Characterisation of CCaMK in symbiosis signalling

Joanna Kate Darch Harrison

Thesis submitted to the University of East Anglia for the degree of Doctor of Philosophy

Bornemann Lab

Department of Biological Chemistry

John Innes Centre

Norwich

September 2015

2

Abstract

Many plants form important and beneficial symbioses with microbes including arbuscular mycorrhization and the legume-rhizobial interaction. The establishment of both these symbioses involves common genes which encode the Sym pathway proteins, of which calcium- and calmodulin-dependent protein kinase (CCaMK) is a central component. CCaMK has been genetically positioned downstream of a nuclear calcium spiking signal and is widely thought to detect and transduce this signal. In the presence of calcium CCaMK undergoes a conformational change, but the nature of this was previously unknown. Here I show that this change likely corresponds to an elongation of the visinin-like domain and this exposes a hydrophobic patch in the full-length protein. This hydrophobic patch could facilitate the binding of CCaMK to target proteins. Calcium-binding to the visinin-like domain stimulates autophosphorylation, in particular at T271 (in *Medicago truncatula* CCaMK). However, the molecular mechanism by which this occurs was not well-established. By contrast, the mechanism of autophosphorylation of CaMKII, a similar animal protein, has been determined to be intra-oligomeric and inter-subunit. Work presented here shows for the first time that MBP-CCaMK forms an oligomer of 16-18 subunits when purified from *E. coli* and experimental evidence is most consistent with an intra-oligomeric, inter-subunit mechanism of autophosphorylation. This is expected to facilitate the stabilisation of a robust "off-state" of the protein. Finally, CCaMK was assessed *in planta*. Preliminary data suggested the presence of potential alternative splice forms of CCaMK, and I undertook an investigation to determine if these were present during symbiosis at the RNA and protein level. Very low levels of splice variants were detected at the RNA level only, suggesting that they do not play a key role during symbiosis. A multi-faceted study is therefore presented, and this provides deeper insights into the function of CCaMK in the Sym pathway.

List of contents

5

6

List of tables

List of figures

9

Acknowledgements

I would like to warmly thank the many people who have helped me during my four years at JIC. First and foremost, my sincere thanks go to my primary supervisor, Dr Stephen Bornemann. Steph has been an extremely supportive supervisor, who has always made time for discussions with me. I have learnt so much from him over the last four years, and he has provided excellent feedback on all aspects of my project including my experiments, presentations and written work. I couldn't have asked for a better PhD supervisor and I can't thank him enough. His encouragement was definitely welcomed when I felt like aspects of my project were not going to plan! I would also like to thank my other supervisors, Professor Giles Oldroyd and Professor Richard Morris. Sincere thanks go to them for the many discussions regarding my project and their many helpful suggestions.

Specific experimental thanks go to several people who collaborated on aspects of this project. These include: Dr Gerhard Saalbach for performing mass spectrometry for me, Dr Patrick Gunning and Andrew Kirby for the atomic force microscopy at IFR, and Dr Dmitri Svergun for the analysis of small angle X-ray scattering data.

There are also many people to thank at JIC for their help and support. Firstly, to those who also work(ed) on CCaMK. I would like to thank previous members of the Bornemann lab, Dr David Swainsbury and particularly Dr Liang Zhou, for their help setting up the project. I also owe a huge debt of gratitude to former Oldroyd lab members Dr Ben Miller and Dr Sylvia Singh, both of whom have been hugely supportive and helpful. Thank you both for letting me ask lots of silly questions! Thank you also to Ben and Professor Allan Downie for their feedback on my thesis. More generally, thanks go to the other members of the Bornemann lab, particularly Dr Farzana Miah, for making it such a good working environment. I also really enjoyed being able to call two labs home, so I would like to thank the Oldroyd lab for welcoming me too. It was really nice to be able to work amongst so many people who find symbiosis fascinating.

I would also like to thank my fellow PhD students at JIC for making this such a fun place to work. It was a nice escape to serve as the student voice committee president for a year, and it was great to be surrounded by so many people who were as keen as I was to get involved. Particular thanks go to the other PhD students in my year, especially Dr Rachel Goddard and Dr Christopher Judge who have definitely helped me to get through this!

List of abbreviations

ABA – Abscisic acid

AFM – Atomic force microscopy

AIM – Auto-induction medium

AI-VLD - Auto-inhibitory (domain) and visinin-like domain

ANS - 8-Anilino-1-naphthalenesulfonic acid

ATP - Adenosine 5'-triphosphate

BLAST – Basic local alignment search tool

BNM – Buffered nodulation medium

CaM – Calmodulin

Cam - Chloramphenicol

CaMKI – Ca^{2+}/CaM-dependent protein kinase I

CaMKII – Ca^{2+}/CaM-dependent protein kinase II

Carb - Carbenicillin

HEPES **-** 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid

CBB – Coomassie brilliant blue

CBL – Calcineurin B-like (protein)

CCaMK – Calcium- and calmodulin-dependent protein kinase

cDNA – Complementary deoxyribonucleic acid

CDPK – Calcium-dependent protein kinase

CIPK – CBL-interacting protein kinase

CML – CaM-like protein

CRK – CDPK-related kinase

DEAE - Diethylaminoethyl

DLS – Dynamic light scattering

DMI – Does not make infections

DNA – Deoxyribonucleic acid

Dpi – Days post inoculation

dsRED - *Discosoma* sp. red fluorescent protein

DTT - DL-Dithiothreitol

DWA – Distilled water agar

EDTA - Ethylenediaminetetraacetic acid

EF – Elongation factor (hand)

gDNA – Genomic deoxyribonucleic acid

GFP – Green fluorescent protein

IMAC – Immobilised metal affinity chromatography

IPD3 – Interacting protein of DMI3

IPTG - Isopropyl-beta-D-thiogalactopyranoside

Kbp – Kilobase pair

K_d – Dissociation constant

kDa - Kilodalton

L – Lysogeny broth - Lennox

LB – Lysogeny broth

LCO - Lipochitooligosacchararide

LRR – Leucine-rich repeat

MBP – Maltose binding protein

MOD FP – Modified Fahraeus Medium

MYA – Million years ago

OD – Optical density

PAGE – Polyacrylamide gel electrophoresis

PCR – Polymerase chain reaction

PEG – Poly(ethylene glycol)

Rif - Rifampicin

RNA – Ribonucleic acid

RT-PCR – Reverse-transcription polymerase chain reaction

SAXS – Small angle X-ray scattering

SBP – Streptavidin-binding peptide (tag)

SDS – Sodium dodecyl sulfate

SEC - Size exclusion chromatography

SOC – Super orbital broth with catabolite repression

Strep - Streptomycin

Sym – Common symbiosis pathway

Tris - Tris(hydroxymethyl)aminomethane

TY – Rhizobium complete medium

v/v – Volume/volume

VLD – Visinin-like domain

w/v – Weight/volume

WT – Wild type

Chapter 1: Introduction

This thesis covers work regarding the *Medicago truncatula* protein calcium- and calmodulin-dependent protein kinase (CCaMK). This protein is a key component of the common symbiosis signalling pathway necessary for the establishment of both nodulation in legumes and arbuscular mycorrhization. This general introduction places the protein into context by describing root symbioses, the common Sym pathway, calcium signalling and binding proteins in plants, and previous literature relating to CCaMK. In addition, background information on an animal protein which shows significant sequence similarity to CCaMK, CaMKII, is detailed. A brief discussion of alternative splicing concludes the introduction, which relates to the experimental work in Chapter 6. Finally, the aims of this thesis project are detailed in section 1.9.

1.1 Introduction to root symbioses

Many plants form important and beneficial plant-microbe interactions known as symbioses. The most ancient form of plant-root symbiosis is the arbuscular mycorrhiza, formed between 70-90% of all land plants and the monophyletic fungal phylem, *Glomeromycota* (Parniske, 2008). The more recently evolved plant symbiotic interaction is the nitrogen-fixing root nodulation symbiosis. This symbiosis is formed between some plants in the orders *Fabales, Fagales, Cucurbitalis,* and *Rosales* and the bacterial family *Rhizobiaceae* or the genus *Frankia* (Popp and Ott, 2011). These plant species form a monophyletic nitrogen-fixing clade, in which the root nodulation symbiosis is thought to have evolved multiple times (Delaux *et al.*, 2015). In particular, the nitrogen-fixation interaction formed between plants of the *Fabaceae* family and rhizobia bacteria has been extensively studied, and is termed the legume-rhizobial symbiosis.

Plants benefit from both mycorrhization and nodulation. From their fungal partners they gain enhanced water and nutrient uptake (most notably phosphate but also nitrogen, micro and macro-nutrients (Parniske, 2008), (Hodge *et al.*, 2001)) and protection from root diseases (Bonfante and Requena, 2011). From bacterial symbionts, plants crucially gain nitrogen in the form of ammonia that can be easily utilised (Oldroyd *et al.*, 2011). In return, their fungal and bacterial partners gain carbon (Harrison, 2005) and energy in the form of di-carboxylic acids (Lodwig *et al.*, 2003), respectively. Therefore mycorrhization and nodulation are agriculturally and economically important. Mycorrhization plays a key role in the improvement of plant growth because it provides phosphorus which is required in

large quantities by plants. Furthermore, since mycorrhization is so widespread amongst plants, it also plays a key role in global carbon cycling (Harrison, 2005).

In turn, despite its restriction to a single clade of plants, the nitrogen-fixing root nodulation symbiosis impacts the global nitrogen cycle. Moreover, nitrogen is essential for plant growth and the presence of nitrogen-fixing bacteria allows the plant to access the atmospheric pool of nitrogen which it otherwise cannot use. Crop plants which cannot form nodules therefore require the application of large quantities of nitrogenous fertiliser for optimum yield. Since nitrogen-availability is therefore growth-limiting, if nitrogen-fixation could be introduced to key crop plants, for example cereals, this would reduce fertiliser-dependency and potentially greatly increase yield. Work has begun to engineer the nitrogen-fixing symbiosis in non-legume crops and this will become extremely important in the future as one of the ways to feed the ever-increasing world population (Rogers and Oldroyd, 2014), (Delaux *et al.*, 2015), (Charpentier and Oldroyd, 2010). Therefore, increasing our understanding of symbiotic interactions is an important goal.

1.1.1 Mycorrhization

Mycorrhization is thought to have originated more than 400 million years ago (MYA) in the Early Devonian period (Remy *et al.*, 1994). It has been theorised that the development of this mutualism played a key role in colonisation of land by green plants because of the widespread distribution of mycorrhization, the fossil record and the presence of highly conserved key genes required for mycorrhization in the common ancestor of land plants (Wang *et al.*, 2010). There are two types of mycorrhiza: ecto- and endo-mycorrhiza. Ecto-mycorrhization is characterised by the presence of the fungal partner outside of the plant cell, and has evolved more recently, approximately 130-180 MYA. This symbiosis is observed for example in boreal and temperate forest trees, and assists the plant in the uptake of nitrogen, phosphorus and magnesium (Venkateshwaran *et al.*, 2013). The more common type of mycorrhization is the endo-mycorrhiza, where the fungus penetrates the plant cell. Arbuscular mycorrhization is a type of endo-mycorrhization which is characterised by the formation of fungal tree-like projections within plant cells. This is likely to be the most common terrestrial symbiosis (Parniske, 2008).

Arbuscular mycorrhizal development begins when plant roots exude strigolactones, which are plant hormones that stimulates hyphal branching, spore germination and mitochondrial activity in the fungal partner (Akiyama *et al.*, 2005), (Besserer *et al.*, 2006). Strigolactones break down quickly when released, and this is theorised to indicate the proximity of the root to the fungal partner. This is the pre-symbiotic phase. Arbuscular mycorrhizal fungi are obligate biotrophs, and therefore to maximise their chance of forming a symbiosis they release diffusible compounds called Myc factors. The structure of the Myc factors has eluded scientists for many years. However, the structure of several Myc factor has been elucidated recently. This is discussed in section 1.2.1. Myc factors induce plant responses including symbiotic gene expression, calcium oscillations and root growth and branching (Kosuta *et al.*, 2008), (Maillet *et al.*, 2011). The fungus forms a hyphopodium developed from mature hyphae and this attaches to the root surface. The plant responds to this by migration of the nucleus towards the anticipated fungal entry point and by the formation of the pre-penetration apparatus. This consists of a column of cytoplasm across the host-cell vacuole and this guides the growth of the fungal hyphae through the epidermis to the cortex. Here, the fungus branches in the apoplast before re-entering the inner cortical cells where it forms developed arbuscules (Figure 1.1) (Parniske, 2008).
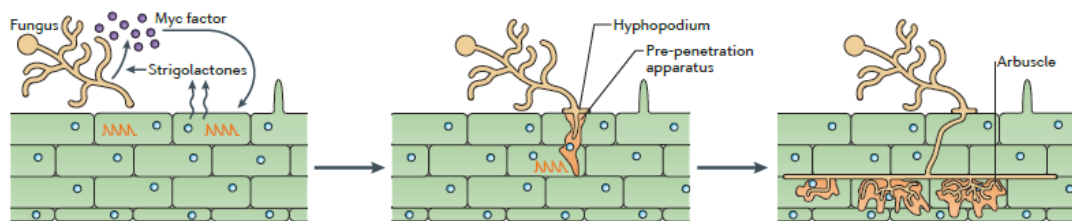
Figure 1.1. Mycorrhizal signalling, fungal penetration and mycorrhizal development. Figure adapted from (Oldroyd, 2013).

1.1.2 Nodulation

In contrast, the legume-rhizobial symbiosis forms an entirely new organ on plant roots, known as nodules. These specialised plant organs provide the right environment for the nitrogen-fixing bacteria; since the rate of ammonia production by the bacterial enzyme nitrogenase is inhibited by oxygen and limited by carbon availability, the nodule regulates the levels of both in order to maximise the amount of nitrogen that is fixed (Oldroyd and Downie, 2004). Two types of nodules can be formed depending on the plant species

involved. Indeterminate nodules are formed by, for example, *M. truncatula*, and are characterised by a zonation of the cells within the nodule, including: the meristem, infection zone, interzone, fixation zone and the senescence zone. The development of these nodules begins with cell divisions in the pericyle and then in the inner cortical cells. Alternatively, determinate nodules are formed by, for example, *Lotus japonicus,* and consist of one central nitrogen fixing tissue with no zonation. In contrast to indeterminate nodules, they originate by cell divisions which occur in the outer root cortex (Popp and Ott, 2011). The two types of nodule are further differentiated by the types of meristem present; indeterminate nodules have a tip-growing meristem, whereas determinate nodules only have a transient meristem, which is lost in mature nodules (Oldroyd *et al.*, 2011).

Root nodule development begins when flavonoids are secreted by the roots of legumes. These are plant secondary metabolites that are detected by rhizobial NodD proteins. These proteins activate gene expression of nodulation-related (*nod*) genes, some of which encode genes which produce Nod factor proteins (Long, 1996). Nod factors consist of decorated lipo-chitooliogsaccharides, which will be discussed in more detail in section 1.2.1. Plants perceive these signals at concentrations as low as 10^{-12} M and symbiotic gene expression is activated. In addition, two calcium signals are detected: calcium flux and nuclear-associated calcium oscillations (Section 1.5). Rhizobia attach to the root, which induces root-hair deformation that entraps the bacteria in a root-hair curl. Rhizobia invade the plant cells via an infection thread, and as they grow down through the epidermis into the cortex they are preceded by calcium oscillations. The direction of this growth is determined by the pre-infection thread structure. Nodules initiate by formation of the meristem in the cortical cell layer and cell divisions occur to form the nodule. Bacteria divide and differentiate in these cells, where they are now able to fix nitrogen and are known as bacteroids. Subsequently, when they are surrounded by plant-derived-membranes in the developed nodule they are known as symbiosomes (Figure 1.2) (Oldroyd and Downie, 2004), (Oldroyd, 2013).
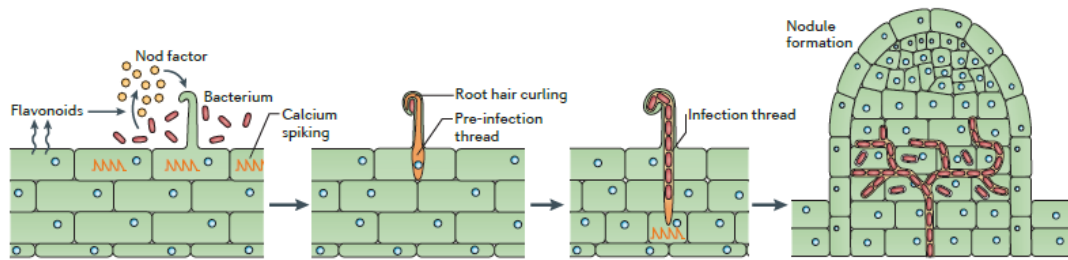
Figure 1.2. Nodulation signalling, infection and nodule organogenesis. Figure adapted from (Oldroyd, 2013).

1.2 Common symbiosis (Sym) pathway

It is widely accepted that the root nodulation symbiosis actually evolved from the more ancient arbuscular mycorrhizal symbiosis, and this is thought to have occurred approximately 58-65 MYA (Lagunas *et al.*, 2015), (Kistner and Parniske, 2002). Mycorrhization is present in all major plant lineages; however, nodulation has only evolved in one clade within Eurosid I. Despite the restricted number of species that can form nodules, nodulation has actually evolved multiple times within this clade. This implies that these species are potentially "pre-disposed to nodulate" (Kistner and Parniske, 2002). Recently, evidence has been found in support of this hypothesis by the identification of a single precursor to nodulation in the phylogeny of angiosperms. This event most likely emerged only once over 100 MYA, and nodulation has subsequently been lost and gained several times in the nitrogen-fixing clade. Furthermore, this pre-disposition is further thought to have been evolutionarily maintained in some angiosperm species (Werner *et al.*, 2014). These data suggest that nodulation may have emerged earlier than previously theorised.

Interestingly, several pieces of evidence have emerged recently which suggest that elements of the signalling pathways for both mycorrhization and nodulation may pre-date the emergence of mycorrhization. First, several components of a "symbiotic toolkit" required for the formation of symbiosis have been identified in Charophytes and pre-date the colonisation of land by plants (Delaux *et al.*, 2013b). Furthermore, there is evidence that the perception of lipo-chitooligosaccharides evolved from earlier plant defence signalling due to similarities in receptor structure and phylogenetic analysis (Liang *et al.*, 2014). Therefore it appears that root nodulation evolved from mycorrhization, which in turn utilised more ancient signalling components.

There is strong evidence for the evolution of the nodulation symbiosis from mycorrhization, the most persuasive of which is the presence of at least seven common genes necessary for the establishment of both symbioses. This is known as the common Sym pathway. The Sym pathway has been characterised in several different species by mutant analysis including *Pisum sativum, M. truncatula* and *L. japonicus.* In *P. sativum* three genes were identified: *Sym19*, *Sym8* and *Sym9* (Schneider *et al.*, 1999), (Ané *et al.*, 2004), (Schneider *et al.*, 2002). In *M. truncatula,* three components common to both pathways were also identified: *DMI1* (DOES NOT MAKE INFECTIONS 1), *DMI2* and *DMI3* (Catoira *et al.*, 2000). Interestingly, in *L. japonicus,* seven genes common to both pathways were identified: *SYMRK, CASTOR, POLLUX, NUP133, CYCLOPS, CCaMK* and *NUP85* (Kistner *et al.*, 2005). It was later determined that the genes identified in *P. sativum* are homologues of those identified in *M. truncatula* and three genes from *L. japonicus*. The common Sym pathway, and the corresponding nodulation and mycorrhization pathways as currently known, is shown in Figure 1.3.
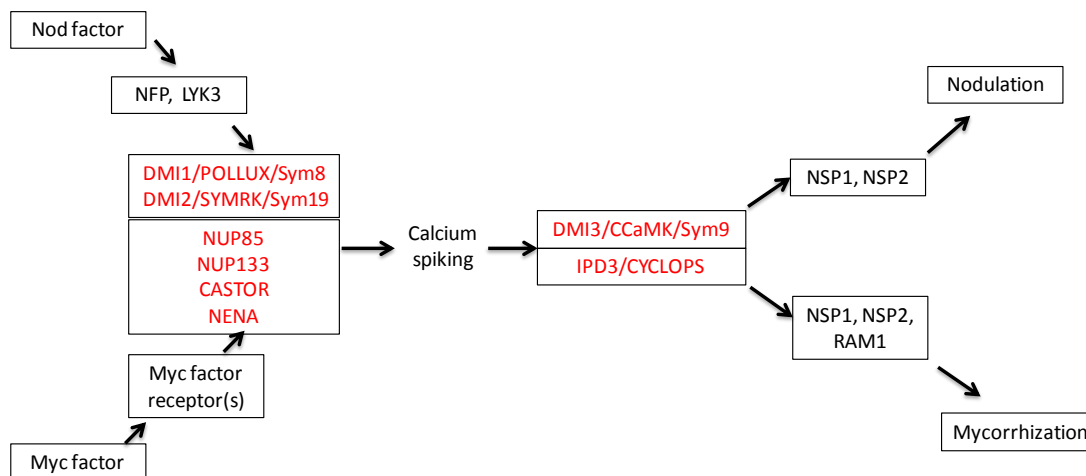
Figure 1.3. Major components of the signalling pathway for the establishment of nodulation and mycorrhization. The common genetic elements of the Sym pathway are marked in red. Arrows indicate progression through the pathway and not interactions. Elements from *M. truncatula* are listed where known, and homologues from *L. japonicus* and *P. sativum* respectively are indicated after a slash. NUP133, NUP85 and CASTOR have no known homologues. This is based on data from (Parniske, 2008), (Oldroyd, 2013) and personal communication.

1.2.1 Nod and Myc factors

Nod factors are lipo-chitooligosaccharides which consist of four or five *N*-acetyl-D-glucosamine residues and are β1-4 linked. An *N*-acyl group is attached to the sugar at the non-reducing end. Interestingly, specificity of the bacterial-plant interaction is encoded in specific modifications which can be present on the reducing and non-reducing terminal sugar residues. These additions include methyl, fucosyl, acetyl and sulphate groups. Further specificity can be conferred by the number of acetyl-glucosamine residues and the mix of Nod factors produced by the bacteria (Oldroyd, 2013), (Geurts and Bisseling, 2002). The first Nod factor which was characterised was from *Sinorhizobium meliloti* and consists of four glucosamine units with a 16 carbon acyl chain and acetyl and sulphate modifications (Figure 1.4) (Lerouge *et al.*, 1990). Nod factors are required for the majority of rhizobia-legume interactions, however there are some examples of interactions without Nod factor including some *Bradyrhizobium* strains (Giraud *et al.*, 2007). Recently, it was demonstrated that a further level of specificity is determined by bacterial exopolysaccharides which are recognised after Nod factor perception by a specific receptor-like kinase in *L. japonicus, Epr3* (Kawaharada *et al.*, 2015).
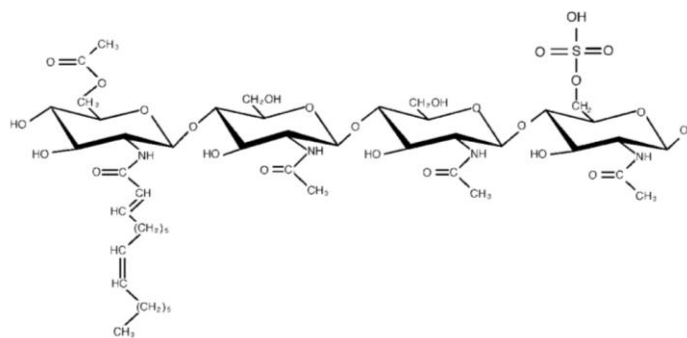
Figure 1.4. The structure of *S. meliloti* Nod factor. Figure adapted from (Geurts and Bisseling, 2002) and (Lerouge *et al.*, 1990).

The structure of Myc factors was elusive for many years, but in 2011 the structure of Myc factors from *Glomus intraradices* were characterised (Maillet *et al.*, 2011). Interestingly, these consisted of sulfated and non-sulfated lipo-chitooligosaccharides similar to those observed in nodulation signalling. Two of these Myc factors are shown in Figure 1.5. Interestingly, further fungal signals have been identified which are non-decorated short-chain chito-oligosaccharides e.g. tetra-acetyl chitotetraose. These were able to elicit a calcium spiking response in *M. truncatula* root organ cultures, which is a marker of the

activation of genes required for mycorrhization (Genre *et al.*, 2013). Therefore, mycorrhizal fungi exude a variety of signalling molecules as part of the molecular dialogue. In contrast to the legume-rhizobia interaction, mycorrhizal fungi are able to interact with many species which this mixture of signals may facilitate. Interestingly, legumes and non-legumes have been shown to respond differently to lipo-chitooligosaccharides and chito-oligosaccharides, therefore showing potential different functions of these signals in activating mycorrhization in different species (Sun *et al.*, 2015).
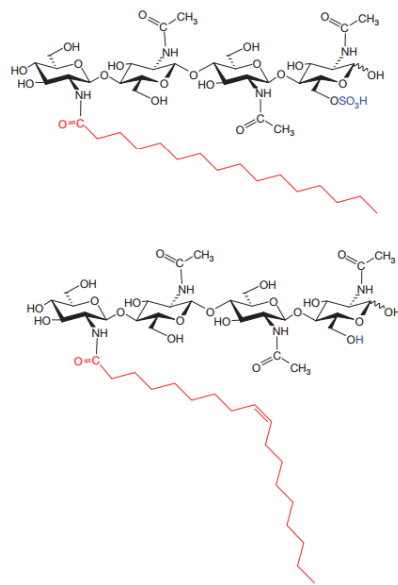
Figure 1.5. The structure of a sulfated and non-sulfated Myc factor from *G. intraradices.* Figure adapted from (Maillet *et al.*, 2011).

1.2.2 Functions of the common Sym pathway components

DMI2 encodes a plasma membrane receptor-like kinase with three extra-cytoplasmic leucine-rich repeats (LRR), which was first identified in *Medicago sativa* as NORK (Endre *et al.*, 2002). This gene is a homolog of *L. japonicus* SYMBIOSIS RECEPTOR-LIKE KINASE (*SYMRK)* (Stracke *et al.*, 2002) which has recently been characterised. SYMRK was demonstrated to interact with Nod factor receptor 5 (NFR5) in the absence of a symbiotic signal. This was promoted by the cleavage of a malectin-like domain (MLD) in the extracellular region SYMRK. The MLD (and LRR) further regulates the protein by enhancing its own degradation. The importance of this domain to the regulation of nodulation was demonstrated by the deletion of the MLD and LRR region, which initiated excess infection thread development (Antolín-Llovera *et al.*, 2014). *DMI1* encodes a cation channel which

was first characterised in *M. truncatula* (Ané *et al.*, 2004) and which localises to the nuclear membrane (Riely *et al.*, 2007). Interestingly, whilst *POLLUX* is the putative orthologue of *DMI1* in *L. japonicus,* both *CASTOR* and *POLLUX* are cation channels. Furthermore, the *castor pollux* double mutant can be rescued by *DMI1* alone, whilst both *CASTOR* and *POLLUX* are required to rescue the *dmi1* mutant. This suggests that *DMI1* has evolved to provide the roles of both *CASTOR* and *POLLUX* in *M. truncatula* (Venkateshwaran *et al.*, 2012). The role of DMI1 in the generation of the oscillatory calcium signal is discussed in section 1.5.2. NUP85 and NUP133 are both nucleoporins and along with NENA form a nuclear pore which is required for the generation of the calcium spiking signal (Groth *et al.*, 2010). It is theorised that this nuclear pore could be involved in the correct localisation of, for example, DMI1 (Oldroyd, 2013).

Downstream of these common genes is a nuclear calcium spiking signal (Section 1.5.2) and two proteins which are thought to interpret and transduce this signal: CCaMK and CYCLOPS. Consistent with this both proteins are nuclear-localised (Takeda *et al.*, 2012), (Yano *et al.*, 2008). *DMI3* (CCaMK) encodes a calcium- and calmodulin-dependent protein kinase and *IPD3* (CYCLOPS) is a binding partner and phosphorylation target of CCaMK, and a recently characterised transcription factor. These two common Sym pathway components are described in detail in Sections 1.7 and 1.7.8, respectively. The nature of the Sym pathway requires transduction of the signal from the plasma membrane components (e.g. DMI2) to the nucleus (e.g. DMI1). Recently, a secondary messenger, mevalonate, was proposed as the signalling component required for this role. Mevalonate can be produced by 3-hydroxy-3-methylglutaryl CoA reductase 1 (HMGR1), and HMGR1 has been shown to interact with DMI2. Crucially, mevalonate is able to activate both nuclear calcium spiking and symbiotic gene expression in a DMI1-dependent manner. Therefore mevalonate (or a metabolite of the mevalonate pathway) is a possible candidate to transduce the symbiotic signal from the plasma membrane to the nucleus via HMGR1 (Venkateshwaran *et al.*, 2015).

1.2.3 Nod and Myc pathway genes

In addition to the common Sym pathway components, nodulation- and mycorrhization-specific genes have also been characterised (Figure 1.3). Two Nod factor receptors have been characterised in *M. truncatula* namely NOD FACTOR PERCEPTION (NFP) and LYSIN

MOTIF RECEPTOR-LIKE KINASE3 (LYK3) (Limpens *et al.*, 2003), (Arrighi *et al.*, 2006). NFP and LYK3 are LysM-domain containing receptor-like kinases, however only LYK3 possesses an active kinase domain. Interestingly, they have recently been shown to form heteromeric complexes *in planta* in cell layers where Nod factor signalling occurs (Moling *et al.*, 2014). By contrast, Myc factor receptors are not well characterised. Interestingly, an *O. sativa* receptor kinase known as CERK1 has recently been shown to play both an essential role in chitin defence signalling and in mycorrhization (Miyata *et al.*, 2014). Furthermore, MtLYK3 was demonstrated to be a close homolog of OsCERK1, and LYK3 but not NFP was shown to be necessary for mycorrhizal infection in *M. truncatula* (Zhang *et al.*, 2015). This therefore suggests further overlap between nodulation, mycorrhization and even defence signalling and thus it is now less clear how specificity of these pathways is achieved.

Downstream of the common Sym pathway are Nodulation Signalling Pathway1 (NSP1) and NSP2. These are nuclear-localised GRAS-type transcription factors and were demonstrated to function downstream of calcium spiking and CCaMK in nodulation signalling (Smit *et al.*, 2005), (Kalo *et al.*, 2005). Furthermore, these receptors were shown to form a complex which directly binds to the promoters of early nodulin genes (Hirsch *et al.*, 2009). In addition, a mycorrhizal-specific GRAS-type transcription factor has been identified: Required for arbuscular mycorrhization1 (RAM1). Mutants affected in this gene are unable to be colonised by mycorrhizal fungi (Gobbato *et al.*, 2012). Recently, it has been shown that there is also crossover in the nodulation and mycorrhization pathways at the transcription factor level. NSP1 has been shown to play a role in mycorrhization, because in the *nsp1* mutant of *M. truncatula* both the frequency of mycorrhizal colonisations and the response to mycorrhizal lipo-chitooligosaccharides were reduced (Delaux *et al.*, 2013a). Furthermore, new data reveal a potential role of both NSP1 and NSP2 in addition to RAM1 in mycorrhization (Maillet *et al.*, 2011), (Leonie Luginbuehl, 2015, personal communication).

Many further genetic elements of both nodulation and mycorrhization signalling have been identified, including NIN, ENOD11, ERN1, NF-YA1/2, VAPYRIN, CERBERUS, HCL, SYMREM etc (Marsh *et al.*, 2007), (Journet *et al.*, 2001), (Cerri *et al.*, 2012), (Laloum *et al.*, 2014), (Murray *et al.*, 2011), (Takeda *et al.*, 2013), (Catoira *et al.*, 2001), (Lefebvre *et al.*, 2010). However, it is not within the scope of this literature review to cover them all because they function downstream, or in conjunction with, the Sym pathway.

1.3 Calcium signalling in plants

1.3.1 Calcium as a universal secondary messenger

Calcium ions are an important secondary messenger for both animal and plant cell signalling. Calcium ions act as an intracellular messenger for a large number of processes, relaying specific information by its ability to produce a wide variety of different signals. In animals, calcium ions are involved in processes throughout the life-span of the organism, right the way from fertilisation to cell death (Berridge *et al.*, 1998). In plants, calcium ion signalling is equally versatile, and occurs via changes in calcium which form distinct stimulus-specific "signatures". Plant calcium signals are involved in a diverse range of processes including: stress signalling e.g. osmotic/ salt/ oxidative stresses, pathogen signalling, circadian clock regulation, hormonal signalling and of course bacterial/fungal signal transduction. Specificity of the signal is therefore extremely important and can be conferred in a variety of ways including by the frequency, amplitude, duration and number of calcium waves or oscillations (although not all calcium signals are oscillatory). Furthermore, specificity can be conferred by comparmentation of the signal, either by variation of spatial localisation within the cytosol, or within organelles including the nucleus, mitochondria and chloroplast (McAinsh and Pittman, 2009). For example, calcium signals in the mitochondria have been implicated in response to touch stimuli in *Arabidopsis thaliana* (Logan and Knight, 2003) and chloroplast calcium signals are involved in response to dark as a circadian indicator in *Nicotiana plumbaginifolia* (Johnson *et al.*, 1995). Further signal specificity can be conferred by signals in different tissue types, cell types and during different developmental stages.

Calcium signals are generated by a number of channels and pumps. Calcium channels in plants are, in general, calcium-permeable rather than calcium-selective and include plasma membrane channels (e.g. hypersensitive- and depololarisation-activated calcium channels, and mechanosensitive channels) and endomembrane channels (e.g. slow-vacuolar, voltage-gated and cADP-ribose channels). Furthermore, calcium pumps are required for rapid removal of calcium because at high levels it is toxic. Calcium efflux transporters include high affinity (e.g. ER-type Ca^{2+}-ATPases and autoinhibited Ca^{2+}-ATPases) and low affinity, high capacity calcium exchangers. Therefore, these channels and pumps serve to modulate

the calcium signal specifically depending on the stimulus and then return calcium ion levels back to the resting state after a signal (McAinsh and Pittman, 2009).

Interestingly, oscillations of calcium are theorised to encode pathway-specific information. For example, in animal cells, it was demonstrated that calcium oscillations are not just a by-product of calcium homeostasis but encode information and increase the efficiency of the signal. This is because specific gene expression was induced by the frequency of the oscillations and oscillations reduced the threshold needed for the activation of particular transcription factors (Dolmetsch *et al.*, 1998). This is also the case in plant systems. Calcium oscillations have been studied in several systems including pollen tube growth, stomatal guard cell regulation and in Nod-factor signalling (section 1.5.2). In stomatal responses, it has been shown that the strength of stomatal closure correlates with the stimulus and the resultant calcium oscillation (Evans *et al.*, 2001). This system is further discussed in section 1.3.2. Therefore, calcium is an important secondary messenger in many signalling pathways and plants are able to precisely control calcium oscillations to produce a variety of downstream responses.

1.3.2 Case study: Cytosolic calcium changes in stomata

One of the most important pathways involving cytosolic calcium ion signalling is the regulation of stomatal opening and closure. Stomata are key to regulating gas exchange and water-loss in the plant, therefore stomatal aperture must be tightly regulated. Cytosolic calcium oscillations and transients in guard cells are triggered in response to a wide variety of stimuli including ABA, H_2O_2, cold, pathogens and external calcium (Hetherington and Brownlee, 2004). Interestingly, different calcium signatures are produced in response to these stimuli which are thought to encode information about the type and strength of the stimulus perceived. It has been demonstrated in *A. thaliana* that different concentrations of external calcium induce different calcium oscillations. For example, oscillations induced by 1 mM external calcium showed an average amplitude of 160 nM but by contrast, 10 mM external calcium-induced oscillations had an average amplitude of 1020 nM (Allen *et al.*, 2000). Furthermore, different stimuli likely access different mechanisms for the generation of these different calcium oscillatory signatures. For example, the *de-etiolated 3* (*det3*) mutant of *A. thaliana* corresponds to a 60% reduction in the expression of the C-subunit of the V-type H1–adenosine triphosphatase (V-

ATPase). In the *det3* mutant, calcium oscillations and stomatal closure were abolished in response to external calcium and oxidative stress. However, stomatal closure and corresponding calcium oscillations still occurred in response to cold stress and ABA, suggesting different mechanisms for the generation of the calcium signatures (Allen *et al.*, 2000).

Different stomatal responses are also encoded by the cytosolic calcium signatures. First, stomatal closure can occur in either a short-term or long-term manner. Short-term closure is encoded in the increase of cytosolic calcium to a required threshold. By contrast, long-term programmed stomatal closure occurs in response to specific calcium oscillatory signals. This was demonstrated to be encoded in the frequency, spike number, duration and amplitude of the calcium signature (Allen *et al.*, 2001). Furthermore, calcium has also been shown to play a role in stomatal opening in response to low atmospheric CO_2 (Young *et al.*, 2006). This demonstrates some of the complexity involved in stomatal signalling, because calcium signatures can mediate both stomatal opening and closure. The mechanisms involved in these different stomatal responses have begun to be elucidated, including the role of calcium-dependent protein kinases (CDPKs). For example, *cpk3* and *cpk6* mutants of *A. thaliana* show impaired ABA activation of plasma membrane calcium-permeable channels and slow-type anion channels. Interestingly, the double mutant of *cpk3* and *cpk6* was impaired in only the short-term closure response and not the long-term programmed response (Mori *et al.*, 2006). These data demonstrate an important role for CDPKs in stomatal calcium signalling and show that the two observed calcium-induced closure responses are mediated in different ways. This system is an example of the importance of calcium signatures in signalling pathways and the precise control plant cells have over calcium oscillations to produce specific downstream responses.

1.4 Calcium signalling in the Sym pathway

Two calcium responses have been observed in legume roots in response to Nod factors: calcium flux and calcium oscillations (Figure 1.6). Calcium influx occurs rapidly after the signal and is followed by membrane depolarisation (Felle *et al.*, 1999), (Ehrhardt *et al.*, 1992). Calcium oscillations are observed approximately 10 minutes after the initial signal and these oscillations are known as calcium spiking (Wais *et al.*, 2002). Calcium spiking is also observed in response to Myc factors (Kosuta *et al.*, 2008).
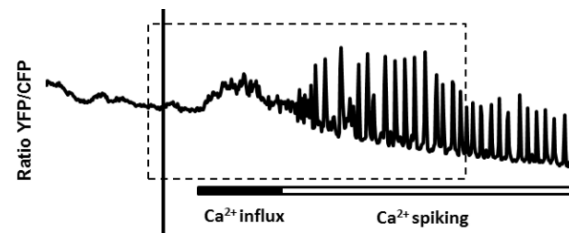
Figure 1.6 Nod factor-induced calcium responses observed in WT *M. truncatula* root hair cells, measured using Cameleon YC3.6. The horizontal black line represents the addition of 10 nM Nod factor. Figure courtesy of Sarah Shailes.

1.4.1 Calcium flux

The calcium flux occurs within one minute of Nod factor application and has been observed with both ion-specific micro-electrodes and calcium sensitive dyes (Felle *et al.*, 1999), (Shaw and Long, 2003). This calcium flux response is widespread amongst legumes and has been detected in the root hair cells of *M. truncatula, Vicia sativa* and *P. sativum,* amongst others. Root hairs have an inherent calcium gradient during growth, with the highest concentration present at the tip of the root hair. When Nod factor is applied, a wave of calcium influx begins at the tip and moves along the length of the root hair cell towards the nucleus. The calcium which participates in this influx is thought to be derived from both external and internal stores. Following this, the efflux of chloride and potassium occurs, with subsequent cytosolic alkalisation (Oldroyd and Downie, 2004).

Shaw and colleagues (Shaw and Long, 2003) determined that calcium flux and calcium spiking are in fact separable. First, different Nod factor requirements are necessary for each response. At 1 nM Nod factor, root cells showed only the calcium spiking response after approximately 10 min. However, on application of 10 nM Nod factor, calcium influx was observed, followed by calcium spiking as expected. These were also differentially localised in the cell. Furthermore, mutants which were impaired in calcium spiking were still capable of producing the calcium influx response. These included the *dmi1* and *dmi2* mutants of *M. truncatula* (Shaw and Long, 2003) and *nup133* and *sym24* of *L. japonicus* (Miwa *et al.*, 2006b). This suggests that calcium spiking does not require prior calcium flux. It is therefore unclear as to what potential role calcium flux might play in symbiosis signalling, since it requires such a high Nod factor concentration. One hypothesis which has been proposed is that they could function in the opposite order than they have been

observed when Nod factor is artificially applied. When bacteria encounter the legume root in the soil, the calcium flux would be activated after calcium spiking and root hair deformation had occurred because when the bacteria is encased in the root hair curl the concentration of Nod factor experienced by the plant would be sufficiently high. The calcium spiking would then play a role in infection thread development, after calcium spiking (Miwa *et al.*, 2006b). Consistent with this, Morieri and colleagues (Morieri *et al.*, 2013) demonstrated the importance of the calcium influx for efficient nodule infection.

1.4.2 Calcium spiking

Calcium spiking is observed approximately 10 minutes after application of Nod factor or rhizobia, and this delay is thought to allow calcium homeostasis changes to occur. Individual calcium spikes are characterised by a sharp increase in calcium and a slower decline. The calcium spiking response was first characterised by Erhardt and colleagues (Ehrhardt *et al.*, 1996) in *M. sativa* roots injected with dextran-linked calcium-selective dyes. They demonstrated that the calcium spiking signal is localised to the nuclear region of the cell and that key structural features of Nod factors were essential to stimulating the response. Furthermore, the calcium changes were quantified to occur between approximately 150 – 800 nM (Ehrhardt *et al.*, 1996), (Felle *et al.*, 1999). Subsequently, the localisation of calcium spiking was confirmed as occurring within the nucleus by Sieberer and colleagues (Sieberer *et al.*, 2009) in *M. truncatula* using a nuclear localised form of Cameleon (a calcium responsive protein, that undergoes fluorescence resonance energy transfer on Ca^{2+} binding that can subsequently be detected). They also demonstrated that some spiking occurs in the nuclear region of the cytosol within the cell. Calcium spiking has also been shown to stimulate gene expression, in particular 36 spikes were shown to be sufficient to induce ENOD11 expression (Miwa *et al.*, 2006a). All nodulating species, present in the nitrogen-fixing clade, have been shown to possess this same calcium response (Granqvist *et al.*, 2015).

Calcium spiking has also been characterised in mycorrhization signalling. Nuclear-localised Cameleon was again used with *M. truncatula* and the non-legume *Daucus carota*. Calcium spiking was observed when a root organ culture came into contact with hyphopodia and when treated with spore exudate. Therefore this shows that calcium spiking is also present in the arbuscular mycorrhizal symbiosis signalling pathway for both legumes and non-

legumes (Chabaud *et al.*, 2011). The calcium spiking signature in both the nodulation and mycorrhization pathway are chaotic in nature, which leads to a robust but flexible signalling system (Hazledine *et al.*, 2009). Interestingly, it has previously been theorised that specificity in the Sym pathway may be encoded by different calcium spiking signatures in the nodulation and mycorrhization pathways. Nod factor-induced calcium spiking was shown to have a periodicity of 90 s and Myc factor-induced signals were more variable, with a shorter periodicity (30 s) and lower amplitude (Kosuta *et al.*, 2008). However, more recent data have called these conclusions into question. Sun and colleagues (Sun *et al.*, 2015) analysed calcium responses in *M. truncatula* to various different lipo-chitoligosaccharides. They concluded that there was in fact a high degree of similarity between the Myc and Nod responses detected, but some differences in the frequency of spiking were still observed. However, these frequency differences can now be explained by work carried out by Sieberer and colleagues (Sieberer *et al.*, 2012). They observed a switch in the frequency of calcium oscillations when both microbes infected the cortex, and prior to this calcium spiking was less regular. This could therefore account for the differences previously observed between mycorrhizal- and rhizobial-induced calcium spiking. Whilst calcium-spiking is therefore not likely to show differences relating to mycorrhization or nodulation, it has been theorised that differences observed could encode information about the specific mix of symbiosis signals which are perceived by the cell (Charpentier and Oldroyd, 2013).

The position of calcium spiking in the sym pathway has been determined genetically. This was analysed using mutants of *dmi1, dmi2* and *dmi3* in *M. truncatula* (Wais *et al.*, 2000). Mutants in all three genes do not develop nodules or infection threads, they show abnormal root hair responses to Nod factor and they don't show cortical cell divisions. However, whilst both *dmi1* and *dmi2* mutants are defective for calcium spiking, *dmi3* mutants showed WT calcium spiking responses to Nod factor application. Therefore, this places calcium spiking downstream of both *DMI1* and *DMI2* but upstream of *DMI3* (Figure 1.3). Furthermore, this was confirmed by work using *P. sativum* mutants (Walker *et al.*, 2000) and this placed *SYM8* (*DMI1*), *SYM10* and *SYM19* (*DMI2*) upstream of calcium spiking but *SYM2*[A], *SYM7*, *SYM9* (*DMI3*) and *SYM30* downstream (homologues of the three equivalent *M. truncatula* genes mentioned above are indicated in brackets). Therefore CCaMK (*DMI3)* is well-placed to interpret and transduce the calcium spiking signal in the

sym pathway, and it is widely accepted as the decoder of this signal. CCaMK is discussed in detail in section 1.7.

The role of different proteins in the generation of the symbiotic calcium spiking signal has also been somewhat elucidated. *DMI1* is essential for the generation of the calcium spiking signal and it encodes an ion channel with similarity to a potassium channel of *Methanobacterium thermoautothrophicum.* It was demonstrated that whilst DMI1 is not a calcium channel, it is important for the regulation of calcium release and therefore it regulates calcium spiking (Peiter *et al.*, 2007), (Charpentier *et al.*, 2008). It is localised to the inner nuclear membrane. Another component which is essential for calcium spiking is the sarco/endoplasmic reticulum calcium ATPase (SERCA) MCA8 (Capoen *et al.*, 2011), which is a calcium pump localised to both the inner and outer nuclear membranes and the endoplasmic reticulum. Mathematical modelling of the system has revealed how these proteins may function together to generate the calcium spiking signal. The store of calcium is hypothesised to be the nuclear envelope which is continuous with the endoplasmic reticulum. Interestingly, the model predicts a crucial third component in the generation of the calcium spiking signal which has not yet been identified: a calcium channel. This channel is theorised to be voltage-gated. The model further predicts the importance of calcium buffer proteins (for example calmodulin (CaM)) for the shaping of and the termination of calcium spiking (Granqvist *et al.*, 2012). Future experimental work will be crucial to validate this model; for example, by identification of the calcium channel. Furthermore, the role of DMI1 in the generation of the calcium signal requires further experimental characterisation. Modelling reveals that it is likely that DMI1 functions as a counter ion channel and as an activator of the potential calcium ion channel, in a calcium-induced manner (Charpentier *et al.*, 2013). Experimental work will also be required to validate this theory.

1.5 Calcium-binding proteins in plants

Plants possess a number of different calcium binding proteins, which can be divided into two groups: those containing EF-hands and those which do not possess EF-hands. Non EF-hand proteins show a lower affinity for calcium and are less well characterised in general. The calcium binding regions of the protein are structurally varied but possess similar acidic amino acids. Examples of non EF-hand proteins include annexins and C2 domain-containing

proteins. The annexin family of proteins are characterised by four 70 amino acid annexin folds that form a cyclic array. They bind calcium on the convex membrane-binding side of the protein. The C2 domain-containing proteins are more thoroughly characterised, and these contain the C2 130-145 amino acid region that consists of three conserved subdomains. Calcium-binding in the C2 domain occurs at loops in between β-sheets and does not confer a conformational change. Instead, binding calcium allows an interaction with phospholipids. Examples of C2 domain proteins include phospholipase C, protein kinase C and cytosolic phospholipase A2 (Sathyanarayanan and Poovaiah, 2004), (Reddy and Reddy, 2004).

EF-hand proteins are much more common and for example the genome of *A. thaliana* encodes over 250 EF-hand-containing proteins (Day *et al.*, 2002). The EF-hand motif is highly conserved, and is found in plants as well as animal proteins. They are generally found in pairs within proteins and allow calcium ions to bind with a high affinity. Each EF-hand consists of a helix-loop-helix structure, where one calcium ion is coordinated by the loop region, inducing a conformational change. Each helix consists of approximately 10 residues, and the loop region contains 12. Calcium is co-ordinated by seven ligands, five of which are provided by the loop region and two of which are supplied by an amino acid of one of the helices. This forms the preferred co-ordination geometry for calcium: a pentagonal bipyramid. This means that several of the residues in the EF-hand loop region are highly conserved. Sometimes, the coordination sphere of calcium is completed with a water molecule. The conformation described corresponds to the canonical EF-hand domain structure, but there are also several different non-canonical EF-hand structures (Gifford *et al.*, 2007).

There are three main types of EF-hand proteins found in plants: CaM (and CaM-like proteins, CML), calcium-dependent protein kinases (CDPKs) and calcineurin B-like (CBL) proteins. CaM is found in all eukaryotes, whereas the other types of proteins are unique to plants and some protists. These proteins can also be further classified by their function; they either function as "sensor relay" or "sensor responder" proteins. For example, CaM and CML proteins bind calcium but don't have an enzymatic function and therefore are classified as sensor relay components. By contrast, CDPKs bind calcium and function as kinases and therefore are sensor responder proteins. CBLs in fact interact with CBL-

interacting protein kinases (CIPKs) and therefore function as bimolecular sensor responders (Batistic and Kudla, 2012). The structures of these proteins are summarised in Figure 1.7.
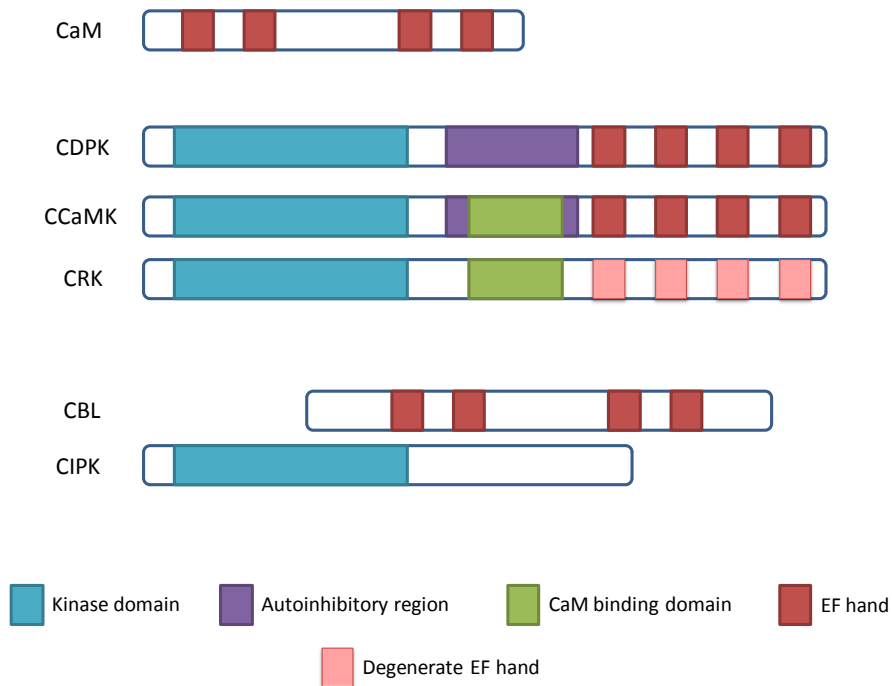
Figure 1.7. Structures of plant calcium binding proteins. Figure based on (Batistic and Kudla, 2012).
CaM – Calmodulin, CDPK – Calcium-dependent protein kinase, CCaMK - Calcium- and calmodulin-dependent protein kinase, CRK – CDPK-related kinase, CBL – Calcineurin B-like protein, CIPK – CBL-interacting protein kinase.

1.5.1 Calmodulin (CaM)

CaM is the most well characterised EF-hand protein from plants. Whilst animal genomes encode a small number of CaM isoforms, typically plant genomes encode many more CaM genes for example *A. thaliana* has seven highly related forms. Furthermore, there are also many more CMLs; in *A. thaliana* there are 50 (DeFalco *et al.*, 2010). This shows both the importance of CaM as a calcium sensor in plants and also indicates that it is involved in many different processes.

CaM is a small (~150 amino acids,~17 kDa) acidic protein that consists of two globular domains, connected by a flexible helical linker region. Each globular domain contains a pair of EF-hands. By contrast, CML proteins are more variable and contain variable numbers of EF-hands. When calcium binds to CaM at each EF-hand, a hydrophobic patch is exposed via a conformational change that allows CaM to interact with downstream targets. The

conformational flexibility of the linker region confers target binding specificity. The CaM binding domains of target proteins do not have a consensus sequence, however they show a consensus secondary structure of an amphipathic α helix (Perochon *et al.*, 2011), (DeFalco *et al.*, 2010). Furthermore, CaM binding domains can also feature a net positive charge and two hydrophobic anchor residues. Interestingly, in some cases CaM can also bind to target proteins in a calcium-independent manner (Tidow and Nissen, 2013). These features allow specific and high affinity binding in a variety of signalling pathways.

CaM can bind to a wide variety of target proteins including kinases, phosphatases, transcription factors, channels, pumps and many others. Furthermore, AtCaM7 has been shown to bind directly to DNA at the Z- and G-box motifs. This induces the expression of light-regulated genes (Kushwaha *et al.*, 2008). In addition, CaM is implicated in a wide range of processes and is regulated by a wide variety of stimuli. For example CaM functions in growth and development, and stress responses including drought, salt, temperature shock, wounding and immune responses. Specificity of CaM and CMLs in these responses is achieved partially due to comparmentation of the different isoforms (e.g. cytosol, nucleus and peroxisome) and due to specific expression patterns during development (Perochon *et al.*, 2011), (DeFalco *et al.*, 2010).

1.5.2 Calcium-dependent protein kinases (CDPKs)

CDPKs consist of an N-terminal domain followed by a kinase domain, junction domain, calmodulin-like domain bearing four EF-hands and a C-terminal domain (Figure 1.7). They are theorised to have originated from gene fusion of a Ser/Thr kinase with an EF-hand domain derived from CaM (Zhang and Choi, 2001). CDPKs are also encoded by large families of genes in plants; in *A. thaliana* there are 34 CDPK genes. They are variable in size and this is conferred by variation in the N-terminal domain. CDPKs are maintained in the inactive state by the interaction of the junction domain, which forms a pseudo-substrate and blocks the active site whilst simultaneously interacting with the CaM-like domain. When calcium binds to the EF-hands, a large conformational change occurs, along with autophosphorylation, and the protein is active for target phosphorylation. Plants can also possess two related kinases: CDPK-related kinases (CRKs) and calcium- and calmodulin-dependent kinase (CCaMK). CRKs have a similar domain structure to CDPKs except the EF-hands in the CaM-like domain are degenerate and some isoforms can directly bind

Ca^{2+}/CaM instead. CCaMKs are dually regulated by calcium and Ca^{2+}/CaM, and their structure and activation is described in detail in section 1.7 (Batistic and Kudla, 2012), (Harper and Harmon, 2005), (DeFalco *et al.*, 2010).

Calcium-dependent protein kinases all transduce calcium signatures through phosphorylation of target proteins. Ser/Thr kinases catalyse the transfer of a phosphoryl group from the γ-phosphate of ATP to the hydroxyl group of a target residue, yielding a phosphorylated target protein and ADP. This is an example of covalent modification and can occur on the kinase itself (known as autophosphorylation) and on other target proteins. Phosphorylation can occur at serine, threonine or tyrosine residues however all plant calcium-dependent kinases phosphorylate just serine and threonine. Regulation by phosphorylation can impact negatively or positively on the activity of the protein, can create recognition sites for other proteins, and can facilitate transitions between ordered and disordered protein states (Endicott *et al.*, 2012). Further regulation can be provided by de-phosphorylation for example by the action of phosphatases. Protein phosphorylation in response to calcium changes occurs in many different processes in plants, for example with CCaMK during symbiosis signalling (discussed in section 1.6).

CPDKs function in diverse areas of cell signalling including: root and pollen tube growth, salt and drought stress, regulation of calcium and potassium homeostasis, and pathogen defence. Indeed, since there are so many different calcium-regulated kinases in plants it is theorised they will be involved in some way in the majority of developmental processes (Harper and Harmon, 2005), (Batistic and Kudla, 2012). Therefore specificity of signalling is important, and one of the ways in which this is achieved is by comparmentation. For example, CDPKs can be specifically found in the cytoplasm, nucleus, mitochondria and peroxisome (Harper and Harmon, 2005). Furthermore, many isoforms are membrane-associated or can be either membrane associated or soluble e.g. AtCPK3 (Dammann *et al.*, 2003). In addition, some CDPKs can change subcellular localisation in response to stress stimuli; for example, *Mesembryanthemum crystallinum* CPK1 shifts from the plasma membrane to the nucleus in response to salt stress (Chehab *et al.*, 2004).

1.5.3 Calcineurin B-like proteins (CBLs)

CBLs fall within the wider CaM superfamily, but are slightly larger than CaM at 23-26 kDa. In general they possess four EF-hands like CaM, but it is not clear if they are all functional. The *A. thaliana* genome encodes ten CBL genes. As described above, CBLs function by interacting with CIPK proteins, of which there are a further 26 encoded by *A. thaliana*. CIPKs are structurally related to sucrose-non-fermenting and AMP-activated kinases from yeast and animals respectively, although they are characterised by a unique C-terminal domain required for interaction with CBLs. Since there are many possible CBL-CIPK interaction combinations they have the potential to participate in a wide variety of pathways (Batistic and Kudla, 2012) (DeFalco *et al.*, 2010). Similarly to the previously described plant calcium binding proteins, the CBLs and CIPKs function in a wide variety of processes and in response to many stimuli. These include salt, cold, wounding, pH, low oxygen and other stresses. Furthermore, there is also cross-talk between the CBL-CIPK network and other pathways including SOS, reactive oxygen species, potassium sensing, ABA and nitrate sensing pathways (Yu *et al.*, 2014).

1.6 CCaMK

Calcium- and calmodulin-dependent protein kinase (CCaMK) is a serine/threonine protein kinase that is unique to plants but not ubiquitous amongst them. It was first identified in *Lilium longiflorum* where it has an open reading frame of 520 amino acids encoded by a single gene and is approximately 56 kDa in size when expressed in *Escherichia coli* (Patil *et al.*, 1995). It was determined to contain three domains: catalytic domain, CaM-binding domain and a visinin-like Ca^{2+} binding domain. CCaMK is regulated in two ways: *via* 3 functional EF-hands present on the visinin-like domain or, *via* CaM binding CCaMK when 4 Ca^{2+} ions are bound to CaM. The fact that this protein can bind calcium directly through the visinin-like domain and indirectly, *via* calmodulin, sets it apart from the CDPKs. CCaMK is also controlled by autophosphorylation, which is stimulated by Ca^{2+} ions binding to the EF-hands of the visinin-like domain, but inhibited by Ca/CaM binding to the CaM binding domain. The kinase domain can phosphorylate downstream targets when autophosphorylated and Ca^{2+}/CaM is bound to CCaMK (Takezawa *et al.*, 1996).

The function of CCaMK in plants remained elusive for several years despite biochemical characterisation of the protein. The first proposed function for the protein in *L. longiflorum* and *Nicotiana tobacum* was to detect transient changes in free Ca^{2+} and control developmental changes in the anther (Poovaiah *et al.*, 1999). This was due to the detection of a stage- and tissue-specific expression pattern of CCaMK mRNA and protein during anther development. However, the key role of CCaMK was not elucidated until 2004, and surprisingly this role was in root and not flower tissue. A mutant of *M. truncatula* which failed to form root nodules (Nod$^-$) known as *dmi3* had been previously identified by Catoira and colleagues (Catoira *et al.*, 2000). Mitra and colleagues (Mitra *et al.*, 2004) and Levy and colleagues (Levy *et al.*, 2004) determined that the *DMI3* gene in fact corresponds to a CCaMK protein due to the high level of sequence similarity. Furthermore, they postulated that since the *dmi3* mutant retains the nuclear calcium spiking signal it is therefore genetically downstream and so it is consequently the best candidate to interpret and transduce this signal in both nodulation and mycorrhization signalling (Levy *et al.*, 2004). This was further confirmed by comparison with the *sym9 P. sativum* gene which has a similar phenotype and has a mutation in the orthologous gene. Furthermore, the *DMI3* gene has high sequence similarity to proteins in species which can form arbuscular mycorrhization, for example *Oryza sativa*. However, there is no *CCaMK* gene in the genome of *A. thaliana*, which cannot form arbuscules.

1.6.1 Early biochemical characterisation and the domain structure of CCaMK

Early biochemical characterisation of CCaMK was carried out on *L. longiflorum* CCaMK. The domain structure of the CCaMK protein is shown in Figure 1.8. This schematic focuses on MtCCaMK because the results presented later in this thesis pertain to that protein, however the only difference pertains to the residue of autophosphorylation: T271 in *M. truncatula* and T267 in *L. longiflorum.*
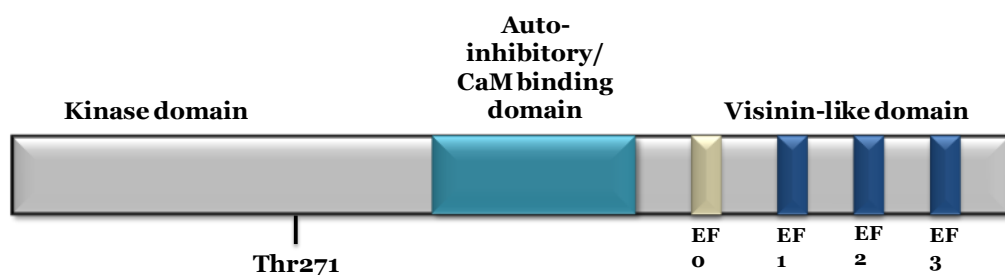
Figure 1.8. Domain structure of the CCaMK gene of *M. truncatula*, with key features marked.

