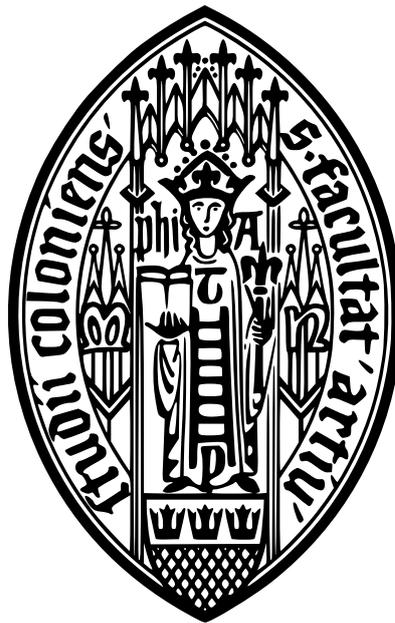# Analysing Novel Structures of Protein Complexes Requiring the Central Plant Immune Regulator EDS1

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Stephan Wagner

aus

Köln

im

März 2013

Referees/Gutachter: PD Dr. Karsten Niefind

Institute of Biochemistry/Institut für Biochemie

University of Cologne/Universität zu Köln

Prof. Dr. Kay Hofmann

Institute for Genetics/Institut für Genetik

University of Cologne/Universität zu Köln

Prof. Dr. Thorsten Nürnberger

Center for Plant Molecular Biology Plant Biochemistry/

Zentrum für Molekularbiologie der Pflanzen

University of Tübingen/Universität Tübingen

Chair of Defence/

Prüfungsvorsitzende: Prof. Dr. Ute Höcker

Botanical Institute/Botanisches Institut

University of Cologne/Universität zu Köln

Tag der letzten mündlichen Prüfung: 14.05.2013

The work described in this thesis was conducted during March 2009 and March 2013 under supervision of PD Dr. Karsten Niefind at the Institute of Biochemistry (Department of Chemistry) at the University of Cologne, Zülpicher Straße 47, D-50674 Cologne, Germany.

Die in dieser Dissertation beschriebenen Arbeiten wurden zwischen März 2009 und März 2013 unter der Anleitung von PD Dr. Karsten Niefind am Institut für Biochemie (Department für Chemie) der Universität zu Köln, Zülpicher Straße 47, D-50674 Köln, Deutschland, durchgeführt.

Parts of this thesis have been published/Teile dieser Arbeit wurden bereits veröffentlicht:

Rietz, S., Stamm, A., Malonek, S., Wagner, S., Becker, D., Medina-Escobar, N., Vlot, A. C., Feys, B. J., Niefind, K., and Parker J. E. (2011). Different roles of Enhanced Disease Susceptibility1 (EDS1) bound to and dissociated from Phytoalexin Deficient4 (PAD4) in Arabidopsis immunity. *The New Phytologist*, 191(1), 107-19. doi:10.1111/j.1469-8137.2011.03675.x

Wagner, S., Rietz, S., Parker, J. E., and Niefind, K. (2011). Crystallization and preliminary crystallographic analysis of Arabidopsis thaliana EDS1, a key component of plant immunity, in complex with its signalling partner SAG101. *Acta crystallographica. Section F, Structural biology and crystallization communications*, 67(Pt 2), 245-8. doi:10.1107/S1744309110051249

Acknowledgements

Contents

Abbreviations

Amino acids

Ala	A	Alanine
Cys	C	Cysteine
Asp	D	Aspartate
Glu	E	Glutamate
Phe	F	Phenylalanine
Gly	G	Glycine
His	H	Histidine
Ile	I	Isoleucine
Lys	K	Lysine
Leu	L	Leucine
Met	M	Methionine
Asn	N	Asparagine
Pro	P	Proline
Gln	Q	Glutamine
Arg	R	Arginine
Ser	S	Serine
Thr	T	Threonine
Val	V	Valine
Trp	W	Tryptophan
Tyr	Y	Tyrosine

Further abbreviations

-	Fused to (in the context of protein fusion constructs)
% (v/v)	Volume percent
% (w/v)	Weight/volume percent
% (w/w)	Weight percent
35S	35S promoter of cauliflower mosaic virus
Å	Ångström (1 Å = 0.1 nm)
Aa	*Arabis alpina* (alpine rock-cress)
AChE	Acetylcholinesterase
AEC	Anion exchange chromatography
AMP/ADP/ATP	Adenosine mono/di/triphosphate
At	*Arabidopsis thaliana* (mouse-ear cress)
AU	Absorption units at a wavelength of 280 nm
β-ME	2-Mercaptoethanol
Bicine	2-(Bis(2-hydroxyethyl)amino)acetic acid; a buffer substance
BS3	Bissulfosuccinimidyl suberate

BSA	Bovine serum albumine
°C	Degree Celsius
CAPS	N-cyclohexyl-3-aminopropanesulfonic acid; a buffer substance
CBB	Coomassie Brilliant Blue
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1- propanesulfonate; a detergent
CHES	N-Cyclohexyl-2-aminoethanesulfonic acid; a buffer substance
Col-0	*Arabidopsis thaliana* accession Columbia
C-terminal	Carboxy-terminal
d	Day(s)
Da	Dalton (1 Da = 1 g/mol)
DDAO	Dodecyldimethylamine oxide; a detergent
DHBA	Dihydroxybenzoic acid
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDS1	ENHANCED DISEASE SUSCEPTIBILITY1
EDTA	Ethylendiamintetraacetate
et al.	*et alii* (and others)
Fig.	Figure
FPLC	Fast performance liquid chromatography
g	Standard gravity ($1\,g \approx 9.81\,\text{m/s}^2$)
g	Gramme(s)
GID1	GIBBERELIC ACID INSENSITIVE DWARF1
GST	Glutathione S-transferase
GuHCl	Guanidine hydrochloride
h	Hour(s)
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid; a buffer substance
His	Polyhistidine tag (in the context of a protein fusion)
Hpa	*Hyaloperonospora arabidopsidis*
HR	Hypersensitive response
I	Intensity
IMAC	Immobilized metal-affinity chromatography
in vitro	"within the glass"; experiments performed in an artificial environment
in vivo	"within the living"; experiments performed in a living cell or organism
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
K	Kelvin ([°C]=[K]-273.15)
kb	kilobase
l	Litre
Ler	*Arabidopsis thaliana* accession Landsberg *erecta*

λ	Wavelength (nm)
LB	Lysogeny broth (nutritionally rich medium)
m	Metre
M	Molar (1 M = 1 mol/l)
MAD	Multi-wavelength anomalous dispersion
MAP	Mitogen-activated protein
MBP	Maltose binding protein
MES	2-(N-morpholino)ethanesulfonic acid; a buffer substance
min	Minute
MML	*Mucor miehei* lipase
mol	Mole; the amount of substance (1 mole=6.022×10^{23} particles)
MPD	2-Methyl-2,4-pentanediol
MPIPZ	Max-Planck-Institut für Pflanzenzüchtungsforschung
	(Max-Planck-Institute for Plant Breeding Research)
MW	Molecular weight
NAD/NADH	Nicotinamide adenine dinucleotide (oxidised/reduced)
NL-1	NEUROLIGIN-1
NTA	Nitrilotriacetic acid
N-terminal	Amino-terminal
OD_x	Optical density at a wavelength of x nm
PAD4	PHYTOALEXIN DEFICIENT4
PAGE	Polyacrylamide gel electrophoresis
PAMP	Pathogen-associated molecular pattern
PCR	Polymerase chain reaction
PDB	Protein Data Bank
pH	*pondus Hydrogenii*
pNP	para-Nitrophenyl (4-nitrophenyl)
R_{free}	Free R-factor
ρ	Mass concentration
rmsd	Root mean square deviation
rpm	Revolutions per minute
RPS	RESISTANCE TO PSEUDOMONAS SYRINGAE
RT	Room temperature
R_{work}	Working R-factor
SABP2	SALICYLIC ACID BINDING PROTEIN2
SAG101	SENESCENE ASSOCIATED GENE101
SD	Synthetic defined
SDS	Sodium dodecyl sulfate; a detergent
sec	Second(s)
SEC	Size exclusion chromatography
σ(I)	Error of the intensity
St	*Solanum tuberosum* (potato)

TAE	Tris-acetate-EDTA buffer
TBS	Tris-buffered saline; a buffer
TBS-T	TBS buffer supplemented with Tween-20
TE	Tris-EDTA buffer
TEA	Triethanolamine; a buffer substance
TCEP	Tris(2-carboxyethyl)phosphine; a reducing agent
T_m	Transition midpoint
Tris	Tris(hydroxymethyl)aminomethane; a buffer substance
Trx	Thioredoxin
U	Unit (1 u = the amount of the enzyme that catalyzes the conversion of 1 micromole of substrate per minute)
UV	Ultraviolet light
V	Volt(s)
wt	Wildtype

Summary

Plants as sessile organisms are constantly attacked by pathogens and have evolved a complex immune system to cope with this challenge. The first inducible layer of plant immunity recognises pathogens outside each plant cell and activates a resistance response that is usually sufficient to halt pathogen infection. Many pathogens, however, produce effector molecules that suppress these defences and thus allow successful invasion of the plant. A second layer of plant immunity relies on the specific recognition of these effectors through intracellular immune receptors, the so-called R (resistance) proteins. Effector recognition triggers a fast and strong immune response that is often associated with localised cell death at the site of infection. To prevent spreading cell death and to reduce the cellular costs of an activated immune system, these activities have to be tightly regulated.

The nucleo-cytoplasmic protein EDS1, which exists in all higher plant species, has emerged as a central regulator downstream of effector recognition. EDS1 physically interacts with two sequence-related proteins, PAD4 and SAG101, that share an organisation into an N-terminal domain that is related to eukaryotic lipases and a C-terminal domain that has no obvious sequence homologs. The EDS1 protein family, as defined by this organisation, is involved in multiple plant stress signalling pathways. While recent studies have helped to genetically position EDS1, PAD4 and SAG101, their biochemical mode of action remained elusive.

To gain mechanistic insights into their functions, the structures of EDS1-containing complexes should be elucidated by means of X-ray crystallography. This thesis describes the preparation of recombinant proteins, their characterisation and the *de novo* structure solution of a functional complex between EDS1 and SAG101.

I found that EDS1, despite being closely related to eukaryotic lipases at sequence and structure level, does not utilise conserved lipase features to bind or process a (lipid-derived) signalling molecule in central immune pathways of the model plant *Arabidopsis*. The lipase-like half of the protein is rather optimised for protein-protein interactions within the EDS1 family.

The interactions between EDS1 and SAG101 were characterised *in vitro* and *in vivo* and used to infer a homology model of PAD4. EDS1, PAD4 and SAG101 loss-of-interaction variants were proposed on the basis of this structural information and could be verified experimentally. These will help to further discriminate between the functions of single components and protein complexes within EDS1-dependent pathways.

While interactions within the EDS1 protein family are largely driven by the lipase-like half, the C-terminal EP domain is sequentially and structurally unique but reveals a distant relation to proteins that exist in multi-protein complexes. This domain potentially acts as a binding platform for proteins outside the EDS1 family and could gain importance as the number of experimentally characterised EDS1 interactors is constantly growing.

Zusammenfassung

Als sessile Organismen sind Pflanzen ständigen Attacken durch Pathogene ausgesetzt und haben als Antwort darauf ein komplexes Immunsystem entwickelt. Die erste Ebene pflanzlicher Immunität erkennt Pathogene noch außerhalb der Zelle und aktiviert Resistenzmechanismen die in der Regel ausreichen, einen Befall durch das Pathogen zu verhindern. Eine Vielzahl von Pathogenen synthetisiert allerdings sogenannte Effektormoleküle, die diese Resistenzmechanismen aushebeln und so den Befall der Pflanze ermöglichen. Eine zweite Ebene des pflanzlichen Immunsystems erkennt solche Effektoren mittels intrazellulärer Immunrezeptoren, den sogenannten R (Resistenz) Proteinen. Erkennung des Effektors führt zu einer Immunantwort, die häufig mit lokalem Zelltod am Ort der Infektion assoziiert ist und deshalb strikt reguliert werden muss.

Das nukleo-zytoplasmatische Protein EDS1, das ist allen höheren Pflanzen existiert, übernimmt eine zentrale regulatorische Rolle nach Effektor-Erkennung. EDS1 wechselwirkt mit den sequenz-verwandten Proteinen PAD4 und SAG101, die, ebenso wie EDS1, in eine N-terminale Domäne mit Ähnlichkeit zu eukaryotischen Lipasen und eine C-terminale Domäne organisiert sind, für die sich keine offensichtlichen Sequenzhomologe finden. Die EDS1 Proteinfamilie ist durch diese Organisation charakterisiert und spielt eine zentrale Rolle in unterschiedlichsten Stress-Signalwegen der Pflanze. Während genetische Analysen geholfen haben, die EDS1 Familie in diese Signalwege einzuordnen, blieb ihre biochemische Funktion unbekannt.

Um mechanistische Einblicke in ihre Funktion zu erlangen, sollten die Strukturen von EDS1-assoziierten Proteinkomplexen mittels Röntgenstrukturanalyse gelöst werden. In dieser Arbeit ist die Präparation von rekombinanten Proteinen, deren Charakterisierung und eine neuartige Struktur des funktionalen Komplexes aus EDS1 und SAG101 beschrieben.

Ich konnte zeigen, dass EDS1 trotz seiner strukturellen Verwandtschaft zu Lipasen in zentralen Immunsignalwegen der Modellpflanze *Arabidopsis* weder ein (lipidähnliches) Kleinmolekül bindet noch umsetzt. Die lipase-ähnliche Domäne ist vielmehr optimiert, um Protein-Protein Wechselwirkungen innerhalb der EDS1 Familie zu vermitteln. Die Wechselwirkungen zwischen EDS1 und SAG101 wurden *in vitro* und *in vivo* charakterisiert und dazu benutzt, ein Homologie-Modell von PAD4 zu erstellen. Interaktionsdefiziente Varianten von EDS1, PAD4 und SAG101 wurden auf Basis dieser Strukturen vorgeschlagen und experimentell validiert. Diese stellen wertvolle Werkzeuge dar, um zwischen den biologischen Funktionen von einzelnen Proteinen und Protein-Komplexen innerhalb EDS1-abhängiger Signalwege zu unterscheiden.

Während die Interaktionen innerhalb der EDS1 Proteinfamilie vor allem von der lipase-ähnlichen Domäne bestimmt werden, scheint die C-terminale EP Domäne auf Sequenz- und Strukturebene einzigartig zu sein. Eine entfernte strukturelle Verwandtschaft zu Proteinen aus Multiprotein-Komplexen bedeutet möglicherweise, dass diese Domäne Interaktionen außerhalb der EDS1-Familie vermittelt. Dieser Aspekt könnte in Zukunft an Bedeutung gewinnen da zunehmend mehr EDS1 Interaktoren charakterisiert werden.

