Profiling protein kinases and other ATP-binding proteins in Arabidopsis using acyl-ATP probes

Inaugural-Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen
Fakultät der Universität zu Köln

vorgelegt von
Joji Grace Villamor-Sernestrand
aus Philippinen

Köln, November 2014

Diese Arbeit wurde am Max-Planck-Institut für Züchtungsforschung in Köln, in der Abteilung Molekulare Phytopathologie (Direktor: Prof. Dr. Paul Schulze-Lefert) angefertigt.

Berichterstatter:	Prof. Dr. Jane Parker
	Prof. Dr. Ulf-Ingo-Flügge
	Dr. Erik Andreasson
Prüfungsvorsitzender:	Prof. Dr. Maria Albani
Tag der Disputation:	22 Januar 2015

SUMMARY

Many protein activities are driven by ATP binding and hydrolysis. Here, we explore the ATP binding proteome of the model plant *Arabidopsis thaliana* using acyl ATP (AcATP) probes. AcATP targets ATP binding sites and covalently labels lysine residues in the ATP binding pocket. Gel-based profiling using biotinylated AcATP showed that labelling is dependent on pH and divalent ions, and can be competed by nucleotides. The vast majority of the AcATP-labelled proteins are known ATP binding proteins. Our search for labelled peptides upon in-gel digest led to the discovery that the biotin moiety of the labelled peptides is oxidized. Also, the in-gel analysis displayed kinase domains of two receptor-like kinases (RLKs) at a molecular weight lower than expected, indicating that these RLKs lost the extracellular domain, possibly as a result of receptor shedding. This putative shedding mechanism is further investigated in a case study of receptor-like kinase RLK902. We found that heterologous expression of RLK902 in *N. benthamiana* leaves causes RLK902 processing that can be inhibited by protease inhibitors E64d and TLCK. Sequence analysis of RLK902 reveals four putative internalization motifs suggesting endocytosis of RLK902 into vesicles. A nuclear localization signal in the kinase domain hints at a possible nuclear transport upon cleavage of RLK902. However, deeper study of the mechanism underlying the RLK902 processing is needed.

Gel-free profiling using desthiobiotinylated AcATP identified 242 different AcATP labelling sites in the Arabidopsis proteome. Analysis of labelling sites revealed a preference of labelling in ATP binding pockets for a broad diversity of ATP binding proteins. Of these, 24 labelled peptides were from a diverse range of protein kinases (PKs), including RLKs, mitogen-activated protein kinases (MPKs), and calcium-dependent protein kinases. A significant portion of the labelling sites could not be assigned to known nucleotide binding sites. However, the fact that labelling could be competed with ATP indicates that these labelling sites might represent previously uncharacterized nucleotide binding sites. A plot of spectral counts against expression levels illustrates the high specificity of AcATP probes for PKs and known ATP binding proteins. AcATP profiling, coupled to targeted MS/MS, of proteomes of Arabidopsis cell cultures treated with or without flg22 did not show differential AcATP

labelling of differentially activated MPKs. This was further investigated using phosphomimic and phosphomutant versions of MPKs that were heterologously expressed in bacteria. These assays confirmed that AcATP labelling is not affected by the phosphorylation (activation) state of the proteins. Overall, this work reveals important opportunities and limitations of profiling ATP binding activities of a large diversity of proteins in plant proteomes.

ZUSAMMENFASSUNG

Die Aktivität vieler Proteine hängt von deren Bindung und Hydrolyse von ATP ab. In dieser Studie wurde das ATP-bindende Proteom in der Modellpflanze *Arabidopsis thaliana* mit Hilfe molekularer Sonden, welche Acyl-ATP (AcATP) enthalten, untersucht. Bei entsprechenden Ziel-Proteinen bindet AcATP kovalent an Lysin-Reste innerhalb der ATP-Bindungsstellen. Untersuchungen mit biotinyliertem AcATP zeigen hierbei, dass der Grad der Markierung sowohl vom pH-Wert, als auch vom Gehalt an zweiwertigen Ionen abhängt, und durch Nukleotide verhindert werden kann. Den größten Teil der AcATP-markierten Proteine bilden bereits bekannte ATP-bindende Proteine. Proteolytische Untersuchungen der markierten Peptide führten hierbei zu der Entdeckung, dass deren Biotin-Gruppe oxidiert vorliegt. Es wurden darüber hinaus zwei Rezeptor-like Kinasen (RLKs) mit einem unerwartet geringen Molekulargewicht identifiziert – möglicherweise als Resultat der Abspaltung der extrazellulären Bereiche der RLKs durch Rezeptor-Shedding. Dieser potentielle Abspaltungsmechanismus wurde anhand der Rezeptor-like Kinase RLK902 weiter untersucht. Die heterologe Expression von RLK902 in Blättern von *Nicotiana benthamiana* zeigt eine Prozessierung der Kinase, was wiederum durch die Protease-Inhibitoren E64d und TLCK verhindert werden kann. Analysen der Sequenz von RLK902 ergeben vier Abschnitte für eine mögliche endozytotische Internalisierung von RLK902 in Vesikel. Darüber hinaus konnte innerhalb der Kinase-Domäne eine Sequenz zur Lokalisierung im Zellkern identifiziert werden, was auf einen möglichen Zellkern-Transport von RLK902 nach der Spaltung hindeutet. Zum Verständnis der Prozessierung von RLK902 sind jedoch weitere Versuche notwendig.

Weitere Untersuchungen des *A. thaliana*-Proteoms mit desthiobiotinyliertem AcATP führten zur Identifizierung von 242 verschiedenen AcATP-Markierungsstellen in der Modellpflanze. Für eine Vielzahl verschiedener ATP-bindender Proteine konnte hierbei eine Präferenz der Markierung an den ATP-Bindungsstellen gezeigt werden. 24 dieser Proteine gehören zu einer diversen Gruppe von Proteinkinasen (PKs), bestehend aus RLKs, Mitogen-aktivierten Proteinkinasen (MAP-Kinasen) und Kalzium-abhängigen Proteinkinasen. Eine signifikante Anzahl der AcATP-Markierungsstellen kann dagegen keiner bekannten Nukleotid-Bindesequenz

zugeordnet werden. Es könnte sich hierbei jedoch um bisher nicht identifizierte Nukleotid-Bindestellen handeln, da die Markierung mit AcATP durch ATP verhindert wird. Generell zeigen molekulare AcATP-Sonden eine hohe Spezifität für PKs und weitere bekannte ATP-bindende Proteine, wie eine graphische Darstellung der spektralen Zählerstände gegen das Expressionslevel demonstriert. Weitere Versuche deuten darauf hin, dass die Markierung mit AcATP in *A*. *thaliana* nicht mit dem Phosphorylierungs- bzw. Aktivierungs-Zustand eines Proteins zusammenhängt. So zeigen MPKs in verschiedenen Aktivierungszuständen aus Flg22-behandelten und unbehandelten Zellkulturen keinen Unterschied bezüglich der AcATP-Markierung. Dies bestätigten auch Untersuchungen nach heterologer Expression von phospho-mimetischen und unphosphorylierten Mutanten von MPKs in Bakterienzellen. Insgesamt zeigt diese Arbeit bedeutende Möglichkeiten und auch Limitierungen bei der Untersuchung von ATP-Bindungsaktivitäten innerhalb des sehr diversen pflanzlichen Proteoms.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

1D	one-dimensional
2D	two-dimensional
Å	Angstrom
AcATP	acyl adenosine-5'-triphosphate
ADP	adenosine diphosphate
AF	apoplastic fluid
ATP	adenosine-5'-triphosphate
BAc	Biotin acyl
Bodipy	boron-dipyrromethene
BSA	bovine serum albumin
BTH	benzothiadiazole
CBB	Coomassie brilliant blue
CID	Collision-induced dissociation
CDK	Calcium dependent kinase
C-terminal	carboxyl terminal
DMF	dimethylformamide
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dpi	day-post-infection
DTT	dithiothreitol
E-64	(L-3-trans-Carboxyoxiran-2-Carbonyl)-L-Leucyl-Admat
EDTA	ethylene diamine tetraacetic acid
EGTA	ethylene glycol tetraacetic acid
ER	endoplasmic reticulum
FDR	false discovery rate
Flg22	flagellin 22
GFP	green fluorescent protein
GO	Gene ontology
GST	glutatione S-transferase
H_2O_2	hydrogen peroxide
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	horseradish peroxidase
IEF	isoelectric focusing
kDa	kilo Dalton
LB	Luria-Bertani
LC	liquid chromatography
LRR	leucine-rich repeat
MES	4-morpholineethansulfonic acid monohydrate
MPK	Mitogen-activated protein kinase
MS	mass spectrometry
MW	molecular weight
NAD	nicotinamide adenine dinucleotide
NADP	nicotinamide adenine dinucleotide phosphate
NPC	no-probe-control

N-terminal	amino terminal
OBAc	Oxidized BAc
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PAMP	pathogen-associated molecular pattern
PBS	phosphate buffered-saline
PCR	polymerase chain reaction
PK	protein kinase
Psy	Pseudomonas syringae pv syringae
Pto	Pseudomonas syringae pathovar (pv) tomato
RFP	red fluorescent protein
Rh	rhodamine
RLK	Receptor-like protein kinase
RuBisCO	ribulose-1,5-bisphosphate carboxylase/oxygenase
SA	salicylic acid
SDS	sodium dodecyl sulfate
SP	signal peptide
TBS	Tris-buffered saline
TFA	trifluoroacetic acid
TM	transmembrane domain
WT	wild-type
YFP	yellow fluorescent protein

CHAPTER 1: INTRODUCTION

ATP and ATP-binding proteins

Biological systems require continuous supply of energy to function and survive. The best known energy molecule in the cell is adenosine triphosphate (ATP). The energy stored in ATP lies in the phosphoanhydride bonds that connect the phosphate groups. The hydrolysis of these bonds release energy needed to do work. ATP, which contains three phosphates, is a high energy molecule, while adenosine diphosphate (ADP) having two phosphates is a low energy molecule. ATP reserves in cells are constantly used and regenerated in the process called respiration. ATP is also the key intermediate product of photosynthesis in plants, where light energy is converted into sugar.

The activities and functions of many proteins are regulated by small molecules or ligands such as ATP. In essence, ATP binding and hydrolysis is the engine in all living organisms. Hundreds of cellular proteins are able to bind and hydrolyse ATP, e.g. to unfold proteins, transport molecules over membranes, or phosphorylate small molecules or proteins. Proteins with very different structures are able to bind ATP. In fact, there is no common motif found in ATP-binding proteins (Chauhan et al, 2009). However, since ATP is a universally important coenzyme and enzyme regulator, understanding the mechanisms of protein-ATP interactions is vital. Importantly, kinases, particularly protein kinases (PKs), play pivotal roles in signal transduction processes and protein function regulation. Kinases are ATP-binding proteins which transfer the gamma phosphate of ATP to substrates.

The genome of the model plant *Arabidopsis thaliana* encodes for over 1099 PKs and hundreds of other ATP-binding proteins including ATP-binding cassette (ABC) transporters, heat shock proteins, structural proteins actin and tubulin, DNA/RNA polymerases and ribosomal proteins (The Arabidopsis Genome Initiative, 2000) (Lamesch et al, 2012). ATP-binding proteins are essential in many physiological processes that cells need to do work and survive. PKs, for example, are involved in nearly all signalling cascades, and regulate processes ranging from cell cycle to flowering and from immunity to germination. Most of the PKs in plants are

membrane-bound receptor-like kinases (RLKs), often carrying extracellular domains with leucine-rich repeats (LRRs), and intracellular kinase domains. The RLK class contains at least 610 members (Shiu & Bleecker, 2001), including receptors involved in development (e.g. BRI1, ER, CLV1) and immunity (e.g. FLS2, EFR). Other important PK classes are mitogen-activated protein (MAP) kinases (MPKs, 20 different members), MPK kinase kinase kinases (MAP3Ks, 60 different members) (Rasmussen et al, 2012) and calcium-dependent protein kinases (CPKs, 34 different members) (Cheng et al, 2002). Because of their diverse and important roles, PKs and ATP-binding proteins have been intensively studied in plant science, however, often individually rather than on a large scale. New approaches are necessary to study PKs and other ATP-binding proteins globally rather than individually. In addition to the complexity brought about by post translational modifications (PTMs), activation regulation, low expression levels, and overlapping binding capabilities to small molecules and ligands of PKs and other ATP-binding proteins, assessing functions and substrate selectivity of these proteins is difficult. However, with new technologies being constantly developed, we expect continued increase in power in studying proteins.

Proteomics: challenges and advances

Proteomics is a field of science that is very dependent and firmly established on technologies. Technical advances are vital to study the complex proteome. Undeniably, it is in proteomics where one can attest the fast improvements and continuing evolution of newer and more powerful tools. To monitor protein dynamics including synthesis/degradation, transport, processing/maturation, PTMs and activation/deactivation in a variety of cellular conditions and perturbations, is crucial to understand the mechanisms that govern animal and plant biology, and to increase insight into cellular signalling networks. However, this is a very daunting task. Indeed, proteomics is a very challenging field, not to mention very dynamic and complex.

The thousands of different protein species making up the proteome of an organism warrant the complexity of proteomics. Hence, high resolution technologies are required to separate the large number of proteins. Common methods used

nowadays are 1D and 2D gel electrophoresis. However, these techniques can not cater to increased demand and expectation of the scientific community to a better resolution, larger proteome coverage and overall more in-depth proteomics, hence the birth of better proteomic technologies. Some noteworthy technological advances in e.g. kinome and ATP-binding protein proteomics, are as follows. First, Zhu and colleagues have first developed protein microarray chips of 119 GST-protein kinase fusions to study yeast protein kinases *en masse*, where high throughput screening for tyrosine phosphorylation activities of PKs were investigated (Lawrence, 2001; Zhu et al, 2000). Second, different enrichment technologies have been developed to increase depth in kinome proteomics. ATP and other kinase inhibitors immobilized to purification columns or beads are now commonly used to enrich for PKs from a complex sample (Graves et al, 2013; Nita-Lazar et al, 2008). Likewise, immobilized metal affinity chromatography (IMAC) is a phosphopeptide enrichment method, which captures phosphorylated peptides of activated PKs (Andersson & Porath, 1986; Nuhse et al, 2003). Third, improved quantification methods in mass spectrometry (MS) is invaluable for proteomics. In addition to label-free MS-based quantification, stable isotope labelling by amino acids in cell culture (SILAC) or isobaric tag for relative and absolute quantitation (iTRAQ) were developed, where two or more samples can be pooled and quantified together in a single MS run (Nita-Lazar et al, 2008). Moreover, considerable upgrades in instrumentations are far-reaching technologies. The development of prefractionation methods coupled to MS led to improved detection and quantification of proteins, e.g. liquid chromatography (LC)-tandem MS (MS/MS), or LC coupled to gas phase fractionation and tandem MS (Davis et al, 2001). In addition, targeted MS/MS which incorporates inclusion and/or exclusion lists, increases sensitivity and reproducibility in detecting target proteins in a complex sample (Schmidt et al, 2009; Wang & Li, 2008). Targeted MS/MS circumvents undersampling of low abundant protein species, and together with having replicate runs of the same conditions, can further increase detection and proteome coverage. Also, mass cytometry has been relatively recently introduced to study kinases and kinase inhibitors in single cells rather than in tissues. However, because of the unique capacities of each of these techniques, more often than not, researchers use a combination of different methods to address one problem/question. Holmes-Davis and colleagues, for example, used a vigorous multiple protein extraction techniques coupled to 2DE and electrospray ionization

(ESI)-MS/MS to study the Arabidopsis pollen proteome (Holmes-Davis et al, 2005). Likewise, the use of multiplexed inhibitor beads (MIBs) to purify kinases, together with iTRAQ or SILAC labelling and MS analysis seem advantageous to study kinases in different cancer cell lines (Cooper et al, 2013; Duncan et al, 2012). But even through such technological advances in proteomics, only a portion of the complex proteome can be monitored at a time.

Another challenge in proteomics, particularly kinome and ATP-binding protein proteomics, is to be able to do functional or activity characterizations of proteins. Large scale proteomics do not usually provide functional information, but the mere presence or absence of proteins and/or their abundance. An added challenge in functional characterizations of proteins is their temporal dynamics as affected by PTMs which reflects on changing protein activity state and conformations. Moreover, kinome and ABP reprogramming upon perturbations/elicitation is difficult to monitor especially for the low abundant proteins. Furthermore, understudied PKs for example, remain elusive especially with methods that use inhibitor-based enrichment techniques, as many understudied PKs have no known inhibitor and/or substrate. Therefore, methods for unbiased study and monitoring of understudied PKs and other ATP-binding proteins are ideal. In this regard, chemoproteomics seems fit. Chemoproteomics relies on the use of small molecule probes to tag proteins indiscriminately and independently of abundance. In addition, chemoproteomics is directed to monitor only a certain class or group of proteins rather than the whole proteome, making analysis less complex and increases depth (Cravatt et al, 2008; Patricelli et al, 2011; van der Hoorn et al, 2011). Chemoproteomics also circumvents elaborate and often difficult prefractionation methods commonly used in kinome (phospho)proteomics. More importantly, chemoproteomics of PKs and other ATP-binding proteins can be directed to monitor PK networks and signalling cascades as it can detect global changes and reprogramming of the kinome i.e. upon stimulus.

Towards mining proteins with ATP-binding activities

ATP binding activities of PKs and other ATP-binding proteins can be detected globally by chemoproteomics using acyl ATP (AcATP) probes (Fig. 1B) (Patricelli et al, 2011; Patricelli et al, 2007). AcATP uses ATP as the binding moiety designed to bind the ATP-binding pockets of ATP-binding proteins. Upon binding, it places the

acyl group in close proximity to conserved lysine residues in the ATP-binding pocket. The acyl phosphonate moiety serves as an electrophilic warhead that can be nucleophilically attacked by the amino group of the lysine, resulting in a covalent attachment of the acyl reporter of AcATP on the lysine and a concomitant release of ATP (Fig. 1B). The reporter tag is usually a biotin (in BAcATP) to capture and identify the labelled proteins. Labelled proteins can be displayed on protein blots using streptavidin-HRP. However, since AcATP label many ATP-binding proteins, and PKs are of relatively low abundance, MS is more often used to identify and quantify labelling with AcATP probes. The analysis is preferably done using Xsite, a procedure that involves trypsination of the entire labeled proteome, followed by analysis of the biotinylated peptides rather than the biotinylated proteins (Okerberg et al, 2005). This 'KiNativ' approach provides enough depth and resolving power to monitor ~160 PKs in a crude mammalian proteome (Patricelli et al, 2011). Of the 518 human PKs (Manning et al, 2002), 394 (=76%) have been detected by AcATP labeling (Patricelli et al, 2007).

KiNativ has mostly been used to validate targets of human drugs that target PKs using competitive labelling experiments. This approach identified selective inhibitors of e.g. Parkinson's disease protein kinase LRRK2 (Deng et al, 2011); the BMK1 and JNK MAP kinases (Yang et al, 2010; Zhang et al, 2012); and the mTOR kinase (Liu et al, 2012). Importantly, the correlation of biological activity of PK-inhibiting drugs with inhibitor affinity detected using KiNativ is better than when compared with affinities determined by assays using heterologously expressed PKs (Patricelli et al, 2011). This improved correlation illustrates that assays in the native environment provide a more realistic measure of PK function.

In addition to characterizing inhibitor selectively, AcATP probes can also display differential ATP binding activities of PKs. Since AcATP labels the conserved Lys in ATP-binding pockets of PKs and other ATP-binding proteins, it indicates on the ATP binding activity of these proteins. Because AcATP reports on the availability of ATP binding sites in ATP-binding proteins, any change in the protein activity that affects ATP binding affinity would be displayed. For example, labelling with AcATP probes during infection with dengue virus detected a 2-8-fold activation of a DNA-dependent PK (Vetter et al, 2012). Likewise, AcATP labelling revealed an unexpected Raf

kinase activation in extracts upon PK inhibitor treatment (Patricelli et al, 2011). In conclusion, profiling with AcATP probes is a powerful approach to monitor PKs and other ATP-binding proteins, and offers unprecedented opportunities to identify selective PK inhibitors and discover PKs with differential ATP binding activities.

Aim and outline of the thesis

This work aims to introduce AcATP profiling of plant proteomes, identify PKs and other ATP-binding proteins labelled by AcATP, sequence AcATP labelling sites, and gain deeper insight in the ATP binding mechanisms of ATP-binding proteins.

Chapter 1 introduces the importance of understanding the mechanisms of ATP-protein interactions, and reviews the different proteomic technologies that can be applied to detect this. Here, the advantages of using a chemoproteomic approach (e.g. AcATP profiling), are discussed.

Chapter 2 shows the results of AcATP labelling experiments. Part 2.1 shows an in-depth characterization of AcATP labelling as affected by probe and proteome concentrations, labelling time, pH, divalent cations, chelating agents, and nucleotide and kinase inhibitors. In addition, application of AcATP labelling in plant science was demonstrated. Furthermore, protein targets of AcATP were identified by MS analysis in both gel-based and gel-free AcATP profiling platforms. In addition to the analysis of labelled peptides, we discovered that biotin is often oxidized in this procedure. We also performed a first in-depth analysis of labelling sites in proteins other than PKs. Part 2.2 introduces the targeted MS/MS approach in AcATP profiling to detect differential AcATP labelling upon flg22-induced immune signalling. Here, we confirmed with an in-depth investigation of AcATP labelling of differentially activated MPKs. Part 2.3 launches a study on a putative receptor shedding event in RLK902, which was originally inferred from gel-based AcATP labelling of Arabidopsis leaf proteome in Part 2.1. We further investigated the putative shedding mechanism by heterologous expression of RLK902 to monitor the processing of this protein in *N. benthamiana* leaves.

Chapter 3 interprets and discusses the results presented in Chapter 2. I highlight different factors that affect AcATP labelling, a comparison of gel-based and gel-free platforms, and the advantages of targeted over data-dependent MS/MS in AcATP profiling. In part 3.1, I describe 63 labelling sites of known nucleotide binding pockets, of which 24 represent a remarkable diversity of PKs, including several LRR-RLKs. In part 3.2, I show that targeted MS/MS increased detection of PKs. Although no differential AcATP labelling was observed for differentially activated MPKs, possible reasons were discussed, and a deeper insight on the mechanism of ATP-protein interaction was made. In Part 3.3, I discuss that processing of RLK902 could be a function of either putative receptor shedding or receptor-mediated endocytosis or both, and suggest future experiments to elucidate this processing further.

Chapter 4 describes the materials used and methods followed in the different experiments presented here. Overall, this work launches a new approach to study ATP-binding proteins in plant science.

CHAPTER 2: RESULTS

2.1 Chemoproteomics with AcATP

2.1.1 Structures and mechanisms of labelling of AcATP probes

The AcATP probes are composed of an ATP, an acyl phosphate linker, and a reporter tag. We used four different AcATP probes that differ in their reporter tags (Fig. 1A). Reporter tags can be a biotin (in BAcATP) or a desthiobiotin (in DTBAcATP) for purification, bodipy (in Bodipy-AcATP) for visualization by fluorescence scanning, or an alkyne minitag (in ≡AcATP) that can be coupled to an N_3-biotin (N_3-B), N_3-rhodamine (N_3-Rh), or N_3-biotin-rhodamine (N_3-BRh) using click chemistry.

AcATP was designed to label the conserved lysines (Lys) of kinases and other ATP-binding proteins (Patricelli et al, 2007). When bound to the ATP binding pocket of ATP-binding proteins, the acyl group of BAcATP is positioned in close proximity to one or more conserved Lys in the ATP binding pocket (Patricelli et al, 2007). The electrophilic acyl group of AcATP provokes a nucleophilic attack of Lys resulting in the release of ATP and labelling of Lys of ATP-binding proteins through an irreversible amide bond (Fig. 1B). Thus, AcATP probes report on the ATP binding activity of ATP-binding proteins.

In this study, AcATP was used to label Arabidopsis leaf or cell culture extracts. Incubation of leaf proteomes with AcATP followed by fluorescence scanning or detection of protein blots with streptavidin-HRP revealed numerous labelled proteins when compared to the no-probe-control (Fig. 1C). As expected, the labelling profile of Arabidopsis leaf extracts using BAcATP, ≡AcATP and Bodipy-AcATP are similar, with two of the most prominent signals indicated by arrows (Fig 1C, left panel). Similar protein profiles, but of different labelling intensities, were observed for BAcATP and DTBAcATP labelling when using cell culture extracts (Fig. 1C, right panel). However, labelling of leaf extracts gave very similar profiles (data not shown).

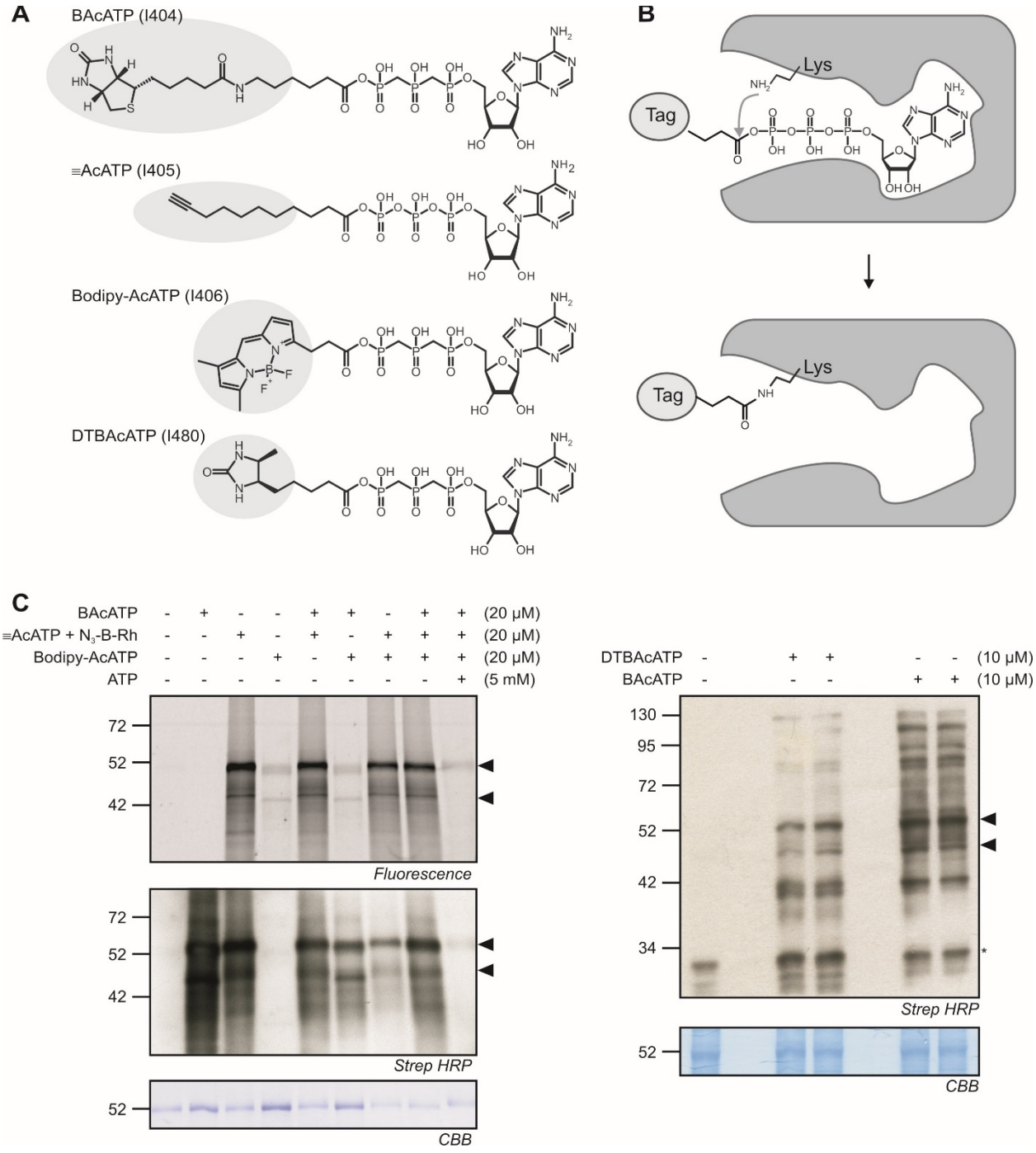


Figure 1. Structure and mechanism of labelling with AcATP probes.

(A) The different AcATP probes all carry an ATP moiety, with the third phosphate group attached to an acyl phosphate reactive group, linked to biotin, alkyne, bodipy, or desthiobiotin for BAcATP, ≡AcATP, Bodipy-AcATP and DTBAcATP, respectively. **(B)** When AcATP binds to the ATP binding pocket of a protein, the amino group of the nearby Lys reacts with the carbonyl carbon, which results in the covalent binding of the tag to the protein while ATP is released. **(C)** Comparison of AcATP labelling of Arabidopsis leaf proteome (left panel) and Arabidopsis cell culture proteome (right panel). Protein extracts were labelled with AcATP and the labelled proteins were detected either by fluorescence scanning or on protein blots using streptavidin-HRP. The ≡AcATP-labelled samples were clicked to an N_3-BRh which can be visualized by both fluorescence scanning and blotting with streptavidin-HRP. For ATP competition, ATP was added to the reaction 30 minutes before adding the probe for labelling. Coomassie Brilliant Blue (CBB) shows loading control. Asterisk indicates endogenously biotinylated proteins.

2.1.2 Characterization of AcATP labelling

Different labelling conditions were tested in order to further characterize AcATP labelling. Here, BAcATP was mainly used for 30-minute labelling reactions of fresh Arabidopsis leaf extracts, followed by detection of labelled proteins from protein blots using streptavidin-HRP.

2.1.2a Labelling is not affected by proteome and probe concentration but is increased by $MgCl_2$

Gradually increasing the proteome concentration for labelling reactions increases the labelled protein signals detected on the blot (Fig. 2A, left panel). However, diluting the labelled high-proteome (4 mg/ml) sample to match the labelled low-proteome (0.5 mg/ml) sample reveals that the labelling profile for both samples are very similar (Fig. 2A, right). The same can be observed with increasing the probe concentrations and diluting it back to a low probe concentration (Fig. 2B). Hence, the proteome and probe concentrations present during labelling do not affect the labelling profiles of Arabidopsis leaf extract.

The extraction in water or 50 mM Tris at pH7.5, or the presence or absence of 0.1 mM EDTA in the BAcATP labelling reaction did not significantly change the labelling profile (Fig. 2C). However, the presence of 10 mM $MgCl_2$ did increase the overall labelling intensity on the protein blot (Fig. 2C). This might be explained by the fact that $MgCl_2$ is a cofactor needed to stabilize the binding of ATP in the ATP-binding pockets of ATP-binding proteins.

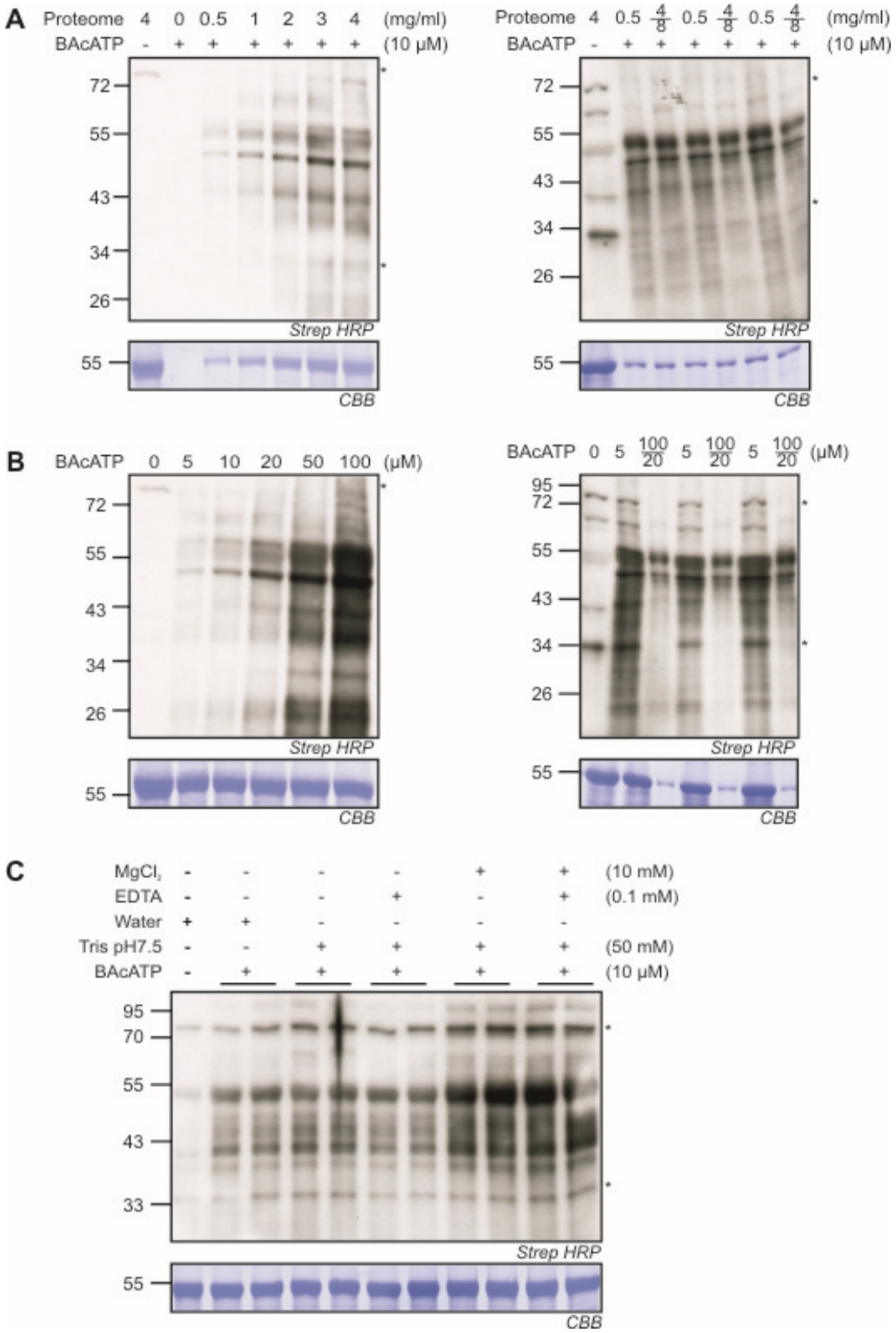


Figure 2. Characterization of BAcATP labelling as affected by proteome concentration, probe concentration and labelling buffer composition.

(A) BAcATP labelling profile of increasing proteome concentrations (left panel). The right panel shows a protein blot where the labelled high proteome sample (4 mg/ml) was diluted to match the labelled low proteome sample (0.5 mg/ml) for better comparison. **(B)** BAcATP labelling profile of increasing probe concentrations (left panel) and dilutions of the labelled high probe sample as compared to the labelled low probe samples (right panel). **(C)** Labelling was done in different buffers with or without 50 mM Tris at pH7.5, $MgCl_2$, and EDTA. CBB shows loading control. Asterisks indicate endogenously biotinylated proteins.

2.1.2b Labelling increases as labelling time increases and is dependent on pH

Labelling gradually increased over time and occurred faster at 25°C than at 4°C (Fig. 3A). All subsequent labelling reactions were then done for 15-30 minutes at 25°C. In addition, labelling with BAcATP depends on pH. Weak or no labelling occurs at acidic pH (pH3–5), and strong labelling occurs at neutral to basic pH (pH7–10; Fig. 3B). We chose pH7.5 as the standard pH during labelling to mimic the conditions of the plant cytoplasm.

The endogenously biotinylated proteins 3-methylcrotonyl-CoA carboxylase (MCCA) and biotin carboxyl carrier protein (BCCP) appeared as background signals at 95 and 34-kDa, respectively, in both labelled samples and no-probe-control. However, both these proteins are also ATP-binding proteins so their intensities increase upon labelling (Fig. 3).

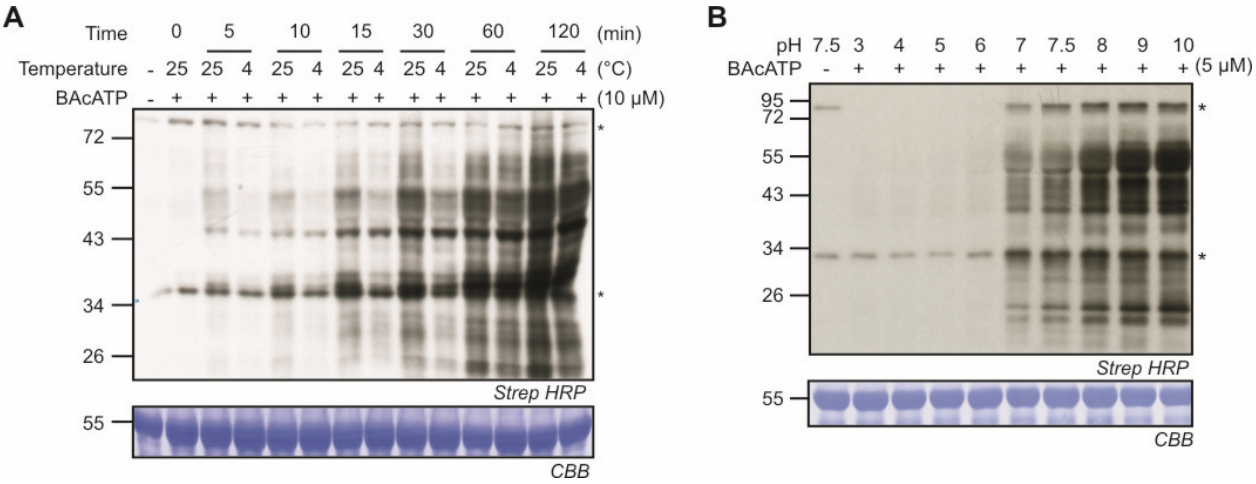


Figure 3. BAcATP labelling is affected by time of labelling and pH

(A) BAcATP labelling profiles of Arabidopsis leaf proteome with increasing labelling time either in 4°C or 25°C. Labelling of each sample was stopped by adding gel loading buffer. **(B)** Arabidopsis leaf extracts were labelled with BAcATP at various pH values, and the biotinylated proteins were detected from protein blots using streptavidin-HRP. CBB shows loading control. Asterisks indicate endogenously biotinylated proteins.

2.1.2c Labelling is affected by the type and amount of divalent metal cations and chelating agents

The observation that $MgCl_2$ increased labelling in Fig. 2C prompted us to test other divalent metal cations like Mn^{2+} and Ca^{2+} in BAcATP labelling. Similar to $MgCl_2$, the addition of $CaCl_2$ or $MnCl_2$ greatly increased labelling as compared to the control (Fig. 4A), while adding chelating agents EDTA and EGTA decreased labelling (Fig. 4B). Interestingly, additional 65 and 45 kDa signals appeared when $CaCl_2$ was added; additional 60 and 70 kDa signals appeared in $MnCl_2$ sample, and two 65 kDa signals appeared in $MgCl_2$ sample (Fig. 4B, black circles). The $MnCl_2$-induced signals were suppressed more effectively by the addition of EGTA than of EDTA (Fig. 4B, grey circles), whereas the $CaCl_2$- and $MgCl_2$-induced signals were suppressed equally by both EDTA and EGTA. These results are in agreement with the fact that EGTA has a higher affinity for Mn^{2+} than EDTA, and suggests that EDTA and EGTA has similar affinity to both Ca^{2+} and Mg^{2+}. In all cases, addition of 50 mM divalent metal cations resulted in stronger labelling compared to 10 mM. The increased labelling upon the addition of divalent metal cations and the reduced labelling in the presence of chelating agents is consistent with the role of divalent ions in ATP binding to proteins and enzymes.

To better assess the effect of divalent metal cations, especially $CaCl_2$, which gave the strongest increase in labelling, we performed a pulldown purification of biotinylated proteins by streptavidin beads, coupled to 2D gel electrophoresis (Fig. 4C). A large number of protein spots can be detected on the protein blot for $CaCl_2$-treated sample which are absent in the no-$CaCl_2$ control. The purification of labelled proteins using streptavidin beads worked well as there were no signals on the blot of the no-probe-control.

