Biochemical and Structural Characterization of RCR3-AVR2: a model for Protease-Inhibitor Interactions at the Plant-Pathogen Interface

Inaugural-Dissertation

zur Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen

Fakultät der Universität zu Köln

vorgelegt von

Selva Kumari Ramasubramanian

aus Indien

Köln, Mai 2012

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

Berichterstatter:	Prof. Dr. Paul Schulze-Lefert
	Prof. Dr. Ulf-Ingo-Fluegge
	Dr. Gunther Döhlemann
Prüfungsvorsitzender:	Prof. Dr. Marcel Bucher
Tag der Disputation:	22. Juni 2012

ACKNOWLEDGEMENT

"Gratitude is the memory of heart"

Feeling such gratitude and not expressing it would be like wrapping a gift and not giving it. This thesis is truly the synergistic product of a team of people near and far. It is said that a mediocre mentor tells, a good one explains, a superior one demonstrates and a great mentor inspires. One such great mentor was Dr. Renier van der Hoorn who constantly motivated, inspired and encouraged me to be the scientist I am today. Thank you for the incredible opportunity, your optimism and relentless support for all these years.

I would also like to thank my second supervisor Dr. Imre Somssich, for his critical comments and suggestions during the annual IMPRS progress meetings. My appreciation likewise extends to Prof. Dr. Paul Schulze-Lefert for accepting to be my "Doctor father" and also providing me an opportunity for great learning experience in the MPIPZ.

My heartfelt thanks to the P.hD thesis committee: Prof. Dr. Marcel Bucher, Dr. Ulf-Ingo Flugge, Dr. Gunther Dohlemann and Dr. Wim Soppe for taking the time to be a part of my defence.

I especially want to thank Robert Kolodziejczyk, Rolf Rose, Shabab for a motivating collaboration and learning.

I treasured all the precious moments I shared with the plant chemetics lab members, friends and colleagues from the MPIPZ. Many thanks to Ilyas and Shaista who have been defining hospitality and care much better than the word actually could mean. Iza, Johanna, Gabby and Adriana for all the nice evening and weekend get-togethers, Joji for the marathon runs through the fields and reena was a blessing in disguise during the last weeks of my thesis submission. And thanks to the Indian community at the MPIPZ – Tripta, Ganga, Geo, Sushy, Vimal, Vivek, Shachi, Arun for the nice TATA BAR and your support. And thanks to Anja, Farnusch, Christian, Haibin, Oscar,

Tram, Leo, Bala and all the others who have been in the plant chemetics lab for a fantastic working ambience.

Just saying thank you to shiva anna and Priya akka will never repay their kindness shown during my initial days very far from home. If the world had more people like you it would be a better place. You do make a beautiful difference.

My deepest gratitude goes to both my family in India for their unconditional love and support throughout my life; this dissertation is simply impossible without them. Appa and Amma have been my greatest strength and role models. I am what I am today because of these two beautiful people in my life. Special thanks to my brother and sharmila anni, you have always made life easy for us and the great marathon formatting for the thesis was simply worth mentioning. Both M.Selvi and R.Selvi, thank you so much for staying beside me in all phases of my life providing the support and unconditional love. Vishra you have been a real stress buster everytime I visited India and rejuvenated me with so much positive energy! You are definitely a MAGIC.

Words cannot express my feelings, nor my thanks to my second parents gita aunty, nagarajan uncle for a lovely journey so far. Their unusual ability in weaving family relationships with commitment, humility, love, patience and the ability to forgive one another truly sings. My special thanks to prethivee for his unconditional love and support. Prashy your unflagging positive energy, encouragement and patience has always infused me with an undying spirit of hope, motivation and art of appreciating failures as much as celebrating success. I know our first anniversary shall pass admist busy schedules for the defense day and most of all across continents but Im sure we will make it up in many beautiful anniversaries to come.

I would also like to thank my mentors, teachers, scholars, authors who have influenced my thinking over the years. And, finally, I express my appreciation for the goodness of an overriding Providence in my life.

ABSTRACT

The tomato apoplast is a molecular battle ground for proteases and inhibitors during plant-pathogen interactions. The interaction between structurally diverse pathogen-derived inhibitors (EPIC1, EPIC2B, AVR2 and RIP1) and their target proteases (RCR3, PIP1 and C14) is an ideal system to study molecular arms-races. First we generated a collection of 52 isoforms of proteases and inhibitors in the pFLAG-ATS expression vector for expression in *Escherichia coli*. We summarise the expression of both the proteases and inhibitors and show that the inhibitor proteins were produced successfully, unlike the proteases.

Second we expressed and purified AVR2 on a large scale and show that purified AVR2 is capable of inhibiting RCR3 and triggering the Cf2-mediated hypersensitive response (HR) in tomato demonstrating that recombinant AVR2 is functional. We next implement biophysical and structural biology tools to elucidate the secondary and tertiary structure of AVR2. The major findings are that: (i) AVR2 is a beta protein determined by circular dichroism (CD) (ii) At apoplastic pH, AVR2 exerts a conformational change associated with RCR3 inhibition determined by CD and tyrosine fluorescent spectroscopy (iii) AVR2 is a potent inhibitor of papain determined by enzyme assay using BODIPY FL casein.

Attempts to crystallize AVR2 with and without epitope tags and at different concentrations and conditions failed, possibly because AVR2 is a heavily charged basic protein. Next all 13 lysines and the *N*-terminus of epitope-tagged AVR2 were methylated but no crystals were obtained. This methylated AVR2 still triggers HR in Cf2 tomato plants. Preliminary NMR experiments of non-methylated AVR2 showed good resolution for future structure elucidation of AVR2.

Third, we tested four different heterologous expression systems (plant, bacterium, insect and yeast) to generate high quantities of active and soluble RCR3. We conclude that yeast is the best expression system to produce high amounts of soluble proRCR3 and proRCR3 is fully converted into mature RCR3 in the presence of reducing agents and acidic pH buffer.

Finally, we study the role of double cysteine (Cys24, Cys25) in the catalytic site of RCR3 which is common to PLCP subclass 6 of plant PLCPs. Using agroinfiltration in *Nicotiana benthamiana* we produced C24A, C25A and C24AC25A mutant RCR3 proteins and we discovered that (i) Cys25 but not Cys24 is the essential catalytic residue labelled by activity-based probe MV201. (ii) Surprisingly maturation of RCR3 does not require the catalytic Cys25, indicating that other endogeneous proteases activate RCR3. (iii) Proteolytically inactive RCR3 mutants triggers Cf-2-mediated HR in the presence of AVR2 and (iv) Interestingly, ascorbate enhances the activity of the C24A mutant but not wild-type RCR3, suggesting that Cys24 in RCR3 might have a role in sensing redox potential.

The autocatalytic activation of proRCR3 produced in yeast combined with the maturation of the C25A mutant *inplanta* suggests that RCR3 can be activated by both intramolecular and intermolecular processing.

ZUSAMMENFASSUNG

Während einer Infektion wird der Interzellularraum (Apoplast) einer Tomatenpflanze zum Kampfplatz zwischen Proteasen und deren Inhibitoren. Die Interaktion zwischen strukturell diversen Inhibitoren, die vom Pathogen sekretiert werden (EPIC1, EPIC2B, AVR2 und RIP1) und deren Zielmolekülen, wirtseigene Proteasen wie RCR3, PIP1 und C14, sind ein ideales Modellsystem um das molekulare Wettrüsten zwischen Wirt und Pathogen zu untersuchen. Für diese Studie haben wir zunächst einen Datensatz bestehend aus insgesamt 52 Isoformen dieser Proteasen und deren Inhibitoren im pFLAG-ATS Expressions system, das zur Expression von Proteinen in *Escherichia coli* geeignet ist, erstellt. In dieser Arbeit fassen wir die Ergebnisse dieser Expressionstudie sowohl der Proteasen als auch der Inhibitoren zusammen und zeigen, dass die Inhibitoren erfolgreich exprimiert und sekretiert werden konnten, die Proteasen dagegen nicht.

Desweiteren haben wir AVR2 im großen Maße exprimiert und aufgereinigt und zeigen, dass unser aufgereinigtes AVR2 ein aktiver Inhibitor ist, da es dazu in der Lage ist, RCR3 zu hemmen und die hypersensitive Immunantwort (HR) in Tomaten zu aktivieren. Im Weiteren untersuchen wir die Sekundär- und Tertiärstruktur von AVR2 mit Hilfe von biophysikalischen und strukturbiologischen Methoden. Unsere wichtigsten Ergebnisse zeigen, dass: (i) AVR2 ein Betaprotein ist (gemessen durch Zirkulardichroismus (CD)), (ii) AVR2 bei apoplastischem pH eine Konformationsänderung auslöst, die mit der Hemmung von RCR3 assoziiert ist (ebenfalls gemessen durch CD) und (iii) AVR2 ein wirksamer Inhibitor von Papain ist, was durch enzymatische Analysen mit BODIPY FL Casein untersucht wurde.

Es wurden verschiedene Versuche durchgeführt, AVR2 mit oder ohne Epitopenmarkierung und unter verschiedenen Ausgangskonditionen sowie in verschiedenen Konzentrationen zu kristallisieren. Diese Versuche waren jedoch nicht erfolgreich – vermutlich weil AVR2 ein stark geladenes, basisches Protein ist. Auch Methylierung sämtlicher 13 Lysine sowie des *N*-Terminus von epitopenmarkiertem AVR2 konnte keine Kristallisierung auslösen. Methyliertes AVR2 ist jedoch immer noch in der Lage, in Tomatenpflanzen, die Cf2 besitzen, eine Immunreaktion

hervorzurufen. Erste Ergebnisse einer Kernspinresonanzspektroskopie mit unmethyliertem AVR2 erzielten eine gute Auflösung und begründen eine gute Ausgangsposition für zukünftige Strukturanalysen mit AVR2.

Im Folgenden wurden verschiedene heterologe Expressionssysteme (Pflanze, Bakterium, Insekt und Hefe) getestet, um eine möglichst effiziente Expression von aktivem und löslichem proRCR3 zu erreichen. proRCR3 wird bei saurem pH und in der Gegenwart von Reduktionsmitteln zur Gänze in reifes und funktionsfähiges RCR3 umgewandelt.

Schließlich untersuchen wir die Rolle des Doppelcysteins im katalytischen Zentrum von RCR3, welches der Unterklasse 6 innerhalb der pflanzlichen papainartigen Cysteinproteasen (PLCPs) gemein ist. Mit Hilfe von Agrobakterium-vermittelter, instabiler Transformation von *Nicotiana benthamiana* haben wir folgende RCR3-Varianten produziert: C24A, C25A und C24AC25A. Mit Hilfe dieser Varianten konnten wir herausfinden, dass (i) Cys25 nicht aber Cys24 die essentielle, katalytische Aminosäure ist, die durch die aktivitätsbasierende Sonde MV201 markiert wird, (ii) zum Reifungsprozess von RCR3 überraschenderweise das katalytische Cys25 nicht benötigt wird, was dadurch zu erklären ist, dass andere endogene Proteasen RCR3 aktivieren, (iii) proteolytisch inaktive RCR3-Varianten eine Cf2 vermittelte Immunantwort auslösen können und (iv) Ascorbat interessanterweise die Aktivität der C24A-Variante, nicht aber von Wildtyp-RCR3 erhöht, was darauf schließen lässt, dass Cys24 eine Rolle bei der Detektion des Redoxpotentials spielt.

Die autokatalytische Aktivierung von in Hefe produziertem proRCR3 zusammen mit dem Reifungsprozess der C25A-Variante *in planta* deuten darauf hin, dass RCR3 durch intra- sowie intermolekulare Prozessierung aktiviert werden kann.

TABLE OF CONTENTS

ACKNOWLEDGEMENT I
ABSTRACT III
ZUSAMMENFASSUNG V
TABLE OF CONTENTS VII
INDEX OF FIGURES X
INDEX OF TABLES XI
ABBREVIATIONS XII
CHAPTER 1: INTRODUCTION 1
1.1 The Plant immune system 1
1.2 Interaction of pathogen-derived inhibitors with proteases in the tomato apoplast 2
1.3 Tomato - Cladosporium fulvum interaction 3
1.4 Defense-related Papain-like Cysteine Proteases (PLCPs); from evolution to mechanism 6
1.5 Molecular-arms race at the plant-pathogen interface 7
1.6 Activity Profiling of PLCPs 10
1.7 Structure studies on protease-inhibitor complexes 11
Key Research questions 14
Research Objectives 15
CHAPTER 2: RESULTS 17
2.1 EXPRESSION OF PATHOGEN-DERIVED INHIBITORS AND PLANT PROTEASES 17
2.1.1 Proteases and inhibitors 17
2.2 EXPRESSION, PURIFICATION AND STRUCTURAL STUDIES OF AVR2 21
2.2.1 Expression and Purification of AVR2 21
2.2.2 Removal of epitope tags from FLAG-HIS_6-TEV-AVR2 23
2.2.3 Purified AVR2 blocks labelling of RCR3 by MV201 24
2.2.4 Purified AVR2 inhibits RCR3 and triggers Cf2-dependent HR 25
2.2.5 AVR2 inhibits papain 26
2.2.6 Secondary structure characterization of AVR2 by Circular Dichroism 28
2.2.7 Tyrosine fluorescence spectroscopy confirms conformational change in AVR2. 30
2.2.8 Crystallisation screens of AVR2 31
2.2.9 Nuclear Magnetic Resonance (NMR) of AVR2 34

2.3 HETEROLOGOUS EXPRESSION OF RCR3 IN DIFFERENT EXPRESSION SYSTEMS 36
2.3.1 Expression in plants (*Nicotiana benthamiana)* 36
2.3.2 Expression in bacteria (*Escherichia coli)* 37
2.3.3 Expression in insects (*Spodoptera frugiperda)* 38
2.3.4 Expression in yeast (*Pichia pastoris):* 39
2.4 ROLE OF CYS 24 IN THE ACTIVE SITE OF RCR3 45
2.4.1 A double Cysteine (Cys) in the active site is common in PLCP subclass 6 45
2.4.2 Maturation of RCR3 does not require catalytic Cys25 47
2.4.3 Cys25 in RCR3 is the only target of MV201 47
2.4.4 Proteolytically inactive RCR3 triggers Cf2-mediated HR in the presence of AVR2. 48
2.4.5 Ascorbate enhances activity of RCR3 catalytic mutant C24A 49
CHAPTER 3: DISCUSSION 51
3.1 Structural Characterization of AVR2 51
3.1.1 *Escherichia coli* is an ideal expression system for pathogen-derived inhibitors .. 51
3.1.2 Beta sheet structure of AVR2 is pH sensitive 52
3.1.3 NMR spectroscopy can complement X-ray crystallography for AVR2 structural studies 55
3.1.4 AVR2 a potent inhibitor of Papain 58
3.2 Biochemical Characterization of RCR3 59
3.2.1 Production of high quantities of active, soluble RCR3. 60
3.2.2 Intramolecular and Intermolecular processing of proRCR3 63
3.2.3 Role of double cysteine in active site of RCR3 – Cys24 as redox sensor? 65
CHAPTER 4: MATERIALS AND METHODS 67
4.1 Materials 67
4.1.1 Chemicals and Biochemicals 67
4.2 Methods 70
4.2.1 Recombinant expression and affinity purification of inhibitors and proteases in bacteria 70
4.2.2 Recombinant expression and affinity purification of RCR3 in insects 71
4.2.3 Recombinant expression of proteases in *yeast* and affinity purification 73
4.2.4 Activation of the proRCR3 into mature RCR3 73
4.2.5 Agroinfiltration of *Nicotiana benthamiana* 73
4.2.6 Apoplastic fluid isolation from *Nicotiana benthamiana* leaves 74
4.2.7 Quantification of protein concentration 74

4.2.8 Site-directed mutagenesis 74
4.2.9 Activity-based labelling 75
4.2.10 In-gel fluorescence scanning 75
4.2.11 Western blotting 75
4.2.12 BTH treatments 76
4.2.13 RNA isolation, cDNA synthesis and analysis 76
4.2.14 CD spectra 77
4.2.15 Fluorescence spectroscopy 77
4.2.16 Enzyme kinetics 77
4.2.17 Crystallization screens 78
4.2.18 Lysine-methylation 78
4.2.19 Hypersensitive Response assay 78
REFERENCES 81
ERKLÄRUNG XIII
LEBENSLAUF XIV

INDEX OF FIGURES

Figure 1-1: Selective inhibition of secreted PLCPs by structurally unrelated pathogen effector molecules in the tomato apoplast…….. 3

Figure 1-2: Disulphide bond pattern in AVR2 and RCR3-AVR2 complex trigger Cf-2-mediated defense response……. 5

Figure 1-3: Maturation and Structure of Papain-Like-cysteine proteases (PLCPs)……. 7

Figure 1-4: Selection pressure at the plant-pathogen interface…… 9

Figure 1-5: Activity-based probes; MV201 and DCG-04 …….. 11

Figure 2-0: Expression and secretion of N-terminal FLAG-HIS6-FXa /TEV fusion proteins…… 19

Figure 2-1: Purified FLAG-HIS6-TEV-AVR2 and confirmation by MS Analysis……. 22

Figure 2-2: TEV digestion of FLAG-HIS6-TEV-AVR2 to remove the epitope tags……..... 24

Figure 2-3: Structure of MV201 and inhibition of MV201 labelling by purified AVR2……. 25

Figure 2-4: Purified AVR2 triggers hypersensitive response……. 26

Figure 2-5: AVR2 inhibits papain…….. 27

Figure 2-6: Beta sheet structure of AVR2 is pH sensitive……. 29

Figure 2-7: pH-dependent conformational change of AVR2…….. 31

Figure 2-8: Crystallisation screens of AVR2…….. 32

Figure 2-9: MS analysis and HR assay of Lys-met AVR2 …… 33

Figure 2-10: HSQC spectrum of AVR2……. 35

Figure 3-1: Expression of RCR3 in plants, bacteria, insects and yeast… 39

Figure 3-2: Prediction results of putative O- and N- Glycosylation RCR3 and RCR3 is active……..... 41

Figure 3-3: Maturation of proRCR3 and activity of mature RCR3…… 43

Figure 4-1: Occurrence of double Cysteine in RCR3 and other plant PLCPs……... 46

Figure 4-2: Mutant RCR3 proteins accumulate as mature proteins …… 47

Figure 4-3: MV201 profiling of the catalytic mutants of RCR3……. 48

Figure 4-4: RCR3 mutants trigger Cf2-mediated HR……. 49

Figure 4-5: Upregulated activity of C24A RCR3 mutant in the presence of ascorbate …….. 49

Figure 5-1: Model: Translocation of AVR2 via pH gradient Accompanying conformational changes…… 54

INDEX OF TABLES

Table 1:Proteases and inhibitors cloned into pPLAG-ATS vector for expression in E.coli …… 20

Table 2: Summary of crystallization screens…… 34

Table 3: Summary of RCR3 leaves in different expression systems…….. 40

Table 4: List of Primers …… 79

Table 5: List of RCR3 mutant Agrobacterium constructs……. 80

ABBREVIATIONS

-	fused to (in the context of gene/protein fusion constructs)
%	percentage
°C	degree Celsius
1D	one-dimensional
1DE	one-dimensional gel electrophoresis
2D	two-dimensional
2DE	two-dimensional gel electrophoresis
3’	three prime end of a DNA fragment
5’	five prime end of a DNA fragment
aa	amino acid
ABP	activity-based probe
ABPP	activity-based protein profiling
AF	apoplastic fluid
ALP	Aleurain-like protease
Asn	Asparagine
Asp	Aspartic acid
Avr	avirulence
AVR2	avirulence protein 2
Bio	biotin
BODIPY	boron-dipyrromethene
bp	base pair (s)
BSA	bovine serum albumin
BTH	benzo(1,2,3)thiadiazole-7-carbothioic acid S-methyl ester
C14	RD21 like protease with C-terminal granulin domain
CatB1,B2	Cathepsin B1,B2
CD	Circular dichroism
cDNA	complementary DNA
Cf2	*Cladosporium fulvum* resistance gene
C. fulvum	*Cladosporium fulvum*
C-terminal	carboxy terminal
CYP1,3	Cysteine protease 1,3
Cys	Cysteine
C24, C25	cysteine 24, cysteine 25
DNA	deoxyribonucleic acid
dpi	day(s) post inoculation
dpt	day-post-treatment
DTT	dithiothreitol
E-64	(L-3-trans-Carboxyoxiran-2-Carbonyl)-L-Leucyl-Admat
EDTA	ethylenediaminetetraacetic acid
EPIC1	Extracellular Protease Inhibitor for Cysteine Proteases-1
EPIC2B	Extracellular Protease Inhibitor for Cysteine Proteases-2B
ER	endoplasmic reticulum

ETS	Effector-triggered Susceptibility
Fig.	figure
flg22	22-amino acid peptide fragment of flagellin
h	hour(s)
H2O2	hydrogen peroxide
hpi	hour(s) post inoculation
hr	Hour
HR	hypersensitive response
HRP	horseradish peroxidase
IEF	isoelectric focusing
JCSG	Joint Centre for Structure Genomics Core suites for crystallization screens
kDa	kiloDalton(s)
K_i	inhibition constant
l	litre(s)
LB	Luria-Bertani
LC	liquid chromatography
LE	leaf extract
LRR	leucine-rich repeat
M	milli
M	molar (mol/l)
MAMPs	Microbe-associated Molecular patterns
MgCl2	magnesium chloride
min	minute(s)
mRNA	messenger RNA
MS	mass spectrometry
MV201	Martin Verdoes 201
MW	molecular weight
N.	*Nicotiana benthamiana*
NMR	Nuclear Magnetic Resonance
NPC	no probe control
N-terminal	amino terminal
OD	optical density
ORF	open-reading frame
PAGE	polyacrylamide gel-electrophoresis
PAMPs	Pathogen-associated Molecular Patterns
PBS	phosphate buffered-saline
PCD	programmed cell-death
PCR	polymerase chain reaction
PEG	polyethylene glycol
pH	negative decimal logarithm of the H+ concentration
P.infestans	*Phytophthora infestans*
P. pastoris	*Pichia pastoris*
PIP1	Phytophthora Inhibited Protease I
PLCP	papain-like cysteine protease
PMF	peptide-mass fingerprint

PR	pathogenesis-related
P. syringae	*Pseudomonas syringae*
PTI	PAMP-triggered Immunity
R	resistance
RCR3	REQUIRED FOR Cladosporium fulvum RESISTANCE-3
$RCR3^{pim}$	RCR3 from Lycopersicon pimpinellifolium
$RCR3^{lyc}$	RCR3 from Solanum Lycopersicum
rcr3-3	Mutant of RCR3 lacking activity
RD21	Responsive to Dessication-21
Rh	Rhodamine
RIP1	RD21-inhibiting protein-1
RNA	ribonucleic acid
ROS	reactive oxygen species
rpm	revolutions per minute
RuBisCO	ribulose-1,5-bisphosphate carboxylase/oxygenase
SA	salicylic acid
SDS	sodium dodecyl sulphate
SP	signal peptide
T3SS	type three secretion system
TBS	Tris-buffered saline
TCEP	tris(2-carboxyethyl)phosphine
TEV	tobacco etch virus
Tris	Tris-(hydroxymethyl)-aminomethane
V	Volt
WT	wild-type
Z-L-R-AMC	Z-Leu-Arg-AMC.(Z: N-carbobenzyloxy; 7-Amino-4-methylcoumarin)
µ	micro

CHAPTER 1: INTRODUCTION

1.1 The Plant immune system

Plant pathogens use different strategies to invade the host plant. Nematodes and aphids insert a stylet into a plant cell to feed on the nutrients, pathogenic bacteria inject proteins into the host cell through the type-III secretion system and many oomycetes and fungi use haustoria to invaginate the host cell plasma membrane (Jones and Dangl, 2006). To circumvent pathogen feeding, plants have evolved sophisticated mechanisms to perceive pathogen attack and trigger effective innate immune response by secreting enzymes and small molecules. One such well-characterized perception mechanism is based on the gene-for-gene hypothesis introduced by Flor in the 1940s where plant recognize pathogens by resistance genes *(R)* in plants and their cognate avirulence genes *(Avr)* in the pathogen (Flor, 1942).

Pathogen-associated Molecular Patterns (PAMPs) or Microbe-associated Molecular Patterns (MAMPs) are recognised by the plant surveillance system that consists of receptors that activate an immune response. For instance, a 22-amino acid peptide fragment of flagellin *(flg22)* from bacteria like *Pseudomonas syringae* induces transcription of more than 1100 genes in *Arabidopsis thaliana* (Chinchilla et al., 2006; Felix et al., 1999; Zipfel et al., 2004). This response leads to PAMP-triggered Immunity (PTI) (Chisholm et al., 2006; Jones and Dangl, 2006). Successful pathogens can avoid PTI through the secretion of effector molecules and this phenomenon is called Effector-triggered Susceptibility (ETS) which facilitates the colonization of the host. At this juncture, there exists a molecular arms-race or co-evolution of the host targets that evade manipulation by the effectors and effectors that adapt to new host targets. This adaptation phenomenon results in positive selection for variation of residues at the interaction surface of the effector target resulting in diversification of effectors and their targets (Misas-Villamil and van der Hoorn, 2008).

1.2 Interaction of pathogen-derived inhibitors with proteases in the tomato apoplast

The apoplast of leaves is the first site of pathogen colonization and considered as a molecular battle ground during plant-pathogen interactions. The tomato apoplast contains seven different proteases – RCR3, PIP1, C14, CYP3, ALP, CatB1 and CatB2 (Shabab et al., 2008). Although PIP1 and RCR3 are closely related Papain-Like cysteine proteases (PLCPs), PIP1 dominates the induced proteolytic activity in the tomato apoplast during defense (Shabab et al., 2008). PIP1 is a salicyclic acid-induced defense protease. During infection, three unrelated pathogens secrete structurally diverse inhibitors to counteract the defense mechanism elicited by proteases in tomato (Figure 1-1).

The first pathogen is *Cladosporium fulvum* which secretes AVR2, a small protein not homologous to any known protein. AVR2 inhibits both RCR3 and PIP1, a close relative of RCR3 (Rooney et al., 2005; Shabab et al., 2008; van Esse et al., 2008) (Figure 1-1).

The second pathogen is the oomycete *Phytophthora infestans*, which secretes EPIC1 and EPIC2B, two closely related cystatin-like proteins. EPICs are expressed at early biotrophic stages during infection of tomato or potato (Haas et al., 2009; Tian et al., 2007). Similar to AVR2, EPIC1 and EPIC2B inhibit both RCR3 and PIP1 (Figure 1-1) (Song et al., 2009; Tian et al., 2007) but they have an even higher affinity to the C14 proteases of tomato and potato (Figure 1-1) (Kaschani et al., 2010). The C14 protease carries a *C*-terminal granulin domain which is similar to the Arabidopsis orthologue called RD21 (Yamada et al., 2001). C14-like proteases accumulate in the vacuole and in vesicles and can act as a peptide ligase (Hayashi et al., 2001; Wang et al., 2008; Yamada et al., 2001). The tomato C14 is an abundant protease and has been studied under the names TDI-65, SENU1 and CYP1 and is transcriptionally induced by heat, cold, drought and senescence (Schaffer and Fischer, 1988; Schaffer and Fischer, 1990). The silencing of C14 proteases in *N. benthamiana* increases the susceptibility to *P. infestans* (Kaschani et al., 2010).

The third pathogen is the bacterial pathogen *Pseudomonas syringae* which produces RIP1, a chagasin-like protein that targets the tomato C14 protease (Figure 1-1) (Kaschani and van der Hoorn, unpublished).

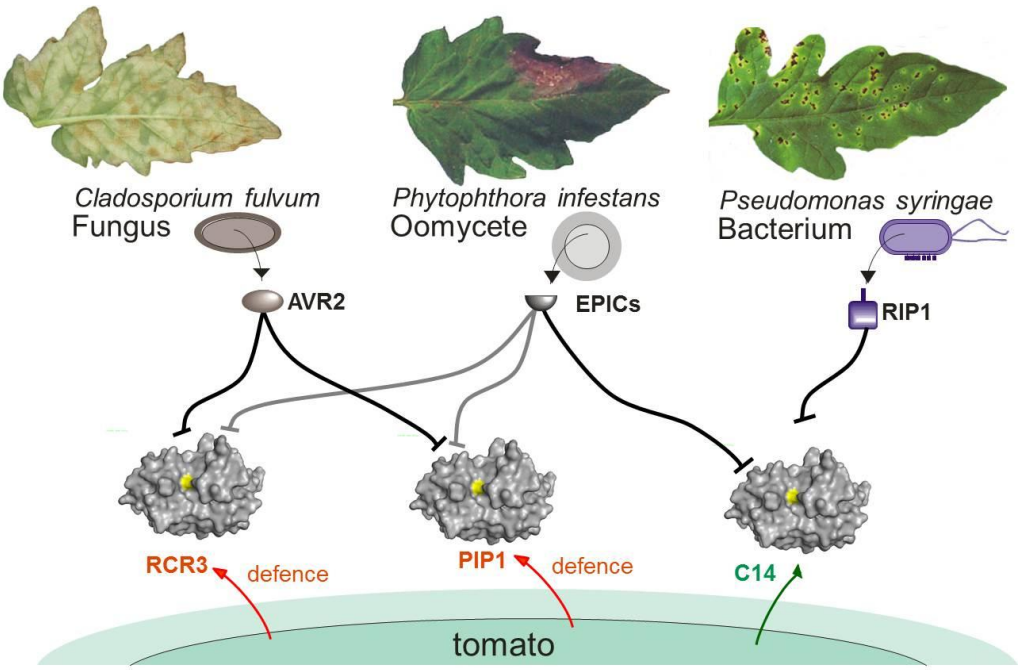


Figure 1-1: Selective inhibition of secreted PLCPs by structurally unrelated pathogen effector molecules in the tomato apoplast

Three different pathogens; a fungus, an oomycete and bacteria attack tomato. During infection there is increased secretion of defense-related PLCPs RCR3, PIP1 and C14. The fungus *Cladosporium fulvum* secretes AVR2 which targets RCR3 and PIP1. Interaction of RCR3 and AVR2 triggers a Cf2-mediated HR in tomato carrying the *Cf-2* resistance gene. The oomycete *Phytophthora infestans* secretes EPIC1 and EPIC2B which inhibits RCR3, PIP1 and C14. The bacterium *Pseudomonas syringae* produces RIP1, which targets C14. Protease-inhibitor interactions are represented with black (strong) and grey (weak) lines. Models of RCR3 and PIP1 are based on crystal structure of 9PAP and C14 on crystal structure of 1S4V. The catalytic Cys (yellow) is in the middle of the substrate binding cleft.

1.3 Tomato - Cladosporium fulvum interaction

Cladosporium fulvum (syn. Passalora fulva) is a biotrophic pathogen on tomato that causes leaf mold disease. Durable resistance against *C. fulvum* has been a major objective for tomato breeders (Joosten and de Wit, 1999; Rivas and Thomas, 2002).

During leaf colonization, *C. fulvum* secretes many low-molecular weight effector proteins into apoplast, and several of these proteins like AVR2, AVR4, AVR4E and AVR9 function as avirulence (AVR2) determinants on specific tomato genotypes mediated by the cognate Cf *(C. fulvum)* resistance (R) proteins like Cf-2, Cf-4, Cf-4E and Cf-9 respectively (de Wit et al., 1997; Joosten and de Wit, 1999; Thomma et al., 2005). Some additional extracellular proteins (Ecps) like Ecp1, Ecp2, Ecp4 and Ecp5 trigger HR in tomato lines carrying cognate Cf-Ecp genes (de Kock et al., 2005; Lauge et al., 2000).

Cladosporium fulvum encodes AVR2, a preprotein of 78 amino acids which matures into a 58 amino acid protein with eight cysteine residues. AVR2 contains disulphide bridges between Cys7-Cys33, Cys12-Cys52, Cys29-Cys43 and Cys 53-57 of which the first three bridges (Cys7-Cys33, Cys12-Cys52, Cys29-Cys43) provide a very compact and stable structure (Figure 1-2A) (Van't Klooster et al., 2011). It has been hypothesised that digestion of this stable structure might be difficult for the extracellular plant proteases in the tomato apoplast. The disulphide pattern for AVR2 is different from that of AVR4 and AVR9 (van den Burg et al., 2003; van den Hooven et al., 2001). AVR2 inhibits several tomato Papain-like Cysteine Proteases (PLCPs) like RCR3 and PIP1 during infection and supresses the defense proteases important for host defense (Shabab et al., 2008; van Esse et al., 2008). AVR2 targets the defense-related protease RCR3 that is required for resistance against the fungus to trigger a Cf-2 mediated HR (Figure 1-2B) (Kruger et al., 2002). RCR3 activity is not dependent on pH, but inhibition by AVR2 occurs only at acidic pH (Rooney et al., 2005). Since AVR2 is also an effective inhibitor of the more abundant PIP1, PIP1 is thought to be virulence target of AVR2 and RCR3 is thought to act as a decoy to trap the fungus into a recognition event in plants carrying Cf-2 receptor protein (Shabab et al., 2008; van der Hoorn and Kamoun, 2008) (Figure 1-2). The virulence role of AVR2 is demonstrated not only in *C. fulvum* by RNAi but also using Arabidopsis expressing AVR2, which became more susceptible to *Botrytis cinerea* and *Verticillium dahlia* (van Esse et al., 2008). However, in tomato carrying *Cf-2*, AVR2 behaves as an avirulence factor and its recognition requires $RCR3^{pim}$ introgressed from *L. pimpinellifolium.*

Although the mechanism of AVR2 perception is unknown, structural modifications of RCR3 by AVR2 is thought to be triggering Cf2-mediated defense signalling, because

(i) a natural variant of RCR3 that occurs in *L. esculentum* causes Cf2-mediated HR in an AVR2-independent manner (Kruger et al., 2002) (ii) inhibition of RCR3 by the irreversible inhibitor E-64 does not trigger a Cf-2-mediated HR unlike the inhibition of RCR3 by AVR2 (Rooney et al., 2005) (iii) rcr3 mutant which also lacks RCR3 activity did not trigger HR (Rooney et al., 2005) (iv) AVR2 behaves as a uncompetitive inhibitor of RCR3 (Van't Klooster et al., 2011).

