STUDY ON CYSTEINE PROTEASE AND THEIR INHIBITORS

A THESIS

SUBMITTED TO THE

UNIVERSITY OF PUNE

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

(IN BIOTECHNOLOGY)

BY

MOHAMMED SHABAB

UNDER THE GUIDANCE OF

DR. M. I. KHAN

DIVISION OF BIOCHEMICAL SCIENCES

NATIONAL CHEMICAL LABORATORY

PUNE - 411008 (INDIA)

MARCH 2009

In The Name Of ALLAH, Most Gracious, Most Merciful

IN FOND MEMORY OF MY MOTHER AND GRAND-MOTHER

CONTENTS

		Page No.
DECLARATION		**ii**
ACKNOWLEDGEMENTS		**iii**
ABSTRACT		**v**
LIST OF ABBREVIATIONS		**xi**
Chapter 1:	**General Introduction**	**1-37**
	Definition	1
	Historical perspective	2
	Proteases	3
	Cysteine proteases	7
	Structure of papain	8
	Families of cysteine peptidases	11
	Protease inhibitors	15
	Cysteine protease inhibitor	18
	Cystatins	22
	Chagasin family	26
	Cysteine proteases in plant – pathogen interactions	28
	Future perspectives	30
	Present investigation	31
	References	32
Chapter 2:	**Fungal effector protein AVR2 targets diversifying, defence-related cysteine proteases of tomato**	**38-80**
	Summary	38
	Introduction	38
	Materials and methods	41
	Results	48
	Discussion	67
	References	72

Chapter 3: Contrasting selectivity of fungal and oomycete effectors targeting tomato cysteine proteases **81-96**

Summary 81
Introduction 82
Materials and methods 83
Results 85
Discussion 92
References 95

Chapter 4: Targeted mutagenesis of RD21, a cysteine protease in *Arabidopsis* thaliana **97-115**

Summary 97
Introduction 97
Materials and methods 100
Results and discussion 104
References 113

Chapter 5: RIP1-Chagasin like cysteine protease inhibitor from *Pseudomonas syringae* DC 3000; its biophysical studies **116-129**

Summary 116
Introduction 117
Materials and methods 119
Results and discussion 122
References 129

Chapter 6: Identification, purification and partial characterization of cysteine protease inhibitor (CPI) from salt tolerant plant pathogenic fungus *Penicillium chrysogenum* **130-156**

Part 1: Isolation and identification of plant pathogenic *Penicillium* sp. SDBF1

	Summary	130
	Introduction	131
	Materials and methods	133
	Results	136
	References	142
	Part 2: Purification and partial characterization of CPI from *Penicillium* sp. SDBF1	140-153
	Summary	143
	Introduction	143
	Materials and methods	145
	Results and discussion	148
	Conclusion	155
	References	156
Chapter 7:	**General Discussion and Conclusion**	**157-167**
	Discussion	157
	Conclusions	163
	List of Publications	164
	References	165
Appendix 1	**Thermodynamics of galactose and its analogue inhibitor 1-de-oxy-galacto homonojirimycin binding to α-galactosidase**	**168-181**

CERTIFICATE

Certified that the work incorporated in the thesis entitled **"Study on cysteine protease and their inhibitors"** submitted by Mr. Mohammed Shabab was carried out under my supervision. Such material as has been obtained from other sources has been duly acknowledged in the thesis.

Dr. M. I. Khan
Research Guide **Date:**

DECLARATION OF THE CANDIDATE

I declare that the thesis entitled **"Study on cysteine protease and their inhibitors"** submitted by me for the degree of Doctor of Philosophy is the record of work carried out by me during the period from **6th Feb, 2004 to 26th March, 2009** under the guidance of **Dr. M. I. Khan** and has not formed the basis for the award of any degree, diploma, associateship, fellowship, titles in this or any other university or other institute of higher learning.

I further declare that the material obtained from other sources has been duly acknowledged in the thesis.

Mohammed Shabab **Date:**

Acknowledgement

I take this opportunity to gratefully thank my research guide Dr. M.I. Khan for his guidance and keen interest during the course of this investigation. He has given me the freedom to think, and work and I shall cherish my learning experience under him. Apart from research, I am also thankful to him for sharing his experiences, which have taught me subtle facts of life. I am also grateful to Mrs. Khan for love, care and treating me like part of the family.

I am equally obliged to Dr. Renier van der Hoorn for giving me chance to work under his guidance and invitating me to work in Max-Plank Institute of plant breading research, Germany.

I am grateful to Dr. (Mrs.) S. M. Gaikwad for her valuable suggestions and help in learning biophysical techniques, Dr. Mahesh Kulkarni for teaching me mass spectrometry and Dr. Farnusch Kaschani for overall supervision at MPIZ.

I am also grateful to Dr Ganesh Pandey and Dr. Shrininvas Dumbery for giving me permission to work on other synthesized compounds.

I am also grateful Dr. M. V. Krishnasastry (NCCS), Dr. C. G. Suresh, Dr. Aditi Pant, Dr. Absar Ahmad, Dr. Asish Kumar Bhattacharya, Dr Asmita Prabhune and Dr. V. Shankar for their valuable suggestions during the course of investigation.

I also thank Dr. (Mrs.) S.S. Deshmukh for caring me like her son.

I would also like to extend my sincere thanks to Dr.Vaijayanti for allowing me to work on C.D spectrophotometer and lyophilizer, Dr. Dhiman Sarkar for fluorescence plate reader.

I express my deep feelings and love for my labmates and friends Ansari, Asad, Atul, Anil, Ajit, Anish, Avinsh, Ashutosh, Jayprakash, Maggie, Manish, Nagraj, Poorva,

Rohtas, Tanpreet, Feroz, Shadab, Siddharath, Shrikanth, Sharmili, Sajid, Shashidhar, Shweta, Sofia, Satya, Sana, Senapati,, Uma, Vaishali, Vaibhav, Kalpesh, Medha, Upasana, Abhishek, Shamim, Harish, Hemeder, Daneshwer, Imran, Padma, Santola, Madhurima,Varsha, Prabha, Farnusch, Takayuki, Johana, Zheming, Christian, Izabella, Andrea, Twinkel, Rafique, Naim, Zaheer, Imtiyaz, Neeraj, Ravi, Noor bhai, Asif, Nadeem bhai, Amjad, Dr. Kerstin Richau, Dr. Tom Colby, Leonard, Raju Chintha, Ravi, Nithya, Nimisha, Neha, Stephano, Kishor, Rafique Gour, Mohinish, Fazal, Cheetan, Jayshree, Shree devi, Vishnu, Poonam, Santosh Gupta, Santosh Chavan, Sahir Masuri, Bhushan bhai, Roshan, Manisha, Shivshakar, Prabhash, Yashwant, Gyan prakash, Ram Singh, Faisal, Aparna, Afzar, Akbar bhai, Akthar bhai, Asgar bhai, Madhu, Patrick, Sharad, Ratika, Atul faranzi, Amrish, Anjali, Indira Mohandas, Rekha and everyone who have helped me, the list is too long to be mentioned.

Special thanks to Namita, my best friend for her constant support and help.

I find no words for my parents who have been a constant inspiration for me. I am also thankful to my brothers (Salman, Shahid and Ovais) and sister (Shabana), cousins and friends, who have been a moral support to me, standing by me and boosting my morale in times of stress.

Finally, I thank Head, Division of Biochemical Sciences and the Director, National Chemical Laboratory, for permitting me to submit this work in the form of the thesis. Financial assistance from Council of Scientific and Industrial Research, India and by Max-plank society, Germany is gratefully acknowledged.

Mohammed Shabab

ABSTRACT

CHAPTER I

General Introduction

This chapter deals with general introduction of protease and their inhibitors with special reference to papain like cysteine proteases, their classification, role and application. Also the cysteine protease inhibitors of natural and synthetic origins are discussed with their involvement in various processes and applications.

The best known cysteine protease is papain, often seen to be typical of cysteine endopeptidases that a newly discovered enzyme of this type is automatically described as "papain like" whether or not it is a homolog! The papain family contains peptidases with a wide variety of activities, including endopeptidases with broad specificity (such as papain), endopeptidases with very narrow specificity (such as glycyl endopeptidases), aminopeptidases, a dipeptidyl-peptidase , and peptidase with both endopeptidases and exopeptidases activities (such as cathepsins B and H) There are also family members that show no catalytic activity. Enzymes of this family have been identified in very diverse biological systems: like baculovirus, eubacteria (*Porphyromonas* and *Lactococuus*), yeast, and probably all protozoa, plants and animals (8).

Papain homologs are generally either lysosomal (vacuolar) or secreted proteins. In plants, they are found in the vacuoles, but are also extracellular in the latex of papaya or fig, and in arthropods such as lobsters and mites they are among the digestive enzymes. Exceptionally, bleomycin hydrolase is a cytosolic enzyme in fungi and mammals (9).

The general mechanism of cysteine protease action has been very well studied, with papain as the model enzyme. The basic features include the formation of a covalent intermediate, the acyl-enzyme, resulting from nucleophillic attack of the active site thiol group on the carbonyl carbon of the scissile amide or ester bond of the bound substrate. The enzymatic activity of papain is exerted by a catalytic dyad formed by Cys25 and His159 residues, which in the pH interval 3.5-8.0 form an ion-pair .These residue are located at the interface of cleft on opposite domains of the enzyme (10).

Papain family cysteine proteases are key factors in the pathogenesis of microbial infections, cancer invasion, arthritis and osteoporosis. Targeting this enzyme family is

one of the strategies in the development of new chemotherapy for a number of diseases. The disparate nature of parasite cysteine protease compared to the host orthologous proteins has opened novel opportunities for chemotherapy. There is emerging evidence that the proteolytic machinery of plants plays important roles in defense against pathogens. Plant proteases being involved in defence was known recently with the identification of RCR3, a secreted cysteine protease that is required for the function of the resistance gene Cf-2 (for *Cladosporium fulvum* resistance-2) and CDR1, a secreted aspartic protease that regulates defence responses (6). The role of cysteine protease and their inhibitors in this context is discussed. Various protease and their interacting counterparts are explained.

CHAPTER II

Fungal effector protein AVR2 targets diversifying defense-related cys proteases of tomato

The interaction between the fungal pathogen *Cladosporium fulvum* and its host tomato (*Solanum lycopersicum*) is an ideal model to study suppression of extracellular host defenses by pathogens. Secretion of protease inhibitor AVR2 by *C. fulvum* during infection suggests that tomato papain like cysteine proteases (PLCPs) are part of the tomato defense response. We show that the tomato apoplast contains a remarkable diversity of PLCP activities with seven PLCPs that fall into four different subfamilies. Of these PLCPs, transcription of only PIP1 and RCR3 is induced by treatment with benzothiadiazole, which triggers the salicylic acid–regulated defense pathway. Sequencing of PLCP alleles of tomato relatives revealed that only PIP1 and RCR3 are under strong diversifying selection, resulting in variant residues around the substrate binding groove. The doubled number of variant residues in RCR3 suggests that RCR3 is under additional adaptive selection, probably to prevent autoimmune responses. AVR2 selectively inhibits only PIP1 and RCR3, and one of the naturally occurring variant residues in RCR3 affects AVR2 inhibition. The higher accumulation of PIP1 protein levels compared with RCR3 indicates that PIP1 might be the real virulence target of

AVR2 and that RCR3 acts as a decoy for AVR2 perception in plants carrying the Cf-2 resistance gene.

CHAPTER III

Contrasting selectivity of fungal and oomycete effectors targeting tomato cysteine proteases

Suppression of host defense responses is an important strategy of adapted plant pathogens. It is known that the fungal effector protein AVR2 targets two defense-related tomato proteases (RCR3 and PIP1) that are under diversifying selection. So we tested which of the seven secreted tomato proteases is targeted by the EPIC1 and EPIC2B effector proteins of the oomycete pathogen *Phytophthora infestans*. In contrast to AVR2, both EPIC proteins inhibit C14, an abundant, ubiquitous, stress-related protease that is typified by an additional C-terminal granulin domain and resides mostly inside the host cell. The fact that effectors from different pathogens target different host proteases might reflect different infection strategies employed by these pathogens.

CHAPTER IV

Structure-function analysis of RD21, a stress induced cysteine protease in *Arabidopsis thaliana*

RD21 is an abundant papain-like cysteine protease in tissues of the model plant *Arabidopsis thaliana*. We studied the structure-function relationship of RD21. The protease is encoded as a preproprotein with a C-terminal extension that shares homology to granulins, which are animal growth hormones released upon wounding. RD21 exists in two active forms: with and without the granulin domain. The processing, activity, targeting and stability of RD21 were studied using processing site, active site and deletion mutants. Localization and processing of RD21 was also studied using YFP tagged RD21.

CHAPTER V

RIP1-Chagasin like cysteine protease inhibitor from *Pseudomonas syringae* DC 3000, its biophysical studies.

In search for new cysteine protease inhibitor the genome of *Pseudomonas syringae* DC 3000 was mined; one gene was found with Chagasin like domain. The gene was cloned and protein was expressed in *E.coli* with FLAG and His tag (Kaschani and Von der Hoorn unpublished). The protein was purified under denaturation conditions and found to be active against papain and cathepsin K. Assay, kinetics and thermodynamics of CPI by Van't Hoff plot was studied. The Thermal denaturation studies were also carried out by CD spectroscopy and fluorescence spectroscopy.

Chapter VI

Identification, Purification and partial characterization of cysteine protease inhibitor (CPI) from salt tolerant plant pathogenic fungus *Penicillium chrysogenum.*

The pathogenic fungus was isolated in the laboratory from the mangrove plant *Avicennia marina*. The culture was purified by the single colony plating technique on MGYP agar plates. On nutrient agar at 28°C, the mycelia were large, irregular, sticky and cream colored. The conidium looks blue to blue-green, and the mold sometimes exudes a yellow pigment. The strain was aerobic, spore forming fungus. The organism was salt tolerant with a broad growth range of 28°C-50°C with optimum growth at 28°C and pH 7.The isolated organism was identified to be *Penicillium sp. SDBF1* based on internal transcribed spacer/5.8S ribosomal DNA (rDNA) sequencing. Sequences was submitted to gene bank under accession number FJ403589.

The extracellular culture filtrate of *Penicillium sp. SDBF1* was subjected to ultra filtration and gel filtration to remove the high molecular weight impurities and salts. The gel filtration fractions showing anti papain activity were concentrated and loaded on C18 manually packed column for reverse phase-FPLC. The anti papain activity was associated with the peak A, having a retention time of 27.726 min and other eluted peaks showed no inhibitory activity. The active fraction was lyophilized and loaded on reverse phase-

HPLC. Homogeneity of the active fractions was indicated by the single peak as analyzed on rp-HPLC and mass spectrometry (ESI-MS) showed Mr of 437 Da.

A very high inhibitory activity of CPI was recorded against papain and no inhibition for other cysteine proteases like cathepsin K and cathepsin L. The inhibitor did not show any inhibitory activity against other classes of the proteases like trypsin, chymotrypsin, pepsin and clostripain.

Chapter VII

General Discussion and Conclusion:

This chapter deals with the general discussion of the results obtained from all the working chapters.

Interaction between the fungal pathogen *Cladosporium fulvum* and its host, *Solanum lycopersicum* is an ideal model to study suppression of extracellular host defenses by pathogens. We carried out the cloning of the PLCPs (Papain like cysteine protease) and then transiently over expressed in *Nicotiana bentiamana* plants. Similarly pathogen inhibitors were cloned and expressed in ***E.coli.*** Two of the cysteine protease RCR3 and PIP1 were found to be inhibited by AVR2. The binding is pH dependent with high affinity. Co-immunoprecipitation experiments showed that these two proteins actually physically interact with each other. Sequencing of PLCPs alleles of tomato relatives revealed that only PIP1 and RCR3 are under strong diversifying selection, resulting in variant residues around the substrate binding groove. The selection pressure on RCR3 showed that one of the naturally occurring variant residues in RCR3 affects AVR2 inhibition.

Phytophthora infestans causes blight disease in *Solanum tuberosa*. The interaction of EPIC1 and EPIC2B, protease inhibitors produced by *Phytophthora infestans*, with PLCPs during infection of potato was studied. Interestingly, inhibition of C14 by EPIC1 and EPIC2B occurs at pH 4-7 C14 interacts with both EPIC1 and EPIC2B. The interaction is prevented by preincubation of C14 with an excess E-64 .Thus, C14 physically interacts with EPIC proteins, presumably at the active site. This complex is tight, but not covalent.

RD21 is subjected to targeted mutagenesis and deletion analysis to study the role of various conserved features, such as its putative glycosylation and nitrosylation sites,

catalytic residues, granulin domain and cleavage sites. Mutant RD21 proteins were overexpressed in tobacco by agroinfiltration and investigated by western blot and activity-based profiling using biotinylated E-64 (DCG-04). This revealed, amongst others, that the granulin domain is required for prodomain removal. Putative roles of the granulin domain and the complicated activation mechanism of RD21 were discussed.

The thermodynamic parameters of RIP1 have shown that the binding of this inhibitor to papain and cathepsin K, is an enthalpy driven process, which means that the binding of these inhibitors is favorable in terms of bond energies. Thermal denaturation studies were also carried out with CD spectroscopy and it was found that there is sharp collapse in secondary structure over temperature range from 40 -50ºC.

The pathogenic fungus was isolated in the laboratory from the mangrove plant *Avicennia marina*. This fungus was found to be a salt tolerant *Penicillium sp. SDBF1,* based on ITS sequencing. Gene bank accession number FJ403589.It was screened for cysteine protease inhibitor (CPI) production. The CPI was purified with various chromatographic techniques mainly by reverse phase HPLC and FPLC. CPI was stable in a broad range of pH (2-11) and temperature (37-80°C). Homogeneity of the active fractions was indicated by the single peak as analyzed on rp-HPLC and mass spectrometry (ESI-MS) showed Mr. of 437 Da.

LIST OF ABBREVIATIONS USED

°C	Degree Celsius
Å	Angstrom
BAPNA	Nα-Benzoyl-DL-Arginine p-Nitroanilide
BLAST	Basic Local Alignment Search Tool
CD	Circular dichroism
CH_3CN	Acetonitrile
CP	Cysteine Protease
CPI	Cysteine Protease Inhibitor
Da	Dalton
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
E	Enzyme
E-64	L-trans-epoxysuccinyl-leucyl-amido (4-guanidino)-butane)
EC	Enzymes classification
EDTA	Ethylene diamine tetra acetic acid
EI*	Enzyme inhibitor complex
EMIS	Electromagnetic induction spectroscopy
g	Gram
h	Hours
I	Inhibitor
IC_{50}	Inhibition constant
IGS	Intergenic spacer
It	Total inhibitor
ITS	Internal transcribed spacer
kDa	Kilo Dalton
K_i	Inhibition constant
K_m	Michaelis Menton Constant
L	Liter

LB	Luria Bertani
M	Molar
MALDI	Matrix-assisted laser desorption/ionization
min	Minutes
ml	Milliliter
mM	Milli molar
MS	Mass spectroscopy
NCBI	National Center for Biotechnology Information
nm	Nanometers
OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PCR	Polymerase Chain Reaction
pI	Isoelectric point
RD21	Responsive to desiccation-21
rDNA	Ribosomal Deoxyribonucleic acid
REA	Relative Enzyme Activity
RNA	Ribonucleic acid
RNAase	Ribonuclease
Rp-FPLC	Reverse phase Fast Protein Liquid Chromatography
Rp-HPLC	Reverse phase High Performance Liquid Chromatography
rpm	Revolutions per minute
rRNA	Ribosomal ribonucleic acid
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
TBS	Tris buffer saline
TCA	Trichloric acid
TE	Tris-Ethylene diamine tetra acetic acid
TFA	Trifluoroacetate
V_{max}	Maximum velocity
µl	Micro liter

CHAPTER 1

GENERAL INTRODUCTION

Definition

Hydrolytic enzymes (hydrolases) catalyze the hydrolysis of a chemical bond and are classified as EC 3 in the EC number classification of enzymes. Hydrolases can be further classified into several subclasses, based on the bonds they act upon: e.g. EC 3.2 glycosylases and EC 3.4 acting on peptide bonds (proteases).

Proteases (EC 3.4) are one of the largest and most diverse families of enzymes. They catalyze the addition of water across amide (and ester) bonds to affect cleavage using a reaction involving nucleophillic attack on the carbonyl carbon of the scissile bond. Proteases play critical roles in many physiological and pathological processes such as protein catabolism, blood coagulation, cell growth, cell migration, tissue arrangement, morphogenesis in development, inflammation, tumor growth, metastasis and activation of zymogens, release of hormones and transport of secretory proteins across membranes. The importance of peptide bond cleavage in biological system is reflected by the finding that nature has separately invented the necessary catalytic machinery. Through evolution, proteases have adapted to the wide range of conditions found in complex organisms (variations in pH, reductive environment and many more) and use different catalytic mechanisms for substrate hydrolysis. The mechanism of cleavage and the active site substituent's vary widely among different protease subtypes and provide the basis for the classification of proteases into the aspartic proteases, serine proteases, cysteine proteases and metallo proteases (1). There are two recently discovered members, threonine proteases and glutamic acid protease (2). In addition, there are a few miscellaneous proteases that do not precisely fit into the standard classification, for example, ATP-dependent proteases, which require ATP for their activity (3).

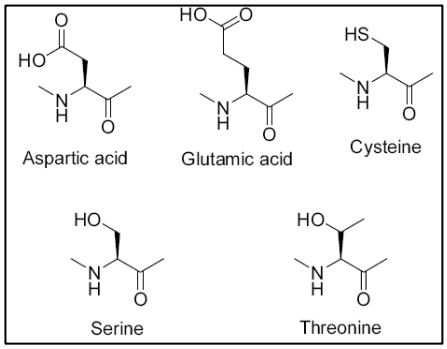


Figure 1. Protease catalytic residues

The following are major metabolic events executed by proteases.

- Gain of function (for example, activation by limited proteolysis).
- Loss of function (for example, cleavage of inhibitor of caspase-activated DNAse (ICAD) by caspase 3 during apoptosis, thereby releasing caspase-activated DNAse (CAD) and enabling DNA fragmentation (4).
- Switch of function (for example, chemokine processing by matrix metalloproteinases and other proteases (5).

Modes of substrate cleavage by peptidases (cathepsins were used as examples): endopeptidase (cathepsin L) and exopeptidases (left, cathepsin H, an aminopeptidase, and right, cathepsin X, a carboxypeptidase) (Figure2). Peptide substrate (schematically represented by cyan balls) runs through the entire length of the active site of an endopeptidase framework (blue) and is cleaved in the middle of the molecule (scissile bond marked yellow). In exopeptidases, substrate binding is structurally constrained (mini-chain in cathepsin H, orange; mini-loop in cathepsin X, green). In cathepsin exopeptidases these additional structural elements also provide negative charge (cathepsin H) to bind to the positively charged amino terminus (blue) of the substrate, or positive charge (cathepsin X) to bind the negatively charged carboxyl terminus (red) of the substrate.

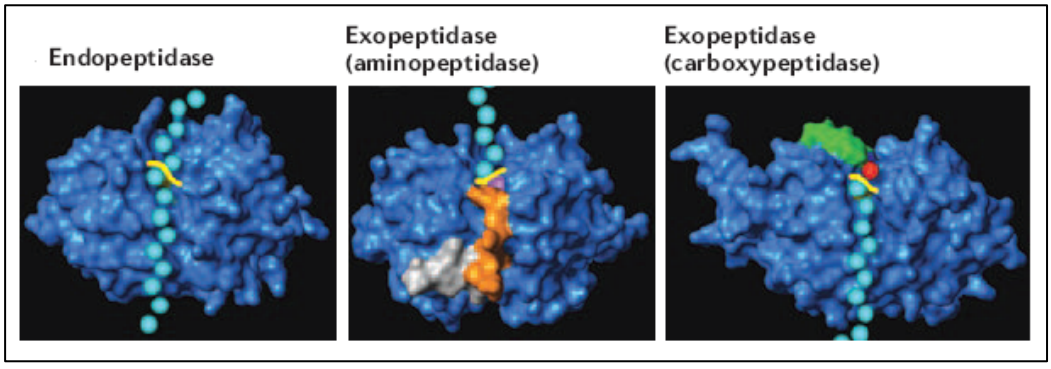


Figure 2. Simplified mode of substrate cleavage by cathepsins.

Catalytic mechanism of proteases

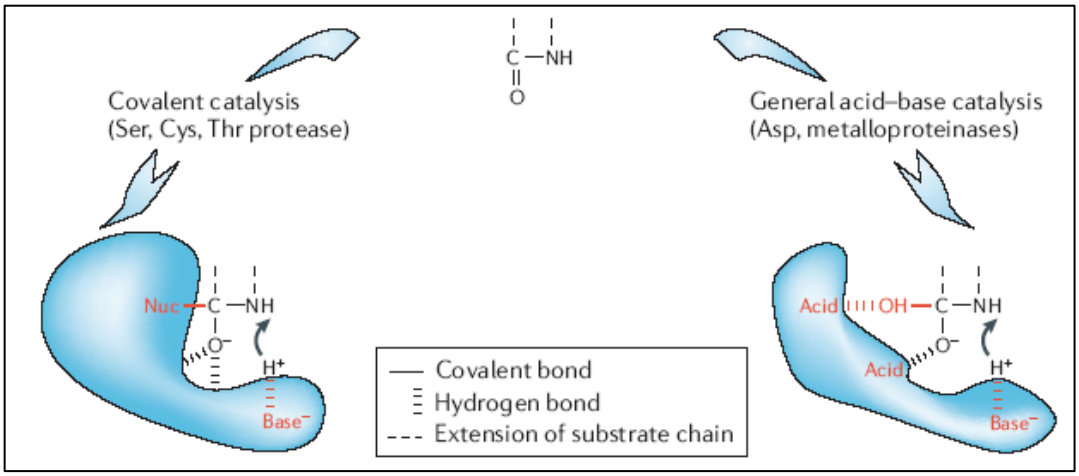


Figure 3*.Catalytic mechanisms of proteases. The five major catalytic classes of proteases use two fundamentally different catalytic mechanisms to stabilize the tetrahedral intermediate.

***Figure adapted from Nature Reviews Drug Discovery (2006) 5: 785-799.**

In serine, cysteine and threonine proteases the nucleophile of the catalytic site is part of an amino acid (covalent catalysis), whereas in the metalloproteinases and aspartic proteases the nucleophile is an activated water molecule (non-covalent catalysis). In covalent catalysis, histidines normally function as a base, whereas in non-covalent catalysis Asp or Glu residues and zinc (metalloproteinases) serve as acids and bases. A further difference between the two groups is in the formation of the reaction products from the tetrahedral intermediate, which for cysteine and serine proteases requires an additional intermediate step (acyl-enzyme intermediate).

Schematic representation of a protein substrate binding to a protease

Proteases cleave their substrates at specific protease recognition sequences. The length and composition of these recognition sequences varies from protease to protease, although there are some similarities in specificity between different proteases belonging to the same clan. From the scissile bond, going towards the amino terminus, the residues of the peptide recognition sequence (substrate) are designated P1, P2, P3, etc.

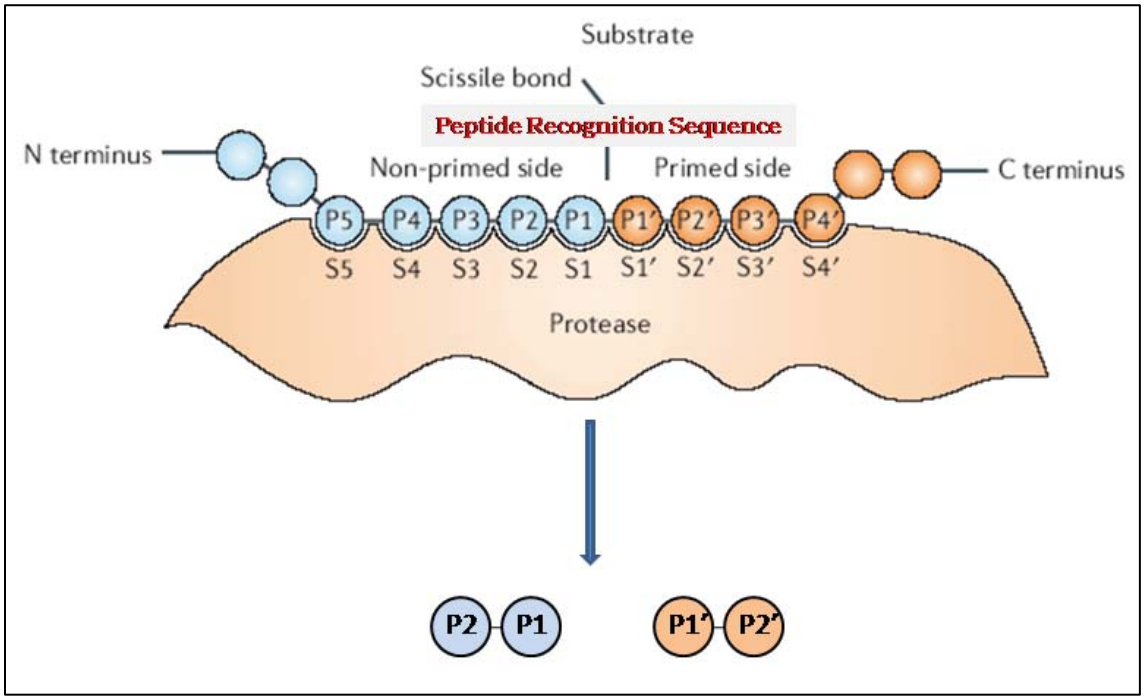


Figure 4. Schematic representation of a protein substrate binding to a protease. The surface of the protease that is able to accommodate a single side chain of a substrate residue is called the subsite (2).

Likewise, the residues going towards the carboxy terminus are designated P1′, P2′, P3′, etc. The corresponding subsites in the active site of the protease are designated S1, S2, S3, etc. and S1′, S2′, S3′, etc.

Aspartic proteases

Aspartyl proteases are present in vertebrates, plants, plant viruses, as well as in retroviruses. The subfamily of aspartyl proteases is characterized by having the highly conserved sequence of Asp- Thr- Gly (1). Because of their two-fold symmetry, it is the general consensus that these domains may have arisen through ancestral gene duplication. While a number of different mechanisms for aspartyl proteases have been proposed, the most widely accepted is a general acid-base mechanism involving coordination of a water

molecule between the two highly-conserved aspartate residues. Aspartyl proteases play an important role in several aspects of our overall health and physiology, including blood pressure (renin), digestion (pepsin and chymosin), and in the maturation of the Human Immunodeficiency Virus (HIV I protease). Aspartyl proteases are a highly specific family of proteases they tend to cleave dipeptide bonds that have hydrophobic residues as well as a beta-methylene group (3).

Serine proteases

This class comprises two distinct families, the chymotrypsin family, which includes the mammalian enzymes such as chymotrypsin, trypsin, elastase or kallikrein and the subtilisin family, which includes the bacterial enzyme such as subtilisin. Although, the three-dimensional structure is different in the two families, they possess similar active site geometry and catalytic mechanism. The serine proteases exhibit different substrate specificities, which are related to amino acid substitutions in the various enzyme subsites interacting with the substrate residues. Some enzymes have an extended interaction site with the substrate whereas others have a specificity restricted to the P1 substrate residue. The catalytic triad of the serine protease, essential in the catalytic process is His, Asp, and Ser.

Metallo proteases

All known metalloproteases use a Zn^{2+} ion to catalyze the hydrolysis of a peptide bond. The metal is tetrahedrally coordinated to three donor groups from the enzyme and a water molecule that is also hydrogen bonded to the carboxylate side chain of glutamic acid, which activates it for nucleophilic attack. Most metalloprotease inhibitors contain a zinc-binding ligand like hydroxamate, carboxylate, phosphinate, or a thiol and target only one side of the active site. The inhibitors are mostly in an extended strand conformation, except for those of thermolysin and one inhibitor of neutrophil collagenase (MMP-8). Aminopeptidase N is important for the inactivation in the kidney of blood-borne polypeptidases such as enkephalins, substance P and interleukin 8, and contributes to proteolysis in the small intestine. The insect aminopeptidase A (M01.013) is the receptor for the insecticidal CrylAc toxin of *Bacillus thuringiensis* (6).

Threonine proteases

The discovery of a fifth peptidase catalytic type was evident when the crystallographic structure of the eukaryotic proteasome from *Saccharomyces cerevisiae* was determined (7), showing an N-terminal threonine at the active site. The N-terminal threonine residues of some of the beta subunits are the nucleophiles in catalysis (8). No natural protein inhibitors are known for this class of proteases. The small, naturally occurring compounds lactacystin , epoxomycin and eponemycin from actinomycete bacteria inactivate by interaction with the catalytic threonine, blocking all three catalytic activities. The proteasome complex molecule consists of four rings of seven subunits, and takes the form of a hollow cylinder with the active sites on the walls of the inner chamber. The proteasome is involved in the turnover of intracellular proteins, including proteins specifically targeted for degradation by polyubiquitination.

Glutamic acid proteases

The peptidases in family G1 form a subset of what were formerly termed "pepstatin-insensitive carboxyl proteinases". After its discovery in about 1970, the pentapeptide pepstatin soon came to be thought of as a very general inhibitor of the endopeptidases that are active at acidic pH. The tertiary structure of scytalidoglutamic peptidase with a bound tripeptide product has been interpreted as showing that Glu136 is the primary catalytic residue (9). The most likely mechanism is suggested to be nucleophilic attack by a water molecule activated by the Glu136 side chain on the *cis*-face of the scissile peptide bond carbon atom to form the tetrahedral intermediate. Electrophilic assistance, and oxyanion stabilization, is provided by the side-chain amide of Gln53. Both scytalidoglutamic peptidase and aspergilloglutamic peptidase cleave (amongst others) the Tyr26 + Thr27 bond in the B chain of oxidized insulin; a bond not cleaved by other acid-acting endopeptidases. Scytalidoglutamic peptidase is most active on casein at pH 2. Scytalidoglutamic peptidase is inhibited by 1, 2-epoxy-3-(nitrophenoxy) propane (EPNP) (10), a compound that also inhibits pepsin.

Cysteine proteases

Cysteine peptidases catalyze the hydrolysis of peptide, amide, ester, thiol ester and thiono ester bonds. Cysteine proteases participate in varied biological processes. The cathepsins alone are involved in protein breakdown in lysosomes, antigen presentation, proteolytic processing of proenzymes and prohormones, fertilization, cell proliferation, differentiation and apoptosis. Imbalanced activity of endogenous cysteine proteases may lead to numerous pathologies such as rheumatoid arthritis, multiple sclerosis, neurological disorders, tumors and osteoporosis (1-3). The cysteine proteases produced by pathogenic microbes are considered important virulence factors e.g.clostripain in gagarin and falcipain in malaria.

The best known cysteine protease is papain, often seen to be typical of cysteine endopeptidases that a newly discovered enzyme of this type is automatically described as "papain like" whether or not it is a homolog! The papain family contains peptidases with a wide variety of activities, including endopeptidases with broad specificity (such as papain), endopeptidases with very narrow specificity (such as glycyl endopeptidases), aminopeptidases, a dipeptidyl-peptidase , and peptidase with both endopeptidases and exopeptidases activities (such as cathepsins B and H) There are also family members that show no catalytic activity. Enzymes of this family have been identified in very diverse biological systems: like baculovirus, eubacteria (*Porphyromonas* and *Lactococuus*), yeast, and probably all protozoa, plants and animals (11).

Papain homologs are generally either lysosomal (vacuolar) or secreted proteins. In plants, they are found in the vacuoles, but are also extracellular in the latex of papaya or fig and in arthropods such as lobsters and mites they are amongst the digestive enzymes. Exceptionally, bleomycin hydrolase is a cytosolic enzyme in fungi and mammals (12).

Structure of papain

Papain has 212 amino acid residues and is folded into two domains: an L-domain consisting of residues 10-11 and 207-212 and R-domain consisting of the remaining residues (Fig 5). The Cys- 25 active centre of papain is positioned at the entrance of α

helix (residues 24-42) in the L-domain. The domain is mostly having helix structure. The R domain contains two helix and an anti parallel β-sheet, which is hydrophobic (13).

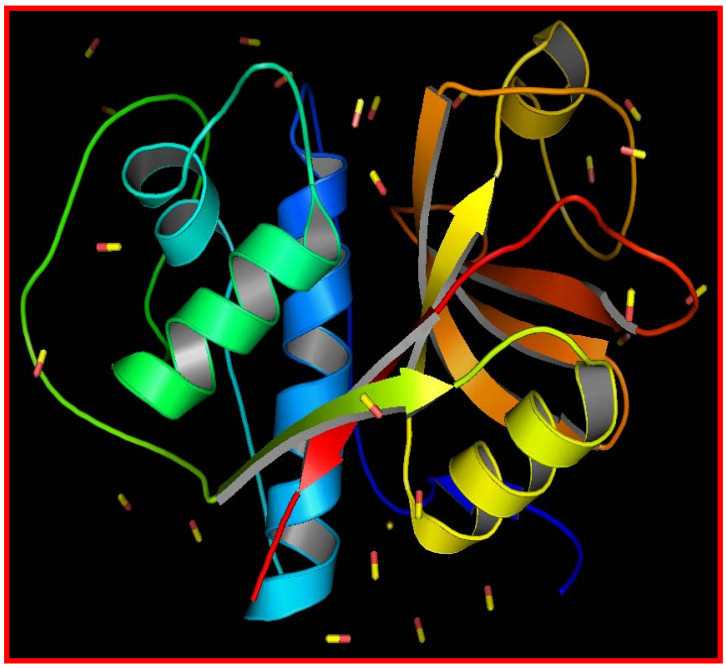


Figure 5. Ribbon representation of papain was created using PyMOL, using papain (PDB code 9papA).Summary of the papain structure is given at right side of the figure.

Catalytic mechanism in papain family of cysteine peptidases:

The general mechanism of cysteine protease action has been very well studied, with papain as the model enzyme (10). The enzymatic activity of papain is exerted by a catalytic dyad formed by Cys25 and His159 residues, which in the pH interval 3.5-8.0 forms an ion-pair (Fig. 6).These residues are located at the interface of cleft on opposite domains of the enzyme. The existence of two ionizable groups as essential catalytic residues in papain is consistent with the bell-shape from for the pH dependency of the activity of the enzyme. The acid limb, with a pK_a of ≈ 4, is usually attributed to the ionization of the active site Cys-25, whereas the basic limb, with a pK_a ≈ 8.5, is considered to reflect ionization of His-159.

Aside from the helix dipole and the geometry of the active site, specific neighboring residues play an important role to stabilize the thiolate – imidazolium ion pair. Asn^{175} is important for orientation of the imidazolium ring of the histidine in the catalytic cleft. The Asn^{175} -His^{159} hydrogen bond is approximately collinear with the C_β-C_γ bond, allowing the imidazole side chain to rotate about the C_β-C_γ bond without disruption of the Asn^{175} -His^{159} hydrogen bond. Gln-19 is proposed to play an important role in proteolysis, through the participation in an oxyanion hole similar to serine peptidase.

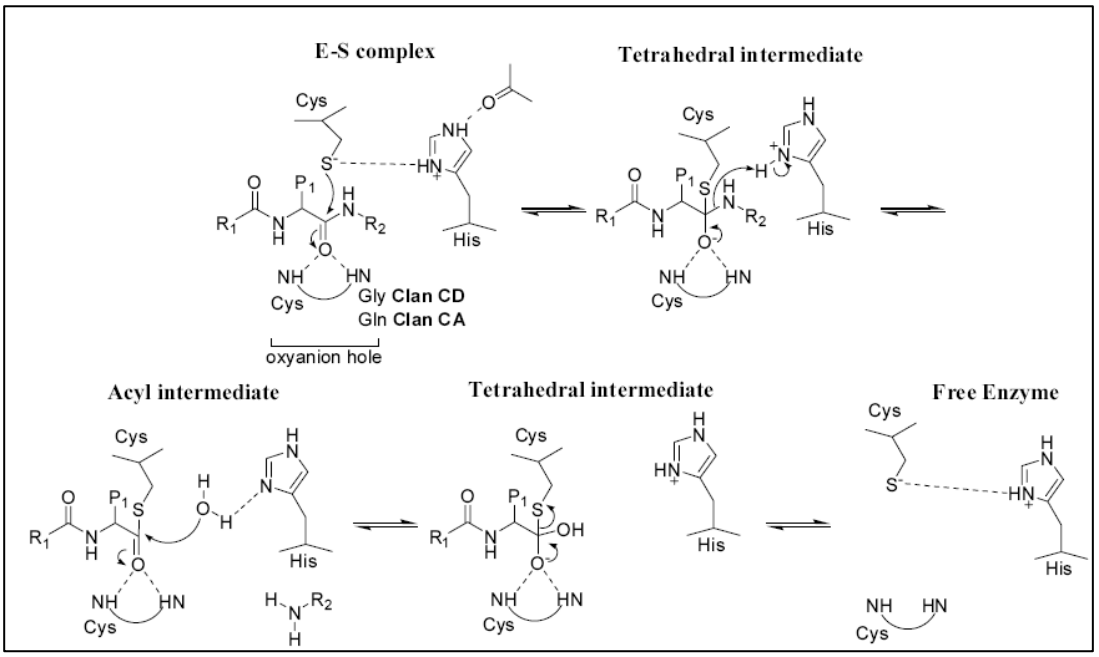


Figure 6. Schematic representation of the various steps and putative intermediates involved in the reaction pathway for hydrolysis of an amide substrate by a cysteine peptidase. Details are given in the text (15).

Following the binding of the substrate to a cysteine protease, the carbonyl carbon of the scissile bond undergoes nucleophilic attack by the Cys-25 thiolate anion of the enzyme. This is facilitated by the carbonyl oxygen binding in the oxyanion hole that is defined by the Cys-25 NH and the side chain amide of Gln-19. As a result, the first anionic tetrahedral is formed. Following the formation of TH1 in cysteine peptidase, the protonated active site histidine, His-159, rotates in order to donate its proton to the amide

nitrogen, promoting the breakdown of the tetrahedral intermediate giving free amine product and an acyl-enzyme intermediate. This intermediate collapses, via acid-assisted catalysis, to the thioester intermediate with release of the C-terminal substrate fragment. At the end, water hydrolysis gives the regenerated active site and the N-terminal substrate fragment (14).

Papain Family (C1)

The papain family contains peptidases with a wide variety of activities, including endopeptidases with broad specificity (such as papain), endopeptidases with very narrow specificity (such as glycyl endopeptidases), aminopeptidases , a dipeptidyl-peptidase and peptidase with both endopeptidases and exopeptidases activities (such as cathepsins B and H) There are also family members that show no catalytic activity. Enzymes of this family have been identified in very diverse biological systems: like baculovirus, eubacteria (*Porphyromonas* and *Lactococuus*), yeast and probably all protozoa, plants and animals. The catalytic domains of papain-like cysteine proteases are between 220 and 260 amino acids in length. Exceptions are several parasite-derived cysteine proteases which contain C-terminal extension of unknown function.

In addition to the active site residues, several other amino acid sequence regions show a high degree of conservation. Among those are several cysteine residues which are involved in the formation of disulfide bridges, Pro2 which might prevent the N-terminal truncation and thus inactivation of the mature protease by aminopeptidases, the Gly-Cys-X-Gly-Gly motif forming a wall along the nonprime subsite binding area and the Gly-Pro motif separating a β and α domain at the interface between the major L and R domains. The classification of papain family (C1) is given in subsequent table.

TABLE I*

CLANS AND FAMILIES OF CYSTEINE PEPTIDASES

Clan	Family	Representative enzyme	Identified catalytic Residues
CA	C1	Papain	Cys/His
CA	C2	Calpain	Cys
CA	C10	Streptopain	Cys/His
CB	C3	Polio virus picornain 3C	His/Cys
CB	C4	Tobacco etch virus NIa endopeptidase	His/Cys
–	C18	Hepatitis C virus endopeptidase	His/Cys
–	C5	Adenovirus endopeptidases	His/Cys
CC	C6	Tobacco etch virus HC-proteinase	Cys/His
CC	C7	Chestnut blight virus p29 endopeptidase	Cys/His
–	C8	Chestnut blight virus p48 endopeptidase	Cys/His
–	C9	Sindbis virus nsP2 endopeptidase	Cys/His
–	C16	Mouse hepatitis virus endopeptidase	Cys/His
–	C21	Turnip yellow mosaic virus endopeptidases	Cys/His
–	C11	Clostripain	Cys
–	C12	Deubiquitinating peptidase Yuh1	Cys/His
–	C19	Deubiquitinating peptidase Ubp1	Cys/His
–	C13	Hemoglobinase	–
–	C14	Interleukin 1β converting enzyme	Cys
–	C15	Pyroglutamyl-peptidase I	Cys
–	C17	Microsomal ER60 endopeptidase	–
–	C20	Type IV prepilin leader peptidase	–

*Methods in Enzymology: Serine and Cysteine Peptidase (1994) 244: 1-765

TABLE II*
PEPTIDASES OF PAPAIN (FAMILY C1)

Peptidase	EC	Database code
Family C1: Papain		
Actinidain	3.4.22.14	ACTN_ACTCH
Aleurain (barley)	–	ALEU_HORVU
Allergen (*Dermatophagoides*)	–	MMAL_DERPT
Allergen (*Eutoglyphus*)	–	EUM1_EURMA
Bleomycin hydrolase	–	BLMH_YEAST
Calotropin (*Calotropis*)	–	CAL1_CALGI
Caricain	3.4.22.30	PAP3_CARPA
Cathepsin B	3.4.22.1	CATB_*,CYSP_SCHMA
Cathepsin H	3.4.22.16	CATH_*
Cathepsin L	3.4.22.15	CATL_*
Cathepsin S	3.4.22.26	CATS_*
Chymopapain	3.4.22.6	PAP2_CARPA
Cysteine aminopeptidase (*Lactococcus*)	–	(M86245)
Cysteine endopeptidase2 and 3 (barley)	–	CYS1_HORVU, CYS2 CYS2_HORVU
Cysteine endopeptidase (*Brassica napus*)	–	CYS4_BRANA
Cysteine endopeptidase (*Caenorhabditis*)	–	CYS1_CAEEL
Cysteine endopeptidase 1 and 2 (*Dictyostelium*)	–	CYS_DICDI,CYS2_DICDI
Cysteine endopeptidase (*Entamoeba*)	–	(M27307), (M64712), (M64721), M(94163)
Cysteine endopeptidase 1 and 2 CYS2_DICDI (*Haemonchus*)	–	(CYS1_HAECO, (M80385),(M80386)
Cysteine endopeptidase (*Hemerocallis*)	–	(X74406)
Cysteine endopeptidase 1,2 and 3 (*Homarus*)	–	CYS1_HOMAM,CYS2_HOMAM CYS3_HOMAM
Cysteine endopeptidase (*Leishmania*)	–	LCPA_LEIME, (M97695), (Z14061)
Cysteine endopeptidase (mung bean)	–	CYSP_VIGMU

Cysteine endopeptidase (*Ostertagia*)	–	(M88505)
Cysteine endopeptidase (pea)	–	CYSP_PEA, (X66061)
Cysteine endopeptidase (*Plasmodium*)	–	CSP_PLACM,CYSP_PLAFA (L08500), (L26362)
Cysteine endopeptidase (*Porphyromonas*)	–	TPR_PORGI
Cysteine endopeptidase (*Tertrahymena*)	–	(L03212)
Cysteine endopeptidase (*Theileria*)	–	CYSP_THEPA, CYSP_THEAN
Cysteine endopeptidase (tobacco)	–	(Z13959), (Z13964)
Cysteine endopeptidase (tomato)	–	CYSL_LYCES, (Z14028)
Cysteine endopeptidase (*Trypanosoma*)	–	CYSP_TRYBR, (L25130), (M90067)
Dipeptidyl peptidase I	3.4.14.1	CATC_RAT
Endopeptidase (baculovirus of *Autographa*)	–	CATV_NPVAC
Endopeptidase EP-C1 (*Phaseolus vulgaris*)	–	CYSP_PHAVU
Glycyl endopeptidase	3.4.22.25	PAP4_CARPA
Oryzain (inculdes froms α, β and γ) ORYB_ORYSA	–	ORYA_ORYSA, ORYC_URYSA
Papain	3.4.22.2	PAPA_CARPA
Stem bromelain	3.4.22.32	BROM_ANACO
Thaumatopain (*Thaumatococcus*)	–	THPA_THADA

*Methods in Enzymology: Serine and Cysteine Peptidase (1994) 244: 1-765

Protease inhibitors

The action of proteases is tightly controlled to prevent improper cleavage of signalling molecules. Protease activities are regulated at the transcriptional level by differential expression and at the protein level by activation of inactive zymogens and by the binding of inhibitors and cofactors. During cell and tissue development and organism homeostasis, the protease signalling pathways work normally and are tightly controlled. But what happens when the regulation of protease signalling fails? At the substrate-cleavage level, there is either too little or too much proteolysis. Diminished proteolysis as a result of insufficient protease activity mostly originates from genetic irregularities (endogenous proteases) excessive inhibitory activity or insufficient activation is carried out often by pathogens. By contrast, excessive or inappropriate proteolysis is seldom a result of genetic aberrations. But most often results from numerous endogenous and/or exogenous factors, which result in unwanted activation of protease signalling pathways, such as the effect of atherosclerotic plaque formation or blood vessel injury on the blood coagulation cascade, which leads to the appearance of intravascular thrombin. So far, inappropriate proteolysis has been found to have a major role in cancer as well as cardiovascular, inflammatory, neurodegenerative, bacterial, viral and parasitic diseases (15).