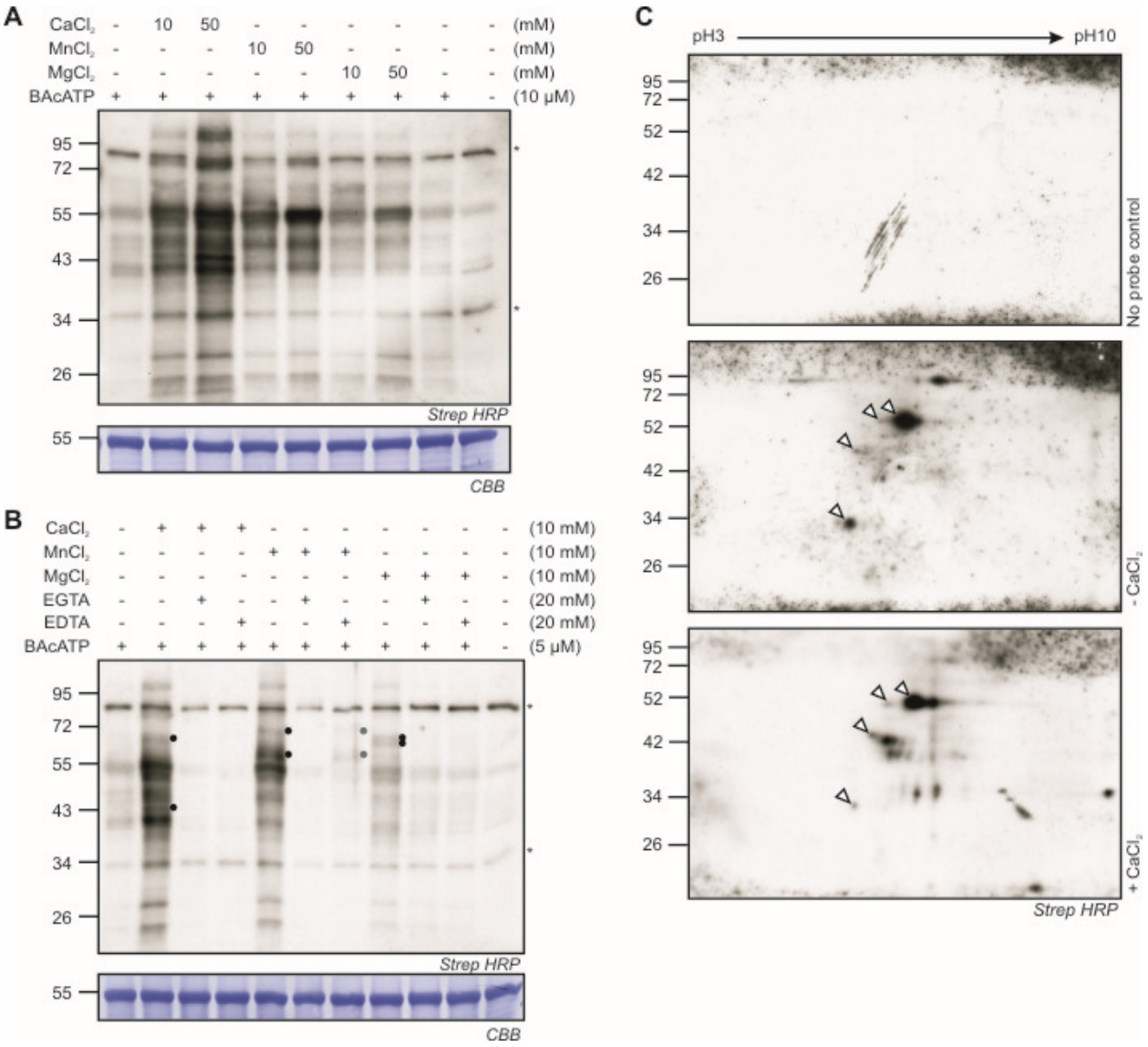


Figure 4. Characterization of BAcATP labelling as affected by divalent metal ions and chelating agents.

(A) Labelling in the presence or absence of $CaCl_2$, $MnCl_2$, and $MgCl_2$ at 10 and 50 mM concentrations. **(B)** Labelling in the presence of different divalent metal ions, with or without chelating agents EDTA and EGTA. Prior to labelling, protein extracts were gel filtered into a DG10 size exclusion columns to remove small molecules before adding the divalent metal ions and/or chelating agents. Biotinylated proteins were detected from a protein blot using streptavidin-HRP. CBB shows loading control. Asterisks indicate endogenously biotinylated proteins. Black circles indicate signals appearing in each treatment, while grey circles indicate signals not inhibited by EDTA. **(C)** A large scale labelling with or without 50 mM $CaCl_2$, followed by purification of biotinylated proteins with streptavidin beads. Purified proteins where resolved with 2D gel electrophoresis, and detected in a protein blot using streptavidin-HRP. Arrows indicate protein spots common to labelling with and without $CaCl_2$.

2.1.2d Labelling is suppressed with ATP-like nucleotides

We next tested whether labelling would be suppressed by different nucleotides (summarized in Fig. 5A). Pre-incubation with ATP or ADP prior to labelling strongly suppressed BAcATP labelling, whereas AMP did not suppress labelling (Fig. 5B). This suggests that phosphates at the β and γ positions are essential for the suppression of BAcATP labelling. In addition, CTP and to a lesser extent, also GTP and TTP also suppressed BAcATP labelling (Fig. 5B). Notably, NADP, but not NAD, also suppressed BAcATP labelling (Fig. 5B), indicating the importance of a phosphate moiety at the 2' position in the nucleotide to suppress labelling (Fig. 5A). Interestingly, nucleotide triphosphates (NTPs) were more potent inhibitors than their deoxynucleotide counterparts (dNTPs; Fig. 5B), suggesting that the 2' hydroxyl group on the sugar moiety of the nucleotide also plays a role in the ability to compete for AcATP labelling (Fig. 5A). In conclusion, these data indicate that labelling is specific because it can be competed with ATP analogues.

2.1.2e Labelling is not significantly affected by kinase inhibitors

We also tested whether labelling would be suppressed by some reported kinase inhibitors that target the ATP binding pocket. Pre-incubation with toyocamycin, suramin, wortmannin, DTNB and K252a (Fig. 6A) prior to labelling did not significantly affect labelling as compared to the no-inhibitor control (Fig. 6B). Independent labelling inhibition experiments with other inhibitor concentrations (up to 100 mM) also did not affect labelling significantly (data not shown). We speculate that these results may be partly (if not completely) due to the limitations of our detection method of protein blotting with streptavidin-HRP in detecting differences in labelling of a small number of protein kinases, rather than looking at the overall inhibition of all ATP-binding proteins in a labelling profile. Moreover, these kinase inhibitors do not target all ATP binding pockets as ATP does, and this is difficult to detect as we see mostly non-protein kinases in our assays (reported in Section 2.1.5c).

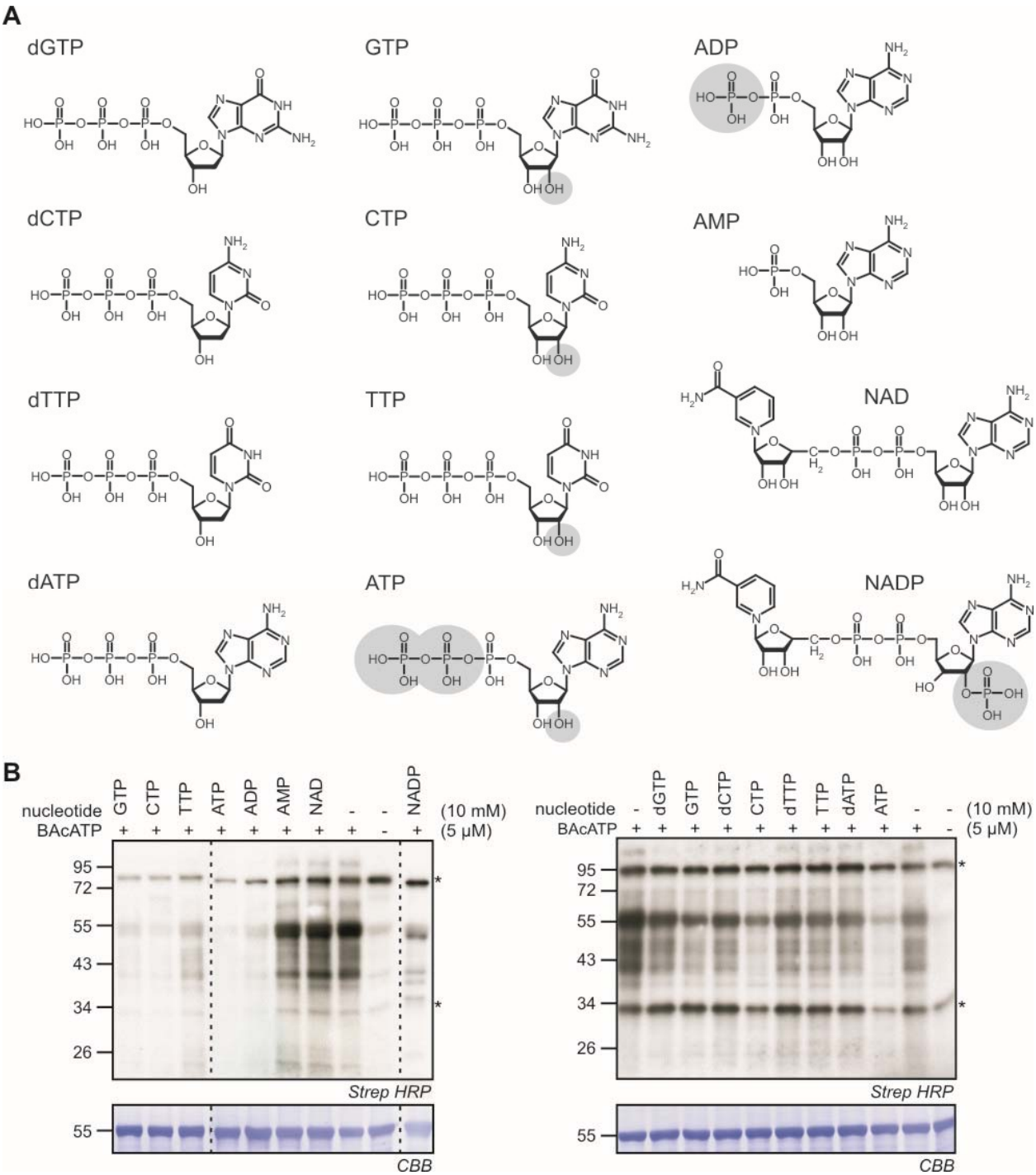


Figure 5. BAcATP labelling competition with nucleotides and nucleotide analogues.

(A) Structures of nucleotides and nucleotide analogues used in BAcATP labelling competition experiments. The OH^- and PO_{4-} groups which are important in competing for BAcATP labelling are highlighted in the structures. **(B)** Protein extracts were gel-filtered with DG10 size exclusion column and pre-incubated for 30 minutes with nucleotide inhibitors prior to labelling. CBB shows loading control. Asterisks indicate endogenously biotinylated proteins.

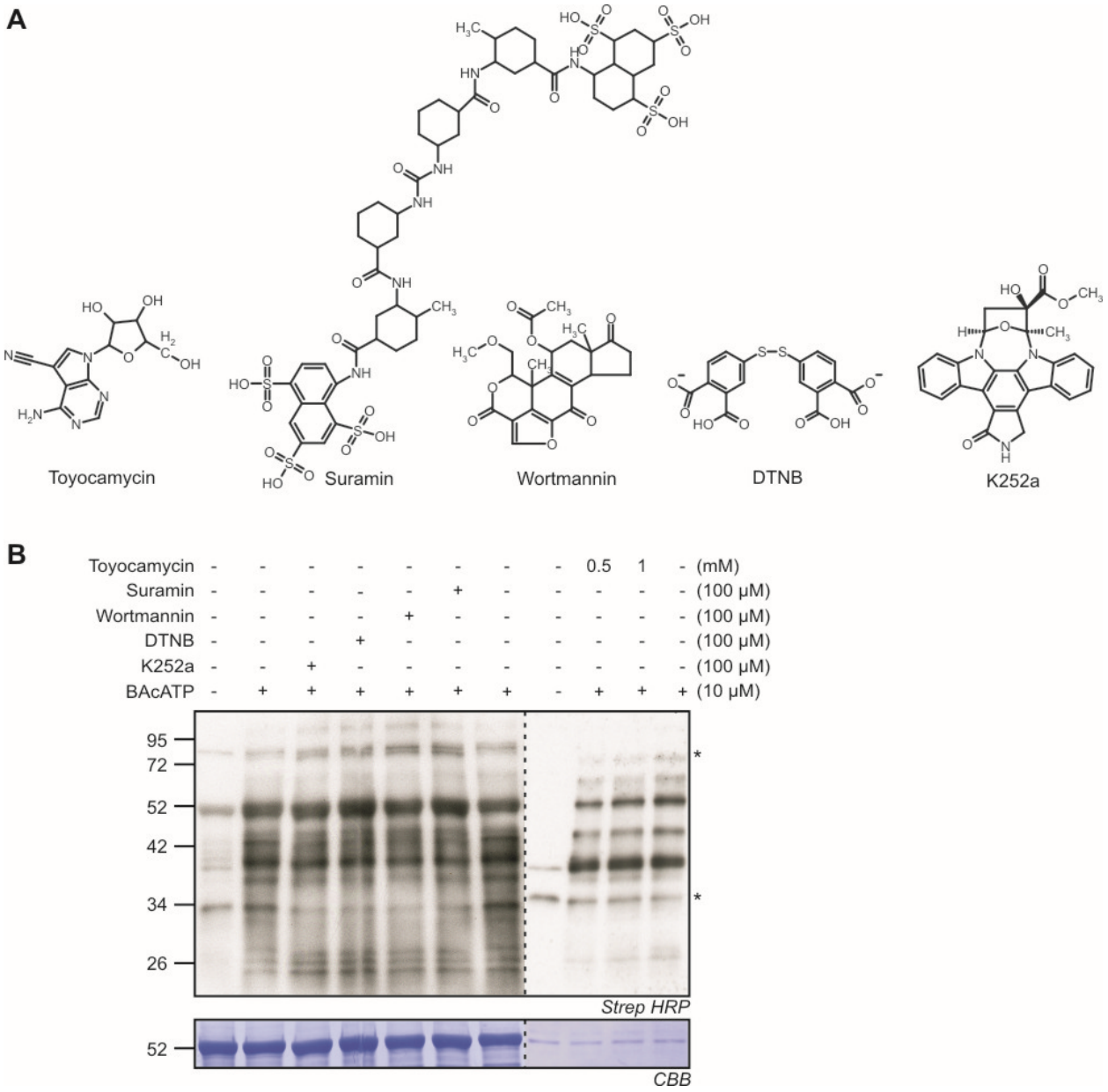


Figure 6. BAcATP labelling competition with known kinase inhibitors.

(A) Structures of selected kinase inhibitors used in BAcATP labelling competition experiments. DTNB stands for 5,5′-Dithiobis(2-nitrobenzoic acid). **(B)** Protein extracts were gel-filtered with DG10 size exclusion column and pre-incubated for 30 minutes with kinase inhibitors prior to BAcATP labelling. CBB shows loading control. Asterisks indicate endogenously biotinylated proteins.

2.1.3 Differential AcATP labelling of various proteomes: Applications of AcATP labelling in plant science

2.1.3a AcATP probes do not label in vivo

In vivo labelling has the advantage that proteins can be monitored in their natural state, active or inactive, depending on conditions of their microenvironment. *In vivo* assays allow the probes to go inside living cells rather than making a lysate. Small molecule probes are often not cell-permeable when tagged with biotin or desthiobiotin, so we tested ≡AcATP and bodipy-AcATP in *in vivo*labelling experiments. However, we did not find labelling after syringe-infiltration of ≡AcATP or bodipy-AcATP into fully expanded Arabidopsis leaves and incubation for two hours (Fig. 7A). Likewise, incubation of bodipy-AcATP with cell cultures did not result in any labelling (Fig. 7B). Finally, incubation of leaves with petioles in ≡AcATP and Bodipy-AcATP probe solutions for two hours also did not result in any labelling (data not shown). Hence, we conclude that AcATP probes probably do not cross the cell membrane of plant cells, and can not be used for *in vivo* labelling experiments.

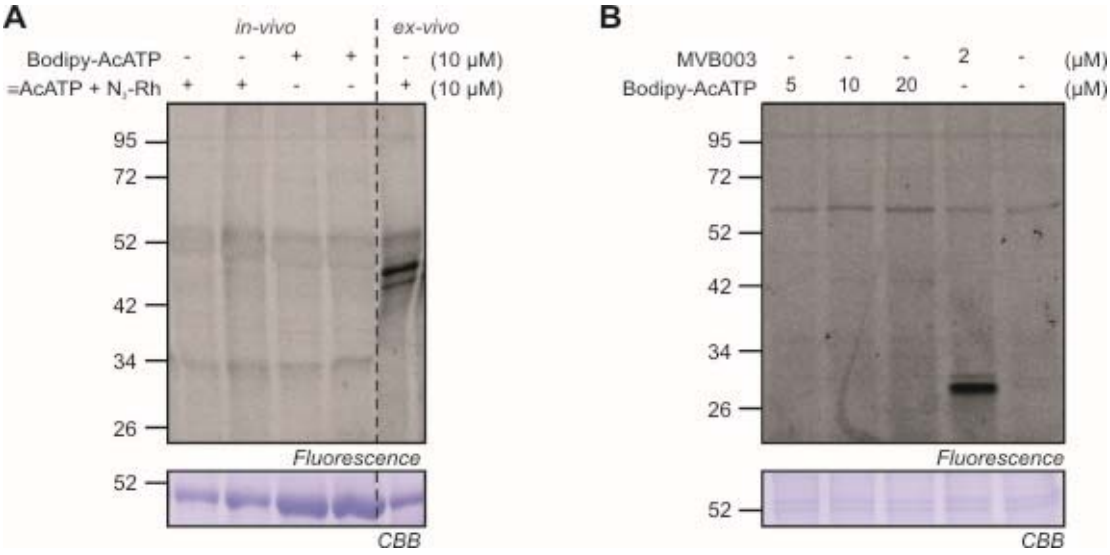


Figure 7. ≡AcATP and Bodipy-AcATP can not be used for in-vivo labelling experiments

(A) *In vivo* labelling by syringe-infiltration of ≡AcATP and Bodipy-AcATP into Arabidopsis leaves. Total proteome from probe-infiltrated leaves were extracted in denaturing buffer containing SDS and separated on acrylamide gels and detected by fluorescence scanning. *Ex vivo* labelling with ≡AcATP, coupled to rhodamine (N_3-Rh) using click chemistry, was used as a control. **(B)** Bodipy-AcATP was incubated with Arabidopsis cell cultures for two hours at different concentrations. Total proteome was extracted in denaturing buffer, separated on acrylamide gel and detected by fluorescence scanning. MVB003 labelling of the proteasome was used as a positive control for *in vivo* labelling. CBB shows loading control.

2.1.3b Searches for differential in vitro labelling

Although AcATP probes cannot be used for *in vivo* studies, they have great potential for *in vitro* labelling studies to compare different proteome samples. Small scale labelling of lysates, followed by detection of labelled proteins by fluorescence scanning or on protein blots using streptavidin-HRP, can be used to detect differential ATP-binding activities of proteins. To test differential ATP-binding activities of proteins, AcATP labelling was done to different proteome preparations from Arabidopsis cell cultures treated with H_2O_2, heat, flg22, or benzothiadiazole (BTH), and *A. thaliana* and *N. benthamiana* plants treated with different bacterial strains *Pseudomonas syringae* pathovar *(pv.) tomato* (Pto) DC3000, PtoDC3000 lacking the HopQ1-1 effector (PtoDC300ΔhQ), and *Pseudomonas syringae* pv. *syringae* (Psy) B728a (Fig. 8A-D).

H_2O_2 treatment causes oxidative stress in Arabidopsis cell cultures. Under natural conditions, oxidative stress in plants is usually caused by many different environmental factors such as pathogen attack, extreme temperatures, wounding, drought, salt, light, and heavy metals. Oxidative stress triggers many defense responses in plants by activating defense-related proteins such as kinases. Moreover, heat treatment also causes oxidative stress and protein degradation, prompting up defense responses in Arabidopsis cell cultures. Labelling with DTBAcATP reveals differential signals at 80, 100 and 150 kDa upon H_2O_2 treatment (Fig. 8A). A similar 100-kDa differential was detected by Bodipy-AcATP in H_2O_2-treated cell cultures as compared to the control (marked with white circles, Fig. 8B). Whereas heat-treating cell cultures at 45°C or 53°C for five minutes results in a weaker 30-kDa signal (Fig. 8B, dark circles). Note however, that this 30-kDa signal is coming from an endogenously fluorescent protein as it is also present in the no-probe-control (Fig. 8B, asterisk).

In addition, flg22, a peptide epitope of bacterial flagellin, triggers PAMP-innate immunity in Arabidopsis activating a downstream mitogen-activated protein kinase (MPK) signalling pathway, and induces the expression of numerous defense-related genes (Newman et al, 2013). BTH, on the other hand, is a functional analog of salicylic acid (SA), triggers SA-dependent defense responses (Lawton et al, 1996).

However, treatments of cell cultures with flg22 or BTH did not result in any differential labelling (Fig. 8B).

Moreover, *Pseudomonas syringae* are bacterial pathogens of diverse plant species. These pathogens carry effectors that trigger immune response in infected plants. Labelling of membrane and soluble fractions obtained from PtoDC3000ΔhQ-treated Arabidopsis leaves was also done using BAcATP. The different labelling profiles of membrane and soluble fractions confirm that these proteomes have different ATP-binding activities (Fig. 8C). With membrane fractionation, one can also enrich for receptor-like kinases (RLKs) located in plasma membranes. Despite the detection of many signals, we did not detect any robust differential labelling upon infection with PtoDC3000 (Fig. 8C). We next investigated the ATP-binding activities of apoplastic proteins during plant-microbe interactions. Infection of *N. benthamiana* with PtoDC3000, PtoDC3000ΔhQ or PsyB728a causes an increase of apoplastic proteins, as detected on Coomassie stained gels (CBB) (Fig. 8D). Interestingly, several of these proteins are labelled with DTBAcATP, and this labelling is competed by pre-incubation with ATP (Fig. 8D). This indicates that some secreted proteins have ATP-binding activities. Importantly, these ATP-binding apoplastic proteins are absent in non-infected controls (Fig. 8D).

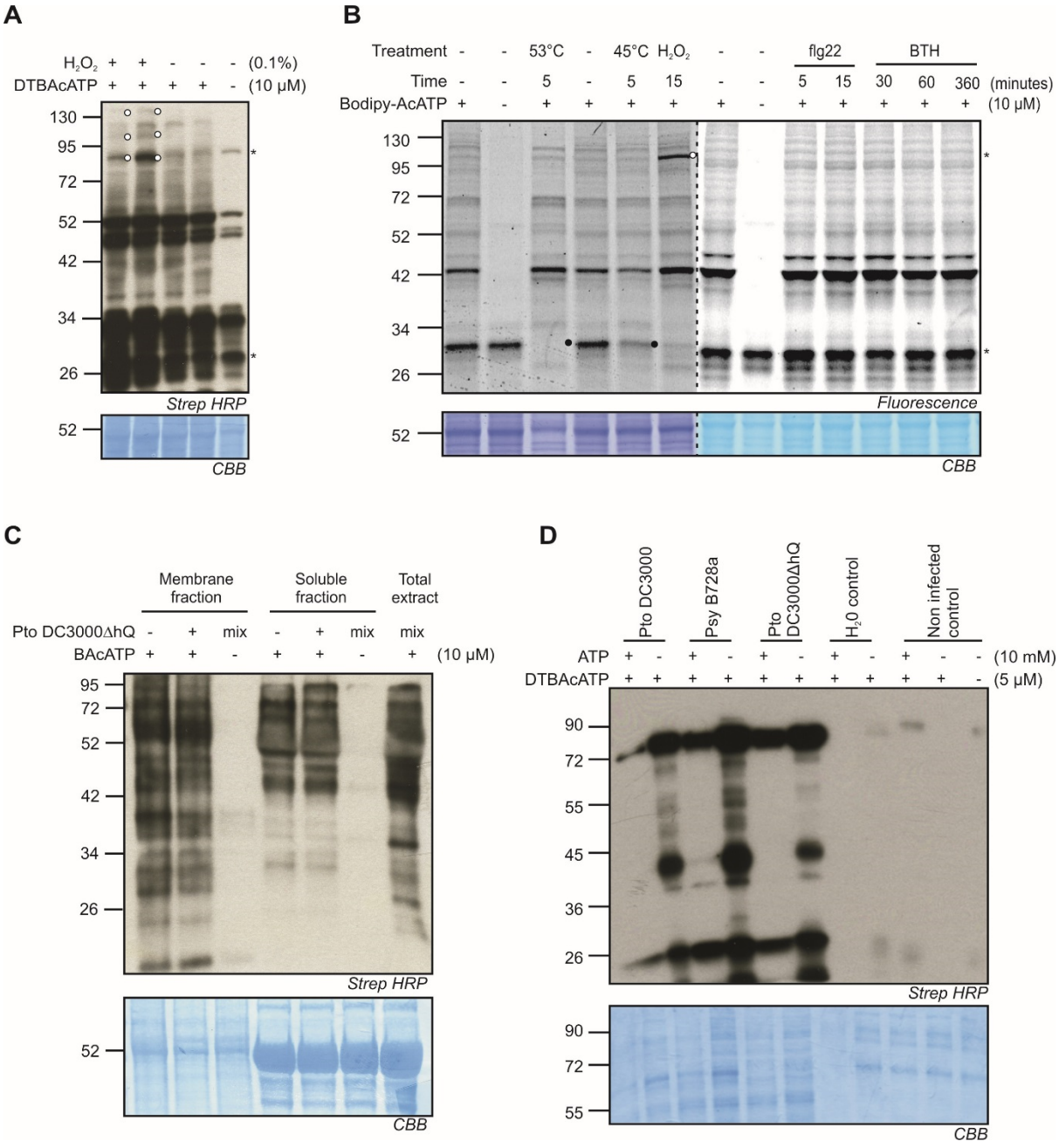


Figure 8. Differential labelling of ATP-binding proteins in various proteomes

Labelling of extracts from treated Arabidopsis cell cultures with DTBAcATP **(A)** and Bodipy-AcATP **(B)**. Total lysate was extracted from cell cultures in kinase buffer with $MnCl_2$ and labelled for 30 minutes. Asterisks indicate endogenously biotinylated (in A) or endogenously fluorescent (in B) proteins. Observed differential protein signals on gels are indicated by white and black circles for H_2O_2-treated and heat-treated samples, respectively. **(C)** Comparative BAcATP labelling of membrane and soluble fractions from Arabidopsis leaves treated with or without PtoDC3000ΔhQ (OD600 of 0.3) at two days post infection (dpi). Total proteome was extracted in non-denaturing membrane buffer and membranes and soluble fraction was separated by ultracentrifugating the lysate. Both membrane and soluble fractions were gel-filtered with DG10 size exclusion column prior to labelling. **(D)** *N. benthamiana* was infiltrated with or without different bacterial strains or water for two dpi and apoplastic fluids were extracted and labelled with DTBAcATP. Labelled proteins were detected on a protein blot using streptavidin-HRP. CBB shows presence of differential proteins accumulating in the apoplast of treated plants as compared to non-infected or water controls. CBB shows loading control.

2.1.3c Large scale labelling and purification of labelled proteins

BAcATP, DTBAcATP or ≡AcATP clicked to N_3-B or N_3-BRh can be used in pulldown experiments to identify labelled proteins. For pulldown experiments, a larger amount of proteome was labelled with 20 µM probe, followed by desalting through a DG10 column to remove salts and the unreacted probe. Biotinylated proteins were captured under denaturing conditions using streptavidin beads. Beads were stringently washed to remove non-biotinylated proteins. Proteins were eluted using gel loading buffer (for 1D gel electrophoresis) or 50% acetic acid (for 2D gel electrophoresis) and digested with trypsin for mass spectrometry (MS) identification. Figure 9A shows an efficient pulldown experiment monitoring the purification of BAcATP-labelled proteins through the different steps in the purification procedure. BAcATP labelling in the presence of 5 mM ATP reduces the purified labelled proteins (Fig. 9B). Protein extracts from Arabidopsis seedlings and cell cultures treated with or without flg22 were also subjected to labelling and pulldown experiments with BAcATP (Fig. 9C). However, there are no obvious differentials between purified proteins from flg22 samples as compared to the controls (Fig. 9C). We followed a similar purification procedure upon ≡AcATP labelling. ≡AcATP labelling might target a different subproteome because this probe does not carry a bulky reporter tag. However, the labelling procedure includes click chemistry to biotinylate the labelled proteins. A 2D gel electrophoresis was done to purified proteins for better resolution and/or more efficient identification of labelled proteins (Fig. 9D).

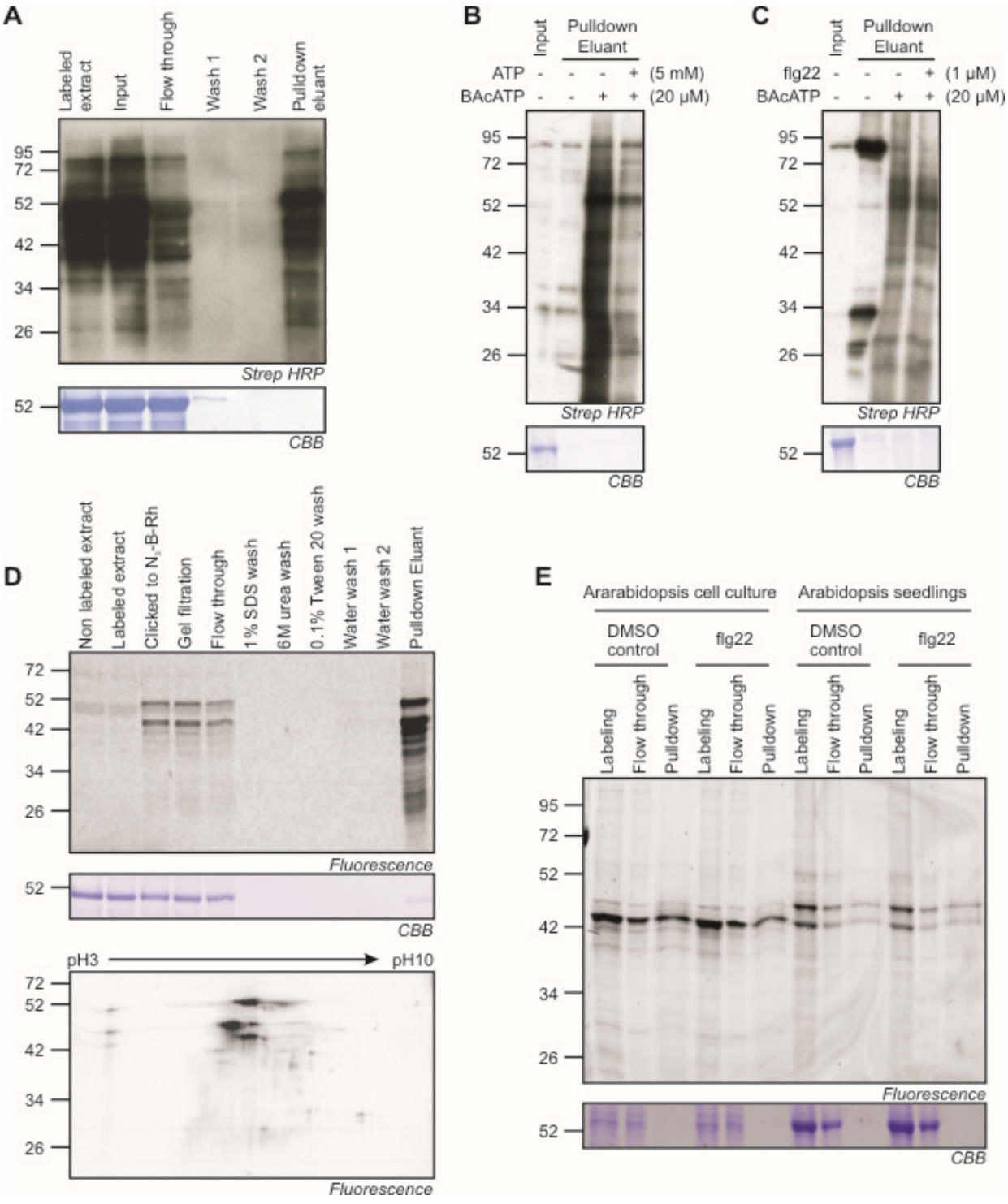


Figure 9. AcATP probes were used to label and purify labelled proteins from a lysate

(A) Protein blot showing the labelling and purification of proteins from a BAcATP-labelled Arabidopsis leaf extract. **(B)** Protein blot of a BAcATP labelling and pulldown of two-week old seedlings with or without ATP competition, and **(C)** Pulldown of BAcATP labelled proteins from proteomes of seedlings treated for 15 minutes with and without flg22. **(D)** Large scale labelling and pulldown with ≡AcATP coupled to N_3-BRh using click chemistry. Lower panel shows separation of purified labelled proteins on a 2D protein gel. **(E)** ≡AcATP labelling and pulldown of labelled proteins from Arabidopsis seedlings and cell cultures with or without flg22 elicitation.

2.1.4 Identification of AcATP targets in Arabidopsis: Gel-based approach

This part of the thesis is published in: Villamor JG, Kaschani F, Colby T, Oeljeklaus J, Zhao D, Kaiser M, Patricelli MP, and van der Hoorn RA (2013) Profiling protein kinases and other ATP-binding proteins in Arabidopsis using acyl-ATP probes. Mol Cell Proteomics. 12. 2481-2496

2.1.4a BAcATP labels ATP-binding proteins

To identify the proteins labelled by BAcATP, biotinylated proteins were affinity purified from BAcATP-labelled leaf extracts using streptavidin beads, separated on protein gel and stained (Fig. 10A). Bands were excised, treated with trypsin, and analyzed by liquid chromatography-tandem mass spectrometry (LC-MS/MS). Peptides with high Mascot scores (>41) and proteins with at least two unique peptides were retained. A total of 112 proteins were identified in the BAcATP-labelled sample while only 13 of these proteins were also identified in the no-probe-control (supplemental Tables S1 and S2). The no-probe-control contains proteins such as the endogenously biotinylated protein MCCA and abundant proteins like RBCL.

We made a selection of the 112 proteins by selecting proteins with the highest spectral counts, supplemented with detected protein kinases (PKs) and proteins for which a labelling site was identified. The resulting selection consists of 38 proteins that could be assigned to the 13 different bands (supplemental Table S2 and Fig. 10A-B). These data indicate that the abundant 55 kDa signal (band 5) is caused by labelling of a subunit of the chloroplastic ATPase (ATPB) and the large subunit of ribulose-1,5-bisphophate carboxylase oxygenase (RBCL). This signal also contains peptides from the β subunit of the mitochondrial ATP synthase (ATPS); glycinamideribonucleotidesynthetase (GARS) and calcium-dependent kinase-9 (CPK9). The 45 kDa signal (band 7) is predominantly caused by phosphoglycerate kinase (PGK1), though this signal also contains spectra from other proteins. The 40 kDa signal (band 8) contains a mixture of the chloroplastic sedoheptulose-1,7-bisphosphatase (SBPASE); chloroplast RNA binding protein (CRB); fructose-biphosphate aldolase-2 (FBA2); subunit A of glyceraldehyde 3-phosphate dehydrogenase (GAPA); and several other proteins including two PKs (LRR1 and PK); an inositol-1,3,4-trisphosphate-5,6-kinase (ITPK1) and an uncharacterized carbohydrate kinase (CK). Weaker signals in the upper 70-90 kDa region of the gel

contain methionine synthase (MS1, band 1); methylcrotonyl-CoA carboxylase (MCCA, band 1); formyltetrahydrofolatesynthetase (THFS, band 2); transketolase (TK, band 2); heat shock proteins 70 (HSP70, band 2) and 60 (HSP60, band 3); thioglucosideglucohydrolase 2 (TGG2, band 3); F-box protein GRH1 (band 3); Ser/Thr/Tyr kinase STY8 (band 3); and subunit B of glutamyl-tRNA aminotransferase (GATB, band 4). Weak signals at 50 kDa (band 6) contain a subunit of a biotin carboxylase (CAC2), a MAP3K protein kinase (VIK1), a monodehydroascorbate reductase (MDAR6) and a calcium-dependent protein kinase (CPK11). Weak signals in the 35-40 kDa region contains peroxisomal NAD-malate dehydrogenase-2 (PMDH2, band 9); an leucine-rich repeat (LRR)-RLK (band 9); and phosphopantothenate-cysteine ligase (COAB, band 10). The signal at 33 kDa (band 11) contains predominantly avidin, but also subunit O2 of photosystem II (PSBO2) and the non-intrinsic ABC protein-7 (NAP7). Finally the weak signals at the bottom of the gel (27-30 kDa) contain chloroplastic glutamine synthetase (GS2, band 11); chloroplast RNA-binding protein (RBFP, band 12); carbonic anhydrase (CA1, band 12); glutathione transferase (GSTF2, band 12), and a UMP/CMP kinase (PYR6, band 12).

In conclusion, the vast majority of the identified proteins in the AcATP-labelled samples are ATP-binding proteins (Fig. 10B, column 8). These ATP-binding proteins include a variety of PKs (STY8, CPK9, VIK1, CPK11, RLK-LRR and LRR1), ATP-based transporters (ATPB, ATPS and NAP7), metabolite kinases (CK, PGK1, ITPK1, PYR6), and metabolic enzymes and chaperones using ATP (GARS, MCCA, THFS, HSP70, HSP60, CAC2, MDAR6, PMDH2, GS2).

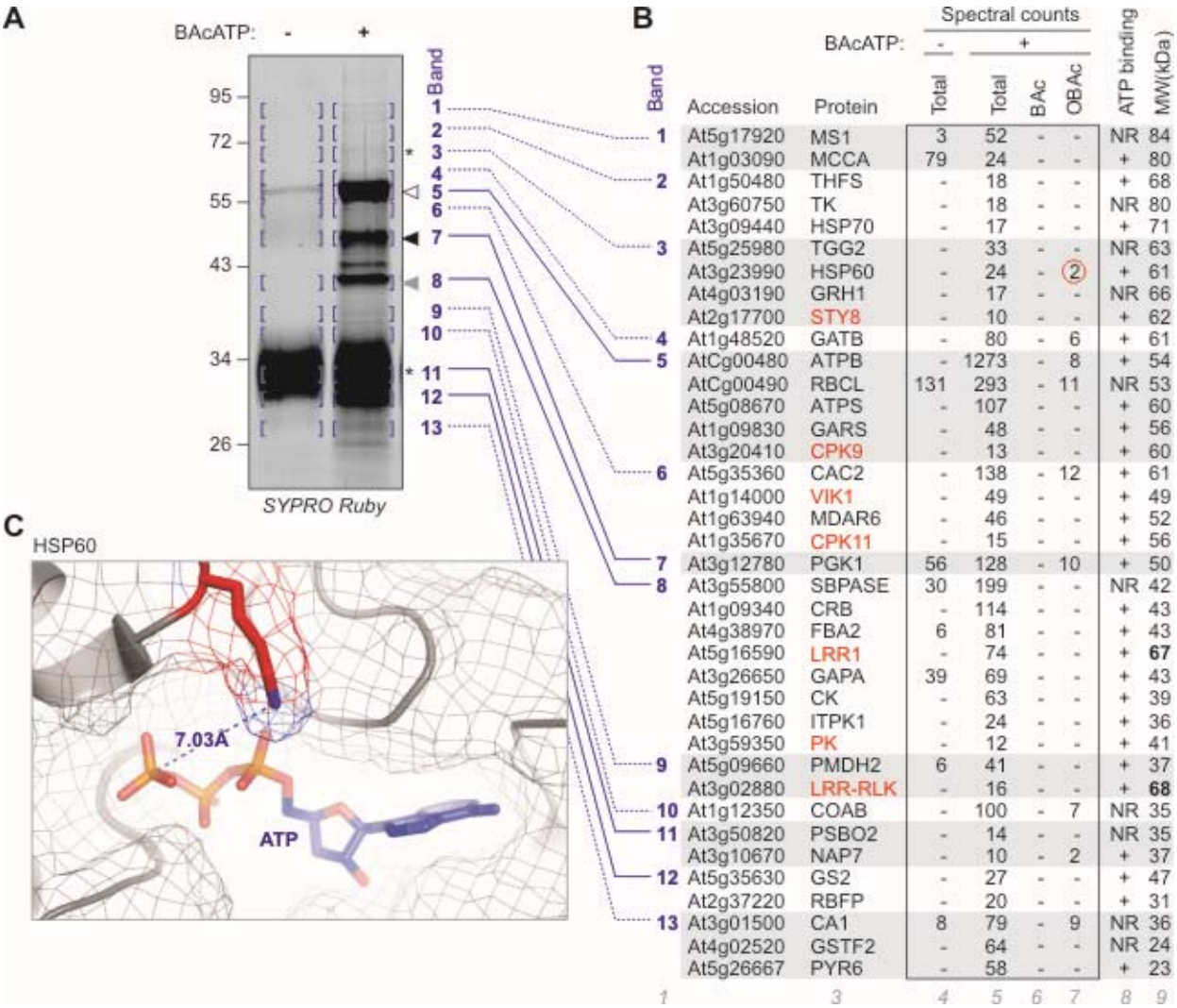


Figure 10. MS analysis of BAcATP-labelled proteins

(A) Arabidopsis leaf extracts were incubated with and without 10 µM BAcATP and biotinylated proteins were purified, separated on protein gel, and detected using SYPRO Ruby staining. Thirteen bands were excised from each lane, treated with trypsin and peptides were eluted and analyzed by LC-MS/MS. Identified proteins are indicated in (B). See supplemental Tables S1 and S2 for more details. **(B)** Spectral counts of proteins in each of the 13 bands. Summarized as: band numbers (column 1); accession code (column 2), protein names (column 3), spectral counts in no-probe-control (column 4) and the AcATP-labelled sample (column 5). The spectral counts are also shown for BAc- and OBAc-modified peptides, indicated separately in columns 6 and 7. Expected ATP binding, based on literature data is indicated in column 8 (ATP binding activity reported (+) or not reported (NR)). The theoretical MW for each protein is given in column 9. Deviations with the found MW are printed in bold. PKs are printed in red. This table is a selection of the most abundant spectral counts per band, supplemented with detected protein kinases and proteins for which labelled peptides were detected. Proteins that were detected in more than one band are only shown in the band with the highest spectral counts. For a complete list of identified proteins, see supplemental Tables S1 and S2. **(C)** BAcATP labelling site in HSP60. The orthologous labelling site of AcATP in HSP60 of Arabidopsis is indicated in the crystal structure of the highly homologous GroEL of *E. coli* (1sx3) containing ATP. The labelled Lys was located from the aligned protein sequences and highlighted in the structure using PyMol. The Lys is at 7.03 Å from the gamma phosphate of ATP.