The kinase domain is located at the N-terminus of the protein, and possesses all 11 conserved sub-domains usually detected in Ser/Thr kinases (Patil *et al.*, 1995). Substrate-phosphorylation was determined to be Ca^{2+}/CaM-dependent *in vitro* and autophosphorylation was shown to occur at threonine residues in response to calcium-binding to the EF-hands of the visinin-like domain (Takezawa *et al.*, 1996). Specifically, T267 in the kinase domain (equivalent to T271 in MtCCaMK) was determined to be the main site of autophosphorylation by MALDI-TOF mass spectrometry. Mutation of this residue to phospho-ablative alanine abolished calcium-stimulated autophosphorylation and Ca^{2+}/CaM-dependent kinase activity. Furthermore, homology modelling with Ca^{2+}/CaM-dependent protein kinase I (CaMKI) was carried out and it was determined that the T267 residue was most likely surface exposed and located distantly from the active site of the kinase. This therefore suggests an inter-molecular mechanism of autophosphorylation occurs (Sathyanarayanan *et al.*, 2001). An inactivation mechanism of the protein was also suggested, because it was demonstrated that Ca^{2+}-stimulated autophosphorylation led to the loss of kinase activity *in vitro* due to the formation of a sedimentable enzyme (Sathyanarayanan and Poovaiah, 2002). However, the relevance of this finding has not been validated *in planta.*

The CaM binding domain contains a single binding site for CaM and shows high sequence identity (79%) to the α subunit of the animal CaMKII protein, with both forming basic amphiphilic α helices (Takezawa *et al.*, 1996). The CaM binding domain also overlaps with an auto-inhibitory (AI) domain. This was identified because a truncation of the protein containing only the kinase domain is constitutively active *in vitro,* and a calmodulin-binding domain peptide was shown to inhibit kinase activity (Ramachandiran *et al.*, 1997). Whilst calcium-binding stimulates autophosphorylation 3.4 times basal level, Ca^{2+}/CaM binding inhibits autophosphorylation to basal levels (Takezawa *et al.*, 1996). Furthermore, Ca^{2+}/CaM affinity is also affected by Ca^{2+}-stimulated autophosphorylation; it was shown that after autophosphorylation, Ca^{2+}/CaM binding was eight-fold tighter, releasing autoinhibition (Sathyanarayanan *et al.*, 2000). Therefore the kinase activity of CCaMK is regulated by both calcium and calmodulin in a complex manner. This is further complicated by the potential for differential regulation of CCaMK by CaM isoforms, however this is not universal as it was only detected in *N. tabaccum* CCaMK but not *L. longiflorum* CCaMK (Liu *et al.*, 1998).

The visinin-like domain (VLD) is located at the C-terminus of the protein and directly binds calcium ions at three functional EF-hands. Using point mutants for each EF-hand which prevented Ca^{2+} binding, it was shown that EF-hands 2 and 3 are essential for a mobility shift suggesting that a conformational change occurs on Ca^{2+} binding to the VLD (Takezawa *et al.*, 1996). The three EF-hands are essential for full kinase activity. This was further dissected using deletion mutants of the EF-hands; as each EF-hand was sequentially deleted, kinase activity decreased. When all three EF-hands were removed, kinase activity was just 4% (Ramachandiran *et al.*, 1997). Therefore, the EF-hands also play a key regulatory role in the function of CCaMK. The presence of a visinin like domain in a plant protein is, in fact, unique to CCaMK. It is so-named due to the structural and sequence homology to the animal visinin protein, which is a photoreceptor-specific calcium-binding protein that functions in the outer nuclear layers of the retina (Yamagata *et al.*, 1990).

These findings likely also apply to the other CCaMK proteins, including *M. truncatula* CCaMK. For example, Gleason and colleagues (Gleason *et al.*, 2006) confirmed that DMI3 of *M. truncatula* is equivalent to CCaMK from *L. longiflorum*. Firstly they share 72% sequence identity, and secondly despite the fact that *L. longiflorum* forms only the association with mycorrhizal fungi and does not form nodules, LlCCaMK can complement the *dmi3-1* mutant of *M. truncatula*. Therefore, it seems that the protein performs the same functions in both species, and in both associations.

1.6.2 Gain-of-function mutants

Important discoveries about the role of CCaMK in the sym pathway were made using gain-of-function mutants. Tirichine and colleagues (Tirichine *et al.*, 2006) characterised the *snf1* (spontaneous nodule formation) mutant of *L. japonicus* which corresponds to a point mutation of T265I. T265 corresponds to the previously characterised autophosphorylation residue of T267 in *L. longiflorum* and T271 in *M. truncatula* CCaMK. When T265 was mutated to isoleucine, spontaneous nodules were formed in the absence of calcium spiking, Nod factor or rhizobial infection. Furthermore, normal bacterially infected nodules were formed in the *snf1* mutant when inoculated with the bacterial symbiont *M. loti*. These data therefore shows that WT CCaMK is negatively regulated and the loss of the ability to be autophosphorylated releases this negative suppression.

Gleason and colleagues (Gleason *et al.*, 2006) characterised mutants of *M. truncatula* CCaMK which lacked the auto-inhibitory domain. When expressed under the native promoter, spontaneous nodulation and concurrent nodulation-gene expression was observed, in a manner analogous to the *snf1* mutant phenotype. Similarly, no bacterial stimulus was needed and nodules developed with no rhizobial infection. However, in contrast to the *snf1* phenotype, no infected nodules were developed on inoculation with the rhizobial partner, *S. meliloti*, indicating that the regulatory domains of CCaMK are essential for the protein's full function. They also demonstrated that point mutations of the autophosphorylation site T271 to alanine (phospho-ablative) and aspartate (phospho-mimetic) led to the activation of nodulation gene expression in the absence of Nod factors. In contrast to the truncation mutant, both of the point mutations were able to complement the *dmi3-1* mutant and produced infected nodules. They are therefore fully functional proteins. This result was unexpected as usually phospho-ablative and phospho-mimetic versions of a protein would yield different biological effects. Furthermore, it implies that the mechanisms by which the gain-of-function phenotype occurs is different for the truncation and point mutants.

The significance of this work is profound. It shows that CCaMK is upstream of nodule organogenesis and furthermore activation of CCaMK alone is sufficient to induce nodulation. Further work by Hayashi and colleagues (Hayashi *et al.*, 2010) demonstrated, in addition, that upstream Sym pathway elements are required entirely to activate CCaMK in the pathway. This is because the gain-of-function T265D mutation complements loss of function mutations in upstream elements (*symrk*, *castor, pollux, nup85*) for both nodule organogenesis and rhizobial and mycorrhizal infection. This emphasizes the central role that CCaMK plays in the Sym pathway, and therefore the potential for engineering nitrogen-fixing crops by the autoactivation of one protein in the pathway.

1.6.3 Biochemical characterisation of calcium and calmodulin binding

The calcium and CaM binding properties of MtCCaMK were further characterised by Swainsbury and colleagues (Swainsbury *et al.*, 2012). The experiments were carried out with shorter constructs of CCaMK, containing either the VLD or the AI and VLD, which were monomers in solution. The conformational change that occurs when calcium ions bind to the EF-hands was determined to correspond to a change in only the tertiary structure of

the protein. This corresponded to an exposure of a hydrophobic cleft on the protein, in a manner analogous to CaM. Since no change in quaternary structure was detected in the presence of calcium, the function of the exposure of this hydrophobic patch and conformational change is most likely to facilitate the binding to target proteins and not to other CCaMK subunits. However, it is currently unknown as to whether this hydrophobic patch is also exposed in the full-length protein.

It was further determined that the three functional EF-hands have different affinities for calcium; EF-hand 3 is a high affinity site ($K_d \leq 20$ nM) and EF-hands 1 and 2 have a lower affinity for calcium ($K_d = 200$ nM). CaM-binding to CCaMK was determined to occur in a 1:1 ratio, with an affinity of 55 nM in the presence of calcium. The calcium-binding affinities for the EF-hands of CaM were determined to be 1 and 15 µM (although these are expected to decrease in the presence of a target peptide i.e. CCaMK).

This therefore facilitates the prediction of how CCaMK functions during calcium-spiking. Basal calcium levels in the nucleus are estimated at 125 – 150 nM, and at this concentration the high affinity EF-hand 3 is occupied, but not EF-hands 1 and 2. During a calcium spike (625 – 850 nM), calcium can then bind to the lower affinity EF-hands, and to CaM which can then bind to CCaMK. Since association rates of calcium and CaM were determined to be near diffusion limited and dissociation rates relatively high, it can also be concluded that the formation of the CCaMK-CaM complex (and binding and release of calcium from the VLD and CaM) mirrors the calcium spiking signal.

1.6.4 CCaMK structure

No crystal structure for CCaMK from any plant species has yet been elucidated and therefore this remains an extremely important question regarding CCaMK. Previous attempts to produce crystals of CCaMK or truncations of the protein in this lab have also been unsuccessful. The best structural data available for CCaMK currently is based on homology modelling. However, homology modelling of CCaMK has limitations due to low sequence identity with potential template proteins, and therefore any conclusions drawn must be considered in this light. It will become extremely important therefore in future work to obtain a high resolution crystal structure of CCaMK in order to validate conclusions.

Homology modelling of CCaMK has been carried out with rat brain CaMKI (Sathyanarayanan *et al.*, 2001) and more recently with *Caenorhabditis elegans* CaMKII (Shimoda *et al.*, 2012). Despite the relatively low sequence identity of CCaMK with these target proteins, some useful structural information was gleaned by this method. Conclusions from the CaMKI model were discussed above (Section 1.6.1). From the CaMKII model, Shimoda and colleagues predicted a hydrogen bonding network involving the key autophosphorylation residue in *L. japonicus,* T265. This network is predicted to form between the un-phosphorylated form of T265 and residues S237, K264, E313 and R317. This network was predicted to be disrupted by autophosphorylation at T265 and in gain-of-function mutants T265D, T265A and T265I. In order to validate this model, they mutated one of the predicted residues, R317, to His and spontaneous nodulation was observed as predicted. Therefore they postulated that this hydrogen bond network acts as a "molecular brake" which would maintain CCaMK in an inactive state. Models for the function of CCaMK were therefore proposed based on this data, which are detailed in (Singh and Parniske, 2012) and (Poovaiah *et al.*, 2013). These models will not be discussed in this review however because subsequent homology modelling of CCaMK was carried out by (Miller *et al.*, 2013) using CDPK structure scaffolds in the presence and absence of calcium. This model is more consistent with recent *in vitro* and *in vivo* data pertaining to the autophosphorylation of CCaMK, and therefore is discussed in detail below.

1.6.5 Model of function

Shimoda and colleagues (Shimoda *et al.*, 2012) predicted the existence of a hydrogen bond network in LjCCaMK as described above. This network would then be disrupted by autophosphorylation at T265 (equivalent to *M. truncatula* T271) which would allow the protein to be active for downstream phosphorylation. However, this model is inconsistent with data obtained by Miller and colleagues (Miller *et al.*, 2013) which suggests that autophosphorylation at T271 in fact negatively regulates the protein. Several pieces of evidence are consistent with this representing the inactive state. First, the mutation or deletion of any of the EF-hands leads to a gain-of-function phenotype and therefore the EF-hands must negatively regulate the protein. Since calcium-binding to the EF-hands promotes T271 phosphorylation, it follows that this may be a negative regulator. These experiments were carried out using the native promoter and not an over-expression promoter (e.g. 35S) (Shimoda *et al.*, 2012) which could account for the differences

observed. Alternatively, the differences observed could be explained by the strength of the inoculum used (Ben Miller, personal communication). Second, the homology model of *M. truncatula* CCaMK (Miller *et al.*, 2013) predicts the hydrogen bond network involving the T271 residue similar to that predicted for LjCCaMK, but this model is instead stabilised by the presence of phosphorylation at T271. This involves hydrogen bonds between R323 and pT271 (2 bonds) and S322 and pT271 (1 bond). This difference between the two models occurs because the model in (Shimoda *et al.*, 2012) did not take phosphorylation into account. Finally, truncations of the protein which are autoactive showed low levels of phosphorylation at T271, therefore suggesting that pT271 is associated with the inactive state of CCaMK. Therefore, the stable inactive state of the protein involves pT271 and a hydrogen bond network formed between pT271 and the CaM-binding domain.

The model proposed by Miller and colleagues (Miller *et al.*, 2013) takes these data into account, along with the kinetics of calcium and CaM binding to CCaMK as described above (Section 1.6.3). This model is shown in Figure 1.9. At basal calcium levels, at least one EF-hand is predicted to bind calcium and this is sufficient for phosphorylation at T271 over time. The mechanism by which this autophosphorylation occurs is currently unknown. In contrast to the model proposed by Shimoda and colleagues, this state is the basal, inactive state of CCaMK. This predicted inactive state of CCaMK at basal calcium (125-150 nM) is stabilised by the hydrogen bond network described above including pT271 and any mutation to this residue would perturb this state and lead to a gain-of-function phenotype. This is because, in this particular case, a previously assumed phospho-mimetic residue (e.g. Asp) would not appropriately mimic the phosphorylated state. Therefore it would not form the same predicted hydrogen bonds, and consequently any mutation of this residue is predicted to disrupt the hydrogen bond network.

During a calcium-spike (625 – 850 nM), calcium binds to the four EF-hands of CaM and CaM can subsequently bind to CCaMK. Due to the proximity of the CaM binding site to the hydrogen bonding residues, this disrupts the hydrogen bonding network. This also overrides the negative regulation of phosphorylation at T271. In this configuration, CCaMK is active for target phosphorylation. In fact, calcium binding to the EF-hands primes the protein for activation because the resultant autophosphorylation enhances the affinity of CCaMK for calmodulin. This dual regulation of CCaMK by calcium and CaM binding allows

CCaMK to distinguish between basal calcium levels and calcium spiking, and therefore CCaMK is uniquely placed to interpret the calcium spiking signal in a robust manner.

A mechanism for the inactivation of CCaMK has also been proposed. Two residues in the CaM-binding domain have been identified which negatively regulate CCaMK function. Routray and colleagues (Routray *et al.*, 2013) and Liao and colleagues (Liao *et al.*, 2012) identified S344 and S343, respectively, as autophosphorylated residues. Phosphorylation of either of these residues has a negative impact on CCaMK function by inhibition of CaM binding, and therefore these residues could represent the "off switch" for the protein. However, the mechanism by which this occurs is currently unclear.
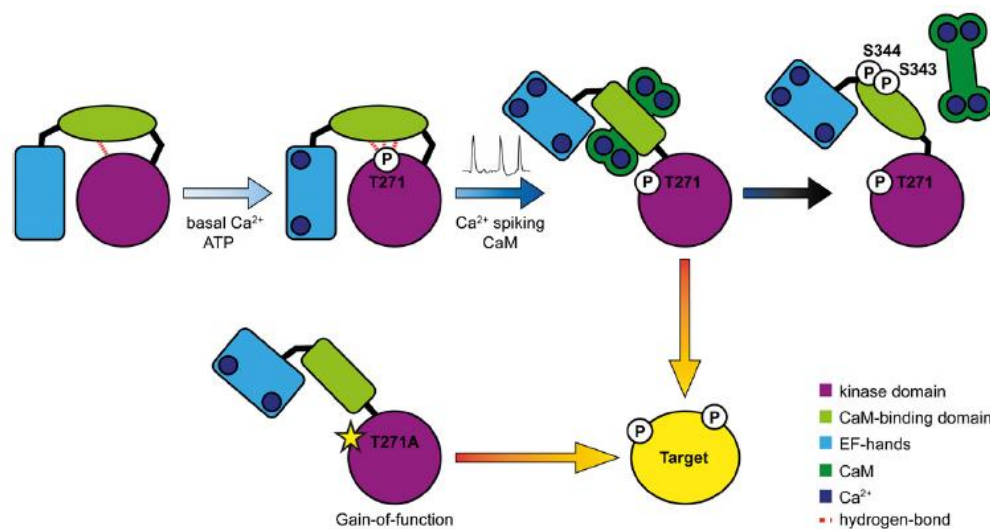
Figure 1.9. Schematic of MtCCaMK activation. Figure adapted from (Miller *et al.*, 2013).

1.6.6 CCaMK in different cell types and different plant species

CCaMK has been shown to have particular functions in different cell types in *M. truncatula*. *DMI3* is expressed in both the epidermis and cortex during nodulation, with up-regulation in dividing cortical cells prior to infection thread growth. Tissue-specific promoters were used to unpick the specific roles for CCaMK in these different cell types. It was demonstrated that epidermal expression of *DMI3* (in the *dmi3* mutant) was sufficient for infection thread formation, but this expression alone did not allow cortical cell division. Interestingly, combined expression of *DMI3* in the epidermis and cortex was sufficient for both cortical cell divisions and subsequent nodule organogenesis in addition to infection thread formation. However, nodule infection was not restored, perhaps due to inappropriate levels of the protein provided by the specific promoters. Despite this, these data show that CCaMK plays important roles in both early and late symbiosis signalling

(Rival *et al.*, 2012). In addition, the cell-specific roles of CCaMK during mycorrhization have been explored. The phenotypes observed when tissue-specific expression of *DMI3* was controlled in mycorrhization were analogous to those observed during nodulation; cortical expression alone did not rescue the mutant phenotype, however penetration of the fungus was observed with epidermal expression (Rival *et al.*, 2013). Therefore, CCaMK also has cell-autonomous roles in mycorrhization.

CCaMK has also been characterised in a number of species. CCaMK has been characterised in *O. sativa,* a species which can form mycorrhizae but not nodules. Godfroy and colleagues (Godfroy *et al.*, 2006) theorised that because nodulation likely evolved from the more ancient mycorrhization symbiosis, that CCaMK may be able to discriminate and interpret both rhizobial and mycorrhizal signals. Interestingly, they demonstrated that the *dmi3* mutant of *M. truncatula* can be recovered for nodule organogenesis by the CCaMK gene from *O. sativa.* This indicates that CCaMK is able to function in both these pathways despite the transfer from a non-legume to a legume, and suggests functional conservation in the evolution of the pathway. However, this conservation is not absolute because rhizobial infection was not complemented. On the other hand, the conservation of CCaMK was further confirmed by mutant analysis in *L. japonicus*. *OsCCaMK* was able to fully complement the equivalent mutant in *L. japonicus* for both mycorrhization and nodulation, including for infection (Banba *et al.*, 2008). This again indicates strong conservation between the pathways, and perhaps indicates a greater similarity between the *O. sativa* symbiotic pathway and that of *L. japonicus* as opposed to *M. truncatula.*

CCaMK is also found in the tropical, semi-aquatic legume species *Sesbania rostrata,* and its role in symbiotic signalling in this species has been characterised (Capoen *et al.*, 2009). *S. rostrata* shows a novel mechanism for bacterial invasion; intracellularly at cracks in the root epidermis. This is also characterised by an altered calcium spiking signal in which calcium spikes have a higher frequency and are more symmetrical. Interestingly, whilst CCaMK is conserved in *S. rostrata* it is not required for the pathway involving crack entry of bacteria. This was demonstrated because knockdown of CCaMK prevented nodule development but not intercellular infection. Therefore, whilst CCaMK is conserved in this species, it seems likely that it functions only in the conventional infection pathway and not the alternate infection pathway.

Chapter 1: Introduction

Interestingly, CCaMK has also been implicated in roles outside of symbiosis. As discussed above, the first role of CCaMK that was discovered was in anther development (section 1.7). However, subsequently further roles of CCaMK have been indicated in *O. sativa* and *Zea mays* in abscisic acid- and brassinosteroid-induced antioxidant defence ((Shi *et al.*, 2014), (Yan *et al.*, 2015), respectively), amongst others. This points to a potentially wider role of CCaMK in plants.

1.6.7 Specificity in nodulation and mycorrhization

The way in which specificity is encoded in symbiosis signalling is an important but as yet unanswered question. As described previously, recent work has shown it likely that the calcium spiking signatures during nodulation and mycorrhization are similar and therefore may not encode specificity. However, one possible difference between the pathways is in CCaMK regulation. Work by Shimoda and colleagues (Shimoda *et al.*, 2012) using a mutant of *L. japonicus* CCaMK which abolishes CaM binding (CCaMK^{FN-ED}) observed spontaneous nodulation, but these were un-infected nodules. However, surprisingly, the same mutant could be infected by mycorrhizal fungi. They concluded that the requirement for CaM binding to CCaMK differs between nodulation and mycorrhization, where it is dispensable for mycorrhization. However, several pieces of evidence call this conclusion into question. First, the CaM binding domain is conserved in all CCaMK proteins of non-legumes which have been identified, and therefore it follows that it most likely plays an important role in mycorrhization. Second, the CCaMK^{FN-ED} mutation autoactivates the protein and leads to a lack of CaM binding and yet conclusions about the relevance of CaM binding have been drawn. This appears to be counter-intuitive because the mutation has rendered CaM binding irrelevant for protein function (Miller *et al.*, 2013).

Further to this work, the same group (Takeda *et al.*, 2012) used a kinase-only form of *L. japonicus* CCaMK, which is a gain-of-function mutant. They showed that this mutant didn't complement bacterial entry (consistent with previous reports by Gleason and colleagues (Gleason *et al.*, 2006)) however the formation of pre-penetration apparatus in root tissue was retained. This was indicated because cytoplasmic reprogramming and nuclear enlargement were both detected. This suggests that arbuscular mycorrhizal fungi colonisation could be complemented by this mutant, although this was not directly demonstrated. If this is the case, it suggests a potential divergence in the regulation of the

pathways by CCaMK, where the kinase domain is sufficient in mycorrhizal signalling whereas the regulatory domains are required for nodulation.

Two alternative theories are that specificity in symbiosis signalling is encoded downstream of CCaMK or by an alternative pathway.

1.6.8 Interaction partners of CCaMK – CYCLOPS and CIP73

IPD3 (interacting protein of DMI3) was first identified by Messinese and colleagues (Messinese *et al.*, 2007) using a yeast-two hybrid system with CCaMK from *M. truncatula* as bait. They determined IPD3 to be a 58 kDa protein and showed that it co-localises with CCaMK to the nucleus and that they physically interact. Furthermore, the *IPD3* gene is highly conserved amongst angiosperms. However *A. thaliana* lacks a homologue in a manner analogous to *DMI3*. Yano and colleagues (Yano *et al.*, 2008) characterised the *cyclops* mutant of *L. japonicus* which shows a symbiotic phenotype of aborted infection with both rhizobia and mycorrhizal fungi. They determined that the *CYCLOPS* gene corresponded to the *IPD3* gene of *M. truncatula* and confirmed that the physical interaction of CCaMK and CYCLOPS in the nucleus occurs constitutively. Furthermore, they demonstrated that CYCLOPS is a phosphorylation target of CCaMK. They also showed that the *CYCLOPS* gene from *O. sativa* complements the *L. japonicus* mutant and therefore its role in nodulation and mycorrhization is highly conserved. Additionally, CYCLOPS mutants cannot be complemented either for rhizobial or mycorrhizal infection by CCaMK gain-of-function mutant, T265D, or WT CCaMK suggesting that CYCLOPS works in concert with CCaMK in the sym pathway (Hayashi *et al.*, 2010).

Further characterisation of *IPD3* mutants in *P. sativum* and *M. truncatula* (Ovchinnikova *et al.*, 2011), (Horvath *et al.*, 2011) confirmed that mutation of CYCLOPS leads to "leaky" phenotypes, i.e. nodulation/mycorrhization are initiated but then aborted. This is in contrast to the other Sym pathway mutants and therefore suggests that mutants of CYCLOPS are less penetrant than those of the others components (Capoen and Oldroyd, 2008). These data taken together show that CYCLOPS plays an important, but unknown, role in symbiosis signalling and further indicates that there may be genetic redundancy at this level in the pathway.

Recently, the function of CYCLOPS was elucidated by Singh and colleagues (Singh *et al.*, 2014). Interestingly, CYCLOPS functions directly as a transcription factor in a DNA sequence specific manner. Five phosphorylation target sites of CCaMK were identified on the CYCLOPS protein, two of which (S50 and S154) are needed for symbiosis signalling. When these two residues are mutated to Asp (phospho-mimetic), known as CYCLOPS-DD, this activates the protein to bind to the *NIN* promoter, specifically at the newly identified CYC-BOX sequence. This DNA-binding most likely occurs by a conformational change that occurs when the protein is phosphorylated at S50 and S154. The C-terminus of CYCLOPS was determined to contain both the activation domain and the DNA binding domain, and the N-terminus was determined to negatively regulate the protein activity. Crucially however, CYCLOPS-DD was shown to lead to a gain-of-function phenotype because in a *ccamk* mutant background, CYCLOPS-DD triggered spontaneous nodules. This therefore shows that phosphorylated CYCLOPS is able to bypass the requirement for CCaMK for nodule organogenesis and indicates that it is a key regulator of nodulation.

CIP73 (CCaMK-interacting protein of approximately 73 kDa) is also an interaction partner of CCaMK in symbiosis signalling (Kang *et al.*, 2011). This protein was similarly identified by yeast-two-hybrid screen, however in this case with just the kinase domain of *L. japonicus* CCaMK. CIP73 is a Scythe_N_ubiquitin-like protein and is also phosphorylated by CCaMK in the presence of Ca^{2+}/CaM. Furthermore, knockdown of CIP73 decreased nodule formation. However, the function of the protein is currently unknown. Recently, an interaction between CIP73 and a homologue of an animal HSC/HSP70-interacting protein (HIP) was identified in *L. japonicus,* and knockdown of the HIP gene led to an increase in nodule number. Therefore HIP represents a potential negative regulator of nodulation (Kang *et al.*, 2015).

It is likely that there are more, as yet unidentified, interaction partners of CCaMK because there is complexity in the system, i.e. multiple roles of CCaMK, which cannot be explained yet. For example, mutants of *cyclops* still retain Nod factor-induced root curling and cell division, and therefore CCaMK likely acts through additional target proteins during nodulation (Limpens and Bisseling, 2014). Furthermore, as discussed above, CYCLOPS binds to the NIN promoter after phosphorylation by CCaMK, however NIN is only induced during nodulation (Marsh *et al.*, 2007). Therefore, additional factors, which are currently not understood, must be involved upstream of NIN to confer downstream specificity.

1.7 CaMKII

Ca^{2+}/CaM-dependent protein kinase II (CaMKII) is a well-characterised and broadly-distributed metazoan Ser/Thr protein kinase. It has been used as a model for CCaMK due to the significant sequence and structural similarities between the two proteins. However, there are also significant differences in the mechanism of function between the two proteins. Key features of the mechanism of function and the regulation of CaMKII are therefore discussed in this section.

CaMKII has been identified in a wide range of species, from lower organisms e.g. *C. elegans* to mammals e.g. *Homo sapiens*. It performs a wide variety of roles in different tissues including, but not limited to, the regulation of carbohydrate, lipid and amino acid metabolism, calcium homeostasis, cytoskeleton regulation and transcription and translation (Hudmon and Schulman, 2002b). It is of particular importance in brain tissue and it is one of the most highly expressed proteins within neurons. Specifically, there is strong evidence that CaMKII is crucial for long term potentiation (LTP). LTP is the process by which long-lasting increases in the strength of synapses are triggered and this is thought to be a mechanism by which memory and learning occur within the brain. An increase in calcium is thought to activate CaMKII which in turn initiates the signalling cascade which leads to LTP (Lisman *et al.*, 2012). CaMKII also plays crucial roles in cardiac tissue and it has been shown to have a role in hypertrophy and heart failure (Zhang *et al.*, 2007).

1.7.1 Structure

CaMKII proteins consist of three main domains: an N-terminal catalytic domain, an autoregulatory domain which overlaps with a CaM binding domain and a C-terminal association domain (Figure 1.10). In mammals, different isoforms of CaMKII are found for example in humans there are four denoted α, β, γ and δ. These isoforms are all closely related and show 80-95% sequence similarity. These isoforms are differentially localised, with α and β situated in brain tissue and γ and δ localised in other regions of the body. Furthermore, these isoforms can be alternatively spliced which produces approximately 40 variants in total (Hudmon and Schulman, 2002a), (Hudmon and Schulman, 2002b), (Stratton *et al.*, 2013). One particular example of alternative splicing of CaMKII is discussed in section 1.8.3.
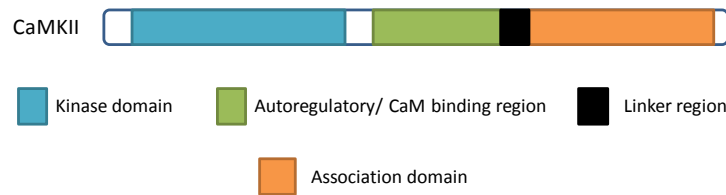
Figure 1.10 Domain structure of CaMKII. Figure adapted from (Hudmon and Schulman, 2002b).

The structure of CaMKII has been studied using a variety of methods. Early characterisation was carried out with transmission electron microscopy, which revealed that CaMKII forms an oligomer, which is usually dodecameric. The association domain is essential for this oligomerisation; whilst the association domains cluster together, the kinase domains and autoregulatory domains are projected to allow CaM-binding and kinase activity. The 12 subunits form into two rings consisting of hexamers of subunits, which sit on top of each other (Kolodziej *et al.*, 2000). Both hetero- and homo-oligomers of the CaMKII isoforms have been observed. More recently, crystal structures of the kinase and autoinhibitory domains separately have been resolved ((Rosenberg *et al.*, 2005), (Hoelz *et al.*, 2003) respectively) and remarkably a crystal structure of the entire ~700 kDa holoenzyme from *H. sapiens* has been resolved (Chao *et al.*, 2011). These structures, alongside small-angle X-ray scattering data, confirmed previous data suggesting the structure consisting of two hexameric rings with projected kinase domains. They also give further insight into the mechanism of function of CaMKII. It is currently unknown as to whether CCaMK forms a similar oligomeric structure.

1.7.2 Activation

CaMKII is maintained in an inactive state by interaction between the catalytic domain and the autoregulatory domain. This basal state is 100-1000 fold less than its activated state and occurs at basal calcium concentrations. The protein is maintained in this inactive state by binding of the α-helical autoregulatory domain which forms a pseudo-substrate and blocks both the substrate binding and ATP-binding sites. When this autoregulatory domain is removed, the kinase is constitutively active.

CaMKII is activated by Ca^{2+}/CaM binding because, in contrast to CCaMK, calcium ions cannot bind directly to the protein. Ca^{2+}/CaM binding to CaMKII releases the autoinhibition

by interrupting the interaction of the pseudo-substrate and the kinase domain. This activates the kinase. CaM binds to CaMKII in a 1:1 ratio, and both lobes of the protein wrap around the CaMKII subunit in a head to tail orientation. The conformational change that occurs when calcium binds to the EF-hands of CaM further increases its affinity for CaMKII. Further regulation involves the interaction of ATP and CaM binding; Ca^{2+}/CaM binding enhances the affinity for ATP and ATP binding stabilises Ca^{2+}/CaM binding, therefore creating a positive feedback loop (Hudmon and Schulman, 2002a), (Hudmon and Schulman, 2002b).

1.7.3 Autophosphorylation

Similarly to CCaMK, autophosphorylation is also key to the regulation of CaMKII. Phosphorylation at T286 in the autoregulatory domain is well characterised and has dramatic effects on the regulation of CaMKII. Phosphorylation at this residue is also Ca^{2+}/CaM-induced and interestingly renders the protein Ca^{2+}/CaM-independent, yielding the autonomous form of CaMKII. Phosphorylation at T286 is both necessary and sufficient for this. Concurrently, autophosphorylation at T286 increases the affinity of CaMKII for Ca^{2+}/CaM 1000-fold, known as CaM-trapping. This provides the protein with a "memory" of past spikes. Interestingly, the structure of the CaMKII oligomer is crucial to autophosphorylation at T286. It has been demonstrated that this autophosphorylation occurs in an inter-subunit but intra-oligomeric manner. This was characterised by co-expression of kinase dead and kinase active subunits within an oligomer (Mukherji and Soderling, 1994). Furthermore, this autophosphorylation can also only occur at adjacent subunits within the oligomer when CaM binds to both partners. This is because CaM binding is necessary to both activate the kinase activity of one subunit and to expose the autophosphorylation site on a neighbouring subunit (Hudmon and Schulman, 2002a), (Hudmon and Schulman, 2002b).

Autophosphorylation can also occur at residues in the CaM binding domain, including T305 and T306. Interestingly, phosphorylation at either of these residues is sufficient to block CaM binding and therefore also block Ca^{2+}/CaM-induced activation of CaMKII (Colbran and Soderling, 1990), (Hanson and Schulman, 1992). Therefore inactivation of CaMKII could occur in a manner analogous to that suggested for CCaMK, by phosphorylation of residues in the CaM binding domain.

1.7.4 CaMKII and calcium spiking

CaMKII, similarly to CCaMK, interprets a calcium spiking signal. Strikingly, the regulation of CaMKII (including CaM-trapping and autophosphorylation) means that the activation of CaMKII is dependent on calcium spiking frequency. This allows specificity of activation in response to the correct stimuli and enables the correct downstream responses to be produced. This is due to the ability of the protein to form an autonomous state when phosphorylated at T286. At low frequencies of spiking, CaM will bind but also dissociate and phosphorylation will be removed by the action of phosphatases. However, at a higher frequency of spiking, there is more likelihood of the binding of CaM to adjacent subunits and therefore autophosphorylation at T286 generating the autonomous state. This propagates the signal by positive feedback enhancing the CaM binding and leading to maximal activation. Therefore, CaMKII activity can be regulated by calcium spike frequency and in particular is controlled by the individual holoenzyme's rate of autophosphorylation at T286 (De Koninck and Schulman, 1998). Specificity of the signal can also be conferred by differential sensitivities of the isoforms to spiking and differential localisation of these isoforms.

1.8 Alternative splicing

Alternative splicing is thought to occur in over 60% of plant intron-containing genes, with that number likely to rise as more examples are discovered. In particular, it is thought that this number will rise as its roles in plant growth and development, physiology, metabolism and responses to environmental stresses and pathogens are explored (Syed *et al.*, 2012). Alternative splicing in plants, and especially potential roles in symbiosis signalling and establishment, has not been extensively studied. Due to the widespread occurrence of alternative splicing in plants and in many functional gene groups, it seems possible that it may play a role in symbiosis signalling.

1.8.1 Mechanism and common types of alternative splicing in plants

Splicing of pre-mRNAs to mature mRNAs involves the removal of introns and joining of exons. This is achieved by a large complex consisting of small nuclear ribonucleoprotein particles (snRNPs) and over 200 proteins, together known as the spliceosome. This has

been well studied in animal systems; however the composition of plant spliceosomes has not yet been elucidated. Introns each possess a 5′ and 3′ splice site, a branch point and polypyrimidine tract near the 3′ site which interact with core splicing factors. Introns are further characterised by their high AU content, in comparison to exons which are GC rich. Alternative splicing of pre-mRNAs is determined by splicing factors, the most common of which include Ser/Arg-rich (SR) proteins and heterogeneous nuclear ribo-nucleoprotein particle (hnRNP) proteins. These are trans-acting regulators, and they bind splice regulatory cis-elements in the pre-mRNA sequences which include exonic and intronic splice enhancers and silencers. These sequences have not been well-characterised in plants, however they function by recruiting and stabilising the spliceosome or recruiting inhibitory proteins in an analogous manner to animal systems. Alternative splicing has two main, important consequences. First, it can regulate transcript levels by nonsense-mediated decay by the incorporation of premature termination codons. Second, if the mRNA forms are reflected at the protein level, truncated proteins which lack regulatory domains can act as dominant negative protein regulators. This occurs, for example, by dimerization interactions which can affect protein function and these can impact on signalling pathways (Reddy *et al.*, 2013), (Staiger and Brown, 2013).

There are four main types of alternative splicing which are observed, and they are summarised in Figure 1.11. Exon skipping (also known as a cassette exon) involves the inclusion or exclusion of an exon in the final transcript. This is the most common form of splicing in *H. sapiens.* An alternative 5′ splice site can be used, which changes the 3′ boundary of the exon upstream. Similarly, an alternative 3′ splice site can be used which conversely changes the 5′ boundary of the downstream exon. Finally, intron retention can occur. This is the most common form of splicing observed in plant species, for example it occurs in approximately 45 and 39% of *A. thaliana* and *M. truncatula* splice events, respectively (Baek *et al.*, 2008). This difference between the predominant types of splicing in animals and plants can perhaps be explained by intron size difference; animal introns are on average 5 kbp, whereas plant introns are on average 160 bp. Different types of splicing can be observed for any given gene and splice variant (Reddy *et al.*, 2013), (Syed *et al.*, 2012).

Chapter 1: Introduction
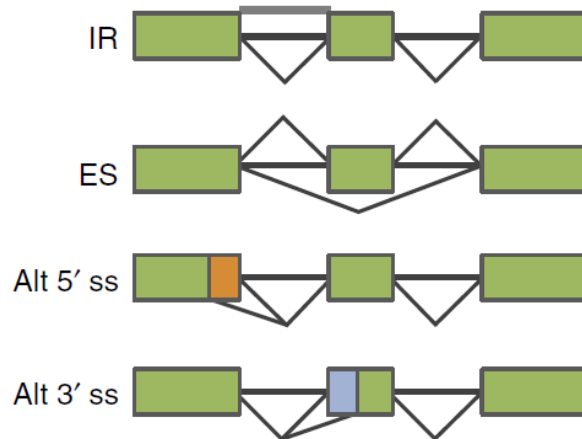
Figure 1.11. Types of alternative splicing observed in plants. Figure adapted from (Syed *et al.*, 2012).
IR – Intron retention, ES – exon skipping, Alt 5′ ss = alternative 5′ splice site and Alt 3′ ss = alternative 3′ splice site.

1.8.2 Alternative splicing in *M. truncatula* and symbiosis

RNA sequencing (RNA-seq) has been carried out for *M. truncatula* at various stages during the symbiosis with *S. meliloti.* From this, 6,587 genes were identified with two or more isoforms which suggests extensive alternative splicing in *M. truncatula*. These were often co-expressed in the same organ but were differentially regulated during nodulation (Boscari *et al.*, 2013). Furthermore, extensive RNA-seq has been carried out by Roux and colleagues (Roux *et al.*, 2014) using laser-capture microdissection which allows further sensitivity of the technique. *DMI3* and other symbiosis genes were shown to be expressed in mature nodules, therefore suggesting that Nod-factor signalling is still necessary at this late symbiotic stage.

Despite the fact that extensive characterisation of the role of alternative splicing in root symbioses has not been carried out, it has been observed in *M. truncatula.* MtHAP2-1 is a CCAAT-binding transcription factor and plays a key role in nodule development, possibly in the meristem. Expression of the conventionally-spliced protein peaks in nodules at 4 days post inoculation (dpi) however an alternatively-spliced form, which contains the first intron, was highly expressed at 21 dpi. This was shown to result in the expression of an upstream open reading frame (uORF) which is able to bind to the 5′ leader sequence of the *MtHAP2-1* transcript. Furthermore, this induces *MtHAP2-1* transcript degradation by an as yet unidentified mechanism therefore regulating its expression. The action of uORF and the

other regulator of MtHAP2-1, miR169, both act to restrict the expression of MtHAP2-1 to the nodule meristem and therefore regulate nodule development (Bayer *et al.*, 2002). This is the first known example of a uORF acting as a *trans* regulator of gene expression, but they have previously been demonstrated to act in a *cis* manner in other, particularly animal, systems. It is likely that further examples of alternative splice forms playing a role in root symbioses will be identified.

1.8.3 Alternative splicing of CaMKII

Alternative splicing has also been observed to have powerful effects on the function of CaMKII, the protein which shows significant sequence similarity to CCaMK. For example, the β subunit has several alternative splice forms including β_M, β and β_e'. These forms contain identical N-terminal kinase domains; however differ in the linker region between the kinase and the C-terminal association domains. Interestingly, despite the fact that the specific Ca^{2+}/CaM activity of the three isoforms are very similar, they show significantly different responses to calcium oscillations. The β_e' form has a higher activation constant for CaM than both β_M and β, which is thought to be due to the shorter length of the linker region. By contrast, the β_M and β have very similar activation constants for CaM despite differences in linker length. However, instead the β_M isoform has a significantly higher initial rate of autophosphorylation at T287 (equivalent to T286 of the α isoform) than the β isoform, and this is thought to be due to the fact that β_M form has a longer linker facilitating the spatial mobility of the kinase domain. These differences mean that the isoforms differ in their response to the frequency of calcium oscillations. Furthermore, the different isoforms show different expression patterns in developing and mature neural tissue, thereby allowing cell-specific differential responses to calcium oscillation signatures (Bayer *et al.*, 2002).

It is not currently know whether CCaMK is also regulated by alternative splicing. However, if this is the case the mechanism by which this occurs would likely be different.

1.9 Aims of this project

The mechanism by which CCaMK functions in the sym pathway of nodulation and mycorrhization has been characterised to some extent, as described above, however key questions remain. The aim of this project, therefore, was to further characterise *M. truncatula* CCaMK. *M. truncatula* was selected for study because it has previously been characterised at the John Innes Centre (Swainsbury *et al.*, 2012), (Miller *et al.*, 2013) and therefore there are established protocols for its use. Furthermore, *M. truncatula* is a widely used model legume plant for the study of symbiosis signalling, selected for several reasons including its small sequenced diploid genome, short generation time, ease of transformation and its interaction with the well-characterised symbiotic partner, *S. meliloti* (Cook, 1999).

First, an *in vitro* biochemical approach was taken in order to investigate the ways in which CCaMK is activated and de-activated. Two key ways in which CCaMK is regulated are autophosphorylation and calcium-binding. Autophosphorylation, specifically at T271, is a key regulator of CCaMK function not least because mutation of this residue to a non-phosphorylatable residue (Tirichine *et al.*, 2006) creates a gain-of-function mutation leading to spontaneous nodulation. However, the mechanism by which this occurs is not well established. By contrast, the mechanism of autophosphorylation of CaMKII, a similar animal protein, is well-established and key to the regulation of the protein: CaMKII autophosphorylates with an intra-oligomer inter-subunit mechanism in response to Ca^{2+}/CaM binding. In order to address this in CCaMK, a full-length form of the protein with an MBP-tag for stability was expressed and the oligomerisation status was determined. Subsequently a western blot approach was used with a state specific antibody for phosphorylation at T271 along with a combination of wild type and kinase-dead versions of the protein. The possibilities of inter/intra-monomeric and inter/intra-oligomeric phosphorylation were then explored. Furthermore, this has important consequences for the kinetics and protein concentration-dependence of the reaction.

As previously discussed, no structural data is currently available for CCaMK despite many attempts. Therefore, in order to further investigate the role that calcium-ion binding to the EF-hands plays in CCaMK function, two techniques were used to characterise the structural change which occurs. These were ANS-fluorescence and small angle X-ray scattering.

Furthermore, crystallisation trials were attempted with the aim of obtaining further, crucial structural information about CCaMK.

Second, the role of CCaMK *in planta* prior to, and during, symbiosis signalling was explored further. Initial experiments were used to determine whether purification of CCaMK was possible from the native host, when under the control of the native promoter, and, if so, with what yield. Subsequent work was used to determine if an over-expression line for CCaMK could be generated or if a GFP-CCaMK-SBP fusion protein could be purified from *M. truncatula.* Protein purified by this method could then be used to determine if the oligomerisation status established *in vitro* was also true *in planta*. Furthermore, protein purified by this method could potentially be more stable than that produced when expressed in *E. coli* and could therefore be used for structural work.

Finally, the role of CCaMK during the progression of symbiosis signalling was analysed. Preliminary data from a collaborator suggested that alternative splicing of *DMI3* occurs during the establishment of nodulation, and if this is reflected at the protein level *in planta* the consequences are potentially profound. For example, truncations of the protein can lead to its autoactivation and spontaneous nodulation (Gleason *et al.*, 2006). Furthermore, alternative splice-forms could potentially facilitate differential regulation of the protein in response to calcium spiking. Therefore a time-course analysis during nodulation and mycorrhization was carried out to determine if alternative splicing occurred at any time, and this was analysed at both the RNA and protein levels.

Chapter 1: Introduction

Chapter 2: Materials and Methods

2.1 Growth media

Compositions of bacterial, plant and fungal media are detailed in table 2.1.

Medium	Constituents /L
Lysogeny broth-Lennox (L)	Tryptone 10 g, yeast extract 5 g, NaCl 5 g, D-glucose 1 g For solid medium: agar 10 g
Lysogeny broth (Luria-Bertani) (LB)	Tryptone 10 g, yeast extract 5 g, NaCl 10 g For solid medium: agar 10 g
Auto-induction medium (AIM)	Yeast extract 10 g, tryptone 5 g, $(NH_4)_2SO_4$ 3.3 g, KH_2PO_4 6.8 g Na_2HPO_4 7.1 g, glucose 0.5 g, α-lactose 2 g, $MgSO_4$ 0.15 g, auto-induction medium micro element mix 0.35 g
Super orbital broth with catabolite repression (SOC)	Tryptone 20 g, yeast extract 5 g, NaCl 0.58 g, KCl 0.19 g, $MgCl_2$ 2.03 g, $MgSO_4.7H_2O$ 2.46 g, glucose 3.6 g
Buffered nodulation medium (BNM)	$CaSO_4.2H_2O$ 344 mg, MES buffer 390 mg, $MgSO_4.7H_2O$ 122 mg, KH_2PO_4 68 mg, $ZnSO_4.7H_2O$ 4.6 mg, H_3BO_3 3.1 g, $MnSO_4.4H_2O$ 2.9 mg, $NaMoO_4.2H_2O$ 0.25 mg, $CuSO_4.5H_2O$ 0.025 mg, $CoCl_2.6H_2O$ 0.025 mg, Na_2EDTA 18.65 mg, $FeSO_4.7H_2O$ 13.9 mg, pH 6
Modified FP (MOD FP)	$CaCl_2.2H_2O$ 0.1 g, $MgSO_4.7H_2O$ 0.1 g, KH_2PO_4 0.075 g, $Na_2HPO_4.12H_2O$ 0.1125 g, $FeC_6H_5O_7$ 6.25 mg, H_3BO_3 2.86 mg, $MnSO_4.4H_2O$ 2.03 mg, $ZnSO_4.7H_2O$ 0.22 mg, $CuSO_4.5H_2O$ 0.4 mg, H_2MoO_4 0.4 mg, pH 6 For solid medium: agar 5 g
Distilled water agar (DWA)	Agar 30 g
Rhizobium complete medium (TY)	Tryptone 5 g, yeast extract 3 g, $CaCl_2.6H_2O$ 1.325 g

Table 2.1. Components of media used in this study.

2.2 Bacterial strains

Bacterial strains used for cloning, protein expression, plant transformation and nodulation are listed in table 2.2.

Strain	Antibiotic resistance	Genotype	Source
E. coli BL21 (DE3)	-	F⁻ *ompT hsdS*$_B$ (r$_B$⁻m$_B$⁻) *gal dcm* (DE3)	Life Technologies
E. coli BL21 (DE3) pLysS	Cam^R	F- *ompT hsdS*$_B$ (r$_B$⁻m$_B$⁻) *gal dcm* (DE3) pLysS (Cam^R)	Life Technologies
E. coli BL21* (DE3)	-	F⁻ *ompT hsdS*$_B$ (r$_B$⁻m$_B$⁻) *gal dcm rne131* (DE3)	Life Technologies
E. coli Top 10	-	F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*) 7697 *gal*U *gal*K *rps*L (Str^R) *end*A1 *nup*G λ-	Life Technologies
E. coli DH5α	-	F⁻ φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (r$_k$-, m$_k$+) *pho*A *sup*E44 λ- *thi*-1 *gyr*A96 *rel*A1	Life Technologies
Agrobacterium rhizogenes 1193	Rif^R, Carb^R	-	John Innes Centre
Agobacterium tumefaciens AGL-1	Rif^R, Carb^R	-	John Innes Centre
S. meliloti 1021	Strep^R	-	John Innes Centre

Table 2.2. Bacterial strains used in this study.

2.3 Antibiotics used

Antibiotics used for the growth selection of *E. coli*, *A. rhizogenes, A. tumefaciens* and *S. meliloti* are listed in table 2.3.

Antibiotic	Solvent	Final concentration (µg/ml)
Carbenicillin	dH_2O	100
Spectinomycin	dH_2O	50
Rifampicin	Methanol	25
Chloramphenicol	Ethanol	34
Streptomycin	dH_2O	100
Kanamycin	dH_2O	20

Table 2.3. Antibiotics used in this study.

2.4 *In vitro* methods

2.4.1 Constructs used for protein production

Construct name	Plasmid backbone	Tag	CCaMK amino acids present	Antibiotic resistance	Reference
CCaMK	pET21a(+)	None	1-523	Carb[R]	(Swainsbury, 2011)
MBP-CCaMK	pMAL-c2X	MBP	1-523	Carb[R]	(Miller *et al.*, 2013)
MBP-CCaMK K47E	pMAL-c2X	MBP	1-523 (K47E mutation)	Carb[R]	This study
AI-VLD	pET101	None	318-523	Carb[R]	(Swainsbury *et al.*, 2012)
His-VLD	pET101	His	347-523	Carb[R]	(Swainsbury *et al.*, 2012)
MBP-CCaMK K47E and CCaMK	pCDFDuet-1	MBP (kinase dead only)	1-523 for both genes (K47E mutation in tagged form)	Spec[R]	This study
His-SUMO-CCaMK peptide (Version 1 and 2)	pOPINS3C	His-SUMO	261-344	Carb[R]	This study
SUMO D192N	pOPINS3C	His-SUMO	1-523 (D192N mutation)	Carb[R]	Sylvia Singh

Table 2.4. Constructs of CCaMK used in this study for protein expression.
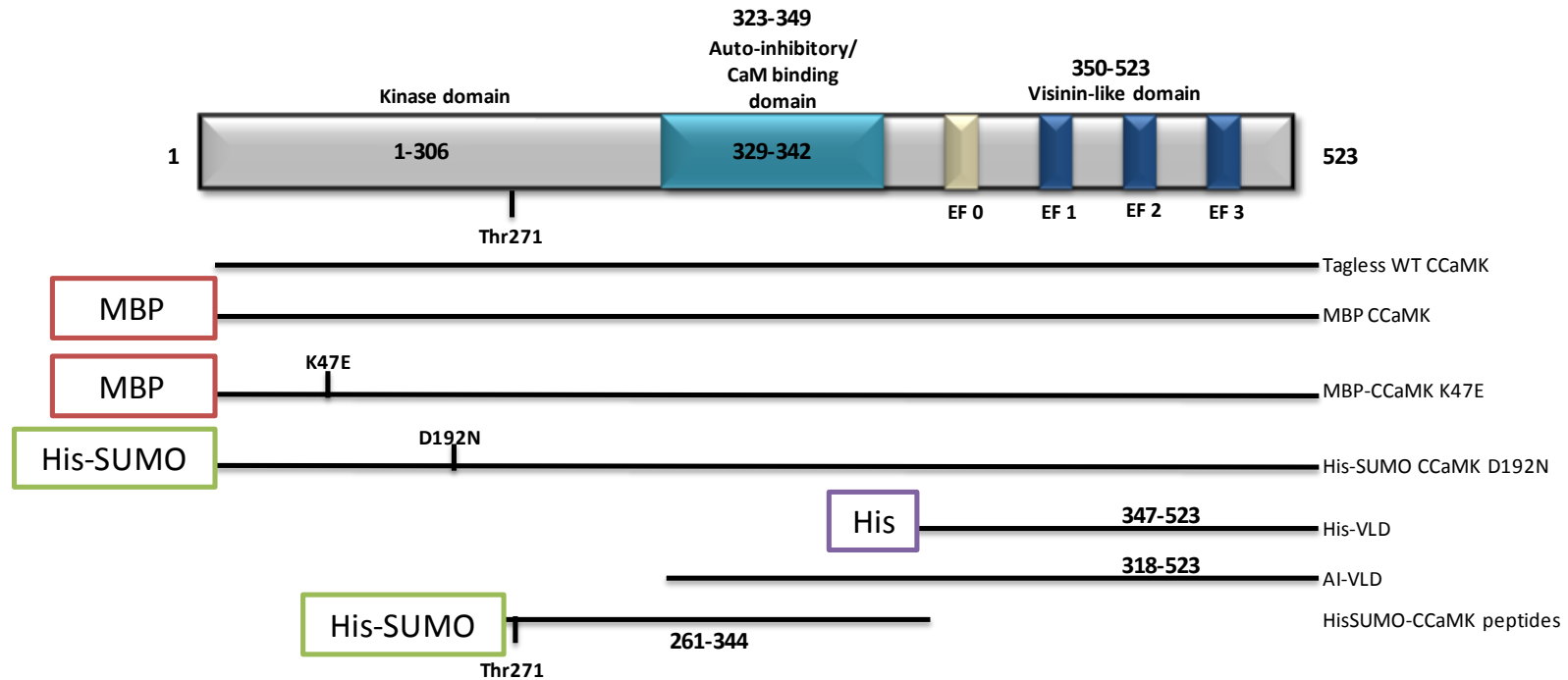
Figure 2.1. Schematic of CCaMK constructs used in this study.
The domains and main features of CCaMK are marked on the uppermost schematic. The constructs of CCaMK used for protein expression in *E. coli* are shown below, with any tags and mutations marked if present. For those constructs which did not contain the full sequence of CCaMK, the amino acids present are marked. The residues that correspond to the AI domain are denoted above the appropriate region, the residues that correspond to the CaM binding domain are notated on the box that represents this domain.

2.4.2 Agarose gel electrophoresis for DNA

DNA gel electrophoresis was performed in one of two ways. A 1% (w/v) Tris/acetic acid/EDTA agarose gel was made with ethidium bromide added to the molten agarose at a concentration of 0.25 µg/ml. Samples were mixed with 1 x MassRuler DNA Loading Dye (Life Technologies) and loaded into wells, alongside one well containing MassRuler DNA Ladder (Life Technologies). A potential of 90 V was applied for 45 min and the DNA was visualised by exposure to UV light at a wavelength of 365 nm.

Alternatively, a 1% (w/v) Tris/borate/EDTA agarose gel was made and stained after electrophoresis with 0.5 µg/ml ethidium bromide for 15 min. The first method was typically used for *E. coli* recombinant work, and the second method was used for RT-PCR experiments. The differences between the two methods did not affect the results.

2.4.3 Plasmid preparation

Plasmid DNA was purified using the QiaPrep Spin Miniprep Kit (Qiagen) according to the manufacturer's instructions. In order to obtain material for use with this kit, single colonies of *E. coli* containing the plasmid of interest were grown overnight in LB medium with the correct antibiotic selection. Cells were harvested by centrifugation at 3500 x g for 10 min and the cell pellet was used for plasmid preparation. Purified DNA was stored at -20 °C and concentrations were determined using a NanoDrop (Thermo Scientific).

2.4.4 Sequencing

Sequencing was performed in one of two ways. Sequencing was carried out with construct-specific primers with the following reaction mix: 1 µl plasmid DNA, 1.5 µl sequencing buffer, 3.5 µl of 1 µM primer, 1 µl 30% dimethyl sulfoxide, 2 µl dH$_2$O and 1 µl Big Dye v3.1 (Life Technologies). PCR cycling conditions were as follows: 96 °C for 2 min, 25 cycles of 96 °C for 10 seconds, 45 °C for 30 seconds and 60 °C for 4 min, and 1 minute at 15 °C. Samples were typically submitted to Eurofins for analysis.

Alternatively, if standard sequencing primers were used, purified plasmid DNA was directly submitted to Eurofins in a final volume of 15 µl with a concentration of 50-100 ng/µl. All sequencing data was visualised for analysis using Chromas2 (Technelysium) software.

2.4.5 Transformation of chemically competent cells

One vial of commercial chemically competent cells (Life technologies) was thawed on ice per reaction and subsequently incubated with 1 µl (5-10 ng) plasmid DNA for 30 min on ice. Cells were heat-shocked for 30 seconds at 42 °C, briefly incubated on ice and 250 µl pre-warmed SOC media was added. Cells were then incubated at 37 °C with agitation for one hour. An appropriate aliquot of the mixture was spread on LB/L agar with the appropriate antibiotic selection, and incubated overnight at 37 °C.

2.4.6 Cloning of CCaMK expression constructs

The *M. truncatula* CCaMK gene was codon-optimised for expression in *E. coli* (Swainsbury *et al.*, 2012) and this sequence was sub-cloned for use in the constructs in this study.

2.4.6.1 K47E mutation in MBP-CCaMK

The K47E mutation was generated in the MBP-CCaMK (pMAL-c2X) plasmid using the QuikChange Lightning site-directed mutagenesis kit (Agilent Technologies) according to manufacturer's instructions. The primers used are listed in section 2.6 and were generated with the QuikChange Primer Design program (Agilent Technologies).

2.4.6.2 MBP-CCaMK K47E and CCaMK WT in pCDFDuet-1 co-expression vector

Wild-type (WT) CCaMK and MBP-CCaMK K47E genes were amplified using PfuUltra High-Fidelity DNA Polymerase (Agilent Technologies) with primers detailed in section 2.6 and cloned into pCDFDuet-1 multiple cloning sites (MCS) one and two, respectively. Firstly, the WT CCaMK amplification product was digested with Nco1 and Not1 (Roche) according to manufacturer's instructions, and purified using the QiaQuick PCR Purification Kit (Qiagen). The pCDFDet-1 vector was digested by the same method, subjected to gel electrophoresis

and purified using the QiaQuick gel extraction kit (Qiagen). Ligation of the two purified products was carried out with T4 DNA ligase (New England Biolabs) according to manufacturer's instructions. *E. coli* Top 10 chemically competent cells were transformed with 2.5 µl of the ligation reaction, and plasmid DNA was isolated as previously described. The presence of the correct insert size was verified by a restriction digest with Nco1 and Not1 and subsequent gel electrophoresis. A vector containing the correct insert was identified by sequencing as previously described (primers listed in section 2.6).

The MBP-CCaMK K47E amplification product was digested with Xho1 and Nde1 (Roche) sequentially according to the manufacturer's instructions, and purified with the QiaQuick PCR Purification Kit (Qiagen). The pCDFDuet-1 vector containing the WT CCaMK insert in MCS1 was digested by the same method, subjected to gel electrophoresis and purified using the QiaQuick gel extraction kit (Qiagen). Ligation, transformation and plasmid purification were carried out as described above. A double digest was carried out with Xho1 and Nde1 to verify the presence of the correct insert size, and the vector was sequenced to confirm the presence of the correct insert as described above.

2.4.6.3 His-SUMO peptides in pOPINS3C

Two His-SUMO peptide constructs were made for use in crystallisation trials; version 1 without a cleavable His-SUMO tag, and version 2 with a 3C cleavage site. Both constructs were made by the following method with different forward primers (detailed in section 2.6). The sequence was amplified using PfuUltra High-Fidelity DNA Polymerase (Agilent Technologies) and purified using the QiaQuick gel extraction kit (Qiagen). The In-Fusion reaction was carried out by incubation of 1 µl linearised pOPINS3C vector, 3.5 µl (100 ng) purified insert, 3.5 µl dH$_2$O and 2 µl In-Fusion enzyme mix for 30 min at 42 °C. DH5α *E. coli* competent cells were transformed with 5 µl of the reaction mixture as previously described. Plasmid DNA was purified as before, and sequenced as described.

2.4.7 Expression and purification of CCaMK constructs

2.4.7.1 Tagless CCaMK

Full-length CCaMK was expressed in *E. coli* BL21 (DE3) pLysS cells grown in LB at 37 °C until the OD_{600} reached 0.3-0.6. Expression was induced with 0.5 mM IPTG for three hours at 28 °C. Cells were harvested by centrifugation at 7950 x g and stored at -20 °C. Cells were thawed on ice and resuspended in 50 mM Tris pH 7.5, 100 mM NaCl, 2 mM $CaCl_2$, 1 mM DTT, ~1 μg/ml DNase 1, lysozyme and a cOmplete™ protease inhibitor cocktail tablet (Roche). The cells were disrupted with a TS Series Benchtop 1.1 kW cell disruptor (Constant Systems Ltd.) at 25 psi and centrifuged at 38 760 x g for 20 min at 4 °C. The supernatant was filtered with a 0.45 μm Minisart high flow syringe filter (Sartorius), and applied to a CaM-Sepharose 4B column (GE Healthcare) pre-chilled to 4 °C and pre-equilibrated with 50 mM Tris pH 7.5, 100 mM NaCl, 2 mM $CaCl_2$ and 1 mM DTT. The column was washed with five column volumes of the same buffer, and bound CCaMK was eluted with 50 mM Tris pH 7.5, 100 mM NaCl, 2 mM EDTA and 1 mM DTT.

2.4.7.2 MBP-CCaMK and MBP-CCaMK K47E

MBP-CCaMK (and MBP-CCaMK K47E) was expressed in *E. coli* BL21 (DE3) cells grown in L with 0.1% glucose (v/v) at 37 °C until the OD_{600} reached 0.3-0.6. Expression was induced with 0.5 mM IPTG for two hours at 28 °C. Cells were harvested as before and resuspended in 20 mM Tris pH 7.4, 200 mM NaCl, 1 mM DTT, ~1 μg/ml DNase 1, lysozyme and a cOmplete™ protease inhibitor cocktail tablet (Roche). Cells were disrupted and centrifuged and the supernatant filtered as described in section 2.4.7.1. This was applied to an MBPTrap HP 1 ml column (GE Healthcare) pre-chilled to 4 °C and pre-equilibrated with 20 mM Tris pH 7.4, 200 mM NaCl, 1 mM DTT. The column was washed with ten column volumes of the same buffer and MBP-CCaMK was eluted with 20 mM Tris pH 7.4, 200 mM NaCl, 1 mM DTT and 10 mM maltose.

When a greater purity was required (for dynamic light scattering, atomic force microscopy and 8-anilino-1-naphthalenesulfonic acid-fluorescence), size exclusion chromatography (SEC) was also performed with a HiLoad 16/60 Superdex 75 prep grade column (GE Healthcare) with buffer containing 50 mM Tris pH 7.5, 5 mM $MgCl_2$, 2 mM EDTA, 200 mM NaCl, 10 % glycerol (v/v) and 1 mM DTT. For these three methods, the MBPTrap buffers were as described above, except they were used at pH 7.5.