1. Introduction

"The extent of land is limited — that is perfectly true. But the labour power to be employed on this area increases along with the population; and even if we assume that the increase in yield due to this increase does not always rise in proportion to the labour, there still remains a third element — which the economists, however, never consider as important — namely, science, the progress of which is just as unceasing and at least as rapid as that of populations."

— Friedrich Engels, Outlines of a Critique of Political Economy, 1843

1.1. The Need For Resistant Plants

Almost 50 years before Friedrich Engels published his "Outlines of a Critique of Political Economy", the British reverend Thomas Malthus already discussed the dangers of population growth in his "Essays on the Principle of Population" [Malthus, 1798]. With a world population that was only reaching one billion people at the beginning 19th century, Malthus could have hardly imagined the world would accommodate seven times more people by 2012. Pointed out by Engels, scientific progress in agriculture has ever since impacted the attainable yields of food crops and provides the basis for today's world nutrition. Yet, given that the world population is projected to reach more than 9 billion people by 2050, food security remains a challenge. As agricultural crops have further gained importance as a source of renewable energy, the optimisation of crop yields is of worldwide concern.

Recently reviewed by Oerke [2006], crop losses are due to abiotic (irradiation, water supply, temperature, nutrients) or biotic factors. Analysing 19 worldwide regions, he estimated that the potential loss to competing weeds, animal pests and pathogens exceeds 50 % of the attainable crop yield. While established crop protection such as mechanical weeding and the use of pesticides improved the situation, farmers still loose about one third of their yield as a consequence of biotic stress [Oerke, 2006].

1.1.1. Genetically Engineered Crop Resistance

Since the first plant genome of *Arabidopsis thaliana* (hereafter referred to as *Arabidopsis*) was published in 2000 [Arabidopsis Genome Initiative, 2000], sequencing techniques have become more sophisticated and led to an exploding number of available genomes in recent years. Although heavily debated, this information is and will be used to genetically engineer plants with desired traits including the resistance to drought and pathogens.

Genetically modified crops have been dominated by two major traits: (1) the introduction of tolerance against the herbicide glyphosate and (2) an endotoxin that confers resistance against insect pests.

Glyphosate (or N-(phosphonomethyl)glycine) is a broad range herbicide that kills weeds which would otherwise compete with crop plants. It inhibits the largely plant specific enzyme 5-enol-pyruvylshikimate-3-phosphate synthase and thereby interferes with the synthesis of essential

amino acids. Glyphosate became extremely popular in the United States after the commercialisation of plants that carry a resistance against them in 1996. Since then, glyphosate has become one of the most widely used herbicides (Fig. 1.1). The specificity of glyphosate led to the anticipated decrease of herbicide application on U.S. acres during the first years after its introduction. Yet, from 2001 until today the amount of herbicides used on U.S. acres has increased ever since and was ultimately 70 % higher in 2011 than it was in 1996 [Benbrook, 2012] (Fig. 1.1).
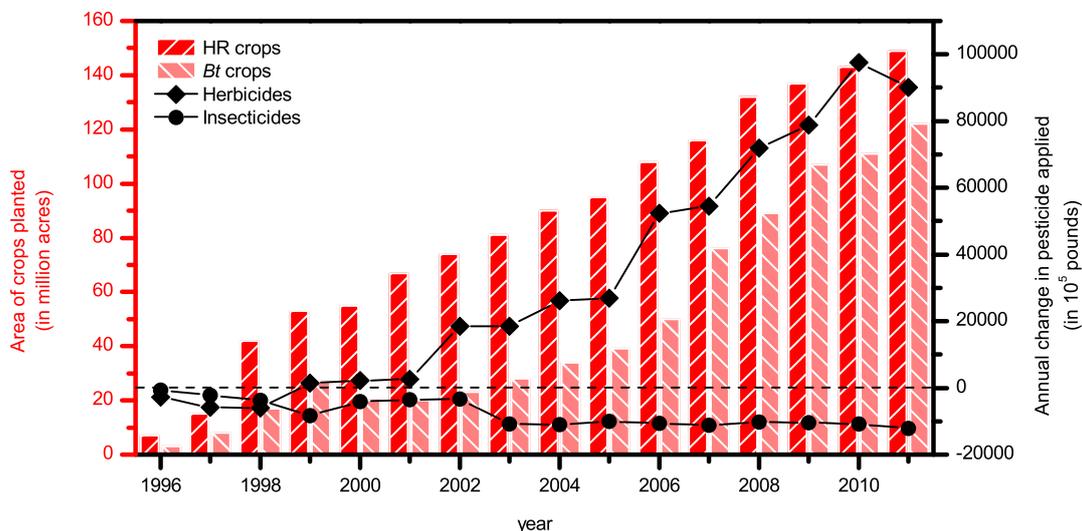
Figure 1.1 | Popularity of genetically modified crops in the U.S. and their impact on pesticide use. The area in the United States on which HR (herbicide resistant) and *Bt* (*Bacillus thuringiensis*) crops were planted (red bars) and the annual in- or decrease of the amount of pesticides (herbicides or insecticides) used on all U.S. acres (black lines) are plotted. Data are taken from Benbrook [2012].

The most important factor driving this trend is the occurrence of glyphosate-resistant weeds[1]. Reliance on a single herbicide creates a strong selection pressure towards the corresponding resistance in weeds that these can develop relatively fast due to short generation times. Farmers responded by increasing the herbicide application rates per year, leading to a doubling of glyphosate consumption between 1996 and 2006 [Benbrook, 2012].

Herbicide resistant weeds promoted the popularity of a different genetically engineered trait that enables crop plants to produce endotoxins that originate from the soil bacterium *Bacillus thuringiensis* (Fig. 1.1). The transgenic plants are therefore termed *Bt* crops. These endotoxins are insecticidal in that they, once activated, form pores in the membranes of insect gut cells that lead to death of these cells and ultimately of the insect. The widespread use of *Bt* crops in the United States indeed helped to constantly decrease the amount of insecticides sprayed on U.S. acres since 1996 (Fig. 1.1). Although different insects species had been already shown to develop a resistance against *Bt* endotoxins in the laboratory (those are reviewed in [Tabashnik *et al.*, 2003]), it was not until 2011 that also a field-evolved resistance was reported [Gassmann *et al.*, 2011].

The rise of pesticide ineffectiveness due to pathogen resistance underlines the importance to develop new crop protection methods. Intensive research is necessary to introduce resistance at

---

[1]http://www.weedscience.org/ tries to keep track of emerging resistances. (request date: January, 2013)

a faster rate than that by which the pests evolve counter-resistance. Instead of traits conferring herbicide resistance or such originating from non-plant organisms, these may as well rely on the innate immune systems that plants possess. For that, a comprehensive understanding of plant immunity and its components is necessary.

1.2. How Plants Defend Themselves

Plants, like humans, have evolved a multi-layered immune system. Despite striking similarities between them, they are differently organised.

The human immune system consists of three layers: (1) physical, chemical and biological barriers, (2) an innate branch of immunity and (3) the adaptive immune system.

Physical, chemical and biological barriers like our skin, antibacterial enzymes or the gut flora that competes with pathogenic bacteria, efficiently fend off attacks of most pathogens (see for example [Turvey and Broide, 2010]).

If this first line of defence fails, invading pathogens are recognised by cells of the innate immune system. These utilise a limited set of pattern recognition receptors that perceive microbial components, which are conserved among large groups of pathogens, the so-called pathogen-associated molecular patterns or PAMPs[2]. Upon PAMP recognition, complex signalling pathways become activated that culminate in the induction of a protective inflammatory response.

These pathways are also essential for the initiation of the third line of defence, the adaptive immune system. This layer of immunity utilises highly specialised cells that travel through the blood and lymphatic system and display structurally related receptors that are, however, extremely diverse regarding their capability to detect non-self molecules, the so-called antigens. Antigen recognition does not only lead to the elimination of the pathogen but also establishes an immunological memory that render the immune response upon recurrence of an antigen faster and more efficient.

Plants as well have three layers of immunity that do partially overlap: (1) preformed barriers, (2) a PAMP-triggered immune response and (3) the effector-triggered immunity [Jones and Dangl, 2006]. Unlike humans, plants lack an adaptive immune system but rely solely on an innate immune system whose components are present in every cell of the plant. Within the next sections, I will present molecular and structural details of plant immunity and will position the central immune regulator ENHANCED DISEASE SUSCEPTIBILITY1 (EDS1) within plant resistance signalling.

1.2.1. A First Layer of Inducible Plant Immunity Restricts Infection by Most Pathogens

The first obstacle an invading plant pathogen has to overcome are preformed barriers. Like the intact skin of humans, those include physical barriers such as the cell wall, leaf hairs (trichomes) and the plant cuticles (waxy layers that cover most of the plants surface). The mechanical

---

[2]Instead of PAMPs the term microbial-associated molecular patterns or MAMPs has been suggested since both pathogenic and non-pathogenic microbes produce MAMPs that would activate the immune response [Ausubel, 2005].

barriers are complemented with a chemical barrier made of phytoanticipins [Nürnberger and Lipka, 2005][3].

Pathogens that are able to overcome preformed defence mechanisms activate the first inducible layer of plant defence. Similar to the innate immune system of humans, the invading pathogen is recognised through conserved microbial molecules, the already introduced PAMPs. They include bacterial flagellin, chitin from the fungal cell wall or specific glucans from oomycetes [Altenbach and Robatzek, 2007]. The activated immune responses are collectively termed PAMP-triggered immunity.

In plants, 10 pairs of PAMPs and their respective pattern recognition receptors have been identified [Robatzek and Wirthmueller, 2012]. The probably best studied pair is bacterial flagellin that is recognised by the FLAGELLIN SENSING 2 (FLS2) receptor. The membrane-anchored FLS2 receptor comprises three domains: (1) an extracellular domain that consists of leucine-rich repeats (LRRs), (2) a membrane-spanning domain and (3) an intracellular kinase domain. FLS2 perceives flagellin through the horseshoe-shaped LRR domain [Gòmez-Gòmez and Boller, 2000]. This leads to a dimerisation of FLS2 with the receptor kinase BAK1 and mutual phosphorylation of BAK1 and the FLS2 kinase domains is thought to result in downstream signalling (Fig. 1.2 A) although Robatzek and Wirthmueller [2012] point out that "the quest for FLS2 crystal structures is open" and the precise mechanisms of FLS2 signal transduction are still elusive.

While the molecular details of pattern recognition receptor signalling are thus poorly understood, the resulting responses are comparatively well studied. Reviewed by Nicaise *et al.* [2009], these include strengthening of the preformed barriers through the deposition of callose into the plant cell wall and the closure of stomatas that constitute an entry point for the invading pathogens. Additionally, PAMP recognition leads to a complex genetic reprogramming of the plant cell through miRNA-mediated gene silencing and an activation of MAP kinase cascades that, in turn, regulate transcription factors of the WRKY-type. The sum of these outputs is in most cases sufficient to stop pathogen invasion. They further prime the plant's immune system through the establishment of a systemic acquired resistance that helps to effectively stop subsequent pathogen infection [Mishina and Zeier, 2007].

1.2.2. The Second Layer of Activated Plant Immunity Needs to Be Tightly Regulated

In a simplified view, plant pathogens have phylogenetically evolved so-called effector molecules that allow them to overcome a PAMP-triggered immunity through different strategies. One is to directly target the pattern recognition receptors or their associated proteins.

The model bacterium *Pseudomonas syringae* pv. tomato (*Pto*) DC3000 for instance produces a small protein, AvrPto, that directly interacts with the FLS2 receptor (see above) and prevents receptor phosphorylation, thereby suppressing an FLS2-mediated immune response [Xiang *et al.*, 2008, 2011]. This leads to an effector-triggered susceptibility of the plant (Fig. 1.2 B).

---

[3]VanEtten *et al.* [1994] propose to define phytoanticipins as molecules that are present in plants before challenge by microorganisms whereas phytoalexins are antimicrobial compounds that are both synthesised by and accumulated in plants after exposure to microorganisms.

Plants, in turn, have co-evolved a class of intracellular receptors to specifically recognise these effector molecules and activate another layer of immunity, the effector-triggered immune response. The involved immune receptors are termed R (resistance) proteins. They exist in every plant cell and belong almost exclusively to the NB-LRR class that generally comprises three domains: a C-terminal leucine-rich repeat (LRR) domain that we already know from the membranous pattern recognition receptors, a central nucleotide binding (NB) domain and N-terminally either a coiled-coil (CC) or a Toll/Interleukin-1 receptor (TIR) domain.

While a CC- and TIR-domain from barley and flax have been analysed structurally [Maekawa *et al.*, 2011; Bernoux *et al.*, 2011], the nucleotide binding and leucine-rich repeat domains of plant R proteins still need to be computationally modelled. Takken and Goverse [2012] have collected this fragmentary knowledge to propose a model for a functional plant R protein: the C-terminal LRR domain is thought to recognise the effector molecule and accordingly exhibits hypervariable amino acids that are under positive selection for effector perception [Seeholzer *et al.*, 2010]. This recognition can happen directly through physical interaction or indirectly by detecting the modification(s) caused by the effector on its host target.

Although cases of direct effector recognition have been reported by Dodds *et al.* [2006] and others, it is only hard to conceive that this is a general mechanism as *Arabidopsis* encodes only 149 NB-LRR proteins [Meyers *et al.*, 2003] while the introduced *Pseudomonas syringae* isolate *Pto* DC3000 already produces 39 effectors [Baltrus *et al.*, 2011]. One of these *Pto* effectors, the already introduced AvrPto, physically interacts with the intracellular protein kinase Pto that, in turn, needs the NB-LRR R protein Prf to activate defence signalling (Fig. 1.2 C).

This indirect type of effector recognition by an R protein can be explained through different models. In the guard model, Pto would be a general component of plant defence which is directly attacked by the AvrPto effector that tries to shut down Pto-mediated immunity. The guarding NB-LRR R protein Prf, however, would recognise the actions of AvrPto (instead of AvrPto itself) and trigger an immune response. This hypothesis was proposed by Van der Biezen and Jones [1998] and helps to explain how a limited set of R proteins can recognise a wide range of effectors.

The guard model was modified by Van der Hoorn and Kamoun [2008] who argue that it results in an evolutionary unstable situation as soon as populations of a plant species either possess or lack an R protein for successful effector recognition. This would lead to an opposing selection pressure on the guarded effector target (the guardee) if it is required for immune signalling: in the plant population lacking the R protein, evolutionary selection would drive the guardee to decrease affinity for the effector in order to evade detection and inhibition by the effector's action. Contrary, in a population where the R protein is present, evolutionary selection would lead to an increased affinity between the guardee and effector to enhance pathogen perception. They thus propose a decoy model in which the effector target (the decoy) is essential for effector recognition, yet not for immune signalling. It rather mimics the actual effector target and activates downstream immune pathways upon effector recognition. Although experimental evidence is lacking, the Pto kinase could in such a model act as a mimic for the FLS2 kinase domain.