The enzyme inhibitors serve as probes for kinetic and chemical processes during catalysis that led to a detailed understanding of enzyme catalytic mechanisms and has provided effective therapeutic agents for the treatment of diseases. They help in elucidating the mode of ligand binding, where the ligand may be an inhibitor, substrate, or substrate analogue. Inhibitors are also used for *in vivo* studies to localize and quantify enzymes in organs or to mimic certain genetic diseases that involve the absence of an enzyme in given biosynthetic pathway.

Classification of enzyme inhibitors

The initial classification of inhibitors based on the binding of inhibitors with enzymes is given by

(1) **Active site directed** -When inhibitor binds with active site of enzyme

(2) **Allosteric effector**-When inhibitor binds other than active site of enzyme

The inhibitors which attack the active site can be divided according to the type of interaction with **covalent/ noncovalent** and **irreversible/reversible** inhibitors.

The **reversible inhibitors** bind to enzymes with non-covalent interactions such as hydrogen bonds, hydrophobic interactions and ionic bonds. Multiple weak bonds between the inhibitor and the active site combine to produce strong and specific binding. In contrast to substrates and irreversible inhibitors, reversible inhibitors generally do not undergo chemical reactions when bound to the enzyme and can be easily removed by dilution or dialysis.

There are three kinds of reversible enzyme inhibitors. They are classified according to the effect of varying the concentration of the enzyme's substrate on the inhibitor.

In **competitive inhibition**, the substrate and inhibitor compete for access to the enzyme's active site. The substrate and inhibitor cannot bind to the enzyme at the same time. This type of inhibition can be overcome by sufficiently high concentrations of substrate, i.e., by out-competing the inhibitor. Competitive inhibitors are often similar in structure to the real substrate.

In **mixed inhibition**, the inhibitor binds to a different site on an enzyme. Inhibitor binding to this allosteric site changes the conformation (i.e., tertiary structure or three-dimensional shape) of the enzyme so that the affinity of the substrate for the active site is reduced. However, the binding of the inhibitor affects the binding of the substrate, and vice versa. This type of inhibition can be reduced, but not overcome by increasing concentrations of substrate. Although it is possible for mixed-type inhibitors to bind in the active site, this type of inhibition generally results from an allosteric effect where the inhibitor binds to a different site on an enzyme.

Non-competitive inhibition is a form of mixed inhibition where the binding of the inhibitor to the enzyme reduces its activity but does not affect the binding of substrate. As a result, the extent of inhibition depends only on the concentration of the inhibitor.

Type	[a]Binding determinants	[b]Chemically reactive group	Minimal mechanism	Rational	Examples
Substrate analog	**S**	**NO**	**E+S'↔E.S'**	**Structural mimicry**	**Malonate/succinate dehydrogense**
Group I. Tight Binding Enzyme Complements (Non covalent interactions)					
Transition state analog	**t.s**	**NO**	**E+l_{ts}↔E.l_{ts}**	**E binds T.S. tightly**	**Phosphonyl peptides/ metallo proteases**
Multisubstrate analog	**S_1+S_2**	**NO**	**E+S_1-S_2↔ E.S_1-S_2**	**Multiple binding effect**	**Phosphono acetyl-L-aspartate/ carbamoyl transferase**
Ground state analog	**S+R**	**NO**	**E+S'-R↔E.S'-R**	**Multiple binding effect**	**---**
Group II. Mechanism-based inhibitors (Covalent interactions)					
Transition state analogs	**t.s**	**NO**	**E+l_{ts}↔E.l_{ts}**	**E binds T.S. irreversively**	**Peptidyl aldehydes/cysteine & serine proteases**
Suicide inhibitor	**S**	**latent**	**E+S'↔E.S'↔ E.I_r→E.-l_r**	**Unmasking of latent reactive group**	**Sulbalactam/β - lactamases**
Dead-end inhibitor	**S**	**NO**	**E+S'↔E.S'→ E'-S'**	**S' lacks structural element for turn over**	**Peptidyl nitriles/cysteine proteases**
Alternate substrate inhibitor	**S**	**NO**	**E+S"↔E.S"↔ E'-S"↔E'-P↔E.P↔E+P**	**Stabilize E'-S", E'-P or E-P**	**Physostigmine/ acetylcholinesterase**
Affinity Labels (Covalent interactions)					
Classical affinity label	**S**	**Yes**	**E+S_{rx}↔E. S_{rx}→E-S'+X^-**	**High effective concentration of reactive group at active site**	**Chloromethylketones/ cysteine & serine proteases**
Quiescent affinity label	**S**	**NO**	**E+S_x↔E.S_x→ E-S'+X^-**	**Different reactivity: Chemical Vs. enzymatic**	**Acyloxymethylketones/ cysteine proteases**

Table 1: Classification of enzyme inhibitors (15)

[a] refer to initial reagent; **S**=Substrate; **t.s.** =transition state; **S_1+S_2** = Cosubstrate; R= entity with affinity for non catalytic site. [b] **S'**=Pseudosubstrate; **l_{ts}** =Transition state analog; **l_r**=reactive intermediate; **E-l_r.E-S'**, **E'-S'**, **E'-S'**, and **E'-P** = enzyme –inhibitor adduct; **S'**=alternate substrate inhibitor; **P**= product; **S_{rx}**=reactive affinity label; **S_x**= quiescent affinity label; **S_1-S_2**= multi-substrate analog; **S'-R**= ground state analog

Irreversible inhibitors (inactivators) always bind to the enzyme covalently. Irreversible inhibitors react with the enzyme via a noncovalent transition state and thus lead to rapid reduction of enzyme activity. In practice it is often difficult to differentiate between reversible and irreversible inhibitors, for example if a reversible inhibitor binds to the enzyme with such high affinity that the enzyme-inhibitor complex only dissociates very slowly and thus appears irreversible. This type of inhibitor is known as "tight binding". Normally a rapid equilibrium is observed with reversible inhibitors, whereas reactions which result in modification of the enzyme take place relatively slowly.

Cysteine protease inhibitor

Due to its significant contribution in variety of processes, there is always a demand of low molecular weight as well as high molecular weight cysteine protease inhibitors.

Low-molecular weight inhibitors

The design and synthesis of cysteine protease inhibitors has a long history and has been extensively reviewed in recent years (15-20). Compounds synthesized included a wide range of peptide aldehydes, methyl ketones, nitriles as reversibly acting inhibitors and diazomethanes, halomethyl ketones, acyloxymethyl ketones, *O*-acylhydroxamates, and epoxysuccinyl derivatives as irreversible inhibitors. Whereas early developments of cysteine protease inhibitors provided useful tools to study the protease activity. Only recently significant progress has been accomplished to develop cysteine protease inhibitors into drugs. This development was mostly driven by the dramatic increase in understanding of papain like cysteine proteases as pharmaceutically valid targets. Cysteine protease inhibitors have been used to develop new drugs.

Aldehydes

Aldehydes and its analogues continue to be attractive moieties despite their well-established chemical reactivity. "Slow binding" inhibition is observed by aldehydes, i.e. the slow attainment of steady state during inhibition with a lag phase with a half life of several minutes, is not due to induction of conformational changes in the enzyme ("slow

binding" in the usual sense) but is rather a result of the low concentration of free aldehyde in solution(21). Development of peptidyl aldehydes as inhibitors of cysteine and serine proteases is based on the assumption that a tetrahedral intermediate is involved in enzymatic hydrolysis has led to investigation of the effect of carbonyl compounds on these proteases, with the intention of developing analogs of this transition state (e.g. Ac-Phe-Glyal)(22-23). These are leupeptines, chymostatins, antipain, elastinal, and α- MAPI (24).

Acyclic and cyclic Ketones

Investigations of inhibition of the serine protease elastase by peptidyl aldehydes *in vivo* have shown that rapid loss of activity occurs if the aldehyde is oxidized to a carboxylic acid. For this reason, the aldehyde group was replaced by the metabolically more stable trifluoromethyl ketone function (TFMK). This produced significantly more potent inhibitors of serine proteases, also *in vitro*. It has been shown that a tetrahedral hemiketal is formed as a covalent enzyme-inhibitor adduct. Many derivatives have been developed which inhibit different serine proteases (25-28)

Epoxysuccinyl analogues

In 1978, Hanada *et al.* succeeded in isolating a highly active, irreversible inhibitor of papain from culture extract of *Aspergillus japonicus* (29).The substance was identified as 1-[[*N*-(L-3-*trans*-carboxyoxiran- 2-carbonyl)-L-leucyl]amino]-4-guanidinobutane, E-64 (30). Systematic studies were carried out to investigate the role of the different structural components of the inhibitor in enzyme inhibition and the *trans*-L-(*S*,*S*)-epoxysuccinic acid was discovered to be the reactive group essential for inhibition (31). A change of configuration of the epoxide residue or the neighboring amino acids reduces the activity by a factor of 10-100 (32). Structural information about the reverse mode of binding of the propeptide to active cathepsin B was utilized in designing novel, extended active site spanning E-64 peptidyl analogues. Selective cathepsin B or cathepsin L inhibitors were synthesized with the most active compounds spanning over six subsites (P4 to P2' residues) (33). Peptidyl epoxides with a tyrosine and biotin moieties allowing iodination

and streptavidin-based detection proved very valuable as functional proteomics tools (34).

Vinyl sulfones

Peptidyl vinyl sulfone inhibitors are remarkably potent irreversible inhibitors of cathepsins (35). They have been shown to be effective in mice arthritis models by significantly reducing inflammation as well as bone and cartilage erosion (36). However, due to the irreversible mode of action, vinyl sulfone inhibitors are unlikely to be developed as therapeutic drugs for chronic diseases such as osteoporosis. Structural analysis of the cruzipain inhibitor complex revealed a covalent Michael adduct with the active site cysteine residue and strong hydrogen bonding interactions in the S1' subsite (37). The same compound was also orally effective in a mouse model of malaria (38, 39).

Nitriles

Recently, nonpeptidyl derivatives of nitriles employing pyrrolidine or azetidine rings have been demonstrated to be potent cysteine protease inhibitors. Interestingly, a four-membered ring derivative was approximately 10-fold more potent than the five-membered ring analogue which was possibly due to the increased chemical reactivity of the azetidine ring. Nonpeptidyl nitrile acts as reversible, but time dependent inhibitors by forming a cleavable isothiourea ester link with the enzyme (40).

***β*-Lactams**

The development of *β*-lactams as cysteine protease inhibitors is, however, very recent. Single rings as well as bicyclic ring β-lactam moieties have been evaluated as inhibitors of cysteine proteases. The *β*-lactam moiety serves as thiol reactive species and is linked with nonpeptidyl or amino acid or peptide portions targeting binding site pockets of relevant cathepsins. 2-Substituted oxapenams employing a nonpeptidyl aromatic or alkyl moiety as subsite motif displayed surprisingly potent inhibitory efficacy in the mid-nanomolar range. As expected, lacking a specific targeting moiety, these compounds were not selective. The mode of inhibition was time dependent and no recovery of enzymatic activity was observed. On the other hand, the incorporation of amino acid or peptidyl moieties as targeting sequence for individual proteases at the other side of the

ring resulted in a significant enhancement of the potency and specificity of the compounds (41-43).

Diacyl Bis Hydrazides

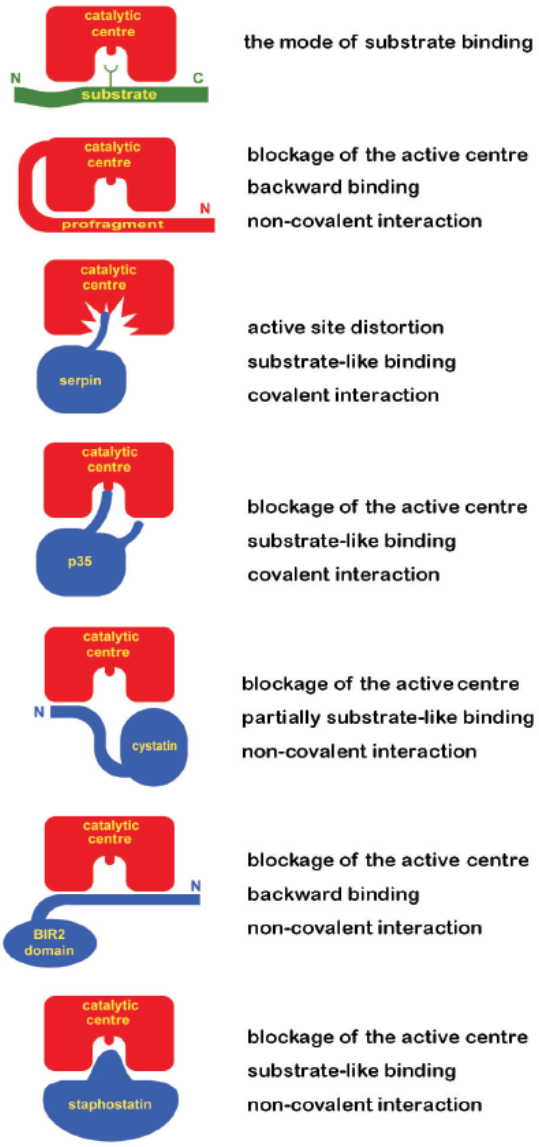
Diacyl bis hydrazides evolved from previously reported diaminopropanones which have been developed as potent and selective cathepsin K inhibitors spanning both S and S' subsites of the substrate binding cleft (44). The incorporation of bis aza analogues increased the potency while maintaining the selective profile for cathepsin K. The most potent compounds possess a leucine residue either in the P2 or P2' position or in both. Cyclic reversible cysteine protease inhibitor contains a peptidomimetic thiazole ring in place of an amide bond and may form an acyl adduct with the enzyme.

Proteinaceous cysteine protease inhibitors

The recent decade has witnessed tremendous development in the field of proteinaceous cysteine protease inhibitors. Though the prototype cystatins discovered in the 1960s remain the best-characterized group, several new large and a few smaller families are now recognized. Currently 10 are described, and the accompanying growth in a number of known inhibitors and processes involving these molecules is even higher. Most, if not all, aspects of such an important activity as proteolysis are accompanied by inhibitors exerting regulatory or protective functions. Advances in genomic studies have already enabled rough estimates of the number of inhibitors and proteases produced by many organisms, however much remains to be done to understand the detailed function of these proteins. Moreover, some inhibitors constituting yet unknown new families might have been easily overlooked in such comparative studies. It is probable that such families still await discovery, especially since known proteinaceous inhibitors of cysteine proteases outside the papain family are scarce.

Two major inhibitor classes are known: (i) cystatins and (ii) serpins, which are active toward cysteine proteases. Cystatins have been implicated in a wide range of regulatory and disease-related processes such as in immune responses, cancer and immune evasion by parasites (45). In plants, cystatin inhibits papain activity and is one of six identified

Arabidopsis cystatin genes. Over expression of this cystatin in *Arabidopsis* cell cultures blocked cell death in response to avirulent bacteria and NO (46).



Cystatins

The cystatins are tight and reversibly binding inhibitors of the papain-like cysteine protease. They are present from protozoa to mammals. They are divided into three groups (types) based on distinct structural details, but also reflecting their distribution in the body and physiological roles in humans (47). **Type 1** cystatins are polypeptides with ~100 amino acid residues, which possess neither disulphide bonds nor carbohydrate side chains (cystatins A and B). They are found mainly intracellularly, but can also appear in body fluids at significant concentrations. They are also known as stefins. **Type 2** encompasses cystatins C, D, E/M, F, G (CRES), S, SN and SA characterized by two conserved disulphide bridges, larger size (~120 residues), and a presence of a signal peptide for extracellular targeting.

Figure 7. Schematic representation of inhibitory mechanisms directed against cysteine proteinases (93). Inhibitors are shown in blue, enzymes in red and substrate in green.

Cystatins belonging to this group are found in most body fluids. Finally, **type 3** contains Kininogens, high molecular mass proteins with three tandemly repeated type 2`-like

cystatin domains, only two of which are able to inhibit cysteine proteases. Kininogens are intravascular proteins of blood plasma.

Phytocystatins

Studies on phyto-cystatins started in 1987 with the cloning of the first phyto-cystatin from rice seed (48). Subsequent studies in other species were based on this oryzacystatin probe and focused mainly on seeds. As a consequence, most of the newly discovered phyto-cystatins were highly homologous to oryzacystatin and very few results concerned the occurrence of phyto-cystatins in other parts of the plant. Generally, their amino acid sequences show higher similarity with the family II cystatins. A phyto-cystatin-specific consensus sequence has also been defined (49) 17 amino acids after the signature of the cystatin super-family: [LVI]-[AGT]-[RKE]-[FY]-[AS]-[VI]-X-[EDQV]-[HYFQ]-N.

Phyto-cystatins present an interest not only for fundamental studies aiming at the understanding of their role in plant development and defence against pathogens, but also for applied research in a variety of fields. The use of cystatins as bio-pesticides has been successfully tested against a large collection of plant pathogens (50).Oryzacystatin I has been tested for its antiviral action against the HSV-1 virus and its antiherpetic effect was similar to that of acyclovir (51). Therefore, phyto-cystatins can be considered as promising candidates in crop protection as well as in human and veterinary medicine.

Other in-depth investigations concerned the role of phyto-cystatins in plant defence mechanisms. They revealed an induction of phyto-cystatins by cell damage, fungal infection or jasmonic acid treatment (52). A role of the phytocystatins in the defence against pathogens has been described for virus (53), fungus (54), nematodes, mites (55) and insects (56, 57). *Pernas et al.* also found that the transcript level of a chestnut cystatin was induced by abiotic stress, namely by cold stress, saline shock or heat stress. Consequently, a key function of phyto-cystatins lies in the defence against biotic and abiotic stress. Another documented role of phyto-cystatin concerns programmed cell death. *Solomon et al*. (58) demonstrated the regulation of programmed cell death (PCD)

by phytocystatins. The results suggest their potential implication in xylogenesis, senescence and hypersensitive response.

Extracellular protease inhibitor with cystatin like domain (EPIC)

Motif searches of *P. infestans* unigenes with predicted signal peptides revealed novel families of putative protease inhibitors with cystatin-like domains (EPIC1 to EPIC4, Inter-Pro IPR000010 and/or IPR003243, MEROPS family I25).These proteins are involved in pathogenicity of *P. infestans.* These cystatins are shown to interacts with host proteases (59).

Inhibition of legumain and calpain

In general, cystatins are characterized as inhibitors of C1 family proteases; however, separate cases of cystatin fold adaptation to the inhibition of cysteine proteases belonging to other families have been reported. Legumain, a C13 family cysteine protease, is potently inhibited by cystatins C, E/M and, with a higher Ki by cystatin F. Interestingly, it has been demonstrated that the binding site for legumain is distinct from that of papain. A ternary complex may be formed allowing the cystatin to inhibit both proteases simultaneously (60). Moreover, the main physiological function of cystatin E/M seems to be legumain and not C1 family protease inhibition (see above). The calcium-dependent cysteine proteases – calpains (family C2) are inhibited by the second cystatin domain of Kininogens. Due to exclusively intracellular localization of calpains, the ability of intravascular Kininogens to inhibit those enzymes is thought to serve as a protection against their activity at accidental release in pathological states.

Clitocypin

Clitocypin has been isolated in large quantities from fruit bodies of an edible mushroom *Clitocybe nebularis* by affinity chromatography on carboxymethylpapain Sepharose. It is the first proteinaceous inhibitor characterized from higher fungi.

Staphostatins

Specific inhibitors of staphylococcal cysteine proteases (staphostatins) constitute yet another, recently described novel inhibitor family. Unlike chagasins and clitocypin, however, detailed structural characterization has been accomplished, and clues for the function are available. Three members of the family have been described staphostatins A and B from *Staphylococcus aureus* and staphostatin A from *Staphylococcus epidermidis* (61).

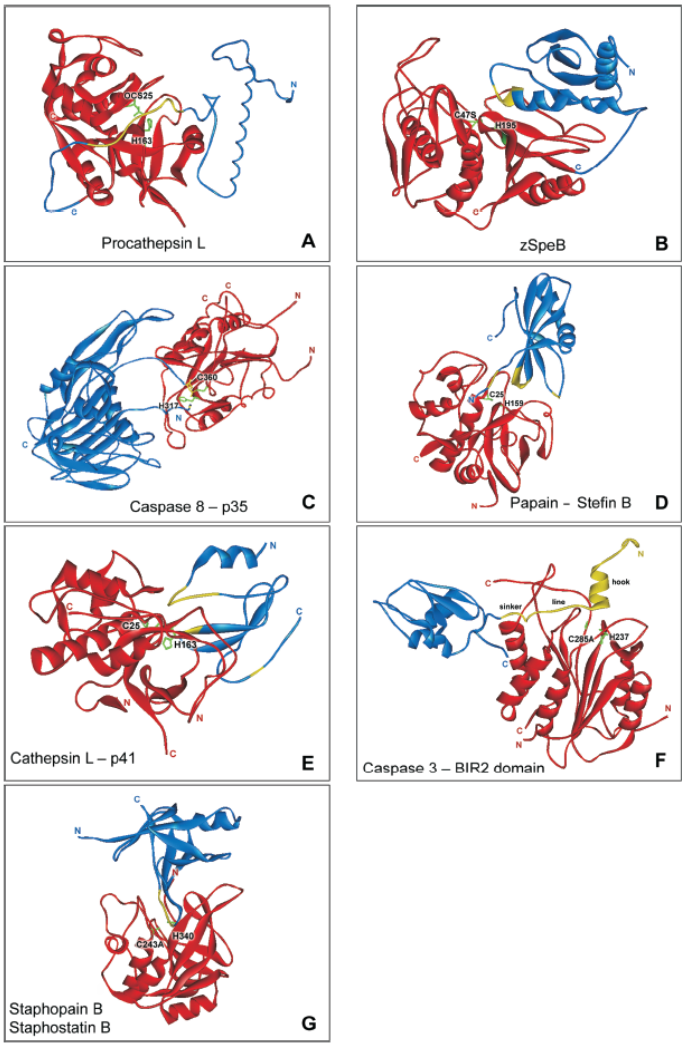


Figure 8*. Ribbon drawing of inhibitory complexes formed by A, procathepsin L (PDB code: 1cs8); B, SpeB zymogen (1dki); C, p35 and caspase 8 (1i4e); D, stefin B and papain (1stf); E, p41 and cathepsin L (1icf); F, XIAP and caspase 3 (1i3o); G, staphostatin B and staphopain B (1pxv).

Inhibitors are shown in blue, and enzymes in red. The location of the catalytic Cys and His residues is shown in green; regions of inhibitor that are most important for interaction with enzyme are labeled yellow.

*Adapted from, Rzychon, M., Chmiel D. and Stec-Niemczyk J. Acta. Biochemical Polonica (2004) 51:861–873.

Chagasin family

Protozoan parasites transmitted by insect bites, such as *Trypanosoma cruzi* (a causative agent of Chagas' heart disease in Latin America), *Trypanosoma brucei* (responsible for sleeping sickness in sub-Saharan Africa), *Leishmania* sp. (*leishmaniasis*) and *Plasmodium* sp. (malaria) produce large quantities of papain-like (family C1) lysosomal proteases. In plants and mammals, proteases like – cathepsins are regulated by endogenous inhibitors of the cystatin family. A search for similar regulators of parasite cysteine proteases resulted in the discovery of a novel inhibitor family distinct from cystatins and other known groups.

Chagasin is a *T. cruzi* inhibitor of the endogenous cysteine protease cruzipain (62). Subsequently, homologues were identified in the genomes of all mentioned parasites as well as in bacteria (63). While genomes of all but one of those organisms encode cruzipain-like proteins, the function of *Pseudomonas aeruginosa* chagasin homologue remains puzzling since the bacterium does not seem to produce any family C1 proteases. Some of the identified proteins were recombinantly expressed and demonstrated to inhibit papain like cysteine proteases (64).

Calpastatin

Calpastatin is a specific inhibitor of calcium-dependent cysteine proteases – calpains. It does not inhibit any other cysteine proteases tested to date. Likewise, activity of calpains, is regulated only by calpastatin. (Though Kininogens are also able to inhibit calpain proteases, their extracellular localization excludes them as regulators of intracellular calpains, and they are thought to serve only as protectants against accidental calpain release (65).

Thyropins (thyroglobulin type-1 domain proteinase inhibitors)

Thyroglobulin type-1 domain (Thyr-1) is a cysteine-rich structural element found either as a single or repetitive module in a variety of functionally unrelated proteins such as a precursor of thyroid hormones, thyroglobulin, basement membrane protein, nidogen (entactin), testicular proteoglycan, testican, insulin-like growth factor binding proteins (IGFBPs), pancreatic carcinoma marker proteins (GA733), major histocompatibility

complex class II (MHC II)-associated p41 invariant chain, equistatin, chum salmon egg cysteine protease inhibitor (ECI) and saxiphilin (66-69).

MHC class II-associated p41-invariant chain fragment

A crucial function of MHC II is to present bound peptides, derived mainly from extracellular antigens, to CD4+ T lymphocytes. The invariant chain (Ii) associates with MHC class II molecules in endoplasmic reticulum, its main function being to block the peptide-binding site during early stages of intracellular transport. Upon arrival at endosomes/lysosomes, Ii is proteolytically degraded, thus allowing the MHC II to bind peptides derived from endocytosed extracellular antigens. Both antigen processing by means of limited proteolysis and Ii degradation are mediated by endosomal proteases. A protein fragment corresponding to p41 Thyr-1 domain was demonstrated to efficiently inhibit cathepsin L, cruzipain and to a lesser extent cathepsin H as well as papain (70). The inhibitory properties towards antigen processing proteases implicate possible modulation of antigen presentation by alternative splicing of Ii. Different lines of evidence in favor of the invariant chain exerting inhibitory and modulatory activity in vivo are delineated below. Proportions of p41 in relation to p31 vary in different cells, suggesting a probable way to regulate antigen presentation. The former can modify the degradation pattern of the latter, demonstrating direct interference with proteolytic pathways.

IAP family

Inhibitors of apoptosis (IAPs) are a family of proteins distinguished by encompassing one or more characteristic, ~70-residue zinc binding BIR (baculovirus IAP repeat) domains. They were primarily characterized as inhibitors of apoptosis (71), though currently some BIR-containing proteins are known which do not seem to confer such a function. Apoptosis (programmed cell death) is one of the crucial events controlling the number of cells in multicellular eucaryotes.

Propeptide-like inhibitors

Papain-like cysteine proteases are produced as inactive precursors. The lack of activity is due to potent inhibition by the N-terminal propeptides. Only at the target location does limited proteolysis of the propeptide occur, allowing the enzyme to exert its action. Fox and colleagues (72) were the first to show that synthetic proregions inhibit their cognate proteases. Later studies demonstrated that propeptides are competitive, slow-binding inhibitors and unlike the broad-spectrum cystatins, possess high selectivity for the enzymes from which they originate. The inhibition is pH dependent (acidification decreases the affinity) being in good agreement with the process of enzyme activation, since pH is one of the most common environmental parameters in triggering the process. Based on structural data the mode of inhibition was elucidated. The propeptide spans the protease active site in an orientation opposite to that of a substrate, thus escaping cleavage. Besides being inhibitors, the propeptides assist in proper folding and targeting, and stabilize the cognate enzymes (73).

Cysteine proteases involved in plant – pathogen interactions

Plants use PLCPs (Papain-like cysteine proteases) to protect themselves against pests and pathogen attack. Examples are papain from papaya and Mir1 (Maize inbred resistance 1) from maize, both acting against insect larvae. Tomato Rcr3 (Required for *Cladosporium* resistance-3) and PIP1 (*Phytophthora* inhibited protease-1) are produced upon pathogen attack and inhibited by pathogen-derived inhibitors.

Papain is a component of latex of papaya trees, which pours out of wounds, presumably as a defence response against herbivores (74). Papain is produced as a preproprotein, and mechanical wounding of papaya fruit enhances papain accumulation and activation (75, 76). However, the mechanism of its accumulation and how it is activated are not yet clear. A role of papain in insect defence has been described only recently. Different lepidopteran caterpillars (*Samia ricini*, *Mamestra brassicae* and *Spodoptera litura*) showed reduced larval weight when fed with leaves containing papain (77). This reduced growth was not observed when the latex was washed out or when the leaves were treated with the cysteine protease inhibitor E-64 (77). This indicates that papain contributes to defence against herbivores.

Mir1 was identified because it was encoded by an abundant transcript in the callus of resistant, but not susceptible, maize when challenged with armyworms (*Spodoptera fugiperda*) (78, 79). Like papain, Mir1 is translated as a preproprotein, suggesting that it is secreted or localized in vesicles. Mir1 protein accumulation occurs rapidly one hour after larval feeding, continues for 7 days and is most abundant at the feeding site (79). Tobacco budworm (*Heliothis virescens*) larvae fed with transgenic maize callus overexpressing the Mir1 gene were significantly smaller than those fed with callus from control plants (80, 81). Feeding on Mir1-producing plants causes severe damage of the caterpillar peritrophic matrix, which is the chitin structure covering the insect gut surface, protecting it from chemical and physical damage (82). It has been suggested that Mir1 can bind to chitin, localizing the proteolytic activity to the insect gut (82).

PIP1 and Rcr3 are two secreted tomato PLCPs that accumulate in the apoplast (83, 84). The Pip1and Rcr3 genes map at the same genetic locus of tomato and are transcriptionally up-regulated during pathogen challenge (83, 84). Both proteases are inhibited by pathogen-derived inhibitors. PIP1 is inhibited by EPIC2B, a cystatin-like protease inhibitor secreted during infection by the oomycete *Phytophthora infestans* (84). Rcr3 is inhibited by Avr2, a secreted, cysteine-rich protein produced by the leaf mould fungus *Cladosporium fulvum* (85, 86). The Rcr3–Avr2 complex, and not Rcr3 inhibition itself, triggers the hypersensitive response mediated by the tomato resistance gene Cf-2 (86). However, the specificity of inhibition by Avr2 and EPIC2B for PLCPs and how Cf-2 recognizes the Rcr3–Avr2 complex are not yet fully understood.

RD21 from *Arabidopsis* is a papain-like protease (C1) that accumulates in vesicles that originate from the endoplasmic reticulum (ER; hence called ER bodies) (87). An RD21-like protease is also induced in potato upon infection with *Phytophthora infestans* (88). A role for ER bodies in defence against insects was suggested by the observation that ER bodies are induced by wounding and insect feeding (89), and that they fuse with the vacuole upon stress (90). ER bodies may therefore be involved in defence against insects and this suggests that RD21 may also play a role in this defence (87).

Future perspectives

The major areas of interest for protease-targeted therapies are likely to remain the cardiovascular, inflammatory and infectious diseases areas, but discovery efforts will probably increase for cancer and neurodegenerative disorders. A more detailed look at the proteases currently considered as potential targets shows that endogenous proteases are often linked to chronic diseases and are therefore attractive to pharmaceutical companies. However, there has recently been a revived interest in protease-targeted drugs for infectious diseases, although only a few (AIDS, hepatitis C) were considered seriously until recently. With the emergence of new technologies, and as the structure and physiological function of more proteases are revealed and updated, it should hopefully become easier both to identify and validate proteases as relevant drug targets and to develop effective and safe drugs against them.

Exciting as recent data are, we only have just begun to appreciate the different roles of proteases in plant defence. Proteases are implicated in perception, signalling and execution leading to plant defence. Papain-like cysteine proteases (clan CA) may be involved in each of these steps, whereas caspase-like cysteine proteases (metacaspases and VPEs [clan CD]) might play a role in signalling.

Studies with protease inhibitors, despite their seemingly contradictory results, strongly indicate crucial roles for (especially cysteine) proteases in signalling, even though the corresponding protease genes have not yet been identified. The expected sites of action of the protease inhibitors that have been used, as well as the localization of well-studied proteases, indicate roles for proteases present at different subcellular locations, such as the cytosol, apoplast and organelles.

The huge number of proteases, their redundancy and their posttranscriptional regulation makes this field a challenge to explore. A new functional proteomic technology, called protease-activity profiling, which displays multiple protease activities rather than their transcript or protein abundance (91), provides an excellent tool with which to study the activities of proteases during defence responses, and has recently been introduced in plants (92). This technology, combined with detailed studies of candidate proteases and other approaches, will uncover proteases that are involved in defence. Another exiting

question that remains to be solved is how these proteases are regulated and what their substrates are. It can be expected that research in the next decade will bring us astonishing discoveries on the roles of the proteolytic machinery in plant defence responses.

While considerable progress has been made with the structure, activity and importance (biological and commercial) of the cysteine proteases much remains to be learned about the distinctions of the various architecture and how these are reflected in the functions of the various enzymes. Inhibitors (naturally occurring and synthetic) have permitted detailed biochemical and crystallographic investigations to be made but an understanding of the selectivity of such inhibitors may be of just as much importance for the design and synthesis of specific inhibitors for use therapeutically in controlling individual cysteine proteases.

Present Investigation

The present work details various aspects of cysteine proteases and their inhibition. The findings of the investigations have been presented in the following chapters:

1. General introduction.
2. Fungal effector protein AVR2 targets diversifying defense-related cys proteases of tomato.
3. Contrasting selectivity of fungal and oomycete effectors targeting tomato cysteine proteases.
4. Structure-function analysis of RD21, a stress induced cysteine protease in *Arabidopsis thaliana.*
5. RIP1-Chagasin like cysteine protease inhibitor from *Pseudomonas syringae* DC3000, its biophysical studies.
6. Purification and partial characterization of cysteine protease inhibitor (CPI) from salt tolerant pathogenic fungus *Penicillium chrysogenum.*
7. General discussion and conclusion.

REFERENCES

1. Barrett, A. J. and Rawlings, N. D. (1995) Arch. Biochem. Biophys. 318 : 247-250.
2. Boris, T. (2006) Nature Reviews Drug Discovery 5: 785-799.
3. Menon, A.S., and Goldberg, A.L. (1987) J. Biol. Chem. 262: 14921-14928.
4. Enari, M., Sakahira, H., Yokoyama, H., Okawa, K., Iwamatsu, A., Nagata, S. (1998) Nature 391: 43-50.
5. Moser, B., Wolf, M., Walz, A. and Loetscher, P. (2004) Trends Immuno. 25 :75-84.
6. Knight, W.B., Green, B.G., Chabin, R., Gale, P., Maycock, A., Weston, H. (1991) Abs. Pap. Am. Chem. Soc. 202: 50-54.
7. Lowe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W. and Huber, R. *Science* (1995) 268: 533–539.
8. Seemuller, E., Lupas, A., Zuhl, F., Zwickl, P. and Baumeister, W. (1995) FEBS Lett. 359 : 173-178
9. Fujinaga, M.M., Cherney, H. Oyama, K. Oda and M.N. James, (2004) Proc. Natl. Acad. Sci. U.S.A. 101: 3364–3369.
10. Oda, K. Sedolisin. In Handbook of Proteolytic Enzymes, 2 Edn (Barrett,A.J., Rawlings, N.D. & Woessner, J.F. Eds), p.1886-1887, Elsevier, London (2004) V
11. Tai, J.Y., Kortt, A.A., Liu, T.Y. and Elliott, S.D. (1976) J. Biol. Chem. 251: 1955-1958.
12. Kortt, A. A. and Liu, T.Y. (1973) Biochemistry 12: 328-337
13. Kamphuis, I.G., Kalk, K.H., Swarte, M.B., Drenth, J. (1984) J. Mol. Biol. 179: 233-256.
14. In Handbook of Proteolytic Enzymes, 2 Edn (Barrett, A.J., Rawlings, N.D. & Woessner, J.F. eds), p.1886-1887, Elsevier, London (2004) V
15. Lecaille, F., Kaleta, J. and Bro¨mme, D. (2002) Chem. Rev. 102: 4459-4488
16. Jasmer, D. P., Roth, J., and Myler, P. (2001) J. Mol. Biochem. Parasitol. 116: 159-162.
17. Falcone, F. H., Tetteh, K. K., Hunt, P., Blaxter, M. L., Loukas, A., and Maizels, R. M. (2000) Exp. Parasitol. 94: 201-205

18. Pierre, P., and Mellman, I. (1998) Cell 93: 1135-1137.
19. Kos, J., Lah, T. T. (1998) Oncol. Rep. 5: 1349.
20. Maizels, R. M., Gomez-Escobar, N., Gregory, W. F., Murray, J., Zang, X. (2001) Int. J. Parasitol. 31: 889-894.
21. Schultz, R., Varma-Nelson, P., Oritz, R., Kozlowski, K., Orawski, A., Pagast, P., Frankfater, A. (1989) J. Biol. Chem 264 : 1497- 1507.
22. Westerik, J., Wolfenden, R. (1972) J. Biol. Chem. 247: 8195- 8197.
23. Mattis, J., Henes, J., Fruton, J. (1977) J. Biol. Chem. 252: 6776- 6782.
24. Frommer, W., Junge, B., Muöller, L., Schmidt, D., Truscheit, E. (1979) Planta Med. 35: 195-215.
25. Imperiali, B., Abeles, R. (1986) Biochemistry 25: 3760-3767.
26. Dunlap, R., Stone, P., Abeles, R. (1987) Biochem. Biophys. Res. Commun. 145: 509-513.
27. Stein, R., Strimpler, A., Edwards, P., Lewis, J., Mauger, R., Schwartz, J., Stein, M., Trainor, D., Wildonger, R., Zottola, M. (1987) Biochemistry 26: 2682-2689.
28. Peet, N., Burkhardt, J., Angelastro, M., Giroux, E., Mehdi, S. Bey, P., Kolb, M., Neises, B., Schirlin, D. (1990) J. Med. Chem. 33: 394-407.
29. Hanada, K., Tamai, M., Yamagishi, M., Ohmura, S., Sawada, J., Tanaka, I. (1978) Agric. Biol. Chem. 42: 523-528.
30. Hanada, K., Tamai, M., Ohmura, S., Sawada, J., Seki, T., Tanaka, I. (1978) Agric. Biol. Chem. 42: 529-536.
31. Hanada, K., Tamai, M., Morimoto, S., Adachi, T., Ohmura, S., Sawada, J., Tanaka, I. (1987) Agric. Biol. Chem. 42: 537-541.
32. Barrett, A., Kembhavi, A., Brown, M., Kirschke, H., Knight, C., Tamai, M., Hanada, K. (1982) Biochem. J. 201: 189-198.
33. Schaschke, N., Assfalg-Machleidt, I., Machleidt, W., Moroder, L. (1998) FEBS Lett. 421: 80-84.
34. Greenbaum, D., Baruch, A., Hayrapetian, L., Darula, Z., Burlingame, A., Medzihradszky, K.F., and Bogyo, M. (2002) Mol. Cell. Proteomics 1: 60–68.
35. Brinen, L. S., Hansell, E., Cheng, J., Roush, W. R., McKerrow, J. H., Fletterick, R. (2000) J. Struct. Fold Des. 8: 831-834.

36. Rosenthal, P. J., Olson, J. E., Lee, G. K., Palmer, J. T., Klaus, J. L., Rasnick, D. (1996) Antimicrob. Agents Chemother. 40: 1600-1607.
37. Olson, J. E., Lee, G. K., Semenov, A., Rosenthal, P. J. (1999) Bioorg. Med. Chem. 7: 633-637.
38. Falgueyret, J. P., Oballa, R. M., Okamoto, O., Wesolowski, G., Aubin, Y., Rydzewski, R. M., Prasit, P., Riendeau, D., Rodan, S. B., Percival, M. D.(2001) J. Med. Chem. 44: 94-96.
39. Greenspan, P. D., Clark, K. L., Tommasi, R. A., Cowen, S. D., McQuire, L. W., Farley, D. L., van Duzer, J. H., Goldberg, R. L., Zhou, H., Du, Z., Fitt, J. J., Coppa, D. E., Fang, Z., Macchia, W., Zhu, L., Capparelli, M. P., Goldstein, R., Wigg, A. M., Doughty, J. R., Bohacek, R. S., Knap, A. K. (2001) J. Med. Chem. 44 : 4524-4529.
40. Moon, J., Coleman, R. and Hanzlik, R. (1986) J. Am. Chem. Soc. 108 : 1350-1351.
41. Singh, R., Zhou, N. E., Guo, D., Kaleta, J., Cameron, A., Purisima, E,; Menard, R., Micetich, R. G. WO 97/38008; 1997.
42. Singh, R., Zhou, N. E., Guo, D., Kaleta, J., Micetich, R. G. US 5,925,633 A 1999.
43. Singh, R., Zhou, N. E., Guo, D., Micetich, R. G. US 6,232,305 B1 2001.
44. Yamashita, D. S., Smith, W. W., Zhao, B., Janson, C. A., Tomaszek, T. A., Bossard, M. J., Levy, M. A., Oh, H.-J., Carr, T. J., Thomson, S. K., Ijames, C. F., McQueney, M., D'Alessio, K. J., Amegadzie, B. Y., Hanning, C. R., Abdel-Meguid, S., Desjarlais, R. L., Gleason, J. G., Veber, D. F. (1997) J. Am. Chem. Soc. 119: 11351-11354.
45. Abrahamson, M., Alvarez-Fernandez, M. and Nathanson, C.M. (2003) Biochem. Soc. Symp. 70:179-199.
46. Belenghi, B., Acconcia, F., Trovato, M., Perazzolli, M., Bocedi, A., Poticelli, F., Ascenzi, P., Delledonne M. (2003) Eur. J. Biochem. 270:2593-2604.
47. Dublin, G. (2005) Cell.Mol.Life Sci.62: 653-669.
48. Kondo, H., Abe, K., Emori, Y. and Arai, S. (1991) FEBS Lett. 278: 87-90.
49. Margis, R., Reis, E.M. and Villeret, V. (1998) Arch. Biochem. Biophys. 359: 24-30.

50. Solomon, M., Urwin, P., Atkinson, H.J., Waller, D.A. and McPherson, M.J. (1995) Plant J. 8: 121-131
51. Arai, S., Matsumoto, I., Emori, Y., and Abe, K. J.Agric. (2002) Food Chem., 50: 6612-6617
52. M. Pernas, R. Sanchez-Mong, G. Salcedo, (2000) FEBS Lett. 467:206– 210.
53. H. Aoki, T. Akaike, K. Abe, M. Kuroda, S. Arai, R. Okamura, A. Negi, H. Maeda, (1995) Antimicrob. Agents Chemother. 39: 846–849.
54. Soares-Costa, A., Beltramini, L.M., Thiemann, O.H., Henrique-Silva, F. (2002) Biochem. Biophys. Res. Commun. 296 1194– 1199.
55. Pernas, M. R., Sanchez-Monge, L., Gomez, G. Salcedo, (1998) Plant Mol. Biol. 38: 1235–1242.
56. Koiwa, H. R., Shade, E., Zhu-Salzman, K. D'Urzo, M.P., Murdock, L.L., Bressan, R.A., Hasegawa, P.M., (2000) FEBS Lett. 471: 67–70.
57. Oppert, B. T.D., Morgan, K., Hartzer, B. Lenarcic, K. Galesa, J. Brzin, V. Turk, K. Yoza, K. Ohtsubo, K.J. Kramer, Comp. Biochem. Physiol. (2003) 134: 481–490.
58. M. Solomon, B. Belenghi, M. Delledonne, E. Menachem, A. Levine, Plant Cell (1999) 11: 431–443.
59. Tian, M., Win, J., Van der Hoorn, R., Van der Knaap, E., and Kamoun, S. (2007) Plant Physiol. 143: 364–377.
60. Strobl, S., Fernandez-Catalan, C., Braun, M., Huber, R., Masumoto, H., Nakagawa, K. (2000) Proc. Natl. Acad. Sci. USA 97**:** 588–592.
61. Rzychon M., Sabat A., Kosowska K., Potempa J. and Dubin A. (2003) Mol. Microbiol. 49**:** 1051–1066.
62. Monteiro, A. C. S., Abrahamson, M., Lima, A. P. C. A., Vannier- Santo, M. A. and Scharfstein, J. (2001) J. Cell Sci. 114: 3933–3942.
63. Rigden, D. J., Mosolov, V. V. and Galperin, M. Y. (2002) Prot. Sci. 11: 1971–1977.
64. Sanderson, S. J., Westrop, G. D., Scharfstein, J., Mottram, J. C. and Coombs, G. H. (2003) FEBS Lett. 542**:** 12–16.

65. Hao, L. Y., Kameyama, A., Kuroki, S., Takano, J., Takano, E., Maki M. et al. (2000) Biochem. Biophys. Res. Comm. 279: 756–761.
66. Ogrinc, T., Dolenc, I., Ritonja, A. and Turk, V. (1993) FEBS Lett. 336**:** 555–559.
67. Yamashita, M. and Konagaya, S. (1996) J. Biol. Chem. 271**:** 1282–1284.
68. Lenarˇciˇc, B., Ritonja, A., Strukelj, B., Turk, B. and Turk, V. (1997) J. Biol. Chem. 272: 13899–13903.
69. Lenarˇciˇc, B., Krishan, G., Borukhovic, R., Ruck, B., Turk, V. and Moczydlowski, E. (2000) J. Biol. Chem. 275: 15572–15577.
70. Bevec, T., Stoka, V., Pungerˇciˇc, G., Dolenc, I. and Turk V. (1996) J. Exp. Med. 183: 1331–1338.
71. Crook, N. E., Clem, R. J. and Miller, L. K. (1993) J. Virol. 67: 2168–2174.
72. Fox, T., de Miguel, E., Mort, J. S. and Storer, A. C. (1992) Biochemistry 31: 12571–12576.
73. Wiederanders, B. (2003) Acta. Biochim. Pol. 50: 691–713.
74. A. El Moussaoui, M. Nijs, C. Paul, R. Wintjens, J. Vincentrelli and M. Azarkan et al., (2001) Cell Mol Life Sci. 58: 556–570.
75. Azarkan, M., Dibiani, R., Baulard, C. and Baeyens-Volant, D. (2006) Int. J. Biol. Macromol. 38: 216–224.
76. Moutim, V., Silva, L.G., Lopes, M.T.P., Wilson Fernandes, G. and Salas, C.E. (1999) Plant Sci. 142: 115–121.
77. Konno, K., Hirayama, C., Nakamura, M., Tateishi, K., Tamura, Y., Hattori, M. and Kohno, K. (2004) Plant J. 37: 370–378.
78. Jiang, B., Siregar, U., Willeford, K.O., Luthe, D.S. and Williams, W.P.(1995). Plant Physiol. 108: 1631–1640.
79. Pechan, T., Ye, L., Chang, Y., Mitra, A., Lin, L., Davis, F.M., Williams, W.P. and Luthe, D.S. (2000) Plant Cell 12: 1031–1040.
80. Chang, Y.M., Luthe, D.S., Davis, F.M. and Williams, W.P. (2000) J. Econ. Entomol. 93: 477–483.
81. Pechan, T., Ye, L., Chang, Y., Mitra, A., Lin, L., Davis, F.M., Williams, W.P.and Luthe, D.S. (2000) Plant Cell 12: 1031–1040.