2.4.7.3 AI-VLD

The auto-inhibitory and visinin-like domain (AI-VLD) was expressed in *E. coli* BL21* (DE3) cells grown in autoinduction medium (AIM) at 37 °C for two hours, followed by 20 °C for 20 hours. Cells were harvested as above and resuspended in 50 mM Tris pH 7.5, 1 M NaCl, 2 mM $CaCl_2$, 1 mM DTT, ~1 µg/ml DNase 1, lysozyme and a cOmplete™ protease inhibitor cocktail tablet (Roche). Cells were disrupted and centrifuged and the supernatant was filtered as described in section 2.4.7.1. This was applied to a CaM-Sepharose column pre-equilibrated with 50 mM Tris pH 7.5, 1 M NaCl, 2 mM $CaCl_2$ and 1 mM DTT. The column was washed with five column volumes of the same buffer and AI-VLD protein was eluted with buffer containing 50 mM Tris pH 7.5, 2 mM EDTA and 1 mM DTT. Additional purification by SEC was carried out with a HiLoad 16/60 Superdex 75 prep grade column (GE Healthcare) with buffer containing 50 mM Tris pH 7.5, 100 mM NaCl and 1 mM DTT.

2.4.7.4 His-VLD

His tagged visinin-like domain (VLD) protein was expressed as for AI-VLD (section 2.4.7.3), with the following exceptions. Bacterial cell pellet was resuspended in buffer containing 50 mM Tris pH 7.5, 300 mM NaCl, 10 mM imidazole, ~1 µg/ml DNase 1, lysozyme and a cOmplete™ protease inhibitor cocktail tablet (Roche). Supernatant was applied to a HisTrap HP 1 ml column (GE Healthcare) which was pre-equilibrated with 50 mM Tris pH 7.5, 300 mM NaCl and 10 mM imidazole. The column was washed with 10 column volumes of the same buffer, and His-VLD protein was eluted with 50 mM Tris pH 7.5, 300 mM NaCl and 500 mM imidazole. SEC was performed as for AI-VLD.

2.4.7.5 MBP-CCaMK-K47E and wt CCaMK co-expression

MBP-CCaMK-K47E and wild-type CCaMK were co-expressed in *E. coli* BL21 (DE3) pLysS cells grown in LB until the OD_{600} reached 0.3-0.6. Expression was induced with 1 mM IPTG for 2.5 hours at 28 °C. Cells were harvested as above. The composition of the buffers and the purification method was the same as for MBP-CCaMK (section 2.4.7.2) with the exception that buffers were used at pH 7.5.

2.4.7.6 His-SUMO-CCaMK peptides

Both His-SUMO-CCaMK peptides were expressed in *E. coli* BL21* (DE3) cells in AIM and were grown and induced for 7.5 hours at 37 °C. Cells were harvested as above and the bacterial cell pellet was resuspended in 20 mM Tris pH 8, 100 mM NaCl, 10 mM imidazole, ~1 μg/ml DNase 1, lysozyme and a cOmplete™ protease inhibitor cocktail tablet (Roche). Cells were disrupted and centrifuged and the supernatant was filtered as described in section 2.4.7.1. This was applied to a HisTrap HP 1 ml column (GE Healthcare) which was pre-equilibrated with 20 mM Tris pH 8, 100 mM NaCl and 10 mM imidazole. The column was washed with 10 column volumes of the same buffer, and bound protein was eluted with 20 mM Tris pH 8, 100 mM NaCl and 500 mM imidazole. No SEC step was necessary.

2.4.7.7 His-SUMO-CCaMK D192N

His-SUMO-CCaMK D192N was expressed in *E. coli* BL21 (DE3) pLysS cells grown in LB at 37 °C until the OD_{600} reached 0.3-0.6. Expression was induced with 0.5 mM IPTG for four hours at 28 °C. Cells were harvested, disrupted and centrifuged and the supernatant was filtered as described in section 2.4.7.1. The composition of the buffers used and the purification method was the same as for His-VLD protein (section 2.4.7.4), except that all purification was carried out at 4 °C and no SEC step was necessary.

2.4.8 SDS-PAGE of proteins

Pre-cast SDS-PAGE gels (Expedeon Ltd.) at 10, 12 and 16% SDS were used to separate proteins according to size. The percentage of the gel was selected depending on the expected size of protein in question. Samples were prepared by the addition of loading dye (200 mM Tris-HCl, pH 6.8, 8% SDS, 4% (v/v) glycerol, 400 mM DTT and 6 mM bromophenol blue) and were heated to 95 °C for five min. In general, 16 μl of each sample was loaded per well and a 2-color SDS marker (Expedeon) protein ladder was added in one well. Gels were run in RunBlue Tris-Glycine SDS Run buffer (Expedeon) for 1 hour at 185 V. Gels were stained for 15 min in Instant Blue protein stain (Expedeon) when not required for subsequent western blotting.

To confirm protein identities when necessary, gel bands were excised with a sterile scalpel, prepared according to standard procedures (Shevchenko *et al.*, 2006) and subjected to a trypsin digest for mass spectrometric characterisation using either a nanoLC-MSMS on an Orbitrap Fusion™ Tribrid™ Mass Spectrometer coupled to an UltiMate® 3000 RSLCnano LC system (Thermo Scientific, Hemel Hempstead, UK) or MALDI-TOF on a Bruker Ultraflex TOF/TOF. Mass spectrometric analysis was performed by Gerhard Saalbach (JIC).

2.4.9 Protein concentration determination

Protein concentration was determined with either DirectDetect (MerckMillipore) according to manufacturer's instructions, or a Lambda 25 UV/Vis spectrophotometer (Perkin Elmer). For the Lambda 25 determination, a zero reading was taken with a buffer only sample and the spectrophotometer was used in scanning mode. The absorbance value for protein concentration determination was taken at 280 nm, and the concentration was calculated using the Beer-Lambert law. Where necessary, the extinction coefficient was calculated using ProtParam.

2.4.10 Dynamic light scattering (DLS)

Fresh MBP-CCaMK protein was used immediately after purification prior to freezing. The protein sample was filtered with a Minisart high flow 0.2 μm syringe filter (Sartorius). A quartz cuvette was thoroughly cleaned using sterile water and isopropanol, and dried with compressed air. The DynaPro Titan (Wyatt Technology Corporation) was used with Dynamics V6 software. Laser power was increased to 100% and the quartz cuvette was confirmed clean with a Count Rate of less than 40,000 for a 13 μl sterile filtered water sample. The cuvette was cleaned and dried again and a 13 μl protein sample was added. Ten measurements were taken of 10 s each with 2 replicates at 20 °C.

2.4.11 8-Anilino-1-naphthalenesulfonic acid (ANS) fluorescence

MBP-CCaMK protein (made with Liang Zhou) was prepared with SEC as described in section 2.4.7.2. The protein sample was dialysed overnight at 4 °C with stirring in 10 mM Tris pH 7.5, 50 mM NaCl and 1 mM TCEP. A 10 ml 150 mM solution of ANS was made with 50 mg/ml NaOH and 50 mg/ml dimethylformamide and the concentration of MBP-CCaMK was

determined to be 10 µM. Four samples were used: 5 µM MBP-CCaMK + 2 mM CaCl$_2$, 5 µM MBP-CCaMK + 2 mM EDTA, ANS buffer + 2 mM CaCl$_2$ and ANS buffer + 2 mM EDTA.

ANS was added to give a final concentration of 150 µM in each sample. Fluorescence analysis was carried out with an LS55 luminescence spectrophotometer (Perkin Elmer) with the following parameters: excitation wavelength 480 nm, wavelength range 400 – 680 nm, slit size 10 nm, scan speed 500 nm/min and cuvette size 100 µl. The cuvette was cleaned in between each sample thoroughly with water and ethanol, and then dried with compressed air. Data were plotted using Microsoft Excel.

2.4.12 Atomic force microscopy (AFM)

MBP-CCaMK protein was prepared as described in section 2.4.7.2 and stored in the gel filtration buffer at 4 °C without freezing. Two different conditions for AFM were tested: no glycerol on a mica surface coated with poly-L-lysine or with 10% (v/v) glycerol in air on untreated mica. The conditions without glycerol were obtained via buffer exchange using a PD-10 column. Protein sample preparation was carried out with Liang Zhou and AFM was performed by Patrick Gunning and Andrew Kirby (IFR). Only the data corresponding to the condition with glycerol is presented because it was of higher quality. The average diameter of MBP-CCaMK was determined from a representative section of an enlarged version of the AFM image. All particles in the selected area were measured in mm and the axes were used to convert these measurements in µm. An average of these measurements was calculated.

2.4.13 Kinase and phosphatase assays and αpT271 western blotting

Kinase assays were performed for 30 min at 30 °C in a total reaction volume of 20 µl. Reaction mixtures for the negative control ("none") contained 50 mM HEPES, pH 7.5 and 1 mM DTT, and the calcium-induced reaction mixture in addition contained: 200 µM ATP, 10 mM MgCl$_2$ and 0.2 mM CaCl$_2$.. All proteins were purified at 4 °C, de-salted (if necessary, to 20 mM Tris pH 7.5, 200 mM NaCl, 1 mM DTT) and diluted to approximately 400 ng/µl. Protein was added in a volume of 3 µl, with 1.5 µl of each protein if two proteins were mixed in one assay. By contrast, when co-expressed MBP-CCaMK-K47E and WT CCaMK protein was used in the assay, 10 µl of freshly purified protein with no freeze-thaw step was used.

When required, a phosphatase assay was carried out with Lambda protein phosphatase (NEB) according to the manufacturer's instructions with a 30 min incubation at 30 °C. Reactions were terminated by the addition of 4 x SDS-PAGE sample buffer and incubation at 95 °C for 5 min.

Prepared protein samples from the kinase assay were separated using a 1.5 mm 10% SDS-PAGE gel and transferred to polyvinylidene difluoride membrane (Bio-Rad). A phosphorylation state-specific antibody for Thr271 CCaMK was used which was produced against the following sequence: CSFYEKpTWKGISQ (Eurogentec), and is referred to as αpT271. This antibody was produced by Akira Miyahara (Miller *et al.*, 2013). Immunodetection of proteins was performed using αpT271 (1:150) and a secondary antibody, horseradish peroxidase–conjugated anti-rabbit antibody (1:10 000). A chemiluminescent signal was generated by Pierce ECL Substrate (Life Technologies) and detected by autoradiography with Fuji Medical x-ray film (FUJIFILM).

Protein staining was carried out for western blots with 0.1% (w/v) PonceauS in 5% (v/v) acetic acid in water. The membrane was incubated with agitation in approximately 20 ml PonceauS stain, de-stained by washing with water prior to visualisation with a G:BOX (Syngene).

2.4.14 Small angle X-ray scattering (SAXS)

Small angle X-ray scattering was carried out for the AI-VLD and VLD in the presence and absence of calcium. However aggregation of both the proteins occurred to an un-satisfactory level. Therefore the data presented in this thesis is based on a previous data set obtained by David Swainsbury for the His-VLD only. However, new data was consistent with the previous data set. Data were collected at the X33 beam line on the DORIS III synchrotron at EMBL, DESY in Hamburg, Germany and data analysis was carried out by Dmitri Svergun.

His-VLD protein was dialysed into buffer containing 10 mM Tris pH 7.5, 50 mM NaCl, 1 mM TCEP with either 2 mM $CaCl_2$ or 0.1 mM EDTA. Readings were taken at various protein concentrations within a range of 0.5 - 7.4 mg/ml and 1.9 - 7.4 mg/ml with $CaCl_2$ and EDTA

respectively. Scattering curves were measured for the buffer, then the protein solution, then a second buffer blank. To monitor for the radiation damage, eight 15 s exposures of protein solutions were compared and no significant changes were observed. The data were normalized to the intensity of the transmitted beam and radially averaged; the scattering of the buffer was subtracted and the difference curves were scaled for protein concentration. Data processing was carried out with the program package PRIMUS. A homology model of the His-VLD was generated with SWISSMODEL against a human neuronal calcium sensor (Heidarsson *et al.*, 2012) (29% sequence identity) and further refined against the SAXS data using BUNCH.

2.4.15 Crystallography methods

2.4.15.1 VLD protein preparation with DAPase

His-VLD protein was purified as described in section 2.4.7.4 but without SEC. Protein was dialysed overnight at 7 °C with stirring in 10 mM Tris pH 7.5, 50 mM NaCl and 1 mM DTT. His-tag-removal was carried out with TAGZyme DAPase Enzyme (Qiagen) according to the manufacturer's instructions. In order to ensure complete digestion, protein was incubated with DAPase for 4.5 hours at 37 °C and then for 16 hours at 4 °C. The DAPase enzyme and co-purified contaminants were removed from the digested VLD protein by subtractive immobilized-metal affinity chromatography. Digestion and purity of the sample were confirmed by SDS-PAGE. Protein was concentrated to 15 mg/ml with Amicon Ultra-4 Centrifugal Filter Unit with Ultracel-10 membrane (Merck Millipore), frozen in aliquots with liquid nitrogen and stored at -80 °C until used.

Crystallisation trials were set up in 96 well MRC plates (Molecular Dimensions) with 50 µl of each screen condition added to each well by the Freedom evo liquid handling robot (Tecan). Each crystallisation drop consisted of 0.3 µl protein and 0.3 µl of the screen liquid, and was mixed in the sitting drop format by the OryxNano robot (Douglas Instruments Ltd.). The following screens were set up: PEG, ammonium sulphate (Qiagen), PACT, JCSG, Structure, Morpheus and Midas (Molecular Dimensions). Plates were covered and stored at 20 °C. Conditions that yielded potential protein crystals were followed up with optimisation screens in the 24-well hanging drop format in VDX optimisation plates (Molecular Dimensions). Wells were filled with 1 ml of the screen condition, and drops consisted of 1

µl of each condition mixed with 1 µl protein from the original stock (and, if appropriate, with an additional 0.2 µl crystal seed-stock). Further screens were set up for VLD protein in the presence of either $CaCl_2$ or EDTA. Protein preparation was as described above, except digested VLD protein was dialysed in 10 mM Tris pH 7.5, 50 mM NaCl, 1 mM DTT with either 1 µM $CaCl_2$ or 1 mM EDTA. Protein was concentrated to 11 and 12 mg/ml respectively and the same seven crystallisation screens were set up.

2.4.15.2 Surface entropy reduction of the VLD

Potential surface-exposed, high entropy amino acids in the VLD were selected for mutation to small, low entropy residues for potential crystallisation enhancement by the SERp server (Goldschmidt *et al.*, 2007). Three clusters were identified by the server, and the cluster with the highest SERp score was selected for mutation; specifically cluster 1 KEEEIE mutation to AAAEIE. Site-directed mutagenesis was carried out with the QuikChange Lightning site-directed mutagenesis kit (Agilent Technologies) according to the manufacturer's instructions with primers designed using the QuikChange Primer Design software (Section 2.6). Sequencing was carried out as described above to confirm the presence of the correct mutations.

SER VLD was expressed, purified and digested as described above for the VLD. Protein was concentrated to 19 mg/ml and the following 96-well screens set up as previously described: KISS (Custom screen comprising PEG and ammonium sulphate, Clare Stevenson), JCSG, Morpheus and Structure (Molecular Dimensions). Appropriate 24-well optimisation screens were set up as previously described.

2.4.15.3 SUMO peptide crystallography

HisSUMO-CCaMK peptide version 1 (with no 3C cleavage site) was purified as described in section 2.4.7.6, and dialysed overnight into 20 mM Tris pH 8 and 100 mM NaCl. Protein was concentrated to 15 mg/ml and the following 96-well crystallisation screens were set up: KISS (Custom screen comprising PEG and ammonium sulphate, Clare Stevenson), PEG (Qiagen), Structure, PACT and JCSG (Molecular Dimensions). The procedure was as previously described except screens were transferred to 96-well plates with a Liquidator™ 96 Manual Pipetting System (Mettler Toledo). Proteins were digested with 3C protease

(Banfield lab) at a final concentration of 10 µg/µl for 16 hours at 4 °C. Where necessary, 5 mM DTT was added or digestion was carried out at room temperature (~20 °C).

2.5 *In planta* methods

2.5.1 *M. truncatula* plant lines used

Plant line	Description	Reference
A17	Wild type line derived from the cultivar Jemalong	(Barker *et al.*, 1990)
Jester	Wild type ecotype ,~90% similar to Jemalong (Fertiprado, Portugal)	(Hill, 2000)
R108	Wild type ecotype	(Hoffmann *et al.*, 1997)
dmi3-1 (TRV25)	Mutant of CCaMK consisting of a 14 bp deletion at position 594. This leads to a premature stop codon and truncated protein (CCaMK 1-208). Derived from an A17 background.	(Levy *et al.*, 2004)
GFP-CCaMK-SBP 694/1	Coding sequence of *DMI3* under the control of the native promoter in a *dmi3-1* background, with an N-terminal GFP tag and a C-terminal SBP tag.	Developed by Akira Miyahara

Table 2.5. *M. truncatula* lines used in this study.

2.5.2 Hydroponic growth method

M. truncatula Jester seeds were surface-sterilised with 10% sodium hypochlorite solution for 2 -3 min then washed six times with dH$_2$O, and imbibed at ~20 °C for at least five hours. When other *M. truncatula* lines were grown hydroponically, an initial 1-2 minute scarification step with sand paper was added prior to sterilisation. The tank and mesh used for growth was sterilised with 10% sodium hypochlorite solution, rinsed with water six times, cleaned with 70% ethanol and rinsed with dH$_2$O. It was then filled with 16 litres of sterile BNM medium and two aeration bricks were spaced evenly throughout the medium and attached to an external air pump with tubing. Hydroponic tanks were located in a controlled environment room (22 °C, 75% relative humidity, ~250-300 µmol m^{-2} s^{-1}, 16-hour photoperiod). After imbibition, seeds were spread evenly across the mesh with a sterile spatula and moistened with dH$_2$O. The use of the airbricks also ensured the mesh was sufficiently moistened for germination; if necessary two transparent lids were used in addition to keep germinating seedlings moist for the first three days of seedling growth. Seedlings were grown for approximately three weeks, until substantial root growth

occurred with minimal over-crowding of the mesh. If necessary the level of the medium was replenished with dH_2O. Root material was harvested with a sterile scalpel, blotted briefly and frozen immediately in liquid nitrogen. Root material was stored at -80 °C until use.

A nodulation test was carried out as follows: ~100 growing A17 seedlings entrapped in the mesh were inoculated by immersion in 500 ml BNM medium containing rhizobia at OD_{600} = 0.001 for approximately three days. They were then returned to the main un-inoculated tank.

2.5.3 αMtCCaMK antibody generation

AI-VLD protein was prepared as described in section 2.4.7.3 and stored at -80 °C. This protein was concentrated appropriately with a pre-rinsed Amicon Ultra-15 Centrifugal Filter Unit with Ultracel-10 membrane (Merck Millipore). Protein was used for immunization in the SEC buffer consisting of 50 mM Tris pH 7.5, 100 mM NaCl and 1 mM DTT. Eight aliquots of 500 µl at 260 µg/ml were sent to Eurogentec on dry ice for immunization. The anti-protein speedy 28-day polyclonal programme was carried out, with two rabbit hosts. Antibody serum was stored at -80 °C on arrival, with working aliquots stored at -20 °C.

2.5.4 Western blotting with αMtCCaMK and αT271

Pre-cast 10% (w/v) SDS-PAGE gels (Expedeon Ltd.) were used to separate protein samples which were then transferred to nitrocellulose membrane (GE Healthcare). Immunodetection of proteins was performed using αMtCCaMK (1:10 000) and a secondary antibody, horseradish peroxidase–conjugated anti-rabbit antibody (1:10 000). A chemiluminescent signal was generated by Pierce ECL Substrate (Life Technologies) or SuperSignal West Femto Substrate and SuperSignal West Pico Substrate used together in a 1:10 ratio (Life Technologies). This was visualised using the ImageQuant™ LAS 500 (GE Healthcare).

The αT271 antibody corresponded to the non-specific fraction from the purification of the phospho-antibody, produced against CSFYEKpTWKGISQ (Eurogentec). This antibody

recognised both the phosphorylated and non-phosphorylated forms of T271. Western blotting with αT271 was carried out using the method above with the following exception: αT271 was used at a ratio of 1:150.

2.5.5 Generation and testing of an over-expression line

2.5.5.1 Golden Gate cloning of EC10265 binary vector

Cloning of the EC10265 vector was carried out using the Golden Gate system (Engler *et al.*, 2009) with components synthesised by GeneArt (Life Technologies). The level one vector was digested and ligated with the following reaction mix: 1 µl linearised backbone vector (100 ng/µl), 1 µl pLjUB (100 ng/µl), 1 µl MtCCaMK (100 ng/µl), 1 µl t35S (100 ng/µl), 0.15 µl 100 x BSA, 1 µl Bsa1, 1.5 µl 10 x T4 buffer, 1 µl T4 DNA ligase (New England Biolabs) and 7.35 µl dH$_2$O. Cycling conditions were as follows: 25 cycles of 3 min at 37 °C, 4 min at 16 °C, followed by 5 min at 50 °C and 5 min at 80 °C. *E. coli* DH5α cells were transformed with 1 µl of the ligation reaction as described in section 2.4.5. Colony PCR was performed with GoTaQ Green and the following reaction mix: 1 µl each FOR- DMI3 and REV-DMI3 primers (20 µM), 3 µl dH$_2$O, 5 µl GoTaQ Green and one colony per reaction. PCR cycling conditions were as follows: 95 °C for 10 min, 30 cycles of 95 °C for 10 seconds, 53 °C for 20 seconds and 72 °C for 3 min, with a final extension step of 5 min at 72 °C. The corresponding *E. coli* colonies were grown overnight in LB with appropriate antibiotic selection. Agarose gel electrophoresis, plasmid preparation and sequencing were carried out as described in sections 2.4.2, 2.4.3 and 2.4.4.

The level two vector was constructed using the same method, with the following exception. The digestion and ligation mix was as follows: 1 µl linearised backbone vector pAGM 4723 (100 ng/µl), 1 µl BASTA (100 ng/µl), 1 µl level 1 vector (100 ng/µl), 1 µl dsRED (100 ng/µl), 1 µl ELE3 end linker, 0.15 µl 100 x BSA, 1 µl Bpi1, 1.5 µl 10 x T4 buffer, 1 µl T4 DNA ligase (New England Biolabs) and 6.35 µl dH$_2$O. Sequencing primers used for this construct are detailed in section 2.6.5.

2.5.5.2 Hairy root transformation

A 20 μl aliquot of *A. rhizogenes* strain AR1193 was transformed with 1.5 μl (150 ng) purified plasmid DNA by the following procedure. DNA and cells were incubated on ice for 5 min, transferred to an electroporation cuvette (Cell Projects, UK) and subjected to an electroshock of 1.25 V, 25 μ F and 200 Ω (Gene-Pulser, Bio-rad). Cells were re-suspended in 500 μl SOC medium, grown for 1 hour at 28 °C with agitation and plated onto TY medium containing the appropriate antibiotic selection. Plates were incubated for two days at 28 °C.

Single colonies were grown for two or three days in TY medium containing the appropriate antibiotic at 28 °C with agitation. Two ml culture was centrifuged at 12 470 x g for 30 seconds, and the cell pellet was resuspended in 400 μl TY medium. Germinating seedlings of *dmi3-1* were prepared by removal of 3 mm of the radical with a sterile scalpel. Cut seedlings were dipped briefly into the *A. rhizogenes* culture, and replanted on MOD FP plates. These were sealed with 3M Micropore tape (Slaughter Ltd, R & L) and grown for three weeks in a controlled environment room (22 °C, ~300 μmol m^{-2} s^{-1}, 16-hour photoperiod).

2.5.5.3 dsRED fluorescence and nodulation assay tests

Roots were scored for successful transformation by fluorescence microscopy (Leica MZFLIII) with a green filter for visualisation of the dsRED marker. Plants containing hairy roots with strong fluorescence were used for a nodulation complementation assay with *S. meliloti* strain 1021. *S. meliloti* was cultured in liquid TY media with the appropriate antibiotic selection two days prior to use, and then diluted to OD$_{600}$ = 0.02-0.05 in 400 ml BNM medium. Seedlings were dipped in rhizobia prior to planting in Phytratray II growth pods (Sigma-Aldrich) containing a 1:1 mix of sand:terragreen (Oil-dry company, UK). Plants were grown in a controlled environment room (22 °C, ~300 μmol m^{-2} s^{-1}, 16-hour photoperiod) and nodulation was scored three weeks after inoculation. Images of nodulation with dsRED fluorescence were taken with an Infinity 1 camera (Luminera).

2.5.5.4 Stable line production in R108

A. tumefaciens strain AGL-1 was transformed with 1.5 µl (100 ng/µl) purified plasmid DNA in section 2.5.5.2. Transformation of *M. truncatula* R108 was carried out according to standard methods by Matthew Smoker (TSL).

2.5.5.5 Testing of potential transformants

Plantlets which developed from embryos grown on antibiotic selective media were re-planted in soil by horticultural services at the John Innes. Genomic DNA extractions were carried out on leaf tissue using the DNeasy Plant Mini Kit (Qiagen) by Richard Gorem. Concentrations of the DNA extractions were determined with a NanoDrop (Thermo Scientific). PCR was performed with GoTaQ Green and the following reaction mix: 2 µl gDNA, 6 µl dH$_2$O, 1 µl of each DMI3-F new and DMI3-R-nostop primer (20 µM) and 10 µl GoTaq Green 2x Master Mix. PCR cycling conditions were as follows: 95 °C for 10 min, 35 cycles of 95 °C for 10 seconds, 53 °C for 20 seconds and 72 °C for 2 min, with a final extension step of 5 min at 72 °C. Agarose gel electrophoresis was performed as described in section 2.4.2.

2.5.6 GFP-CCaMK-SBP stable line testing

A potentially transgenic line in a *dmi3-1* background was tested for successful transformation with pCCaMK:GFP-CCaMK-SBP. Seeds were harvested from 16 T$_1$ generation plants and subjected to a nodulation complementation test. Seeds were sterilised as described in section 2.5.2, with an additional scarification step with sandpaper for 1-2 min prior to sterilisation. They were then plated onto DWA, stratified in the dark at 4 °C for two days and then germinated overnight at room temperature. Twelve seedlings for each line were planted in sand:terragreen (Oil-dry company, UK) in 60-well trays and were covered with transparent lids for three days. After seven days, seedlings were inoculated with rhizobia at OD$_{600}$ = 0.04 by direct application of 2 ml culture to the base of the seedling. They were then scored three weeks post inoculation for the presence of nodules.

One line from the T_2 generation was selected for experiments on the basis of successful nodulation and a strong western blot protein band at the correct size for the fusion protein, probed with the αMtCCaMK antibody. This line was subjected to a copy number analysis by iDNA Genetics using a TaqMan-based analysis which confirmed the line was heterozygous for the insert. Seeds were bulked for the line by horticultural services at the John Innes.

2.5.7 Purification methods used for MtCCaMK

2.5.7.1 Buffer trials – RIPA/Li buffers and YS/Li

To compare the extraction of RIPA and Li buffers (method 1), frozen Jester root tissue was ground to a fine powder in a pre-chilled pestle and mortar with liquid nitrogen and divided equally between two buffers. Either RIPA buffer (50 mM Tris pH 7.5, 150 mM NaCl, 1 mM EDTA, 1% (v/v) Triton X-100, 0.5% (v/v) sodium deoxycholate, 0.05% (v/v) SDS, 5 mM DTT, 1 mM EGTA, cOmplete™ protease inhibitor cocktail tablet (Roche)) or Li buffer (150 mM Tris pH 7.8, 50 mM potassium acetate, 20 % (v/v) glycerol, 4 mM DTT, ~1 µg/ml DNase 1, cOmplete™ protease inhibitor cocktail tablet (Roche)) (Li *et al.*, 1996) was added at approximately 2 ml/g tissue. Samples were subjected to cell disruption at 20 psi, incubated on ice for 5 min and centrifuged at 38 760 g for 20 min at 4 °C. Samples were prepared for SDS-PAGE as previously described, and separated using a pre-cast 16% (w/v) SDS-PAGE gels (Expedeon Ltd.). Western blotting probed with αMtCCaMK was performed as described in section 2.5.4.

To compare the extraction of YS and Li buffers (method 2), frozen Jester root tissue was ground to a fine powder in a pre-chilled pestle and mortar with liquid nitrogen and divided equally between two buffers. Either YS buffer (50 mM Tris pH 7.4, 150 mM NaCl, 1 mM EDTA, 5 mM DTT, 1% (v/v) IGEPAL CA-630, 20% (v/v) glycerol, cOmplete™ protease inhibitor cocktail tablet (Roche)) or Li buffer was added at 2 ml/g tissue. Samples were incubated with agitation at 7 °C for 1 hour, prepared for SDS-PAGE as before and subjected to western blot probed with αMtCCaMK as described in section 2.5.4.

2.5.7.2 MonoQ purification

Jester extract was prepared as in method 1 (section 2.5.7.1) with a Li buffer extraction and dialysed overnight at 4 °C with stirring in 20 mM Tris pH 7.8, 10% (v/v) glycerol and 1 mM DTT. The sample was filtered with a Minisart high flow syringe filter, 0.45 μm (Sartorius), and applied to a Mono Q 5/50 GL column (GE Healthcare) pre-equilibrated with 20 mM Tris pH 7.8, 10% (v/v) glycerol and 1 mM DTT. The column was washed with five column volumes of the same buffer, and bound proteins were eluted with 20 mM Tris pH 7.8, 1 M NaCl, 10% (v/v) glycerol and 1 mM DTT.

2.5.7.3 DEAE purification

Jester extract was ground to a fine powder in a pre-chilled pestle and mortar with liquid nitrogen and resuspended in Li buffer at 4 °C. In this case, the sample was ground further with a 1 L ball mill (Capco Test Equipment) run at 7 °C for 18 hours. The sample was then subjected to cell disruption at 20 psi and centrifuged at 38 760 g for 20 min at 4 °C. The sample was dialysed twice at 7 °C with stirring in 20 mM Tris pH 7.8 and 1 mM DTT. The sample was filtered with a 0.45 μm Minisart high flow syringe filter (Sartorius), and applied to a HiTrap DEAE Sepharose FF column (GE Healthcare) pre-chilled at 4 °C pre-equilibrated with 20 mM Tris pH 7.8 and 1 mM DTT. The column was washed with five column volumes of the same buffer, and bound proteins were eluted with 20 mM Tris pH 7.8, 1 M NaCl and 1 mM DTT. For each new purification a new DEAE column was used that was cleaned according to manufacturer's instructions prior to use.

2.5.7.4 CaM-Sepharose purification

Jester extract was ground as described in section 2.5.7.3 (without ball milling) and resuspended in 1 ml/g Li buffer at 4 °C. The sample was subjected to cell disruption and centrifugation as previously described. The supernatant was dialysed overnight at 7 °C with stirring in 50 Tris pH 7.5, 100 mM NaCl, 2 mM CaCl$_2$ and 1 mM DTT. The sample was filtered as before and applied to a CaM-Sepharose 4B column (GE Healthcare) pre-equilibrated with 50 Tris pH 7.5, 100 mM NaCl, 2 mM CaCl$_2$ and 1 mM DTT. The column was washed with five column volumes of the same buffer, and bound proteins were eluted with 50 Tris pH 7.5, 100 mM NaCl, 2 mM EDTA and 1 mM DTT.

2.5.7.5 DEAE followed by CaM-Sepharose purifications

Jester extract was prepared and DEAE purification carried out as described in section 2.5.7.3 with the following exceptions. Ground tissue was resuspended in buffer containing 20 mM Tris pH 7.8, 5% (v/v) glycerol, 1 mM DTT, ~1 µg/ml DNase 1 and cOmplete™ protease inhibitor cocktail tablet (Roche) and the supernatant was not subjected to a dialysis step prior to application to the HiTrap DEAE Sepharose FF column (GE Healthcare).

Appropriate protein fractions from the DEAE purification were dialysed overnight at 7 °C with stirring in 50 Tris pH 7.5, 100 mM NaCl, 2 mM $CaCl_2$ and 1 mM DTT. CaM-Sepharose purification was carried out as described in section 2.5.7.4 with the following exception: purification was carried out at 4 °C.

2.5.7.6 GFP-CCaMK-SBP purification trials (DEAE, Streptavidin, CaM-Sepharose and SEC)

DEAE purification was carried out for GFP-CCaMK-SBP protein as described in section 2.5.7.3 with the following exceptions. Ground tissue was resuspended in buffer at 4 ml/g and resuspended tissue was extracted for 1 hour at 7 °C with stirring, instead of a ball mill extraction.

GFP-CCaMK-SBP root tissue was prepared for streptavidin-column purification using the same method as for DEAE purification in this section. The supernatant was dialysed overnight at 7 °C with stirring in 20 mM Tris pH 7.5, 100 mM NaCl and 1 mM DTT. The sample was filtered as before and applied to a HiTrap Streptavidin HP column (GE Healthcare) pre-chilled to 4 °C and pre-equilibrated with buffer containing 20 mM Tris pH 7.5, 100 mM NaCl and 1 mM DTT. The column was washed with ten column volumes of the same buffer, and bound proteins were eluted with 20 mM Tris pH 7.5, 100 mM NaCl, 10 mM d-Desthiobiotin and 1 mM DTT.

CaM-Sepharose purification of GFP-CCaMK-SBP was performed as described in section 2.5.7.4 with the following exceptions. Li buffer was added to ground tissue at 4 ml/g, resuspended tissue was extracted for 1 hour at 7 °C with spinning prior to cell disruption and purification was carried out at 4 °C.

SEC chromatography was carried out on appropriate fractions from the CaM-Sepharose purification. Purification was carried out at room temperature (approximately 20 °C) with a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare) with buffer containing 50 mM Tris pH 7.5, 200 mM NaCl and 1 mM DTT.

2.5.8 Growth method for nodulation and mycorrhization time-course

Seeds for Jester, A17 and GFP-CCaMK-SBP 694/1 seeds were used for time-course experiments. A17 and GFP-CCaMK-SBP 694/1 were scarified with sandpaper for 1-2 min, and seeds for all three lines were subsequently surface sterilised with 10% sodium hypochlorite solution for 2 -3 min. Seeds were then washed six times with dH$_2$O, and imbibed for at least three hours. They were then plated onto DWA, stratified in the dark at 4 °C for three days and then germinated overnight at room temperature. Germinated seedlings were transplanted to MOD FP pH 6 agar plates (6/plate) on a layer of sterile filter paper and sealed with 3M micropore tape. Seedlings were grown on plates for 10 days at (22 °C, ~300 µmol m^{-2} s^{-1}, 16-hour photoperiod).

For nodulation and mycorrhization time-courses, seedlings from the MOD FP plates were transferred to an autoclaved 1:1 mix of sand:terragreen (Oil-dry company, UK) in 40-well trays. Day 0 root samples for each line were harvested in liquid nitrogen prior to potting out in terragreen/sand. Seedlings for both nodulation and mycorrhization time-courses were grown in a controlled environment growth room (22 °C, 75% relative humidity, ~250-300 µmol m^{-2} s^{-1}, 16-hour photoperiod). Transparent lids were used to cover seedlings for three days after their transfer to terragreen:sand, and they were subsequently watered with RO water.

Nodulation of seedlings was carried out with *S. meliloti* strain 1021. *S. meliloti* was cultured in liquid TY media with appropriate antibiotic selection two days prior to use, and then diluted to OD 600 = 0.03-0.05 in 400 ml BNM medium. Seedlings were dipped in rhizobia prior to pricking out, and root samples were harvested four, 10 and 21 days post inoculation. Seedlings were rinsed, blotted and roots removed with a sterile scalpel prior to immediate freezing in liquid nitrogen. Typically roots from five plants were pooled and frozen together.

Mycorrhization of seedlings was carried out with SolRize Pro *G. intraradices* inoculum (Agrauxine) used at 1:12 (v/v spore:terragreen/sand). Seedlings were planted directly into the terragreen/sand containing the inoculum. The progression of mycorrhization was monitored by staining of representative root samples. Roots were pre-cleared by incubation in 10% (w/v) KOH at 95 °C for 10 min, rinsed three times with water and stained with 5% ink (with 5% acetic acid) (v/v) for 6 min at 95 °C. Roots were de-stained with water, and visualised with a Leica M80 Stereo microscope (Leica) and images taken with a Retiga-2000R CCD camera (QImaging). Colonisation was scored with the gridline intersect method (Giovannetti and Mosse, 1980). Root samples were therefore harvested seven, 13 and 21 days post inoculation.

2.5.9 Immunoprecipitation

Frozen root tissue was ground in liquid nitrogen and incubated with 2 ml/g extraction buffer (Li) for one hour at 7 °C with agitation. Samples were taken at this stage for crude extract western blot analysis with αMtCCaMK. In some cases, cell disruption of samples at 20 psi was carried out (not for time-course experiments). Samples were then centrifuged at 38 760 x g for 20 min at 4 °C and filtered through two layers of Miracloth (Calbiochem). The supernatant was centrifuged again at 38 760 g for 10 min at 4 °C and filtered through two layers of Miracloth. The supernatant was used for immunoprecipitation.

αMtCCaMK antibodies were covalently immobilized to Pierce™ NHS-Activated Magnetic Beads (Life Technologies) according to the manufacturer's instructions. Prepared samples (or a purified protein fraction) were incubated with the appropriate volume of conjugated beads (40 – 80 µl) proportional to the sample volume. The storage buffer was removed from the beads and samples were incubated for five hours with rotation in a falcon tube (Corning) using a Techne Hybridiser HB-2D at 7 °C, after an input sample was taken. Antibody coupled beads were collected using a magnet, and the supernatant was stored as the flow-through sample. Beads were washed three times with Li buffer, and protein was eluted with 100 µl 0.1 M glycine pH 2 for five min with agitation. Eluted protein samples were neutralised with 20 µl 1 M pH 10 glycine, and 20 µl of each sample was prepared for SDS-PAGE and western blotting as described in section 2.5.3.

2.5.10 RNA preparation and cDNA synthesis

Plant root tissue was stored at -80 °C and ground in an autoclaved and pre-chilled pestle and mortar with liquid nitrogen. Ground tissue was placed into RNA-free tubes pre-chilled on dry ice, and RNA was extracted with the RNeasy plant mini kit (Qiagen). RNA concentration was determined using a Nanodrop-1000 (Thermo Scientific), and the A260/280 ratio was checked for quality (1.8-2.2). Three micrograms of RNA for each sample was treated with DNase-1 (Sigma) in a 30 µl reaction volume according to the manufacturer's instructions. The quality of the DNase-1 treated RNA was checked by agarose gel electrophoresis. A 1% agarose gel was loaded with 1 µg DNase-1 treated RNA, run for 20 min at 120 V and post-stained with 0.5 µg/ml ethidium bromide for 15 min. RNA was tested for contaminating DNA by PCR with *DMI3* primers with reagents as follows: 1 µl RNA, 1 µl each DMI3-F new and DMI3-R-nostop (20 µM), 7 µl dH$_2$O and 10 µl GoTaq Green 2x Master Mix (Promega). Cycling was as follows: 95 °C for 10 min, 30 cycles of 95 °C for 10 seconds, 53 °C for 20 seconds and 72 °C for 4 min, with a final extension step of 5 min at 72 °C.

cDNA synthesis was carried out with the iScript cDNA synthesis kit (Bio-rad) for 1 µg RNA for each sample according to manufacturer's instructions. RNA samples were stored at -80 °C and cDNA samples were stored at -20 °C.

2.5.11 RT-PCR

RT-PCR was performed with GoTaQ Green and the following reaction mix: 2 µl cDNA, 6 µl dH$_2$O, 1 µl of each DMI3-F new and DMI3-R-nostop primer (20 µM) and 10 µl GoTaq Green 2x Master Mix. PCR cycling conditions for 2 minute extension time experiments were as follows: 95 °C for 10 min, 30 cycles of 95 °C for 10 seconds, 53 °C for 20 seconds and 72 °C for 2 min, with a final extension step of 5 min at 72 °C. PCR cycling conditions for 1 minute extension times were as follows: 95 °C for 10 min, 35 cycles of 95 °C for 10 seconds, 58 °C for 20 seconds and 72 °C for 1 minute, with a final extension step of 5 min at 72 °C. PCR cycling conditions for 30 second extension time experiments were as follows: 95 °C for 10 min, 35 cycles of 95 °C for 10 seconds, 58 °C for 20 seconds and 72 °C for 30 seconds, with a final extension step of 5 min at 72 °C.

2.5.12 PCR product cloning and sequencing

PCR product cloning was carried out with the PCR product cloning kit (Qiagen) according to manufacturer's instructions. PCR products were used directly in the ligation reaction, with no prior purification. *E. coli* Top 10 chemically competent cells were transformed with the ligation reaction as described in section 2.4.5. Colonies were screened using blue-white selection and subsets of white colonies were selected for colony PCR. Colony PCR was performed with GoTaQ Green and the following reaction mix: 1 µl each DMI3-F new and DMI3-R-nostop (20 µM), 3 µl dH$_2$O, 5 µl GoTaQ Green and one colony per reaction. PCR cycling conditions were as follows: 95 °C for 10 min, 30 cycles of 95 °C for 10 seconds, 53 °C for 20 seconds and 72 °C for 2 min, with a final extension step of 5 min at 72 °C. Corresponding *E. coli* colonies were grown overnight in LB medium with appropriate antibiotic selection. Agarose gel electrophoresis was carried out as described in section 2.4.2 to identify the appropriate vectors for sequencing. Plasmid DNA was purified as described in sections 2.4.3 and sequenced using the standard T7 and SP6 promoter primer as described in 2.4.4.

2.6 Oligonucleotide primer sequences (5′-3′)

Where necessary, standard sequencing primers (T7 promoter, T7 terminator and SP6) were used for sequencing reactions and are not listed here.

2.6.1 Quikchange primers for K47E mutation

Primer name	Primer sequence
A139G_CCaMK_F	TCACAAAGCCAGGTTGCCATTGAAACCCTGCGCC
A139G_CCaMK_R	GGCGCAGGGTTTCAATGGCAACCTGGCTTTGTGA

Table 2.6. Primers used for generation of the Quikchange K47E mutation.

2.6.2 pMal sequencing primers to check for K47E mutation

Primer name	Primer sequence
MalE-F	GGTCGTCAGACTGTCGATGAAGCC
M13_pUC-R	CGCCAGGGTTTTCCCAGTCACGAC

Table 2.7. Primers used for sequencing of the pMal vector.

2.6.3 pCDF duet vector cloning and sequencing

Primer name	Primer sequence
CCaMK-MCS1-F	CGTTCAGAGCCATGGGCTATGGTACA
CCaMK-MCS1-R	AATTATTGCGGCCGCTTACGGGCGGAT
MBP-CCamK-F	GGCGCACGCATATGAAAATCGAAG
MBP-CCamK-R	GATTGATACTCGAGTTACGGGCGGATT
DuetUP2	TTGTACACGGCCGCATAATC
pACYCDuet-1UP	GGATCTCGACGCTCTCCCT
DuetDOWN-1	GATTATGCGGCCGTGTACAA
Seq_in_MBP-CCamK-F	CCAACATCGACACCAGCA
Seq_in_MBP-CCamK-R	GACCAACAACAGGGTCGGT
pCDF-upMCS1-F	GTTTTCGCAGAAACGTGGCT
pCDF-upMCS1-R	CAGCAGAGTATCACTCACGG

Table 2.8. Primers used for duet vector generation and sequencing.

2.6.4 His-SUMO-CCaMK peptide cloning primers

Primer name	Primer sequence
POPINS3C_CCaMKpep_F	GCGAACAGATCGGTGGTATGAACGGTAACTTTAGCTTTTACG
S3C_new_CCaMK-F	AAGTTCTGTTTCAGGGCCCGATGAACGGTAACTTTAGCTTTTACG
POPINS3C_CCaMKpep_R	ATGGTCTAGAAAGCTTTAGCTAGACCAAACCGACGCAATGGC

Table 2.9. Primers used for generation of two HisSUMO-CCaMK peptide vectors.

2.6.5 Quikchange primers for surface entropy reduction of the VLD

Primer name	Primer sequence
a49g_a50c_a53c_a56c	CCTGGTCGGCTCATATGATCTGGCAGCAGCAGAAATCGA AAACCTGCGCATG
a49g_a50c_a53c_a56c_antisense	CATGCGCAGGTTTTCGATTTCTGCTGCTGCCAGATCATAT GAGCCGACCAGG

Table 2.10. Primers used for surface entropy reduction of the VLD.

2.6.5 Golden Gate cloning and sequencing primers for EC10265

Primer name	Primer sequence
proLjUB_end_F	CAGTTCCCAACCCTAACAATTCTTG
MtCCaMK_sta_F	CACAATAGGGTTCCCTACAATGAGAC
DsRED_end_F	GGGACGCCACCATCTGTTCCTTTAAGC
Pro-DsRED-end_F	GCCAATTTTCAGCTCCACCGTAT
DsRED_sta_F	CAGATGCAGATCTTAATGGGGTCATC
BASTAseq_F	GACGCGGGACAAGCCGTTTTACGTTTG
GoldenGate-3	CCCGCCAATATATCCTGTC
GoldenGate-4	GCGGACGTTTTTAATGTACTG
MtCCaMK_sta_F	CACAATAGGGTTCCCTACAATGAGAC
REV-DMI3	TTATGGACGAATAGAAGAGAGAAC
REV-DMI3.1	CTTCATATTCATCTGAGAG
FOR-DMI3	ATGGGATATGGAACAAGAAAAC

Table 2.11. Primers used for Golden Gate cloning of CCaMK over-expression construct.

2.6.7 *DMI3* PCR and RT-PCR primers

Primer name	Primer sequence
Mt-DMI3-F	ATGGGATATGGAACAAGAAAACTC
Mt-DMI3-R-nostop	TGGACGAATAGAAGAGAGAACTAC

Table 2.12. DMI3 primers used for PCR and RT-PCR experiments.

Chapter 3: Mechanism of autophosphorylation at T271

3.1 Introduction

Autophosphorylation, particularly at T271, is a key regulator of CCaMK function. This residue is particularly important because mutation to a non-phosphorylatable residue (Tirichine *et al.*, 2006) creates a gain-of-function mutation leading to spontaneous nodulation. This is thought to be due to the disruption of a predicted hydrogen-bond network, involving pT271, which autoactivates the protein for downstream signalling (Miller *et al.*, 2013). Therefore, autophosphorylation at T271 is key to the regulation of CCaMK. However, the mechanism by which this occurs is not well understood.

In contrast, the mechanism by which autophosphorylation at T286 in CaMKII occurs is well characterised. Autophosphorylation at this particular residue occurs in an inter-subunit, intra-oligomeric mechanism (Mukherji and Soderling, 1994). Since CCaMK shows significant sequence and structural similarities to CaMKII, this begs the question of whether autophosphorylation of CCaMK occurs with an analogous mechanism. This was first addressed with an investigation into the oligomeric status of the protein using a non-aggregating, full length construct of CCaMK. This construct consisted of an *E. coli* codon-optimised CCaMK with an N-terminal MBP tag for stability.

The mechanism of autophosphorylation at T271 has been studied previously, however the results were contradictory. Sathyanarayanan and colleagues (Sathyanarayanan *et al.*, 2001) studied the protein concentration-dependence of the rate of autophosphorylation at T271. They observed a saturation curve and concluded an inter-monomeric/oligomeric mechanism; however this trend does not show the required second order kinetics for this conclusion. The mechanism of autophosphorylation at T271 was also studied by Tirichine and colleagues (Tirichine *et al.*, 2006). By contrast, they observed an exponential increase in phosphate incorporation in a protein-concentration dependent manner, which is consistent with second order kinetics and an inter-monomeric/oligomeric mechanism of autophosphorylation. Therefore despite the observation of different trends of phosphorylation dependence on protein concentration in two studies, the same conclusion was reached on the mechanism of autophosphorylation. In order to clarify this discrepancy, a thorough investigation was undertaken to test all possible mechanisms of autophosphorylation at T271: intra-monomeric, inter-monomeric, inter-oligomeric and intra-oligomeric.

3.2 MBP-CCaMK purification

In order to characterise biochemical properties of CCaMK, it was necessary to purify the *M. truncatula* CCaMK protein using an *E. coli* expression system. Previous work by David Swainsbury established that when full-length CCaMK was purified with or without His or GST tags, it formed large aggregates. Therefore, previous biochemical characterisation of CCaMK was carried out using truncations of the protein which lacked the kinase domain that were present in solution as monomers (Swainsbury *et al.*, 2012). However, in order to characterise the oligomeric state of CCaMK, it was necessary to establish an expression system for the full-length protein. Therefore, a maltose binding protein (MBP) fusion was utilised, in order to aid protein stability and prevent aggregation. This was developed using the codon-optimised sequence for expression in *E. coil* by Akira Miyahara (Miller *et al.*, 2013). Previous work by Liang Zhou established that this fusion protein was a stable oligomer with low levels of aggregation and therefore was suitable for biochemical experiments (section 3.2.1).

In order to isolate protein of sufficient purity, a two-step purification was used. MBP-affinity chromatography (Figure 3.1A) was utilised followed by size exclusion chromatography (SEC) (Figure 3.1B) to remove contaminants. The size and purity of the protein was confirmed using SDS-PAGE (Figure 3.1C). The later SEC peak was also analysed by SDS-PAGE, and was shown to not contain MBP-CCaMK protein (data not shown). Some minor contaminants remained in the MBP-CCaMK sample after SEC however the protein was of sufficient purity for subsequent experiments.
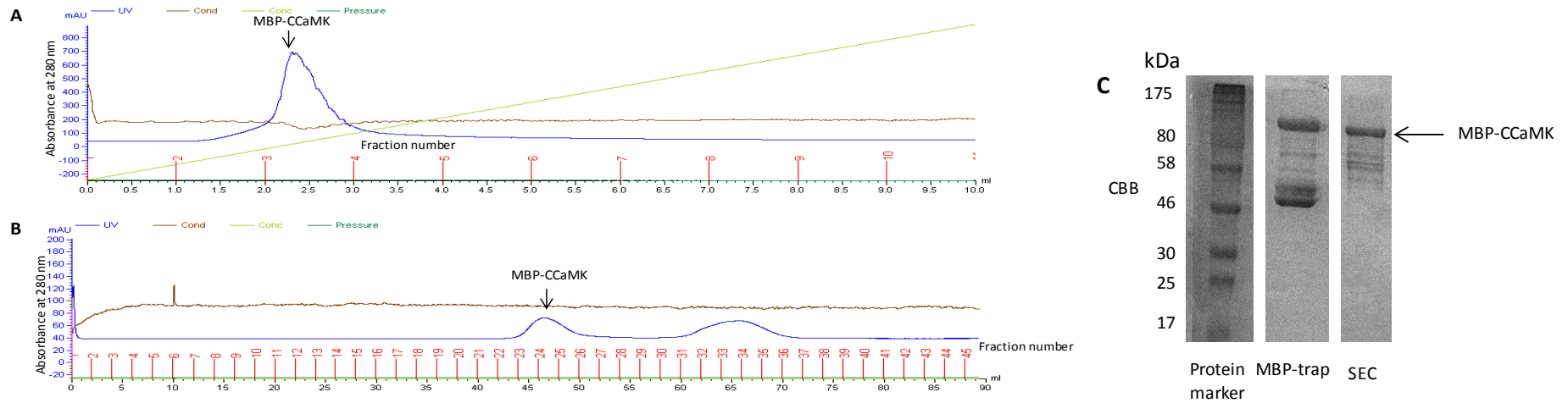
Figure 3.1. Purification of recombinant MBP-CCaMK using MBP affinity chromatography and SEC.
(A) Chromatogram for the purification of MBP-CCaMK using an MBPTrap HP column. Fractions corresponding to the marked peak were collected, pooled and subjected to size exclusion chromatography, shown in **(B)**.
(B) Size exclusion chromatogram for MBP-CCaMK using a HiLoad 16/60 Superdex 75 prep grade column. The marked peak was confirmed to contain the correct protein by SDS-PAGE shown in **(C)**.
(C) SDS-PAGE for MBP-trap and SEC purification of MBP-CCaMK. The observed size of MBP-CCaMK was approximately 98 kDa, as predicted. The second elution peak did not contain MBP-CCaMK, and this was also confirmed by SDS-PAGE (data not shown).

Chapter 3: Mechanism of autophosphorylation at T271

3.2.1 Size estimates of the MBP-CCaMK oligomer with SEC, DLS and AFM

In order to characterise the size of the MBP-CCaMK oligomer analytical SEC, dynamic light scattering (DLS) and atomic force microscopy (AFM) were used. First, analytical SEC was carried out by Liang Zhou of MBP-CCaMK with a HiPrep 16/60 Sephacryl S-300 HR column (Figure 3.2A). MBP-CCaMK eluted in a single peak at 39.3 ml which indicated the presence of a large oligomer. A calibration curve for the column was established with monodisperse proteins of known molecular weights, and from this a size estimate of the MBP-CCaMK oligomer was calculated of 1607 ± 1.3 (S_x) kDa. The predicted size of a monomer of MBP-CCaMK is 100 kDa so therefore the oligomer size observed during SEC corresponded to approximately 16 subunits.

DLS was also carried out on freshly purified MBP-CCaMK and three readings are shown (Figure 3.2B). Each graph corresponded to an average of 10 individual measurements carried out on the same protein sample. The most abundant species detected in the first reading (left panel) was 1341 kDa which corresponded to an oligomer size of approximately 13-14 subunits. A much larger species was present (Radius = 34.3 nm, Size = 13201 kDa, Mass % = 7.3) and was presumably an aggregate of MBP-CCaMK. For both measurement two and three (middle and right panel, respectively), the most abundant species detected was formed of 20-21 subunits (molecular weights detected were 2072 and 2014 kDa, respectively). The diameter of the oligomer was also measured and the three readings gave a mean diameter of 29 nm ± 3 (SD). Substantial variation was detected between several measurements taken on the same protein in quick succession. This variability between and within measurements is somewhat expected because DLS is affected by the polydispersity of the samples. Samples were filtered with an 0.2 μm filter in order to reduce the presence of large species which may have affected this. However, at least in the first measurement a larger aggregate was detected, and this polydispersity may have affected the variability of these data. This is why many readings were taken, and then the average was recorded. In addition, assumptions were made when carrying out DLS since DLS assumes that macromolecules are spherical, which is most-likely not the case for MBP-CCaMK. Therefore the mean size of the MBP-CCaMK oligomer of 18 ± 4 (SD) subunits from DLS was an estimate only.

AFM was carried out on MBP-CCaMK by Patrick Gunning and Andrew Kirby (Figure 3.2C). Distinct particles which corresponded to oligomers of MBP-CCaMK were detected using this method. Their size was variable, where the larger particles most likely corresponded to aggregates of the protein. However, from the smaller particles it was possible to determine that oligomers of MBP-CCaMK were globular in shape with a mean diameter of 33 nm ± 12 (SD). From discussions with Andrew Kirby and Patrick Gunning it was determined that this resolution would be the limit of this technique, so therefore a rough estimate of the size and shape was the maximum level of certainty that can be established from this technique. However, this size was consistent with the DLS data obtained.

Taken together, these techniques provided an estimate of the size of the MBP-CCaMK oligomer. Recombinant MBP-CCaMK forms an oligomer of 16-18 subunits with a diameter of ~31 nm (Table 3.1).
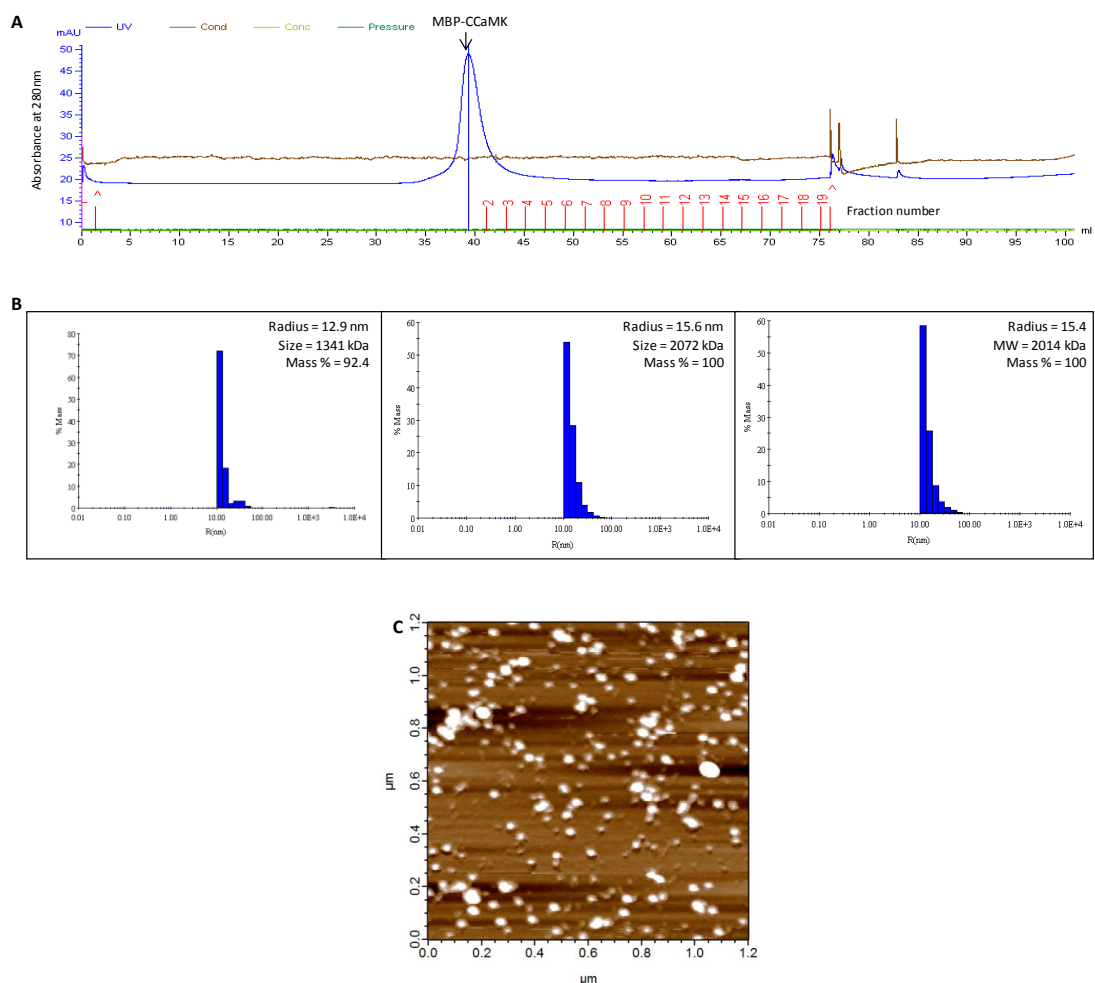
Figure 3.2. Recombinant MBP-CCaMK is an oligomer.

Chapter 3: Mechanism of autophosphorylation at T271

(A) Analytical gel filtration chromatogram for freshly prepared MBP-CCaMK protein carried out by Liang Zhou with a HiPrep 16/60 Sephacryl S-300 HR column. MBP-CCaMK eluted at 39.29 ml and is marked on the chromatogram. Calibration of the column was performed using the following standards: blue dextran (2000 kDa), apoferritin (669 kDa), apoferritin (443 kDa), β-amylase (200 kDa), albumin (66 kDa) and carbonic anhydrase (29 kDa).
(B) DLS for freshly-prepared MBP-CCaMK showing percentage mass against radius. Insignificant peaks were removed in order to prevent data skew. Each graph corresponded to an average of ten 10 s measurements, taken sequentially on the same protein sample. Insets show the radius, mass and percentage mass of the main species detected.
(C) AFM image for freshly-prepared MBP-CCaMK in air on untreated mica with glycerol. Protein sample preparation was carried out with Liang Zhou and AFM was performed by Patrick Gunning and Andrew Kirby (IFR). To determine the average diameter of MBP-CCaMK, the particles in a representative section were measured in mm on an enlarged version of the AFM image. These were converted to μm using the axes and an average was taken of these values.

Technique	Size (kDa)	Number of subunits	Diameter (nm)
Gel filtration	1607 ± 1.3 (S_x)	16	-
Dynamic light scattering	1809 ± 406 (SD)	18 ± 4 (SD)	29 ± 3 (SD)
Atomic force microscopy	-	-	33 ± 12 (SD)

Table 3.1. MBP-CCaMK is an oligomer of 16-18 subunits with a diameter of ~31 nm.

3.3 Mechanism of autophosphorylation at Thr271

Phosphorylation at T271 stimulated by calcium binding to the EF-hands is crucial to the function of CCaMK as previously described. However the mechanism by which this autophosphorylation occurs is unclear. This has important kinetic implications (Section 3.3.1) for how the protein decodes calcium spiking in the sym pathway. There are four possible mechanisms by which autophosphorylation can occur: intra-monomeric, inter-monomeric, intra-oligomeric and inter-oligomeric as shown in Figure 3.3.

In order to determine whether intra-monomeric autophosphorylation was possible, a homology model of the kinase domain with the autoinhibitory domain was generated with Phyre2 (Figure 3.4). The predicted spatial positioning of the T271 residue and the active site suggested that intra-monomeric phosphorylation is unlikely because they were located on two different sides of the protein. The kinase domain is well-conserved and therefore this is likely to be reliable despite the use of a homology model. Furthermore, there are no indications on the model that a significant conformational change to facilitate intra-monomeric autophosphorylation is possible. Additionally, the autoinhibitory helix appeared to be located over the active site. However, even if this helix is removed (which would be necessary for phosphorylation to occur) intra-monomeric autophosphorylation is unlikely. This conclusion is consistent with homology modelling carried out by Sathyanarayanan and colleagues (Sathyanarayanan *et al.*, 2001).

Evidence described above in section 3.2.1 showed that CCaMK is likely present as an oligomer. This therefore suggested that an inter-monomeric mechanism of autophosphorylation was unlikely due to the presence of 16-18 subunits of CCaMK in a complex. Therefore it seems likely that autophosphorylation occurs either within or between oligomers. Furthermore, preliminary data obtained by Liang Zhou suggested that the oligomeric state of CCaMK may be crucial to its ability to autophosphorylate. A mutant of CCaMK which lacked the first five amino acids was unable to form a stable oligomer. Interestingly, this mutant was also unable to autophosphorylate itself, suggesting that the presence of neighbouring subunits was essential to autophosphorylation. In addition, data from Akira Miyahara suggested that inter-oligomeric autophosphorylation does not occur (Figure 3.5). These data comprise an autoradiogram showing total protein phosphorylation and the corresponding protein stained gel for GST-CCaMK (wild type, WT), GST-CCaMK

K47E (a point mutation which rendered the kinase inactive) and WT CCaMK. When the proteins were present singly, autophosphorylation was detected in WT GST-CCaMK and not GST-CCaMK K47E after a kinase assay, as expected. When GST-CCaMK K47E and WT CCaMK were mixed and then subjected to a kinase assay, autophosphorylation was detected only in the WT protein. Therefore no inter-oligomeric autophosphorylation was detected. However, this experiment was carried out using GST-CCaMK, without calcium-induction and autophosphorylation at all potential phosphorylation sites was detected. This construct is also not ideal for study because it does not require calcium for autophosphorylation so it has limited relevance to CCaMK *in planta.* Therefore it remained to be determined if the outcome of the experiment would be the same with calcium-induced autophosphorylation at T271, using MBP-CCaMK.