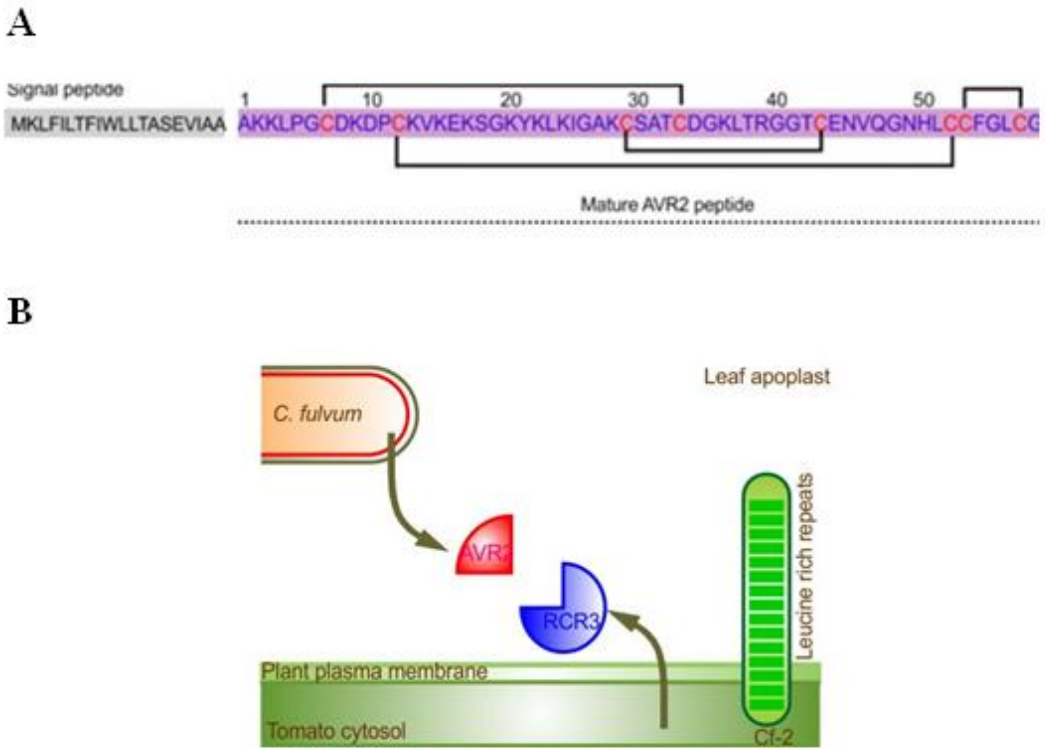


Figure 1-2: Disulphide bond pattern in AVR2 and RCR3-AVR2 complex trigger Cf-2-mediated defense response.

(A)AVR2 protein sequence. The amino acid sequence of AVR2 (78 amino acid residues), consists of 20 amino acid signal peptide (grey) and the 58 amino acid mature AVR2 peptide (violet). The disulphide bridges (black full lines) between the cysteine residues (red) (Cys7-Cys33, Cys12-Cys52, Cys29 –Cys43 and Cys53- Cys57 respectively). The figure was modified and adapted from (Van't Klooster et al., 2011). ***(B)***The fungus *Cladosporium fulvum* secretes AVR2 which targets RCR3 in the tomato apoplast. Interaction of RCR3 (blue) and AVR2 (red) is recognised by Cf-2 receptor protein (green) which triggers a Cf2-mediated HR in tomato carrying the *Cf-2* resistance gene.

1.4 Defense-related Papain-like Cysteine Proteases (PLCPs); from evolution to mechanism

RCR3, PIP1 and C14 are PLCPs that belong to family C1A of clan CA in the MEROPS protease database *(http://merops.sanger.ac.uk)* (Rawlings et al., 2006). The *Arabidopsi*s genome encodes over 800 proteases which are distributed over almost 60 families belonging to 30 different clans. PLCPs belong to the family C1 of Clan CA. The C1 family can be subdivided into i) apoplastic PLCPs (subfamily C1A) which comprises proteases that contain disulphide bridges and accumulate in vesicles or vacuole or apoplast and ii) cytoplasmic PLCPs (subfamily C1B) which comprises proteases that are located in the cytoplasm and lack disulphide bridges (Rawlings et al., 2006). Here, we focus on plant papain-like cysteine proteases (PLCPs) that belong to subfamily C1A. Subfamily C1B proteases do not exist in plants and shall not be discussed further.

PLCPs are expressed as pre-proenzyme precursors which contains a signal peptide (SP); an auto-inhibitory prodomain or propeptide and a mature catalytic domain (Cygler and Mort, 1997; Turk et al., 2000; Wiederanders et al., 2003) (Figure 1-3A, left). The SP targets the proteins into the secretory pathway and is removed during the translocation into the endoplasmic reticulum (Figure 1-3A, left). The propeptide at the *N*-terminus of the protease acts as an intramolecular chaperone facilitating proper folding of the protease and inhibits the protease to prevent premature activity (Santamaria et al., 1998; Taylor et al., 1995; Velasco et al., 1994; Wiederanders et al., 2003). The inactive proenzyme or zymogen undergoes proteolytic processing, whereby the prodomain is removed from the N-terminus thereby producing an active mature protease (Wiederanders et al., 2003) (Figure 1-3A). PLCPs can be activated by other proteases or by autocatalytic processing in the acidic environment, Cathepsin B for example, can be activated by both (Turk et al., 2000; Turk et al., 2001). In some proteases like RD21 and C14, a *C*-terminal granulin domain is also removed by processing (Shindo et al., 2012; van der Hoorn et al., 2004; Yamada et al., 2001) (Figure 1-3A, right). The mature proteases are usually 23-30 kDa in size, and the catalytic triad consists of Cys-His-Asn where the thiol group of catalytic cysteine cleaves peptide bonds in the substrate (Shindo and Van der Hoorn, 2008) (Figure 1-2B). An additional Cys residue in the catalytic site of subclass 6 of plant PLCPs are

thought to have a functional role (Richau et al., 2012). Also RCR3 and PIP1 are some of the proteases that contain this double cysteine in the active site. In this thesis, we study the potential role of extra cysteine residue (Cys 24) in RCR3.

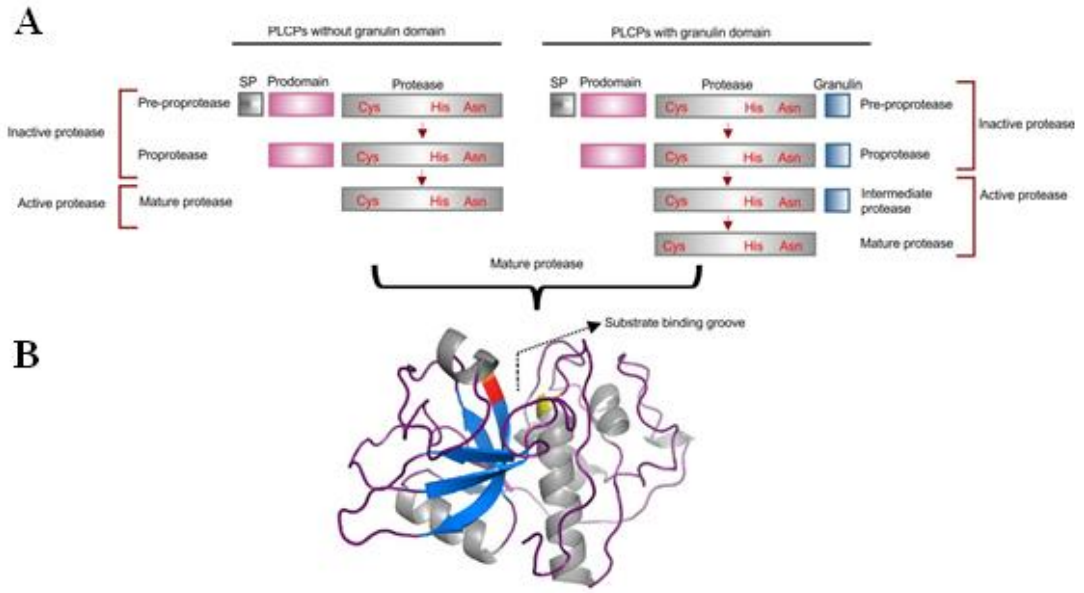


Figure 1-3: Maturation and Structure of Papain-Like-cysteine proteases (PLCPs)

(A)Domain structure of the open reading frame of PLCPs subfamily C1A. The PLCP gene product contains a signal peptide (SP, grey), an autoinhibitory prodomain (prodomain, magenta), a protease domain (protease, dark grey) with catalytic triad Cys, His, Asn residues (red) (Left). PLCPs are produced as pre-proprotease which converts to the proprotease by removing the SP. The proprotease further activates into mature protease by removing the prodomain. Some PLCPs like RD21 and C14, contain additional C-terminal extension domain (granulin, dark green, right) (Yamada et al., 2001, Shabab et al., 2008). The granulin domain is removed from the intermediate protease during conversion into the mature protease. The pre-proprotease, proprotease are the inactive isoforms of PLCPs. Intermediate and mature proteases are active isoforms. ***(B)***Cartoon model of papain (9PAP) with the alpha-helix domain (grey, right), beta-sheet domain (blue, left), loops (magenta) and the catalytic triad with Cys 25 (yellow), His 159 (red) and Asn 158 (orange). The substrate binding cleft is marked in the enzyme (black dotted arrow).

1.5 Molecular-arms race at the plant-pathogen interface

Selection pressure on the effector-targeted proteases to evade the inhibition and on the pathogen-derived effectors to adapt to the new proteases has caused variant residues

at the interaction surface of the protease-inhibitor complex. Traces of the molecular-arms race of the enzyme-inhibitor complexes at the plant-pathogen interface has been reviewed with different examples and model (Figure 1-4A) (Misas-Villamil and van der Hoorn, 2008).

Interestingly, the structural models of RCR3 and PIP1 based on papain (9PAP) revealed that natural variation in wild tomato resides in amino acids at the surface of the protein around the substrate binding groove. In contrast, amino acids located inside the protein are under conservative selection (Shabab et al., 2008). Such variation in amino acids on the surface of the protease might affect the interaction or inhibition by the pathogen effectors. For example, a naturally occurring N194D mutation on the surface of RCR3 prevents inhibition by AVR2 (Figure 1-4B). This selection pressure leads to diversifying selection or molecular arms race paving way for evolvement of new enzymes and inhibitors to withstand the selection pressure.

C14 in wild potato exhibits a pattern of diversifying selection which is not present in wild tomato, which correlates with co-evolving plant-pathogen interactions (Kaschani et al., 2010). Furthermore, a 3D model of C14–EPICs based on papain (9PAP) and cystatin (3IMA) predicted that variant residues in potato C14 are located at the putative interaction surface with EPICs (Kaschani et al., 2010; Kaschani and Van der Hoorn, 2011). It was proposed that three variant residues (K186ND, T210I, E62Q) in C14 might be the result of arms-race at the plant-pathogen interface (Figure 1-4C).

One of the main objectives of my thesis is to implement structural biology to understand the positive selection on defense proteases RCR3 and PIP1. Structural biology is essential to understand the mode of inhibition, the evolutionary origin of inhibitors and the selective pressures on the interaction surface.

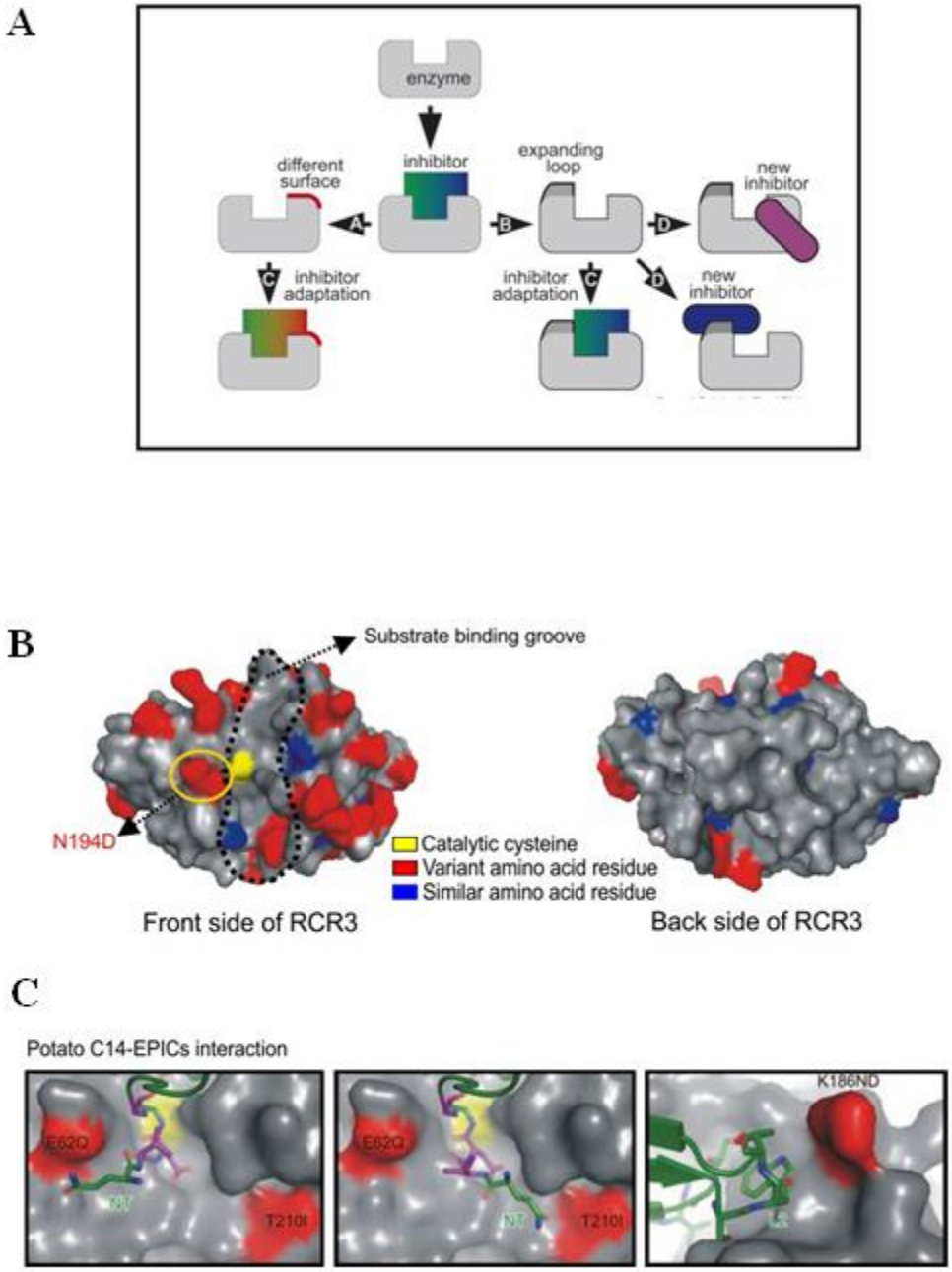


Figure 1-4: Selection pressure at the plant-pathogen interface

(A) Different modes of selective pressure on enzymes to evade inhibition leads to adaptations for example surface decoration (A) or loop expansion (B) Subsequent inhibitor adaptation may reside in adjustments (C) or the selection of entirely new inhibitors (D) Adaptations in both inhibitors and enzymes results in positive selection for variance at solvent-exposed residues at the interacting surface. Figure adapted from (Misas-Villamil and van der Hoorn, 2008). ***(B)*** Front and back side of RCR3 model based on papain (9PAP), with variant residues found in wild tomato. The catalytic residue (yellow), substrate binding groove (black dotted line), position with nonsimilar variance and similar variance are indicated in red and blue respectively

(Shabab et al., 2008). The naturally occurring N194D variant (orange circle) is insensitive to AVR2 inhibition. The figure is modified and adapted from (Shabab et al., 2008). ***(C)***Model of C14-EPICs interaction. The EPICs bind to the substrate binding groove of C14. Residues in the N-terminus and loops LI and L2 are shown as sticks. Polymorphic residues in EPICs are shown in purple. The EPICs may stretch into the left (Figure 1-4C, left) or right (Figure 1-4C, right) cavities of non-prime substrate binding site of C14. The L2 loop of EPICs may interact with K168ND polymorphic site in C14. The figure was adapted from (Kaschani and Van der Hoorn, 2011).

1.6 Activity Profiling of PLCPs

Activity-based Protein Profiling (ABPP) will be frequently used in this thesis to detect PLCPs activity and inhibition. ABPP is the use of small reactive molecules called Activity-based Probes (ABPs) that react with the active site residues of proteins in an activity-dependent manner resulting in a covalent, irreversible bond between the probe and the protein (Campbell and Szardenings, 2003). The labelled proteins can be detected on protein gels or blots or identified by mass spectrometry. ABPP displays activities rather than the abundance of the proteases and is a simple method to demonstrate the activities of proteases in complex proteomes.

ABPP was launched one decade ago by chemical biologists in the field of medicine, mainly by the laboratories of Dr. Cravatt (Scripps Institute, San Diego, CA, USA) and Dr. Bogyo (Stanford Medical School, CA, USA). This technology was implemented in plants by Dr. Van der Hoorn and coworkers initially by monitoring the activities of PLCPs in leaf proteomes of *Arabidopsis thaliana* with the biotin version of DCG-04 (van der Hoorn et al., 2004).

ABPs are composed of three major elements: reactive group or warhead, reporter tag and a linker. The warhead is often an irreversible inhibitor that acts as a substrate but locks the mechanism of the enzyme in the covalent intermediate state. Other warheads can be based on reversible inhibitors coupled to photoreactive groups, or suicide substrates (Cravatt et al., 2008; Kolodziejek and van der Hoorn, 2010). Using different warheads, probes have been generated for phosphatases, PLCPs, lipases, methylesterases, kinases and many other protein families (Cravatt et al., 2008; Evans and Cravatt, 2006). The reporter is either biotin for purification and detection or a

fluorophore like bodipy or rhodamine for quick detection. Finally, the linker connects the reactive group to the reporter tag at an appropriate distance avoiding steric congestion between the components.

Design of DCG-04 and MV201

In this work, ABPP has been used to monitor the activity of wild-type and mutant RCR3 proteins using labelled derivatives of E-64 called DCG-04 and MV201 (Figure 1-5). E-64 is a promiscuous irreversible inhibitor cysteine protease inhibitor that is broadly reactive towards PLCPs (Barrett et al., 1982). DCG-04 is a biotinylated derivative of E-64 (Figure 1-5A) (Greenbaum et al., 2000) where MV201 carries bodipy and an azide minitag (Figure 1-5B) (Richau et al., 2012) .

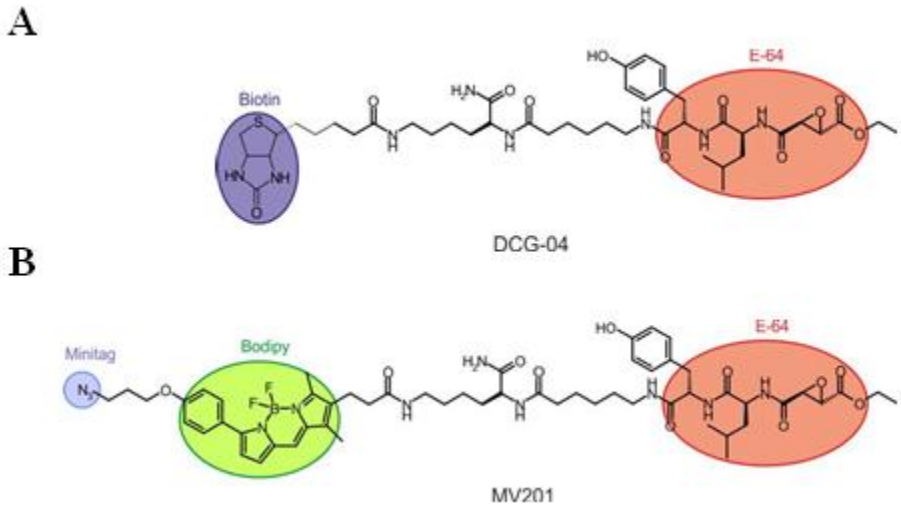


Figure 1-5: Activity-based probes; MV201 and DCG-04

DCG-04 ***(A)*** and MV201 ***(B)*** have E-64 (red), bodipy reporter tag (green) and biotin reporter tag (violet) or an azide minitag (blue). Both ABPs label the thiol group of the catalytic cysteine residue of PLCPs in an activity-dependent manner.

1.7 Structure studies on protease-inhibitor complexes

Structural information on protease-inhibitor complexes will greatly enhance our understanding on for (e.g) the perception of AVR2 and the role of variant residues. The structure of proteins can be elucidated at two different levels i) determining the secondary structure, folding, conformational stability and binding properties of protein in solution and ii) determining the tertiary structure of proteins.

Circular dichroism (CD) is an excellent tool for rapidly evaluating the secondary structure of proteins. Briefly, circular dichroism measures the differences in the absorption of left-handed polarized light versus right-handed polarized light as a result of structural asymmetry. An ordered structure results in a spectrum that contains both positive and negative signals, while the absence of a regular structure in a protein corresponds to zero intensity (Greenfield, 2006). In this thesis, we used CD measurements to determine how AVR2 is folded, and if the structure of AVR2 is pH dependent. Inhibition of RCR3 by AVR2 but not RCR3 activity is pH-dependent because the inhibition was observed only at apoplastic pH (Rooney et al., 2005). Hence, we also study the conformational stability of AVR2 at different pH.

Fluorescence spectroscopy determines the folding and refolding of proteins. This technique is dependent on the content of aromatic amino acids with intrinsic fluorescent properties (Tyr, Trp or Phe). The conformational changes exerted by the proteins can be monitored by the aromatic residues because of their fluorescent properties which are sensitive to their environment when protein folds or unfolds (Yan and Marriott, 2003). In the native form, Tyr and Trp have a high quantum yield and high fluorescence intensity as they are buried within the core of the protein. However, the unfolding of protein could be monitored by the decrease in fluorescence intensity when exposed to solvents (Yan and Marriott, 2003). We implemented this tool for AVR2 to study if conformational changes in AVR2 were pH- dependent.

The three-dimensional (3D) structure of proteins can be studied by various methods, of which significant applications include Nuclear Magnetic Resonance (NMR) and X-ray crystallography. NMR and X-ray crystallography remain the standards of structure determination.

NMR provides information on the structural, thermodynamic and kinetic properties of proteins. This method has been used for molecular characterization of the 3D structure and protein dynamics. In principle, NMR relates to the spinning of electrically charged atomic nuclei in the macromolecule or protein with a static magnetic field. This magnetic field makes the spin-states of the atomic nuclei vary in energy level and NMR is used to measure this transition in energy at different spin state. For example, common NMR active nuclei are 1H, 13C, 15N, 31P, 29Si and others. Advancements in NMR have made structural analyses of proteins or protein

complexes in 30-40 kDa range more routine (Homans, 2004) unlike the larger proteins (greater than 40 kDa). This remains the major limitation of NMR.

X-ray crystallography is the scattering of X-rays by electrons in the crystallized macromolecule. This technique has been used extensively to study the interactions of protease-inhibitor complexes like Chagasin-Cathepsin L (Ljunggren et al., 2007), Chagasin-Papain (Redzynia et al., 2009), Papain-E64 (Kim et al., 1992; Yamamoto et al., 1991), Chagasin-Human Cathepsin B (Redzynia et al., 2008) and Stefin B-Papain (Stubbs et al., 1990).

The determination of a 3D- structure by X-ray crystallography involves seven different steps.

1. ***Protein purification*:** High purity and quantity of the purified protein (milligrams) is a prerequisite factor for protein crystallization. Affinity tags (His_6, GST, MBP) and other genetically engineered fusion cleavage tags (TEV, Factor Xa) are employed to facilitate the purification of the protein and removing the epitope tags.

2. ***Protein crystallization*:** The critical part towards determining the structure of a protein is to obtain large and highly diffracting crystals. However during the initial screening, mostly small crystals are obtained which requires further optimisation to grow them further under suitable conditions. The most often employed method of growing protein crystals is vapour diffusion, where the protein is present in the well with a lower reagent concentration than the reservoir. Water from the protein drop vapourises and ends up in the buffer reservoir thereby slowly increasing the supersaturation level for the protein.

3. ***X-ray diffraction:*** The protein crystals are mounted and placed on the detector under the X-ray beam either at room temperature or mounted frozen in a small loop in a stream of liquid nitrogen at 100K. A powerful X-ray accelerator like a synchrotron is used to scatter the electrons in the beam. This involves measuring a large number of reflection intensities where position of each atom in the crystal structure influences the intensities of reflections.

4. ***Data processing****:* The primary result of an X-ray diffraction experiment is an electron density map which represents the distribution of electrons in the molecule. The processing of the diffraction data involves many established algorithms and software that convert the data into an electron density map.

5. ***Structure solution and refinement****:* The quality of the electron density map obtained during data processing could be improved by refinement. Refinement will improve phases and produce precise and clear maps and thereby generating better models using different software and algorithms.

6. ***Structure analysis****:* A protein structure with an R-factor below 25% is a good model. However at this stage, knowledge of the protein at the biological level is necessary to interpret the details in the structure.

7. ***Deposition in the Protein Data Bank (PDB)*****:** The PDB database (http://www.pdb.org) is a collection of different protein molecules with and without its counterpart ligands. The structure is deposited in the PDB where the structure is assigned to a code (e.g. Papain 9PAP).

Key Research questions

- A natural variant of RCR3 that occurs in *Solanum lycopersicum (*RCR3^{lyc}) causes Cf2- mediated HR in an AVR2-independent manner (Kruger et al., 2002).
  Is the structure of RCR3^{lyc} different from the RCR3^{pim} structure?

- It was proposed that the inhibition of RCR3 by AVR2 induces a conformational change in RCR3 thereby triggering a HR (Kruger et al., 2002). Inhibition of RCR3 by the irreversible inhibitor E-64 does not trigger a Cf2-mediated HR unlike the inhibition of RCR3 by AVR2 (Rooney et al., 2005).
  Is there a conformational change in RCR3 upon AVR2 binding or structural change in AVR2?

- Inhibition of RCR3 by AVR2 and not RCR3 activity itself is dependent on apoplastic pH (Rooney et al., 2005).
  Does apoplastic pH influence the conformation of AVR2 to bind and inhibit RCR3*?*

- It has been reported that AVR2 is a non-competitive inhibitor of RCR3 (Van't Klooster et al., 2011) which implies that AVR2 may be an allosteric inhibitor that changes conformation of RCR3 upon binding.
  Where is the binding site of AVR2 on RCR3?

- In nature, RCR3, PIP1 and C14 carry variant residues on their surface (Misas-Villamil and van der Hoorn, 2008). Single variant residues at the surfaces may determine outcome of interactions with pathogen-derived inhibitors (Kaschani et al., 2010; Kaschani and Van der Hoorn, 2011; Shabab et al., 2008).
  How do variant residues affect the interactions with inhibitors? Why is the N194D mutant of RCR3 insensitive to AVR2 inhibition? Are these variant residues resulting from a co-evolutionary arms-race?

- PLCPs can be activated both by intramolecular and intermolecular processing.
  How is RCR3 activated? Is the prodomain autocatalytically removed?

- PLCPs of subclass 6 contain plant PLCPs RCR3 and PIP1 that contain a double cysteine in the active site (Richau et al., 2012).
  What is the role of the extra cysteine Cys24 in RCR3 or PIP1?

Research Objectives

The objective of this project is to study protease-inhibitor interactions at the plant-pathogen interface. To achieve this objective, we will follow the following approaches:

First, a collection of tomato proteases and pathogen-derived inhibitors will be generated using a novel secretory expression system (pFLAG-ATS) in *E.coli*. The

protein expression of the proteases and inhibitors will be tested and the protein expression level will be optimised to generate sufficient protein quantities.

Second, from the library, we prioritize RCR3 and AVR2 for large scale protein production as they are of primary interest. We show that AVR2 is produced and purified in sufficient amount. And RCR3 expression is insufficient in *E.coli* even after further optimisation. Different heterologous expression systems (plant, insect and yeast) were tested to produce large, soluble quantities of RCR3. RCR3 produced in *Pichia* will be studied for maturation of the proRCR3 to mature and active RCR3.

Third, after successful expression of AVR2 in *E.coli* and RCR3 in *Pichia*, we test the purified recombinant AVR2 for inhibitor activity by ABPP and also determine if AVR2 is active in an HR assay to determine if AVR2 is in native conformation. We use enzyme assays to show that AVR2 inhibits papain. Finally we implement CD spectroscopy and tyrosine fluorescence for secondary structure determination at various pH.

Fourth, we setup the crystallization screens for three variant forms of AVR2 (i) untagged AVR2 (ii) FLAG-HIS_6-TEV-AVR2 and (iii) methylated FLAG-HIS_6-TEV-AVR2 for three-dimensional (3D) structure determination of protein.

Finally, we study the role of the double Cys (Cys24 and Cys25) in the active site of RCR3 by site-directed mutagenesis, agroinfiltration in *Nicotiana benthamiana* and ABPP.

CHAPTER 2: RESULTS

2.1 EXPRESSION OF PATHOGEN-DERIVED INHIBITORS AND PLANT PROTEASES

2.1.1 Proteases and inhibitors

To generate a collection of proteases and inhibitors, the pFLAG-ATS expression vector system was implemented to clone the pathogen-derived inhibitors (EPIC1, EPIC2B, AVR2 and RIP1) and the proteases (RCR3, PIP1 and C14). The strong *tac* promoter (a hybrid of trytophan and lactose promoter) in this vector drives high protein expression levels when using IPTG as a de-repressor. The OmpA signal sequence translocates the processed and soluble proteins into the periplasm of *E.coli* cells. In addition, pFLAG-ATS encodes an *N*-terminal FLAG epitope tag, (Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys). Due to the hydrophilic nature, the FLAG peptide will be on the surface of the fusion protein and serves as a binding site for anti-FLAG antibody for detection and purification of the fusion protein. In addition to the above mentioned modifications, we introduced a polyhistidine (His_6) epitope tag *C*-terminal to the FLAG epitope as cost-effective measure for purification. We also introduced the Ile-Glu-Gly-Arg peptide as a cleavage site for Factor Xa (FXa) processing or the Glu-Asn-Leu-Tyr-Phe-Gln-Gly peptide for processing by the TEV protease to remove the epitope tags after protein purification which might otherwise hamper the crystallization process. To improve the efficiency of TEV cleavage, we also introduced a spacer sequence (Asp-Tyr-Asp-Ile-Pro-Thr-Thr) before the TEV cleavage site (Figure 2-0A). Hence the recombinant fusion proteins were expressed as FLAG-HIS_6-FXa-inhibitor/protease or FLAG-HIS_6-TEV-inhibitor/protease.

We generated two different forms of the proteases: with the prodomain (PD) and without the prodomain (pd). The importance of the prodomain in protein folding and activation has already been explained (see Chapter 1; section 1.4). C14 was generated either in the presence or absence of the *C*-terminal granulin domain. Furthermore we

generated catalytic mutants of RCR3, PIP1 and C14 because mutation of the catalytic cysteine (Cys25) residue into alanine is necessary to prevent the autocatalytic degradation of cysteine proteases (Eakin et al., 1993). In addition to the C25A mutant in RCR3 we also generated C24A and the double mutant C24AC25A to study the role of the Cys24 (see Chapter 2; section 2.4).

Following the construction of the plasmids, the proteins were expressed and protein accumulation in the medium was analyzed on protein gels with coomassie staining. Inhibitors were all detected in the medium in relatively high amounts, and with the expected molecular weight (Figure 2-0B). The low expression of EPIC2B could be improved by IPTG induction at 28˚C instead of 37 ˚C (data not shown).

Unfortunately, none of the proteases were detected in the medium (summarized in Table 1). We tested protein levels in the soluble cell fraction (SCF), and could only detect very low amounts of the proteases (Table 1). However, when the insoluble cell fraction (ICF) was dissolved in urea and analysed on protein gel, high amounts of proteases were detected (Table 1). Thus all inhibitors were secreted as soluble proteins in high amounts but the proteases could not be produced as soluble proteins using this *E.coli* expression system.

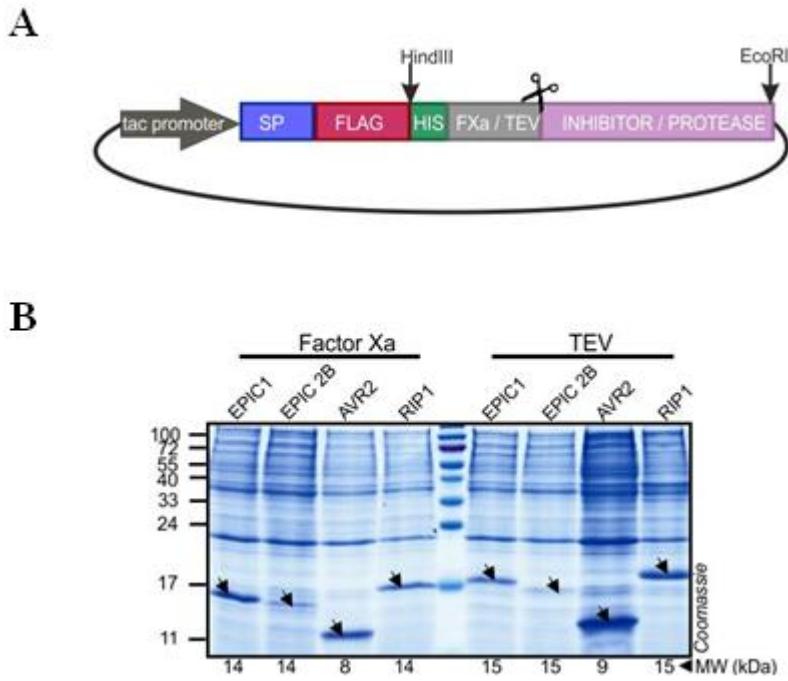


Figure 2-0: Expression and secretion of *N*-terminal FLAG-HIS$_6$-FXa /TEV fusion proteins.