Insights from a AvrPto/Pto complex crystal structure [Xing *et al.*, 2007] have helped to understand the subsequent signalling events. Pto, either as a guardee or a decoy, is assumed
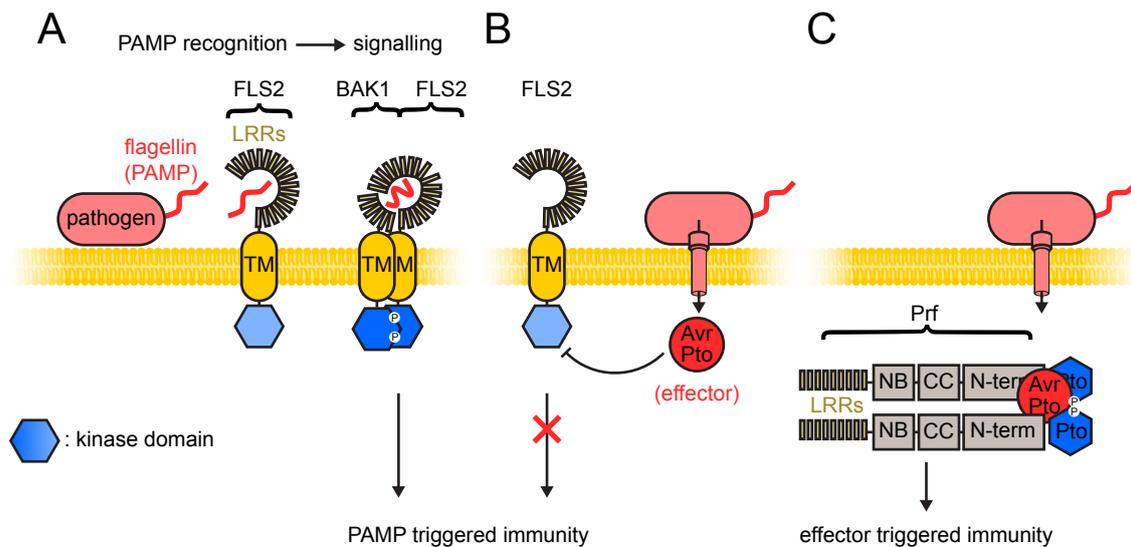
Figure 1.2 | PAMPs and effectors trigger different plant immune responses. (A) Bacterial flagellin can be recognised as a PAMP by the pattern recognition receptor FLS2. This leads to a dimerisation of the receptor with the receptor kinase BAK1 and mutual phosphorylation events that activate pathways leading to a PAMP-triggered immunity in an unknown way. (B) Pathogens have evolutionary evolved effector molecules like AvrPto that are injected into the plant cell and inhibit components of PAMP-triggered immunity (e.g. the FLS2 kinase domain). (C) These effectors are recognised by intracellular immune receptors that usually contain a C-terminal leucine-rich repeat (LRR) domain, a central nucleotide binding (NB) domain and varying types of N-terminal domains. AvrPro gets recognised through the concerted actions of the Pro kinase and the NB-LRR protein Pfr that together trigger an immune response. Figure adapted from [Chinchilla *et al.*, 2009] and [Ntoukakis *et al.*, 2013].

to act as a cofactor stabilising the inactive conformation of the NB-LRR protein Pfr. Upon the formation of an AvrPto-Pto complex, a structural change within the Pto kinase domain leads to a conformational rearrangement and activation of Pfr. Because there is neither ultimate evidence for the guard model nor the decoy model, this bait-and-switch model, with Pto acting as a bait that switches on Pfr upon AvrPto perception, has been proposed as a general mechanism for indirect effector recognition [Collier and Moffett, 2009]. The different modes of effector recognition are summarised in Figure 1.3.

The immune output following effector recognition differs significantly from a PAMP-triggered immune response in being generally faster and more prolonged [Tsuda and Katagiri, 2010]. Although both pathways share common features and cellular machineries [Kim *et al.*, 2012], effector-trigger immunity is usually associated with a so-called hypersensitive response that is rather uncommon (but possible, see [Naito *et al.*, 2008]) for an immune response triggered by PAMPs.

The hypersensitive response is characterised by local cell death at the site of pathogen infection, a mechanism comparable to the scorched earth policy used in warfare: to fend off an attack, anything that might be useful for the invader is destroyed. A hypersensitive response is thus only effective against biotrophic (such living on host tissue) but not necrotrophic (such killing cells and living from the remains) pathogens. Although plants are modular organisms that can kill single leaves without further consequences, a hypersensitive response needs tight regulation to
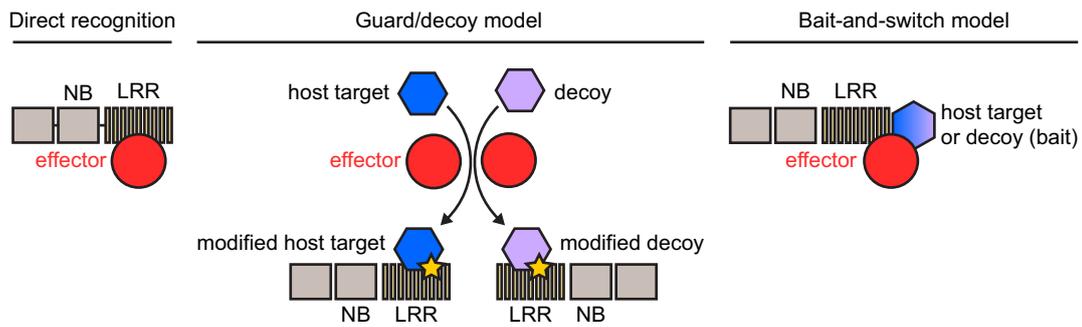
Figure 1.3 | Effector recognition models. Pathogen effectors can be recognised differently by intra-cellular plant immune receptors of the NB-LRR class. As direct recognition is assumed to be rather rare, the guard and decoy model have been proposed. In the guard model, the host target (blue) has a function in plant immunity and is thus attacked by an effector. The effector's activity (e.g. a modification of the host target) is perceived by an NB-LRR receptor that is able to guard more than one guardee. In the decoy model, the guarded molecule is not active in immune signalling but rather mimics a component of plant immunity luring the effector away from its actual target. In the bait-and-switch model, a bait molecule (either a guardee or a decoy) resides in complex with an immune receptors and resistance signalling is activated upon effector interaction.

prevent spreading cell death and to reduce the cellular costs of an induced resistance [McDowell and Simon, 2006].

1.3. How EDS1 Is Involved in Plant Defence Signalling

As described above, the 149 *Arabidopsis* NB-LRR R proteins can be divided into two groups: while CC-NB-LRR R proteins signal through the membrane-associated protein NDR1 (NON-RACE SPECIFIC DISEASE RESISTANCE 1), TIR-NB-LRR R proteins generally use the nucleo-cytoplasmatic protein EDS1 (ENHANCED DISEASE SUSCEPTIBILITY 1) as a signal transducer [Aarts *et al.*, 1998].

1.3.1. EDS1 Forms Molecularly and Spatially Distinct Complexes Within Plant Cells

EDS1 has been characterised in chemically mutagenised *Arabidopsis* plants that showed an enhanced disease susceptibility phenotype against an oomycete pathogen that causes downy mildew [Parker *et al.*, 1996], a plant disease common for grapevine, hop and other agricultural important plants. In a similar screen, Glazebrook *et al.* [1996] identified PAD4 (PHYTOALEXIN DEFICIENT 4) as another central immune component of *Arabdidopsis* against multiple pathogens and Feys *et al.* [2001] could demonstrate that EDS1 and PAD4 interact directly in yeast two-hybrid assays and in plant cells. Almost ten years after the identification of EDS1 and PAD4, another protein of then unknown function, SAG101 (SENESCENCE ASSOCIATED GENE 101), was shown to interact and function together with EDS1 in plant immunity [Feys *et al.*, 2005].

Complexes of EDS1 with PAD4 are detectable inside the cytoplasm and nuclei of plant cells while such of EDS1 and SAG101 exists predominantly in the nucleus [Feys *et al.*, 2005; Zhu *et al.*, 2011]. EDS1 is further able to interact with itself *in planta* and does so primarily in the cytoplasm as resolved by measuring FRET signals between transiently expressed fluorescent

protein-tagged proteins in individual *Arabidopsis* cell compartments [Wiermer, 2005; Feys *et al.*, 2005]. Analysis of protein accumulation in plants that lack either EDS1, SAG101, PAD4 or combinations of them revealed that the proteins are stabilised through their interaction partners [Feys *et al.*, 2005].

Analytical size exclusion chromatography using pathogen-unchallenged *Arabidopsis* leaf extracts suggests that EDS1 forms preferentially homodimers with itself (~ 143 kDa) and heterodimers with SAG101 (~ 133 kDa) [Feys *et al.*, 2005]. Judging from these experiments, EDS1 exists additionally as a monomer (~ 72 kDa) and in a higher molecular mass complex with PAD4 (~ 200 kDa) that could consist of two EDS1 and one PAD4 molecules (or could include yet unknown components).

A recent study further proposes a ternary complex in which EDS1 bridges between PAD4 and SAG101 if the proteins are transiently expressed in tobacco plants [Zhu *et al.*, 2011]. In contrast, a ternary complex could not be shown in yeast two-hybrid assays or *in vitro* between recombinantly produced proteins [Rietz *et al.*, 2011]. This illustrates that there are still open questions concerning the assembly of EDS1-requiring complexes. The next paragraphs will illustrate that these complexes, but additionally the single components, are involved in various plant signalling pathways.

1.3.2. EDS1 Complexes Fulfil Different Functions in Plant Immunity

EDS1 Acts in Multiple Plant Resistance Types Both EDS1 and PAD4 were shown to be required for *Arabidopsis* resistance conferred by R proteins of the TIR-NB-LRR type (e.g. RPP2, RPP4, RPS4) but dispensable for immune signalling through CC-NB-LRR R proteins (e.g. RPS2, RPS5) [Glazebrook *et al.*, 1997; Parker *et al.*, 1996; Zhou, 1998; Aarts *et al.*, 1998]. This seems to be conserved among higher plants as EDS1 is also required for immunity that is mediated by the TIR-type R protein N but not the CC-type R protein Rx in tobacco plants [Peart *et al.*, 2002]. As a CC-NB-LRR protein, which confers resistance against the biotrophic turnip crinkle virus, and an atypical R protein, which protects *Arabidopsis* from the powdery mildew plant disease, additionally depend on EDS1, this discrimination might be an oversimplification [Xiao *et al.*, 2005; Chandra-Shekara *et al.*, 2004].

RPP (RESISTANCE TO PERONOSPORA PARASITICA) and RPS (RESISTANCE TO PSEU-DOMONAS SYRINGAE) TIR-NB-LRR R proteins recognise effectors of the biotrophic oomycete *Hyaloperonospora arabidopsidis* (formerly *Peronospora parasitica* and hereafter referred to as *Hyaloperonospora*) and of the biotrophic bacterium *Pseudomonas syringae*.

Phenotypically, *Arabidopsis* plants lacking EDS1 (*eds1*) show a complete loss of TIR-NB-LRR-mediated resistance characterised by the absence of localised cell death after *Hyaloperonospora* infection (Fig. 1.4). In contrast, mutant plants missing PAD4 (*pad4*) retain a limited hypersensitive response that is, however, insufficient to halt pathogen growth completely. This leads to a trailing necrosis along the path of pathogen growth (Fig. 1.4 and [Feys *et al.*, 2001]).

An *Arabidopsis pad4* mutant phenotype that is an intermediate between an immune response of wild-type plants and such that lack EDS1 (*eds1*) can also be observed if the plants are infected with *Pseudomonas syringae* bacteria expressing the effector AvrRps4 that is recognised by the EDS1-dependent TIR-NB-LRR protein RPS4 (Fig. 1.4).

| | Oomycete infection (Hyaloperonospora) | | Bacterial infection (Pseudomonas) |
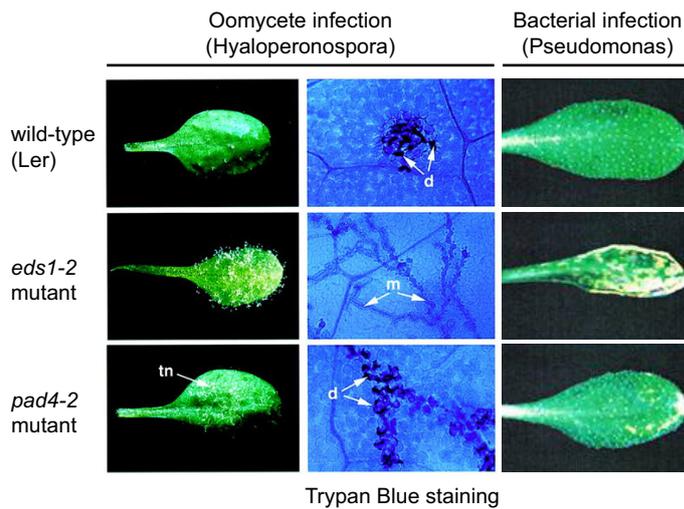
Figure 1.4 | Phenotypes of eds1 and pad4 mutant plants. Wild-type *Arabidopsis* plants of accession Landsberg *erecta* (Ler) were inoculated with a *Hyaloperonospora* cell suspension. Leaves were photographed 6 d after inoculation. Trailing necrosis (tn) in *pad4-2* is indicated by an arrow. Leaves were stained with Trypan Blue 5 days after inoculation to visualise pathogen mycelium (m) and dead plant cells (d). For bacterial infection, leafs were dipped in a cell suspension of *Pseudomonas syringae* pv. tomato (*Pto*) DC3000 expressing the effector AvrRps4 and photographed after 6 d. Figures taken from [Feys *et al.*, 2001].

In contrast, *sag101* mutant plants show a wild-type like hypersensitive response in similar assays [Feys *et al.*, 2005]. However, a *pad4/sag101* double mutant plant is as susceptible as an *eds1* plant arguing for a redundancy between PAD4 and SAG101 in TIR-NB-LRR R protein mediated immune signalling [Feys *et al.*, 2005]. The same is true for so-called basal resistance [Feys *et al.*, 2005].

As this will become important, I will briefly explain the different modes of interaction between *Arabidopsis* and the oomycete *Hyaloperonospora*. As reviewed by Coates and Beynon [2010], the interaction between *Arabidopsis* accessions and *Hyaloperonospora* isolates show extreme levels of genotypic and phenotypic variation due to a constant co-evolution between host and pathogen. One can discriminate between compatible and incompatible interactions and (if an *Arabidopsis* plant from the same accession is used) between compatible and incompatible *Hyaloperonospora* isolates. An incompatible isolate activates a hypersensitive response through the recognition of an effector by the corresponding R protein and will be efficiently restricted in growth. The interaction between *Hyaloperonospora* isolate Noco2 and *Arabidopsis* accession *Landsberg erecta* (Ler) in Figure 1.4 is a perfect example. Absence of effector recognition through the respective R protein results in a compatible interaction and allows pathogen growth. Importantly, residual defence mechanisms remain active during an compatible interaction and result in a basal resistance of the plant. Jones and Dangl [2006] define basal resistance as the additive effects of a PAMP-triggered and weak effector-triggered immunity minus those of an effector-triggered susceptibility.

