1988) Chemical cleavage of proteins. Methods Mol Biol **3:** 71-88

Snow JW (2016) A fluorescent method for visualization of Nosema infection in whole-mount honey bee tissues. J Invertebr Pathol **135:** 10-14

Song W, Liu L, Wang J, Wu Z, Zhang H, Tang J, Lin G, Wang Y, Wen X, Li W, Han Z, Guo H, Chai J (2016) Signature motif-guided identification of receptors for peptide hormones essential for root meristem growth. Cell Res **26:** 674-685

Song W, Wang B, Li X, Wei J, Chen L, Zhang D, Zhang W, Li R (2015) Identification of Immune Related LRR-Containing Genes in Maize (Zea mays L.) by Genome-Wide Sequence Analysis. International Journal of Genomics **2015:** 231358

Sonnhammer EL, von Heijne G, Krogh A (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol **6:** 175-182

Sperschneider J, Catanzariti AM, DeBoer K, Petre B, Gardiner DM, Singh KB, Dodds PN, Taylor JM (2017) LOCALIZER: subcellular localization prediction of both plant and effector proteins in the plant cell. Sci Rep **7:** 44598

Stahl Y, Simon R (2012) Peptides and receptors controlling root development. Philosophical transactions of the Royal Society of London. Series B, Biological sciences **367:** 1453-1460

Stynen B, Tournu H, Tavernier J, Van Dijck P (2012) Diversity in genetic in vivo methods for protein-protein interaction studies: from the yeast two-hybrid system to the mammalian split-luciferase system. Microbiol Mol Biol Rev **76:** 331-382

Sun Y, Han Z, Tang J, Hu Z, Chai C, Zhou B, Chai J (2013) Structure reveals that BAK1 as a co-receptor recognizes the BRI1-bound brassinolide. Cell Research **23:** 1326

Sun Y, Li L, Macho AP, Han Z, Hu Z, Zipfel C, Zhou J-M, Chai J (2013) Structural Basis for flg22-Induced Activation of the Arabidopsis FLS2-BAK1 Immune Complex. Science **342:** 624

Sun Y, Rombola C, Jyothikumar V, Periasamy A (2013) Forster resonance energy transfer microscopy and spectroscopy for localizing protein-protein interactions in living cells. Cytometry A **83:** 780-793

Sutherland BW, Toews J, Kast J (2008) Utility of formaldehyde cross-linking and mass spectrometry in the study of protein-protein interactions. J Mass Spectrom **43:** 699-715

Szablewski L (2017) Glucose Transporters in Brain: In Health and in Alzheimer's Disease. J Alzheimers Dis **55:** 1307-1320

Tabata R, Kamiya T, Shigenobu S, Yamaguchi K, Yamada M, Hasebe M, Fujiwara T, Sawa S (2013) Identification of an EMS-induced causal mutation in a gene required for boron-mediated root development by low-coverage genome re-sequencing in Arabidopsis. Plant Signal Behav **8:** e22534

Tabata R, Sumida K, Yoshii T, Ohyama K, Shinohara H, Matsubayashi Y (2014) Perception of root-derived peptides by shoot LRR-RKs mediates systemic N-demand signaling. Science **346:** 343-346

Tager HS, Steiner DF (1974) Peptide hormones. Annu Rev Biochem **43:** 509-538

Taleski M, Chapman K, Imin N, Djordjevic MA, Groszmann M (2020) The peptide hormone receptor CEPR1 functions in the reproductive tissue to control seed size and yield. Plant Physiology**:** pp.00172.02020

Taleski M, Imin N, Djordjevic MA (2018) CEP peptide hormones: key players in orchestrating nitrogen-demand signalling, root nodulation, and lateral root development. J Exp Bot **69:** 1829-1836

Tautermann CS (2014) GPCR structures in drug design, emerging opportunities with new structures. Bioorg Med Chem Lett **24:** 4073-4079

Tavormina P, De Coninck B, Nikonorova N, De Smet I, Cammue BPA (2015) The Plant Peptidome: An Expanding Repertoire of Structural Features and Biological Functions. The Plant Cell **27:** 2095-2118

Ten Hove CA, Bochdanovits Z, Jansweijer VMA, Koning FG, Berke L, Sanchez-Perez GF, Scheres B, Heidstra R (2011) Probing the roles of LRR RLK genes in Arabidopsis thaliana roots using a custom T-DNA insertion set. Plant Mol Biol **76:** 69-83

ten Hove CA, Bochdanovits Z, Jansweijer VMA, Koning FG, Berke L, Sanchez-Perez GF, Scheres B, Heidstra R (2011) Probing the roles of LRR RLK genes in *Arabidopsis thaliana* roots using a custom T-DNA insertion set. Plant Molecular Biology **76:** 69-83