***(A)*Schematic diagram of pFLAG-ATS vector system.** The tac promoter (dark grey arrow), OmpA signal peptide (SP, blue), FLAG peptide (red), HIS$_6$ tag(green), FXa or TEV recognition site (grey) and the inhibitor or protease insert (purple). The PCR products of the proteases and inhibitors were cloned into the pFLAG-ATS system as a *HindIII/XhoI* or as *HindIII/EcoRI* respectively. The scissor represents the cleavage site of FXa or TEV proteases to remove the epitope tags. ***(B)* SDS-PAGE analysis of the inhibitor proteins secreted into the medium.** Recombinant inhibitors with FLAG-HIS$_6$-FXa/TEV epitope tags at the *N*-terminus were ectopically expressed in *E.coli.* The crude protein extracts were analysed by SDS-PAGE to detect the expression level of the proteins (black arrow heads). The numbers below the protein gel indicate the theoretical molecular weight of the inhibitors (EPIC1, EPIC2B, AVR2 and RIP1) with *N*-terminal tags.

Table1

Proteases and Inhibitors cloned into pFLAG-ATS vector for expression in *E.coli*

Protein[1] *(species)*	Constructs[2]	Description[3]	Theoretical MW* (kDa)	Expression [4]		
				Medium[a]	Cell fraction [b] (soluble:SCF)	Cell fraction[c] (Insoluble:ICF)
Pathogen-derived inhibitors						
EPIC1 (*P. infestans*)	pSK 001.01	FLAG-HIS$_6$-FXa-EPIC1	14	+++	ND	ND
EPIC1 (*P. infestans*)	pSK 001.02	FLAG-HIS$_6$-FXa-EPIC1	14	+++	ND	ND
EPIC2B (*P. infestans*)	pSK 002.01	FLAG-HIS$_6$-FXa-EPIC2B	14	+++	ND	ND
EPIC2B (*P. infestans*)	pSK 002.02	FLAG-HIS$_6$-FXa-EPIC2B	14	+++	ND	ND
AVR2 (*C. fulvum*)	pSK 003.01	FLAG-HIS$_6$-FXa-AVR2	8	+++	ND	ND
AVR2 (*C. fulvum*)	pSK 003.02	FLAG-HIS$_6$-FXa-AVR2	8	+++	ND	ND
RIP1 (*P. syringae*)	pSK 004.01	FLAG-HIS$_6$-FXa-RIP1	14	+++	ND	ND
RIP1 (*P. syringae*)	pSK 004.02	FLAG-HIS$_6$-FXa-RIP1	14	+++	ND	ND
EPIC1 (*P. infestans*)	pSK 005.01	FLAG-HIS$_6$-TEV-EPIC1	15	-	ND	ND
EPIC2B (*P. infestans*)	pSK 006.01	FLAG-HIS$_6$-TEV-EPIC2B	15	+++	ND	ND
AVR2 (*C. fulvum*)	pSK 007.01	FLAG-HIS$_6$-TEV-AVR2	10	+++	ND	ND
AVR2 (*C. fulvum*)	pSK 007.02	FLAG-HIS$_6$-TEV-AVR2	10	+++	ND	ND
RIP1 (*P. syringae*)	pSK 008.01	FLAG-HIS$_6$-TEV-RIP1	16	+++	ND	ND
RIP1 (*P. syringae*)	pSK 008.02	FLAG-HIS$_6$-TEV-RIP1	16	+++	ND	ND
Plant proteases						
C14 (*S.lycopersicum*)	pSK 009.01	FLAG-HIS$_6$-FXa-pd-C14 Granulin	38	-	+	ND
C14 (*S.lycopersicum*)	pSK 010.01	FLAG-HIS$_6$-TEV-pd-C14 Granulin	39	-	+	ND
C14 (*S.lycopersicum*)	pSK 011.01	FLAG-HIS$_6$-TEV-PD-C14 Granulin	52	-	+	ND
C14 (*S.lycopersicum*)	pSK 011.02	FLAG-HIS$_6$-TEV-PD-C14 Granulin	52	-	+	ND
PIP1 (*S.lycopersicum*)	pSK 012.01	FLAG-HIS$_6$-FXa-pd-PIP1	25	-	+	ND
PIP1 (*S.lycopersicum*)	pSK 013.01	FLAG-HIS$_6$-FXa-PD-PIP1	37	-	+	ND
PIP1 (*S.lycopersicum*)	pSK 013.02	FLAG-HIS$_6$-FXa-PD-PIP1	37	-	+	ND
PIP1 (*S.lycopersicum*)	pSK 014.01	FLAG-HIS$_6$-TEV-pd-PIP1	26	-	+	ND
PIP1 (*S.lycopersicum*)	pSK 015.01	FLAG-HIS$_6$-TEV-PD-PIP1	38	-	+	ND
RCR3 (*S.lycopersicum*)	pSK 016.01	FLAG-HIS$_6$-FXa-pd-RCR3	26	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 016.02	FLAG-HIS$_6$-FXa-pd-RCR3	26	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 017.01	FLAG-HIS$_6$-FXa-PD-RCR3	38	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 018.01	FLAG-HIS$_6$-TEV-pd-RCR3	27	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 019.01	FLAG-HIS$_6$-TEV-PD-RCR3	39	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 019.02	FLAG-HIS$_6$-TEV-PD-RCR3	39	-	+	+++
C14 (*S.lycopersicum*)	pSK 020.01	FLAG-HIS$_6$-FXa-pd-C14 granulin	28	-	+	ND
C14 (*S.lycopersicum*)	pSK 021.01	FLAG-HIS$_6$-TEV-pd-C14 granulin	29	-	+	ND
C14 (*S.lycopersicum*)	pSK 022.01	FLAG-HIS$_6$-TEV-PD-C14 granulin	42	-	+	ND
RCR3 (*S.lycopersicum*)	pSK 023.01	FLAG-HIS$_6$-FXa-pd-RCR3 (C25A)	26	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 024.01	FLAG-HIS$_6$-FXa-PD-RCR3 (C25A)	38	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 025.01	FLAG-HIS$_6$-TEV-pd-RCR3 (C25A)	27	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 026.01	FLAG-HIS$_6$-TEV-PD-RCR3 (C25A)	39	-	+	+++
PIP1 (*S.lycopersicum*)	pSK 027.01	FLAG-HIS$_6$-FXa-pd-PIP1 (C25A)	25	-	ND	ND
PIP1 (*S.lycopersicum*)	pSK 028.01	FLAG-HIS$_6$-FXa-PD-PIP1 (C25A)	37	-	ND	ND
PIP1 (*S.lycopersicum*)	pSK 029.01	FLAG-HIS$_6$-TEV-pd-PIP1 (C25A)	26	-	ND	ND
PIP1 (*S.lycopersicum*)	pSK 030.01	FLAG-HIS$_6$-TEV-PD-PIP1 (C25A)	38	-	ND	ND
C14 (*S.lycopersicum*)	pSK 031.01	FLAG-HIS$_6$-FXa-pd-C14 granulin(C25A)	28	-	ND	ND
C14 (*S.lycopersicum*)	pSK 032.01	FLAG-HIS$_6$-TEV-pd-C14 granulin(C25A)	29	-	ND	ND
C14 (*S.lycopersicum*)	pSK 033.01	FLAG-HIS$_6$-TEV-PD-C14 granulin (C25A)	42	-	ND	ND
C14 (*S.lycopersicum*)	pSK 034.01	FLAG-HIS$_6$-FXa-pd-C14 Granulin (C25A)	38	-	ND	ND
C14 (*S.lycopersicum*)	pSK 035.01	FLAG-HIS$_6$-TEV-pd-C14 Granulin (C25A)	39	-	ND	ND
C14 (*S.lycopersicum*)	pSK 036.01	FLAG-HIS$_6$-TEV-PD-C14 Granulin (C25A)	52	-	ND	ND
RCR3 (*S.lycopersicum*)	pSK 037.01	FLAG-HIS$_6$-TEV-pd-RCR3 (C24A,C25A)	27	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 037.02	FLAG-HIS$_6$-TEV-pd-RCR3 (C24A,C25A)	27	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 037.03	FLAG-HIS$_6$-TEV-pd-RCR3 (C24A,C25A)	27	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 037.04	FLAG-HIS$_6$-TEV-pd-RCR3 (C24A,C25A)	27	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 038.01	FLAG-HIS$_6$-TEV-pd-RCR3 (C24A)	27	-	+	+++
RCR3 (*S.lycopersicum*)	pSK 038.02	FLAG-HIS$_6$-TEV-pd-RCR3 (C24A)	27	-	+	+++

[1]Source organism of effector or protease; [2] Plasmid library code; [3] All proteins are *N*-terminally tagged with FLAG and His_6 epitope tags, followed by a TEV or Factor Xa cleavage site. Proteases were produced with (PD) and without prodomain (pd). In addition, Cys residues were mutated (C25A and C24A) and the granulin domain of C14 was present (Granulin) or absent (granulin); [*] Theoretical molecular weight; [4]The proteins accumulating in the [a]medium, [b]soluble cell fraction (SCF) and [c]insoluble cell fraction (ICF). Detection of protein expression is based on western blot analysis using anti-FLAG or anti RCR3 antibody using RCR3 expressed in *E.coli* (ICF) as a control. Quantification of proteins by Bradford assay gives a read out for the accumulation of protein: > 3mg/ml (+++), < 0.2mg/ml (+), undetected (-) and not determined (ND).

2.2 EXPRESSION, PURIFICATION AND STRUCTURAL STUDIES OF AVR2

2.2.1 Expression and Purification of AVR2

After confirming the expression of recombinant FLAG-HIS-TEV-AVR2 (pSK 007.02) in *E.coli*, we expressed and purified FLAG-HIS-TEV-AVR2 to generate large amounts for structural studies. First culture media containing AVR2 was precipitated by ammonium sulphate. This technique is often used to concentrate protein from a larger volume without affecting the protein structure and can also help to remove contaminant proteins. The precipitated FLAG-HIS-TEV-AVR2 was dissolved and bound to Ni^{2+}-NTA agarose (see Chapter 4; section 4.2.1) and eluted by an imidazole gradient. The fractions analyzed on SDS-PAGE showed that nickel affinity chromatography yielded FLAG-HIS-TEV-AVR2 with < 70% purity and no significant loss (Figure 2-1A, lane 2).

Since the protein purity is a prerequisite factor in crystallisation experiments, we further purified FLAG-HIS-TEV-AVR2 by size exclusion chromatography (SEC) to remove high molecular weight (HMW) proteins. Protein separation occurs on the basis of molecular size and conformation. Separation is achieved using a porous matrix to which the proteins, for steric reasons, have different degrees of access; smaller molecules have greater access and larger molecules are excluded from the matrix causing HMW proteins to elute first. The chromatogram showed elution of

proteins in decreasing order of size and contained a major symmetrical peak that corresponds to FLAG-HIS-TEV-AVR2 evident when the protein fractions were analysed on the SDS-PAGE (Figure 2-1B, right). Thus, by combining nickel affinity chromatography with size exclusion chromatography, we can purify AVR2 to homogeneity in high yield.

We next confirmed the peptide mass of AVR2 by mass spectrometry (MS). The MS measurement was performed in the Proteomics department, Max Planck Institute for Plant breeding research (MPIPZ), Cologne. The MS profile showed a single peptide peak of 9901.190 Da, close to the expected theoretical mass of 9900.22 Da (Figure 2-2, left).

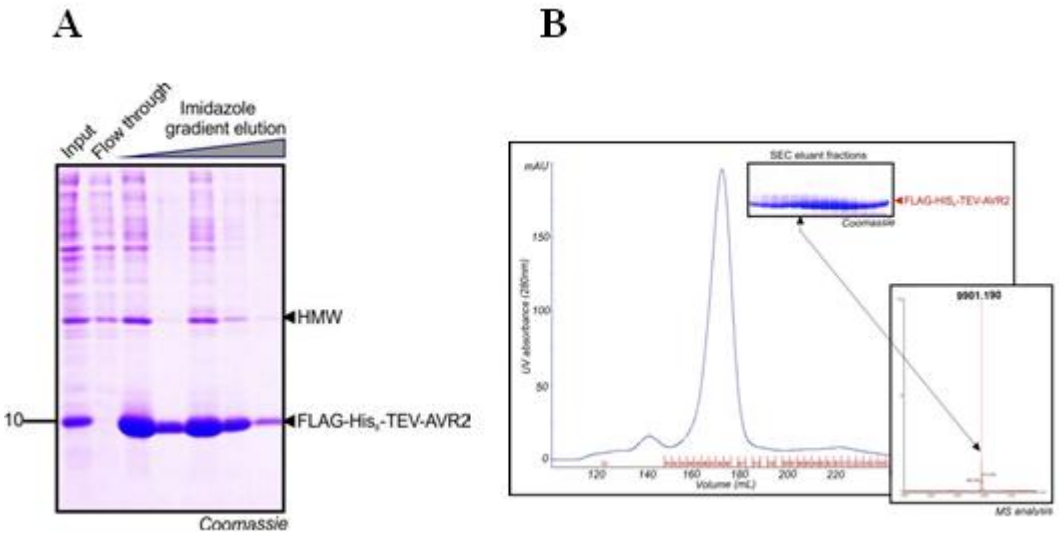


Figure 2-1: Purified FLAG-HIS$_6$-TEV-AVR2 and confirmation by MS analysis

***(A)* Ectopic expression and affinity purification of FLAG-HIS$_6$-TEV-AVR2**. AVR2 with hexa histidine tag at the N-terminus was ectopically expressed in *E.coli*, and purified using Ni^{2+}-NTA agarose matrix by affinity chromatography from the ammonium sulphate precipitated culture medium. Coomassie staining shows the purity of the eluted protein. ***(B)* Chromatogram of size exclusion chromatography (SEC).** Affinity purified AVR2 was separated on Hiload superdex 75pg and eluting proteins were detected at 280 nm. Fractions A3-B12 corresponding to the major peak were analysed on protein gel (upper inset). MS analysis on fraction A7 revealed a single molecular weight (MW) of 9901.190 Da, close to the theoretical MW of

FLAG-HIS_6-TEV-AVR2 (9900.22 Da). The UV absorbance at 280nm is indicated on the y-axis and the elution volume is indicated on the x-axis.

2.2.2 Removal of epitope tags from FLAG-HIS_6-TEV-AVR2

All tags, whether large or small, have the potential to interfere with the biological activity of a protein, impede its crystallisation, or otherwise influence its behaviour. Consequently, it is desirable to remove the purification tag. To remove the fusion tags from FLAG-HIS_6-TEV-AVR2, we used a highly site-specific cysteine protease, the TEV protease from *Tobacco Etch Virus* (TEV). A construct encoding HIS_6-TEV protease was provided by Institute for Bioorganic chemistry (IBCH), Poznan, Poland. The TEV protease recognises the Glu-Asn-Leu-Tyr-Phe-Gln-Gly in the TEV recognition site of FLAG-HIS_6-TEV-AVR2 with high efficiency and cleaves between the Gln and Gly to remove the tags, leaving a single non-native glycine residue on the N-terminus of AVR2 (Figure 2-2A). We produced and purified the HIS_6-tagged TEV protease in *E.coli* to cleave the epitope tags. TEV digestion was performed with the FLAG-HIS_6-TEV-AVR2 for 18 hours at 4˚C followed by nickel- affinity chromatography. The 10 kDa signal in the flow through and wash fractions corresponded to the untagged AVR2 (Figure 2-2B). However, the HIS_6-TEV protease, the epitope tag FLAG-HIS_6-TEV and uncleaved FLAG-HIS_6-TEV-AVR2 remain bound to the nickel column and was later eluted with a higher concentration of imidazole (data not shown). These data demonstrate an efficient and complete removal of the purification tags from AVR2.

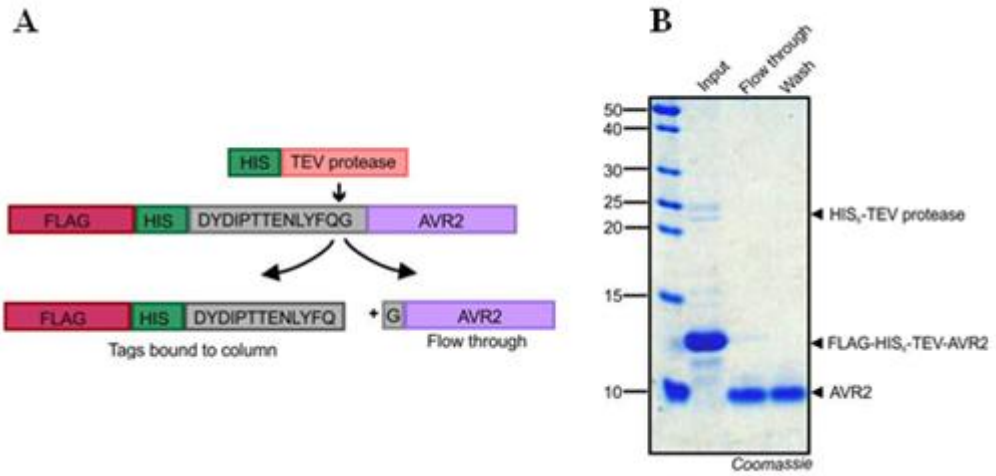


Figure 2-2: TEV digestion of FLAG-HIS$_6$-TEV-AVR2 to remove the epitope tags.

(A) Schematic diagram of cleavage of the epitope tags in recombinant FLAG-HIS$_6$-TEV-AVR2 by recombinant HIS$_6$-TEV protease. The purified HIS$_6$-TEV protease recognises a seven amino acid recognition site (Glu-Asn-Leu-Tyr-Phe-Gln-Gly) in the linker region of FLAG-HIS$_6$-TEV-AVR2. The cleavage occurs between Gln and Gly removing the epitope tags from AVR2. ***(B)*** Cleavage of FLAG-HIS$_6$-TEV-AVR2 by His$_6$-TEV protease. 2.5 mg purified FLAG-HIS$_6$-TEV-AVR2 (input) was incubated for 18 hours at 4˚C in the presence 200 µg of His$_6$-TEV protease. Nickel affinity chromatography was performed with the reaction mixture and the fractions were separated by SDS-PAGE and visualised by coomassie brilliant blue staining. The digested and unbound AVR2 elutes in the flow through and the His$_6$-TEV protease and the epitope tags FLAG-HIS$_6$-TEV were bound to the nickel column.

2.2.3 Purified AVR2 blocks labelling of RCR3 by MV201

To determine if the purified recombinant AVR2 can bind and inhibit RCR3 protease, we applied competitive ABPP using MV201 (Figure 2-3A) (Richau et al, 2012). MV201 is a fluorescent derivative of E-64 that irreversibly reacts with the active site cysteine residue of PLCPs in a mechanism-dependent manner. Protease activity profiling with MV201 was performed with apoplastic fluids (AFs) from *Nicotiana benthamiana* plants agroinfiltrated with Rcr3^{pim} in the presence and/or absence of AVR2. The commercially available PLCP inhibitor E-64 was used as positive inhibition control. AVR2 blocks labelling of RCR3 demonstrating that purified AVR2 inhibits RCR3 (Figure 2-3B).

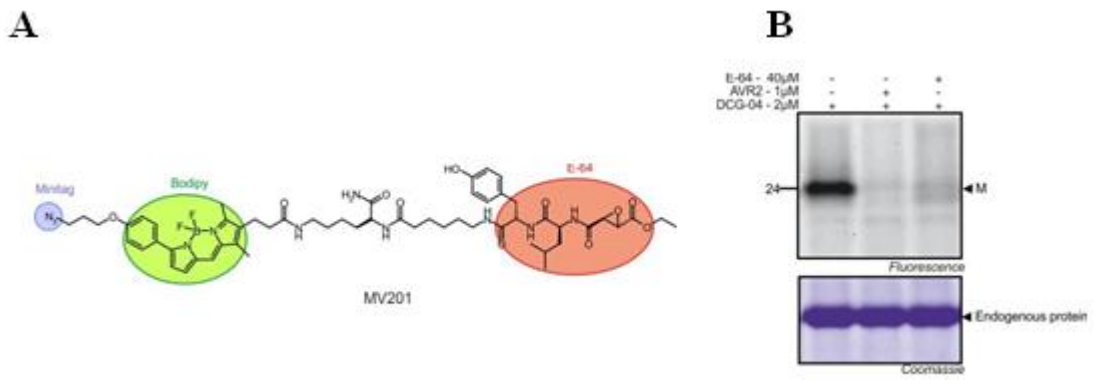


Figure 2-3: Structure of MV201 and inhibition of MV201 labelling by purified AVR2:*(A)* MV201 has an epoxide reactive group (red), azide minitag (blue) and a bodipy reporter tag (green) that enables detection of signals by fluorescent imaging. ***(B)*** Apoplastic fluids of *N. benthamiana* plants transiently expressing RCR3^{pim} AFs were preincubated in the presence (+) or absence (-) AVR2 (1 µM) or E-64 (40 µM) for 30 minutes and labelled with 2 µM MV201 for four hours. The labelled proteins were analysed on protein gels. Detection by fluorescent imaging revealed a 24 kDa signal corresponding to mature (M) RCR3^{pim} in the absence of AVR2 and E-64, whereas in the presence of inhibitors the signal diminished, indicating that both AVR2 and E-64 inhibit RCR3^{pim} activity.

2.2.4 Purified AVR2 inhibits RCR3 and triggers Cf2-dependent HR

To determine if AVR2 produced in bacteria triggers RCR3/Cf2-dependent HR, we infiltrated leaves of Cf2/RCR3, Cf0/RCR3 and Cf2/rcr3-3tomato plants with different concentrations of purified AVR2. Infiltration of AVR2 alone and not water or buffer triggered HR only in Cf-2/RCR3and not in Cf0/RCR3 or Cf2/rcr3-3 tomato leaves (Figure 2-4). We found that 100 nM of AVR2 is sufficient to trigger HR although the onset of the response was observed only after five-day post infiltration (5dpi). At 500 nM and higher concentrations of AVR2, HR was triggered within 1dpi.

We also observed that the lesion area of 500 nM AVR2 was similar to 1000 and 5000 nM AVR2. This data demonstrate that the purified AVR2 is active and existed in a native conformation because it can trigger RCR3 and Cf2-dependent cell death in tomato.

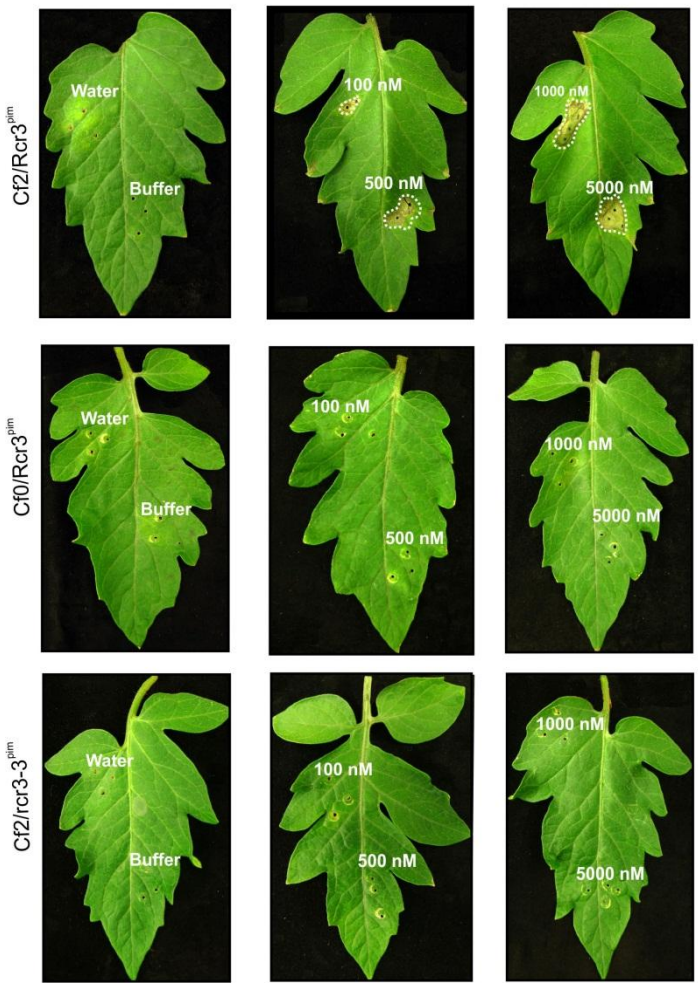


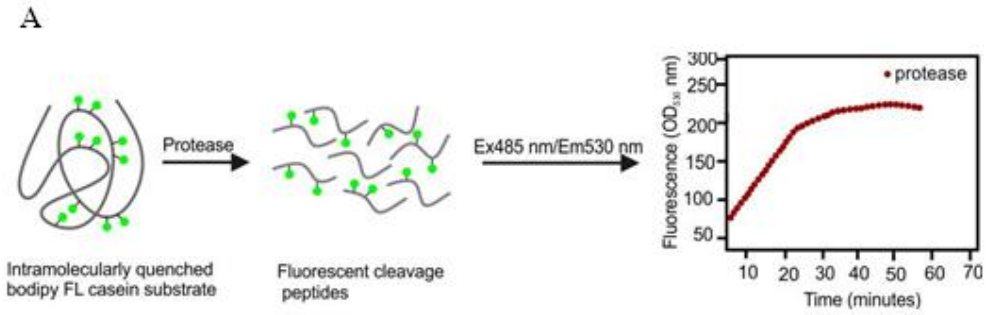
Figure 2-4: Purified AVR2 triggers hypersensitive response

Leaflets of 5-week old Cf-2/RCR3, Cf-0/RCR3 and Cf-2/rcr3-3tomato plants (*Lycopersicon pimpinellifolium*) were infiltrated with purified AVR2 generated in *E.coli* at different concentrations (100, 500, 1000 and 5000 nM). Water and buffer (50 mM Tris-HCl pH 7.5) were infiltrated as negative controls. HR is observed only in Cf2/RCR3 tomato leaves. The infiltrated sectors showing cell death symptoms are outlined and the infiltrated solutions indicated. Leaflets were photographed at 5-day post infiltration.

2.2.5 AVR2 inhibits papain

We anticipated that AVR2 could also interact with papain, a well-studied cysteine protease from the C1A subfamily. Papain is close to RCR3 in the phylogenetic tree of plant PLCPs (Richau et al., 2012) and they presumably also share close homology at

the structural level (Shabab et al., 2008). We studied the kinetics of the interaction between papain and recombinant AVR2 using green fluorescent casein (Bodipy FL-casein) as substrate. Hydrolysis of Bodipy FL-casein releases highly fluorescent Bodipy FL dye labelled peptides (Figure 2-5A). The increase in fluorescence is measured by a fluorescence reader and is proportional to the enzyme activity. In this analysis, we used constant papain and substrate concentrations and varying concentrations of AVR2 (0.1, 1, 25, 100, 250 µM). At increasing concentrations, AVR2 inhibited papain unlike in the absence of the inhibitor. The inhibition observed for 1 µM AVR2 was fairly similar to 25, 100 and 250 µM AVR2 suggesting that nearly all the binding sites of papain is occupied with AVR2 at a concentration of 1 µM (Figure 2-5B top). It has been reported that papain is rapidly inactivated by E-64 (K_i=10 µM). Hence E-64 at different concentrations (40, 250, 500 µM) was included as positive control. All E-64 concentrations completely inhibit the activity of papain (Figure 2-5B bottom).



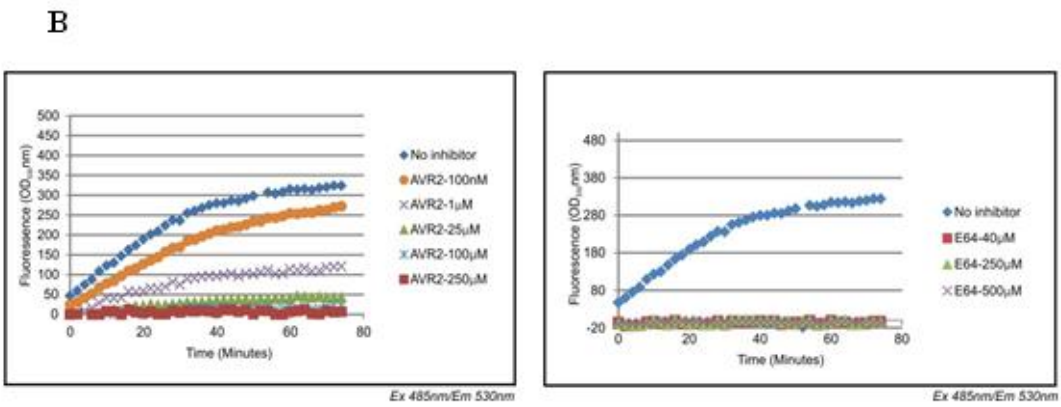


Figure 2-5: AVR2 inhibits papain

(A) Principle of detecting protease activity using bodipy FL-casein (modified from Molecular probes, 2004). Casein labelled with the pH-insensitive green-fluorescent BODIPY FL dyes are hydrolysed in the presence of protease, releasing highly fluorescent BODIPY FL peptides which can be detected at Ex485/Em530 nm by a fluorescence reader. ***(B)*** AVR2 inhibits papain. 427.24 nM of papain was preincubated in sodium acetate buffer pH 5.0 with different concentrations of purified AVR2 (left) or E-64 (right) for 30 minutes. Hydrolysis of 400 nM of Bodipy FL casein substrate was measured every two minutes for 74 minutes at 530 nm. The fluorescence is plotted on the y-axis as fluorescence (OD_{530}nm) and the time interval (minutes) on the x-axis. Each point is an average of two measurements with similar results. The experiment was repeated with similar results.

2.2.6 Secondary structure characterization of AVR2 by Circular Dichroism

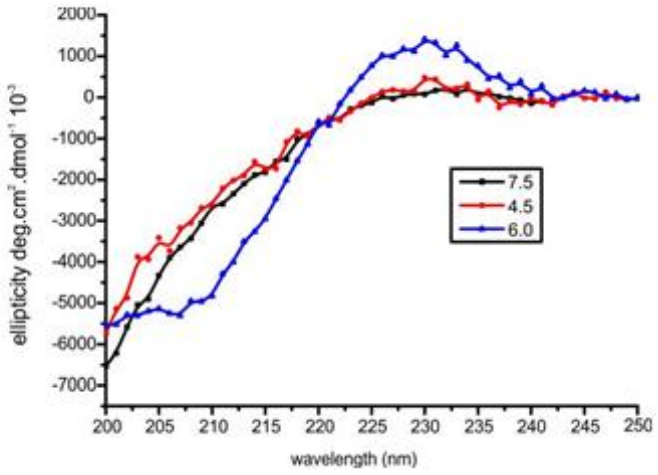
After confirming that the purified AVR2 inhibits RCR3 and papain and also triggers HR in Cf2/RCR3tomato plants, we used a biophysical approach to investigate the secondary structure of purified AVR2. This work was done in collaboration with Dr. Mohammed Shabab, Department of Biorganic chemistry, Max Planck Institute for Chemical Ecology. It has been previously reported that AVR2 inhibits RCR3 at pH 4.5-5.5 but not at pH 6.0 or higher (Rooney et al., 2005). Here we investigated if this pH-sensitivity is intrinsic to the AVR2 structure.

Circular dichroism (CD) spectroscopy measures differences in the absorption of left-handed polarized light versus right-handed polarized light, which arises due to structural asymmetry. The absence of regular structure results in zero CD intensity, while an ordered structure results in a spectrum which can contain both positive and negative signals. The spectrum can be interpreted by the CONTINLL software of CDpro package for quantitative estimation of structural elements.

We used CD spectroscopy (CD) in the far-UV spectral region (190-250 nm) to study changes in the secondary structure of AVR2 without epitope tag at pH 4.5, 6.0 and 7.5. The CD analysis predicts that AVR2 predominantly contains beta sheets. Interestingly, overlay of the CD spectra at different pH shows a loss of more than one third of the beta sheet structure of AVR2 at pH 7.5 and 6.0 compared to pH 4.5 (Figure 2-7A, B). Thus, CD analysis predicts a conformational change in AVR2 at

acidic pH, resulting in increased beta sheet structure associated with the interaction with RCR3.

A



B

	% conformational		
Structural conformation	pH 4.5	pH 6.0	pH 7.5
Alpha-helix	4.3	4.1	4.3
Beta-sheet	63.8	37.3	37.3
Turns	22.1	21.9	21.2
Unordered structure	34.1	36.3	37.2

Figure 2-6: Beta sheet structure of AVR2 is pH sensitive.

(A) CD spectra of conformational change in AVR2 at different pH (10 mM acetate buffer pH 4.5, or 10 mM phosphate buffer pH 7.5) and low salt (10 mM sodium chloride). All solutions contained 122 µg/ml AVR2 without tags. The ellipticity per residue ($deg.cm^2.dmol^{-1}10^{-3}$) was plotted on the y-axis against wavelength (nm) indicated on the x-axis. ***(B)*** Secondary structure of AVR2 at different pH predicted by CONTINLL program of CDPro software. The predominant conformational change in beta-sheet with the change in pH is highlighted in grey.

2.2.7 Tyrosine fluorescence spectroscopy confirms conformational change in AVR2.

AVR2 has two aromatic amino acid residues (Tyr^{21}) and (Phe^{54}) in mature AVR2 peptide. To confirm the conformational change exhibited by AVR2 at different pH, we used Tyr fluorescence spectroscopy. Phenylalanine was not preferred due to its weak fluorescent property. In the native folded state, tyrosine has a low quantum yield as this residue is generally located within the core of the protein, whereas in a partially folded or unfolded state, tyrosines become exposed to solvent which increases the fluorescence.

In our experiment, the intrinsic fluorescence of Tyr^{21} was monitored at different pH at 274 nm excitation wavelength. The emission spectra show the sum of two peaks: The peak at 303 nm is the expected peak for Tyr fluorescence. This peak increases at increasing pH, suggesting that Tyr becomes solvent exposed at higher pH. A second, broader peak at 330-350 nm increases from pH 4.5 to 6.5, but then decreases at pH 7.5.