Both EDS1 and PAD4 have been shown to be required for the accumulation of the plant hormone salicylic acid in an effector-triggered immune response mediated by TIR-NB-LRR (but not CC-NB-LRR) R proteins [Falk *et al.*, 1999; Feys *et al.*, 2001; Zhou, 1998] and in basal resistance. As transcription of both proteins is, in turn, upregulated by the application of salicylic acid [Falk *et al.*, 1999; Jirage *et al.*, 1999], EDS1 and PAD4 are thought to act in a positive feedback loop leading to increased salicylic acid levels. Salicylic acid is known to act in the establishment of the systemic acquired resistance that primes the plant's immune system against

further pathogen attacks [Métraux *et al.*, 1990; Vlot *et al.*, 2009]. EDS1 thus functions in effector-triggered, basal and systemic acquired resistance.

Spatially and Molecularly Distinct EDS1 Complexes Discriminate Between Functional States Recent studies suggest that EDS1 and PAD4 exist in a complex and as separate entities and that these states discriminate between individual functions [Rietz *et al.*, 2011]. An EDS1 variant (EDS1L262P) that is unable to interact with PAD4 but still able to bind SAG101 was therein used to show that physical interaction between EDS1 and PAD4 is necessary for a basal and systemic immune response but not for TIR-NB-LRR-triggered resistance. The ability of individual EDS1 and PAD4 to establish a TIR-NB-LRR-triggered cell death is not due to SAG101 compensation as the same phenotypes have been shown in a *sag101* mutant background [Rietz *et al.*, 2011].

Functions within EDS1-related complexes are further discriminated by cellular localisation: García *et al.* [2010] demonstrated that EDS1 is actively shuttled between the cytoplasm and the nucleus of plant cells through components of the nuclear pore complex. The authors could show that EDS1 is accumulated in the nucleus upon pathogen infection and that nuclear EDS1 drives a TIR-NB-LRR-mediated transcriptional reprogramming, including genes that are essential for the synthesis and signalling of salicylic acid. While the role of a cytoplasmic EDS1 pool is less clear, it may act in counterbalancing the actions of nuclear EDS1 [García *et al.*, 2010]. Active transport of EDS1 between both compartments and the restriction of SAG101 to the nucleus however imply that proper protein localisation is required for full functionality. The inability of SAG101 to fully compensate for the loss of PAD4 in TIR-NB-LRR-mediated and basal resistance (compared to the *vice versa*, see above) might result from a cellular imbalance of EDS1 heterocomplexes and further supports this claim. Proposed EDS1-containing complexes and their putative role in TIR-NB-LRR-dependent immune signalling is schematically depicted in Figure 1.5.

1.3.3. Manifold Stress Pathways Rely on EDS1, PAD4 and SAG101

The role of localised cell death that is mediated by EDS1 and PAD4 extends beyond a plant's resistance against biotrophic pathogens (*Pseudomonas*, *Hyaloperonospora* and others). When herbivorous insects deposit their eggs on *Arabidopsis* and other plants, these use a cell death mechanism that shares similarities with a hypersensitive response triggered by effector recognition [Little *et al.*, 2007] and these mechanisms have recently been shown to depend on EDS1 [Gouhier-Darimont *et al.*, 2013].

Mateo *et al.* [2004], Rustérucci *et al.* [2001] and others observed that EDS1 and PAD4 are moreover required in the regulation of cell death upon plant photo-oxidative stress that can be caused by high light. The authors analysed *Arabdiopsis lsd1* (lesions simulating disease 1) mutant plants that fail to stop the initiation and propagation of cell death upon light stress or application of superoxide ions and found that all traits (e.g. reduced catalase activity and stomatal conductance) were dependent on EDS1 and PAD4. As Straus *et al.* [2010] could reveal that cell death upon photo-oxidative stress in another *Arabdiopsis* mutant (*nudt7*) is as well EDS1-dependent, they proposed that EDS1 has a "master" role in coordinating biotic but also abiotic stress stimuli.
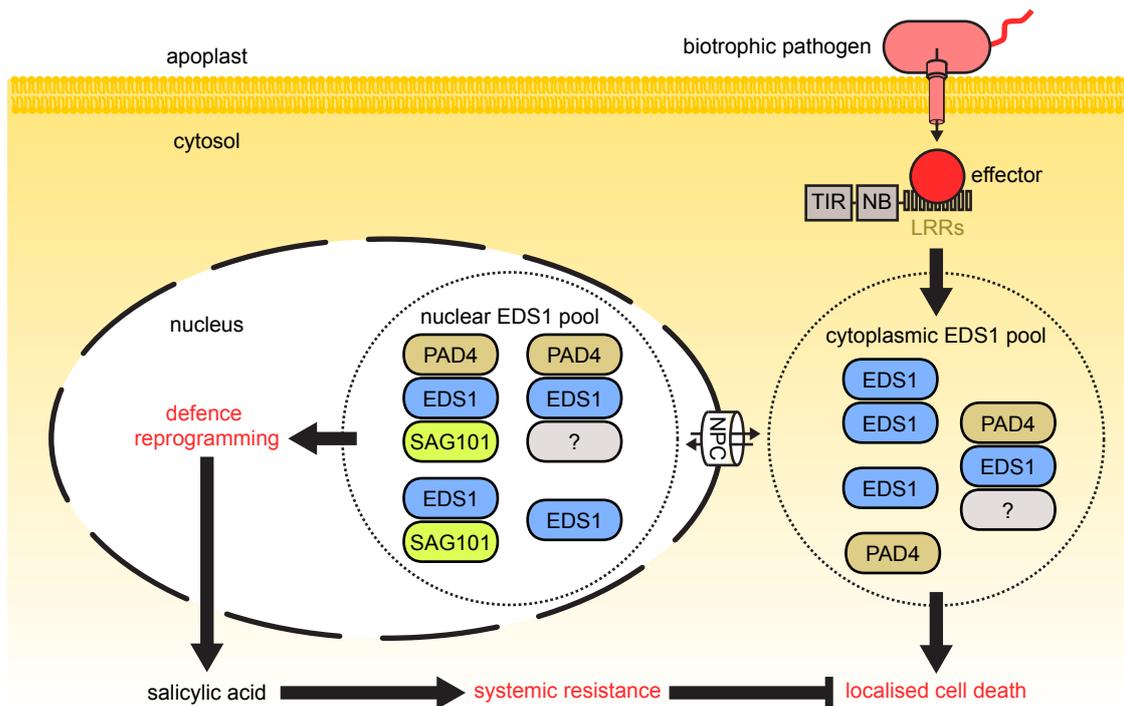
Figure 1.5 | Hypothetical model of EDS1-related protein functions in TIR-NB-LRR-mediated resistance. EDS1 is a nucleo-cytoplasmatic protein. Proposed by Feys *et al.* [2005], EDS1 dimers exist alongside heterocomplexes with PAD4 inside the cytosol while the plant cell nucleus contains EDS1/SAG101 and EDS1/PAD4 heterocomplexes. The authors further found EDS1 monomers but did not analyse their cellular localisation. Zhu *et al.* [2011] proposed a nuclear EDS1/PAD4/SAG101 ternary complex for that, however, neither Feys *et al.* [2005] nor Rietz *et al.* [2011] found further evidence. I tried to integrate the current literature into a very schematic EDS1 working model. Upon effector recognition through an TIR-NB-LRR R protein, EDS1 and PAD4 (probably separated from another) "kick-start" the resistance response and trigger a localised cell death at the infection site [Rietz *et al.*, 2011]. EDS1 in an unknown state is then actively shuttled into the nucleus through the nuclear pore complex (NPC) and triggers transcriptional reprogramming of resistance genes [García *et al.*, 2010]. This leads to an accumulation of the plant hormone salicylic acid which establishes a systemic acquired resistance in neighbouring cells that do not undergo programmed cell death [Rietz *et al.*, 2011].

In addition to EDS1-requiring complexes, the EDS1 binding partners SAG101 and PAD4 seem to fulfil functions independent of EDS1. SAG101 has been shown to be essential for a defence response of *Arabidopsis* against the biotrophic turnip crinkle virus [Zhu *et al.*, 2011] and PAD4 acts independently of EDS1 in the resistance of *Arabidopsis* plants against feeding of the green peach aphid, an insect that is considered an agricultural pest as it transmits plant viruses [Pegadaraju *et al.*, 2007, 2005].

1.3.4. The Biochemical Role of the EDS1 Protein Family Remains Elusive

Interestingly, Louis *et al.* [2012] observed that some, yet not all, resistance functions of PAD4 against green peach aphid feeding seem to depend on potentially catalytic residues within the N-terminal half of PAD4. Both, PAD4 and EDS1, had long before been shown to be sequentially related to hydrolases (or, more precisely, lipases) [Falk *et al.*, 1999; Jirage *et al.*, 1999]. The similarity of N-terminal EDS1 and PAD4 to eukaryotic lipases encompasses three conserved

residues in *Arabdiopsis* EDS1 and PAD4 that may represent a catalytic serine-aspartate-histidine (S-D-H) triad that is essential for the function of many hydrolytic enzymes. Louis *et al.* [2012] exchanged the serine within the PAD4 triad to an alanine and expressed this protein variant stably in *Arabidopsis* instead of wildtype PAD4. The authors observed that pathogenic green peach aphid insects spent significantly more time feeding on these mutant plants, arguing for a hydrolytic activity of PAD4 in this particular response. Apart from this finding, the role of EDS1, PAD4 and SAG101 enzymatic activity is largely unknown.

In addition to the N-terminal lipase-like domain, Falk *et al.* [1999] and Jirage *et al.* [1999] reported that EDS1 and PAD4 possess a C-terminal domain for which they found no sequence homologues. Feys *et al.* [2001] recognised that the C-terminal domains of EDS1 and PAD4 are closely related to each other and coined the term EP-domain (for EDS1-PAD4 defined). Already then, they reported that SAG101 is the only other *Arabidopsis* protein that is likewise organised and after showing a physical interaction between EDS1 and SAG101, they propose a domain organisation as depicted in Figure 1.6 [Feys *et al.*, 2005]. As the EP domain seems to be unique among plant and non-plant proteins, EDS1, PAD4 and SAG101 can be considered to form a protein family and they will be thus collectively referred to as the EDS1 (protein) family throughout this thesis.
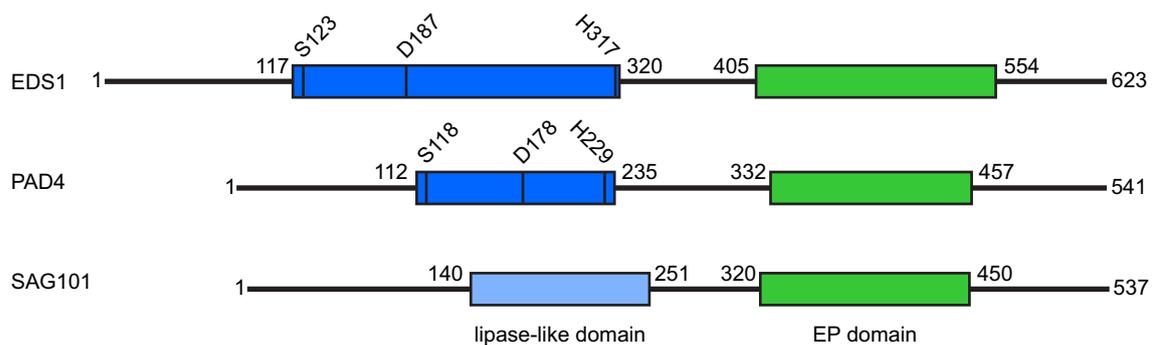
Figure 1.6 | Domain organisation within the EDS1 protein family. Boundaries of the lipase-like and EP domain are shown for the *Arabidopsis* variants of EDS1, PAD4 and SAG101. Within the lipase-like domains of EDS1 and PAD4, three conserved residues that might correspond to a Serine-Aspartate-Histidine catalytic triad are marked. Figure adapted from [Feys *et al.*, 2001].

1.3.5. Emergence of Further EDS1 Interactors

While the interactions within the EDS1 protein family have been intensively studied by yeast two-hybrid assays and *in planta* [Feys *et al.*, 2001, 2005; Rietz *et al.*, 2011; Zhu *et al.*, 2011], recent studies have piled evidence that the family members could additionally interact with other proteins. Heidrich *et al.* [2011] and Bhattacharjee *et al.* [2011] demonstrated that EDS1 physically interacts with different TIR-NB-LRR immune receptors (RPS4, RPS6 and SNC1) and two corresponding, yet sequence-unrelated, effector proteins, AvrRps4 (perceived by RPS4) and HopA1 (perceived by RPS6) that disrupt the EDS1/TIR-NB-LRR complexes. The authors suggest that EDS1, possibly as an immune component in PAMP-triggered immunity (and thus an effector target), is guarded by the interacting TIR-NB-LRR R protein (see Fig. 1.3) that initiates an effector-triggered immune response upon EDS1/effector interaction.

While the immune receptors RPS4 and RPS6 did not interact directly or indirectly with PAD4 [Bhattacharjee *et al.*, 2011], another study by Kim *et al.* [2012] found that the TIR-NB-LRR R protein VICTR (VARIATION IN COMPOUND TRIGGERED ROOT growth response) does not only interact with EDS1 but also PAD4 (probably in a co-complex). Intriguingly, the interactions between VICTR, EDS1 and PAD4 are important in the arrest of *Arabidopsis* root growth upon perception of a small molecule that the authors identified by chemical screening. Although the precise role of this small molecule needs to be determined, their work suggests that root growth (probably in response to soil-born pathogens) is yet another plant signalling pathway regulated in dependence of the EDS1 protein family.

1.4. Thesis Aims

The preceding paragraphs have shown that the EDS1 protein family is a central signalling hub involved in a variety of plant processes including (1) effector-triggered immunity, (2) basal immunity, (3) systemic acquired resistance, (4) regulation of cell death upon photo-oxidative stress, (5) insect resistance and (6) root growth.

With a single exception – the role of PAD4 in insect resistance [Louis *et al.*, 2012] – these functions have not been linked to molecular details. Although the relation to lipases and the possibility that any of the proteins transduces a lipid-derived molecule has been widely discussed in the literature, this is not further supported by experimental evidence [Wiermer *et al.*, 2005]. While sequence analysis additionally allowed to predict functional EDS1 family features including nuclear localisation and nuclear export signals [Falk *et al.*, 1999; García *et al.*, 2010], it could not further explain the biochemical mode of action and it is yet unclear how the EDS1 protein family is activated (e.g. after effector recognition by TIR-NB-LRR R proteins) and how they transduce downstream signals.