2.1.4b Putative receptor shedding of RLKs

We next examined the experimental molecular weights (MW) of proteins identified from the protein gel. We noticed a good correlation for nearly all the detected proteins with the striking exception of two leucine rich repeat-receptor like kinases (LRR-RLKs), which have a calculated MW of 68-kDa, but migrate in the region of 40-kDa (Fig. 10B). These are two highly homologous LRR-RLKs carrying six extracellular LRRs, a transmembrane domain and a cytoplasmic kinase domain (Fig. 11). The extracellular domain also contains multiple putative glycosylation sites, which cause LRR-RLK proteins to migrate with an apparent MW that is typically 20-30 kDa larger than their calculated MW. Peptide coverage (Fig. 11 and supplemental Table S1) suggests that these labelled proteins consist of the full protein kinase domain, but lack the extracellular domain. The detection of half-tryptic peptides of the C-termini of both these LRR-RLK proteins indicates that the C-terminus is intact and that a significant truncation must have occurred from the N-terminus. Interestingly, the calculated MW of the cytoplasmic domain is 37 and 38 kDa for At3g02880 and At5g16590, respectively, which coincides with the observed MW of these proteins. Taken together, these data indicate that these LRR-RLKs exist in Arabidopsis leaf extracts as kinase domains lacking the extracellular domain.

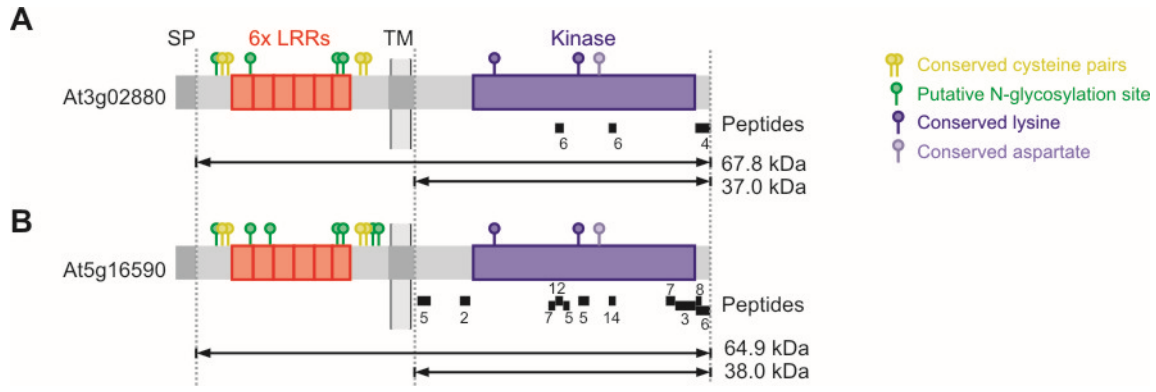


Figure 11. Putative receptor shedding of two receptor-like kinases.

Both RLKs At3g02880 and At5g16590 carry a signal peptide (SP), an N-terminal LRR-flanking domain with a conserved Cys pair; six leucine-rich repeats (LRRs, red), an extracellular juxtamembrane domain with a conserved Cys pair (yellow); a transmembrane domain (TM), an intracellular juxtamembrane domain; a protein kinase domain (blue) and a C-terminal tail. The extracellular domain carries several putative *N*-glycosylation sites (green). Identified peptides and spectral counts are indicated with black bars below the proteins. The theoretical molecular weight is calculated from the plain amino acid sequence for the cytoplasmic domain, and the full mature protein without signal peptide.

2.1.4c Detection of an oxidized biotin moiety

The searches mentioned above included a variable modification of Lys residues with the probe, whilst allowing two trypsin miscleavages. Searches with the theoretical biotin acyl (BAc) modification on the Lys (339.16 Da) did not lead to hits with Mascot scores above the threshold (Fig. 10B, column 6). However, during more liberal searches we noticed that BAc modifications appeared, but only in peptides which included an oxidation associated with the presence of a methionine in the peptide sequence (data not shown). This observation sparked the idea that the modification of the probe is associated with additional oxygen. Indeed, when we searched the data by adding an oxygen to the BAc modification (OBAc, 355.16 Da), we could identify modified peptides for nine different proteins with relatively high Mascot scores (41-71) (Fig. 10B, column 7).

One of those labelled peptides is from HSP60 (At3g23990), a well-described ATP-BINDING PROTEIN. To determine the location of the labelling site in the protein relative to ATP, we searched the protein database for proteins homologous to HSP60 and identified GroEL, an HSP60 protein from *Escherichia coli*, for which an ATP-bound crystal structure is available (1sx3, Chaudhry et. al., 2004). HSP60 and GroEL proteins share 56% amino acid identity including the region where the labelled Lys is located. The amino group of the labelled Lys is located at 7.0 Å from the gamma phosphate of the bound ATP in GroEL (Fig. 10C). Thus, labelling in HSP60 occurred at the expected labelling site.

We next studied the fragmentation spectrum of the labelled peptide of HSP60 further to identify the location of the oxygen within the labelled peptide. The labelled peptide in HSP60 has the sequence VTKDGVTVAK and was detected with a Mascot score of 41.4 with a parental ion of 686.93 *m/z* (z=2) (Fig. 12). Nearly all b- and y-ions were detected in the fragmentation spectrum, locating the labelled Lys including the oxygen at the third position in the peptide: VTK*DGVTVAK. The spectrum also contains three unexplained peaks in the low masses: 243.15, 356.31 and 439.35 Da. Importantly, the 243.15 Da ion corresponds to an oxidized biotin (expected mass: 243.3). An additional 113 Da for the aminohexanoic acid linker predicts an ion of 356.5 Da, which corresponds to the second signal at 356.31 Da. The third ion,

observed 83 Da further at 439.35 Da, corresponds to the labelled version of the commonly-observed immonium ion variant for Lys. High-energy CID fragmentation of Lys residues typically results in an unsaturated cyclic 84 Da fragment, known as a piperidine or tetrahydropyridine ion (Shek et al, 2006). Since the Lys has lost a proton during the labelling reaction, the observed 83 Da distance is explained, indicating that such fragmentation and subsequent cyclization of Lys do not cause a loss of the biotin. Taken together, these data demonstrate that the additional oxygen is located at the biotin, presumably on the sulfur. The 243.2 and 356.5 and 439.3 Da ions are also found in spectra of the other labelled peptides, indicating that oxygen is consistently present on the biotin moiety.

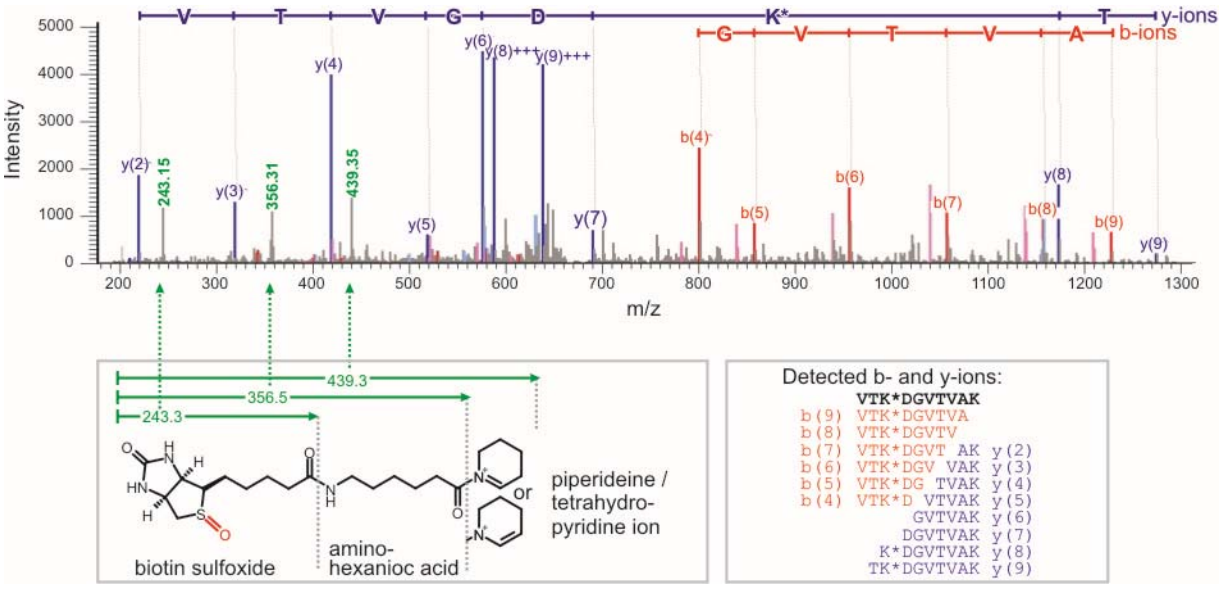


Figure 12. Fragmentation spectrum of HSP60 peptide labelled with oxidized biotin.

The b- and y-ion series is annotated as explained in the bottom right. The putative structure of the modified Lys is shown on bottom left, explaining the three observed ions. The oxide (red) is presumably on the sulfur of the biotin, resulting in biotin sulfoxide.

2.1.5 Identification of AcATP labelling sites: Gel-free approach

2.1.5a Xsite identifies 242 labelling sites

The observation that most of the 112 BAcATP-labelled proteins are ATP binding reflects the potency of the probe to target ATP-binding proteins in a crude proteome. However, only a few of these 112 proteins are PKs. Furthermore, some of the

BAcATP-labelled proteins are not annotated as ATP-binding proteins, and therefore the labelling sites on these proteins are uncertain. To address both issues, we analyzed the labelled proteome deeper by analyzing the labelled peptides. For this we used DTBAcATP (Fig. 1A) (Patricelli et al, 2011) which carries a desthiobiotin tag. Desthiobiotin cannot be oxidized and has a lower affinity for streptavidin, resulting in a more efficient elution of peptides for identification. After labelling with DTBAcATP, the proteomes were digested with trypsin and the desthiobiotinylated peptides were purified using streptavidin beads, and analyzed by LC-MS/MS (Fig. 13A). We recorded 10465 peptide spectra, corresponding to 567 different peptides carrying a modified Lys residue (Fig. 13B). Peptides with >95% probability score (Adam et. al. 2004) and those that were identified in two or more reactions were selected, resulting in 6992 spectra corresponding to 242 peptides (Fig. 13B, supplemental Table S4).

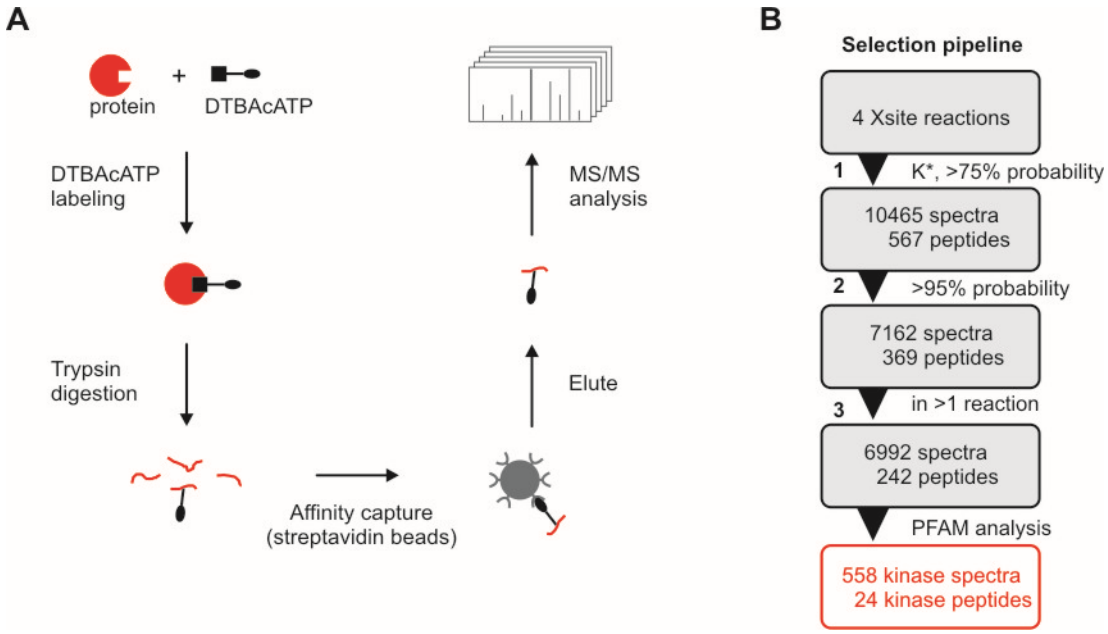


Figure 13. Identification of labelled peptides by Xsite.

(A) Arabidopsis leaf extracts were labelled with DTBAcATP, trypsin-digested and the labelled peptides affinity purified and analyzed by LC-MS/MS. **(B)** Pipeline for analysis of Xsite reactions. A total of four Xsite reactions were analyzed. Peptides were selected for having: 1) labelled Lys with probability score >75%; 2) >95% probability; and 3) found in at least two of the four Xsite reactions. To select peptides from PKs, PFAM analysis was performed on the list of labelled proteins, resulting in a list of 558 spectra corresponding to 24 different peptides from PKs.

2.1.5b DTBAcATP labels a diversity of protein kinases

We first analyzed the data for peptides derived from PKs. Annotation through the PFAM (http://pfam.sanger.ac.uk/) database revealed that 24 of the 242 labelled peptides originate from at least 21 different PKs (Fig. 14A; supplemental Table S4). This list contains seven RLKs of which six carry extracellular LRRs. The detection of RLKs is remarkable since we did not enrich for membranes. In addition, we identified PK peptides from several mitogen-activated protein kinases (MPKs), calcium-dependent protein kinases (CPKs), PTI-like kinases and other Ser/Thr protein kinases. We also detected AvrPphB susceptible 1 (PBS1) (Swiderski & Innes, 2001), VH1-interacting kinase (VIK) (Wingenter et al, 2011), AT6 protein kinase (Gao & Xiang, 2008), and proline-rich extension-like receptor kinase-1 (PERK1) (Silva & Goring, 2002). Of the 24 PK peptides, 19 are unambiguous and are derived from only one PK (black accession numbers in Fig. 14A). Four protein kinase-derived peptides are ambiguous since the peptide sequence is identical in two or more PKs in the TAIR10 protein database (gray accession numbers in Fig. 14A). Two PK-derived peptides were identified for two PKs (top in Fig. 14A), indicating that DTBAcATP labels these PKs at two positions. In case of PERK1, two overlapping peptides containing the same labelled lysine were detected.

We marked the identified PKs in the *Arabidopsis thaliana* kinase database (http://bioinformatics.cau.edu.cn/athKD/index.htm) which classifies all the 1099 PKs of Arabidopsis into classes, groups, and families based on sequence similarities. We identified representatives of 11 different PK families belonging to 10 different groups and the three major PK classes (Fig. 14B). This indicates that AcATP probes are not family specific and can be used to profile ATP binding activities of the majority of Arabidopsis PKs.

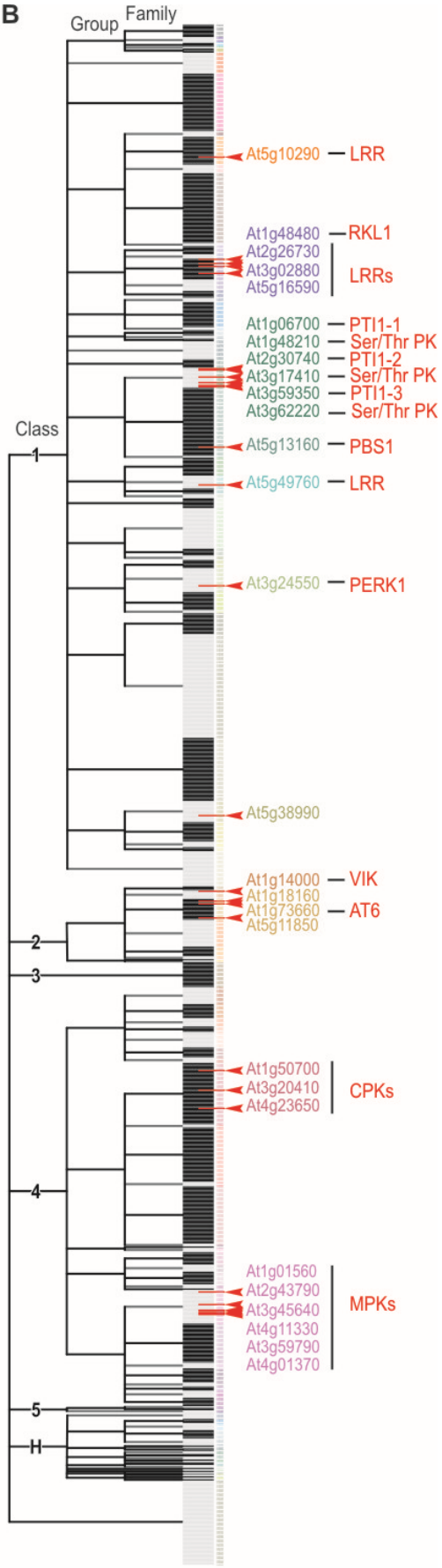
Alignment of the identified PK-derived peptides showed that these peptides fall into four groups containing four different labelled Lys residues: Lys1-4 (Fig. 14C). When mapped on the protein kinase domains, Lys1 is positioned in the beginning of the protein kinase domain, Lys2 in the middle, Lys3 in the second quarter, and Lys4 in the third quarter of the protein kinase domain (Fig. 14A). Lys1 and Lys2 are also part of the conserved protein kinase motifs II and VIb, respectively (Hanks & Hunter,

1995). These two Lys are also present in the Pto kinase, a tomato PK involved in immunity, for which a structure is available (2qkw) (Xing et al, 2007). Both Lys1 and Lys2 reside in the ATP binding pocket of the Pto kinase, with a distance of 7.43 and 4.32 Å from the gamma phosphate of the bound ATP (Fig. 15A). The closer proximity of Lys2 seems reflected in the higher spectral count of Lys2-containing peptides, when compared to Lys1-containing peptides. Modelling of PTI1-1 indicates that Lys3 locates in an extended loop that might fold back on the ATP binding pocket (Fig. 15B) whereas modelling of VIK1 indicates that Lys4 locates still in close proximity to the gamma phosphate of bound ATP (Fig. 15C).

A

Protein	Spectra	Domain structure
At5g16590	35+67	LRR1
At3g02880	57+20	LRR
At1g48480	50	RKL1
At2g26730	9	LRR
At5g49760	11	LRR
At5g10290	4	LRR
At5g38990	3	
At4g01370 +3	46	MPK3/4/6/11
At4g11330	37	MPK5
At3g59790	5	MPK10
At5g13160	14	PBS1
At3g20410 +1	14	CPK9/33
At4g23650	2	CPK3
At1g06700	6	PTI1-1
At2g30740	7	PTI1-2
At3g59350	18	PTI1-3
At3g62220	43	Ser/Thr PK
At3g17410 +1	34	Ser/Thr PK
At1g14000	16	VIK
At3g24550	25	PERK1
At3g24550	23	PERK1
At1g73660 +2	12	AT6

B



C

Accession	Gene name	DBAcATP-labelled peptide
Lys1 (motif II)		
At5g16590	LRR1	ASFDHGLVVAV**K**R
At3g02880	LRR	ASFEHGLVVAV**K**R
At5g38990		GRIDGGATLVAV**K**R
At5g10290	LRR	GVLPDNTKVAV**K**R
Lys2 (motif VIb)		
At5g16590	LRR1	DATTSHGNI**K**SSNILLSESFEAK
At3g02880	LRR	DGTTSHGNI**K**SSNILLSDSYEAK
At1g48480	RKL1	LDY...GNI**K**SSNILLTK
At2g26730	LRR	LVHGNI**K**ASNILLHP...PPNR
At5g49760	LRR	DI**K**SNNILLDENLTAK
At4g01370+3	MPK3/4/6/11	DL**K**PSNLLLNANCDLK
At5g11330	MPK5	DL**K**PSNLLLNSNCDLK
At3g59790	MPK10	DL**K**PSNLLLSTQCDLK
At5g13160	PBS1	DF**K**SSNILLDEGFHPK
At3g20410+1	CPK9/33	DL**K**PENFLLSSKDEK
At4g23650	CPK3	DL**K**PENFLFLSK
At3g62220	Ser/Thr PK	DI**K**SSNVLIFDNDVAK
At3g17410+1	Ser/Thr PK	DI**K**SSNVLLFDDDVAK
At3g24550	PERK1	IIHRDI**K**ASNILIDFKFEAK
At3g24550	PERK1	DI**K**ASNILIDFKFEAK
At1g73660+2	AT6	DL**K**SPNLLVDKNWVVK
Lys3		
At1g06700	PTI1-1	**K**GVQGAQPGPTLDWITR
At2g30740	PTI1-2	**K**GVQGAQPGPTLDWLTR
At3g59350	PTI1-3	**K**GVQGAQPGPTLDWIQR
Lys4		
At1g14000	VIK	LI**K**VQNSHDVYK

Figure 14. Protein kinases labelled with DTBAcATP.

(A) Protein domains are indicated as boxes, with protein kinase domains containing bars representing identified labelled peptides: containing Lys1 (blue), Lys2 (red), Lys3 (green), and Lys4 (black). PKs that match the same peptides are indicated with gray accession numbers and summarized as At4g01370 +3 = At4g01370, At3g45640, At2g43790 and/or At1g01560; At3g20410 +1 = At3g20410 and/or At1g50700; At3g17410 +1 = At3g17410 and/or At1g48210; At1g73660 +2 = At1g73660, At5g11850 and/or At1g18160. Common names of PKs are indicated, as well as corresponding spectral counts. **(B)** Classification of the Arabidopsis PKs according to the Arabidopsis thaliana Kinase Database (AthKD). The identified labelled PKs are indicated with red arrowheads. Accession numbers and common names are shown. Detected PKs represent eleven families that belong to ten groups and three classes. **(C)** Alignment of identified labelled peptides containing four different labelled Lys. Lys1 and Lys2 belong to PK motifs II and VIb, respectively. Colour coding is the same as in (A). Labelled Lys are in bold.

To determine how common these labelled Lys residues are within the PK family of Arabidopsis, we counted these Lys in the alignments of each PK family. We found that Lys1 and Lys2 are conserved in 70% and 47% of the 1099 protein kinases, respectively. Lys3 in PTI kinases and Lys4 in VIK1 are less conserved among PKs (3.5% and 5.0%, respectively). Thus, the frequencies of the theoretically labelled Lys in PKs correspond to the frequencies of the different labelling sites detected in PKs.

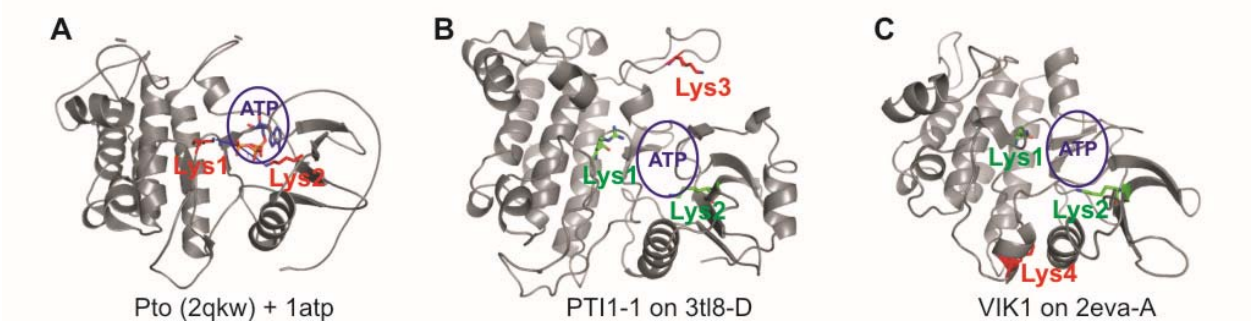


Figure 15. Models of labelled protein kinases show labelling sites close to ATP binding pocket.

(A) Lys1 and Lys2 are in the ATP binding pocket. The structure of the Pto protein kinase (2qkw) was aligned with a protein kinase carrying ATP (1atp). The labelled Lys (red) are close to the ATP (blue) in the Pto kinase (grey cartoon). **(B)** Model of PTI1-1 indicating the position of Lys3 (red). PTI1 1 was modelled on 3tl8-D using SWISS-MODEL. The positions of Lys1 (conserved) and Lys2 (Arg in PTI1-1) are shown in green. Lys3 (red) locates in an extended loop that might fold back on the ATP binding pocket. **(C)** Model of VIK1 indicating the position of Lys4. VIK1 was modelled on 2eva using SWISS-MODEL. Lys1 and Lys2 (conserved in VIK1) are indicated in green. Lys4 (red) is nearby the ATP binding pocket (blue). All models were visualized using PyMol and are shown from the same angle.

2.1.5c DTBAcATP also labels other ATP-binding proteins in their ATP-binding pocket

When ranked according to peptide frequency, peptides from PKs are not among the top 20 most frequent peptides (Fig. 16). The most frequent peptide from a PK is that from LRR1 (At5g16590) on position 22 with 67 spectral counts (Fig. 16 and Supplemental Table S4). The most frequently detected labelled peptide is from phosphoglycerate kinase 1 (PGK1) with 968 spectral counts. PGK1 is also labelled

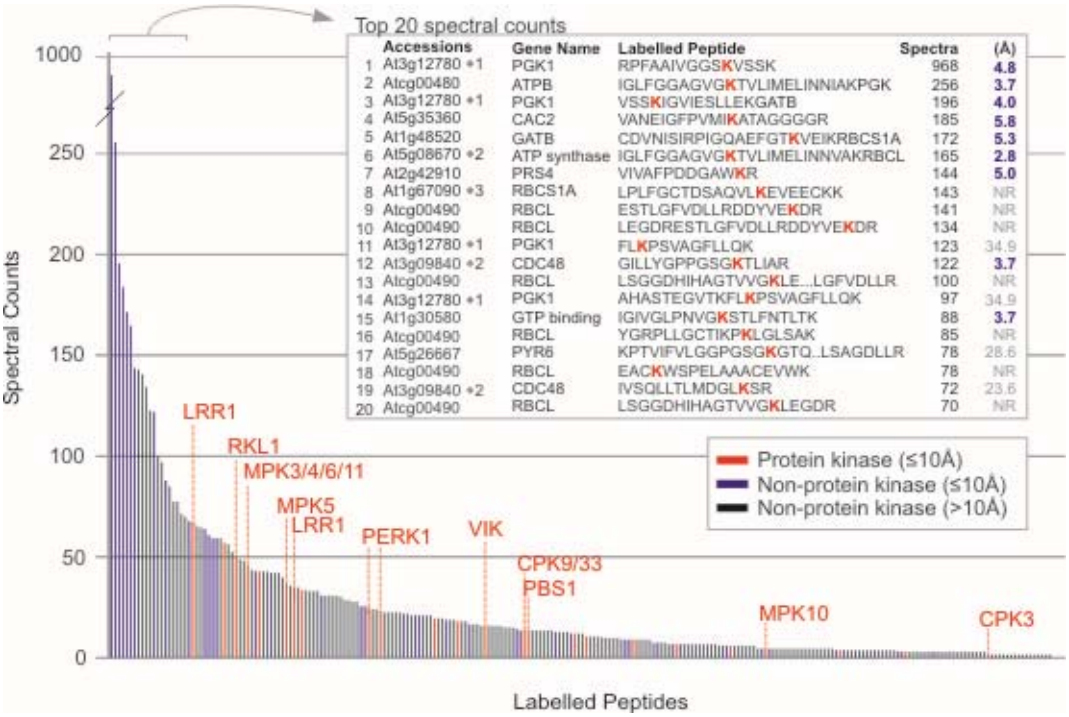


Figure 16. Spectral count analysis of labelled peptides.

Labelled peptides ranked on spectral count frequencies. Labelling sites in the ATP binding pocket of protein kinases and other nucleotide binding proteins are indicated in red and blue, respectively. The top 20 peptides with highest spectral counts are shown in the box, including their accession numbers and their common names, the peptide sequences (labelled Lys, bold) and their spectral counts. Distances of the nucleotide to the labelled Lys are indicated in Å and are derived from models of nucleotide-bound structures. NR, not reported nucleotide binding site. Ambiguous peptides are summarized as follows: At3g12780+1 = At3g12780 and/or At1g56190; At5g08670+2 = At5g08670, At5g08680 and/or At5g08690; At1g67090+3 = At1g67090, At5g38430, At5g38420, and/or At5g38410; At3g09840+2 = At3g09840, At3g53230, and/or At5g03340.

at two additional Lys residues with 196 and 97 spectral counts, respectively. The second most frequent labelled peptide is from chloroplast ATP synthase subunit beta (ATPB) with 256 spectral counts. Other proteins in the top 20 are acetyl coenzyme A (CAC2), subunit B of Glu-tRNA aminotransferase (GATB), an ATP synthase (ATPS),

ribose-phosphate pyrophosphokinase 4 (PRS4), the small subunit of ribulose bisphosphate carboxylase (RBCS1A), cell division cycle 48 (CDC48), a GTP binding protein (At1g30580), and UMP/CMP kinase (PYR6) (Fig. 16). Most, but not all of these proteins were also detected abundantly using the gel-based approach (Fig. 9). Although these proteins are unrelated and catalyze very different reactions, ATP is a common substrate, consistent with the reactivity of the probe (Table 1). Protein structures of homologs of PGK1, ATPB, CAC2, and PRS4 contain the labelled Lys in close proximity to the terminal phosphate of the bound nucleotide (Fig. 17). Note that CAC2 is a biotin binding protein, but the labelled Lys is close to the bound ADP (Fig. 17). This demonstrates that the labelling of these proteins is in accordance to the labelling mechanism of AcATP probes. However, not all proteins in this list are known to bind nucleotides. Most evident is the large subunit of Rubisco (RBCL), which is labelled at 13 different Lys that are scattered on the surface of the RBCL protein (Fig. 18).

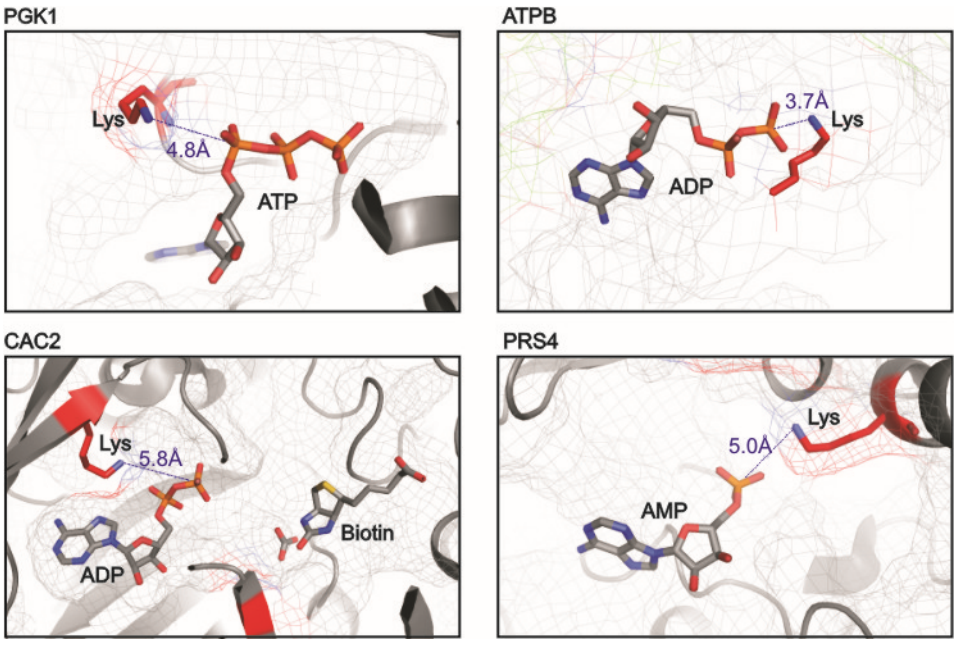


Figure 17. Structural models of four representative proteins PGK1, ATPB, CAC2, and PRS4 showing their nucleotide binding activity.

Structure models were extracted from Protein Data Bank (PDB; http://www.rcsb.org/pdb/home/home.do) database (PGK1, 2PAA; ATPB, 1H8E; CAC2, 3G8C; PRS4, 2HCR). When bound to the proteins, the nucleotide phosphates are positioned in close proximity to the labelled Lys residues. The distances of the labelled Lys to the terminal phosphate of the bound nucleotide are indicated in Å.

Table 1. Biochemistry of reactions catalyzed by some of the most frequently labelled proteins.

These data were extracted from the BRENDA database (http://www.brenda-enzymes.info/index.php4). ATP (printed in bold), is a common substrate for these proteins.

Accession	Gene Name	Spectra	Biochemistry of Reaction
At1g56190	PGK1	968	**ATP** + 3-phospho-D-glycerate → ADP + 3-phospho-D-glyceroyl phosphate
Atcg00480	ATPB	256	**ATP** + H_2O → ADP + phosphate
At5g35360	CAC2	185	**ATP** + biotin + apo-[acetyl-CoA:carbon dioxide ligase] → AMP + diphosphate +[acetyl-CoA:carbon dioxide ligase]
At1g48520	GATB	172	**ATP** + glutamine + tRNAGln → AMP + diphosphate + L-glutaminyl-tRNAGln
At5g08670	ATP synthase	165	**ATP** + H_2O + H^+/in → ADP + phosphate + H^+/out
At2g42910	PRS4	144	**ATP** + D-ribose 5-phosphate → AMP + 5-phospho-alpha-D-ribose 1-diphosphate
At3g09840	CDC48	122	**ATP** + H_2O → ADP + phosphate

2.1.5d Potential of DTBAcATP labelling in Arabidopsis proteome

We extended the analysis of labelling sites in homologous protein structures and found that 63 of the 242 labelled peptides result from labelling of a Lys in known nucleotide binding pockets at ≤10Å from the phosphate of the bound nucleotide (Fig.19, blue in Fig. 16). To estimate the potential of DTBAcATP labelling in the Arabidopsis proteome, we retrieved for each of the 63 labelled peptides the corresponding PFAM domain. A total of 26 different PFAM domains corresponding to 13 different protein clans plus six uncategorized PFAM families were identified (Table 2). The Arabidopsis genome encodes for a total of 1683 proteins carrying these 26 different PFAM domains. The largest PFAM domains are those of PKs (PF00069 = 813 members and PF07714 = 388 members), followed by AAA ATPases (147 members). Since all these 1683 proteins are probably binding ATP and representatives were found to be labelled by DTBAcATP, we predict that the Arabidopsis proteome contain at least 1683 putative targets that could be labelled in the ATP binding pocket by DTBAcATP (Table 2).

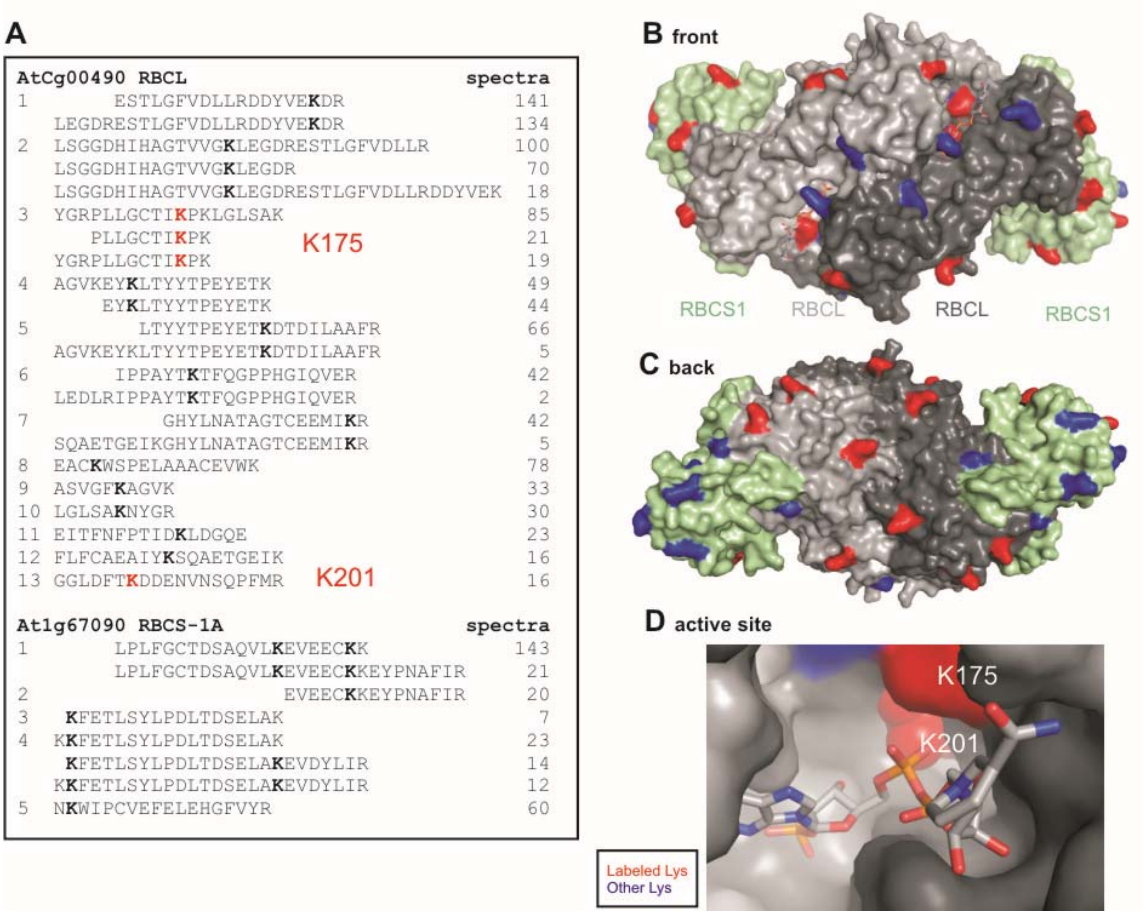


Figure 18. Labelling sites of DTBAcATP on Rubisco.