This peak is possibly caused by deprotonated or hydrogen-bonded Tyr. Thus Tyr fluorescence spectroscopy data show conformational changes in AVR2 at different pH similar to CD data.

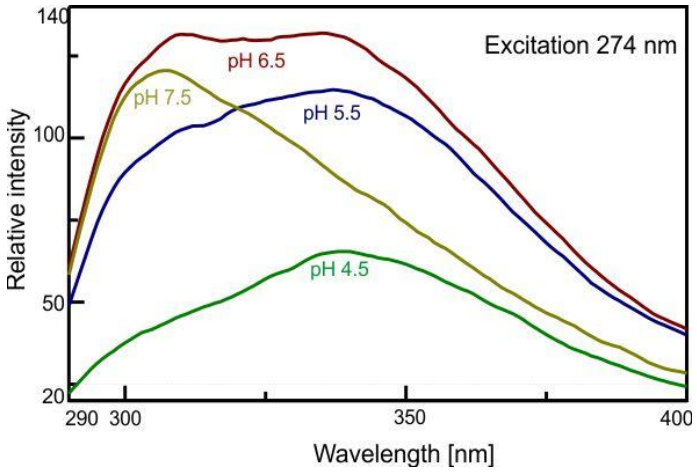


Figure 2-7: pH-dependent conformational change of AVR2.

Fluorescence spectra of AVR2 (122 µg/ml) incubated at different pH; 4.5 (green), 5.5 (blue), 7.5 (sap green) and 6.5 (maroon). The tyrosine emission scans at excitation 274 nm was monitored for change in intrinsic fluorescence. The relative fluorescence intensity is plotted against the wavelengths (nm).

2.2.8 Crystallisation screens of AVR2

Any X-ray structure analysis has to overcome two main obstacles. First, to provide sufficient amounts of pure protein and second to obtain well-ordered three-dimensional crystals. The crystallographic work was done in collaboration with Dr Rolf Rose and Dr. Christian Ottmann, Chemical Genomics Centre, Dortmund.

The quantity of AVR2 was not a limiting factor, so we screened 3840 conditions for obtaining crystals of AVR2 with and without the epitope tags. These screening conditions with the grid screens systematically evaluate different factors, such as a precipitant (salts, polymers and solvents); various AVR2 concentrations and various pH values. Our crystallisation attempts of different AVR2 concentrations with and without the epitope tags had one of the following outcomes: (i) clear drops and no crystals (Figure 2-8A); (ii) phase separation (Figure 2-8B); (iii) amorphous precipitate (Figure 2-8C); and (iv) large salt crystals (Figure 2-8D). No protein crystals were detected in any of these conditions. These data indicate that AVR2 did not reach the nucleation phase as it remained soluble, irrespective of very high concentrations. This high solubility of AVR2 is thought to be correlated with the high lysine content (ten lysines) in the protein. Next we modified the lysines of AVR2 by reductive methylation where the large charged lysine residues are masked with two methyl groups (Walter et al., 2006). The reductive methylation process was performed with 25 mg/ml FLAG-HIS_6-TEV-AVR2 protein. The reaction lead to a significant amount of precipitated protein and the final recovery was only 11mg/ml after purification by size exclusion chromatography.

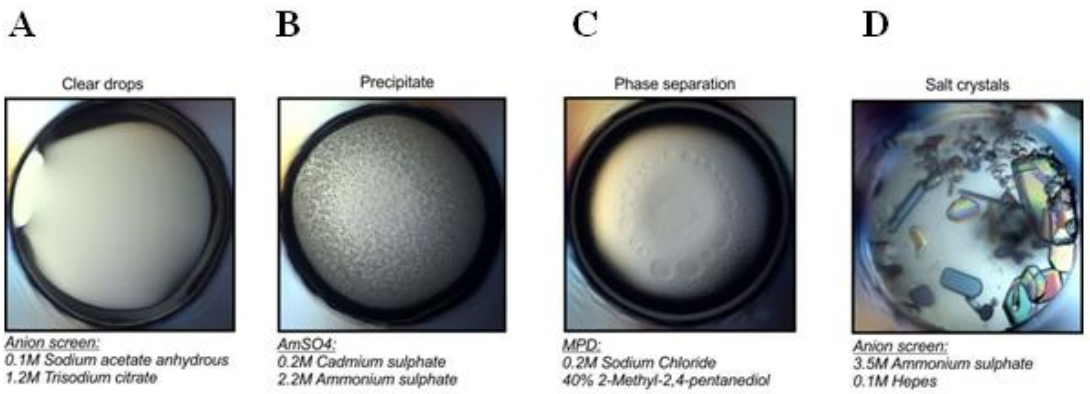


Figure 2-8: Crystallisation screens of AVR2

Images of crystallization experiments for the four different screening conditions of AVR2. 3840 different conditions were tested (various pH conditions and precipitants at 18˚C and some at 4 ˚C) in the presence and absence of epitope tags in AVR2. Clear drops ***(A)***, precipitates ***(B)***, phase separation ***(C)*** and sometimes salt crystals ***(D)*** were observed. Specific conditions for the selected picture are shown on the bottom.

Mass spectrometry (MS) analysis was performed with the pre- and post-reaction samples of the reductive methylation process to evaluate the extent of lysine methylation. The MS data revealed a major peak at 10322.172 Da (Figure- 2-9C) which is 420,982 Da more than non-methylated AVR2 (9901.190 Da) (Figure- 2-9D). This mass increase corresponds to 30 methylation events (each methyl group is 14 Da) per AVR2 protein consistent with the presence of one N-terminus and 14 lysine residues (Figure- 2-9A) in FLAG-HIS_6-TEV-AVR2. Enlargement of the major peak shows a series of seven peaks with 14 Da distance. The first peak at 10237.906 Da indicates that AVR2 accepted at least 24 methylation events (Figure- 2-9D). To investigate whether reductive methylation affected AVR2 function, we infiltrated methylated FLAG-HIS_6-TEV-AVR2 and a control FLAG-HIS_6-TEV-AVR2 into Cf2/RCR3 tomato leaves. Cell death phenotype was observed in both the methylated and control samples indicating that the lysine-methylated AVR2 was still recognised by RCR3, triggering the Cf2-mediated defense response (Figure 2-9E). This methylated FLAG-HIS_6-TEV-AVR2 protein was tested in crystallization screens but

no crystals were detected. The summary of different crystallization screens of AVR2 is listed in Table 2.

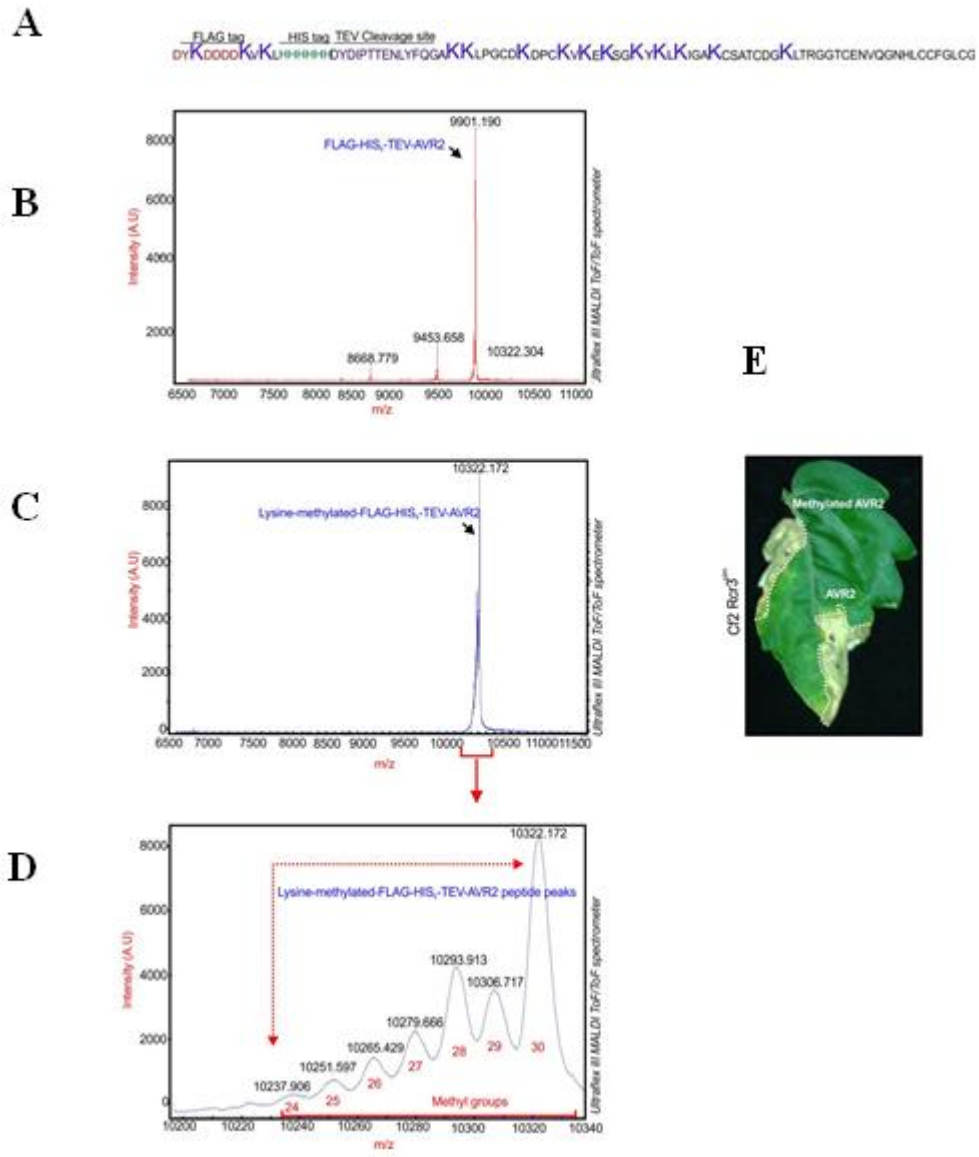


Figure 2-9: MS analysis and HR assay of Lys-met AVR2

(A)Sequence of FLAG-HIS_6-TEV-AVR2 showing the 13 lysines in blue, bold and the epitope tags in color: FLAG (red), HIS_6 (green) and TEV recognition site (magenta) ***(B)***MS analysis of the non-methylated sample revealed a MW of 9901.190 Da. ***(C)***MS analysis of the methylated sample revealed a major peak with MW of 10322.172 Da. ***(D)***Closer view of methylated FLAG-HIS_6-TEV-AVR2 show different peptide peaks with an increase of 14 Da. The numbers indicate the number of methylation events (red) ***(E)*** Cf2-dependent HR response triggered by methylated FLAG-HIS_6-TEV-AVR2. Leaflets of 5-week-old Cf-2/RCR3 tomato were infiltrated with 1 µM

methylated or unmethylated FLAG-HIS_6-TEV-AVR2. Leaves were photographed five days post infiltration.

Table 2

Summary of crystallization screens

Protein[1]	Concentration[2]	Conditions[3]	Crystallization screens[4]
AVR2	15	384	JCSG I, JCSG II, JCSG III, JCSG IV
AVR2	60	480	JCSG II, JCSG plus, MPD, Classic CRYOS
FLAG-HIS_6-TEV-AVR2	9, 8, 36	2592	JCSG I, JCSG II, JCSG III, JCSG IV MPD, Classic, AmSO4, Cation, Anion
Methylated FLAG-HIS_6-TEV-AVR2	11	384	JCSG I, JCSG II, JCSG III, JCSG IV

[1]AVR2 with and without the epitope tags or methylated; [2] Protein concentration in mg/mL; [3]Number of screening conditions tested for each protein concentration; [4]Names of ready- made crystallization screens from Qiagen used for screening.

2.2.9 Nuclear Magnetic Resonance (NMR) of AVR2

Efforts to crystallize AVR2 without the protease were not successful. Hence we utilized NMR spectroscopy to determine the structure of AVR2 and also study pH-dependent structural changes in AVR2. The NMR work was done in collaboration with Professor Dr. Marcellus Ubbink, Protein chemistry section, Leiden Institute of Chemistry, Leiden University.

In principle, NMR detects the spinning of electrically charged atomic nuclei in the macromolecule or protein with a static magnetic field. This magnetic field makes the spin-states of the atomic nuclei vary in energy level and NMR is used to measure this transition in energy at different spin state. Common NMR active nuclei are 1H, 13C, 31P, 15N and 29Si. An initial measurement of unlabelled untagged AVR2 with proton NMR suggested that AVR2 has a rigid structure (data not shown). Next AVR2 protein was produced as 15N- labelled form to make a Heteronuclear Single Quantum Coherence (HSQC) spectrum of AVR2. IH-15N-HSQC displays N-H bonds that are found in the peptide backbone, and in Lys, Arg, Trp, His, Asn, Gln residues. AVR2

contains 58 N-H groups in the protein backbone, one *N*-terminal amino group, 11 lysines, one arginine, ten glycine, three alanine, five leucine, two proline, eight cysteine, three aspartic acid, two valine, two glutamic acid, two serine, one tyrosine, one isoleucine, three threonine, two asparagine, one glutamine and one histidine; thus 59 different N-H bonds. The peaks in the 1H-15N HSQC spectrum disperse well and the number is similar to the expected number of N-H bonds. According to NMR experts this structure is solvable by NMR. Labelling of AVR2 with 13C to generate the 1H-13C spectrum of AVR2 is underway.

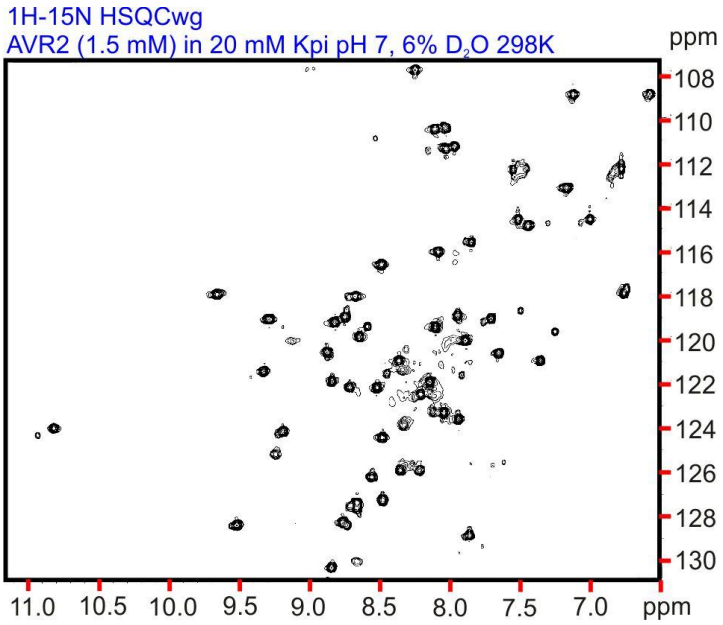


Figure 2-10: HSQC spectrum of AVR2. 1H-15N NMR spectrum between 11.0 – 7.0 ppm of AVR2 labelled with 15N isotope. The chemical shift represented as ppm in the axis. Each spot or peak in the spectrum represents each amide in AVR2 protein.

2.3 HETEROLOGOUS EXPRESSION OF RCR3 IN DIFFERENT EXPRESSION SYSTEMS

2.3.1 Expression in plants (*Nicotiana benthamiana)*

Transient expression of genes through infiltration of *Agrobacterium tumefaciens* (agroinfiltration) is a quick and easy method to study genes of interest (Kapila et al., 1997; Rossi et al., 1993). The advantage of this expression system is that *N. benthamiana* is a close relative of tomato, and has been reported that RCR3 produced in this system is active and inhibited by AVR2 (Rooney et al., 2005). In addition, apoplastic fluids (AFs) contain only few endogenous plant proteins, which eases the purification procedure. Therefore, we tested RCR3 expression in plants by transient over expression of RCR3 through agroinfiltration of *Nicotiana benthamiana.*

N. benthamiana leaves of 4-5 week old plants were infiltrated with a mixture in a 1:1 ratio of *A. tumefaciens* cultures carrying the binary plasmid containing either $RCR3^{pim}$or *P19* genes (Song et al., 2009). *P19* expresses a viral suppressor protein that prevents the onset of post-transcriptional gene silencing (PTGS) in the infiltrated tissue and allows high level of transient protein expression (Voinnet et al., 2000). As negative control, a few leaves were infiltrated only with *P19* expressing *A. tumefaciens* cultures. The AFs of the adult leaves were extracted in water at five days-post-infiltration (5dpi) and the expression of RCR3 was detected and analysed on protein blots using the RCR3 antibody. A 24-kDa signal corresponding to mature RCR3 was detected in the AFs from the leaves agroinfiltrated with $RCR3^{pim}$ and P19 but not in the leaves agroinfiltrated with P19 alone (Figure 3-1A). Comparison of RCR3 detected in the AFs with a known quantity of RCR3 expressed in *E.coli* showed that RCR3 quantity was less than 3 mg/ml in the AFs. Thus, to produce 30 mg RCR3 needed for crystallisation, we would need to isolate AFs from at least 300 leaves. These data indicate that RCR3 production by agroinfiltration is not feasible because of the inadequate quantity of RCR3 protein production.

2.3.2 Expression in bacteria (*Escherichia coli)*

As a second expression system, we tested a secretion system based on plasmid pFLAG-ATS in *E.coli*. Because RCR3 contains three putative disulphide bridges, the *E.coli* secretion system is attractive for the production of secreted eukaryotic proteins because the periplasm is an oxidising environment and contains enzymes facilitating the formation and rearrangement of disulphide bonds (Missiakas and Raina, 1997). Bacteria carrying pFLAG-ATS encoding a secreted *N*-terminal FLAG-HIS_6-TEV-tagged RCR3 without prodomain (pSK 018.01) were grown and induced with IPTG at different temperatures. Sample aliquots of RCR3 protease expressed in medium, soluble and insoluble cell fractions after IPTG induction were analyzed by western blotting using RCR3 antibody.

This analysis revealed that RCR3 cannot be detected in the medium irrespective of the induction time and temperature (Figure 3-1B, top). In contrast, large RCR3 quantities were detected in the supernatant of the Insoluble Cell Fraction (ICF) but only when the lysed bacterial cells were dissolved in 8M urea (Figure 3-1B, bottom). Less than 1.5 mg/L RCR3 was recovered in the soluble cell fraction (SCF) from the lysed bacterial cells without urea treatment (Figure 3-1B, middle). Thus, despite the secretory expression system tested under various induction conditions, RCR3 was found to accumulate mostly as insoluble protein inside the cells, presumably in inclusion bodies.

2.3.2.1 Refolding of RCR3 produced in E.coli

We next tested if we could refold RCR3 from the insoluble cell fraction. Refolding is a critical step for proteases with disulphide bridges. The 8M urea fraction containing RCR3 was diluted 100-fold into the refolding buffer containing arginine to increase the polarity and refolding yield, and EDTA to chelate metal ions (D'Alessio K et al., 1999; Hwang and Chung, 2002; Kopitar et al., 1996; Tobbell et al., 2002). In addition, a 10:1 ratio of reduced and oxidised glutathione in the refolding buffer facilitates efficient refolding and disulphide formation. This refolding procedure was performed by dialysis at 4˚C for 20 hours. The refolded protease was tested for activity by a casein substrate assay after removing EDTA, arginine and the redox components by dialysis. In contrast to RCR3 produced by agroinfiltration, RCR3 treated with the refolding buffer did not degrade casein, indicating that the refolding of RCR3 failed (data not shown). We next tested a gradient refolding procedure

by dialyzing the protease in refolding buffer in which the urea concentration was gradually reduced from 8M, 6M, 4M, 2M and 0M. Proteolytic analysis revealed that the activity of RCR3 was still 70 fold less in comparison to the RCR3 expressed in the AFs by agroinfiltration (data not shown).

In addition to RCR3, also PIP1 and C14 were difficult to express in the bacterial system. Furthermore, prodomain containing RCR3 (pSK 019.01) and the catalytic RCR3 mutant (C25A) with and without the prodomain (pSK 026.01 and pSK 025.01) respectively were also tested in *E.coli* but the levels of protein expressions were not different from wild-type RCR3 (pSK 018.01). These data indicate that heterologous expression of PLCPs in *E.coli* failed to produce soluble proteases, even though sufficient protein was produced.

2.3.3 Expression in insects (*Spodoptera frugiperda)*

As a third expression system, we tested the baculovirus-mediated expression system. This expression system has been used for the production of several mammalian proteases, including cathepsins and other papain-like cysteine proteases. The expression of proRCR3 in insect cells by the baculo virus system was done in collaboration with Professor Lukas Mach, University of Natural Resources and Life Sciences (BOKU), Vienna.

The transfer vector (Pvt-Bac-His_1) derived from *Autographa californica* Multiple Nuclear Polyhedrosis Virus (AcMNPV) was used for gene expression of *proRCR3* in insect cells with a co-transfection protocol to initiate the homologous recombination of the transfer vector with the viral genome in the insect cells. The transfection of the proRCR3 into viral genome was confirmed by microscopic examination where the infected cells appeared larger than the uninfected cells and had enlarged nuclei. After expression, proRCR3 in the supernatant was purified using Ni^{2+} affinity matrix. The 37 kDa signal on the coomassie stained protein gel indicated proRCR3 expression. However proRCR3 was less than 3 mg/L (Figure 3-1C). These data indicate that this expression system is not suitable to generate large quantities of proRCR3 protease.

2.3.4 Expression in yeast (*Pichia pastoris):*

Finally, the *Pichia pastoris* expression system was implemented to overcome the disadvantages of the other three expression systems mentioned in this chapter and generate high yield of the proRCR3 protein. This work was done in collaboration with Dr. Boquian Wu, Carlsberg laboratory, Denmark. Yeast systems have been extensively used to express papain-like cysteine proteases (Bromme et al., 2004). RCR3 was also expressed in *Pichia pastoris* though in low amounts (Rooney et al., 2005). A major advantage of *P. pastoris* is its low level of secretion of endogenous proteins. The heterologously expressed protein is secreted and therefore the only major protein in the medium, which dramatically simplifies the purification procedure. Similar to the baculovirus-mediated expression in insect cells, the yeast expression system also allows post-translational modifications such as disulphide formation, proteolytic processing, and glycosylation.

The methanol-mediated induction initiated the heterologous expression in the yeast PichiaPink^{TM} strain 4 (Invitrogen), which resulted in the production of proRCR3 with a C-terminal His_6 epitope tag. To achieve high protein expression, we maintained the pH of the culture neutral or alkaline because acidic conditions can sometimes inititate the activation of papain-like cysteine proteases. Following expression, Nickel (Ni^{2+}) affinity chromatography was implemented to purify proRCR3 from the crude extract. This purified sample was analysed by coomassie staining of the protein gel. A high amount of a 37 kDa signal was detected indicating a relatively high yield and purity of the secreted proRCR3 (Figure 3-1D).

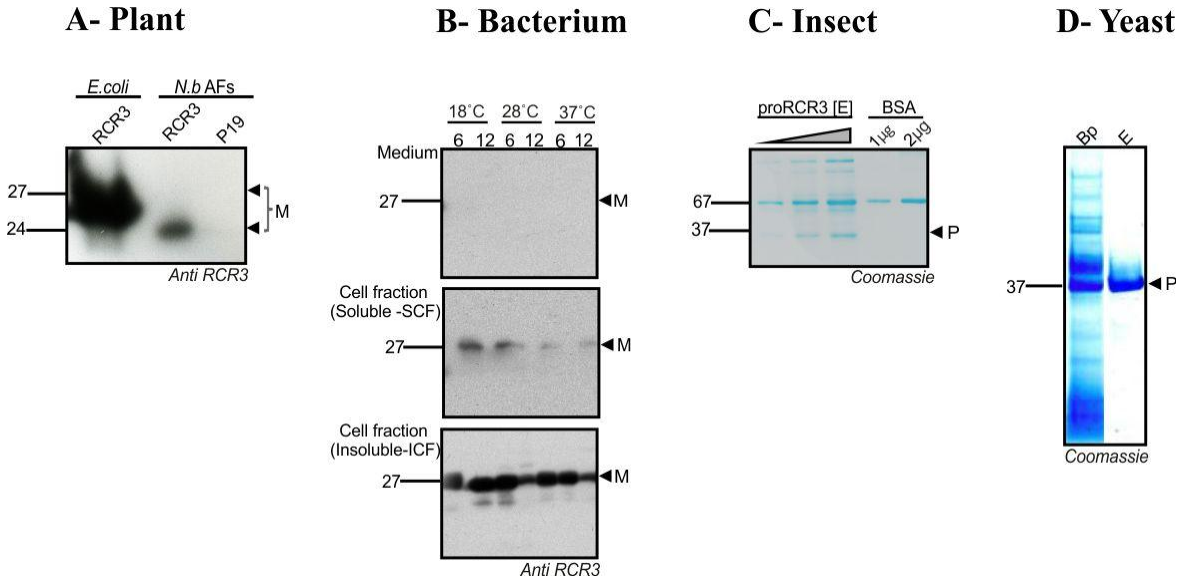


Figure 3-1: Expression of RCR3 in plants, bacteria, insects and yeast.

***(A)* RCR3 production in plants**:*Nicotiana benthamiana* plants were agroinfiltrated with bacterium construct carrying binary plasmids encoding *RCR3* and *P19* in the ratio 1:1. Apoplastic fluids (AFs) of RCR3 were extracted from the agroinfiltrated leaves at five-day post infiltration. Both signals (27 and 24 kDa) reveal mature RCR3 (M). The size difference is due to the purification tags on RCR3 expressed in *E.coli*. The RCR3 protein in the AFs is insufficient for large protein extraction although it is less contaminated with the endogeneous proteins. ***(B)*RCR3 production in bacteria**: *E.coli* cells expressing RCR3 without prodomain (pSK 018.01) were induced with IPTG and grown at different temperatures (18˚C, 28˚C and 37˚C) and time points (6 and 12 hours). Proteins were extracted from the medium, soluble cell fraction (SCF) and insoluble cell fraction (ICF). The western blot analysis using anti-RCR3 shows no protein secretion into the medium (top) and high amount of protein in the insoluble cell fraction (ICF, bottom) and a low amount in the soluble cell fraction (SCF, middle). ***(C)* RCR3 production in insects**:The baculovirus transfer vector carrying *proRCR3* was transfected into 2×10^6 insect cells. The proRCR3 (P) was expressed in the supernatant of the culture medium. Coomassie protein gel analysis of proRCR3 elutions (E) after Ni^{2+} affinity shows 37 kDa secreted protein. However, protein levels are insufficient in comparison to 1 and 2µg of BSA used for quantitation. ***(D)* RCR3 production in yeast:** Expression of *P. pastoris* strain carrying the transfer vector expressing proRCR3. The 37 kDa signal corresponding to the secreted proRCR3 (P) in crude medium (Bp) on the left lane of the gel was purified by Ni^{2+} affinity purification and right lane is pooled eluant (E) containing purified proRCR3.

Table 3

Summary of RCR3 levels in different expression systems

	Yield (mg/L)[2]	
Expression system[1]	Soluble RCR3[a]	Insoluble RCR3[b]
Plant (*Nicotiana benthamiana*)	<3 (AFs)	
Bacterium (*Escherichia coli*)	<1 (medium), <1.5 (SCF)	<10 (ICF)
Insect (*Spodoptera frugiperda*)	<3	
Yeast (*Pichia pastoris*)	30-50	

[1]Different expression systems for RCR3 protein expression; [2] Yield of RCR3 protein expressed (mg/L); RCR3 expressed in [a] soluble fraction (medium and soluble cell fraction) and [b] insoluble fraction. Detection of protein expression based on the western blot analysis using anti-FLAG and anti-RCR3 antibody. Abbreviations: Apoplastic Fluid (AFs), Soluble Cell Fraction (SCF), Insoluble Cell Fraction (ICF).

2.3.4.1 Putative Glycosylation sites of the expressed RCR3 in *Pichia*

O- and *N*-linked glycosylations are common post-translational modifications performed by *P. pastoris* (Eckart et al., 1996). Pichia can glycosylate heterologous proteins, even when those proteins are not normally glycosylated. Also when the protein is glycosylated in the native host, Pichia may not glycosylate it on the same positions (serine or threonine residues). We used NetGly 3.1 server software to predict potential glycosylation sites in proRCR3. The prediction results suggested 46 putative *O*-glycosylation sites and one putative *N*-glycosylation site in the proRCR3. However, the *N*-glycosylation site is present in the prodomain of proRCR3, which will be eventually removed during the activation process (Figure 3-2A). The apparent MW of 37 kDa for RCR3 produced in *P. pastoris* suggest no significant glycosylation since this is close to the expected molecular weight of native protein (37 kDa). Furthermore co-infiltration of proRCR3 together with AVR2 triggers HR only in Cf2-RCR3 and not Cf0-RCR3 plants, indicating that RCR3 is functional. MS analysis is required to confirm the exact mass.

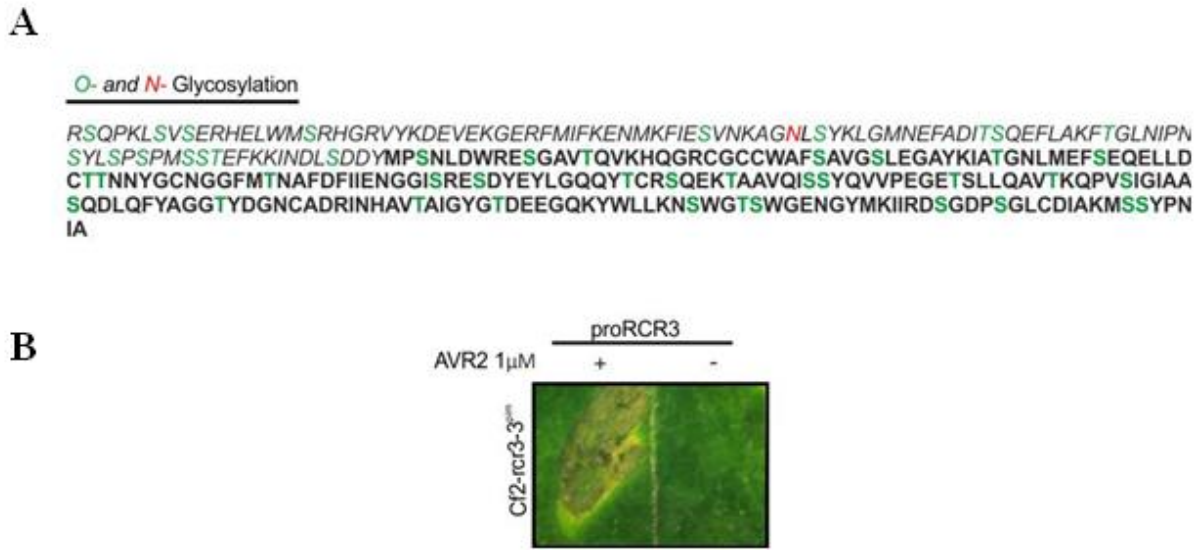


Figure 3-2: Prediction results of putative *O*- and *N*- Glycosylation sites in RCR3 and RCR3 is active

(A)RCR3 protease sequence; prodomain (italics) and protease domain (bold). Putative *O*-glycosylation (predicted by NetOGlyc 3.1 Server) sites are shown in green and putative *N*-glycosylation (predicted by NetNGlyc 3.1 Server) site is indicated in red. ***(B)*** RCR3 produced in Pichia triggers HR when co-infiltrated with AVR2 in Cf2-rcr3-3^{pim}plants. Fully expanded leaves of 5-week old Cf2-rcr3-3^{pim}plants were infiltrated with proRCR3, with and without 1

µM AVR2. HR is observed only in the proRCR3 co-infiltrated with AVR2. Leaves were photographed at five-days post infiltration.

2.3.4.2 Maturation of RCR3 produced in *Pichia pastoris*

The propeptide region or prodomain must be removed to generate proteolytically active RCR3. Several parameters (reducing agents, pH conditions and incubation time) needed to be tested for reproducible processing. ProRCR3 was incubated at 25˚C in different reducing agents (DTT and L-cysteine) and at different pH (pH 5.0 and 8.0). RCR3 conversion was analyzed at various timepoints (30, 90 and 300 minutes) using protein blot with RCR3 antibody.

At 300 minutes, 24 kDa mature RCR3 was detected, when incubated at low pH in the presence of reducing agents (Figure 3-2A, red, lanes 2-4). The 37 kDa proRCR3 even disappeared at 300 minutes, indicating that the conversion to mature RCR3 is complete. At 90 minutes, proRCR3 was partially converted mature RCR3, and no mature RCR3 was detected after 30 minutes. The conversion to mature RCR3 is absent when proRCR3 is incubated at pH 8.0 with reducing agent and also when incubated at pH 5.0 without reducing agent suggesting that both the reducing agent and acidic pH are required for the maturation.

Enzymatic assays with Z-Leu-Arg-AMC as substrate was used to monitor the activation of the protease during the maturation process. ProRCR3 in the presence or absence of DTT, acidic pH buffer (pH 5.0) or alkaline pH buffer (pH 8.0) was treated with 6.66 µM Z-Leu-Arg-AMC and the conversion of proRCR3 to mature RCR3 was immediately monitored by measuring protease activity at 460 nm. The proRCR3 treated with both DTT and acidic pH buffer was activated quickly (Figure 3-2B). In contrast no RCR3 activity was detected if RCR3 was incubated at pH 5.0 without DTT or at pH 8.0 with DTT (Figure 3-2B). The protein blot analysis also confirms the observation (Figure 3-2B, inset). These data confirm that both DTT and acidic pH buffer (pH 5.0) are required and sufficient for the activation of proRCR3 into mature RCR3.