By analysing a limited set of proteins, Chothia and Lesk [1986] showed what has become a dogma in structural biology, namely, that structure is generally more conserved than sequence. This holds for more comprehensive analyses [Illergård *et al.*, 2009] and implies that proteins can adopt similar folds despite having different sequences. My ultimate goal was therefore a comprehensive structure-function relationship of EDS1 and its signalling partners PAD4 and SAG101. These structures should answer how closely the proteins are related to lipases at the structural level and should yield insights into the unique EP domain that has no obvious sequence homologues. To solve the structures, methods for the preparation of the proteins had to be established. Stable variants of each protein and combinations of them should be used in protein crystallisation experiments as a prerequisite for structure determination by means of X-ray diffraction methods. Furthermore, the complexes should be characterised biophysically to reveal insights into how they behave in solution and how they dynamically interact with each other.

2. Material and Methods

"Every honest researcher I know admits he's just a professional amateur. He's doing whatever he's doing for the first time. That makes him an amateur. He has sense enough to know that he's going to have a lot of trouble, so that makes him a professional."

— Charles F. Kettering

2.1. Materials

2.1.1. Chemicals

Laboratory grade chemicals and reagents were purchased from Sigma-Aldrich, Roth, Merck, Serva or Applichem unless otherwise stated.

2.1.2. Enzymes

All enzymes used are listed in Table 2.1. These were stored at -20 °C and freshly thawed before use.

Table 2.1 | Enzymes and corresponding buffers.

Enzyme	Buffer	Supplier
DNase I	See section 2.3.1	Boehringer
Gateway® BP Clonase™ II Mixture	Supplied 10 × reaction buffer	Invitrogen/Life Technologies
Gateway® LR Clonase™ II Mixture	Supplied 10 × reaction buffer	Invitrogen/Life Technologies
Lysozyme	See section 2.3.1	Merck
*Pfu*Turbo-DNA-Polymerase	10 × PCR reaction buffer	Stratagene
Mucor miehei lipase	See section 2.4.8	Sigma-Aldrich
Restriction endonucleases	Supplied 10 × reaction buffer	New England Biolabs
T4 DNA ligase	Supplied T4 DNA ligation buffer	Roche
Taq DNA polymerase	10 × PCR reaction buffer	MPIPZ, Köln
Trypsin	Various	Merck

2.1.3. Antibodies

Table 2.2 lists the primary and secondary antibodies used for immunoblot detection. Stock solutions of antibodies that were not enzyme-conjugated were stored at -20 °C and freshly thawed before use while enzyme-conjugates of antibodies were stored in dark vials at 4 °C. Dilutions of antibodies were kept at 4 °C and used for several days while such that were not conjugated with horseradish peroxidase (HRP) were preserved through the addition of sodium azide at a final concentration of 0.02 % (w/v).

2.1.4. Oligonucleotides

Oligonucleotides were purchased from Sigma-Aldrich, Operon or Metabion. Lyophilised primers were resuspended in bidistilled water to a final concentration of 100 µM. Working solutions were diluted to 10 pmol/µl (=10 µM).

Table 2.2 | Used antibodies.

Primary antibody	Source	Dilution	Supplier
α-c-Myc	Mouse monoclonal	1:5000	Santa Cruz Biotechnology
α-EDS1	Rabbit polyclonal	1:1000	S. Rietz; J. Parker (MPIPZ)
α-HA 3F10	Rat monoclonal	1:5000	Roche
α-His	Mouse monoclonal	1:2000	Roche

Secondary antibody	Conjugated with	Dilution	Supplier
Goat anti-mouse IgG	Alkaline phosphatase	1:5000-1:10000	Santa Cruz Biotechnology
Goat anti-rabbit IgG	Alkaline phosphatase	1:5000-1:10000	Santa Cruz Biotechnology
Goat anti-rat IgG	Alkaline phosphatase	1:5000-1:10000	Santa Cruz Biotechnology
Goat anti-mouse IgG	Horseradish peroxidase	1:5000-1:10000	Santa Cruz Biotechnology
Goat anti-rabbit IgG	Horseradish peroxidase	1:5000-1:10000	Santa Cruz Biotechnology
Goat anti-rat IgG	Horseradish peroxidase	1:5000-1:10000	Santa Cruz Biotechnology

2.1.5. Buffers and Solutions

Buffers and solutions frequently used in this thesis are listed in Table 2.3 or given in the corresponding text section if this is more appropriate. All buffers and solutions were prepared with bidistilled water and such used for molecular biological experiments were generally autoclaved or sterile-filtered.

Table 2.3 | Solutions and buffers.

DNA gelelectrophoresis	
$10 \times$ running buffer	0.4 M Tris, 0.2 M acetic acid, 10 mM EDTA, pH 8.5
$6 \times$ loading buffer	40 % (w/v) sucrose, 0.5 M EDTA,
	0.2 % (w/v) bromophenol blue
DNA molecular weight marker	10 % (v/v) $6 \times$ loading buffer, 5 % (v/v) 1 Kb DNA
	ladder (Roth)
PCR	
dNTP mix	2.5 mM dATP, 2.5 mM dCTP, 2.5 mM dGTP,
	2.5 mM dTTP
$10 \times$ PCR reaction buffer	500 mM KCl, 100 mM Tris, 15 mM $MgCl_2$,
	1 % (v/v) Triton X-100, pH 9.0
Protein gelelectrophoresis	
$20 \times$ MES running buffer	1 M MES, 1 M Tris, 2 % (w/v) SDS, 0.6 % (w/v) EDTA
$10 \times$ Tris-glycine running buffer	250 mM Tris, 1.92 M glycine, 1 % (w/v) SDS
$2 \times$ SDS sample buffer	60 mM Tris pH 6.8, 4 % (w/v) SDS, 200 mM DTT,
	20 % (v/v) glycerol, 0.2 % (w/v) bromophenol blue
$6 \times$ SDS sample buffer	60 mM Tris pH 6.8, 12 % (w/v) SDS, 600 mM DTT,
	47 % (v/v) glycerol, 0.6 % (w/v) bromophenol blue
Staining solution [Fairbanks *et al.*, 1971]	25 % (v/v) isopropanol, 10 % (v/v) acetic acid,
	0.04 % (w/v) Coomassie Brilliant Blue G-250
Destaining solution	25 % (v/v) ethanol, 8 % (v/v) acetic acid

Staining solution [Kang *et al.*, 2003]	10 % (v/v) ethanol, 5 % (w/v) aluminium sulfate, 2 % (v/v) orthophosphoric acid, 0.02 % (w/v) Coomassie Brilliant Blue G-250
Destaining solution	10 % (v/v) ethanol, 2 % (v/v) orthophosphoric acid

Immunoblotting

TBS buffer	10 mM Tris, 150 mM NaCl, pH 7.5
TBS-T buffer	10 mM Tris, 150 mM NaCl, 0.05 % (v/v) Tween 20, pH 7.5
10 × transfer buffer	250 mM Tris, 1.92 M glycine, 1 % (w/v) SDS, 10 % (v/v) Methanol
Ponceau S	Dilution of ATX Ponceau concentrate (Fluka) 1:5 in water
AP reaction buffer	50 mM Tris, pH 9.5, 100 mM NaCl, 0.5 mM $MgCl_2$

Protein purification

Anion exchange chromatography

AEC buffer A	50 mM buffer substance (e.g. HEPES or BICINE), 1 mM DTT, 0.5 mM EDTA, pH 8.0
AEC buffer B	50 mM buffer substance (e.g. HEPES or BICINE), 1 mM DTT, 0.5 mM EDTA, 1 M NaCl, pH 8.0

Lysis and immobilized metal-affinity chromatography

IMAC lysis buffer	50 mM HEPES, 150 mM NaCl, 20 mM imidazole, 1 mM DTT, 10 % (v/v) glycerol, pH 8.0
IMAC elution buffer	50 mM HEPES, 150 mM NaCl, 250 mM imidazole, 1 mM DTT, 10 % (v/v) glycerol, pH 8.0

Protein purification from plants (adapted from [Witte *et al.*, 2004])

Ex-Strep buffer	100 mM Tris, pH 8.0, 5 mM EGTA, 5 mM EDTA, 150 mM NaCl, 10 mM DTT, plant protease inhibitor cocktail (MPIPZ), 0.5 % (v/v) Triton X-100, 100 µg/ml avidin
W-Strep buffer	50 mM Tris, pH 8.0, 2.5 mM EDTA, 150 mM NaCl, 2 mM DTT, 0.05 % (v/v) Triton X-100
Elu-Strep buffer	10 mM Tris pH 8.0, 10 mM desthiobiotin, 2 mM DTT, 0.05 % (v/v) Triton X-100

2.1.6. Recombinant Expression

Different strains of *E. coli* were used for plasmid transformation and protein expression. Frequently used strains are given in Table 2.5.

Genes of interest were cloned into commercially available plasmids that are listed in Table 2.6 and often allow to encode affinity-tag fusions of the corresponding protein product.

Table 2.4 | Crystallisation screens.

Name	No. of conditions	Supplier
Crystal Screen	48	Hampton Research
Crystal Screen 2	48	Hampton Research
Crystal Screen Cryo	48	Hampton Research
Crystal Screen HT	48	Hampton Research
Grid Screen Ammonium sulfate	24	Hampton Research
Grid Screen MPD	24	Hampton Research
Grid Screen PEG/LiCl	24	Hampton Research
Grid Screen PEG6000	24	Hampton Research
Grid Screen Sodium chloride	24	Hampton Research
Grid Screen Sodium malonate	24	Hampton Research
JBScreen Classic 1-10	240	Jena Bioscience
JBScreen Kinase	96	Jena Bioscience
JCSG Core Suite I	96	Qiagen
JCSG Core Suite II	96	Qiagen
JCSG Core Suite III	96	Qiagen
JCSG Core Suite IV	96	Qiagen
PEG/Ion	48	Hampton Research
SaltRx	96	Hampton Research

Table 2.5 | Untransformed *E.coli* strains.

Name	Supplier	Genotype
E. coli BL21(DE3)	Invitrogen	F⁻ *ompT hsdS*(r$_B^-$ m$_B^-$) *gal dcm* (DE3)
E. coli DH5α	Invitrogen	F⁻ *end*A1 *thi*-1 *rec*A1 *rel*A1 *gyr*A96 φ80d*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *pho*A *sup*E44 *hsd*R17(r$_K^-$) m$_K^+$ λ-
E. coli DH10B	Invitrogen	F⁻ *end*A1 *rec*A1 *gal*U *gal*K *nup*G *rps*L Δ*lac*X74 φ80*lac*ZΔM15 *ara*D139 Δ(*ara leu*)7697 *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) λ-
E. coli Rosetta™(DE3)	Novagen	F⁻ *ompT hsdS*$_B$(r$_B^-$ m$_B^-$) *gal dcm* (DE3) pRARE (Cam^R)

Table 2.6 | Empty plasmids for *E.coli* transformation.

Name	Supplier	Used for	Resistance
pCDFDuet-1	Novagen	Co-expression; expression as an N-terminal 6His-fusion	Streptomycin
pDEST™15	Invitrogen	Expression as an N-terminal GST-fusion	Ampicillin
pDEST™17	Invitrogen	Expression as an N-terminal 6His-fusion	Ampicillin
pET22b	Novagen	Expression of a C-terminal 6His-fusion	Ampicillin
pET42a(+)	Novagen	Expression of a C-terminal 6His-fusion	Kanamycin
pET59-DEST™	Invitrogen	Expression as an N-terminal 6His-Trx-fusion	Ampicillin
pMAL-c2x	New England Biolabs	Expression as an N-terminal MBP-FactorXa-fusion	Ampicillin
pRSFDuet-1	Novagen	Co-expression; expression as an N-terminal 6His-fusion	Kanamycin

Successfully transformed strains that were used to express frequently needed protein variants are summarised in Table 2.7. The number (No.) refers to the "Stammsammlung *E. coli*" list used by the Niefind group.

Table 2.7 | Selected transformed *E. coli* strains.

No.	Strain	Plasmid	Insert	Resistance
33	Rosetta™(DE3)	pET22b	EDS1$^{1\text{-}623}$-8His	Chloramphenicol, ampicillin
35	Rosetta™(DE3)	pET42a(+)	EDS1$^{1\text{-}384}$-8His	Chloramphenicol, kanamycin
45	Rosetta™(DE3)	pRSF-Duet	6His-SAG101$^{1\text{-}537}$	Chloramphenicol, kanamycin
47	Rosetta™(DE3)	pRSF-Duet	EDS1$^{1\text{-}623}$ w/o tag	Chloramphenicol, kanamycin
56	Rosetta™(DE3)	pET42a(+)	EDS1$^{1\text{-}623}$L262P-8His	Chloramphenicol, kanamycin
59	Rosetta™(DE3)	pET42a(+)	EDS1$^{1\text{-}616}$-8His	Chloramphenicol, kanamycin
61	Rosetta™(DE3)	pET42a(+)	SAG101$^{291\text{-}537}$-8His	Chloramphenicol, kanamycin
64	Rosetta™(DE3)	pCDF-Duet	SAG101$^{1\text{-}537}$ w/o tag	Chloramphenicol, streptomycin
68	Rosetta™(DE3)	pET59-DEST	6His-Trx-EDS1$^{1\text{-}623}$	Chloramphenicol, ampicillin
112	Rosetta™(DE3)	pET42a(+)	*Aa*EDS1$^{1\text{-}623}$-8His	Chloramphenicol, kanamycin
114	Rosetta™(DE3)	pET22b	EDS1$^{1\text{-}362}$-8His	Chloramphenicol, ampicillin
118	Rosetta™(DE3)	pET22b	EDS1$^{1\text{-}384}$ w/o tag	Chloramphenicol, ampicillin

Media that were used to cultivate *E. coli* cells are given in Table 2.8. All media were prepared with deionised water and autoclaved before use. Non-autoclavable components were sterile-filtered. In most cases, these media were supplemented with antibiotics that are listed in Table 2.9 and were stored at -20 °C as 1000-fold concentrated stock solutions.

Table 2.8 | Media for *E. coli* cultivation.

Name	Recipe for 1 litre
Lysogeny broth (LB)	10 g peptone, 10 g NaCl, 5 g yeast extract, pH 7.5
LB agar	10 g peptone, 10 g NaCl, 5 g yeast extract, 15 g agar, pH 7.5
Terrific broth (TB)	12 g peptone, 24 g yeast extract, 4 ml glycerol, 100 ml 0.17 M KH_2PO_4 and 0.72 M K_2HPO_4 (add sterile after autoclaving)

Table 2.9 | Antibiotics used in *E.coli* media.

Name	Stock concentration
1000 × ampicillin	100 mg/ml ampicillin in water
1000 × chloramphenicol	50 mg/ml chloramphenicol in ethanol
1000 × kanamycin	50 mg/ml kanamycin in water
1000 × streptomycin	50 mg/ml streptomycin in water

For yeast two-hybrid studies, genes of interest were cloned into empty plasmids that were made Gateway®-compatible by Johannes Stuttmann (MPIPZ, Köln) (Table 2.10). Successfully cloned plasmids were transformed into yeast strain AH109 (Table 2.11).

Table 2.10 | Empty plasmids for yeast transformation.

Name	Supplier	Used for	Selection type	Epitope
pGADT7-Rec	Clontech	Expression as a GAL4 AD fusion	*LEU2*	HA
pGBKT7	Clontech	Expression as a GAL4 BD fusion	*TRP1*	c-Myc

Table 2.11 | Untransformed yeast strains.