82. Pechan, T., Cohen, A., Williams, W.P. and Luthe, D.S. (2002) Proc. Natl. Acad. Sci. USA 99: 13319–13323
83. Kruger, J., Thomas, C.M., Golstein, C., Dixon, M.S., Smoker, M., Tang, S., Mulder, L. and Jones, J.D.G. (2002) Science 296: 744–747.
84. Tian, M., Win, J., Song, J., van der Hoorn, R.A.L., van der Knaap, E. and Kamoun, S. (2007). Plant Physiol. 143: 364–377.
85. Luderer, R., Takken, F.L.W., De Wit, P.J.G.M. and Joosten, M.H.A.J. (2002). Mol. Microbiol. 45: 875–884.
86. Rooney, H.C.E., Van't Klooster, J.W., van der Hoorn, R.A.L., Joosten, M.H., Jones, J.D.A.J. and de Wit, P.J.G.M. (2005) Science 308: 1783–1786.
87. Hara-Nishimura I, Matsushima R (2003) Curr. Opin. Plant Biol. 6:583-588.
88. Avrova, A.O., Stewart, H.E., De Jong, W.D., Heilbronn, J., Lyon, G.D., Birch, P.R. (1999) Mol. Plant Microbe Interact. 12:1114-1119.
89. Matsushima, R, Hayashi Y, Kondo M, Shimada T, Nishimura M, Hara-Nishimura I Plant Physiol. (2002) 130:1807-1814.
90. Hayashi, Y., Yamada K., Shimada, T., Matsushima, R., Nishizawa, N.K, Nishimura, M., Hara-Nishimura, I. (2001) Plant Cell Physiol. 42:894-899.
91. Campbell, D.A., Szardenings, A.K. (2003) Curr. Opin. Chem. Biol. 7:296-303.
92. Van der Hoorn, R.A.L., Leeuwenburgh, M.A., Bogyo, M., Joosten, M.H.A.J., and Peck, S.C. (2004) Plant Physiol. 135: 1170–1178.
93. Rzychon, M., Chmiel, D. and Stec-Niemczyk, J. (2004) Acta. Biochemical Polonica 51:861–873.

CHAPTER 2

FUNGAL EFFECTOR PROTEIN AVR2 TARGETS DIVERSIFYING DEFENSE-RELATED CYS PROTEASES OF TOMATO

SUMMARY

The interaction between the fungal pathogen *Cladosporium fulvum* and its host tomato (*Solanum lycopersicum*) is an ideal model to study suppression of extracellular host defenses by pathogens. Secretion of protease inhibitor AVR2 by *C. fulvum* during infection suggests that tomato papain-like cysteine proteases (PLCPs) are part of the tomato defense response. We show that the tomato apoplast contains a remarkable diversity of PLCP activities with seven PLCPs that fall into four different subfamilies. Of these PLCPs, transcription of only PIP1 and RCR3 is induced by treatment with benzothiadiazole, which triggers the salicylic acid–regulated defense pathway. Sequencing of PLCP alleles of tomato relatives revealed that only PIP1 and RCR3 are under strong diversifying selection, resulting in variant residues around the substrate-binding groove. The doubled number of variant residues in RCR3 suggests that RCR3 is under additional adaptive selection, probably to prevent autoimmune responses. AVR2 selectively inhibits only PIP1 and RCR3, and one of the naturally occurring variant residues in RCR3 affects AVR2 inhibition. The higher accumulation of PIP1 protein levels compared with RCR3 indicates that PIP1 might be the real virulence target of AVR2 and that RCR3 acts as a decoy for AVR2 perception in plants carrying the Cf-2 resistance gene.

INTRODUCTION

Considering that invasion of the plant apoplast is a critical phase of the infection cycle of numerous pathogens, it can be postulated that this compartment serves as a molecular battlefield that contributes to the success of pathogen infection and plant resistance. This battlefield is likely to be ancient, predating the evolution of translocation mechanisms for effector proteins by which pathogens manipulate the host cytoplasm and suppress host defense responses (1). Therefore, understanding the nature of plant defenses in the apoplast and the counter defense mechanisms that pathogens evolved to overcome these defenses is essential for a comprehensive understanding of host–pathogen interactions and should complement the body of knowledge that has emerged on cytoplasmic effectors and defense mechanisms.

The apoplast of tomato (*Solanum lycopersicum*) is easily accessible for biochemical experiments and ideal for studying apoplastic molecular plant–pathogen interactions. Tomato is the only host for the well-studied fungal pathogen *Cladosporium fulvum* and an important host for the devastating solanaceous oomycete pathogen *Phytophthora infestans* (2, 3, 4). Infection of tomato with these pathogens triggers the accumulation of a large amount of pathogenesis-related (PR) proteins in the apoplast (5, 6). For example, PR proteins such as β-1,3-glucanase (PR2), chitinases (PR3), and subtilase-like serine protease (P69B) directly target these pathogens by degrading their cell wall components (7,8). Pathogens, on the other hand, secrete different effector proteins during infection to prevent this degradation. *P. infestans* secretes glucanase and subtilase inhibitors, whereas *C. fulvum* secretes chitin binding Avr4 proteins to protect its cell wall (9, 10, 11, 12, 13). Selection pressure on host enzymes to evade inhibition and pathogen inhibitors to adapt to new host enzymes cause variant residues at the interaction surface for many of these enzyme– inhibitor interactions (14).

Although apoplast-localized plant proteases can play a role in defense, they might also act in signaling or perception upon pathogen infection (15). A role in signaling has been indicated for secreted proteases CDR1 and CathB. CDR1 is an aspartic protease that probably releases systemic signaling molecules that trigger defense responses in *Arabidopsis thaliana* (16). CathB is a secreted papain-like cysteine protease (PLCP) required for the development of the hypersensitive response (HR) in *Nicotiana benthamiana* (17). HR is an effective defense response at the site of pathogen infection that includes programmed cell death.

A very interesting role has been documented for RCR3, a secreted PLCP of tomato. RCR3 is essential for the function of the tomato resistance gene Cf-2, which mediates recognition of the pathogenic fungus *C. fulvum* carrying the avirulence gene Avr2 (18). This gene-for-gene recognition event triggers an effective defense response that includes HR, making RCR3 an essential component of a pathogen perception system. However, RCR3 is transcriptionally regulated as a PR protein, and the secreted AVR2 protein binds and inhibits RCR3 (18, 19). This mechanism of perception suggests that RCR3 is rather a

virulence target of AVR2 that became guarded by the Cf-2 resistance protein to monitor pathogen entry (1, 19, 20).

The hypothesis that PLCPs can be virulence targets is also supported by the discovery that *P. infestans* secretes PLCP inhibitors during infection of tomato (21). However, the cystatin-like proteins EPIC1 and EPIC2B are not homologous to AVR2 bind and inhibit PIP1, a PLCP that is closely related to RCR3 and accumulates in the apoplast as a PR protein (21). These data indicate that tomato plants secrete PLCPs during defense to create a proteolytic apoplast that is harmful to pathogens. However, it is unknown what the full content of apoplastic PLCP activities is, to what extent these secreted PLCPs are inhibited by AVR2, and other pathogen derived inhibitors. Description of these defensive proteases and their inhibition would therefore uncover a new layer of molecular host–pathogen interactions that would allow us to study coevolution of proteases, substrates, and inhibitors as well as resistance proteins that monitor the inhibition of these proteases.

In this study, we used a biochemical approach to investigate which PLCPs are active in the tomato apoplast during the benzothiadiazole (BTH)-triggered defense response and to determine the extent to which these PLCPs are inhibited by AVR2. Our results suggest a remarkable diversity of PLCP activities in the apoplast. Tomato secretes seven different active PLCPs during defense that belong to four phylogenetic subfamilies. Among these, PIP1 and RCR3 are transcriptionally upregulated during BTH-induced defense, with PIP1 dominating apoplastic PLCP activities. Sequencing PLCP alleles of wild tomato species showed that PIP1 and RCR3 are under diversifying selection, suggesting their participation in coevolutionary arms races with plant pathogens. Finally, by expressing the proteases separately through transient agroinfiltration assays, we demonstrated that AVR2 inhibits PIP1 and RCR3 and that a naturally occurring variant residue in RCR3 alters its susceptibility to being inhibited.

MATERIALS AND METHODS

Plant Materials

Nicotiana benthamiana and tomato (*Solanum lycopersicum* Money Maker [MM] Cf0, Cf2, and Cf2/rcr3-3) were grown in a climate chamber at a 14-h light regime at 18°C (night) and 22°C (day). Four- to six-week old plants were used for experiments. BTH treatment was done by watering 5-week-old tomato plants with 25 mg/mL BTH (Actigard; Syngenta) or water every second day. Samples were taken at 5 d after starting the BTH treatment, unless otherwise indicated.

Production and Purification of AVR2 Protein

All DNA manipulations were performed using standard protocols (22). The ORF of Avr2 (23) was amplified from *Cladosporium fulvum* race 5 by RT-PCR using primers 5'TCCCCGCGGGCCAAAAAACTACCTGGCTGCGA-3' and 5' CGGGGTACCTCAACCGCAAAGACCAAAACAGCA- 3' and cloned into pFLAG-ATS (Sigma-Aldrich) using SacII and KpnI restriction enzymes, resulting in pFLAG-AVR2 (pFK3). FLAG-AVR2 was expressed and purified on an anti-FLAG matrix as described previously (11). Protein purities were checked on Coomassie Brilliant Blue–stained 17% protein gels and quantified using the Bradford assay (Bio Rad).

Protein Work

AFs (Apoplastic fluid) were isolated by vacuum infiltration of tomato leaves with ice-cold water. Infiltrated leaves were dried on the surface and centrifuged (10 min, 1600g) in a tube with holes in the bottom. The AF was collected below in a larger collection tube. Equal volumes (200 to 450 mL) were used for protease activity profiling, protein gel blot analysis, or Coomassie blue staining. In each case, samples were 10-fold concentrated by precipitating with 67% ice-cold acetone and dissolving the protein pellet in protein gel-loading buffer. TEs were generated by grinding six randomly taken 28-mm^2 leaf discs into 1 mL of water. TEs were centrifuged (1 min, 16,100g), and the supernatant was analyzed as described for AFs.

Protease activity profiling was performed as described previously (18). Unless otherwise indicated, 200 to 450 mL of TE or AF was labeled for 5 h at room temperature with 2 mM DCG-04 (18) in the presence of 50mM sodium acetate, pH 5.0, and 1 mM L-cysteine. Competition with E-64 was done by adding 20 mM E-64 (Sigma-Aldrich) 30 min before adding DCG-04. Proteins were precipitated by adding 1mL ice-cold acetone and centrifugation (1min, 16,100g). Pellets were dissolved in SDS gel loading buffer. Detection of biotinylated proteins was done as previously described using streptavidin-HRP polymer (Sigma-Aldrich) (18).

Protein gel blot analysis was performed as described for detection of biotinylated proteins, using RCR3 and PIP1 antibodies (19, 21), followed by detection with HRP-conjugated anti-rabbit antibodies (Amersham). For large-scale labeling of AF, 100 mL of AF was isolated as described above from 6-week-old tomato plants, 7 d after BTH treatment, and labeled with 2 mM DCG-04 at pH 5.0 (25 mM sodium acetate and 1 mM L-cysteine). Large-scale in situ labeling of tomato leaves by DCG-04 infiltration was done by vacuum infiltrating; 183 tomato leaflets with water containing 2 mM DCG-04. Leaves were incubated for 5 h at room temperature. Ice-cold water was vacuum infiltrated, and AF was quickly isolated as explained above.

AF proteins were precipitated using 67% ice-cold acetone, dissolved in TBS (50 mM Tris, pH 7.5, and 150 mM NaCl) containing a protease inhibitor cocktail (Complete tablets; Roche), and incubated for 1 h at room temperature. Undissolved proteins were removed by centrifugation (10 min, 3000g), and 500 mL of washed magnetic streptavidin beads (Promega) were added. Biotinylated proteins were captured by rotating overnight at room temperature. Beads were collected in a magnetic holder (Promega) and washed three times with washing buffer (1% Triton X-100, 1M NaCl, and 50 mM Tris, pH 7.5), three times with TBS (Tris buffer saline), and three times with water. Beads were boiled in gel loading buffer for 1D analysis or eluted with 50% formic acid for 2D analysis. For 2D analysis of purified biotinylated PLCPs, the formic acid eluate was dried and dissolved 160 mL of IEF buffer (7 M urea, 2 M thiourea, 2% CHAPS, 0.5% IPG buffer

[pI 3 to 10; Amersham], 0.002% bromephenol blue, and 20 mM DTT). The sample was impregnated overnight into an IPG ready strip (7 cm, 3-6L; Bio-Rad).

Isoelectric focusing was done in ZOOM IPG runner cassettes (Invitrogen) (15 min, 175 V; 45 min, 175 to 2000 V ramp; 25 min, 2000 V). After IEF, strips were incubated for 15 min in 4.5 mL LDS sample buffer (Invitrogen) supplemented with 0.5 mL sample reducing agent (Invitrogen) and then incubated for 15 min in 5 mL of LDS containing 116 mg iodoacetamide (Sigma-Aldrich). Strips were loaded on precasted NuPage 4 to 12% ZOOM gradient gels (Invitrogen) for 50 min at 200 V, and the gel was stained with SYPRO Ruby (Invitrogen).

After 2D electrophoresis, spots of various intensities were automatically excised and tryptically digested using the DP Chemical 96 kits for fully automated in-gel digestion (Proteineer spII/dp; Bruker). Aliquots of the digests were automatically prepared (Proteineer dp; Bruker) for subsequent matrix-assisted laser-desorption ionization time of flight (MALDI-TOF) MS analysis on Anchor Chip targets (Bruker) according to Gobom et al. (24). Mass spectra of tryptic peptides were taken with a Bruker Reflex IV MALDI-TOF MS. The obtained peptide mass fingerprints were processed in Xmass 5.1.16 (Bruker) and used to identify the corresponding proteins in the ProteinScape 1.3 database system (Protagen), which triggered MASCOT (Matrix Science) and Profound (Genomic Solutions) searches against the National Center for Biotechnology Information (NCBI) and TIGR genome databases.

For fluorescent protease activity profiling, 565 μL of AF isolated from water- or BTH-treated tomato plants was preincubated for 30 min with and without 83 nM AVR2 in 50 mM NaAc, pH 5. The remaining noninhibited proteases were labeled for 5 h with 3 mM DCG-04-TMR (25). Proteins were precipitated with trichloroacetic acid and acetone, washed, dissolved in IEF buffer, and separated on 2D gels as described above. Fluorescent proteins were detected using the Typhoon fluorescent imager (Molecular Dynamics). Gels were scanned simultaneously, and the standards had equal intensities.

For coimmunoprecipitation experiments, PIP1-His (21) was expressed by agroinfiltration of *N. benthamiana* (see below), and AF was isolated at 5 dpi. Forty microliters of PIP1-His–containing AFs were incubated with 50 mL of Ni-NTA agarose beads (Qiagen) for 2 h in 50 mM NaPO4, pH 8, 150 mM NaCl, and 1 mM imidazole. Beads were washed with 50 mM NaAc, pH 6.0, and incubated overnight in a cold room in the presence of absence of 10 mM E-64. Beads were washed with 50 mM NaAc, pH 6.0, and incubated for 20 min with or without 50 mL 260 nM AVR2. Beads were quickly washed with ice-cold 50 mM NaAc, pH 6, and eluted with 50 mL of 50 mM $NaPO_4$, pH 8.0 150 mM NaCl, and 500 mM imidazole. Eluted proteins were analyzed on protein blots using anti-PIP1 and anti-FLAG antibodies.

Agroinfiltration of Tomato Proteases

The pFK26 cloning vector was constructed as follows: the multiple cloning site (mcs) of pBluescript KSþ (Stratagene) was modified by digesting with KpnI and HindIII, treating with Klenow polymerase, and self-ligation, resulting in pFK18. The promoter-terminator (35S-TPI) from pRH80 (26) was amplified by PCR using primers r014 and r015 (Table1). Digesting the PCR product with PstI XhoI resulted in a 300- and 500-bp fragment. The 500-bp fragment (containing the 35S promoter of pRH80) was cloned into pFK18 using PstI and XbaI restriction sites, resulting in pFK19. The terminator region (TPI) was amplified from pRH80 using primers r013 and r014 (Table1) and cloned into pGEM-T (Promega), resulting in pFK22.

The TPI fragment was cloned from pFK22 into pFK19 using NcoI and EcoRI restriction sites, resulting in pFK26. The pTP5 binary vector was generated by cloning the 35S-mcs-TPI cassette of pFK26 into pRH385 (27) using XbaI and EcoRI restriction enzymes. The tomato protease ORFs were amplified by RT-PCR from RNA extracted from tomato leaves, cloned into pFK26, and sequenced. Expression cassettes containing correct ORF sequences were shuttled into binary vector pTP5 as follows.

Details of these cloning procedures are summarized in Table 3.RCR3 mutant proteins were generated by introducing mutations into pTP28 (35S:RCR3:term) using the Quick

change site-directed mutagenesis kit (Stratagene) and primers (see Supplemental Table 1 online). Mutations were confirmed by sequencing, and correct expression cassettes were shuttled into pTP5 using XbaI-SalI, resulting in pMS34 ($RCR3^{H148N}$), pMS35 ($RCR3^{R151Q}$), pMS36 ($RCR3^{H148N, R151Q}$), pMS37 ($RCR3^{N194D}$), and pMS38 ($RCR3^{Q267H}$).

Agrobacterium tumefaciens strain GV3101 was transformed with binary vectors and used for agroinfiltration as described previously (26). *Agrobacterium* was grown overnight at 28°C in 10mL Luria-Bertani medium containing 50 mg/mL kanamycin and 50 mg/mL rifampicin. The culture was centrifuged (10 min, 3000g) and the bacterial pellet resuspended in 10 mM MES, pH 5, 10 mM $MgCl_2$, and 1 mM acetosyringone to a final OD of 2. *Agrobacterium* cultures containing binary protease expression vectors were mixed with vector for silencing inhibitor p19 (28). *Agrobacterium* cultures were infiltrated into 5-week old *N. benthamiana* plants using a syringe without a needle. AF was isolated and the remaining tissue was ground in a mortar with water. This TE was centrifuged (10 min, 16,100g) and the supernatant stored in aliquots until further use. Both AF and TE were stored in aliquots at 20°C.

Inhibition assays with agroinfiltrated proteases were done as follows. AFs or total extracts from agroinfiltrated *N. benthamiana* were isolated at 5 dpi. In a total volume of 400 mL, 40 mL extract was preincubated in 50 mM NaAc, pH 5.0, 1 mM L-cysteine with 65 nM (or less) AVR2, or 10 mM E-64 for 30 min at room temperature. DCG-04 was added to a final concentration of 2 mM, and the labeling was performed for another 5 h at room temperature. Proteins were analyzed on protein blots using streptavidin-HRP or anti-RCR3 antibodies.

Phylogenetic Analysis

Sequences of 143 PLCPs (List1) were aligned using ClustalW 1.83 (29). The alignment was optimized manually, resulting in a file with a minimum number of gaps. Protein distances were calculated using the program Protdist from the PHYLIP suite of programs (30) with the JTT matrix on 1000 bootstraps. This data set (List1) was then used to generate phylogenetic trees based on neighbor joining. The resulting trees were then

joined to a consensus tree using the Consense program and plotted with Drawtree (30). The output from Drawtree was further edited in Adobe Illustrator.

Bioinformatic Analysis

For comparative tBLAST searches through the solanaceae TIGR database, the TIGR database of each of the solanaceous species was searched with queries (below), and the percentage of identity from the highest tBLASTp score was taken and averaged for all solanaceous species. The TIGR database contained 26,918 *N. benthamiana*, 29,894 pepper (*Capsicum annuum*), 8336 petunia (*Petunia hybrida*), 219,407 potato (*Solanum tuberosum*), 72,850 tobacco, and 213,947 tomato EST sequences (August, 2007). The accession codes used for the queries are in List 1. If known, only the domain of the mature protein was taken for tBLASTp searches.

Structural models of RCR3 and PIP1 were created using SWISSMODEL (31), using papain (PDB code 9papA) as a template. This template was chosen since it was one of the top hits given by the template search by SWISS-MODEL and because it is close to RCR3 and PIP1 in phylogenetic tree (32). The models were evaluated using Ramachandran plots and other outlier-detecting software of the MOE package (Molecular Operating Environment; Chemical Computing Group), and amino acid environment was evaluated using VERIFY-3D (33). All parameters were within the acceptable ranges, and no disturbing outliers were detected. Sequencing, Analysis of Variant Nucleotides and Residues, and Quantitative RT-PCR

For sequencing protease alleles from wild tomato species, RNA was isolated from (old) leaves (BTH-treated) using the Qiagen RNeasy mini kit, and cDNA was synthesized using Superscript II reverse transcriptase (Invitrogen) using an oligo(dT) 20 primer (Invitrogen) according to the instructions of the manufacturer. cDNA was used as a template for PCR using primers at the start and stop codons (Table 1). PCR products were sequenced using forward sequencing primers (Table 1). Sequences were aligned using Multialign (34), and variant nucleotides were verified in the trace data. Except for one nucleotide in CatB1 and one in CatB2, all generated sequences from *S. lycopersicum*

MM-Cf0 and *S. lycopersicum var cerasiforme* were identical to those at the NCBI database.

The new sequence of the *RCR3*^{pim} allele differs at one position compared with the sequence at the NCBI database. The *RCR3*^{pen} allele was identical to the previously published sequence at the NCBI database. In addition, except for one nucleotide in PIP1, all sequences from the two cheesmaniae species were identical. The polymorphism causing the N194D mutation in *RCR3*^{chi} was confirmed by sequencing genomic DNA. Taken together, these data underline the confidence of the generated sequences and indicate the presence of some polymorphic sites within some tomato species. Variant codons were analyzed in translation frames for their effect on the encoded amino acid. The ratio of nonsimilar/ similar residues was calculated from these data. To superimpose the variant residues onto the papain structure, an alignment between PIP1 and papain and between RCR3 and papain was used to select residues that are variant in PIP1 and RCR3. These residues were colored in the papain structure (1PPP) using PYMOL (62, 35).

For real-time RT-PCR, cDNA was synthesized as described above. Gene-specific primers were designed using Pearl Primer software and validated for specificity. Reaction mixtures for SYBR green (Roche)- based real-time RT-PCR was made as described previously (36). DNA synthesis was recorded with the IQ5 Multicolor Real Time PCR detection system (Bio-Rad). Threshold cycles (Ct) were recorded in triplicate over five independent biological samples, corrected for the Ct of ubiquitin (37), and subjected to statistical analysis following the manufacturer's guidelines (Bio-Rad).

Accession Numbers

Accession numbers for the sequences used to construct the phylogenetic tree shown in Figure 3B can be found in List 1 online and those used for Figure 3C are in List 2. Accession numbers for tomato relatives used as a template for sequencing are as follows: *S. lycopersicum* MM-Cf0; LA0927 (*S. cheesmaniae*), LA1407 (*S. cheesmaniae*), LA0442 (*S. pimpinellifolium*); LA1930 (*S. chilense*); LA0716 (*S. pennellii*); LA1777 (*S.*

habrochates/hirsutum); (*S. peruvianum*); LA1028 (*S. schiewlskii*); LA1322 (*S. parviflorum*); and LA1204 (*S. lycopersicum var cerasiforme*).

RESULTS

BTH Treatment Results in Increased PLCP Activity in the Tomato Apoplast

To investigate secreted PLCP activities during the defense response of tomato, we isolated apoplastic fluids (AFs) 5 d after treating tomato plants with water or with the salicylic acid (SA) analog BTH. The AF of BTH-treated tomato contains the expected PR proteins, demonstrating that the BTH treatment was successful (Figure 1A). We employed protease activity profiling as a tool to display activities of PLCPs in the AFs (38). Protease activity profiling is based on the use of a biotinylated, broad-range PLCP inhibitor E-64, which reacts with the catalytic Cys residue of the protease in an activity-dependent manner (25). E-64 lacks specificity-determining binding groups and is often used for diagnostic purposes since it is reactive to the entire range of PLCPs. Although the readout does not contain information on substrate specificity and conversion rates, signals represent the availability and abundance of active sites of PLCPs. We previously used protease activity profiling with biotinylated E-64 (called DCG-04) to show that AVR2 inhibits RCR3 and that EPIC2B inhibits PIP1 (19, 21).

In this work, protease activity profiling with DCG-04 revealed three biotinylated signals in AF of water-treated tomato plants: a strong signal at 25 kD, a weaker signal at 30 kD, and a weak signal at 35 kD (Figure 1B, lane 5). All three of these signals can be competed by adding an excess of E-64 during labeling (Figure 1B, lane 6). These three signals are upregulated in AF of BTH treated tomato plants (Figure 1B, lane 7). Quantification of these signals from films showed that the overall upregulation of signal intensity was 1.93-fold (60.38) (n ¼ 7). The differential activity is less pronounced in the total extracts (TE), suggesting that most of the differential PLCP activities upon BTH treatment reside in the apoplast and not inside the cells (Figure 1B, lanes 1 and 3). The difference in PLCP activities upon BTH treatment is stronger in AFs isolated from younger leaves of BTH-treated plants (Figure 1C) when compared with AFs isolated

from all the leaves (Figure 1B). This is explained by a lower PLCP activity in AFs in young leaves when compared with old leaves from the same non treated tomato plant (Figure 1D). Thus, tomato plants create a proteolytic apoplast by secreting active PLCPs upon BTH treatment, especially in young leaves.

Tomato Apoplast Contains Activities of Different PLCPs

To identify the PLCPs in the tomato apoplast, AFs were labeled with DCG-04 and biotinylated proteins were purified, separated on two-dimensional (2D) gels, and identified by tandem mass spectrometry (MS). The 28 protein spots on the 2D gel represent five different PLCPs called C14, PIP1, CYP3, ALP, and CatB1 (Figure 2A) (discussed below). The experimental pI and molecular weight of the identified proteases roughly coincide with that predicted for the mature proteins (Figures 2A and 3A). The horizontal rows of spots for most of the proteases probably come from different isoforms that exist in planta or are generated during extraction. Peptides were found only for the protease domains of these proteins, consistent with the prediction that the autoinhibitory prodomain should be cleaved off to activate the protease (Figure 3A). Peptides containing the active site Cys were missing from the spectra since these peptides are modified by DCG-04.

To confirm the identity of these apoplastic PLCPs independently, we infiltrated tomato plants with DCG-04 and isolated AF after in situ labeling. Since DCG-04 is not membrane permeable (39), this approach ensures that only apoplastic proteases are labeled and not proteases that might leak from cells during AF isolation. Biotinylated proteins were purified and analyzed on a one-dimensional (1D) protein gel (Figure 2B, left). Analysis of the proteins by tandem MS revealed the same five proteases as were found by 2D gel analysis (Figure 2B, right). An additional cathepsin B–like protease (CatB2) was identified from the 30-kD region. This protease has a calculated pI of 7.2 and was therefore missed in the 2D analysis. PIP1 was identified in every band, probably because this protein is highly abundant and contaminated the other signals during excision of the bands.

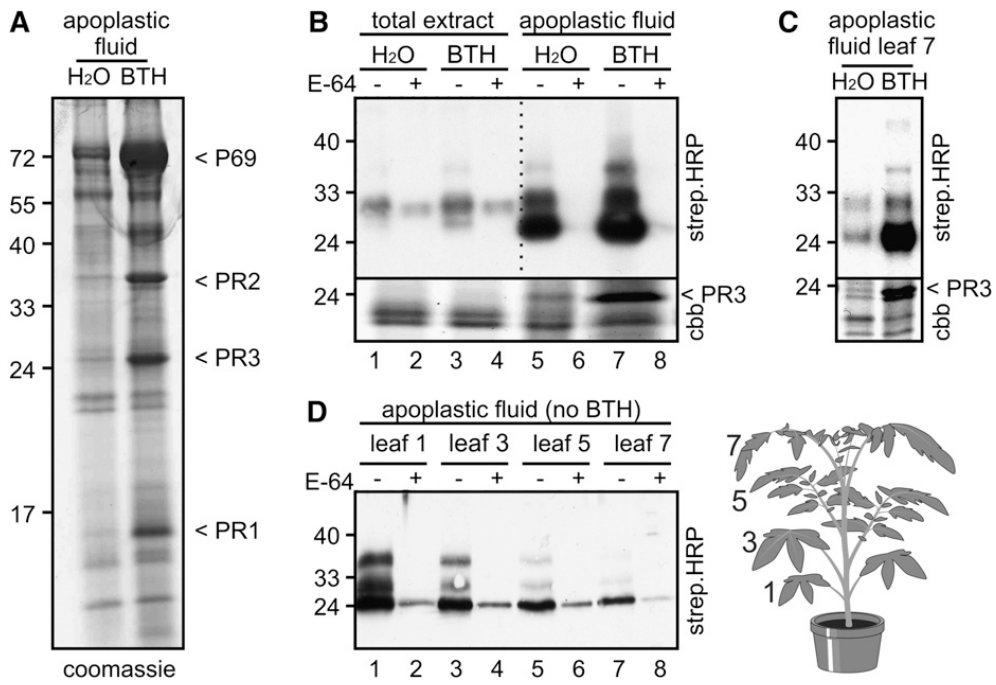


Figure 1. BTH-Induced Protease Activities in the Tomato Apoplast.

(A) Apoplastic PR accumulation in BTH-treated tomato plants. Five-week-old tomato plants were treated with water or BTH. AFs were isolated after 5 d, and equal volumes were separated on a 17% protein gel to visualize PR protein accumulation in the AF.

(B) Protease activity profiling on equal volumes of total extract and AF of combined leaves from water- or BTH-treated tomato plants. Extracts were labeled at pH 5.0 with DCG-04 in the absence or presence of an excess of E-64. Proteins were separated on 12% protein gels, and biotinylated proteins were detected on a protein blot with streptavidin HRP. Total extracts corresponds to;0.17 cm^2 leaf area/lane; AF corresponds to;7.4 cm^2 leaf area/ lane. A representative of seven independent experiments is shown. An enlarged Coomassie blue–stained gel (cbb) shows equal loading and the accumulation of PR3 proteins.

(C) Protease activity profiling of AFs from young leaves of water- and BTH-treated tomato plants. A Coomassie blue–stained gel (cbb) shows equal loading and the accumulation of PR3 proteins.

(D) Protease activity profiling on AFs isolated from leaves of different age of the same water-treated plant (indicated in illustration on the right). E-64 competition was incomplete, resulting in some remaining 25-kD signal.

Importantly, RCR3 was missing from both 1D and 2D analyses. The pI/molecular weight of RCR3 is identical to that of PIP1, which indicates that the abundant PIP1 probably hides the presence of RCR3 on 1D and 2D gels and therefore interferes with its detection by MS. To investigate the contribution of RCR3 to the overall activity profile, we labeled AF from wild-type and rcr3-3 tomato plants. The rcr3-3 mutation results in a premature stop codon that prevents RCR3 protein accumulation (15). The apoplastic PLCP activity profile of the rcr3-3 mutant plants is identical to that of wild-type plants, indicating that the amount of active RCR3 is minor when compared with PIP1 and other apoplastic PLCPs (Figure 2C).

Secreted Tomato PLCPs Belong to Four Different Classes

A phylogenetic tree constructed with 145 different plant PLCPs (31), including the seven identified tomato PLCPs, shows that the seven tomato PLCPs (Figure 3A) belong to four different subfamilies (Figure 3B). C14/TDI-65/CYP1/SENU2 belongs to subfamily 1, which includes PLCPs that are encoded by stress-inducible (heat, cold, or senescence) genes (40, 41,). Characteristic for many PLCPs of subfamily 1 is that they carry a C-terminal granulin domain (59, 61). PIP1 and RCR3 are closely related and belong to subfamily 5 (Figure 3B). These two proteases were described in the introduction. CYP3 and ALP are two aleurain-like proteases of subfamily 7 that both carry the vacuolar targeting signal NPIR in their prodomain (Figure 3A) (42). CatB1 and CatB2 are two cathepsin-B like proteases of subfamily 8 (Figure 3B). The Cathepsin B–like protease of *N. benthamiana* (CathB) is secreted as an active protease, and gene silencing of CathB in *N. benthamiana* prevents HR induced by non host bacteria and gene-for-gene interactions (17). In summary, the seven identified apoplastic PLCPs represent four subfamilies that probably fulfill different biological functions.

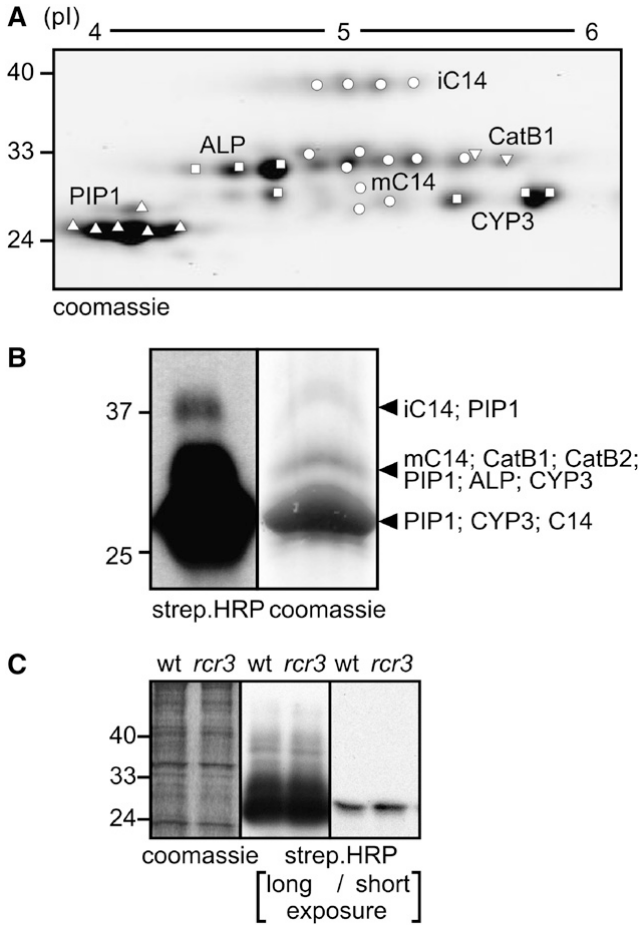


Figure 2. Identification of BTH-Induced Tomato Apoplastic Proteases.

(A) Identification of in vitro–labeled tomato PLCPs from AFs. AF of BTH treated leaves was labeled with DCG-04. After labeling, biotinylated proteins were purified and separated on a 2D gel. Proteins were isolated from Coomassie blue–stained spots/bands and analyzed by tandem MS.

(B) Identification of in vivo–labeled tomato PLCPs from AFs. Tomato leaflets (183) were vacuum infiltrated with 2 mM DCG-04 and labeled in situ. AF was isolated, and biotinylated proteins were purified and separated on protein gel. A small portion of the sample was analyzed on protein blots with streptavidin-HRP to visualize the biotinylated proteins (left panel). Proteins were isolated from Coomassie blue–stained spots/bands and analyzed by tandem MS.

(C) RCR3 is not a major protease in AFs of BTH-treated tomato plants. Cf2/Rcr3 and Cf2/rcr3-3 plants were treated with BTH for 5 d, and AF was isolated. AF was labeled with DCG-04, and biotinylated proteins were displayed on a protein blot using streptavidin-HRP. A representative of three independent experiments is shown.

Only PIP1 and RCR3 are induced by BTH treatment, we used fluorescent protease activity profiling using TMR-DCG- 04 (25) to investigate which of the PLCPs is differentially active in tomato AFs upon BTH treatment. Fluorescent protease activity profiling facilitates a direct, quantitative readout and reduces background caused, for example, by binding of streptavidin-horseradish peroxidase (HRP) to other proteins.

The signals detected by fluorescent protease activity profiling on 1D gels are similar to those observed using streptavidin-HRP (Figure 4A, top). However, in contrast with streptavidin-HRP detection, BTH treatment results in an upregulation of only the 25-kD signal (Figure 4A, top). The fluorescent signals were further analyzed on 2D gels. Comparison of the fluorescent 2D map (Figure 4A, bottom) with the Coomassie blue–stained 2D gel containing purified PLCPs (Figure 2A) indicates that we can detect the activities of PIP1, C14, CatB1, CYP3, and perhaps even ALP simultaneously using fluorescent protease activity profiling. Comparison of the fluorescent 2D maps of water- and BTH-treated tomato plants indicates that only the signal corresponding to PIP1 is upregulated upon BTH treatment (Figure 4A, bottom). The signals corresponding to the other PLCPs remain the same when compared with the standard.

These data indicate that PIP1 dominates the increased PLCP activity during BTH treatment. To investigate the transcriptional regulation of the protease encoding genes during BTH-induced defense responses, we performed quantitative real-time RT-PCR on BTH- and water treated tomato plants. This analysis shows that PIP1 and RCR3 transcripts are eightfold upregulated upon BTH treatment, whereas the other PLCP transcript levels are not significantly induced by BTH treatment (Figure 4B). The induction by BTH treatment is similar to those of PR genes (43), indicating that PIP1 and RCR3 gene expression is regulated through the SA pathway and that they can be considered as being PR genes.

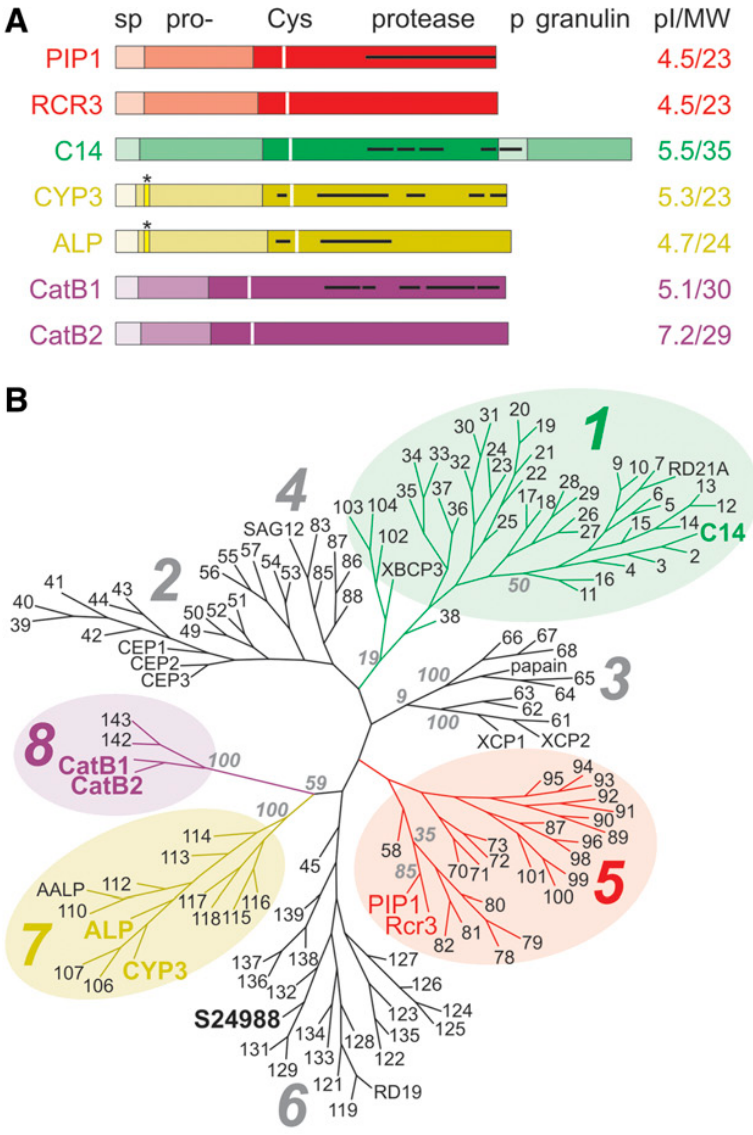
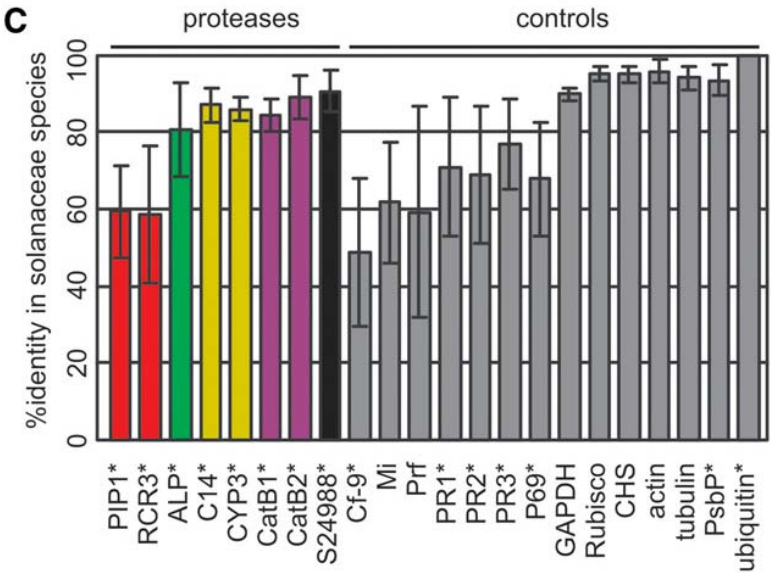


Figure 3. In Silico analysis of Identified Secreted Tomato Proteases.

(A) Domains encoded by the open reading frames (ORFs) of the identified secreted tomato proteases. The peptides that are reproducibly found in MS analysis are indicated with black bars. sp, signal peptide; pro-, autoinhibitory prodomain; Cys, catalytic Cys residue; protease, mature protease domain; p, Pro-rich domain; granulin, granulin-like domain; *, vacuolar targeting signal NPIR; pI/MW, isoelectric point and molecular weight of the protease after removal of the signal peptide and prodomain. The color coding represents the similarity-colored subfamilies in (B).

(B) Unrooted phylogenetic tree of 145 plant papain-like Cys proteases showing the diversity of the identified tomato proteases (colored names). The classes are subdivided into eight subfamilies (31). Accession codes of other plant PLCPs (numbered) are in Set 1. S24988 is a tomato protease of group 6 that was not identified in AFs in this study. Bootstrap values for critical nodes are indicated in gray bold italics.



(C) Conservation of proteases and control proteins within solanaceous species. tBLASTp searches of the tomato proteins to translated EST libraries of tobacco, pepper, *N. benthamiana*, potato, and petunia resulted in the identification of close homologs from each of these solanaceous species. The amino acid identity over the protease domain was scored and combined for the different species. Resistance proteins, PR proteins, and various household proteins were used as controls (gray; see Methods). Asterisk indicates mature proteins used for tBLASTp analysis. Error bars represent SD

PIP1 and RCR3 Are Less Conserved Than the Other Proteases

To investigate the conservation of PLCP sequences in other solanaceous species, we searched for homologous genes in The Institute for Genomic Research (TIGR) database. tBLASTp searches of the seven tomato proteases to translated EST libraries of tobacco, pepper (*Capsicum annuum*), *N. benthamiana*, potato (*Solanum tuberosum*), and petunia (*Petunia hybrida*) resulted in the identification of close homologs from each of these solanaceous species. The amino acid identity over the protease domain was scored and combined for the different species.

This in silico analysis indicates that the proteases fall into two groups: a conserved group, consisting of C14, CYP3, ALP, CatB1, and CatB2 proteins; and a diverse group, containing PIP1 and RCR3 (Figure 3C). The degree of sequence identity of the conserved proteases among the examined species is similar to housekeeping proteins with known,

conserved functions (Figure 3C, right columns). By contrast, the degree of identity of diverse proteases PIP1 and RCR3 is as low as that of resistance proteins (Cf9, Mi, and Prf) or PR proteins (PR1, PR2, PR3, or P69A) (Figure 3C, middle columns). Although this analysis is limited by the absence of complete genome sequences for the examined species and by the size of the EST libraries, the relative expression levels of each gene, and the tissues from which these EST libraries were generated, these data suggest that the protease domains of PIP1 and RCR3 are more diverse and rapidly evolving within solanaceous species than those of the other PLCPs.

PIP1 and RCR3 Are Subject to Diversifying Selection

The observation that PIP1 and RCR3 are probably more diverse inspired us to investigate if these genes are subject to diversifying selection. We therefore sequenced the region encoding the protease domain of eight wild tomato relatives: *S. cheesmanniae*, *S. pimpinellifolium, S. chilense*, *S. pennellii*, *S. habrochates* (*hirsutum*), *S. peruvianum*, *S. schiewlskii*, and *S. parviflorum*. Sequences of theses alleles were validated (see Methods) and found to be 98% identical to the reported tomato sequences.

Amino acids encoded by the polymorphic codons of all the protease domains are shown in Figure 5A. The protease-coding part of each gene contains; 20 variant nucleotides, except for RCR3, which has 41 variant nucleotides (Figure 5B). Some of the variant nucleotides are shared among different species, indicating that part of the variation predates speciation (Figure 5A). Most of the polymorphic nucleotides, however, are allele specific. The consequence of these variant residues at amino acid level is striking. Variant codons hardly change the encoded amino acids in C14, CYP3, ALP, CatB1, and CatB2 (Figure 5A, bottom, white and gray residues).

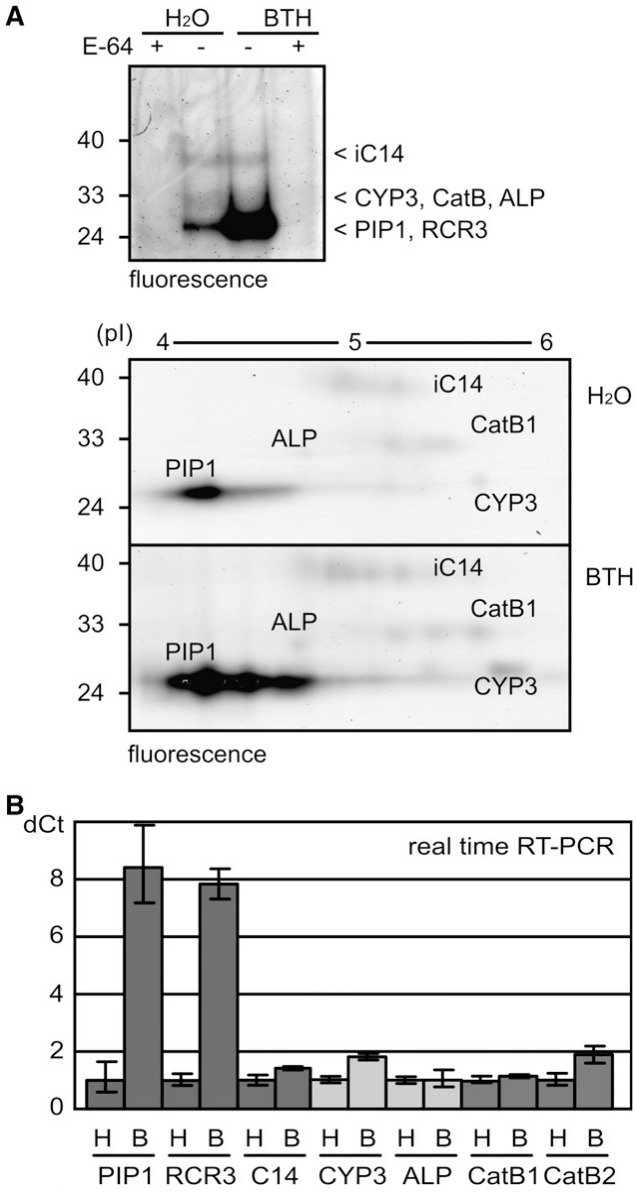


Figure 4. Induction of PIP1 and RCR3 upon BTH Treatment by Fluorescent Protease Activity Profiling and Real-Time RT-PCR.

(A) Fluorescent protease activity profiling of AFs. Plants were treated with BTH for 5 d, and AF was isolated. Equal volumes of AF were used for protease activity profiling with TMR-DCG-04. Proteins were separated by 1D (top) and 2D (bottom) gel electrophoresis, and fluorescent proteins were detected by fluorescent protease activity profiling. Please note that only signals coinciding with PIP1 are significantly stronger from BTH treated plants. A representative of three independent experiments is shown.

(B) Induction of transcript levels by BTH treatment. Tomato leaves were harvested at 5 d after water (H) or BTH (B) treatment. Quantitative real time RT-PCR was performed using gene-specific primers. The difference in threshold cycles (dCt) between the protease transcript and ubiquitin transcripts was calculated from three independent samples. Error bars represent SD. A representative of five independent biological experiments is shown.