Thavarajah R, Mudimbaimannar VK, Elizabeth J, Rao UK, Ranganathan K (2012) Chemical and physical basics of routine formaldehyde fixation. Journal of oral and maxillofacial pathology : JOMFP **16:** 400-405

Thimm O, Blasing O, Gibon Y, Nagel A, Meyer S, Kruger P, Selbig J, Muller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J **37:** 914-939

Toth J, Biggin MD (2000) The specificity of protein-DNA crosslinking by formaldehyde: in vitro and in drosophila embryos. Nucleic acids research **28:** e4-e4

van der Knaap E, Song WY, Ruan DL, Sauter M, Ronald PC, Kende H (1999) Expression of a gibberellin-induced leucine-rich repeat receptor-like protein kinase in deepwater rice and its interaction with kinase-associated protein phosphatase. Plant Physiol **120:** 559-570

Vit O, Petrak J (2017) Integral membrane proteins in proteomics. How to break open the black box? J Proteomics **153:** 8-20

Waduwara CI, Walcott SE, Peterson CA (2008) Suberin lamellae of the onion root endodermis: their pattern of development and continuity. Botany **86:** 623-632

Wang J, Li H, Han Z, Zhang H, Wang T, Lin G, Chang J, Yang W, Chai J (2015) Allosteric receptor activation by the plant peptide hormone phytosulfokine. Nature **525:** 265

Willemse MTM (1989) Cell Wall Autofluorescence. *In* A Chesson, ER Ørskov, eds, Physico-Chemical Characterisation of Plant Residues for Industrial and Feed

Use. Springer Netherlands, Dordrecht, pp 50-57
Wilson JL, Scott IM, McMurry JL (2010) Optical biosensing: Kinetics of protein A-IGG binding using biolayer interferometry. Biochem Mol Biol Educ **38:** 400-407
Wong S-H, King C-Y (2015) Amino Acid Proximities in Two Sup35 Prion Strains Revealed by Chemical Cross-linking. The Journal of biological chemistry **290:** 25062-25071
Wong SH, King CY (2015) Amino Acid Proximities in Two Sup35 Prion Strains Revealed by Chemical Cross-linking. J Biol Chem **290:** 25062-25071
Xiao Y, Stegmann M, Han Z, DeFalco TA, Parys K, Xu L, Belkhadir Y, Zipfel C, Chai J (2019) Mechanisms of RALF peptide perception by a heterotypic receptor complex. Nature **572:** 270-274
Xing S, Wallmeroth N, Berendzen KW, Grefen C (2016) Techniques for the Analysis of Protein-Protein Interactions in Vivo. Plant physiology **171:** 727-758
Yeung Y-G, Stanley ER (2010) Rapid detergent removal from peptide samples with ethyl acetate for mass spectrometry analysis. Current Protocols In Protein Science **Chapter 16:** Unit-16.12
Yoshida S, Uemura M, Niki T, Sakai A, Gusta LV (1983) Partition of Membrane Particles in Aqueous Two-Polymer Phase System and Its Practical Use for Purification of Plasma Membranes from Plants. Plant Physiology **72:** 105-114
Zhang H, Lin X, Han Z, Qu L-J, Chai J (2016) Crystal structure of PXY-TDIF complex reveals a conserved recognition mechanism among CLE peptide-receptor pairs. Cell Research **26:** 543
Zhang K, McKinlay C, Hocart CH, Djordjevic MA (2006) The *Medicago truncatula* small protein proteome and peptidome. J Proteome Res **5:** 3355-3367
Zhang Y, Song G, Lal NK, Nagalakshmi U, Li Y, Zheng W, Huang P-j, Branon TC, Ting AY, Walley JW, Dinesh-Kumar SP (2019) TurboID-based proximity labeling reveals that UBR7 is a regulator of N NLR immune receptor-mediated immunity. Nature Communications **10:** 3252
Zhao Q, Liang Y, Yuan H, Sui Z, Wu Q, Liang Z, Zhang L, Zhang Y (2013) Biphasic microreactor for efficient membrane protein pretreatment with a combination of formic acid assisted solubilization, on-column pH adjustment, reduction, alkylation, and tryptic digestion. Anal Chem **85:** 8507-8512
Zheng S, Doucette AA (2016) Preventing N- and O-formylation of proteins when incubated in concentrated formic acid. Proteomics **16:** 1059-1068
Zhu F, Isaacs NW, Hecht L, Barron LD (2005) Raman optical activity: a tool for protein structure analysis. Structure **13:** 1409-1419