Therefore these preliminary data indicated that most likely phosphorylation does not occur inter-oligomerically and that oligomerisation most likely plays a role in autophosphorylation. However they did not demonstrate definitively that autophosphorylation in CCaMK occurs intra-oligomerically. Therefore an experiment was designed to distinguish between the three most likely mechanisms of autophosphorylation at T271: inter-monomeric, inter-oligomeric or intra-oligomeric mechanism.
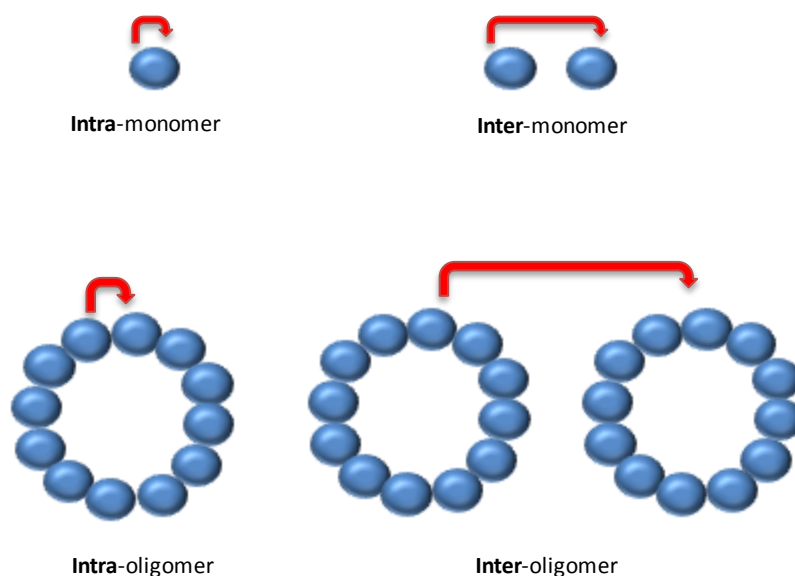
Figure 3.3. Schematic of possible mechanisms of autophosphorylation at Thr271.
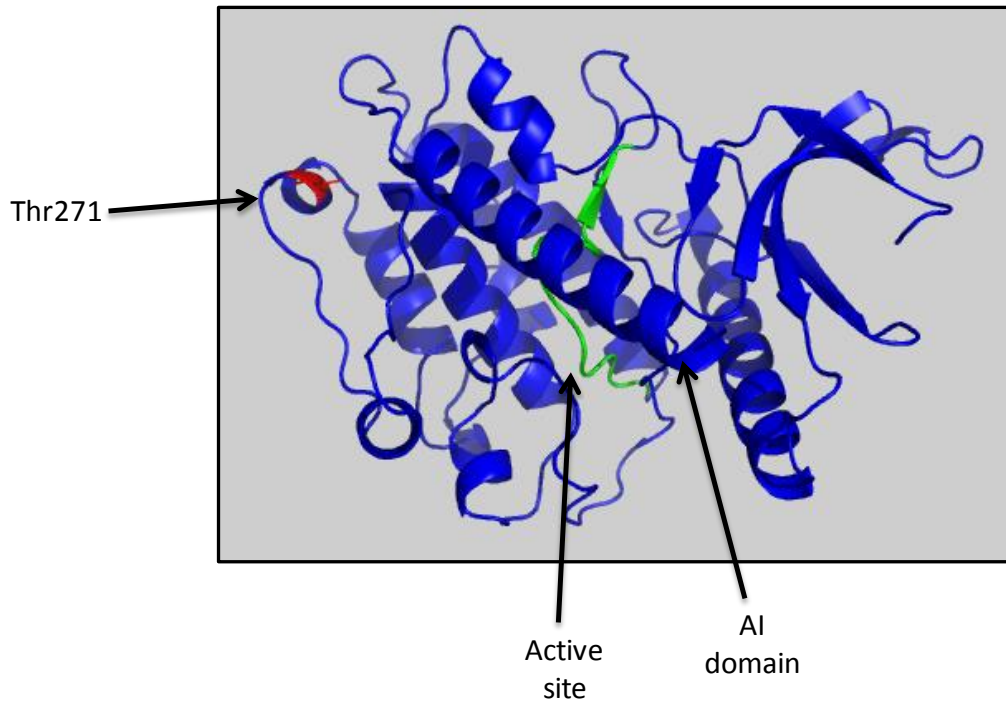
Figure 3.4. Intra-monomeric phosphorylation is unlikely due to the locations of Thr271 and the active site.

This model of the kinase domain with the auto-inhibition domain was produced with Phyre2 and the locations of the active site and Thr271 residues are marked in green and red respectively. The active site and Thr271 residue are predicted to be located on different faces of the well-conserved kinase domain and the AI domain is predicted to be located over the active site (marked with an arrow). Therefore, intra-monomeric phosphorylation is unlikely.
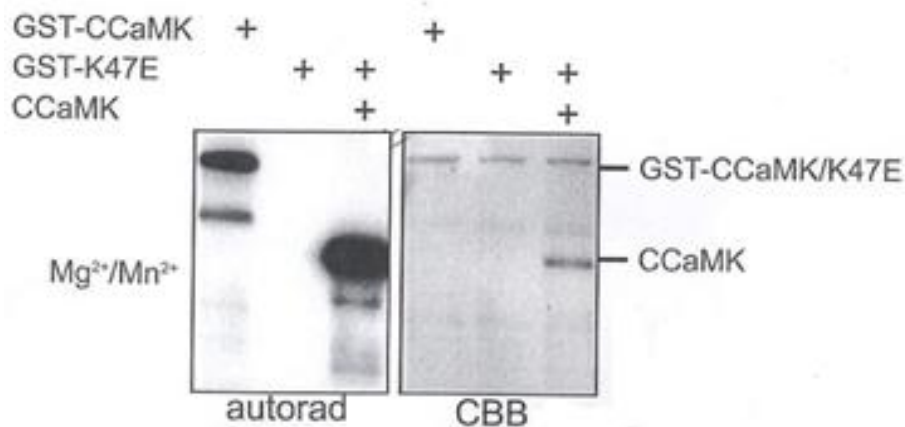
Figure 3.5. Autophosphorylation does not occur between separately expressed and subsequently mixed WT CCaMK and GST-CCaMK K47E. Data are courtesy of Akira Miyahara.

Separately expressed constructs of GST-CCaMK, GST-CCaMK K47E and WT CCaMK mixed with GST-CCaMK K47E were subjected to a kinase assay followed by autoradiography to detect total autophosphorylation. Autophosphorylation was observed for GST-CCaMK and

not for GST-CCaMK K47E, as expected. Autophosphorylation was observed only with WT CCaMK and not GST-CCaMK K47E when both proteins were mixed in the kinase assay. All kinase assays were performed in the presence of Mg^{2+} and Mn^{2+}, and without calcium because this construct did not require it. This therefore implies that this CCaMK construct is not fully functional.

3.3.1 Kinetic implications of the mechanism of autophosphorylation

The mechanism of autophosphorylation at T271 has important implications on the kinetics of the phosphorylation reaction which justify its study. Figure 3.6 shows examples of the extent of normalised phosphorylation over time. The mechanism by which phosphorylation occurs has implications for the protein concentration-dependence of the reaction as shown. If the reaction occurs intra-oligomerically (or intra-monomerically), alteration of the concentration of the protein does not affect the rate at which the reaction proceeds. This reaction proceeds with first order kinetics and eventually levels off. Conversely, an inter-oligomeric (or inter-monomeric) reaction would be concentration-dependent, with an increase in reaction rate when the concentration of the protein increases. This is due to the increasing likelihood of two molecules colliding and reacting with second order kinetics.
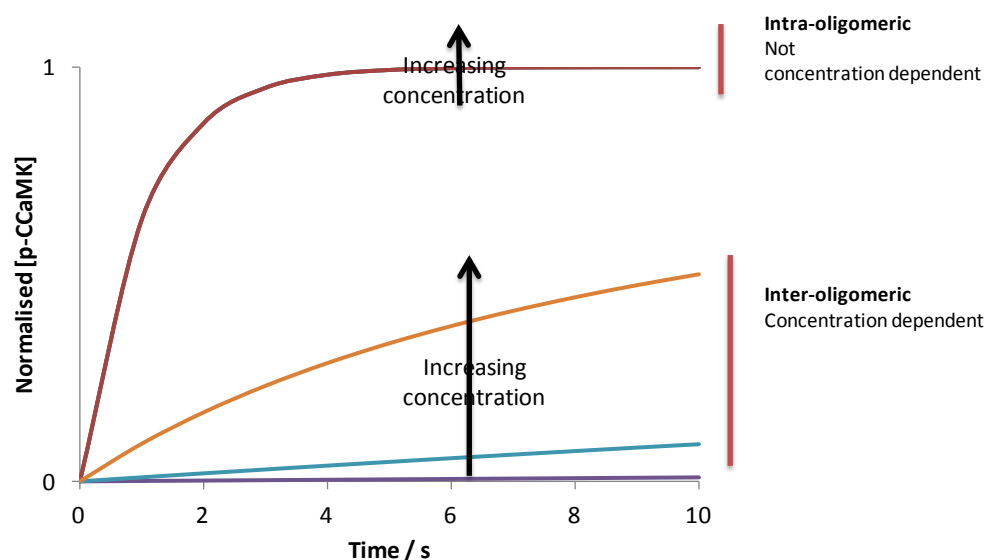
Figure 3.6 Kinetics of the inter- and intra-oligomeric mechanisms. Figure courtesy of Stephen Bornemann. Three sample curves are shown for each mechanism type and in the case of intra-oligomeric phosphorylation these three curves overlap.

Chapter 3: Mechanism of autophosphorylation at T271

3.3.2 Experimental design

An experiment was devised in order to distinguish between the mechanisms of autophosphorylation (Mukherji and Soderling, 1994) (Figure 3.7). Two forms of CCaMK were used for the experiment: tagless WT CCaMK and MBP-CCaMK K47E (kinase-dead point mutation). These forms were co-expressed in *E. coli* using a pCDFDuet-1 vector to form mixed oligomers consisting of kinase-active and kinase-dead subunits. These mixed oligomers were purified using MBP affinity chromatography and the presence of both proteins was confirmed by SDS-PAGE and subsequent mass spectrometry. Therefore, the proteins must have been present in mixed oligomers. These proteins were also expressed and purified as separate oligomers.

In order to determine if inter-oligomeric phosphorylation occurred, the two proteins were separately expressed and then mixed before a kinase assay and subjected to western blotting probed with αpT271, a state-specific antibody which specifically recognised phosphorylation at T271. Due to the presence of an MBP tag on the kinase dead form, it could be determined whether the WT form could phosphorylate the kinase dead form when present in separate oligomers using SDS-PAGE prior to western blotting. This approach assumed that once oligomers were formed, subunits were not exchanged between them during the kinase assay. This experiment would also determine whether inter-monomeric phosphorylation could happen, since some of the proteins could be present as monomers rather than oligomers.

By contrast, to determine if intra-oligomeric phosphorylation occurred, the co-expressed mixed oligomers were subjected to a kinase assay and a subsequent western blot with αpT271. Due to the presence of the MBP tag on the kinase-dead form, it could be determined whether the WT subunits present in the same oligomer could phosphorylate the kinase dead subunits in an intra-oligomeric manner using SDS-PAGE prior to western blotting.
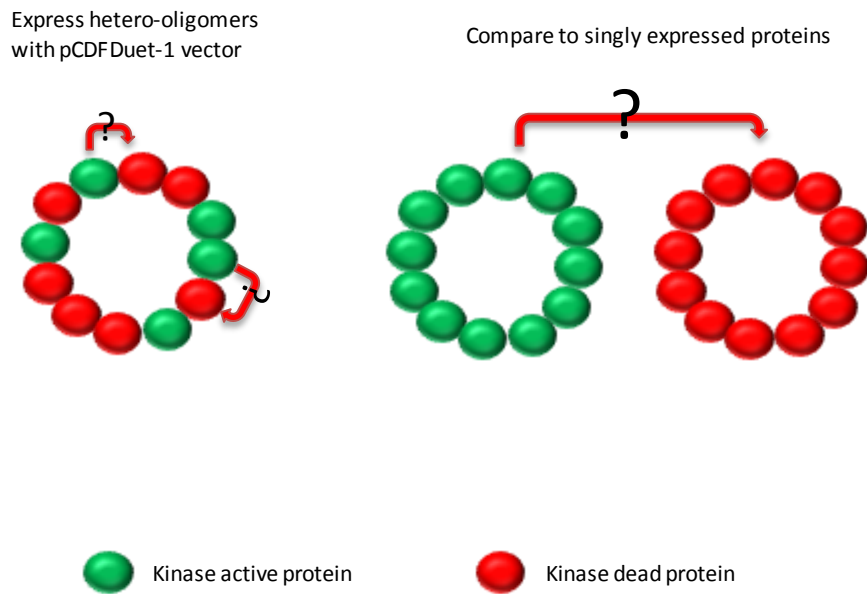
Figure 3.7. Experimental design to distinguish between an intra- or inter-oligomeric mechanism of autophosphorylation at Thr271.

Constructs of CCaMK were co-expressed in a pCDFDuet -1 vector or expressed separately and mixed, and then subjected to a kinase assay. In each case a kinase dead (K47E/D192N) and kinase active form (WT) of CCaMK was used. The kinase dead form was tagged in each case for separation by SDS-PAGE from the WT form. Co-expressed mixed oligomers were confirmed by purification using a tag (MBP) and a mass spectrometric analysis.

3.3.3 Inter-oligomeric autophosphorylation

In order to determine if autophosphorylation occurred inter-oligomerically, MBP-CCaMK K47E and WT CCaMK were separately expressed and purified, and then mixed during a kinase assay. MBP-CCaMK K47E was purified as described in section 3.2, except no SEC was performed. WT tagless CCaMK was purified as shown in Figure 3.8. CaM-Sepharose affinity chromatography was used and, as with MBP-CCaMK K47E, no subsequent SEC purification was required. This was because both proteins were of sufficient purity for use in the kinase assay despite the presence of some co-purified contaminants.

MBP-CCaMK K47E and WT CCaMK were both diluted to approximately 400 ng/μl and equal volumes of each protein were mixed in a kinase assay with the addition of ATP, Ca^{2+} and Mg^{2+} and then subjected to western blotting probed with αpT271 (Figure 3.9). Phosphorylation at T271 was observed only in the WT protein and no phosphorylation was detected in the kinase dead protein. The presence of the kinase dead protein was confirmed by staining of the blot after development with Ponceau S. Therefore, under these conditions no inter-oligomeric autophosphorylation was detected. It follows therefore that no inter-monomeric phosphorylation was detected in this experiment either. A control WT MBP-CCaMK protein was also subjected to a kinase assay, and strong autophosphorylation was detected. This did not appear to be exclusively calcium-induced. However, this was most likely due to the presence of background phosphorylation that sometimes occurred during the expression of the protein in *E. coli.*
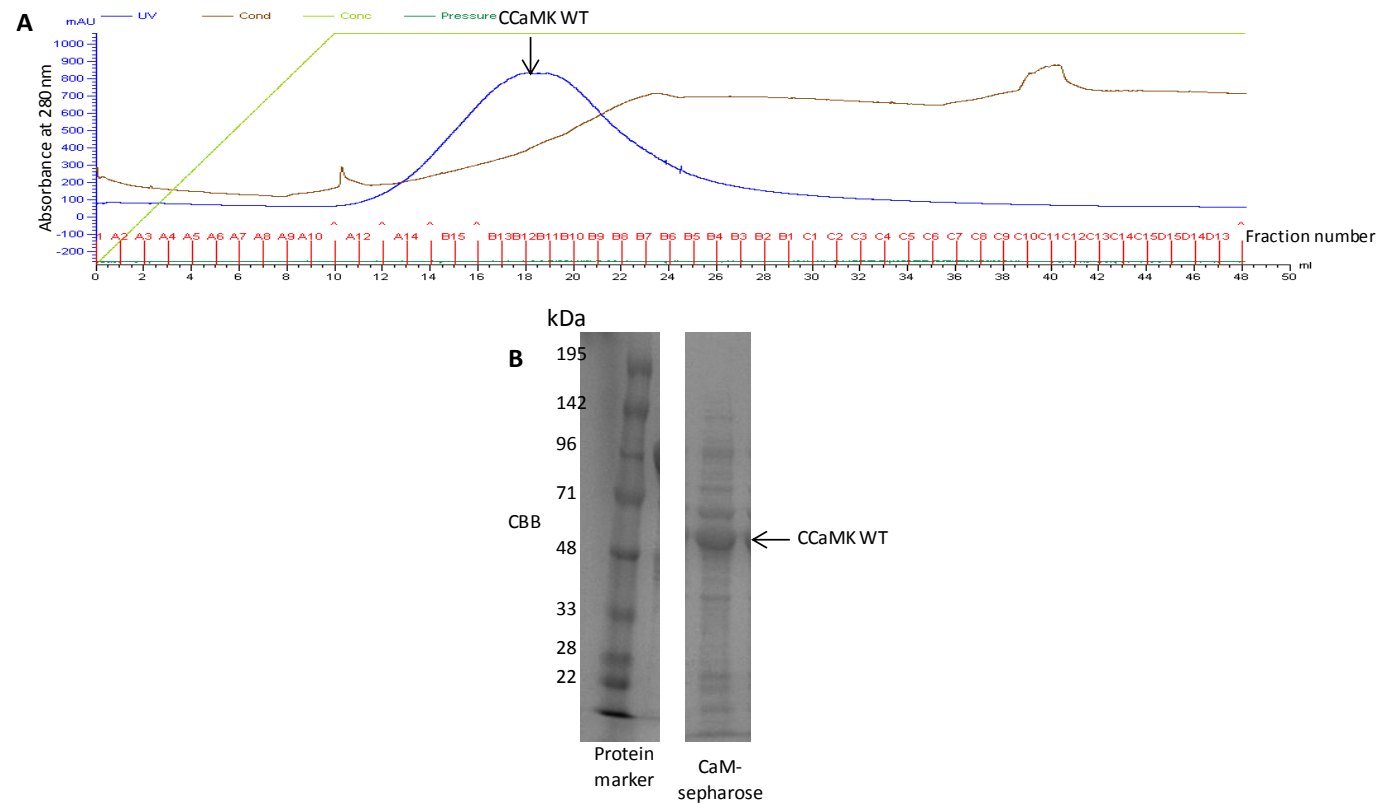
Figure 3.8. Purification of recombinant WT, tagless MtCCaMK.
(A) Chromatogram for CCaMK purified using the CaM-Sepharose column at 4 °C. Fractions corresponding to the marked peak were collected and verified by SDS-PAGE as shown in **(B)**.
(B) SDS-PAGE of CaM-Sepharose purification of WT CCaMK. The observed size of CCaMK was 58 kDa, as predicted.

Chapter 3: Mechanism of autophosphorylation at T271
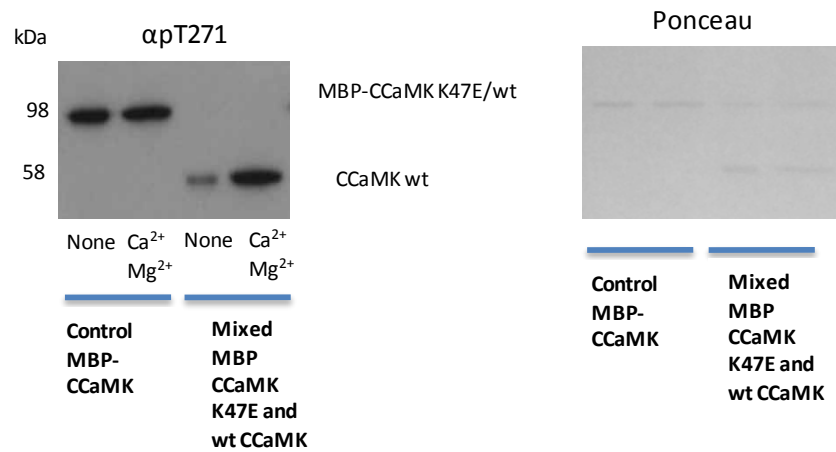
Figure 3.9. Autophosphorylation at Thr271 does not occur *via* an inter-oligomeric mechanism.
MBP-CCaMK K47E and WT CCaMK were separately expressed and subsequently mixed before a kinase assay, and subjected to western blotting probed with αpT271.
Autophosphorylation was induced with the addition of ATP, Ca^{2+} and Mg^{2+}.
Autophosphorylation was detected only in the WT CCaMK oligomers and not in the kinase-dead oligomers. MBP-CCaMK was present in the first two lanes as a control.

Chapter 3: Mechanism of autophosphorylation at T271

3.3.4 Intra-oligomeric autophosphorylation

In order to determine if intra-oligomeric phosphorylation occurred at T271, MBP-CCaMK K47E and WT CCaMK were co-expressed to generate mixed oligomers and subjected to a kinase assay. Co-expression was carried out with the pCDFDuet-1 vector and purification was carried out with MBP-affinity chromatography (Figure 3.10). Purified protein was analysed by SDS-PAGE and the presence of both of the proteins was confirmed by a mass spectrometric analysis. SEC was not performed so that a higher concentration of the protein was retained for use in the assay, because it was anticipated that it might not be possible to concentrate the protein complex due to potential aggregation after extra time and manipulations. Due to the purification method used, a higher proportion of MBP-CCaMK K47E was present in the mixed oligomers compared to the concentration of WT protein present in the complexes (Figure 3.10B) and it was not possible to alter this ratio. However, due to the higher concentration of the MBP-K47E subunits it followed that the WT subunits would most likely be directly adjacent to a kinase-dead subunit instead of another WT subunit. Since intra-oligomer inter-subunit phosphorylation would most likely occur at adjacent subunits, this protein preparation is ideally suited to detect this autophosphorylation. Therefore, if phosphorylation did occur within the oligomers, it would most likely be able to occur in this protein conformation.

Freshly co-purified MBP-CCaMK K47E and WT CCaMK oligomers were subjected to a kinase assay induced with the addition of ATP, Ca^{2+} and Mg^{2+} and subsequent western blotting probed with αpT271 (Figure 3.11). In contrast to the inter-oligomeric experiment, low levels of phosphorylation were detected in both the WT subunits and the kinase dead subunits. Surprisingly, however, no calcium-induction of autophosphorylation was detected in either the WT or the kinase dead protein. Therefore the phosphorylation detected corresponded either to intra-oligomeric autophosphorylation which occurred during expression and not during the assay, or the bands detected corresponded to non-specific background signal. Separately expressed WT CCaMK showed calcium-induced autophosphorylation and separately expressed MBP-CCaMK K47E showed no phosphorylation at T271 under the same conditions. The control proteins therefore showed the expected trend under these conditions. Taken together, these data suggest that intra-oligomeric autophosphorylation may occur in this protein configuration.

In order to determine if the potential intra-oligomeric phosphorylation detected was real or just a background signal, lambda protein phosphatase was used to test if the signal was specifically removed (Figure 3.12). As with Figure 3.11, low levels of phosphorylation were apparently detected in co-expressed oligomers of both WT and kinase dead subunits. However, when the lambda phosphatase-treatment was used, the signal was mostly lost from the WT subunits but was not from kinase-dead protein. A control MBP-CCaMK WT protein subjected to the same treatment showed almost complete removal of the phosphorylation signal detected. Therefore, it can be concluded that the signal detected in the kinase dead subunits, and the residual signal detected for WT subunits and MBP-CCaMK protein after the phosphatase treatment, most likely corresponded to a background signal and not phosphorylation. The signal detected for the mixed oligomers was significantly lower than for the separately expressed proteins, since the exposure time of the western blot membrane necessary to produce the image in Figures 3.11 and 3.12 was significantly longer. It therefore follows that the signal detected corresponded to background signal and not to autophosphorylation. Therefore, it could not be determined from this experiment whether autophosphorylation at T271 can occur intra-oligomerically.
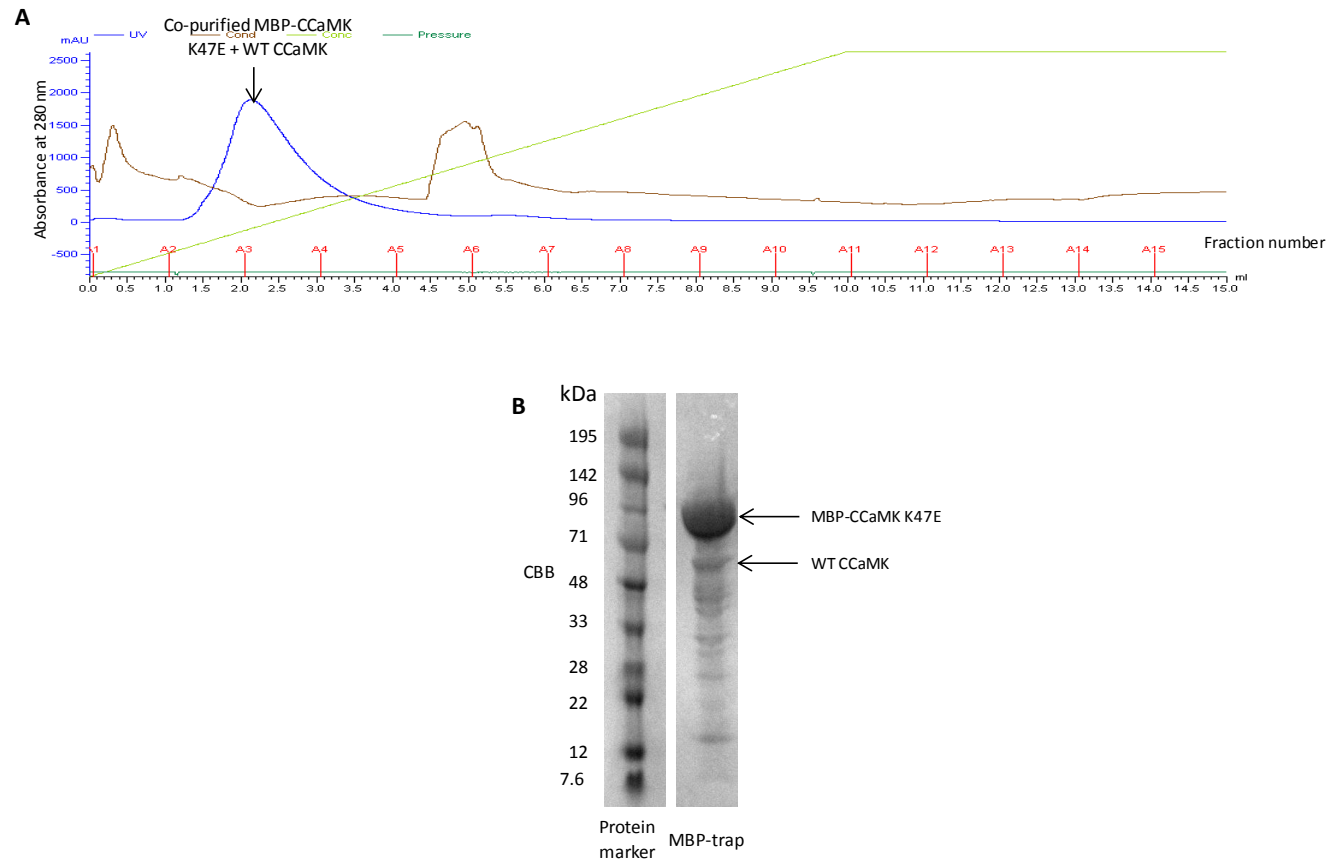
Figure 3.10. Co-purification of mixed oligomers containing WT tagless CCaMK and MBP-CCaMK K47E using MBP affinity chromatography.
(A) Chromatogram for MBP-CCaMK K47E and CCaMK WT mixed oligomers purified using an MBPTrap column at 4 °C. Fractions corresponding to the marked peak were collected and verified by SDS-PAGE as shown in **(B)**.
(B) SDS-PAGE of co-expressed MBP-CCaMK K47E and WT CCaMK. The observed sizes of MBP-CCaMK K47E and WT CCaMK were 98 and 58 kDa, respectively, as predicted. The presence of both proteins was verified by a mass spectrometric analysis.

Chapter 3: Mechanism of autophosphorylation at T271

kDa

αpT271

98 ········· MBP-CCaMK K47E

58 ········· CCaMK wt

| None | Ca²⁺ Mg²⁺ | None | Ca²⁺ Mg²⁺ | Ca²⁺ Mg²⁺ |

Co-expressed MBP CCaMK K47E and WT CCaMK   **Control WT CCaMK**   **Control MBP-CCaMK K47E**

kDa

Ponceau

98

58

Co-expressed MBP CCaMK K47E and WT CCaMK   **Control WT CCaMK**   **Control MBP-CCaMK K47E**
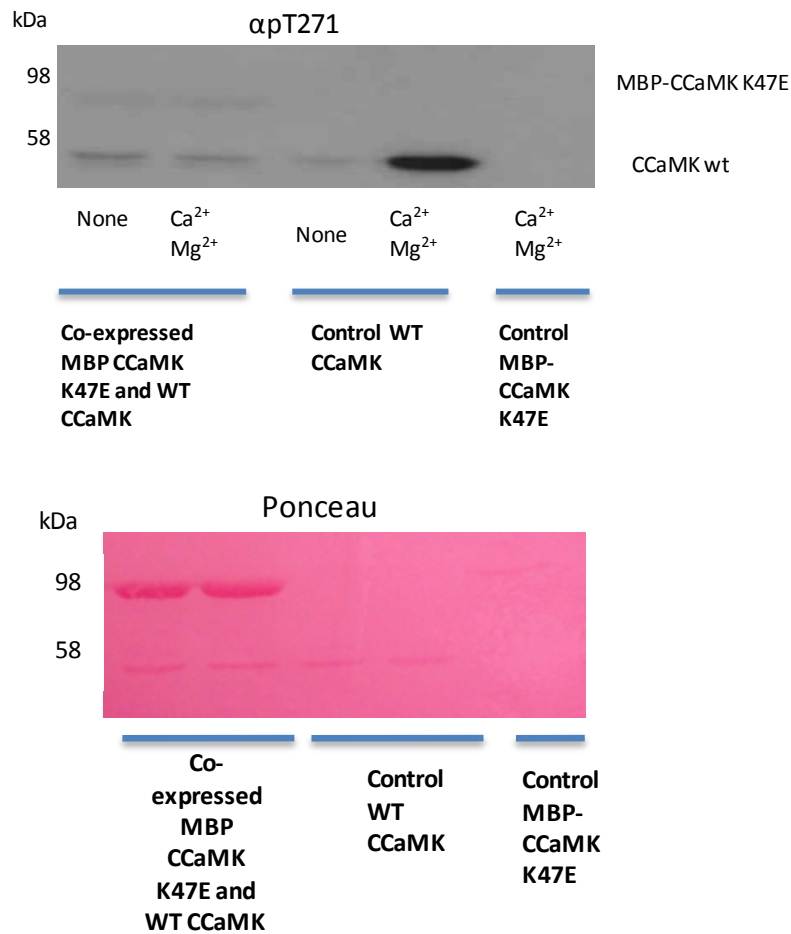
Figure 3.11. Low levels of intra-oligomeric autophosphorylation with no calcium-induction are observed for co-expressed MBP-CCaMK K47E and WT CCaMK.

Co-expressed MBP-CCaMK K47E and WT CCaMK were subjected to a kinase assay and analysed by western blotting probed with αpT271. Autophosphorylation was induced with the addition of ATP, Ca^{2+} and Mg^{2+}. Phosphorylation was detected in both the WT CCaMK subunits and in the kinase dead subunits. However, levels of autophosphorylation were low and no calcium-induction was observed. Separately-expressed WT CCaMK and MBP-CCaMK K47E were present as control proteins and calcium-induced autophosphorylation was detected in WT CCaMK but not MBP-CCaMK K47E, as expected.
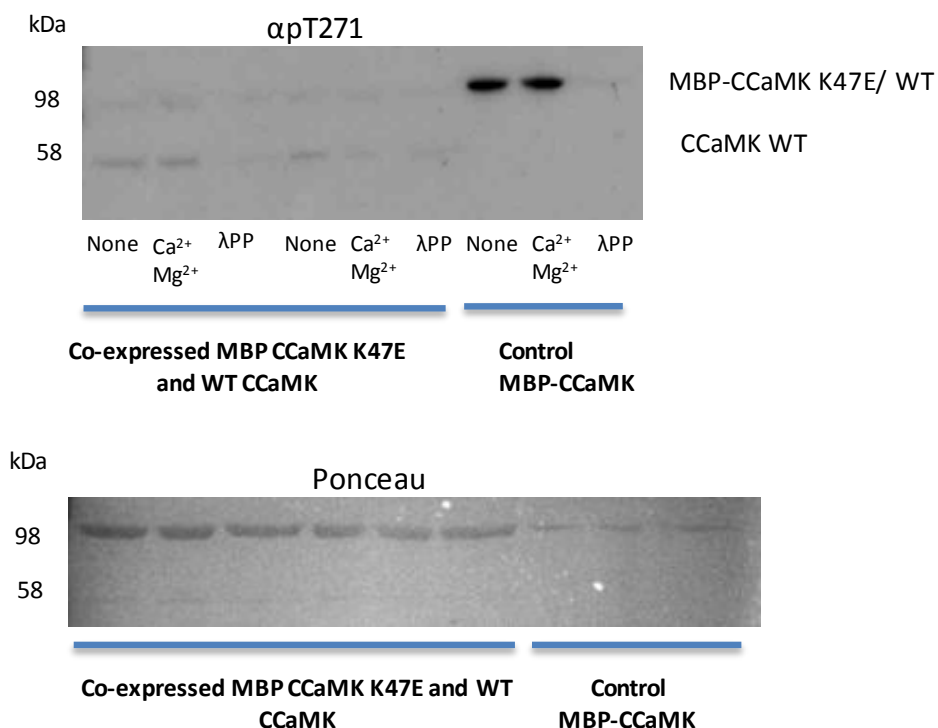
Figure 3.12. Intra-oligomeric phosphorylation of MBP-CCaMK K47E is not removed by a λ phosphatase treatment.

Co-expressed oligomers comprising MBP-CCaMK K47E and WT CCaMK were subjected to a kinase assay followed by a λ phosphatase (λPP) treatment which was analysed by western blotting probed with αpT271. Low levels of phosphorylation with no calcium-induction were detected in both WT and kinase dead subunits following the kinase assay, as observed in Figure 3.11. However, the band corresponding to MBP-CCaMK K47E phosphorylation was not removed by the phosphatase treatment. The same three treatments were repeated in lanes 3-6, which confirmed this trend. However, under the same conditions, phosphorylation of the MBP-CCaMK WT control protein was removed by the phosphatase treatment. Background signal was detected in the phosphatase-treated MBP-CCaMK and therefore the signal detected for co-expressed MBP-CCaMK K47E most likely also corresponded to background signal and not to intra-oligomeric phosphorylation.

3.3.5 Inter-oligomeric autophosphorylation experiment with an alternative kinase dead mutation (D192N)

It was theorised that the presence of the K47E kinase dead mutation could have affected the ability of CCaMK to be autophosphorylated at T271, and therefore affect whether the protein could be either an acceptor or donor of phosphate. In order to determine if the spatial location of the mutation could have affected T271, the mutation was located on the Phyre2 homology model of the kinase domain with the autoinhibitory domain (Figure 3.13). This showed that the Lys47 residue was located distantly from T271. In addition to the location of the K47E mutation, the position of an alternative kinase dead mutation, D192N, was located on the model. This residue was also predicted to be distantly located from the T271 residue. This most likely means that the mutations would not affect the ability of T271 to phosphorylated for example by blocking it. However, distant mutations can also affect the protein in unknown ways and therefore it could not be determined with certainty from this model what global affects the mutations could have on the protein.

In order to determine whether the use of an alternative kinase dead mutation altered the ability of T271 to be phosphorylated, a form of CCaMK possessing the D192N mutation was expressed with a His-SUMO tag using the pOPINS3C vector. Purification was carried out with a HisTrap column and the presence of the correct protein was verified by SDS-PAGE (Figure 3.14). No SEC purification was necessary. His-SUMO-CCaMK D192N and WT CCaMK were both diluted to approximately 400 ng/μl and equal volumes of each protein were mixed in a kinase assay induced with the addition of ATP, Ca^{2+} and Mg^{2+} and subsequently subjected to western blotting with αpT271 (Figure 3.15). Calcium-induced autophosphorylation was detected in WT protein and a background signal only was detected in the kinase-dead protein. This was determined to correspond to a background signal because there was no calcium-induction of autophosphorylation and the His-SUMO-CCaMK D192N protein subjected to a kinase assay alone showed the same low signal. Separately expressed WT CCaMK showed calcium induced autophosphorylation as expected under the same assay conditions. Therefore inter-oligomeric autophosphorylation was also not detected when an alternative kinase dead mutation was used, and therefore it can be concluded that the specific kinase-inactivation mutant most likely did not affect the ability of the protein to be phosphorylated if an inter-oligomeric mechanism were to occur.

The data presented in this section showed that autophosphorylation at T271 most likely does not occur by an inter-oligomeric (or inter-momomeric) mechanism and it follows therefore that it must occur via an intra-oligomeric mechanism. This is because inter-oligomeric phosphorylation did not occur with GST-CCaMK K47E, MBP-CCaMK K47E or His-SUMO-CCaMK D192N and previous data suggested that CCaMK is not able to autophosphorylate when it did not form oligomers. Intra-oligomeric autophosphorylation was not able to be shown conclusively with the mixed oligomer experiment. However, this is most likely due to interference from the tag and insufficient WT protein in the mixed oligomer experiment. A further possibility is that the kinase dead mutation affected the ability of the protein to be phosphorylated, and this cannot be ruled out for both experiments. However, taken together, these data indicated that the most likely mechanism of autophosphorylation at T271 is intra-oligomeric.
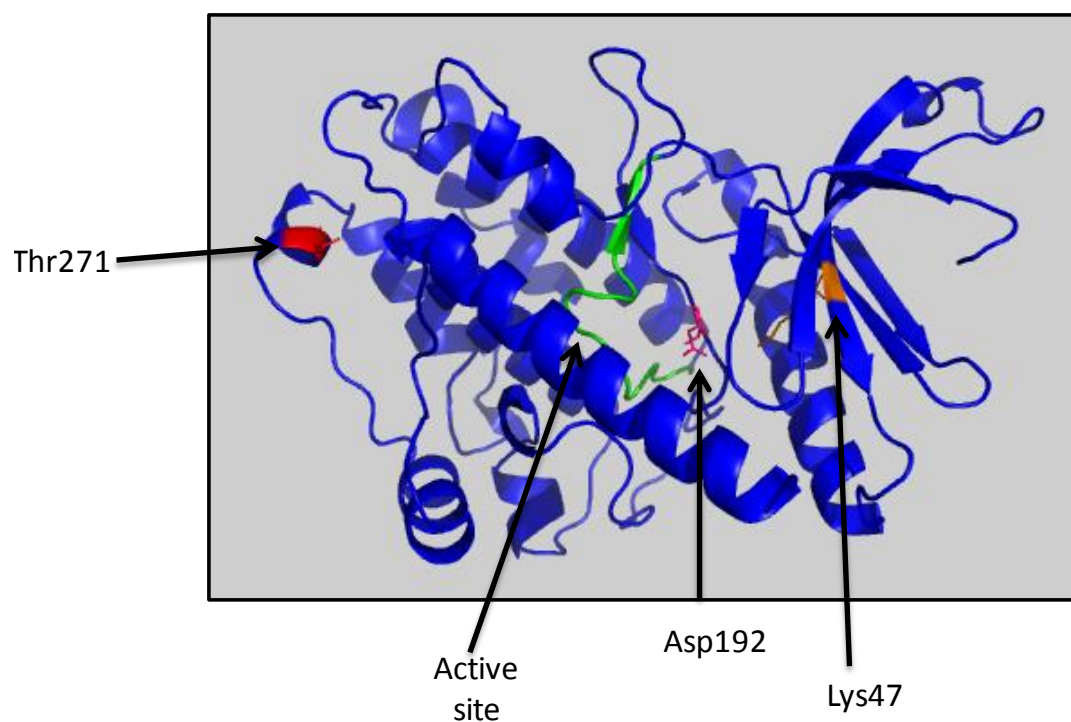
Figure 3.13. The kinase dead mutations are predicted to be located distantly from the Thr271 residue.
A model of the kinase and auto-inhibitory domain was produced with Phyre2, and the locations of Asp192 and Lys47 are marked in pink and orange respectively. Mutation of either Asp192Asn or Lys47Glu confers a classical kinase-dead phenotype. The location of the key residue of autophosphorylation, Thr271, is marked in red and is predicted to be distantly located from both kinase-dead mutations.
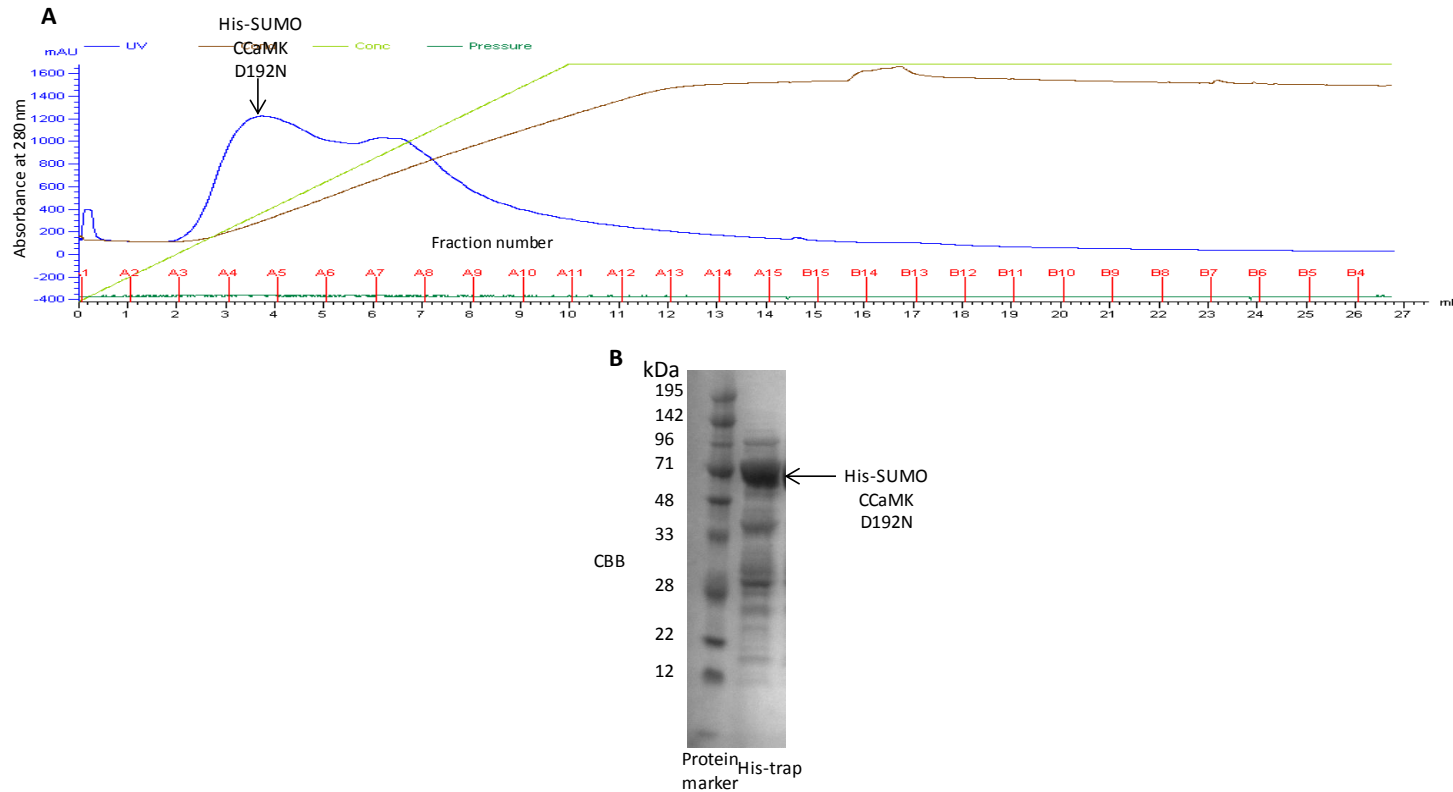
Figure 3.14. Purification of recombinant His-SUMO CCaMK D192N protein.
(A) Chromatogram for His-SUMO CCaMK D192N purified using the HisTrap column at 4 °C. The marked peak was confirmed to contain the correct protein by SDS-PAGE, as shown in **(B)**.
(B) SDS-PAGE of His-SUMO CCaMK D192N protein. The protein was observed at 72 kDa, as predicted. Non-specifically bound proteins which co-eluted with CCaMK D192N protein were also observed.

Chapter 3: Mechanism of autophosphorylation at T271
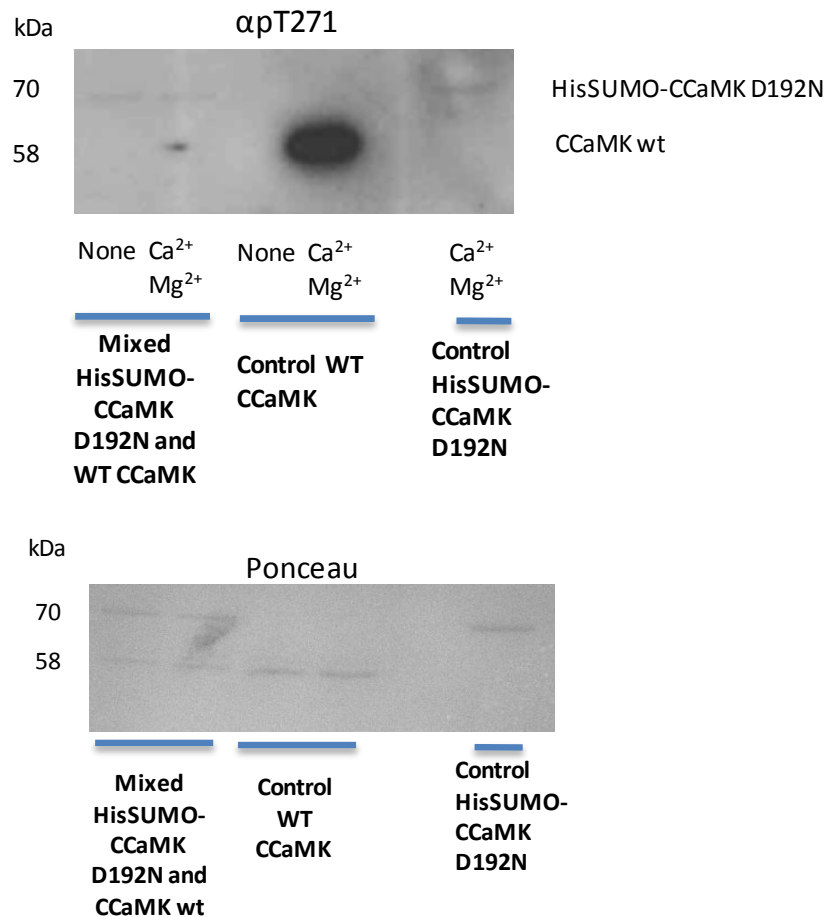
Figure 3.15. Inter-oligomeric autophosphorylation is not observed between His-SUMO CCaMK D192N and WT CCaMK.

His-SUMO-CCaMK D192N and WT CCaMK were separately expressed and mixed before a kinase assay, and subsequently subjected to western blotting probed with αpThr271. Autophosphorylation was induced with the addition of ATP, Ca^{2+} and Mg^{2+}. Phosphorylation was detected in the WT CCaMK protein and a background signal only was detected in the kinase-dead His-SUMO-CCaMK D192N protein, indicating no inter-oligomeric phosphorylation had occurred. WT tagless CCaMK showed calcium-induced autophosphorylation as expected and a background signal only was detected for His-SUMO D192N protein.

3.4 Discussion

The aim of this Chapter was to elucidate the mechanism of autophosphorylation at T271. First, it was necessary to determine the oligomeric status of CCaMK. Previous biochemical characterisation of MtCCaMK has been carried out using truncations of the protein lacking the kinase domain, which were present as monomers (Swainsbury *et al.*, 2012). In order to characterise the full-length protein, an MBP-CCaMK fusion was generated (Miller *et al.*, 2013). Previous work by Liang Zhou established the stability of this fusion protein by SEC before and after freeze-thawing of the protein. This fusion was therefore a stable, non-aggregating full-length construct of CCaMK and the most suitable candidate for oligomeric status experiments. Using SEC, AFM and DLS it was determined that MBP-CCaMK forms an oligomer of 16-18 subunits with a diameter of ~31 nm. Whilst these data provided an estimate only of the oligomer size, it nevertheless indicates that CCaMK forms a large oligomer when expressed in *E. coli*. Furthermore, previous data obtained by Liang Zhou indicates that a construct of CCaMK lacking the first five amino acids of the kinase domain does not form stable oligomers. This finding, in addition to that of Swainsbury and colleagues, indicates a likely role of the kinase domain in oligomer formation.

The ability of CCaMK to form oligomers leads to the question of whether the oligomeric state of the protein impacts on the mechanism of autophosphorylation at T271. Data presented in this Chapter indicates that this likely occurs in an intra-oligomeric manner. Whilst it was not possible to determine this with certainty, several pieces of evidence likely rule out the other possible mechanisms. Homology modelling carried out on the kinase and AI domains indicates that intra-monomeric autophosphorylation is unlikely due to the locations of T271 and the active site in the well-conserved kinase domain. In addition, three experiments carried out to determine if inter-monomeric or inter-oligomeric autophosphorylation occurs showed no evidence of the occurrence of either of these mechanisms. Finally, the lack of intra-oligomeric autophosphorylation detected in this particular experiment could be due to a number of experimental factors. These include interference of the MBP-tag due to the presence of un-tagged and tagged protein in the complex and insufficient WT CCaMK present to detect the autophosphorylation signal. The interference of the MBP-tag itself can be ruled out because this protein shows calcium-induced autophosphorylation when expressed separately (Miller *et al.*, 2013). A further possibility is that the kinase-dead mutation prevents the subunit from being

phosphorylated. However, two kinase-dead mutations were tested in the separately expressed experiment as a control and both showed the same result. However, such mutations cannot be ruled out as interfering factors.

The likely intra-oligomeric mechanism of autophosphorylation at T271 has important kinetic implications. As described above, the rate of an inter-oligomeric (or inter-monomeric) mechanism is dependent on protein concentration, whereas an intra-oligomeric (or intra-monomeric) mechanism is not. These correspond to second and first order kinetics, respectively. Therefore, it follows that an intra-oligomeric mechanism is more robust and more efficient than an inter-oligomeric mechanism. As this is most likely the case, CCaMK follows a similar mechanism of autophosphorylation to that of CaMKII at T286.

3.5 Summary

MBP-CCaMK is an oligomer of 16-18 subunits with a diameter of ~31 nm when expressed in *E. coli.* This has implications for the mechanism of autophosphorylation at T271. Homology modelling indicates this likely doesn't occur intra-monomerically, and three experiments using different CCaMK constructs do not show inter-monomer or inter-oligomeric autophosphorylation. The likely mechanism is therefore intra-oligomeric, which is kinetically more efficient and protein concentration-independent.

Chapter 4: Structural studies of CCaMK

4.1 Introduction

CCaMK is controlled by calcium in two ways: directly by binding to three EF-hands in the visinin-like domain (VLD) and indirectly, when four calcium ions bind to CaM which binds to the CaM-binding domain. Interestingly, previous data have shown that CCaMK undergoes a conformational change when calcium binds to the EF-hands of the VLD. This was first characterised in *L. longiflorum* CCaMK by Takezawa and colleagues (Takezawa *et al.*, 1996) by mobility shift assay. They demonstrated, using point mutations at each EF-hand, that EF-hands 2 and 3 are predominantly responsible for this conformational change. The conformational change has been further characterised by mobility shift assay by Shimoda and colleagues (Shimoda *et al.*, 2012). They determined that *L. japonicus* CCaMK also undergoes a conformational change, and under these circumstances EF-hand 3 is predominantly responsible. Finally, the conformational change of *M. truncatula* CCaMK was characterised by Swainsbury and colleagues (Swainsbury *et al.*, 2012) using monomeric truncations of the protein for stability. They determined that the calcium-induced conformational change corresponds to a change in tertiary structure, and reveals a hydrophobic cleft on the protein. Therefore, the conformational change that CCaMK undergoes in the presence of calcium has been characterised to some extent, however key questions remained. Consequently, two strategies were undertaken in this Chapter to further characterise this important conformational change. First, a non-aggregating full-length construct of CCaMK, MBP-CCaMK, was used to determine if the hydrophobic patch is also exposed when the kinase domain is present. Second, small-angle X-ray scattering (SAXS) was used to characterise the conformational change which occurs in the VLD.

In addition to SAXS analysis of the VLD, crystallography trials for this protein are also detailed in this Chapter. One of the key remaining challenges for CCaMK is to obtain an atomic-resolution structure for the protein. Currently, the best structural data available for CCaMK consists of homology modelling based on CaMKI (Sathyanarayanan *et al.*, 2001), CaMKII (Shimoda *et al.*, 2012) and a CDPK scaffold (Miller *et al.*, 2013). However, homology modelling of CCaMK has limitations because of low sequence similarity to potential templates so conclusions drawn using homology modelling must therefore be considered in this light. Extensive previous attempts to crystallise CCaMK have been unsuccessful due to the tendency of CCaMK to aggregate, particularly when the kinase domain is present.

Therefore, crystal trials of two truncations of CCaMK are detailed in this Chapter: the VLD and a peptide containing T271 and the CaM-binding domain.

4.2 CCaMK elongates when calcium binds

4.2.1 MBP-CCaMK exposes a hydrophobic patch when Ca^{2+} binds to the EF-hands

In order to characterise the conformational change that occurs when calcium ions bind to the EF-hands of CCaMK a hydrophobic dye, 8-anilino-1-naphthalenesulfonic acid (ANS), was used. ANS is a hydrophobic dye that increases in fluorescence when it binds to hydrophobic regions of proteins. Therefore fluorescence was recorded in the presence of ANS dye with either Ca^{2+} ions or EDTA. EF-hand containing proteins, for example CaM, expose hydrophobic residues when calcium binds which interact with target proteins (Hoeflich and Ikura, 2002). It is currently unknown as to whether full-length CCaMK exposes a hydrophobic patch, and, if it does, whether this facilitates the binding of CCaMK to itself or target proteins.

Previous work carried out by David Swainsbury showed that both the autoinhibitory domain and VLD (AI-VLD) and the His-VLD underwent a calcium-dependent increase in fluorescence in the presence of ANS (Swainsbury *et al.*, 2012). This corresponded to the exposure of a hydrophobic patch on each protein and therefore a calcium-dependent conformational change occurred in each. Furthermore, baseline fluorescence in the presence of ANS and EDTA was higher for the AI-VLD compared to the His-VLD. This indicated that hydrophobic residues were exposed when Ca^{2+} was not present in the AI-VLD and these were not present in the VLD.

In order to determine whether a hydrophobic patch was also exposed when calcium ions bound to the full-length protein, full-length MBP-CCaMK was tested (Figure 4.1). MBP-CCaMK protein was purified as described in Figure 3.1. Similarly to the AI-VLD and His-VLD, an increase in fluorescence was detected in the presence of calcium for MBP-CCaMK. This indicated that full-length CCaMK also underwent a conformational change leading to an exposure of hydrophobic residues when calcium ions bound to the EF-hands. In a manner analogous to the AI-VLD, background fluorescence was increased compared to baseline

when protein was added in the presence of EDTA. This suggested that the same hydrophobic residues exposed on the AI-VLD were also exposed in the full length protein.

Due to the nature of this technique, it was not possible to quantify the fluorescence change in any meaningful way. However, since the qualitative change in fluorescence was observed with full-length CCaMK, it can be deduced that the function of this hydrophobic patch was not to interact with other CCaMK subunits. This is because no increase in fluorescence would have been detected if the patch was already blocked by interaction with CCaMK subunits. Therefore, it follows that the function of this patch exposure could be to bind a target signalling protein which could potentially be CYCLOPS or another as yet unknown target binding partner of CCaMK.
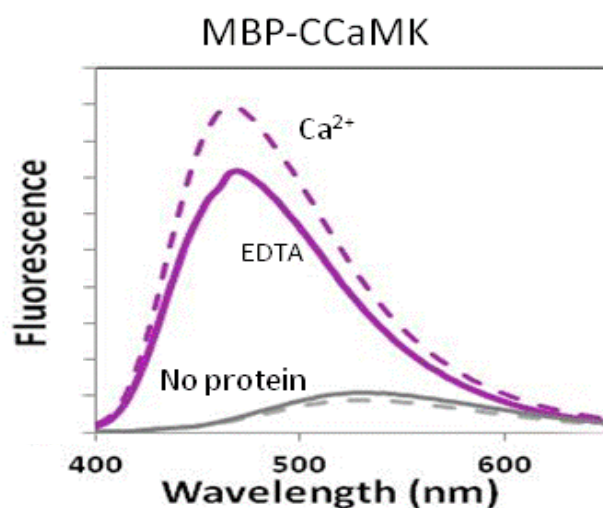
Figure 4.1. MBP-CCaMK exposes a hydrophobic patch when Ca^{2+} ions bind to the EF-hands. Fluorescence spectra of 8-anilino-1-naphthalenesulfonic acid (ANS) in the presence of MBP-CCaMK with either 2 mM $CaCl_2$ (purple dotted line) or 2 mM EDTA (purple solid line). Equivalent spectra for ANS fluorescence with no protein and either 2 mM $CaCl_2$ or 2 mM EDTA are shown with a dashed grey line and a solid grey line, respectively. The increase in fluorescence detected in the presence of calcium indicated the exposure of hydrophobic residues. This experiment was carried out in conjunction with Liang Zhou.

4.2.2 His-VLD elongates when Ca^{2+} binds to the EF-hands

In order to further characterise the conformational change that occurs when calcium binds to the EF-hands, small angle X-ray scattering (SAXS) was carried out. This technique involves the illumination of a sample in solution with X-rays and the measurement of the scattered radiation. This yields information on the size and structure of the macromolecules in solution, and can be refined to produce a SAXS envelope. Crucially, samples are required to be highly monodisperse in order to yield high quality scattering data, so therefore no full length construct of CCaMK was suitable. Hence, the AI-VLD and the His-VLD proteins were used. However, the AI-VLD protein showed unsatisfactory levels of aggregation during the scattering experiments and therefore the data was not of sufficient quality for analysis, so it was not considered further.

His-VLD protein was purified as shown in Figure 4.2. In order to isolate protein of sufficient purity, a two-step purification was used. Nickel-affinity chromatography (Figure 4.2A) was utilised followed by SEC (Figure 4.2B) to remove contaminants. The size and purity of the protein was confirmed using SDS-PAGE (Figure 4.2C). SAXS data were collected at the X33 beam line on the DORIS III synchrotron at EMBL, DESY in Hamburg, Germany. An original set of data was obtained by David Swainsbury, but there were concerns with these data due to low statistical significance of the difference observed between the forms in the presence and absence of calcium. This is an issue because of the low-resolution of this technique. Therefore, I carried out two further attempts to improve upon the quality of the data obtained. Strategies to improve the data included the use of DTT to reduce aggregation and the use of the same protein and buffer solutions divided between calcium and EDTA treatments. However, an improvement on the original data was not possible due to a little aggregation which occurred in both new datasets in a manner analogous to AI-VLD. The data presented in Figure 4.3, therefore, corresponds to a previously acquired data set for the His-VLD obtained by David Swainsbury, which was subsequently re-analysed by Dmitri Svergun. This data analysis was carried out subsequently to my newly obtained data, and therefore is presented in this thesis for the first time. The newly-analysed data are consistent with both datasets that I acquired.

Scattering profiles were obtained for the His-VLD at various protein concentrations within a range of 0.5 - 7.4 mg/ml and 1.9 - 7.4 mg/ml with CaCl$_2$ and EDTA, respectively (Figure 4.3). The *ab initio* models constructed from the SAXS data were fitted to the calcium-bound

structure of a human neuronal calcium sensor (Heidarsson *et al.*, 2012). This neuronal calcium sensor contains four EF-hands of which three are able to bind calcium, analogous to the VLD of CCaMK (Swainsbury *et al.*, 2012). Furthermore, this template protein undergoes a conformational change on calcium-binding, revealing a hydrophobic crevice, allowing the protein to bind to targets (Ames and Lim, 2012). Overall, the SAXS envelopes for the His-VLD were broadly similar with and without calcium. However, a conformational change was observed leading to an overall extension of the structure when calcium was present. Surprisingly, although the template structure from which the model was derived was calcium-bound, the template showed a better fit to the envelope in the presence of EDTA. Due to the limit of resolution, the exact structural change that occurred could not be modelled. Furthermore, the statistical difference observed between treatments was marginal. Nevertheless, these data indicate the likely structural change occurring in the VLD. Moreover, these data correlate with the observation in section 4.2.1 that when calcium binds to the EF-hands, a hydrophobic patch is exposed on full-length CCaMK indicating a conformational change. Therefore, this change likely corresponds to an elongation of the EF-hand domain.
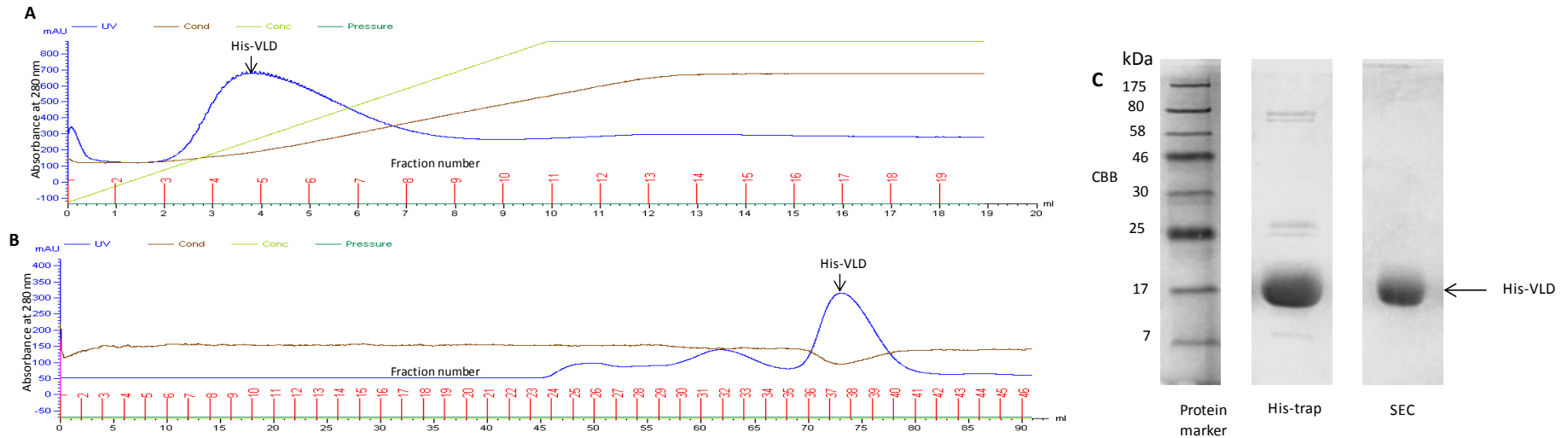
Figure 4.2. Purification of recombinant His-VLD for small angle X-ray scattering (SAXS).