Name	Supplier	Genotype
AH109	Clontech	MATa, trp1-901, leu2-3, 112, ura3-52, his3-200, gal4Δ, gal80Δ, LYS2 : : GAL1$_{UAS}$-GAL1$_{TATA}$-HIS3, GAL2$_{UAS}$-GAL2$_{TATA}$-ADE2, URA3 : : MEL1$_{UAS}$-MEL1$_{TATA}$-lacZ, MEL1

Yeast media (Table 2.12) were prepared using bidistilled water and were autoclaved before use. Non-autoclavable components were added after being sterile-filtered.

Table 2.12 | Media for yeast cultivation.

Name	Recipe for 1 litre
Synthetic defined (SD)	6.7 g nitrogen base w/o amino acids (Clontech), 20 g dexrose (add sterile after autoclaving), pH 5.8
SD selection media	6.7 g nitrogen base w/o amino acids (Clontech), 20 g dexrose (add sterile after autoclaving), a combination of dropout supplements (Clontech) and amino acid stock solutions, pH 5.8
SD plates	6.7 g nitrogen base w/o amino acids (Clontech), 20 g dexrose (add sterile after autoclaving), 15 g agar, pH 5.8
YPD	20 g peptone, 10 g yeast extract, 20 g dextrose (add sterile after autoclaving)

2.1.7. Software

2.2. Molecular Biological Methods

2.2.1. Amplification of DNA

DNA molecules were amplified *in vitro* by polymerase chain reaction (PCR) adapted from Mullis *et al.* [1986]. A PCR involves annealing of two oligonucleotide primers that are complementary to the sense and anti-sense strands of a DNA template. Repeated cycles of DNA denaturing at 94 °C, primer annealing and primer elongation through a thermostable DNA polymerase allow synthesis of major amounts of a desired DNA fragment with sizes of up to several kilobase pairs.

A typical reaction mix and thermal settings are shown below:

2.2.2. Purification of PCR Products

To remove contaminating proteins, components of the PCR reaction buffer, excessive dNTPs etc., the PCR product was optionally cleaned using the QIAquick PCR Purification Kit (Qiagen)

Table 2.13 | Used software.

Used for	Software (package)	Developer/Reference
Preparation of figures	Adobe Illustrator	Adobe Systems
Preparation of figures	Adobe Photoshop	Adobe Systems
Preparation of figures	Inkscape	`http://inkscape.org/`
Preparation of protein structure figures	PyMOL	[Schrödinger, LLC, 2010]
Preparation of graphs	OriginPro	Origin Lab
Preparation of text	LaTeX Editor	Adam Skórczyński and Sebastian Deorowicz
Reference manager	Mendeley Desktop	`http://www.mendeley.com/`
DNA related software		
Sequence data analysis	LASERGENE® SeqMan Pro™	DNASTAR
Cloning	LASERGENE® SeqBuilder™	DNASTAR
Protein related software		
CD spectra analysis	Spectra Manager™	JASCO
Chromatography systems	FPLC/UNICORN Manager	Pharmacia/GE Healthcare
Dynamic light scattering analysis	DYNAMICS®	Wyatt Technology
Gel documentation	Image Lab™	Bio-Rad
Sequence alignments	T-Coffee	[Notredame *et al.*, 2000]
Sequence alignments and phylogenetic tree-building	SeaView	[Gouy *et al.*, 2010]
Thermofluor analysis	CFX Manager™	Bio-Rad
Structural biology software		
Data processing	XDS	[Kabsch, 2010]
Model building	Coot	[Emsley *et al.*, 2010]
Molecular dynamics	GROMACS	[Van der Spoel *et al.*, 2005]
Phasing	CCP4 suite	[CCP4, 1994]
Refinement	Phenix suite	[Adams *et al.*, 2011]
Topology diagrams	TopDraw	[Bond, 2003]

Table 2.14 | A 50 µl PCR mix.

Component	Volume
10 × PCR reaction buffer	5 µl
dNTP mix	5 µl
Forward primer (10 µM)	1 µl
Reverse primer (10 µM)	1 µl
Taq DNA polymerase (4U/µl)	0.5 µl
Template DNA	1 - 50 ng
Water	37.5 µl

Table 2.15 | Thermal setting for a standard PCR.

Stage	Temperature (°C)	Time	No. of cycles
Initial denaturation	95	3 min	1
Denaturation	95	30 sec	
Annealing	50-60	30 sec	30-40
Extension	72	1 min per kb	
Final extension	72	3 min	1

according to the manufacturer's instructions. Purified DNA was eluted in bidistilled water or TE buffer at pH 8.

2.2.3. Restriction Endonuclease Digestion of DNA

DNA was cut by endonucleases at specific recognition sites according to the manufacturer's instructions. All digests were carried out at the appropriate temperature for a minimum of 60 minutes.

2.2.4. Gel Electrophoresis of DNA Molecules

DNA molecules can be separated in agarose gels according to their size as they carry a negatively charged phosphate backbone and migrate towards the positively charged anode in an electric field.

To prepare a 1 % agarose gel that is suitable to resolve DNA fragments of a size between 500 and 10000 base pairs, 1 % (w/v) agarose was boiled in TAE buffer, pH 8. Molten agarose was cooled to approximately 50 °C before ethidium bromide was added at a final concentration of 0.5 µg/ml. The agarose was poured and allowed to solidify before being placed in an electrophoresis tank filled with TAE buffer, pH 8. Each DNA sample was mixed with 6 × loading dye before loading them onto the gel. After applying a voltage of approximately 5 V/cm distance between tank electrodes and separation of DNA fragments, these were visualised by placing the gel onto a UV transilluminator set to a wavelength of 312 nm.

2.2.5. Isolation of DNA Fragments from Agarose Gels

DNA fragments visualised on a UV transilluminator were excised from a separating agarose gel using a cleaned razor blade and commercially available extraction kits following the manufacturer's instructions.

2.2.6. Site-Specific Recombination of DNA in Gateway-Compatible Vectors

The Gateway® cloning system (Invitrogen/Life Technologies) allows fast and efficient transfer of DNA fragments between plasmids. This makes it extremely helpful if many clonings are done in parallel and in between different expression hosts of the desired gene product. It was developed in the 1990s based on the recombination system that the λ phage uses naturally to integrate DNA into the genome of *E. coli* cells [Hartley *et al.*, 2000].

In a first step, attB1 and attB2 (Gateway® nomenclature) recombination sites of 25 base pairs have to be added by PCR to the 5' and 3' ends of the gene of interest. The PCR products are then mixed with "Donor Vectors" and a "BP Clonase™" enzyme mix that catalyses the recombination and insertion of the PCR product into the attP recombination sites within the "Donor Vector" resulting in an "Entry Clone". The efficiency of recombination is dramatically increased through the use of a *ccdB* gene that is lethal for *E. coli* as its gene product is a toxin that targets the bacterial DNA gyrase. Successful recombination leads to an exchange of the *ccdB* gene with the gene of interest (Fig. 2.1). When the reaction mixture is transformed into *E. coli* (see section 2.2.12), cells that take up an unreacted vector carrying the *ccdB* gene or by-product molecules retaining the *ccdB* gene will fail to grow.

The "Entry Clone" gene cassette with flanking attL recombination sites can then be efficiently transferred into a set of commercially available "Destination Vectors" carrying attR recognition sites by using an "LR Clonase™" mix. These allow to fuse the protein of interest to different affinity tags including polyhistidine, thioredoxin, the maltose binding protein and many more.

The *ccdB* gene is again used to increase efficiency. The concept of Gateway® cloning is illustrated in Figure 2.1.
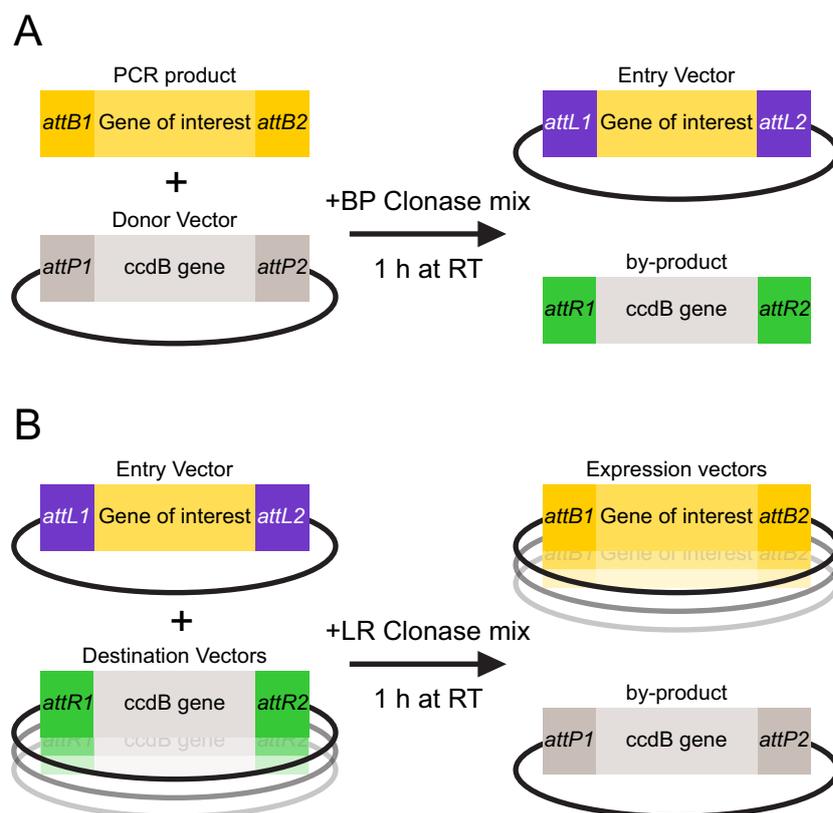
Figure 2.1 | The principle of Gateway® cloning. (A) An "Entry Vector" is created through site-specific recombination that is catalysed by a "BP Clonase™" enzyme mix as described in the text. (B) The "Entry Vector" is used in a second step to recombine the gene of interest into one of many commercially available "Destination Vectors", yielding Expression vectors in which the protein of interest is usually fused to different affinity-tags.

2.2.7. DNA Sequencing and Sequence Analysis

DNA fragments were sequenced either by the ADIS (Automatische DNA Isolierung und Sequenzierung) service at the MPIPZ (Köln) on Applied Biosystems Abi Prism® 377 and 3700 sequencers using the BigDye® Terminator technology or by the GATC Sequencing Service (GATC Biotech AG) on Applied Biosystems 3730xl sequencers. Sequences were analysed with the SeqMan™ program within the DNASTAR Lasergene® software suite.

2.2.8. Maintenance of Bacterial Cells

Laboratory strains of the gut bacterium *E. coli* have become a powerful tool for the heterologous production of proteins as they are easily cultured and genetically manipulated. Genetic information is usually transferred into and between *E. coli* cells through plasmids, small molecules of DNA that are physically independent of chromosomal DNA and replicated autonomously. For heterologous expression of a protein in bacterial cells, plasmids usually carry (1) an origin of replication, (2) a multiple cloning site for the insertion of any gene that encodes for the protein of interest, (3) elements required for protein expression including a promoter and terminator, a ribosomal binding site (the Shine-Dalgarno sequence in *E. coli*) and system-specific elements and (4) a selection marker. Genes encoding for proteins that provide resistance against otherwise lethal antibiotics are commonly used as markers that allow to specifically select cells containing the desired plasmids and to avoid contamination of bacterial cultures.

Unless otherwise stated, all liquid and solid media used for growth and maintenance of *E. coli* were supplemented with antibiotics whenever cells carried the corresponding resistance genes at final concentrations of:

- Ampicillin: 100 µg/ml
- Chloramphenicol: 50 µg/ml
- Kanamycin: 50 µg/ml
- Streptomycin: 50 µg/ml

For long term storage, glycerol stocks of *E. coli* cells were prepared by inoculating 5 ml LB medium with cells from a single colony grown on a selective LB-agar plate. This culture was incubated at 30 °C over night while shaking at 220 rpm. On the next morning, 1 ml of this culture was mixed with 600 µl sterile 80 % (v/v) glycerol and frozen at -80 °C.

For mid term storage, 5 ml LB medium were inoculated with either a single colony grown on selective LB-agar or with cells from a glycerol stock. After incubation at 30 °C and 220 rpm over night, a glass rod was dipped into this culture and cells were streaked out onto selective LB-agar plates that were incubated over night at 37 °C and finally stored for several weeks at 4 °C.

2.2.9. Plasmid Preparation

For preparation of DNA plasmids from *E. coli* cells, 5 ml LB medium were inoculated with a single colony grown on a selective LB-agar plate. This culture was incubated at 37 °C on a shaker at

220 rpm and cells were harvested on the next day by 30 sec of centrifugation at $11000 \times g$ and room temperature.

Cells were then treated with different commercially available plasmid extraction kits that generally involve an alkaline lysis of cells, the precipitation of cellular debris, proteins and genomic DNA followed by adsorption of plasmid DNA to a silica membrane. Membrane-bound plasmid DNA was then washed and finally eluted in either water or TE buffer, pH 8.

2.2.10. Determination of DNA Concentration

The concentration of DNA in solution can be measured photometrically as the aromatic nucleobases absorb light with a maximum absorbance near 260 nm wavelength. Using the Lambert-Beer law that relates the absorption of light to the properties of the material through which the light is travelling, the concentration of DNA can be calculated using the following formula:

$$c = \frac{A_{260}}{\varepsilon_{260} \times d}$$

where c is the concentration of DNA (in M), ε_{260} is the absorption coefficient (in M^{-1} cm^{-1}) and d is the cell pathlength (in cm). Using the specific absorption coefficient for double-stranded DNA ($0.02 \, (\mu g/ml)^{-1}$ cm^{-1}), the concentration can be calculated using this simplified formula:

$$\rho = A_{260} \times 50 \, \mu g/ml$$

While other specific absorption coefficients can be used for single-stranded DNA ($0.027 \, (\mu g/ml)^{-1}$ cm^{-1}) and single-stranded RNA ($0.025 \, (\mu g/ml)^{-1}$ cm^{-1}), the coefficient for short oligonucleotides should be calculated using tools like OligoCalc[4] as it depends critically on the nucleobase composition. Proteins, in contrast, absorb light with a maximum absorbance near 280 nm and the A_{260}/A_{280}-ratio can be used to analyse protein contamination. For double-stranded DNA, a ratio lower than 1.8 indicates the presence of proteins and a ratio higher than 2.0 indicates a probable contamination (e.g. by phenols).

2.2.11. Preparation of Chemically Competent Cells

E. coli cells can be treated such that they efficiently pick up plasmids from their environment. A widespread method involves twovalent cations (calcium and magnesium) that alter the permeability of bacterial membranes and allow DNA plasmids to enter the cells, especially when they are shocked by heating.