Having established the conditions for activation to mature RCR3, competitive ABPP was implemented to determine if the mature RCR3 can be inhibited by AVR2. Labelling of mature RCR3 with 2 µM MV201 (bodipy derivative of E-64) in the presence of 1 mM DTT

and sodium acetate pH 5.0 reveals a strong 24 kDa fluorescent signal confirming that mature RCR3 is proteolytically active (Figure 3-2C). Importantly, labelling of RCR3 is suppressed upon preincubation with AVR2 or E-64 (Figure 3-2C). Thus RCR3 produced in yeast is active and interacts with AVR2. Suprisingly, when proRCR3 was preincubated with E-64 or AVR2 in the presence of reducing agent and acidic pH buffer at 25˚C, the maturation of the proRCR3 was blocked (Figure 3-2D). The activation mechanism of proRCR3 is discussed (Chapter 1 section 1.3).

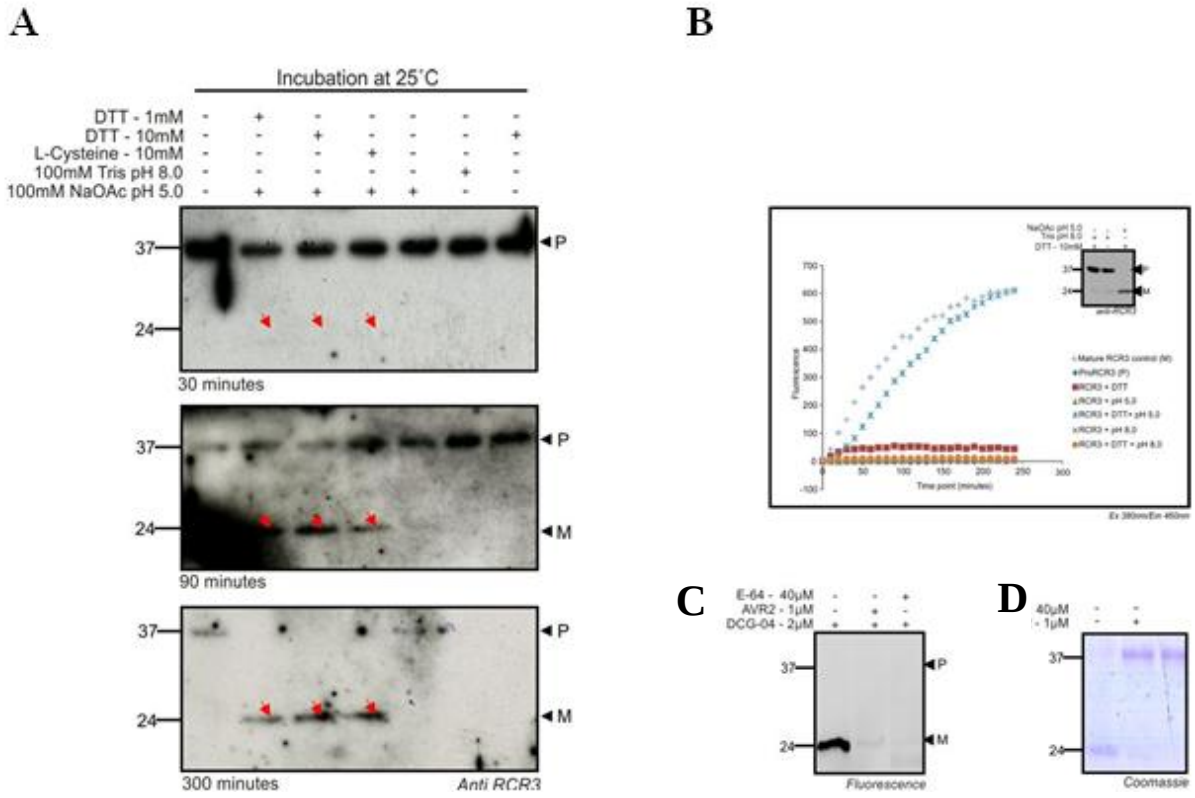


Figure 3-3: Maturation of proRCR3 and activity of mature RCR3

***(A)*Acidic pH and reducing agent promote maturation/activation of RCR3 *in vitro*;** ProRCR3 (P) was incubated with different reducing agents and with different pH buffers at 25˚C. The protein sample aliquots were taken at different time points (30, 90 and 300 minutes) to analyse the conversion into mature RCR3 (M), (red arrow heads) by protein blot using RCR3 antibody. ***(B)* Protease activity monitors the maturation of RCR3.** ProRCR3 (P) was incubated with or without 10 mM DTT in the presence of Tris pH 8.0 or sodium acetate pH 5.0. 6.66 µM Z-Leu-Arg-AMC (Z: N-carbobenzyloxy-7-Amino-4-methylcoumarin) was included to monitor the activity of RCR3 by the hydrolysis of the substrate measured at Ex 360 nm/Em 490 nm. Protein blot using anti-RCR3 antibody shows maturation of proRCR3 only in the presence of DTT and NaOAc pH 5.0 buffer (300 minutes)

(C) **E-64 and AVR2 inhibit RCR3.** Mature RCR3 (M) produced in yeast preincubated with or without 1 µM AVR2 or 40 µM E-64 for 30 minutes in the presence of 1 mM DTT and 25 mM sodium acetate pH 5.0 and labelled by MV201 for 4 hours. Labelled proteins were detected from protein gels using fluorescent scanning. ***(D)*** **Maturation of proRCR3 can be blocked by AVR2 or E-64**. The proRCR3 produced in yeast was incubated in the presence or absence of 1 µM AVR2 or 40 µM E-64 in 10 mM DTT and 100 mM sodium acetate pH 5.0 at 25˚C. Samples were analysed by coomassie protein gel after 90 minutes.

2.4 ROLE OF CYS 24 IN THE ACTIVE SITE OF RCR3

2.4.1 A double Cysteine (Cys) in the active site is common in PLCP subclass 6

Nearly all PLCPs possess the catalytic triad (Cys-His-Asn) of which Cys25 (papain numbering) is the main residue responsible for the proteolytic activity of the enzyme. Plant PLCPs are subdivided into nine subfamilies (Richau et al., 2012) (Figure 4-1A). RCR3 and PIP1 belong to subfamily 6 together with senescence-associated PLCP SAG12 from *Arabidopsis thaliana*. Interestingly, SAG12-like PLCPs are typified by carrying an additional Cys residue (Cys24) adjacent to the catalytic cysteine (Cys25). This double Cys occurs in 85% of this subfamily but hardly in other families (Figure 4-1B). Also RCR3 carries a double Cys in the active site. The protease domain of RCR3 contains eight cysteine residues of which six cysteines (Cys22, Cys63, Cys96, Cys151 and Cys204) form three putative disulphide bridges. Since Cys25 is the catalytic cysteine, Cys24 must be an unpaired cysteine. A model of RCR3 using papain (9PAP) as a template indicates that both Cys24 and the catalytic Cys25 residue are located in the substrate binding groove (Figure 4-1C).

These observations lead to the hypothesis that Cys24 might have a role in RCR3. To investigate the role of Cys24 in RCR3, we generated several mutants of RCR3 (C24A, C25A and C24AC25A) by substituting the cysteine residues into alanine because alanine is electrostatically neutral (Figure 4-1D).

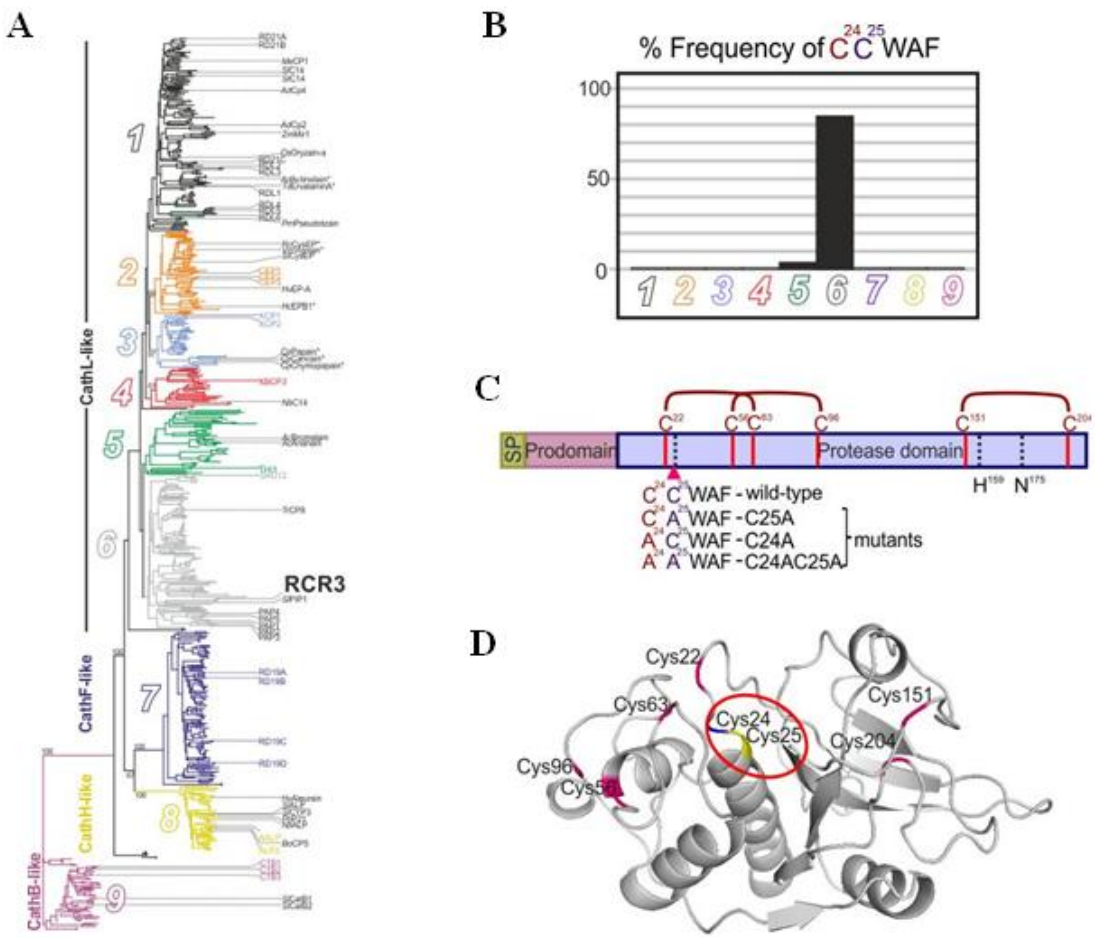


Figure 4-1: Occurrence of double Cysteine in RCR3 and other plant PLCPs.

(A) Phylogenetic tree analysis subclassifies 723 plant PLCPs into nine subfamilies (adapted and modified Richau et al., 2011). RCR3 belongs to the subfamily 6 (grey). ***(B)***Frequency of the double Cys in the active site of different PLCP subfamilies observed in subclass 5 and predominantly (more than 85%) in the subclass 6. ***(C)***Schematic diagram of RCR3 mutants. The catalytic triad (Cys-His-Asn) is indicated (bold letters, dotted lines). RCR3 mutants C25A (CAWAF), C24A (ACWAF) and double mutant C24AC25A (AAWAF) were generated from wild-type RCR3 (CCWAF) by site-directed mutagenesis. Proteins were produced by *Agrobacterium* mediated expression *in planta* (agroinfiltration). The signal peptide (SP, green), prodomain (violet), protease domain (blue) and putative disulphide bridges (Cys22- Cys63, Cys56-Cys96 and Cys151-Cys204) (red) are indicated. ***(D)*** Structural model of RCR3 was modelled on papain (9PAP). RCR3 is depicted in gray ribbon structure with eight cysteine residues (colored). Cys25 is the catalytic cysteine (yellow) and six other cysteine residues (Cys22, Cys56, Cys63, Cys96, Cys151 and Cys204) form three putative disulphide bridges (magenta) and Cys24 is the unpaired cysteine (blue). The double Cysteine (Cys24 and Cys25) in the catalytic groove is indicated (red circle).

2.4.2 Maturation of RCR3 does not require catalytic Cys25

To test the accumulation and activity of mutant RCR3, mutant and wild-type RCR3 were produced *in planta* by agroinfiltration of *N. benthamiana*. The analysis of apoplastic fluid (AFs) by protein blot using anti-RCR3 revealed that all mutant proteins accumulate as mature proteins (M) similar to wild-type RCR3 (Figure 4-2). These data show that the maturation of RCR3 is independent of the Cys25 and Cys24 residues. We also noted that the level of protein accumulation of C25A and C24AC25A proteins was consistently higher than the wild-type and C24A mutant (Figure 4-2).

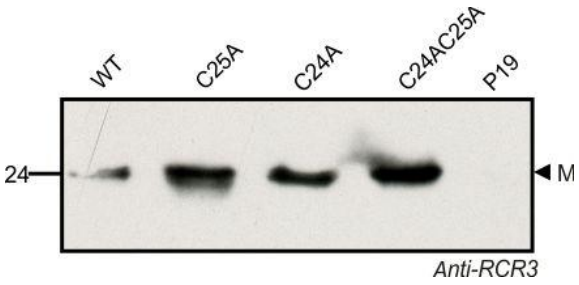


Figure 4-2: Mutant RCR3 proteins accumulate as mature proteins

Apoplastic fluid (AFs) from *N. benthamiana* leaves overexpressing wild-type and mutant RCR3 were analyzed on protein blot using anti-RCR3. The 24 kDa signal shows accumulation of mature (M) RCR3 in all mutants, similar to wild-type RCR3.

2.4.3 Cys25 in RCR3 is the only target of MV201

To determine if the mutants that accumulate as mature RCR3 are also proteolytically active, we labelled apoplastic fluids (AFs) containing wild-type and mutant RCR3 with MV201, (Figure 1-6B). Interestingly, MV201 did not label C25A and the double mutant C24AC25A, suggesting that the catalytic Cys25 is the only residue labelled by MV201 and that these mutants are proteolytically inactive (Figure 4-3). In contrast, wild-type and C24A mutant RCR3 can be labelled by MV201, but not when the samples were pre-incubated with E-64, leupeptin, chicken cystatin or AVR2, indicating that the C24A mutant is proteolytically active and sensitive for PLCP inhibitors including AVR2.

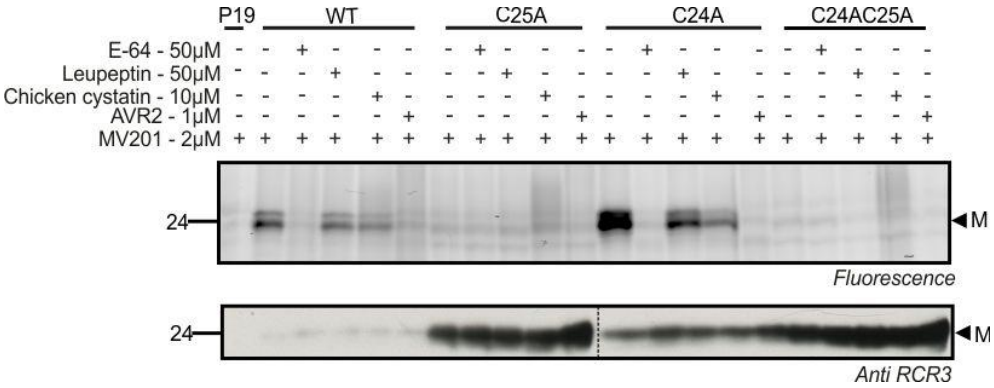


Figure 4-3: MV201 profiling of the catalytic mutants of RCR3.

Wild-type and mutants RCR3 were agroinfiltrated into Nicotiana benthamiana and the apoplastic fluid (AFs) containing the mutant proteins was isolated at five-day post infiltration (5dpi) from the leaves. AFs were preincubated for 1 hour with E-64 (50 µM), leupeptin (50 µM), chicken cystain (10 µM) or AVR2 (1 µM) in the presence of 1 mM DTT and 25 mM sodium acetate pH 5.0 and labelled with 2 µM MV201 for four hours. Proteins were separated on protein gels and analysed by fluorescence scanning and western blot with anti-RCR3. Abbreviation: M, mature RCR3.

2.4.4 Proteolytically inactive RCR3 triggers Cf2-mediated HR in the presence of AVR2.

To determine if the RCR3 mutants can trigger a Cf2-mediated hypersensitive response (HR), we infiltrated leaves of Cf2/rcr3-3 and Cf0/rcr3-3 tomato plants with AFs containing mutant RCR3 together with purified AVR2. Interestingly, all the mutants triggered HR only in Cf2/rcr3-3 tomato leaves in the presence of AVR2 and not with buffer or water or AVR2 alone or in Cf0/rcr3-3 (Figure 4-4). The HR response was observed at 3dpi for all RCR3 mutants indistinguishable from wild-type RCR3. The control infiltration of only RCR3 does not mount HR responses. These data show that, although the C25A and the C24AC25A mutant proteins are inactive, they are still recognised by Cf2 plants when co-infiltrated with AVR2.

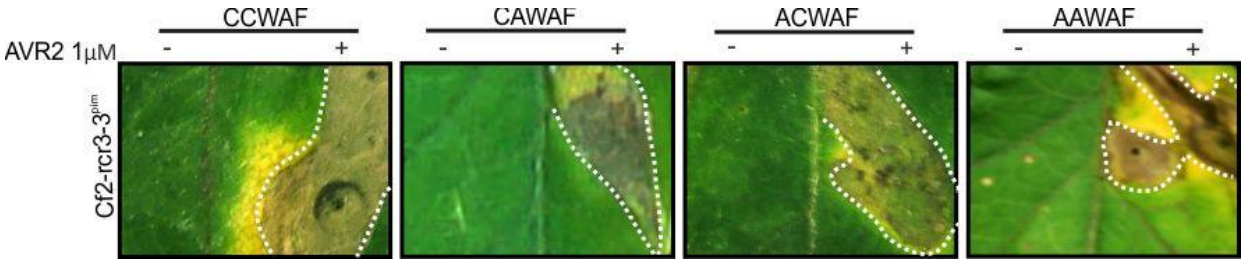


Figure 4-4: RCR3 mutants trigger Cf2-mediated HR

AFs containing mutant RCR3 were co-infiltrated with 1 µM AVR2 into leaflets of 5 week-old Cf2/rcr3-3 tomato plants. The infiltrated sectors showing cell death symptom are outlined and the infiltrated proteins indicated. Leaflets were photographed at 5 days post infiltration.

2.4.5 Ascorbate enhances activity of RCR3 catalytic mutant C24A

Cysteine residues may act as redox-sensitive regulatory switches which may regulate oxidative signalling in plants. To investigate if Cys24 has any role in redox function of the protease, we labelled the wild-type and C24A mutant proteins with MV201 in the presence of different redox agents like TCEP, DTT, L-cysteine, ascorbate and hydrogen peroxide. We anticipated upregulation of labelling in the presence of reducing agents. Surprisingly, ascorbate induces labelling of C24A mutant RCR3 significantly stronger than WT RCR3 (Figure 4-5). In contrast, DTT, TCEP and L-Cys induce RCR3 labelling of WT and C24A equally well (Figure 4-5). No labelling occurred in the presence of H_2O_2.

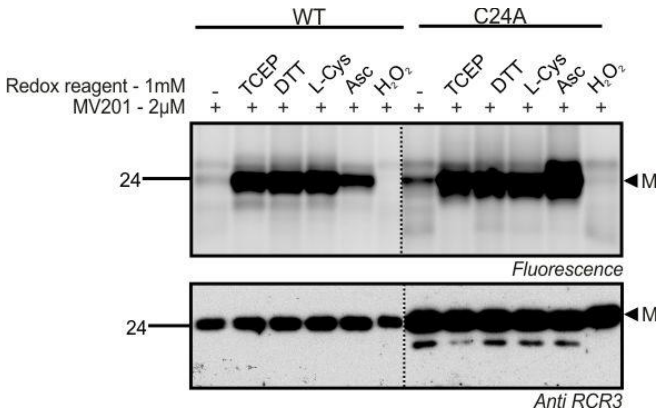


Figure 4-5: Upregulated activity of C24A RCR3 mutant in the presence of ascorbate

AFs containing wild-type and C24A RCR3 were labelled with MV201 in the presence of 25 mM sodium acetate pH 5.0 and 1 mM redox agents (TCEP, DTT, L-cysteine (L-cys), ascorbate (Asc) and hydrogen peroxide (H_2O_2) for 4 hours. Proteins were separated on protein gels and detected by fluorescence scanning and western blotting using RCR3 antibody. Western blot with anti-RCR3 confirms equal RCR3 protein levels. C24A is more labelled than WT in the absence of redox reagents possibly because of higher C24A protein levels.

CHAPTER 3: DISCUSSION

3.1 Structural Characterization of AVR2

We cloned a collection of 52 isoforms of proteases (RCR3, PIP1 and C14) and inhibitors (EPIC1, EPIC2B, AVR2 and RIP1) into the pFLAG-ATS vector for expression in *Escherichia coli*. The inhibitors were successfully expressed in *E.coli*, unlike the proteases. Among the collection of inhibitors, we focussed on AVR2 for structural characterization as no known homologous protein of AVR2 has been reported so far. Circular dichroism indicates AVR2 is a beta sheet protein and that it loses a part of its beta structure during a shift from acidic to neutral pH. Tyrosine fluorescence spectroscopy supports this conformational change. This conformational change coincides with RCR3 inhibition at apoplastic but not cytoplasmic pH. We also show that AVR2 is a potent inhibitor of papain in enzymatic assays. Attempts to crystallize AVR2 with and without epitope tags at different concentrations and conditions failed, possibly because AVR2 is a heavily charged basic protein. However, methylation of the 13 lysine residues and the *N*-terminus of epitope-tagged AVR2 generated no crystals either. Surprisingly, the methylated AVR2 still triggered HR in Cf-2 tomato plants. We also show preliminary NMR experiments of AVR2 that promises future 3D-structure elucidation of AVR2.

3.1.1 *Escherichia coli* is an ideal expression system for pathogen-derived inhibitors

For heterologous expression of proteins, *Escherichia coli* is generally the first choice due to the ease with which it is manipulated and its rapid growth in inexpensive media. Alternative systems are usually considered only after the *E.coli* system has been explored (Graslund et al., 2008). Similarly, our results also show that the expression of pathogen-derived inhibitors EPIC1, AVR2 and RIP1 in *E.coli*, generated sufficient amounts of the protein without optimisation except for EPIC2B. The expression of EPIC2B into the medium was tremendously improved by lowering the expression temperature and IPTG concentration. We presume that induction at high IPTG concentrations at 37˚C leads to high rates of transcription and translation

which resulted in accumulation of EPIC2B in the insoluble cell fraction (inclusion bodies) (Pugsley, 1993). Inclusion bodies are common in bacterial cells expressing mammalian proteins (Kane and Hartley, 1988; Marston, 1986; Mitraki and King, 1989). High-level expression of even *E.coli* proteins itself can result in the accumulation of inclusion bodies (Cheng, 1983; Georgiou et al., 1986; Scherrer et al., 1994; Yoshimura et al., 1993).

3.1.2 Beta sheet structure of AVR2 is pH sensitive

The ability of proteins to interconvert between different forms of secondary structure can occur due to changes in pH, temperature, detergent, ligand binding, etc and can be monitored by Circular Dichroism (CD) and intrinsic fluorescence spectroscopy. The estimation of the extent of conformational changes in the secondary structure of proteins is important; CD has been frequently used for this purpose (Kelly and Price, 1997; Pace, 1986). For instance, β-sheet structures can act as a precursor to the formation of amyloid fibrils which accompany degenerative diseases such as scrapie, bovine spongiform encephalopathy (BSE) and Alzheimer's disease (Hope et al., 1996; Jackson et al., 1999). To monitor the transitions from α-helical to β-sheet structures in such proteins and peptides, far UV CD has been an invaluable tool.

The CONTIN/LL program from the CDPro package (Provencher and Glockner, 1981) was used to interpret the CD spectrum information for quantitative estimation of structural elements in AVR2. Far UV CD analysis of AVR2 without epitope tags indicated that beta sheets are the major structural element in AVR2. CD spectra also indicate a conformational change from 37.3% to 63.8% beta sheets of AVR2 when lowering pH from 7.5-4.5 (Figure 2-6B). A change in conformation was consistent with the inhibition of RCR3 by AVR2, which only occurs at apoplastic pH (pH 4.5-5.5). It is important to note that the pH shift from 7.5 to 6 did not show significant changes.

To confirm the CD data, we performed intrinsic fluorescence spectroscopy of AVR2. Aromatic amino acids are useful indicators of the local environment in proteins and also monitor the conformational changes of proteins. AVR2 has two aromatic amino acid residues (Tyr^{21}) and (Phe^{54}) which can be used to study conformational changes by monitoring their intrinsic fluorescence. In our study we implemented tyrosine fluorescent spectroscopy, because phenylalanine fluorescence is relatively weak. In

our study Tyr fluorescence (excitation 274 nm and emission 303 nm) indicated conformational change exhibited by AVR2 that correlated with the unfolding of AVR2: increasing fluorescence intensity with increasing pH. However we also observed a (i) reduced tyrosine 303 nm emission fluorescence at pH 7.5 and (ii) an anomalous tyrosine emission at~ 330-350 nm inaddition to the normal tyrosine emission at 303 nm (Figure 2-7)(Raychaudhuri et al., 2008). This anamolous behaviour at wavelength other than 303 nm has also been observed in other tyrosine containing and tryptophan free proteins: (e.g) cow adrenodoxin at 331 nm (Lim and Kimura, 1980), human serum albumin at 345 nm (Longworth, 1981), parsley plastocyanin at 315 nm (Graziani et al., 1974), ocomodulin at 345 nm (MacManus et al., 1984), Indian-cobra-venom cytotoxins at 345 nm (Szabo et al., 1978), ceratitus histone HI at 325-350 nm (Jordano et al., 1983) and HYPK at around 340 nm (Raychaudhuri et al., 2011). Although the origin of this shifted tyrosine emission is not fully understood, we can definitely exclude any tryptophan fluorescence because only the SP of AVR2 contains a tryptophan residue. One possible explanation for the 330-350 nm emission and reduced 303 nm fluorescence intensity at pH 7.5 could be that the pKa value of tyrosine is~ 10.3. Hence when the Tyr residue is at higher pH, there is increased population of ionized and stabilized tyrosine in the excited state due to loss of proton from the aromatic hydroxyl group. This would lead to quenching of tyrosine fluorescence and possible local hydrogen bonding, influenced by neighbouring charges (Raychaudhuri et al., 2011).

In summary, the excited state proton transfer from the phenolic hydroxyl group to a concluded proton acceptor in the local microenvironment around the tyrosine residue may lead to the anomalous intrinsic fluorescence property of Tyr^{21} at 340 nm in AVR2. The dynamic equilibrium of this single fluorophor switching between tyrosine and tyrosinate forms in a conformation-driven environment may attribute to various functional properties of AVR2. Although the chemical origin of the anomaly is still unclear, these results unequivocally demonstrate a conformation-determined microenvironment of the tyrosine residue in AVR2.

pH sensitive conformations are not uncommon for effector proteins. For example, AvrPto secreted by *Pseudomonas syringae* exhibits a conformational change during a pH gradient inorder to translocate from the bacterial cytoplasm (mildly acidic) into the host cell (neutral). This pH gradient is required as folded proteins will not be able

to pass through the narrow pilus of the TTSS. Hence unfolding of the effector is required for translocation at acidic pH in the pilus. The effector refolds upon delivery in the host cytoplasm (Dawson et al., 2009). Similar hypothesis or model could be proposed where AVR2 (pI 9.13) encounters a pH gradient during secretion from the fungal cell into the acidic apoplast where the effector exhibits a structural change in the beta sheets to desirable extent in order to bind and inhibit RCR3 (Figure 5-1).

The speed and convenience of CD are the main advantage of the technique when compared to X-ray crystallography and NMR. However the main limitation of CD is that it only provides relatively low resolution structural information. Thus, although far UV CD can give reasonably reliable estimates of the secondary structure content of a protein, it must be noted that these are overall figures and do not indicate which regions of the protein are of which structural type.

Furthermore although CD spectra and tyrosine spectroscopy indicated structural changes in AVR2, these studies do not exclude that addition of interacting proteins or molecules can lead to structural changes which are essential for their function (Price and Stevens, 1999). For example addition of the competitive inhibitor N-acetylglucosamine to lysozyme led to large increase in the CD peaks at 285 nm and 295 nm, characteristic of Trp chains (Glazer and Simmons, 1996). Hence the CD analysis with both RCR3 and AVR2 at different pH is required to make further biological relevance.

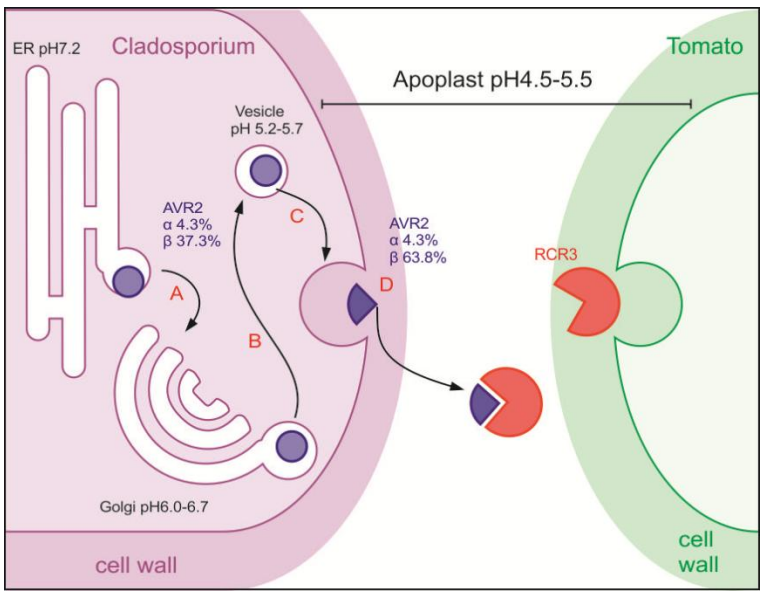


Figure 5-1: Model: Translocation of AVR2 via pH gradient accompanying conformational changes.

AVR2 (purple) is secreted by *Cladosporium fulvum* and passes through the endoplasmic reticulum (ER, pH 7.2) (A), the golgi complex (pH 6.0-6.7) (B) and through the secretory granules (C) to be released into the apoplast. Upon delivery, in the acidic apoplast (D) AVR2 exerts structural change in the beta sheet region (more beta sheets). AVR2 binds and inhibits RCR3 (red) at the apoplastic pH (4.5-5.5). Percentage conformational change indicated (violet).

3.1.3 NMR spectroscopy can complement X-ray crystallography for AVR2 structural studies

Three variants of AVR2 were used in crystallization studies: (i) native wild-type AVR2 (58-residue protein), (ii) epitope-tagged AVR2 (89-residue protein) and (iii) methylated epitope- tagged AVR2. At first, we screened different concentrations up to 50 mg/ml AVR2 without epitope tags using 864 conditions. Attempts to crystallize AVR2 without epitope tags at different concentrations and conditions yielded mostly clear drops and some phase separation and precipitates. Hence we continued screening using AVR2 with epitope tag (FLAG-HIS_6-TEV) under 2592 screening conditions because fusion tags in some cases provides contacts for crystal formation. For example, ph0828 (1v30), a 116-residue protein from *Pyrococcus horikoshii* was crystallized in the presence of His tag, where the His-tagged derivative enhanced crystallization unlike the wild-type (Klose et al., 2004). Apart from the presence itself, length of the His tag can be crucial to the crystallization process, as it induces additional and important crystal contacts. For instance His_5 tag not His_6 tag induced a stable conformation of OcDH (Octopine dehydrogenase) by creating crystal contacts (Smits et al., 2008). However, there are also few reports on the adverse effects of a His tag on the structure-function relationships. In our experiments, both epitope-tagged and untagged AVR2 did not crystallize.

From the above crystallization screen results, one could attribute the failure of crystal formation to the high pI of AVR2: 9.13 for untagged AVR2 and 8.15 for tagged AVR2 caused by eleven lysines. High solubility of AVR2 could destabilize the crystal lattice through entrophic effects. Ingeneral lysine residues are predominantly located on the surface, with 68% exposed, 26% partly exposed and only 6% buried (Baud and

Karlin, 1999). This makes the lysine residues the most solvent-exposed residues in proteins. Even in proteins with average lysine content (5.8%), these amino acids constitute a large fraction of the surface area because most are solvent-accessible. It has been calculated that lysines account on an average for 12-15% of solvent-accessible surface (DoConte and Chothia, 1999). We addressed the the charge problem of increased solubility by reductive methylation. This method was originally used as a method of isotope labelling (Means and Feeney, 1968; Rice and Means, 1971) where free amino groups in primary amines (lysine residues and *N*- terminus) are modified to tertiary amines. This method has enhanced the ability of proteins to crystallize and resulted in diffracting crystals for maltooligosyl trehalose synthase (Kobayashi et al., 1999), pokeweed antiviral protein-II (Kurinov et al., 2000), myosin subfragment-1(Rayment et al., 1993), Cathepsin Z (Saxena et al., 2006) and a ternary complex consisting of *Yersinia pestis* virulence factors YopN, SycN and YscB (Schubot and Waugh, 2004). Therefore, in our study we implemented reductive methylation to enhance the crystallisability of epitope-tagged AVR2. We observed precipitation after methylation of AVR2 and the AVR2 concentration reduced 4-fold. Mass spectrometry proved that most AVR2 was completely methylated. We also observed a series of ion species 14 Da apart that corresponded to addition of single methyl groups (Figure 2-9C, D) and that AVR2 carries at least 24 methyl groups. This incomplete methylation is observed frequently and could be due to perished reagents (Derewenda and Vekilov, 2006).

As expected, methylated proteins usually display reduced solubility (Schubot and Waugh, 2004). Similarly, our 384 crystallization screen tests of methylated AVR2 exhibited more precipitates, phase separation and less clear drops suggesting reduced solubility of methylated AVR2 unlike non-methylated AVR2. Overall methylation of AVR2 did not improve the crystal formation. Interestingly, methylated AVR2 still triggered Cf-2-mediated HR in tomato similar to unmethylated AVR2 (Figure 2-9E). Van't Klooster et al (Van't Klooster et al., 2011) showed that AVR2 mutants K20A and K17A was still able to trigger a Cf2-mediated HR. In agreement with (Van't Klooster et al., 2011) thus stability or functionality of AVR2 was not decreased by reductive lysine methylation eventhough the charge of AVR2 was altered. However, Lys residues can be important for effector recognition. For instance an effector protein PopP2 of *Ralstonia solanacearum* requires a specific lysine residue to trigger

immunity in Arabidopsis by the *RRS1*-R resistance gene perception (Tasset et al., 2010).