(A) Chromatogram for the purification of His-VLD by nickel affinity chromatography using a HisTrap column. Fractions corresponding to the marked peak were collected, pooled and subjected to size exclusion chromatography as shown in **(B)**.

(B) A size exclusion chromatogram for the His-VLD using a HiLoad Superdex 75 prep grade 16/60 column. The marked peak was confirmed to contain the correct protein by SDS-PAGE shown in **(C)**.

(C) SDS-PAGE of HisTrap and SEC fractions of His-VLD. His-VLD was observed at 21 kDa, as predicted, and was highly pure. Fractions from the SEC purification were pooled and dialysed, concentrated appropriately and sent on dry ice for SAXS analysis.
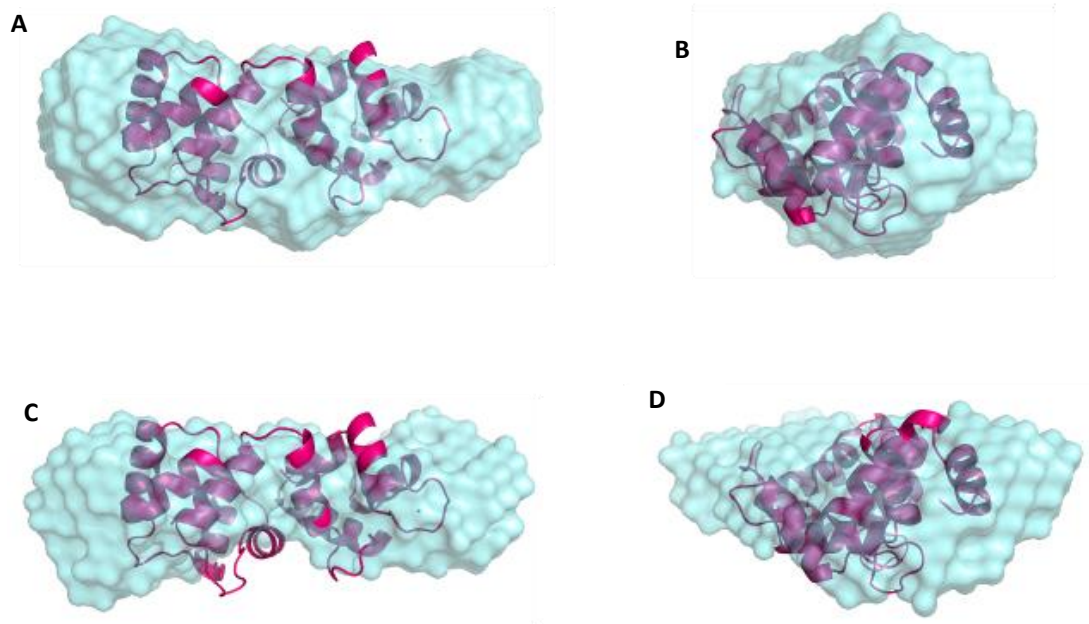
Figure 4.3. *ab initio* SAXS envelopes of the His-VLD show an elongation of the protein when calcium binds to the EF-hands (Figure courtesy of David Swainsbury and Dmitri Svergun).

(A) SAXS envelope of the His-VLD (blue) fitted to a homology model of the protein generated against a calcium-bound human neuronal calcium sensor (Heidarsson *et al.*, 2012) using SWISSMODEL (magenta), in the presence of 0.1 mM EDTA. Readings were taken at various protein concentrations within a range of 1.9 - 7.4 mg/ml (90 - 352 µM).

(B) 90° rotation of the His-VLD SAXS envelope in the presence of 0.1 mM EDTA.

(C) SAXS envelope of the His-VLD (blue) fitted to the SWISSMODEL homology model of the protein (magenta) in the presence of 2 mM CaCl$_2$. Readings were taken at various protein concentrations within a range of 0.5 - 7.4 mg/ml (24 - 352 µM).

(D) 90° rotation of the His-VLD SAXS envelope in the presence of 2 mM CaCl$_2$.

4.3 Crystallography trials of His-VLD with DAPase digestion

To date, no crystal structure of CCaMK has been elucidated. A crystal structure of CCaMK would clarify many questions which still remain as to the mechanism by which CCaMK functions. Full-length CCaMK is unsuitable for crystallisation due to its tendency to aggregate. Therefore constructs of CCaMK were considered which were potentially more stable than the full-length form, in particular those that excluded the problematic kinase domain. The AI-VLD which lacked the kinase domain only was not suitable due to low yields. Previous work performed by David Swainsbury using the His-VLD demonstrated that the stability and solubility of the protein was significantly increased by the removal of the His tag using DAPase digestion. Therefore, crystallisation trials were attempted with this construct.

His-VLD was purified as shown in Figure 4.2A with no SEC purification. Protein was dialysed overnight in buffer containing in 10 mM Tris pH 7.5, 50 mM NaCl, 1 mM DTT. Purified protein was digested with DAPase which removed the His tag and yielded the protein shown in Figure 4.4A from the red arrow onwards. Successful digestion of the protein was confirmed using a 17 % SDS-PAGE gel that was sufficient to show a gel shift corresponding to the tag removal (Figure 4.4B). Contaminants which co-purified with the His-VLD, un-digested His-VLD protein, the cleaved tag and the DAPase enzyme (which possessed a C-terminal His tag) were removed with reverse immobilised metal affinity chromatography (IMAC). The purity of the digested protein was confirmed by SDS-PAGE (Figure 4.4C). The protein was concentrated to 15 mg/ml and frozen in liquid nitrogen followed by storage in aliquots at -80 °C prior to use.

Crystallisation screens were set up for digested VLD protein as described in Table 4.1 and plates were stored at 20 °C. Screens were monitored for approximately one month, and 24-well optimisation screens were set up for conditions that yielded potential protein crystals, namely condition C2 in the JCSG screen (0.2 M ammonium dihydrogen phosphate, 0.1 M Tris pH 8.5, 50% (v/v) MPD) and H9 in the structure screen (0.1 M MES pH 6.5, 12% (w/v) PEG 20 000). Two crystals which formed in the C2 JCSG optimisation screen were frozen in liquid nitrogen and data was collected for one crystal using the *Rigaku* RU-H3R rotating anode X-ray generator at the John Innes Centre by Clare Stevenson. However, the diffraction pattern corresponded to that of a salt crystal and not a protein crystal.

Therefore no conditions from these screens yielded protein crystals for structure determination.

In order to determine if the presence or absence of calcium stabilised the VLD, further crystallisation trials were carried out using digested VLD protein in the presence of either 1 µM $CaCl_2$ or 1 mM EDTA. Protein was concentrated to 11 and 12 mg/ml respectively and crystallisation screens were set up as described in table 4.1. Screens were monitored for approximately one month and 24-well optimisation screens were set up for two conditions: B11 in the PEG screen (0.2 M potassium sulfate, 20% (w/v) PEG 3350) and C4 in the structure screen (0.1 M HEPES pH 7.5, 1.5 M lithium sulfate), both in the presence of calcium. The B11 PEG condition optimisation was set up with an additional 0.2 µl crystal seed-stock from the original screen present in the drop. However, none of the conditions which were tested yielded protein crystals in the presence of either calcium or EDTA.

A

MHHHHHHFLRTKKLKSLVGSYDLKEEEIENLRMHFKKICADRDNATLSEFEEVLKAMNMLSLIPFASRIFDLF
DNNRDGTVDMREILCGFSSLKNSKGEDALRLCFQMYDTDRSGCISKEEVASMLRALPYDCLPTDITEPGKLD
EIFDLMDANNDGKVTFDEFKAAMQRDSSLQDVVLSSIRP
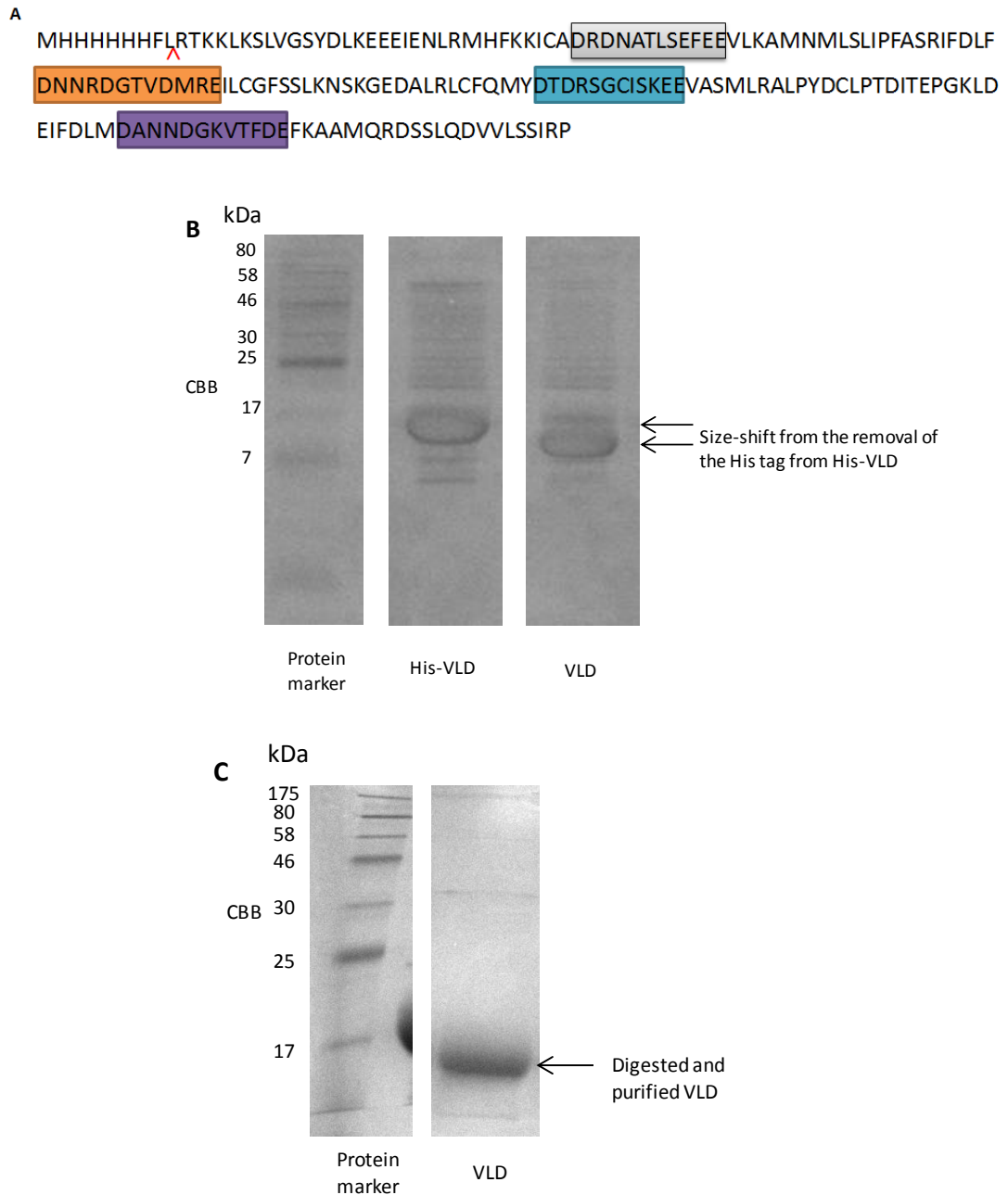
Figure 4.4. His-VLD digestion with DAPase for VLD crystallography trials.
(A) The amino acid sequence of the His-VLD construct. The three functional EF-hands are marked with coloured boxes, and the non-functional EF-hand is marked with a grey box. The site of DAPase cleavage is marked with a red arrow.
(B) SDS-PAGE (17%) of His-VLD protein purified by nickel affinity chromatography using a HisTrap column and the VLD protein after removal of the His tag by digestion with DAPase.
(C) SDS-PAGE of DAPase-digested VLD after reverse IMAC purification to remove the DAPase enzyme and contaminants, which co-purified with the His-VLD. The VLD protein was subsequently concentrated and used for crystallography trials.

Protein construct	Protein concentration (mg/ml)	Additional conditions	Screens set up
VLD	15	-	PEG, ammonium sulphate, PACT, JCSG, Structure, Morpheus, Midas
VLD	11	1 μM $CaCl_2$	PEG, ammonium sulphate, PACT, JCSG, Structure, Morpheus, Midas
VLD	12	1 mM EDTA	PEG, ammonium sulphate, PACT, JCSG, Structure, Morpheus, Midas

Table 4.1. Crystallisation trial conditions for the VLD.

4.3.1 Crystallisation trials of the VLD with surface entropy reduction

Successful crystallisation of a protein involves the development of intermolecular contacts and the subsequent formation of a crystalline lattice. The side chains that are surface exposed on a protein can potentially affect the ability to crystallise. Surface entropy reduction (SER) involves the selection of high-entropy surface-localised residues for mutation to residues of lower entropy. This involves the mutation of a small cluster of high entropy residues, e.g. Lys, Glu or Gln, to Ala and has been shown to increase the ability of some proteins to crystallise (Goldschmidt *et al.*, 2007).

Therefore, in order to thermodynamically favour the crystallisation of the VLD, residues were selected for mutation for SER using the SERp server (Figure 4.5). Three clusters were predicted for mutation, and ranked by SERp score, where the highest score corresponds to the best candidate for mutation. A homology model of the VLD was generated using Phyre2 to map the spatial locations of the three selected clusters. Since cluster 2 overlapped with an EF-hand domain, this cluster was not selected. Conserved sequences were not recommended for mutation because mutation could negatively affect the structure of the protein. Cluster 1 and 3 were both surface exposed in the homology model, however it was not recommended to mutate more than one cluster. Therefore cluster 1 was selected for mutation due to the higher SERp score.

Mutation of the cluster from KEEEIE to AAAEIE as recommended by the SERp server was carried out using the QuikChange Lightning site-directed mutagenesis kit and confirmed by sequencing. The protein was expressed, purified and digested as described for the VLD (Section 4.3). Protein was concentrated to 19 mg/ml and crystallisation screens were set up as described in table 4.2, with plates stored at 20 °C. Screens were monitored for approximately one month, and 24-well optimisation screens were set up with conditions that yielded potential protein crystals, namely condition A9 from the customised KISS screen (0.1 M MES pH 6, 10% (w/v) PEG 3350) and A10 from the KISS screen (0.1 M MES pH 6, 20% (w/v) PEG 3350). A further optimisation screen was set up using crystal seed stock from the A9 optimisation screen, which was added to a 96-well PEG screen and monitored. Despite these efforts, potential crystals corresponded to fungal contamination and SER of the VLD yielded no protein crystals for structure determination.
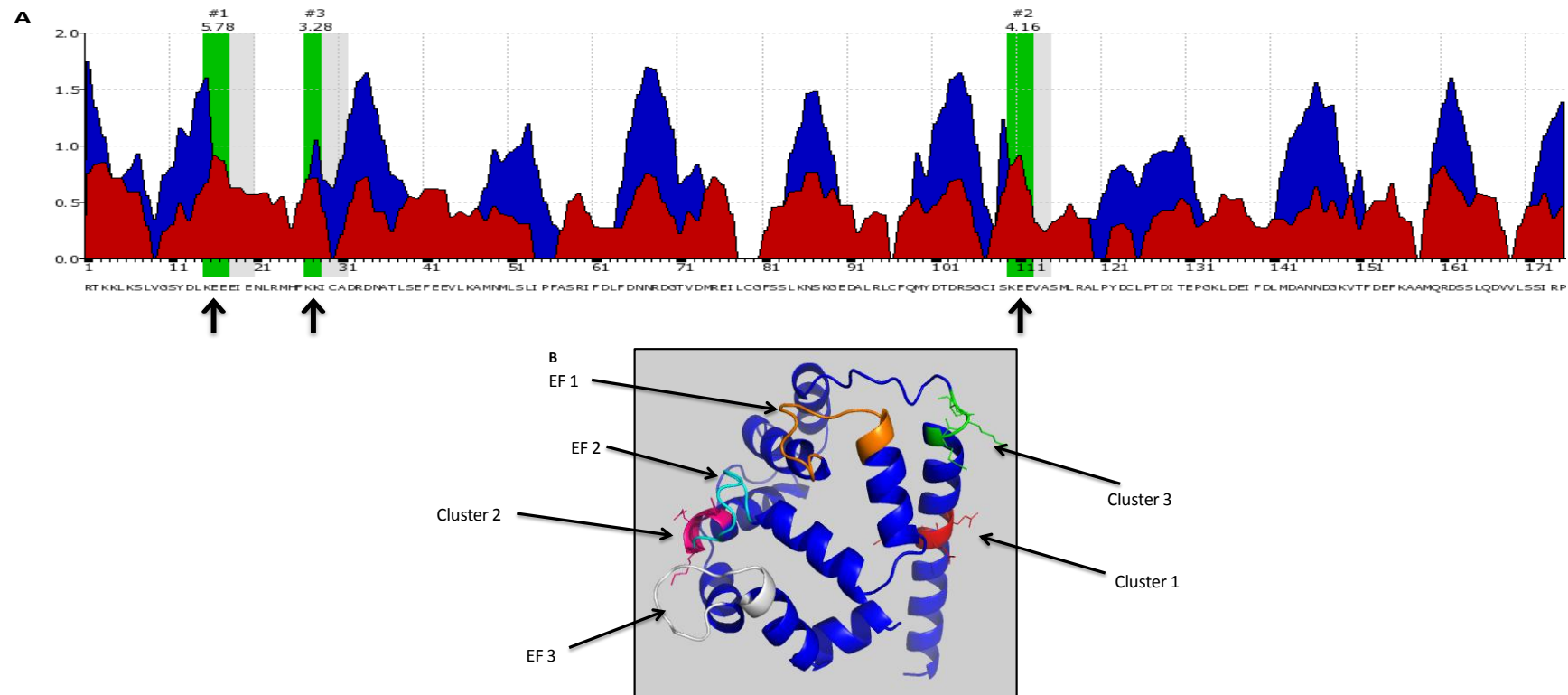
Figure 4.5. Three high-entropy residue clusters are predicted for surface entropy reduction of the VLD by the SERp server.
(A) Surface-exposed and high-entropy residues for potential mutation were predicted for the VLD protein with the SERp server. The three predicted clusters are marked 1-3, with the corresponding SERp scores shown above. Green bands show predicted residues for mutation in each cluster, with low entropy residues marked in grey. Blue peaks correspond to coiled coil predictions and red peaks correspond to the predicted average entropy of residues.
(B) Homology model of the VLD produced with Phyre 2 showing the locations of EF-hands 1, 2 and 3 in orange, cyan and white respectively and the three proposed clusters marked in red, pink and green, respectively. Cluster 2 overlapped with EF-hand 2 and was therefore disregarded, and cluster 3 had a markedly lower SERp score than cluster 1 or 2. Therefore cluster 1 was selected for mutation due the highest SERp score and residues 15-17 (KEE) were mutated to alanines.

Protein construct	Protein concentration (mg/ml)	Additional conditions	Screens set up
VLD with SER	19	-	KISS, JCSG, Morpheus, Structure

Table 4.2. Crystallisation trial conditions for the VLD with SER.

4.4 Crystallography trials of His-SUMO-CCaMK peptides

Attempts to crystallise the VLD of CCaMK yielded no protein crystals for structural studies, therefore alternative truncations of CCaMK were considered for crystallisation trials. Previous work carried out by Sathyanarayanan and colleagues (Sathyanarayanan *et al.*, 2001) utilised an 84 amino-acid peptide from *L. longiflorum* CCaMK for CaM-binding experiments. This peptide possessed the T267 residue (equivalent to T271 in MtCCaMK) and the CaM-binding site. Interestingly, this peptide showed a five-fold increase in affinity for calmodulin when phosphorylated at T267, compared to a seven-fold increase detected when full-length CCaMK was used. Therefore the peptide retained some of the functionality of the full-length protein and it follows that it retained some of the same structural characteristics as when it was present in the full-length protein. Since acceptable homology models can be produced of the kinase domain with the autoinhibition domain (Figure 3.4) and the VLD (Figure 4.5), the data provided by a crystal structure of the link between the kinase and VLD would provide crucial missing information about how CCaMK is activated and deactivated.

Therefore, the equivalent MtCCaMK peptide (Figure 4.6A) was sub-cloned for expression with a His-SUMO tag in the pOPINS3C vector. Two versions of this construct were produced for use in structural experiments: version 1 which possessed a non-cleavable His-SUMO tag and version 2 which possessed a 3C cleavage site for removal of the tag. Both His-SUMO-CCaMK peptides were purified using nickel affinity chromatography with a HisTrap column (Figure 4.6B) and the presence of the correct protein was verified by SDS-PAGE (Figure 4.6C and D, version 1 and 2, respectively). The presence of both proteins was also confirmed by mass spectrometric analysis. His-SUMO-CCaMK peptide version 1 was significantly purer after purification. However, since version 2 of the peptide possessed a cleavage site, the protein would be subsequently purified further by subtractive IMAC. Therefore both peptides were sufficiently pure for structural studies.

His-SUMO-CCaMK peptide (version 1) was prepared for crystallisation trials by buffer exchange to 20 mM Tris pH 8 and 100 mM NaCl and subsequent concentration to 15 mg/ml. Crystallisation screens were set up for this protein as described in Table 4.3. Screens were monitored for approximately one month however no potential protein crystals were observed.

A

MNGNFSFYEK**T**WKGISQPAKNLISSLLTVDPSKR
PSALELLSDPWVKGEKAKDVQMDPEIVSRLQSF
NARRKLRAAAIASVWSS

B

C

D
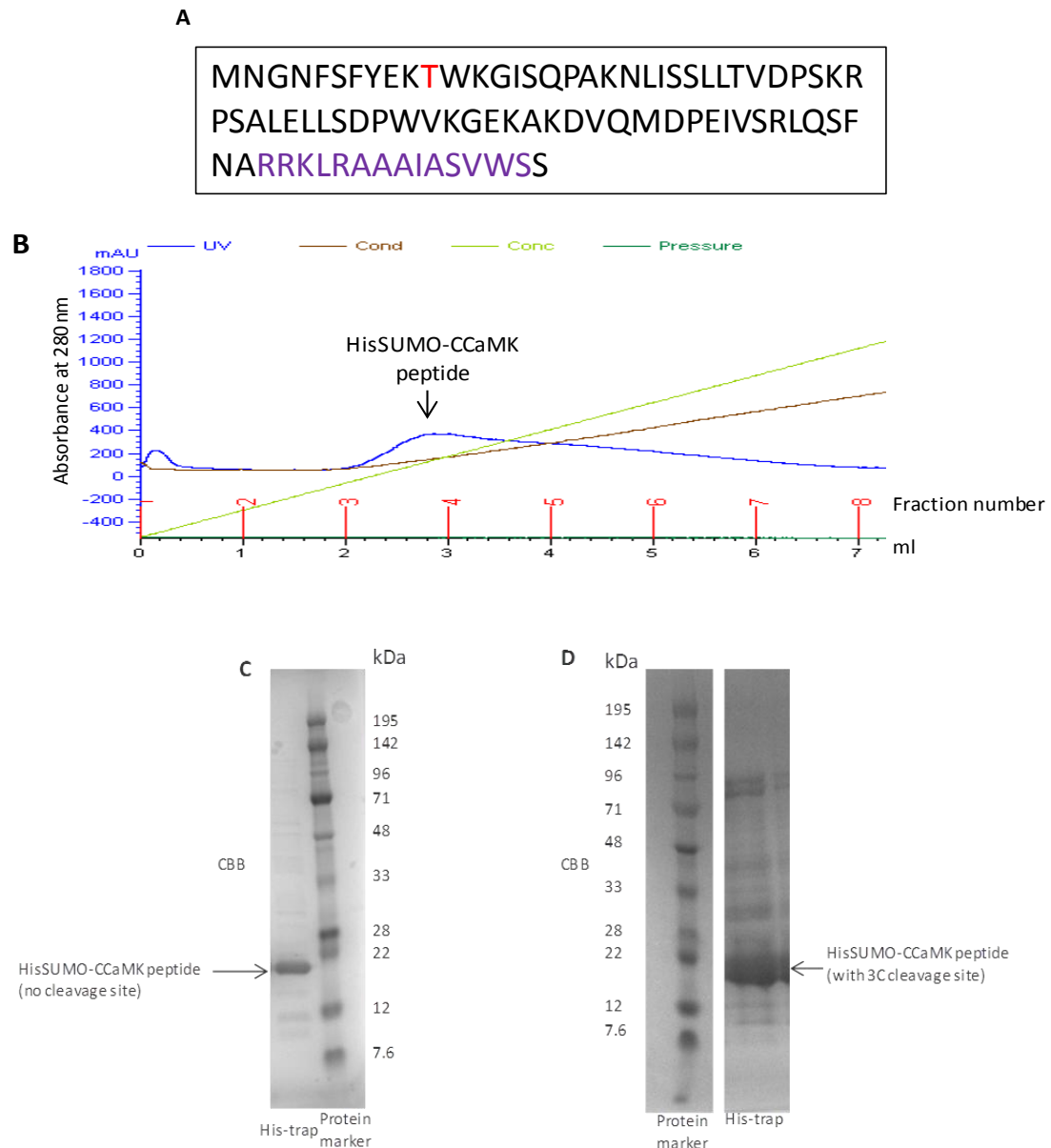
Figure 4.6. Purification of the recombinant His-SUMO CCaMK peptides.
(A) Sequence of the CCaMK peptide present in the pOPINS3C vectors. The Thr271 residue is marked in red, and the CaM-binding site is marked in purple. The same amino acid sequence was present in both versions of the His-SUMO-CCaMK peptide vectors. Version 1 of the vector lacked a 3C protease cleavage site, and version 2 possessed a 3C protease cleavage site for removal of the His-SUMO tag from the purified protein.
(B) Representative chromatogram for purification of the His-SUMO-CCaMK peptides purified using the HisTrap column. Fractions corresponding to the marked peak were collected and verified by SDS-PAGE as shown in **(B)** and **(C)**.
(C) SDS-PAGE of purified His-SUMO CCaMK peptide without a 3C cleavage site (version 1). The protein was predicted to be 23.7 kDa (peptide = 9.4 kDa, His-SUMO tag = 14.3 kDa) and observed at ~21 kDa. The presence of the correct protein was verified by mass spectrometric analysis. This purification was performed by Karl Syson.
(D) SDS-PAGE of purified His-SUMO CCaMK peptide with a 3C cleavage site (version 2). The protein was predicted to be 23.7 kDa (peptide = 9.4 kDa, His-SUMO tag = 14.3 kDa) and

observed at ~21 kDa. Non-specifically bound proteins which co-eluted with the SUMO peptide were also observed. The presence of the correct protein was verified by mass spectrometric analysis.

Protein construct	Protein concentration (mg/ml)	Additional conditions	Screens set up
His-SUMO-CCaMK peptide (version 1, no cleavage site)	15	-	KISS, PEG, Structure, PACT and JCSG

Table 4.3. Crystallisation trial conditions for the His-SUMO-CCaMK peptide lacking a cleavage site (version 1).

4.4.1 Digestion trials of His-SUMO-CCaMK peptide with the 3C protease

His-SUMO-CCaMK peptide (version 2) was subjected to 3C protease cleavage trials in an attempt to produce a tagless version of the peptide for further crystallisation trials (Figure 4.7). Digestion was carried out at 4 °C for a minimum of 16 hours with 10 μg/μl 3C protease. Surprisingly, no digestion of the His-SUMO-CCaMK peptide was observed with these typical digestion conditions (Figure 4.7A) whereas digestion of a control protein was observed (Figure 4.7B). The size shift of the protein, which corresponded to removal of the tag, was 14.3 kDa, as expected. Despite the presence of the 3C protease at approximately the same size as the His-SUMO-CCaMK peptide, it was established by mass spectrometric analysis that the His-SUMO-CCaMK peptide was present. Further digestion trials of the His-SUMO-CCaMK peptide (version 2) were carried out in the presence of 5 mM DTT (in an attempt to stimulate the activity of the cysteine protease) and at room temperature (approximately 20 °C) however no digestion was observed under these alternative conditions (data not shown).

Since the cleaved His-SUMO tag from the control protein was observed at approximately the same size as the His-SUMO-CCaMK peptide (Figure 4.7A-B), IMAC was carried out to determine if the protein was digested however no gel shift was observed (Figure 4.7C). No evidence of digested CCaMK peptide was seen in the flow-through fractions which represented proteins which did not bind to the HisTrap column. However, a strong band corresponding to the undigested His-SUMO-CCaMK peptide was detected in the fraction that corresponded to proteins that were eluted from the column using 500 mM imidazole. Therefore, under the conditions tested no evidence of the digestion of the His-SUMO-

CCaMK peptide was observed, which was most likely due to obstruction of the cleavage site by the conformation of the protein.

In addition, purification trials were attempted with the same peptide fused to a His-tag in the Topo100 vector. However, protein was not expressed to a sufficient level to attempt crystal trials using this expression system.

Therefore, no protein crystals of CCaMK were obtained by these experiments and therefore no structural elucidation could be carried out.
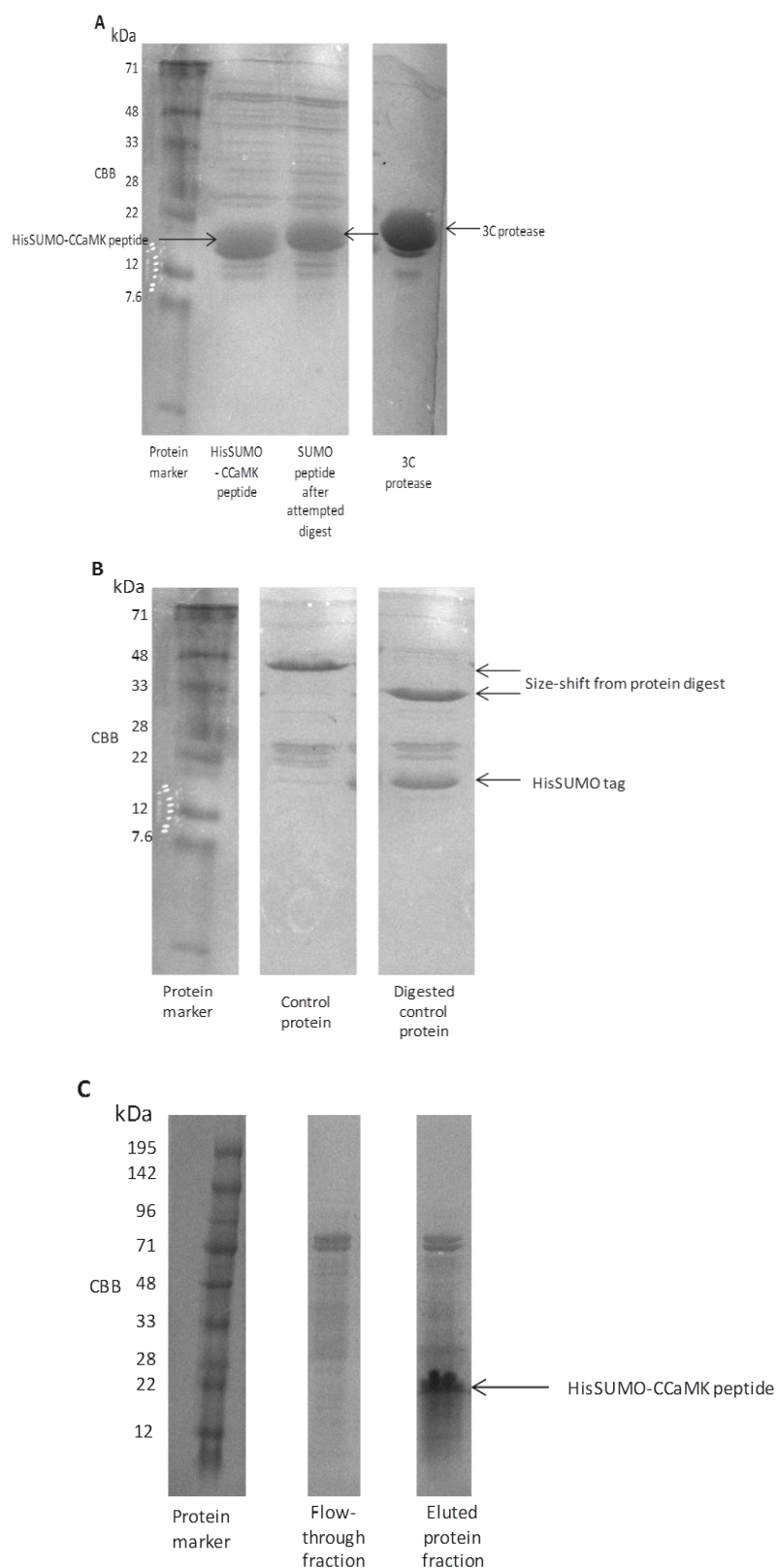
Figure 4.7. 3C protease digestion trials of the His-SUMO-CCaMK peptide (version 2).
(A) SDS-PAGE of a 3C protease digestion trial for His-SUMO-CCaMK peptide where no size shift corresponding to His-SUMO tag cleavage was observed. The 3C protease was observed at 22 kDa. The predicted size of SUMO-peptide is 23.7 kDa (Peptide = 9.4 kDa,

Chapter 4: Structural studies of CCaMK

His-SUMO tag = 14.3 kDa), and the observed size was smaller at approximately 21 kDa. The presence of the His-SUMO-CCaMK peptide was confirmed by mass spectrometric analysis. **(B)** Digestion of a control His-SUMO tagged protein (52 kDa predicted, 47 kDa observed) that showed the expected size shift of 14.3 kDa which corresponded to the successful removal of the His-SUMO tag. This digestion was carried out under the same conditions as used for the His-SUMO-CCaMK peptide digestion trial shown in **(A)**.

(C) SDS-PAGE of the subtractive IMAC for the His-SUMO-CCaMK peptide after a 3C protease digestion trial. The correct protein was observed only in eluted protein fractions and not in the flow-through wash fractions and therefore the protein was not digested.

4.5 Discussion

The first aim of experimental work presented in this Chapter was to characterise the conformational change that occurs when calcium binds to the EF-hands of the VLD. EF-hand proteins typically undergo a calcium-induced conformational change in order to bind target proteins (Gifford *et al.*, 2007). CaM is a calcium binding protein that contains four EF-hands present in pairs on two globular domains, separated by a helical linker region. When calcium ions bind, CaM also undergoes a conformational change, and interestingly this corresponds to the exposure of hydrophobic residues in the linker region which bind to targets (Bhattacharya *et al.*, 2004). Data presented in this Chapter demonstrates that this exposure of hydrophobic residues in the presence of calcium also occurs in full-length CCaMK. This had previously been demonstrated as the case for truncations of the protein only, lacking the kinase domain or both the kinase and AI domains (Swainsbury *et al.*, 2012). Furthermore, data presented in this Chapter showed the conformational change as an elongation of the VLD in the presence of calcium ions, using SAXS.

The exposure of a hydrophobic patch in response to calcium binding in CCaMK has now been observed in the VLD, AI-VLD (Swainsbury *et al.*, 2012) and full length MBP-CCaMK. This demonstrates that the conformational change that occurs in the VLD still occurs in the context of the whole protein. Furthermore, this indicates that the likely binding target of this hydrophobic patch is not the interaction with other CCaMK monomers. Instead, it is likely that the function of this patch is to bind another target of CCaMK, perhaps CYCLOPS or another phosphorylation target. The SAXS data obtained demonstrated further that this hydrophobic patch exposure corresponds to a likely elongation of the VLD. The SAXS data obtained is clearly consistent with the ANS data presented in this Chapter and with the conformational change data presented in (Swainsbury *et al.*, 2012). Therefore, these data taken together provide strong evidence for the hydrophobic patch corresponding to elongation of the VLD.

It is currently unclear as to exactly which EF-hands are responsible for this structural change, and most likely it requires the functionality of all three EF-hands. First, it was demonstrated that functionality of EF-hand 2 and 3 was necessary for this conformational change (Takezawa *et al.*, 1996), and it was further confirmed that functionality of EF-hand 3 was necessary by Shimoda and colleagues (Shimoda *et al.*, 2012). Interestingly, analysis of

calcium affinities by Swainsbury and colleagues (Swainsbury *et al.*, 2012) showed that EF-hand 3 corresponds to a high affinity site which likely has calcium bound at basal calcium levels. Therefore, it seems possible that the conformational change recorded in the experiments presented in this Chapter correspond to the binding of calcium to EF-hands 1 and 2, but that the calcium-binding functionality of EF-hand 3 is also necessary for this change. Further work will be required to unpick this complex regulation.

Crystal trials for two constructs of CCaMK were also presented in this Chapter. Two truncations of CCaMK were selected for this work because full-length CCaMK has a tendency to aggregate with or without a tag over longer time scales (Swainsbury *et al.*, 2012). The unsuitability of full-length CCaMK for crystallography was confirmed by a crystallisation prediction using the XtalPred server (Slabinski *et al.*, 2007). Predictions are based on a variety of factors including length, predicted structural disorder and isoelectric point. Two predictions carried out by this server determined that CCaMK is in the "difficult" and "least promising" categories for crystallisation. Therefore, extensive crystal trials were carried out for truncations of CCaMK. For the VLD, the full available range of precipitant conditions was trialled. In addition, the presence of calcium or EDTA and SER was tested in an attempt to improve the ability of the VLD domain to crystallise. In addition, two different tags were tested for the expression of a central peptide, and the His-SUMO-tagged peptide was selected for use in crystal trials. A range of precipitant conditions were tested and removal of the tag was unsuccessfully trialled. The tag was most likely unable to be removed due to the obstruction of the cleavage site when fused to the peptide because a control protein was digested under identical conditions. After these extensive crystal trials, no crystals of CCaMK were obtained. It therefore remains an important future goal to obtain high-resolution structural data for CCaMK in order to validate the model of function (Miller *et al.*, 2013).

4.6 Summary

CCaMK undergoes a conformational change when calcium binds to the EF-hands. This corresponds to an elongation of the VLD and this conformational change is also present in the full-length protein. Furthermore, this corresponds to the exposure of a hydrophobic patch which is likely involved in the binding of CCaMK to phosphorylation targets. Crystal

trials for truncations of CCaMK are detailed; however these were unsuccessful in the production of crystals for structure elucidation.

Chapter 4: Structural studies of CCaMK

Chapter 5: CCaMK *in planta*

5.1 Introduction

The aim of the experimental work in this chapter was to determine to what extent it is possible to purify CCaMK from *M. truncatula.* This involved the development of a multi-stage purification strategy. First, a growth strategy for *M. truncatula* was required in order to maximise the yield of root tissue obtained. This was necessary because the protein is likely present in cells at a low concentration. Commercial wild-type *M. truncatula* seeds were obtained in order to have enough material for large-scale growth trials, because it was anticipated that root material in the order of 500 g would be required for protein on a milligram scale, based on studies on other plant proteins (Requena, 2000). *M. truncatula* Jester seeds were selected for this, because the cultivar is ~90 % similar to Jemalong (Kamphuis *et al.*, 2013), from which the commonly used A17 line is derived. Alternative sources of protein were also explored and are detailed in this section, namely an over-expression line for tagless WT protein and an *M. truncatula* line with CCaMK under the native promoter and dual-tagged with GFP and SBP.

Second, it was necessary to develop an antibody in order to monitor the purification strategies tested. For this, a construct of CCaMK which lacked the kinase domain was purified in *E. coli* for use as an antigen, and tested for specificity against various constructs of CCaMK. Once this system was established, a wide variety of purification strategies were tested in order to determine to what extent MtCCaMK can be extracted and characterised. The first aim of these trials was to establish if it was possible to obtain a single band on a western blot probed with the developed antibody that corresponded to MtCCaMK. This was carried out to determine if the size of CCaMK in *M. truncatula* corresponds to the expected protein size, for example as determined in *L. longiflorum* at approximately 56 kDa (Patil *et al.*, 1995). Subsequently, purification trials were carried out in order to determine whether it was possible to purify sufficient protein for biochemical characterisation. Protein purified from *M. truncatula* could be used for key experiments to validate experiments carried out with constructs of CCaMK expressed in *E. coli.* For example, as detailed in Chapter 3, recombinant MBP-CCaMK forms an oligomer of 16-18 subunits which has important consequences for the mechanism of autophosphorylation. Therefore, the number of subunits of CCaMK present in this oligomer *in planta* could be addressed. Furthermore, as detailed in Chapter 4, previous attempts to crystallise CCaMK when purified in *E. coli* have been unsuccessful, however it is possible that CCaMK is more stable

when purified from the native host. Therefore, this Chapter details the purification trials carried out for MtCCaMK.

5.2 Development of a growth strategy for Jester

In order to purify as much MtCCaMK as possible, a growth strategy was required which maximised root growth. Growth on agar plates was deemed too labour intensive and low yielding, while growth on soil would have required extra cleaning steps to be viable and so was not considered further. Two possible growth strategies were therefore identified and trialled: aeroponics and hydroponics. Figure 5.1A shows a representative trial for an aeroponic growth of A17 in a caisson chamber. Seedlings were individually planted in a perforated plate located at the top of chamber (left panel), and roots grew down towards a pump located at the bottom of the chamber, which misted the growing roots with BNM medium. Root growth was excellent with this method, with roots extending significantly further than when grown by other methods. However, when scaled up and planted onto a mesh instead of a plate, growth was less optimal – plants either became infected because they were excessively moist, or dried out too much. It was also theorised that whilst aeroponics allowed extensive root growth, this could be sub-optimal for protein extraction since it may be the case that the extension corresponds to higher water content and not more protein.

Therefore, hydroponics was trialled as an alternative. A system established in the Bornemann group for the purification of a barley enzyme was trialled due to its previous success (Figure 5.1D). This was trialled initially with 1000 A17 seeds, and root growth was determined to be healthy and substantial without excessive root elongation (Figure 5.1B). Seeds were bleached, scarified and imbibed for approximately 5 hours and then spread in a single layer on the mesh. Sixteen litres of BNM medium was added to the tank, and aeration was provided by two air bricks per tank to minimise stress caused by anaerobicity, which can lead to infection. This also allowed the medium to moisten the mesh and seeds and provide sufficient moisture for seed germination. In this set up, it was also established that despite the fact that roots were present in liquid, nodulation was possible (Figure 5.1C). A subset of the seedlings growing in the mesh were inoculated by immersion in 500 ml BNM medium containing rhizobia (OD_{600} = 0.001) for several days. They were then

returned to the main un-inoculated tank, and pink nitrogen-fixing nodules were observed after approximately three weeks.

In order to produce enough root material for preparative protein extraction, growth in the hydroponics system was scaled up ten-fold and two tank systems were used simultaneously (Figure 5.1E). Jester seeds were used instead of A17 due to the low availability of A17, and after a trial it was revealed that higher germination rates were achieved by eliminating the scarification step. The tanks were also re-located to a controlled environment room to further increase yield. Seedlings were grown for approximately 3 weeks, which produced large quantities of root material and prevented overcrowding of the individual plants in the tank. Although other nitrogen-rich media were considered, growth was excellent with BNM nitrogen-limiting medium so this was used for all subsequent hydroponics growths. This also allowed the potential for nodulation experiments if required. From 20 000 seeds planted in one batch, approximately 800 g of root material was harvested, blotted briefly and flash frozen in liquid nitrogen for future experiments.

A

Lid

Location of seedling plate

Location of media pump

Complete aeroponic set-up   Top view of seedling plate   Bottom view of seedling plate

B

C

D

Mesh

BNM medium

Tray

Airbrick

M. truncatula seeds

Air pump tubing

Air pump

E
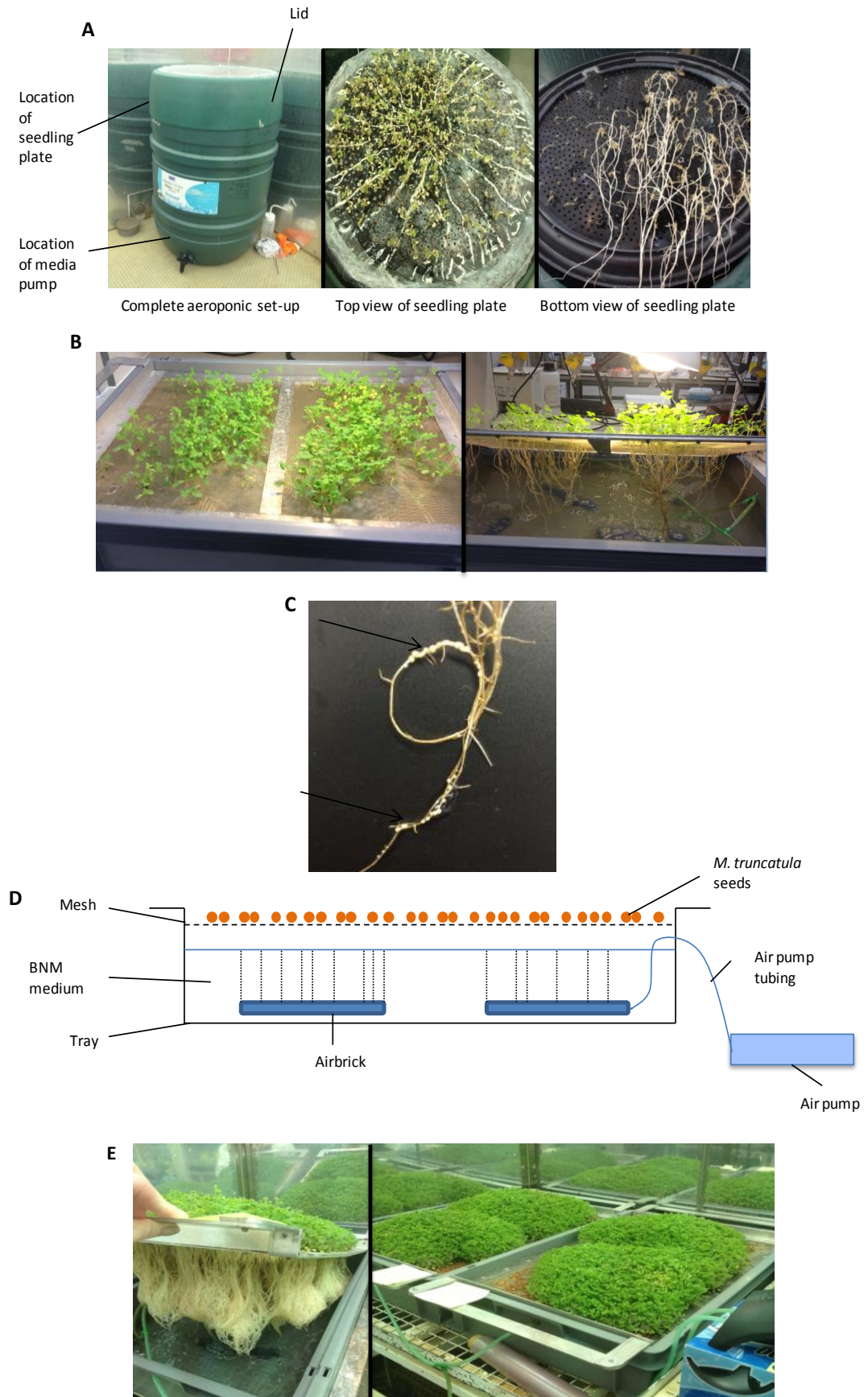
Figure 5.1. Strategy for maximising root growth in *M. truncatula* for protein purification.

Chapter 5: CCaMK *in planta*

(A) Trial of *M. truncatula* A17 growth in an aeroponic caisson chamber.
(B) Small-scale trial of a hydroponic growth method, planted with approximately 1000 *M. truncatula* A17 seeds.
(C) A representative A17 root from the trial hydroponic set up (**B**) showing fully-developed nodules.
(D) Schematic of the hydroponic growth system.
(E) Large-scale preparative growth of *M. truncatula* Jester in two hydroponic tanks planted with approximately 20 000 seeds in total.

5.3 Generation and testing of αMtCCaMK

In order to establish a purification method for MtCCaMK, it was most practical to generate an antibody to assess different purification strategies. After an extensive search for a suitable peptide antigen for immunization, it was decided that the sequence similarity of target peptides was too high to many other *M. truncatula* proteins. This was based on the sequence similarity of more than 4 consecutive amino acids between the peptide and other host proteins. Therefore, a protein antigen strategy was used instead. Previous work in the Bornemann lab with CCaMK *in vitro* constructs has determined that full-length, tagless CCaMK was unstable and difficult to purify in quantities suitable for antibody production. Specifically the kinase domain was found to cause solubility issues (Swainsbury *et al.*, 2012). Therefore, two existing constructs were considered: AI-VLD and His-VLD. The AI-VLD was selected since using a larger section of the protein would increase coverage of the target protein and potentially improve specificity (especially since EF-hands have a consensus sequence common to many proteins) and the lack of a tag would reduce artefactual cross-reactivity.

In order to isolate protein of sufficient purity, a two-step purification was used. Affinity chromatography with a CaM-Sepharose column (Figure 5.2A) was utilised and followed by a SEC step to remove any remaining contaminating proteins from the sample (Figure 5.2B). The size and purity of the protein was confirmed by SDS-PAGE (Figure 5.2C). Since the yield of protein from the CaM-Sepharose column was comparatively low, this process was repeated twice, and the resulting protein was pooled and dialysed in the buffer required for immunization. The protein was diluted to the appropriate concentration (260 µg/ml), and aliquots were sent for a one-month immunization programme in two rabbit hosts (SY05 and SY06).

Sera from both rabbits were tested at various concentrations to test for the affinity and specificity of the polyclonal antibodies to CCaMK (Figure 5.2D). Initial trials established that serum dilutions of 1:500 and 1:1000 were too low (data not shown), therefore dilutions of 1:2000 and 1:10 000 were used. Both sera were tested with three constructs of CCaMK: MBP-CCaMK (full-length), AI-VLD and His-VLD and all forms were recognised by both sera. At the 1:2000 dilution, both sera showed some band smearing, which was not present at the 1:10 000 dilution therefore the higher dilution was selected. At this dilution, the MBP-CCaMK lane contained two main bands, the higher of which is MBP-CCaMK and the lower of which is presumably a degradation product. His-VLD shows one strong and one very weak band, therefore this is also a specific interaction. AI-VLD shows a strong band for the correct protein, and also some cross-reactivity. However, it is likely that these were *E. coli*-derived contaminants that were present at very low concentrations in the AI-VLD samples administered to the rabbits. Such contaminants would therefore not be present in plant tissue. SY06 serum was selected for future use due to it giving slightly less band smearing, particularly with the AI-VLD protein. There were no bands detected in the pre-immune sera when tested against the same three proteins consistent with the antibodies being raised against AI-VLD. Therefore, SY06 serum was deemed to recognise CCaMK strongly and specifically and no further purification of the serum was necessary. This serum was used throughout the rest of this study and is henceforth referred to as αMtCCaMK.
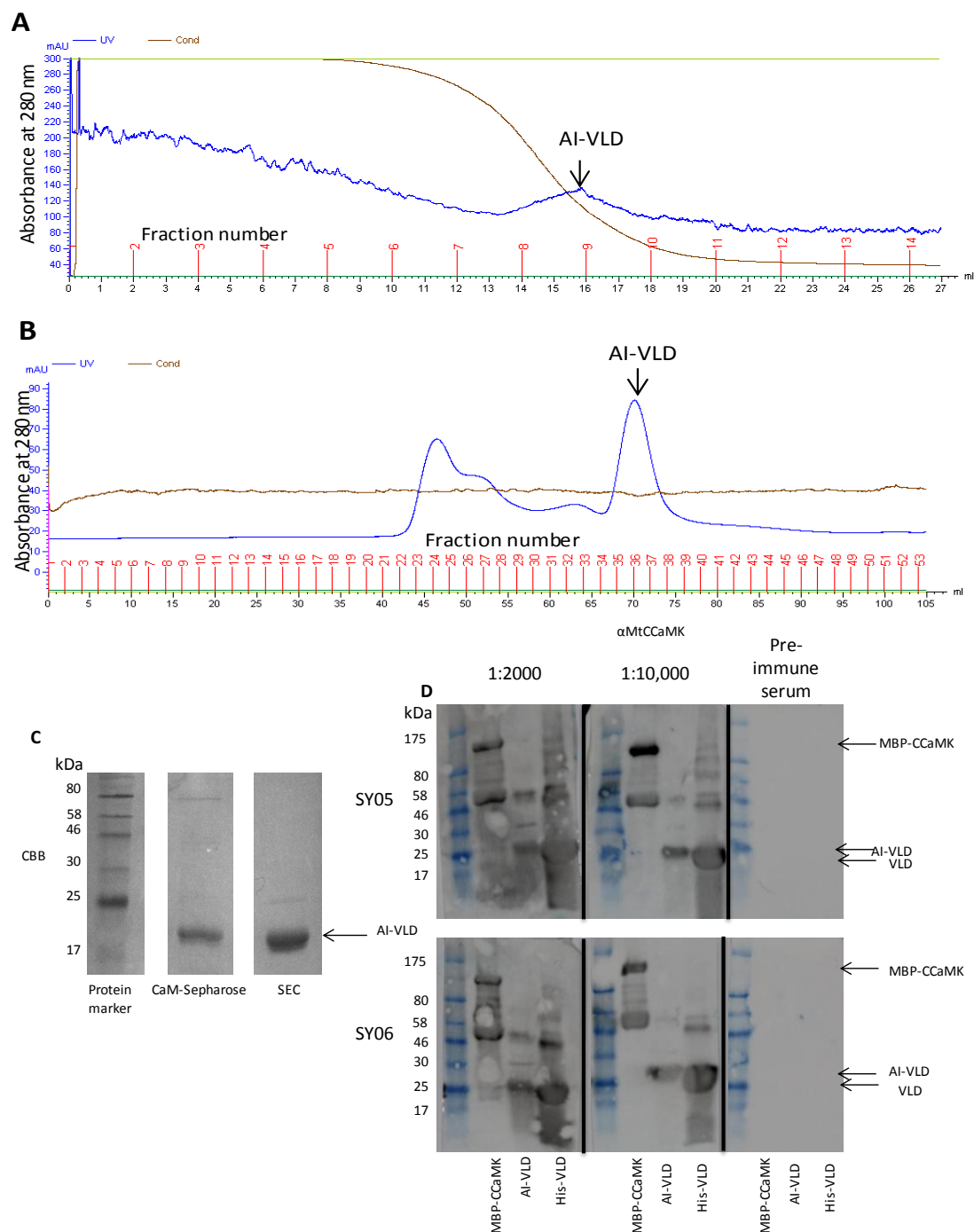
Figure 5.2. αMtCCaMK antibody generation and testing.

(A) Chromatogram for AI-VLD purified using the CaM-Sepharose column. Fractions corresponding to the marked peak were collected, pooled and subjected to size exclusion chromatography as shown in **(B)**.

(B) A size exclusion chromatogram for the AI-VLD using a HiLoad Superdex 75 prep grade 16/60 column. The marked peak was confirmed to contain the correct protein by SDS-PAGE shown in **(C)**.

(C) SDS-PAGE of CaM-Sepharose and SEC fractions of AI-VLD. The observed size of AI-VLD was 24 kDa as predicted. Appropriate fractions from SEC were pooled, diluted to 260 μg/ml and sent for antibody production in 500 μl aliquots.

(D) Western blots to test appropriate concentrations of sera (using purified AI-VLD antigen) derived from rabbits SY05 and SY06. The observed sizes of MBP-CCaMK, AI-VLD and VLD recombinant proteins were 98, 24 and 21 kDa, respectively, as predicted. SY06 at 1:10,000 dilution was selected for future use due to its strong affinity for CCaMK in various forms and its lower cross-reactivity.

5.4 Development of a purification strategy for Jester MtCCaMK

In order to develop a purification strategy for Jester MtCCaMK, a buffer trial was first undertaken. Briefly, root tissue was ground in liquid nitrogen, re-suspended in buffer, passed through a cell disruptor at 20 psi, incubated on ice for extraction and centrifuged. The resulting supernatant was blotted for comparison of two buffers (RIPA and Li) (Figure 5.3A). Extraction was superior with the Li buffer, especially regarding a 58 kDa protein (marked with an arrow) which corresponds to full-length CCaMK, so therefore this buffer was selected for further purification experiments. Also notable from this extraction is a strong band recognised by αMtCCaMK at just below 48 kDa (henceforth denoted as the 48 kDa protein).

The purification strategy subsequently trialled with Jester extract involved ion-exchange chromatography to remove many contaminants and CaM-Sepharose affinity chromatography to purify CCaMK further. Firstly, two columns were trialled for ion-exchange: MonoQ and DEAE Sepharose. The fractions corresponding to eluted protein were subjected to western blotting and probed with αMtCCaMK (Figure 5.3B and C, respectively). In both cases, two main proteins were detected by the antibody at 96 and 48 kDa. DEAE purification was selected for future experiments because it proved more robust and high-throughput than the MonoQ purification. Many attempts were made to increase the yield of these proteins, especially the 48 kDa protein, to give enough protein for mass spectrometric analysis. These attempts included the use of a ball mill, pulverisette machine and blender to extract more protein. This also involved alteration of the pressure and the number of extractions with the cell disruptor. Despite this effort, the yield of the 48 kDa protein was never high enough to see an equivalent band on an SDS gel and therefore a mass spectrometric analysis of this protein was not possible.

Since the 48 kDa protein was recognised strongly by the αMtCCaMK antibody, it was either a form of CCaMK that was smaller than expected (e.g. due to alternative splicing or proteolysis) or a protein to which the antibody cross-reacts. In order to explore this further,

crude Jester extract from 20 g of starting root material was applied directly to a CaM-Sepharose column. The resulting eluted protein fractions (Figure 5.3D) and flow-through wash fractions (Figure 5.3E) were subjected to western blotting probed with αMtCCaMK. Six main proteins that were eluted from the column were detected, including a 48 kDa protein and potential full-length 58 kDa CCaMK. Again, there was not enough protein purified by this method for subsequent experiments. A 48 kDa protein was also detected in the wash fractions which could either indicate that it doesn't bind to CaM-Sepharose or the capacity of the column was exceeded.

In an attempt to purify the 48 kDa protein on a preparative scale, a large scale extraction of Jester was attempted with 100 g of starting root material (Figure 5.3F). The 48 kDa protein was first purified in batches using the DEAE step, and this protein was collected, dialysed, and loaded onto a CaM-Sepharose column. Surprisingly, this protein did not bind to this column and was therefore not purifiable using this method. Therefore, it appears that this protein is either not CCaMK-derived or is a splice-variant of CCaMK that lacks the CaM-binding domain. Furthermore, since no full-length protein was purified by these methods, a new approach was needed to both clarify if the 48 kDa protein corresponds to a variant of CCaMK and if it is possible to purify full-length CCaMK from Jester.
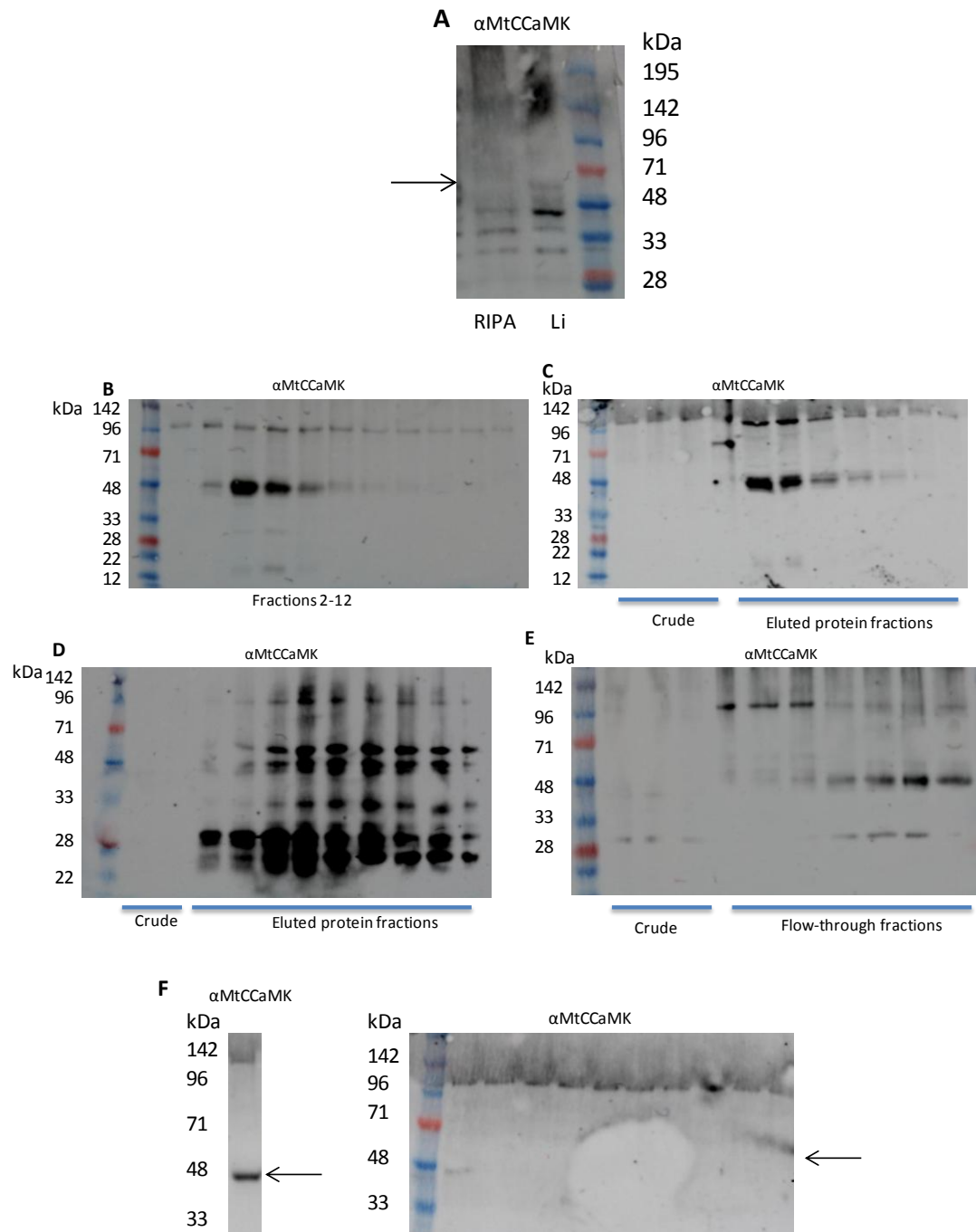
Figure 5.3. Purification trials for *M. truncatula* Jester CCaMK show predominantly a smaller protein (approx. 48 kDa) according to western blots probed with αMtCCaMK.
(A) Blot showing a comparison of RIPA and Li extraction buffers for extraction of CCaMK from *M. truncatula* Jester. Briefly, tissue was ground in liquid nitrogen, subjected to cell disruption at 20,000 psi and centrifuged. Li buffer was selected for further use due to superior extraction, particularly of full-length CCaMK protein (58 kDa).
(B) Blot showing eluted protein fractions from MonoQ ion exchange chromatography with *M. truncatula* Jester extract. Two main proteins were detected using this method at approximately 96 and 48 kDa.
(C) Blot showing eluted protein fractions from ion exchange chromatography using a DEAE column with *M. truncatula* Jester extract. Two main proteins were detected using this

method at approximately 96 and 48 kDa. Lanes 1-4 correspond to crude Jester extract at stages during the extraction process.