By contrast, nearly all variant codons of PIP1 and RCR3 cause nonsimilar amino acid substitutions (Figure 5A, bottom, red residues). The ratio between nonsimilar and similar amino acids indicates that C14, CYP3, ALP, CatB1, and CatB2 are under conservative

selection, whereas PIP1 and RCR3 are under diversifying selection (Figure 5C). These data are consistent with the indication from tBLASTp searches that PIP1 and RCR3 proteases are less conserved in solanaceous species than the other PLCPs (Figure 3C). Taken together, these observations demonstrate that PIP1 and RCR3 are under diversifying selection, possibly to adapt to diversifying substrates or inhibitors, whereas the other proteases are under conservative selection.

Diversified Residues in PIP1 and RCR3 Reside around the substrate Binding Groove

To investigate where within the protease structure the variation of PIP1 and RCR3 resides, we generated structural models for PIP1 and RCR3 to indicate the position of the variant amino acids (Figure 5D). Like with all PLCPs, the PIP1 and RCR3 models consists of two lobes with a substrate binding groove in between (34). The active site Cys is in the middle of this groove. All of the nonsimilar residue variance of PIP1 and RCR3 are on the surface of the protein (red in Figure 5D). By contrast, similar residue variance is mostly located inside the protein structure (blue in Figure 5D). Five of the variant amino acids of PIP1 and RCR3 coincide at a similar location, but the other variant residues are unique to either PIP1 or RCR3. These residues are nearly all on the front side, around the substrate binding groove, but not in the groove itself. The location of the variant residues could coincide with a binding region for substrates or inhibitors.

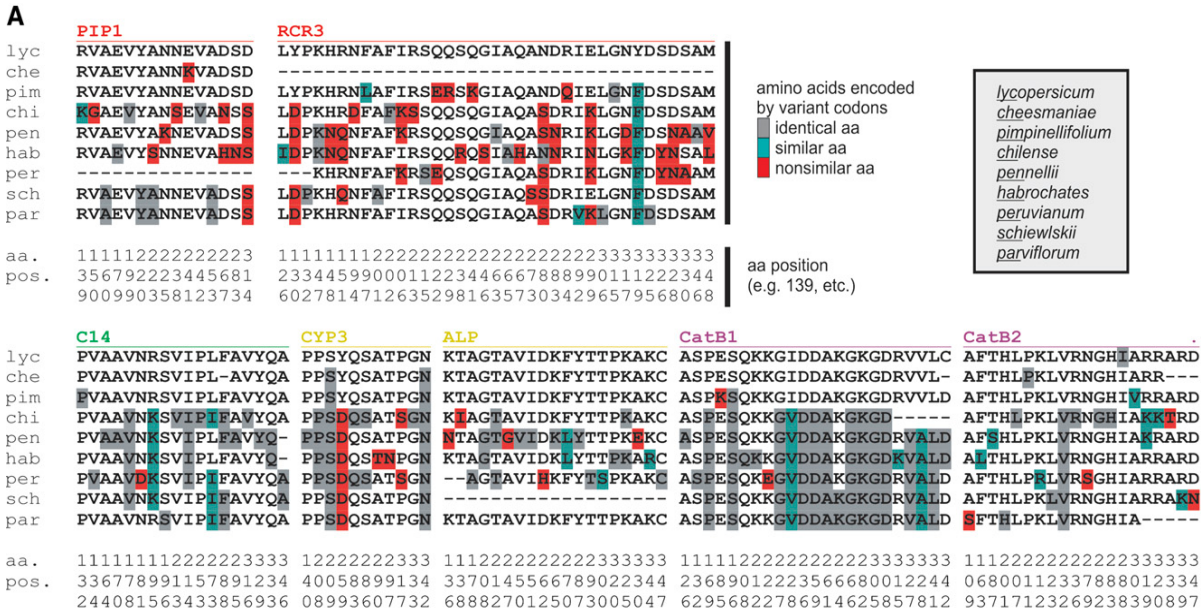
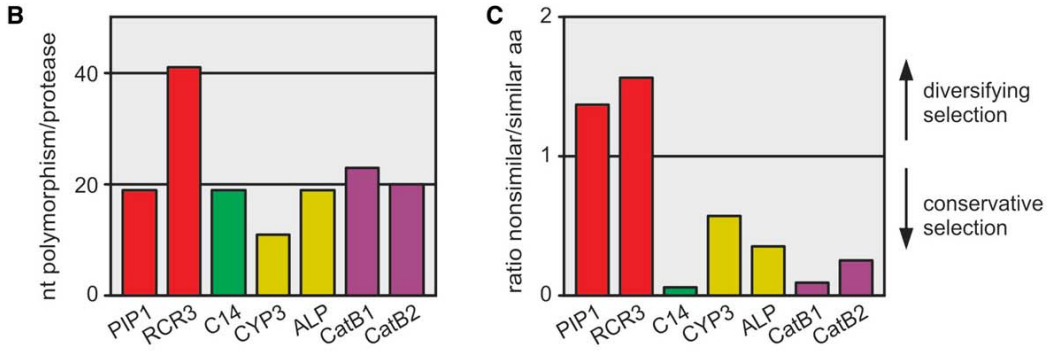


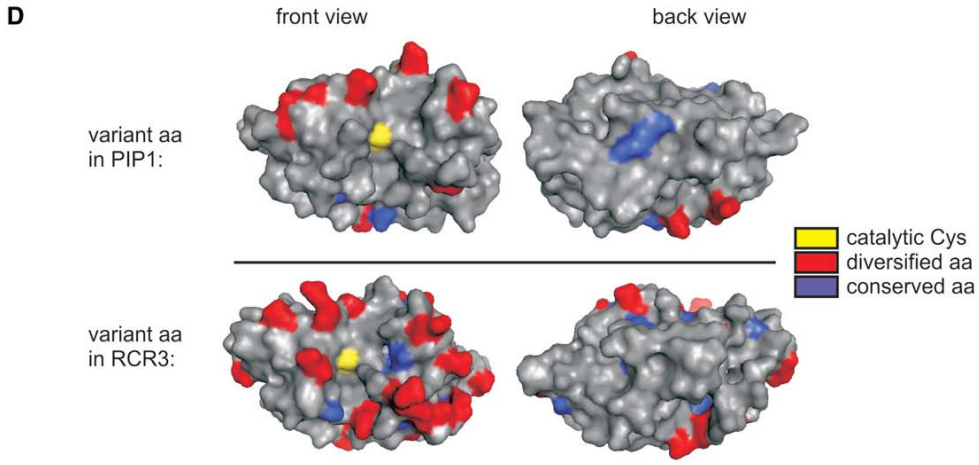
Figure 5. Sequence Analyses of Proteases from Tomato Relatives.

(A) Summary of amino acids encoded by variant codons in the protease domains of C14, PIP1, RCR3, CYP3, ALP, CatB1, and CatB2 alleles sequenced from various wild tomato relatives (indicated top right). Amino acids encoded by the variant codons are summarized by leaving out the amino acids of nonvariant codons from the protein alignment. Amino acids encoded by codons different from the *S. lycopersicum* (lyc) allele are indicated with gray, blue, and red residues if they are identical, similar, or nonsimilar, respectively, compared with the lyc sequence. Dashes indicate missing sequence information. RCR3 of *S. cheesmanniae* is not shown since it contained a premature stop codon and could be amplified from genomic DNA and not from cDNA .



(B) Number of single nucleotide (nt) polymorphisms per protease.

(C) Ratio of nonsimilar/similar amino acid (aa) substitutions calculated from (A). PIP1 and RCR3 are under diversifying selection; the other proteases are under conservative selection.



(D) Position of variant residues in structural models of PIP1 and RCR3. Positions with nonsimilar variance and similar variance are indicated in red and blue, respectively.

AVR2 Targets Specific PLCPs in the Tomato Apoplast

Having identified the PLCPs that pathogens encounter when they invade the tomato apoplast, we next investigated which PLCPs are inhibited by AVR2 of the fungus *C. fulvum* (19). AVR2 was produced as an N-terminally FLAG tagged protein in *Escherichia coli.* This FLAG-AVR2 triggers hypersensitive cell death when injected into tomato plants carrying the Cf-2 resistance gene, demonstrating that AVR2 is folded properly (Figure 6A). To investigate the inhibition of PLCPs by AVR2 in crude AFs of BTH-treated tomato plants, we preincubated these AFs for 30 min with AVR2 and then added DCG-04 to label the remaining, noninhibited proteases. These assays revealed that AVR2 prevents the labeling of most of the 25-kD proteases, whereas the other signals remain unaltered (Figure 6B).

To determine the selectivity of AVR2 inhibition in more detail, we performed fluorescent protease activity profiling using TMR-DCG-04 on AF of BTH-treated tomato plants with and without preincubation with AVR2. The 2D images were normalized using the standard and compared. Comparing the 2D fluorescent maps in the absence or presence of AVR2 shows that fluorescent labeling of C14, CYP3, and CatB1 is unchanged, whereas labeling of PIP1 is significantly reduced in the presence of AVR2 (Figure 6C). These data indicate that AVR2 selectively targets PIP1 and does not inhibit other C14, CYP3, or CatB1 in the apoplast.

AVR2 Inhibits PIP1 and RCR3

To investigate the inhibition of each protease separately, we overexpressed each of the seven tomato proteases in planta through agroinfiltration of *N. benthamiana* (26, 28). Protease activity profiling with DCG- 04 on total extracts or AFs from these agroinfiltrated leaves revealed strong additional biotinylated signals in addition to weaker signals of the endogenous proteases of *N. benthamiana*. The intensities of these signals were used to dilute the extracts to contain comparable concentrations of active proteases. The sizes of the signals are consistent with the sizes of these proteases observed in tomato AFs and the sizes predicted for the mature protease domains (Figures 2A and 2B).

This transient expression system allowed us to monitor activities of in planta–produced tomato proteases separately.

To monitor the inhibition by pathogen effector proteins, we preincubated the protease-containing extracts with AVR2 and then added DCG-04 to label the noninhibited proteases. This revealed that AVR2 inhibits labeling of PIP1 and RCR3 but not C14, CYP3, ALP, or CatB1 (Figure 7A). These inhibition data are consistent with the data observed for the tomato apoplastic PLCPs (Figure 6C). The inhibition assays described above were done at 66 nM inhibitor concentrations. To investigate inhibition at lower inhibitor concentrations, PIP1 and RCR3 were incubated with or without various AVR2 concentrations and then incubated with DCG-04. This revealed that AVR2 inhibits both PIP1 and RCR3 at concentrations above 7.2 nM (Figure 7B), indicating that AVR2 inhibits both RCR3 and PIP1 to a similar level. However, precise comparisons of inhibition strengths were not possible with these assays. The inhibition experiments were so far performed at the pH of the apoplast (pH 5). To investigate if the inhibition could also take place under a wider pH range, inhibition assays were performed at pH 4 to 7. Although PIP1 and RCR3 are similarly active at all tested pH, inhibition by AVR2 only occurs at pH 4 to 5.5 (Figure 7C). This shows that inhibition of both PIP1 and RCR3 by AVR2 is a pH-dependent process that occurs at apoplastic pH.

Coimmunoprecipitation experiments were performed to investigate if inhibition is caused by a physical interaction between AVR2 and PIP1. C-terminally His-tagged PIP1 (21) was produced by agroinfiltration and immobilized on nickel nitrilotriacetic acid agarose (Ni-NTA) columns. These PIP1- coated beads were incubated with FLAG-AVR2, washed, and eluted, and the eluate was used for protein blots using anti-PIP1 and anti-FLAG antibodies. This revealed that AVR2 is in the eluate of immobilized PIP1 columns (Figure 7D, lane 2). The interaction is prevented by preincubation of PIP1-His with an excess E-64 before immobilization (Figure 7D, lane 3), indicating that the interaction is at the active site. Taken together, these data show that AVR2 physically interacts with PIP1, similar to what was demonstrated previously for RCR3 (19).

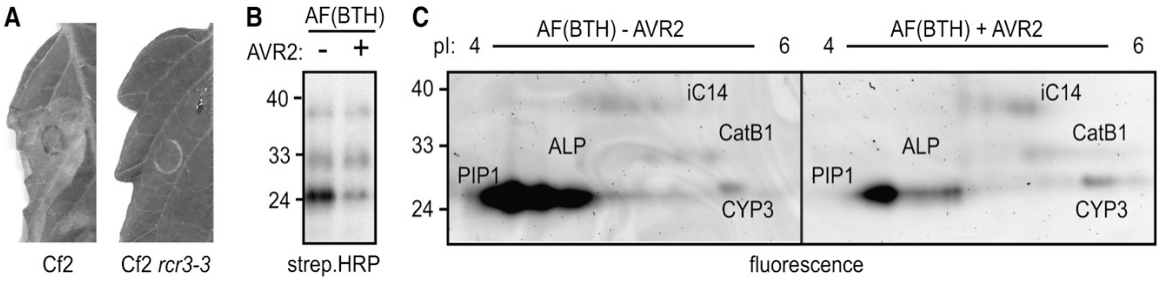


Figure 6. Production of AVR2 and Inhibition in AF.

(A) Recombinant AVR2 triggers hypersensitive cell death in tomato leaves of MM-Cf2 plants. Recombinant AVR2 was injected into leaves of MM-Cf2 and MM-Cf2/rcr3-3 plants. Picture was taken 4 d after injection.

(B) Inhibition of PLCPs in AFs by AVR2. AF of BTH-treated tomato plants were preincubated with 66 nM AVR2 before adding DCG-04 to label the remaining noninhibited proteases. Biotinylated proteins were detected using streptavidin-HRP. A representative of three independent experiments is shown.

(C) Fluorescent protease activity profiling of PLCPs in AFs in presence or absence of AVR2. AFs of BTH-treated tomato plants were preincubated with 83 nM AVR2 before adding TMR-DCG-04 to label the remaining noninhibited proteases. Proteins were separated by 2D gel electrophoresis, and fluorescent proteins were visualized. A representative of three independent experiments is shown.

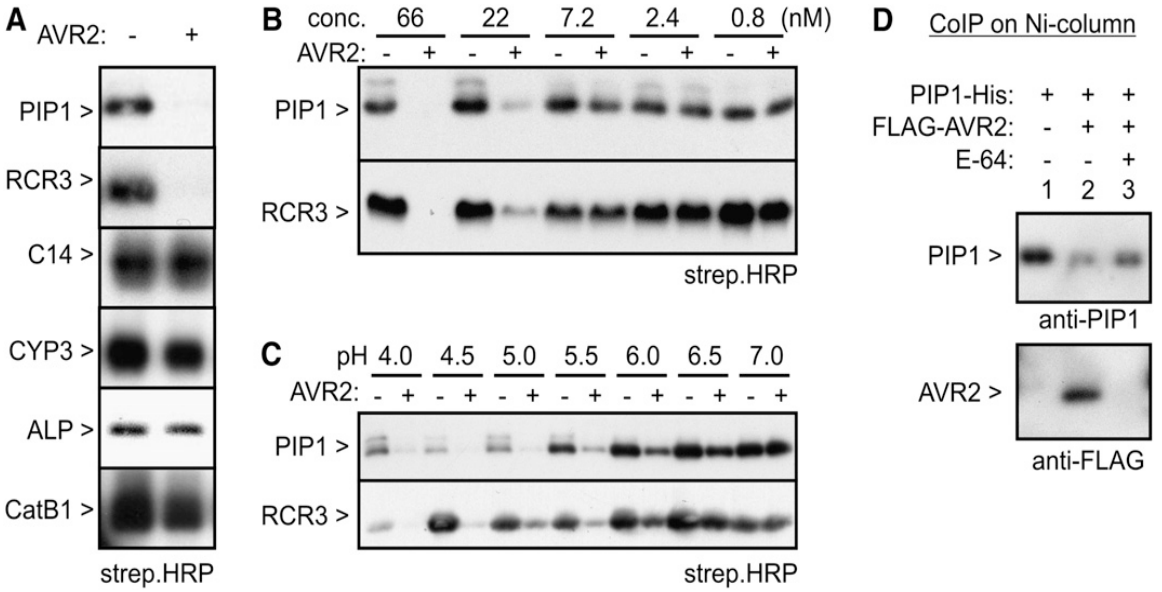


Figure 7. Inhibition of Agroinfiltrated Tomato Proteases by AVR2.

(A) Extracts from agroinfiltrated *N. benthamiana* leaves overexpressing different tomato proteases (indicated on the left) were preincubated for 30 min with 66 nM AVR2. DCG-04 was added after preincubation to label the noninhibited proteases. Biotinylated proteins were visualized on protein blots using streptavidin-HRP. Representatives of at least three independent experiments are shown.

(B) Concentration dependency of inhibition by AVR2. Protease-containing extracts were incubated with different AVR2 concentrations for 30min. DCG- 04 was added after preincubation to label the noninhibited proteases. Biotinylated proteins were visualized on protein blots using streptavidin-HRP. A representative of three independent experiments is shown.

(C) pH dependency of inhibition by AVR2. Extracts from agroinfiltrated leaves overexpressing the proteases were preincubated for 30 min at different pH with or without 66 nM AVR2. DCG-04 was added after preincubation to label the noninhibited proteases. Biotinylated proteins were visualized on protein blots using streptavidin-HRP. A representative of three independent experiments is shown.

(D) AVR2 physically interacts with PIP1. Extracts from agroinfiltrated leaves expressing His-tagged PIP1 were preincubated with or without an excess E-64 and incubated with FLAG-AVR2 at pH 5. Protein complexes containing PIP1-His were purified using Ni-NTA columns and analyzed using anti-PIP and anti-FLAG antibodies. A representative of three independent experiments is shown.

A Naturally Occurring Variant Residue in RCR3 Affects Inhibition by AVR2

To investigate if variant residues in RCR3 can affect its inhibition by AVR2, we selected variant residues in RCR3 that are close to the active site. Four of the variant residues in RCR3 (orange in Figure 8A) were not tested since these variant residues are present in *RCR3*^{pim}, and this protein is inhibited by AVR2 to a similar extent as *RCR3*^{lyc} (19). Four other variant residues (H148N, R151Q, N194D, and Q267H; red in Figure 8A) were selected and introduced into RCR3 using site-directed mutagenesis. In addition, one double mutant was generated (H148N-R151Q). Mutant RCR3 proteins were produced by agroinfiltration and subjected to protein blot analysis with anti-RCR3 antibodies and protease activity profiling with DCG-04. All mutant RCR3 proteins accumulate to similar levels and react with DCG-04, indicating that these proteins are stable, active proteases (Figure 8B). Preincubation of the (mutant) RCR3 proteins with AVR2 showed that mutants H148N, R151Q, and N194D and the double mutant H148N-R151Q can be inhibited by AVR2 to a similar extent as wild-type RCR3 (Figure 8B). By contrast, the N194D mutant RCR3 protein is less sensitive to AVR2 inhibition, even though it can be inhibited by E-64 (Figure 8B). These data demonstrate that the variance of a single amino acid can affect the inhibition by AVR2. However, in addition to N194D variance, the *S. chilense* RCR3 isoform carries another five variant residues. The effect of these residues on AVR2 inhibition will be investigated in future studies.

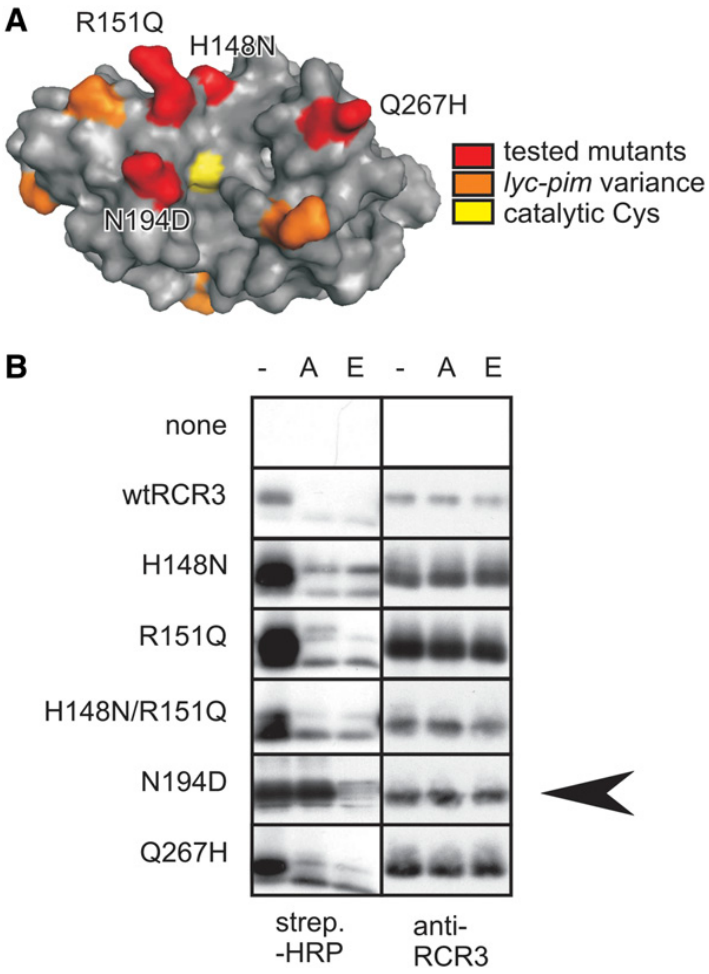


Figure 8. The Naturally Occurring N194D Mutation in RCR3 Affects Inhibition by AVR2.

(A) Position of variant residues in a structural model of RCR3. Differences between the *RCR3*^{lyc} and *RCR3*^{pim} alleles (orange) were not affecting AVR2 inhibition (19). Four variant residues (red) were selected for targeted mutagenesis.

(B) The RCR3^{N194D} mutant is less sensitive for AVR2 inhibition. Extracts from agroinfiltrated leaves expressing (mutant) RCR3 proteins were preincubated at pH 5.0 with or without 65 nM AVR2 (A) or 20 mM E-64 (E). DCG-04 was added to label the remaining noninhibited proteases, and biotinylated proteins were detected on protein blots using streptavidin- HRP. The experimental differences in accumulation of RCR3 is shown using RCR3 antibodies. A representative of four independent assays is shown.

DISCUSSION

This work demonstrates that the apoplast of BTH-treated tomato contains at least seven active PLCPs that fall into four different subfamilies based on their homology, structural features, and putative cellular functions. PIP1 and RCR3 are the only proteases that are strongly transcriptionally upregulated upon BTH treatment.

In tomato species, these proteases are rapidly evolving and are under strong diversifying selection, resulting in variant residues around the substrate binding groove of the protease. The proteases of other subfamilies are not induced by BTH treatment and are under conservative selection. The *Cladosporium* AVR2 effector protein specifically targets the defense related proteases PIP1 and RCR3, and one of the naturally occurring variant residues in RCR3 affects AVR2 inhibition. The described proteases are probably only a subset of the proteases that are present in the apoplast of BTH-treated tomato plants. P69B, for example, is another predominant PR protein and an active subtilisin-like Ser protease (44, Figure 1A). P69B is inhibited by Kazal-like EPI1 and EPI10 proteins of *P. infestans* (10, 11). We focused on papain-like proteases by applying protease activity profiling using the broad-range probe DCG-04, which reacts with PLCPs with limited selectively (15). The seven PLCPs that we identified are probably not the full set of secreted PLCPs but certainly comprise the majority of active PLCPs in the tomato apoplast. It is likely that extracellular pathogens will encounter these secreted PLCPs during infection.

We identified proteases in the apoplast by careful analysis that excludes leakage from internal sources. However, we did find aleurain-like proteases that carry vacuolar targeting signals. In fact, many of the apoplastic proteases we identified have Arabidopsis orthologs that were identified in the vegetative vacuole proteome (45). These Arabidopsis vacuolar PLCPs contain representatives of group 1 (RD21A and XBCP3), group 6 (number 132), group 7 (aleurain-like, AALP), and group 8 (CatB like, numbers 142 and 143) (indicated in Figure 3B) (44). In addition, AALP is used as a vacuolar marker protein, and RD21 was found in stress-induced protease storage vesicles (46, 47, 48).

The data presented here indicate that pools of some of these proteases are also partly secreted. This partial secretion probably occurs with RD21-like (group 1), aleurain-like (group 7), and cathepsin B-like (group 8) proteases and may result from a default secretion pathway that is followed if sorting by the subcellular targeting machinery is reduced. However, it remains to be investigated how these proteases are secreted and why. Transcript analysis showed that only PIP1 and RCR3 are induced by BTH treatment. This indicates that these genes are under control of the SA signaling pathway and that PIP1 and RCR3 belong to the class of PR proteins that accumulate during the basal immune response.

The other PLCP genes are not induced by BTH treatment. Consistent with BTH induction, RCR3 is also induced during infection with C. fulvum (18), and PIP1 is upregulated during infection with *Pseudomonas syringae* and *P. infestans* (21, 48). However, since SA signaling is not the only pathway that is regulated during infection, the final expression levels of PLCPs will strongly depend on the time point, cell type, and pathogen. Infection of tomato with *P. syringae*, for example, induces the expression of C14 but not CatB2 (49). Furthermore, infection of potato plants with avirulent *P. infestans* results in upregulation of both C14 and CatB2 (50, 51). Thus, although the other PLCP genes are not induced by the SA analog BTH, their expression can be induced during infection, probably through other pathways.

However, a distinction based on BTH induction appears to be relevant given the fact that only BTH-induced proteases are under diversifying selection and targeted by AVR2. Allele sequencing revealed that PIP1 and RCR3 are under diversifying selection, whereas the other PLCPs are under conservative selection. This indicates that BTH-induced proteases are under diversifying selection to coevolve with a changing environment imposed by pathogens. The variant residues within the PIP1 and RCR3 proteins are at the surface, around the substrate binding groove, but not within. This region could coincide with a binding region for pathogen-derived substrates and inhibitors.

Diversifying selection at interaction surfaces between enzyme and their inhibitors are common in plant–pathogen interactions (reviewed in Misas-Villamil and Van der Hoorn, 2008). Similar diversifying residues at the interaction surfaces have been reported for xylanase, polygalacturonase, and glucanase inhibiting proteins and their target enzymes (52, 53, 54, 55). That AVR2 selectively targets PIP1 and RCR3 among the apoplastic PLCPs was demonstrated by fluorescent protease activity profiling and by studying PLCPs separately through agroinfiltration. The fact that only BTH-induced PLCPs are inhibited is an interesting observation, although it cannot be excluded that *C. fulvum* secretes other proteins that target other PLCPs. Although these data suggest that RCR3 and PIP1 play a role in defense in the absence of Cf-2, this remains to be demonstrated. Conversely, a contribution of AVR2 to virulence remains to be shown.

The hypothesis that some of the variant residues in PIP1 and RCR3 might affect inhibition by pathogen-derived inhibitors was tested with AVR2 on a series of RCR3 mutants. Although weak inhibition cannot be measured with these assays, we showed that one mutant (RCR3N194D) is less sensitive to inhibition by AVR2. This variant N194D residue is close to the active site of RCR3, which indicates that this surface is important for binding AVR2. The other tested residues and those present in RCR3pim (orange and red in Figure 8A) do not alter AVR2 inhibition. These experiments indicate that single variant residues can affect inhibition by AVR2 and other pathogen-derived inhibitors, such as the cystatin-like inhibitors EPIC1 and EPIC2B of *P. infestans*, which also inhibit PIP1 and RCR3 (21; S. Kamoun personal communication). The contribution of single variant residues in the context of other variance in the different RCR3 alleles remains to be investigated since some variant residues may compensate others. Diversifying selection on PLCPs and their inhibitors might be common to plant–pathogen/pest interactions since PLCPs are frequently employed on the battlefield in these interactions, for example, in the apoplast, plant cytoplasm, and herbivore gut (56).

RCR3 has accumulated twice as many variant residues in comparison to PIP1. This is an interesting observation since it probably results from the adaptation of RCR3 to Cf resistance proteins. The *RCR3*^{lyc} allele triggers necrosis in the presence of the Cf-2

resistance gene, explaining why the Cf2 resistance locus was introgressed from *S. pimpinellifolium* together with the RCR3pim allele (18). The RCR3esc and RCR3pim proteases differ in four nonsimilar residues in the protease domain, located around the active site (Figure 8A, orange residues). The adaptation of RCR3 to Cf proteins is probably essential to prevent self-recognition. A similar accidental self recognition is probably also the basis of HR triggered by expression of recombinant tomato Cf proteins and Cf proteins from *S. peruvianum* in tobacco (57). The responses of these auto-activators differed between different tobacco species, suggesting that an endogenous polymorphic factor contributes to the degree of autoimmune responses.

These data are consistent with the model that non-self-recognition in plants causes necrosis that is often observed in hybrids (58). In Arabidopsis, for example, an NB-LRR resistance gene was found to trigger immune responses in hybrids, when combined with a specific allele at a second locus (58). These observations suggest that some of the variation in RCR3 is to adjust its interaction to Cf resistance proteins preventing accidental necrosis induced by Cf-2 in the absence of the pathogen.

These genetic data indicate that non-self RCR3 proteins could trigger accidental immune responses by an inadvertent physical interaction with Cf-2 proteins. Cf proteins consist predominantly of Leu-rich repeats that are folded as a horseshoe, having the inner concave surface with solvent-exposed residues available for specific protein–protein interactions. A direct interaction between Cf-2 and RCR3 remains to be shown, but docking studies of RCR3 with Cf proteins (27) suggest that RCR3 could fit inside the Leu-rich repeat curvature of Cf-2 with the variant amino acids of RCR3 interacting with the solvent-exposed residues of Cf-2. In this model, AVR2 could act as a molecular glue to enhance interactions of RCR3 with Cf-2.

PIP1 is the most dominant PLCP in the apoplast of BTH treated tomato. RCR3 accumulates to a much lower level and is even hard to detect, partly because its presence is over shadowed by PIP1. This observation was further substantiated with the fact that there is no significant difference in overall protease activity profiles between BTH-

treated wild-type plants and rcr3-3 mutants (Figure 3D). The additional adaptation of RCR3, and its low abundance compared with PIP1, suggests that RCR3 could be a decoy to trap the fungus into a recognition event, rather than a defensive protease itself. The real target of AVR2 would then be the abundant PIP1 protease. The approach taken in this study to identify (defense-related) secreted proteases and analyze their inhibition by pathogen-derived proteins has revealed interesting aspects of an apoplastic molecular battlefield between plants and pathogens. Given these observations, it seems likely that many pathogens secrete PLCP inhibitors during infection and that these inhibitors, the proteases, and their substrates are involved in a continuous coevolutionary battle. How coevolution shaped the apoplastic defense, how these defense-related proteases discriminate between self and non-self, and how RCR3 became part of a pathogen perception mechanism involving the Cf-2 resistance gene remain exciting questions to resolve.

Acknowledgements

We are grateful to Takayuki, Christian, Dr. Farnusch Kaschani, Twinkal, Raju and all other members of Plant Chemetics group (MPIZ, Cologne, Germany) for their valuable participation in the experiments. Especially Takayuki for sequencing experiments and sequence analysis part, Dr. Farnusch Kaschani for making phylogenetic tree and last , but not least Christian for 2D electrophoresis experiments.

REFERENCES

1. Jones, J.D.G., and Dangl, J.L. (2006). Nature 444: 323–329.
2. Kamoun, S., and Smart, C.D. (2005). Plant Dis. 89: 692–699.
3. Rivas, S., and Thomas, C.M. (2005). Annu. Rev. Phytopathol. 43: 395–436.
4. Thomma, B.P.H.J., Van Esse, H.P., Vrous, P.W., and De Wit, P.J.G.M. (2005) Mol. Plant Pathol. 6: 379–393.
5. Kombrink, E., Schro¨ der, M., and Halbrock, K. (1988) Proc. Natl. Acad. Sci. USA 85: 782–786.
6. Joosten, M.H.A.J., and De Wit, P.J.G.M. (1989) Plant Physiol. 89: 945–951.
7. Van Loon, L.C., Rep, M., and Pieterse, C.M.J. (2006) Annu. Rev. Phytopathol. 44: 135–162.
8. Ferreira, R.B., Monteiro, S., Freitas, R., Santos, C.N., Chen, Z., Batista, L.M., Duarte, J., Borges, A., and Teixeira, A.R. (2007). Mol. Plant Pathol. 8: 677–700.
9. Rose, J.K., Ham, K.S., Darvill, A.G., and Albersheim, P. (2002) Molecular cloning and characterisation of glucanase inhibitor proteins: coevolution of a counterdefense mechanism by plant pathogens. Plant Cell 14: 1329–1345.
10. Tian, M., Benetti, B., and Kamoun, S. (2005) Plant Physiol. 138: 1785–1793.
11. Tian, M., Huitema, E., Da Cunha, L., Torto-Alalibo, T., and Kamoun, S. (2004)J. Biol. Chem. 279: 26370–26377.
12. Van den Burg, H.A., Harrison, S.J., Joosten, M.H.A.J., Vervoort, J., and De Wit, P.J.G.M. (2006) Mol. Plant Microbe Interact. 19: 1420–1430.
13. Van Esse, H.P., Bolton, M.D., Stergiopoulos, I., De Wit, P.J.G.M., and Thomma, B.P.H.J. (2007) Mol. Plant Microbe Interact. 20: 1092–1101.
14. Misas-Villamil, J.C., and Van der Hoorn, R.A.L. (2008) Curr. Opin. Plant Biol. 11: in press.
15. Van der Hoorn, R.A.L., and Jones, J.D.G. (2004) Curr. Opin. Plant Biol. 7:400–407.
16. Xia, Y., Suzuki, H., Borevitz, J., Blount, J., Guo, Z., Patel, K., Dixon, R.A., and Lamb, C. (2004) EMBO J. 23: 980–988.
17. Gilroy, E.M., et al. (2007) Plant J. 52: 1–13.

18. Kruöger, J., Thomas, C.M., Golstein, C., Dixon, M.S., Smoker, M., Tang, S., Mulder, L., and Jones, J.D.G. (2002) Science 296: 744–747.
19. Rooney, H.C.E., Van't Klooster, J.W., van der Hoorn, R.A.L., Joosten, M.H.A.J., Jones, J.D., and De Wit, P.J.G.M. (2005) Science 308: 1783–1786.
20. Van der Hoorn, R.A.L., De Wit, P.J.G.M., and Joosten, M.H.A.J. (2002) Trends Plant Sci. 6: 67–71
21. Tian, M., Win, J., Van der Hoorn, R., Van der Knaap, E., and Kamoun, S. (2007) Plant Physiol. 143: 364–377.
22. Sambrook, J., Fritsch, E.F., and Maniatis, T.A. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).
23. Luderer, R., Takken, F.L.W., De Wit, P.J.G.M., and Joosten, M.H.A.J. (2002) Mol. Microbiol. 45: 875–884
24. Gobom, J., Schuerenberg, M., Mueller, M., Theiss, D., Lehrach, H., and Nordhoff, E. (2001) Anal. Chem. 73: 434–438.
25. Greenbaum, D., Baruch, A., Hayrapetian, L., Darula, Z., Burlingame, A., Medzihradszky, K.F., and Bogyo, M. (2002) Mol. Cell. Proteomics 1: 60–68.
26. Van der Hoorn, R.A.L., Laurent, F., Roth, R., and De Wit, P.J.G.M. (2000) Mol. Plant Microbe Interact. 13: 439–446.
27. Van der Hoorn, R.A.L., Wulff, B., Rivas, S., Durrant, M.C., Van der Ploeg, A., De Wit, P.J.G.M., and Jones, J.D.G. (2005) Plant Cell 17: 1000–1015.
28. Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003) Plant J. 33: 949–956.
29. Higgins, D., Thompson, J., and Gibson, T. (1994) Nucleic Acids Res. 22: 4673–4680.
30. Felsenstein, J. (1989). PHYLIP - Phylogeny Inference Package (Version 3.2). Cladistics 5: 164–166.
31. Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C. (2003) SWISSMODEL: An automated protein homology-modeling server. Nucleic Acids Res. 31: 3381–3385.

32. Beers, E.P., Jones, A.M., and Dickerman, A.W. (2004) Phytochemistry 65: 43–58.
33. Lu¨ thy, R., Bowie, J.U., and Eisenberg, D. (1992) Nature 356: 83–85.
34. Corpet, F. (1988) Nucleic Acids Res. 16: 10881–10890.
35. Kim, M.-J., Yamamoto, D., Matsumoto, K., Inoue, M., Ishida, T., Mizuno, H., Sumiya, S., and Katamura, K. (1992) Biochem. J. 287: 797–803.
36. Karsai, A., Mu¨ ller, S., Platz, S., and Hauser, M.T. (2002) Biotechniques 32: 790–796.
37. Rotenberg, D., Thompson, T.S., German, T.L., and Willis, D.K. (2006) J. Virol. Methods 138: 49–59.
38. Van der Hoorn, R.A.L., Leeuwenburgh, M.A., Bogyo, M., Joosten, M.H.A.J., and Peck, S.C. (2004) Plant Physiol. 135: 1170–1178.
39. Lennon-Dumenil, A.M., Bakker, A.H.,Maer, R., Fiebiger, E., Overkleeft, H.S., Rosemblatt, M., Ploegh, H.L., and Lagaudriere-Gesbert, C. (2002) J. Exp. Med. 196: 529–539.
40. Schaffer, M.A., and Fischer, R.L. (1988) Plant Physiol. 87: 431–436.
41. Harrak, H., Azelmat, S., Baker, E.N., and Tabaeizadeh, Z. (2001) Genome 44: 368–374.
42. Holwerda, B.C., Padgett, H.S., and Rogers, J.C. (1992). Plant Cell 4: 307–318.
43. Sanz-Alferez, S., Mateos, B., Alvarado, R., and Sanches, M. (2008) Eur. J. Plant Pathol. 120: 417–425.
44. Tornero, P., Conejero, V., and Vera, P. (1997) J. Biol. Chem. 272: 14412–14419.
45. Carter, C., Pan, S., Zouhar, J., Avila, E.L., Girke, T., and Raikhel, N.V. (2004) Plant Cell 16: 3285–3303.
46. Ahmed, S.U., Rojo, E., Kovaleva, V., Venkataraman, S., Dombrowski, J.E., Matsuoka, K., and Raikhel, N.V. (2000) J. Cell Biol. 149: 1335–1344.
47. Hayashi, Y., Yamada, K., Shimada, T., Matsushima, R., Nishizawa, N.K., Nishimura, M., and Hara-Nishimura, I. (2001) Plant Cell Physiol. 42: 894–899.
48. Watanabe, E., Shimada, T., Tamura, K., Matsushima, R., Koumoto, Y., Nishimura, M., and Hara-Nishimura, I. (2004) Plant Cell Physiol. 45: 9–17.

49. Zhao, Y., Thilmony, R., Bender, C.L., Schaller, A., He, S.Y., and Howe, G.A. (2003) Plant J. 36: 485–499
50. Avrova, A.O., Stewart, H.E., De Jong, W.D., Heilbronn, J., Lyon, G.D., and Birch, P.R. (1999) Mol. Plant Microbe Interact. 12: 1114–1119.
51. Avrova, A.O., et al. (2004) Mol. Plant Pathol. 5: 45–56.
52. Stotz, H.U., Bishop, J.G., Bergmann, C.W., Koch, M., Albersheim, P., Darvill, A.G., and Labavitch, J.M. (2000) Physiol. Mol. Plant Pathol. 56: 117–130.
53. Bishop, J.G., Ripoll, D.R., Bashir, S., Damasceno, C.M.B., Seeds, J.D., and Rose, J.K.C. (2005) Genetics 169: 1009–1019.
54. Bomblies, K., Lempe, J., Epple, P., Warthmann, N., Lanz, C., Dangl, J.L., and Weigel, D. (2007) PLoS Biol. 5: 1962–1972.
55. Raiola, A., Sella, L., Castiglioni, C., Balmas, V., and Favaron, F. (2008) Fungal Genet. Biol. 45: 776–789.
56. Shindo, T., and Van der Hoorn, R.A.L. (2008) Mol. Plant Pathol. 9: 119–125.
57. Wulff, B.B.H., Kruijt, M., Collins, P.L., Thomas, C.M., Ludwig, A.A., De Wit, P.J.G.M., and Jones, J.D.G. (2004) Plant J. 40: 942–956.
58. Bomblies, K., and Weigel, D. (2007) Nat. Rev. Genet. 8: 382–393.
59. Bateman, A., and Bennett, H.P.J. (1998) J. Endocrinol. 158: 145–151.
60. Yamada, K., Matsushima, R., Nishimura, M., and Hara-Nishimura, I. (2001) Plant Physiol. 127: 1626–1634.
61. Drake, R., John, I., Farrell, A., Cooper, W., Schuch, W., and Grierson, D. (1996) Plant Mol. Biol. 30: 755–767.
62. DeLano, W.L. (2002) The PyMOL User's Manual. (Palo Alto, CA: DeLano Scientific).

.

CHAPTER 3

CONTRASTING SELECTIVITY OF FUNGAL AND OOMYCETE EFFECTORS TARGETING TOMATO CYSTEINE PROTEASES.

SUMMARY

Suppression of host defense responses is an important strategy of adapted plant pathogens. In previous chapter we have showed that the fungal effector protein AVR2 targets two defense-related tomato proteases (RCR3 and PIP1) that are under diversifying selection. In this chapter we tested which of the seven secreted tomato proteases is targeted by the EPIC1 and EPIC2B (**E**xtracellular **P**rotease **I**nhibitor having **C**ystatin like domain) effector proteins of the oomycete pathogen *Phytophthora infestans*. In contrast to AVR2, both EPIC proteins inhibit C14, an abundant, ubiquitous, stress-related protease that is typified by an additional C-terminal granulin domain and resides mostly inside the host cell. The fact that effectors from different pathogens target different host proteases might reflect different infection strategies employed by these pathogens.

INTRODUCTION

Tomato is a host for leaf pathogens with contrasting lifestyles. The fungal pathogen *Cladosporium fulvum* is a tomato-adapted biotrophic pathogen, which grows intercellular on tomato leaves, without killing the host tissue (1). The oomycete pathogen *Phytophthora infestans*, on the other hand, is a hemibiotrophic pathogen that requires living cells during the initial stage of infection, but then causes extensive necrosis at a later stage (2).

The defense response of tomato is universal and includes the accumulation of proteins that are potentially harmful for these pathogens, such as β-1, 3-glucanases, chitinases, subtilases (e.g. P69B), and cysteine proteases (e.g. PIP1) (3- 6). These defense-related enzymes are thought to directly target the pathogens e.g. by degrading their cell wall components. Successful tomato pathogens evolved means to suppress these defense responses. *P. infestans*, for example, secretes glucanase and subtilase inhibitors whereas *C. fulvum* secretes chitin-binding proteins to protect its cell wall (5, 7-11). *P. infestans* and *C. fulvum* also secrete cystatin-like EPIC1 and EPIC2B (**E**xtracellular **P**rotease **I**nhibitor having **C**ystatin like domain) proteins, and AVR2, respectively, to inhibit tomato secreted papain-like cysteine proteases RCR3 and PIP1 (6, 12-14, J. Song and S. Kamoun, unpublished). Taken together, these observations are consistent with the hypothesis that pathogen effectors evolve to inhibit harmful secreted proteases.

However, the emerging concept is that most effectors have multiple targets (15), and PIP1 and RCR3 are not the only papain-like cysteine proteases (PLCPs) that are encountered by these pathogens during infection. Tomato secretes seven different PLCPs into the apoplast that belong to four distinct classes (12). Besides RCR3 and PIP1 (subfamily 5), tomato secretes five other PLCPs that belong to different subfamilies. These PLCPs are C14 (subfamily 1), aleurain-like ALP and CYP3 (subfamily 7) and cathepsin B-like CatB1 and CatB2 (subfamily 8) (12, 16). C14 is unique amongst these proteases since it carries a C terminal granulin domain with unknown function (17).

In this study, we investigated which of the seven tomato proteases are targeted by the EPIC1 and EPIC2B effectors of *P. infestans* and compared this selectivity to that of AVR2. This reveals that, at high stringency, EPIC and AVR2 effectors target different host proteases, which might reflect the different lifestyles of these pathogens.

MATERIALS AND METHODS

Materials

L-cysteine, SDS, Coomassie Brilliant Blue, M2 FLAG antisera, streptavidin-HRP polymer and FLAG M2 affinity matrix were purchased form Sigma-Aldrich. PVDF membrane was from Millipore. HRP conjugated anti-rabbit antibodies were commercially obtained from Amersham. BTH was purchased from Actigard, Syngenta. All other chemicals were of analytical grade.

Plant growth and BTH treatment

All plants were grown in a climate chamber at a 14-hour light regime at 18ºC (night) and 22ºC (day). Four- to six-week old plants were used for experiments. BTH treatment was carried out by watering 5- week-old tomato plants (cultivar Money Maker 2) with 25 µg/mL BTH very second day. Samples were taken 5 days after starting the BTH treatment, unless otherwise indicated.

EPIC1, EPIC2B and AVR2 proteins

Expression of AVR2, EPIC1 and EPIC2B in *E. coli* was conducted as described previously (6, 12). Protein purities were checked on 17% protein gels by staining with Coomassie Brilliant Blue or by silver staining and proteins were quantified using the Bradford assay (Bio-Rad), following the manufacturers guidelines.

Inhibition assays with agroinfiltrated proteases.

Tomato proteases were overexpressed by agroinfiltration as described in previous chapter. (12). Apoplastic fluid or total extracts from agroinfiltrated *Nicotiana benthamiana* were isolated at 5 dpi. In a total volume of 400 µL, 40 µL extract was preincubated in 50 mM NaAc pH 5, 1 mM L-cysteine, with or without AVR2, EPIC1

and EPIC2B or 10 μM E- 64 for 30 min at room temperature. DCG-04 was added to a final concentration of 2 μM, and the labeling was performed for another 5 h at room temperature. Proteins were precipitated by adding 1 mL ice-cold acetone and centrifugation (1 min, 16,000g). Pellets were dissolved in SDS gel loading buffer. Detection of biotinylated proteins was done as previously described using streptavidin-HRP polymer (18).

Activity-based profiling on tomato extracts

Apoplastic fluid (AF) was isolated by vacuum infiltration of tomato leaves as described earlier (Chapter 2). Equal volumes (450 μL) were used for protease activity profiling, protein gel blot analysis, or Coomassie staining. Samples were 10-fold concentrated by precipitating with two volumes ice-cold acetone and dissolving in the protein pellet in protein gel loading buffer. Total extract (TE) were generated by grinding six randomly taken 28 mm^2 leaf discs into 1 mL of water. TEs were centrifuged (1 min, 16,100g), and the supernatant was analyzed as described for AFs.

Western analysis

Western blot analysis was performed as described for detection of biotinylated proteins, using C14 and PIP1 antibodies (6), followed by detection with HRP conjugated anti-rabbit antibodies (Amersham). C14 antibodies were raised in rabbits using peptides DTEEDYPYKERNGVC and DQYRKNAKVVKIDSYC (Eurogentec, Belgium) and tested on extracts with and without C14.

Immuno precipitation

For Coimmunoprecipitations (Co-IPs), 180 μl of C14 containing extracts was preincubated with or without 100 μM E-64 for 30 min. at pH 6.5; incubated with 66 nM FLAG-tagged EPIC proteins (30 min); and purified on FLAG M2 affinity matrix according to the manufacturer's instructions. The matrix was washed with TBS and eluted with 50 μl gel loading buffer. Eluates were analyzed on western blots with C14 and M2 FLAG antisera.

RESULTS

EPICs and AVR2 target different host proteases.

To investigate which of the secreted PLCPs of tomato are inhibited by EPICs, we produced each of the PLCPs by agroinfiltration and used extracts of agroinfiltrated leaves for activity-based protein profiling (ABPP) in the absence and presence of inhibitors. ABPP of PLCPs is based on the use of DCG-04, which is a biotinylated derivative of PLCP inhibitor E-64 that irreversibly reacts with the active site cysteine residue in a mechanism dependent manner (19). This technique was used to show that AVR2 inhibits RCR3 and PIP1 (12- 14), and EPIC2B inhibits PIP1 (6). The advantage of using ABPP is that proteases can be produced *in planta* and tested without purification, allowing us to test selectivity in the presence of other proteases. To test which of the seven tomato proteases are inhibited by AVR2, EPIC1 and EPIC2B, extracts containing the proteases were pre-incubated with these inhibitors and then incubated with DCG-04 to label the non-inhibited proteases. We choose conditions to select for strong interacting inhibitors by using long labeling times (5 hours), at relatively high DCG-04 concentrations (2 μM), and low inhibitor concentrations (66 nM).