Following a protocol established by Dagert and Ehrlich [1979], 100 ml LB medium were inoculated with a single colony of an appropriate *E. coli* strain grown on a LB-agar plate. The culture was incubated on a shaker at 37 °C until the cells reach an OD of 0.3 to 0.4 and then harvested by centrifugation in pre-cooled tubes at $2500 \times g$ and 4 °C for 10 min. Cell pellets were resuspended on ice in 25 ml (1/4 culture volume) of ice-cold $MgCl_2$, taking 3 to 5 minutes for this procedure. After centrifugation at 4 °C and $2000 \times g$, pelleted cells were resuspended on ice

---

[4]http://www.basic.northwestern.edu/biotools/oligocalc.html

in 5 ml (1/20 culture volume) $CaCl_2$ before an additional 45 ml (9/20 culture volume) ice-cold $CaCl_2$ were added. This suspension was kept on ice for 20 minutes and subsequently pelleted by 10 min of centrifugation at $2000 \times g$ and 4 °C in pre-cooled tubes. Finally, cells were resuspended in 2 ml (1/50 culture volume) of ice-cold 85 mM $CaCl_2$ in 15 % (v/v) glycerol, dispensed in 100 µl aliquots and frozen at -80 °C.

2.2.12. Transformation of Chemically Competent Cells

For a single transformation, 50 µl chemically competent *E. coli* cells were briefly thawed on ice before adding 1 to 10 ng of plasmid DNA. The cells were kept on ice for 20 minutes before heatshocking them in a 42 °C water bath for 75 seconds. After putting the cells back on ice for another 2 minutes, 600 µl LB medium were added and this small culture was incubated at 37 °C for 45-90 minutes on a shaker. Different volumes (25 to 100 µl) of this culture were then plated on LB-agar and incubated over night at 37 °C.

2.2.13. Colony PCR

To check for positive transformants, single colonies were directly picked with pipette tips into a PCR reaction mix:

- 3 µl $10 \times$ PCR reaction buffer
- 3 µl dNTP mix
- 0.5 µl *Taq* polymerase
- 0.6 µl forward primer
- 0.6 µl reverse primer
- 23 µl H_2O

2.3. Protein Biochemical Methods

2.3.1. Small-scale Expression of Recombinant Proteins

To test the level of expression dependent on different parameters (*E. coli* strain, composition of the medium, inducer concentration, expression temperature and length of expression time), single colonies for each *E. coli* strain were picked into 5 ml LB medium and grown over night while shaking at 37 °C. On the next day, 10 ml expression medium were inoculated in a sterile tube with 200 µl of the starter culture and incubated at 37 °C until the cells reached an OD_{600} of 0.7 to 0.9. Expression of the target protein was then induced by the addition of IPTG in varying concentrations and the cells were incubated at different temperatures between 18 °C and 37 °C for several hours or over night. IPTG (Isopropyl β-D-1-thiogalactopyranoside) is a non-metabolised mimic of the disaccharide allolactose that triggers transcription of the bacterial *lac*-operon. All expression plasmids used in this thesis do contain components of the *lac*-operon and thus allow to turn on expression by the addition of IPTG. Following expression, cells were yielded by centrifugation at $2500 \times g$ for 10 minutes at 4 °C and the cell pellets were either stored

at -80 °C or used directly. To check for protein expression, each pellet was resuspended in 1 ml lysis buffer supplemented with 1 mg/ml lysozyme, 10 µg/ml DNaseI and 1 mM DTT and put on ice for 30 minutes. Final cell lysis was achieved through sonication on ice applying 2 min pulse separated by pauses (40 % on a 500 Watt VCX 500 sonicator; Sonics®). To pellet cell debris, each cell suspension was centrifuged for 30 min at 4 °C and $20.000 \times g$. The supernatant (soluble fraction) was kept in fresh tubes while the pellet (insoluble fraction) was resuspended in 1 ml lysis buffer. Samples of each fraction were prepared for analysis on SDS-PAGE.

2.3.2. Small-scale Purification of His-tagged Proteins

Following the protocol for small-scale expression of recombinant proteins (see previous section), 50 µl preequilibrated Ni-NTA matrix (Qiagen) were added to the soluble fraction of proteins and incubated at 4 °C for 1 h while gently shaking to allow for binding of His-tagged proteins to the nickel ions. The matrix was then pelleted by centrifugation at $3000 \times g$ for 30 sec at 4 °C. The supernatant ("flow-through") containing unbound proteins was collected carefully and the Ni-NTA resin was washed with 500 µl wash buffer and pelleted by centrifugation as described above. This was repeated three times collecting the supernatants for later analysis by SDS-PAGE. To elute His-tagged proteins, the Ni-NTA resin was either resuspended multiple times in 50 µl of elution buffer (incubating 5 min at 4 °C after adding the buffer, then spinning down the resin by centrifugation and collecting the supernatants) or boiled after adding 100 µl $2 \times$ SDS sample buffer. Samples of each fraction were prepared for SDS-PAGE analysis.

2.3.3. Large-scale Expression of Recombinant Proteins

Cells from LB-agar plates were used for the inoculation of a starter culture that was incubated in a shaker at 37 °C over night using either TB or LB medium. For the expression culture, 500 ml medium were inoculated with 10 ml starter culture and incubated at 37 °C while shaking until they reached an OD_{600} of ~ 0.2. The incubator was then set to cool to 18 °C and cells were induced by the addition of IPTG at a final concentration of 0.2 mM when they had reached an OD_{600} of 0.7 to 0.9. These cultures were incubated over night at 18 °C and 180 rpm as low expression temperatures have been shown to promote the production of soluble protein [Vera et al., 2007]. On the next day, cells were harvested by centrifugation at $3300 \times g$ and 4 °C for 20 minutes, washed in 0.9 % (w/v) NaCl and pelleted again by centrifuging them at $2500 \times g$ and 4 °C for 10 minutes. The supernatant was discarded and cells were frozen at -80 °C or used directly.

2.3.4. Large-scale Purification of His-tagged Proteins

Cell Lysis   Cells from the freezer were thawed on ice for at least 30 min and afterwards resuspended in an appropriate volume (approximately 5 ml/g pellet) of lysis buffer supplemented with 1 mg/ml lysozyme, 10 µg/ml DNaseI and 1 mM DTT and EDTA-free Complete™ Protease Inhibitor Cocktail tablets (Roche). This suspension was incubated on ice for 30 min and finally lysed by ultrasonication for a total of 2 minutes at 40 % output, separated by pauses on ice. Cell debris was spun down by 30 min of centrifugation at 4 °C and almost $120000 \times g$ in a XL70

ultracentrifuge (Beckman Coulter). The lysate was cleared through a folded filter (Schleicher & Schuell No. 589/5) before further use.

Immobilised Metal-Affinity Chromatography (IMAC)   Metal-affinity chromatography was first used by Porath *et al.* [1975] to purify recombinant proteins and uses that cysteine and histidine residues form chelating complexes with certain metals. In the most popular form, proteins are N- or C-terminally fused to a polyhistidine tag consisting of mostly 6 to 10 consecutive histidine residues. These form a complex with immobilised nickel- or cobalt ions through their imidazole sidechain. Free imidazole at high concentrations displaces His-tagged proteins by competing for metal coordination sites and can be used to elute the protein after it has been stringently washed.

Cleared lysate was directly applied to a manually packed Ni-NTA column of appropriate size (the matrix binds at least 10 mg protein per ml of resin) through a superloop connected to either a FPLC (Pharmacia) or ÄKTA (GE Healthcare) chromatography system. To facilitate binding of His-tagged protein to the Ni-NTA matrix, this was performed at a flow-rate of only 0.3 ml/min. After protein application, the column was washed with 10-20 column volumes of lysis buffer until the absorption at 280 nm reached baseline level. His-tagged protein was finally eluted through a 30 ml gradient between 20 mM (100 % lysis buffer) and 190 mM imidazole (25 % lysis buffer, 75 % IMAC elution buffer). Elution fractions containing major amounts of protein were pooled and concentrated for further experiments.

Anion Exchange Chromatography (AEC)   Ion exchange chromatography allows to separate proteins according to their charge. Binding to the matrix depends on the pH and ion strength of the buffer and the isoelectric point of the proteins to be separated. To allow for binding to a positively charged anion exchange matrix, the buffer pH should be at least one pH unit above the isoelectric point of the protein of interest. Separation of bound proteins is achieved by increasing the buffer ion strength gradually or stepwise.

Optionally, proteins purified by IMAC were applied to a Mono Q HR 10/10 column (Pharmacia) that had been preequilibrated with AEC buffer A. After washing the column with an additional 20 ml AEC buffer A, bound proteins were eluted by an 80 ml gradient to a salt concentration of 1 M using AEC buffer B. Elution fractions containing the protein of interest were pooled and concentrated for further use.

Size Exclusion Chromatography (Gelfiltration)   For preparative gelfiltration, a concentrated protein sample was injected into a loop of two times the sample volume. The sample was then applied onto a prequilibrated gelfiltration column and eluted isocratically. HiLoad 16/60 columns (Pharmacia/GE Healthcare) packed with two different Superdex matrices were used. While Superdex 75 material separates proteins of a molecular weight between 3 and 70 kDa, columns packed with a Superdex 200 matrix reliably separate proteins in a range between 10 and 600 kDa.

2.3.5. Protein Concentration by Ultrafiltration

All protein solutions were concentrated in Amicon® Ultra filtration units (Millipore) with an appropriate molecular weight cut-off (at least 20 % below the molecular weight of the protein to be concentrated) by centrifugation at $4000 \times g$.

2.3.6. Purification of Strep-Tagged Proteins from Plants

To prepare transiently or stably expressed proteins carrying a Strep-tag from plant material, leaves of *Arabidopsis* or *Nicotiana* plants (2 g fresh weight) were grinded in liquid nitrogen using mortar an pestle. Homogenised material was then resuspended in 2-3 ml Ex-Strep buffer and briefly vortexed. Cell debris was pelleted through centrifugation at 4 °C and $20.000 \times g$ for 30 minutes. The supernatant was filtered through a sterile filter with a pore size of 0.22 µm before adding 120 µl of 50 % Strep-Tactin High Capacity matrix (IBA Life Sciences) to 2 ml. After incubation at 4 °C for 40 min, this was poured onto a gravity-flow column (Bio-Rad) and the matrix was washed with at least 20 column volumes W-Strep buffer in several steps before the proteins were eluted by the application of either Elu-Strep buffer or by boiling the matrix with $2 \times$ SDS sample buffer.

2.4. Protein Analytical Methods

2.4.1. Denaturing SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE [Laemmli, 1970] is used to separate proteins according to their molecular weight. Protein samples have to be prepared in presence of the anionic detergent sodium dodecyl sulfate (SDS) and a reducing agent (e.g. DTT) and boiled for 5 min at 95 °C to break disulfide bridges and denature the protein. Anionic SDS then binds to the linearised protein in a 1.4:1 g ratio resulting in a negative charge that is correlated with the protein weight. Charged proteins are then applied onto gels containing varying amounts of polyacrylamide and separated in an electric field. The migration behaviour depends on the polyacrylamide concentration: gels with high concentrations allow small proteins to migrate faster compared to large proteins. This sieving effect is used to adjust the molecular weight range of proteins that are separated in a gel.

Recipies for gels used in this thesis are presented in Table 2.16.

Table 2.16 | Recipies for SDS-PAGE gel solutions.

kDa range	50-200	30-95	20-80	12-60	10-43	
	10 ml Resolving gel					5 ml Stacking gel
Water (ml)	5.8	5.3	4.8	4.3	3.55	3.7
40 % acrylamide mix (ml)	1.5	2	2.5	3	3.75	0.63
1.5 M Tris, pH 8.8 (ml)	2.5	2.5	2.5	2.5	2.5	-
1 M Tris, pH 6.8 (ml)	-	-	-	-	-	0.63
10 % SDS (µl)	100	100	100	100	100	50
10 % APS (µl)	100	100	100	100	100	50
TEMED (µl)	10	7.5	5	5	5	5

Self-cast gels were prepared by pouring resolving gel solutions given in table 2.16 between glass plates separated by a 0.75 mm spacer and were overlayed with 2-propanol. After polymerisation of the resolving gel, the 2-propanol layer was discarded and a stacking gel solution was poured between the glass plates on top of the resolving gel. After adding combs, the stacking gel was allowed to polymerise at room temperature. These gels were used in Mini-PROTEAN systems (Bio-Rad) with Tris-glycine running buffer, applying voltages between 100 V and 140 V until proteins were separated.

Alternatively, pre-cast 4-12 % gradient gels (Invitrogen/Life Technologies) were used in MES running buffer, applying voltages between 150 V and 200 V.

Gels were documented using the ChemiDoc™ imaging system (Bio-Rad).

2.4.2. Protein Staining on Gels

After separating SDS-PAGE, proteins were stained using the triphenylmethane dye Coomassie Brilliant Blue (CBB). Originally used as a textile dye, CBB binds unspecifically to positively charged amino groups of protein side-chains and has become one of the most widely used chemicals for protein detection since de St Groth *et al.* [1963] first used it.

Usually, a staining solution additionally contains an acidic component and either methanol or ethanol to fix the proteins in the gel. During this thesis, two variants of staining/destaining solutions were used. A staining solution based on Fairbanks *et al.* [1971] and composed of CBB, ethanol and acetic acid was used to detect protein bands down to 100 ng protein. The gel was stained for approximately 60 min before it was destained in an ethanol/acetic acid solution while heating up in a microwave. An alternative staining protocol described by Kang *et al.* [2003] allows detection of proteins down to 10 ng protein per band. After separation, the gel is washed twice with water for 5 to 10 minutes before a staining solution composed of CBB, aluminium sulfate, ortho-phosphoric acid and ethanol is poured over. After staining for at least one hour, the gel was destained over night in a mixture of ethanol and ortho-phosphoric acid.

2.4.3. Immunoblot Analysis (Western Blotting)

Alternatively to unspecific dye-staining, proteins can be visualised specifically after resolving SDS-PAGE through western blotting developed by Towbin *et al.* [1979]. For that, proteins were transferred from the gel to a Hybond™-ECL nitrocellulose membrane (GE Healthcare) using $1\times$ transfer buffer and a blotting apparatus (TE 22 Mini Tank Transfer Unit (GE Healthcare) or Mini Trans-Blot® Module (Bio-Rad)) according to the manufacturer's manual. Transfer was carried out in the cold room for 90 min and 90 V. Membranes were then washed on an orbital shaker for 5 min in TBS-T before blocked for 1 h at room temperature in TBS-T containing 5 % (w/v) non-fat dry milk (Roth) or 5 % (w/v) BSA. The blocking solution was discarded and the membranes were rinsed with TBS-T. Blocked membranes were then incubated on an orbital shaker in TBS-T supplemented with non-fat dry milk or BSA at 2.5 % (w/v) and the primary antibody at an appropriate concentration either 1 h at room temperature or at 4 °C over night. The primary antibody solution was afterwards removed, the membrane rinsed twice with TBS-T before washing it three times for 15 minutes in TBS-T while shaking at room temperature. After

washing, the membrane was incubated in TBS-T containing non-fat dry milk or BSA at 2 % (w/v) and the secondary antibody at an appropriate concentration for 1 h at room temperature. This solution was subsequently removed and the membrane was rinsed and washed as described above. Depending on the secondary antibody, proteins were detected either by chemiluminescence or chromogenically.