(A) There are 22 different labelled RBCL peptides identified, which correspond to at least 13 different labelled Lys (bold). Likewise, eight different RBCS-1A peptides correspond to at least 5 different labelling sites. Spectral counts for each labelled peptide are indicated on the right. **(B-D)** Crystal structure of the tetrameric Rubisco comlex of rice with NADPH (3axk). Labelled Lys identified in Arabidopsis are indicated in red and other Lys in blue. Shown are the front (B), back (C) and active site (D) of the tetramer containing two large subunits (RBCL, light and dark grey) and two small subunits (RBCS1, light green), bound to NADPH (sticks). Labelled K175 and K201 (red) are at 3.97 and 6.46 Å from the phosphates of NADPH, respectively.

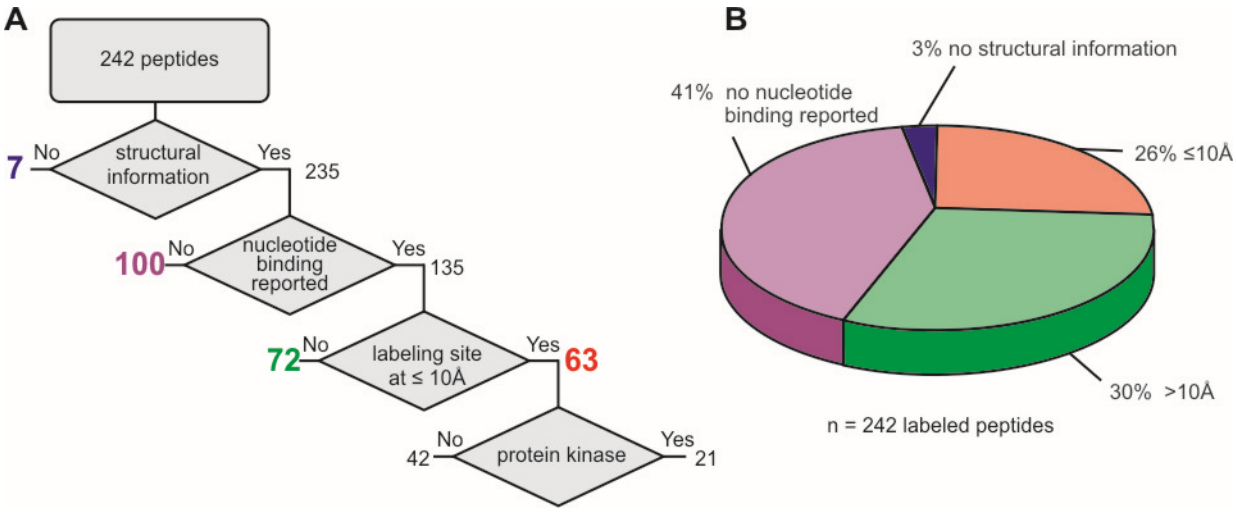


Figure 19. Analysis of labelling sites.

(A) Selection pipeline showing the stepwise analysis and decision channel in determining the expected labelling of peptides. **(B)** Pie distribution diagram of the labelled peptides categorized into four groups: no structural information available; no nucleotide binding reported; and labelling detected at ≤10Å; or >10Å from the phosphate of the nucleotide.

2.1.5e Preferential labelling of ATP-binding proteins

To investigate if PKs and other ATP-binding proteins are preferentially labelled when compared to highly abundant proteins, we plotted the spectral counts of the labelled peptides against the spectral counts detected for these proteins in leaf proteomes of the AtProteome database (Baerenfaller et al, 2011) (Fig. 20A). We chose this approach to ensure the inclusion of low abundance PKs, which were detected by Baerenfaller and colleagues with over 1300 MS runs (Baerenfaller et al, 2008; Baerenfaller et al, 2011). The leaf proteomes that they used are virtually the same as we used for our studies. Interestingly, this splits the labelled peptides into three groups: peptides from PKs cluster in group 1 at low protein levels and medium labelling. Peptides from other ATP-binding proteins are in group 2 and show medium protein levels and medium to high labelling. Finally, the third major group contains proteins with high protein levels and low to medium labelling (Fig. 20A). Importantly, the vast majority of these proteins were not labelled in ATP binding pockets. This illustrates that DTBAcATP shows a high selectivity for labelling ATP-binding proteins because only highly abundant proteins are labelled outside known ATP binding pockets.

Table 2. Clan and domain classification of peptides labelled at the ATP-binding pockets of ATP-binding proteins. PFAM domains were identified for the 63 labelling sites located at ≤10Å to the bound nucleotide. The number of identified labelled peptides and the number of sequence occurrences in the Arabidopsis thaliana genome are indicated in the table. PDB and BRENDA databases were searched to assign whether (+) or not (-) members of the PFAM were reported to bind nucleotides. GO codes assign molecular functions for each PFAM and confirm nucleotide binding activity.

Clan	PFAM no.	PFAM name	No. of labelled peptides	No. of sequences in Arabidopsis	Nucleotide binding	Gene Ontology
CL0016	PF01633.14	Choline_kinase	1	6	+	GO:0016773
	PF00069.19	Pkinase	11	813	+	GO:0005524; GO:0004672; GO:0006468
	PF07714.1	Pkinase_Tyr	10	388	+	GO:0004672; GO:0006468
CL0021	PF08206.5	OB_RNB	1	ND	ND	
CL0023	PF00004.23	AAA	3	147	+	GO:0005524
	PF00406.16	ADK	1	15	+	GO:0005524; GO:0019205
	PF00006.19	ATP-synt_ab	8	11	+	GO:0015991; GO:0016820
	PF01926.17	MMR_HSR1	1	32	+	GO:0005525
CL0039	PF00582.20	Usp	1	53	-	GO:0006950
CL0040	PF03099.13	BPL_LplA_LipB	1	10	-	GO:0003824
CL0063	PF02826.13	2-Hacid_dh_C	1	11	+	GO:0016616; GO:0048037
CL0099	PF00171.1	Aldedh	3	23	-	GO:0016491
CL0110	PF00483.17	NTP_transferase	1	17	+	GO:0016779
CL0137	PF03767.8	Acid_phosphat_B	1	12	+	GO:0003993
CL0179	PF08442.4	ATP-grasp_2	2	4	+	GO:0005524
	PF02786.11	CPSase_L_D2	3	7	+	GO:0005524; GO:0003824
	PF01071.13	GARS_A	1	1	ND	
CL0254	PF00676.14	E1_dh	1	15	-	GO:0016624
CL0266	PF00169.23	PH	1	26	-	GO:0005515
CL0270	PF00180.14	Iso_dh	1	15	ND	
No_clan	PF00571.22	CBS	1	38	-	GO:0005515
	PF02934.9	GatB_N	2	4	+	GO:0016874
	PF00334.13	NDK	1	7	+	GO:0005524; GO:0004550
	PF00162.13	PGK	4	5	+	GO:0004618
	PF00156.21	Pribosyltran	1	22	+	GO:0009116
	PF01259.12	SAICAR_synt	1	1	+	GO:0004639
		Total	**63**	**1683**		

Interestingly, PKs were detected at a moderate labelling/protein ratio compared to labelling events in ATP binding pockets of non-PKs [0.5±0.57(n=15) vs. 2.86±0.68(n=34)] (Fig. 20B). These data suggest that PKs have a lower affinity to ATP or that part of the PK pool cannot be labelled. To investigate this in more detail, we extracted spectral counts for 494 PKs from leaf proteomes of the AtProteome database (Baerenfaller et al, 2008). Interestingly, of the top-20 PKs that are most frequently detected in leaf proteomes, only eight were detected upon labelling (Fig. 21), and these data show no correlation between spectral counts in the AtProteome database and spectral count frequency upon DTBAcATP labelling. Furthermore, inspection of PK sequences revealed that most of the undetected labelled peptides fall in the MW range that should have been detected in our assay. These data re-enforce the idea that not all kinases can be labelled in a given extract, possibly because they do not bind ATP. However, such data should be interpreted with extreme caution since different peptides have different ionization potential and the absence of a labelled peptide does not mean that the protein is not labelled.

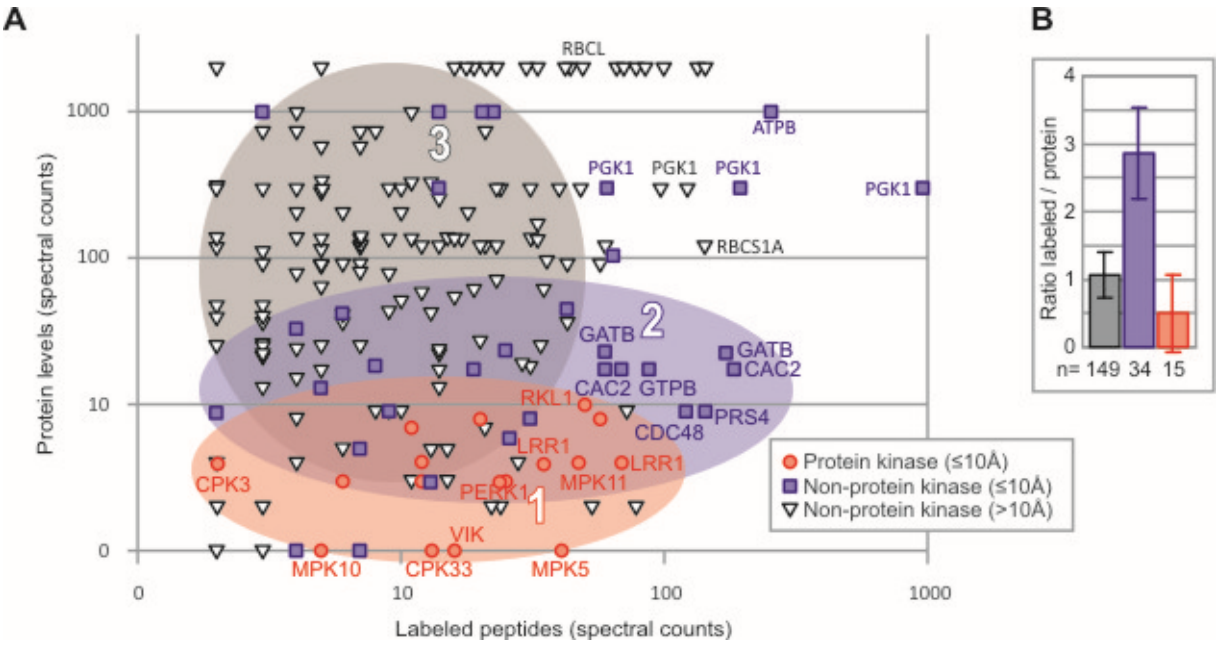


Figure 20. Correlation analysis between ATP binding activity and protein abundance.

(A) Correlations between ATP binding activity and protein abundance splits the labelled peptides into three groups. The frequency of spectral counts of the labelled peptides is plotted against the spectral counts of the corresponding proteins detected in leaf proteomes (extracted from the AtProteome database). For ambiguous labelled peptides, the protein with the highest spectral count was selected from the AtProteome database. Note that for some proteins (e.g. RBCL) multiple labelled peptides were detected. Group numbers 1-3 are explained in the main text. **(B)** Average spectral count ratio between labelling (by DTBAcATP) and abundance (extracted from AtProteome database) for the three groups of labelled proteins. Error bars represent SEM for n proteins, indicated below the bars.

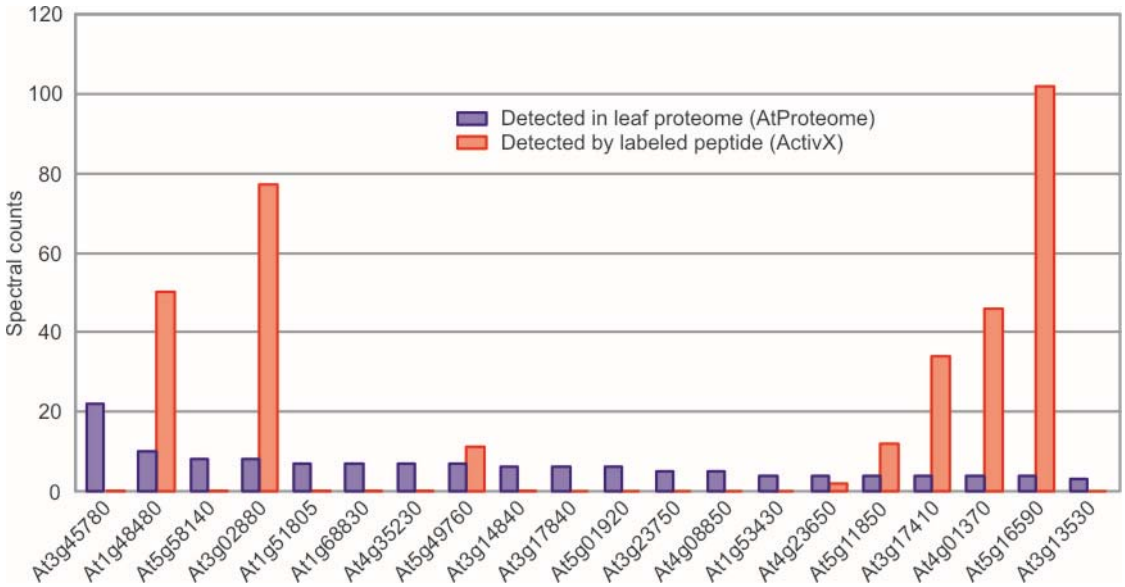


Figure 21. Not all protein kinases detected in leaf proteomes are detected upon AcATP labelling.

Spectral counts for 494 protein kinases were extracted from the AtProteome database. The 20 PKs with highest spectral counts in leaf proteomes are shown in blue. The number of spectral counts of detected DTBAcATP-labelled peptides is shown in red.

2.2 Targeted MS/MS analysis of DTBAcATP-labelled peptides

AcATP labelling proved to be a powerful tool in detecting ATP-binding activities of ATP-binding proteins including kinases, as illustrated above (Chapter 2.1). We hypothesize that labelling not only reflects the abundance of proteins with ATP activity, but also reflects the active state (active or inactive) of ATP-binding proteins. ATP binding is dynamic, and is dependent on the cellular state. The inactive Insulin receptor kinase (IRK) for example, has a collapsed activation loop conformation so that the ATP nucleotide cannot bind the P-loop and the peptide substrate cannot bind the active site. However, upon activation by phosphorylation of the three tyrosine residues in the activation loop, the activation loop opens up allowing ATP and substrates to bind (Huse & Kuriyan, 2002). We therefore reason that active kinases bind ATP more effectively than inactive kinases, and that the turnover of ATP to ADP, and thus continuous binding to new ATPs, is more rapid. To test this hypothesis, we study the activation of PKs upon external stimuli, together with targeted MS/MS analysis to boost the detection of PKs. In targeted MS/MS, the sensitivity of detecting a particular ion is increased tremendously by using an inclusion list coupled to retention times and searches for marker ions in the MS2 mode. Patricelli and colleagues (2011) for example, have increased their detection of PKs to about 160 PKs per cell line using targeted MS/MS analysis.

Here we apply this strategy to detect plant kinases that show differential labelling upon external stimulus. To induce a kinase signaling pathway leading to immunity, we used flg22, a 22-amino acid peptide present in bacterial flagellin. Flg22 binds to the RLK Flagellin Sensitive 2 (FLS2) on the plant plasma membrane. This induces the heterodimerization of FLS2 with the LRR-RLK Brassinosteroid Insensitive 1-Associated Receptor Kinase (BAK1), which triggers their reciprocal phosphorylation. The activated FLS2-BAK1 complex then phosphorylates the bound receptor-like cytoplasmic kinase (RLCK) Botrytis-induced Kinase 1 (BIK1), which results in transphosphorylation of FLS2-BAK1 and further phosphorylation of BIK1 and its concomitant release from the complex (Gao & Xiang, 2008; Newman et al, 2013). These signaling events further proceed to the activation of MAP kinase kinase kinase MEKK1, which activates MAP kinase kinases MKK1/2/4/5, which then activates MAP kinases MPK3/4/6/11 (Fig. 22A-B) (Asai et al, 2002; Bethke et al,

2012; Gao & Xiang, 2008). This cascade of tightly regulated events happens almost instantly upon elicitation. It has been reported that FLS2 and BAK1 activation happens as early as 5 seconds (Schulze et al, 2010), and MPK3/6 activation within 5-30 minutes (Bethke et al, 2012; Galletti et al, 2011) after elicitation.

In this study, Arabidopsis cell cultures were treated with or without flg22 for 15 minutes, when MPKs are expected to be activated. After treatment, the medium was removed and the cells were immediately frozen in liquid nitrogen. Cell cultures were used to detect more kinases as they have less Rubisco proteins than Arabidopsis plants or seedlings. Our previous experiment suggests that the abundant Rubisco in the samples has reduced the detection of low abundant kinases. To confirm the activation of MPK3 and MPK6 upon flg22 elicitation, we detected phosphorylated MPK3/6 on a protein blot using phosphopeptide antibodies. As expected, these MPKs are differentially phosphorylated in the flg22-treated sample when compared to the control (Fig. 22C). DTBAcATP labelling was done in duplicate for each sample, followed by trypsin digestion, and purification of desthiobiotinylated peptides using streptavidin beads. In the targeted MS/MS approach, the samples were first run on the MS to detect as many labelled PK peptides as possible, adding labelled peptides from other ATP-binding proteins to fill an inclusion list of 250 peptides. The peptides were then quantified in a second run using the inclusion list. Supplementary Table 5 shows the results of the targeted MS/MS analysis.

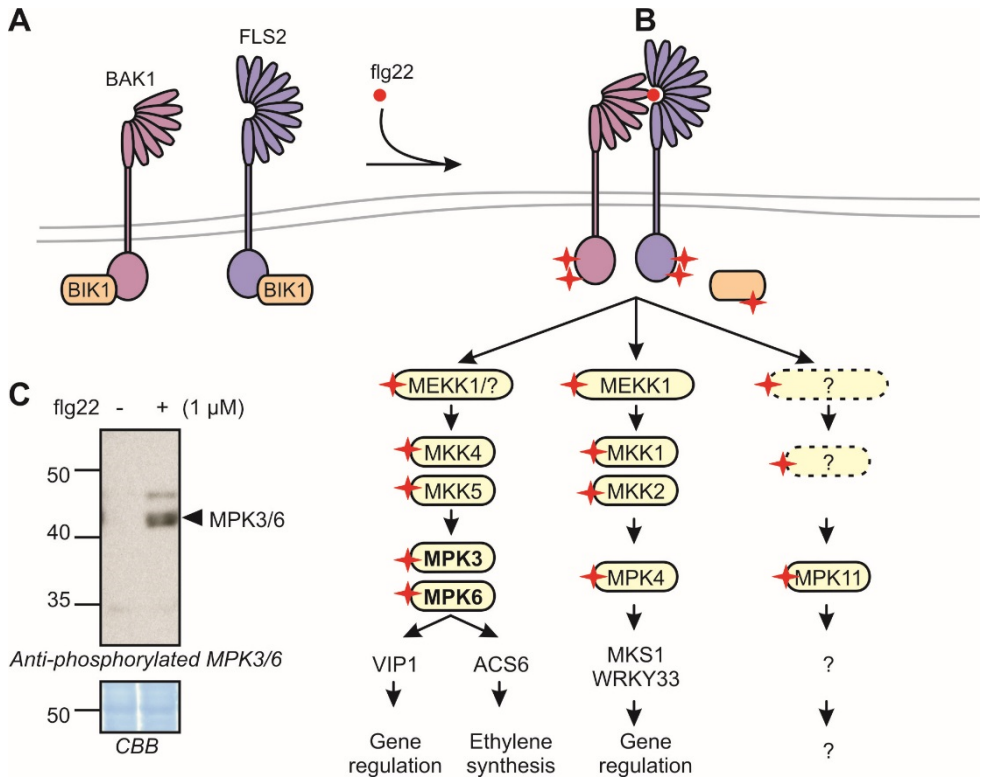


Figure 22. Elicitation with flg22 in Arabidopsis activates many kinases.

(A) In the absence of flg22, BIK1 associates with inactive FLS2 and BAK1. **(B)** Upon flg22 elicitation, flg22 binds and activates FLS2. The activated FLS2 then forms a complex with BAK1, and induces phosphorylation of BAK1. The activated BAK1 phosphorylates BIK1, which in turn transphosphorylates the FLS2–BAK1 complex, which may further phosphorylate BIK1. The activated BIK1 is released from the FLS2-BAK1 complex to activate downstream signalling cascade. The downstream MAPK cascade involves MAPK kinase kinases (MEKKs) which activate downstream MAPK kinases (MKKs), which then activate downstream MAP kinases (MPKs). The activated MPKs then phosphorylate downstream substrates that regulate gene expression. There have been four MPKs shown to be activated by flg22. These are MPK3 and MPK6 which are activated by MKK4/MKK5 (Asai et al, 2002), MPK4 which are activated by MKK1/MKK2 (Gao et al, 2008), and MPK11 (Bethke et al, 2012). Red circle represents flg22 and red star indicates phosphorylation. MAPKs in bold are used as a reference in the targeted MS/MS analysis. **(C)** Western blot analysis of Arabidopsis cell culture treated with or without flg22 for 15 minutes, using anti-phospho p44/p42 MAPK antibody.

2.2.1 Targeted MS/MS analysis increased the detection of protein kinases

MS analysis, summarized in Supplementary Table 5, reveals that the targeted approach have doubled the identified kinase peptides to 46 as compared to the data-dependent approach used previously (Supplementary Table 4), which displayed 24 kinase peptides. Among the 22 PKs that were newly detected in the targeted approach are EDR1 (ENHANCED DISEASE RESISTANCE 1), lectin kinases AT1G15530, AT5G60320 and AT3G45430, Ser/Thr PK AT1G20650, CPK19

(CALMODULIN-DOMAIN PROTEIN KINASE 9), TMK1 (TRANSMEMBRANE KINASE 1), LRR proteins AT1G68400 and AT2G24230, CBL-interacting protein kinases CIPK10 and CIPK12, ABC1 protein AT4G24810, and uncharacterized PKs AT1G70520, AT2G07180, AT2G17700, AT3G46920, AT3G58640, AT4G10730, AT5G14720, AT5G18500, AT5G18610, and AT5G54590 (Supplementary Table 5). Moreover, all the 46 PK peptides identified in the targeted MS/MS analysis are mapped to 5 different positions in their protein kinase domains. Among these, 12 are classified as Lys1, 25 Lys2, three Lys3, one Lys4, three are other than Lys1-4 within the kinase domain, and two are Lys within the activation loop. (Supplementary Table 5). Having been mapped to the kinase domains explains the labelling of these peptides.

Similar to the previous experiment, apart from detecting PKs, we also detected other ATP-binding proteins. The majority of these other ATP-binding proteins have been previously detected, with only some newly labelled proteins. However, there are less identified labelled peptides from ATP-binding proteins other than PKs in the targeted approach (93 peptides) compared to the data-dependent approach (218 peptides) (Supplementary Tables 5 and 4, respectively). In the data-dependent approach, there were five and 13 labelling sites in RBCS1A and RBCL, respectively (Fig. 18), but only one labelled peptide for each of these proteins were detected in the targeted approach (Supplementary Table 5). Alignment of Arabidopsis RBCS1A and RBCL proteins with that of the structure of rice Rubisco (PDB entry: 3axk), shows that these labelled peptides are on the surface of the protein, at 28 and 38.7 Å respectively, from the bound NADPH (data not shown). Obviously, the reduced labelling sites observed in the targeted MS/MS experiment is probably due to the use of Arabidopsis cell cultures which contains less RBCL protein when compared to Arabidopsis seedlings.

2.2.2 DTBAcATP labelling is not affected by the phosphorylation state of proteins

In this study, all quantitation was performed by extracting characteristic fragment ion signals from targeted MS/MS spectra and comparing signals in control and treated samples. Two technical replicates were done, and for each peptide quantitated, the change in MS signal for the treated samples relative to the MS signal for the control samples is expressed as fold change. The ratio between flg22 and control samples and their P values were also calculated (Supplementary Table 5). Supplementary Table 5 shows that the vast majority of the ion intensities of the identified peptides in the targeted MS/MS analysis are very similar in both control and flg22 samples, indicating that these biological samples are very similar. To identify changes between the control and flg22 samples, a volcano plot was generated, which plots significance against fold-change in the dataset. Significance thresholds P=0.05 and P=0.01 were indicated as dotted lines and the fold-change threshold is indicated as grey area. Labelled peptides deemed to be significantly labelled lies outside these threshold marks. Volcano plot analysis reveals that there are no significant differences between the control and flg22 treatments. Unfortunately, the high similarity also extends to proteins like MPK3 and MPK6, where we expected differential labelling because these proteins are differentially phosphorylated upon flg22 elicitation compared to the control, based on the protein blot using phosphopeptide antibodies (Fig. 22C). Among the other kinases known to be involved in flg22 signalling (Fig. 22B), only MPK4 and MPK11 were detected, however not significantly changed between samples. These data tell that changes in the phosphorylation state of the MPK proteins do not significantly affect AcATP labelling.

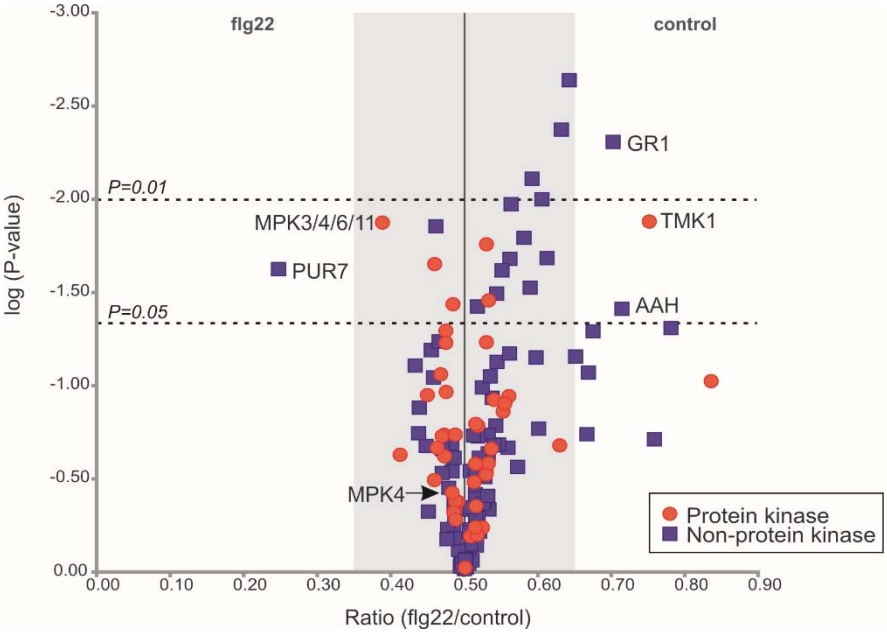


Figure 23. Phosphorylation state of proteins do not significantly affect DTBAcATP labelling.

Volcano plot analysis of peak intensities of labelled peptides identified by targeted MS/MS analysis. The graph plots the ratio of flg22 treatment over control against the logarithm of P-values. Gray area indicates threshold ratio between flg22 and control, while dotted lines indicate P=0.05 and P=0.01 significance thresholds. Proteins with peptide intensities above the set threshold values are highlighted, as well as detected proteins known to be involved in flg22 signalling. MPK3/4/6/11 represents the labelled peptide DLKPSNLLLNANCDLK (Supplementary Table 5) which is an ambiguous peptide shared by the four MAPKs. A second labelled peptide for MPK4 is detected and indicated by an arrow. Abbreviations: MPK3/4/6/11, Mitogen Activated Protein Kinase 3/4/6/11; AAH, Adenine nucleotide alpha hydrolases-like superfamily protein; TMK1, TRANSMEMBRANE KINASE 1;GR1, Glutathione-disulfide Reductase 1; PUR7 or PURIN7, Phosphoribosylaminoimidazolesuccinocarboxamide Synthase 7.

On the other hand, AcATP labelling of Glutathione-disulfide Reductase 1 (GR1), Transmembrane kinase 1 (TMK1), and Adenine nucleotide alpha hydrolases-like superfamily protein (AAH) were significantly increased with 2-, 3- and 2.5-fold, respectively, upon flg22 treatment, whereas Phosphoribosylaminoimidazolesuccinocarboxamide Synthase 7 (PUR7) labelling is 3-fold reduced upon flg22 treatment. AAH may be involved in stress responses (Interpro electronic annotation, http://www.ebi.ac.uk/interpro/). TMK1 is an RLK, with an extracellular LRR domain and a serine-threonine autophosphorylation activity (Chang et al, 1992; Schaller & Bleecker, 1993). Moreover, GR1 may be involved in the photoperiod-influenced responses in Arabidopsis plants subjected to ozone such as lesion formation, SA accumulation, and Pathogenesis-related 1 (PR1) gene induction (Dghim et al, 2013). And PUR7 is expressed in mitotic tissues in

Arabidopsis and encodes for 5'-phosphoribosyl-4-(N-succinocarboxamide)-5-aminoimidazole synthetase involved in purine biosynthesis (Senecoff et al, 1996). Differential labelling suggests that these four proteins may be involved in flg22-influenced responses in plant immunity, and could be further investigated.

2.2.3 DTBAcATP labels both MPK3/6 phosphomimic and phosphomutants

We next investigated the rather puzzling finding that phosphorylation of proteins does not affect AcATP labelling. We reasoned that active kinases have a higher ATP turnover rate when compared to inactive kinases because they phosphorylate proteins, and intuitively, we expected that active kinases are more likely to react with DTBAcATP than kinases that do not hydrolyse the bound ATP. However, during our *ex vivo* labelling, the absence of differential MPK labelling might be because we remove the ATP from both active and inactive kinases by gel filtration prior to labelling. We therefore designed an experiment where both ATP and a substrate are present during labelling with DTBAcATP. We focused this case study on MPK3 and MPK6 because they are well studied MPKs, and that expression constructs have been previously made and reported (Sorensson et al, 2012; Ulm et al, 2002).

The GST-MPK3 and GST-MPK6 were expressed in BL21 *E. coli* strain by IPTG induction, and bacterial cells were lysed. We first tested DTBAcATP labelling of lysates and results show the expected stronger labelling in induced samples, when compared to uninduced samples (data not shown). However, to better assess the labelling potential of differentially active MPK3 and MPK6, we applied phosphomimetics where an amino acid replacement chemically mimics a phosphorylated residue or unphosphorylated residue resulting to an active or inactive proteins, respectively. We generated different protein variants of GST-MPK3 and GST-MPK6 by site-directed mutagenesis (Fig. 24A). Phosphomimic and phosphomutant versions of MPK3 and MPK6 were generated by exchanging the two phosphorylation sites threonine (T) and tyrosine (Y) to aspartate (D) and alanine (A), respectively (Beckers et al, 2009; Matsuoka et al, 2002; Yang et al, 2001). In addition, kinase inactive GST-MPK fusion proteins were generated by exchanging the conserved lysine (K) residue in the ATP binding domains to alanine (A) (Cho et

al, 2008; Zhang & Liu, 2001). This conserved lysine is also targeted by AcATP (Fig. 1B).

The GST-MPK fusion proteins were then expressed in BL21 *E. coli* strain. Coomassie brilliant blue staining of proteins shows induction of expression of the GST-MPK3 and GST-MPK6 proteins and their mutated versions with IPTG as compared to no IPTG controls (Fig. 24B). DTBAcATP labelling of the GST-MPK fusion proteins was detected using streptavidin-HRP (Fig. 24C). To quantify labelling differences between samples, signal intensities were measured using ImageJ and the ratio between streptavidin-HRP and coomasie signals were calculated, subtracting the background signal derived from the no IPTG samples from the signals of IPTG-treated samples and these were plotted in a graph (Fig. 24C). Signal analysis shows no significant labelling differences between MPK3 WT, MPK3 phosphomimic T196D T198D, and MPK3 phosphomutant T196A T198A, as well as in MPK6 WT compared to MPK6 phosphomimic T221D T223D (Fig. 24C). More importantly, there is reduction of labelling of MPK3 and MPK6 Lys mutants compared to the WT controls (Fig. 24C). This confirms that AcATP labelling requires the conserved Lys in the ATP binding domain of kinases and that AcATP labelling is insensitive to the phosphorylation state of proteins.

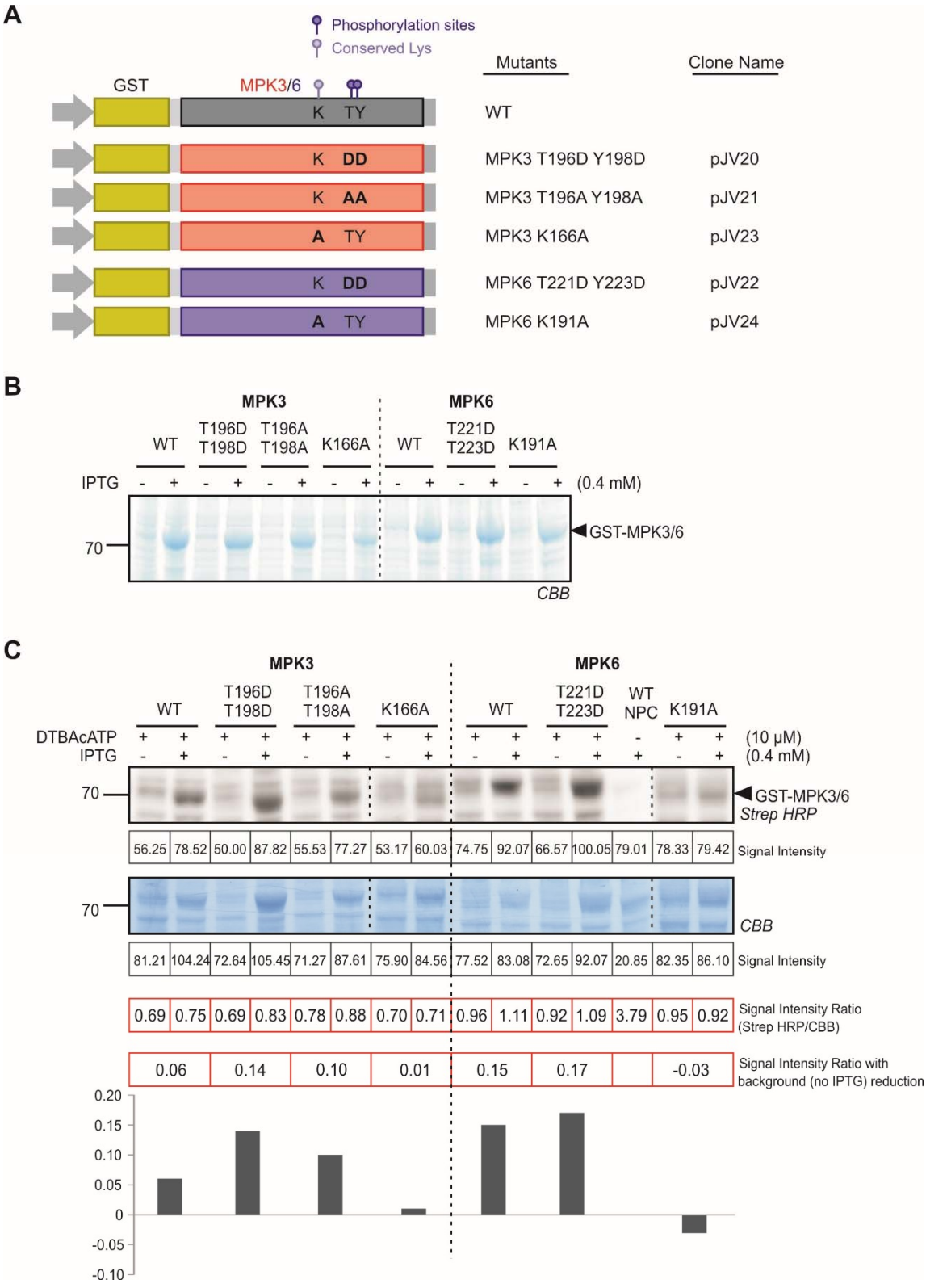


Figure 24. DTBAcATP labelling of MPK3/6 wildtype and mutant proteins

(A) Site-directed mutagenesis generating different mutant variants of MPK3 (red) and MPK6 (blue) proteins in IPTG-inducible pGEX-3X cloning and expression vector. MPK3 or MPK6 is fused to a GST tag (yellow) in their N-terminus behind a Tac promoter.The proteins are expressed in BL21 *E. coli*. The mutated residues are in bold letters. **(B)** Expression and induction of GST-MPK3/6 fusion proteins with and without 0.4 mM IPTG for three hours at 28 ºC. **(C)** Labelling of GST-MPK3 and

GST-MPK6 proteins with DTBAcATPfor one hour. Labelled proteins were detected in a protein blot using streptavidin-HRP. Coomassie brilliant blue (CBB) shows loading control. Signal intensities on the blot and CBB loading control were quantified using ImageJ and the ratios between streptavidin-HRP signals and CBB signals were indicated below each gel lane. Calculated signal intensities of no IPTG control samples were subtracted from IPTG-treated samples, and were plotted in a graph. NPC indicates no-probe-control.

2.3 Putative receptor shedding of RLK902

We noticed from our AcGTP labelling experiments (data not shown) that the experimental MW of RLK902 (At3g17840) does not correspond to its theoretical MW (70.4 kDa). This observation is similar to what has been observed previously in AcATP labelling experiments (Fig. 11 of Chapter 2.1). Like At3g02880 and At5g16590, RLK902 is a member of the LRR-RLK family of proteins, which carry an extracellular LRR domain and an intracellular kinase domain (Fig. 25). Noteworthy, these three LRR-RLKs are close homologues all belonging to LRR III subfamily (Fig. 26, red arrowheads; Arabidopsis thaliana kinase database, http://bioinformatics.cau.edu.cn/athKD/index.htm; iTAK - Plant Transcription factor & Protein Kinase Identifier and Classifier, http://bioinfo.bti.cornell.edu/cgi-bin/itak/index.cgi). From our MS analyses, peptide coverage suggests that the labelled protein is a full kinase domain lacking the LRR domain. Interestingly, the size of the kinase domain together with the transmembrane domain and the intracellular juxtamembrane domain coincides with the observed 42 kDa MW of the protein (Fig. 25). This suggests that putative receptor shedding may be a common phenomenon for this LRR-RLK subfamily.

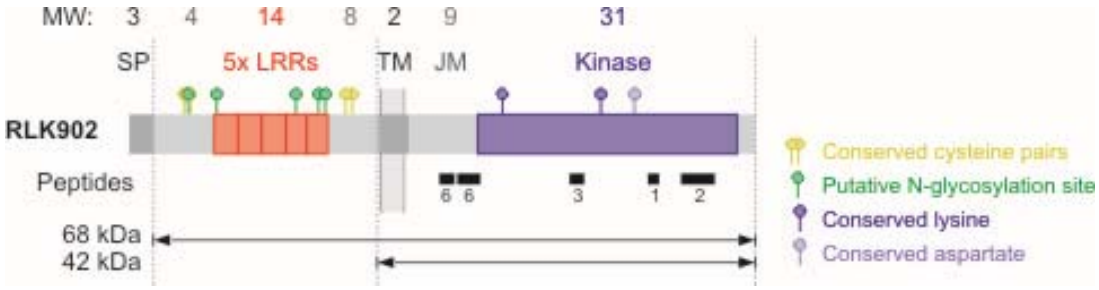


Figure 25. Putative receptor shedding of RLK902.

RLK902 carries a signal peptide (SP), an N-terminal LRR-flanking domain with a conserved Cys pair; five LRRs (red), an extracellular juxtamembrane domain with a conserved Cys pair (yellow); a transmembrane domain (TM), an intracellular juxtamembrane domain (JM); a protein kinase domain (blue) and a C-terminal tail. The extracellular domain carries five putative *N*-glycosylation sites (green). Identified peptides and spectral counts are indicated with black bars. The theoretical MW is calculated from the plain amino acid sequence for the cytoplasmic domain, and the full mature protein without SP.