(D) Blot showing eluted protein fractions from affinity chromatography using a CaM-Sepharose column with *M. truncatula* Jester extract. Lanes 1-2 correspond to crude Jester extract prior to CaM-Sepharose purification.

(E) Blot showing fractions from the CaM-Sepharose column wash step (prior to elution) from the same affinity purification as blot **(D)**. Lanes 1-3 correspond to crude Jester extract prior to CaM-Sepharose purification.

(F) Eluted protein from DEAE purification (left panel) was purified by CaM-Sepharose affinity purification and the protein fractions were blotted (right panel). The 48 kDa protein which was detected after the DEAE purification was not detected in fractions corresponding to the protein elution peak after CaM-Sepharose purification and subsequent western blot analysis.

5.5 Characterisation of a 48 kDa protein and 58 kDa MtCCaMK from Jester

In order to determine whether αMtCCaMK also recognised the 48 kDa protein in a native environment, immunoprecipitation experiments were carried out. Firstly, crude Jester extract was incubated with αMtCCaMK-cross-linked beads for 5 hours, and the resulting protein was blotted and probed with αMtCCaMK (Figure 5.4A). Surprisingly, although no protein bands were visible in the input extract because the concentration was presumably too low, a single band was visible in the eluted output protein. This protein corresponds to full-length CCaMK at 58 kDa and no 48 kDa protein was detected in the eluted protein. Therefore either the 48 kDa did not bind to αMtCCaMK during the immunoprecipitation, or the concentration of the 48 kDa protein was too low in crude extract to be successfully enriched by this method.

In order to determine if the 48 kDa was able to be detected when present at a higher concentration in the IP input, the appropriate fraction from the DEAE purification of Jester extract (Figure 5.3C) was subjected to immunoprecipitation with the αMtCCaMK beads (Figure 5.4B). In both the input and flow-through fractions, the 48 and 96 kDa proteins were detected with αMtCCaMK, as expected. However, the eluted output protein showed a single band at the full-length size of 58 kDa, with no sign of the 48 kDa protein. It therefore appeared that the 48 kDa protein did not bind to αMtCCaMK in a native environment and therefore may not be CCaMK-derived.

Finally, in order to establish if the 48 kDa protein functioned as a kinase and therefore was CCaMK-derived, an aliquot of the appropriate DEAE fraction (Figure 5.3C) was also

subjected to kinase and phosphatase assays. These samples were then analysed by western blot with both αMtCCaMK (Figure 5.4C) and αpT271 (Figure 5.4D). One fraction from the DEAE purification was selected, concentrated and subjected to either a kinase assay, phosphatase assay, or a kinase assay followed by a phosphatase treatment. When probed with αMtCCaMK, the 48 and 96 kDa proteins were detected, as expected. Surprisingly, when probed with αpT271, the predominant protein detected corresponded to full-length 58 kDa CCaMK. It was unclear whether this protein behaves as a kinase because when subjected to a kinase assay; the strength of the band detected appears weaker when compared to untreated protein. This indicated a surprising reduction in phosphorylation instead of the expected increase. When compared to the αMtCCaMK blot of the same samples, it can be determined that the protein concentrations present in the first two lanes were approximately equal, so this result was not a consequence of protein loading. It would therefore appear that the kinase assay was unsuccessful. However it has been previously established that CCaMK function can be disrupted by freeze-thawing and this protein sample had been freeze-thawed before use in the assay. It therefore seems likely that, despite not behaving as expected for a kinase, the 58 kDa protein detected by αpT271 likely corresponded to a full-length form of CCaMK.

Taken together, these data suggest the 58 kDa full-length CCaMK protein was purifiable (e.g. with DEAE chromatography) but the concentration of the protein was too low to detect by western blotting without concentration by IP. These data also suggest that the 48 kDa protein is in fact not CCaMK-derived, because it is not pulled down by IP and is not the predominant form detected by the αpT271 antibody. Some cross-reactivity with other *M. truncatula* proteins is seen with the αpT271 antibody, however this is likely because the antibody was raised against a peptide which may have similarity to other proteins.

It seems likely therefore that the epitope recognised by αMtCCaMK in the 48 kDa protein is internal to the protein and is therefore only recognised in denaturing conditions (i.e. SDS PAGE and western blotting) but not recognised by the antibody in conditions where the protein is in the native state (i.e. during IP). A mass spectrometric analysis was therefore attempted of the 48 kDa protein, by the excision of the appropriate section of the SDS-PAGE of the DEAE fraction (despite there being no visible bands with coomassie staining of the gel). This revealed a large number of *M. truncatula* proteins, not including CCaMK. The proteins that also showed some sequence similarity to CCaMK in a BLAST analysis, and

could therefore have been recognised by the antibody, were: phosphopyruvate hydratase (GenBank: AES75861.2) and pyridine nucleotide-disulfide oxido-reductase (GenBank: KEH42067.1). In addition, these proteins were approximately 48 kDa and therefore were good candidates. Nevertheless, there were also many other proteins identified by the mass spectrometric analysis which the antibody could have cross-reacted with.

It was also therefore clear that these purification methods for full-length MtCCaMK from Jester yielded too little protein for biochemical characterisation.
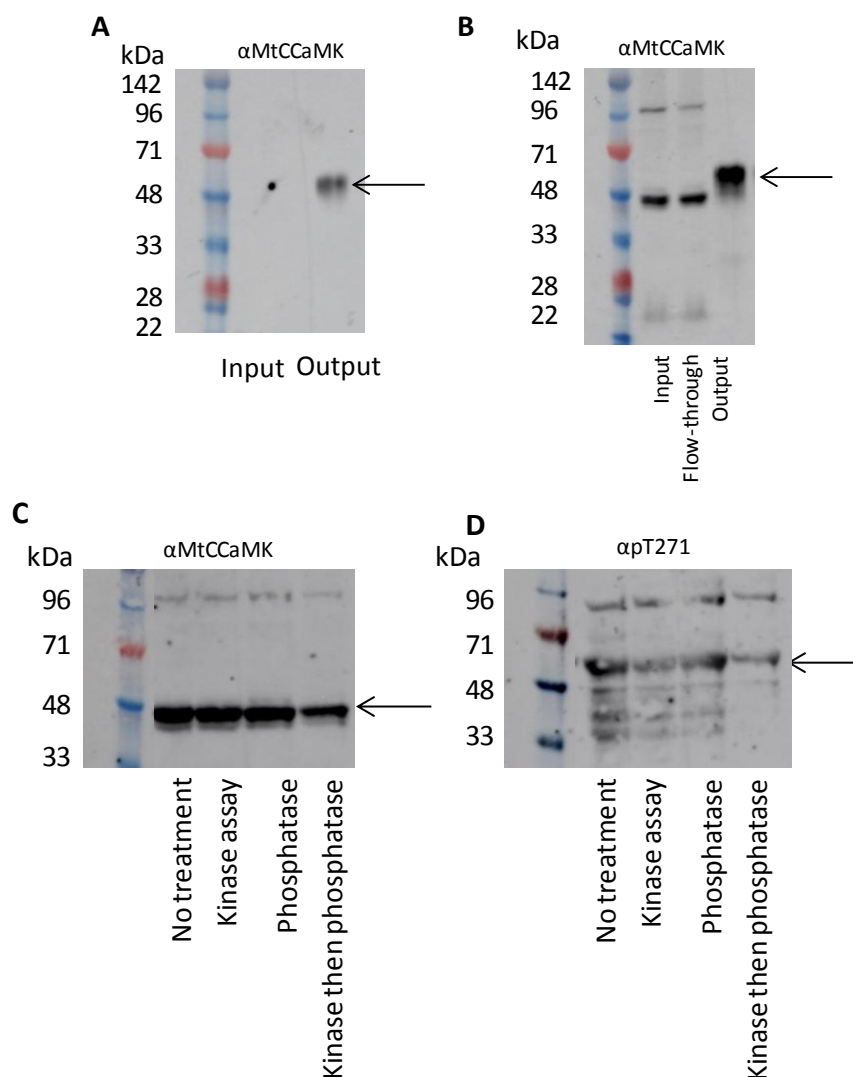
Figure 5.4. Immunoprecipitation and kinase assay of Jester extracts reveal full-length MtCCaMK protein (58 kDa).
(A) A western blot with αMtCCaMK of immunoprecipitate generated from *M. truncatula* Jester root extract and magnetic beads cross-linked with αMtCCaMK gave a single protein band of approximately 58 kDa, which corresponded to full-length CCaMK protein.

(B) A western blot with αMtCCaMK of immunoprecipitate from a 48 kDa-containing fraction of *M. truncatula* Jester protein purified using the DEAE column (**Figure 3(C)**). Input and flow-through fractions from the IP confirmed the presence of the 48 kDa protein band, whereas the IP output sample showed a single protein band of approximately 58 kDa indicative of full-length CCaMK.

(C) A fraction of Jester protein purified using a DEAE column (**Figure 5.3C**) was subjected to a kinase assay followed by λ-phosphatase-treatment to assess whether the proteins function as kinases. The resulting fractions were assessed with a western blot probed with αMtCCaMK to verify relative protein concentrations. Lanes contained protein after either no treatment, protein subjected to a kinase assay with 0.2 mM Ca^{2+} and 10 mM Mg^{2+}, λ-phosphatase treated protein or finally protein subjected to a kinase assay followed by a λ-phosphatase treatment. The predominant protein detected was approximately 48 kDa in all cases.

(D) The same samples as in **(C)** were subjected to western blot with a primary antibody αpT271**.** In contrast to **(C)** the predominant protein detected was approximately 58 kDa which corresponds to full-length CCaMK protein. However, the 58 kDa does not show evidence of autophosphorylation or phosphate removal because the protein fraction had been freeze-thawed prior to analysis.

5.6 Generation and testing of an over-expression line

Since levels of MtCCaMK in *M. truncatula* Jester were extremely low, an alternative source of protein for biochemical characterisation of CCaMK was sought. Therefore the development of an over-expression line was attempted (Figure 5.5). A binary-vector was generated using Golden Gate cloning (Patron *et al.*, 2015) for *A. tumefaciens*-mediated stable transformation. This vector contained MtCCaMK cDNA with an *L. japonicus* ubiquitin over-expression promoter, and a dsRed reporter to check for successful transformation (Figure 5.5A). This construct was tested for functional complementation by hairy root transformation of *dmi3-1* seedlings and successful transformation was confirmed by both dsRed fluorescence (Figure 5.5B) and a nodulation complementation assay.

The construct was then used for stable transformation of R108 *M. truncatula* by Matthew Smoker. This line was selected because of the ease of transformation and the high number of potential transformants which are usually generated. Transformation was carried out according to standard procedures, however the incorrect plant selection marker was used due to a miscommunication (kanamycin instead of BASTA) (Figure 5.5C). Since kanamycin was the bacterial selection marker (and was not contained within the right and left border) and BASTA was the plant selection marker, this kanamycin selection should not have been functional for plant transformation. Despite this, approximately 30 potential transformants were generated after approximately 6 months and were therefore tested for the presence

of the correct transgene in case the marker was also functional for plant selection due to a leaky left border transformation.

PCR was carried out with *DMI3* primers on DNA extracted from potential transformants, and an extension time of two minutes was used to amplify any 1.5 kbp transgene if present (Figure 5.5D). This would distinguish between the transgene and the genomic DNA because the band amplified from gDNA would contain introns giving a 4 kbp amplification product. Four samples showed amplification of the appropriately sized band (marked with arrows) and so were potentially transgenic. However, this PCR experiment was repeated twice more with these samples and the bands were not reproduced. In addition, a TaqMan-based copy number analysis for the BASTA gene was performed by iDNA genetics and no BASTA gene was detected in any of the samples. If transformation had been successful despite use of the incorrect marker, the BASTA cassette should still have been present in the inserted DNA. Therefore the plants were deemed most likely non-transgenic and work was focussed on other strategies.
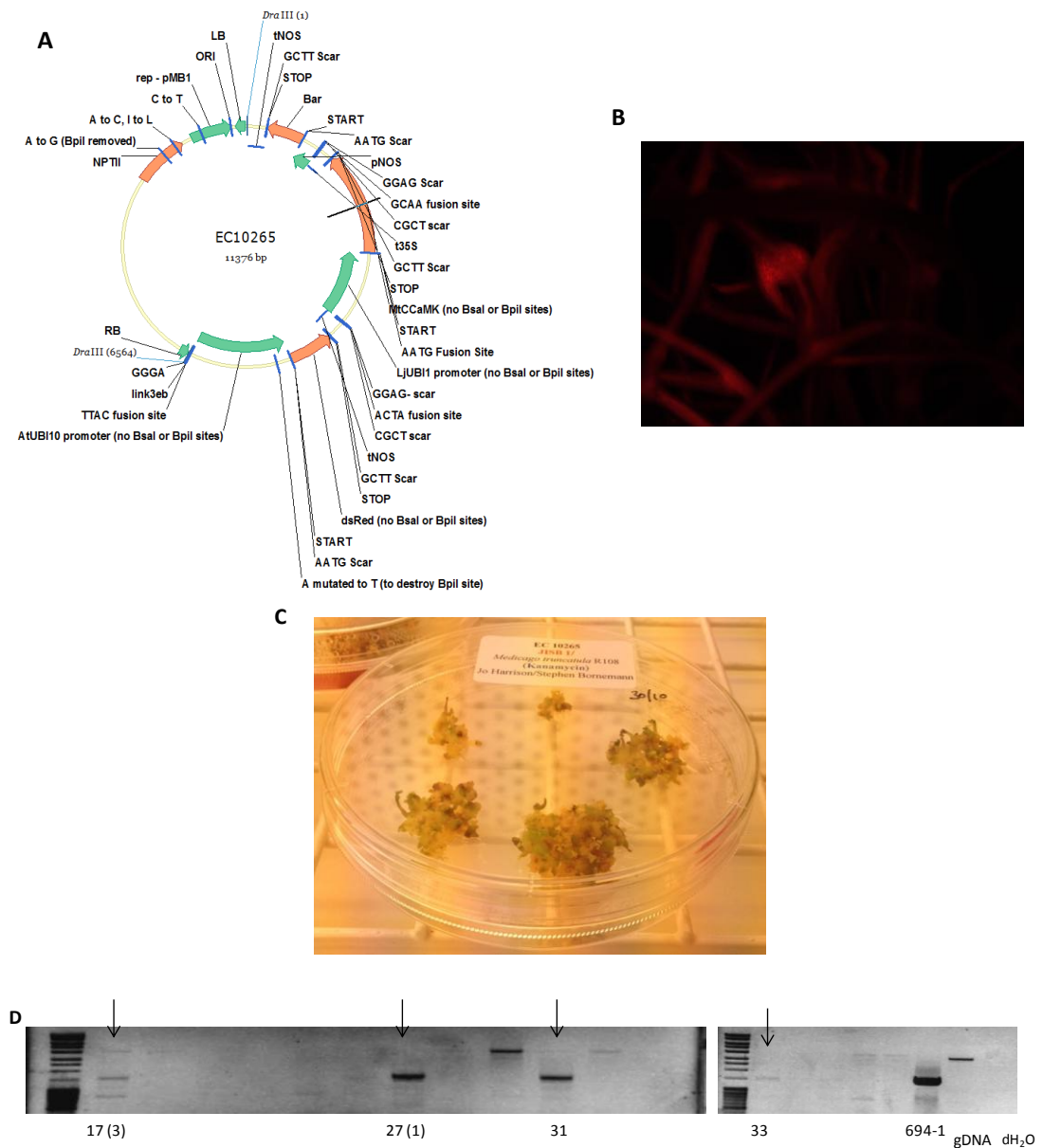
Figure 5.5. Attempted development of stable *M. truncatula* CCaMK over-expression line.
(A) Map for the binary vector (EC10265) which was used to transform *M. truncatula* R108 for over-expression of CCaMK. Key features included a dsRED reporter gene, *L. japonicus* ubiquitin over-expression promoter and MtCCaMK coding sequence (lacking introns).
(B) Image showing dsRed fluorescence of a representative *dmi3-1* root transformed with EC10265 by hairy root transformation.
(C) Potentially transgenic calli (transformed with EC10265) with developing embryos grown on kanamycin medium. Transformation was carried out by Matthew Smoker (TSL).
(D) PCR analysis, using *DMI3* primers, of a representative selection of potential transformants. The arrows mark four potentially transgenic plant lines with bands of the correct size for *DMI3* cDNA (approx. 1.5 kbp), however subsequent PCR experiments were unable to replicate these bands. The last three lanes contain control PCR samples including a confirmed transgenic line (GFP-CCaMK-SBP 694/1), purified genomic DNA (gDNA) and a negative water control.

5.7 Testing of GFP-CCaMK-SBP stable line and subsequent purification attempts

Since protein levels are extremely low in wt *M. truncatula* lines, alternative transgenic lines were explored. A stable line of *M. truncatula* had previously been produced in the Oldroyd lab although its successful transformation had not yet been confirmed. This line was generated in a *dmi3-1* mutant background with a GFP-CCaMK-SBP construct (under the control of the native pDMI3 promoter). GFP was present at the N-terminus and SBP was a streptavidin-binding peptide tag present at the C-terminus. Seeds from the T_1 generation were tested and 16 lines were selected on the basis of successful complementation of nodulation. They were also subjected to western blot analysis and probed with either αGFP or αMtCCaMK to identify lines with a good level of CCaMK protein. In particular, line 694 showed good levels of the correct fusion protein (Figure 5.6A) with an observed size of 87 kDa, as expected. A T_2 progeny (denoted as 694/1) of this line was also confirmed to contain a single copy of the insertion by a TaqMan analysis carried out by iDNA genetics, which was sufficient for protein studies. This line was therefore selected for use in further experiments.

Firstly, the previous approach of DEAE chromatography was trialled to see if it was effective for GFP-CCaMK-SBP since the predicted pI was very close to that for CCaMK (5.64 and 5.57, respectively). The chromatogram for this experiment shows an eluted protein peak which is marked by a blue line (Figure 5.6B) which was tested by western blotting with αMtCCaMK (Figure 5.6C). Surprisingly, the eluted protein fractions did not contain the expected protein, however the protein was clearly visible in the crude extract used for this purification. In addition, no fusion protein was detected in the wash fractions for this purification when subjected to western blot with with αMtCCaMK (data not shown). Therefore it appeared that the 1 M NaCl elution step was insufficient to elute the fusion protein from the column, since the protein was clearly visible in the crude extract and the purification step would be a concentrating step. Due to the low availability of plant material for this line, other methods of purification were trialled instead of attempted optimisations for this purification method.

Secondly, a Streptavidin Hitrap column was used to purify GFP-CCaMK-SBP based on the affinity of the streptavidin-binding peptide for streptavidin and utilising the biotin analogue, desthiobiotin, for elution. The chromatogram shows a potential protein peak

marked by a blue line (Figure 5.6D) which was tested by western blotting with αMtCCaMK. Despite the presence of a protein peak, again no GFP-CCaMK-SBP was detected with αMtCCaMK (data not shown) and therefore it seems likely that the concentration of 10 mM desthiobiotin was not sufficient to elute the protein. An extraction was also trialled with harvested aerial plant material, since the transgene will be expressed throughout the plant. However, a crude extract blotted and probed with αMtCCaMK did not reveal an appropriate band, most likely due to low concentrations, and therefore this material was not used for further experiments.

A final attempt was made with the all the remaining root tissue for this line (approximately 10 g). This particular plant line did not grow as well as Jester so despite growing ~4200 seeds in a hydroponic system, a total yield of 14 g of root tissue was harvested and used for all trials. A CaM-Sepharose affinity column was used and the predicted protein containing fractions (Figure 5.6E) were pooled and applied directly to a pre-equilibrated size exclusion column (Figure 5.6F) to estimate the size of the oligomer. Despite the presence of a protein peak, a subsequent western blot probed with αMtCCaMK revealed no GFP-CCaMK-SBP protein was purified after using both columns. This is most likely because there was insufficient material for the lower affinity column, as opposed to tag-based purification systems. Therefore, the successful purification of this CCaMK fusion protein required further optimisation which was not possible due to limited availability of material.
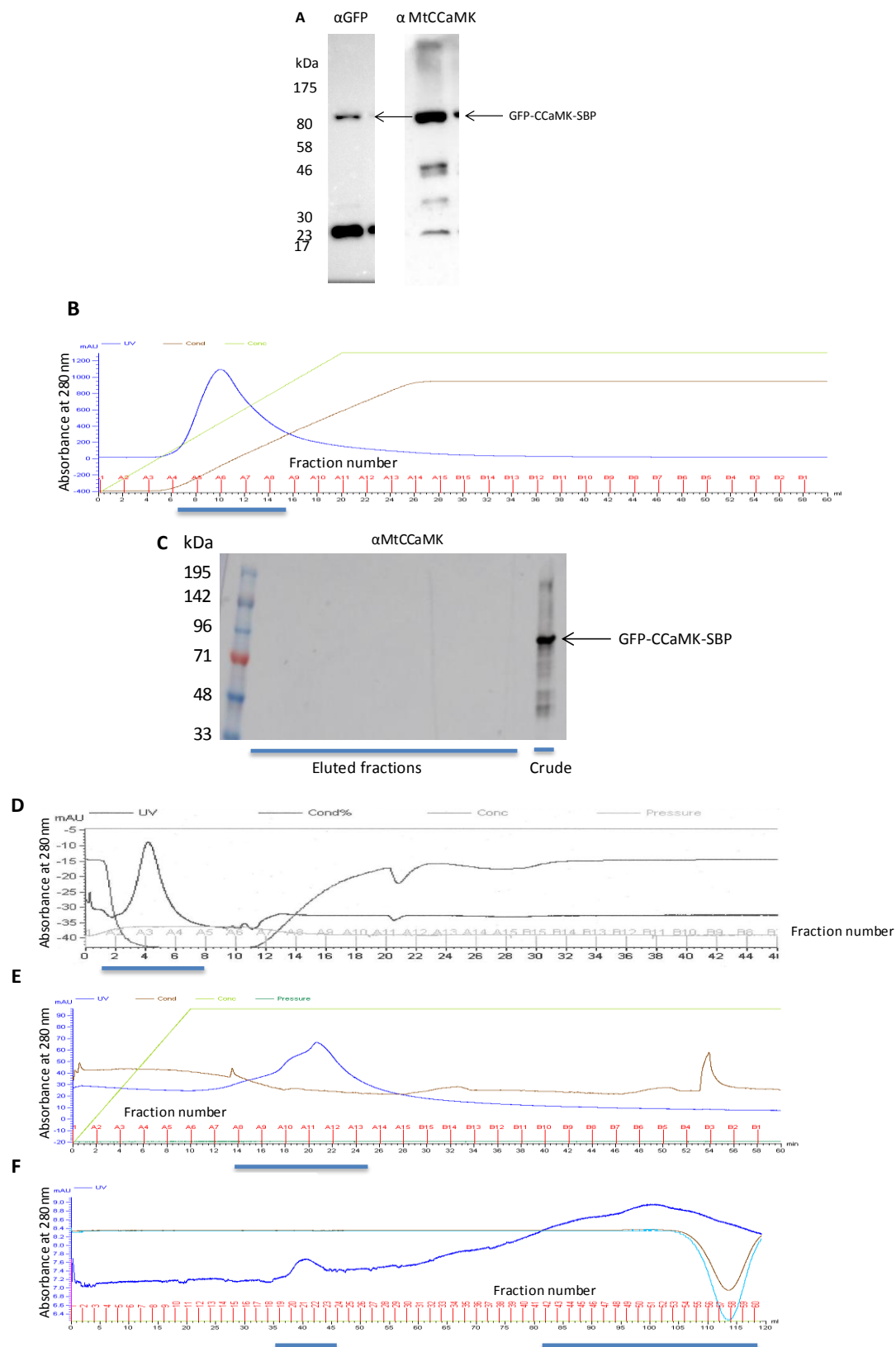
Figure 5.6. Purification attempts of GFP-CCaMK-SBP from stably transformed *M. truncatula* line 694/1.
(A) Crude protein extracts of GFP-CCaMK-SBP line 694/1 were analysed by western blots probed with αGFP and αMtCCaMK. The target protein of GFP-CCaMK-SBP is predicted to be

Chapter 5: CCaMK *in planta*

87 kDa and was detected at the appropriate size (see marked band). This figure is courtesy of Sylvia Singh.

(B) Chromatogram for the attempted purification of GFP-CCaMK-SBP by ion exchange chromatography with a DEAE column performed at 4 °C. The potential protein-containing fractions which eluted from the column are marked on the chromatogram with a blue bar.

(C) Eluted protein fractions from the attempted purification of GFP-CCaMK-SBP with a DEAE column **(B)** were subjected to a western blot and probed with αMtCCaMK. The fractions eluted from the column were present in lanes 1-10 while lane 11 contained crude extract. The target protein of 87 kDa is marked on the blot in the lane which contained crude extract.

(D) Chromatogram for the attempted purification of GFP-CCaMK-SBP by affinity chromatography using a streptavidin Hitrap column performed at 4 °C. The potential protein-containing eluted fractions are marked with a blue bar.

(E) Chromatogram for the attempted purification of GFP-CCaMK-SBP by affinity chromatography using a CaM-Sepharose column performed at 4 °C. The potential protein-containing eluted fractions are marked with a blue bar.

(F) The predicted protein-containing fractions from the CaM-Sepharose purification **(E)** were pooled, and subjected to size-exclusion chromatography with a Sephacryl S300 HR 16/60 column. The resulting chromatogram is shown and the potential protein-containing fractions are marked with blue bars.

5.8 Discussion

An investigation was undertaken to determine to what extent it is possible to purify CCaMK from *M. truncatula*. First, a purification strategy was established to generate significant quantities of root material for protein extraction. A hydroponic growth strategy was selected and was advantageous for several reasons including a high yield of root tissue, minimal cleaning of tissue so it was possible to freeze it quickly with low levels of root contamination. Therefore, a successful growth strategy was developed during this project that is optimal for protein purification from *M. truncatula*.

Second, an antibody was generated against CCaMK using AI-VLD protein expressed in *E. coli*. Testing of this antibody against *E. coli*-expressed constructs of CCaMK (MBP-CCaMK, AI-VLD and His-VLD) showed that it recognised both full-length and truncated forms of the protein strongly and specifically. Subsequent experiments with plant protein extracts demonstrated that the antibody also showed some cross-reactivity. However, use of the protein for IP experiments demonstrated that the antibody does specifically recognise MtCCaMK and was able to enrich the protein during IP so that it is within the detection limits of αMtCCaMK. Importantly, the IP experiments demonstrated that these cross-reactive proteins (including the oft-purified 48 kDa protein) could be distinguished from MtCCaMK (or MtCCaMK-derived proteins). This is because they were recognised by the antibody under denaturing conditions (e.g. SDS-PAGE) and not in a native environment (i.e. IP), and therefore the epitopes recognised by the antibody in these proteins were likely internal to the proteins. Overall, this demonstrates that αMtCCaMK is a useful tool for the detection of MtCCaMK in plant protein preparations.

Purification of MtCCaMK was therefore possible, to an extent, because a single band was observed on a western blot probed with αMtCCaMK after IP. However, it was not possible to purify sufficient protein for biochemical experiments including oligomer-size determination and structural work. In an attempt to obtain sufficient protein yields, many different extraction procedures, buffers and protein purification methods were trialled, the most important of which are detailed in this Chapter. Extraction procedures which were trialled included the use of different tissue extraction machines namely a ball mill, liquid-nitrogen cooled pulverisette, cell disruptor and blender. Only the cell disruptor showed improved extraction (data not shown) and therefore this was utilised in subsequent experiments. As detailed in this Chapter, both one-step and two-step affinity column

purification methods were utilised in an attempt to obtain the correct protein, including MonoQ, DEAE-Sepharose and CaM-Sepharose chromatography. However, these did not yield a sufficient quantity of the correct protein. This was due to the extremely low concentration of CCaMK in root tissue. This was determined because the protein was only detectable *via* western blot after enrichment using IP. Furthermore, the use of over-expression or transgenic *M. truncatula* lines for the purification of CCaMK was also unsuccessful. The over-expression line was unsuccessful due to the use of the incorrect antibiotic resistance marker; however the use of the GFP-CCaMK-SBP line for CCaMK purification was again unsuccessful due to low protein levels. Therefore, extensive trials were inadequate in this case to yield sufficient protein for biochemical work.

Despite the low protein yield of CCaMK discussed above, the work described in this Chapter was necessary and important for several reasons. First, it was possible to establish that the 48 kDa protein which was predominantly purified did not correspond to either a splice variant or degradation product of CCaMK. Furthermore, it was established that CCaMK was present *in planta* at the predicted and expected size of 58 kDa. Finally, procedures were developed during the course of these investigations which facilitated further investigations. These included the development of the αMtCCaMK and the corresponding western blot technique for this antibody. Furthermore, the IP procedure for αMtCCaMK was developed which allows cross-reactive proteins and MtCCaMK to be distinguished. These procedures were integral to the work described in Chapter 6.

5.9 Summary

Experimental work described in this Chapter detail the successful establishment of a large-scale growth system for *M. truncatula* and the development and testing of an antibody for CCaMK, αMtCCaMK. MtCCaMK was determined to be the expected size when purified from *M. truncatula* and was distinguishable by IP from cross-reactive proteins. It was not possible to purify sufficient MtCCaMK protein for biochemistry from either WT or transgenic *M. truncatula* lines. However, the techniques developed in this Chapter underpin some of the work carried out in Chapter 6.

Chapter 6: Alternative splicing of MtCCaMK

6.1 Introduction

Protein purification trials from *M. truncatula* Jester revealed a 48 kDa protein which the αMtCCaMK antibody strongly recognised and was detected when purified by several methods (Chapter 5). Whilst subsequent experiments demonstrated that this protein was not CCaMK-derived, the presence of multiple protein bands detected by western blot probed with αMtCCaMK raised the question of whether any of these proteins corresponded to alternative splice forms of CCaMK. Interestingly, preliminary data from Pierre-Marc Delaux indicate that splicing of *DMI3* at the RNA-level may occur (Figure 6.2). In addition, alternative splicing is thought to be common in plants with over 60% of intron-containing genes predicted to be alternatively spliced (Syed *et al.*, 2012). Therefore work was undertaken to ascertain whether CCaMK is alternatively spliced in *M. truncatula.*

An experiment was therefore devised to investigate whether alternative splicing of CCaMK occurs as part of its function in the nodulation and mycorrhization pathways. Previous experiments have shown that truncation of the protein can cause its auto-activation (Gleason *et al.*, 2006) and spontaneous nodulation in the absence of rhizobia, therefore alternative-splicing of CCaMK is potentially biologically significant. Furthermore, alternative splicing has been demonstrated to have profound effects on CaMKII and the interpretation of calcium spiking. Changes within a linker region of different splice-forms of the β subunit affect the activation constant for CaM-binding. This affects the response of the protein to the frequency of calcium oscillations (Bayer *et al.*, 2002). Splicing of CCaMK could therefore be extremely powerful, and could play a key role in the interpretation of the calcium spiking signal in the Sym pathway. Therefore, this Chapter details an experimental analysis of *DMI3* cDNA and CCaMK protein during nodulation and mycorrhization in order to determine whether splicing of CCaMK plays a role in the Sym pathway.

6.2 Exon structure of *DMI3*

First, the exon structure of the *DMI3* gene was determined (Figure 6.1). *DMI3* is made up of 7 exons, with a nucleotide sequence length of 1572 bp and a total predicted protein size of 58 kDa. Primers were designed for RT-PCR experiments to characterise potential splicing, with the forward primer at the beginning of exon 1 and the reverse primer at the end of exon 7. This method would detect the main forms of splicing that retain the first and last

exons. These include exon skipping, intron retention, and alternative 5′ and 3′ splice sites at the exon/intron borders (Syed *et al.*, 2012). Splice forms that lack the first exon would require the presence of alternate start sites, and similarly those lacking the final exon would require alternate polyadenylation sites which are difficult to predict due to the current under-development of bioinformatic tools for *M. truncatula*. In addition, exploration of these forms also dramatically increases the number of combinations that require testing. Therefore these experiments were carried out with primers spanning the full sequence to identify the main forms of splicing at the RNA level, if present. Furthermore, the use of a protein-level approach in addition would potentially identify all the different types of splice-forms discussed above, if present, and therefore this approach was comprehensive.
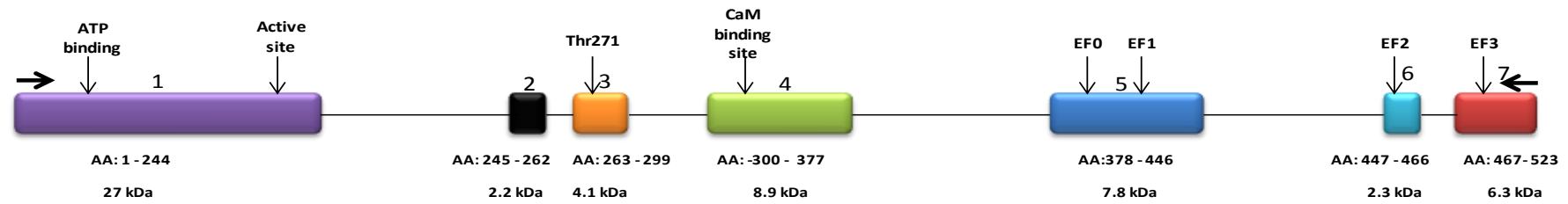
Figure 6.1. Exon structure of *M. truncatula DMI3*.

Exons 1-7 of the *DMI3* gene are shown with the corresponding amino acid numbers shown at the bottom together with the predicted sizes of each exon. Key protein features are marked at the appropriate locations at the top (where EF = EF-hands). The location of the primers used for subsequent RT-PCR experiments are marked by horizontal arrows on exon 1 and 7.

Chapter 6: Alternative splicing of MtCCaMK

6.3 Previous evidence of splicing

RT-PCR data provided by Pierre-Marc Delaux suggested that splicing of the *DMI3* gene may occur (Figure 6.2). An RT-PCR time-course analysis of *DMI3* during nodulation in A17 was carried out with samples taken at one, two and three wpi with rhizobia, alongside a mock inoculated sample at each time-point (Figure 6.2A). A total of five different splice-forms of *DMI3* were identified in addition to the full-length *DMI3* gene (1572 bp). Interestingly, the full-length form of *DMI3* was the least-abundant form detected at all time-points. The presence of different splice-forms also changed during the progression of nodulation. However, this was surprisingly also the case with the mock inoculated samples. The most predominant form detected at all time-points was 1173 bp and its expression altered over the three week experiment for both rhizobia- and mock-inoculated samples. Similarly the expression of a 993 bp form altered over time, and was detected only two wpi for both rhizobia and mock-inoculated samples. By contrast, a much smaller form of 498 bp only appears at a late stage of nodulation (3 wpi) and was not detected in any of the mock inoculated samples.

Three of the most strongly expressed forms were also sequenced to identify which exons and partial exons of CCaMK were present (Figure 6.2B). The 1173 bp form consisted of a portion of the kinase domain and the entire VLD. It could be theorised that this form, if reflected at the protein level *in planta,* would be an auto-active form of CCaMK since the Thr271 residue, which negatively regulates the protein, is missing. In contrast the 993 bp form, which lacks a large proportion of the kinase domain, would be inactive for target phosphorylation because it lacked the active site of the enzyme. Finally, the function of the 498 bp form, which contained a portion of the kinase domain and one EF-hand, is unclear since it lacks most of the regulatory elements of the protein and also the active site, so would therefore not encode a functional kinase. If reflected in the protein forms present the different forms of CCaMK detected in this experiment could perhaps be present to fine tune the progress of symbiosis signalling with a balance of an autoactive form and inactive forms.

The data presented in Figure 6.2 came from a single experiment and therefore required verification. Three important questions were therefore invoked by these data: are the splice-forms detected real or artefactual, are they reflected at the protein level *in planta*

and are there alternative splice-forms present during mycorrhization? This chapter addresses these three questions.
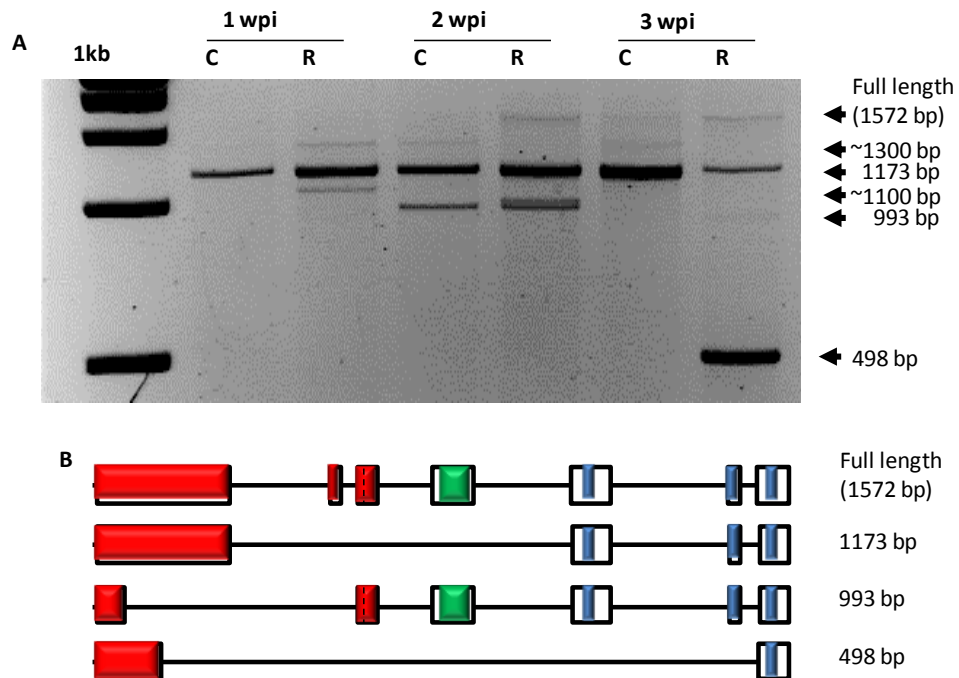
Figure 6.2. Characterisation of splice events in *DMI3* (Pierre-Marc Delaux and Jean-Michel Ané, personal communication).
(A) Time course of the expression of *DMI3* full length and splice variants several weeks post inoculation (wpi) with *S. meliloti* (R) or a mock solution (C).
(B) Schematic representation of the splice variants and the encoded protein domains present, as determined by sequencing. Red, green and blue squares indicate the kinase domain, the calmodulin binding domain and the EF-hands respectively. The vertical dashed lines in the kinase domain indicate the autophosphorylation site.

6.4 Experimental design

A time-course analysis was devised to investigate alternative splicing of CCaMK during nodulation and mycorrhization. Three lines were selected for analysis: Jester, A17 and GFP-CCaMK-SBP 694/1 (henceforth referred to as GFP). Jester and A17 were selected as wild-type lines and the GFP line was selected as a control line, since this was generated with the coding sequence of *DMI3* and therefore lacks introns which are necessary for splicing. Seedlings were grown on MOD FP agar for 10 days after stratification and replanted in terragreen/sand at Time 0. Seedlings were then inoculated with either rhizobia or mycorrhizal inocula at time 0, and an un-inoculated control for each line was harvested at this point.

Time-points were selected for both symbioses at an early, middle and later stage of the interaction (Figure 6.3A). For nodulation, the time-points used were: four days (some nodule primordia visible), 10 days (small white nodules visible) and 21 days post inoculation (dpi) (fully-developed nitrogen-fixing nodules) (Figure 6.3 B). The progression of mycorrhization was established by ink and vinegar staining of representative Jester roots at appropriate time-points post-inoculation, as show in Figure 6.3C. At three and four dpi with mycorrhizal inoculum, no penetration of infection hyphae was visible. At seven dpi, stained Jester roots had either zero, one or two infection units (Figure 6.3C, left panel) per plant (Table 6.1). At 13 dpi, roots were scored with the gridline intersect method, and colonisation was approximately 20% (data not shown). At 21 dpi, six root samples were scored by the same method and total colonisation was approximately 24-46% (Table 6.2), with fully-developed arbuscules (Figure 6.3C, right panel). Therefore, roots were harvested for the mycorrhizal time-course at 7, 13 and 21 dpi.

At each of the time-points, 20 roots were harvested, blotted and flash frozen in liquid nitrogen. They were then used for either RNA extractions and subsequent RT-PCR analysis or protein extractions and subsequent western blotting and IP with αMtCCaMK.
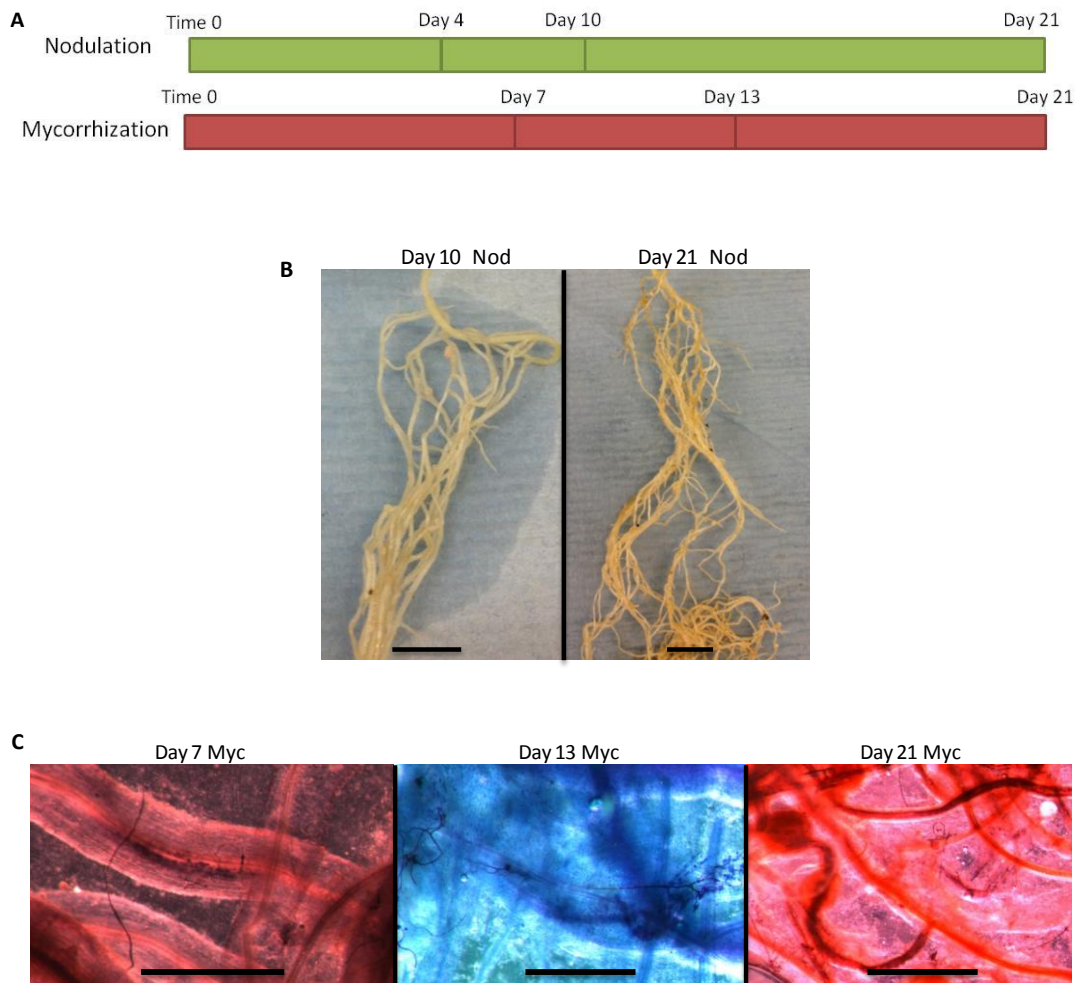
Figure 6.3. Experimental set up of time-course experiment for characterisation of CCaMK during nodulation and mycorrhization.

(A) A schematic representation of the time-points investigated for the characterisation of CCaMK during nodulation and mycorrhization. Ten roots were harvested for each line (A17, Jester and GFP-CCaMK-SBP 694/1) at each of the seven time-points. Time 0: 10 day-old un-inoculated roots, grown on Mod FP. Day 4, 10 and 21 nodulation: 4, 10 and 21 days post-inoculation with rhizobia, transferred as 10-day old seedlings on day 0 from Mod FP to terragreen sand. Day 7, 13 and 21 mycorrhization: 7, 13 and 21 days post-inoculation with mycorrhizal inoculum, transferred as 10-day old seedlings on day 0 from Mod FP.
(B) Images show representative Jester roots from 7 and 10 dpi with rhizobia, with approximate scale bars of 1 cm.
(C) Images show representative Jester roots from 7, 13 and 21 dpi with mycorrhizal inoculum, with approximate scale bars of 1 mm.

Plant number	Myc event per plant
1	1
2	2
3	0
4	1

Table 6.1. Total number of mycorrhization infection events (scored by visual microscope analysis of ink and vinegar stained roots) at 7 dpi with mycorrhizal inoculum, in four representative Jester root samples.

Plant number	Percentage colonisation (%)
1	38
2	46
3	24
4	27
5	27
6	42

Table 6.2. Percentage colonisation of six representative Jester root samples scored by the gridline intersect method (Giovannetti and Mosse, 1980), 21 dpi with mycorrhizal inoculum. Error is likely to be ±5-10%.

6.5 RNA quality verification

RNA was extracted from A17, Jester and GFP samples at each of the seven time-points for mycorrhization and nodulation (Figure 6.3A). RNA was extracted from approximately 100 mg of tissue, which consisted of tissue from approximately 2-10 roots (depending on the age and size of roots harvested) from one experimental batch of plants. RNA quality was verified by gel electrophoresis for each sample (Figure 6.4A-C). RNA quality was sufficient for all samples because two strong rRNA bands (25S and 18S) were present at each time-point for each line.

Chapter 6: Alternative splicing of MtCCaMK

RNA was treated with DNase 1 to remove any contaminating DNA. This was verified by PCR analysis using *DMI3* primers and a four minute extension time, in order to amplify genomic *DMI3* at ~4 kbp if present. There were no bands detected for any of the 21 samples (Figure 6.4D-F) and a positive control with gDNA showed the 4 kbp amplification product as expected. Therefore all the RNA was deemed of sufficient quality for use in subsequent RT-PCR experiments. cDNA synthesis was carried out with the Biorad iScript cDNA synthesis kit, which utilises a mixture of oligo(dT) and random primers.
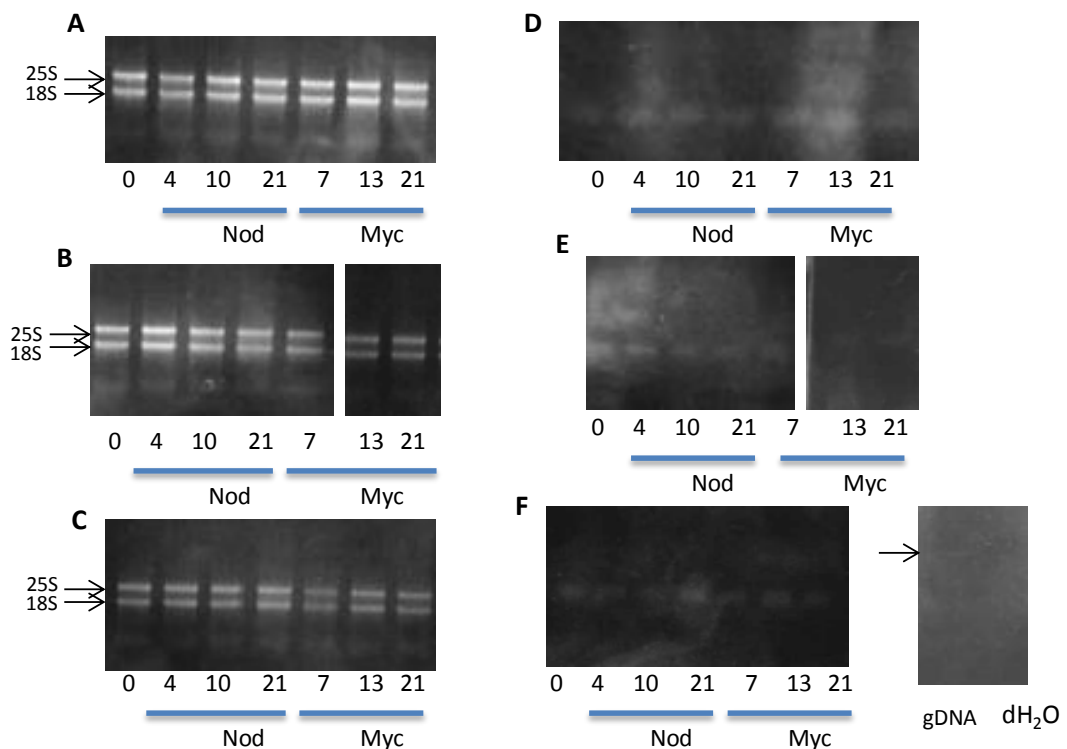
Figure 6.4. Quality verification for RNA extractions in time-course experiment.
(A) RNA from samples of Jester roots at 7 time-points (Figure 6.3A) was analysed by gel electrophoresis. The two key bands detected for all samples at all time-points were 25S and 18S rRNA and this indicates intact RNA extractions.
(B) RNA from samples of A17 roots were analysed, as in **(A)**.
(C) RNA from samples of GFP-CCaMK-SBP 694/1 roots were analysed, as in **(A)**.
(D) DNA contamination was analysed for all RNA samples of Jester, treated with DNase, by PCR using *DMI3* primers. No bands were detected at 4 kbp (genomic *DMI3*) and therefore no contaminating DNA was present.
(E) A17 RNA samples were analysed for DNA contamination, as in **(D)**.
(F) GFP-CCaMK-SBP 694/1 RNA samples were analysed for DNA contamination, as in **(D)**. The final two lanes contained PCR samples for a positive control genomic DNA sample (gDNA) in which the 4 kbp genomic *DMI3* was detected (marked with an arrow) and a negative control with water.

6.6 Full-length *DMI3* is the predominant form during nodulation and mycorrhization

An RT-PCR experiment was carried out for Jester, A17 and GFP cDNA from a time-course of nodulation and mycorrhization (Figure 6.3A).The cDNA synthesis was carried out using iScript cDNA synthesis kit and *DMI3* primers with a two minute PCR extension time were used in order to amplify the full-length *DMI3* gene (1.5 kbp) and any smaller splice-variants if present. At all time points for WT samples (Jester and A17) for nodulation and mycorrhization, the only band detected was 1.5 kbp corresponding to full-length *DMI3* (Figure 6.5A-B). This therefore showed no evidence of alternative splicing at these time-points in either A17 or Jester. This was also the case for the control line GFP which cannot be spliced such that only a full-length *DMI3* amplification product was detected at all time-points, as expected (Figure 6.5C). No band was detected in the genomic DNA lane because the extension time was too short (2 minute extension time, instead of 4 min required for the 4 kbp gene), and a water control was negative as expected. An alternative reverse transcriptase (Superscript III) with oligodT primers was also trialled with representative samples (A17 and Jester, 21 dpi with nodulation/mycorrhization) to determine if splice variants were detected when an alternative cDNA synthesis strategy was used. Under these conditions, full-length *DMI3* and no splice variants were detected (data not shown). Therefore, for all experiments presented, the iScript cDNA synthesis kit was used.

RT-PCR with *DMI3* primers was also carried out with an independent source of cDNA provided by Ben Miller to determine if alternative splice forms were present. cDNA corresponded to A17 roots at 0, 4 and 10 dpi with rhizobia and RT-PCR with these samples (and a two minute extension time) again showed full-length *DMI3* and no alternative splice-forms (Figure 6.5D). Therefore under these conditions no splicing of *DMI3* occurs.

In order to determine whether splicing occurred specifically during early stages of symbiosis signalling, an early time-course of nodulation was investigated. cDNA samples corresponding to Jester roots at 0, 2, 4, 8 and 16 hours post inoculation with a mock solution or rhizobia were used in an RT-PCR experiment with a two minute extension time (Figure 6.5E-F). Similarly, at all time-points tested and for both treatments the only form of *DMI3* detected corresponded to a full length *DMI3* amplification product and no alternative splice forms were detected.

In addition, an RNA-Seq analysis was undertaken for a comparison of un-inoculated and inoculated A17 root hairs (Martin Trick and Andrew Breakspear, personal communication). Under both conditions, the full-length form of *DMI3* was detected with no detected splicing (data not shown). Therefore, under these conditions no evidence of alternative splicing of *DMI3* was detected.
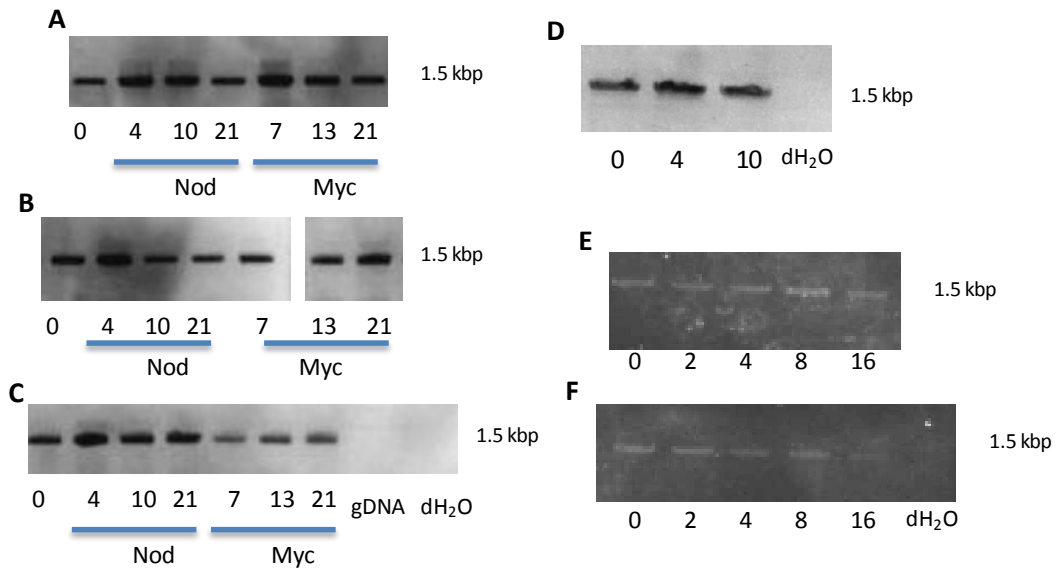
Figure 6.5. Full-length *DMI3* cDNA is the predominant form detected during nodulation and mycorrhization. All panels depict RT-PCR experiments carried out with a 2 minute extension time.
(A) RT-PCR using *DMI3* primers for Jester samples at 7 time-points during nodulation and mycorrhization (Figure 6.3A). The only band detected was approximately 1.5 kbp, which corresponds to full-length *DMI3*.
(B) RT-PCR for A17 samples, as in **(A)**.
(C) RT-PCR for GFP-CCaMK-SBP 694/1 samples, as in **(A)**. A positive control with genomic DNA (gDNA) and negative control with water were included in the final two lanes.
(D) RT-PCR was also carried out using *DMI3* primers with A17 cDNA 0, 4 and 10 dpi with rhizobia as provided by Ben Miller. The only band detected at all time-points is marked and corresponds to 1.5 kbp.
(E) RT-PCR was also carried out using *DMI3* primers with Jester cDNA 0, 2, 4, 8, and 16 hours post inoculation with a mock solution as provided by Jodi Lilley. The only band detected at all time-points was also at 1.5 kbp.
(F) RT-PCR for Jester cDNA 0, 2, 4, 8, and 16 hours post inoculation with rhizobia, as in **(E)**. The only band detected at all time-points was also at 1.5 kbp. A negative control with water was included in the final lane **(F)**.

6.7 Reduction of the PCR extension time reveals low levels of splice variants

An RT-PCR experiment for Jester and A17 over a time-course of nodulation and mycorrhization revealed full-length *DMI3* and no alternative splicing. However splice-variants had previously been detected in A17 during the establishment of nodulation by Pierre-Marc Delaux (Figure 6.2). In order to determine if splice variants were present during the time-course of nodulation and mycorrhization but were not detected (Figure 6.5) due to low expression levels, the elongation time in the RT-PCR experiment was reduced to favour their amplification.

Firstly, the elongation time was reduced to one minute (Figure 6.6), and RT-PCR was repeated for all time-points for Jester, A17 and GFP. Interestingly, low levels of potential splice variants were revealed by the reduction in the elongation time. The amplification product corresponding to full-length *DMI3* was still the strongest for all lines and during every stage of nodulation and mycorrhization. However, in all Jester (Figure 6.6A) and A17 (Figure 6.6B) samples an approximately 500 bp band was also present, which was not detected with the 2 minute extension time. However this 500 bp band was potentially also detected in one GFP sample at day 0, and therefore this may not correspond to a genuine splice-variant. In addition, in Jester day 0 and day 13 Myc, and A17 day 0 samples, another potential splice variant was detected at approximately 1 kbp. This 1 kbp band was absent from the GFP control line and therefore it appeared that this may be a genuine splice variant.

In order to determine if weak potential splice-variants were also present during the early stages of nodulation, the elongation time was also reduced to one minute for the Jester cDNA corresponding to 0-16 hours post inoculation (Figure 6.6D-E). As in Figure 5.6A-C, the full-length *DMI3* PCR product was still the most abundant at all time-points. Interestingly, the smaller amplification products that were detected were not the same as with the later time-course. Many very weak bands were detected in these samples and unexpectedly these were more prominent in the mock inoculated samples (Figure 6.6D) than in the rhizobia-inoculated samples (Figure 6.6E). These were extremely weak in comparison to those detected in the later time-course so it seems likely that they do not correspond to genuine splice-variants and instead may be artefacts of the PCR. They were therefore not considered further, and work instead focussed on those low-level splice-forms detected during later stages of nodulation and mycorrhization.
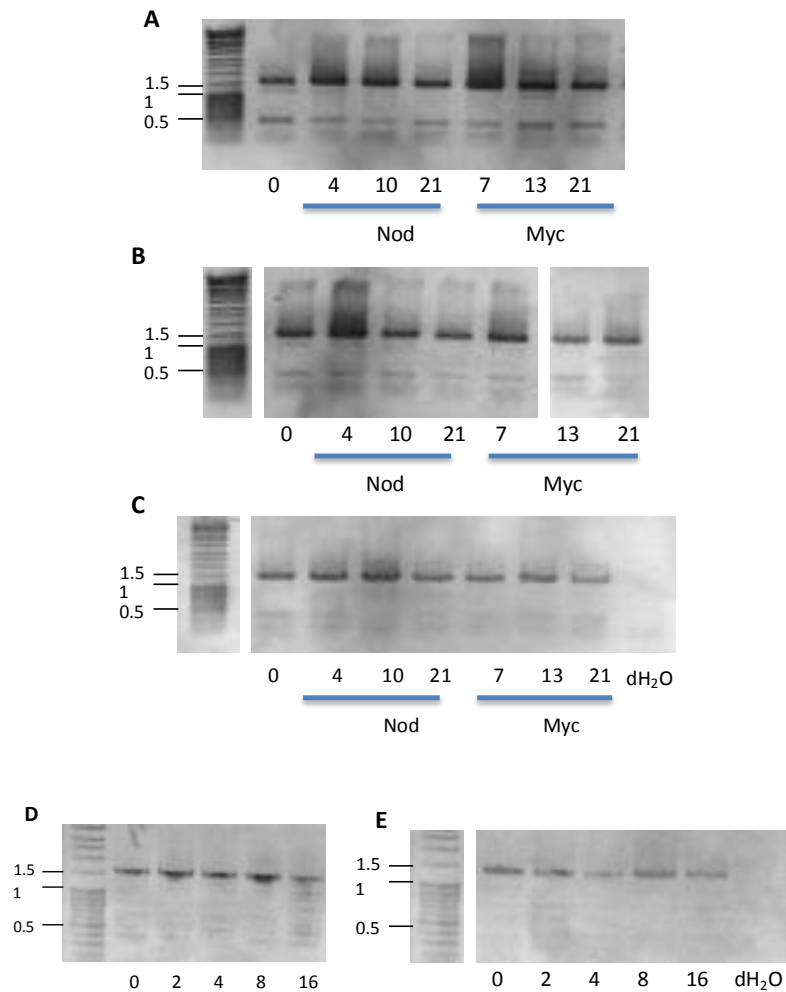
Figure 6.6. Potential splice-forms are enhanced with shorter extension times during RT-PCR. All panels correspond to RT-PCR experiments carried out with a 1 minute extension time.