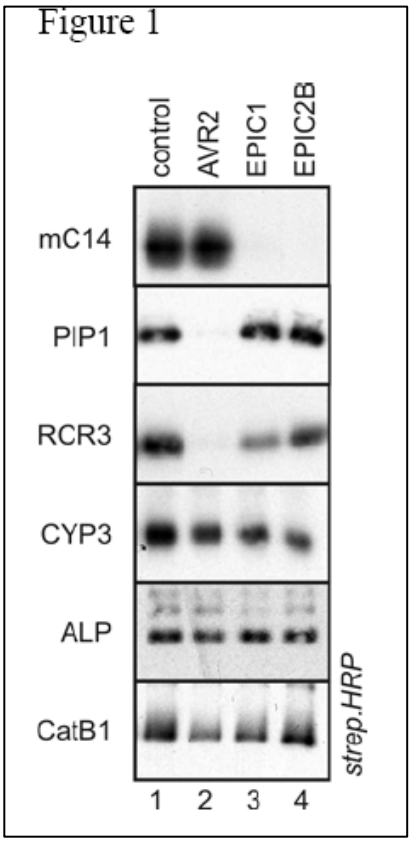


Figure 1. Selective inhibition of tomato proteases. Extracts from agroinfiltrated *N. benthamiana* leaves overexpressing different proteases (indicated on the left) were preincubated for 30 min with 66 nM AVR2, EPIC1 or EPIC2B. DCG-04 was added after preincubation to label the noninhibited proteases. Biotinylated proteases were visualized on protein blots using streptavidin-HRP. Representatives of at least three independent experiments are shown.

Under these conditions, weak, reversible interactions will not be detected since DCG-04 reacts irreversibly and would eventually label all proteases. Preincubation of the protease-containing extracts with the inhibitors, followed by labeling with DCG-04 revealed that EPIC1 and EPIC2B prevents DCG-04 labeling of only C14, whereas preincubation with AVR2 prevents biotinylation of only RCR3 and PIP1 (**Fig. 1A**). This remarkable selectivity indicates that, under stringent conditions, these inhibitors target different host proteases, forming tight complexes that persist over long incubation times.

EPIC and AVR2 have different inhibition characteristics.

The inhibition assays described above were done at 66 nM inhibitor concentrations. To investigate inhibition at lower inhibitor concentrations, C14 and PIP1 were incubated with or without various concentrations of EPIC1/EPIC2B or AVR2, respectively, and then incubated with DCG-04. This revealed that EPIC1 and EPIC2B inhibit C14 at concentrations above 2.4 nM (**Fig. 2A**), indicating that EPIC1 and EPIC2B have a similar affinity for C14. The concentrations required for inhibition of C14 by EPIC1 and EPIC2B is lower when compared to the inhibition of PIP1 by AVR2 (**Fig. 2A**), suggesting that EPIC proteins are relatively strong inhibitors of C14.

However, precise comparisons of inhibition strengths were not possible with these assays. The inhibition experiments described above were performed at pH of the apoplast (pH 5). To investigate if the inhibition could also take place at a wider pH range, inhibition assays were performed at pH 4 to pH 7. Both C14 and PIP1 can be labeled throughout this pH range, indicating that these proteases are active (**Fig. 2B**). Interestingly, inhibition of C14 by EPIC1 and EPIC2B occurs at pH 4-7, in contrast to the inhibition of PIP1 by AVR2, which only occurs at acidic pH (**Fig. 2B**).

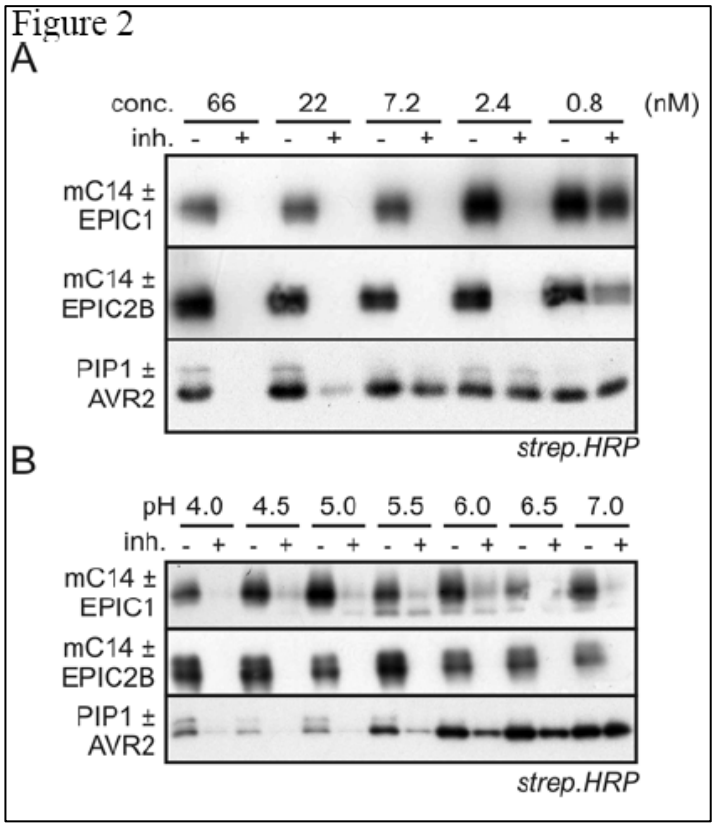


Figure 2. Characteristics of EPIC-C14 and AVR2-PIP1 interactions.

A, Lowest inhibitor concentrations. Protease-containing extracts were incubated with different inhibitor concentrations for 30 minutes. DCG-04 was added after preincubation to label the non-inhibited proteases. Biotinylated proteins were visualized on protein blots using streptavidin HRP.

B, pH-dependency of inhibition. Extracts from agroinfiltrated leaves overexpressing the proteases were preincubated for 30 min. at different pH with or without 66 nM inhibitor. DCG-04 was added after preincubation to label the non-inhibited proteases. Biotinylated proteins were visualized on protein blots using streptavidin HRP.

EPIC and AVR2 target different proteases in apoplastic fluids.

To investigate if selective inhibition also occurs in secreted proteomes, we preincubated apoplastic fluids isolated from tomato with AVR2, EPIC1 and EPIC2B, and then labeled with DCG-04. Like its Arabidopsis ortholog (RD21, Ref.17), tomato C14 exists in two active isoforms, depending on the presence or absence of a C-terminal granulin domain (**Fig. 3A**). The intermediate isoform (iC14) carries the granulin domain and is 37 KDa, whereas the mature isoform (mC14) is 30 kDa and lacks the granulin domain. The experiments described above were done with mC14, because iC14 tends to precipitate and was therefore not present in extracts from agroinfiltrated leaves. In activity profiles of crude apoplastic fluids, however, iC14 causes a unique 40kD signal, in addition to signals at 30 kDa (mC14, CatB1, CatB2, CYP3 and ALP) and 25 kDa (PIP1 and RCR3) (12). Activity-based profiling of apoplastic fluids preincubated with EPICs or AVR2 demonstrate that both EPIC1 and EPIC2B prevent biotinylation of iC14 (**Fig. 3B**, lanes 3 and 4). Signals in the region containing mC14 are also suppressed in the presence of EPICs. EPIC2B also suppressed the biotinylation of the lowest strong signal, which predominantly represents PIP1. This is consistent with previous findings that EPIC2B but not EPIC1 can inhibit, and indicates that this assay also facilitates the detection of also weaker interactions (6). In contrast, AVR2 does not prevent biotinylation of iC14, but suppresses biotinylation of the lowest signal, which predominantly consists of PIP1 (**Fig. 3B**, lane 2, Ref.12). These data indicate that EPIC and AVR2 selectively target different proteases also in apoplastic fluids in the presence of other abundant proteases.

C14 and PIP1 accumulate differently during defense and at different locations.

To investigate C14 and PIP1 accumulation during the defense responses of tomato, we isolated apoplastic fluids (AFs) and total extracts (TE) at five days after treating tomato plants with water or with BTH (benzothiadiazole). BTH is a salicylic acid analog that is routinely used to induce defense response in plants (20). Antibodies raised against C14 detect two signals on the immunoblot, which correspond to the intermediate (i) and mature (m) isoforms of C14. Since we loaded apoplastic fluid extracts from 44-times more leaf material than total extracts, we conclude that most of the C14 protein is in the

total extract while only a low amount is present in apoplastic fluids (Fig. 4, lanes 1 and 3). This is in contrast to PIP1, which is fully secreted (Fig. 4, lanes 1 and 3). In BTH-treated plants, C14 protein levels in total tissue extract and apoplastic fluid are only slightly increased when compared to the loading controls (**Fig. 4**, lanes 2 and 4). In contrast, PIP1 levels are strongly induced upon BTH-treatment. These data are consistent with our previous data that transcript levels and activity of C14 and PIP1 are slightly and strongly induced by BTH treatment, respectively (12). Thus, in contrast to the secreted and BTH-induced accumulation of PIP1, C14 accumulation is only slightly induced by BTH-treatment and only partially secreted.

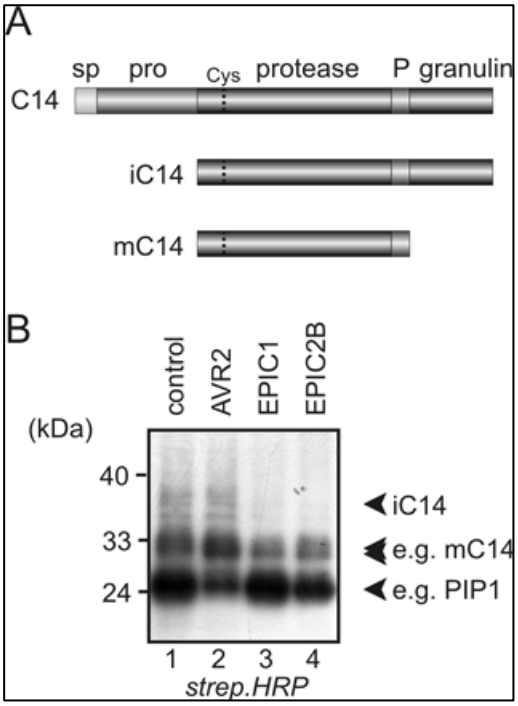


Figure. 3. Selective inhibition in apoplastic fluids.

A. Isoforms of C14. The C14 gene product contains a signal peptide (sp); protease domain with catalytic cysteine (Cys); proline-rich domain (P); and a granulin domain. As with *Arabidopsis* RD21 (17), C14 exists in two active forms: intermediate C14 (iC14) and mature C14 (mC14) (12).

B. Selective inhibition in apoplastic fluids. Tomato apoplastic fluids were preincubated with 65 nM AVR2, EPIC1 or EPIC2B, before adding DCG-04 to label the remaining non-inhibited proteases. Biotinylated proteins were detected using streptavidin-HRP.

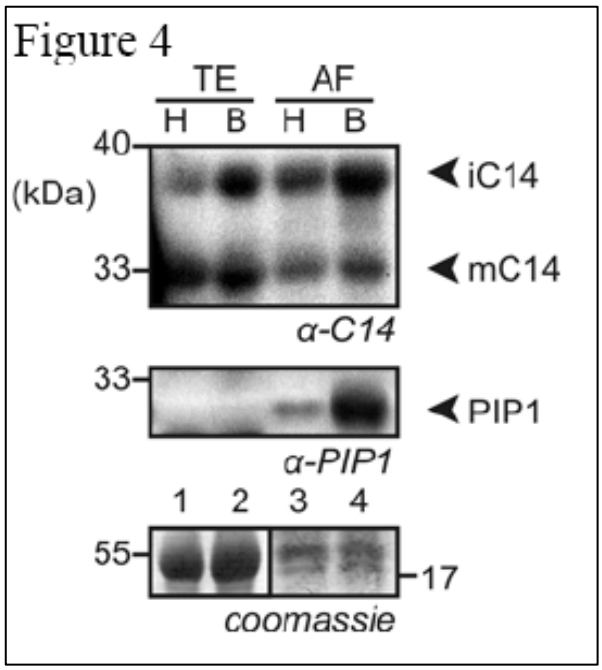


Figure. 4. Induced accumulation and location of C14 and PIP1. Total extracts (TE) and apoplastic fluids (AF) were generated from tomato plant treated for five days with water (H) or BTH (B) and protein blots were probed with antibodies against C14 (top) and PIP1 (middle). Coommassie (cbb)-stained gels (bottom) show protein loading. The AF loaded on gel correlates to ~7.4 cm^2 leaf, whereas the TE sample correlates to ~0.17 cm^2 leaf. A representative of three independent experiments is shown.

EPIC1 and EPIC2B physically interact with C14.

In the previous chapter, we showed a physical interaction between PIP1 and AVR2 (12). We performed coimmunoprecipitation to investigate if there was also a physical interaction between C14 and EPIC1s. C14 containing extracts were preincubated with FLAG-EPIC1 or FLAG- EPIC2B and FLAG-tagged protein complexes were pulled down using anti FLAG beads. Proteins interacting with the beads were analyzed on protein blots using anti-C14 and anti-FLAG antibodies. This revealed that C14 interacts with both EPIC1 and EPIC2B (Fig. 5, lane 2 and 5). The interaction is prevented by preincubation of C14 with an excess E-64 (Fig. 5, lane 3 and 6). Thus, C14 physically

interacts with EPIC proteins, presumably at the active site. This complex is tight, but not covalent.

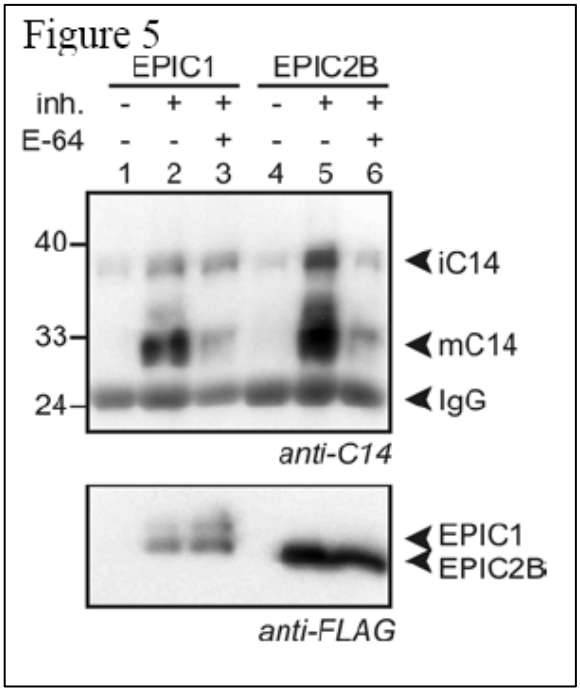


Figure. 5. EPIC1 and EPIC2B physically interact with C14. Extracts containing C14 were preincubated for 30 minutes with or without 100 µM E-64. FLAG-EPIC1 or FLAG-EPIC2B (66nM) was added and incubated for 30 minutes. Protein complexes were immuno- precipitated using an anti-FLAG matrix.

The matrix was washed and boiled in SDS sample buffer. Eluted proteins extra space detected on protein blot using C14 and FLAG antibodies. iC14 represents the intermediate isoform of C14, which tends to precipitate. mC14 is the mature C14 isoform. IgG is the light chain anti-FLAG antibody eluted from the anti-FLAG column.

DISCUSSION

This study revealed a contrasting selectivity of fungal and oomycete inhibitors targeting host proteases. The fungal AVR2 effector targets defense-related proteases PIP1 and RCR3 (12). In contrast, the oomycete EPIC proteins are strong inhibitors of C14, a partially secreted papain-like cysteine protease of tomato carrying a granulin domain. The EPIC-C14 interaction is specific, occurs at acidic and neutral pH and at low inhibitor concentrations. Besides different selectivity, also the characteristics of the inhibition and the targeted proteases are different between EPICs and AVR2. First, in contrast to AVR2-RCR3/PIP1 interactions, which occur only at apoplastic (acidic) pH, EPIC-C14 interactions also occur at neutral pH (Fig. 2B). This is probably caused by structural properties of the interaction surface of the protease and/or inhibitor and indicates that EPIC-C14 interactions can occur at non-apoplastic conditions. Second, in contrast to the fully secreted PIP1 and RCR3, C14 is only partially secreted (Fig. 4).

This indicates that C14 functions inside the host cell rather than extracellularly and is consistent with immunocytological studies where C14-like proteases were found in vesicles, the nucleus and chloroplasts (17, 21). Third, in contrast to the strong accumulation of PIP1 upon BTH treatment, C14 protein levels are only moderately induced by BTH (Fig. 4). These data are consistent with the fact that transcript levels of C14 and PIP1 and are slightly and strongly in BTH-treated plants, respectively (12).

Finally, C14 and PIP1 belong to distinct phylogenetic clades and C14 carries a unique C-terminal granulin domain, indicating that these proteases have different functions. C14 is not the only target of EPICs. This is consistent with the emerging concept in effector biology that most effectors have multiple targets in the host (15). It was previously shown that EPIC2B but not EPIC1 inhibits PIP1 (6). In addition, EPIC1 and EPIC2B both inhibit RCR3, which is closely related to PIP1 (J. Song and S. Kamoun, unpublished). Short labeling times and different probe and inhibitor concentrations were used in these assays to show that EPICs prevent biotinylation of PIP1 and RCR3. By using long labeling times, high DCG-04 concentrations and low inhibitor concentrations, we selected for strong inhibitors since DCG-04 is an irreversible probe that would eventually

label all proteases that are reversibly inhibited, unless the interaction with the inhibitor is strong enough.

Under these conditions, only biotinylation of C14 can be prevented by EPICs, indicating that C14 is the target with the highest affinity for EPICs. However, in the experiment described in Fig. 3B, we confirmed the finding that EPIC2B but not EPIC1 inhibits PIP1 (6). Collectively, these results suggest that pathogen inhibitors have multiple targets, with the C14-EPIC and the AVR2- RCR3/PIP1 interactions being relatively strong since they occur at nanomolar concentrations and persist for hours.

C14 is a highly conserved protease that occurs throughout the plant kingdom. C14-like proteases are characterized by a unique, C terminal granulin-like domain that shares homology to animal growth hormones that are released upon wounding (22). The tomato C14 is relatively abundant and has been studied many times under the names TDI-65, CYP1 and SENU1 and is known to be transcriptionally induced by heat, cold, drought and senescence (23- 26). The potato ortholog of C14 is called CYP1 and is transcriptionally induced in resistant potato cultivars early during infection of *P. infestans* (27).

The Arabidopsis ortholog is named RD21, accumulates in vesicles and can act as a peptide ligase (17, 28, 29). Although the function of C14- like proteases is currently unknown, these data suggest that C14-like proteases have an intrinsic function inside the cell, related to general stress responses. The distinct target-selectivity between EPICs and AVR2 might reflect different infection strategies between biotrophic and hemibiotrophic pathogens. PIP1 and RCR3 are secreted enzymes that accumulate in the apoplast, which is the compartment where *C. fulvum* grows without causing cell death of the host. AVR2 produced by the fungus inhibits RCR3 and PIP1 only at apoplastic pH, presumably creating a nonproteolytic niche around the pathogen. In contrast, C14 accumulates mostly inside plant cells. We speculate that *P. infestans* encounters C14 mainly during the necrotrophic phase of infection, when cell death occurs and internal proteases become

exposed to the pathogen. The property of EPIC to inhibit C14 at both acidic and neutral pH might be an important strategy to prevent damage caused by C14.

In conclusion, the fact that pathogen-derived effectors target different host proteases is an intriguing finding that points at distinct infection strategies employed by these pathogens. Assays to study these proteases by reverse genetics and during pathogen infection are underway and could reveal unique infection strategies of biotrophic fungi and hemibiotrophic oomycetes.

REFERENCES

1. Thomma, B. P. H. J., Van Esse, H. P., Crous, P. W., and De Wit, P. J. G. M. (2005) Mol. Plant Pathol. 6: 379-393.
2. Kamoun, S., and Smart, C. D. (2005) Plant Disease 89: 692-699.
3. Van Loon, L. C., Rep, M., and Pieterse, C. M. J. (2006) Annu. Rev. Phytopathol. 4 : 135-162.
4. Ferreira, R. B., Monteiro, S., Freitas, R., Santos, C. N., Chen, Z., Batista, L. M., Duarte, J.,Borges, A., and Teixeira, A. R. (2007) Mol. Plant Pathol. 8 : 677-700.
5. Tian, M., Benetti, B., and Kamoun, S. (2005) Plant Physiol. 138 : 1785-1793.
6. Tian M., Win, J., Van der Hoorn, R., Van der Knaap, E., and Kamoun, S. (2007) Plant Physiol. 143 : 364-377.
7. Rose, J. K., Ham, K. S., Darvill, A. G., and Albersheim, P. (2002) Plant Cell 14: 1329-1345.
8. Tian, M., Huitema, E., Da Cunha, L., Torto-Alalibo, T., and Kamoun, S. (2004) J. Biol. Chem. 279 : 26370-26377.
9. Van den Burg, H. A., Harrison, S. J., Joosten, M. H. A. J., Vervoort, J., and De Wit, P. J. G. M. (2006) Mol. Plant-Microbe Interact. 19 : 1420-1430.
10. Van Esse, H. P., Bolton, M. D., Stergiopoulos, I., De Wit, P. J. G. M., and Thomma, B. P. H. J. (2007) Mol. Plant-Microbe Interact. 20 : 1092-1101.
11. Bolton, M. D., Van Esse, H. P., Vossen, J. H., De Jong, R., Stergiopoulos, I., Stulemijer, I. J., Van den Berg, G. C., Borras-Hidalgo, O., Dekker, H. L., De Koster, C. G., De Wit, P. J. G., M., Joosten, M. H. A. J., and Thomma, B. P. H. J. (2008) Mol. Microbiol. 69: 119-136.
12. Shabab, M., Shindo, T., Gu, C., Kaschani, F., Pansuriya, T., Chintha, R., Harzen, A., Colby, T., Kamoun, S., Van der Hoorn, R.A.L. (2008). Plant Cell 20 : 1169-1183.
13. Rooney, H. C. E., Van't Klooster, J. W., van der Hoorn, R. A. L., Joosten, M. H. A. J., Jones, J. D., and De Wit, P. J. G. M. (2005) Science 308 : 1783-1786.

14. Van Esse, H. P., Van ‘t Klooster, J. W., Bolton, M. D., Yadeta, K., Van Baarlen, P., Boeren, S., Vervoort, J., De Wit, P. J. G. M., and Thomma, B. P. H. J. (2008). Plant Cell 20 : 1948- 1963.
15. Hogenhout, S. A., Van der Hoorn, R. A. L., Terauchi, R., and Kamoun, S. (2008) Mol. Plant- Microbe Interact. in press.
16. Beers E. P., Jones A. M., and Dickerman A. W. (2004) Phytochemistry 65: 43-58.
17. Yamada, K., Matsushima, R., Nishimura, M., and Hara-Nishimura, I. (2001) Plant Physiol. 127: 1626-1634.
18. Van der Hoorn, R. A. L., Leeuwenburgh, M. A., Bogyo, M., Joosten, M. H. A. J., and Peck, S.C. (2004) Plant Physiol. 135 : 1170-1178.
19. Greenbaum, D., Medzihradszky, K. F., Burlingame, A., and Bogyo, M. (2000) Chem. Biol. 7: 569-581.
20. Kohler, A., Schwindling, S., and Conrath, U. (2002) Plant Physiol. 128: 1046-1056.
21. Tabaeizadeh, Z., Chamberland, H., Chen, R.D., Yu, L.X., Bellemare, G., and Lanfontaine, J. G. (1995) Protoplasma 186: 208-219.
22. Bateman, A., and Bennett, H. P. J. (1998) J. Endocrin. 158: 145-151.
23. Schaffer, M. A., and Fischer, R. L. (1990) Plant Physiol. 93: 1486-1491.
24. Schaffer, M. A., and Fischer, R. L. (1988) Plant Physiol. 87: 431-436.
25. Drake, R., John, I., Farrell, A., Cooper, W., Schuch, W., and Grierson, D. (1996) Plant Mol. Biol. 30 : 755-767.
26. Harrak, H., Azelmat, S., Baker, E. N., and Tabaeizadeh, Z. (2001) Genome 44: 368-374.
27. Avrova, A.O., Stewart, H.E., De Jong, W.D., Heilbronn, J., Lyon, G.D., and Birch, P.R.(1999) Mol. Plant Microbe Interact. 12: 1114–1119.
28. Hayashi, Y., Yamada, K., Shimada, T., Matsushima, R., Nishizawa, N. K., Nishimura, M., and Hara-Nishimura, I. (2001) Plant Cell Physiol. 42: 894-899.
29. Wang, Z., Gu, C., Colby, T., Shindo, T., Balamurugan, R., Waldmann, H., Kaiser, M., and Van der Hoorn, R. A. L. (2008) Nat. Chem. Biol. 4 : 557-563.

CHAPTER 4

TARGETED MUTAGENESIS OF RD21, A CYSTEINE PROTEASE IN *ARABIDOPSIS THALIANA*

SUMMARY

RD21 is an abundant papain-like cysteine protease in tissues of the model plant *Arabidopsis thaliana*. We have studied the structure-function relationship of RD21. The protease is encoded as a preproprotein with a C-terminal extension that shares homology with granulins, which are animal growth hormones released upon wounding. It exists in two active forms: with and without the granulin domain. The protease processing, activity, targeting and stability of RD21 were studied using processing site, active site and deletion mutants respectively. Localization and processing of RD21 was also studied using YFP tagged RD21.

INTRODUCTION

Vacuoles are main compartments for degradation of cellular materials such as proteins, nucleic acids, and oligosaccharides (1). To degrade the macromolecules, the vacuoles accumulate various hydrolytic enzymes. Vacuolar proteins, including hydrolytic enzymes are synthesized on the rough endoplasmic reticulum and are transported to the vacuoles. These proteins are often synthesized as larger precursors, including N-terminal propeptide (NTPP) and/or a C-terminal propeptide (CTPP). NTPPs and CTPPs are post-translationally processed just after arriving at the vacuoles to make the respective mature proteins. NTPPs and CTPPs of vacuolar proteins are known to function as vacuolar targeting signal and/or a modulator of enzyme activities. The Barley (*Hordeum vulgare*) aleurain (2) and sweet potato (*Ipomoea batatas*) sporamin (3) have a vacuolar targeting signal, Asn-Pro-Ile-Arg (NPIR), in their NTPPs (4, 5).

An *rd21* gene encoding a papain family protease, RD21 (D13043), was found to be up regulated during dehydration of *Arabidopsis* plants (6). Papain family proteases have NTPP that includes an ERFNIN motif, which is conserved in rat cathepsin H (Y00708), cathepsin L (Y00697) and human cathepsin S (7, 8). The NTPPs seem to function as an autoinhibitory domain of papain family proteases. In contrast to no CTPP in most members of papain family, RD21 (**R**esponse to **D**esiccation 21) has a

long extension sequence at the C terminus. Such an extension sequence has been found in 14 members of papain family (9), including rice (*Oryza sativa*) oryzains α (D90406) and β (D90407) (10).

The C-terminal extension sequences are composed of two domains, a Pro-rich domain and a domain that has a high homology to animal proteins of the epithelin/ granulin family (11, 12). The epithelins and granulins are approximately 6-kDa proteins that stimulate or inhibit the growth of animal cells (12). The crystal structure of carp granulin 1 has seven disulfide bonds and four stacked β hairpins (13). Compared with the animal granulins, plant granulins have an approximately 4-kD insertion with two Cys residues, which might cause a structural difference between plant and animal granulins. Plant granulins are a subclass of the papain family with a single granulin domain in the C terminal region. On the other hand, animal granulins are synthesized as a larger precursor containing seven-tandem repeats of the granulin domain, and the precursor is processed to make multiple mature granulins (11).Details of the RD21 domains and protein sequence is given in Figure 1.

RD21 is also interesting because it exists in two active forms [with and without the granulin domain] (14) and it localizes in the vesicles, those contents are probably released into the vacuole or apoplast upon pathogen infection. (15 and Van der Hoorn, unpublished).

Large-scale site-directed mutagenesis is a simple and reliable tool to get information on the structure–function relationship of proteins (16). This approach is facilitated by synthesis of oligonucleotides, PCR-based cloning and rapid DNA sequencing. To understand reaction mechanisms or to map the interaction sites, analyses of the roles of numerous amino acid positions in a given protein are often necessary. This study is aimed to study the structure-function relationship of RD21 in detail. This will increase our understanding of how RD21 functions during plant-pathogen interactions and other stress responses.

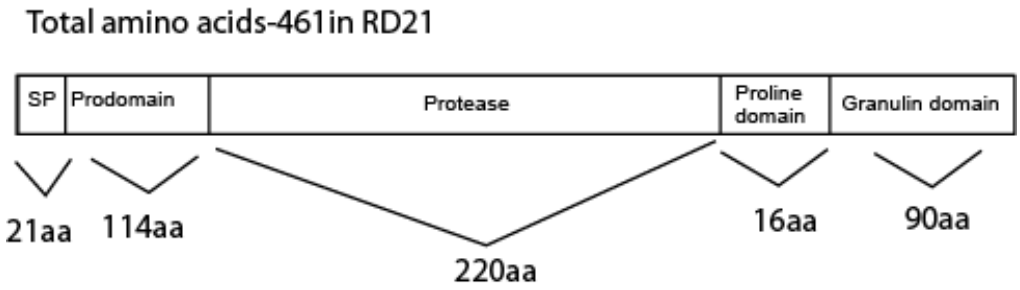


Figure 1 Schematic representation of RD21, showing different domains.RD21 has typical five domains namely SP-signal peptide, prodomain, protease domain, proline-rich domain and granulin domain.

UniProt/Swiss-Prot|P43297|RD21A_ARATH Cysteine proteinase RD21a precursor

```
10         20         30         40         50         60
MGFLKPTMAI LFLAMVAVSS AVDMSIISYD EKHGVSTTGG RSEAEVMSIY EAWLVKHGKA

        70         80         90        100        110        120
QSQNSLVEKD RRFEIFKDNL RFVDEHNEKN LSYRLGLTRF ADLTNDEYRS KYLGAKMEKK

       130        140        150        160        170        180
GERRTSLRYE ARVGDELPES IDWRKKGAVA EVKDQGGCGS CWAFSTIGAV EGINQIVTGD

       190        200        210        220        230        240
LITLSEQELV DCDTSYNEGC NGGLMDYAFE FIIKNGGIDT DKDYPYKGVD GTCDQIRKNA

       250        260        270        280        290        300
KVVTIDSYED VPTYSEESLK KAVAHQPISI AIEAGGRAFQ LYDSGIFDGS CGTQLDHGVV

       310        320        330        340        350        360
AVGYGTENGK DYWIVRNSWG KSWGESGYLR MARNIASSSG KCGIAIEPSY PIKNGENPPN

       370        380        390        400        410        420
PGPSPPSPIK PPTQCDSYYT CPESNTCCCL FEYGKYCFAW GCCPLEAATC CDDNYSCCPH

       430        440        450        460
EYPVCDLDQG TCLLSKNSPF SVKALKRKPA TPFWSQGRKN IA
```

SIGNAL	1 21	Potential.
PROPEP	22 136	Activation peptide (Potential).
CHAIN	137 356	Cysteine proteinase RD21a.
PROPEP	357 462	Removed in mature form
ACT_SITE	161 161	By similarity.
ACT_SITE	297 297	By similarity.
ACT_SITE	317 317	By similarity.
CARBOHYD	90 90	N-linked (GlcNAc...) (Potential).
CARBOHYD	414 414	N-linked (GlcNAc...) (Potential).
DISULFID	158 200	By similarity.
DISULFID	192 233	By similarity.
DISULFID	291 342	By similarity.
DISULFID	375 387	By similarity.
DISULFID	381 402	By similarity.

MATERIALS AND METHODS

Materials

DCG-04 was provided by Dr. Van Swieten, Leiden University. Streptavidin horseradish peroxidase (HRP) polymer conjugates, Tris, dithiothreitol (DTT), Triton X-100, NaCl, L-cysteine, sodium acetate, acetone, MES buffer were obtained from Sigma, USA. All the other reagents used were of analytical grade.

Plant materials

Nicotiana benthamiana and *Arabidopsis thaliana* were grown in a climate chamber at 14-h light regime at 18°C (night) and 22°C (day). Four- to six-week old plants were used for experiments.

Site-directed mutagenesis and other mutants

RD21 mutant proteins were generated by introducing mutations into pRH628 (35S:RD21: term) using the Quick-change site-directed mutagenesis kit (Stratagene); the primers are listed in Table 1. 20-base oligonucleotide primers were constructed such that one or two base changes would produce a codon change that resulted in an amino acid substitution. Mutations were confirmed by sequencing and correct expression cassettes were shuttled into pTP5 using XbaI and EcoRI resulting in pMS14, pMS15, pMS16, pMS17, pMS20, pMS21, pMS22, pMS23 and pMS24. Deletion mutants and YFP-tagged RD21 mutants were generated by standard procedures (17).

A. tumefaciens-mediated transient expression

Agrobacterium tumefaciens strain GV3101 was transformed with binary vectors and used for agroinfiltration as described previously (18). *Agrobacterium* was grown overnight at 28°C in 10mL Luria-Bertani medium containing 50 mg/mL kanamycin and 50 mg/mL rifampicin. The culture was centrifuged (10 min, 3000g) and the bacterial pellet was resuspended in 10 mM MES, pH 5, 10 mM $MgCl_2$, and 1 mM acetosyringone to a final OD_{600} of 2. *Agrobacterium* cultures containing binary

protease expression vectors were mixed with vector for silencing inhibitor p19 (19). *Agrobacterium* cultures were infiltrated into 5-week old *N. benthamiana* plants using a syringe without a needle.

Activity-based protein profiling

After 4 days, proteins were extracted by grinding a 2 cm^2 leaf disc in Eppendorf tubes, with 0.5ml water. Protease activity profiling was performed as described previously (18). Unless otherwise indicated, 200 to 450 mL of TE or AF was labeled for 5 h at room temperature with 2 mM DCG-04 (18) in the presence of 50mM sodium acetate, pH 6.0, and 1mM L-cysteine. Competition with E-64 was done by adding 20 mM E-64 (Sigma-Aldrich) 30 min before adding DCG-04. Proteins were precipitated by adding 1mL ice-cold acetone and centrifugation (1min, 16,100g). Pellets were dissolved in SDS gel loading buffer. Detection of biotinylated proteins was done as previously described using streptavidin-HRP polymer (Sigma-Aldrich) (18).

SDS-PAGE and Western-Blot Analyses

Proteins were subjected to 10% to 15% SDS-PAGE as described previously (17). Following electrophoresis, proteins were visualized by silver nitrate staining following the method of Morrissey (21), Coomassie Brilliant Blue staining (17), Colloidal Coomassie Blue staining, or the proteins were transferred to supported nitrocellulose membranes (Bio-Rad Laboratories) using a Mini Trans-Blot apparatus (Bio-Rad Laboratories). Detection of antigen-antibody complexes was carried out with a western-blot alkaline phosphatase kit (Bio-Rad Laboratories). Antiserum to RD21 was kindly provided by Dr. Hara-Nishimura (Kyoto University Japan). In western-blot analyses of tomato intercellular fluids, the antisera to RD21 reacted only with two bands of molecular weight of 33 kD and 38 kD.

Table 1. List of primers

Construct Number	Name	Primer sequence
pMS14	Catalytic cysteine mutant C161A	CAGGGTGGTTGCGGGAGTgcTTGGGCGTTTTCAACCATTGG> r121 C161A forward CCAATGGTTGAAAACGCCCAAgcACTCCCGCAACCACCCTG> r122 C161A reverse
pMS15	Catalytic residue mutant N317A	GATTACTGGATTGTGAGAgcCTCATGGGGTAAAAGCTGGGGAG> r129 N317A forward CTCCCCAGCTTTTACCCCATGAGgcTCTCACAATCCAGTAATC> r130 N317A reverse
pMS16	Putative nitrosylation site C192S	GTCTGAACAAGAGTTGGTCGATaGTGACACTTCATACAACGAAGG> r125 C192S forward CCTTCGTTGTATGAAGTGTCACtATCGACCAACTCTTGTTCAGAC> r126 C192S reverse
pMS17	Deletion of Granulin domain	ccccctgcagctATTGGGTTGGAGGCTTGATGGGAGATGG> r135 GDM reverse
pMS20	ERNIN mutant N79A	CGAGATCTTTAAAGACgcTCTTCGTTTCGTCGATGAACATAACG> r138 N79A Forward CGTTATGTTCATCGACGAAACGAAGAgcGTCTTTAAAGATCTCG> r139 N79A reverse
pMS21	Prodomain cleavage mutant R148A	GAGAAGGACTAGCCTAgcGTACGAGGCTCGTGTCGGTGATG> r140 R131A Forward CATCACCGACACGAGCCTCGTACgcTAGGCTAGTCCTTCTC> r141 R131A reverse
pMS22	TQ2AA granulin cleavage mutant R132S	CCATCTCCCATCAAGCCTCCAgCCgcATGTGACAGTTACTACACTTG> r131 TQ2AA Forward CAAGTGTAGTAACTGTCACATgcGGcTGGAGGCTTGATGGGAGATGG> r132 TQ2AA reverse
pMS23	CCC3AAA destabilizing	CACTTGTCCTGAGAGCAACACTgcTgcTgcTCTGTTTGAGTA

	granulin domain:	TGGCAAG> r136 3C3A Forward CTTGCCATACTCAAACAGAgcAgcAgcAGTGTTGCTCTCAGG ACAAGTG> r137 3C3A reverse
pMS24	AGDE Random mutagenesis of VGDE into AAAA	GCCTACGGTACGAGGCTCGTGyCGsTGmTGmGCTACCGGAGT CTATTGACTGG> r117 VGDE Forward CCAGTCAATAGACTCCGGTAGCkCAkCAsCGrCACGAGCCTC GTACCGTAGGC> r118 VGDE reverse

Localization of the YFP-tagged RD21 under the confocal microscope

Plant tissues and subcellular fractions were inspected under a laser-scanning confocal microscope (LSM510, Carl Zeiss) with a krypton-argon laser and an YFP filter set, and a fluorescence microscope (Axiphot 2, Carl Zeiss). The latter included a filter set (an excitation filter; BP450-490, a dichroic mirror; FT510, a barrier filter;BP515-565, Carl Zeiss), a CCD camera (CoolSNAP, RS Photometric, Chiba, Japan) and a light source (Arc HBO 100W, Atto, Tokyo, Japan).

RESULTS AND DISCUSSION

Catalytic cysteine residues required for RD21activity but not essential for RD21 processing

Papain-like cysteine proteases often activate themselves by pH-induced autocatalytic removal of the prodomain (22). However, when preproRD21 was produced in insect cells, removal of the prodomain and the granulin domain was achieved only when a plant extract was added (14).This indicates that there are other factors involved in the proper processing of preproRD21. To investigate if the catalytic activity of RD21 is required for prodomain and granulin domain removal, mutations of catalytic residues were introduced in RD21 and the accumulation, processing and activity of mutated RD21, were investigated. The catalytic cysteine (RD21 numbering 161) has been changed to alanine.

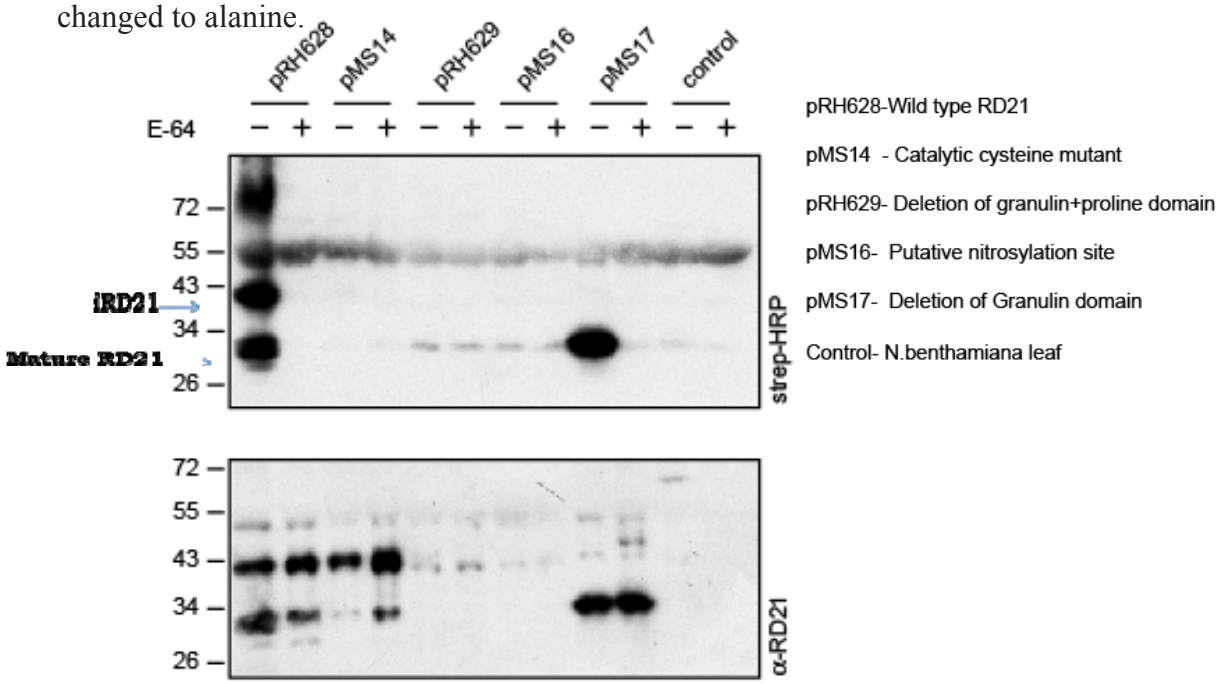


Figure 2. Extracts from agroinfiltrated leaves expressing (mutant and wild type) RD21 proteins were preincubated at pH 5.0 with or without 20 mM E-64. DCG-04 was added to label the remaining noninhibited proteases and biotinylated proteins were detected on protein blots using streptavidin-HRP. The presence of RD21 is shown using RD21 antibodies. A representative of four independent experiments is shown.

We have preferred point mutations (amino acids) most of the time into alanine because of its electronically neutral nature. As expected mutated RD21 was not active as it was not able to bind DCG04 (it binds only to active cysteine protease). However, interestingly the protein was processed and accumulated as both intermediate as well as mature form of RD21 (Figure 2, Lane2). Even in the presence of E-64, accumulation was not affected, which gives an important clue about the maturation of the protein, that any cysteine proteases are not needed for the conversion of preproRD21to mature RD21. It should be noted that no 10-kD free granulin protein was detected in the leaves.

Granulin domain is not required for the processing and activity of RD21 and destabilizing the granulin domain results in high accumulation of protein

Granulins have been isolated from several sources. All are approximately 6 kDa and are characterized by a unique and highly conserved motif of 12 cysteines consisting of four cysteine pairs flanked by two singletons towards both the amino and the carboxy termini. The granulin-like domain may serve to regulate thiol protease activity in plants (24). In RD21 granulin domain is associated with the proline-rich domain and the exact functions of both the domains are not known. RD21 with and without granulin domain are active.

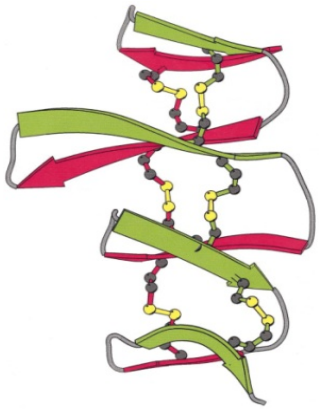


Figure 3 The three-dimensional structure of carp granulin, a novel molecular architecture of four stacked β-hairpins (23).

Deletion of granulin domain from RD21 leads to almost no effect on the processing, activity and stability of the protein (Figure 2). Deletion of granulin domain with proline-

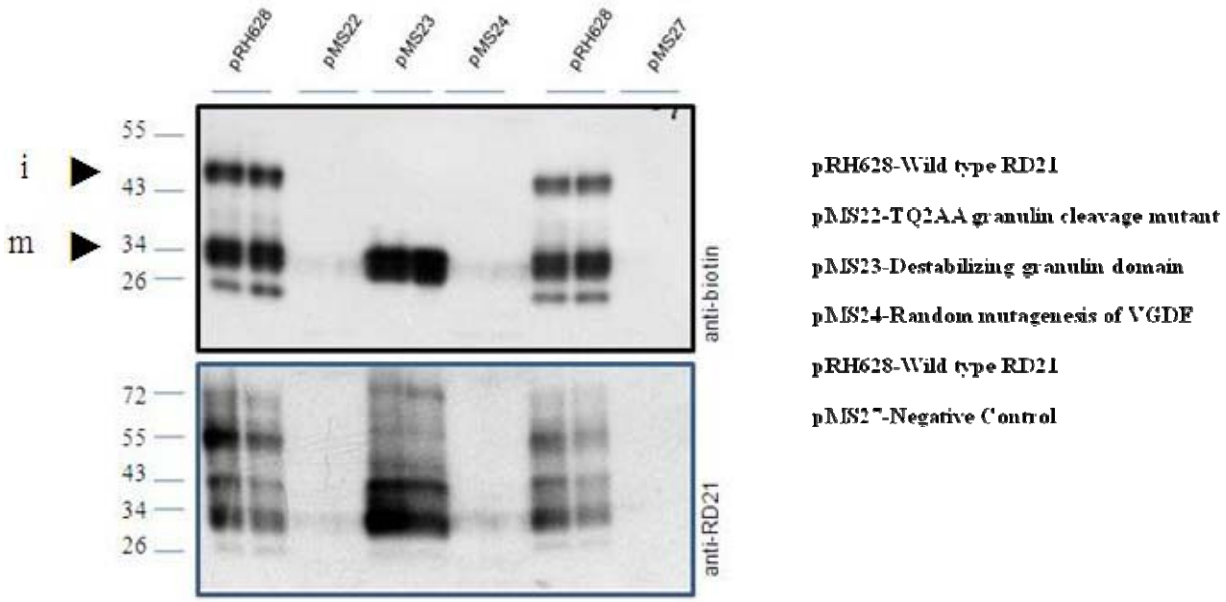


Figure 4. Extracts from agroinfiltrated leaves expressing (mutant and wild type) RD21 proteins were preincubated at pH 5.0. DCG-04 was added to label the proteases and biotinylated proteins were detected on protein blots using streptavidin- HRP. The presence of RD21 is shown using RD21 antibodies. A representative of four independent experiments is shown.

-rich domain(pRH629) leads to complete disappearance of the RD21 protein, that may be due to the formation of unfolded or misfolded protein. This leads to degradation of the protein by housekeeping enzymes in the cell (Figure 2). In eukaryotic cells, to ensure the quality of intracellular proteins, a quality-control system exists that rapidly eliminates misfolded or damaged proteins whose accumulation would interfere with normal cell function and viability (23).

The secondary structure of carp granulin has been solved using two-dimensional nuclear magnetic resonance (2D NMR) spectroscopy (24). This has revealed novel protein architecture of four stacked β-hairpins with an axial rod of disulphide bridges and with the peptide chain twisting through a left-handed super-helix (Figure 3, Lane 3). The disulphide bonds present in the domain have a very important role in the stability of the protein. We mutated the 3 cysteine (CCC) in a row to alanine (pMS23). This mutation resulted in the destabilization of granulin domain leading to high accumulation of protein. However, the mutated RD21 (mRD21) retained its

activity. (Figure 4, lane 3). Interestingly no intermediate form of RD21 was detected on both immuno as well as anti- biotin blots.

Change in putative nitrosylation site influences activity of RD21

Activity of many proteins (including proteases) can be regulated by nitrosylation of cysteine that are flanked by charged residues (26). RD21 has one cysteine, which is probably not involved in the formation of disulphide bridges but is exposed at the surface and flanked by acidic residues. A mutation of this residue was introduced in RD21 and the accumulation, processing and activity of this mutated RD21 was investigated. It was found that this cysteine is critical for the protein integrity, as it does not accumulate; moreover, it seems that the protein is not stable.