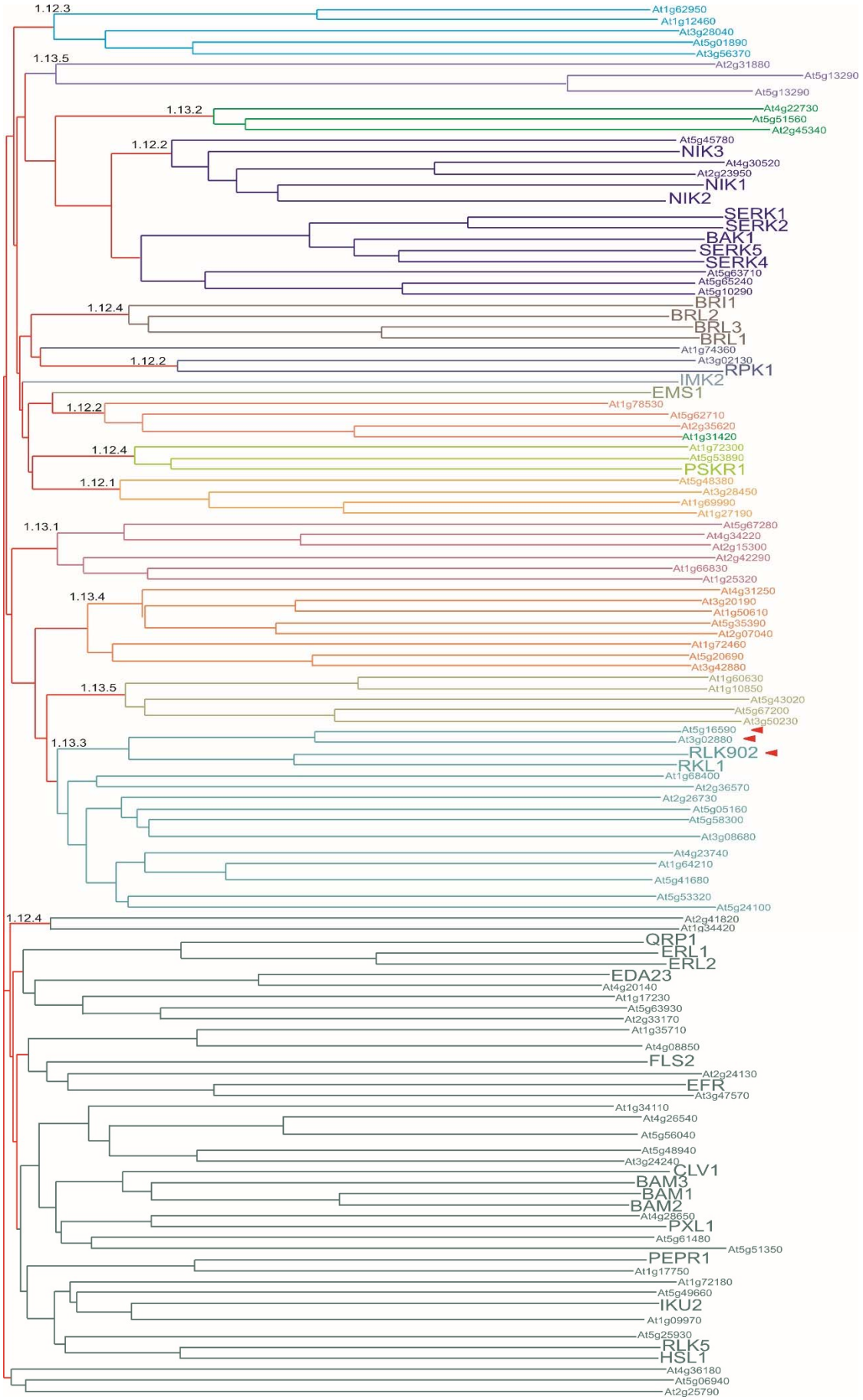

1.12.3
At1g62950
At1g12460
At3g28040
At5g01890
At3g56370
1.13.5
At2g31880
At5g13290
At5g13290
1.13.2
At4g22730
At5g51560
At2g45340
1.12.2
At5g45780
NIK3
At4g30520
At2g23950
NIK1
NIK2
SERK1
SERK2
BAK1
SERK5
SERK4
At5g63710
At5g65240
At5g10290
1.12.4
BRI1
BRL2
BRL3
BRL1
At1g74360
1.12.2
At3g02130
RPK1
IMK2
EMS1
1.12.2
At1g78530
At5g62710
At2g35620
At1g31420
1.12.4
At1g72300
At5g53890
PSKR1
1.12.1
At5g48380
At3g28450
At1g69990
At1g27190
1.13.1
At5g67280
At4g34220
At2g15300
At2g42290
At1g66830
At1g25320
1.13.4
At4g31250
At3g20190
At1g50610
At5g35390
At2g07040
At1g72460
At5g20690
At3g42880
At1g60630
1.13.5
At1g10850
At5g43020
At5g67200
At3g50230
At5g16590
1.13.3
At3g02880
RLK902
RKL1
At1g68400
At2g36570
At2g26730
At5g05160
At5g58300
At3g08680
At4g23740
At1g64210
At5g41680
At5g53320
At5g24100
1.12.4
At2g41820
At1g34420
QRP1
ERL1
ERL2
EDA23
At4g20140
At1g17230
At5g63930
At2g33170
At1g35710
At4g08850
FLS2
At2g24130
EFR
At3g47570
At1g34110
At4g26540
At5g56040
At5g48940
At3g24240
CLV1
BAM3
BAM1
BAM2
At4g28650
PXL1
At5g61480
At5g51350
PEPR1
At1g17750
At1g72180
At5g49660
IKU2
At1g09970
At5g25930
RLK5
HSL1
At4g36180
At5g06940
At2g25790

Figure 26. Phylogenetic tree of LRR-RLKs in Arabidopsis.

Protein sequences of kinases belonging to groups 1.12 and 1.13 were retrieved from the Arabidopsis thaliana kinase database (http://bioinformatics.cau.edu.cn/athKD/index.htm) and phylogenetic tree was constructed with ClustalX2. Red arrowheads mark the three LRR-RLKs observed to undergo putative receptor shedding events in our assays.

Receptor shedding refers to the release of extracellular domains of transmembrane proteins or receptor proteins like cytokines and growth factors. It is an essential post-translational modification reported in animals and humans, commonly called ectodomain shedding. Proteolytic enzymes belonging to the ADAM (A Disintegrin and Metalloproteinase) family are responsible for the shedding of a number of transmembrane proteins. Ectodomain shedding activates the transmembrane proteins and is implicated in development, homeostasis, and pathological conditions and inflammation (Horiuchi, 2013; Weber & Saftig, 2012; Xu & Derynck, 2010).

Receptor shedding has not been well studied in plants. We therefore investigated possible receptor shedding events of the three LRR-RLKs mentioned above. We focused our study of receptor shedding on RLK902, an RLK capable of autophosphorylation, expressed strongly in root meristem, lateral root primordia, stipules and floral organ abscission zones, and whose expression is positively regulated by wounding and negatively regulated by SA and *Pseudomonas syringae* (Tarutani et al, 2004a). RLK902 and its closest homologue RKL1, interacts with the uncharacterized proteins At5g05190, At3g27210 and At3g17950 in yeast two-hybrid assays (Tarutani et al, 2004b).

2.3.1 RLK902 accumulates in plasma membranes of *N. benthamiana* leaves

The full length RLK902 gene was cloned into a binary vector behind a 35S promoter for overexpression in *N. benthamiana* leaves by agroinfiltration (Fig. 27A). The RLK902 was tagged with a 4x Myc tag at the N-terminus and a YFP tag at the C-terminus to allow us to monitor both N-terminal and C-terminal processing products. The Myc tag used contains an intron to prevent expression of RLK902 in Agrobacterium (Fig. 27A). Transient expression of RLK902 in *N. benthamiana* by agroinfiltration was analyzed by fluorescence microscopy and western blot. At four

days post infiltration (dpi), YFP fluorescence was observed in the plasma membrane of leaf epidermal cells (Fig. 27B). This expression pattern is distinct from the cytoplasmic-YFP used as a control, where cytoplasmic strands were observed through the vacuole (Fig. 27B), typical to cytoplasmic protein expression. Co-localization studies using a membrane marker protein remorin (Bozkurt et al, 2014) fused to RFP further confirmed the localization of RLK902-YFP in the plasma membrane (Fig. 27B).

To further assess the localization of RLK902, we isolated membranes from leaf extracts by ultracentrifugation, followed by western blotting against anti-cMyc and anti-GFP antibodies (Fig. 28A). The full length RLK902 protein was detected in the membrane fraction (pellet fraction 1) and not in the soluble fraction using both anti-cMyc and anti-GFP antibodies (Fig. 28B, black arrows). This confirms the above fluorescence microscopy findings. Addition of 1% NP40 to the membrane fraction, coupled with vigorous vortexing failed to solubilize the RLK902 signal from the membrane (Fig. 28B, pellet fraction 2). However, adding 1% SDS to the NP40-resistant pellet fraction 2 solubilized RLK902 (Fig. 28B, soluble fraction 3). We therefore conclude that RLK902 is in an NP40-resistant membrane fraction.

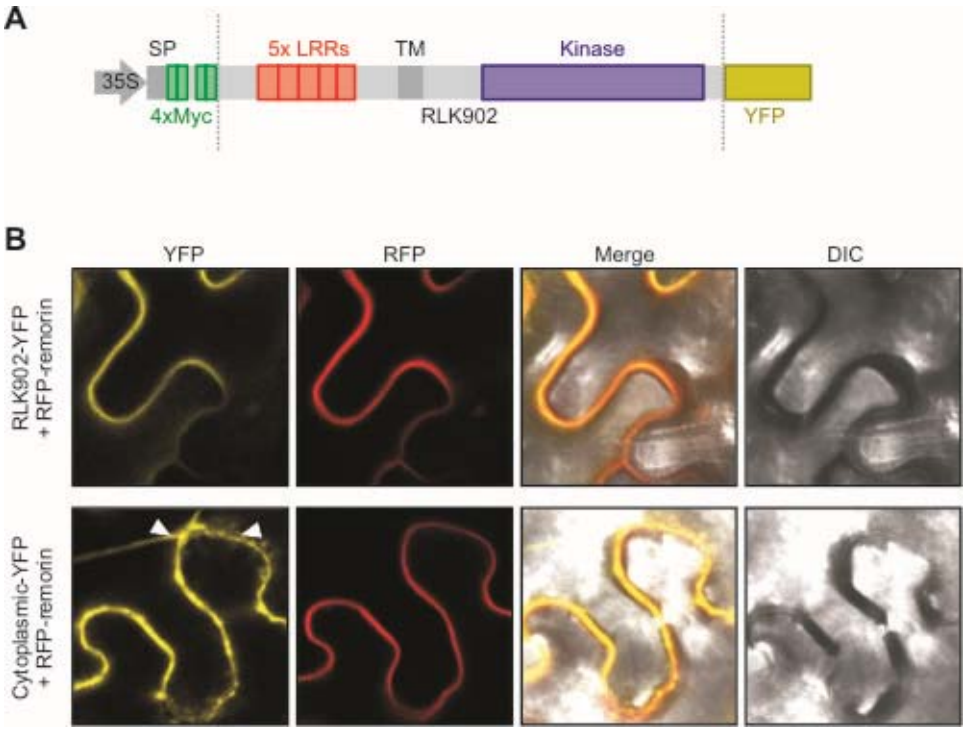


Figure 27. Cloning and expression of RLK902 in N. benthamiana leaves by agroinfiltration.

(A) RLK902 was cloned in a binary expression vector behind the 35S promoter, fused to a Myc tag at the N-terminal and YFP at the C-terminal of the mature protein. The signal peptide (SP) is from the tobacco PR1a gene. Agrobacterium carrying the construct, together with Agrobacterium carrying silencing construct P19, was agroinfiltrated into *N. benthamiana* leaves using a needleless syringe. **(B)** At 3-6 dpi, leaves were subjected to fluorescence microscopy to assess the subcellular localizations of RLK902. Cytoplasmic-YFP and the membrane marker protein remorin fused to RFP were used as controls. White arrows indicate cytoplasmic strands.

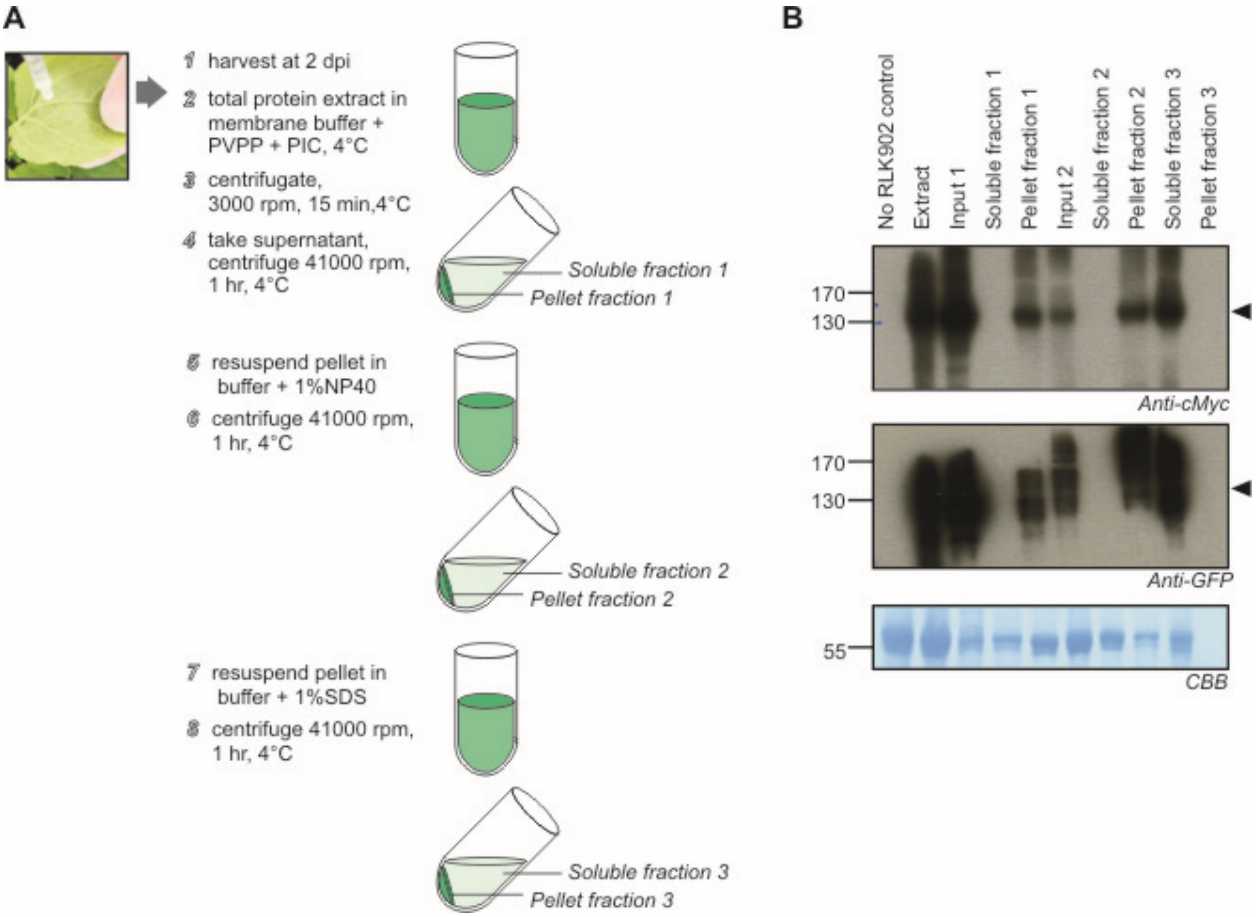


Figure 28. Subcellular fractionation reveals that RLK902 is in an NP40 resistant membrane fraction.

(A) Steps in membrane fractionation experiment. Membrane fraction (pellet fraction 1) was isolated by ultracentrifugation of an extract from agroinfiltrated leaves at 2 dpi. The pellet was treated with NP40 and SDS to dissolve the membrane. Soluble and pellet (membrane) fractions 1, 2 and 3 used for western blot analysis in (B) are also indicated in the different steps. The membrane was enriched particularly in the pellet fraction 1. **(B)** Western blot analysis of protein samples taken from each step during membrane fractionation, using anti-cMyc and anti-GFP antibodies. Input 1 is soluble fraction after step 3 in (A). Input 2 is pellet fraction 1 after adding 1% NP40. Arrow heads indicate the full length RLK902 protein. CBB shows coomasie stained gel.

2.3.2 RLK902 is processed *in vivo*

We next followed the expression of RLK902 in *N. benthamiana* leaves at later time points by western blot and found that RLK902 is processed *in vivo* (Fig. 29A). To circumvent possible *ex vivo* processing or degradation of the full length protein, extraction of protein from fresh leaf samples expressing RLK902 was done directly into denaturing gel loading buffer. In the time course experiment, there is no detectable RLK902 at one dpi, and a full length 150 kDa protein detected at two dpi. Processing of full length RLK902 into cleavage products was observed beginning at three dpi and it is even more pronounced at four and five dpi (Fig. 29A). Cleavage products detected with either anti-cMyc or anti-GFP antibodies were mapped into the full length protein according to their MWs (Fig. 29C). Signal A represents the 150 kDa full length protein, without the signal peptide (SP), containing both Myc and YFP tags, detected in both anti-cMyc and anti-GFP blots (Fig. 29A-B). Signal B is a 130 kDa without Myc tag and must therefore have resulted from N-terminal processing. Signals C, D, and E at 90, 70, and 50 kDa, respectively, are terminally truncated in the LRR domain, but all carry the intact kinase domain and the C-terminal YFP tag (Fig. 29A-B). Signal F is 50 kDa and is detected by anti-cMyc antibody indicating that this protein carries only the N-terminal LRR domain plus the Myc tag (Fig. 29C).

2.3.3 Deglycosylation confirms molecular weight of RLK902 fragments

Signals A, B, C and F have higher apparent MW than predicted from their calculated MW. This could be due to glycosylation at the five putative N-glycosylation sites (NxS/T) in the ectodomain of RLK902 (Fig. 25). To confirm MW of the non-glycosylated proteins, samples were treated with Peptide-*N*-Glycosidase F (PNGaseF), which remove glycans from Asp residues (Maley et al, 1989). Western blot analysis shows a 20-30 kDa shift in the migration of signals A, B, C and F which carries the extracellular domain or a part of it, and no shift was observed for signals D and E, which do not carry N-glycosylation sites (Fig. 29B-C).

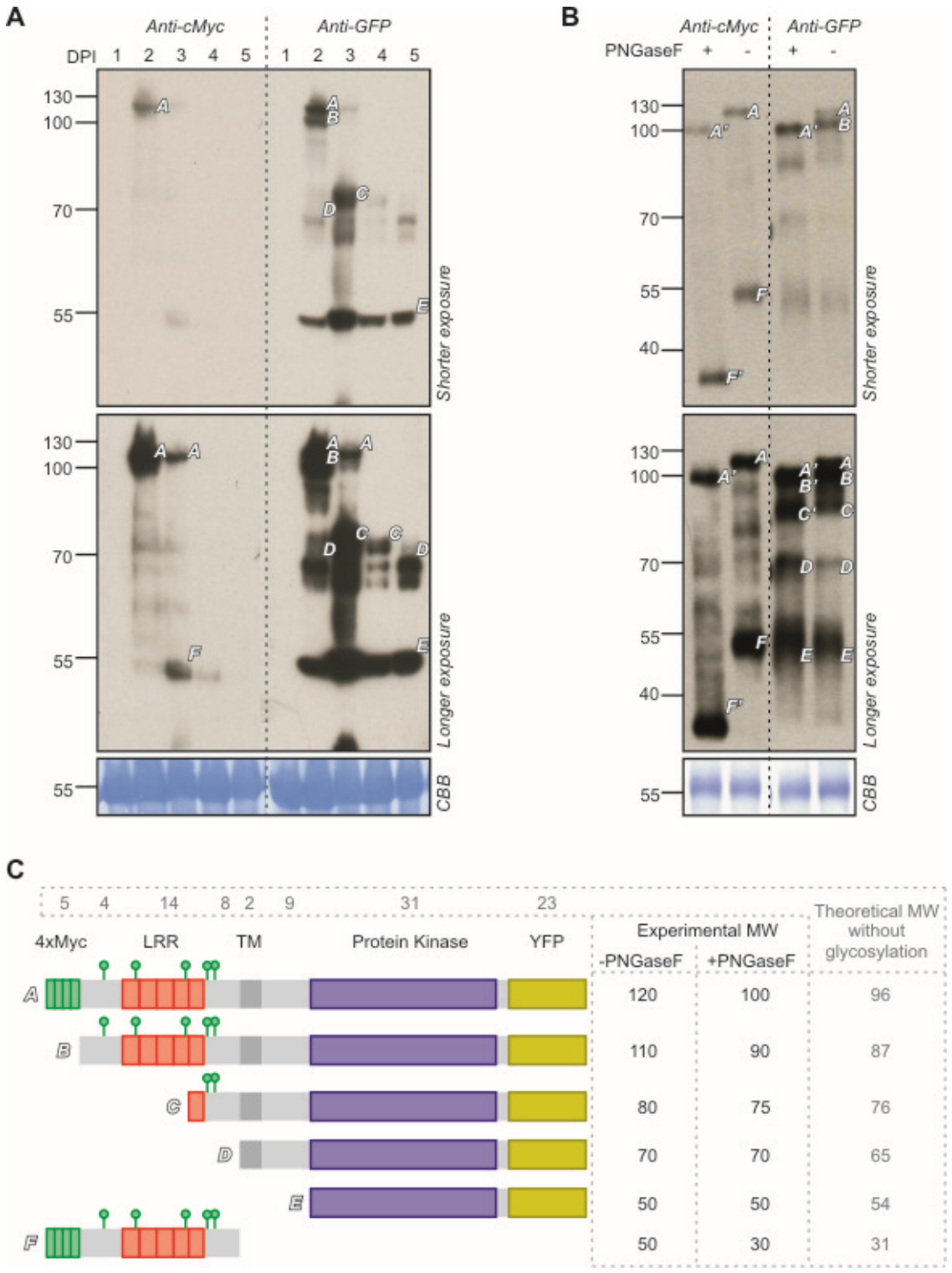


Figure 29. *In vivo* processing of RLK902.

(A) RLK902 were transiently expressed in *N. benthamiana* leaves by agroinfiltration and proteome from agroinfiltrated leaf disks at one to five dpi were extracted in gel loading buffer and separated on acrylamide gels for western blot analysis using anti-cMyc and anti-GFP antibodies. The different cleavage products observed over time are marked with letters A-F corresponding to the drawings of presumed RLK902 fragments in (C). **(B)** Leaf extracts at two dpi from RLK902-agroinfiltrated leaves, treated with or without PNGaseF, were analyzed by western blot using anti-cMyc and anti-GFP antibodies. RLK902 degradation products are marked as in (A). Apostrophes after the letter mark a

shift in the MW of the protein band upon PNGaseF treatment. **(C)** Schematic diagram showing RLK902 cleavage products and their corresponding experimental MWs with or without PNGaseF treatment and theoretical MWs in kDa.

2.3.4 RLK902 *in vivo* processing is inhibited by E64d and TLCK

We next tested several protease inhibitors for the inhibition of RLK902 processing. Agroinfiltrated leaves were injected at four dpi with cell permeable protease inhibitors and samples were collected after six hours. We hypothesized that at four dpi, inhibition of processing should cause accumulation of the full length RLK902 protein. Interestingly, E64d treatment caused a strong signal of the full length protein and also signals B, C and D but signal E did not accumulate (Fig. 30). In addition, MG132, a potent, reversible, cell permeable proteasome inhibitor, and the protease inhibitor cocktail caused weak signals of the full length protein (Fig. 30). In contrast, no full length RLK902 protein was detected upon pre-treatment with soybean kunitz, epoxomicin, phenylmethylsulfonyl fluoride (PMSF), marimastat and phenantroline (Fig. 30). For most of these, signals C, D, and E were detected except for phenantroline, where only signal E is detected.

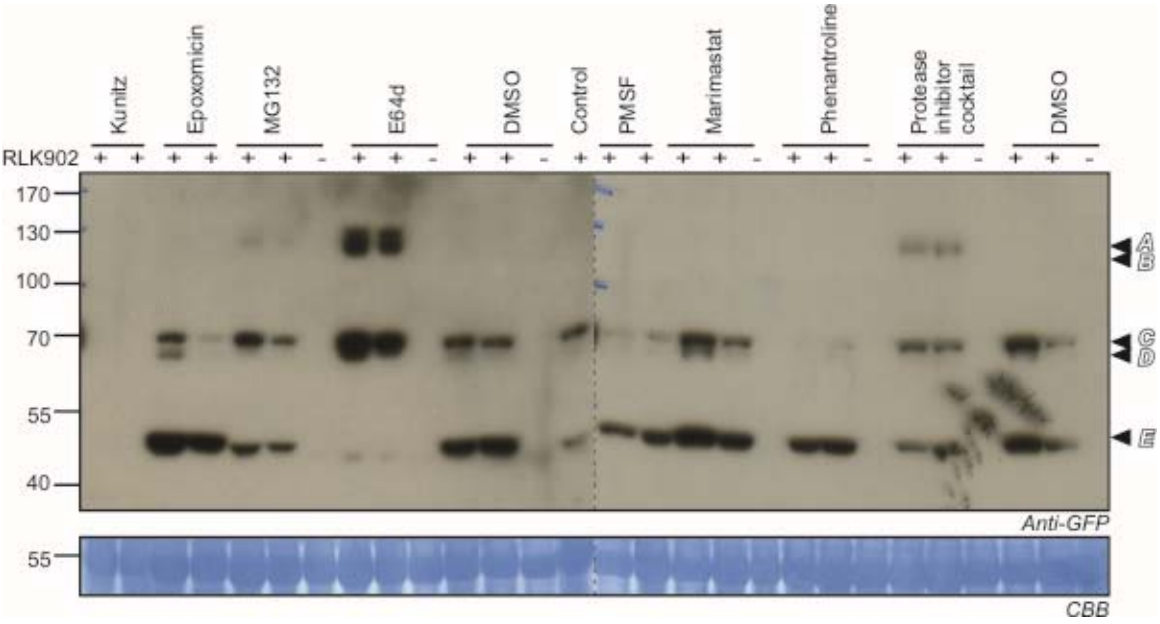


Figure 30. *In vivo* processing of RLK902 is inhibited by E64d.

RLK902 was transiently expressed in *N. benthamiana* by agroinfiltration. At four dpi when no full length protein is detected, protease inhibitors (100 µM) were infiltrated in duplicate. Six hours later, total proteomes were extracted and analyzed by western blotting. Letter annotations on the right are the same as in Fig. 29. DMSO is used as a negative control for inhibitors, and control means no infiltration done.

These data suggests that E64d may inhibit the protease that cleaves RLK902. We next tested known protease inhibitors Tosyllysine chloromethyl ketone hydrochloride (TLCK) and Odanacatib in our assays. TLCK is a serine protease inhibitor and blocks degradation of endocytosed proteins like epidermal growth factor (EGF) (Schultz et al., 1996). While, odanacatib, a cathepsin K inhibitor, blocks bone turnover and transcytosis, and is used as drug to prevent osteoporosis (Leung et al, 2011). Results show that TLCK and E64d inhibit processing of RLK902 at six hours after infiltration (Fig. 31A). However, unlike E64d, TLCK did not block the accumulation of product E (Fig. 31A). In contrast, odanacatib did not cause accumulation of full length RLK902, even at earlier time points (1, 2 and 3 hrs) after infiltration (data not shown). In addition, fluorescence microscopy studies show accumulation of endosomal-like structures in both E64d and TLCK infiltrated samples, as compared to the DMSO control (Fig. 31B). These results suggest that E64d and TLCK inhibit processing of RLK902 by blocking its vacuolar degradation pathway.

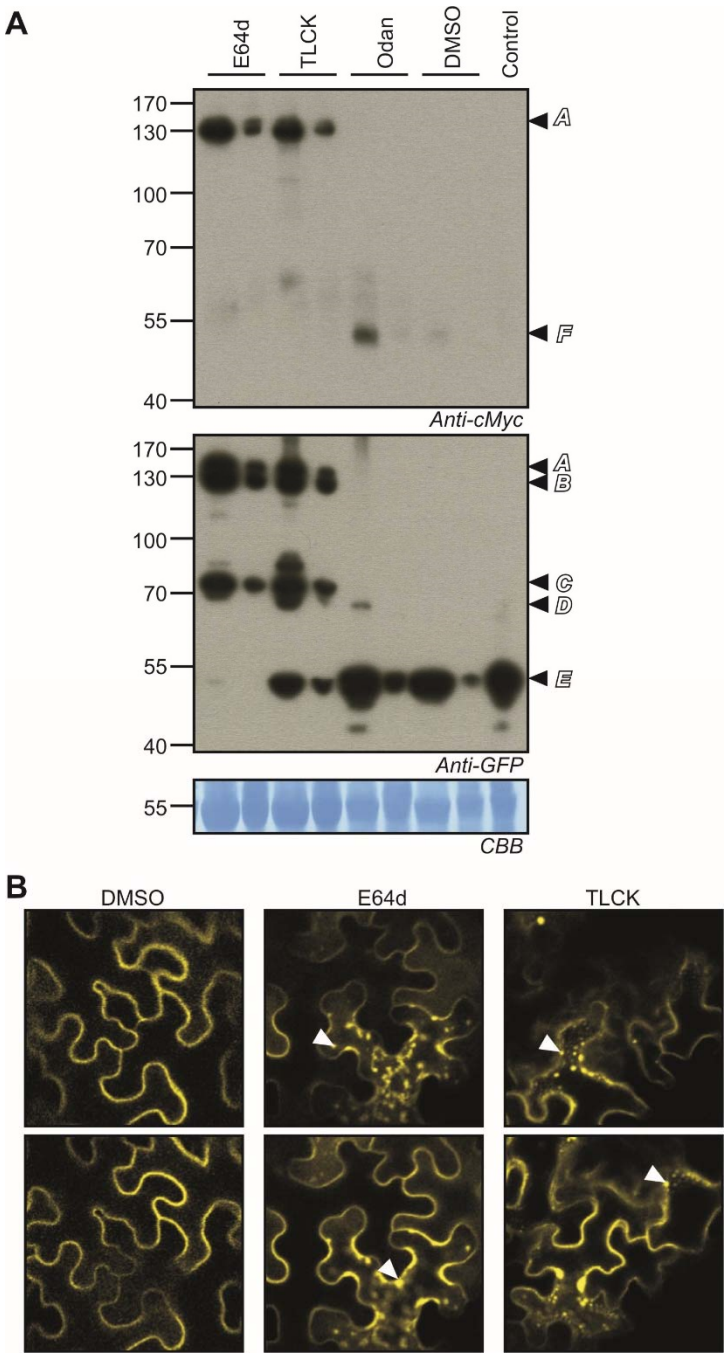


Figure 31. E64d and TLCK prevent degradation of RLK902 possibly by inhibiting the vacuolar degradation pathway.

(A) Testing of protease inhibitors TLCK and Odanacatib in causing accumulation of RLK902 in *N. benthamiana* leaves. After four days of transiently expressing RLK902 into *N. benthamiana* leaves by agroinfiltration, the protease inhibitors were syringe-infiltrated into leaf samples, collected after six hours, and analyzed by western blot using anti-cMyc and anti-GFP antibodies. Letter annotations are the same in Fig. 29. **(B)** Fluorescence microscopy of leaf samples infiltrated with E64d and TLCK. White arrows indicate endosomal bodies.

Plasma membrane receptors are involved in various environmental and developmental responses in plants. These protein receptors pass through subcellular trafficking pathways from their synthesis in endoplasmic reticulum, to their processing and maturation in golgi complexes, and finally to their recruitment and employment in plasma membranes. From the plasma membranes, receptors like FLS2, undergo endocytosis for either recycling back to plasma membrane or targeting to late endosomes and subsequent degradation in vacuoles (Robatzek, 2007).

Interestingly, E64d can prevent vacuolar degradation of proteins in the vacuole by blocking trafficking of endocytic vesicles (Moriyasu & Ohsumi, 1996; Yano et al, 2004). We therefore speculated that inhibition of processing of RLK902 may be due to inhibition of endocytotic internalization. Hence, we tested several inhibitors of vacuolar targeting pathways such as Brefeldin A (BFA), Wortmannin, Tyrphostin A23, Concanamycin A, Latrunculin B, and Cantharidin (Fig. 32). BFA inhibits protein transport from the ER to golgi by releasing cytosolic coat proteins β-COP (Klausner et al, 1992). Wortmannin is an inhibitor of phosphatidyl-inositol 3-kinase and disrupts presynaptic actin causing inhibition of vesicle recycling (Richards et al, 2004). In addition, Tyrphostin A23 is a protein tyrosine kinase inhibitor and blocks the interaction of the YXXΦ internalization motif in transmembrane proteins and the μ2 subunit of AP-2 adaptor complex involved in clathrin-mediated endocytosis (Banbury et al, 2003). Concanamycin A, also called Folimycin, inhibits acidification of organelles involved intracellular trafficking like lysosomes and golgi complex by inhibiting vacuolar type H^+-ATPases (Muroi et al, 1993; Palokangas et al, 1994). Latrunculin B prevents polymerization of actin required for receptor-mediated endocytosis (Lamaze et al, 1997; Moscatelli et al, 2012). Furthermore, cantharidin is a protein phosphatase 2A inhibitor and blocks internalization of the plasma membrane receptor kinase FLS2 (Serrano et al, 2007). However, results show that none of targeting inhibitors affected RLK902 processing (Fig. 32). In addition, we tested targeting inhibitors at earlier time points (1, 2, and 3 hrs) but the same results as Fig. 32 were observed (data not shown).

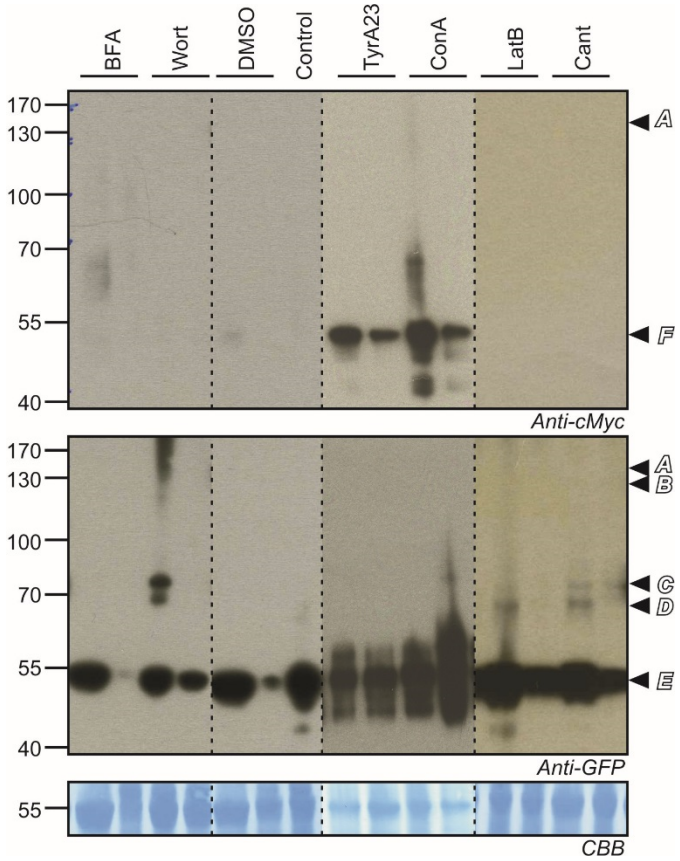


Figure 32. Trafficking inhibitors do not prevent degradation of RLK902 in N. benthamiana leaves.

Screening of trafficking inhibitors implicated in vacuolar degradation pathway of membrane proteins. Among screened were Brefeldin A (BFA), Wortmannin (Wort), Tyrphostin A23 (TyrA23), Concanamycin A (ConA), Latrunculin B (LatB) and Cantharidin (Cant). After days of transiently expressing RLK902 into *N. benthamiana* leaves by agroinfiltration, vacuolar degradation inhibitors were syringe-infiltrated into leaf samples, collected after six hours, and analyzed by western blot using anti-cMyc and anti-GFP antibodies. Letter annotations are the same in Fig. 29.

2.3.5 Mutations in the kinase domain do not affect processing of RLK902

To test if mutations on the conserved residues of the RLK902 kinase domain will stabilize RLK902 and prevent its processing, mutations of the conserved lysines (K393 and K492) and the conserved aspartate (D508) were mutated into alanine (A; Fig. 33A). K393 and K492 in the catalytic domain of the kinase are in close proximity to a bound ATP and is involved in the phosphotransfer reaction important for phosphorylation events (Carrera et al, 1993), while D508 is in the activating DFG (Asp-Phe-Gly) motif and is involved in the chelating and positioning the Mg ion in the active site of the kinase (Zheng et al, 1993). Mutating these conserved residues renders the kinase inactive (Carrera et al, 1993; Kornev et al, 2006). The RLK902

mutants generated were expressed by agroinfiltration in *N. benthamiana* leaves and samples were collected at two, three, four and five dpi. Western blot analysis shows RLK902 processing for all mutants, comparable to wildtype (WT) (Fig. 33B). Moreover, only the K393A mutant was detected as full length protein at five dpi. Intermediate products C, D, E and F were also detected for the K393A mutant. No full length protein was detected from the K492A and D508A mutants, but intermediates E were detected in K492A, and intermediates C, D, and E were detected in D508A (Fig 33B). Interestingly, the full length proteins, as well as intermediates B, C, D, and E, from all three RLK902 mutants accumulated upon E64d inhibition at four dpi. These data suggests that the conserved lysine and aspartate residues do not affect processing of RLK902, but a protease that is sensitive to E64d inhibition.

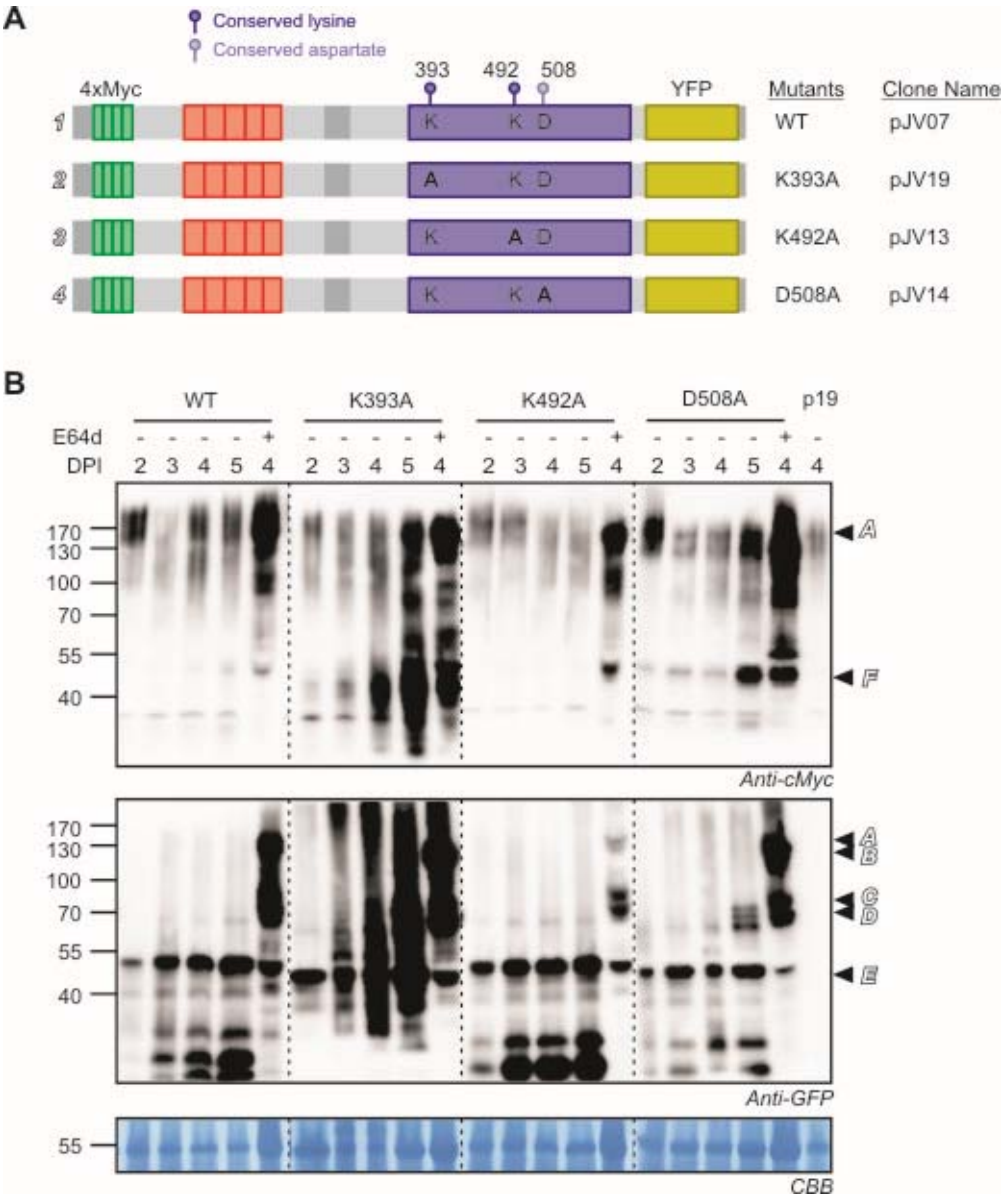


Figure 33. Cloning and expression of RLK902 kinase domain mutants.