(A) RT-PCR using *DMI3* primers for Jester samples at 7 time-points during nodulation and mycorrhization (Figure 6.3A). The sizes marked in the ladder are in kbp.

(B) RT-PCR for A17 samples, as in **(A)**.

(C) RT-PCR for GFP-CCaMK-SBP 694/1 samples, as in **(A)**. A negative control with water was included in the final lane.

(D) RT-PCR was also carried out using *DMI3* primers with Jester cDNA 0, 2, 4, 8, and 16 hours post inoculation with a mock solution as provided by Jodi Lilley.

(E) RT-PCR for Jester cDNA 0, 2, 4, 8, and 16 hours post inoculation with Rhizobia, as in **(D)**. A negative control with water was included in the final lane.

6.8 Splice variants are further enhanced with a 30 second extension time

In order to further enhance the potential splice variants detected in the RT-PCR with a one minute extension time, the RT-PCR was repeated with a further reduction of the extension time to 30 seconds. This was carried out for all Jester, A17 and GFP samples in the nodulation and mycorrhization time-courses (Figure 6.7). This experiment revealed further potential splice variants in both the Jester time-course (Figure 6.7A) and A17 time-course (Figure 6.7B) in addition to those previously detected. Consistent with the one minute extension time experiment, in all cases full-length *DMI3* was the most predominant form, including the control GFP line (Figure 6.7C). The 500 bp band was again the second most abundant, and was present in all Jester and A17 samples and strangely five out of six of the GFP line samples. Therefore this amplification product corresponded to either a non-specific PCR product or several different amplification products which were approximately 500 bp. Therefore those products detected in the GFP line samples would correspond to a non-specific product, and those detected in the WT lines would likely correspond to a splice variant.

In the Jester time-course, the 1 kbp amplification product was now detected at all time-points. In addition, six out of seven of the A17 samples also possessed this amplification product. Furthermore, several Jester time-points also had additional potential splice-variants; namely 4 dpi with rhizobia, and 7 and 13 dpi with mycorrhization. Similarly, the same three samples in the A17 time-course contained low levels of potential splice forms, and in addition 21 dpi with mycorrhization showed several potential forms. It was not possible from this particular experiment to determine the sizes of all the detected forms due to poor resolution of the DNA ladder and these potential splice-forms therefore required further characterisation.

Therefore shorter extension times revealed several potential splice-forms in both Jester and A17. They were extremely low in abundance but, since several of these forms were not detected in the GFP control line samples, they were potentially genuine splice forms.
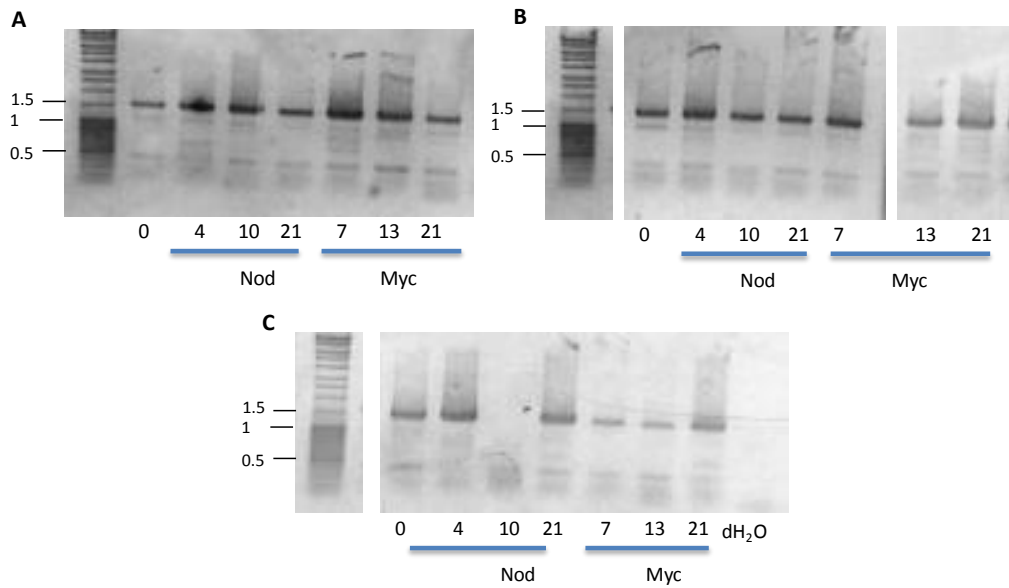
Figure 6.7. Potential splice forms are enhanced further by halving the elongation time during RT-PCR again. All panels correspond to RT-PCR experiments carried out with a 30 second extension time.

(A) RT-PCR using *DMI3* primers for Jester samples at 7 time-points during nodulation and mycorrhization (Figure 6.3A). The sizes marked in the ladder are in kbp.

(B) RT-PCR for A17 samples, as in **(A)**.

(C) RT-PCR for GFP-CCaMK-SBP 694/1 samples, as in **(A)**. A negative control with water was included in the final lane. One PCR reaction did not work due to a potential pipetting error at 10 dpi with nodulation, however this does not reflect on the cDNA quality because this sample has been successfully amplified during other experiments (e.g. Figure 4.12C).

6.9 Cloning of splice variants

In order to sequence the potential splice forms detected during the RT-PCR experiments, two PCR product samples were selected for cloning (Figure 6.8A). These two samples were from the 30 second elongation RT-PCR experiment, and corresponded to Jester at four dpi with rhizobia and seven dpi with mycorrhizal inoculum. They were selected because they contained the common potential splice-variants detected in all the A17 and Jester time-points tested (500 bp and 1 kbp PCR products) and they also contained other potential splice variants, the size of which was unclear from preliminary experiments.

These bands were cloned into the pDRIVE sequencing vector using the PCR Cloning Kit (Qiagen) and colonies were screened by colony PCR and subsequent gel electrophoresis (Figure 6.8B-C). A large proportion of the PCR products which were cloned corresponded with the full-length *DMI3* amplification product at 1.5 kbp, as expected. However several of the potential splice-variants were also successfully cloned by this method. All amplification products which were either larger or smaller than full-length *DMI3* and were therefore potential splice variants were sequenced. These are marked on Figure 6.8B-C.
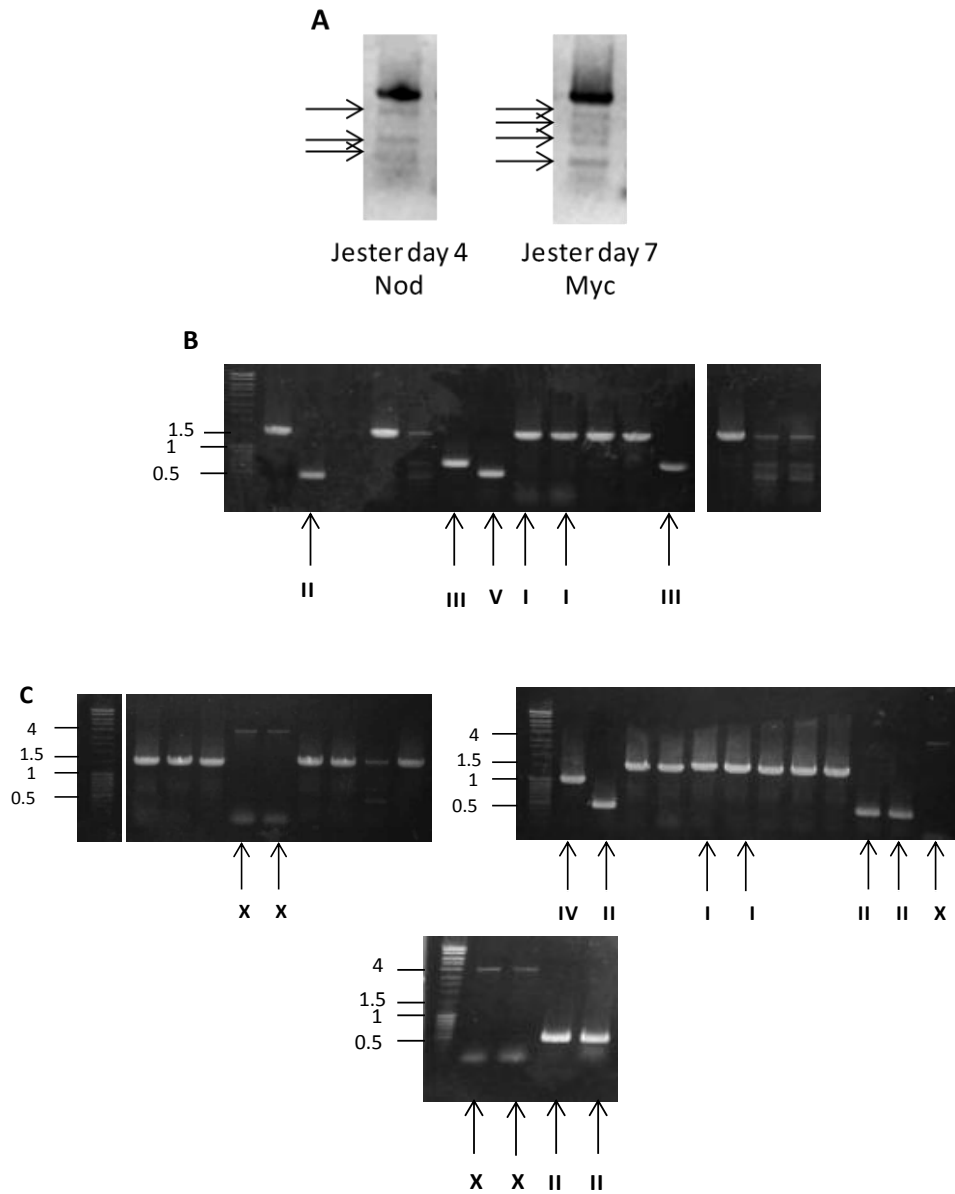
Figure 6.8. Cloning of potential splice-forms of *DMI3*.
(A) Two samples from the RT-PCR with a 30 second elongation step were selected for identification of splice variants due to the presence of distinct bands, which are marked by horizontal arrows. These were used for cloning and verified by colony PCR.
The successfully cloned variants from Jester at 4 dpi with nodulation and 7 dpi with mycorrhization are shown in **(B)** and **(C)** respectively. The indicated bands were selected for sequencing, and the corresponding splice variants are named with roman numerals. The sizes marked in the ladder are in kbp.

6.10 Sequencing of splice variants

The splice variants were sequenced (Figure 6.8) and this confirmed that they were all *DMI3*-derived, except for the variant marked X. All sequences which corresponded to band X were subjected to a BLAST analysis against the *M. truncatula* nucleotide database which revealed no matches. However, a BLAST analysis with the complete nucleotide database revealed band X corresponded to the cloning vector used (pDRIVE). This may have occurred due to the presence of primer dimers that were present in the PCR reactions, which were subsequently ligated into the cloning vector. These bands were therefore not considered further since they did not correspond to potential splice-variants.

Two potentially full-length *DMI3* bands at 1.5 kbp were selected from each sample (Jester 4 dpi and 7 dpi with rhizobia/mycorrhizal inoculum respectively) and all four of these amplification products were confirmed as full-length *DMI3* with all 7 exons present (Figure 6.9 I).

The other four potential variants that were successfully cloned corresponded with splice variants containing whole or partial *DMI3* exons (Figure 6.9, II - V). From the Jester 4 dpi with rhizobia sample, three alternative splice forms were cloned and were marked as II, III and V. II corresponded to the first few amino acids of the kinase domain and the EF-hands, with a partial deletion at the beginning of exon 5 and exon 6 and 7 present in their entirety. A protein form containing these regions would bind calcium, however would not be a functional kinase because it lacks the vast majority of the kinase domain. However, this fusion was not in frame and therefore would likely not produce a corresponding protein form. Form III corresponded to a partial kinase domain (which possessed the ATP-binding site but not the active site) and a partial VLD (which included only the final two EF-hands). Therefore this protein would be able to bind ATP and calcium; but it would again not be active for target-phosphorylation due to the lack of a kinase domain. However, again, the fusion of the partial VLD was not in frame and therefore this would likely not be reflected in protein form. Form V corresponded to a partial kinase domain (containing the ATP-binding site but no active site) with the final EF-hand only. This potential protein would be able to bind ATP and calcium; but it would again not be active for target-phosphorylation due to the lack of a kinase domain. Interestingly, this fusion was in-frame and therefore could potentially be reflected at the protein level.

Splice form II was also cloned from the Jester 7 dpi with mycorrhization sample, and another novel form, IV, was also cloned. This form contained some of the kinase domain (including both the ATP-binding and the active site), the end of the auto-inhibitory domain (which lacked the CaM binding site) and all of the EF-hands. It is unclear how a protein containing these domains would function since it is possible it would be an active kinase because this splice-variant possessed the active site, in contrast with the other forms identified. However the Thr271 site was absent which negatively regulates the protein, and the CaM binding domain was also lacking which provides positive regulation of the protein. Therefore it is not clear whether a protein containing these regions would be active or inactive. However, this fusion was again not in-frame and so would likely not form a corresponding protein.

Consequently, it is unclear if these splice forms are biologically relevant or if any of these forms, or those previously detected (Figure 6.2), are reflected at the protein level. Therefore, it was important to determine whether any splice variants are present at the protein level *in planta.*

184
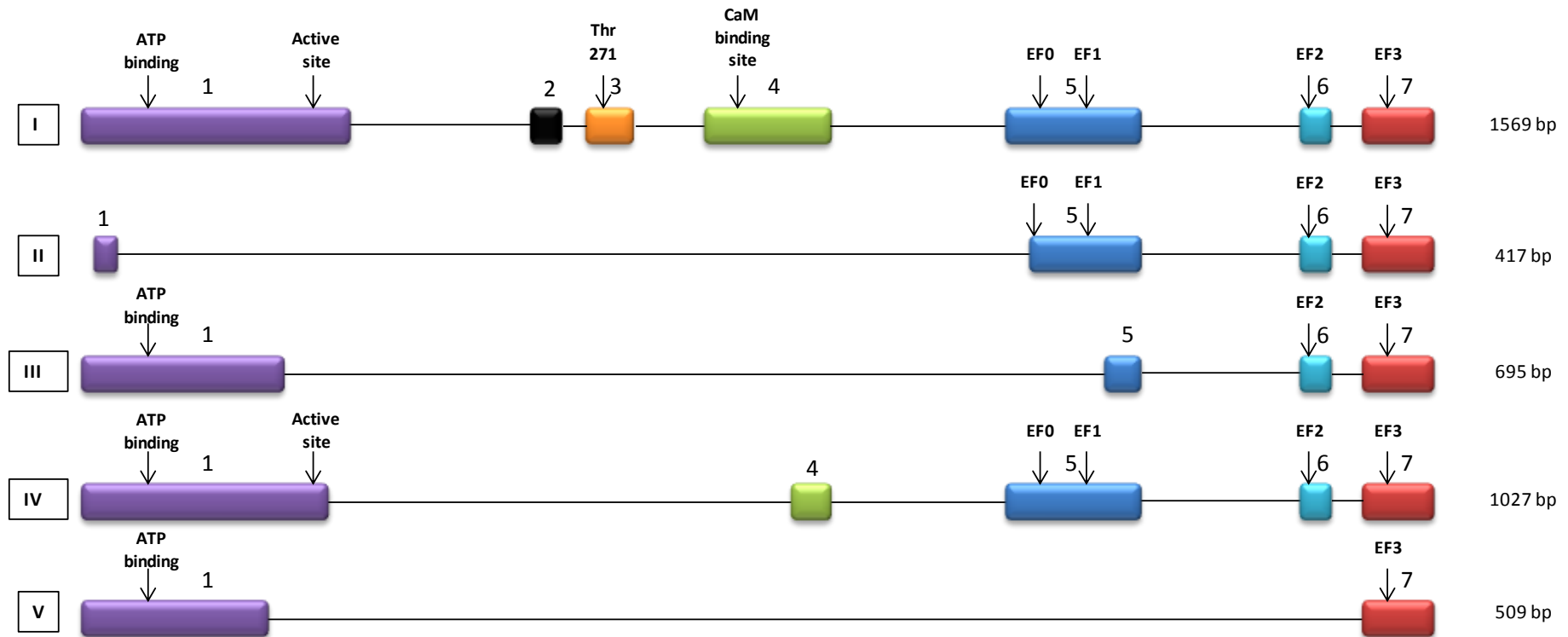
Figure 6.9. Schematic representation of splice variants of *DMI3*.
The splice variants detected during cloning (Figure 6.8) were sequenced and the corresponding complete or partial exons are shown in each. The total lengths of the splice-variants are indicated. Form I corresponds to the exon structure of the full-length *DMI3* gene.

6.11 Western blotting of Jester nodulation and mycorrhization time-courses

In order to determine if any splice forms were present at the protein level, a western blot analysis of Jester root samples was undertaken.

First, to establish the optimum extraction method for the root samples, two buffers were trialled with un-inoculated Jester root samples (Figure 6.10A). Jester samples grown on plates for 7, 14 or 21 days with no rhizobial inoculation were extracted using either YS or Li buffer (for YS buffer see Materials and Methods, for Li buffer refer to (Li *et al.*, 1996)). The same tissue for each time-point was divided equally by weight between each buffer. Under these conditions, six main protein bands were detected by αMtCCaMK with each extraction buffer; however, the proportion of each band was different with the two treatments. Li buffer was selected for all further western blotting experiments due to a superior extraction of proteins, particularly of full-length CCaMK at 58 kDa. Although the weight of tissue used between 7, 14 and 21 days varied (1.1, 1.4 and 3.1 g, respectively) it was possible to conclude that the levels of each protein detected did not vary greatly with the age of the tissue. Since the level of protein extraction from plant tissue was extremely low (as detailed in Chapter 5) it was not possible to determine the identity of the six proteins identified with αMtCCaMK by mass spectrometry. Since there was previous evidence of cross-reactivity of this antibody, it was unclear if smaller proteins and the protein at 96 kDa corresponded to CCaMK-derivatives or to other proteins to which the antibody cross-reacted. However it did seem likely that the band seen at 58 kDa corresponded to full-length CCaMK.

In order to determine if CCaMK protein splice-forms were present during nodulation or mycorrhization, Li buffer was used to extract Jester samples from one batch of plants for the time-course of nodulation and mycorrhization (Figure 6.10B). As seen in Figure 6.10A, six main proteins were detected with αMtCCaMK in most samples (except Nod 4, 10 and 21, which had seven). As before, there was insufficient protein to verify by mass spectrometry if those proteins detected corresponded to CCaMK, however it was likely that the band at 58 kDa corresponded to full-length CCaMK. It appeared that the 58 kDa protein stayed at the same level at each time-point, except 4 dpi with rhizobia. This analysis was somewhat quantitative because buffer was added to 2 ml/g of ground tissue but, due to the viscosity of the crude samples which potentially led to small variations in the volumes

pipetted, the results were not fully quantitative. Due to the known cross-reactivity of the αMtCCaMK antibody, it was not possible to determine if there were splice-variants present or if the 58 kDa band definitively corresponded to CCaMK from this experiment.
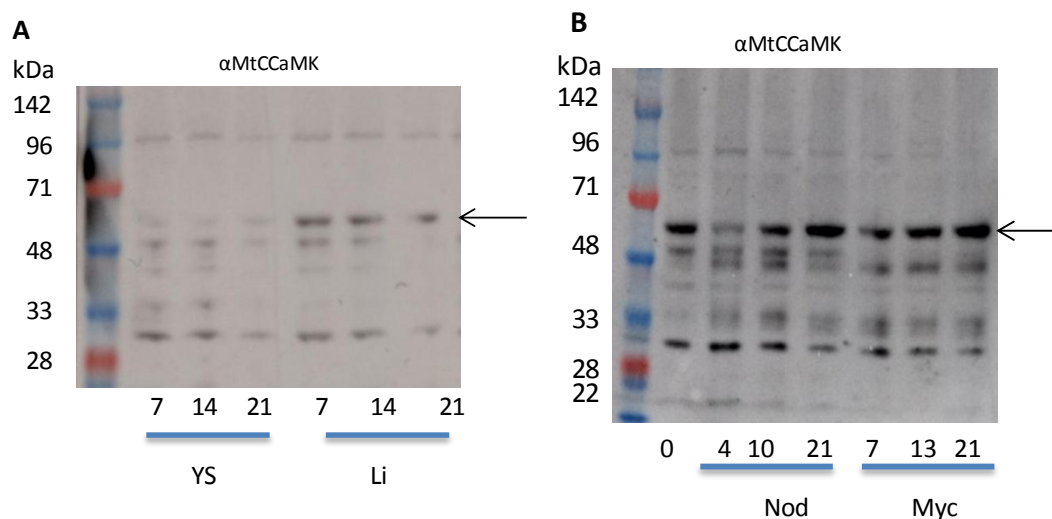
Figure 6.10. Characterisation of CCaMK protein during nodulation and mycorrhization. Both panels are western blots developed with αMtCCaMK.
(A) Two buffers were trialled to establish the optimal method for protein extractions for the plant root time-course samples. YS and Li buffers were used with *M. truncatula* Jester root samples grown on plates for 7, 14 or 21 days with no inoculation. Li buffer was selected for use in time-course western blot and immunoprecipitation (IP) experiments due to the better extraction of proteins, particularly full-length CCaMK (58 kDa).
B) Western blot analysis of Jester samples at all 7 time-points studied (Figure 6.3A). Briefly, samples were ground to a fine powder in liquid nitrogen, extracted with 2 ml/g Li buffer and prepared for SDS-PAGE. This extraction revealed several potential different forms of CCaMK in addition to full-length protein, but the antibody is known to cross-react with other proteins.

6.12 IP reveals the full-length CCaMK protein form

In order to determine whether proteins detected with αMtCCaMK corresponded to splice-forms of CCaMK or were cross-reactive proteins, IP was carried out for the same Jester root samples. Since this procedure was more involved, it was only carried out for Jester samples at the seven time-points during nodulation and mycorrhization (and not for A17 or GFP samples). Interestingly, at all seven time-points the predominant protein that immunoprecipitated with αMtCCaMK was present at 58 kDa and therefore most likely corresponded to full-length CCaMK (Figure 6.11A). The appropriate gel slice for each sample was analysed by mass spectrometry, and CCaMK was detected in six samples,

indicating that full-length CCaMK was likely immunoprecipitated by this method. This result coincided with the predominant form detected during RT-PCR experiments which was the full-length transcript at all time-points. Two smaller proteins were also detected by this method at approximately 43 and 30 kDa, however these contributed to a tiny percentage of the total protein detected and therefore were unlikely to be biologically significant.

Due to the nature of the IP experiment the output bands were not quantitative and therefore it was not possible to determine if concentrations of full-length CCaMK protein changed during the symbiosis signalling process. It follows therefore that the weaker full-length band of the day 13, mycorrhization sample does not necessarily suggest that levels of CCaMK were lower at this stage of signalling.

The result of the IP experiment was in contrast to the western blot of crude protein extract with αMtCCaMK (Figure 6.10B), which showed six bands that potentially corresponded to CCaMK splice-forms. The IP result showed therefore that the five proteins (excluding full length 58 kDa) were most likely cross-reactive proteins, and the epitope(s) which was/were recognised by the antibody in each of these proteins may have been internal to the proteins. The epitope(s) would therefore be recognised by the antibody when the proteins were denatured, for example during SDS-PAGE and subsequent western blotting. However, the epitope(s) would not be detected when the protein was in a native state, for example during IP. Therefore, it can be concluded from the IP experiment that the predominant form of CCaMK protein during symbiosis signalling is full-length and no significant splice-forms were detected.

The input and flow-through fractions for each of these IP samples are shown in Figure 6.11B-C. All five of the detected proteins from the western blot (Figure 6.10B) (excluding full-length CCaMK) were very weakly detected in the input and flow-through samples. This confirmed that these were cross-reactive proteins which did not bind to the antibody when in the native state, since they were detected before and after IP. The 58 kDa band is also detected very weakly in the wash and flow-through; this is either because binding to the antibody did not occur completely or there was another 58 kDa cross-reactive protein.
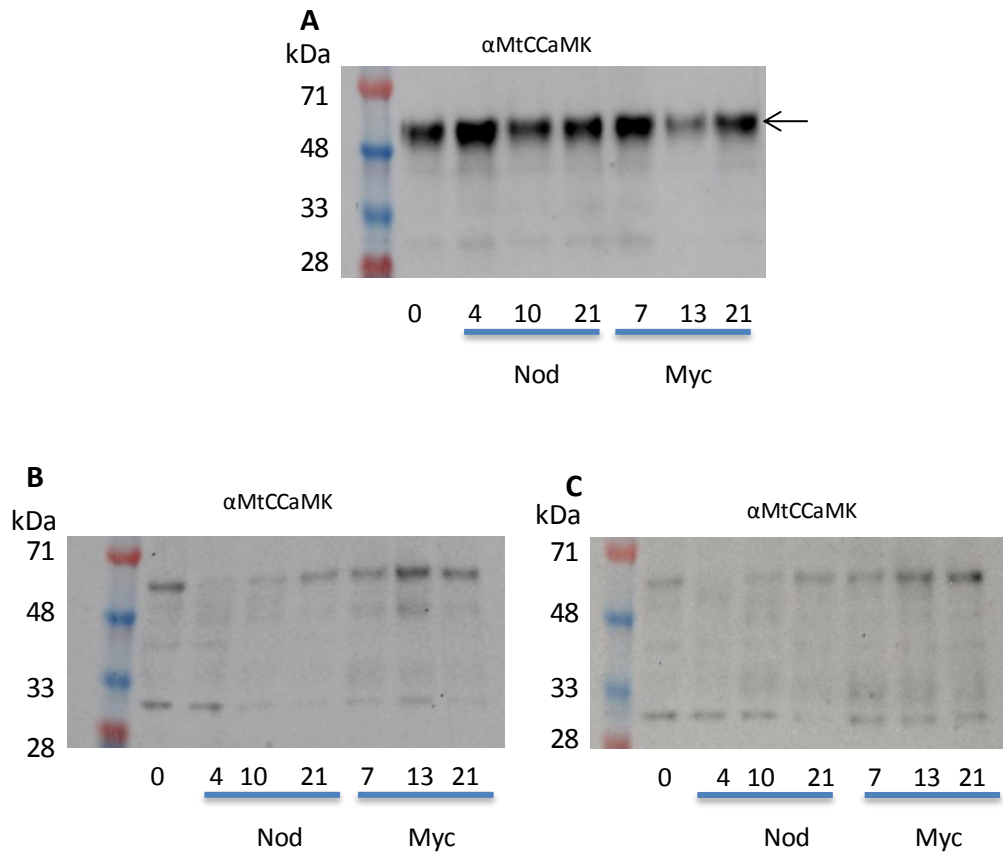
Figure 6.11. IP of CCaMK reveals full-length protein is the predominant form at key stages of nodulation and mycorrhization. All panels show western blots developed with αMtCCaMK.

(A) The immunoprecipitates of Jester extracts at 7 time-points (Figure 6.3A) are shown. This was carried out using magnetic beads cross-linked to αMtCCaMK and the predominant band (marked with an arrow) corresponds with full-length CCaMK protein (approximately 58 kDa).

(B) The input samples from the IP of Jester extracts at 7 time-points **(A)** using magnetic beads cross-linked to αMtCCaMK are shown.

(C) The flow-through fractions from the IP of Jester extracts **(A)** are shown.

Chapter 6: Alternative splicing of MtCCaMK

6.13 Kinase domain antibodies do not reveal CCaMK splice-forms

Since the αMtCCaMK antibody was produced against the AI-VLD antigen, protein splice-variants could have been absent in the IP experiment if they contained only the kinase domain (33.3 kDa). Therefore antibodies with affinity to a region of the kinase domain were utilised with representative samples to determine if it was possible to detect other forms of CCaMK (Figure 6.12). Samples used for this experiment corresponded to the final time-point in the time-course experiment for A17 and Jester, with mycorrhization or nodulation at 21 dpi. First, a western blot for the samples was carried out and probed with αMtCCaMK (Figure 6.12A). The resulting protein bands detected corresponded with the bands seen during a previous time-course extraction (Figure 6.10B) where subsequent IP showed that full-length protein was predominant. It can therefore be assumed that this form was also predominant in this extraction when detected with the antibody with affinity for the AI-VLD.

Two antibodies were tested with affinity to the kinase domain of CCaMK. αpT271 (previously used in Chapter 3) has affinity for the phospho-form of Thr271. The antibody annotated as αT271 corresponds to the non-specific fraction from the purification of the anti-phospho antibody, which has previously been demonstrated to bind to both the phospho- and non phospho-forms of CCaMK (data not shown). Therefore, using both of these antibodies would indicate if either phosphorylated or non-phosphorylated forms of CCaMK were present. The same A17 and Jester 21 dpi samples were subjected to western blotting and probed with either the αT271 or the αpT271 antibody. Figure 6.12B showed the western blot probed with αT271 and all samples possessed a single strong protein band at 48 kDa. This was most likely not a splice variant and instead corresponded to a cross-reactive protein. This is because this protein was not pulled-down with the previous IP experiment (6.11A) and a kinase-only splice-variant would correspond to a size of 33.3 kDa or less. A BLAST analysis was carried out in order to identify potential candidates for this cross-reactive protein, with the peptide used for the production of this antibody (CSFYEKTWKGISQ). The three most likely candidates identified were: ERG2 and Sigma1 receptor-like protein (NCBI Ref.: XP_013469825.1), tRNA pseudouridine synthase (NCBI Ref.: XP_003630601.2) and a B3 DNA binding domain protein. These candidates were identified because they contained more than four consecutive amino acids in common with the peptide. Of these, the most likely protein for cross-reactivity was the B3 DNA binding

domain protein (NCBI Ref.: XP_013466763.1), because it has a predicted size of approximately 48 kDa. This protein may correspond to the 48 kDa protein which was previously detected by αMtCCaMK, but this is unlikely since the epitopes of the antibodies which detected these proteins were different.

Figure 6.12C corresponds to the same four Jester and A17 samples probed with the anti-phospho specific antibody, αpT271. No bands were detected despite a long exposure time. This is most likely because the affinity of the αpT271 antibody was too low to detect proteins at this concentration. Alternatively, CCaMK or splice-forms under these conditions were entirely un-phosphorylated at Thr271 and therefore were not detected.

It was therefore not possible to determine if kinase-only forms of CCaMK are present during symbiosis establishment using these antibodies, but the RT-PCR and IP data taken together suggests that it is highly likely that the predominant form of CCaMK is full-length.

A αMtCCaMK

Myc	+	-	+	-
Nod	-	+	-	+

kDa
96
71
48
33
28
22

A17   Jester

B αT271

Myc	+	-	+	-
Nod	-	+	-	+

kDa 96
71
48
33
28
22

A17   Jester

C αpT271

Myc	+	-	+	-
Nod	-	+	-	+

kDa 96
71
48
33
28
22

A17   Jester
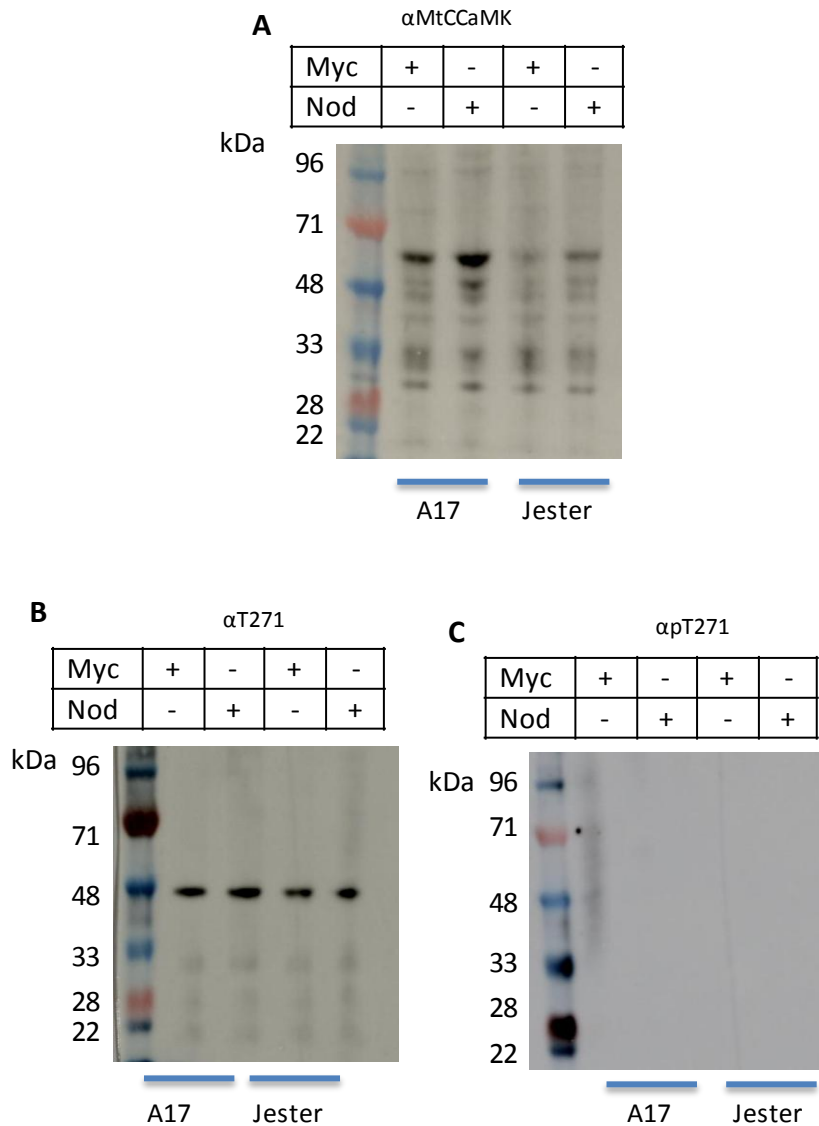
Figure 6.12. Use of kinase domain antibodies does not provide further insight into CCaMK splicing.
(A) A western blot using αMtCCaMK with A17 and Jester samples, 21 dpi with mycorrhization or nodulation as indicated. Samples were inoculated as 10 day old seedlings, and re-planted from Mod FP agar to terragreen sand.
(B) A western blot using αT271 with the same A17 and Jester samples as **(A)**, 21 dpi with mycorrhization or nodulation as marked. One main protein band was detected at approximately 48 kDa with αT271 however this is most likely not a CCaMK-derived protein.
(C) A western blot using αpT271 with the same A17 and Jester samples as **(A)**, 21 dpi with mycorrhization or nodulation as marked. Despite a long exposure time of the blot, no bands were detected.

6.14 Discussion

An investigation was undertaken in order to determine if splicing plays a role in how CCaMK functions in the Sym pathway. Very low abundance splice-variants of *DMI3* were revealed at the RNA level under short-extension time PCR conditions for both Jester and A17. Interestingly, only one of these forms, form V, could potentially be reflected at the protein level due to an in-frame fusion. This form contains a partial kinase domain and one EF-hand, EF-hand 3. Therefore, if this form was reflected in a corresponding protein form, it would not encode a functional kinase but it could bind ATP and calcium. The remaining three *DMI3* splice-forms detected contained regions which were spliced out-of-frame and therefore would likely not be translated into corresponding protein forms.

Surprisingly, none of the forms sequenced matched exactly to those detected previously by Pierre-Marc Delaux. Pierre-Marc detected three main splice forms: 1173, 993 and 498 bp. The 1173 bp form contained a partial kinase domain and the VLD (Figure 6.2). This form was most similar to form IV, and, if functional, it would most likely be a calcium-regulated kinase. However, the exact sequences present were different. The 993 bp splice-form had no likely counterpart in the experiments detailed in this Chapter. This form lacked key regions of the kinase domain necessary for functionality, but retained both the CaM binding domain and EF-hands. None of the forms detected in the studies subsequently presented in this Chapter retained both calcium-responsive elements of the protein. However, the 498 bp form is very similar to form V detected in these experiments. They both contain a kinase domain with only the ATP binding site, and one EF-hand of the VLD. Therefore, this protein could act as a calcium-binding protein only. Surprisingly though, the exact sequences encoding these similar splice-forms were different. The fact that none of the splice forms detected was identical to those previously identified suggests that they do not have key roles in the Sym pathway.

Unexpectedly, it also appeared that one or more of these forms were present in the control GFP-CCaMK-SBP line. This line lacks splicing of *DMI3* because it only possesses the coding sequence of *DMI3*. Therefore, these amplification products most-likely do not correspond to *DMI3* splice-forms and are most likely non-specific bands or artefacts.

Crucially, as demonstrated in this chapter, no significant splice-variants were reflected at the protein level for all nodulation and mycorrhization Jester time-points tested. This is

because the predominant protein form at all time-points corresponded to full-length protein. It seems likely that this is also the case for the A17 samples, because a similar pattern of amplification products was detected in the RT-PCR for this line and the most-abundant of these was also full-length in all samples. In fact it could be predicted that the splice-variants would not be reflected in the predominant protein forms detected, because they reflect an extremely low proportion of the total *DMI3* amplification products detected. In addition, several forms contained a premature stop codon and therefore would likely not be reflected at the protein level. One possibility is that the low abundance splice-variants detected at the RNA level correspond to errors that occur during normal mRNA processing.

Under these conditions it therefore appears that alternative-splicing of CCaMK does not play a significant role in symbiosis signalling. Crucially, as demonstrated in chapter 5 and by others, for example (Shimoda *et al.*, 2012), the coding sequence of *DMI3* complements the *dmi3-1* mutant for symbiosis establishment. Therefore any role of alternative splicing of CCaMK during symbiosis would be minor or of no significance. However, since some splice-forms of *DMI3* were detected, it is possible that the splicing of CCaMK could play a role in other signalling pathways or in other tissues.

6.15 Summary

The predominant form of CCaMK in nodulation and mycorrhization signalling is full-length, at both RNA and protein levels. Very low abundance splice-variants of *DMI3* were revealed under specific PCR conditions but these did not match previously detected variants and were not reflected at the protein level. Therefore, it can be concluded that alternative splicing of CCaMK does not play a major role in symbiosis signalling.

Chapter 7: General Discussion

The aim of this project was to characterise CCaMK in the symbiosis signalling pathway. In order to do this a multi-faceted study was undertaken in order to address key questions. This thesis has provided insight into how the protein responds to calcium in the nucleus, including the size of the oligomer and the likely mechanism of autophosphorylation at the important T271 site. Both of these factors impact on the kinetics of CCaMK function and potential modelling of the system. Furthermore, the conformational change when calcium binds was characterised and this provides insights into the possible roles this plays in CCaMK as a signal transducer. Finally the wider role of CCaMK in the pathway *in planta* was addressed by studying the changes in RNA and protein forms of CCaMK prior to symbiosis formation and during the progression of nodulation and mycorrhization.

The CCaMK oligomer

Autophosphorylation at T271 has recently been revealed to be crucial for the regulation of CCaMK function by the maintenance of the inactive state of the protein. This occurs via a hydrogen bonding network predicted to form between pT271 and two residues of the CaM binding domain at basal calcium levels (Miller *et al.*, 2013). Autophosphorylation is also key to the regulation of the animal protein CaMKII and, in contrast to CCaMK, the mechanism by which this occurs is well characterised. Interestingly, phosphorylation in CaMKII occurs at T286 within the dodecameric oligomers at adjacent subunits when both of the subunits involved are bound to Ca^{2+}/CaM (Hudmon and Schulman, 2002b). Therefore, in order to address how autophosphorylation occurs in CCaMK, it was first necessary to identify the oligomeric status of CCaMK which was currently unknown. By the use of three different methods described in Chapter 3, it was determined that MBP-CCaMK when purified from *E. coli* forms an oligomer. Furthermore, an estimate of the oligomer size was produced of 16-18 subunits with a diameter of ~31 nm.

These data therefore show that CCaMK is capable of forming an oligomer similar to that of CaMKII. However, it could form an oligomer with more than 12 subunits. This corresponds to a large protein complex of between ~1600 – 1800 kDa, since each monomer of MBP-CCaMK is ~100 kDa. However, the mechanism by which this oligomerisation occurs is likely to be different to that of CaMKII. CaMKII possesses a C-terminal association domain required for the establishment of the oligomer (Hudmon and Schulman, 2002a); however CCaMK does not possess an analogous domain. There was also more variability detected in

the oligomer size of CCaMK. Whilst the SEC gave a single value for the oligomer size of 1607 kDa due to the presence of a sharp elution peak, both the DLS and AFM data showed more variability. Variability is expected when using DLS with a polydisperse sample due to the measurement of the size of particles in solution where they are in Brownian motion. Some variability can be observed with AFM due to artefacts that form during the deposition of the protein. However, the variability in sizes detected during AFM is most likely indicative of the presence of different particle sizes. This is most likely due to larger aggregates of MBP-CCaMK which formed after purification, and therefore they were only detected in the AFM and occasionally during DLS. Whilst MBP-CCaMK was the best available non-aggregating full-length construct of CCaMK, attempts to concentrate the protein led to aggregation. Therefore, it was not possible to carry out, for example, analytical ultracentrifugation to obtain a more precise oligomer size. Furthermore, techniques to obtain a more detailed image of the oligomer, for example cryo-EM, were not possible due to the aggregation observed with AFM.

Despite the variability in the exact size of the CCaMK oligomers which were detected, the evidence shows that CCaMK is predominantly present as oligomers as opposed to monomers. Crucially, no evidence of MBP-CCaMK monomers was detected in any of the experiments. These experiments were carried out using an N-terminal MBP tag, and it could be theorised that the tag used could potentially affect the oligomeric state of the protein. However, the MBP-tag is widely used as a fusion tag to aid solubility and prevent aggregation, whilst also allowing ease of purification (Kapust and Waugh, 1999). Furthermore, there is little evidence of the MBP-tag actually causing this aggregation, because preliminary data regarding a tagless form of MtCCaMK also shows the presence of a large oligomer, before the onset of significant aggregation (Liang Zhou). The fact that CCaMK forms an oligomer therefore has implications for the common Sym pathway, namely on how CCaMK interacts with downstream proteins to relay the symbiotic signal. Of main interest is CYCLOPS, which is a key interacting protein of CCaMK. CCaMK interacts with CYCLOPS in the absence of symbiotic signalling and CYCLOPS is a phosphorylation target of CCaMK (Yano *et al.*, 2008), suggesting that both the autophosphorylated inactive form of CCaMK and the form which is active for downstream phosphorylation interact with CYCLOPS. Furthermore, CYCLOPS has also been shown to form homodimers (Singh *et al.*, 2014). Together, this indicates that CCaMK-CYCLOPS may form a large hetero-complex before and during calcium spiking in order to detect and transduce the signal. However, the

Chapter 7: General Discussion

number of subunits present in this complex is currently unclear. In addition, it is likely that CCaMK interacts with further target proteins, for example CIP73 and other as yet unknown proteins, and their role in this complex is also currently unclear.

Intra-oligomeric autophosphorylation at T271

The mechanism by which autophosphorylation at T271 occurs has previously been studied kinetically. Despite the presence of two different trends for the protein concentration-dependence of autophosphorylation at T271, the authors concluded an inter-monomeric/oligomeric mechanism in both studies (Sathyanarayanan *et al.*, 2001), (Tirichine *et al.*, 2006). The data from these two papers are contradictory, therefore, and required further clarification. In addition, the data presented in Chapter 3 showing the presence of CCaMK in oligomers has implications for the mechanism of autophosphorylation. Therefore it was necessary to undertake a thorough investigation to distinguish between the four possible mechanisms of autophosphorylation: intra-monomeric, inter-monomeric, inter-oligomeric and intra-oligomeric.

The data presented in Chapter 3 does not conclusively define the mechanism of autophosphorylation; but it does strongly indicate an intra-oligomeric mechanism. This is due to the reasonable elimination of the other possible mechanisms, rather than the positive detection of intra-oligomeric phosphorylation. First, intra-monomeric autophosphorylation was deemed unlikely due to the relative locations of the active site and T271 on a homology model, which concurs with findings from homology modelling carried out by Sathyanarayanan and colleagues (Sathyanarayanan *et al.*, 2001). Second, three independent experiments showed no evidence of inter-oligomeric/monomeric autophosphorylation (Figures 3.6, 3.9 and 3.15), utilising three constructs of kinase dead CCaMK protein (GST-CCaMK K47E, MBP-CCaMK K47E and HisSUMO-CCaMK D192N). These were expressed separately, mixed with WT CCaMK, and then subjected to a kinase assay and a western blot probed with αpT271. These three experiments therefore determined whether inter-oligomeric or inter-monomeric autophosphorylation were possible. This is because, whilst the majority of the proteins were present as oligomers, it is likely some protein was present as monomers too. Furthermore, since three different fusion tags and two different kinase dead mutations were used it seems unlikely that these factors interfered with inter-monomer phosphorylation. The same experiment was used to

demonstrate that autophosphorylation of SYMRK occurs with an intermolecular mechanism (Yoshida and Parniske, 2005) therefore it seems likely that this experiment would allow inter-molecular autophosphorylation of CCaMK if possible. As this is a negative result it cannot be unambiguously verified, however taken together it strongly suggests that inter-oligomeric/monomeric autophosphorylation at T271 in CCaMK does not occur.

The remaining possible mechanism is intra-oligomeric autophosphorylation. In order to determine if this occurs in CCaMK oligomers, mixed oligomers were produced containing MBP-CCaMK K47E and WT CCaMK. These proteins were co-expressed and purified, and phosphorylation at T271 in each type of subunit was determined after a kinase assay using a western blot probed with αpT271. Whilst potential intra-oligomeric autophosphorylation was detected, a subsequent λ phosphatase treatment did not remove this signal. Therefore, it must be concluded that the signal detected was non-specific background. This background likely corresponds to the binding of αpT271 to non-phosphorylated protein occurring at a low level, and detected in this case due to the long exposure time used during the development of the western blot. However it seems likely that this indicates not that the hypothesis of intra-oligomeric autophosphorylation is wrong, instead that the strategy used gave an ambiguous result.

This strategy has previously been utilised to determine the mechanism of autophosphorylation for CaMKII, which was demonstrated to occur in an inter-subunit intra-oligomeric manner (Mukherji and Soderling, 1994). There are several reasons why this method may not have worked in this particular experiment for CCaMK. First, as discussed above, the way in which CCaMK forms oligomers is likely different to CaMKII due to the lack of an identified analogous association domain. It could be possible in this case that the presence of a kinase-dead mutation could have caused a conformational change and therefore aberrantly affected the way in which the mixed oligomers formed, preventing cross-phosphorylation. Furthermore, the MBP-tag is a large fusion tag of approximately 40 kDa (Terpe, 2003) and it is possible that the presence of the tagged and un-tagged version of the protein when assembled into a complex do not form a favourable conformation permitting *trans*-phosphorylation. Alternatively, perhaps ATP must bind to both the donor and recipient to allow inter-oligomeric phosphorylation, and, if this is the case, the kinase-dead mutation would also prevent its own autophosphorylation. In addition, the kinase

dead mutation could prevent the phosphorylation of inactive subunits perhaps by a structural change in the protein, which would affect both the inter- and intra-oligomeric experiments. Such effects of the kinase-dead mutations cannot be ruled out.

These data taken together therefore suggest, but do not conclusively prove, an intra-oligomeric mechanism of autophosphorylation at T271. This is similar to the intra-oligomeric mechanism observed for CaMKII; however, there are also substantial differences between the two mechanisms. Whilst CaMKII is autophosphorylated at T286 by Ca^{2+}/CaM-binding, CCaMK is instead autophosphorylated in response to Ca^{2+} ions binding to the EF-hands. In addition, autophosphorylation appears to regulate the proteins in opposite ways: CaMKII is activated by phosphorylation at T286 and CCaMK is inactivated by phosphorylation at T271 (Miller *et al.*, 2013). Furthermore, it has been shown that it is necessary for Ca^{2+}/CaM to bind to adjacent subunits in CaMKII in order for this autophosphorylation to occur (Hanson *et al.*, 1994). Therefore, there are key differences between these two mechanisms.

An intra-oligomeric mechanism of autophosphorylation at T271 has implications for the wider mechanism of CCaMK regulation and activation. First, it suggests that autophosphorylation occurs in a concentration-independent manner with first order kinetics. This would be favourable because it is more robust than a concentration-dependent reaction, due to a lack of CCaMK-gene expression and protein level effects. As previously discussed, CCaMK is thought to be autophosphorylated at T271 at the basal, inactive state (Miller *et al.*, 2013). The presence of a stable, inactive state of CCaMK is important because it allows tight regulation of symbiosis. Downstream components leading to nodule organogenesis, for example, must be carefully regulated so that they are only induced in response to the correct signal, i.e. nuclear calcium spiking which peaks at 625-800 nm nuclear calcium (Ehrhardt *et al.*, 1996). Therefore, the favourable kinetics of the intra-oligomeric autophosphorylation reaction would help to enhance this inactive state stabilised by the hydrogen bond network, and therefore boost the stringent regulation observed during symbiotic signalling.

CCaMK undergoes a calcium-dependent conformational change

To study the conformational change that CCaMK undergoes when calcium ions bind to the EF-hands of the VLD, two complementary approaches were undertaken. The VLD construct was used for SAXS, which showed an elongation of the domain in the presence of calcium. Furthermore, the conformational change was also analysed using the full length MBP-tagged version of CCaMK. This revealed that a hydrophobic patch is exposed when calcium binds to the VLD. These data are consistent with previous data regarding the conformational change; however it also extends our understanding of this occurrence. The conformational change was initially characterised by mobility shift assay (Takezawa *et al.*, 1996), (Shimoda *et al.*, 2012) however there are potential errors with this method including the loss of bound calcium ions during gel electrophoresis. The conformational change was further analysed by Swainsbury and colleagues (Swainsbury *et al.*, 2012) who confirmed that it corresponded to a change in tertiary structure that also reveals a hydrophobic cleft. However, these data were obtained using constructs lacking either the kinase domain or both the kinase and AI domains, and therefore could only be assumed to apply to the protein as a whole. Therefore, the data in this thesis confirm the previously detected conformational change. Furthermore, it was demonstrated that the hydrophobic patch is exposed in the full-length protein and corresponds to an elongation of the VLD.

The conformational change has important potential implications for the function of CCaMK in the symbiotic pathway. First, this conformational change can be rationalised in terms of the function of the protein proposed by Miller and colleagues (Miller *et al.*, 2013) (Figure 1.9). Binding of calcium to the EF-hands has been shown to induce autophosphorylation at T271 (Takezawa *et al.*, 1996), and it follows that this could be facilitated by the aforementioned calcium-induced conformational change. This conformational change likely facilitates a structural rearrangement in the full-length protein that is required for intra-oligomeric autophosphorylation at T271. The phosphorylated form corresponds to the basal state of the protein. At higher calcium concentrations, it seems likely that the protein undergoes a further conformational change when Ca^{2+}/CaM binds to the protein, which disrupts the hydrogen bond network and allows activation for downstream signalling (Miller *et al.*, 2013). This would be a separate conformational change to that which was characterised above and future work would be required to determine the structural changes involved.

Furthermore, the exposure of a calcium-dependent hydrophobic cleft in the full-length CCaMK protein has important consequences for downstream signalling. Hydrophobic patches generally function in the binding of target proteins and this is observed in, for example, CaM binding to target proteins (Bhattacharya *et al.*, 2004). Since this patch exposure is observed with full-length CCaMK, it is likely that it does not facilitate interaction of CCaMK monomers. This evidence, combined with the data presented in Chapter 3, suggest the formation of an oligomer of CCaMK which is not modulated by calcium binding. This follows because the oligomeric state of MBP-CCaMK was addressed in the presence of EDTA, and the exposure of the hydrophobic patch in the presence of calcium suggests that the conformational change does not mediate further calcium-specific CCaMK:CCaMK interactions. Thus, the hydrophobic patch likely functions in target protein binding. One protein which binds to CCaMK is CaM itself, and theoretically the hydrophobic patch exposure could enhance CaM binding. This could explain how binding of CaM is observed to be 8-fold stronger after calcium-induced autophosphorylation (Sathyanarayanan *et al.*, 2000), a phenomenon which is thought to prime the protein for activation. However, CaM has a well-defined binding site in the CaM binding domain of CCaMK, and both the conformational change and hydrophobic patch exposure have been observed when only the VLD is present (Swainsbury *et al.*, 2012). Furthermore, this phenomenon was still observed, albeit reduced from 8-fold to 5-fold, when a peptide of CCaMK lacking the VLD was used (Sathyanarayanan *et al.*, 2001). Therefore if seems likely that the hydrophobic patch exposure does not contribute significantly to this phenomenon, even if the conformational change itself facilitates CaM binding.

Another binding protein of CCaMK is CYCLOPS. CYCLOPS has been shown to bind to CCaMK in the absence of symbiotic signalling and it is also a phosphorylation target of CCaMK (Yano *et al.*, 2008). Since calcium-induced autophosphorylation occurs in the basal state of the protein, the patch would therefore be exposed at basal calcium. Therefore the function could be to bind to CYCLOPS, because it is currently unknown as to whether CCaMK can bind in the absence of calcium. There are also other interaction proteins of CCaMK, for example Cip73, and there are likely to be more identified in the future. Therefore it is also plausible that the function could be to bind one of these target proteins.

Whilst the data obtained in this project regarding the calcium-dependent conformational change are robust, the SAXS envelopes are not ideal. This is because there are regions in

the SAXS envelope which lack density from the fitted homology model, showing that they are not a perfect match. The SAXS data is further limited because it provides an envelope structure only and doesn't show atomic-level resolution. It is important for this reason and, crucially, for the understanding of CCaMK during signalling, that a high resolution crystal structure of the protein is obtained. It would be ideal to obtain a crystal structure of the full-length protein; however previous work with CCaMK showed aggregation which would adversely affect its ability to crystallise (Swainsbury *et al.*, 2012). Furthermore, attempts to concentrate a full-length construct of CCaMK which shows minimal aggregation, such as MBP-CCaMK, were unsuccessful and so could not be used for crystal trials at this time. Therefore crystal trials were carried out for the VLD and a central peptide containing T271 and the CaM binding domain (Chapter 4).

Despite extensive attempts to produce crystals of these two constructs, none were obtained. Crystallisation of a protein requires the formation of precise inter-molecular interactions which allows the formation of a crystalline lattice. There are therefore a myriad of different factors which can affect a protein's ability to crystallise including, but not limited to, the temperature, precipitant, pH, buffer components, sample purity and homogeneity, metal ions, co-factors or ligands, aggregation and concentration used (Russo Krauss *et al.*, 2013). Whilst it was possible to vary some of these (e.g. metal ions and precipitants) it was obviously not possible to vary all possible factors which may have affected the constructs used. Therefore, under the conditions tested, it was not possible to crystallise constructs of CCaMK and this remains an important future goal.

Characterisation and purification trials of MtCCaMK

A study was undertaken in order to determine to what extent it is possible to extract and purify CCaMK from *M. truncatula*. A growth strategy for the commercial *M. truncatula* WT line, Jester, was first developed in order to obtain as much root tissue as possible. This hydroponic growth method was based on a method used to purify an unrelated oxalate oxidase from *Hordeum vulgare* (Requena and Bornemann, 1999) which yielded sufficient protein for biochemical studies. This method proved successful for the growth of large quantities of *M. truncatula* (20 000 seeds = 800 g root material) and furthermore roots were shown to be capable of forming nodules, suggesting that growth in this system was sufficiently similar to a terrestrial environment. In addition, it was also possible to develop

an antibody with strong affinity to CCaMK constructs expressed in *E. coli,* by using the AI-VLD protein as the antigen for immunization. This facilitated the evaluation of different purification strategies.

Many different purification strategies were tested in an attempt to yield MtCCaMK protein for characterisation. The first aim was to purify sufficient protein to obtain a clear band on a western blot when probed with αMtCCaMK to confirm the size of CCaMK *in planta*, which was first determined as approximately 56 kDa in *L. longiflorum* (Patil *et al.*, 1995). Subsequent experiments were attempted to determine if it is possible to purify sufficient CCaMK protein for biochemical experiments. The first purification strategy tested was based on that which was used in (Requena and Bornemann, 1999), however this strategy was for an apoplastic protein and did not prove successful for CCaMK (data not shown). It was therefore determined that a strategy which specifically included nuclear extraction would likely be more successful, due to the nuclear localisation of CCaMK (despite the lack of a canonical nuclear localisation signal) (Smit *et al.*, 2005), (Takeda *et al.*, 2012).

There are two main ways in which nuclear-localised proteins can be purified: isolation of nuclei and subsequent purification or whole cell extract purification. Both methods were trialled for the purification of MtCCaMK, with methods based on (Palma *et al.*, 2007) and (Backstrom *et al.*, 2007) respectively. The isolation of nuclei prior to purification did not produce a good protein yield with respect to CCaMK, as verified by western blotting, and therefore was not trialled further (data not shown). The final purification method utilised was based on the method utilised in (Backstrom *et al.*, 2007), with buffers derived from (Li *et al.*, 1996). This was chosen because the extraction was superior to those previously tested (including RIPA buffer, shown in Figure 5.3A) and because some CCaMK protein has been reported to be also present in the cytosol (Sylvia Singh, personal communication). In addition, buffers including non-ionic detergents were included in order to maintain protein-protein interactions, in case these stabilised the purification of CCaMK.

After these extensive purification trials, it was determined that it is possible to obtain a protein band on a western blot probed with αMtCCaMK after IP, which corresponds to full-length CCaMK protein at 58 kDa. This corresponds with the previously observed size of CCaMK. Further attempts to purify CCaMK, and obtain a corresponding protein band on SDS-PAGE stained with coomassie, were unsuccessful including the use of overexpression

plant lines. This is most likely due to the presence of CCaMK at extremely low levels *in planta,* which is consistent with detection of the full length protein during western blotting only after enrichment with IP and not in crude plant extracts.

Alternative splicing of CCaMK

Analysis of CCaMK at a variety of key developmental stages during the establishment of nodulation and mycorrhization revealed very low levels of splicing at the RNA level which were not reflected in corresponding protein splice forms. This suggests that whilst CCaMK is capable of being spliced this does not play a major role during symbiosis. This result was unexpected, because previous data from Pierre-Marc Delaux showed the detection of at least three different splice forms of *DMI3* during nodulation (Figure 6.2). Furthermore, the full-length *DMI3* transcript was less abundant than the three main alternatively spliced forms which were detected. There are several potential reasons for this discrepancy including the plant growth conditions, the reverse transcriptase and the particular PCR conditions used. It may, indeed, be a combination of these factors. However, for this project, a range of time-points during nodulation and mycorrhization were tested, with two WT lines (A17 and Jester), two different reverse transcriptases, different PCR conditions and different sources of cDNA. Crucially, the IP experiment showed that the splice variants which were detected at low levels at the RNA level were not reflected at the protein level for Jester (Figure 6.11A). Therefore, it seems likely that if significant splice variants were present during symbiosis they would have been detected by these methods and this indicates that splicing of CCaMK does not play a key role in symbiosis.

One potential type of splice variant which could have been missed during the experiments at the protein level is a variant containing the kinase domain only, which would be 33.3 kDa or smaller. This form would not have been detected during experiments at the RNA-level due the primers used. Attempts to utilise kinase domain-specific antibodies (αT271 and αpT271, Figure 6.12B-C) did not identify any kinase splice variants, and further did not identify the predominant, full-length CCaMK protein. It seems most likely that this is due to low affinity of the antibodies coupled with the low levels of the proteins in crude extract. A further possibility is that no full-length protein was detected with the αpT271 antibody because all the protein present was not phosphorylated at T271. This is a possibility because all samples were taken from roots at 21 days post infection, and therefore CCaMK

205

may have been fully de-phosphorylated and active for signalling at this point (Miller *et al.*, 2013). Alternatively, phosphorylation could have been removed by the action of a phosphatase during purification. Therefore, even if low levels were phosphorylated, these may have been below the threshold for detection by the αpT271 antibody. However, both the phosphorylated and non-phosphorylated forms of T271 should have been detected by the αT271 antibody, and therefore the most likely explanation is that the detection limit was too low for both antibodies. Hence, whilst it is not possible to rule out the presence of kinase only protein forms, the RT-PCR and IP data taken together strongly point to the presence of the predominantly full-length form of CCaMK at these symbiotic stages. Furthermore, these data were corroborated by RNA-Seq analysis which detected only the full length form in un-inoculated and roots inoculated with rhizobia. Finally, the coding sequence of *DMI3* has been shown to complement the *dmi3-1* mutant for symbiosis by myself (Chapter 5) and others including (Shimoda *et al.*, 2012). Therefore, the conclusion that splice variants do not play a major role during symbiosis is robust from the data presented.

In addition, the potential biological roles of the low-level splice forms detected are unclear. Form V, the only in-frame alternative transcript detected during the new nodulation and mycorrhization experiments, lacked a functional kinase domain but retained one EF-hand. Therefore, if present at the protein level this splice-form would not be an active kinase protein. It is possible, however, that it could play a role in calcium buffering in the nucleus, thereby potentially affecting the calcium spiking signal perceived by functional CCaMK. The other three splice-forms detected (II, III and IV) contained out-of-frame regions and therefore would likely not have a function at the protein level. Further evidence that the weakly expressed splice variants would not play a major biological role during symbiosis signalling is that the RNA forms detected and sequenced from the PCR with short extension times do not exactly match those previously sequenced by Pierre-Marc Delaux. It follows that if splice variants were essential for particular roles, it is likely that there would be evolutionary pressure to maintain them. Therefore the existence of a variety of splice variants does not suggest this evolutionary pressure.

Despite the lack of clear symbiotic roles of these low-level splice forms detected during symbiosis they could potentially have biological roles which were not identified from this study. For example, two of the main roles of alternative splicing are transcript regulation

or, if they are reflected in protein forms, they can act as dominant negative protein regulators which repress full-length CCaMK (Reddy *et al.*, 2013), (Staiger and Brown, 2013). Therefore, the splice forms identified could play a role in the fine-tuning of CCaMK regulation during symbiosis signalling even if they do not have clearly identifiable biological functions as discussed above. For example, the three forms identified with premature termination codons could be degraded by nonsense-mediated decay and therefore function in the regulation of transcript levels. However, crucially, in these particular experiments they were only detected at very low levels implying a non-essential function during symbiosis.