Another problem for AVR2 crystallization might be the flexibility of domains or polypeptides. Flexibile motions in protein can be reduced or eliminated by binding to crystallization chaperones such as chemical inhibitors (Ren et al., 1995) or antibody fragments (Iwata et al., 1995). Hence a future strategy is to co-crystallize AVR2 with RCR3 to stabilize and improve the contacts in the crystal lattice.

Although crystallization remains to be the leading method for solving structures of macromolecules and their complexes in atomic details, the success rests upon obtaining diffraction-quality crystals. At present the success rate remains no more than 50% (Canaves et al., 2004; Fogg et al., 2006). NMR spectroscopy, despite its protein size limitation can complement x-ray crystallography in structural studies (Christendat et al., 2000; Savchenko et al., 2003; Yee et al., 2002). X-ray crystallography and NMR studies have shown similar overall structure for cysteine protease inhibitors chicken cystatin C (Bode et al., 1988; Dieckmann et al., 1993; Engh et al., 1993), cystatin A (Martin et al., 1995) and cystatin B in complex with papain (Stubbs et al., 1990) by.

Although extensive NMR experiments of AVR2 have not been performed, the initial ^{15}N-HSQC NMR spectrum of AVR2 indicates that AVR2 is amenable to structure determination by NMR spectroscopy (Figure 2-10). The HSQC spectra indicated folded protein and contained well-dispersed peaks of roughly equal intensity and an expected number of N-H peaks. These spectra suggested that the process of determining the 3D-structure of AVR2 by NMR should be relatively straightforward.

Hence, the application of both X-ray crystallography and NMR spectroscopy to validate the structures independently and also probe a wide range of structural and dynamic properties of proteins will produce a better understanding of the biological function involved.

3.1.4 AVR2 a potent inhibitor of Papain

Papain is a classic family C1A cysteine protease has been widely used and studied both in crystallization and enzymatic studies in complex with inhibitors, small peptides etc. Papain is inhibited by chagasin (K_i = 36 pM) (Redzynia et al., 2009); RIPI from *P. syringae*(K_i = 2.13nM) (Shabab et al., 2008), cystatins ($K_i = 10^{-9}$ -10^{-11}) and chicken cystatin (K_d = 60 fM) (Anastasi et al., 1983; Bjork et al., 1989; Lindahl et al., 1988; Nicklin and Barrett, 1984). The K_i in extreme cases can be as low as 10^{-14}M for the human cystatin C-papain complex (Lindahl et al., 1992). Papain is also inhibited reversibly by a series of peptidyl α-keto esters, α-keto amides, α-keto acids, and α-diketones (Bendall et al., 1977). Papain is inhibited by methyl 3-(N-benzyloxycarbonyl-L-phenylalanyl)amino-2-oxopropionate (a dicarbonyl compound; K_i = 1 µM) and 3-(N-acetyl-L-phenylalanyl)aminopropanone (a monocarbonyl compound; K_i = 1.5 mM) (Bendall et al., 1977). E-64 inhibits papain with (K_i= 10µM). Other well-known potent reversible inhibitors for papain are peptidyl aldehydes (K_i = 1-46nM) (Bendall et al., 1977; Rich, 1986), leupeptin (K_i = 2nM) (Rich, 1986) and peptidyl nitriles (K_i = 6-7µM) (Liang and Abeles, 1987).

In our study, we show that AVR2 is another potent inhibitor of papain. Papain (9PAP) is very close to RCR3 and PIP1 in phylogenetic tree (Beers et al., 2004) and also one of the top hits given by a template search by SWISS-MODEL for RCR3 sequence (Schwede et al., 2003) And papain from Carica papaya is highly homologous (43%) to the 344 amino acids protein encoded by Rcr3 (Kruger et al., 2002). We performed inhibition assays for different AVR2 concentrations at constant papain and bodipy substrate concentrations. We show that AVR2 begins to inhibit papain at 25 µM and completely saturates papain at 250 µM (Figure 2-5A). Although we clearly observe an inhibition pattern in the presence of AVR2, we cannot conclude the mode of inhibition without Ki values. At 40 µM E-64, most of the binding sites of papain are occupied resulting in no activity (Figure 2-5B).

Van'T Klooster et al reported that AVR2 behaves as an non-competitive inhibitor of RCR3 using bodipy-FL casein as substrate (Van't Klooster et al., 2011). These data indicate that AVR2 binds outside the substrate binding site in RCR3. In such non-competitive inhibiton, the inhibitor reduces the activity of the enzyme upon binding but it does not affect the binding of the substrate. It is also possible that PLCPs can

undergo a conformational change forced by inhibitor binding as observed in Human cathepsin B-Chagasin interaction (Redzynia et al., 2008) . The occluding loop of cathepsin B is differently displaced indicating a large range of movement and adoption of a conformation forced by the inhibitor (Redzynia et al., 2008).

Another phenomenon in kinetics is the correlation of inhibition rate with external factors like pH, temperature etc. Interestingly, the Papain-RIP1 interaction increases with increased in temperatures (Shabab et al., 2008). Papain also shows different $K_i$values with the inhibitor Gly-Gly-Tyr (Bzl)-Arg at different pH ($K_i$of 10 µM at pH 4.3 and $K_i$of 150 µM at pH 6.0) (Blumberg et al., 1970). The impact of RCR3-AVR2 interaction on inhibition constant (K_i) values at different pH should coincide with ABPP data (Samali et al., 1999; Shabab et al., 2008) and may be supported by conformational changes observed by CD assays and intrinsic Tyr fluorescence spectroscopy data.

The determination of K_i of RCR3 with AVR2, EPIC1 and EPIC2B are planned for future experiments. These assays will present a better understanding of the mode of inhibition (competitive or uncompetitive), and the relative strength of inhibition.

3.2 Biochemical Characterization of RCR3

Four different expression systems (plant, bacterium, insect and yeast) were tested for high amounts of active and soluble RCR3. We show that *Pichia pastoris* expression system is the best of the four expression systems. Further, we show that the activation of proRCR3 produced in yeast into mature RCR3 is triggered by reducing agents and acidic pH buffer *invitro*. This observation indicates that RCR3 can be activated by intramolecular processing. However, inhibition of RCR3 maturation by E-64 or AVR2 may also support a model of intermolecular processing. Finally we study the role of double cysteine (Cys24, Cys25) in the catalytic site of RCR3 which is a common feature in subfamily 6 of plant PLCPs. Using agroinfiltration in *Nicotiana benthamiana* we produced C24A, C25A and C24AC25A mutant RCR3 proteins and we show that Cys25 but not Cys24 is the essential catalytic residue labelled by activity-based probe MV201. Surprisingly, maturation of RCR3 does not require the catalytic Cys25, indicating that other endogeneous proteases can activate RCR3 *in*

vivo by intermolecular processing. Further we show that proteolytically inactive RCR3 mutants trigger Cf-2-mediated HR in the presence of AVR2. Interestingly, we found that ascorbate enhances the activity of the C24A mutant but not wild-type RCR3, suggesting that Cys24 in RCR3 might have a role in sensing redox potential.

3.2.1 Production of high quantities of active, soluble RCR3.

Several expression systems were tested for producing RCR3. Transient expression of genes through agroinfiltration is a quick and easy method to study genes of interest(Kapila et al. 1997; Rossi et al. 1993). Especially *Nicotiana benthamiana* has been exploited as an expression host for recombinant proteins of clinical interest (Jackson et al., 1999; Price and Stevens, 1999). This system presents several advantages over other plant expression systems, including: (i) short expression time in mature leaves tempering the growth interference by the transgene products or the risks of toxicity; (ii) production of significant amount of recombinant protein in short time (days); (iii) easy recovery of proteins secreted in the apoplast by easy vacuum infiltration protocols. Isolation of apoplastic fluids (AFs) is a very simple and elegant method to study secreted proteins (Joosten, 2012). The presence of viral suppressors further enhances the expression of proteins in *Nicotiana benthamiana*during agroinfiltration (Voinnet et al., 2003). *N. benthamiana* has been used for transient expression of seven different tomato PLCPs (Shabab et al., 2008) and ten Arabidopsis PLCPs (Richau et al., 2012). Since RCR3 is present in the apoplast, we agroinfiltrated *N. benthamiana* leaves to produce RCR3 *inplanta*. Even though the protein was soluble, active and relatively pure due to low endogeneous proteins in the apoplast, we found that only low amounts of RCR3 was present in the apoplast (Figure 3-1A). The < 3 mg/ml RCR3 in AF from 40 leaves is sufficient for labelling experiments and other biological assays but not for structural biology studies where milligrams of pure protein are required.

We next tested the *E.coli* system for the production of RCR3. The first plant proteases successfully expressed and produced in high quantities in *E.coli* were papain (Taylor et al., 1995) and caricain (Revell et al., 1993). Later also mammalian cathepsins such as cathepsin X (Santamaria et al., 1998), cathepsin O (Velasco et al., 1994), cathepsin S (Kopitar et al., 1996; Tobbell et al., 2002), cathepsin L (Barlic-Maganja et al., 1998; Dolinar et al., 1995), cathepsin K (D'Alessio K et al., 1999; Hwang and Chung, 2002)

and cathepsin B (Kuhelj et al., 1995) have been produced in *E.coli*. However, in each case the protease was produced in inclusion bodies and refolded back into an active protein. Similarly, we show that the expression of RCR3, PIP1 and C14 in different isoforms in *E.coli* accumulated in high quantities in the insoluble fraction even though a secretion system was used. RCR3 could not be refolded in sufficient quantities. Likewise, Cathepsin S was expressed without its prodomain but with a low yield when compared with the expression of the proenzyme (Tobbell et al., 2002). Similar results were observed with Cathepsin L (Kelly and Price, 1997). One exception is falcipain-2, which requires only the additional *N*-terminal folding domain but not the prodomain for proper folding (Pace, 1986). From the knowledge obtained from the above mentioned research, we produced RCR3 with and without the propeptide.

However, refolding is the most critical step after producing PLCPs in *E. coli*. We initially implemented the refolding of RCR3 by rapidly diluting the denatured protein directly into the refolding buffer (Maeda et al., 1995) . Such a dilution refolding method has been used for example mild removal of urea from denatured lysozyme by dialysis effectively renatured the protein. However, this dilution method did not restore RCR3 activity. Protein aggregation at relatively high protein concentration can be suppressed by a linear decrease in urea concentration (Yoshii et al., 2000). We therefore used a modified gradient refolding protocol used for falcipain (Singh et al., 2006). Falcipain, like RCR3, belongs to subfamily C1A of PLCPs. The resulting RCR3 fractions showed an increase in proteolytic activity when the denaturant was decreased (4M to 0M). However, gradiently refolded RCR3 degraded casein 100-fold less than RCR3 expressed *inplanta*. The low activity of RCR3 both with and without the prodomain after refolding could be due to scrambled disulphide bridges and/or improper associations with correct domains in the protein.

We next tested the baculovirus expression system. This system has been used for production of large quantities of various parasite-derived and mammalian PLCPs (Giancotti et al., 1980; Hammerschmidt, 2010; Lohaus et al., 2001). Insect cells allow post-translational modifications similar to mammalian cells. Human cathepsin B (Hammerschmidt, 2010) and cathepsin C (Cheng, 1983) have been successfully produced in the insect cells. However a distinct disadvantage for expression of cysteine proteases in insect cells is the expression of several endogeneous cathepsin-like proteases like cathepsin B which might interfere in the analysis of the

recombinant protease (Bromme and Okamoto, 1995). We used baculovirus expression vectors in insect cells for the production and processing of proRCR3. Unlike the successful production of cathepsins and other cysteine proteases mentioned above, proRCR3 was secreted only in micrograms per liter culture, indicating that the insect cell expression system is not an ideal system for RCR3 expression (Figure 3-1C).

We next tested the *Pichia pastoris* expression system for expression of proRCR3. *P.pastoris* expression system has been successfully used to produce highly disulphide bonded proteins (White et al., 1994). The reducing environment of the cytoplasm renders prokaryotic expression systems less successful for production of disulphide-bonded proteins as there is a necessity to refold disulphide-bonded proteins from inclusion bodies, or to secrete the proteins into the periplasmic space. A specific additional advantage of the *P. pastoris* expression system is a very low secretion of endogeneous proteins, causing the recombinant protein being the major protein in the media, which dramatically simplifies the purification procedure (Cregg et al., 1993). This espression system has been used for the production of recombinant PLCPs like human cathepsin L and X (Carmona et al., 1996; Khouri et al., 1991; Nagler et al., 1999), cathepsin B from *Leishmania major*(Chan et al., 1999), cathepsin B2 from *Schistosoma mansoni(Caffrey et al., 2002)* and cathepsin L from *Camellia sinensis*(Park et al., 2001) and also RCR3 (Rooney et al., 2005; Van't Klooster et al., 2011).

We found large quantities of proRCR3 (50 mg/L) and low endogenous proteins in the medium, making proRCR3 the predominant protein which facilitated an easy purification (Figure 3-1D). In *P. pastoris*, unintended glycosylation is a concern since it is a common post-translational modification (Eckart and Bussineau, 1996). It is possible that *Pichia* glycosylates heterologous proteins, even when those proteins are normally not glycosylated by the native host. When the protein is glycosylated in the native host, *Pichia* may glycosylate the protein not on the same residues (Cereghino and Cregg, 2000). The extent of *O*-glycosylation by *P. pastoris* has been studied on the catalytic domain of glucoamylase from *Aspergillus awamori*(Heimo et al., 1997). The molecular weight of this protein was 20 kDa heavier when expressed in *P. pastoris*, when compared to the native protein. About 10 kDa of this weight could be attributed to *N*-glycosylation, and the rest was attributed to *O*-linked glycosides

(Heimo et al., 1997). When used as pharmaceuticals, *Pichia*-derived glycosylated proteins have the potential to trigger inappropriate immune responses. For example, *O*-glycosylation near the reactive site of recombinant human antithrombin III resulted in a recombinant protein exhibiting only half the inihibitory activity against thrombin when compared to the native protein (Mochizuki et al., 2001). Although bioinformatic analysis revealed 46 putative *O*-glycosylation sites and one *N*-glycosylation site in proRCR3, the molecular weight (MW) of proRCR3 was close to the expected MW (Figure 3-2A). Furthermore, RCR3 expressed in yeast was functional in the HR assay (Figure 3-2B) and in AVR2 inhibition assays (Figure 3-3C). However, MS analysis is required to confirm the exact mass upon expression in *Pichia*.

3.2.2 Intramolecular and Intermolecular processing of proRCR3

Similar to most other proteases, RCR3 is synthesized as inactive precursor or zymogen containing *N*-terminal propeptide. RCR3 is activated by proteolytic removal of the propeptide. The 3D structures of procathepsin L (Coulombe et al., 1996), rat procathepsin B (Cygler et al., 1996), procaricain (Groves et al., 1996), human procathepsin S (Kaulmann et al., 2006) and human procathepsin K (LaLonde et al., 1999) reveal that despite having a low homology in their protein sequence, the propeptide adopts a specific structural conformation that is conserved between propeptides of the PLCP family. Previous reports suggest that a conformational change in the propeptide domain is triggered by acidic pH, resulting in an intermediate state which reduces the association affinity towards the protease domain, leading to prodomain cleavage and resulting in a mature active enzyme (Gutierrez-Gonzalez et al., 2006). Processing of PLCPs is therefore thought to occur autocatalytically at acidic pH (Bromme et al., 2004). For example, recombinant propapain was one of the first cysteine protease precursors successfully activated as *invitro* at pH 4.0 under reducing conditions (20 mM cysteine) and elevated temperature (60˚C) (Vernet et al., 1991). Because mutation of the active site C25S in papain prevented the processing reaction, this activation was suggested to be initiated and to proceed intramolecularly. We observed a similar type of activation in the processing of proRCR3. A time course experiment for processing of proRCR3 expressed in *Pichia*(Figure 3-3A, B) shows that recombinant proRCR3 can be

activated at acidic pH and reducing conditions, mimicking the environment of the endosomal/prelysosomal compartments (Figure 3-3A, B). Low pH and reducing conditions may induce a structural transition required for intramolecular cleavage, similar to the one described during processing of pepsinogen to pepsin (Glick et al., 1989). The results described above are consistent with the hypothesis that the propeptide processing occurs intramolecularly. However, *in vitro* processing of the C25A catalytic mutant of RCR3 is required to confirm the mechanism of intramolecular processing. This requires the expression of C24A, C25A and C24AC25A mutants in *Pichia*.

However, we also found evidence for intermolecular processing. The first reason we observe is that *in vitro* removal of propetide from RCR3 can be inhibited by E-64 or AVR2 (Figure 3-3C, D). Since the active site is not accessible for E-64 in proproteases this observation suggests that one mature protease could trigger a chain reaction by intermolecular processing. Inhibition of the mature protease could therefore block the activation. Such mechanism has also been proposed for cathepsin B: one mature protease can activate the others in a chain reaction creating a positive feedback loop (Turk et al., 2000). A second observation supporting intermolecular processing is that the catalytic mutants C25A and C24AC25A accumulated as mature proteases upon agrofiltration (Figure 4-2).A similar observation was made for the RD21 protease of Arabidopsis. These observations imply that the catalytic cysteine is not required for removal of prodomain *in planta.* The accumulation of mature RCR3 mutants in *N. benthamiana* could be because other endogenous proteases in the lysosomal/ endosomal compartments process RCR3. Proteolytic cascades have not been described in plants but are common in animals (Mason and Joyce, 2011). RD21 catalytic mutant C161A accumulates as processed RD21 even though the mutant was inactive in DCG-04 labelling and rubisco degradation (Gu et al., 2012).

In future experiments, we aim to to use class specific inhibitors to identify the endogenous protease. A similar method was used to identify proteases in the activation of procathepsins B, H and L (Nishimura et al., 1989; Nishimura et al., 1988; Wiederanders and Kirschke, 1989). To determine the subcellular localization of activation of the proteases in plants, the secretory pathway inhibitor brefeldin can be used. This exocytosis inhibitor was used to demonstrate that cathepsin B is processed during secretion (Gilroy et al., 2007).

In summary our study implies that proRCR3 can be processed by both intramolecular and intermolecular processing. This mechanism is similar to recombinant procathepsin L where both intra- and intermolecular reactions might be involved (Mach et al., 1994; McDonald and Emerick, 1995; Nomura and Fujisawa, 1997).

3.2.3 Role of double cysteine in active site of RCR3 – Cys24 as redox sensor?

In this study we set out to investigate the role of Cys24 in the active site of PLCPs in subfamily 6. We speculate that the unpaired Cys24 has a functional relevance and hence we focused on RCR3 mutants for our study. In proteins, cysteine is the most intrinsically nucleophilic amino acid and its reactivity is tuned to perform diverse biochemical functions. During initial studies, we hypothesized that this extra cysteine may act as an alternative catalytic cysteine or as a redox sensor (Richau et al., 2012).

We presumed that Cys24 might either perform a catalytic role together or in the absence of Cys25. Hence we expressed different RCR3 mutants (C24A, C25A and C24AC25A) proteins *in planta* and tested the mutants for MV201 labelling in the presence or absence of different inhibitors. We detected labelling of a 24 kDa signal corresponding mature RCR3 only in the wild-type and C24A and this signal was competed in the presence of inhibitors (Figure 4-3). This observation implies that Cys25 is the only target of MV201. We conclude that C25A and C24AC25A mutants are proteolytically inactive. Interestingly, all the RCR3 mutants triggered HR in the presence of AVR2 in Cf2/rcr3-3 tomato (Figure 4-4). These data demonstrate that the mutations did not abolish AVR2/Cf2 recognition in tomato. We also observed that the C25A and C24AC25A mutants accumulated as mature forms upon agroinfiltration. Enzymatic assays with the RCR3 mutants using bodipy casein as substrate did not work because of endogenous proteins in AFs even in the P19 control. To produce purified RCR3 mutant proteins, the expression of C24A, C25A and C24AC25A mutants in Pichia are planned for future experiments. We next explained if the Cys24 has a role in redox regulation in plants. We observed that H_2O_2 prevented the labelling of RCR3 wild-type and mutant C24A using MV201. This observation is in agreement with (Borutaite and Brown, 2001) where H_2O_2 inactivates caspases and DTT could reverse the H_2O_2-induced inactivation. All reducing agents activate both wild-type and C24A. Interestingly, ascorbate activates C24A more than the other reducing

agents, whereas it activates wild-type RCR3 less than the other reducing agents. Ascorbate is a mild reducing agent that is mechanistically distinct from the other agents. DTT and Cysteine carry sulfhydryl groups to donate hydrogen; TCEP contains a phosphate to accept oxygen and donate hydrogen; and ascorbate contains hydroxyl groups that can donate hydrogens. We were unable to find an adequate explanation for the effect that ascorbate has on the activity of wild-type and C224A RCR3. Importantly, ascorbate exists in the apoplast in millimolar concentrations (Foyer et al., 1983). Thus the conditions tested here are physiological.

To address the redox function further several questions should be addressed. At first it is necessary to identify if Cys24 is a hyper-reactive cysteine because susceptibility of cysteines to modifications largely depends on the characteristics of each thiol. Exposure and Pk_a play a significant role in determining cysteine reactivity. Hyper-reactive cysteines can be identified with iodoacetate probes (Weerapana et al., 2010). Next we should extend this study by MS analysis of the wild-type and C24A proteins in the presence of ascorbate to determine modifications of Cys24. A future 3D structure of RCR3 will also help in deciphering the biological relevance to the extra cysteine.

CHAPTER 4: MATERIALS AND METHODS

4.1 Materials

4.1.1 Chemicals and Biochemicals

All chemicals and antibiotics were supplied by Sigma (Deisenhofen, Germany), Roth (Karlsruhe, Germany), Merck (Darmstadt, Germany) and Duchefa (Haarlem, Germany). E-64, leupeptin, epoxomicin, were purchased from Sigma. All inhibitors were dissolved in DMSO and stored at -20 °C. BTH 50WG (Actigard/Bion, 50% wettable granule) was ordered from Ciba-Geigy AG. Dimethylamine-borane complex from Fluka (ABC). Z-Leu-Arg-AMC (Z: N-carbobenzyloxy-7-Amino-4-methylcoumarin) was from Bachem (Germany).

Antibodies

Anti-FLAG M2 antibody was purchased from Sigma-Aldrich and anti-RCR3 antibody was provided by Dr. Jonathan jones (Sainsbury laboratory, U.K). Secondary antibodies (HRP-conjugated anti-rabbit antibody and HRP-conjugated anti-mouse antibody) were purchased from Thermo Scientific.

Enzymes

Restriction enzymes were from Fermentas (St.Leon-Rot, Germany) and New England Biolabs (Frankfurt/Main, Germany). Taq polymerase for standard PCR was either from Promega (Mannheim, Germany) or BioBudget (Krefeld, Germany) and high-fidelity polymerase was from Roche (Karlsruhe, Germany). Reverse-transcriptase was from Invitrogen (Karlsruhe, Germany). Ligases were either from Promega (Mannheim, Germany) or Fermentas (St.Leon-Rot, Germany). DNase and RNase were from Roth (Karlsruhe, Germany).

Activity-based probes

Epoxide probes (DCG-04 and MV201) were provided by Dr. Herman Overkleeft (Leiden University, Netherlands) and Dr. Matthew Bogyo (Stanford Medical School, USA) and synthesized as described previously (Greenbaum et al., 2002, Richau et al., 2012). All probes were solved in DMSO to a stock concentration of 1 mM and stored at -20 °C. Fluorescent probes were kept in the dark.

Kits and primers

Oligonucleotide primers were synthesized by Invitrogen (Karlsruhe, Germany). HPLC purified primers were generated by Sigma (Deisenhofen, Germany). Oligo(dT) primers were from Invitrogen (Karlsruhe, Germany) or Roche (Karlsruhe, Germany). Kits for isolating DNA or RNA were supplied from Qiagen (Hilden, Germany). Plasmid isolation was carried out using kits of Peqlab (Erlangen, Germany) or Macherey-Nagel (Duren, Germany). The Enzcheck assay kit for enzyme kinetics was supplied from Molecular Probes (Paisley, UK). Site-directed mutagenesis was carried out by Quick-change site-directed mutagenesis kit from Stratagene. Size exclusion chromatography was carried out by HiLoad 26/60 Superdex 75 pg prepacked column supplied by GE Health care. Amicon centrifugal filters (MWCO 3 kDa) from Millipore were used for concentrating and desalting proteins.

All primers are listed in Table 4 at the end of the Material and Methods section. Site-directed mutagenesis primers are indicated in red.

Crystallization Screens

Standard initial crystallization screens (JCSG Core Suites (I-IV), JCSG+ Suite, Classics, Cryos suites, AmSO4 Suite, MPD Suite, Anion and Cation Suites and PEG Suites) were supplied by Qiagen.

4.1.2 Biological materials

Plant material

Nicotiana benthamiana (310A) and tomato (*Solanum lycopersicum* cv. Money-Maker) used in this work were grown at the MPIPZ (Cologne, Germany). *N.*

benthamiana and tomato were grown in a climate chamber at a 14 hour light regime at 22 °C (day) and 18 °C (night). Usually, 4 to 6 week-old plants were used for experiments. *N. benthamiana* leaf proteins were extracted from the agroinfiltrated leaves at 3 days post-infiltration (3dpi) or from fully expanded leaves of a 5 week-old plant. Tomato leaf proteins were extracted from the apoplast of fully expanded compound leaf of a 4 week-old plant. The apoplastic fluid extraction method is discussed in the methods section 4.2.6.

Bacterial strains

Escherichia coli strain DH10B and BLyS strain BL21-(DE3) F-Omp T hsdSB (rB-mB-) Gal dcm (DE3) were used for standard cloning and protein expression respectively. *Agrobacterium tumefaciens* strain GV3101 was used for agroinfiltration and plant transformation.

Yeast strains

Pichia pastoris strain 4 (Pichiapink expression system) supplied by Invitrogen was used for protein expression.

Insect strains

Insect cell line Sf9 (Cat. No. 21300L) established from ovarian tissues of *Spodoptera frugiperda* larvae were supplied by BD Pharmingen (San Diego) for protein expression.

Plasmid vectors

The pFLAG-ATS plasmid was supplied by Sigma Aldrich. Agrobacterium carrying the $Rcr3^{pim}$expression cassette with C-terminal His_6 epitope tag was provided by Dr. Sophien Kamoun, Sainsbury laboratory, UK (Song et al., 2009) and Pichia carrying the expression cassette $RCR3^{esc}$ with C-terminal His_6 tag was provided by Dr. Boquian Wu, Carlsberg laboratory, Denmark (unpublished). pFK26, pTP25and binary vector pTP05 were described previously (Shabab et al., 2008).

All collection of pathogen-derived inhibitors and proteases cloned into pFLAG-ATS vector for *E.coli* expression are listed in Table 1 (Chapter 2, section 2.1). All

collection of mutant RCR3 plasmids for agrobacterium mediated expression *in planta* is listed in Table 5 at the end of the Material and Methods section.

4.2 Methods

4.2.1 Recombinant expression and affinity purification of inhibitors and proteases in bacteria

Medium

E. coli strain BLyS (BL21-(DE3) F- omp T hsdSB (rB-mB-) Gal dcm (DE3) transformed with plasmid pFLAG carrying the inhibitor or protease expression cassette was grown overnight at 37 ℃ in 20 ml LB medium containing 100 µg/ml ampicillin. The bacterial culture was induced with 0.8 mM IPTG once the OD_{600} reached 0.6 and incubated at 37 ℃ for 12 hr with vigorous agitation at 200 rpm on a shaker. The bacterial culture was centrifugated at 6000 g for 30 min with Sorvall RC-5B centrifuge (Du Pont Instruments), and the supernatant containing secreted proteins was filtered through 0.22 µm filters (Steritop, Millipore). The supernatant was precipitated at 4℃ with 95% ammonium sulphate (527.52 g/L) determined by http://www.encorbio.com/protocols/AM-SO4.htm. Proteins in the supernatant were pelleted by centrifugation at 17000 g for 45 min at 4℃ and dissolved in TBS buffer by gentle agitation on a roller mixer (SRT2, Stuart). Proteins recovered from the culture supernatant were captured and purified by His_6 affinity using a Poly-prep C-column (Bio-Rad) packed with 1000 µl Nickel agarose affinity gel (Qiagen) at 4 ℃. The column was equilibrated with two column volumes of binding buffer (50 mM Tris pH 7.5, 150 mM NaCl) and after passing through the protein sample, the column was washed three times with 2 ml binding buffer and the bound proteins were eluted with increasing concentrations of imidazole (gradient elution) and immediately kept on ice and at -80°C for long term storage.

Soluble Cell Fraction (SCF)

The bacterial cells expressing protease or inhibitor proteins were expressed as mentioned above. After induction, the bacterial culture was centrifugated at 6000 g for 30 minutes with Sorvall RC-5B centrifuge (Du Pont Instruments), and the

bacterial pellet was lysed by freeze-thaw cycles (using liquid nitrogen), lysozyme treatment (30 µg/ml) and sonication (branson sonifier, Germany). The lysed cell fraction was centrifuged at 15,000 g for 30 minutes and the soluble fraction from the bacterial pellet was analysed by protein blot and used for purification.

Insoluble Cell Fraction (ICF)

The lysed bacterial pellet was solubilised using 8M Urea in binding buffer (50 mM Tris pH 7.5, 150 mM NaCl). For refolding, the denatured protein was 100-fold diluted into the refolding buffer (50 mM Tris pH 7.5, 5 mM EDTA, 10% glycerol, 10 mM reduced glutathione, 1 mM oxidised glutathione and Arginine) (Kopitar et al., 1996; Tobbell et al., 2002; D'Alessio et al., 2002 and Hwang et al., 2002). The refolding procedure was performed by dialysis at 4°C for 20 hours. After dialysis, the protein sample was concentrated by amicon centrifugal filters and dialysed against the refolding buffer without the redox pair components.

4.2.2 Recombinant expression and affinity purification of RCR3 in insects

Co-transfection

2 x 10^6 *Sf9* cells were seeded onto 6 cm plates, one plate for the lipofectin transfection and one as a negative control. The cells were allowed to attach to the plate surface at 27°C for 30 minutes. In the meantime the lipofectin mix (30 µl lipofectin in 1.5 ml IPL-41 medium w/o Fetal Calf Serum (FCS)) was added to the DNA mix (50 ng baculogold DNA and 1 µg transfer vector hosting the gene of interest in 1.5 ml IPL-41 w/o FCS) and incubated at room temperature for 30 minutes.

The medium was removed from the 6 cm plates and the cells were washed three times with 4 ml medium w/o FCS. 3 ml lipofectin-DNA mix was added to one plate, 3 ml IPL-41 w/o FCS to the negative control and both were incubated at 27°C for 5 hours. The supernatant was removed and the plates were washed twice with medium containing 5% (v/v) FCS and incubated in IPL-41 with 5% (v/v) FCS at 27°C for 5 to 6 days. The cells were checked for infection under the microscope. Infected cells are larger and their nuclei are enlarged in comparison to uninfected ones. After incubation the supernatant was transferred into a new tube and centrifugated at 1000 g at room temperature for 10 minutes. The supernatant was transferred to a new tube and stored

at 4°C. 5 ml ice-cold PBS was added to the plate and the attached cells were removed with a sterile cell scraper. This suspension of cells in Phosphate Buffer Saline (PBS) was transferred to the cell pellet from the first centrifugation. The cells were now centrifuged at 1000 g at room temperature for 5 min and the pellet was resuspended in 1 ml PBS and transferred to a microcentrifuge tube.

Virus amplification

1.8 x 10^6 *Sf21* cells were seeded onto 25 cm^3 tissue culture flasks. After 30 minutes incubation at 27°C, the medium was discarded and the cells were infected with 500 µl supernatant from the co-transfection. The sample as well as a negative control of cells in 500 µl IPL-41 with 5% (v/v) FCS were incubated at room temperature for 60 minutes. The supernatant was removed and the incubation was continued with 5 ml IPL-41 with 5% (v/v) FCS at 27°C for 3 days. Cells and supernatant were harvested as described above.

Intermediate stock

5 x 10^6 *Sf21* cells were seeded onto 75 cm^3 tissue culture flasks. After 30 minutes incubation at 27°C, the medium was discarded and the cells were infected with 300 µl virus amplification supernatant and 1.2 ml IPL-41 with 5% (v/v) FCS were added. The sample as well as a negative control of cells in 1.5 ml IPL-41 with 5% (v/v) FCS were incubated at room temperature for 60 minutes. The supernatant was removed and the incubation was continued with 15 ml IPL-41 with 5% (v/v) FCS at 27°C for 3 days. Cells and supernatant were harvested as described above.

Production

1 x 10^8 *Sf21* cells were infected with 10 ml intermediate stock supernatant in a falcon tube. The solution was incubated at room temperature for 60 minutes. During infection, the tube was agitated every 15 minutes very carefully to prevent cells from attaching to the wall. The cells were centrifuged at 1000 rpm at room temperature for 5 min and the supernatant was discarded. Cells were transferred with IPL-41 with 5% (v/v) FCS and 0.1% (w/v) Pluronic F-68 into the spinner flask and filled up to 200 ml with medium containing 5% (v/v) FCS and 0.1% (w/v) Pluronic F-68. The cells were

incubated in a spinner flask at 27°C and at 80 rpm for 3 days. The supernatant was harvested as described above and purified by Nickel affinity chromatography.