(A) Site-directed mutagenesis of the two conserved lysines (K393 and K492) and the conserved aspartate (D508) generated three different mutant versions of RLK902. These residues are exchanged to alanine (A), as indicated in bold letters. **(B)** Expression and western blot analysis of the RLK902 mutants expressed in *N. benthamiana* upon agroinfiltration. Samples were taken at two to five dpi. In addition, E64d was also infiltrated at four dpi and leaf samples were collected after six hours of infiltration, for western blot experiments. Letter annotations are the same in Fig. 29.

CHAPTER 3: DISCUSSION

3.1 Chemoproteomics with AcATP

3.1.1 Characterization of AcATP labelling

Four different AcATP probes that differ in their reporter tags were used in labelling Arabidopsis leaf or cell culture extracts. The different reporter tags require different detection methodologies, e.g. protein blotting using streptavidin-HRP for BAcATP and DTBAcATP, and fluorescence scanning for bodipy-AcATP, while the ≡AcATP needs to be conjugated with biotin or rhodamine, or both, using click chemistry. The different AcATP probes gave similar labelling profiles, although labelling intensities vary, which may be due to the sensitivity of detection methods and/or to the labelling capability of each probe type. The four probes not only differ in their reporter tags but also in the lengths of their linkers. The bodipy-AcATP has the shortest linker containing two carbons, followed by DTBAcATP, alkyne-AcATP and then BAcATP with four, eight and ten carbons in the linker, respectively. The differences in their linker length were reflected in the labelling capacity of the probes where the shorter linker gave less labelling when compared to those with longer linkers. This is especially remarkable in the stronger labelling of Arabidopsis cell culture extracts using BAcATP when compared to DTBAcATP, particularly the high MW proteins (above 55 kDa; Fig. 1C, right panel). Noteworthy, the resulting labelling profile using bodipy-AcATP is the weakest among the four probes, while BAcATP and alkyne-AcATP having the longest linkers gave the strongest labelling intensities (Fig. 1C, left panel). These data suggest that longer linkers provides more access of probes to allosteric ATP binding pockets of ATP-binding proteins.

We have optimized AcATP labelling of Arabidopsis leaf extracts by testing different factors that affect labelling. We found that labelling greatly depends on pH. This is because pH strongly affects conformation of proteins, by protonation (or deprotonation) of important residues, affecting their stability and (enzymatic) activity (Di Russo et al, 2012). In fact, many proteins can act as intracellular pH sensors, with activities and ligand-binding capabilities changing upon small shift in their environmental pH (Srivastava et al, 2007). Moreover, strong acidic or basic pH

denatures proteins and many proteins have a neutral physiological pH (pH 7-7.5), where they are mostly active. We observed that no labelling was detected at acidic pH and strong labelling at neutral to basic pH. All subsequent labelling experiments were done at pH 7.5. In addition, divalent metal cations Mg^{2+}, Ca^{2+} and Mn^{2+} increased labelling, while chelating agents EDTA and EGTA decreased labelling. This is because divalent metal cations stabilizes the binding of ATP in the ATP binding loop of protein kinases and ATP-binding proteins which allows phosphorylation to occur (Adams & Taylor, 1993; Gonzalez et al, 2006; Saylor et al, 1998). Chelating agents, on the other hand, sequesters divalent metal ions (Aylward & Findlay, 2002) making them unavailable for binding to protein kinases and ATP-binding proteins. We also observed that labelling is suppressed with ATP-like nucleotides such as GTP, CTP, TTP and NADP. These nucleotides share similar structural features to ATP, hence, they compete for binding to the ATP binding loop of ATP-binding proteins. We also discovered that AcATP probes do not label *in vivo*, and can only be used in *in vitro* labelling experiments. This is because ATP has high negative charge which makes it difficult to pass through the cell membrane (Stryer et al, 2007). All these findings are in accordance to what is expected of the AcATP probes.

Searches for differential AcATP labelling in different proteomes seem to be promising, even in low resolution gel-based assays. Differentials were found for H_2O_2 and heat treated Arabidopsis cell cultures, and in bacteria-induced accumulation of ATP-binding proteins in *N. benthamiana* apoplastic fluids (Fig. 9A, B, D). In contrast, treatments with flg22, BTH, and PtoDC3000ΔhQ did not result to differential AcATP labelling in extracts from Arabidopsis cell cultures and leaves (Fig. 9B, C), Purification of AcATP-labelled proteins prior to gel-based detection also did not display differential AcATP labelling of flg22 treated Arabidopsis leaf or cell culture extracts (Fig. 9C, E). This is because virtually all the >1200 protein kinases bind ATP, and labelled proteins can not be effectively resolved in a gel-based assay. Moreover, differentials may not be detected when the labelled kinase(s) of interest have a very low abundance. Patricelli and colleagues (2007) pointed out that only methods of extremely high resolution such as quantitative proteomics by MS are suitable for analysis of native kinases with AcATP probes. One effective method of increasing detection of differential AcATP labelling is coupling the gel-based

resolution of proteins with LC-MS/MS analysis of gel bands cut from the gel (illustrated in Chapter 2.1.4). And even more powerful method is the bead-based enrichment of labelled proteins, followed by Xsite where labelling sites are quantitated by MS (illustrated in Chapter 2.1.5). Comparisons of these methods are discussed below.

3.1.2 Identification of AcATP targets in Arabidopsis

This part of the thesis is published in: Villamor JG, Kaschani F, Colby T, Oeljeklaus J, Zhao D, Kaiser M, Patricelli MP, and van der Hoorn RA (2013) Profiling protein kinases and other ATP-binding proteins in Arabidopsis using acyl-ATP probes. Mol Cell Proteomics. 12. 2481-2496

We report the first in-depth analysis of targets of an AcATP probe in a plant proteome. By comparing a gel-based identification platform of labelled proteins with a gel-free platform for labelled peptides, we demonstrate the advantages and limitations offered by these complementary approaches. The analysis of labelling sites using the protein database shows that AcATP probes target predominantly but not exclusively the ATP binding pockets of a broad range of protein kinases and other unrelated ATP-binding proteins. The labelled proteins include a few well-characterized ATP-binding proteins, and many ATP-binding proteins that have not been studied so far.

3.1.2a Gel-based vs Gel-free approach

The data generated by the gel-based and gel-free platforms are complementary and each platform has its advantages and disadvantages. Although many proteins were detected using both approaches, several were detected by only one approach and not the other, and *vice versa*. HSP60 and HSP70, for example, were only detected using the in-gel approach, whereas CDC48 and MPKs were only detected by the gel-free approach. Gel-based identification resulted in the detection of 112 proteins of which only nine were protein kinases. Although the vast majority of the detected proteins were absent in the no-probe control, the labelling site could only be determined for nine proteins, based on the detection of labelled peptides. This number would even be lower in case we did not realize that the biotin is somehow oxidized during this procedure. By contrast, sequencing the labelled peptides using Xsite/KiNativ by the gel-free approach identified more protein kinases and also

determines labelling sites in each of the labelled proteins. Although this gel-free approach seems more powerful, it can not discriminate between proteins if the labelled peptide is identical. For example, the labelled peptides do not discriminate between CPK9 or CPK11 because it is identical, whereas both proteins were distinguished using the gel-based approach. The gel-free approach also does not provide the MW of the labelled protein, and this can be very important information. For example, we noticed that two labelled LRR-RLKs do not migrate at the expected 67-68 kDa, but at 40 kDa, indicating that they have lost the extracellular LRR domain. At this stage we can not exclude that this processing occurred upon protein extraction. It is interesting to point out that XA21, an LRR-RLK from rice, is proteolytically cleaved during signalling to release a cytosolic protein kinase domain that migrates to the nucleus to phosphorylate transcription factors (Park & Ronald, 2012). A similar receptor cleavage mechanism might be at work for the two LRR-RLKs detected on our study.

3.1.2b Detection of an oxidized biotin moiety

Our analysis of biotinylated proteins by in-gel digest and LC-MS/MS revealed no peptides that were labelled with biotin-acyl (BAc), but several peptides that were labelled with OBAc, the oxidized version of BAc. To our knowledge, this type of modified biotin has not been reported before during proteomics and this discovery may have important implications for the detection of other modifications containing biotin. We speculate that the oxygen is located on the sulphur of biotin, resulting in a biotin sulfoxide (Melville, 1954). This modification was also observed e.g. during the synthesis of oligonucleotides (Upadhya et al, 2005), and would not occur on desthiobiotin. At this stage we do not know when during the in-gel procedure the oxidation occurs. Searches on the occurrence of oxidized biotin on proteins labelled with other probes will tell how common this modification is in the future.

3.1.2c How to detect more protein kinases?

We detected 24 labelled peptides from protein kinases, but there are 1099 protein kinases encoded by the Arabidopsis genome (The Arabidopsis Genome Initiative, 2000). That we do not detect more protein kinases has several reasons. First, most of the protein kinases are not expressed in tissues that we used for the labelling

experiments. Second, peptides from some protein kinases are too small or too large to be detected with the chosen MS settings. However, our *in silico* analysis indicates that the majority (95%) of the theoretically labelled peptides of Arabidopsis protein kinases are in the range of 5-30 amino acids. Third, some protein kinases can not be discriminated based on the sequence of the labelled peptides. Nonetheless, our *in silico* analysis showed that 64% of the Arabidopsis protein kinases contain a putative labelled lysine in a peptide sequence that is unique in the Arabidopsis proteome (Fig. 34). Fourth, mass spectrometry is not sensitive enough to detect labelled peptides from protein kinases in an unbiased mode. Indeed, inclusion lists coupled to retention times, and searches for marker ions in the MS2 mode can tremendously increase the number of identified protein kinases. The added search modes have increased the number of robustly detected protein kinases to 160 protein kinases per cell line (Patricelli et al, 2011). Thus, we expect that we will be able to detect labelling of more protein kinases from a single proteome with the implementation of empirical searches (shown in Chapter 2.2 of this thesis). The wider range of detection will increase the power of protein kinase profiling in plants tremendously.

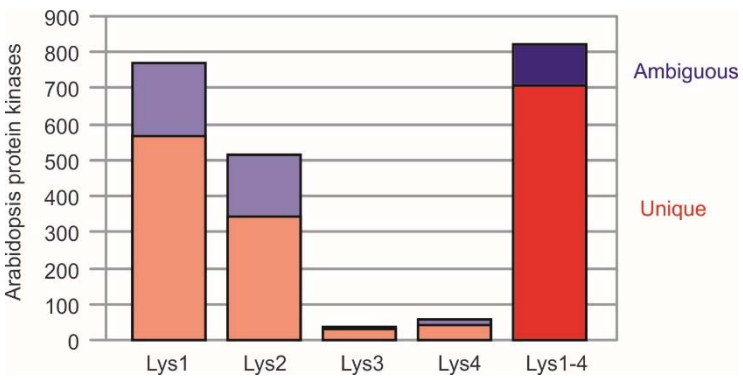


Figure 34. In silico analysis of resolving potential of labelled peptides.

Sequences for each subfamily in the protein kinase database were retrieved and aligned using MultiAlign. The labelled peptide that was experimentally identified was highlighted in the alignment and the orthologus Lys residues were indicated. *In silico* tryptic digest (no missed cleavages except the orthologous lysine and if the Lys/Arg is followed by Pro), resulted in theoretically labelled peptides. The selected peptides were ranked on sequence and redundancy for each peptide sequence was counted (n proteins encode the same peptide). Ile/Leu residues were not discriminated in this analysis. Protein kinases with a unique labelled lysine (n=1) were counted and plotted (red). The other protein kinases contain ambiguous peptides (n>1). The same analysis was done for Lys2, Lys3 and Lys4. 820 protein kinases have at least one of the four lysine residues identified in the alignment, and 705 of these proteases have a unique peptide sequence containing either Lys1, Lys2, Lys3 or Lys4.

3.1.2d Profiling Arabidopsis protein kinases

We have detected 24 labelled peptides that originate from protein kinases. Analysis of the labelling sites in these protein kinases showed that AcATP probes react with lysines that reside at four distinct positions. Lys1 and Lys2 reside in the ATP binding pocket in motifs II and VIb, respectively, and were previously described to be labelled by AcATP probes (Patricelli et al, 2007). Labelling at Lys3 and Lys4 was detected in PTI1-like kinases and MAP3K VIK, respectively, and was not characterized before. Although crystal structures are not available for representatives of these protein kinase subfamilies, modelling indicates that regions carrying Lys3 and Lys4 might indeed be able to fold back on the ATP binding pocket. The reactivity of Lys3 and Lys4 indicates that, even though these lysines are poorly conserved, they might be in close proximity to ATP and might play important roles in PTI1- and VIK-like protein kinases. The absence of any other labelled peptide from the detected protein kinases illustrates the remarkable selectivity of the AcATP probe to target the ATP-binding pocket of these proteins.

The labelled protein kinases are remarkably diverse and contain representatives of all major classes of Arabidopsis protein kinases. The list includes RLKs, Pelle-like kinases, MAP kinases, MAP3K kinases, calcium-dependent kinases, PTI-like kinases, and a proline-rich, extension-like receptor kinase PERK1. The diversity of the protein kinases illustrates the wide range of protein kinases that can be labelled with AcATP, and is consistent with the fact that all protein kinases (by definition) bind ATP and most of them carry lysine residues to stabilize the phosphate. The data are also consistent with studies of AcATP probes on animal proteomes, where >75% of the human protein kinases have been detected (Patricelli et al, 2007). Thus, AcATP probes have the remarkable potential to label nearly all protein kinases not only of mammals but also other organisms. The detection of several RLKs with relatively high spectral counts is noteworthy because these proteins are thought to be low in abundance and we did not enrich for membrane proteins.

Functions have been described for several of the detected protein kinases. MPK3, 4, 6 and 11 are involved in immune signalling, and control gene expression through WRKY transcription factors (Bethke et al, 2012; Li et al, 2012; Rasmussen et al,

2012). PBS1 is involved in perception of bacterial pathogens secreting the AvrPphB effector (DeYoung et al, 2012; Shao et al, 2003). AvrPphB cleaves PBS1 and this activates resistance protein RPS5. Calcium-dependent kinase CPK3 is involved in stomatal closure, herbivore defence, salt stress acclimation and is a positive regulator of sphingolipid-induced cell death (Lachaud et al, 2013; Mehlmer et al, 2010; Mori et al, 2006). PTI1-2 interacts with oxidative stress-response protein kinase OXI1 and integrates phosphatidic acid signalling with reactive oxygen signalling (Anthony et al, 2006). VIK is a MAP3K involved in the uptake of glucose into the vacuole (Wingenter et al, 2011), whereas AT6 is a MAP3K that acts as a negative regulator of salt tolerance (Gao & Xiang, 2008). PERK1 is thought to be involved in sensing cell wall stress during wounding and pathogen infection (Haffani et al, 2006; Silva & Goring, 2002). The detection of these characterized protein kinases by AcATP labelling is consistent with their presence and activities in the Arabidopsis rosette used for our studies. The other two-third of the detected protein kinases have only been described by transcript levels and phylogeny. Our data demonstrate that these uncharacterized protein kinases are present and able to bind ATP.

3.1.2e Labelling of other ATP-binding proteins

Previous studies with AcATP probes have been mostly focused on the labelling of protein kinases. In this study we also examined the labelling of proteins other than protein kinases. An in-depth analysis of the labelled peptides revealed that AcATP probes label an astonishing diversity of protein families. Most of the labelled proteins are able to bind nucleotides, and analysis of structural data indicates that AcATP probes label these nucleotide-binding proteins predominantly inside ATP binding pockets. The structural diversity of the nucleotide binding proteins is remarkably diverse, illustrated the fact that they belong to 13 different protein clans in the PFAM database (Finn et al, 2010).

Several different ATP-binding proteins have been identified with high spectral count frequencies. Phosphoglycerate kinase (PGK1) is the top hit among the labelled proteins. PGK degrades ATP during glycolysis (Plaxton, 1996). GATB is an Asn/Gln-tRNA amidotransferase involved in the transamidation of misacylated Asp-tRNA or

Glu-tRNA forming a correctly charged Asn-tRNA or Gln-tRNA, a reaction that requires ATP (Ibba et al, 2000; Pujol et al, 2008). CDC48 encodes for cell division cycle protein-48 and is important in cell division, expansion and differentiation. Arabidopsis plants carrying mutations in CDC48-encoding genes show impaired seedling development and defective pollen and embryo development (Park et al, 2008). CDC48 is an AAA-type ATPase, and belongs to a large superfamily containing a highly conserved triple-A domain which binds and hydrolyses ATP (Ogura & Wilkinson, 2001). CAC2 is the biotin carboxylase subunit of the heterodimeric, biotin-containing enzyme ACCase (chloroplastic acetyl-coenzyme A carboxylase). ACCase is involved in the synthesis of fatty acids in Arabidopsis (Sun et al, 1997) and uses ATP for the conjugation of a carboxylate containing molecule to an amino or thiol group-containing molecule (Galperin & Koonin, 1997). CAC2 contains a phosphate-binding loop and a Mg^{2+}-binding site and has two alpha-beta subdomains that grasp the ATP molecule. This 'ATP-grasp' domain is common to a superfamily that also contains D-alanine:D-alanine ligase (DD-ligase), glutathione synthetase (GSHase) and carbamoyl phosphate synthase. PYR6 is a uridine 5´-monophosphate (UMP)/cytidine 5´-monophosphate (CMP) kinase, which converts UMP/CMPs to uridine and cytidine diphosphates, respectively. PYR6 orthologs in bacteria and yeast are required for cellular proliferation and RNA and protein synthesis (Liljelund & Lacroute, 1986; Yamanaka et al, 1992). PYR6 has also been detected in the mature pollen of Arabidopsis, consistent to its role in cell proliferation and division (Holmes-Davis et al, 2005). Taken together, these five structurally and functionally diverse examples illustrate the broad range and relevance of ATP-binding enzymes that we can monitor using AcATP probes.

3.1.2f Labelling outside known ATP binding pockets

Our analysis of labelling sites also showed that a large number of labelling sites are not located in known nucleotide binding pockets. It is possible that several of these labelling sites are nucleotide binding sites that have not yet been described. However, the correlation between spectral counts with transcript levels indicates that labelling outside known nucleotide binding pockets is typically detected for highly abundant proteins, such as Rubisco. An obvious explanation for the labelling of abundant proteins outside nucleotide binding sites would be that solvent-exposed

lysine residues have a low basal reactivity towards AcATP and that this causes aspecific labelling of highly abundant proteins. Strangely enough, however, AcATP labelling of abundant proteins nevertheless requires divalent ions and can be competed with nucleotides that do not carry an acyl reactive group. These data support the exciting idea that ATP-competable AcATP labelling sites might identify novel nucleotide binding sites in proteins. Interestingly, a recent crystal structure of rice Rubisco with NADPH revealed that two of the labelled lysines are in close proximity to the phosphate in NADPH (supplemental Fig. S4D) (Matsumura et al, 2012). Thus, labelling outside known nucleotide binding sites is an interesting topic for further studies.

Our above study shows that AcATP labelling is a powerful way to monitor ATP binding activities in native proteomes. Besides the majority of the 1099 protein kinases of Arabidopsis, AcATP can potentially label another 482 ATP-binding proteins, representing another eleven different protein families. Although AcATP probes also label outside known nucleotide binding sites, the preferential labelling of nucleotide-binding proteins in ATP binding pockets demonstrates the high selectivity of the AcATP probe. With these findings, we find it interesting to study the changes of ATP binding activities of proteins upon an external stimulus. Our data suggest that labelling does not only reflect the abundance of the ATP-binding proteins, but is also affected by the ATP binding activity of the protein. This theory is further investigated in the next chapter, where we applied targeted MS/MS analysis to monitor changes of ATP binding activities upon a stimulus.

3.2 Targeted MS/MS analysis of DTBAcATP-labelled peptides

Our targeted MS/MS analysis had successfully increased the detection of protein kinases when compared to the data-dependent approach we used in the previous chapter, from 24 protein kinases to 46. The first detected 24 protein kinases were also detected in the targeted approach demonstrating the reproducibility of our methods. The tremendous increase in our detection method is due to increased sensitivity of the MS by incorporating an inclusion list. In data-dependent MS analysis, the selection of a particular precursor ion is dependent on its signal intensity in the MS scan, generating a bias against low abundant protein because of oversampling of the high intensity precursor ions (Picotti et al, 2007; Schmidt et al, 2009). This is particularly limiting in high complexity protein samples that we use in our studies. However, in targeted approach, also called directed MS analysis, preselecting ions for MS analysis not only increases sensitivity of MS in detecting these ions, but also prevents redundancy and oversampling of high intensity ions, and increases reproducibility of quantitation of data sets in two or more sample runs (Schiess et al, 2009; Schmidt et al, 2009). Nonetheless, the success of proteomic profiling by targeted MS/MS depends on the accuracy and reproducibility of detecting precursor ion retention times and mass-to-charge (m/z). This requires several number of runs of the same and/or different sample, collecting as much features as possible until no new target ion masses are identified. In the same manner, one can use an exclusion list, in lieu, or in addition to inclusion list, for added power in MS detection (Wang & Li, 2008).

We applied targeted MS/MS to samples where activation of proteins was stimulated by flg22 elicitation resulting in the activation of a kinase signalling pathway. Curiously, our data suggests that DTBAcATP labelling is not affected by the phosphorylation state of proteins. There is no differential MPK3/6 labelling between the control and the flg22-treated sample, whereas MPK3/6 is phosphorylated upon flg22 treatment, when compared to the control, where MPK3/6 is not phosphorylated (Fig. 22C). Phosphorylation of MPK3/6 makes the protein kinase active (Asai et al, 2002). Although this result may have reflected the true character of the probe, meaning that protein kinase labeling is not activity-based, this may also be because

we have saturated the sample with AcATP, so that inactive kinases were labelled to the same extent as the active kinases. Note that irreversible probes like AcATP are greatly affected by proteome concentration, probe concentration and length of labelling time during labelling. Here, we labelled 6 mg/ml of Arabidopsis cell culture extract with 5 µM DTBAcATP for 10 minutes. Previous labelling experiments suggested that these conditions could not have saturated the system with labelling. However, this may not be the case when investigating the active state of protein kinases especially when labelling with BAcATP occurs as fast as one minute (data not shown). Further investigation of AcATP labelling on differentially activated MPK3/6 in a time course experiment can be done in the future.

Unexpectedly, MS results also show that the protein kinases FLS2, BAK1, BIK1, MEKK1, and MKK1,2,4 and 5, which are activated during flg22 elicitation, were not detected in our analysis. That they were not detected in our MS scans could be because of a number of reasons. First, the abundance of these proteins may be too low for detection, even though our MS analysis is directed to detecting these low abundant protein kinases. Second, the activation of these protein kinases is very transient, occurring only in seconds up to a few minutes. The flg22-induced signaling cascade occurs almost instantly upon elicitation. FLS2 and BAK1 activation occurs as early as five seconds (Schulze et al, 2010), and MPK3/6 activation is within five to 30 minutes (Bethke et al, 2012; Galletti et al, 2011) after elicitation. In addition, FLS2 and BAK1 receptor kinases are quickly degraded upon activation (Geldner & Robatzek, 2008; Robatzek, 2007; Robatzek et al, 2006). Since we sampled the flg22-treated Arabidopsis cell cultures at 15 minutes after elicitation, we may have missed the narrow time window where FLS2 and BAK1 are active and available for ATP binding. Third, these proteins may have been labelled, but the labelled peptides are too small or too large to be detected in the MS scans. However, *in silico* analysis of labelling sites suggests otherwise. All eight of these kinases carry both Lys1 and Lys2 labelling sites in tryptic peptides that should be easily detected in MS if labelled, except for the Lys2 (5.1 kDa) and Lys1 (4.7 kDa) labelling sites for FLS2 and MEKK1, respectively, which may be too large for efficient MS detection. To detect these protein kinases, subcellular fractionation could have been done to enrich for receptor kinases in membrane fractions.

A more in-depth investigation of differentially activated MPK3/6 was pursued by labelling GST-MPK3/6 fusion proteins produced in bacteria. Our gel-based DTBAcATP labelling experiment shows no significant labelling differences between WT, phosphomimic (MPK3 T196D T198D; MPK6 T221D T223D) and phosphomutant (MPK3 T196A T198A) GST-MPK fusion proteins (Fig. 24C). This is rather intriguing because we expected more labelling in the active phosphomimic MPK3/6 mutant, and less labelling in the inactive MPK3/6 phosphomutant, when compared to the WT MPK3/6. Nonetheless, Lys (K) mutants (MPK3 K166A; MPK6 K191A) show weak or no labelling compared to WT.

The invariant Lys residues in the ATP binding loop of protein kinases have been implicated in anchoring ATP. However, it has been suggested that these highly conserved Lysines are also involved in the phosphotransfer reaction of the γ-phosphate of ATP, rather than in positioning ATP in the ATP binding pocket (Carrera et al, 1993). Site-directed mutagenesis of these Lys residues renders the kinase inactive (Kamps & Sefton, 1986; Knighton et al, 1991). A conserved Asp (D) residue in the Asp-Phe-Gly (DFG) motif is important for the catalytic activity of the kinase and in chelating the MgATP, hence positioning the ATP in its place (Knighton et al, 1991). This indicates that the conserved Lys and Asp are in close proximity to the γ-phosphate of ATP. Phosphorylation of protein kinase A (PKA) rearranges the magnesium-binding loop of the protein, positioning the DFG motif such that Asp can interact with the bound ATP (Beckers et al, 2009; Kornev et al, 2006). This suggests that the conserved Lys, in both active and inactive protein kinases, and the conserved Asp, in active protein kinases, play major roles in securing the ATP. In our assay, the Lys mutant MPKs were weakly labelled with DTBAcATP when compared to WT. This is in accordance to the labelling mechanism of AcATP, that is, the conserved Lysines stabilize the acyl group of AcATP, and are labelled by the probe (Fig. 1). However this does not explain our observation that the phosphomimic and phosphomutant MPKs were equally (or more strongly) labelled with DTBAcATP when compared to WT. In fact, many reports indicate that the dual phosphorylation of the conserved Thr and Tyr in MPKs enhance ATP and substrate binding affinities. As in the case of extracellular signal-regulated kinase 2 (ERK2), monophosphorylation of Tyr185 configures the active site for binding ATP, while the subsequent phosphorylation of Thr183 enhances myelin basic protein (MBP)

substrate binding (Lew, 2003). In addition, active protein kinases have generally higher affinity to ATP and peptide substrates compared to their inactive unphosphorylated conformations (Knight & Shokat, 2005; Wodicka et al, 2010) and this helped shape the growing numbers of type I protein kinase inhibitors used as drugs in cancer therapy, for example. Type I inhibitors of protein kinases target the ATP-binding loop of protein kinases, and are ATP competable. However and noteworthy, these type I inhibitors do not discriminate between active and inactive conformations of kinases, that is, they bind both indiscriminately (Liu & Gray, 2006). This may reflect on DTBAcATP labelling of both phosphomimic and phosphomutant MPKs in our assays, suggesting AcATP labelling is insensitive to the phosphorylation state of proteins. Since both type I inhibitors of protein kinases and DTBAcATP target the ATP binding loop of protein kinases, we infer that both have a similar binding mechanism. This somehow indicates that DTBAcATP, like type I inhibitors, also binds both active and inactive protein kinase conformations indiscriminately. We can not completely rule out, however, that other factors, besides the conformational changes resulting in the activation or deactivation of protein kinases, also influence ATP binding affinities. Eathiraj and colleagues reported non-conserved residues around the ATP binding loop that contribute to the overall polarity of the loop, which affects the accessibility of the pocket to nucleotides and/or inhibitors (Eathiraj et al, 2011). Moreover, ATP binding affinities to the ATP binding pockets of protein kinases is sensitive to small perturbations on assay conditions, in the type and amount of substrate used, and in divalent cations used (Knight & Shokat, 2005). All these factors are to be considered in examining protein kinase binding affinities to ligands such as AcATP. In this regard, further investigations of AcATP labelling of differentially activated MPKs in various labelling conditions should be done. One can test different labelling buffers, varying amounts of ATP and probe, and different type and amount of substrates and divalent cations.

In summary, our results confirm that AcATP labelling requires the conserved Lys in the ATP binding domain of MPKs. In addition, AcATP labelling of MPKs is not dependent on phosphorylation state. Taken together, the above results suggest that AcATP labelling reports on the availability of ATP-binding sites, rather than the activation state of protein kinases. This implicates that also inactive kinases bind ATP. Furthermore, DTBAcATP labelling of Asp mutants of GST-MPK3/6 fusion

proteins could be investigated to further characterize ATP binding and AcATP labelling in differentially activated Arabidopsis MPKs, since the conserved Asp in the DFG motif is implicated in binding and stabilizing ATP.

3.3 Putative receptor shedding of RLK902

Plants and animals respond to a variety of environmental stimuli including pathogens, through receptor-like kinases (RLK). Arabidopsis encodes for over 600 RLKs, which reflects the importance of these proteins in the overall development and health of the plant (The Arabidopsis kinase database http://bioinformatics.cau.edu.cn/athKD/index.htm). RLKs usually contain an N-terminal signal peptide, an extracellular domain containing a putative ligand-binding site, a single transmembrane spanning domain, and a cytoplasmic protein kinase catalytic domain (Shiu & Bleecker, 2001). In plants, the largest RLK class is the LRR-RLKs, to which RLK902 belongs. LRR-RLKs carry characteristic leucine-rich repeats (LRRs) of 20-30 amino acids spanning the extracellular domain (Kajava, 1998).

RLKs are grouped according to the presence (RD) or absence (non-RD) of Arg (R) residue before the catalytic Asp (D) residue in the subdomain VI of the kinase. RD kinases are activated through autophosphorylation of their activation loop while non-RD kinases do not undergo autophosphorylation, but require another protein for their activation and regulation (Dardick & Ronald, 2006). Some of non-RD kinases are constitutively active (Johnson et al, 1996). Most RLKs belong to the RD class. RLKs that act as pattern-recognition receptors (PRRs) involved in plant immunity, e.g. *Arabidopsis* flagellin sensitive 2 (FLS2) and rice *Xanthomonas* resistance 21 (XA21), are categorized as non-RD and their activation indeed involves other kinases. RLK902, which lacks the conserved Arg and Asp residues in subdomain VI, does not belong to either RD or non-RD, but rather to the alternative catalytic function (ACF) class (Dardick & Ronald, 2006). Although RLK902 carries the other two conserved catalytic residues namely Lys in subdomain II and Asp of the DFG motif in subdomain VII (Dardick & Ronald, 2006), this indicates that functional regulation of RLK902 is distinct. RLK902 is a phosphoprotein that is phosphorylated on at least four Ser residues - three in juxtamembrane region and one in the less conserved region between subdomains VII and VIII in the kinase domain (Fig. 35) (Nuhse et al, 2004) (http://phosphat.uni-hohenheim.de/).

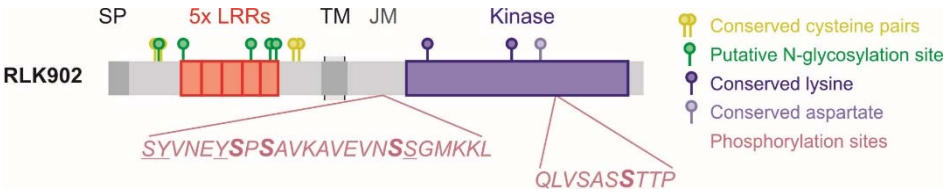


Figure 35. Phosphorylation sites in RLK902.

The two phosphorylation hotspots in RLK902 are found in the less conserved regions of the protein: (1) the juxtamembrane domain, and (2) the region between kinase subdomains VII and VIII. Underlined residues represent ambiguous phosphorylation sites. Residues in bold are phosphorylation sites experimentally validated and detected in phosphoproteomics analysis (Nuehse et al, 2004; http://phosphat.uni-hohenheim.de/).

Cell surface-localized RLKs in animals and plants are responsible to relay information from the cell´s outside environment by extracellularly binding to ligands and transmitting these signals intracellularly, or extracellularly onto neighbouring cells. An example is the Arabidopsis FLS2, which binds to bacterial flagellin-derived elicitor flg22. The binding of FLS2 to flg22 triggers its interaction with BAK1, resulting to a series of phosphorylation events activating an MPK (Asai et al, 2002; Bethke et al, 2012; Gao et al, 2008) and CPK (Boudsocq et al, 2010) signalling pathways in the cytoplasm. These signalling events lead to the activation and expression regulation of downstream target genes in the nucleus.

In animals, signal perception and transduction involves ectodomain shedding of extracellular domain of several protein receptors. Ectodomain shedding is a function of proteolytic enzymes called sheddases (Chow & Fernandez-Patron, 2007), which results to the release of soluble ligand-binding domains that can bind and activate proteins in the cell or in neighbouring cells, thereby controlling expression of downstream genes (Arribas & Borroto, 2002). Hence, ectodomain shedding regulates the expression and function of transmembrane proteins, and this is implicated in many cellular and physiological processes involved in development, homeostasis, inflammation and pathology (Horiuchi, 2013; Weber & Saftig, 2012; Xu & Derynck, 2010). In some cases, ectodomain shedding leads to further processing in the transmembrane domain by other proteases, releasing cytoplasmic domains which can associate with other cytoplasmic proteins, and/or travel to the nucleus to regulate transcription of target genes (Arribas & Borroto, 2002; Heldin & Ericsson, 2001; Ni et al, 2001).

In addition to ectodomain shedding, receptors are also regulated by internalization into vesicles. This mechanistic regulation of receptor proteins is called receptor-mediated endocytosis (RME). In RME, membrane-bound receptors are internalized into vesicles, either to recycle back to the cell membrane, or transport into lysosomes for degradation. Endocytosis regulates RLK signalling through downregulation of RLKs, or target RLKs into endosomes where their downstream signalling effectors are located (Teis et al, 2002). For example, receptor tyrosine kinases (RTKs) like epidermal growth factor receptor (EGFR), neurotrophic tyrosine kinase receptor type 1 (TrkA), and insulin receptor (IR) are functionally regulated through RME. In this chapter, we investigated an LRR-RLK called RLK902, and discuss the possible regulation of its function through a putative shedding mechanism similar to ectodomain shedding, or a putative endocytotic internalization and eventual degradation.

Western blot analysis showed the accumulation of the Myc-RLK902-YFP in the membrane fraction of *N. benthamiana* leaves (Fig. 28). In addition, we showed that Myc-RLK902-YFP colocalizes with the remorin-RFP plasma membrane marker (Fig. 27). These results support the previous finding that RLK902-GFP fusion protein was detected on the cell surface after plasmolysis using mannitol (Tarutani et al, 2004a) 2004a) indicating that it is a plasma membrane protein. These validate, first of all, that RLK902 is a receptor-like protein carrying a signal peptide and a transmembrane domain. Moreover, the detection of N-glycosylation of transiently expressed RLK902 in *N. benthamiana* leaves is consistent with the presence of the predicted five N-glycosylation sites in the extracellular domain of RLK902. N-glycosylation is typical for the ectodomain of membrane-bound proteins. Furthermore, we have observed that RLK902 is processed *in vivo* through time. With full length RLK902 detected at two dpi, and processed products observed beginning at three dpi that gone more pronounced at four and five dpi (Fig. 29A). Moreover, substitution of the conserved residues K393, K492 and D508 in the intracellular kinase domain into Ala (A) did not affect the processing of RLK902 (Fig. 33), suggesting that kinase activity is not required for processing. K393 in subdomain II and K492 in subdomain VI are the conserved Lys residues predicted to be labelled by AcATP/AcGTP, while D508 is the conserved Asp in the DFG motif of subdomain

VII. The labelling and detection of a cleaved/processed RLK902 protein is interesting. This indicates that it is not immediately degraded, and has a time window, during when it can be functional in the cell. In case the processed RLK902 is still active and functional, it would make sense to assess the activity state of uncleaved protein, and whether or not activation of RLK902 requires processing. The processing sites can be identified by Ala scanning mutagenesis of putative cleavage sites.

A number of inhibitors were screened for inhibition of processing of RLK902. Interestingly, protease inhibitors E64d and TLCK inhibited the *in vivo* processing of RLK902. Moreover, E64d and TLCK triggered the accumulation of endosomal-like structures in RLK902-agroinfiltrated *N. benthamiana* leaves (Fig. 31). These results suggest that E64d and TLCK may inhibit the protease that cleaves RLK902 and/or blocks RLK902 vacuolar degradation pathway. However, inhibitors of the vacuolar targeting pathways (BFA, wortmannin, Tyr A23, concanamycin A, latrunculin B, and cantharidin) did not inhibit RLK902 processing (Fig. 32).

Internalization of plasma membrane-localized protein receptors by RME is required for efficient receptor-mediated signal transduction. This is exemplified by EGFR and Toll-like receptor-mediated cell responses in animals (Johnsen et al, 2006; Teis et al, 2002). In plants for example, signal transduction mediated by FLS2 involves internalization into endosomes and degradation in vacuoles upon activation by flg22 binding (Beck et al, 2012; Robatzek et al, 2006). Evidently, the endocytotic trafficking of protein receptors is also a function of certain sequence motifs associated with internalization. Some reported internalization signals include the dileucine motifs D/ExxxLL/I or DxxLL, the tyrosine-based motifs NPxY, YxxΦ, YxxxΦN, FxNPxY, YRAL, and YFIPIN (where X stands for any amino acid and Φ represents any bulky hydrophobic amino acid), and the ubiquitination signals for vacuolar degradation of proteins (Goh & Sorkin, 2013; Kozik et al, 2010; Wang et al, 2007). Interestingly, sequence analysis reveals that RLK902 carry at least four putative internalization motifs. These are two dileucine internalization motifs DLEDLL and DEKLL, which follows the consensus D/ExxxLL/I and DxxLL, respectively, and two tyrosine-based internalization motifs YKAV and YDFM, both follows the consensus YxxΦ (Fig. 35) (Goh & Sorkin, 2013; Kozik et al, 2010). Moreover, it carries lysines, serine and

threonines which can be ubiquitinated, which signals for endocytosis and eventual protein degradation in vacuoles (Goh & Sorkin, 2013; Kozik et al, 2010; Wang et al, 2007)

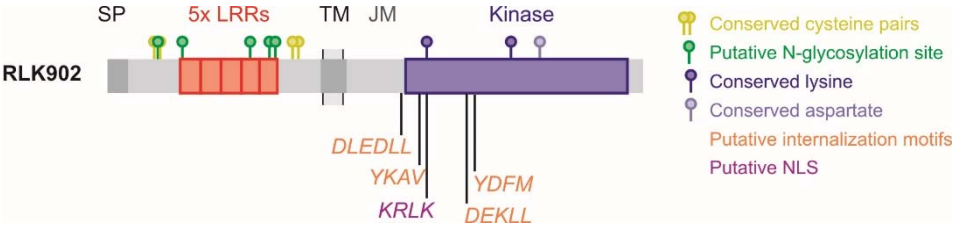


Figure 36. Sequence analysis of RLK902.

Protein sequence analysis reveals putative short sequence motifs in RLK902 that signals for its internalization into vacuoles or its transport into the nucleus. Four putative internalization signals (orange) and one nuclear localization signal (NLS, purple) were found. There are two dileucine internalization motifs namely DLEDLL and DEKLL following the consensus D/ExxxLL/I and DxxLL, respectively, and two tyrosine based internalization motifs namely YKAV and YDFM, both following the consensus YxxΦ. X stands for any amino acid and Φ represents any bulky hydrophobic amino acid. KRLK is a putative NLS, following the consensus K(K/R)x(K/R).