Interestingly, one reason why alternative splice-forms of CCaMK could have been detected at low levels is due to the dilution effect introduced by including multiple tissue types within each sample. For example, CCaMK has been shown to play roles in both infection and cell division during the establishment of nodulation, which occur in the epidermis and cortex respectively (Rival *et al.*, 2012). An attractive possibility is that differential splicing of CCaMK in different cell types could facilitate different cellular processes. This would be analogous to splicing seen in CaMKII, where differential splice forms of the β form which vary in linker length facilitates different responses to calcium spiking frequency in different cell types (Bayer *et al.*, 2002). However, none of the splice variants identified showed clear evidence of these subtle differences in the ability of CCaMK to perceive the calcium spiking signal, so this theory is speculation only.

A further possibility is that whilst splicing of CCaMK does not play a role in symbiosis signalling, it instead plays a role in the regulation of the protein in other pathways. For example, CCaMK was first identified in the anthers of *L. longiflorum* and *N. tabaccum.* CCaMK was detected throughout the development of anthers, from the pollen mother cells to the tetrad stage of meiosis, where protein levels peaked (Poovaiah *et al.*, 1999). Therefore, it is possible that the ability of CCaMK to be spliced could be crucial in alternate tissue types e.g. flowers instead of roots. Furthermore, other roles for CCaMK have recently been described, in particular during ABA signalling in *O. sativa, Z. mays* and *Triticum aestivum* (Shi *et al.*, 2014), (Ma *et al.*, 2012), (Yang *et al.*, 2011). In wheat, for example, the D-genome CCaMK was shown to contain upstream ABA-responsive elements, and was predominantly expressed in root tissue. In addition, ABA down-regulated CCaMK expression which indicates a potential role for CCaMK in stress responses (Yang *et al.*,

2011). It is therefore theoretically possible that the ability of CCaMK to be spliced in root tissue in fact enables it to function optimally in alternative pathways, for example stress responses.

Chapter 8: Outlook and Future Work

This project has addressed key questions relating to the function of CCaMK in the Sym pathway using a combination of approaches. An investigation into the autophosphorylation of CCaMK revealed that CCaMK forms an oligomer that likely facilitates intra-oligomeric autophosphorylation at T271. This study could be extended in a number of ways. First, since the oligomeric state was determined using MBP-tagged and *E. coli*-expressed CCaMK, it would be pertinent to validate this finding *in planta.* Extensive attempts to carry this out during this project as detailed above were unsuccessful, so this is not a trivial undertaking. An alternative to a purification strategy is to utilise fluorescence correlation spectroscopy, perhaps with a GFP-tagged CCaMK form, to determine the oligomeric state *in planta.* One of the best courses of action seems to be the exploration of alternative overexpression lines, however this has the additional risk that the overexpression of the protein may influence the oligomeric state of the protein by potentially causing aggregation. A further possibility, which is less attractive is the use of plant expression systems, for example *N. benthamiana.* However, the *HyperTrans* plant transient expression system has been previously attempted for CCaMK by Nicola Gilberthorpe and was unsuccessful.

Further interesting areas for future oligomeric state research regarding CCaMK include unpicking the involvement of CYCLOPS in the oligomer structure and identification of the residues necessary for oligomerisation. The interaction with CYCLOPS could be investigated by co-expression and co-purification by SEC. Identification of the CCaMK residues involved in the oligomerisation process present a challenge. However, we have evidence that the N-terminus may be involved because the AI-VLD is present as a monomer (Swainsbury *et al.,* 2012) and data by Liang Zhou showed that removal of the first five amino acids in the N-terminus disrupted oligomer formation. Furthermore, the kinase domain has been implicated in the interaction between CCaMK and CYCLOPS (Yano *et al.*, 2008). Therefore, future work should focus on the kinase domain for the identification of residues responsible for oligomerisation.

The data presented in this thesis strongly suggest that autophosphorylation at T271 occurs intra-oligomerically. However, this was not conclusively demonstrated due to the presumably unfavourable conformation formed in the oligomers of tagged and tagless protein. In order to determine if this is the case, the co-expressed mixed oligomer experiment could be repeated utilising different tags on the kinase dead protein. For example, this could be carried out using His-SUMO-CCaMK D192N protein instead of MBP-

CCaMK K47E. Alternatively, the question could be addressed kinetically by assessing the degree of autophosphorylation at T271 as a function of time and protein concentration, analogous to the experiments carried out in (Tirichine *et al.*, 2006) and (Takezawa *et al.*, 1996). The results from these experiments could be used to test mathematical models of the reaction to determine the mechanism. However, these experiments are challenging because they require precise control of the calcium concentrations in order to reach physiologically relevant calcium concentrations lower than 10 μM.

One of the most crucial experiments remaining for CCaMK is the elucidation of a complete structure. The importance of this has been mentioned in many recent studies, including work presented by Miller and colleagues (Miller *et al.*, 2013) and Shimoda and colleagues (Shimoda *et al.*, 2012). The challenge that this presents, however, cannot be overstated. There are obviously many, many different strategies which could be undertaken to attempt to produce crystals of CCaMK but some alternatives present more attractive options. For example, in order to obtain the crystal structure of the complete CaMKII oligomer, point mutations were made at key functional residues for example the autophosphorylation site and kinase-inactivating sites (Chao *et al.*, 2011). A similar strategy could be taken for CCaMK, and other potential mutation sites include inactivating mutations affecting the EF-hands and CaM binding domain. However, crystal trials utilising full-length CCaMK with strategic mutations have been initiated by Sylvia Singh, and have so far been unsuccessful. A further possibility for exploration is cryo-EM, which has been used recently for high resolution structures with no size-limit (Frauenfeld *et al.*, 2011). In addition, alternative truncations of the protein could be explored, perhaps for NMR structural determination (provided they are smaller than 40 kDa), which also yields atomic level resolution. However, the ultimate goal remains a structure of the full-length protein. Furthermore a characterisation of the conformational changes that the full-length protein undergoes by SAXS, in the presence of calcium and CaM, would be ideal. However, there are currently no suitable full-length candidate constructs.

Interestingly, whilst the data obtained for this thesis indicated that alternative splicing is not required for symbiosis, it does appear that low levels of CCaMK splice variants are produced at the RNA level. As previously suggested these could correspond to the function of CCaMK in different pathways, for example in ABA signalling. Therefore a similar study could be undertaken in different tissue types, during different developmental stages or

during stress responses in *M. truncatula* in order to determine if splicing of CCaMK is important in different contexts. However, since there is currently no strong evidence for this involvement this is unlikely to be a priority for future work.

In addition to the topics covered in this thesis, there are more unanswered questions relating to CCaMK. For example, certain elements of the model of CCaMK function described in (Miller *et al.*, 2013) remain to be confirmed *in planta,* for example that the basal and inactive state of CCaMK is with phosphorylation at T271. Furthermore, the roles of the other phosphorylation sites in CCaMK (including S343 and S344) require further clarification *in planta.* The hydrogen-bonding network has thus-far been identified by homology modelling alone, despite its crucial predicted role in the mechanism of CCaMK inactivation. Therefore, structural confirmation of the relative positions of the amino acids involved in this network is necessary to confirm its importance. In addition, since phosphorylation is crucial to the mechanism of CCaMK function, it follows that one or more phosphatases may be involved in the system, which have yet to be identified. Finally, it seems likely that further interaction proteins of CCaMK exist which perhaps serve to specifically transduce the signal to produce either nodules or arbuscules. Therefore this will also be an important goal. A wider question regarding the Sym pathway, and potentially CCaMK, remains as to how specificity is conferred in the production of either arbuscules or nodules. This could potentially be mediated by CCaMK; but other possibilities include a parallel pathway and downstream pathway elements.

CCaMK is, and remains, an extremely important but challenging protein for study. A full and thorough understanding of CCaMK will become increasingly important as we progress towards the goal of engineering nitrogen-fixing crop plants. However, it is also a protein worthy of study in its own right due to the important role it plays in this pathway and the fascinating complexity surrounding its mode of action.

References

Akiyama, K., Matsuzaki, K. and Hayashi, H. (2005). Plant sesquiterpenes induce hyphal branching in arbuscular mycorrhizal fungi. Nature **435**: 824-827.

Allen, G.J., Chu, S.P., Harrington, C.L., Schumacher, K., Hoffman, T., Tang, Y.Y., Grill, E. and Schroeder, J.I. (2001). A defined range of guard cell calcium oscillation parameters encodes stomatal movements. Nature **411**: 1053-1057.

Allen, G.J., Chu, S.P., Schumacher, K., Shimazaki, C.T., Vafeados, D., Kemper, A., Hawke, S.D., Tallman, G., Tsien, R.Y., Harper, J.F., Chory, J. and Schroeder, J.I. (2000). Alteration of stimulus-specific guard cell calcium oscillations and stomatal closing in *Arabidopsis det3* mutant. Science **289**: 2338-2342.

Ames, J.B. and Lim, S. (2012). Molecular structure and target recognition of neuronal calcium sensor proteins. Biochimica Et Biophysica Acta **1820**: 1205-1213.

Ané, J.M., Kiss, G.B., Riely, B.K., Penmetsa, R.V., Oldroyd, G.E.D., Ayax, C., Levy, J., Debelle, F., Baek, J.M., Kalo, P., Rosenberg, C., Roe, B.A., Long, S.R., Denarie, J. and Cook, D.R. (2004). *Medicago truncatula DMI1* required for bacterial and fungal symbioses in legumes. Science **303**: 1364-1367.

Antolín-Llovera, M., Ried, M.K. and Parniske, M. (2014). Cleavage of the SYMBIOSIS RECEPTOR-LIKE KINASE Ectodomain Promotes Complex Formation with Nod Factor Receptor 5. Curr. Biol. **24**: 422-427.

Arrighi, J.-F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L.C., Mirabella, R., de Carvalho-Niebel, F., Journet, E.-P., Gherardi, M., Huguet, T., Geurts, R., Denarie, J., Rouge, P. and Gough, C. (2006). The *Medicago truncatula* Lysine Motif-Receptor-Like Kinase Gene Family Includes *NFP* and New Nodule-Expressed Genes. Plant Physiol. **142**: 265-279.

Backstrom, S., Elfving, N., Nilsson, R., Wingsle, G. and Bjorklund, S. (2007). Purification of a plant mediator from *Arabidopsis thaliana* identifies PFT1 as the Med25 subunit. Mol. Cell **26**: 717-729.

Baek, J.-M., Han, P., Iandolino, A. and Cook, D.R. (2008). Characterization and comparison of intron structure and alternative splicing between *Medicago truncatula*, *Populus trichocarpa*, *Arabidopsis* and rice. Plant Mol. Biol. **67**: 499-510.

Banba, M., Gutjahr, C., Miyao, A., Hirochika, H., Paszkowski, U., Kouchi, H. and Imaizumi-Anraku, H. (2008). Divergence of Evolutionary Ways Among Common *sym* Genes: CASTOR and CCaMK Show Functional Conservation Between Two Symbiosis Systems and Constitute the Root of a Common Signaling Pathway. Plant and Cell Physiology **49**: 1659-1671.

Barker, D.G., Bianchi, S., Blondon, F., Dattee, Y., Duc, G., Essad, S., Flament, P., Gallusci, P., Genier, G., Guy, P., Muel, X., Tourneur, J., Denarie, J. and Huguet, T. (1990). *Medicago truncatula,* a model plant for studying the molecular genetics of the *Rhizobium*-legume symbiosis. Plant Mol. Biol. Rep. **8**: 40-49.

Batistic, O. and Kudla, J. (2012). Analysis of calcium signaling pathways in plants. Biochimica Et Biophysica Acta **1820**: 1283-1293.

Bayer, K.U., De Koninck, P. and Schulman, H. (2002). Alternative splicing modulates the frequency-dependent response of CaMKII to Ca^{2+} oscillations. EMBO J. **21**: 3590-3597.

Berridge, M.J., Bootman, M.D. and Lipp, P. (1998). Calcium - a life and death signal. Nature **395**: 645-648.

Besserer, A., Puech-Pages, V., Kiefer, P., Gomez-Roldan, V., Jauneau, A., Roy, S., Portais, J.-C., Roux, C., Becard, G. and Sejalon-Delmas, N. (2006). Strigolactones stimulate arbuscular mycorrhizal fungi by activating mitochondria. PLoS Biol. **4**: 1239-1247.

Bhattacharya, S., Bunick, C.G. and Chazin, W.J. (2004). Target selectivity in EF-hand calcium binding proteins. Biochimica Et Biophysica Acta **1742**: 69-79.

Bonfante, P. and Requena, N. (2011). Dating in the dark: how roots respond to fungal signals to establish arbuscular mycorrhizal symbiosis. Curr. Opin. Plant Biol. **14**: 451-457.

Boscari, A., del Giudice, J., Ferrarini, A., Venturini, L., Zaffini, A.-L., Delledonne, M. and Puppo, A. (2013). Expression Dynamics of the *Medicago truncatula* Transcriptome during the Symbiotic Interaction with *Sinorhizobium meliloti*: Which Role for Nitric Oxide? Plant Physiol. **161**: 425-439.

Capoen, W., Den Herder, J., Sun, J.H., Verplancke, C., De Keyser, A., De Rycke, R., Goormachtig, S., Oldroyd, G. and Holsters, M. (2009). Calcium Spiking Patterns and the Role of the Calcium/Calmodulin-Dependent Kinase CCaMK in Lateral Root Base Nodulation of *Sesbania rostrata*. Plant Cell **21**: 1526-1540.

Capoen, W. and Oldroyd, G. (2008). How CYCLOPS keeps an eye on plant symbiosis. Proc. Natl. Acad. Sci. USA **105**: 20053-20054.

Capoen, W., Sun, J., Wysham, D., Otegui, M.S., Venkateshwaran, M., Hirsch, S., Miwa, H., Downie, J.A., Morris, R.J., Ane, J.-M. and Oldroyd, G.E.D. (2011). Nuclear membranes control symbiotic calcium signaling of legumes. Proc. Natl. Acad. Sci. USA **108**: 14348-14353.

Catoira, R., Galera, C., de Billy, F., Penmetsa, R.V., Journet, E.P., Maillet, F., Rosenberg, C., Cook, D., Gough, C. and Denarie, J. (2000). Four genes of *Medicago truncatula* Controlling Components of a Nod Factor Transduction Pathway. Plant Cell **12**: 1647-1665.

Catoira, R., Timmers, A.C.J., Maillet, F., Galera, C., Penmetsa, R.V., Cook, D., Denarie, J. and Gough, C. (2001). The *HCL* gene of *Medicago truncatula* controls *Rhizobium*-induced root hair curling. Development **128**: 1507-1518.

Cerri, M.R., Frances, L., Laloum, T., Auriac, M.-C., Niebel, A., Oldroyd, G.E.D., Barker, D.G., Fournier, J. and de Carvalho-Niebel, F. (2012). *Medicago truncatula* ERN Transcription Factors: Regulatory Interplay with NSP1/NSP2 GRAS Factors and Expression Dynamics throughout Rhizobial Infection. Plant Physiol. **160**: 2155-2172.

Chabaud, M., Genre, A., Sieberer, B.J., Faccio, A., Fournier, J., Novero, M., Barker, D.G. and Bonfante, P. (2011). Arbuscular mycorrhizal hyphopodia and germinated spore exudates trigger Ca^{2+} spiking in the legume and nonlegume root epidermis. New Phytol. **189**: 347-355.

Chao, L.H., Stratton, M.M., Lee, I.-H., Rosenberg, O.S., Levitz, J., Mandell, D.J., Kortemme, T., Groves, J.T., Schulman, H. and Kuriyan, J. (2011). A Mechanism for Tunable Autoinhibition in the Structure of a Human Ca^{2+}/Calmodulin-Dependent Kinase II Holoenzyme Cell **146**: 732-745.

Charpentier, M., Bredemeier, R., Wanner, G., Takeda, N., Schleiff, E. and Parniske, M. (2008). *Lotus japonicus* CASTOR and POLLUX Are Ion Channels Essential for Perinuclear Calcium Spiking in Legume Root Endosymbiosis. Plant Cell **20**: 3467-3479.

Charpentier, M., Martins, T.V., Granqvist, E., Oldroyd, G. and Morris, R. (2013). The role of DMI1 in establishing Ca^{2+} oscillations in legume symbioses. Plant Signal. Behav. **8**: 1-3.

Charpentier, M. and Oldroyd, G. (2010). How close are we to nitrogen-fixing cereals? Curr. Opin. Plant Biol. **13**: 556-564.

Charpentier, M. and Oldroyd, G.E.D. (2013). Nuclear Calcium Signaling in Plants. Plant Physiol. **163**: 496-503.

Chehab, E.W., Patharkar, O.R., Hegeman, A.D., Taybi, T. and Cushman, J.C. (2004). Autophosphorylation and Subcellular Localization Dynamics of a Salt- and Water

Deficit-Induced Calcium-Dependent Protein Kinase from Ice Plant. Plant Physiol. **135**: 1430-1446.

Colbran, R.J. and Soderling, T.R. (1990). Calcium/Calmodulin-independent Autophosphorylation Sites of Calcium/Calmodulin-dependent Protein Kinase II. J. Biol. Chem. **265**: 11213-11219.

Cook, D.R. (1999). *Medicago truncatula* - a model in the making! Commentary. Curr. Opin. Plant Biol. **2**: 301-304.

Dammann, C., Ichida, A., Hong, B.M., Romanowsky, S.M., Hrabak, E.M., Harmon, A.C., Pickard, B.G. and Harper, J.F. (2003). Subcellular Targeting of Nine Calcium-Dependent Protein Kinase Isoforms from Arabidopsis. Plant Physiol. **132**: 1840-1848.

Day, I.S., Reddy, V.S., Shad Ali, G. and Reddy, A.S.N. (2002). Analysis of EF-hand-containing proteins in *Arabidopsis*. Genome biology **3**: 1-24.

De Koninck, P. and Schulman, H. (1998). Sensitivity of CaM kinase II to the frequency of Ca^{2+} oscillations. Science **279**: 227-230.

DeFalco, T.A., Bender, K.W. and Snedden, W.A. (2010). Breaking the code: Ca^{2+} sensors in plant signalling. Biochem. J. **425**: 27-40.

Delaux, P.-M., Becard, G. and Combier, J.-P. (2013a). NSP1 is a component of the Myc signaling pathway. New Phytol. **199**: 59-65.

Delaux, P.-M., Radhakrishnan, G. and Oldroyd, G.E.D. (2015). Tracing the evolutionary path to nitrogen-fixing crops. Curr. Opin. Plant Biol. **26**: 95-99.

Delaux, P.-M., Séjalon-Delmas, N., Bécard, G. and Ané, J.-M. (2013b). Evolution of the plant-microbe symbiotic 'toolkit'. Trends Plant Sci. **18**: 298-304.

Dolmetsch, R.E., Xu, K.L. and Lewis, R.S. (1998). Calcium oscillations increase the efficiency and specificity of gene expression. Nature **392**: 933-936.

Ehrhardt, D.W., Atkinson, E.M. and Long, S.R. (1992). Depolarization of Alfalfa Root Hair Membrane Potential by *Rhizobium meliloti* Nod Factors. Science **256**: 998-1000.

Ehrhardt, D.W., Wais, R. and Long, S.R. (1996). Calcium spiking in plant root hairs responding to Rhizobium nodulation signals. Cell **85**: 673-681.

Endicott, J.A., Noble, M.E.M. and Johnson, L.N. (2012). The Structural Basis for Control of Eukaryotic Protein Kinases. Annu. Rev. Biochem. **81**: 587-613.

Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kalo, P. and Kiss, G.B. (2002). A receptor kinase gene regulating symbiotic nodule development. Nature **417**: 962-966.

Engler, C., Gruetzner, R., Kandzia, R. and Marillonnet, S. (2009). Golden Gate Shuffling: A One-Pot DNA Shuffling Method Based on Type IIs Restriction Enzymes. Plos One **4**: 1-9.

Evans, N.H., McAinsh, M.R. and Hetherington, A.M. (2001). Calcium oscillations in higher plants. Curr. Opin. Plant Biol. **4**: 415-420.

Felle, H.H., Kondorosi, E., Kondorosi, A. and Schultze, M. (1999). Elevation of the Cytosolic Free $[Ca^{2+}]$ Is Indispensable for the Transduction of the Nod Factor Signal in Alfalfa. Plant Physiol. **121**: 273-279.

Frauenfeld, J., Gumbart, J., van der Sluis, E.O., Funes, S., Gartmann, M., Beatrix, B., Mielke, T., Berninghausen, O., Becker, T., Schulten, K. and Beckmann, R. (2011). Cryo-EM structure of the ribosome-SecYE complex in the membrane environment. Nat. Struct. Mol. Biol. **18**: 614-621.

Genre, A., Chabaud, M., Balzergue, C., Puech-Pages, V., Novero, M., Rey, T., Fournier, J., Rochange, S., Becard, G., Bonfante, P. and Barker, D.G. (2013). Short-chain chitin oligomers from arbuscular mycorrhizal fungi trigger nuclear Ca^{2+} spiking in *Medicago truncatula* roots and their production is enhanced by strigolactone. New Phytol. **198**: 179-189.

216

Geurts, R. and Bisseling, T. (2002). Rhizobium nod factor perception and signalling. Plant Cell **14**: S239-S249.

Gifford, J.L., Walsh, M.P. and Vogel, H.J. (2007). Structures and metal-ion-binding properties of the Ca^{2+}-binding helix-loop-helix EF-hand motifs. Biochem. J. **405**: 199-221.

Giovannetti, M. and Mosse, B. (1980). An evaluation of techniques for measuring vesicular arbuscular mycorrhizal infection in roots. New Phytol. **84**: 489-500.

Giraud, E., Moulin, L., Vallenet, D., Barbe, V., Cytryn, E., Avarre, J.-C., Jaubert, M., Simon, D., Cartieaux, F., Prin, Y., Bena, G., Hannibal, L., Fardoux, J., Kojadinovic, M., Vuillet, L., Lajus, A., Cruveiller, S., Rouy, Z., Mangenot, S., Segurens, B., Dossat, C., Franck, W.L., Chang, W.-S., Saunders, E., Bruce, D., Richardson, P., Normand, P., Dreyfus, B., Pignol, D., Stacey, G., Emerich, D., Vermeglio, A., Medigue, C. and Sadowsky, M. (2007). Legumes Symbioses: Absence of *Nod* Genes in Photosynthetic Bradyrhizobia. Science **316**: 1307-1312.

Gleason, C., Chaudhuri, S., Yang, T.B., Munoz, A., Poovaiah, B.W. and Oldroyd, G.E.D. (2006). Nodulation independent of rhizobia induced by a calcium-activated kinase lacking autoinhibition. Nature **441**: 1149-1152.

Gobbato, E., Marsh, J.F., Vernie, T., Wang, E., Maillet, F., Kim, J., Miller, J.B., Sun, J., Bano, S.A., Ratet, P., Mysore, K.S., Denarie, J., Schultze, M. and Oldroyd, G.E.D. (2012). A GRAS-Type Transcription Factor with a Specific Function in Mycorrhizal Signaling. Curr. Biol. **22**: 2236-2241.

Godfroy, O., Debelle, F., Timmers, T. and Rosenberg, C. (2006). A rice Calcium- and Calmodulin-Dependent Protein Kinase Restores Nodulation to a Legume Mutant. Mol. Plant-Microbe Interact. **19**: 495-501.

Goldschmidt, L., Cooper, D.R., Derewenda, Z.S. and Eisenberg, D. (2007). Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. Protein Sci. **16**: 1569-1576.

Granqvist, E., Sun, J., Op den Camp, R., Pujic, P., Hill, L., Normand, P., Morris, R.J., Downie, J.A., Geurts, R. and Oldroyd, G.E.D. (2015). Bacterial-induced calcium oscillations are common to nitrogen-fixing associations of nodulating legumes and non-legumes. New Phytol. **207**: 551-558.

Granqvist, E., Wysham, D., Hazledine, S., Kozlowski, W., Sun, J., Charpentier, M., Martins, T.V., Haleux, P., Tsaneva-Atanasova, K., Downie, J.A., Oldroyd, G.E.D. and Morris, R.J. (2012). Buffering Capacity Explains Signal Variation in Symbiotic Calcium Oscillations. Plant Physiol. **160**: 2300-2310.

Groth, M., Takeda, N., Perry, J., Uchida, H., Draexl, S., Brachmann, A., Sato, S., Tabata, S., Kawaguchi, M., Wang, T.L. and Parniske, M. (2010). *NENA*, a *Lotus japonicus* Homolog of *Sec13*, Is Required for Rhizodermal Infection by Arbuscular Mycorrhiza Fungi and Rhizobia but Dispensable for Cortical Endosymbiotic Development. Plant Cell **22**: 2509-2526.

Hanson, P.I., Meyer, T., Stryer, L. and Schulman, H. (1994). Dual role of calmodulin in autophosphorylation of multifunctional cam kinase may underlie decoding of calcium signals. Neuron **12**: 943-956.

Hanson, P.I. and Schulman, H. (1992). Inhibitory Autophosphorylation of Multifunctional Ca^{2+}/Calmodulin-dependent Protein Kinase Analyzed by Site-directed Mutagenesis. J. Biol. Chem. **267**: 17216-17224.

Harper, J.F. and Harmon, A. (2005). Plants, symbiosis and parasites: A calcium signalling connection. Nat. Rev. Mol. Cell Biol. **6**: 555-566.

Harrison, M.J. (2005). Signaling in the arbuscular mycorrhizal symbiosis. Annu. Rev. Microbiol. **59**: 19-42.

Hayashi, T., Banba, M., Shimoda, Y., Kouchi, H., Hayashi, M. and Imaizumi-Anraku, H. (2010). A dominant function of CCaMK in intracellular accommodation of bacterial and fungal endosymbionts. Plant J. **63**: 141-154.

Hazledine, S., Sun, J., Wysham, D., Downie, J.A., Oldroyd, G.E.D. and Morris, R.J. (2009). Nonlinear Time Series Analysis of Nodulation Factor Induced Calcium Oscillations: Evidence for Deterministic Chaos? Plos One **4**: 1-10.

Heidarsson, P.O., Bjerrum-Bohr, I.J., Jensen, G.A., Pongs, O., Finn, B.E., Poulsen, F.M. and Kragelund, B.B. (2012). The C-Terminal Tail of Human Neuronal Calcium Sensor 1 Regulates the Conformational Stability of the Ca^{2+}-Activated State. J. Mol. Biol. **417**: 51-64.

Hetherington, A.M. and Brownlee, C. (2004). The generation of Ca^{2+} signals in plants. Annu. Rev. Plant Biol. **55**: 401-427.

Hill, J.R. (2000). Jester. Plant Varieties Journal **13**: 40-40.

Hirsch, S., Kim, J., Munoz, A., Heckmann, A.B., Downie, J.A. and Oldroyd, G.E.D. (2009). GRAS Proteins Form a DNA Binding Complex to Induce Gene Expression during Nodulation Signaling in *Medicago truncatula*. Plant Cell **21**: 545-557.

Hodge, A., Campbell, C.D. and Fitter, A.H. (2001). An arbuscular mycorrhizal fungus accelerates decomposition and acquires nitrogen directly from organic material. Nature **413**: 297-299.

Hoeflich, K.P. and Ikura, M. (2002). Calmodulin in Action: Diversity in Target Recognition and Activation Mechanisms. Cell **108**: 739-742.

Hoelz, A., Nairn, A.C. and Kuriyan, J. (2003). Crystal Structure of a Tetradecameric Assembly of the Association Domain of Ca^{2+}/Calmodulin-Dependent Kinase II. Mol. Cell **11**: 1241-1251.

Hoffmann, B., Trinh, T.H., Leung, J., Kondorosi, A. and Kondorosi, E. (1997). A New *Medicago truncatula* Line with Superior in Vitro Regeneration, Transformation, and Symbiotic Properties Isolated Through Cell Culture Selection. Mol. Plant-Microbe Interact. **10**: 307-315.

Horvath, B., Yeun, L.H., Domonkos, A., Halasz, G., Gobbato, E., Ayaydin, F., Miro, K., Hirsch, S., Sun, J., Tadege, M., Ratet, P., Mysore, K.S., Ane, J.-M., Oldroyd, G.E.D. and Kalo, P. (2011). *Medicago truncatula IPD3* Is a Member of the Common Symbiotic Signaling Pathway Required for Rhizobial and Mycorrhizal Symbioses. Mol. Plant-Microbe Interact. **24**: 1345-1358.

Hudmon, A. and Schulman, H. (2002a). Neuronal Ca^{2+}/Calmodulin-Dependent Protein Kinase II: The Role of Structure and Autoregulation in Cellular Function. Annu. Rev. Biochem. **71**: 473-510.

Hudmon, A. and Schulman, H. (2002b). Structure-function of the multifunctional Ca^{2+}/calmodulin-dependent protein kinase II. Biochem. J. **364**: 593-611.

Johnson, C.H., Knight, M.R., Kondo, T., Masson, P., Sedbrook, J., Haley, A. and Trewavas, A. (1995). Circadian Oscillations of Cytosolic and Chloroplastic Free Calcium in Plants. Science **269**: 1863-1865.

Journet, E.P., El-Gachtouli, N., Vernoud, V., de Billy, F., Pichon, M., Dedieu, A., Arnould, C., Morandi, D., Barker, D.G. and Gianinazzi-Pearson, V. (2001). *Medicago truncatula ENOD11*: A Novel RPRP-Encoding Early Nodulin Gene Expressed During Mycorrhization in Arbuscule-Containing Cells. Mol. Plant-Microbe Interact. **14**: 737-748.

Kalo, P., Gleason, C., Edwards, A., Marsh, J., Mitra, R.M., Hirsch, S., Jakab, J., Sims, S., Long, S.R., Rogers, J., Kiss, G.B., Downie, J.A. and Oldroyd, G.E.D. (2005). Nodulation signaling in legumes requires NSP2, a member of the GRAS family of transcriptional regulators. Science **308**: 1786-1789.

Kamphuis, L.G., Lichtenzveig, J., Peng, K., Guo, S.-M., Klingler, J.P., Siddique, K.H.M., Gao, L.-L. and Singh, K.B. (2013). Characterization and genetic dissection of resistance to spotted alfalfa aphid (*Therioaphis trifolii*) in *Medicago truncatula*. J. Exp. Bot. **64**: 5157-5172.

Kang, H., Xiao, A., Huang, X., Gao, X., Yu, H., He, X., Zhu, H., Hong, Z. and Zhang, Z. (2015). A *Lotus japonicus* Cochaperone Protein Interacts With the Ubiquitin-Like Domain Protein CIP73 and Plays a Negative Regulatory Role in Nodulation. Mol. Plant-Microbe Interact. **28**: 534-545.

Kang, H., Zhu, H., Chu, X.J., Yang, Z.Z., Yuan, S.L., Yu, D.Q., Wang, C., Hong, Z.L. and Zhang, Z.M. (2011). A Novel Interaction between CCaMK and a Protein Containing the Scythe_N Ubiquitin-Like Domain in *Lotus japonicus*. Plant Physiol. **155**: 1312-1324.

Kapust, R.B. and Waugh, D.S. (1999). *Escherichia coli* maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. Protein Sci. **8**: 1668-1674.

Kawaharada, Y., Kelly, S., Nielsen, M.W., Hjuler, C.T., Gysel, K., Muszynski, A., Carlson, R.W., Thygesen, M.B., Sandal, N., Asmussen, M.H., Vinther, M., Andersen, S.U., Krusell, L., Thirup, S., Jensen, K.J., Ronson, C.W., Blaise, M., Radutoiu, S. and Stougaard, J. (2015). Receptor-mediated exopolysaccharide perception controls bacterial infection. Nature **523**: 308-312.

Kistner, C. and Parniske, M. (2002). Evolution of signal transduction in intracellular symbiosis. Trends Plant Sci. **7**: 511-518.

Kistner, C., Winzer, T., Pitzschke, A., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Webb, K.J., Szczyglowski, K. and Parniske, M. (2005). Seven *Lotus japonicus* genes required for transcriptional reprogramming of the root during fungal and bacterial symbiosis. Plant Cell **17**: 2217-2229.

Kolodziej, S.J., Hudmon, A., Waxham, M.N. and Stoops, J.K. (2000). Three-Dimensional Reconstructions of Calcium/Calmodulin-dependent (CaM) Kinase IIα and Truncated CaM kinase IIα Reveal a Unique Organization for Its Structural Core and Functional Domains. J. Biol. Chem. **275**: 14354-14359.

Kosuta, S., Hazledine, S., Sun, J., Miwa, H., Morris, R.J., Downie, J.A. and Oldroyd, G.E.D. (2008). Differential and chaotic calcium signatures in the symbiosis signaling pathway of legumes. Proc. Natl. Acad. Sci. USA **105**: 9823-9828.

Kushwaha, R., Singh, A. and Chattopadhyay, S. (2008). Calmodulin7 plays an important role as transcriptional regulator in *Arabidopsis* seedling development. Plant Cell **20**: 1747-1759.

Lagunas, B., Schaefer, P. and Gifford, M.L. (2015). Housing helpful invaders: the evolutionary and molecular architecture underlying plant root-mutualist microbe interactions. J. Exp. Bot. **66**: 2177-2186.

Laloum, T., Baudin, M., Frances, L., Lepage, A., Billault-Penneteau, B., Cerri, M.R., Ariel, F., Jardinaud, M.-F., Gamas, P., de Carvalho-Niebel, F. and Niebel, A. (2014). Two CCAAT-box-binding transcription factors redundantly regulate early steps of the legume-rhizobia endosymbiosis. Plant J. **79**: 757-768.

Lefebvre, B., Timmers, T., Mbengue, M., Moreau, S., Herve, C., Toth, K., Bittencourt-Silvestre, J., Klaus, D., Deslandes, L., Godiard, L., Murray, J.D., Udvardi, M.K., Raffaele, S., Mongrand, S., Cullimore, J., Gamas, P., Niebel, A. and Ott, T. (2010). A remorin protein interacts with symbiotic receptors and regulates bacterial infection. Proc. Natl. Acad. Sci. USA **107**: 2343-2348.

Lerouge, P., Roche, P., Faucher, C., Maillet, F., Truchet, G., Prome, J.C. and Denarie, J. (1990). Symbiotic host-specificity of *Rhizobium meliloti* is determined by a sulphated and acylated glucosamine oligosaccharide signal. Nature **344**: 781-784.

Levy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E.P., Ane, J.M., Lauber, E., Bisseling, T., Denarie, J., Rosenberg, C. and Debelle, F. (2004). A Putative Ca^{2+} and Calmodulin-Dependent Protein Kinase Required for Bacterial and Fungal Symbioses. Science **303**: 1361-1364.

Li, Y., Bjorklund, S., Kim, Y.J. and Kornberg, R.D. (1996). Yeast RNA polymerase II holoenzyme. Methods Enzymol. **273**: 172-176.

Liang, Y., Toth, K., Cao, Y., Tanaka, K., Espinoza, C. and Stacey, G. (2014). Lipochitooligosaccharide recognition: an ancient story. New Phytol. **204**: 289-296.

Liao, J., Singh, S., Hossain, M.S., Andersen, S.U., Ross, L., Bonetta, D., Zhou, Y., Sato, S., Tabata, S., Stougaard, J., Szczyglowski, K. and Parniske, M. (2012). Negative regulation of CCaMK is essential for symbiotic infection. Plant J. **72**: 572-584.

Limpens, E. and Bisseling, T. (2014). CYCLOPS: A New Vision on Rhizobium-Induced Nodule Organogenesis. Cell Host & Microbe **15**: 127-129.

Limpens, E., Franken, C., Smit, P., Willemse, J., Bisseling, T. and Geurts, R. (2003). LysM Domain Receptor Kinases Regulating Rhizobial Nod Factor-Induced Infection. Science **302**: 630-633.

Lisman, J., Yasuda, R. and Raghavachari, S. (2012). Mechanisms of CaMKII action in long-term potentiation. Nat. Rev. Neurosci. **13**: 169-182.

Liu, Z.H., Xia, M. and Poovaiah, B.W. (1998). Chimeric calcium/calmodulin-dependent protein kinase in tobacco: differential regulation by calmodulin isoforms. Plant Mol. Biol. **38**: 889-897.

Lodwig, E.M., Hosie, A.H.F., Bordes, A., Findlay, K., Allaway, D., Karunakaran, R., Downie, J.A. and Poole, P.S. (2003). Amino-acid cycling drives nitrogen fixation in the legume-*Rhizobium* symbiosis. Nature **422**: 722-726.

Logan, D.C. and Knight, M.R. (2003). Mitochondrial and Cytosolic Calcium Dynamics Are Differentially Regulated in Plants. Plant Physiol. **133**: 21-24.

Long, S.R. (1996). *Rhizobium* symbiosis: Nod factors in perspective. Plant Cell **8**: 1885-1898.

Ma, F., Lu, R., Liu, H., Shi, B., Zhang, J., Tan, M., Zhang, A. and Jiang, M. (2012). Nitric oxide-activated calcium/calmodulin-dependent protein kinase regulates the abscisic acid-induced antioxidant defence in maize. J. Exp. Bot. **63**: 4835-4847.

Maillet, F., Poinsot, V., Andre, O., Puech-Pages, V., Haouy, A., Gueunier, M., Cromer, L., Giraudet, D., Formey, D., Niebel, A., Martinez, E.A., Driguez, H., Becard, G. and Denarie, J. (2011). Fungal lipochitooligosaccharide symbiotic signals in arbuscular mycorrhiza. Nature **469**: 58-63.

Marsh, J.F., Rakocevic, A., Mitra, R.M., Brocard, L., Sun, J., Eschstruth, A., Long, S.R., Schultze, M., Ratet, P. and Oldroyd, G.E.D. (2007). *Medicago truncatula NIN* is essential for rhizobial-independent nodule organogenesis induced by autoactive calcium/calmodulin-dependent protein kinase. Plant Physiol. **144**: 324-335.

McAinsh, M.R. and Pittman, J.K. (2009). Shaping the calcium signature. New Phytol. **181**: 275-294.

Messinese, E., Mun, J.H., Yeun, L.H., Jayaraman, D., Rouge, P., Barre, A., Lougnon, G., Schornack, S., Bono, J.J., Cook, D.R. and Ane, J.M. (2007). A Novel Nuclear Protein Interacts With the Symbiotic DMI3 Calcium- and Calmodulin-Dependent Protein Kinase of *Medicago truncatula*. Mol. Plant-Microbe Interact. **20**: 912-921.

Miller, J.B., Pratap, A., Miyahara, A., Zhou, L., Bornemann, S., Morris, R.J. and Oldroyd, G.E.D. (2013). Calcium/Calmodulin-Dependent Protein Kinase Is Negatively and Positively Regulated by Calcium, Providing a Mechanism for Decoding Calcium Responses during Symbiosis Signaling. Plant Cell **25**: 5053-5066.

Mitra, R.M., Gleason, C.A., Edwards, A., Hadfield, J., Downie, J.A., Oldroyd, G.E.D. and Long, S.R. (2004). A Ca^{2+}/calmodulin-dependent protein kinase required for

symbiotic nodule development: Gene identification by transcript-based cloning. Proc. Natl. Acad. Sci. USA **101**: 4701-4705.

Miwa, H., Sun, J., Oldroyd, G.E.D. and Downie, J.A. (2006a). Analysis of calcium spiking using a cameleon calcium sensor reveals that nodulation gene expression is regulated by calcium spike number and the developmental status of the cell. Plant J. **48**: 883-894.

Miwa, H., Sun, J., Oldroyd, G.E.D. and Downie, J.A. (2006b). Analysis of Nod-Factor-Induced Calcium Signaling in Root Hairs of Symbiotically Defective Mutants of *Lotus japonicus*. Mol. Plant-Microbe Interact. **19**: 914-923.

Miyata, K., Kozaki, T., Kouzai, Y., Ozawa, K., Ishii, K., Asamizu, E., Okabe, Y., Umehara, Y., Miyamoto, A., Kobae, Y., Akiyama, K., Kaku, H., Nishizawa, Y., Shibuya, N. and Nakagawa, T. (2014). The Bifunctional Plant Receptor, OsCERK1, Regulates Both Chitin-Triggered Immunity and Arbuscular Mycorrhizal Symbiosis in Rice. Plant and Cell Physiology **55**: 1864-1872.

Moling, S., Pietraszewska-Bogiel, A., Postma, M., Fedorova, E., Hink, M.A., Limpens, E., Gadella, T.W.J. and Bisseling, T. (2014). Nod Factor Receptors Form Heteromeric Complexes and Are Essential for Intracellular Infection in *Medicago* Nodules. Plant Cell **26**: 4188-4199.

Mori, I.C., Murata, Y., Yang, Y., Munemasa, S., Wang, Y.-F., Andreoli, S., Tiriac, H., Alonso, J.M., Harper, J.F., Ecker, J.R., Kwak, J.M. and Schroeder, J.I. (2006). CDPKs CPK6 and CPK3 function in ABA regulation of guard cell S-type anion- and Ca^{2+}-permeable channels and stomatal closure. PLoS Biol. **4**: 1749-1762.

Morieri, G., Martinez, E.A., Jarynowski, A., Driguez, H., Morris, R., Oldroyd, G.E.D. and Downie, J.A. (2013). Host-specific Nod-factors associated with *Medicago truncatula* nodule infection differentially induce calcium influx and calcium spiking in root hairs. New Phytol. **200**: 656-662.

Mukherji, S. and Soderling, T.R. (1994). Regulation of Ca^{2+} Calmodulin-dependent Protein-Kinase-II by Intersubunit-catalyzed and Intrasubunit-catalyzed Autophosphorylations. J. Biol. Chem. **269**: 13744-13747.

Murray, J.D., Muni, R.R.D., Torres-Jerez, I., Tang, Y., Allen, S., Andriankaja, M., Li, G., Laxmi, A., Cheng, X., Wen, J., Vaughan, D., Schultze, M., Sun, J., Charpentier, M., Oldroyd, G., Tadege, M., Ratet, P., Mysore, K.S., Chen, R. and Udvardi, M.K. (2011). Vapyrin, a gene essential for intracellular progression of arbuscular mycorrhizal symbiosis, is also essential for infection by rhizobia in the nodule symbiosis of Medicago truncatula. Plant J. **65**: 244-252.

Oldroyd, G.E.D. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. Nat. Rev. Microbiol. **11**: 252-263.

Oldroyd, G.E.D. and Downie, J.A. (2004). Calcium, kinases and nodulation signalling in legumes. Nat. Rev. Mol. Cell Biol. **5**: 566-576.

Oldroyd, G.E.D., Murray, J.D., Poole, P.S. and Downie, J.A. (2011). The Rules of Engagement in the Legume-Rhizobial Symbiosis. Annu. Rev. Genet. **45**: 119-144.

Ovchinnikova, E., Journet, E.P., Chabaud, M., Cosson, V., Ratet, P., Duc, G., Fedorova, E., Liu, W., den Camp, R.O., Zhukov, V., Tikhonovich, I., Borisov, A., Bisseling, T. and Limpens, E. (2011). IPD3 Controls the Formation of Nitrogen-Fixing Symbiosomes in Pea and *Medicago* Spp. Mol. Plant-Microbe Interact. **24**: 1333-1344.

Palma, K., Zhao, Q., Cheng, Y.T., Bi, D., Monaghan, J., Cheng, W., Zhang, Y. and Li, X. (2007). Regulation of plant innate immunity by three proteins in a complex conserved across the plant and animal kingdoms. Genes Dev. **21**: 1484-1493.

Parniske, M. (2008). Arbuscular mycorrhiza: the mother of plant root endosymbioses. Nat. Rev. Microbiol. **6**: 763-775.

Patil, S., Takezawa, D. and Poovaiah, B.W. (1995). Chimeric plant calcium/calmodulin-dependent protein kinase gene with a neural visinin-like calcium-binding domain. Proc. Natl. Acad. Sci. USA **92**: 4897-4901.

Patron, N., Orzaez, D., Marillonnet, S., Warzecha, H., Matthewman, C., Youles, M., Raitskin, O., Leveau, A., Farre, G., Rogers, C., Smith, A., Hibberd, J., Webb, A.A.R., Locke, J., Schornack, S., Ajioka, J., Baulcombe, D.C., Zipfel, C., Kamoun, S., Jones, J.D.G., Kuhn, H., Robatzek, S., Van Esse, H.P., Sanders, D., Oldroyd, G., Martin, C., Field, R., O'Connor, S., Fox, S., Wulff, B., Miller, B., Breakspear, A., Radhakrishnan, G., Delaux, P.-M., Loque, D., Granell, A., Tissier, A., Shih, P., Brutnell, T.P., Quick, W.P., Rischer, H., Fraser, P.D., Aharoni, A., Raines, C., South, P.F., Ane, J.-M., Hamberger, B.R., Langdale, J., Stougaard, J., Bouwmeester, H., Udvardi, M., Murray, J.A.H., Ntoukakis, V., Schaefer, P., Denby, K., Edwards, K.J., Osbourn, A. and Haseloff, J. (2015). Standards for plant synthetic biology: a common syntax for exchange of DNA parts. New Phytol. **208**: 13-19.

Peiter, E., Sun, J., Heckmann, A.B., Venkateshwaran, M., Riely, B.K., Otegui, M.S., Edwards, A., Freshour, G., Hahn, M.G., Cook, D.R., Sanders, D., Oldroyd, G.E.D., Downie, J.A. and Ane, J.-M. (2007). The *Medicago truncatula* DMI1 Protein Modulates Cytosolic Calcium Signaling. Plant Physiol. **145**: 192-203.

Perochon, A., Aldon, D., Galaud, J.-P. and Ranty, B. (2011). Calmodulin and calmodulin-like proteins in plant calcium signaling. Biochimie **93**: 2048-2053.

Poovaiah, B.W., Du, L., Wang, H. and Yang, T. (2013). Recent Advances in Calcium/Calmodulin-Mediated Signaling with an Emphasis on Plant-Microbe Interactions. Plant Physiol. **163**: 531-542.

Poovaiah, B.W., Xia, M., Liu, Z.H., Wang, W.Y., Yang, T.B., Sathyanarayanan, P.V. and Franceschi, V.R. (1999). Developmental regulation of the gene for chimeric calcium/calmodulin-dependent protein kinase in anthers. Planta **209**: 161-171.

Popp, C. and Ott, T. (2011). Regulation of signal transduction and bacterial infection during root nodule symbiosis. Curr. Opin. Plant Biol. **14**: 458-467.

Ramachandiran, S., Takezawa, D., Wang, W. and Poovaiah, B.W. (1997). Functional domains of plant chimeric calcium/calmodulin-dependent protein kinase: Regulation by autoinhibitory and visinin-like domains. J. Biochem. **121**: 984-990.

Reddy, A.S.N., Marquez, Y., Kalyna, M. and Barta, A. (2013). Complexity of the Alternative Splicing Landscape in Plants. Plant Cell **25**: 3657-3683.

Reddy, V.S. and Reddy, A.S.N. (2004). Proteomics of calcium-signaling components in plants. Phytochemistry **65**: 1745-1776.

Remy, W., Taylor, T.N., Hass, H. and Kerp, H. (1994). Four-hundred-million-year-old vesicular-arbuscular mycorrihizae. Proc. Natl. Acad. Sci. USA **91**: 11841-11843.

Requena, L. (2000). The purification and characterisation of barley oxalate oxidase. In Department of Biological Chemistry, John Innes Centre: University of East Anglia.

Requena, L. and Bornemann, S. (1999). Barley (*Hordeum vulgare*) oxalate oxidase is a manganese-containing enzyme. Biochem. J. **343**: 185-190.

Riely, B.K., Lougnon, G., Ane, J.-M. and Cook, D.R. (2007). The symbiotic ion channel homolog DMI1 is localized in the nuclear membrane of *Medicago truncatula* roots. Plant J. **49**: 208-216.

Rival, P., Bono, J.-J., Gough, C., Bensmihen, S. and Rosenberg, C. (2013). Cell autonomous and non-cell autonomous control of rhizobial and mycorrhizal infection in *Medicago truncatula*. Plant Signal. Behav. **8**: 1-4.

Rival, P., de Billy, F., Bono, J.-J., Gough, C., Rosenberg, C. and Bensmihen, S. (2012). Epidermal and cortical roles of *NFP* and *DMI3* in coordinating early steps of nodulation in *Medicago truncatula*. Development **139**: 3383-3391.

Rogers, C. and Oldroyd, G.E.D. (2014). Synthetic biology approaches to engineering the nitrogen symbiosis in cereals. J. Exp. Bot. **65**: 1939-1946.

Rosenberg, O.S., Deindl, S., Sung, R.J., Nairn, A.C. and Kuriyan, J. (2005). Structure of the Autoinhibited Kinase Domain of CaMKII and SAXS Analysis of the Holoenzyme. Cell **123**: 849-860.

Routray, P., Miller, J.B., Du, L., Oldroyd, G.E.D. and Poovaiah, B.W. (2013). Phosphorylation of S344 in the calmodulin-binding domain negatively affects CCaMK function during bacterial and fungal symbioses. Plant J. **76**: 287-296.

Roux, B., Rodde, N., Jardinaud, M.-F., Timmers, T., Sauviac, L., Cottret, L., Carrere, S., Sallet, E., Courcelle, E., Moreau, S., Debelle, F., Capela, D., de Carvalho-Niebel, F., Gouzy, J., Bruand, C. and Gamas, P. (2014). An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. Plant J. **77**: 817-837.

Russo Krauss, I., Merlino, A., Vergara, A. and Sica, F. (2013). An Overview of Biological Macromolecule Crystallization. Int. J. Mol. Sci. **14**: 11643-11691.

Sathyanarayanan, P.V., Cremo, C.R. and Poovaiah, B.W. (2000). Plant chimeric Ca^{2+}/calmodulin-dependent Protein Kinase. J. Biol. Chem. **275**: 30417-30422.

Sathyanarayanan, P.V. and Poovaiah, B.W. (2002). Autophosphorylation-dependent inactivation of plant chimeric calcium/calmodulin-dependent protein kinase. Eur. J. Biochem. **269**: 2457-2463.

Sathyanarayanan, P.V. and Poovaiah, B.W. (2004). Decoding Ca^{2+} signals in plants. Crit. Rev. Plant Sci. **23**: 1-11.

Sathyanarayanan, P.V., Siems, W.F., Jones, J.P. and Poovaiah, B.W. (2001). Calcium-stimulated Autophosphorylation Site of Plant Chimeric Calcium/Calmodulin-dependent Protein Kinase. J. Biol. Chem. **276**: 32940-32947.

Schneider, A, Walker, S.A., Poyser, S., Sagan, M., Ellis, T.H.N. and Downie, J.A. (1999). Genetic mapping and functional analysis of a nodulation-defective mutant (*sym19*) of pea (*Pisum sativum* L.). Mol. Gen. Genet. **262**: 1-11.

Schneider, A., Walker, S.A., Sagan, M., Duc, G., Ellis, T.H.N. and Downie, J.A. (2002). Mapping of the nodulation loci sym9 and sym10 of pea (Pisum sativum L.). Theor. Appl. Genet. **104**: 1312-1316.

Shaw, S.L. and Long, S.R. (2003). Nod Factor Elicits Two Separable Calcium Responses in *Medicago truncatula* Root Hair Cells. Plant Physiol. **131**: 976-984.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J.V. and Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. Nat. Protoc. **1**: 2856-2860.

Shi, B., Ni, L., Liu, Y., Zhang, A., Tan, M. and Jiang, M. (2014). OsDMI3-mediated activation of OsMPK1 regulates the activities of antioxidant enzymes in abscisic acid signalling in rice. Plant Cell Environ. **37**: 341-352.

Shimoda, Y., Han, L., Yamazaki, T., Suzuki, R., Hayashi, M. and Imaizumi-Anraku, H. (2012). Rhizobial and Fungal Symbioses Show Different Requirements for Calmodulin Binding to Calcium Calmodulin-Dependent Protein Kinase in *Lotus japonicus*. Plant Cell **24**: 304-321.

Sieberer, B.J., Chabaud, M., Fournier, J., Timmers, A.C. and Barker, D.G. (2012). A switch in Ca^{2+} spiking signature is concomitant with endosymbiotic microbe entry into cortical root cells of *Medicago truncatula*. Plant J. **69**: 822-830.

Sieberer, B.J., Chabaud, M., Timmers, A.C., Monin, A., Fournier, J. and Barker, D.G. (2009). A Nuclear-Targeted Cameleon Demonstrates Intranuclear Ca^{2+} Spiking in *Medicago truncatula* Root Hairs in Response to Rhizobial Nodulation Factors. Plant Physiol. **151**: 1197-1206.

Singh, S., Katzer, K., Lambert, J., Cerri, M. and Parniske, M. (2014). CYCLOPS, A DNA-Binding Transcriptional Activator, Orchestrates Symbiotic Root Nodule Development. Cell Host & Microbe **15**: 139-152.

Singh, S. and Parniske, M. (2012). Activation of calcium- and calmodulin-dependent protein kinase (CCaMK), the central regulator of plant root endosymbiosis. Curr. Opin. Plant Biol. **15**: 444-453.

Slabinski, L., Jaroszewski, L., Rychlewski, L., Wilson, I.A., Lesley, S.A. and Godzik, A. (2007). XtalPred: a web server for prediction of protein crystallizability. Bioinformatics **23**: 3403-3405.

Smit, P., Raedts, J., Portyanko, V., Debelle, F., Gough, C., Bisseling, T. and Geurts, R. (2005). NSP1 of the GRAS protein family is essential for rhizobial Nod factor-induced transcription. Science **308**: 1789-1791.

Staiger, D. and Brown, J.W.S. (2013). Alternative Splicing at the Intersection of Biological Timing, Development, and Stress Responses. Plant Cell **25**: 3640-3656.

Stracke, S., Kistner, C., Yoshida, S., Mulder, L., Sato, S., Kaneko, T., Tabata, S., Sandal, N., Stougaard, J., Szczyglowski, K. and Parniske, M. (2002). A plant receptor-like kinase required for both bacterial and fungal symbiosis. Nature **417**: 959-962.

Stratton, M.M., Chao, L.H., Schulman, H. and Kuriyan, J. (2013). Structural studies on the regulation of Ca^{2+}/calmodulin dependent protein kinase II. Curr. Opin. Struct. Biol. **23**: 292-301.

Sun, J., Miller, J.B., Granqvist, E., Wiley-Kalil, A., Gobbato, E., Maillet, F., Cottaz, S., Samain, E., Venkateshwaran, M., Fort, S., Morris, R.J., Ane, J.-M., Denarie, J. and Oldroyd, G.E.D. (2015). Activation of Symbiosis Signaling by Arbuscular Mycorrhizal Fungi in Legumes and Rice. Plant Cell **27**: 823-838.

Swainsbury, D.J.K. (2011). Decoding calcium spiking: The calcium ion-binding properties of CCaMK. In Department of Biological Chemistry, John Innes Centre: University of East Anglia.

Swainsbury, D.J.K., Zhou, L., Oldroyd, G.E.D. and Bornemann, S. (2012). Calcium Ion Binding Properties of *Medicago truncatula* Calcium/Calmodulin-Dependent Protein Kinase. Biochemistry **51**: 6895-6907.

Syed, N.H., Kalyna, M., Marquez, Y., Barta, A. and Brown, J.W.S. (2012). Alternative splicing in plants - coming of age. Trends Plant Sci. **17**: 616-623.

Takeda, N., Maekawa, T. and Hayashi, M. (2012). Nuclear-Localized and Deregulated Calcium- and Calmodulin-Dependent Protein Kinase Activates Rhizobial and Mycorrhizal Responses in *Lotus japonicus*. Plant Cell **24**: 810-822.

Takeda, N., Tsuzuki, S., Suzaki, T., Parniske, M. and Kawaguchi, M. (2013). CERBERUS and NSP1 of Lotus japonicus are Common Symbiosis Genes that Modulate Arbuscular Mycorrhiza Development. Plant and Cell Physiology **54**: 1711-1723.

Takezawa, D., Ramachandiran, S., Paranjape, V. and Poovaiah, B.W. (1996). Dual Regulation of a Chimeric Plant Serine/Threonine Kinase by Calcium and Calcium/Calmodulin. J. Biol. Chem. **271**: 8126-8132.

Terpe, K. (2003). Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems. Appl. Microbiol. Biotechnol. **60**: 523-533.

Tidow, H. and Nissen, P. (2013). Structural diversity of calmodulin binding to its target sites. FEBS J. **280**: 5551-5565.

Tirichine, L., Imaizumi-Anraku, H., Yoshida, S., Murakami, Y., Madsen, L.H., Miwa, H., Nakagawa, T., Sandal, N., Albrektsen, A.S., Kawaguchi, M., Downie, A., Sato, S., Tabata, S., Kouchi, H., Parniske, M., Kawasaki, S. and Stougaard, J. (2006). Deregulation of a Ca^{2+}/calmodulin-dependent kinase leads to spontaneous nodule development. Nature **441**: 1153-1156.

Venkateshwaran, M., Cosme, A., Han, L., Banba, M., Satyshur, K.A., Schleiff, E., Parniske, M., Imaizumi-Anraku, H. and Ané, J.-M. (2012). The Recent Evolution of a Symbiotic Ion Channel in the Legume Family Altered Ion Conductance and Improved Functionality in Calcium Signaling. Plant Cell **24**: 2528-2545.

Venkateshwaran, M., Jayaraman, D., Chabaud, M., Genre, A., Balloon, A.J., Maeda, J., Forshey, K., den Os, D., Kwiecien, N.W., Coon, J.J., Barker, D.G. and Ané, J.-M. (2015). A role for the mevalonate pathway in early plant symbiotic signaling. Proceedings of the National Academy of Sciences of the United States of America **112**: 9781-9786.

Venkateshwaran, M., Volkening, J.D., Sussman, M.R. and Ané, J.-M. (2013). Symbiosis and the social network of higher plants. Curr. Opin. Plant Biol. **16**: 118-127.

Wais, R.J., Galera, C., Oldroyd, G., Catoira, R., Penmetsa, R.V., Cook, D., Gough, C., Denarie, J. and Long, S.R. (2000). Genetic analysis of calcium spiking responses in nodulation mutants of *Medicago truncatula*. Proc. Natl. Acad. Sci. USA **97**: 13407-13412.

Wais, R.J., Keating, D.H. and Long, S.R. (2002). Structure-Function Analysis of Nod Factor-Induced Root Hair Calcium Spiking in Rhizobium-Legume Symbiosis. Plant Physiol. **129**: 211-224.

Walker, S.A., Viprey, V. and Downie, J.A. (2000). Dissection of nodulation signaling using pea mutants defective for calcium spiking induced by Nod factors and chitin oligomers. Proc. Natl. Acad. Sci. USA **97**: 13413-13418.

Wang, B., Yeun, L.H., Xue, J.-Y., Liu, Y., Ané, J.-M. and Qiu, Y.-L. (2010). Presence of three mycorrhizal genes in the common ancestor of land plants suggests a key role of mycorrhizas in the colonization of land by plants. New Phytol. **186**: 514-525.

Werner, G.D.A., Cornwell, W.K., Sprent, J.I., Kattge, J. and Kiers, E.T. (2014). A single evolutionary innovation drives the deep evolution of symbiotic N_2-fixation in angiosperms. Nat. Commun. **5**: 1-9.

Yamagata, K., Goto, K., Kuo, C.H., Kondo, H. and Miki, N. (1990). Visinin: A novel calcium-binding protein expressed in retinal cone cells. Neuron **4**: 469-476.

Yan, J., Guan, L., Sun, Y., Zhu, Y., Liu, L., Lu, R., Jiang, M., Tan, M. and Zhang, A. (2015). Calcium and ZmCCaMK are involved in brassinosteroid-induced antioxidant defense in maize leaves. Plant Cell Physiol. **56**: 883-896.

Yang, C., Li, A., Zhao, Y., Zhang, Z., Zhu, Y., Tan, X., Geng, S., Guo, H., Zhang, X., Kang, Z. and Mao, L. (2011). Overexpression of a Wheat CCaMK Gene Reduces ABA Sensitivity of *Arabidopsis thaliana* During Seed Germination and Seedling Growth. Plant Mol. Biol. Rep. **29**: 681-692.

Yano, K., Yoshida, S., Muller, J., Singh, S., Banba, M., Vickers, K., Markmann, K., White, C., Schuller, B., Sato, S., Asamizu, E., Tabata, S., Murooka, Y., Perry, J., Wang, T.L., Kawaguchi, M., Imaizumi-Anraku, H., Hayashi, M. and Parniske, M. (2008). CYCLOPS, a mediator of symbiotic intracellular accommodation. Proc. Natl. Acad. Sci. USA **105**: 20540-20545.

Yoshida, S. and Parniske, M. (2005). Regulation of plant symbiosis receptor kinase through serine and threonine phosphorylation. J. Biol. Chem. **280**: 9203-9209.

Young, J.J., Mehta, S., Israelsson, M., Godoski, J., Grill, E. and Schroeder, J.I. (2006). CO_2 signaling in guard cells: Calcium sensitivity response modulation, a Ca^{2+}-independent phase, and CO_2 insensitivity of the *gca2* mutant. Proc. Natl. Acad. Sci. USA **103**: 7506-7511.

Yu, Q., An, L. and Li, W. (2014). The CBL-CIPK network mediates different signaling pathways in plants. Plant Cell Rep. **33**: 203-214.

Zhang, T., Kohlhaas, M., Backs, J., Mishra, S., Phillips, W., Dybkova, N., Chang, S., Ling, H., Bers, D.M., Maier, L.S., Olson, E.N. and Brown, J.H. (2007). CaMKIIδ Isoforms

Differentially Affect Calcium Handling but Similarly Regulate HDAC/MEF2 Transcriptional Responses. J. Biol. Chem. **282**: 35078-35087.

Zhang, X., Dong, W., Sun, J., Feng, F., Deng, Y., He, Z., Oldroyd, G.E.D. and Wang, E. (2015). The receptor kinase *CERK1* has dual functions in symbiosis and immunity signalling. Plant J. **81**: 258-267.

Zhang, X.R.S. and Choi, J.H. (2001). Molecular Evolution of Calmodulin-Like Domain Protein Kinases (CDPKs) in Plants and Protists. J. Mol. Evol. **53**: 214-224.