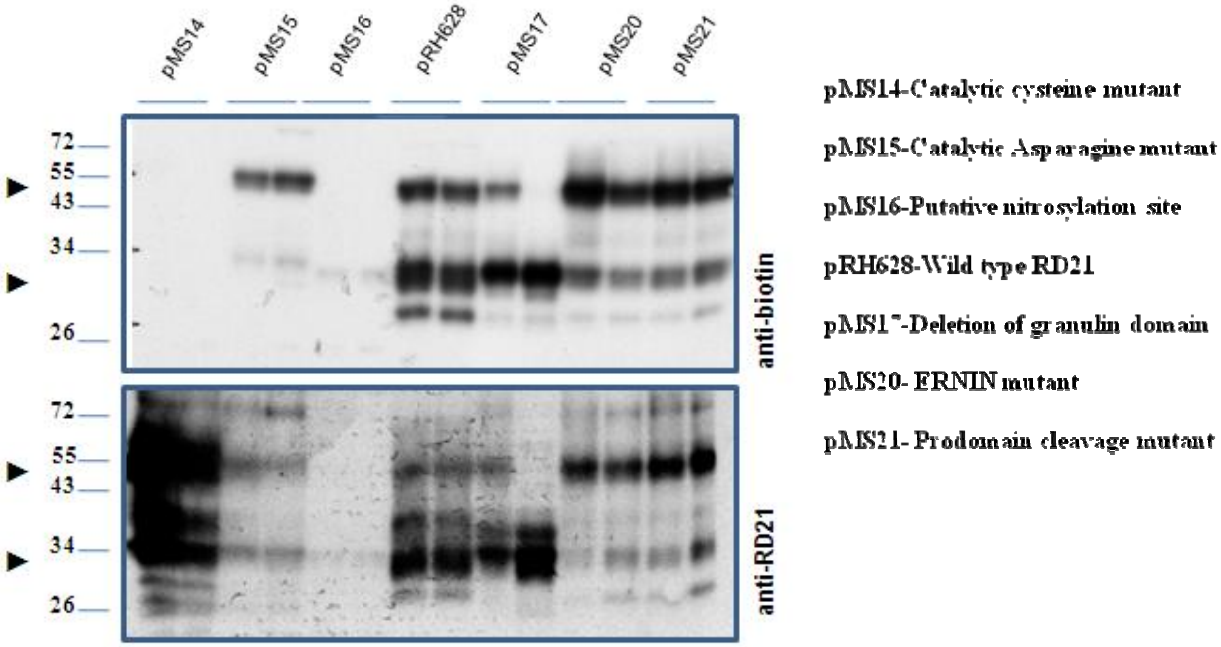


Figure 5 Extracts from agroinfiltrated leaves expressing (mutant and wild type) RD21 proteins were preincubated at pH 5.0. DCG-04 was added to label the proteases and biotinylated proteins were detected on protein blots using streptavidin- HRP. The presence of RD21 is shown using RD21 antibodies.

Catalytic asparagine mutant (pMS15) not required for RD21 activity but essential for RD21 processing

The role of the asparagine residue in the Cys-His-Asn "catalytic triad" of cysteine proteases has been investigated by replacing Asn (N317A) in RD21 with alanine using site-directed mutagenesis. In the previous studies, it was shown that in papain Asn 175 was shown to increase aggregation and proteolytic susceptibility of the proenzymes and to have on effect on the thermal stability of the mature enzyme, reflecting a contribution of the asparagine residue to the structural integrity of papain (27).

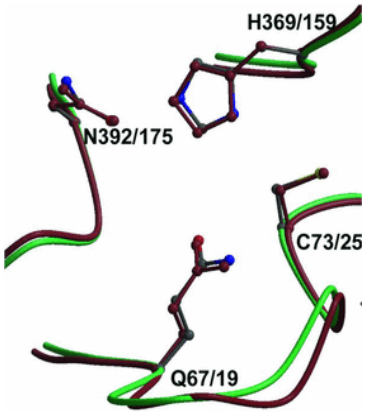


Figure 6 Catalytic triad (Cys-His-Asn) of papain like cysteine proteases, papain used for representation.

In RD21, it was found that Asn-317 is not required for the activity of the enzyme because DCG04 was able to label the protein (Figure5, lane 2). DCG-04 can only bind to active enzyme. Since all the protein was present as intermediate form (iRD21), it seems that pMS15 is essential for RD21 processing. The mature form of RD21 was not seen on antiRD21 immunoblot. Conversion of intermediate form (iRD21) to mature mRD21 requires the catalytic asparagine, which take places with unknown mechanism.

ERFNIN mutant

The ERFNIN is a very common motif found in many papain like cysteine proteases. The propeptide region of RD21 also contains this motif. It consists of conserved

amino acids interspersed with variable ones: EX3RX2 (V/I) FX2NX3IX3N. This "ERFNIN" motif, named for the single letter code of the conserved amino acids, was found in 15 cysteine protease genes identified in a combined search of the literature and the GenBank database (28).

ERFNIN motif in the propeptide serves to inhibit protease activity and the removal of the propeptide converts the protein to the enzymatically active form. Although it has been demonstrated that the propeptide inhibits enzyme activity of cysteine proteases, proregions peptides of an aspartyl protease and carboxypeptidase A specifically inhibit the respective mature proteases (29, 30).

In this study, Asparagine (Asp-79) was converted into alanine by site-directed mutagenesis. The construct pMS20 (N79A) was found to be as iRD21. This indicates that ERFNIN motif does not have much influence on activity, processing and maturation of protein (Figure 5, lane 6). This observation is very surprising, because most of the cysteines proteases are regulated by their conserved motifs in propeptide region.

Residues at the cleavage sites are required for RD21 processing

Mass-spectrometry studies have revealed the real N- and C- terminus of the mature RD21 (lacking the prodomain and granulin domain, Van der Hoorn, unpublished data). Mutations at the putative cleavage sites were introduced to determine the specific requirements for cleavage. In general prodomain removal is necessary for enzyme to be active in the cell. In order to study the mechanism underlying the cleavage of prodomain from the protease and the factors involved in the process, we introduced point mutations at Arg-148.

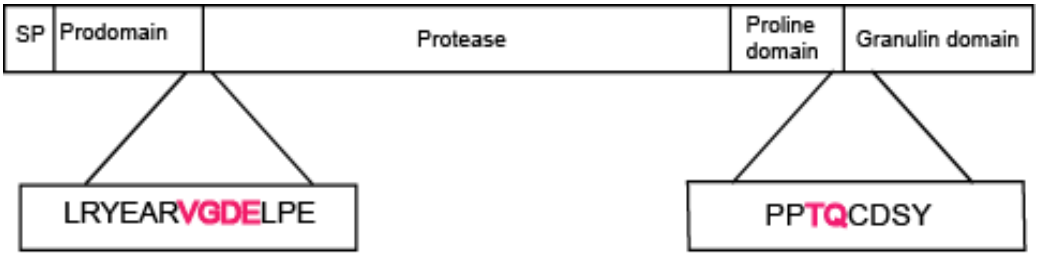


Figure 7 Schematic presentation of various processing sites of RD21

pMS21 ((R148A) prodomain cleavage mutant active was found to be as intermediate RD21.It was found that Arginine mutation to alanine did not effect to the activity and accumulation of the protein (Figure 5, lane 6).

Mutation in the beginning of proline rich region for checking the influence on the removal of granulin domain leads to complete loss of the activity and accumulation of the protein (pMS22) TQ-AA granulin cleavage mutant. This indicates that two amino acids (Threonine, glutamic acid to alanine) are critical for the maturation of intermediate RD21 to mature RD21 (Figure 4 lane 2).

The VGDE region of prodomain was randomly mutagenized into AAAA. In this approach, degenerate primers were used that were able to generate different combinations of alanine in place of VGDE sequence of protein as shown in figure 8.

V	G	D	E
A			
	A		
		A	
			A
A	A		
	A	A	
		A	A
A	A	A	
	A	A	A
A	A	A	A

We were able to generate one successful pMS24 (AGDE) mutant, in which valine has been changed to alanine. This resulted in the complete loss absence of RD21 with no accumulation of the protein in the cells. (Figure 4, lane 4)

Figure 8. Random mutagenesis of VGDE

RD21 accumulate into Endoplasmic reticulum

Immunolocalization studies have revealed that ER-bodies contain RD21 (15). To further investigate the localization of RD21, YFP-tagged fusion proteins were generated and expressed by agroinfiltration. These data confirm that RD21 localizes in mobile vesicles.

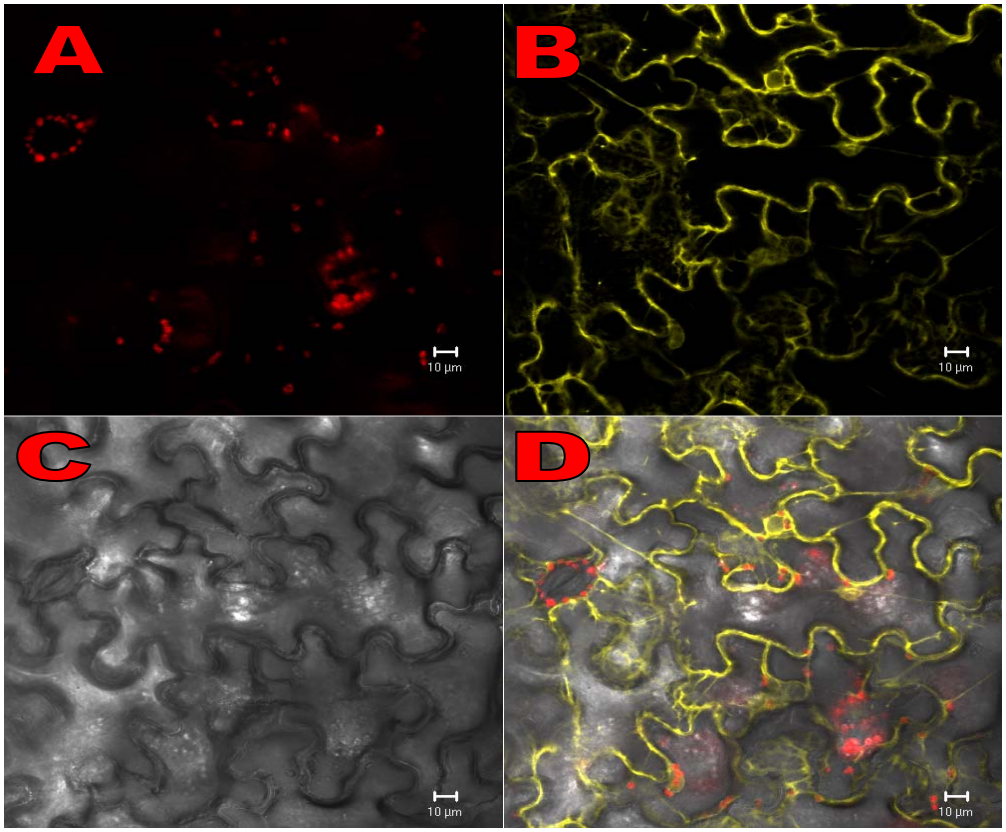


Figure 9 Localization of RD21-YFP under confocal microscope, the net like yellow structures (Figure 9B) shows the accumulation of RD21. Figure 9A shows the control photograph of the tissue, Figure 9C is the phase contrast image of the plant section. Image 9D shows the hybrid of fluorescence and phase-contrast image.

RD21-YFP gave very good images and clearly unravels the localization of protein inside the cell. Recent studies have shown that RD21 functions as a Protein Disulphide Isomerase in the endoplasmic reticum, which is very much consistent with our results (31).

CONCLUSION

In this study, RD21 was subjected to targeted mutagenesis and deletion analysis to study the role of various conserved features, such as its putative glycosylation and nitrosylation sites, catalytic residues, granulin domain and cleavage sites. Mutant RD21 proteins were overexpressed in tobacco by agroinfiltration and investigated by Western blot and activity-based profiling using biotinylated E-64 (DCG-04). This revealed, amongst others, that the granulin domain is required for prodomain removal. Putative roles of the granulin domain and the complicated activation mechanism of RD21 were discussed. The maturation pathways, processing of the RD21 and importance of various residues have been accounted

ACKNOWLEDGEMENTS

pRH628 and pRH629 construct was kindly provided by Dr. Renier von der Hoorn, Max-Plank Institute of Plant Breeding Research, Germany. I also thank to Takayuki for his valuable suggestions and to Dr.Izabella for helping in confocal microscopy.

REFERENCES

1. Matile, P. (1975) The lytic compartment of plant cells. In Cell Biology Monographs, Vol 1. Springer-Verlag, Wien, Austria.

2. Rogers, J.C., Dean, D., Heck, G.R. (1985) Proc. Natl. Acad. Sci. USA 82: 6512–6516.

3. Murakami, S., Hattori, T., Nakamura, K. (1986) Plant Mol. Biol. 7: 343–355.

4. Matsuoka, K., Nakamura, K. (1999) Plant Mol. Biol. 41: 825–835.

5. Chrispeels, M.J., Raikhel, N. (1992) Cell 68: 613–616.

6. Koizumi, M., Yamaguchi-Shinozaki, K., Tsuji, H., Shinozaki, K. (1993) Gene 129: 175–182

7. Bromme, D., Bonneau, P.R., Lachance, P., Wiederanders, B., Kirschke, H., Peters, C., Thomas, D.Y., Storer, A.C., Vernet, T. (1993) J. Biol. Chem. 268: 4832–4838

8. Karrer, K.M., Peiffer, S.L., DiTomas, M.E. (1993). Proc. Natl. Acad. Sci. USA 90: 3063–3067.

9. Gietl, C., Schmid, M., Simpson, D. (2000) Annual Plant Reviews, 5:90–111.

10. Watanabe, H., Abe, K., Emori, Y., Hosoyama, H., Arai, S. (1991) J. Biol. Chem. 266: 16897–16902.

11. Bateman, A., Bennett, H.P.J. (1998) J. Endocrinol. 158: 145–151.

12. Bhandari, V., Palfree, R.G.E., Bateman, A. (1992) Proc. Natl. Acad. Sci. USA 89: 1715–1719.

13. Hrabal, R., Chen, Z., James, S., Bennett, H.P., Ni, F. (1996) Nat. Struct. Biol. 3: 747–752.

14. Yamada, K., Matsushima, R., Nishimura, M., and Hara-Nishimura, I. (2001) Plant Physiol. 127: 1626-1634.

15. Hayashi, Y., Yamada, K., Shimada, T., Matsushima, R., Nishizawa, N. K., Nishimura, M., and Hara-Nishimura, I. (2001) Plant Cell Physiol. 42: 894-899.

16. Greene, J. J., and Rao, V. B. (1998) Recombinant DNA Principles and Methodologies. Marcel Dekker, Inc. New York.

17. Sambrook, J., Fritsch, E.F., and Maniatis, T.A. (1989). Molecular Cloning: A Laboratory Manual, 2nd ed. (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

18. Van der Hoorn, R. A. L., Leeuwenburgh, M. A., Bogyo, M. , Joosten, M. H. A. J., and Peck, S. C. (2004) Plant Physiol. 135 : 1170-1178.

19. Voinnet, O., Rivas, S., Mestre, P., and Baulcombe, D. (2003) Plant J. 33: 949-956.

20. Laemmli, U.K. (1970) Nature 227: 680-685.

21. Morrissey, J. H. (1981) Anal. Biochem.117: 307-310.

22. Twining, S. S. (1994) Crit. Rev. Biochem. Mol. Biol. 29: 315-383.

23. Goldberg, A.L. (2003) Nature 426: 895-899.

24. Hrabal, R., Chen, Z., James, S., Bennett, H.P.J. (1996) Nat. Struct. Biol. 3: 747–752.

25. Berti, P.J. and Storer, A.C. (1995) Journal of Molecular Biology 246: 273–283.

26. Stamler, J. S., Lamas, S., and Fang, F.C. (2001) Cell 106 : 675–683.

27. Vernet, T., Tessier, D.C., Chatellier, J., Ploufee, C., Lee, T., Thomas, D. Storer, A., Menard, R. (1995) J. Biol. Chem. 28 : 16645-16652

28. Karrer, K.M., Peiffer, S.L., DiTomas, M.E., (1993) Proc. Natl. Acad. Sci. USA 90: 3063–3067.

29. San Segundo, B., Martinez, M. C., Vilanova, M., Cuchillo, C. M. and Aviles, F. X. (1982) Biochim. Biophys. Acta 707: 74-80.

30. Evin, G., Devin, J., Castro, B., Menard, J. & Corvol, P. (1984) Proc. Natl. Acad. Sci. USA 81 : 48-52.

31. Christine, A., Ondzighi, D. A., Christopher, E. J. Shu-Choeng, C. and Staehelina, L. A. (2008) The Plant Cell 20: 2205–2220.

CHAPTER 5

RIP1-CHAGASIN LIKE CYSTEINE PROTEASES INHIBITOR FROM *PSEUDOMONAS SYRINGAE* DC3000, ITS BIOPHYSICAL STUDIES

SUMMARY

The conformational stability of RIP1 (**R**D21 **I**nhibiting **P**rotein1) was studied by denaturation at different temperatures. At 50ºC, a major change in the secondary structure was observed leading to a complete loss of activity. Thus, thermal inactivation of RIP1 takes place with major changes in the conformation of the protein. The change in both secondary and tertiary structure of protein leads to the complete loss of inhibitor activity. Comparison of binding of RIP1 separately with papain and cathepsin K was studied. RIP1 inhibits papain 1000 times more efficiently than it does with cathepsin K. The difference could be due to structural difference in P2 binding site of enzymes.

INTRODUCTION

Pseudomonas syringae is a rod shaped, gram-negative bacterium with polar flagella. It is a member of the *Pseudomonas* genus, and based on 16S rRNA analysis, *P. syringae* has been placed in the *P. syringae* group (1). *Pseudomonas syringae* pv. *tomato* DC3000 (*Pto* DC3000) is the cause of bacterial speck on tomato and *Arabidopsis*, and represents an important model in molecular plant pathology (2). It exists as over 50 different pathovars.

Individual *P. syringae* strains usually cause disease only in a limited set of plants. The host range of *P. syringae* strains appears to be controlled in part by the ability of the host to mount a cellular defence response (3, 4). In resistant plants, primary defence responses to *P. syringae* strains are usually rapid, commonly detectable within the first few hours of the interaction, and frequently culminate in programmed cell death (PCD) observable in the responding cells 6 hours after inoculation.

During the course of infection *Pseudomonas syringae* pv. *tomato* DC3000 secretes many effector molecules, which are responsible for pathogenicity. Effectors are pathogen molecules that manipulate host cell structure and function thereby facilitating infection and/or triggering defense responses. Effectors can be elicitors and/or toxins. Unlike these, the term effector is neutral and does not imply a negative or positive impact on the outcome of the disease interaction (5).

Some of the known effectors are enzyme inhibitors (6). The inhibitors are not only important in regulating endogenous protease activity but also many pathogens, including viruses, bacteria and parasites, use cysteine proteases to evade target organism defense mechanisms, which points to the general protective role of host inhibitors. In turn, the pathogens themselves use protease inhibitors to evade the host defense system. Proteinase inhibitors are specialized and diverse group of proteins that have several roles towards serine- and cysteine proteinases (7, 8). RIP1 (**R**D21 **I**nhibiting **P**rotein 1) is possible effector chagasin molecule, which inhibits RD21 (Kaschani and Von der Hoorn unpublished).

Cystatins are the most important inhibitory proteins for cysteine proteases. A search for similar regulators of parasite cysteine proteases resulted in the discovery of a novel inhibitor family distinct from cystatins and other known groups. Chagasin was the first described member of the group after which the family was so named. Chagasin is a *T. cruzi* inhibitor of the endogenous cysteine protease, Cruzipain (9). Subsequently, homologues were identified in the genomes of all mentioned parasites as well as in bacteria (10).

Chagasin is a single 110-residue polypeptide chain with no sequence similarity to other known groups of cysteine protease inhibitors. It potently inactivates papain-like enzymes through the formation of 1:1 tight–binding complexes. Chagasin-like inhibitors that occur in other protozoa and bacteria (10, 11) have been designated as ICP (Inhibitor of Cysteine Peptidases), and together comprise the chagasin family (Clan XI – Family I42).

The chagasin molecule is composed almost exclusively of β-strands and can be considered as a split β-barrel formed by two β-sheets, whose edges do not meet. Four β-strands, β1, β3, β6 and β5, on one side of the barrel, create a sheet with a predominately antiparallel architecture, in which only strand β1 is parallel with its neighbor (12).

The three-dimensional structure assumed by a protein, unique and suitable for its function, can in general be considered to be thermodynamically the most stable conformation adopted by the polypeptide chain. A polypeptide can also adopt a less rigid or more flexible conformation, different from its functional native form, responding to changes in environment. Thus, proteins are only marginally stable because of the functional requirement for their inherent dynamic state and due to delicate balancing of interactions involved in stabilizing or destabilizing particular structure. In this context, detailed denaturation studies of partially folded conformation of a protein or a family of proteins is important for understanding the general principles that govern protein folding pathways (13). By recording changes in

intrinsic tryptophan fluorescence and the secondary and tertiary structural features of protein in response to tailored changes in surroundings, one can establish presence of interesting structural intermediates relevant to structure-function relationship of the protein.

The conformational stability of proteins is studied by denaturation of protein using different agents such as temperature, pH, chaotropic agents and pressure, either individually or in combination of the denaturants. Denaturation is followed by monitoring the changes in any measurable physical property of the protein such as intrinsic fluorescence, CD spectra and corresponding changes in the protein function are checked by measuring its activity. The aim of this study is to understand the importance of a correctly folded structure and active site geometry of RIP1 with respect to its functional stability. In addition, kinetics and thermodynamics of binding of RIP1 with papain and cathepsin K were studied.

MATERIALS AND METHODS

Papain, cathepsin L, cathepsin K, urea, tris, dithiothreitol (DTT), guanidine hydrochloride (Gdn-HCl), N α-Benzoyl-DL-Arginine p-Nitroanilide (BAPNA), Phe-Arg-AMC, trifluoroacetate (TFA) and other chromogenic substrates were obtained from Sigma, USA. All the other reagents used were of analytical grades. Solutions prepared for spectroscopic measurement were in Milli Q^{TM} water.

Purification of Inhibitor

RIP1 from *Pseudomonas syringae* DC3000 was cloned and overexpressed in *E.coli* (Kaschani et al unpublished data). The protein was purified under denaturing conditions using 6M urea. Protease inhibitor cocktail was added for avoiding proteolysis of inhibitor during purification. The protein was stabilized with glycine, DDT and triton X-100. The homogeneity of the protein was determined using SDS-PAGE. Enzyme samples for spectroscopic measurement were centrifuged, filtered through a 0.45 mm filter and the exact pH and protein concentration were determined by Lowry's method before taking measurements.

Papain inhibition assay

The inhibitory activity of RIP1 against papain was determined by assaying the proteolytic activity of 50 µl of papain (1mg/ml) in Tris-HCL buffer, pH 6.5 in the presence of 10 mM DDT and 10 mM EDTA, using BAPNA (1.5 mM) as the substrate in the presence and absence of RIP1.

Assay for inhibitory activity of RIP1 towards Cathepsin K

The inhibitory activity against Cathepsin K was determined by assaying the proteolytic activity of 5 µl cathepsin K (0.1mg/ml) in Acetate buffer, pH 5.5, in the presence 10 mM DTT and 10 mM EDTA using Phe-Arg-AMC as the substrate in the presence and absence of RIP1. The inhibition assay was performed in a 96 well assay plate. The fluorescence of liberated AMC was measured (excitation-360 nm / emission-460 nm).

Initial kinetic analysis for the determination of K_m and K_i

The kinetic parameters for substrate hydrolysis were determined by measuring the initial rate of enzymatic activity. The inhibition constant *K*i was determined by the Lineweaver-Burk equation. The K_i value was also calculated from the double reciprocal equation by fitting the data into the computer software Origin 6.1. For the Lineweaver-Burk, analysis, papain (1µM) was incubated with 100 nM inhibitor and assayed at increasing concentration of BAPNA (0.1-5 mM) at 37^0C for 30 min. The reciprocals of substrate hydrolysis (1/V) for inhibitor concentration were plotted against the reciprocal of the substrate concentration, and the K_i was determined by fitting the resulting data.

Ligand –binding thermodynamics data analysis

Free energy changes of inhibitory constant (ΔG) was determined by the equation;

$$\Delta G = - RT \ln K_a \qquad [1]$$

Temperature dependence of the inhibitor constants was used to determine the thermodynamic parameters. Changes in enthalpy (ΔH) were determined from the

Van't Hoff plots by using the equation;

$$\ln K_a = (-\Delta H/RT) + \Delta S/R \qquad [2]$$

Where R is gas constant, ΔS is entropy change and T is the absolute temperature. The entropy change was obtained from the equation:

$$\Delta G = \Delta H - T\Delta S \qquad [3]$$

The assay was done at different temperatures calculating various K_i of papain and cathepsin K with RIP1.

Fluorescence Measurement

Intrinsic fluorescence of RIP1 was measured using a Perkin-Elmer Luminescence spectrometer LS50B connected to a Julabo F20 circulating water bath. The protein solution of 0.2 mM was excited at 280 nm and the emission was recorded in the range of wavelengths 300 – 400 nm at 30°C. The slitwidths for both the excitation and emission were set at 5 nm, and the spectra were recorded at 100 nm/min. To eliminate the background emission the signal produced by either buffer solution, or buffer containing the appropriate quantity of denaturants was subtracted. Fluorescence spectra were recorded as described above. After taking the scans suitable aliquots were removed from the samples to determine the residual activity.

Thermal Denaturation

Effect of temperature on RIP1 was studied using a thermostatic cuvette holder connected to an external constant temperature circulation water bath. The protein sample (0.2 mM) was incubated for 15 min at specified temperature before taking the scan. The appropriate aliquots were removed from the sample at every specified condition for measuring the residual activity under standard assay conditions. Fluorescence spectra were recorded as described above.

Circular Dichroism Measurement

The CD spectra were recorded on a J-715 spectropolarimeter with a PTC343 Peltier unit (Jasco, Tokyo, Japan) 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 eight scans at 100 nm/min with a 1nm slit width and a time constant of 1s for a nominal resolution of 1 nm. 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. The tertiary structure of the inhibitor (6 mM) was monitored with near UV CD spectra in the wavelength 250 – 300 nm using path length 1 cm.

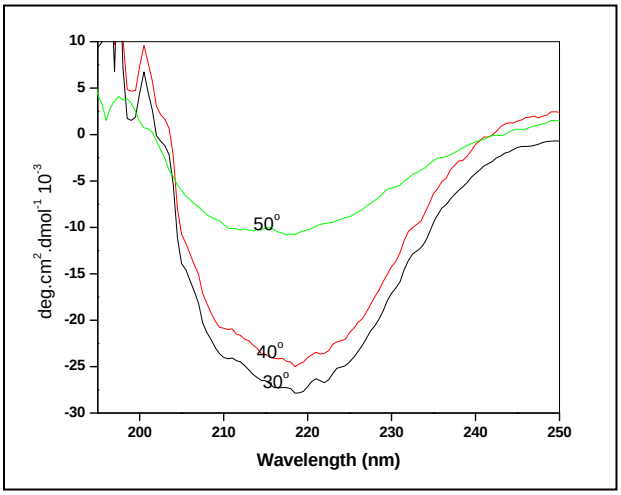
RESULTS AND DISCUSSION

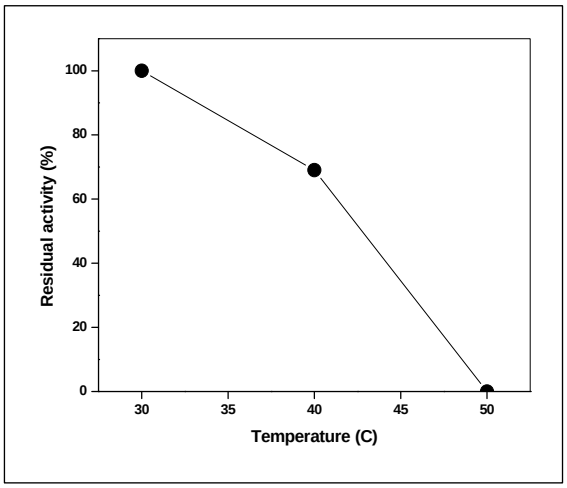
Thermal Denaturation of RIP1

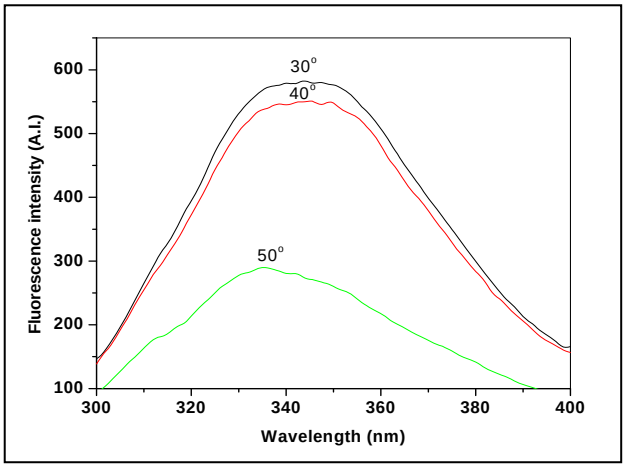
The intrinsic fluorescence of the protein due to the presence Trp residues was monitored at different temperatures. The spectrum of the native inhibitor shows maximum fluorescence intensity between 340-348nm indicating several populations of Trp, differentially exposed to the polar environment (Fig. 1C). The fluorescence intensity of the inhibitor gradually decreased with increasing temperature. This must be due to the deactivation of the singlet-excited state by non-radiative processes. The observed λ_{max} (340-348 nm) remained same (340-348 nm) until 40ºC, while above 50ºC it showed a blue shift to 335 nm presumably due to Trp burying in the interior (Fig. 1C).

There was no loss of inhibitory activity up to 40ºC, but at 50 ºC there was a drastic loss of inhibitory activity of RIP1 (Fig. 1B). The ratio of fluorescence intensity (F340/348 nm), 0.970 at 30ºC increase to that of 1.055 at 50ºC indicating slight changes in the Trp environment with increasing temperature.

A. **B.**







C.

Figure 1 Thermal denaturation of RIP1.

(A) Fluorescence spectra of protein (0.2 mM) incubated at different temperatures for 15 min.

(B) Plot of activity (●) against temperature

(C) Far UV CD spectra of RIP1 at 30°, 40° and 50°C

0. 2 mM protein was used, each sample was incubated for 15 min at the respective temperature and then scans were recorded.

The far UV CD spectra of the inhibitor recorded at different temperatures are shown in Fig. 1C. Raising the temperature up to 40ºC did not lead to any significant loss in activity. At 50ºC, major loss in the negative ellipticity was observed, that leads to complete loss of activity. Thus, thermal inactivation of RIP1 takes place with major changes in the secondary and tertiary structure of the protein. The loss in the secondary structure could be leading to the movement of Trp residues in the hydrophobic environment.

Initial kinetic analysis for the determination of K_i

The Lineweaver-Burk reciprocal plot (Fig.2, 3) shows that both the enzymes were competitively inhibited by RIP1. The inhibition constant K_i, was determined by the classical double reciprocal plot (Fig.1, 2). The Ki was found to be $3.98*10^{-9}$ for papain with RIP1 and $2.91*10^{-6}$ for cathepsin K with RIP1.

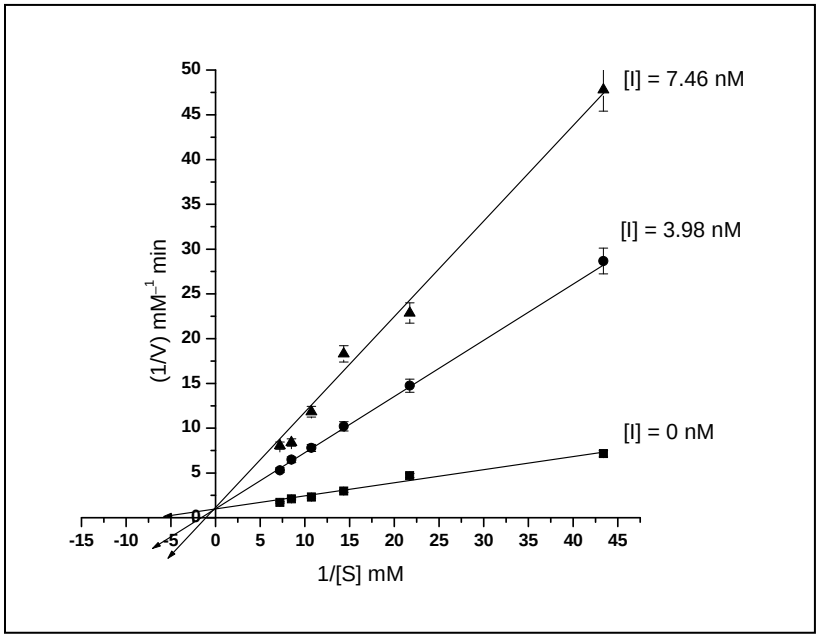


Figure 2 Papain (1mg/ml) was incubated, without (■) or with RIP1 at 7.46 nM (●,▲) concentrations and assayed at increasing concentrations of the substrate for determining K_i *value* from Lineweaver- Burk plot at 37°C. (K_i value=3.98nM)

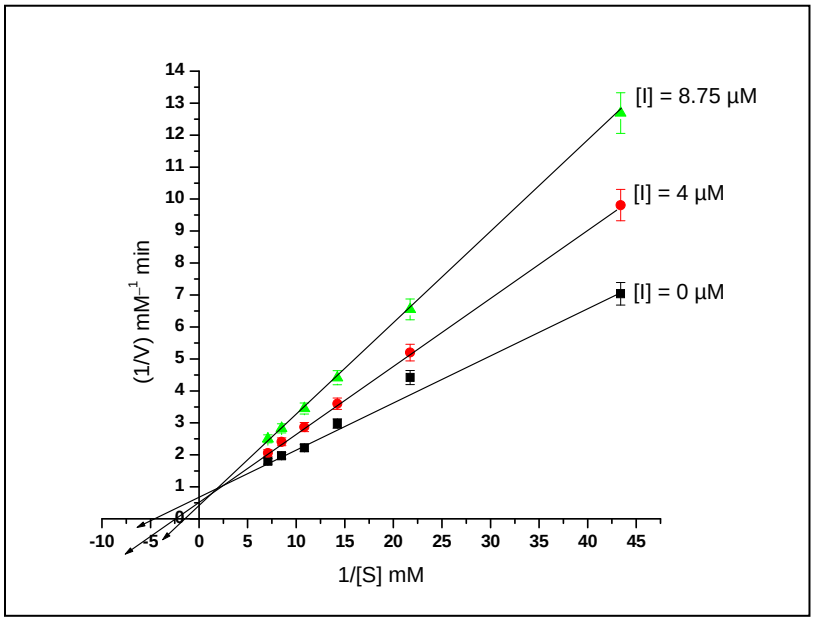


Figure 3 Cathepsin K (0.1mg/ml) was incubated, without (■) or with RIP1 at 8.75 µM (●,▲) concentrations and assayed at increasing concentrations of the substrate for determining K_i *value* from Lineweaver- Burk plot at 25°C. (K_i value=2.91µM)

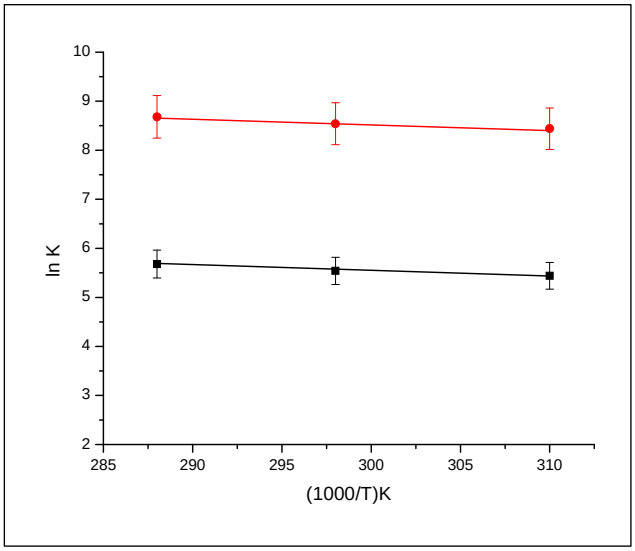


Figure 4 Van't Hoff plots for the inhibitory constant of RIP1 to papain and cathepsin K are drawn according to the regression equation. The symbols used are: (●) papain and (■) cathepsin K. The values of K_i determined by tube assay are listed in Table 1.

Ligand –binding thermodynamics data analysis

The thermodynamic parameters viz. free energy (ΔG), enthalpy (ΔH) and entropy (ΔS) of binding calculated by determining K_i at different temperatures (Table 1). The van't Hoff plots were linear ($r > 0.9$) for both the enzymes in the temperature range studied (Fig.4).

Free energy of the binding (ΔG) was negative at all the temperatures, suggesting the spontaneous nature of the binding. Enthalpy change (ΔH) was negative in both the cases but the entropy (ΔS) change of binding was positive for cathepsin K and negative for papain. It reflects that binding of papain and RIP1 is exothermic and interaction is mostly hydrophilic and polar. In case of binding between cathepsin K and RIP1, it was hydrophobic, exothermic and mostly non-polar interactions (Table 1).

The inhibitory constant (K_i) increases when there is increase in temperature in both the cases viz. papain and cathepsin K. The value of K_i with papain (K_i = 2.13 nM) was 1000 fold higher than that of cathepsin K (K_i = 2.08μM). The higher affinity of papain can be explained in terms of more favorable entropy for RIP1-papain interaction. In addition, papain is plant protein and RIP1 is from bacterium *Pseudomonas syringae*. It can be that this putative effector having high affinity for plant proteases due to similarity with its natural target in host system. The value of enthalpy was almost same in both the interactions (RIP1-papain, RIP1-cathepsin K).

Negative values of ΔS for papain show the greater involvement of the polar interactions in the protein-ligand binding. The data can be well explained in terms of structural difference between both the enzymes (McGrath et al. 1997). Cathepsin K having the overall positive charge that is different from the papain, which is more or less negatively charged molecule. The presence of Leu at 205 in cathepsin K creates a pocket that is too shallow as compare to papain, due to this cathepsin K does not favor Phe in P2 subsite. So overall, there is difference in binding sites of these two enzymes.

Negative ΔH and ΔS values also indicate primarily intermolecular hydrogen bonding and increased polar interactions between papain–RIP1 interactions. The overall free Gibbs energy (ΔG) is more in papain-RIP1 interaction as compared to cathepsin K interaction. This is reflected in binding of RIP1 with papain.

Table 1: Inhibitory constant and thermodynamic parameters for the binding of enzymes to the inhibitor at different temperatures

Enzyme	K_i (M) Temperatures 15	25	37	ΔH kJ.mol^{-1}	ΔG kJ.mol^{-1}	ΔS J.mol^{-1}.K^{-1}
Papain	2.13*10^{-9}	3.28*10^{-9}	3.98*10^{-9}	-20.18	-21.02	- 2.83
Cathepsin K	2.08*10^{-6}	2.91*10^{-6}	3.64*10^{-6}	-20.17	-13.71	21.67

*Values at 25 ^{0}C

Conclusion

Our studies on the thermal stability of RIP1 show that protein is inactivated at 50 ºC due to major changes in the protein structure. The denaturation take place step-wise where the in beginning the protein is unfolded gradually and then there is sharp loss in the three dimensional structure of protein. This leads to complete loss of the inhibitory activity of RIP1. Previous observations with cystatins in the thermal denaturation event showed that a quick drop in the negative ellipticity, reflecting significant lose of the protein native structure, appeared around 60ºC (16). The biological significance of the labile nature of its protein structure may be due to nature of functional environment and regulation of inhibiting properties of this protein.

The kinetic parameters showed that RIP1 was competitive inhibitor of the both enzymes. Moreover, there are about 1000 fold difference in Ki values of RIP1 with papain and cathepsin K. The temperature dependence of the kinetics parameters of RIP1 binding with cathepsin K and papain is quite similar. The thermodynamics parameters for binding of RIP1 with papain and cathepsin K shows this interaction are entropy driven process. Binding is in fact associated with essentially no change in enthalpy.

Acknowledgment:

RIP1 construct was kindly provided by Dr. Farnusch Kaschani and Dr. Renier Van der Hoorn, Max-Plank Institute of Plant Breeding Research, Germany.

REFERENCES

1. Anzai, Y., Kim, H., Park, J.-Y., Wakabayashi, H. & Oyaizu, H. (2000) Int. J. Syst. Evol. Microbiol. 50: 1563–1589.
2. Alfano, J. R., and A. Collmer. (1996) Plant Cell 8:1683-1698.
3. Hutcheson, S.W. (1998) Annu. Rev. Phytopathol. 36: 59–90.
4. Dangl, J., and Jones, J. (2001) Nature 411: 826–833.
5. Sophien Kamoun (2006) Annu. Rev. Phytopathol. 44:41–60.
6. Tian, M., Huitema, E., Da Cunha, L., Torto-Alalibo, T., Kamoun, S. (2004) J. Biol. Chem. 279:26370–77.
7. Barrett, A. J. (1987) Trend Biochem. Sci. 12 : 193-196.
8. Turk, V. and Bode, W. (1991) FEBS Lett. 285 : 213-219.
9. Monteiro, A. C. S., Abrahamson, M., Lima, A. P. C. A., Vannier- Santos, M. A. and Scharfstein, J. (2001) J. Cell Sci. 114: 3933–3942.
10. Rigden, D. J., Mosolov, V. V. and Galperin, M. Y. (2002) Prot. Sci. 11: 1971–1977.
11. Sanderson, S.J., Westrop, J., Scharfstein, J., Mottram, J.C. and Coombs, G.H. (2003). FEBS Lett. 542: 12–16.
12. Ljunggren, A., Redzynia, I., Alvarez-Fernandez, M., Abrahamson, M., Mort, J., Krupa, J., Jaskolski, M. and Bujacz and G. J. Mol. Biol. (2007) 371: 137–153.
13. Kelly, J. W., and Balch, W. E. (2006) Nat. Chem. Biol. 2: 224 – 227.
14. Dr. Farnusch Kaschani and Dr. Renier Van der Hoorn (unpublished data).
15. McGrath, M. E. (1999) Annu. Rev. Biophys. Biochem. 28: 181-204.
16. Elzbieta, J., Wars*a*w, W., Zbigniew, G. (2004) Eur. Biophys. J. 33: 454–461.

CHAPTER 6
Part- 1

ISOLATION AND IDENTIFICATION OF PLANT PATHOGENIC *Penicillium sp. SDBF1*

SUMMARY

The pathogenic fungus was isolated in the laboratory from the mangrove plant *Avicennia marina*. The culture was purified by the single colony plating technique on MGYP agar plates. On nutrient agar at 28°C, the mycelia were large, irregular, sticky and cream colored. The conidium looks blue to blue-green, and the mold sometimes exudes a yellow pigment.

The strain was an aerobic, spore forming fungus. The organism was salt tolerant with a broad growth temperature range of 28°C-50°C with optimum growth at 28°C and pH 7.The isolated organism was identified to be *Penicillium sp. SDBF1* based on internal transcribed spacer/5.8S ribosomal DNA (rDNA) sequencing. The sequence has been submitted to NCBI gene bank, under the accession number FJ403589.

INTRODUCTION

Mangrove forests are generally distributed in tropical and subtropical regions of the world and several mangrove species are valuable source of useful metabolites for medicinal usage (1). One of the reason for the dominance of mangrove plants may be due to fungal endophytes and pathogens associated with host plants. In fact, fungi from mangroves are the second largest group among the marine fungi (2). A large variety of new bioactive compounds have recently been isolated from the marine fungi, especially the mangrove fungi (3,4). In our search for secondary metabolites of mangrove pathogenic fungi from India, we have isolated a novel cysteine protease inhibitor. This was purified from an unidentified pathogenic fungus isolated from the leaf of *Avicennia marina*.

Fungal rRNA genes are tandemly repeated, with each repeat encoding 18S (small-subunit), 5.8S, and 26S (large-subunit) genes. Two other regions exist in each repeat: the internal transcribed spacer (ITS) region and the intergenic spacer (IGS) region (Fig. 1). Although rRNA genes are highly conserved, the ITS regions are divergent and distinctive (5, 6). Ribosomal DNA (rDNA) has been widely utilized for molecular systematics and the identification of microorganisms. The D1/D2 regions of 26S and ITS sequences have been used mainly to identify pathogenic fungi. The analysis of ITS sequences has been carried out mainly for pathogenic yeast and fungal species (7, 8). Peterson & Kurtzman and Sugita *et. al.* demonstrated that a single species showed less than 1% dissimilarity in either the ITS region or D1/D2 26S rDNA. Therefore, IGS and ITS sequences analysis appears to be a powerful tool for identification of fungal strains.

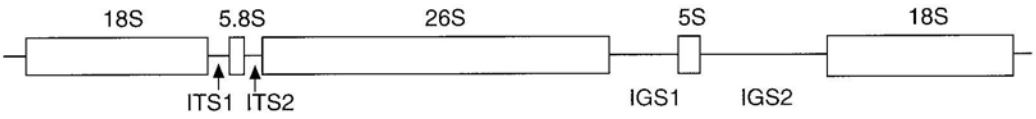


Figure. 1. Schematic representation of the rDNA locus in *Penicillium sp.* Boxes indicate coding regions.

Penicillium chrysogenum is a filamentous, spore-forming, obligate aerobe belonging to the mold group of fungus. They are widely distributed in nature and often found living on foods and in indoor environments. It was previously known as *Penicillium notatum*. It is the source of several β-lactam antibiotics, most significantly penicillin. Other secondary metabolites of *P. chrysogenum* include various different penicillins, roquefortine C, meleagrin, chrysogine, xanthocillins, secalonic acids, sorrentanone, sorbicillin and PR-toxin (9).

Like many other species of the genus *Penicillium*, *P. chrysogenum* reproduces by forming dry chains of spores (or conidia) from brush-shaped conidiophores. The conidia are typically carried by air currents to new colonization sites. In *P. chrysogenum* the conidia are blue to blue-green, and the mold sometimes exudes a yellow pigment. However, *P. chrysogenum* cannot be identified based on color alone. Observations of morphology and microscopic features are needed to confirm its identity. *P. chrysogenum* has been used industrially to produce penicillin and xanthocillin X, to treat pulp mill waste, to produce the enzymes polyamine oxidase, phospho-gluconate dehydrogenase, glucose oxidase, proteases and amylases (10).Various species of *Penicillium* are used in the production of different kinds of cheese.

Penicillium species are usually regarded as unimportant in terms of causing human disease. *Penicillium marneffei*, discovered in 1956, is different. This is the only known thermally dimorphic species of *Penicillium* and it can cause a lethal systemic infection (penicilliosis) with fever and anaemia similar to disseminated cryptococcosis. The airborne spores of *P. chrysogenum* are important human allergens. Vacuolar and alkaline serine proteases have been implicated as the major allergenic proteins (11). *Penicillium expansum*, a widespread filamentous fungus, is a major causative agent of fruit decay and may lead to the production of mycotoxin that causes harmful effects on human health. *Penicillium aurantiogriseum* is a plant pathogen.

This part of the present chapter deals with the isolation and identification of a salt tolerant *Penicillium chrysogenum* plant pathogenic strain that produces a cysteine protease inhibitor.

MATERIALS AND METHODS

MATERIALS

Lysozyme, SDS, EDTA, chloroform, ethanol, agarose, ethidium bromide, proteinase K, RNAase, deoxyribonucleoside triphosphate (dNTP) and *Taq* DNA polymerase were from Sigma Chem. Co. (USA). 18S rDNA primers were commercially obtained from Life technologies, (India). Activated charcoal, NaCl, malt extract, glucose, yeast extract, peptone, ultra marine salt, Tris and agar were purchased from Himedia and Qualigens (India). All other chemicals were of analytical grade.

Isolation and identification of the microorganism

The organism was isolated in the laboratory from the mangrove plant *Avicennia marina*. The plant leaves were collected from Mandovi estuaries, Goa, India .The culture was purified by the single colony plating technique on MGYP agar plates. The organism was identified by carrying out internal transcribed spacer/5.8S ribosomal DNA (rDNA) sequencing. The internal transcribed spacer/5.8S ribosomal DNA (rDNA) sequence was blasted at NCBI.