The *in vivo* processing of RLK902 could function to release its ectodomain to be a soluble extracellular signal molecule for the cell or to a distant neighbouring cell. However, the possibility of a second proteolysis at the intramembrane or juxtamembrane domain can not be excluded. This could release a soluble cytoplasmic molecule for intracellular signalling. A first report on a plant LRR-RLK, the rice XA21, is proteolytically cleaved during signalling to release a cytosolic protein kinase domain that migrates to the nucleus to phosphorylate transcription factors (Park & Ronald, 2012). A similar signalling mechanism could be at work for RLK902. Noteworthy, a putative nuclear localization signal (NLS) KRLK following the classic monopartite three basic amino acid NLS motif K(K/R)x(K/R) (Hicks et al, 1995; Kosugi et al, 2009; Lange et al, 2007) is found in the kinase domain of RLK902 (Fig. 35), which gives a hint to a possible nuclear transport of the protein kinase.

At this point, our motivation is led by the following hypotheses: RLK902 processing is caused by (a) putative receptor shedding event, (b) by RME, (c) or both. However, our speculations are yet to be validated. First of all, what triggers processing and/or internalization of RLK902 should be established. The processing of RLK902 was first observed in an unelicited condition, indicating that it could be a constitutive

processing event. Several experiments can be performed to test this. First, mutagenizing the putative internalization motifs could block processing. Second, biochemical analysis using the NLS mutant of RLK902 in combination with trafficking markers and microscopy studies can be used to investigate possible translocation to the nucleus. Third, MS analysis of cleavage sites in Arabidopsis, where the processing was originally observed can be done, together with biochemical probing of putative sheddase(s), or putative downstream target genes, if any. Moreover, we can express RLK902 by agroinfiltration into protease-depleted *N. benthamiana* leaves. Also importantly, some of these experiments should be repeated in Arabidopsis seedlings to confirm that this processing upon agroinfiltration is not an overexpression artefact.

CHAPTER 4: MATERIALS AND METHODS

4.1 Biological materials

Plant materials and growth conditions

Arabidopsis thaliana ecotype *Col-0* was grown in a climate cabinet in a 10-hour light regime at 22ºC and 60% relative humidity. Plant proteins were extracted from rosette leaves of 4-6 week old Arabidopsis plants.

Arabidopsis cell cultures (ecotype Landsberg) were weekly subcultured in modified MS-medium (4.41% MS-medium, 3% sucrose, 5.4 µM α-Naphtyleneacetic acid, 0.23 µM kinetin, pH 5.6) in light at 22°C at 160 rpm continuous shaking. Five to seven day old cultures were used for experiments. Prior to treatment, cells were suspended in fresh medium. After treatment, medium was discarded and cells were quickly rinsed in fresh medium. The cells were then either frozen in liquid nitrogen for storage, or immediately lysed for labeling experiments. Labelling was always done in fresh lysates. Protein lysate was prepared by grinding cells in labeling buffer, followed by centrifugation to clear the lysate.

Nicotiana benthamiana was grown in a climate chamber at a 14-hr light regime at 22 °C (day) and 18 °C (night). Four to six-week old plants were used for experiments. Proteins were usually extracted by bead-based agitation in SDS gel loading buffer.

Other biological materials

Pseudomonas syringae pathovar (pv.) *tomato* (Pto) DC3000 was from Dr. Jane Parker (Max Planck Institute for Plant Breeding Research, Cologne, Germany). PtoDC3000 lacking HopQ1-1 effector (PtoDC3000ΔhQ) was from Dr. Johana C. Misas-Villamil (Max Planck Institute for Plant Breeding Research, Cologne, Germany). *Pseudomonas syringae* pv. *syringae* (Psy) B728a was provided by R. Dudler (Institute of Plant Biology, University of Zurich, Switzerland).

4.2 Chemical materials

Activity-based probes

AcATP probes (BAcATP, ≡AcATP and Bodipy-AcATP) were synthesized following the procedures described by Patricelli et al. (2007) with some modifications. These were synthesized by Dr. Julian Oeljeklaus in Universität Duisburg Essen, Essen, Germany. Briefly, e.g. BAcATP, was synthesized as follows: *N*-(+)-biotinyl-6-aminohexanoic acid (30 mg, 0.085 mmol) was suspended under an argon atmosphere in anhydrous dioxane/DMF/DMSO (1:1:1, 3 mL) and cooled to 0 °C. To this ice-cold solution were then added triethylamine (47 µL, 0.34 mmol) and *iso*-butyl chloroformate (33 µL, 0.255 mmol), resulting in the appearance of a precipitate. The resulting suspension was stirred at 0 °C for 20 min, allowed to warm up to room temperature, and then stirred for an additional 1.5 h. ATP triethylammonium salt (69 mg, 0.085 mmol), dissolved in anhydrous DMSO (1 mL), was added to the reaction mixture. After 18 h, the reaction was stopped by addition of water (4 mL) and the mixture was extracted with ethyl acetate (3 x 4 mL). The aqueous layer was frozen and subsequently lyophilized. The resulting solid was suspended in water (1 mL), transferred to a RP-C_{18} flash-column (LiChroprep® RP-18, 40-63 µm, Merck) pre-equilibrated with 5% acetonitrile in water and eluted with 30% acetonitrile in water. Product containing fractions were pooled, frozen, lyophilized and subsequently stored at -20 °C (Yield: 4.3 mg, 5.1 µmol, 6%). Identity of the product was confirmed by LC-ESI-MS analysis on a Thermo Scientific LCQ FleetTM ESI-spectrometer, equipped with an Eclipse XDB-C18, 5 µm column from Agilent and by using a linear gradient of solvent B (5 mM NH_4OAc in acetonitrile) in solvent A (5 mM NH_4OAc in H_2O) at 1 mL/min flow rate (gradient program: 0 min / 10% B → 1 min / 10% B → 10 min / 100% B → 12 min / 100% B → 15 min / 10% B.): R_t = 2.16 min, calcd. for $C_{26}H_{40}N_8O_{16}P_3S^-$ $[M-H]^-$ 845.15, found 845.13.

DTBAcATP was purchased from Thermo Scientific. The tags N_3-Rh and N_3-BRh used in click chemistry reactions with ≡AcATP was provided by Dr. Markus Kaiser (Universität Duisburg Essen, Essen, Germany). All probes were dissolved in DMSO and stored at -20 °C. The AcGTP probe used to label RLK902 is from Thermo Scientific.

Chemical inhibitors

Kinase inhibitors Toyocamycin, Suramin, Wortmannin, 5,5'-Dithiobis(2-nitrobenzoic acid) (DTNB), and K252a were from Sigma. The other chemical inhibitors used in the RLK902 study were obtained from different sources. Soybean kunitz, E-64d, MG132, PMSF, Phenantroline, Brefeldin A, and Tosyllysine chloromethyl ketone hydrochloride (TLCK) were from Sigma. Epoxomicin was from ApexBio. Odanacatib was from ChemScene. Tyrphostin A23 and Concanamycin A were from Santa Cruz.

Other chemicals

The nucleotides and nucleotide analogues ATP, GTP, CTP, TTP, ADP, AMP, and NADP were from Sigma. NAD was from Roche. The deoxynucleotides dATP, dGTP, dCTP and dTTP were from Fermentas. General laboratory chemicals and reagents were mainly purchased from Sigma (St.Louis, USA) and Merck (Darmstadt, Germany).

Antibodies

The anti-cMyc (sc-789) antibody was from Santa Cruz and anti-GFP monoclonal antibody (632375) was from Clonetech. The anti-rabbit HRP (1858415) and anti-mouse HRP (1858413) antibodies were from Pierce. The anti-phosphopeptide antibody for the detection of phosphorylated MPK3/6 is from Cell Signalling Technology (Phospho-p44/42 MAPK (Erk1/2) (Thr202/Tyr204) Antibody #9101).

4.3 Methods

Quantification of protein concentration

Protein concentration was quantified using the RCDC protein assay kit (Biorad). Two different concentration dilutions of protein samples were prepared and their concentrations were calculated using a standard curve measured from serial dilutions of a BSA standard (Promega).

Labeling with AcATP

Arabidopsis proteome was extracted from 4-week old Arabidopsis rosette leaves in 50 mM Tris pH 7.5. The lysate was cleared by centrifugation, and subjected to gel filtration using DG10 columns (Biorad). Labelling was performed in 50 mM Tris pH 7.5 with 10 mM $MgCl_2$, unless specified otherwise. The lysate was incubated with 5 or 10 µM AcATP or DMSO for the no-probe-control, at room temperature (RT) for 15 minutes. For inhibition experiments, the lysate was preincubated with the different inhibitors at 10 mM final concentration for 30 minutes before labeling with AcATP for 15 minutes. The labeling reaction was stopped by adding gel loading buffer containing SDS and DTT. For labeling with ≡AcATP, the labelled sample was first acetone precipitated (for 50 µl reaction) with cold 100% acetone or gel filtrated (for large scale labeling, 1 ml to 2 ml reactions) through a DG10 column (Biorad), followed by click chemistry with N_3-Rh or N_3-BRh to attach a rhodamine, or biotin and rhodamine, respectively, alkyne on the labelled proteins. For labeling with DTBAcATP, protein extracts were prepared and labeled in a lysis buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 0.1% Triton-X-100, phosphatase inhibitors [Cocktail II AG Scientific #P-1518]). Proteins were usually separated on 12% acrylamide gel. For biotinylated or desthiobiotinylated protein detection, proteins were transferred onto polyvinylidene fluoride (PVDF) membranes and the protein blot was probed with streptavidin-HRP (Ultrasensitive, Sigma) and detected with enhanced chemiluminescence (SuperSignal West Femto Chemiluminescent Substrate, Thermo Scientific). For rhodamine or bodipy-tagged proteins, the protein gels were scanned for fluorescence on a Typhoon 9000 scanner.

Click Chemistry

After labeling with ≡AcATP, protein samples were acetone precipitated with cold 100% acetone to remove unreacted probes and salts. The samples were then denatured and dissolved in 1% SDS in PBS (80 g NaCl, 2 g KCl, 26.8 g Na_2HPO_4, 2.4 g KH_2PO_4 in 1000 ml water pH 7.4) solution. To the denatured sample, 20 μM N_3-Rh or N_3-BRh were added, followed by 1 mM TCEP, then 34 μM TBTA, and then 1 mM $CuSO_4$, with quick vortexing upon every reagent added. The reaction mix was incubated at RT in the dark for 1 hr., and then acetone precipitated remove copper and unreacted tags. For gel electrophoresis, the samples were dissolved in gel loading buffer. For pull down experiments, the samples were dissolved in 2% SDS before affinity purification with streptavidin beads.

Western blotting

After SDS-PAGE, proteins were transferred onto PVDF membrane (Immobilon-P, Millipore) at 200 mA for 60-70 min using X-Cell II Blot Module system (Invitrogen). After the transfer, the membrane was incubated with 5% BSA (Biomol) in TBS (50 mM Tris, 150 mM NaCl pH 7.5) solution for 20 min with gentle agitation on a shaker. To detect biotinylated proteins, the membrane was incubated with streptavidin-HRP (Ultrasensitive, Sigma) at 1:3000 in 5% BSA in TBST (TBS with 2% Tween-20) solution for 1 hr. Washing of membranes to remove unbound streptavidin-HRP was done 5 times with TBST for 10 min each wash step. To detect non-biotinylated proteins, the membrane was incubated with protein specific primary antibodies for 1 hr followed by 5x washing step, before incubation with HRP-conjugated anti-rabbit or anti-mouse secondary antibodies (Pierce) at 1:5000 in 5% BSA-TBST for 1 hr. Final washing was done 5x with TBST. The membrane was then incubated for 1 minute with chemiluminescent substrates of HRP (SuperSignal West Pico/Femto, Pierce) underneath a piece of overhead-projection transparency, and exposed to X-ray films (BioMax MR, Kodak) in a darkroom. The exposed film was developed by automatic X-ray film processor (Optimax, Protec).

In-gel fluorescence scanning

SDS-PAGE gels containing fluorescent proteins were washed three times with water

and labeled proteins were visualized by in-gel fluorescence scanning using a Typhoon 9000 scanner (Molecular Dynamics) with excitation and emission at 532 and 580 nm, respectively.

Different proteome preparations for in vitro labeling studies

For Arabidopsis cell culture treatments, 1 ml of 5 day old cultures were transferred to a sterile 12-well plate and cell suspensions were acclimatized to fresh media with continuous shaking for 1 hr before each treatment. For heat treatments, the cell suspension was transferred into a 1.5 ml eppendorf tube and incubated in a water bath at 45 or 53 °C for 5 minutes. For flg22 treatment, 100 nM flg22, or DMSO for the control, was added to the cell cultures for 5 or 15 minutes. For BTH treatment, 25 µM benzothiadiazole (BTH) (Bion, Syngenta) was added to the cultures for 30, 60 or 360 minutes. After treatment with flg22 or BTH, cells were quickly rinsed twice with fresh medium and the cells collected by centrifugation for protein extraction.

For Pseudomonas infections of Arabidopsis plants, the PtoDC3000ΔhQ strain was grown overnight 28 °C in 10 ml LB medium in a 50 ml Falcon tube. The culture was centrifuged at 3000 g for 10 min, and the bacterial pellet was washed with 10 mM $MgCl_2$ and diluted to OD 600 of 0.3 (measured by spectrometer Ultrospec II, LKB Biochrom) with 10 mM $MgCl_2$ and 0.01% Silwet. The bacterial suspension was then sprayed with a perfume sprayer onto the leaf surface of 4-6 week old Arabidopsis plants until the droplets ran off. Inoculated plants were kept in trays with transparent covers to maintain high humidity, and grown under standard conditions in a growth chamber. Rosette leaves were harvested at 2 dpi for membrane fractionation and labeling with BAcATP.

For pathogen infections of *N. benthamiana* plants, the bacterial strains PtoDC3000, PtoDC3000ΔhQ and Psy B728a were grown overnight 28 °C in 10 ml LB medium in a 50 ml Falcon tube. The culture was centrifuged at 3000 g for 10 min, and the bacterial pellet was resuspended in water to a final OD 600 of 0.2. These bacterial suspensions were then syringe-infiltrated into fully expanded *N. benthamiana* leaves using needleless 1 ml syringe (Plastipak, BD). Infiltrations with water, and no infiltrations, were used as controls. *N. benthamiana* leaves were collected at 2 dpi for

apoplastic fluid extractions and labeling with DTBAcATP.

For proteome extraction from Arabidopsis seedlings, seeds were sterilized with bleach and 70% ethanol, and rinsed extensively with sterile water. The sterile seeds were sown and initially grown on a sterile MS agar plate for two weeks. The seedlings were then transferred into a fresh liquid MS medium one day before treatment. For flg22 treatments, the seedlings were incubated with 1 μM flg22, or DMSO for control, for 15 minutes. After treatment, the seedlings were rinsed twice with fresh media and then blotted dry with absorbent paper. Protein extract was prepared in the labeling buffer.

Membrane fractionation

Arabidopsis or *N. benthamiana* leaves of 2 g fresh-weight were homogenized with mortar and pestle on ice in 4 ml cold membrane buffer (25 mM Tris pH7.5, 250 mM sucrose) containing protease inhibitor cocktail (Roche) and 3% polyvinylpolypyrrolidone (PVPP), and filtered through a 64 μm nylon mesh. Samples were collected in a 50 ml tube and centrifuged at 4000 rpm for 15 min at 4 °C. The supernatant was then ultracentrifuged at 41000 rpm for 1 hr at 4 °C. The pellet was resuspended in cold membrane buffer and stored at -80°C. The supernatant was the soluble fraction and the pellet was the membrane fraction.

***N. benthamiana* apoplastic fluid isolation**

Leaves were vacuum-infiltrated with ice cold water in a 500 ml beaker, using an exicator. The water-infiltrated leaves were then surface dried using absorbent paper, and centrifuged at 1600 g for 10 min at 4 °C in an AF isolation tube containing holes of 1 mm diameter. The AF was collected in a bottom collection tube.

2-dimensional (2D) protein gel electrophoresis

After purification of labeled proteins, the protein samples were dried under vacuum and dissolved in 160 μl isoelectric focusing (IEF) buffer containing 7 M urea, 2 M thiourea, 2% CHAPS, 1% v/v ampholytes (ZOOM Carrier pH 3-10, Invitrogen), 0.002% bromophenol blue and 1% w/v DTT. This sample was then impregnated into

an IPG strip (ReadyStrip pH 3-10, Bio-Rad) overnight at room temperature. Proteins were separated on the IPG strip in ZOOM IPGRunner cassettes (Invitrogen) attached to a ZOOM IPGRunner Core (Invitrogen) using a voltage program (175 V for 15 min; 175 to 2000 V ramp for 45 min; 2000 V for 30 min) of a voltage ramp-compatible power supply (EPS 3501XL, GE). After IEF, strips were incubated in 5 ml of equilibration buffer (6 M urea, 0.375 M Tris pH8.8, 20% glycerol, 2% SDS) containing 2% w/v DTT for 30 min, followed by incubation in 5 ml equilibration buffer containing 2.5% w/v IAA for another 30 minutes. The strips were then mounted on top of a 12% SDS PAGE gel and the proteins were separated at 200 V using the Novex Minicell system (Invitrogen).

Affinity purification and Identification of BAcATP-labelled proteins

After labelling of Arabidopsis leaf extract (1mL 5 mg/mL protein) for one hour, the extract was desalted using a DG10 column (Biorad). The eluate was incubated with 100 µL neutravidin beads (Thermo Scientific) in PBS solution with 0.2% SDS for one hour. The beads were washed twice with 0.2 % SDS, twice with 6 M urea and twice with water, and eluted with 50% formic acid. The eluent was then dried under vacuum and dissolved in gel loading buffer. Proteins were separated on acrylamide gels and stained with SYPRO Ruby (Invitrogen). Bands of interest were excised from both labelled and control samples and subjected to in-gel digestion with trypsin. LC-MS/MS analysis was performed on an LTQ Velos (Thermo) coupled to an EasyNanoLC (Proxeon/Thermo) using a ‘Top 20’ data-directed acquisition. LC was performed on a C18 column (10 cm x 75 µm) with a gradient of 5-40% acetronitrile in 0.1% formic acid for 30 minutes followed by a wash with 95% acetonitrile. The data-directed acquisition included active exclusion for mass (±40 ppm) and time (60 sec.). Spectra were searched using Mascot 2.3 using a fixed modification of cysteine (57.02 Da for carbamidomethyl), and variable modifications of methionine (15.99 Da for oxidation) and lysine (339.16 and 355.16 Da for BAc and OBAc labeling, respectively). A mass tolerance was set to 0.3 Da for the precursor ions and 0.4 Da for fragment ions. Up to two mis-cleavages for trypsin were permitted since the labeling would prevent cleavage at the labeled lysine. The MS2 spectra were searched against the TAIR10_pep_2012 (version November 2012) containing 35386 proteins of *Arabidopsis thaliana*, supplemented with a small database with 1095

artefact proteins and a reversed decoy database of the same proteins (totals 72962 entries). Peptides were retained with Mascot scores of >41, which is above the Mascot significance level. Hits to the decoy database within this selection there were giving a 0.3% false discovery rate (FDR) for all peptides and 0.7% FDR for non-redundant peptides. Proteins were selected having a minimum of two different peptides, shown in supplemental Table S1. This selection for two peptides per protein removed all the false positive hits to the decoy database. Proteins were ranked on their score and manually selected considering the occurrence of the protein in the rest of the gel and its expected molecular weight and overall score and spectral count (supplemental Table S2). Only robustly identified proteins that are highly enriched in the gel slice compared to other gel slices are reported.

Labelling and MS analysis with DTBAcATP

Analysis of DTBacATP-labelled Arabidopsis leaf extracts was performed as described (Patricelli et. al., 2007, 2011). Briefly, lysates from Arabidopsis leaves were generated by bead-based agitation in kinase buffer (20 mM HEPES pH 7.8, 150 mM NaCl, 0.1% Triton X-100). The resulting lysates were gel filtered into 20 mM HEPES pH7.8, 150 mM NaCl, 1% v/v phosphatase inhibitor cocktail II (Calbiochem) to remove endogenous competing nucleotides prior to labeling with 20 µM DTBAcATP probe for 15 minutes in the presence of 20 mM $MnCl_2$. The labelled extracts were denatured and reduced (6M urea, 10 mM DTT, 65 ^oC for 15 minutes), then alkylated with iodoacetamide (40 mM, 30 minutes at 37 ^oC). Following another gel filtration step (into 10 mM ammonium bicarbonate, 2 M urea) to remove unreacted reagents and lower the urea concentration, the samples were digested with trypsin. DTBAcATP-labelled peptides were enriched on streptavidin resin (Thermo Scientific) and eluted using 0.1% TFA in 50% acetonitrile.

MS analysis: MS analysis was performed on a Thermo-LTQ linear ion trap instrument in a data dependent mode as described for initial characterization (Patricelli, et. al., 2011) using a mass range from 500-1800 *m/z*. The MS2 spectra were searched using SEQUEST against the TAIR9_pep_20090619 database (version June 2009) containing 32,769 proteins of *Arabidopsis thaliana*. MS2 searches included fixed iodoacetamide modification of cysteine (57 Da for

alkylation), and variable modifications of methionine (16 Da for oxidation) and lysine (196 Da for DTBAcATP labeling). Up to three mis-cleavages with trypsin were allowed and non-tryptic or half-tryptic peptides were excluded. A mass tolerance was set to 3 Da for the precursor ions and 1 Da for fragment ions. The high mass tolerance for precursor ions had to be used since most of the precursor ions are detected with multiple charges and mass tolerance in SEQUEST is not charge state specific. A probability score was calculated for each labelled peptide as described previously (Adam, et. al., 2004). This score is based on Xcorr, delta-Cn and peptide mass/length and compared to the distribution of false positive scores generated by searching with incorrect masses of the modification (+10 and -10 Da from the correct value). The resulting MS2 spectra were assembled in one Excell file and all spectra with <95% probability were removed. Peptides that were only detected in one of the four MS runs were discarded. The resulting list of 242 peptides includes peptides that are unique (u) in the Arabidopsis proteome or match multiple genes (ambiguous, a). Some peptides match multiple gene models of a single gene (isoforms, i). These isoforms are indicated in the supplemental table but were treated as identifiers for specific genes. Supplemental Table S3 contains the lists with 10465 spectral counts with individual scores with probability of >75%. Supplemental Table S4 summarizes the spectral counts of the 242 labeled peptides with >95% probability.

Analysis of labelling sites: To map the labeling site onto protein structures, we searched the PDB database for sequences that are homologous to the labeled proteins. Sequences for which a co-crystal with a nucleotide was available were aligned with the identified proteins and the orthologous position of the labeled lysine was indicated in the alignment. This residue was selected in the protein structure using PyMol and the distance to the nearest phosphate of the bound nucleotide was measured.

Comparison with protein levels in the AtProteome database: For every labeled protein of Table S4, we retrieved the number of spectral counts from leaves from the AtProteome database (Baerenfaller et al., 2008; tgcz-atproteome.unizh.ch/). For ambigious peptides, only the protein with the highest spectral counts was selected

In silico analysis of labelling sites: To predict the labeled peptides for all the protein kinases in Arabidopsis, we retrieved protein sequences of all protein kinases from the *Arabidopsis thaliana* kinase database (http://bioinformatics.cau.edu.cn/athKD/index.htm). Sequences for each subfamily in the AthKD database were retrieved and aligned using MultiAlign (http://multalin.toulouse.inra.fr/multalin/multalin.html). The labeled peptide that was experimentally identified was highlighted in the alignment and the orthologous Lys residues were indicated and Lys and Arg residues were highlighted in the alignment. Tryptic peptides containing the labeling site were selected from the alignments. The orthologous labeled Lys, and the Arg/Lys-Pro sites were not considered as trypsin cleavage sites, and miscleavages were ignored. The selected peptides were ranked on sequence and redundancy for each peptide sequence was counted (*n* proteins carrying the same labeled peptide). Ile/Leu residues were not discriminated in this analysis as they have the same mass. Protein kinases with a unique labeled lysine (n=1) were counted. The other protein kinases contain ambiguous peptides (n>1). This analysis was done for peptides carrying Lys1, Lys2, Lys3 and Lys4.

Targeted MS/MS analysis

Arabidopsis cell pellets were lysed by sonication in lysis buffer (50 mM HEPES, pH 7.5, 150 mM NaCl, 0.1% Triton-X-100, phosphatase inhibitors [Cocktail II AG Scientific #P-1518]). After lysis, the samples were cleared by centrifugation, and the supernatant desalted by gel filtration over Sephadex G25. $MnCl_2$ was then added to a final concentration of 20 mM. Final protein concentration of lysates was 6 mg/mL. DTBAcATP probe was added to each sample to a final concentration of 5 µM for 10 minutes. After labeling, samples were prepared for targeted MS analysis using the ActivX standard protocol (Patricelli et al, 2011). Briefly, samples denatured, reduced and alkylated, then digested with trypsin, and the desthiobiotinylated peptides were enriched on streptavidin resin. Enriched labeled peptides were analyzed by liquid chromatography tandem mass spectrometry (LC-MS/MS) on a Thermo-LTQ ion trap mass spectrometer. An inclusion list of about 250 Arabidopsis kinases and ATP-binding proteins was generated and used to quantify signal intensities of detected labeled peptides (Patricelli et al, 2011).

All quantitation was performed by extracting characteristic fragment ion signals from targeted MS/MS spectra and comparing signals in control and treated samples. For each peptide quantitated, the change in MS signal for the treated samples relative to the MS signal for the control samples is expressed as fold-change. In cases where the average MS signal of the treated samples was greater than the average MS signal of the control samples, the following equation was used:

$$Foldchange = -\left(\frac{Average\ MS\ signals\ from\ treated\ samples}{Average\ MS\ signals\ from\ control\ samples}\right)$$

In cases where the average MS signal of the control samples was greater than the average MS signal of the treated samples, the following equation was used:

$$Foldchange = -\left(\frac{Average\ MS\ signals\ from\ control\ samples}{Average\ MS\ signals\ from\ treated\ samples}\right)$$

Generation of mutant GST-MPK3/6 fusion constructs

GST-MPK3 and GST-MPK6 encoding pGEX-3X (GE Healthcare) plasmids was previously reported in (Sorensson et al, 2012) and (Ulm et al, 2002), and was donated by Dr. Erik Andreasson (Lund University, Sweden). These were used as template for site-directed mutagenesis to create the different GST-MPK3 and GST-MPK6 mutant fusion constructs. PCR was done using forward primer (FP) 5'-GAATGATTTTATGGATGAGGATGTTGTTACGAGATGG-3' and reverse primer (RP) 5'- CCATCTCGTAACAACATCCTCATCCATAAAATCATTC-3' for generating MPK3 T196D Y198D, FP 5'-GAATGATTTTATGGCTGAGGCTGTTGTTACGAGATGG-3' and RP 5'- CCATCTCGTAACAACAGCCTCAGCCATAAAATCATTC-3' for MPK3 T196A Y198A, FP 5'- CATTATTCATAGGGATTTAGCGCCGAGCAATCTTCTGTTG-3' and RP 5'- CAACAGAAGATTGCTCGGCGCTAAATCCCTATGAATAATG-3' for MPK3 K166A, FP 5'- GAGTGATTTCATGGATGAAGATGTTGTCACGAGATGG-3' and RP 5'-CCATCTCGTGACAACATCTTCATCCATGAAATCACTC-3' for MPK6 T221D Y223D, and FP 5'-GCTTCACAGGGATTTGGCACCAAGTAATCTCCTCCTG-3' and RP 5'- CAGGAGGAGATTACTTGGTGCCAAATCCCTGTGAAGC-3' for MPK6 K191A. All constructs were electrotransformed into DH5α and selected on LB plates

containing 100 µg/ml ampicillin. Colonies were selected by PCR and the constructs were confirmed by sequencing using the primers FP 5'-GGGCTGGCAAGCCACGTTTGGTG-3' and RP 5'-CCGGGAGCTGCATGTGTCAGAGG-3'.

Recombinant MPK3/6 expression in E. coli and labeling with DTBAcATP

E. coli strain BL21 was transformed with pGEX-3X carrying the WT or mutant GST-MPK3 or GST-MPK6 fusion constructs and grown overnight at 37°C in 5 ml LB medium containing 100 µg/ml ampicillin. The bacterial cultures were transferred into fresh 50 ml LB medium with ampicillin and were allowed to grow to OD 600 of 0.8 for about 2 hours at 37°C. Expression of the recombinant protein was induced by adding 0.4 mM IPTG and incubating at 28°C for 3 hrs. The bacterial cultures were then incubated on ice for 30 minutes, and centrifuged at 4000 rpm for 15 min at 4°C to collect the cells. The supernatant media were discarded and the bacterial pellets resuspended in 2 ml ice-cold non-denaturing lysis buffer. The cells were lysed by tip sonication twice for 30 sec, and centrifuged at 4000 rpm for 15 min at 4°C to clear the lysates. The cleared lysates were gel filtrated into a labeling buffer prior to labeling with DTBAcATP. Labeling was done as previous but the reaction was supplemented with 0.5 µM ATP and 0.1% casein substrate during labeling. Signals were quantified using ImageJ.

Generation of RLK902 WT and mutant fusion constructs

The *RLK902* gene was amplified by PCR from genomic DNA isolated from *Arabidopsis thaliana* ecotype Col-0 using primers FP 5'-AGCTTCTAGACTCGAGAGATCTCGCCGCCGACAAATCCGCTC-3' and RP 5'-AGCTGTCGACCCCACCCGATCTGCACCCGATTG-3'. The 2 kb PCR product was then cloned into pRH604 using XbaI and SalI restriction enzymes to fuse the gene inframe with a YFP tag at its C-terminus. The authenticity of the cloned PCR product was confirmed by sequencing using primers FP 5'-CACACAGGAAACAGCTATGACCATG-3' and RP 5'-GATTAGCATGTCACTATGTGTGCATCC-3'. The confirmed clone was named pJV3. After confirmation of the clone, the RLK902-YFP cassette was shuttled into the

binary vector pRH509 using XhoI and EcoRI restriction sites to fuse RLK902 N-terminally to a 35S promoter followed by a region encoding the signal peptide of PR1a and four Myc tags interrupted by a small intron used to exclude bacterial expression. The resulting construct was then confirmed by restriction digestion with XbaI, EcoRI, and BamHI restriction enzymes, and the confirmed clone was named pJV7.

The different mutant versions of RLK902 were generated by site-directed mutagenesis of clone pJV3 using primers FP 5'-GACATTGGTGGCTGTGGCGAGACTGAAGGATGTAACG-3' and RP 5'-CGTTACATCCTTCAGTCTCGCCACAGCCACCAATGTC-3' for K393A, FP 5'-CTCACGGAAACGTCGCGTCCTCCAATATCCTC-3' and RP 5'-GAGGATATTGGAGGACGCGACGTTTCCGTGAG-3' for K492A, and FP 5'-GACGCACGAGTGTCTGCTTTCGGCCTGGCTCAGC-3' and RP 5'-GCTGAGCCAGGCCGAAAGCAGACACTCGTGCGTC-3' for D508A. After confirming mutagenesis of these clones by sequencing, the clones were then shuttled into binary vector pRH509. The confirmed clones are named pJV19, pJV13, pJV14 for the K393A, K492A, and D508A mutants of RLK902, respectively. All confirmed binary clones were electrotransformed into *Agrobacterium tumefaciens* strain GV3101 for transient transformation into *N. benthamiana* leaves by agroinfiltration.

Agro-infiltration and expression of RLK902 fusion proteins in N. benthamiana

Agrobacterium tumefaciens strains GV3101 carrying binary vectors pJV7, pJV13, pJV14 and pJV19 containing recombinant (mutant) RLK902 expression cassettes were grown overnight at 28 °C in an LB medium with 50 µg/ml kanamycin and 50 µg/ml rifampicin. The bacterial cells were collected by centrifugation, and resuspended in 10 mM MES pH 5, 10 mM $MgCl_2$ and 0.2 µM acetosyringone to a final OD 600 of 2. Agrobacterium suspensions containing the binary RLK902 expression vectors were mixed with Agrobacterium suspensions containing binary expression vector for RNA silencing inhibitor p19 (Voinnet et. al., 2003) at a 1:1 ratio. The Agrobacterium suspensions were syringe-infiltrated into fully expanded *N. benthamiana* leaves using needleless 1 ml syringes (Plastipak, BD).

Fluorescence microscopy imaging of RLK902

N. benthamiana leaves expressing YFP and/or RFP fusion proteins were subjected to fluorescence microscopy using the Zeiss LSM 700. The highest magnification at 40x oil emersion objective was used. YFP signals were detected at 510/530 excitation/emmision, while RFP signals at 580/610.

Deglycosylation of RLK902 with PNGaseF

The PNGaseF deglycosylation kit NEB P0704S (New England Biolabs) was used for N-deglycosylation of proteins. Protein samples were first extracted in SDS denaturing buffer and then acetone precipitated to clean the samples. Acetone precipitated samples were then dissolved in glycoprotein denaturing buffer and heated to 100°C for 10 min. To 10 µl denatured protein sample, 2 µl water, 1x G7 reaction buffer, 1% NP-40 and 2 µl PNGaseF were added. The sample was then incubated at 37°C for 1 hr. The reaction was stopped by adding SDS gel loading buffer.

REFERENCES

Journal Articles

Adams JA, Taylor SS (1993) Divalent metal ions influence catalysis and active-site accessibility in the cAMP-dependent protein kinase. *Protein science : a publication of the Protein Society* **2:** 2177-2186

Andersson L, Porath J (1986) Isolation of phosphoproteins by immobilized metal (Fe3+) affinity chromatography. *Analytical biochemistry* **154:** 250-254

Anthony RG, Khan S, Costa J, Pais MS, Bogre L (2006) The Arabidopsis protein kinase PTI1-2 is activated by convergent phosphatidic acid and oxidative stress signaling pathways downstream of PDK1 and OXI1. *The Journal of biological chemistry* **281:** 37536-37546

Arribas J, Borroto A (2002) Protein ectodomain shedding. *Chemical reviews* **102:** 4627-4638

Asai T, Tena G, Plotnikova J, Willmann MR, Chiu WL, Gomez-Gomez L, Boller T, Ausubel FM, Sheen J (2002) MAP kinase signalling cascade in Arabidopsis innate immunity. *Nature* **415:** 977-983

Baerenfaller K, Grossmann J, Grobei MA, Hull R, Hirsch-Hoffmann M, Yalovsky S, Zimmermann P, Grossniklaus U, Gruissem W, Baginsky S (2008) Genome-scale proteomics reveals Arabidopsis thaliana gene models and proteome dynamics. *Science* **320:** 938-941

Baerenfaller K, Hirsch-Hoffmann M, Svozil J, Hull R, Russenberger D, Bischof S, Lu Q, Gruissem W, Baginsky S (2011) pep2pro: a new tool for comprehensive proteome data analysis to reveal information about organ-specific proteomes in Arabidopsis thaliana. *Integrative biology : quantitative biosciences from nano to macro* **3:** 225-237

Banbury DN, Oakley JD, Sessions RB, Banting G (2003) Tyrphostin A23 inhibits internalization of the transferrin receptor by perturbing the interaction between tyrosine motifs and the medium chain subunit of the AP-2 adaptor complex. *The Journal of biological chemistry* **278:** 12022-12028

Beck M, Zhou J, Faulkner C, MacLean D, Robatzek S (2012) Spatio-temporal cellular dynamics of the Arabidopsis flagellin receptor reveal activation status-dependent endosomal sorting. *The Plant cell* **24:** 4205-4219

Beckers GJ, Jaskiewicz M, Liu Y, Underwood WR, He SY, Zhang S, Conrath U (2009) Mitogen-activated protein kinases 3 and 6 are required for full priming of stress responses in Arabidopsis thaliana. *The Plant cell* **21:** 944-953

Bethke G, Pecher P, Eschen-Lippold L, Tsuda K, Katagiri F, Glazebrook J, Scheel D, Lee J (2012) Activation of the Arabidopsis thaliana mitogen-activated protein kinase MPK11 by the flagellin-derived elicitor peptide, flg22. *Molecular plant-microbe interactions : MPMI* **25:** 471-480

Boudsocq M, Willmann MR, McCormack M, Lee H, Shan L, He P, Bush J, Cheng SH, Sheen J (2010) Differential innate immune signalling via Ca(2+) sensor protein kinases. *Nature* **464:** 418-422

Bozkurt TO, Richardson A, Dagdas YF, Mongrand S, Kamoun S, Raffaele S (2014) The Plant Membrane-Associated REMORIN1.3 Accumulates in Discrete Perihaustorial Domains and Enhances Susceptibility to Phytophthora infestans. *Plant physiology* **165:** 1005-1018

Carrera AC, Alexandrov K, Roberts TM (1993) The conserved lysine of the catalytic domain of protein kinases is actively involved in the phosphotransfer reaction and not required for anchoring ATP. *Proceedings of the National Academy of Sciences of the United States of America* **90:** 442-446

Chang C, Schaller GE, Patterson SE, Kwok SF, Meyerowitz EM, Bleecker AB (1992) The TMK1 gene from Arabidopsis codes for a protein with structural and biochemical characteristics of a receptor protein kinase. *The Plant cell* **4:** 1263-1271

Chauhan JS, Mishra NK, Raghava GP (2009) Identification of ATP binding residues of a protein from its primary sequence. *BMC bioinformatics* **10:** 434

Cheng SH, Willmann MR, Chen HC, Sheen J (2002) Calcium signaling through protein kinases. The Arabidopsis calcium-dependent protein kinase gene family. *Plant physiology* **129:** 469-485

Cho SK, Larue CT, Chevalier D, Wang H, Jinn TL, Zhang S, Walker JC (2008) Regulation of floral organ abscission in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 15629-15634

Chow FL, Fernandez-Patron C (2007) Many membrane proteins undergo ectodomain shedding by proteolytic cleavage. Does one sheddase do the job on all of these proteins? *IUBMB life* **59:** 44-47

Cooper MJ, Cox NJ, Zimmerman EI, Dewar BJ, Duncan JS, Whittle MC, Nguyen TA, Jones LS, Ghose Roy S, Smalley DM, Kuan PF, Richards KL, Christopherson RI, Jin J, Frye SV, Johnson GL, Baldwin AS, Graves LM (2013) Application of multiplexed kinase inhibitor beads to study kinome adaptations in drug-resistant leukemia. *PloS one* **8:** e66755

Cravatt BF, Wright AT, Kozarich JW (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annual review of biochemistry* **77:** 383-414

Dardick C, Ronald P (2006) Plant and animal pathogen recognition receptors signal through non-RD kinases. *PLoS pathogens* **2:** e2

Davis MT, Spahr CS, McGinley MD, Robinson JH, Bures EJ, Beierle J, Mort J, Yu W, Luethy R, Patterson SD (2001) Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry. II. Limitations of complex mixture analyses. *Proteomics* **1:** 108-117

Deng X, Dzamko N, Prescott A, Davies P, Liu Q, Yang Q, Lee JD, Patricelli MP, Nomanbhoy TK, Alessi DR, Gray NS (2011) Characterization of a selective inhibitor of the Parkinson's disease kinase LRRK2. *Nature chemical biology* **7:** 203-205

DeYoung BJ, Qi D, Kim SH, Burke TP, Innes RW (2012) Activation of a plant nucleotide binding-leucine rich repeat disease resistance protein by a modified self protein. *Cellular microbiology* **14:** 1071-1084

Dghim AA, Mhamdi A, Vaultier MN, Hasenfratz-Sauder MP, Le Thiec D, Dizengremel P, Noctor G, Jolivet Y (2013) Analysis of cytosolic isocitrate dehydrogenase and glutathione reductase 1 in photoperiod-influenced responses to ozone using Arabidopsis knockout mutants. *Plant, cell & environment* **36:** 1981-1991

Di Russo NV, Estrin DA, Marti MA, Roitberg AE (2012) pH-Dependent conformational changes in proteins and their effect on experimental pK(a)s: the case of Nitrophorin 4. *PLoS computational biology* **8:** e1002761

Duncan JS, Whittle MC, Nakamura K, Abell AN, Midland AA, Zawistowski JS, Johnson NL, Granger DA, Jordan NV, Darr DB, Usary J, Kuan PF, Smalley DM, Major B, He X, Hoadley KA, Zhou B, Sharpless NE, Perou CM, Kim WY, Gomez SM, Chen X, Jin J, Frye SV, Earp HS, Graves LM, Johnson GL (2012) Dynamic reprogramming of the kinome in response to targeted MEK inhibition in triple-negative breast cancer. *Cell* **149:** 307-321