4.2.3 Recombinant expression of proteases in *yeast* and affinity purification

Pichia pastoris strain (Invitrogen) expressing full length $RCR3^{esc}$ or proRCR3 with C-terminal His_6 epitope tag were grown in BMGY medium (1% (w/v) Yeast extract, 2% (w/v) peptone, 1.34% (w/v) yeast nitrogen base, 1% (v/v) glycerol, 100 mM KH_2PO_4, pH 6.0) for 2 days until the OD_{600} reaches 30-60. After removing the cells by centrifugation at 1500g at room temperature, induction was inititated by introducing cells into the Buffered Methanol-complex Medium (BMMY) (1% (w/v) yeast extract, 2% (w/v) peptone, 1.34% (w/v) yeast nitrogen base, 1% (v/v) methanol, 100 mM KH_2PO_4, pH 6.0). The protein expressed in the medium was concentrated by amicon centrifugal filters. The tagged proteins were subsequently purified on a Ni-NTA superflow column (Qiagen; 30430) according to the manufacturer's protocol. The eluate fractions containing pure proteins were pooled. Protein concentrations were determined by the Bradford protein assay (Biorad).

4.2.4 Activation of the proRCR3 into mature RCR3

The proRCR3 was incubated in the presence of 100 mM NaOAc, pH 5.0 and 10 mM DTT for 60 – 75 minutes at 28°C.

4.2.5 Agroinfiltration of *Nicotiana benthamiana*

Agrobacterium tumefaciens strain GV3101 pMP90 transformed with binary vector containing recombinant tomato PLCP expression cassette was grown overnight at 28°C in 10 ml LB medium containing 50 µg/ml kanamycin and 50 µg/ml rifampicin in a 50 ml Falcon tube. The culture was centrifuged at 3000 g for 10 minutes and the bacterial pellet was 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 binary PLCP expression vectors were mixed at a 1:1 ratio with Agrobacterium suspensions containing binary expression vector for RNA silencing inhibitor p19 (Voinnet et al., 2003). The mixtures were kept at room temperature for 3 hours. The fully expanded leaves of 5 week-old *N. benthamiana* plants were punched with a needle and the

mixed Agrobacterium suspensions were infiltrated into the leaves through the needle-holes using a 1-ml syringe (Plastipak, BD) without a needle. Leaves were harvested after 5 days, and the proteins from the apoplast were extracted by vacuum-mediated apoplastic fluid isolation.

4.2.6 Apoplastic fluid isolation from *Nicotiana benthamiana* leaves

Agroinfiltrated mature leaves from 5 week-old *N.benthamiana* plants were submerged into 200 ml water in a 500 ml beaker underneath a grid and a moderate weight to avoid up-floating of the leaves. Water was vacuum-infiltrated into *N.benthamiana* leaf apoplast by pumping out air in an exicator for 5 minutes and releasing the vacuum slowly. water-infiltrated leaves were dried on the surface using tissue paper and centrifuged at 1600 g for 10 minutes at 4°C in a tube with holes (diameter = 1 mm) in the bottom (generic "AF isolation device"). The AF was collected below in a larger collection tube.

4.2.7 Quantification of protein concentration

Protein concentrations were quantified using the Bio-Rad protein assay. In brief, serial dilutions of BSA standard (Promega) of 0.05-1 mg/ml were prepared, and the protein samples were diluted to a concentration of 0.05-1 mg/ml. 10 µl of each concentration standard and each protein sample were loaded into wells of a 96-well microtiter plate (Microtest, Falcon), and mixed with 200 µl of 1:4-diluted dye reagent concentrate (Bio-Rad) and incubated at room temperature for 5 minutes. The absorbance at 595 nm was measured with the 680 Microplate reader (Bio-Rad), and the concentrations of the protein samples were deduced from the standard curve made from the standards made with different concentrations of BSA (6.25, 12.5, 25, 50, 100, 200, 400, 800 and 1000 µg/µl). Protein solutions were assayed in triplo.

4.2.8 Site-directed mutagenesis

Catalytic mutant proteins of RCR3, PIP1 and C14 were generated by introducing mutations using their parental plasmids (pSK016.01, pSK017.01, pSK018.01 and pSK019.01 for RCR3; pSK012.01, pSK013.01, pSK014.01 and pSK015.01 for PIP1; pSK009.01, pSK010.01, pSK011.01, pSK020.01, pSK021.01 and pSK022.01 for

C14) by using the Quick-change site-directed mutagenesis kit (Stratagene) using primers listed in Table 4.1. The 30-40 bp oligonucleotide primers were constructed such that one or two nucleotide changes result in a codon change that caused an amino acid substitution. Mutations were confirmed by sequencing and the confirmed sequence expression cassettes were shuttled into pFLAG-ATS (Sigma) or pTP5 (Shabab et al., 2008).

4.2.9 Activity-based labelling

Tomato apoplastic proteins

DCG-04 (MV201) labelling for 1D analysis was usually done by incubating 200-300 µl tomato AFs in 0.5 ml reaction volume containing 25 mM NaOAc (pH 5.0), 1 mM DTT and 2 µM DCG-04 or MV201 for 4 hours (in the dark for MV201). All labeling reactions were performed at room temperature under gentle agitation on a rotator (STR4, Stuart). Equal volumes of DMSO were added to the no-probe-controls. Proteins after labelling were precipitated by adding 1 ml -20°C acetone and subsequent centrifugation at 16000g for 1 minute with a tabletop centrifuge (5415D, Eppendorf). The protein pellet was dissolved in 2x SDS-PAGE loading buffer containing β-mercaptoethanol, and the proteins were separated on 12% SDS PAGE gels at 200 V using the Novex Minicell (Invitrogen). Competition or inhibition assays were done by preincubating the apoplastic fluids with competitor or inhibitor molecules for 30 minutes before labeling with activity-based probes.

4.2.10 In-gel fluorescence scanning

SDS-PAGE gel containing fluorescent probe-labeled proteins was washed three times with water and labeled proteins were visualized by in-gel fluorescence scanning using a Typhoon 8600 scanner (Molecular Dynamics) with excitation and emission at 532 and 580 nm, respectively. Fluorescent signals were quantified with ImageQuant 5.2 software (Molecular Dynamics).

4.2.11 Western blotting

After SDS PAGE, proteins were transferred onto polyvinylidene fluoride membrane (PVDF-Immobilon-P, Millipore) at 200 mA for 60-70 minutes using X-Cell II Blot

Modulesystem (Invitrogen). After the transfer, the membrane was moved into a 50 ml Falcon tube and blocked with 5 ml of 3% BSA (Biomol) solution for 5 minutes with gentle agitation on a roller mixer (SRT2, Stuart). For the detection of biotinylated proteins, the membrane was incubated with streptavidin-HRP enzyme (Ultrasensitive, Sigma) at 1:3000 in the presence of 2% Tween-20 for 1 hour, and then washed five times with Tris-buffered-saline (TBS) (50 mM Tris 500 mM NaCl, adjust pH with HCl to pH 7.6) plus0.1% Tween-20 for 5 minutes each time. For the detection of proteins with antibodies, the membrane was incubated with the protein-specific primary antibody (see materials 4.1.) at 1:5000 in the presence of 2% Tween-20 for 1 hour, and then washed 5 times with TBS buffer plus 0.1% Tween-20 for 5 minutes each time. Next, the membrane was incubated with HRP-conjugated anti-rabbit secondary antibodies (Amersham) at 1:5000 in the presence of 2% Tween-20 for 1 hour, and then washed five times with TBS buffer plus 0.1% Tween-20 for 5 minutes each time. For detection, the membrane was covered 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 the darkroom. The exposed film was developed by automatic X-ray film processor (Optimax, Protec).

4.2.12 BTH treatments

300 μM BTH (Bion, Syngenta) was sprayed onto tomato plants with a perfume sprayer (Roth) to the leaf surface until the droplets ran off and the plant samples were harvested at day 5.

4.2.13 RNA isolation, cDNA synthesis and analysis

Total RNA was isolated from tissues frozen in liquid nitrogen using the RNeasy plant mini kit (QIAGEN) according to the manufacturer's guidelines. DNase treatment was done before the RNA concentration was measured. cDNA was synthesised using Superscript II reverse transcriptase and Oligo dT primers according to the instructions of the manufacturer. cDNA was used as a template for PCR using primer pairs for gene-specific primers as summarized in Table 4.1

4.2.14 CD spectra

The CD spectrum was measured on Jasco J-810 at 25˚C in a quartz cuvette. All spectra were corrected for buffer contributions and observed values were converted to molar ellipticity. Each CD spectrum was accumulated from six scans at 195-250 nm with a 1nm slit width and a time constant of 1s for a nominal resolution of 1nm. Far UV CD spectra (2 mM) were collected in the range of wavelengths of 190-250 nm using a cell of path length 0.1 cm for monitoring secondary structure.

4.2.15 Fluorescence spectroscopy

The fluorescence spectra were measured on Jasco FP-750 spectrofluorometer at 25˚C in a quartz cuvette. All spectra were corrected for buffer contributions. The excitation wavelength were 280 nm and 274 nm, the emitted fluorescence were scanned between 290 nm to 450 nm. The protein concentration was 122 µg/ml. The samples were prepared in different pH buffers using Sartorius filters.

4.2.16 Enzyme kinetics

Bodipy-Casein

Papain was pre-incubated for 30 minutes with different concentrations of AVR2 or E-64 in the presence of 50 mM NaOAc, pH 5.0. The substrate Bodipy FL-casein (Molecular Probes) was added and the activity of the enzyme was measured immediately. The relative amount of the released bodipy was measured every two minutes in a fluorescence microplate reader (Synergy4, Biotek, Germany) with standard fluorescein excitation filters at 485±12.5nm and emission at 530±15 nm. All activity measurements were performed in triplicate.

Z-Leu-Arg-AMC (Z: N-carbobenzyloxy-7-Amino-4-methylcoumarin)

Determination of activation of proRCR3 to mature RCR3 was measured by measuring the activity of the enzyme immediately after incubation with the substrate Z-Arg-MCA (Bachem, Germany) in the presence of 100 mM NaOAc, pH 5.0 and 10 mM DTT. The relative amount of the released methyl coumarin was measured in a fluorescence microplate reader (Synergy4, Biotek, Germany) with standard

fluorescein excitation filters at 380 nm and emission at 460 nm. All activity measurements were performed in triplicate.

4.2.17 Crystallization screens

All crystallization experiments were performed at 18˚C or 4˚C using the vapour diffusion method and using 100 nL of protein and different precipitant types from commercial screens from the Joint Centre for Structural Genomics (JCSG I-IV screen) and Qiagen screens including Anions, Cations, Cryos, MPD and $AmSO_4$. The above screens were setup with different concentration of purified rAVR2 with and without epitope tags by using the automated robotic systems.

4.2.18 Lysine-methylation

The methylation reaction was performed overnight in 50 mM HEPES (pH 7.5), 250 mM NaCl at 1 mg/ml AVR2. Twenty micro liters freshly prepared 1M dimethylamine-borane complex (ABC; Fluka product 15584)and 40 µl 1 M formaldehyde (made from 37% stock; Fluka product 33220) were added per 1 mL of protein solution, and the reactions were gently mixed and incubated at 4°C for 2 hours. After the incubation, 20 µl ABC and 40 µl formaldehyde were added and the incubation was continued for 2 hours. This was followed by a final addition of 10 µl ABC and the reaction was incubated overnight at 4°C. In some cases, the reaction leads to a significant amount of precipitated protein which was removed by centrifugation before purification of the soluble methylated protein by size-exclusion chromatography columns (HiLoad 26/60 Superdex 75 pg column) on Akta express FPLC systems; GE Healthcare and purified fractions without aggregates were concentrated in amicon centrifugal concentrators (Vivascience) to appropriate concentrations. This protocol was from Walter, 2006 derived from (Rayment, 1997; Rayment et al., 1993).

4.2.19 Hypersensitive Response assay

Affinity purified AVR2 was concentrated using Amicon centrifugal filter unit (Millipore; 3 kDa MWCO) and diluted to 1 µM final concentration in double distilled

water before infiltration into tomato leaves of 4-5 week old tomato plants (*Cf-2/RCR3*). The plants were grown under regular greenhouse conditions to check for the HR response after three days post-infiltration. Buffer and water were included as negative controls. Other tomato genotypes (*Cf-2/rcr3-3*, *Cf-0/rcr3-3* and*Cf-0/RCR3*) were included.

Table 4 List of Primers

Identification[1]	Code[2]	Length (bp)[3]	Primer sequence[4]
Pathogen- derived inhibitors			
HIS_6-FXa-AVR2 (F)	r310	67	agctAAGCTTCATCACCATCACCATCACATCGAAGGTCGTGCCGCCAAAAAACTACCTGGCTGCGAC
AVR2 (R)	r402	36	agctGAATTCTCATCAACCGCAAAGACCAAAACAGC
HIS_6-FXa-EPIC1 (F)	r312	64	atgcAAGCTTCATCACCATCACCATCACATCGAAGGCCGACAAGTGGACGGCGGATACTCGAAG
EPIC1 (R)	r313	37	GCATGAATTCCTACTTAACTGGGGTAATCGACGTCAC
HIS_6-FXa-EPIC2B (F)	r314	65	atgcAAGCTTCATCACCATCACCATCACATCGAAGGCCGACAACTGAACGGATACTCAAAGAAGG
EPIC2B (R)	r315	37	GCATGAATTCTTAGTTGGCGGGCGTAATCGACGTCAC
HIS_6-FXa-RIP1 (F)	r316	64	atgcAAGCTTCATCACCATCACCATCACATCGAAGGCCGACAAACGCCCAAGAACATCGTTTCG
RIP1 (R)	r317	43	GCATGAATTCTCAGTTCACCGTGATTGCGCACTCGAAGGTCTG
HIS_6-TEV-AVR2 (F)	r347	94	agctAAGCTTCATCACCATCACCATCACGACTACGACATTCCTACTACGGAAAACCTATACTTCCAGGGC GCCAAAAAACTACCTGGCTGCGAC
HIS_6-TEV-EPIC1 (F)	r348	94	atgcAAGCTTCATCACCATCACCATCACGACTACGACATTCCTACTACGGAAAACCTATACTTCCAGGGCC AAGTGGACGGCGGATACTCGAAG
HIS_6-TEV-EPIC2B (F)	r349	95	atgcAAGCTTCATCACCATCACCATCACGACTACGACATTCCTACTACGGAAAACCTATACTTCCAGGGCC AACTGAACGGATACTCAAAGAAGG
HIS_6-TEV-RIP1 (F)	r350	94	atgcAAGCTTCATCACCATCACCATCACGACTACGACATTCCTACTACGGAAAACCTATACTTCCAGGGCC AAACGCCCAAGAACATCGTTTCG
Plant proteases (Tomato)			
HIS_6-FXa-pC14 (F)	r387	64	agctAAGCTTCATCACCATCACCATCACATCGAAGGCCGATCCGATATGTCAATTATTAGCTAC
HIS_6-FXa-C14 (F)	r388	76	agctAAGCTTCATCACCATCACCATCACATCGAAGGCCGATTGCCGGAATCAATTGACTGGAGAGAAAAA GGTGTG
HIS_6-TEV-pC14 (F)	r389	94	agctAAGCTTCATCACCATCACCATCACGACTACGACATTCCTACTACGGAAAACCTATACTTCCAGGGCT TGCCGGAATCAATTGACTGGAGAGAAAAAGGTGTG
HIS_6-TEV-C14 (F)	r390	105	agctAAGCTTCATCACCATCACCATCACGACTACGACATTCCTACTACGGAAAACCTATACTTCCAGGGCT CCGATATGTCAATTATTAGCTAC
HIS_6-FXa-pPIP1 (F)	r391	67	agctAAGCTTCATCACCATCACCATCACATCGAAGGCCGAGGCCGCAACTTAAAAGAATTATCCATG
HIS_6-FXa-PIP1 (F)	r392	64	agctAAGCTTCATCACCATCACCATCACATCGAAGGCCGAGAAGTTCCAAATAGCATGGACTGG
HIS_6-TEV-pPIP1 (F)	r393	94	agctAAGCTTCATCACCATCACCATCACGACTACGACATTCCTACTACGGAAAACCTATACTTCCAGGGCC GCAACTTAAAAGAATTATCCATG
HIS_6-TEV-PIP1 (F)	r394	94	agctAAGCTTCATCACCATCACCATCACGACTACGACATTCCTACTACGGAAAACCTATACTTCCAGGGCGA AGTTCCAAATAGCATGGACTGG
HIS_6-FXa-pRCR3 (F)	r395	67	agctAAGCTTCATCACCATCACCATCACATCGAAGGCCGAGGCCGCAGCCAGCCAAAACTGTCCGTG
HIS_6-FXa-RCR3 (F)	r396	64	agctAAGCTTCATCACCATCACCATCACATCGAAGGCCGAATGCCGTCTAACTTGGACTGGAGA
HIS_6-TEV-pRCR3 (F)	r397	94	agctAAGCTTCATCACCATCACCATCACGACTACGACATTCCTACTACGGAAAACCTATACTTCCAGGGCC GCAGCCAGCCAAAACTGTCCGTG
HIS_6-TEV-RCR3 (F)	r398	94	agctAAGCTTCATCACCATCACCATCACGACTACGACATTCCTACTACGGAAAACCTATACTTCCAGGGCAT GCCGTCTAACTTGGACTGGAGA
C14 (R)	r399	34	agctCTCGAGTCAAGAACTGCTCTTCTTTCCTCC
PIP1 (R)	r400	34	agctCTCGAGTCAAGCAGTAGGGAACGACGCAAC
RCR3 (R)	r401	35	agctCTCGAGCTACGCTATGTTTGGATAAGAAGAc
C14-C25A (F)	r403	38	GGAAGCTGTGGGAGTGCG TGGGCATTCTCTGCTGTTGC
C14-C25A (R)	r404	38	GCAACAGCAGAGAATGCCCACGCACTCCCACAGCTTCC
PIP1-C25A (F)	r405	35	GGTGTATGTGGATGTGCG TGGGCATTTTCTGCGGC
PIP1-C25A (R)	r406	35	GCCGCAGAAAATGCCCACGCACATCCACATACACC
RCR3-C25A (F)	r407	31	CGATGTGGATGTGCG TGGGCGTTTTCTGCAG
RCR3-C25A (R)	r408	31	CTGCAGAAAACGCCCACGCACATCCACATCG
C14-GRAN (R)	r409	40	agct CTCGAG ctaCTCTGTAGGTGGCTTGACCGGAGATGG
RCR3-C24A,C25A (F)	r490	41	CAAGGTCGATGTGGAGCTGCTGGGCGTTTTCTGCAGTTGG
RCR3-C24A,C25A (R)	r491	41	CCAACTGCAGAAAACGCCCAGGCAGCTCCACATCGACCTTG
RCR3-C24A (F)	r492	41	CAAGGTCGATGTGGAGCTTGCTGGGCGTTTTCTGCAGTTGG
RCR3-C24A (R)	r493	41	CCAACTGCAGAAAACGCCCAGCAAGCTCCACATCGACCTTG
RCR3-C24A,C25A (F)**	S04	45	agctTTTGTATTATTAGGAGCTGCCTGGGCGTTTTCTGCAGTTgg
RCR3-C24A,C25A (R)**	S05	45	agctAACTGCAGAAAACGCCCAGGCAGCTCCTAATAATACAAAcc
RCR3-C24A (F)**	S06	45	agctTTTGTATTATTAGGAGCTTGCTGGGCGTTTTCTGCAGTTgg
RCR3-C24A (R)**	S07	45	agctAACTGCAGAAAACGCCCAGCAAGCTCCTAATAATACAAAcc
RCR3-C25A (F)**	S08	45	agctTTTGTATTATTAGGATGTGCCTGGGCGTTTTCTGCAGTTgg
RCR3-C25A (R)**	S09	45	agctAACTGCAGAAAACGCCCAGGCACATCCTAATAATACAAAgg

[1]Forward (F) and reverse primers (R); [2] Primer library code; [3]Length indicated in base pairs

[4] Primer sequences; mutagenesized primers are in red, mutagenesized codons are underlined, [**]Primers formutagenizing agroinfiltration vectors

Table 5 List of RCR3 mutant Agrobacterium constructs

RCR3 mutants cloned into pTP05 vector for expression in *Agrobacterium tumefaciens*		
Description[1]	Intermediate clone in pTP25[2]	Binary clone in pTP05[3]
RCR3		
$RCR3^{tyc}$ (C24AC25A)	pSK 039.01	pSK 042.01
$RCR3^{tyc}$ (C24A)	pSK 040.01	pSK 043.01
$RCR3^{tyc}$ (C24A)	pSK 040.02	pSK 043.02
$RCR3^{tyc}$ (C25A)	pSK 041.01	pSK 044.01

[1]Description of the RCR3 mutant construct (C24AC25A, C24A,C25A) in the background of *Solanum lycospersicum*; [2]Intermediate clone where the pTP 05 was used as shuttling vector for RCR3 mutant plasmids; [3]Binary clone of the RCR3 mutant plasmids using pTP05 for Agrobacterium-mediated expression.

REFERENCES

Anastasi A., Brown M.A., Kembhavi A.A., Nicklin M.J., Sayers C.A., Sunter D.C., Barrett A.J. (1983) Cystatin, a protein inhibitor of cysteine proteinases. Improved purification from egg white, characterization, and detection in chicken serum. *The Biochemical journal***211**:129-38.

Barlic-Maganja D., Dolinar M., Turk V. (1998) The influence of Ala205 on the specificity of cathepsin L produced by dextran sulfate assisted activation of the recombinant proenzyme. *Biological chemistry***379**:1449-52.

Barrett A.J., Kembhavi A.A., Brown M.A., Kirschke H., Knight C.G., Tamai M., Hanada K. (1982) L-trans-Epoxysuccinyl-leucylamido(4-guanidino)butane (E-64) and its analogues as inhibitors of cysteine proteinases including cathepsins B, H and L. *The Biochemical journal***201**:189-98.

Baud F., Karlin S. (1999) Measures of residue density in protein structures. *Proceedings of the National Academy of Sciences of the United States of America***96**:12494-9.

Beers E.P., Jones A.M., Dickerman A.W. (2004) The S8 serine, C1A cysteine and A1 aspartic protease families in Arabidopsis. *Phytochemistry***65**:43-58.

Bendall M.R., Cartwright I.L., Lowe G., Nurse D., Clark P.I. (1977) Inhibition of Papain by N-Acyl-Aminoacetaldehydes and N-Acyl-Aminopropanones - Evidence for Hemithioacetal Formation by a Cross-Saturation Technique in Nmr-Spectroscopy. *European Journal of Biochemistry***79**:201-209.

Bjork I., Alriksson E., Ylinenjarvi K. (1989) Kinetics of binding of chicken cystatin to papain. *Biochemistry***28**:1568-73.

Blumberg S., Schechter I., Berger A. (1970) The purification of papain by affinity chromatography. *European journal of biochemistry / FEBS***15**:97-102.

Bode W., Engh R., Musil D., Thiele U., Huber R., Karshikov A., Brzin J., Kos J., Turk V. (1988) The 2.0 A X-ray crystal structure of chicken egg white cystatin and its possible mode of interaction with cysteine proteinases. *The EMBO journal***7**:2593-9.

Borutaite V., Brown G.C. (2001) Caspases are reversibly inactivated by hydrogen peroxide. *FEBS letters***500**:114-8.

Bromme D., Nallaseth F.S., Turk B. (2004) Production and activation of recombinant papain-like cysteine proteases. *Methods***32**:199-206.

Bromme D., Okamoto K. (1995) The baculovirus cysteine protease has a cathepsin B-like S2-subsite specificity. *Biological chemistry Hoppe-Seyler***376**:611-5.

Caffrey C.R., Salter J.P., Lucas K.D., Khiem D., Hsieh I., Lim K.C., Ruppel A., McKerrow J.H., Sajid M. (2002) SmCB2, a novel tegumental cathepsin B from adult Schistosoma mansoni. *Molecular and biochemical parasitology***121**:49-61.

Campbell D.A., Szardenings A.K. (2003) Functional profiling of the proteome with affinity labels. *Current opinion in chemical biology***7**:296-303.

Canaves J.M., Page R., Wilson I.A., Stevens R.C. (2004) Protein biophysical properties that correlate with crystallization success in Thermotoga maritima: maximum clustering strategy for structural genomics. *Journal of molecular biology***344**:977-91. DOI: 10.1016/j.jmb.2004.09.076.

Carmona E., Dufour E., Plouffe C., Takebe S., Mason P., Mort J.S., Menard R. (1996) Potency and selectivity of the cathepsin L propeptide as an inhibitor of cysteine proteases. *Biochemistry***35**:8149-57. DOI: 10.1021/bi952736s.

Cereghino J.L., Cregg J.M. (2000) Heterologous protein expression in the methylotrophic yeast Pichia pastoris. *FEMS microbiology reviews***24**:45-66.

Chan V.J., Selzer P.M., McKerrow J.H., Sakanari J.A. (1999) Expression and alteration of the S2 subsite of the Leishmania major cathepsin B-like cysteine protease. *The Biochemical journal***340 (Pt 1)**:113-7.

Cheng Y.S. (1983) Increased cell buoyant densities of protein overproducing Escherichia coli cells. *Biochemical and biophysical research communications***111**:104-11.

Chinchilla D., Bauer Z., Regenass M., Boller T., Felix G. (2006) The Arabidopsis receptor kinase FLS2 binds flg22 and determines the specificity of flagellin perception. *The Plant cell***18**:465-76. DOI: 10.1105/tpc.105.036574.

Chisholm S.T., Coaker G., Day B., Staskawicz B.J. (2006) Host-microbe interactions: shaping the evolution of the plant immune response. *Cell***124**:803-14. DOI: 10.1016/j.cell.2006.02.008.

Christendat D., Yee A., Dharamsi A., Kluger Y., Savchenko A., Cort J.R., Booth V., Mackereth C.D., Saridakis V., Ekiel I., Kozlov G., Maxwell K.L., Wu N., McIntosh L.P., Gehring K., Kennedy M.A., Davidson A.R., Pai E.F., Gerstein M., Edwards A.M., Arrowsmith C.H. (2000) Structural proteomics of an archaeon. *Nature structural biology***7**:903-9. DOI: 10.1038/82823.

Coulombe R., Grochulski P., Sivaraman J., Menard R., Mort J.S., Cygler M. (1996) Structure of human procathepsin L reveals the molecular basis of inhibition by the prosegment. *The EMBO journal***15**:5492-503.

Cravatt B.F., Wright A.T., Kozarich J.W. (2008) Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annual review of biochemistry***77**:383-414. DOI: 10.1146/annurev.biochem.75.101304.124125.

Cregg J.M., Vedvick T.S., Raschke W.C. (1993) Recent advances in the expression of foreign genes in Pichia pastoris. *Bio/technology***11**:905-10.

Cygler M., Mort J.S. (1997) Proregion structure of members of the papain superfamily. Mode of inhibition of enzymatic activity. *Biochimie***79**:645-52.

Cygler M., Sivaraman J., Grochulski P., Coulombe R., Storer A.C., Mort J.S. (1996) Structure of rat procathepsin B: model for inhibition of cysteine protease activity by the proregion. *Structure***4**:405-16.

D'Alessio K J., McQueney M.S., Brun K.A., Orsini M.J., Debouck C.M. (1999) Expression in Escherichia coli, refolding, and purification of human procathepsin K, an osteoclast-specific protease. *Protein expression and purification***15**:213-20. DOI: 10.1006/prep.1998.1013.

Dawson J.E., Seckute J., De S., Schueler S.A., Oswald A.B., Nicholson L.K. (2009) Elucidation of a pH-folding switch in the Pseudomonas syringae effector protein AvrPto. *Proceedings of the National Academy of Sciences of the United States of America***106**:8543-8. DOI: 10.1073/pnas.0809138106.

de Kock M.J.D., Brandwagt B.F., Bonnema G., de Wit P.J.G.M., Lindhout P. (2005) The tomato Orion locus comprises a unique class of Hcr9 genes. *Molecular Breeding***15**:409-422. DOI: DOI 10.1007/s11032-005-0386-8.

de Wit P.J., Lauge R., Honee G., Joosten M.H., Vossen P., Kooman-Gersmann M., Vogelsang R., Vervoort J.J. (1997) Molecular and biochemical basis of the interaction between tomato and its fungal pathogen Cladosporium fulvum. *Antonie van Leeuwenhoek***71**:137-41.

Derewenda Z.S., Vekilov P.G. (2006) Entropy and surface engineering in protein crystallization. *Acta crystallographica. Section D, Biological crystallography***62**:116-24. DOI: 10.1107/S0907444905035237.

Dieckmann T., Mitschang L., Hofmann M., Kos J., Turk V., Auerswald E.A., Jaenicke R., Oschkinat H. (1993) The structures of native phosphorylated chicken cystatin and of a recombinant unphosphorylated variant in solution. *Journal of molecular biology***234**:1048-59. DOI: 10.1006/jmbi.1993.1658.

DoConte L.L., Chothia C.J., J. (1999). *Journal of Molecular Biology***285**:2177-2198.

Dolinar M., Maganja D.B., Turk V. (1995) Expression of full-length human procathepsin L cDNA in Escherichia coli and refolding of the expression product. *Biological chemistry Hoppe-Seyler***376**:385-8.

Eakin A.E., McGrath M.E., McKerrow J.H., Fletterick R.J., Craik C.S. (1993) Production of crystallizable cruzain, the major cysteine protease from Trypanosoma cruzi. *The Journal of biological chemistry***268**:6115-8.

Eckart M.R., Bussineau C.M. (1996) Quality and authenticity of heterologous proteins synthesized in yeast. *Current opinion in biotechnology***7**:525-30.

Engh R.A., Dieckmann T., Bode W., Auerswald E.A., Turk V., Huber R., Oschkinat H. (1993) Conformational variability of chicken cystatin. Comparison of structures determined by X-ray diffraction and NMR spectroscopy. *Journal of molecular biology***234**:1060-9. DOI: 10.1006/jmbi.1993.1659.

Evans M.J., Cravatt B.F. (2006) Mechanism-based profiling of enzyme families. *Chemical reviews***106**:3279-301. DOI: 10.1021/cr050288g.

Felix G., Duran J.D., Volko S., Boller T. (1999) Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *The Plant journal : for cell and molecular biology***18**:265-76.

Flor H. (1942) Inheritance of pathogenicity in a cross between physiologic races 22 and 24 of Melampsora lini. *Phytopathology***32**:653-69.

Fogg M.J., Alzari P., Bahar M., Bertini I., Betton J.M., Burmeister W.P., Cambillau C., Canard B., Corrondo M.A., Coll M., Daenke S., Dym O., Egloff M.P., Enguita F.J., Geerlof A., Haouz A., Jones T.A., Ma Q., Manicka S.N., Migliardi M., Nordlund P., Owens R.J., Peleg Y., Schneider G., Schnell R., Stuart D.I., Tarbouriech N., Unge T., Wilkinson A.J., Wilmanns M., Wilson K.S., Zimhony O., Grimes J.M. (2006) Application of the use of high-throughput technologies to the determination of protein structures of bacterial and viral pathogens. *Acta crystallographica. Section D, Biological crystallography***62**:1196-207. DOI: 10.1107/S0907444906030915.

Foyer C., Rowell J., Walker D. (1983) Measurement of the Ascorbate Content of Spinach Leaf Protoplasts and Chloroplasts during Illumination. *Planta***157**:239-244.

Georgiou G., Telford J.N., Shuler M.L., Wilson D.B. (1986) Localization of inclusion bodies in Escherichia coli overproducing beta-lactamase or alkaline phosphatase. *Applied and environmental microbiology***52**:1157-61.

Giancotti V., Quadrifoglio F., Cowgill R.W., Crane-Robinson C. (1980) Fluorescence of buried tyrosine residues in proteins. *Biochimica et biophysica acta***624**:60-5.

Gilroy E.M., Hein I., van der Hoorn R., Boevink P.C., Venter E., McLellan H., Kaffarnik F., Hrubikova K., Shaw J., Holeva M., Lopez E.C., Borras-Hidalgo O., Pritchard L., Loake G.J., Lacomme C., Birch P.R. (2007) Involvement of cathepsin B in the plant disease resistance hypersensitive response. *The Plant journal : for cell and molecular biology***52**:1-13. DOI: 10.1111/j.1365-313X.2007.03226.x.

Glazer A.N., Simmons N.S. (1996). *Journal of Amer.Chem.Soc.*,:2335-2336.

Glick D.M., Shalitin Y., Hilt C.R. (1989) Studies on the Irreversible Step of Pepsinogen Activation. *Biochemistry***28**:2626-2630.

Graslund S., Nordlund P., Weigelt J., Hallberg B.M., Bray J., Gileadi O., Knapp S., Oppermann U., Arrowsmith C., Hui R., Ming J., dhe-Paganon S., Park H.W., Savchenko A., Yee A., Edwards A., Vincentelli R., Cambillau C., Kim R., Kim S.H., Rao Z., Shi Y., Terwilliger T.C., Kim C.Y., Hung L.W., Waldo G.S., Peleg Y., Albeck S., Unger T., Dym O., Prilusky J., Sussman J.L., Stevens R.C., Lesley S.A., Wilson I.A., Joachimiak A., Collart F., Dementieva I., Donnelly M.I., Eschenfeldt W.H., Kim Y., Stols L., Wu R., Zhou M., Burley S.K., Emtage J.S., Sauder J.M., Thompson D., Bain K., Luz J., Gheyi T., Zhang F., Atwell S., Almo S.C., Bonanno J.B., Fiser A., Swaminathan S., Studier F.W., Chance M.R., Sali A., Acton T.B., Xiao R., Zhao L., Ma L.C., Hunt J.F., Tong L., Cunningham K., Inouye M., Anderson S., Janjua H., Shastry R., Ho C.K., Wang D., Wang H., Jiang M., Montelione G.T., Stuart D.I., Owens R.J., Daenke S., Schutz A., Heinemann U., Yokoyama S., Bussow K., Gunsalus K.C. (2008) Protein production and purification. *Nature methods***5**:135-46. DOI: 10.1038/nmeth.f.202.

Graziani M.T., Agro A.F., Rotilio G., Barra D., Mondovi B. (1974) Parsley plastocyanin. The possible presence of sulfhydryl and tyrosine in the copper environment. *Biochemistry***13**:804-9.

Greenbaum D., Medzihradszky K.F., Burlingame A., Bogyo M. (2000) Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chemistry & biology***7**:569-581.