Isolation of genomic DNA

The fungal culture was grown in 100 ml of MGYP broth (pH 7.0) at 37^{o}C for 7 days with continuous shaking at 160 rpm. Genomic DNA of the fungal culture was extracted by using the salting out method (12). 5 g fungal mycelia were ground into a powder using liquid nitrogen and suspended in 5 ml SET buffer (75mM NaCl, 25mM EDTA (pH 8.0) and 20 mM Tris (pH 7.5)). 100 µL lysozyme (final conc. 1mg/ml) was added to the above solution and incubated at 37^{o}C for 60 min. 140 µL of proteinase K (final conc. 0.5 mg/ml) and 600 µL of 10% SDS were added, and incubated for 2 hrs at 55^{o}C with occasional mixing. 2 mL of 5M NaCl (final conc. 1.25M) was added and the mixture was cooled to 37^{o}C. 5mL chloroform was added and mixed for about 30 min at room temperature and then the mixture was centrifuged for 20 min, at 4,500 x g. The supernatant was transferred into a fresh tube and 0.6 vol of isopropanol was added to precipitate the DNA. The DNA was pelleted and washed twice with 70% ethanol, air dried and dissolved in 1-2 mL of TE

buffer (10 mM Tris, 1mM EDTA, pH 8.0) at 28ºC. The quantification of the DNA was done on 0.8% agarose gel stained with ethidium bromide.

PCR Amplification of 18S rDNA

The primers used for the identification of fungal sp. were universal primers for fungal amplification: ITS1 (5'–TCC GTA GGT GAA CCT GCG G-3'), which hybridizes at the end of 18S rDNA, and ITS4 (5' –TCC TCC GCT TAT TGA TAT GC -3') which hybridizes at the beginning of 28S rDNA (13).The 25µl Polymerase chain reaction (PCR) was set using the genomic DNA. The reaction mixture typically contained genomic DNA 50 ng, 10X Buffer 2.50µl, 0.2mM deoxyribonucleoside triphosphate (dNTP) 2.5µl, forward and reverse primers 10- 20pmoles, (1.25µl each), distilled water 16.87µl, and 1unit of Taq DNA polymerase 0.25µl (Bangalore Genei). All the additions were done on ice and the PCR reaction was performed on Gene Amplifier PCR System 9700 (Perkin Elmer, USA). The PCR conditions for 18S rDNA gene amplification were: Initial denaturation 95°C for 3min, followed by 35cycles of 94°C for 1min., 55°C for 1min, 72°C for 1min. and final extension at 72°C for 10 min. 5µl of the above PCR amplified product was used for electrophoresis using 1.0% agarose gel in 1X TBE buffer (Working solution: 0.5X; Stock: 5X, 54 g Tris base, 27.5 g boric acid, 20 ml 0.5M EDTA [pH 8]). The gel was run at 80V for 90min, using 1X TBE as running buffer. The gel was stained in 1% ethidium bromide for 45 min. and was observed under UV illumination.

Purification of PCR amplified product

10 µl of the above PCR amplified product was used for electrophoresis using 1.0% agarose gel in 1X TBE buffer (Working solution: 0.5X; Stock: 5X, 54g Tris base, 27.5g boric acid, 20ml 0.5M EDTA [pH 8]). The gel was run at 80V for 90min, using 1X TBE as running buffer. The gel was observed under UV illumination and the DNA band (450 bp) was cut and purified with Qiagen agarose gel purification kit. Purity of DNA was checked again on gel.

Sequencing of the purified PCR product

The sequencing reactions were carried out using Taq DNA polymerase dye terminator cycle sequencing using the 'ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit' (Perkin Elmer Applied Biosystems Division, Foster City, CA) according to the manufacturer's protocol. This Kit contains the four dNTPs with different fluorescence labels termed as BigDye Terminators. 2 µl PCR product and 3 pmol of the sequencing primer were used in a 20 µl sequencing reaction. The sequencing reaction mixes were subjected to 25 cycles in a Perkin Elmer thermal cycler 9700. Each cycle consisted of 95°C for 10min, 50°C for 5min and 60°C for 4min. DNA sequencing was carried out on ABI 3700 DNA sequencer. The analysis of nucleotide sequence was done at NCBI server using BLAST-n (www.ncbi.ncm.gov/blast).

Growth conditions of the microorganism

The isolate was maintained on MGYP- agar slants and preserved at 4°C. The pHs of the slants were adjusted to pH 7.0. The fungal culture was grown at 37°C for 7 days. The media composition for the growth of the organism was:

Malt extract	:	0.3%
Glucose	:	1.0%
Yeast extract	:	0.3%
Peptone	:	0.5%
Ultra marine salt	:	1.1%

The cultivation of the organism was carried out by inoculating a loop full of freshly grown culture on MGYP-agar slant into 500 ml Erlenmeyer flasks containing 100 ml of the above medium and was incubated under shaking condition (160 rpm) at 37°C for 7 days.

RESULTS

Identification of the microorganism

The isolated organism was identified to be ***Penicillium* sp. *SDBF1*** based on the biochemical characteristics as described in the Bergey's Manual of Determinative Bacteriology. On nutrient agar at 28°C, the mycelia were large, irregular, sticky and white colored (Figure 1). The strain was an aerobic, spore forming fungus (Figure 2). The organism was salt-tolerant showing growth at broad range of salt concentration.

Identification of the microorganism by18S rDNA sequencing

The genomic DNA was isolated and checked on agarose gel, was found to be of high molecular weight and intact. The spectrophotometeric analysis of the DNA showed that the DNA had an A 260: A 280 ratio of 2.0, which indicates the purity of the DNA preparation. Figure 3 shows the 18S rDNA sequence from *Penicillium* sp. On NCBI BLAST, this sequence showed closest homology (99.5%) to *Penicillium chrysogenum*. The blast hits have been shown in Figure 4. It revealed that the isolate of *Penicillium* sp. had maximum similarity to *Penicillium chrysogenum*

Optimum growth conditions for the organism

The optimum fermentation condition for *Penicillium sp. SDBF1* was found to be 28°C and pH 7.0 for 7 days. The growth of the organism in fermentation media was measured by taking O.D. at 600 nm.

Figure. 1. Mycelium was large, irregular, sticky and white colored from *Penicillium* sp. *SDBF1*

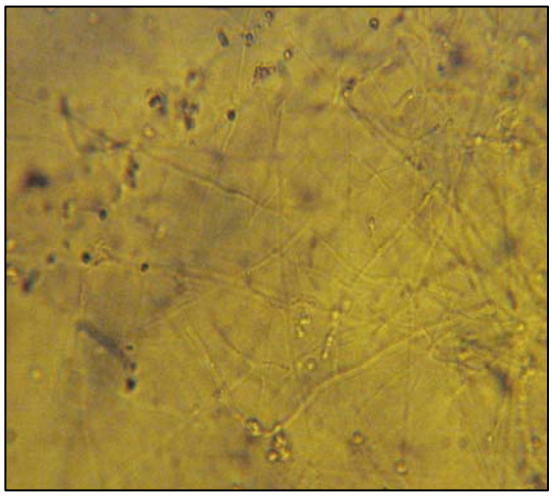


Figure. 2. The strain was an aerobic, spore forming fungus *Penicillium* sp. *SDBF1*

CGGGAACTACCTGATCCGAGGTCACCTGGATAAAAATTTGGGTTGATCGGCAAG
CGCCGGCCGGGCCTACAGAGCGGGTGACAAAGCCCCATACGCTCGAGGACCGGA
CGCGGTGCCGCCGCTGCCTTTCGGGCCCGTCCCCCGGGATCGGAGGACGGGGCC
CAACACACAAGCCGTGCTTGAGGGCAGCAATGACGCTCGGACAGGCATGCCCCC
CGGAATACCAGGGGGCGCAATGTGCGTTCAAAGACTCGATGATTCACTGAATTT
GCAATTCACATTACGTATCGCATTTCGCTGCGTTCTTCATCGATGCCGGAACCAA
GAGATCCGTTGTTGAAAGTTTTAAATAATTTATATTTTCACTCAGACTACAATCTT
CAGACAGAGTTCGAGGGTGTCTTCGGCGGGCGCGGGCCCGGGGGCGTAAGCCCC
CCGGCGGCCAGTTAAGGCGGGCCCGCCGAAGCAACAAGGTAAAATAAACACGG
GTGGGAGGTTGGACCCAGAGGGCCCTCACTCGGTA

Fig. 3. 18S rDNA sequence of *Penicillium* sp. *SDBF1*.

Sequences producing significant alignments:

Accession	Description	Max score	Total score	Query coverage	E value	Max ident
FJ176474.1	*Penicillium sp*. MX9 18S small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 28S large subunit ribosomal RNA gene, partial sequence	950	950	98%	0.0	99%
EU497955.1	*Penicillium sp*. F23 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	950	950	98%	0.0	99%
EU497943.1	*Penicillium sp*. 36 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	950	950	98%	0.0	99%
EU128598.1	*Penicillium chrysogenum* strain 32P 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and	950	950	98%	0.0	99%

	28S ribosomal RNA gene, partial sequence					
EU128591.1	*Penicillium chrysogenum* strain 16P 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	950	950	98%	0.0	99%
EU139860.1	*Penicillium sp*. M1100 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	950	950	98%	0.0	99%
EU139854.1	*Penicillium sp*. M1203 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	950	950	98%	0.0	99%
EU139852.1	*Penicillium sp*. M1135 18S ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	950	950	98%	0.0	99%
DQ681334.1	*Penicillium granulatum* isolate 732 internal transcribed spacer 1, partial sequence; 5.8S ribosomal RNA gene and internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial sequence	950	950	98%	0.0	99%
AY373902.1	*Penicillium chrysogenum* strain FRR 807 18S ribosomal RNA gene, partial sequence; ITS 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and 28S ribosomal RNA gene, partial seq	950	950	98%	0.0	99%

```
Query  1
TACCGAGTGAGGGCCCTCTGGGTCCAACCTCCCACCCGTGTTTATTTTACCTTGTTGCTT
60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  53
TACCGAGTGAGGGCCCTCTGGGTCCAACCTCCCACCCGTGTTTATTTTACCTTGTTGCTT
112

Query  61  cggcgggcccgccttaactggccgccgggggggcttacgccccgggcccgcgcccgccgA  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  113
CGGCGGGCCCGCCTTAACTGGCCGCCGGGGGGCTTACGCCCCCGGGCCCGCGCCCGCCGA
172

Query  121
AGACACCCTCGAACTCTGTCTGAAGATTGTAGTCTGAGTGAAAATATAAATTATTTAAAA
180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  173
AGACACCCTCGAACTCTGTCTGAAGATTGTAGTCTGAGTGAAAATATAAATTATTTAAAA
232

Query  181
CTTTCAACAACGGATCTCTTGGTTCCGGCATCGATGAAGAACGCAGCGAAATGCGATACG
240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  233
CTTTCAACAACGGATCTCTTGGTTCCGGCATCGATGAAGAACGCAGCGAAATGCGATACG
292

Query  241
TAATGTGAATTGCAAATTCAGTGAATCATCGAGTCTTTGAACGCACATTGCGCCCCCTGG
300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  293
TAATGTGAATTGCAAATTCAGTGAATCATCGAGTCTTTGAACGCACATTGCGCCCCCTGG
352

Query  301
TATTCCGGGGGGCATGCCTGTCCGAGCGTCATTGCTGCCCTCAAGCACGGCTTGTGTGTT
360
              |||||||||||||||||||||||||||| ||||||||||||||||||||||
Sbjct  353
TATTCCGGGGGGCATGCCTGTCCGAGCGTCATTTCTGCCCTCAAGCACGGCTTGTGTGTT
412

Query  361
GGGCCCCGTCCTCCGATCCCGGGGGACGGGCCCGAAAGGCAGCGGCGGCACCGCGTCCG
G  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  413
GGGCCCCGTCCTCCGATCCCGGGGGACGGGCCCGAAAGGCAGCGGCGGCACCGCGTCCG
G  472
```

```
Query  421
TCCTCGAGCGTATGGGGCTTTGTCACCCGCTCTGTAGGCCCGGCCGGCGCTTGCCGATCA
480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  473
TCCTCGAGCGTATGGGGCTTTGTCACCCGCTCTGTAGGCCCGGCCGGCGCTTGCCGATCA
532

Query  481  ACCCAAATTTTTATCCAGGTGACCTCGGATCAGGTAG  517
             |||||||||||||||||||||||||||||||||||
Sbjct  533  ACCCAAATTTTTATCCAGGTGACCTCGGATCAGGTAG  569
```

Fig. 4. BLAST result hits obtained from NCBI for *Penicillium* sp. 18SrDNA sequence.

REFERENCES

1. Kathiresan, K. and Bingham, B.L. (2001) Adv. Mar. Biol. 40: 81-251.
2. Zhong-shan, C., Wen-cheng, T., Zhi-jian, S., Cai, Y., Shi-feng S., Qi-jin, C., Wang, F., Lin, Y., She, Z. and Vrijmoed, L.L.P. (2008) J. For. Res.19: 219-224.
3. Lin, Y., Wu, W., Feng, S., Jiang, G., Jiang, G., Zhou, S. Vrijmoed, L.L.P. and Jones, E.B.G. (2001) Tetrahedron Lett. 42: 449-451.
4. Lin, Y., Wu, W., Feng, S., Jiang, G., Jiang, G., Zhou, S. Vrijmoed, L.L.P. and Jones, E.B.G. (2001) J. Org. Chem. 66: 6252-6255.
5. Chen, Y. C., Eisner, J. D., Kattar M.M., Rassoulian-Barret, S. L., Lafe, K., Yarfitz, S.L., Limaye, A.P. and Cookson, B.T. (2000) J. Clin. Microbiol. 38: 2302–2310.
6. Accensi, F., Cano, L., Figuera, M. L., Abarca, and Caban˜es, F. J. (1999) FEMS Microbiol. Lett. 180**:**191–196.
7. Turenne, C. Y., Sanche, S. E., Hoban, D. J., Karlowsky, J. A. and Kabani, A. M. (1999) J. Clin. Microbiol. 37 : 1846–1851.
8. Sugita, T., Nishikawa, A., Ikeda, R. and Shinoda, T. (1999) J. Clin. Microbiol. 37: 1985–1993.
9. De Hoog, G.S., Guarro, J., Gené, J. and Figueras, F. (2000) Atlas of Clinical Fungi - 2nd Edition, Centraal bureau voor Schimmel cultures (Utrecht).
10. Raper, K.B. and Thom, C. (1949) A manual of the Penicillia, Williams & Wilkins Company (Baltimore).
11. Shen, H.D., Chou, H., Tam, M.F., Chang, C.Y., Lai, H.Y. and Wang, S.R. (2003) Allergy 58: 993–1002.
12. Neumann, B., Pospiech, A. and Schairrer, H.U. (1992) Trends Genet. 8: 332-333.
13. White, T. J., Bruns, T., Lee, S., and Tailor, S. (1990) PCR protocols. A guide to methods and applications. Academic Press, Inc., San Diego, Calif. 315–322.

CHAPTER 6
Part- 2

PURIFICATION AND PARTIAL CHARACTERIZATION OF CPI FROM *Penicillium* sp. *SDBF1*

SUMMARY

The indispensable nature of the cysteine proteases in numerous physiological functions has evoked tremendous interest towards isolating new inhibitors from various resources. After extensive screening, a salt tolerant *Penicillium sp. SDBF1* was isolated, which produces a cysteine protease inhibitor (CPI). The organism produced CPI in the peptone, ultra marine salt complex media at 28°C and at pH 7. The extracellular culture filtrate of *Penicillium sp. SDBF1* was subjected to ultra filtration and gel filtration to remove the high molecular weight impurities and salts. The gel filtration fractions showing anti papain activity were concentrated and loaded on manually packed C-18 column for reverse phase-FPLC. The anti papain activity was associated with the peak A, having a retention time of 27.726 min and other eluted peaks showed no inhibitory activity. The active fraction was lyophilized and loaded on reverse phase-HPLC. Homogeneity of the active fraction was indicated by the single peak as analyzed on rp-HPLC and mass spectrometry (ESI-MS) showed a Mr of 437 Da.

CPI was stable in a broad range of pH (2-11) and temperature (37-80°C). CPI was stable for 2 hours and 30 minutes at 90°C and 100°C respectively. A very high inhibitory activity of CPI was recorded against papain and no inhibition was observed for other cysteine proteases like cathepsin K and cathepsin L. The inhibitor did not show any inhibitory activity against other classes of the proteases like trypsin, chymotrypsin, pepsin and clostripain.

INTRODUCTION

Cysteine proteases are important therapeutic targets because of their role in several diseases, such as cathepsin B in tumour growth and cathepsins K in osteoporosis. Calpains are also involved in a variety of disease states, such as Alzheimer's disease, multiple sclerosis, muscular dystrophy, type 2 diabetes, traumatic brain and spinal cord injury, and cerebral ischemia (1-2). Recent studies showed that they play significant role in plant pathogen interactions (3-4). Therefore, there is continuous demand for new chemical scaffolds for cysteine protease inhibitors.

Nature continues to be one of the most significant sources of pharmacologically active compounds in the quest for drugs. Cysteine protease inhibitors are reported from almost all sources like bacteria (5), fungi (6), plants (7) and animals (5). Marine natural products provide a rich source of chemical diversity that can be used to design and develop new, potentially useful therapeutic agents. Large varieties of new bioactive compounds have recently been isolated from the marine fungi, especially from mangrove fungi (8).

The enzyme inhibitors serve as probes for kinetic and chemical processes during catalysis that led to a detailed understanding of enzyme catalytic mechanisms. They help in elucidating the mode of ligand binding, where the ligand may be an inhibitor, substrate, or substrate analogue (9). Inhibitors are also used for *in vivo* studies to localize and quantify enzymes in organs or to mimic certain genetic diseases that involve the absence of an enzyme in a given biosynthetic pathway**.** The other uses of inhibitors are the detection of short-lived enzyme-bound reaction intermediates, or the identification of amino acid residues at the active site that are necessary for the catalytic activity of the enzyme.

Traditional approaches for development of inhibitors have involved natural product screening or replacing the scissile bond of a peptide substrate with a non-cleavable isostere. More recently, peptidomimetic research, which has the ultimate goal of developing inhibitors with improved pharmacokinetic properties, has gained importance (10-11).

Although in the past decade, the rational drug discovery has taken the movement away from the development of enzyme inhibitors by screening of natural products, the biodiversity widespread in soil, water, insects, tropical plants and marine sources still have tremendous potential for the isolation of novel and effective enzyme inhibitors.

There have been plethora of reports of synthetic cysteine protease inhibitors however, there is a paucity of reports of such inhibitors from microorganisms. Furthermore, to

our knowledge, there are no reports of peptidic inhibitors of cysteine proteases from salt tolerant microorganisms.

We have isolated a salt tolerant *Penicillium sp. SDBF1*, which produces a cysteine protease inhibitor, CPI. This part of the present chapter deals with the purification and biochemical characterization of the inhibitor.

MATERIALS AND METHODS

Materials

Papain, N α-Benzoyl-DL-Arginine p-Nitroanilide (BAPNA), Trifluoroacetate (TFA) and other chromogenic substrates were purchased from Sigma Chem. Co. (USA). Acetonitrile was purchased from E-Merck (Germany). NaCl, malt extract, glucose, yeast extract, peptone, ultra marine salt, Tris and agar were purchased from Himedia and Qualigens (India). All other chemicals were of analytical grade. Reverse phase matrix for FPLC (Fast protein liquid chromatography) was from GE healthcare (India).

Methods

Production of the Cysteine Protease Inhibitor

The media composition for the production of Cysteine Protease Inhibitor (CPI) was as given below:

Malt extract	:	0.3%
Glucose	:	0 .1%
Yeast extract	:	0.3%
Peptone	:	0.5%
Ultra marine salt	:	1.1%

The pH of the medium was adjusted to 7.0 before autoclaving. Production of the CPI was carried out by inoculating a loop full of freshly grown culture from above

mentioned agar slants into 500 ml Erlenmeyer flasks containing 100 ml of the above-mentioned broth medium and was incubated under shaking condition (160 rpm) at 28^{o}C for 7 days.

Assay for inhibitory activity of CPI towards papain

The inhibitory activity of CPI against papain was determined by assaying the proteolytic activity of 50 µl of papain (1mg/ml) in Tris-HCL buffer, pH 6.5 in the presence of 10 mM DDT and 10 mM EDTA, using BAPNA (1.5 mM) as the substrate in the presence and absence of CPI.

Time course for the production of CPI

The production of inhibitor at various time intervals was checked by removing samples at 24 hrs. time intervals and assayed for its anti-papain activity. The time course of production of CPI was estimated in terms of percent inhibition against papain activity.

Purification of CPI

The extracellular culture filtrate of the isolate was obtained by centrifugation of growth medium constituents at 5000 rpm, at 4^{o}C for 20 minutes. Further CPI was purified by ultrafiltration (UM-30 and UM-3) and gel filtration (Sephadex LH-20) to remove high molecular weight impurities. The fractions were eluted on isocratic mode with H_2O at a flow rate of 0.5 ml/min and monitored at a wavelength of 280 nm. Then CPI was purified by RP-FPLC, by using rp-column prepared by polystyrene/divinyl benzene monodisperse beads (15µm diameter). The fractions were eluted on a linear gradient of 0-95% acetonitrile with H_2O containing 0.05% TFA at a flow rate of 0.5 ml/min and monitored at wavelength 220 nm. The active fractions obtained from RP-FPLC were pooled, lyophilized and loaded on RP-HPLC using prepacked µBondaPak column. The fractions were eluted on a linear gradient of 0-95% acetonitrile with H_2O containing 0.05% TFA at a flow rate of 0.5 ml/min and monitored on dual wavelengths 220 nm and 280 nm. The eluted sample was lyophilized and dissolved in distilled water to check the anti-papain activity.

Biochemical characterization of CPI

The molecular mass of purified CPI was determined by qudrupole electrospray ionization mass spectrometer. Methanol–water (1:1) system was used as mobile phase. The purity of inhibitor was also checked on UPLC-MS (ultra pressure liquid chromatography) system coupled with mass spectrophotometer system (WATERS). Acquity UPLC BEh C18 column (2.1x 50 mm) was used for analysis. The run was for 8 min in isocratic mode and CPI showed a retention time of 3.2 min.

Assay for inhibitory activity of CPI towards trypsin and Chymotrypsin

The inhibitory activity of CPI against Chymotrypsin and trypsin was determined by assaying the proteolytic activity of 25 µl each of Chymotrypsin and trypsin (1mg/ml) in phosphate buffer, pH 7.0, in the presence of 10 mM EDTA, and using casein (1%) as the substrate. A blank reaction was performed in the absence of CPI.

Assay for inhibitory activity of CPI towards Cathepsin L and K

The inhibitory activity against Cathepsin L and K was determined by assaying the proteolytic activity of 5 µl cathepsin L (0.1mg/ml) in Acetate buffer, pH 5.5, in the presence 10 mM DTT and 10 mM EDTA using Phe-Arg-AMC as the substrate in the presence and absence of CPI. The inhibition assay was performed in a 96 well assay plate. The fluorescence of AMC was measured (excitation /emission; 360 nm / 460 nm).

Temperature and pH stability of CPI

The temperature stability of CPI was determined by incubating CPI (50 µl) at temperatures from 30-100^{o}C for 6 hours and by estimating anti-papain activity at different time intervals. pH stability of CPI was determined by incubation of CPI at different range of pH values in appropriate buffers for 1 hours and its inhibitory activity against papain was estimated.

Preparation of intracellular fraction for antimicrobial activity

Fungal mycelia were washed twice with sterile distilled water, to remove the extracellular impurities. They were frozen in liquid nitrogen and then crushed using a mortar and pestle. Water extract was centrifuged at 5000 rpm for 10 min and supernatant was used for disk diffusion assay.

Anti microbial activity

Microbes *Serratia marcescens* (NCIM no 2078); *Bacillus circulas* (NCIM No.2160) were obtained from National Collection of Industrial Microorganisms (NCIM), NCL, Pune. The antibacterial activities of the obtained fractions were evaluated by the agar disk diffusion technique.

RESULTS AND DISCUSSION

Production of cysteine protease inhibitor (CPI)

The maximum production of CPI was observed from 4 to 7 days of growth using a fermentation media containing malt extract (0.3%), glucose (1%), yeast extract (0.3%), peptone (0.5%) and ultra marine salt (1.1%). The pH was adjusted to 6.8-7.2 before autoclaving. Figure 1 shows time profile for the production of CPI by the organism. There was a sharp decrease in inhibitor production after 7 days.

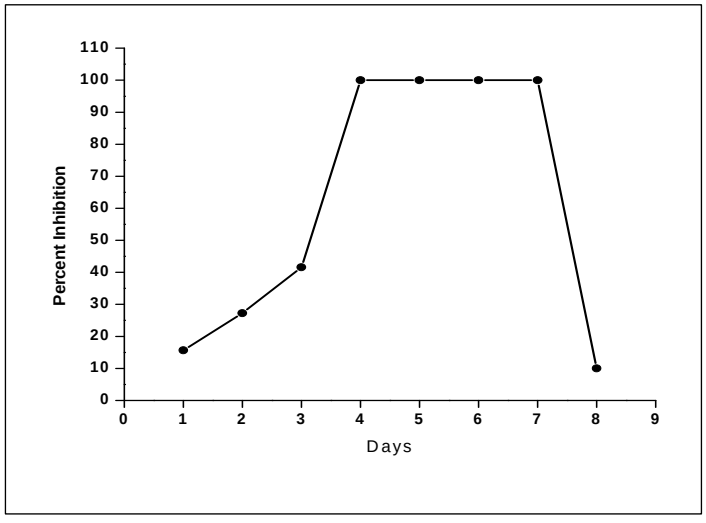


Figure 1. Time profile for the production of CPI.

Ultra marine salt (1.1%) is essential for growth of the fungus *Penicillium chrysogenum*. The production of CPI was also achieved without shaking the fungus; even in the stationary growth condition CPI level was same in the flasks.

Secondary metabolites are compounds produced by an organism that are not required for primary metabolic processes (12). Fungi produce an enormous array of secondary metabolites, some of which are important in industry. They are generally produced in lag phase of the growth of fungus (13). Production of metabolites also depends on the nutrient environment of the fungus. We have standardized the time and nutrient requirements of *Penicillium sp. SDBF1* for maximum production of CPI.

Purification of CPI

The extracellular culture filtrate of *Penicillium chrysogenum* was subjected to centrifugation and ultra filtration (30 KDa cut off and 3 KDa cut off). The filtrate was concentrated and loaded on gel filtration column (Sephadex LH-20). The gel filtration fractions 58-80 showing papain inhibitory activity (Figure 2) were pooled, concentrated and loaded on RP-FPLC column (prepared by polystyrene/divinyl benzene monodisperse beads of 15 μm diameter).

The RP-FPLC fractions 15-16 showing papain inhibitory activity (Figure 3) were concentrated and loaded in μBondaPak prepacked column for RP-HPLC. The anti papain activity was associated with the peak A (labeled as active fraction in figure 4A), having a retention time of 14 min (Figure 4A) and some other eluted peak with differing retention time. The fractions showing the inhibitory activity were pooled and lyophilized. Homogeneity of the active fractions containing peak A was indicated by the single peak as analyzed on RP-HPLC with retention time 14 min (Figure 4B).

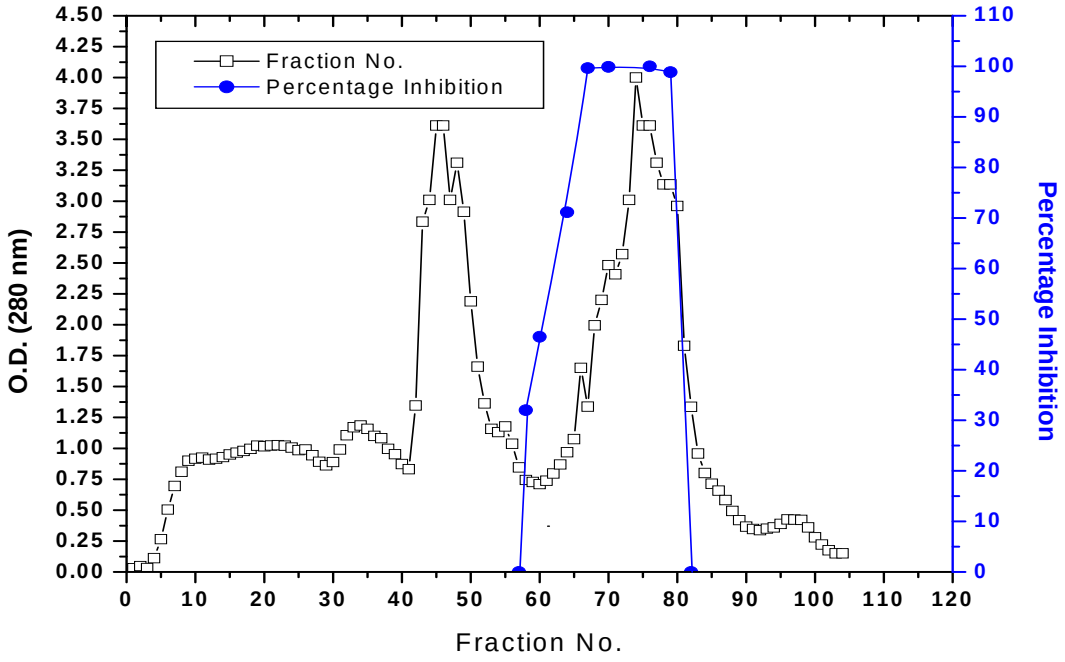


Figure 2. Chromatographic profile of inhibitor on a Sephadex LH-20 matrix column. The column was equilibrated with milli Q water.

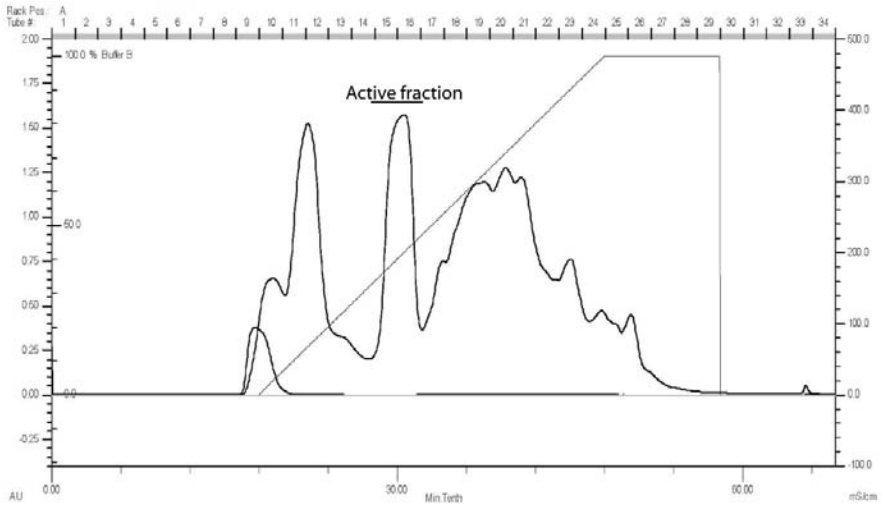


Figure 3. Reverse phase-FPLC purification of CPI. 1ml of the lyophilized CPI sample was loaded on rp-FPLC column prepared by polystyrene/divinyl benzene monodisperse beads of 15µm diameter, pre equilibrated with 5 % acetonitrile (CH_3CN) and 0.05 %trifluoroacetate (TFA).

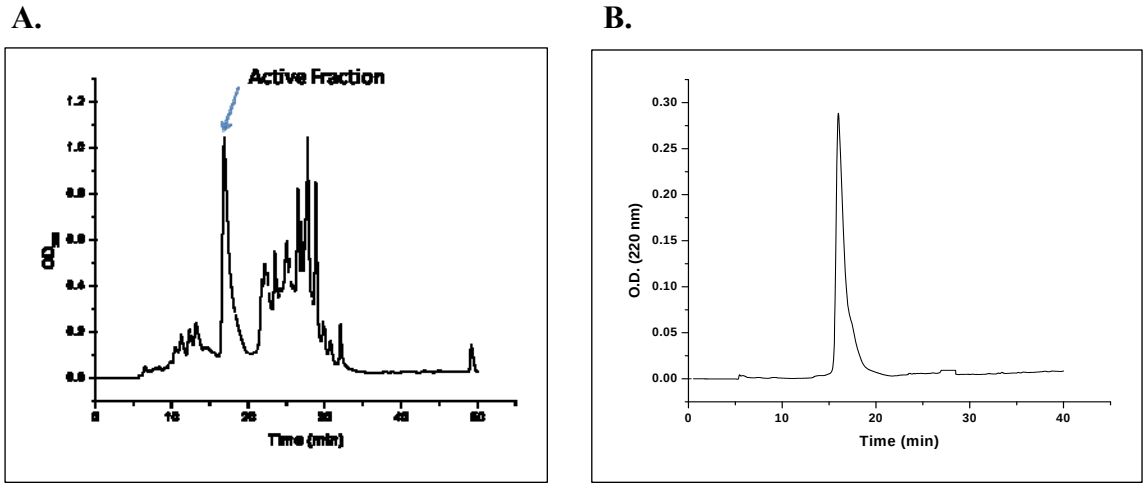


Figure 4. Reverse phase-HPLC purification of CPI.

(A) 40µl of the lyophilized CPI sample was loaded on prepacked µBondaPak column (waters RP-C18) pre equilibrated with acetonitrile (CH_3CN) and trifluoroacetate (TFA). The fractions containing the peaks A were collected manually and assayed for the anti-papain activity.

(B) 50µl of the pooled fractions containing the peak A (associated with the anti papain activity) was reloaded onto the reverse-phase HPLC system under similar experimental conditions. The peak detected showed a retention time of 15.76 min.

Biochemical characterization of CPI

The chromatographic profile of UPLC and HPLC shows that the CPI is a purified compound with almost no impurity. We had used diode array detector for scanning the compounds present in sample, only single peak of CPI was detected (Figure 5). The molecular mass of the inhibitor was found to be 437.312 Da. The mass spectra also shows fragmentation pattern of CPI. The compound was fragmented into small peaks of 139.084, 150.949, 151.101, 183.196 and 197.114. (Figure 6).

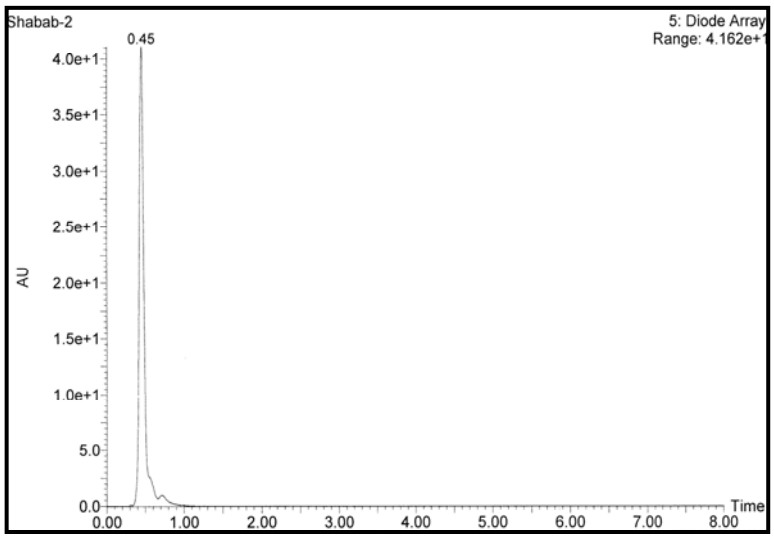


Figure 5. Reverse phase-UPLC profile of CPI. 0.5 ml of the lyophilized CPI sample was loaded on rp-UPLC C18 column. The inhibitor was monitored on diode array detector.

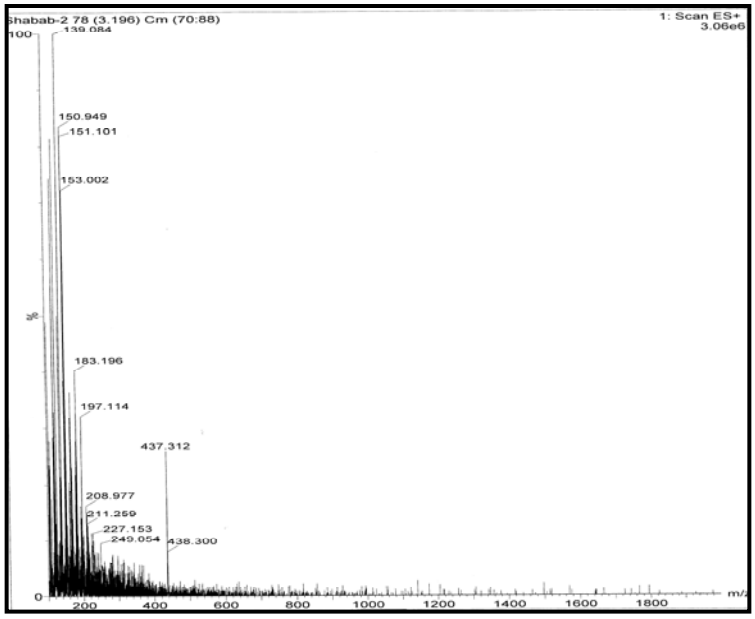


Figure 6. Molecular mass of CPI. The purified CPI was analyzed for the determination of the Mr by electrospray ionization spectra.

Secondary metabolites are generally produced following active growth, and many have an unusual chemical structure (14). Some metabolites are found in a range of related fungi, while others are only found in one or a few species. The restricted distribution implies a lack of general function of secondary metabolites in fungi (15).

Temperature and pH stability of CPI

For the temperature stability experiments, CPI (20 µL) was incubated at temperatures ranging from 25-100°C for 6 hrs and its inhibitory activity was estimated against papain at different intervals of time (Table 1). CPI (25 µL) was treated with buffers of different pH (ranging from 2 to 10) for 2 hrs. No loss in activity of CPI was observed in all pH ranges, indicating the stability of CPI in different ionic environments.

Table 1. Temperature stability of CPI

Temperature (^{o}C) / Time (hrs.)	50	60	100
1	99.64	99.84	99.74
2	100	99.72	99.65
3	100	100	100
4	100	100	99.79
5	100	99.88	100
6	100	100	100

CPI was also tested for inhibition against other classes of enzymes like pepsin (aspartic protease), trypsin and chymotrypsin (serine protease). It did not show any inhibition activity for theses enzymes, indicating that CPI is a cysteine protease inhibitor. Inhibition against other cysteine proteases like cathepsin L and K was also investigated but CPI did not show any inhibition against them.

Anti microbial activity

Fungi are prolific producers of structurally diverse, biologically active secondary metabolites (16). High levels of environmental stress and intense and frequent interactions with other organisms promote the production of metabolically diverse compounds (17). The salt tolerant plant pathogenic fungi are an underexploited group of significant taxonomic diversity and represent one of the largest untapped pools of novel fungi. There are few reports of biologically active metabolites from this class of fungi, like cyclo-(N-MeVal-N-MeAla) from endogenous fungus of *Avicennia marina* (18). Intracellular fraction of *P. sdbf1* was found to contain potent antibacterial activity against human pathogen *Serratia marcescens* and *Bacillus circulans*.

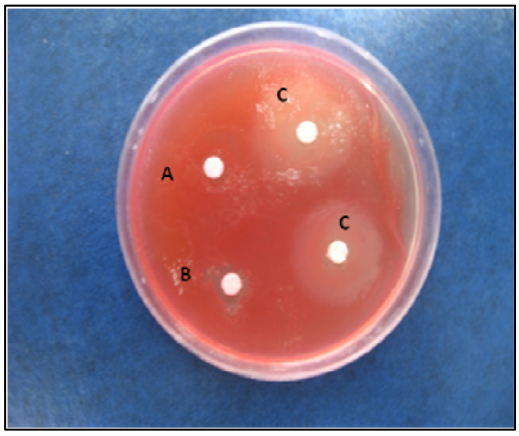


Figure 7. Zone of growth inhibition was observed by intracellular fraction *P. sdbf1* against *Serratia marcescens*.

A: Sterile Distilled water
B: Ampicillin control
C: Intracellular fraction

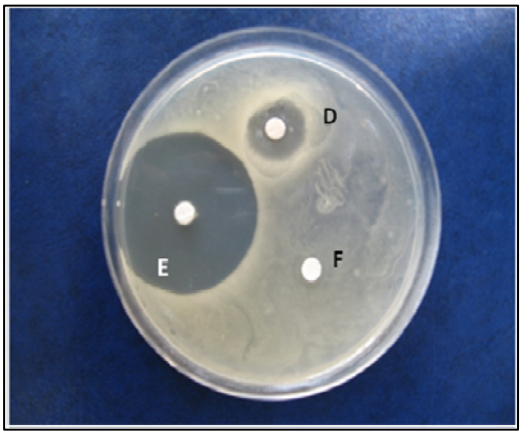


Figure 8. Zone of growth inhibition was observed by intracellular fraction of *P. sdbf1* against *Bacillus circulans*.

D: Intracellular fraction
E: Ampicillin control
F: Sterile Distilled water

Serratia marcesccens is a species of gram-negative rod shaped bacteria in the family enterobacteriaceae. It is involved in nosocomial infections; particularly catheter associated bacterimia, urinary tract infections and wound infections (19). Most *S. marcescens* strains are resistant to several antibiotics because of the presence of R factors which is a type of plasmid that carry one or more genes that encode resistance proteins; all are considered intrinsically resistant to ampicillin macrolides and first generation cephalosporins (19). *Bacillus circulans* are gram-positive, aerobic, spore-forming rods. *Bacillus circulans* is used for producing L-glutamic acid. Although not a major pathogen, *Bacillus circulans* has been reported to cause infections (20).

CONCLUSION

After extensive screening, a salt tolerant fungus *Penicillium sp. SDBF1* was isolated, which produces a cysteine protease inhibitor (CPI). The organism produced CPI in the peptone ultra marine salt complex media at 28°C and pH 7. This culture was isolated from leaves of *Avicennia marina* where it resides as a plant pathogen. The plant leaves were collected from Mandovi estuaries, Goa, India. Sequencing of 18S ribosomal RNA gene and 5.8S ribosomal RNA gene identified fungus as *Penicillium sp.* and was named as *Penicillium sp. SDBF1*. The sequences were submitted to NCBI database under the accession number FJ403589.

Low molecular weight cysteine protease inhibitor was purified from the extracellular broth of *Penicillium sp. SDBF1*, with a molecular weight of 437.312. The inhibitor was found to be specific for papain and did not show any inhibition against trypsin, chymotrypsin and pepsin as well as other cysteine proteases produced such as cathepsin L and K. Intracellular fraction of *P. sdbf1* was found to contain potent antibacterial activity against human pathogen *Serratia marcescens* and *Bacillus circulans*.

REFERENCES

1. Lecaille, F., Kaleta, J., Bromme, D. (2002) Chem. Rev. 102: 4459-4488.
2. Jedeszko, C., Sloane, B. F. (2004) Biol. Chem. 385: 1017-1027.
3. Tian, M., Win, J., Van der Hoorn, R., Van der Knaap, E., and Kamoun, S. (2007) Plant Physiol. 143: 364–377.
4. Shabab, M., Shindo, T., Gu, C., Kaschani, F., Pansuriya, T., Chintha, R., Harzen, A., Colby, T., Kamoun, S., van der Hoorn, R.A.L. (2008) The Plant Cell 20: 1169-1183.
5. Rzychon, M., Sabat, A., Kosowska, K., Potempa, J. and Dubin, A. (2003) Mol. Microbiol. 49: 1051–1066.
6. Chao-Mei, Y., Jonathan, M., Curtis, A., Walter, J., Wright, C., Stephen, W., Jadwiga, K., Lori, Q., and Ziba, F.A. (1996) The J. of Anti. 49: 395-397.
7. Laskowski, M. and Kato, I. (1980) Annu. Rev. Biochem. 49: 593–626.
8. Kathiresan, K. and Bingham, B.L. (2001) Adv. Mar. Biol. 40: 81-251.
9. Hiratake, J. (2005) Chem Rec. 5: 209-28.
10. Ripka, A. S. and Rich, D. H. (1998) Cur. Opin. Chem. Biol. 2: 441–452.
11. Hruby, V.J. (2001) Acc. Chem. Res. 34: 389-393.
12. Vining, L.C. (1990) Annu. Rev. of Micro-biol. 44: 395 – 427.
13. Yu, J.H. and Keller, N. (2005) Annu. Rev. of Phyto. 43: 437-458.
14. Mann, J. (1986) Secondary Metabolism.OUP, Oxford.
15. Yu, J-H. and Keller, N. (2005) Annu. Rev. of Phyto. 43: 437-458.
16. Henkel, T., Brunne, R. M., Muller, H. and Reichel, F. (1999) Angew. Chem., Int. Ed. 38: 643-647.
17. Dreyfuss, M. M., Chapela, I. H. In *The Discovery of Natural Products with Therapeutic Potential*; Gullo, V. P., Ed.; Butterworth-Heinemann: Boston, MA, 1994: 49-80.
18. Wang, S.-Y., Xu, Z.-L., She, Z.-G., Wang, H., Li, C-R and Lin, Y.-C. (2008) Journal of Asian natural products research 10:625-629
19. Hejazi, A., Falkiner, F.R. (1997) J. Med. Microbiol. 46: 903–12.
20. Skerman, V.B.D., McGowan, V., and Sneath, P.H.A. (editors). Approved lists of Int. J. Syst. Bacteriol. (1980) 30:225-420.

CHAPTER 7

GENERAL DISCUSSION AND CONCLUSION

General discussion and Conclusion

DISCUSION

Papain family cysteine proteases are key factors in the pathogenesis of microbial infections, cancer invasion, arthritis and osteoporosis. Targeting this enzyme family is one of the strategies in the development of new chemotherapy for a number of diseases. The disparate nature of parasite cysteine proteases compared to the host orthologous proteins has opened novel opportunities for chemotherapy. There is emerging evidence that the proteolytic machinery of plants plays important roles in defence against pathogens (1). Plant proteases being involved in defence has come to knowledge with the identification of RCR3, a secreted cysteine protease that is required for the function of the resistance gene Cf-2 (for *Cladosporium fulvum* resistance-2) and CDR1, a secreted aspartic protease that regulates defence responses (2). The role of cysteine proteases and their inhibitors in this context is discussed.

Indications of roles of plant proteases in defence came from observations that subtilisin-like proteases (P69 [for 69-kDa PR protein]) accumulate in viroid-infected tomato plants (3) and that acidic leucine aminopeptidase (LapA) activity increases during insect feeding. Association between the induction of protease genes and defence have also been found for genes that encode metallo, aspartic and cysteine proteases (5).

The biotrophic fungus *Cladosporium fulvum* is the causal agent of leaf mold of tomato (*Lycopersicon* sp).The effector protein AVR2 from *Cladosporium fulvum* has been shown to inhibit RCR3, a cysteine protease in tomato plant (6). AVR2 is an unclassified inhibitor, it does not share any homology with known cysteine protease inhibitors (6). We used a biochemical approach to investigate which PLCPs (papain like cysteine proteases) are active in the tomato apoplast during the benzothiadiazole (BTH)-triggered defence response and to determine the extent to which these PLCPs are inhibited by AVR2.

We built the phylogenetic map of these identified proteins with other known cysteine proteases from plants. Cat B1 and Cat B2 have come under one clade, PIP1 and RCR3 show very high sequence homology and the rest are in different clades of our phylogenetic map. It was shown that **P**hytophthora **I**nhibited **P**rotease 1 (PIP1) interacts and inhibits EPIC2B (7). In our studies we were able to found one more biologically relevant interaction of AVR2 and PIP1. This is an important interaction for studying effector–host protein. AVR2-PIP1 binding is pH dependent and takes place only in acidic conditions. Inhibition is indeed very potent, only nano molar concentration of AVR2 is enough for inhibition.

The region encoding the protease domain of various ecotypes of tomato was sequenced. The ratio between nonsimilar and similar amino acids indicated that PIP1 and RCR3 are under diversifying selection and rest of the identified proteases are under conservative selection. We generated structural models for studying the variations of PIP1 and RCR3 residues. A naturally occurring N194D mutation in RCR3 affects inhibition by AVR2. These data demonstrate that the variance of a single amino acid can affect the inhibition by AVR2. It is just the beginning of understanding the molecular interaction within host-pathogens by the role of protease-inhibitor interaction.

Phytophthora infestans is an oomycete that causes the serious potato disease known as late blight. The organism can also infect tomatoes and some other members of the Solanaceae. The effects of *Phytophthora infestans* in Ireland in 1845-57 were one of the factors, which caused over one million to starve to death and forced another two million to emigrate from affected countries. Most commonly referenced is the Great Irish Famine, during the late 1840s. *P. infestans* is still a difficult disease to control today by ordinary methods.