Eathiraj S, Palma R, Hirschi M, Volckova E, Nakuci E, Castro J, Chen CR, Chan TC, France DS, Ashwell MA (2011) A novel mode of protein kinase inhibition exploiting hydrophobic motifs of autoinhibited kinases: discovery of ATP-independent inhibitors of fibroblast growth factor receptor. *The Journal of biological chemistry* **286:** 20677-20687

Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K, Holm L, Sonnhammer EL, Eddy SR, Bateman A (2010) The Pfam protein families database. *Nucleic acids research* **38:** D211-222

Galletti R, Ferrari S, De Lorenzo G (2011) Arabidopsis MPK3 and MPK6 play different roles in basal and oligogalacturonide- or flagellin-induced resistance against Botrytis cinerea. *Plant physiology* **157:** 804-814

Galperin MY, Koonin EV (1997) A diverse superfamily of enzymes with ATP-dependent carboxylate-amine/thiol ligase activity. *Protein science : a publication of the Protein Society* **6:** 2639-2643

Gao L, Xiang CB (2008) The genetic locus At1g73660 encodes a putative MAPKKK and negatively regulates salt tolerance in Arabidopsis. *Plant molecular biology* **67:** 125-134

Gao M, Liu J, Bi D, Zhang Z, Cheng F, Chen S, Zhang Y (2008) MEKK1, MKK1/MKK2 and MPK4 function together in a mitogen-activated protein kinase cascade to regulate innate immunity in plants. *Cell research* **18:** 1190-1198

Geldner N, Robatzek S (2008) Plant receptors go endosomal: a moving view on signal transduction. *Plant physiology* **147:** 1565-1574

Goh LK, Sorkin A (2013) Endocytosis of receptor tyrosine kinases. *Cold Spring Harbor perspectives in biology* **5:** a017459

Gonzalez DA, Ostuni MA, Lacapere JJ, Alonso GL (2006) Stoichiometry of ATP and metal cofactor interaction with the sarcoplasmic reticulum Ca(2+)-ATPase: a binding model accounting for radioisotopic and fluorescence results. *Biophysical chemistry* **124:** 27-34

Graves LM, Duncan JS, Whittle MC, Johnson GL (2013) The dynamic nature of the kinome. *The Biochemical journal* **450:** 1-8

Haffani YZ, Silva-Gagliardi NF, Sewter SK, Grace Aldea M, Zhao Z, Nakhamchik A, Cameron RK, Goring DR (2006) Altered Expression of PERK Receptor Kinases in Arabidopsis Leads to Changes in Growth and Floral Organ Formation. *Plant signaling & behavior* **1:** 251-260

Hanks SK, Hunter T (1995) Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **9:** 576-596

Heldin CH, Ericsson J (2001) Signal transduction. RIPping tyrosine kinase receptors apart. *Science* **294:** 2111-2113

Hicks GR, Smith HM, Shieh M, Raikhel NV (1995) Three classes of nuclear import signals bind to plant nuclei. *Plant physiology* **107:** 1055-1058

Holmes-Davis R, Tanaka CK, Vensel WH, Hurkman WJ, McCormick S (2005) Proteome mapping of mature pollen of Arabidopsis thaliana. *Proteomics* **5:** 4864-4884

Horiuchi K (2013) A brief history of tumor necrosis factor alpha--converting enzyme: an overview of ectodomain shedding. *The Keio journal of medicine* **62:** 29-36

Huse M, Kuriyan J (2002) The conformational plasticity of protein kinases. *Cell* **109:** 275-282

Ibba M, Becker HD, Stathopoulos C, Tumbula DL, Soll D (2000) The adaptor hypothesis revisited. *Trends in biochemical sciences* **25:** 311-316

Johnsen IB, Nguyen TT, Ringdal M, Tryggestad AM, Bakke O, Lien E, Espevik T, Anthonsen MW (2006) Toll-like receptor 3 associates with c-Src tyrosine kinase on endosomes to initiate antiviral signaling. *The EMBO journal* **25:** 3335-3346

Johnson LN, Noble ME, Owen DJ (1996) Active and inactive protein kinases: structural basis for regulation. *Cell* **85:** 149-158

Kajava AV (1998) Structural diversity of leucine-rich repeat proteins. *Journal of molecular biology* **277:** 519-527

Kamps MP, Sefton BM (1986) Neither arginine nor histidine can carry out the function of lysine-295 in the ATP-binding site of p60src. *Molecular and cellular biology* **6:** 751-757

Klausner RD, Donaldson JG, Lippincott-Schwartz J (1992) Brefeldin A: insights into the control of membrane traffic and organelle structure. *The Journal of cell biology* **116:** 1071-1080

Knight ZA, Shokat KM (2005) Features of selective kinase inhibitors. *Chemistry & biology* **12:** 621-637

Knighton DR, Zheng JH, Ten Eyck LF, Ashford VA, Xuong NH, Taylor SS, Sowadski JM (1991) Crystal structure of the catalytic subunit of cyclic adenosine monophosphate-dependent protein kinase. *Science* **253:** 407-414

Kornev AP, Haste NM, Taylor SS, Eyck LF (2006) Surface comparison of active and inactive protein kinases identifies a conserved activation mechanism. *Proceedings of the National Academy of Sciences of the United States of America* **103:** 17783-17788

Kosugi S, Hasebe M, Matsumura N, Takashima H, Miyamoto-Sato E, Tomita M, Yanagawa H (2009) Six classes of nuclear localization signals specific to different binding grooves of importin alpha. *The Journal of biological chemistry* **284:** 478-485

Kozik P, Francis RW, Seaman MN, Robinson MS (2010) A screen for endocytic motifs. *Traffic* **11:** 843-855

Lachaud C, Prigent E, Thuleau P, Grat S, Da Silva D, Briere C, Mazars C, Cotelle V (2013) 14-3-3-regulated Ca(2+)-dependent protein kinase CPK3 is required for sphingolipid-induced cell death in Arabidopsis. *Cell death and differentiation* **20:** 209-217

Lamaze C, Fujimoto LM, Yin HL, Schmid SL (1997) The actin cytoskeleton is required for receptor-mediated endocytosis in mammalian cells. *The Journal of biological chemistry* **272:** 20332-20335

Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, Muller R, Dreher K, Alexander DL, Garcia-Hernandez M, Karthikeyan AS, Lee CH, Nelson WD, Ploetz L, Singh S, Wensel A, Huala E (2012) The Arabidopsis Information

Resource (TAIR): improved gene annotation and new tools. *Nucleic acids research* **40:** D1202-1210

Lange A, Mills RE, Lange CJ, Stewart M, Devine SE, Corbett AH (2007) Classical nuclear localization signals: definition, function, and interaction with importin alpha. *The Journal of biological chemistry* **282:** 5101-5105

Lawrence DS (2001) Functional proteomics: large-scale analysis of protein kinase activity. *Genome biology* **2:** REVIEWS1007

Lawton KA, Friedrich L, Hunt M, Weymann K, Delaney T, Kessmann H, Staub T, Ryals J (1996) Benzothiadiazole induces disease resistance in Arabidopsis by activation of the systemic acquired resistance signal transduction pathway. *The Plant journal : for cell and molecular biology* **10:** 71-82

Leung P, Pickarski M, Zhuo Y, Masarachia PJ, Duong LT (2011) The effects of the cathepsin K inhibitor odanacatib on osteoclastic bone resorption and vesicular trafficking. *Bone* **49:** 623-635

Lew J (2003) MAP kinases and CDKs: kinetic basis for catalytic activation. *Biochemistry* **42:** 849-856

Li G, Meng X, Wang R, Mao G, Han L, Liu Y, Zhang S (2012) Dual-level regulation of ACC synthase activity by MPK3/MPK6 cascade and its downstream WRKY transcription factor during ethylene induction in Arabidopsis. *PLoS genetics* **8:** e1002767

Liljelund P, Lacroute F (1986) Genetic characterization and isolation of the Saccharomyces cerevisiae gene coding for uridine monophosphokinase. *Molecular & general genetics : MGG* **205:** 74-81

Liu Q, Kirubakaran S, Hur W, Niepel M, Westover K, Thoreen CC, Wang J, Ni J, Patricelli MP, Vogel K, Riddle S, Waller DL, Traynor R, Sanda T, Zhao Z, Kang SA, Zhao J, Look AT, Sorger PK, Sabatini DM, Gray NS (2012) Kinome-wide selectivity profiling of ATP-competitive mammalian target of rapamycin (mTOR) inhibitors and characterization of their binding kinetics. *The Journal of biological chemistry* **287:** 9742-9752

Liu Y, Gray NS (2006) Rational design of inhibitors that bind to inactive kinase conformations. *Nature chemical biology* **2:** 358-364

Maley F, Trimble RB, Tarentino AL, Plummer TH, Jr. (1989) Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Analytical biochemistry* **180:** 195-204

Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S (2002) The protein kinase complement of the human genome. *Science* **298:** 1912-1934

Matsumura H, Mizohata E, Ishida H, Kogami A, Ueno T, Makino A, Inoue T, Yokota A, Mae T, Kai Y (2012) Crystal structure of rice Rubisco and implications for

activation induced by positive effectors NADPH and 6-phosphogluconate. *Journal of molecular biology* **422:** 75-86

Matsuoka D, Nanmori T, Sato K, Fukami Y, Kikkawa U, Yasuda T (2002) Activation of AtMEK1, an Arabidopsis mitogen-activated protein kinase kinase, in vitro and in vivo: analysis of active mutants expressed in E. coli and generation of the active form in stress response in seedlings. *The Plant journal : for cell and molecular biology* **29:** 637-647

Mehlmer N, Wurzinger B, Stael S, Hofmann-Rodrigues D, Csaszar E, Pfister B, Bayer R, Teige M (2010) The Ca(2+) -dependent protein kinase CPK3 is required for MAPK-independent salt-stress acclimation in Arabidopsis. *The Plant journal : for cell and molecular biology* **63:** 484-498

Melville DB (1954) Biotin sulfoxide. *The Journal of biological chemistry* **208:** 495-501

Mori IC, Murata Y, Yang Y, Munemasa S, Wang YF, Andreoli S, Tiriac H, Alonso JM, Harper JF, Ecker JR, Kwak JM, Schroeder JI (2006) CDPKs CPK6 and CPK3 function in ABA regulation of guard cell S-type anion- and Ca(2+)-permeable channels and stomatal closure. *PLoS biology* **4:** e327

Moriyasu Y, Ohsumi Y (1996) Autophagy in Tobacco Suspension-Cultured Cells in Response to Sucrose Starvation. *Plant physiology* **111:** 1233-1241

Moscatelli A, Idilli AI, Rodighiero S, Caccianiga M (2012) Inhibition of actin polymerisation by low concentration Latrunculin B affects endocytosis and alters exocytosis in shank and tip of tobacco pollen tubes. *Plant biology*

Muroi M, Shiragami N, Nagao K, Yamasaki M, Takatsuki A (1993) Folimycin (concanamycin A), a specific inhibitor of V-ATPase, blocks intracellular translocation of the glycoprotein of vesicular stomatitis virus before arrival to the Golgi apparatus. *Cell structure and function* **18:** 139-149

Newman MA, Sundelin T, Nielsen JT, Erbs G (2013) MAMP (microbe-associated molecular pattern) triggered immunity in plants. *Frontiers in plant science* **4:** 139

Ni CY, Murphy MP, Golde TE, Carpenter G (2001) gamma -Secretase cleavage and nuclear localization of ErbB-4 receptor tyrosine kinase. *Science* **294:** 2179-2181

Nita-Lazar A, Saito-Benz H, White FM (2008) Quantitative phosphoproteomics by mass spectrometry: past, present, and future. *Proteomics* **8:** 4433-4443

Nuhse TS, Stensballe A, Jensen ON, Peck SC (2003) Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. *Molecular & cellular proteomics : MCP* **2:** 1234-1243

Nuhse TS, Stensballe A, Jensen ON, Peck SC (2004) Phosphoproteomics of the Arabidopsis plasma membrane and a new phosphorylation site database. *The Plant cell* **16:** 2394-2405

Ogura T, Wilkinson AJ (2001) AAA+ superfamily ATPases: common structure--diverse function. *Genes to cells : devoted to molecular & cellular mechanisms* **6:** 575-597

Okerberg ES, Wu J, Zhang B, Samii B, Blackford K, Winn DT, Shreder KR, Burbaum JJ, Patricelli MP (2005) High-resolution functional proteomics by active-site peptide profiling. *Proceedings of the National Academy of Sciences of the United States of America* **102:** 4996-5001

Palokangas H, Metsikko K, Vaananen K (1994) Active vacuolar H+ATPase is required for both endocytic and exocytic processes during viral infection of BHK-21 cells. *The Journal of biological chemistry* **269:** 17577-17585

Park CJ, Ronald PC (2012) Cleavage and nuclear localization of the rice XA21 immune receptor. *Nature communications* **3:** 920

Park S, Rancour DM, Bednarek SY (2008) In planta analysis of the cell cycle-dependent localization of AtCDC48A and its critical roles in cell division, expansion, and differentiation. *Plant physiology* **148:** 246-258

Patricelli MP, Nomanbhoy TK, Wu J, Brown H, Zhou D, Zhang J, Jagannathan S, Aban A, Okerberg E, Herring C, Nordin B, Weissig H, Yang Q, Lee JD, Gray NS, Kozarich JW (2011) In situ kinase profiling reveals functionally relevant properties of native kinases. *Chemistry & biology* **18:** 699-710

Patricelli MP, Szardenings AK, Liyanage M, Nomanbhoy TK, Wu M, Weissig H, Aban A, Chun D, Tanner S, Kozarich JW (2007) Functional interrogation of the kinome using nucleotide acyl phosphates. *Biochemistry* **46:** 350-358

Picotti P, Aebersold R, Domon B (2007) The implications of proteolytic background for shotgun proteomics. *Molecular & cellular proteomics : MCP* **6:** 1589-1598

Plaxton WC (1996) The Organization and Regulation of Plant Glycolysis. *Annual review of plant physiology and plant molecular biology* **47:** 185-214

Pujol C, Bailly M, Kern D, Marechal-Drouard L, Becker H, Duchene AM (2008) Dual-targeted tRNA-dependent amidotransferase ensures both mitochondrial and chloroplastic Gln-tRNAGln synthesis in plants. *Proceedings of the National Academy of Sciences of the United States of America* **105:** 6481-6485

Rasmussen MW, Roux M, Petersen M, Mundy J (2012) MAP Kinase Cascades in Arabidopsis Innate Immunity. *Frontiers in plant science* **3:** 169

Richards DA, Rizzoli SO, Betz WJ (2004) Effects of wortmannin and latrunculin A on slow endocytosis at the frog neuromuscular junction. *The Journal of physiology* **557:** 77-91

Robatzek S (2007) Vesicle trafficking in plant immune responses. *Cellular microbiology* **9:** 1-8

Robatzek S, Chinchilla D, Boller T (2006) Ligand-induced endocytosis of the pattern recognition receptor FLS2 in Arabidopsis. *Genes & development* **20:** 537-542

Saylor P, Wang C, Hirai TJ, Adams JA (1998) A second magnesium ion is critical for ATP binding in the kinase domain of the oncoprotein v-Fps. *Biochemistry* **37:** 12624-12630

Schaller GE, Bleecker AB (1993) Receptor-like kinase activity in membranes of Arabidopsis thaliana. *FEBS letters* **333:** 306-310

Schiess R, Wollscheid B, Aebersold R (2009) Targeted proteomic strategy for clinical biomarker discovery. *Molecular oncology* **3:** 33-44

Schmidt A, Claassen M, Aebersold R (2009) Directed mass spectrometry: towards hypothesis-driven proteomics. *Current opinion in chemical biology* **13:** 510-517

Schulze B, Mentzel T, Jehle AK, Mueller K, Beeler S, Boller T, Felix G, Chinchilla D (2010) Rapid heteromerization and phosphorylation of ligand-activated plant transmembrane receptors and their associated kinase BAK1. *The Journal of biological chemistry* **285:** 9444-9451

Senecoff JF, McKinney EC, Meagher RB (1996) De novo purine synthesis in Arabidopsis thaliana. II. The PUR7 gene encoding 5'-phosphoribosyl-4-(N-succinocarboxamide)-5-aminoimidazole synthetase is expressed in rapidly dividing tissues. *Plant physiology* **112:** 905-917

Serrano M, Robatzek S, Torres M, Kombrink E, Somssich IE, Robinson M, Schulze-Lefert P (2007) Chemical interference of pathogen-associated molecular pattern-triggered immune responses in Arabidopsis reveals a potential role for fatty-acid synthase type II complex-derived lipid signals. *The Journal of biological chemistry* **282:** 6803-6811

Shao F, Golstein C, Ade J, Stoutemyer M, Dixon JE, Innes RW (2003) Cleavage of Arabidopsis PBS1 by a bacterial type III effector. *Science* **301:** 1230-1233

Shek PY, Zhao J, Ke Y, Siu KW, Hopkinson AC (2006) Fragmentations of protonated arginine, lysine and their methylated derivatives: concomitant losses of carbon monoxide or carbon dioxide and an amine. *The journal of physical chemistry A* **110:** 8282-8296

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

Silva NF, Goring DR (2002) The proline-rich, extensin-like receptor kinase-1 (PERK1) gene is rapidly induced by wounding. *Plant molecular biology* **50:** 667-685

Sorensson C, Lenman M, Veide-Vilg J, Schopper S, Ljungdahl T, Grotli M, Tamas MJ, Peck SC, Andreasson E (2012) Determination of primary sequence specificity of

Arabidopsis MAPKs MPK3 and MPK6 leads to identification of new substrates. *The Biochemical journal* **446:** 271-278

Srivastava J, Barber DL, Jacobson MP (2007) Intracellular pH sensors: design principles and functional significance. *Physiology* **22:** 30-39

Sun J, Ke J, Johnson JL, Nikolau BJ, Wurtele ES (1997) Biochemical and molecular biological characterization of CAC2, the Arabidopsis thaliana gene coding for the biotin carboxylase subunit of the plastidic acetyl-coenzyme A carboxylase. *Plant physiology* **115:** 1371-1383

Swiderski MR, Innes RW (2001) The Arabidopsis PBS1 resistance gene encodes a member of a novel protein kinase subfamily. *The Plant journal : for cell and molecular biology* **26:** 101-112

Tarutani Y, Morimoto T, Sasaki A, Yasuda M, Nakashita H, Yoshida S, Yamaguchi I, Suzuki Y (2004a) Molecular characterization of two highly homologous receptor-like kinase genes, RLK902 and RKL1, in Arabidopsis thaliana. *Bioscience, biotechnology, and biochemistry* **68:** 1935-1941

Tarutani Y, Sasaki A, Yasuda M, Nakashita H, Yoshida S, Yamaguchi I, Suzuki Y (2004b) Identification of three clones which commonly interact with the kinase domains of highly homologous two receptor-like kinases, RLK902 and RKL1. *Bioscience, biotechnology, and biochemistry* **68:** 2581-2587

Teis D, Wunderlich W, Huber LA (2002) Localization of the MP1-MAPK scaffold complex to endosomes is mediated by p14 and required for signal transduction. *Developmental cell* **3:** 803-814

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature 408: 796-815

Ulm R, Ichimura K, Mizoguchi T, Peck SC, Zhu T, Wang X, Shinozaki K, Paszkowski J (2002) Distinct regulation of salinity and genotoxic stress responses by Arabidopsis MAP kinase phosphatase 1. *The EMBO journal* **21:** 6483-6493

Upadhya K, Khattak IK, Mullah B (2005) Oxidation of biotin during oligonucleotide synthesis. *Nucleosides, nucleotides & nucleic acids* **24:** 919-922

van der Hoorn RA, Colby T, Nickel S, Richau KH, Schmidt J, Kaiser M (2011) Mining the Active Proteome of Arabidopsis thaliana. *Frontiers in plant science* **2:** 89

Vetter ML, Rodgers MA, Patricelli MP, Yang PL (2012) Chemoproteomic profiling identifies changes in DNA-PK as markers of early dengue virus infection. *ACS chemical biology* **7:** 2019-2026

Wang N, Li L (2008) Exploring the precursor ion exclusion feature of liquid chromatography-electrospray ionization quadrupole time-of-flight mass spectrometry for improving protein identification in shotgun proteome analysis. *Analytical chemistry* **80:** 4696-4710

Wang X, Herr RA, Chua WJ, Lybarger L, Wiertz EJ, Hansen TH (2007) Ubiquitination of serine, threonine, or lysine residues on the cytoplasmic tail can induce ERAD of MHC-I by viral E3 ligase mK3. *The Journal of cell biology* **177:** 613-624

Weber S, Saftig P (2012) Ectodomain shedding and ADAMs in development. *Development* **139:** 3693-3709

Wingenter K, Trentmann O, Winschuh I, Hormiller, II, Heyer AG, Reinders J, Schulz A, Geiger D, Hedrich R, Neuhaus HE (2011) A member of the mitogen-activated protein 3-kinase family is involved in the regulation of plant vacuolar glucose uptake. *The Plant journal : for cell and molecular biology* **68:** 890-900

Wodicka LM, Ciceri P, Davis MI, Hunt JP, Floyd M, Salerno S, Hua XH, Ford JM, Armstrong RC, Zarrinkar PP, Treiber DK (2010) Activation state-dependent binding of small molecule kinase inhibitors: structural insights from biochemistry. *Chemistry & biology* **17:** 1241-1249

Xing W, Zou Y, Liu Q, Liu J, Luo X, Huang Q, Chen S, Zhu L, Bi R, Hao Q, Wu JW, Zhou JM, Chai J (2007) The structural basis for activation of plant immunity by bacterial effector protein AvrPto. *Nature* **449:** 243-247

Xu P, Derynck R (2010) Direct Activation of TACE-Mediated Ectodomain Shedding by p38 MAP Kinase Regulates EGF Receptor-Dependent Cell Proliferation. *Molecular Cell* **37:** 551-566

Yamanaka K, Ogura T, Niki H, Hiraga S (1992) Identification and characterization of the smbA gene, a suppressor of the mukB null mutant of Escherichia coli. *Journal of bacteriology* **174:** 7517-7526

Yang KY, Liu Y, Zhang S (2001) Activation of a mitogen-activated protein kinase pathway is involved in disease resistance in tobacco. *Proceedings of the National Academy of Sciences of the United States of America* **98:** 741-746

Yang Q, Deng X, Lu B, Cameron M, Fearns C, Patricelli MP, Yates JR, 3rd, Gray NS, Lee JD (2010) Pharmacological inhibition of BMK1 suppresses tumor growth through promyelocytic leukemia protein. *Cancer cell* **18:** 258-267

Yano K, Matsui S, Tsuchiya T, Maeshima M, Kutsuna N, Hasezawa S, Moriyasu Y (2004) Contribution of the plasma membrane and central vacuole in the formation of autolysosomes in cultured tobacco cells. *Plant & cell physiology* **45:** 951-957

Zhang S, Liu Y (2001) Activation of salicylic acid-induced protein kinase, a mitogen-activated protein kinase, induces multiple defense responses in tobacco. *The Plant cell* **13:** 1877-1889

Zhang T, Inesta-Vaquera F, Niepel M, Zhang J, Ficarro SB, Machleidt T, Xie T, Marto JA, Kim N, Sim T, Laughlin JD, Park H, LoGrasso PV, Patricelli M,

Nomanbhoy TK, Sorger PK, Alessi DR, Gray NS (2012) Discovery of potent and selective covalent inhibitors of JNK. *Chemistry & biology* **19:** 140-154

Zheng J, Knighton DR, ten Eyck LF, Karlsson R, Xuong N, Taylor SS, Sowadski JM (1993) Crystal structure of the catalytic subunit of cAMP-dependent protein kinase complexed with MgATP and peptide inhibitor. *Biochemistry* **32:** 2154-2161

Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, Klemic KG, Smith D, Gerstein M, Reed MA, Snyder M (2000) Analysis of yeast protein kinases using protein chips. *Nature genetics* **26:** 283-289

Online Sources

The Arabidopsis kinase database, Available:
http://bioinformatics.cau.edu.cn/athKD/index.htm

R.D. Finn, A. Bateman, J. Clements, P. Coggill, R.Y. Eberhardt, S.R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E.L.L. Sonnhammer, J. Tate, M. Punta (2014) *Nucleic Acids Research* Database **42:** D222-D230.
The Pfam protein families database, Available: http://pfam.sanger.ac.uk/

BRENDA, The comprehensive enzyme information center. Available:
http://www.brenda-enzymes.info/index.php4

Baerenfaller K, Grossmann J, Grobei MA, Hull R, Hirsch-Hoffmann M, Yalovsky S, Zimmermann P, Grossniklaus U, Gruissem W, Baginsky S (2008) Genome-scale proteomics reveals Arabidopsis thaliana gene models and proteome dynamics. *Science* **320:** 938-941
AtProteome database, Avaliable: http://fgcz-atproteome.unizh.ch/

iTAK - Plant Transcription factor & Protein Kinase Identifier and Classifier, http://bioinfo.bti.cornell.edu/cgi-bin/itak/index.cgi)

Interpro electronic annotation, Available: http://www.ebi.ac.uk/interpro/

Nuhse TS, Stensballe A, Jensen ON, Peck SC (2004) Phosphoproteomics of the Arabidopsis plasma membrane and a new phosphorylation site database. *The Plant cell* **16:** 2394-2405
The Arabidopsis Protein Phosphorylation Site Database, Available:
http://phosphat.uni-hohenheim.de/

Corpet F (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Research* **16:** 10881-10890
MultiAlign, Available: http://multalin.toulouse.inra.fr/multalin/multalin.html

Books

Schultz SG, Andreoli TE, Brown AM, Fambrough DM, Hoffman JF, Welsh MJ. (1996) Molecular biology of membrane transport disorders, Springer

Aylward GH, Tristan JV, Findlay (2002) SI Chemical Data, Wiley, 5th edition

Stryer L, Berg JM, Tymoczko JL (2007) Biochemistry, San Francisco: WH Freeman

LEGENDS FOR SUPPLEMENTAL TABLES

All supplemental tables are available as soft copies in CDs attached to this manuscript.

Supplemental Table S1: Identified Arabidopsis proteins using the in-gel digest procedure. This report summarizes all Arabidopsis proteins identified with at least two different peptides with Mascot scores >41. The data is ordered per spot ranking the proteins with highest scores first.

Supplemental Table S2: Spectral counts per protein and per band from in-gel digests. ***Sheet 1*** summarizes all spectral counts of the 112 proteins detected by in-gel digest in each of the 30 spots of Table S1, ranked on accession code. Protein kinases are marked red. Proteins with spectral counts indicated with light green boxes were selected for the summary of sheet 2. ***Sheet 2*** summarizes a the spectral counts per band for the selection of the 38 proteins found by in-gel digest, ranked on their occurrence in the gel and the frequencies of the spectral counts. These proteins were selected from sheet 1 using the following criteria: For each band, proteins having the highest spectral counts were selected. Proteins occurring in more than one band in the same lane were only considered for the band where they have the highest count. The list is supplemented with protein kinases and with proteins for which labeling sites were detected. Protein kinases are marked red. Spectral counts indicated with light green boxes are shown in Fig. 10B.

Supplemental Table S3: Detected peptides using gel-free analysis. This table summarizes all the detected labelled peptides of each of the four LC-MS/MS runs, ranked on the protein accession code. Peptides with probability scores <95% are printed in grey. The spectra have later been counted per peptide with probability scores >95% and detected in at least two of the four runs in Table S4.

Supplemental Table S4: Spectral counts of the labelled peptides using gel-free approach. This table summarizes the spectral counts per labelled peptide in each of the four gel-free runs, ranked on the spectral count frequency. Only peptides with a probability score of >95% and detected in two or more of the four runs are shown.

This table also tells if the identified peptide is unique in the Arabidopsis proteome, or common to different isoforms of the same gene, or ambiguous. Two more columns were added summarizing the PDB structure that was used to identify the labeling site in the 3D structure of a homologous protein, and the measured distance between the phosphate of the bound nucleotide and the labelled lysine in Angstrom (Å). NR, no nucleotide binding site reported; NS, no 3D structure available.

Supplemental Table S5: Targeted MS/MS of AcATP-labelled Arabidopsis cell cultures. This table summarizes the signal intensities of AcATP-labelled peptides from control and fgl22-treated Arabidopsis cell cultures. Fold change indicates change in MS signal for the treated samples relative to the MS signal for the control samples.

ERKLÄRUNG

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von untenngegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. Jane Parker betreut worden.

Max Planck Institut für Pflanzenzüchtungsforschung
Carl-von-Linne Weg 10, D-50829 Köln
04 November 2014

Joji Grace Villamor-Sernestrand

Teilpublikationen:

Villamor JG, Kaschani F, Colby T, Oeljeklaus J, Zhao D, Kaiser M, Patricelli MP, and van der Hoorn RA (2013) Profiling protein kinases and other ATP-binding proteins in Arabidopsis using acyl-ATP probes. Mol Cell Proteomics. 12. 2481-2496

ACKNOWLEDGMENT

Doing a PhD is hard. It is full of challenges, heartaches, frustrations, sleepless nights, missed lunches, and a whole lot of life changes, because one has to leave home and venture into an unknown territory all for the sake of science. Hence, doing a PhD takes a whole lot of help, from a whole lot of people. To them, I say my sincere "Thank you".

Thank you, Renier, first of all for believing in me when you took me as your student. Thank you for your supervision throughout my studies. Thank you for your invaluable comments and suggestions. Thank you for your encouragements and constant reassurance when frustrations attack. It is from you that I learn the most how to do science. And I am really grateful that even from a distance starting this year, you have helped me a lot during the writing of this manuscript.

Thank you Dr. Jane Parker for your supervision and guidance. Thank you for sparing time to look at my work and for giving suggestions on 'how to proceed' during my yearly PhD evaluation, together with Dr. Olof Persson. Thank you too for heading my PhD thesis committee.

Thank you Dr. Erik Andreasson, and Dr. Ulf-Ingo-Flügge for being part of my thesis evaluation committee. Thank you Dr. Erich Kombrinck for standing as the beisitzer in my thesis defense and thank you for your comments and suggestions during PSL meetings and joint meetings in Essen. Finally, a big thank you Prof. Dr. Maria Albani for chairing my PhD thesis committee.

Thank you Plant Chemetics group. You have given life to an otherwise (sometimes) tedious and exhausting daily lab works. It is mostly because of you that I get up every morning to go to the institute, where I can talk to you, work with you, laugh with you, drink coffee with you, eat lunch with you, pour out frustrations with you, and rejoice in celebration with you when good results grace us; as well as prepare buffers with you, run gels with you, 'waste' lab consumables with you, do science with you. Thank you Takayuki, Christian, Johanna, Ilyas, Selva, Rohini, Oscar, Anja, Ali, Haibin, Tram, Gerry, Bala, Matthias, Asif and Aranka. A big thank you to Farnusch, who spearheaded the kinase profiling project, and who have helped me get started with my PhD.

And if there is anything of consolation in doing a PhD, it is the people you meet, the places you visit, and the whole lot of learning you experience, both in science and in life. If there´s anything I gained from doing a PhD, it is the people, who once were strangers to me that became a family.

The people of the plant chemetics lab became my first family in Cologne, Germany. I am so grateful that I was given the best lab group, and the best teacher one can possibly have. Thank you, that besides the exciting science we did together, you have shared with me the gift of friendship and love. Thank you for the sweet memories I will forever cherish. Thank you for all that you have taught me, workwise, and not-so-workwise.

The people of the International Baptist Church Cologne, as well as the people of The Connection in Bruehl, is the family I now have in Cologne. You all are such a wonderful blessing to me.

Thank you so much for the friendship, love, and prayers. Thank you for the encouragements and support you have showered me. Thank you for everything dear brothers and sisters. Special mention to Jane, Rita and Pontus, who cooked food for me during my writing days; you fueled me up with ATP, so I can have the energy to continue with the things I need to do. Thank you too Jane, Rita and Maimouna for babysitting my Josef so I can unwind and recharge. A big thank you to Blake, Ailene, Susan and Steven, for all the home cooked meals during the sleepless and tiresome "thesis baby" days.

And most importantly, Pontus, my beloved, thank you so much for all that you do. Thank you for the big leap of faith you took so we can be together, so we can be a family. Thank you for your love and great support for the things that I do. Thank you for your encouragements. You inspire me to be the best I can be. And if I were to do all these over again, I won't mind at all, because otherwise I would not have met you.

And to you my little Josef Nelson, my little baby boy, who I first saw and held last May 1st, thank you so much. You are such a big miracle, such a wonderful blessing, and such a joy to Mama and Papa. Did you know that even without you trying, you inspire Mama to soar higher, to be the best. You motivate Mama so she can finish this (PhD) test.

Thank you too, my beautiful family in Sweden. Thank you for welcoming me so warmly in to your life. Thank you for your kindness, love and support. And most importantly, thank you for raising Pontus to be the man that he is today.

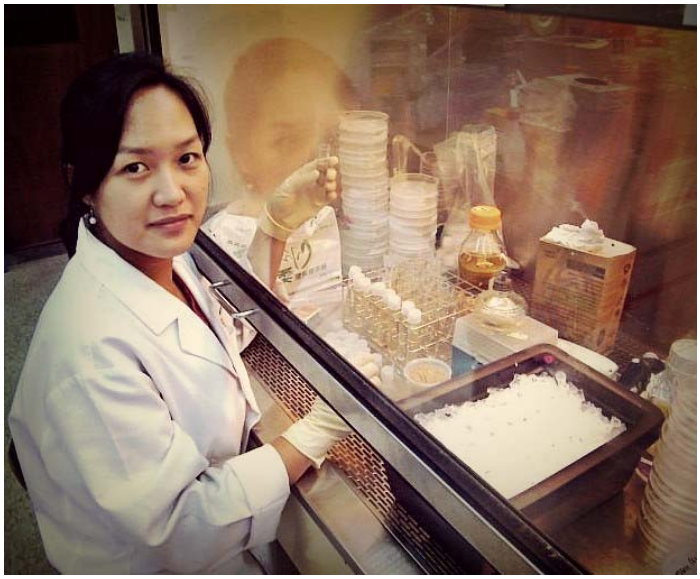
I knew too that in time of (PhD) test, family is best.

Thank you so much for everything Mama, Papa, Shonel, and Bunso. Thank you for your encouragements, prayers and support. You have always been the cheering squad behind me, cheering, yelling, shouting "*Ikaw pa, kaya mo yan! Galingan mo!*" Thank you for always believing in me, trusting in me. You give me confidence in all that I do, knowing that no matter what, whatever happens, you still love and support me. I have a lot to be thankful for, my beautiful family - the memories through the years, the times shared together, full of love and laughter, and sometimes tears. *At para sa* three musketeers, *salamat sa kulitan, kilitian, at kantiyawan*. *Hanggang sa muli!*

And Philippians 4:13 says "I can do all things through Christ who strengthens me." Yes, I acknowledge that all these would have been impossible if not through my Lord, my God and Saviour. Thank You for Your unconditional love, unfailing guidance, overflowing mercy and grace. Thank You for Your wonderful plan for me. Thank You for laying on my heart the desire to pursue science. Thank You for the surprises You showered my path. I can never thank You enough, praise You enough, and worship You enough. May this humble work somehow glorify You.

About the Author (über den Autor)

Joji Grace Villamor-Sernestrand has always been fascinated with plants. As a child, she excitingly watched her parents, both agriculturists, grow a garden in their yard. Her mother, a plant scientist herself, had inspired her to pursue science. In her grade school days, Joji loved visiting her mother´s tissue culture lab. She had always examined the tissue cultures with curious eyes, and had eversince referred to them as tiny green creatures in a bottle. Now, as an adult, she travels the world to pursue science. And in her travels, she would often sample and photograph different plant species. She loves nature, embraces the outdoors, and delights in the flora (and fauna) that grows in it.



Lab 426, Institute of Plant and Microbial Biology, Academia Sinica, Taipei, Taiwan
September 2010

LEBENSLAUF

Persönliche Informationen

Name: Villamor-Sernestrand
Vorname: Joji Grace

Adresse: Max Planck Institut für Pflanzenzüchtungsforschung
Carl-von-Linne Weg 10, D-50829 Köln, Deutschland

Büro 02215062256/276 Mobil 017656554948 villamor@mpipz.mpg.de

Ausbildung

Jan 2011 – Jan 2015	**Promotions-Studium** an der Universität zu Köln, Köln, Deutschland. Max-Planck-Institut fur Pflanzenzüchtungsforschung, unter der Leitung von Prof. Dr. Jane Parker und Dr. Renier van der Hoorn. Thema: Profiling protein kinases and other ATP-binding proteins in Arabidopsis using acyl-ATP probes Stipendium der International Max Planck Research School
Sept 2008 – Sept 2010	**Wissenschaftlicher Mitarbeiter** Institute of Plant and Microbial Biology Academia Sinica, Taipei, Taiwan Betreuer: Dr. Paul Verslues Thema: Quantitative trait loci mapping and characterization of two Arabidopsis ecotypes that differ in proline accumulation in response to low water potential and ABA
Aug 2006 – Aug 2008	**M.Sc. in Molecular Biology** Graduate Institute of Molecular Biology National Chung Hsing University, Taichung, Taiwan Betreuer: Dr. Liang-Jwu Chen Laboratory Thema: Rice functional genomics study using T-DNA insertion mutants: Characterization of rice T-DNA mutants M38333, M69202 and M69217 Stipendium der Taiwan Government Scholarship
Juni 2000 – April 2004	**Bachelor of Science in Biology** University of the Philippines Baguio, Baguio, Philippines

Publikationen

Chandrasekar B, Colby T, Emon AEK, Jiang J, Hong TN, Villamor JG, Harzen A, Overkleeft HS, and van der Hoorn RAL (2014) Broad range glycosidase profiling. Mol. Cell. Proteomics. *in press.*

Zulet A, Gil-Monreal M, Villamor JG, Zabalza A, van der Hoorn, RAL and Royuela M. (2013) Proteolytic pathways induced by herbicides that inhibit amino acid biosynthesis. PLoS One 8, e73847

Villamor JG, Kaschani F, Colby T, Oeljeklaus J, Zhao D, Kaiser M, Patricelli MP, and van der Hoorn RA (2013) Profiling protein kinases and other ATP-binding proteins in Arabidopsis using acyl-ATP probes. Mol Cell Proteomics. 12. 2481-2496

Gu C, Shannon DA, Colby T, Wang Z, Shabab M, Kumari S, Villamor JG, McLaughlin CJ, Weerapana E, Kaiser M, Cravatt BF, van der Hoorn RA (2013) Chemical proteomics with sulfonyl fluoride probes reveals selective labeling of functional tyrosines in glutathione transferases. Chem Biol. 20(4): 541-548

Sharma S, Lin W, Villamor JG, Verslues PE (2013) Divergent low water potential response in Arabidopsis thaliana accessions Landsberg erecta and Shahdara. Plant Cell Environ. 36(5):994-1008

Kesari R, Lasky JR, Villamor JG, Des Marais DL, Chen YJ, Liu TW, Lin W, Juenger TE, Verslues PE (2012) Intron-mediated alternative splicing of Arabidopsis P5CS1 and its association with natural variation in proline and climate adaptation. Proc Natl Acad Sci USA. 109 (23): 9197-9202

Lenger J, Kaschani F, Lenz T, Dalhoff C, Villamor JG, Köster H, Sewald N, van der Hoorn RA (2012) Labeling and enrichment of *Arabidopsis thaliana* matrix metalloproteases using an active-site directed, marimastat-based photoreactive probe. Bioorg Med Chem. 20 (2): 592-596

Sharma S, Villamor JG, Verslues PE (2011) Essential role of tissue-specific proline synthesis and catabolism in growth and redox balance at low water potential. Plant Physiol. 157(1): 292-304

Facundo B. Asia, Joji Grace C. Villamor and Jogel C. Faylogna (2012) The effect of prepared diet on the somatic and gonad growth performance of the sea urschin *Tripneustes gratilla* (Linnaeus, 1758). E-International Scientific Research Journal. 4: 214-228

Wissenschaftliche Leistungen

Speaker	Keystone Symposium on **Plant Immunity: Pathways and Translation** April 7, 2013 - April 12, 2013 Big Sky Resort - Big Sky, Montana, USA
Internship	Proteomics Coursework at the **Ben Cravatt Laboratory** May 28-June 23, 2012 Scripps Research Institute, 10550 North Torrey Pines Road La Jolla, CA 92037, USA
Speaker	**4th West-European Activity-based Proteomics Meeting** June 20-21, 2011 Universität Duisburg-Essen Campus Essen, Germany