Greenfield N.J. (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nature protocols***1**:2876-90. DOI: 10.1038/nprot.2006.202.

Groves M.R., Taylor M.A., Scott M., Cummings N.J., Pickersgill R.W., Jenkins J.A. (1996) The prosequence of procaricain forms an alpha-helical domain that prevents access to the substrate-binding cleft. *Structure***4**:1193-203.

Gu C., Shabab M., Strasser R., Wolters P.J., Shindo T., Niemer M., Kaschani F., Mach L., van der Hoorn R.A. (2012) Post-translational regulation and trafficking of the granulin-containing protease RD21 of Arabidopsis thaliana. *PloS one***7**:e32422. DOI: 10.1371/journal.pone.0032422.

Gutierrez-Gonzalez L.H., Rojo-Dominguez A., Cabrera-Gonzalez N.E., Perez-Montfort R., Padilla-Zuniga A.J. (2006) Loosely packed papain prosegment displays inhibitory activity. *Archives of biochemistry and biophysics***446**:151-60. DOI: 10.1016/j.abb.2005.12.005.

Haas B.J., Kamoun S., Zody M.C., Jiang R.H., Handsaker R.E., Cano L.M., Grabherr M., Kodira C.D., Raffaele S., Torto-Alalibo T., Bozkurt T.O., Ah-Fong A.M., Alvarado L., Anderson V.L., Armstrong M.R., Avrova A., Baxter L., Beynon J., Boevink P.C., Bollmann S.R., Bos J.I., Bulone V., Cai G., Cakir C., Carrington J.C., Chawner M., Conti L., Costanzo S., Ewan R., Fahlgren N., Fischbach M.A., Fugelstad J., Gilroy E.M., Gnerre S., Green P.J., Grenville-Briggs L.J., Griffith J., Grunwald N.J., Horn K., Horner N.R., Hu C.H., Huitema E., Jeong D.H., Jones A.M., Jones J.D., Jones R.W., Karlsson E.K., Kunjeti S.G., Lamour K., Liu Z., Ma L., Maclean D., Chibucos M.C., McDonald H., McWalters J., Meijer H.J., Morgan W., Morris P.F., Munro C.A., O'Neill K., Ospina-Giraldo M., Pinzon A., Pritchard L., Ramsahoye B., Ren Q., Restrepo S., Roy S., Sadanandom A., Savidor A., Schornack S., Schwartz D.C., Schumann U.D., Schwessinger B., Seyer L., Sharpe T., Silvar C., Song J., Studholme D.J., Sykes S., Thines M., van de Vondervoort P.J., Phuntumart V., Wawra S., Weide R., Win J., Young C., Zhou S., Fry W., Meyers B.C., van West P., Ristaino J., Govers F., Birch P.R., Whisson S.C., Judelson H.S., Nusbaum C. (2009) Genome sequence and analysis of the Irish potato famine pathogen Phytophthora infestans. *Nature***461**:393-8. DOI: 10.1038/nature08358.

Hammerschmidt R. (2010) The dynamic apoplast. *Physiological and Molecular Plant Pathology***74**:199-200. DOI: DOI 10.1016/j.pmpp.2010.05.003.

Hayashi Y., Yamada K., Shimada T., Matsushima R., Nishizawa N.K., Nishimura M., Hara-Nishimura I. (2001) A proteinase-storing body that prepares for cell death or stresses in the epidermal cells of Arabidopsis. *Plant & cell physiology***42**:894-9.

Heimo H., Palmu K., Suominen I. (1997) Expression in Pichia pastoris and purification of Aspergillus awamori glucoamylase catalytic domain. *Protein expression and purification***10**:70-9. DOI: 10.1006/prep.1996.0713.

Homans S.W. (2004) NMR spectroscopy tools for structure-aided drug design. *Angewandte Chemie***43**:290-300. DOI: 10.1002/anie.200300581.

Hope J., Shearman M.S., Baxter H.C., Chong A., Kelly S.M., Price N.C. (1996) Cytotoxicity of prion protein peptide (PrP106-126) differs in mechanism from the cytotoxic activity of the Alzheimer's disease amyloid peptide, A beta 25-35. *Neurodegeneration : a journal for neurodegenerative disorders, neuroprotection, and neuroregeneration***5**:1-11.

Hwang H.S., Chung H.S. (2002) Preparation of active recombinant cathepsin K expressed in bacteria as inclusion body. *Protein expression and purification***25**:541-6.

Iwata S., Ostermeier C., Ludwig B., Michel H. (1995) Structure at 2.8 A resolution of cytochrome c oxidase from Paracoccus denitrificans. *Nature***376**:660-9. DOI: 10.1038/376660a0.

Jackson G.S., Hosszu L.L., Power A., Hill A.F., Kenney J., Saibil H., Craven C.J., Waltho J.P., Clarke A.R., Collinge J. (1999) Reversible conversion of monomeric human prion protein between native and fibrilogenic conformations. *Science***283**:1935-7.

Jones J.D., Dangl J.L. (2006) The plant immune system. *Nature***444**:323-9. DOI: 10.1038/nature05286.

Joosten M., de Wit P. (1999) THE TOMATO-CLADOSPORIUM FULVUM INTERACTION: A Versatile Experimental System to Study Plant-Pathogen Interactions. *Annual review of phytopathology***37**:335-367. DOI: 10.1146/annurev.phyto.37.1.335.

Joosten M.H. (2012) Isolation of apoplastic fluid from leaf tissue by the vacuum infiltration-centrifugation technique. *Methods in molecular biology***835**:603-10. DOI: 10.1007/978-1-61779-501-5_38.

Jordano J., Barbero J.L., Montero F., Franco L. (1983) Fluorescence of histones H1. A tyrosinate-like fluorescence emission in Ceratitis capitata H1 at neutral pH values. *The Journal of biological chemistry***258**:315-20.

Kane J.F., Hartley D.L. (1988) Formation of Recombinant Protein Inclusion-Bodies in Escherichia-Coli. *Trends in Biotechnology***6**:95-101.

Kapila J., DeRycke R., VanMontagu M., Angenon G. (1997) An Agrobacterium-mediated transient gene expression system for intact leaves (vol 122, pg 101, 1997). *Plant Science***124**:227-227.

Kaschani F., Shabab M., Bozkurt T., Shindo T., Schornack S., Gu C., Ilyas M., Win J., Kamoun S., van der Hoorn R.A. (2010) An effector-targeted protease contributes to defense against Phytophthora infestans and is under diversifying selection in natural hosts. *Plant physiology***154**:1794-804. DOI: 10.1104/pp.110.158030.

Kaschani F., van der Hoorn R. (unpublished).

Kaschani F., Van der Hoorn R.A. (2011) A model of the C14-EPIC complex indicates hotspots for a protease-inhibitor arms race in the oomycete-potato interaction. *Plant signaling & behavior***6**:109-12.

Kaulmann G., Palm G.J., Schilling K., Hilgenfeld R., Wiederanders B. (2006) The crystal structure of a Cys25 -> Ala mutant of human procathepsin S elucidates enzyme-prosequence interactions. *Protein science : a publication of the Protein Society***15**:2619-29. DOI: 10.1110/ps.062401806.

Kelly S.M., Price N.C. (1997) The application of circular dichroism to studies of protein folding and unfolding. *Biochimica et biophysica acta***1338**:161-85.

Khouri H.E., Plouffe C., Hasnain S., Hirama T., Storer A.C., Menard R. (1991) A model to explain the pH-dependent specificity of cathepsin B-catalysed hydrolyses. *The Biochemical journal***275 (Pt 3)**:751-7.

Kim M.J., Yamamoto D., Matsumoto K., Inoue M., Ishida T., Mizuno H., Sumiya S., Kitamura K. (1992) Crystal structure of papain-E64-c complex. Binding diversity of E64-c to papain S2 and S3 subsites. *The Biochemical journal***287 (Pt 3)**:797-803.

Klose J., Wendt N., Kubald S., Krause E., Fechner K., Beyermann M., Bienert M., Rudolph R., Rothemund S. (2004) Hexa-histidin tag position influences disulfide structure but not binding behavior of in vitro folded N-terminal domain of rat corticotropin-releasing factor receptor type 2a. *Protein science : a publication of the Protein Society***13**:2470-5. DOI: 10.1110/ps.04835904.

Kolodziejek I., van der Hoorn R.A. (2010) Mining the active proteome in plant science and biotechnology. *Current opinion in biotechnology***21**:225-33. DOI: 10.1016/j.copbio.2010.02.003.

Kopitar G., Dolinar M., Strukelj B., Pungercar J., Turk V. (1996) Folding and activation of human procathepsin S from inclusion bodies produced in Escherichia coli. *European journal of biochemistry / FEBS***236**:558-62.

Kruger J., Thomas C.M., Golstein C., Dixon M.S., Smoker M., Tang S., Mulder L., Jones J.D. (2002) A tomato cysteine protease required for Cf-2-dependent disease resistance and suppression of autonecrosis. *Science***296**:744-7. DOI: 10.1126/science.1069288.

Kuhelj R., Dolinar M., Pungercar J., Turk V. (1995) The preparation of catalytically active human cathepsin B from its precursor expressed in Escherichia coli in the form of inclusion bodies. *European journal of biochemistry / FEBS***229**:533-9.

LaLonde J.M., Zhao B., Janson C.A., D'Alessio K.J., McQueney M.S., Orsini M.J., Debouck C.M., Smith W.W. (1999) The crystal structure of human procathepsin K. *Biochemistry***38**:862-9. DOI: 10.1021/bi9822271.

Lauge R., Goodwin P.H., de Wit P.J.G.M., Joosten M.H.A.J. (2000) Specific HR-associated recognition of secreted proteins from Cladosporium fulvum occurs in both host and non-host plants. *Plant Journal***23**:735-745.

Liang T.C., Abeles R.H. (1987) Inhibition of papain by nitriles: mechanistic studies using NMR and kinetic measurements. *Archives of biochemistry and biophysics***252**:626-34.

Lim B.T., Kimura T. (1980) Conformation-associated anomalous tyrosine fluorescence of adrenodoxin. *The Journal of biological chemistry***255**:2440-4.

Lindahl P., Abrahamson M., Bjork I. (1992) Interaction of recombinant human cystatin C with the cysteine proteinases papain and actinidin. *The Biochemical journal***281 (Pt 1)**:49-55.

Lindahl P., Alriksson E., Jornvall H., Bjork I. (1988) Interaction of the cysteine proteinase inhibitor chicken cystatin with papain. *Biochemistry***27**:5074-82.

Ljunggren A., Redzynia I., Alvarez-Fernandez M., Abrahamson M., Mort J.S., Krupa J.C., Jaskolski M., Bujacz G. (2007) Crystal structure of the parasite protease inhibitor chagasin in complex with a host target cysteine protease. *Journal of molecular biology***371**:137-53. DOI: 10.1016/j.jmb.2007.05.005.

Lohaus G., Pennewiss K., Sattelmacher B., Hussmann M., Hermann Muehling K. (2001) Is the infiltration-centrifugation technique appropriate for the isolation of apoplastic fluid? A critical evaluation with different plant species. *Physiologia plantarum***111**:457-465.

Longworth J.W. (1981) A new component in protein fluorescence. *Annals of the New York Academy of Sciences***366**:237-45.

Mach L., Mort J.S., Glossl J. (1994) Maturation of human procathepsin B. Proenzyme activation and proteolytic processing of the precursor to the mature proteinase, in vitro, are primarily unimolecular processes. *The Journal of biological chemistry***269**:13030-5.

MacManus J.P., Szabo A.G., Williams R.E. (1984) Conformational changes induced by binding of bivalent cations to oncomodulin, a paravalbumin-like tumour protein. *The Biochemical journal***220**:261-8.

Maeda Y., Koga H., Yamada H., Ueda T., Imoto T. (1995) Effective renaturation of reduced lysozyme by gentle removal of urea. *Protein engineering***8**:201-5.

Marston F.A.O. (1986) The Purification of Eukaryotic Polypeptides Synthesized in Escherichia-Coli. *Biochemical Journal***240**:1-12.

Martin J.R., Craven C.J., Jerala R., Kroon-Zitko L., Zerovnik E., Turk V., Waltho J.P. (1995) The three-dimensional solution structure of human stefin A. *Journal of molecular biology***246**:331-43.

Mason S.D., Joyce J.A. (2011) Proteolytic networks in cancer. *Trends in cell biology***21**:228-37. DOI: 10.1016/j.tcb.2010.12.002.

McDonald J.K., Emerick J.M. (1995) Purification and characterization of procathepsin L, a self-processing zymogen of guinea pig spermatozoa that acts on a cathepsin D assay substrate. *Archives of biochemistry and biophysics***323**:409-22.

Means G.E., Feeney R.E. (1968) Reductive alkylation of amino groups in proteins. *Biochemistry***7**:2192-201.

Misas-Villamil J.C., van der Hoorn R.A. (2008) Enzyme-inhibitor interactions at the plant-pathogen interface. *Current opinion in plant biology***11**:380-8. DOI: 10.1016/j.pbi.2008.04.007.

Missiakas D., Raina S. (1997) Protein folding in the bacterial periplasm. *Journal of bacteriology***179**:2465-71.

Mitraki A., King J. (1989) Protein Folding Intermediates and Inclusion Body Formation. *Bio-Technology***7**:690-697.

Mochizuki S., Hamato N., Hirose M., Miyano K., Ohtani W., Kameyama S., Kuwae S., Tokuyama T., Ohi H. (2001) Expression and characterization of recombinant human

antithrombin III in Pichia pastoris. *Protein expression and purification***23**:55-65. DOI: 10.1006/prep.2001.1479.

Nagler D.K., Zhang R., Tam W., Sulea T., Purisima E.O., Menard R. (1999) Human cathepsin X: A cysteine protease with unique carboxypeptidase activity. *Biochemistry***38**:12648-54.

Nicklin M.J., Barrett A.J. (1984) Inhibition of cysteine proteinases and dipeptidyl peptidase I by egg-white cystatin. *The Biochemical journal***223**:245-53.

Nishimura Y., Kawabata T., Furuno K., Kato K. (1989) Evidence that aspartic proteinase is involved in the proteolytic processing event of procathepsin L in lysosomes. *Archives of biochemistry and biophysics***271**:400-6.

Nishimura Y., Kawabata T., Kato K. (1988) Identification of latent procathepsins B and L in microsomal lumen: characterization of enzymatic activation and proteolytic processing in vitro. *Archives of biochemistry and biophysics***261**:64-71.

Nomura T., Fujisawa Y. (1997) Processing properties of recombinant human procathepsin L. *Biochemical and biophysical research communications***230**:143-6. DOI: 10.1006/bbrc.1996.5905.

Pace C.N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods in enzymology***131**:266-80.

Park H., Hong K.M., Sakanari J.A., Choi J.H., Park S.K., Kim K.Y., Hwang H.A., Paik M.K., Yun K.J., Shin C.H., Lee J.B., Ryu J.S., Min D.Y. (2001) Paragonimus westermani: cloning of a cathepsin F-like cysteine proteinase from the adult worm. *Experimental parasitology***98**:223-7. DOI: 10.1006/expr.2001.4634.

Price N.C., Stevens L. (1999) Fundamentals of Enzymology :Cell Biology of Catalytic proteins. *3rd edition, Oxford University Press*, :pp. 47-117.

Provencher S.W., Glockner J. (1981) Estimation of globular protein secondary structure from circular dichroism. *Biochemistry***20**:33-7.

Pugsley A.P. (1993) The complete general secretory pathway in gram-negative bacteria. *Microbiological reviews***57**:50-108.

Rawlings N.D., Morton F.R., Barrett A.J. (2006) MEROPS: the peptidase database. *Nucleic acids research***34**:D270-2. DOI: 10.1093/nar/gkj089.

Raychaudhuri S., Choudhury K.R., Palchoudhuri S., Chopra S., Bhattacharyya N.P., Mukhopadhyay D. (2011) Spectroscopic studies reveal conformational flexibility of intrinsically unstructured protein HYPK. *Journal of Biophysical Chemistry***2**:434-442.

Raychaudhuri S., Majumder P., Sarkar S., Giri K., Mukhopadhyay D., Bhattacharyya N.P. (2008) Huntingtin interacting protein HYPK is intrinsically unstructured. *Proteins***71**:1686-98. DOI: 10.1002/prot.21856.

Rayment I., Rypniewski W.R., Schmidt-Base K., Smith R., Tomchick D.R., Benning M.M., Winkelmann D.A., Wesenberg G., Holden H.M. (1993) Three-dimensional structure of myosin subfragment-1: a molecular motor. *Science***261**:50-8.

Redzynia I., Ljunggren A., Abrahamson M., Mort J.S., Krupa J.C., Jaskolski M., Bujacz G. (2008) Displacement of the occluding loop by the parasite protein, chagasin, results in efficient inhibition of human cathepsin B. *The Journal of biological chemistry***283**:22815-25. DOI: 10.1074/jbc.M802064200.

Redzynia I., Ljunggren A., Bujacz A., Abrahamson M., Jaskolski M., Bujacz G. (2009) Crystal structure of the parasite inhibitor chagasin in complex with papain allows identification of structural requirements for broad reactivity and specificity determinants for target proteases. *The FEBS journal***276**:793-806. DOI: 10.1111/j.1742-4658.2008.06824.x.

Ren J., Esnouf R., Garman E., Somers D., Ross C., Kirby I., Keeling J., Darby G., Jones Y., Stuart D., et al. (1995) High resolution structures of HIV-1 RT from four RT-inhibitor complexes. *Nature structural biology***2**:293-302.

Revell D.F., Cummings N.J., Baker K.C., Collins M.E., Taylor M.A., Sumner I.G., Pickersgill R.W., Connerton I.F., Goodenough P.W. (1993) Nucleotide sequence and expression in Escherichia coli of cDNAs encoding papaya proteinase omega from Carica papaya. *Gene***127**:221-5.

Rice R.H., Means G.E. (1971) Radioactive labeling of proteins in vitro. *The Journal of biological chemistry***246**:831-2.

Rich D.H. (1986) in Proteinaceous Inhibitors (Barrett, A. J., and Salvesen, G., Eds). *Elsevier, Amsterdam*:153-178.

Richau K.H., Kaschani F., Verdoes M., Pansuriya T.C., Niessen S., Stuber K., Colby T., Overkleeft H.S., Bogyo M., Van der Hoorn R.A. (2012) Subclassification and biochemical analysis of plant papain-like cysteine proteases displays subfamily-specific characteristics. *Plant physiology***158**:1583-99. DOI: 10.1104/pp.112.194001.

Rivas S., Thomas C.M. (2002) Recent advances in the study of tomato Cf resistance genes. *Molecular plant pathology***3**:277-82.

Rooney H.C., Van't Klooster J.W., van der Hoorn R.A., Joosten M.H., Jones J.D., de Wit P.J. (2005) Cladosporium Avr2 inhibits tomato Rcr3 protease required for Cf-2-dependent disease resistance. *Science***308**:1783-6. DOI: 10.1126/science.1111404.

Rossi L., Hohn B., Tinland B. (1993) The VirD2 protein of Agrobacterium tumefaciens carries nuclear localization signals important for transfer of T-DNA to plant. *Molecular & general genetics : MGG***239**:345-53.

Samali A., Nordgren H., Zhivotovsky B., Peterson E., Orrenius S. (1999) A comparative study of apoptosis and necrosis in HepG2 cells: oxidant-induced caspase inactivation leads to necrosis. *Biochemical and biophysical research communications***255**:6-11. DOI: 10.1006/bbrc.1998.0139.

Santamaria I., Velasco G., Pendas A.M., Fueyo A., Lopez-Otin C. (1998) Cathepsin Z, a novel human cysteine proteinase with a short propeptide domain and a unique chromosomal location. *The Journal of biological chemistry***273**:16816-23.

Savchenko A., Yee A., Khachatryan A., Skarina T., Evdokimova E., Pavlova M., Semesi A., Northey J., Beasley S., Lan N., Das R., Gerstein M., Arrowmith C.H., Edwards A.M. (2003) Strategies for structural proteomics of prokaryotes: Quantifying the advantages of studying orthologous proteins and of using both NMR and X-ray crystallography approaches. *Proteins***50**:392-9. DOI: 10.1002/prot.10282.

Schaffer M.A., Fischer R.L. (1988) Analysis of mRNAs that Accumulate in Response to Low Temperature Identifies a Thiol Protease Gene in Tomato. *Plant physiology***87**:431-6.

Schaffer M.A., Fischer R.L. (1990) Transcriptional activation by heat and cold of a thiol protease gene in tomato. *Plant physiology***93**:1486-91.

Scherrer S., Robas N., Zouheiry H., Branlant G., Branlant C. (1994) Periplasmic aggregation limits the proteolytic maturation of the Escherichia coli penicillin G amidase precursor polypeptide. *Applied microbiology and biotechnology***42**:85-91.

Schubot F.D., Waugh D.S. (2004) A pivotal role for reductive methylation in the de novo crystallization of a ternary complex composed of Yersinia pestis virulence factors YopN, SycN and YscB. *Acta crystallographica. Section D, Biological crystallography***60**:1981-6. DOI: 10.1107/S0907444904023005.

Schwede T., Kopp J., Guex N., Peitsch M.C. (2003) SWISS-MODEL: An automated protein homology-modeling server. *Nucleic acids research***31**:3381-5.

Shabab M., Shindo T., Gu C., Kaschani F., Pansuriya T., Chintha R., Harzen A., Colby T., Kamoun S., van der Hoorn R.A. (2008) Fungal effector protein AVR2 targets diversifying defense-related cys proteases of tomato. *The Plant cell***20**:1169-83. DOI: 10.1105/tpc.107.056325.

Shindo T., Misas-Villamil J.C., Horger A.C., Song J., van der Hoorn R.A. (2012) A role in immunity for Arabidopsis cysteine protease RD21, the ortholog of the tomato immune protease C14. *PloS one***7**:e29317. DOI: 10.1371/journal.pone.0029317.

Shindo T., Van der Hoorn R.A. (2008) Papain-like cysteine proteases: key players at molecular battlefields employed by both plants and their invaders. *Molecular plant pathology***9**:119-25. DOI: 10.1111/j.1364-3703.2007.00439.x.

Singh N., Sijwali P.S., Pandey K.C., Rosenthal P.J. (2006) Plasmodium falciparum: biochemical characterization of the cysteine protease falcipain-2'. *Experimental parasitology***112**:187-92. DOI: 10.1016/j.exppara.2005.10.007.

Smits S.H.J., Mueller A., Grieshaber M.K., Schmitt L. (2008) Coenzyme- and His-tag-induced crystallization of octopine dehydrogenase. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications***64**:836-839. DOI: Doi 10.1107/S1744309108025487.

Song J., Win J., Tian M., Schornack S., Kaschani F., Ilyas M., van der Hoorn R.A., Kamoun S. (2009) Apoplastic effectors secreted by two unrelated eukaryotic plant pathogens target the tomato defense protease Rcr3. *Proceedings of the National Academy of Sciences of the United States of America***106**:1654-9. DOI: 10.1073/pnas.0809201106.

Stubbs M.T., Laber B., Bode W., Huber R., Jerala R., Lenarcic B., Turk V. (1990) The refined 2.4 A X-ray crystal structure of recombinant human stefin B in complex with the cysteine proteinase papain: a novel type of proteinase inhibitor interaction. *The EMBO journal***9**:1939-47.

Szabo A.G., Lynn K.R., Krajcarski D.T., Rayner D.M. (1978) Tyrosinate fluorescence maxima at 345 nm in proteins lacking tryptophan at pH 7. *FEBS letters***94**:249-52.

Tasset C., Bernoux M., Jauneau A., Pouzet C., Briere C., Kieffer-Jacquinod S., Rivas S., Marco Y., Deslandes L. (2010) Autoacetylation of the Ralstonia solanacearum effector PopP2 targets a lysine residue essential for RRS1-R-mediated immunity in Arabidopsis. *PLoS pathogens***6**:e1001202. DOI: 10.1371/journal.ppat.1001202.

Taylor M.A., Baker K.C., Briggs G.S., Connerton I.F., Cummings N.J., Pratt K.A., Revell D.F., Freedman R.B., Goodenough P.W. (1995) Recombinant pro-regions from papain and papaya proteinase IV-are selective high affinity inhibitors of the mature papaya enzymes. *Protein engineering***8**:59-62.

Thomma B.P., HP V.A.N.E., Crous P.W., PJ D.E.W. (2005) Cladosporium fulvum (syn. Passalora fulva), a highly specialized plant pathogen as a model for functional studies on plant pathogenic Mycosphaerellaceae. *Molecular plant pathology***6**:379-93. DOI: 10.1111/j.1364-3703.2005.00292.x.

Tian M., Win J., Song J., van der Hoorn R., van der Knaap E., Kamoun S. (2007) A Phytophthora infestans cystatin-like protein targets a novel tomato papain-like apoplastic protease. *Plant physiology***143**:364-77. DOI: 10.1104/pp.106.090050.

Tobbell D.A., Middleton B.J., Raines S., Needham M.R., Taylor I.W., Beveridge J.Y., Abbott W.M. (2002) Identification of in vitro folding conditions for procathepsin S and cathepsin S using fractional factorial screens. *Protein expression and purification***24**:242-54. DOI: 10.1006/prep.2001.1573.

Turk B., Turk D., Turk V. (2000) Lysosomal cysteine proteases: more than scavengers. *Biochimica et biophysica acta***1477**:98-111.

Turk V., Turk B., Turk D. (2001) Lysosomal cysteine proteases: facts and opportunities. *The EMBO journal***20**:4629-33. DOI: 10.1093/emboj/20.17.4629.

Van't Klooster J.W., Van der Kamp M.W., Vervoort J., Beekwilder J., Boeren S., Joosten M.H., Thomma B.P., De Wit P.J. (2011) Affinity of Avr2 for tomato cysteine protease Rcr3 correlates with the Avr2-triggered Cf-2-mediated hypersensitive response. *Molecular plant pathology***12**:21-30. DOI: 10.1111/j.1364-3703.2010.00647.x.

van den Burg H.A., Westerink N., Francoijs K.J., Roth R., Woestenenk E., Boeren S., de Wit P.J., Joosten M.H., Vervoort J. (2003) Natural disulfide bond-disrupted mutants of AVR4 of the tomato pathogen Cladosporium fulvum are sensitive to proteolysis, circumvent Cf-4-mediated resistance, but retain their chitin binding ability. *The Journal of biological chemistry***278**:27340-6. DOI: 10.1074/jbc.M212196200.

van den Hooven H.W., van den Burg H.A., Vossen P., Boeren S., de Wit P.J., Vervoort J. (2001) Disulfide bond structure of the AVR9 elicitor of the fungal tomato pathogen Cladosporium fulvum: evidence for a cystine knot. *Biochemistry***40**:3458-66.

van der Hoorn R.A., Kamoun S. (2008) From Guard to Decoy: a new model for perception of plant pathogen effectors. *The Plant cell***20**:2009-17. DOI: 10.1105/tpc.108.060194.

van der Hoorn R.A., Leeuwenburgh M.A., Bogyo M., Joosten M.H., Peck S.C. (2004) Activity profiling of papain-like cysteine proteases in plants. *Plant physiology***135**:1170-8. DOI: 10.1104/pp.104.041467.

van Esse H.P., Van't Klooster J.W., Bolton M.D., Yadeta K.A., van Baarlen P., Boeren S., Vervoort J., de Wit P.J., Thomma B.P. (2008) The Cladosporium fulvum virulence protein Avr2 inhibits host proteases required for basal defense. *The Plant cell***20**:1948-63. DOI: 10.1105/tpc.108.059394.

Velasco G., Ferrando A.A., Puente X.S., Sanchez L.M., Lopez-Otin C. (1994) Human cathepsin O. Molecular cloning from a breast carcinoma, production of the active enzyme in Escherichia coli, and expression analysis in human tissues. *The Journal of biological chemistry***269**:27136-42.

Vernet T., Khouri H.E., Laflamme P., Tessier D.C., Musil R., Gour-Salin B.J., Storer A.C., Thomas D.Y. (1991) Processing of the papain precursor. Purification of the zymogen and characterization of its mechanism of processing. *The Journal of biological chemistry***266**:21451-7.

Voinnet O., Lederer C., Baulcombe D.C. (2000) A viral movement protein prevents spread of the gene silencing signal in Nicotiana benthamiana. *Cell***103**:157-67.

Voinnet O., Rivas S., Mestre P., Baulcombe D. (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *The Plant journal : for cell and molecular biology***33**:949-56.

Wang Z., Gu C., Colby T., Shindo T., Balamurugan R., Waldmann H., Kaiser M., van der Hoorn R.A. (2008) Beta-lactone probes identify a papain-like peptide ligase in Arabidopsis thaliana. *Nature chemical biology***4**:557-63. DOI: 10.1038/nchembio.104.

Weerapana E., Wang C., Simon G.M., Richter F., Khare S., Dillon M.B., Bachovchin D.A., Mowen K., Baker D., Cravatt B.F. (2010) Quantitative reactivity profiling predicts functional cysteines in proteomes. *Nature***468**:790-5. DOI: 10.1038/nature09472.

White C.E., Kempi N.M., Komives E.A. (1994) Expression of highly disulfide-bonded proteins in Pichia pastoris. *Structure***2**:1003-5.

Wiederanders B., Kaulmann G., Schilling K. (2003) Functions of propeptide parts in cysteine proteases. *Current protein & peptide science***4**:309-26.

Wiederanders B., Kirschke H. (1989) The processing of a cathepsin L precursor in vitro. *Archives of biochemistry and biophysics***272**:516-21.

Yamada K., Matsushima R., Nishimura M., Hara-Nishimura I. (2001) A slow maturation of a cysteine protease with a granulin domain in the vacuoles of senescing Arabidopsis leaves. *Plant physiology***127**:1626-34.

Yamamoto D., Matsumoto K., Ohishi H., Ishida T., Inoue M., Kitamura K., Mizuno H. (1991) Refined x-ray structure of papain.E-64-c complex at 2.1-A resolution. *The Journal of biological chemistry***266**:14771-7.

Yan Y.L., Marriott G. (2003) Analysis of protein interactions using fluorescence technologies. *Current Opinion in Chemical Biology***7**:635-640. DOI: DOI 10.1016/j.cbpa.2003.08.017.

Yee A., Chang X., Pineda-Lucena A., Wu B., Semesi A., Le B., Ramelot T., Lee G.M., Bhattacharyya S., Gutierrez P., Denisov A., Lee C.H., Cort J.R., Kozlov G., Liao J., Finak G., Chen L., Wishart D., Lee W., McIntosh L.P., Gehring K., Kennedy M.A., Edwards A.M., Arrowsmith C.H. (2002) An NMR approach to structural proteomics. *Proceedings of the National Academy of Sciences of the United States of America***99**:1825-30. DOI: 10.1073/pnas.042684599.

Yoshii H., Furuta T., Yonehara T., Ito D., Linko Y.Y., Linko P. (2000) Refolding of denatured/reduced lysozyme at high concentration with diafiltration. *Bioscience, biotechnology, and biochemistry***64**:1159-65.

Yoshimura T., Ashiuchi M., Esaki N., Kobatake C., Choi S.Y., Soda K. (1993) Expression of glr (murI, dga) gene encoding glutamate racemase in Escherichia coli. *The Journal of biological chemistry***268**:24242-6.

Zipfel C., Robatzek S., Navarro L., Oakeley E.J., Jones J.D., Felix G., Boller T. (2004) Bacterial disease resistance in Arabidopsis through flagellin perception. *Nature***428**:764-7. DOI: 10.1038/nature02485.

ERKLÄRUNG

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt,die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen derArbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken imWortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnungkenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oderUniversität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenenTeilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solcheVeröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde.Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mirvorgelegte Dissertation ist von Prof. Dr. Paul Schulze-Lefert betreut worden.

Max-Planck-Institut für Züchtungsforschung,
Köln, 02/05/2012,

Selva Kumari Ramasubramanian

Teilpublikationen:

Kumari, S., and Van der Hoorn, R. A. L. (2011) A structural biology perspective on bioactive small molecules and their plant targets. **Curr.Opin. Plant Biol**. 14, 480-488

LEBENSLAUF

Angaben zur Person

Name:	Selva Kumari
Vorname:	Ramasubramanian
Geburtsdatum:	25. Februar 1985
Geburtsort:	Chennai, Indien
Nationalität:	Indisch

Ausbildung

Seit 10/2008	**Promotions-Studium** an der Universität zu Koeln, Max Planck Institute für Pflanzenzüchtungsforschung im Plant Chemetics Lab, Abteilung: Molekulare Phytopathologie Angestrebter Abschluss: Doktor rer. nat. in Genetik Thema: Biochemical and Structural Characterization of RCR3-AVR2: a Model for Protease-Inhibitor Interactions at the Plant-Pathogen Interface Betreuer: Prof. Dr. Paul Schulze-Lefert
07/2007-12/2007	**Post M.S Studium (weiterführender Kurs)** an der Madurai Kamaraj University (MKU) Abteilung Biotechnologie der Pflanzen Betreuer: Prof. Dr. K.Veluthambi
12/2006-06/2007	**Masterarbeit** bei Biocon Biopharmaceuticals Abteilung: Bioprozessierung von monoklonalen Antikörpern Thema: Studies in Analytical Methods of Therapeutic Proteins Betreuer: Dr. Jyothi Iyer

06/2005-06/2007 **M.S Studium**am Vellore Institute of Technology, India
Abteilung: Biotechnologie
Erworbener Abschluss: Master of Science (M.Sc) in Biotechnologie

06/2002-04/2005 **B.S Studium**am P.S.G College of Arts and Science
Abteilung: Biotechnologie
Erworbener Abschluss: Bachelor of Science (B.Sc) in Biotechnologie
Betreuer: Prof. Dr. Kannan

Publikationen

Kumari, S., and Van der Hoorn, R. A. L. (2011) A structural biology perspective on bioactive small molecules and their plant targets. **Curr.Opin. Plant Biol**. 14, 480-488

Köln, 2.Mai 2012 Selva Kumari Ramasubramanian