When mined, currently available expressed sequence tags (ESTs) and whole-genome shotgun sequences of *P. infestans* (8, 9) identify genes encoding putative extracellular proteins (10) and similarity and protein motif searches to annotate the predicted

extracellular proteins. Four of the sequences corresponded to cDNAs and one to a random genomic sequence. These new proteins are identified as EPICs (**E**xtracellular **P**rotease **I**nhibitor having **C**ystatin like domain).

We screened the EPIC1 and EPIC2B for inhibition against different cysteine proteases of tomato. Selective inhibition in apoplastic fluids showed that C14 (close homolog of RD21 (11)) is inhibited by this protein. But the role of this protein was not known. C14 was originally found to accumulate rapidly in various stress responses (11). Previously it was shown that a potato cysteine protease (*cyp*) cDNA expressed at an early stage of an incompatible interaction with *Phytophthora infestans.* Both the nucleotide and deduced amino acid sequences are highly homologous to those of a tomato cysteine protease, C14 (12).This gave the clue that C14 could function in plant defence. Inhibition of C14 by EPIC1 and EPIC2B is more or less similar in all the parameters. The binding between theses inhibitors to C14 is very potent and pH having no role in this. This data clearly shows the contrasting selectivity of fungal and oomycete effectors targeting tomato cysteine proteases.

RD21 (Responsive to Desiccation-21) is a papain like cysteine protease in *Arabidopsis thaliana.* An *rd21* gene encoding a papain family protease, RD21 (D13043), was found to be up regulated during dehydration of *Arabidopsis* plants (13). It has a unique maturation mechanism in senescing leaves (14). An intermediate of RD21 was accumulated as an aggregate in the vacuoles, and slowly matured during leaf senescence. Interestingly the intermediate RD21 (iRD21) and mature RD21 (mRD21) both are active. Recently it was shown that RD21 is required for proper seed development and regulates the timing of PCD (programmed cell death) by chaperoning and inhibiting Cys proteases during their trafficking to vacuoles before PCD of the endothelial cells (15). Due to all these novel properties, we were interested in structure-function analysis of this protein.

We generated various constructs of RD21 for point and deletion mutants. Various domains of the protein were deleted for their role and significance on protein. The

granulin-like domain may serve to regulate thiol protease activity in plants (16). The exact function of granulin domain in plant proteins is still unknown (17). At least our data shows that it is not essential for survival of the cells, may be it functions as an interacting domain for proteins under stress conditions. The destabilization of this domain leads to accumulation of only the intermediate RD21 form of protein. Deletion of the granulin domain does not affect the activity, processing and maturation of protein. But if we delete the proline-rich region + granulin domain together, then protein was not accumulate. It seems that proline domain is required for proper folding and processing of protein.

Lysosomal cysteine protease precursor activation was first studied intensively by Nishimura *et al*. (18, 19). They identified the lysosomes not only as the site of cathepsin action but also of their activation. They were able to prove that correctly trimmed carbohydrate chains and an acidic pH are prerequisites for processing of procathepsins to mature enzymes (20, 21). Mutations in the catalytic sites have given very valuable information about processing and maturation of protein. The catalytic cysteine does not play any role in activation of the protein, and more importantly, no cysteine protease in involved in this process as it was expected. Catalytic asparagine is not even important for binding with DCG-04, as DCG-04 only binds to active enzymes. Mutation in conversed domains like in ERFNIN motifs does not cause much change to protein. Processing sites mutants at various domains have indeed given more insight in maturation of protein. Most notably mutation in prodomain processing site leads to complete absence of protein.

Immunocytochemical analyses show that the ER bodies accumulate two stress-inducible cysteine proteases that transported to vacuoles, RD21 and γ-VPE (22). In the vacuoles, the inactive enzyme should be converted to the active one. Thus, the ER bodies appear to be a novel proteinase-storing system that assists in cell death under the stressed conditions. We confirmed the localization of RD21 into endoplasmic reticulum, as we are able to see RD21-YFP into ER.

Pseudomonas syringae pv. *tomato* DC3000 (*Pto* DC3000) is the cause of bacterial speck on tomato and *Arabidopsis*, and represents an important model in molecular plant pathology. During infection *Pseudomonas syringae* pv. *tomato* DC3000 secretes many effector molecules, which are responsible for pathogenicity. RIP1 (RD21 Inhibiting Protein 1) is possible effector chagasin molecule, which inhibits RD21 (Kaschani and von der Hoorn unpublished).

Thermal denaturation of RIP1 has showed that upto 40 ºC, there is no major change in the conformation of the protein. At 50ºC, major change in the structure was observed, that leads to complete loss of activity. Thus, thermal inactivation of RIP1 takes place with major changes in the conformation of the protein. Previous observations with cystatins in the thermal denaturation event showed that a quick drop in the negative ellipticity, reflecting significant lose of the protein native structure, appeared around 60ºC (23). Our data is very much consistent with that. Generally, these inhibitors are thermolabile proteins. Only few proteins are reported to be thermostable, where they have melting point around 115ºC (24).

The RIP1 was found to inhibit papain 1000 times more potently than cathepsin K. This can be well explained in terms of more favorable entropy for RIP1-papain interaction. RIP1 binding with papain is an entropy driven process and a negative entropy shows the hydrophilic and polar interactions. While the RIP1 binding with cathepsin K has positive entropy, which reflects the hydrophobic and non-polar interaction. The thermodynamic study will help us to design better inhibitors, which could interact properly with the protease. The Van't Hoff plots were linear for both enzymes in the temperature range studied. Free energy of binding (ΔG) was negative at all temperatures, suggesting the spontaneous binding.

Marine organisms continue to be excellent sources of bioactive natural products. Large varieties of new bioactive compounds have recently been isolated from the marine fungi, especially the mangrove fungi (25, 26).We have isolated a pathogenic fungus from the leaf of mangrove *Avicennia marina*. The fungus was found to be a

new strain of salt tolerant pathogenic fungus *Penicillium chrysogenum*. Identification of fungus was carried out by ITS and 18S rRNA sequencing and the sequence has been submitted to Gene bank under the accession number FJ403589. *Penicillium chrysogenum* is not a common as plant pathogen and only few species have been identified as plant pathogens. (27). Hence, our isolate can give more insight about plant pathogenic *Penicillium chrysogenum* species.

Target-oriented searches for bioactive substances have become a growing need for small-molecule inhibitors essential for studies of complex processes at the interface of chemistry and biology (28). Inhibitors of specific cysteine proteases can be potential drug leads for many disease states (29, 30). Previous work on parasite cysteine protease inhibitors included analysis of diazomethane inhibitors, fluoromethyl ketones, and, more recently, oxygen-containing heterocycles and vinyl sulfones (31, 32).

Halophytic fungi require high salt concentration for their growth; because of their extreme environment, they tend to produce novel enzymes and metabolites. Many interesting cysteine protease inhibitors were isolated from marine sources. Miraziridine, a novel peptide cysteine protease inhibitor is from the marine sponge *Theonella* aff. *Mirabilis*. Aziridine-2, 3-dicarboxylic acid is a rare natural product, reported only once from a *Streptomyces*. Tokaramide A, a linear peptide having cathepsin B inhibitory activity was isolated from sponge *Theonella* (33).

Sorbicillactone A, the first sorbicillinoid alkaloid from a sponge-derived *Penicillium chrysogenum* strain, exhibiting antileukemic activity is also very interesting (34). In continuation of this, we have screened the culture filtrate of *Penicillium chrysogenum* for inhibitory activity against papain and we were able to get inhibition. The potential inhibitor was purified with different chromatographic techniques. The purified inhibitor was partial characterized by different techniques. The ESI-MS spectrum showed a peak at m/z 437.312 attributing to M^+ ion of inhibitor. Inhibitor was stable at different pH (ranging from 2 to 10) for 2 hours. The inhibitor did not show any

inhibitory activity against other classes of the proteases like trypsin, chymotrypsin, pepsin and clostripain.

CONCLUSIONS

1. Selective inhibition of host papain like cysteine proteases by pathogen inhibitors.
2. Up-regulation of proteases when treated with Benzothiadiazole (S*alicylic acid analogue* in tomato plants.
3. The effectors protein AVR2 form *Cladosporium fulvum* has been shown to inhibit PIP1
4. PIP1/RCR3 under diversifying selection, others conserved in tomato.
5. AVR2 physically interacts with PIP1 and naturally, occurring N194D mutation in RCR3 affects inhibition by AVR2.
6. The effectors proteins EPIC1 and EPIC2B form *Cladosporium fulvum* has been shown to inhibit C14 protein form tomato.
7. C14 and EPICs physically interacts with each other.
8. Granulin domain is not required for processing and activation of RD21.
9. No cysteine protease are involved in the processing of RD21 in *Arabidopsis thaliana*
10. Catalytic cysteine is not required for processing of RD21.
11. Proline rich domain is required for accumulation of RD21.
12. Binding between RIP1 and papain is entropy driven process.
13. A new fungus form leaf of *Avicennia marina was* isolated and identified as *Penicillium chrysogenum* sp. *SDBF1*
14. A new low molecular weight cysteine protease inhibitor was isolated form *Penicillium chrysogenum* sp. *SDBF1*.

List of Publications

1. **M. Shabab,** M. I. Khan, M. J. Kulkarni. Determination of binding constant of cystatin for papain by Intensity fading matrix-assisted laser desorption/ ionization time-of-flight mass spectrometry (2008) **The Protein Journal** 27:7–12.

2. **Mohammed Shabab**, Takayuki Shindo, Farnusch Kaschani, Twinkal Pansuriya, Christian Gu, Anne Harzen, Tom Colby, Sophien Kamoun and Renier A. L. van der Hoorn Fungal Effector Protein AVR2 Targets Diversifying Defense-Related Cys Proteases of Tomato (2008). **Plant Cell** 20: 1169-1183.

3. G. Pandey, S. Dumbre, M. Islam Khan and **M. Shabab**. Convergent Approach towards the synthesis of the stereoisomer of C-6 homologues of 1-deoxynojirimycin and their analogues: Evaluation as Specific Glycosidase Inhibitors. (2006) **The Journal of Organic Chemistry** 71(22) : 8481-8488.

4. G. Pandey, S. Dumbre, M. I. Khan, **M. Shabab** and V. G. Puranik. β-Lactam-azasugar hybrid as a competitive potent galactosidase inhibitor (2006) **Tetrahedron Letters** 47 (45) : 7923-7926.

5. Meera Parthasarathy, Vijayamohanan K. Pillai, Imtiaz S. Mulla, **M.Shabab**, M. I. Khan. Inter-heme interactions from All-Solid State Protein Film Voltammetry: Mimicking Cellular Confinement at Protein-Polymer Interfaces. (2007) **Biochemical and Biophysical Research Communications** 364: 86–91.

6. G. Pandey, S. Dumbre, S.G. Pal, M. I. Khan, **M. Shabab**. Synthesis and evaluation of 1-deoxy-8-epi-castanosperime, 1-deoxy-8-hrdroxymethyl castanospermine and (6S, 7S, 8R, 8aR)-8-amino-octahyroindolizine-6, 7-diol **(**2006) **Tetrahedron Letters** 63, (22): 4756-4761.

7. **M.Shabab**, S. Dumbre, G. Pandey, M. I. Khan. Thermodynamics of galactose and its analogue inhibitor 1-de-oxy-galactohomonojirimycin binding to α-galactosidase. **(Communicated)**.

8. **Mohammed Shabab**, Christian Gu, Farnusch Kaschani, Sophien Kamoun and Renier A. L. van der Hoorn (2008). Contrasting selectivity of fungal and oomycete effectors targeting tomato cysteine proteases. **(Communicated)**.

9. **Mohammed Shabab** and Renier A. L. van der Hoorn Structure-function analysis of RD21, a stress related cysteine protease in *Arabidopsis thaliana.* (**Manuscript under preparation**).

10. **Mohammed Shabab,** Vaibhav Mhaindarkar, Siddharth Bhosale and M. I. Khan Purification and characterization of cysteine protease inhibitor form plant pathogen fungus. (**Manuscript under preparation**).

REFERENCES

1. Van der Hoorn, R.A.L., and Jones, J.D.G. (2004) Curr. Opin. Plant Biol. 7: 400–407.
2. Xia, Y., Suzuki, H., Borevitz, J., Blount, J., Guo, Z., Patel, K., Dixon, R.A., Lamb, C. (2004) EMBO J., 23:980-988.
3. Tornero, P., Conejero, V., and Vera, P. (1997) J. Biol. Chem. 272: 14412–14419.
4. Pautot, V., Holzet, F.M., Reisch, B., Walling, L.L. (1993) Proc Natl. Acad. Sci. USA 90: 9906-9910.
5. Avrova, A.O., Stewart, H.E., De Jong, W.D., Heilbronn, J., Lyon, G.D., and Birch, P.R. (1999) Mol. Plant Microbe Interact. 12: 1114–1119.
6. Rooney, H.C.E., Van't Klooster, J.W., van der Hoorn, R.A.L., Joosten, M.H.A.J., Jones, J.D., and De Wit, P.J.G.M. (2005) Science 308: 1783–1786.

7. Tian, M., Win, J., Van der Hoorn, R., Van der Knaap, E., and Kamoun, S. (2007) Plant Physiol. 143: 364–377.
8. Kamoun, S., Hraber, P., Sobral, B., Nuss, D., Govers, F. (1999) Fungal Genet. Biol. 28: 94–106.
9. Randall, T.A., Dwyer, R.A., Huitema, E., Beyer, K., Cvitanich, C., Kelkar, H., Fong, A.M., Gates, K., Roberts, .S, Yatzkan, E. (2005) Mol Plant Microbe Interact. 18: 229–243.
10. Torto, T.A., Li, S., Styer, A., Huitema, .E, Testa, A., Gow, N.A., van West, .P, Kamoun, S. (2003) Genome Res 13: 1675–1685.
11. Schaffer, M.A., and Fischer, R.L. (1988) Plant Physiol. 87: 431–436.
12. Anna, O. A., Helen, E. S., Walter, D., Jacqueline, H., Gary, D. L., and Birch, P. R. J. M.P.M.I (1999) 12:1114–1119.
13. Koizumi, M., Yamaguchi-Shinozaki, K., Tsuji, H., Shinozaki, K. (1993) Gene 129: 175–182.
14. Kenji, Y., Ryo, M., Mikio, N., and Ikuko, H.-N., (2001) Plant Physiology 127, 1626–1634.
15. Christine, A., Ondzighi, D.A., Christopher, E. J., Shu-Choeng, C. and Staehelina, L.A. (2008)The Plant Cell 20: 2205–2220.
16. Bhandari, V., Palfree, R.G.E., Bateman, A. (1992) Proc. Natl. Acad. Sci. USA 89: 1715–1719.
17. Van der Hoorn, R.A, Jones, J.D. (2004) Curr. Opin. Plant Biol. 7: 400–407.
18. Nishimura, Y., Kato, K. (1987) Biochem. Biophys. Res. Commun. 148: 254–9.
19. Nishimura, Y., Kato, K. (1987) Biochem. Biophys. Res. Commun. 148: 329–34.
20. Nishimura, Y., Amano, J., Sato, H., Tsuji, H., Kato, K. (1988) Arch. Biochem. Biophys. 262: 159–70.
21. Nishimura, Y., Furuno, K., Kato, K., (1988). Arch. Biochem. Biophys. 263: 107–116.
22. Hayashi, Y., Ryo, M., Mikio, N., and Ikuko, H.N. (2001) Plant cell Physiol. 42 : 9894-9899.
23. Elzbieta, J., Wars*a*w, W., Zbigniew, G. (2004) Eur. Biophys J. 33: 454–461

24. Zerovik, E., Cimerman, N., Kos, J., Turk, V., Lohner, K.(1997) Biological Chemistry 378: 1199-1203
25. Zhong-shan, C., Wen-cheng, T., Zhi-jian, S., Cai, Y., Shi-feng, S., Qi-jin, C., Wang, F., Lin Y., She, Z. and Vrijmoed, L.L.P. (2008) J. For. Res.19: 219-224.
26. Yongcheng, L., Xiongyu, W., Shuang, F., Guangce, J., Jinghui, L., Shining, Z., L. L. P. Vrijmoed, E. B. G. Jones, Karsten, K., Klaus, S. and Ferenc, Z. (2001) J. Org. Chem. 66, 6252-6256
27. Samson, R. A., Hadlok, R and Stolk A. C. Antonie van Leeuwenhoek (1977) 43 :169-175.
28. Yoichi, N., Nobuhiro, F. (2007) J. Nat. Prod. 70 : 689 -710.
29. Rosenthal, P.J., Lee, G.K., Smith, R.E. (1993) J. Clin. Invest. 91:1052-1056.
30. McKerrow, J. H. (1999) International Journal for Parasitology 29: 8 33-837.
31. Reddy, M. V. R., Rao, M. R., Rhodes, D., Hansen, M. S. T., Rubins, K., Bushman, F. D., Venkateswarlu, Y., Faulkner, D. J. (1999) J. Med. Chem. 42**:** 1901-1907.
32. Bolger, L., Brittain, R. J., Jack, D., Jackson, M. R., Martin, L. E., Mills, J., Poynter, D., Tyers, M. B. (1972) Nature 238 : 354-356.
33. Fusetani, N., Fujita, M., Nakao, Y., Matsunaga, S. (1999) Bioorg. Med. Chem. Lett. 9: 3397-3402.
34. Yoichi, N., Masaki, F., Kaoru, W., Shigeki, M., and Nobuhiro, F. Miraziridine A, (2000) J. Am. Chem. Soc. 122: 10462-10463.

APPENDIX 1

COMPARISON OF BINDING BETWEEN GALACTOSE AND ITS ANALOGUE INHIBITOR 1-DE-OXY GALACTOHOMONOJIRIMYCIN TO α-GALACTOSIDASE.

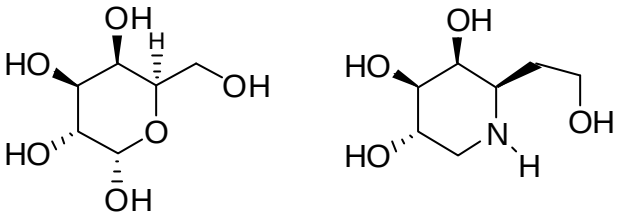
Summary

Evaluation of the kinetic parameters on the inhibition of α-galactosidase by galactose and its analogue viz. 1-deoxygalactohomonojirimycin showed that the inhibition is competitive and the latter is a more potent inhibitor. Thermodynamic parameters indicated that binding of both the inhibitors are enthalpy driven with no change in entropy. Moreover, fluorescence spectroscopic studies revealed no significant changes in the conformation of the enzyme. Although galactose binding did not lead to changes in the mean lifetime of the enzyme, minor changes were observed with 1-deoxygalactohomonojirimycin. These results are discussed with a view to design suitable drugs for the treatment of Fabry's disease.

Introduction

Glycosidases are important drug target and their roles in various lysosomal storage diseases are well known (1). α-Galactosidase (α -D-galactoside galactohydrolase, E.C.3.2.1.22) is an exo-glycosidase which cleaves the terminal α-1-6-linked D-galactosyl residues from wide range of substrates including raffinose, stachyose, melibiose, verbascose, polysaccharides of galactomannans, locust bean gum and guar gum (2). An increased interest in α-galactosidase can be seen in human medicine as it is used in several aspects such as blood group transformation, treatment of Fabry's disease and also in xenotransplantation (3, 4). It has been observed that inhibitors of human lysosomal glycosidase have therapeutic potential for the corresponding lysosomal storage disease. Enzyme inhibition is an established mode of drug action, and is important to understand in order to know how to design new inhibitors. Azasugars are inhibitors of α-galactosidase and these small molecules have been proposed as potential drugs to treat Fabry's disease (5). At sub-inhibitory concentrations, potent competitive inhibitors of the mutated enzyme can act as active-site-specific chaperones that either induce or stabilize the proper conformation of the mutant enzyme (6). How do these azasugars binds and stabilizes these glycosidase is the question of great importance. Here we compare the binding of two different inhibitors of α galactosidase viz galactose and 1-de-oxy-galactohomojirimycin (1-DGJ). Galactose is a product inhibitor of enzyme where as 1-de-oxy-

galactohomojirimycin is a substrate analogue of galactosidase. Both are competitive inhibitor of enzyme hence binding mechanism is expected to be similar.



Galactose 1-DGJ

Comparisons of thermodynamic parameters for binding of various molecules provides on insight into stoichiometry of the binding. The major advantage includes the proper understanding of bond energy and feasibility of the reaction.

It is now well established that change in the quantum yield and emission maxima of intrinsic fluorescence can be used to monitor conformational changes of proteins and their interaction with ligands (7). In this study we report the thermodynamics of binding galactose and 1-de-oxy-galactohomojirimycin with α-galactosidase.

Materials and Methods

α-Galactosidase (green coffee beans), D-galactose and p-nitrophenyl α-D-galactopyranoside are from Sigma (St. Louis, MO U.S.A), Sephadex G-75 was purchase from Amersham Biosciences, Uppsala Sweden. All other chemicals used were of analytical grade.

Purification of α galactosidase

Commercial available α-galactosidase contains bovine serum albumin for long term storage and stabilization, which was removed by gel filtration chromatography on Sephadex G-75 and the purity of the enzyme was confirmed by SDS-PAGE (8). Protein concentrations were determined according to Bradford (9) using BSA as standard.

Synthesis of Inhibitor

Synthesis and testing of anti α-galactosidase activity of the azasugar inhibitor (1-de-oxy-galactohomojirimycin) have been described previously 10).

Enzyme assay

The α-galactosidase assay was performed in 50 mM citrate phosphate buffer at pH 6.5 with *p*-nitrophenyl α-D-galactopyranoside as the substrate. Varying concentration of the substrate and the 1-de-oxy-galactohomojirimycin were employed. The reaction was initiated by the addition of 100 µl of appropriately diluted enzyme and the reaction mixture, which had a final volume of 1 ml, was incubated at 37 ^{0}C, for 20 min. It was quenched by the addition of 2 ml of 1 M Na_2CO_3 solution and the optical density of the resulting solution was read at 405 nm.

Evaluation of kinetic parameters

Lineweaver-Burk analysis of α-galactosidase (5.5 nM) was done with 2.3 µM, 4.6 µM of 1-deoxy-D-galactohomonojirimycin and 1 mM, 1.5 mM of galactose. These inhibitors were incubated with enzyme and assayed at increased concentration of *p*-nitrophenyl α-D-galactopyranoside (0.75 mM to 5 mM µM) at 37 °C for 30 min. The reciprocals of substrate hydrolysis (1/v) for each inhibitor concentration were plotted against the reciprocals of the substrate concentrations, and the K_i was determined by fitting the resulting data. For the Dixon's plot, hydrolytic activity of α-galactosidase (5.5 nM) was measured by incubating with galactose (0.765 µM to 4.6 µM) and 1-deoxy-D-galactohomonojirimycin (0.1 mM to 1 mM) in the presence of p-nitrophenyl α-D-galactopyranoside (0.5 mM and 1.25 mM). The reciprocals of substrate hydrolysis (1/v) were plotted against the inhibitors concentration and the K_i was determined by fitting the data using Microcal ORIGIN 6.1.

Ligand –binding thermodynamics data analysis

Free energy changes of inhibitory constant (ΔG) was determined by the equation;

$$\Delta G = - RT \ln K_a \quad [1]$$

Temperature dependence of the inhibitor constants was used to determine the thermodynamic parameters. Changes in enthalpy (ΔH) were determined from the Van't Hoff plots by using the equation;

$$\ln K_a = (-\Delta H/RT) + \Delta S/R \quad [2]$$

Where R is gas constant, ΔS is entropy change and T is the absolute temperature. The entropy change was obtained from the equation:

$$\Delta G = \Delta H - T\Delta S \quad [3]$$

The assay was done at different temperatures calculating various K_i of α-galactosidase with galactose and 1-de-oxy-galactohomojirimycin.

Steady state and time-resolved fluorescence measurements

Steady state Fluorescence measurements were made using a Perkin Elmer LS-50B Spectrofluorimeter, with slit width of 7 nm for both the monochromators and scan speed was 100 nm/min. A fresh glycosidase solution (5.5 nM) was prepared in 20 mM (pH 6.5) and 150 mM NaCl. This enzyme concentration resulted in a sufficient fluorescence intensity signal and was still low enough to avoid aggregation. α-Galactosidase samples were placed in a quartz cuvette maintained at desired temperature (± 0.1 °C) by means of a Julabo circulating cryobath. The inhibitor solution was added in 10-12 aliquots (5 to 10 µl each). Concentration of the inhibitors stock solutions was in the range of 1mM to 5 mM for 1-deoxy-D-galactohomonojirimycin and 10 mM to100 mM for galactose. Samples were excited at 280 nm and the emission spectra were recorded for wavelength ranging from 300 nm to 400 nm. Each spectrum was an average of 5 accumulations. The fluorescence intensity at 345 nm (λ_{max} of α-galactosidase) was considered for all the data analysis and calculations. Corrections were also made to compensate the dilution effect upon addition of inhibitor to enzyme. At the highest concentration of the inhibitor to α galactosidase, volume change was less than 5 % of the solution in the cuvette. All the

titrations were repeated three times and the average value was considered with a Standard deviation less than 5 %. All fluorescence measurements were recorded at 37 ^{0}C and were background-corrected for buffer (+inhibitor) autofluorescence.

The association constants were calculated according to the method described by (12*)*, the abscissa intercept of the plot of log $[C]_f$ against log $\{(\Delta F)/(F_c - F_\infty)\}$, where $[C]_f$ is the free ligand concentration, yielded pK_a value for enzyme-ligand interaction according to the relationship [4].

$$\log [F_0 - F_c / F_c - F_\infty] = \log K_a + \log \{ [C]_t - [P]_t (\Delta F / \Delta F_\infty) \} \qquad [4]$$

Where F_c is the fluorescence intensity of the enzyme at any point during the titration, $[P]_t$ is the total protein concentration, ΔF_∞ is the change in fluorescence intensity at saturation binding, $[C]_t$ is the total ligand concentration, and $[C]_f$ is the free ligand concentration, given by

$$[C]_f = \{[C]_t - [P]_t (\Delta F / \Delta F_\infty)\} \qquad [5]$$

Time-resolved fluorescence and anisotropy spectra were measured with a time correlated single photon counting FLS920 spectrometer (Edinburgh instruments, Livingston, UK), equipped with M300 excitation and emission monochromators and a red light-sensitive S300-R PMT detector. Excitation at 280 and 295 nm was achieved with an F900 nanosecond flash-lamp (Edinburgh instruments, Livingston, UK), with hydrogen as operating gas. Pulse frequency was 40 kHz. Recorded according to the single-photon counting method (7), as described previously. The resulting fluorescence decays were background- corrected by subtracting the fluorescence decay of the buffer (plus inhibitor) solution recorded under the same conditions as the enzyme (plus inhibitor) decay. The instrument response function was determined by measuring the ultra fast fluorescence decay of *Ludox. Time-resolved data analysis* was performed using the program supplied with the instrument (F900). For the analysis of the fluorescence decay, a distribution of 150 equally

spaced lifetime values τ on a log (τ) scale between 0.01 and 20 ns with amplitudes α were used. The starting distribution models were chosen to be flat since no a priori knowledge about the system was available. In all cases, the χ^2 values were close to 1.0 (see Table 2), and the weighted residuals as well as the autocorrelation of the residuals were randomly distributed around zero indicating an optimal fit. For the measurement of enzyme-inhibitor complexes, the inhibitors were added in a 10-fold molar excess yielding at least 95% complex formation.

The average life times are calculated by the expressions (13)

$$\tau_0 = \frac{\sum_i \alpha_i \tau_I}{\sum_i \alpha_i} \qquad [6]$$

$$<\tau> = \frac{\sum_i \alpha_i \tau^2_I}{\sum_i \alpha_i \tau_i} \qquad [7]$$

Where τ_0 and $<\tau>$ are the average fluorescence lifetimes obtained by the two different approaches.

Results

Initial kinetic analysis for the determination of K_m and K_i

The Lineweaver-Burk reciprocal plot (Fig.1) showed that both the inhibitors were a competitively inhibiting α-galactosidase. The K_m value for the α-galactosidase of *p*-nitrophenyl α-D-galactopyranoside was 0.9 mM. Inhibitors were found to inhibit α-galactosidase with an IC_{50} value (50% inhibitory concentration) of 0.48 μM for azasugar and 0.35 mM for Galactose (Data not shown). The inhibition constant K_i, determined by the classical double reciprocal plot and also by Dixon plot was 1.7 μM for 1-deoxy-D-galactohomonojirimycin and 0.35 mM for galactose (Fig.1).

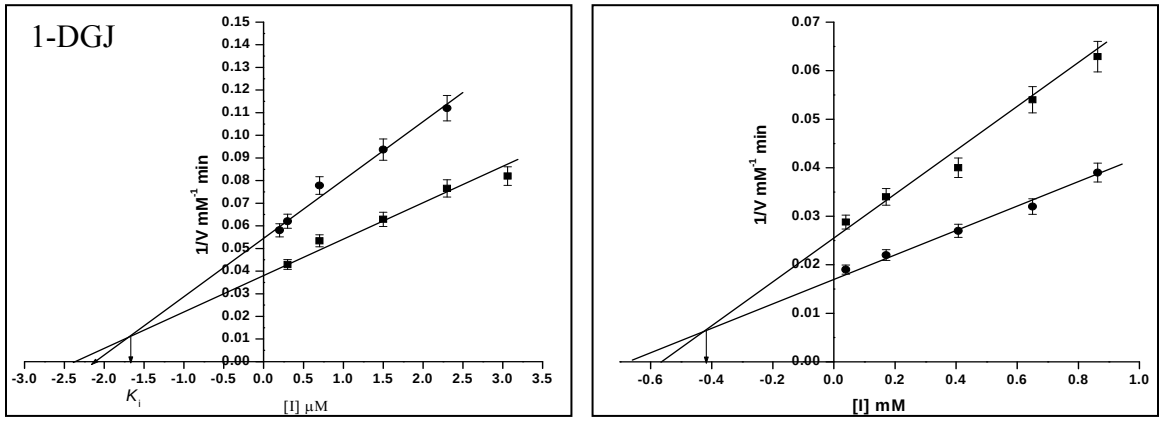
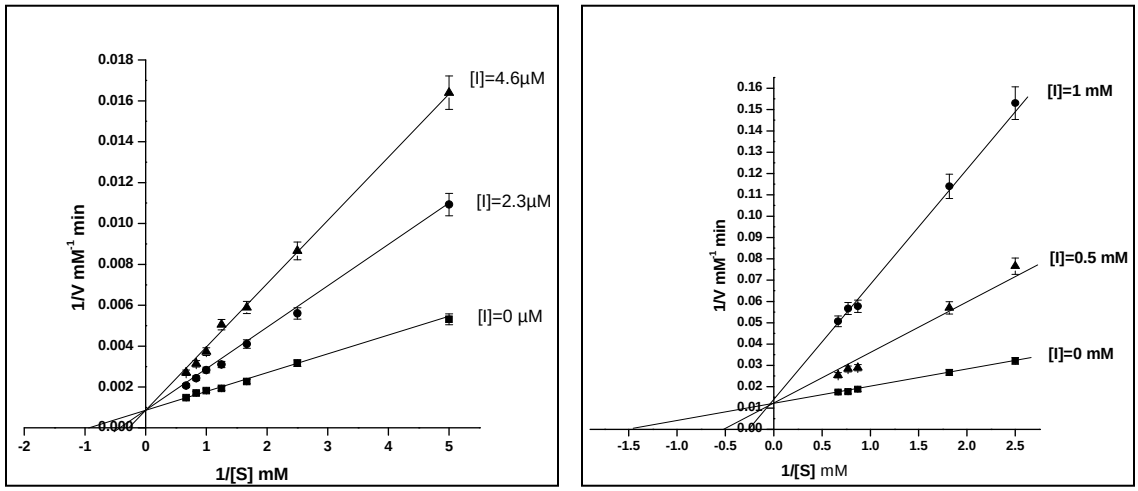


Figure 1. Binding of inhibitor to the enzyme and inhibition kinetics analyses

A) Dixon plot. Enzymatic activity of the α-galactosidase (5.5 nM) was estimated using the substrate *p*-nitrophenyl α-D-galactopyranoside 0.5 mM and 1.25 mM at different concentrations of inhibitors. Reciprocals of the reaction velocity were plotted versus the inhibitor concentration. The straight lines indicated the best fit of the data obtained. The inhibition constant K_i was calculated form the point of the intersection of the plots.



B) α-galactosidase (5.5 nM) was incubated without and with the inhibitor and assayed at increasing concentration of the substrate. The reciprocals of the rate of the substrate hydrolysis for each inhibitor concentration were plotted against the reciprocal of the substrate concentration. K_i was determined from the formula as per the competitive type of inhibition.

Ligand –binding thermodynamics data analysis

The thermodynamic parameters viz. free energy (ΔG), enthalpy (ΔH) and entropy (ΔS) of binding, determined by K_i at different temperatures, are given in Table 1. The van't Hoff plots were linear ($r > 0.9$) for both the inhibitors in the temperature range studied (Fig.2).

Free energy of the binding (ΔG) was negative at all the temperatures, suggesting the spontaneous nature of the binding. Enthalpy change (ΔH) and entropy (ΔS) change of binding was negative for both the inhibitors, indicating exothermic and enthalpically driven nature of the binding (Table 1). Negative ΔH and ΔS values also indicate primarily intermolecular hydrogen bonding and increased polar interactions between enzyme-inhibitor interactions. The inhibitory constant (K_i) increases when there is increase in temperature in the case both the cases viz. galactose and 1-de-oxy-galactohomojirimycin. The value of K_i with galactose (K_i = 0.35 mM) was 1000 fold lower than of the 1-de-oxy-galactohomojirimycin (K_i=1.7 μM). Low affinity of galactose than the 1-de-oxy-galactohomojirimycin is also manifested in lower enthalpy (ΔH= -11.1 kJ mol^{-1}). The larger value for 1-de-oxy-galactohomojirimycin shows the greater involvement of the hydrogen bonding in the protein-ligand binding. The data can be well explained in terms of structural difference between both the inhibitors. The azasugar contains the nitrogen in place of oxygen in the ring. This energy could be received from forming a strong hydrogen bonding or a salt bridge between the catalytic carboxylic acid group of the enzyme and ring nitrogen. The entropy (ΔS) is same for both the inhibitors indicating their no entropy change in the binding pattern of galactose and 1-de-oxy-galactohomojirimycin. The over all free Gibbs energy (ΔG) is more in protein-1-de-oxy-galactohomojirimycin interaction as compare to galactose interaction.

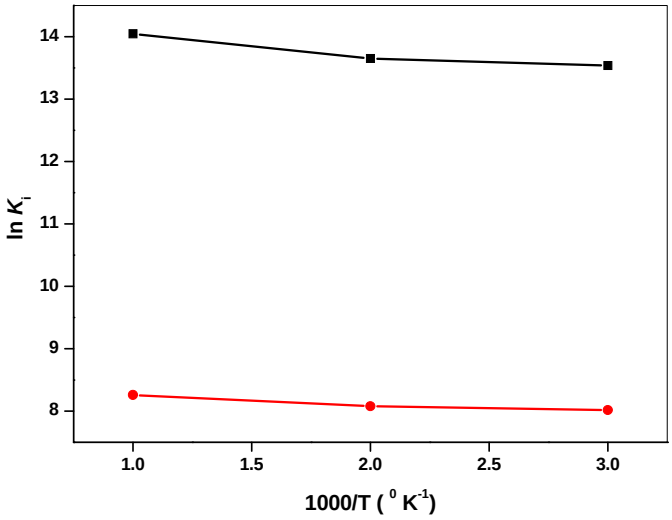


Figure 2 Van't Hoff plots for the inhibitory constant of galactose and 1-de-oxy-galactohomojirimycin to α-galactosidase are drawn according to the regression equation (>0.9, N=3). The symbols used are: (■) galactose and (□) 1-de-oxy-galactohomojirimycin. The values of K_i determined by tube assay are listed in Table 1.

Steady state Fluorescence

The fluorescence emission spectrum of the α-galactosidase is typical for a tryptophan containing protein in aqueous solution with maximum intensity at 345 nm (figure 3). The binding of inhibitor resulted in a concentration –depended quenching of the fluorescence with saturation reaching or above 10 mM for galactose and 1μM for 1-de-oxy-galactohomojirimycin (Fig 3 inset). The absence of blue or red shift in λ_{max} negated any drastic gross conformational changes in the three-dimension structure of the enzyme due to inhibitors binding. The maximum quenching of the intrinsic fluorescence of the protein on binding with inhibitors in both the cases was 25 %.(Fig 3 inset). The slope of the plot of log $[F_0\text{-}F_c/F_c\ \text{-}F_\infty]$ versus $\log[C]_f$ was near unity for both the ligands, indicating one binding site per monomer of the enzyme. The dissociation constant was found to 20000 for 1-de-oxy-galactohomojirimycin and 6000 for Galactose.

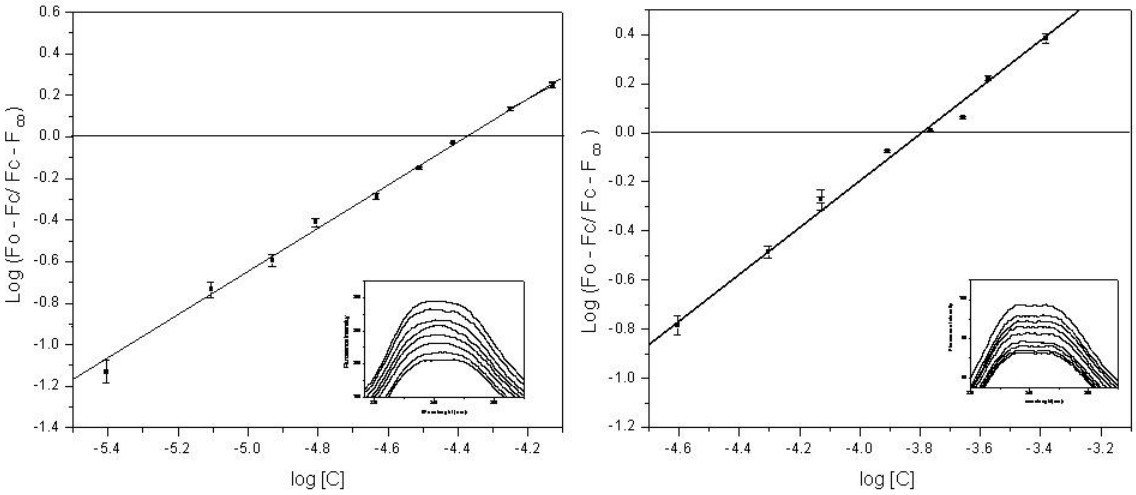


Figure 3A. **Figure 3B.**

Figure 3. Determination of the association constant for galactose and 1-de-oxy-galactohomojirimycin for α-galactosidase interaction by fluorescence spectroscopy. F_0 and F_∞ are the fluorescence intensity of free ligand and that bound to the α-galactosidase at infinite ligand concentration. F_c is the corrected value of fluorescence intensity at any point of titration; [C] is the protein concentration in morality. Inset represents fluorescence quenching of α-galactosidase on addition of aliquots of 1-de-oxy-galactohomojirimycin to the enzyme solution. Inset represents fluorescence quenching of α-galactosidase on addition of aliquots of galactose and 1-de-oxy-galactohomojirimycin to the enzyme solution. Figure 3A represents for galactose and 3B represents for 1-de-oxy-galactohomojirimycin.

Time-resolved fluorescence and anisotropy spectra

A typical time –resolved fluorescence decay of the α-galactosidase is shown in (Fig.4a) collectively with the respective data analysis by the Nf200 software. The decay was recorded at 310 K with the excitation wavelength set at 280 nm, and the emission was collected at 348 nm. It was seen during the life time measurement the α-galactosidase τ1 and τ2 were found to be 5.12ns and 1.14ns respectively (Fig.4a and 4b). The χ^2 was near to 1 for all life time measurements. The value of τ1 and τ2 were similar to native α-galactosidase, on D-galactose binding but τ1 becomes faster and τ2 slower on binding of1-de-oxy-galactohomojirimycin.The τ1 was found to be

contributing about 70 % of the total fluorescence of the protein, while the τ2 have only 30 % of protein fluorescence (Table 2).

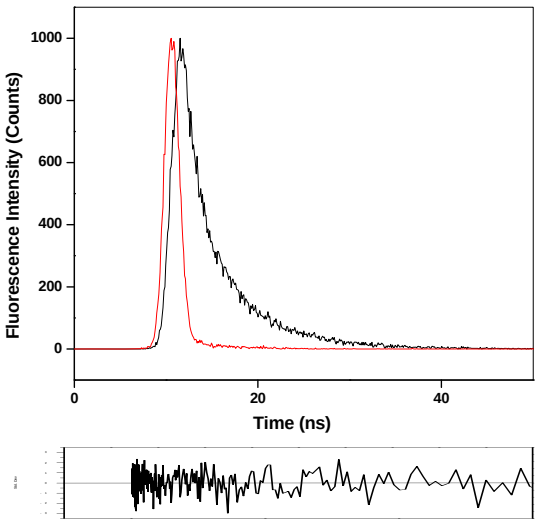


Figure 4A.

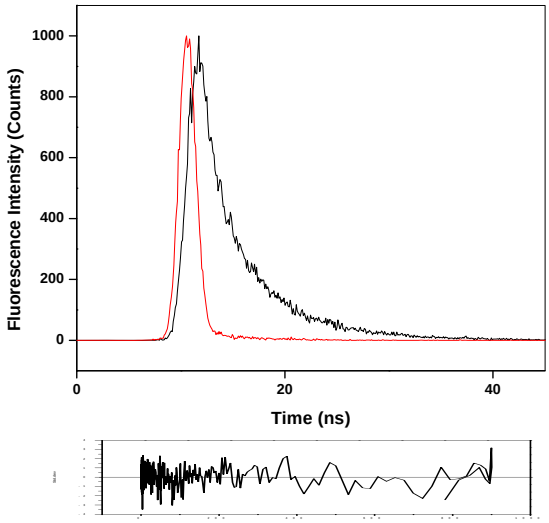


Figure 4B.

Figure 4. Time-resolve fluorescence decay profiles of α-galactosidase in the presence of 2 mM galactose and 2 μM of 1-de-oxy-galactohomojirimycin to the enzyme solution. The solid lines correspond to the nonlinear least square fit of the experimental data to a biexponential function. The lower panel represents the residual. Figure 4A represents for galactose and 4B represents for 1-de-oxy-galactohomojirimycin.

Discussion

The kinetic parameters showed that 1-de-oxy-galactohomojirimycin and galactose are competitive inhibitors of the enzyme. Moreover, there is about 1000 fold differences in their K_i values. The temperature dependences of the kinetics parameters of galactose and its analogue inhibitor 1-de-oxy-galactohomonojirimycin is quiet similar. The thermodynamics parameters of 1-de-oxy-galactohomojirimycin and galactose show that the binding of these inhibitors to α-galactosidase, is an enthalpy driven process. Binding is in fact associated with essentially no change in entropy. This suggests that the both inhibitors bind in fundamentally same way.

The overall retention of the α monomer in the catalytic product is achieved through a double displacement reaction mechanism (14). Two successive nucleophilic attacks on the same anomeric carbon invert the chiral center twice, resulting in the retention of the α-anomer at the completion on the reaction cycle. This mechanism requires two carboxylic acids, one acting as an acid and than a base over the course of the reaction cycle (15). Fluorescence titration of α-galactosidase with the galactose and 1-de-oxy-galactohomojirimycin resulted in quenching of the α-galactosidase fluorescence without any shift in the emission maximum. The fluorescence decay profile obtained in the presence of galactose remains unchanged as compared to that of native α-galactosidase and the fluorescence life times decease only barely compared to the native protein when recorded in the presence of the inhibitor. These results , which are consistent with the steady state fluorescence spectral data as well the quenching experiments, indicate that any conformational changes resulting from the galactose binding in solution do not lead to major changes in the environment of tryptophan residues.

Different fluorescence decay profiles were obtained in the presence of 1-de-oxy-galactohomojirimycin compared to that of native α-galactosidase. The interchange of τ1 and τ2 observed in presence of D-galactose and 1-de-oxy-galactohomojirimycin indicates stronger interaction with Trp residues in presence of 1-de-oxy-galactohomojirimycin.

In conclusion, the major forces for the binding of both the inhibitors to α-galactosidase are hydrogen bonding and Van der Waals interactions coupled with some polar contribution. This suggests that both the inhibitors binds in fundamentally same way. The higher affinity accompanied by a large enthalpy contribution observed with 1-de-oxy-galactohomojirimycin suggested preference of the forming a strong hydrogen bonding or a salt bridge between the catalytic carboxylic acid group of the enzyme and ring nitrogen. Fluorescence experiments show that there is no significant change in the protein conformation during inhibitor binding in both the cases.

Table 1. Inhibitory constant and thermodynamic parameters for the binding of enzyme to the inhibitors at different temperatures.

Inhibitor	K_i (M) Temperatures					
	15	25	30	-*ΔH kJ.mol^{-1}	-* ΔG kJ.mol^{-1}	-*ΔS J.mol^{-1}.K^{-1}
Galactose	$0.259*10^{-3}$	$0.31*10^{-3}$	$0.33*10^{-3}$	11.1kJ	19.99	30
1-de-oxy-galacto-homojirimycin	$0.793*10^{-6}$	$1.18*10^{-6}$	$1.32*10^{-6}$	25.2kJ	33.81	29

*Values at 25 ^{0}C

Table 2. Life times of native α Galactosidase and after mixing with 1-de-oxy-galactohomojirimycin and galactose.

Sample	α_1(%)	τ_1 (ns)	α_2(%)	τ_2 (ns)	τ_{mean}	χ^2
α Gal	66.50	5.44 ns	33.50	1.51 ns	2.66 ns	1.004
Galactose*	67.50	5.31 ns	32.50	1.29 ns	2.70 ns	1.011
1-de-oxy-galato homojirimycin*	31.66	1.11 ns	68.34	5.14 ns	2.40 ns	1.070

* After incubating with 1 μM enzyme

References

1. Butters, T.D., Dwek, R.A. and Platt, F.M. (2005) Glycobio. 15: 43R-52R.
2. Gote, M.M., Khan M.I., Gokhale, D.V., Bastawde, K.B. and Khire, J.M. (2006) Pro. Biochem. 41: 1311-1317.
3. Linthorst, G.E., Hollak, V.W.E., Donker, K., Strijland A. and Aerts, J.M.F.C. (2004). Kidney. Int. 66: 1589-1595.
4. Olsson, M.L., Hill, C.A., Dela, V.H., Liu, Q.P., Stroud, M.R. and J. Valdinocci (2004) Transf. Clin. Et Biol. 11: 33-39.
5. Fan, J.-Q., Ishii S., Asanom, N., and Suzuki, Y. (2006) Nature Med. 5: 112-115.
6. Fan, J.-Q. (2003). Tre. Pharm. Sci. 24: 355-360.
7. Kungl, A.J., Visser, N. V. Hoek, A. V., Visser,A.J.W.G., Billich, A. Gstach, H. and Auer M. (2006) Biochemistry 37: 2786-2789
8. Laemmli, U.K. (1970) Nature: 670-685.
9. Bradford, M.M. (1976) Anal. Biochem. 72: 248-254
10. Pandey, G., Dumbre, S.,Khan, M.I. and Shabab, M. (2006) J.Org.Chem 71: 8481-8488.
11. Pandey, G., Dumbre, S.,Khan, M.I., Shabab, M. and Puranik, V.G. (2006) Tetrahedron Lett. 47: 7923-7926.
12. Chipman, D.M., Grisaro, V. and Sharon, N. (1967) J. Biol. Chem. 242: 4388-4394.
13. Grinvald, A. and Steinberg, I.Z. (1974) Anal. Biochem. 59: 583–598
14. Garman, S.C., Garboczi D.N. (2004) J. Mol. Biol. 337: 319-335.
15. Golubev, A.M., Nagem, V J.R.B., Neto, K.N., Neustroev, E.V. Eneyskaya, A.A. Kulminskaya, K.A. Shabalin, A.N. Savelev, A.N. Savelev and I. Polikarpov (2006). J. Mol. Biol. 339: 413-422.