Antibody Detection by Chemiluminescence   For this detection method, secondary antibodies conjugated with horseradish peroxidase (HRP) were used. HRP accepts various commercially available substrates and, in the presence of peroxide, produces a transient product that decays to a lower energy state while emitting light. This light can be captured on an X-ray film or by CCD detectors. As HRP substrates we used the SuperSignal$^{®}$ West Pico or Femto kits (Pierce) according to the manufacturer's instructions and visualised produced light on photographic BioMax light film (Kodak) in a dark room.

Antibody Detection by Chromogenic Substrates   Instead of HRP, secondary antibodies carried an alkaline phosphatase (AP) conjugate for this detection method. The AP hydrolyses the substrate BCIP (5-bromo-4-chloro-3'-indolylphosphate p-toluidine salt) that subsequently tautomerises and finally produces an indigo dye through dimerisation of the tautomers in an alkaline environment. The dimerisation reaction is coupled to the reduction of a second component, nitro-blue tetrazolium chloride (NBT), to the purple NBT-formazan that enhances the chromogenic signal.

For chromogenic detection, membranes were preequilibrated for 10 minutes in AP reaction buffer while shaking at room temperature. 10 ml of AP reaction buffer were then supplemented with 66 µl of a NBT-stock (0.5 g NBT in 10 ml dimethylformamide) and 33 µl of a BCIP-stock (0.5 g BCIP in 10 ml dimethylformamide). The membrane was incubated in this solution at room temperature and gentle shaking until the desired staining intensity was reached. The reaction was stopped by washing the membrane rigorously in water.

Ponceau Staining   Membrane-bound proteins can additionally be stained unspecifically and reversible by negatively charged Ponceau S dye before blocking with non-fat dry milk. For that, membranes were briefly incubated in a Ponceau S working solution until bands appeared. The dye solution was then removed and membranes were washed in TBS-T until they were completely destained.

2.4.4. Protein Quantification

Bradford Assay   This protein quantification method developed by Bradford [1976] uses the already introduced dye Coomassie Brilliant Blue (CBB). In an acidic environment, CBB maximally absorbs light of a wavelength around 465 nm and appears brownish. Upon formation of protein/dye complexes, the deprotonated, anionic sulfonate form of CBB is stabilised and produces a blue colour. This colour change can be followed photometrically at a wavelength of 595 nm where the anionic CBB form has its absorption maximum. Although CBB binds more efficiently

to basic amino acids, an absorption increase at 595 nm is almost proportional to the protein concentration of a given solution.

To determine protein amounts, usually 2 μl sample were diluted in 798 μl water and mixed with 200 μl Bradford solution (Roth). The absorbance at 595 nm was measured photometrically after 10 minutes of incubation in the dark and at room temperature. Dilutions series of BSA were used as standards to quantify the amount of protein.

Spectroscopic Measurement at 280 nm The protein concentration in a homogeneous and pure protein solution can be directly measured photometrically as aromatic amino acids (especially tryptophan and tyrosine), and less important disulfide bonds, absorb light of 280 nm wavelength. The Lambert-Beer law introduced in section 2.2.10 allows to calculate the concentration of a pure protein solution using following formula:

$$\rho = \frac{A_{280} \times MW}{\varepsilon_{280} \times d}$$

where ρ is the protein's concentration (in mg/ml), A_{280} is the absorption of light at a wavelength of 280 nm, MW is the molecular weight of the protein (in Da), ε_{280} is the molar absorption coefficient (in M^{-1} cm^{-1}) at the same wavelength and d is the cell pathlength (in cm).

The molar absorption coefficient ε_{280} can be estimated as described by Pace *et al.* [1995] if the amino acid composition of the protein is known. According to Grimsley and Pace [2004], the protein concentration in an unpurified mixture can be estimated by assuming that a protein concentration of 1 mg/ml gives an A_{280} value of 1.3 as this is the average value for 116 proteins of various molecular weights and amino acid compositions [Pace *et al.*, 1995].

2.4.5. Circular Dichroism (CD)

Circular dichroism is a spectroscopic technique that allows to study folding and unfolding of proteins in solution. In a CD experiment, circularly polarised light is passed through a protein solution. As reviewed by Kelly *et al.* [2005], any chiral (optically active) chromophore produces a CD signal as they differently absorb left- and right handed components of circularly polarised light. The difference between these components is expressed in terms of ellipticity (θ). In proteins, peptide bonds, disulfide bridges and aromatic side chains yield CD signals if circularly polarised light of a given wavelength is used. Yet most importantly, the arrangement of peptide bonds in secondary structure elements produces CD spectra in the far UV range (180-240 nm) that are characteristic for α-helices, β-sheets and unstructured regions. As extensively reviewed by [Greenfield, 2006b,a], this information can be used to (1) estimate the secondary structure composition of a protein, (2) determine whether a purified protein or a mutant variant is folded, (3) study how binding of a ligand affects protein stability and even to (4) estimate binding constants of protein-protein and protein-ligand interactions.

In this thesis, CD spectroscopy was used to study the folding state of recombinantly expressed protein variants. Purified protein was gelfiltrated in 10 mM NaH$_2$PO$_4$, 50 mM Na$_2$SO$_4$, pH 7.5 and set to a concentration of 0.2 mg/ml before injecting it into a 0.01 cm cell. CD spectra at 10 °C were collected in the range of 190-240 nm using a scanning speed of 100 nm per minute and

10 accumulations per scan. To follow protein unfolding and determine its melting temperature (T_m), a freshly prepared protein solution was heated from 10 to 80 °C at a constant rate of 1 °C per minute while following the CD signal at a wavelength of 222 nm. The ellipticity at this particular wavelength is dominated by a negative CD signal characteristic for α-helices. Thermal unfolding of α-helices goes along with this signal becoming more positive. All CD experiments were performed on a J-715 spectropolarimeter (JASCO) and analysed using the Spectra Manager™ software (JASCO).

2.4.6. Differential Scanning Fluorimetry (Thermofluor)

The Thermofluor method was developed by Pantoliano *et al.* [2001] and allows to screen substances that increase the thermostability of a protein. Generally, the protein is heated in the presence of fluorescent dye that preferentially binds to hydrophobic patches on the protein surface. Minimising the number of hydrophobic amino acid side-chains that are exposed to an aqueous environment is an essential driving force of proper protein folding and therefore most hydrophobic patches are buried within a folded protein's core. However, if the protein is thermally unfolded, these patches become accessible for the fluorescent dye that emits light of a certain wavelength upon binding to hydrophobic regions. The amount of light emitted by the dye is thus directly correlated with the folding state of the protein.

Following protein unfolding in the presence of (putative) ligands, different buffer systems and further additives allows to characterise an environment in which the protein is most thermostable. As described by Ericsson *et al.* [2006], Vedadi *et al.* [2006] and others, thermal stability as derived from Thermofluor experiments has been shown to be correlated with the tendency of a protein to form crystals. This correlation is one of the reasons why the thermophile bacterium *Thermotoga maritima* has been chosen as a model organism in structural genomics [Lesley *et al.*, 2002].

I designed a screen containing 96 different buffers, salts and additives (see Appendix Table A.1). Each protein was purified and used in a standard buffer (mostly 10 mM NaH_2PO_4, pH 7.5) at a concentration of 1 mg/ml. In white 96 well PCR plates (Low Multiplate 96, Bio-Rad), 5 µl protein solution were mixed with 5 µl 10 × SYPRO® Orange dye solution (Sigma-Aldrich) and 40 µl of 96 different components from the screen. This plate was carefully sealed with Microseal 'B' Adhesive Seals (Bio-Rad), centrifuged for 1 minute at 500 × *g* to mix the components and placed into a CFX96™ real-time PCR machine (Bio-Rad). This machine heats up all samples from 4 to 95 °C at a constant rate of 12 °C per minute while reading the emission of fluorescence at 516 nm every 5 seconds (excitation: 492 nm). This yields melting curves for most condition (some additives may quench the fluorescence signal or denature the protein even at 4 °C) of which the first derivative allows to determine the melting temperature T_m. The curves were analysed using the CFX Manager™ software (Bio-Rad).

2.4.7. Dynamic Light Scattering (DLS)

Dynamic light scattering allows to determine the size distribution of small particles (like proteins) in solution. Following complex physics, proteins scatter laser light and the intensity of scattered light fluctuates as the protein molecules move through Brownian motion. As large proteins or

protein assemblies move slower compared to small proteins, the degree of fluctuation in light intensity can be directly used to estimate the hydrodynamic radii of proteins and their assemblies in solution. Prior to protein crystallisation, DLS is widely used to determine whether a protein solution is monodisperse (consists of objects that have the same size, shape or mass) or polydisperse (inconsistent distribution of size, shape and mass) as expressed in a mean polydispersity ratio, % Pd. Zulauf and D'Arcy [1992], Ferré-D'Amaré and Burley [1994], and others could show that a monodisperse protein solution tends to yield crystals with a higher probability compared to polydisperse samples. A large-scale analysis of several hundred crystallisation trials carried out by the Northeast Structural Genomics Consortium [Price *et al.*, 2009] shows that monodisperse samples (normalised polydispersity (% Pd) below 10) yield crystals more frequently than "predominantely monodisperse" or polydisperse samples (% Pd between 10 and 30 or % Pd > 30, respectively). Polydispersity arises from multiple stoichiometric states and/or protein aggregates. In case of polydisperse protein samples, substances that increase monodispersity should by identified by DLS before setting up crystallisation trials.

For DLS experiments, the protein concentration in a purified sample was adjusted to 0.1 - 0.5 mg/ml (smaller proteins need to be at a higher concentration). This solution was centrifuged for 20 min at $20.000 \times g$ at 4 °C to remove larger particles. 100 µl of the supernatant were pipetted into disposable UVettes$^{\circledR}$ (Eppendorf) and the polydispersity was measured in a DynaPro NanoStar$^{™}$ instrument (Wyatt Technology) using laser light at a wavelength of 830 nm. The normalised polydispersity (% Pd) and molecular weight of a sample (assuming globular proteins) were averaged across five consecutive measurements of each 10 acquisitions. All experiments were performed using a DynaPro NanoStar$^{™}$ instrument (Wyatt Technology).

2.4.8. Lipase Assay

Recombinantly expressed EDS1 or its lipase-like domain EDS1$^{1\text{-}384}$ were purified by immobilised nickel-affinity chromatography and gelfiltration according to section 2.3.4. Purified EDS1 and EDS1$^{1\text{-}384}$ were incubated in 50 mM Tris-HCl buffer, pH 8.5 at a concentration of 50 µM and in presence of 0.5 % (w/v) Triton X-100 and 1 mM of different p-nitrophenyl (pNP) esters for 2 minutes at 30 °C. Triton X-100 was used at a concentration where it forms micelles harbouring the p-nitrophenyl substrates that promote "interfacial activation" as explained in more detail in section 3.4.2. The reaction was followed by detecting the photometrically active product p-nitrohenolate at a wavelength of 412 nm (see Fig. 2.2). Bovine serum albumine (BSA) and a lipase from *Mucor miehei* (MML) at concentrations of 50 µM were used as negative and positive controls, respectively.
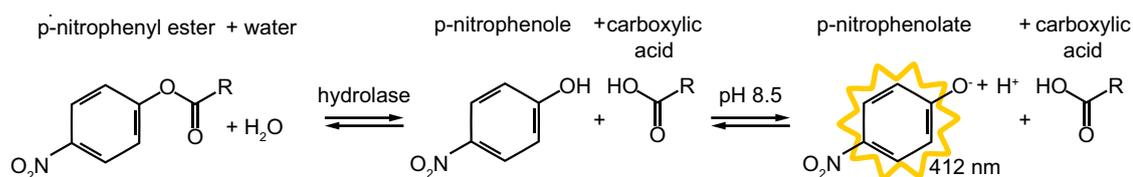
Figure 2.2 | P-nitrophenyl ester based hydrolase assay. Hydrolases catalyse the reaction of p-nitrophenyl esters with varying chain length (R) to p-nitrophenol through hydrolysis of the ester bond. At a basic pH, p-nitrophenol deprotonates to the p-nitrophenolate that produces a yellow colour with an absorbance maximum at approximately 412 nm.

2.5. Structural Methods

2.5.1. Protein Crystallisation

88 % of all protein structures deposited in the Protein Data Bank (PDB [Berman, 2000]) have been solved by X-ray diffraction [5]. Prerequisite, and in many cases the high-hanging fruit, is that the protein needs to form crystals in which the proteins molecules are arranged in ordered pattern in all three dimensions. As pointed out by Doye *et al.* [2004], most proteins in a cellular environment "do not want to crystallise, because a protein that is prone to crystallisation, or in fact any form of aggregation, is potentially deleterious to the cell.". Although structure elucidation from small crystals grown *in vivo* may "open new routes in structural biology" in combination with the recently established technique of serial femtosecond X-ray crystallography (SFX) [Koopmann *et al.*, 2012] this remains an exception and most crystals used for X-ray diffraction data collection are still being produced *in vitro*.

Protein crystallisation takes place in two steps – the formation of a crystal nucleus followed by crystal growth. Both steps require the controlled oversaturation of a pure and homogeneous protein solution. Summarised by McPherson [2004] many methods to achieve oversaturation are available: (1) altering the protein itself (e.g. by pH changes that affect the ionisation state of surface residues), (2) altering the chemical activity of the water (e.g. by addition of salts), (3) altering the degree of attraction of protein molecules for another (e.g. by pH changes or the addition of bridging ions), (4) altering the nature of interactions between the protein and the solvent (e.g. by polymers and ions), (5) altering the temperature or (6) direct removal of water (evaporation). As it is virtually impossible to predict a suitable combination of saturation effects that promotes the formation of crystal nuclei, the sparse matrix sampling method developed by Jancarik and Kim [1991] has become the most widespread method to search for crystallisation conditions. These screens are composed of crystallant solutions (including salts, volatile and non-volatile organic solvents and long-chained polymers) that are empirically derived based on known or published crystallisation conditions of various proteins. Although many different techniques exist in which protein and crystallant are mixed (and gradually oversaturated), the vapour diffusion method has been used in this thesis as it requires only small amounts of protein sample and allows robotic pipetting of crystallisation trials. As depicted in Figure 2.3, the protein solution is mixed with a crystallant solution (often in 2:1, 1:1 or 1:2) ratios and

---

[5]http://www.rcsb.org/pdb/static.do?p=general_information/pdb_statistics/index.html (request date: January, 2013)

placed on top of a reservoir containing only the crystallant. As the crystallant concentration within the protein drop is lower than that in the reservoir, water will diffuse in the vapour phase leading to an increasing concentration of both protein and crystallant in the drop. The processes yielding macroscopic protein crystals can be visualised by the phase diagram shown in Figure 2.4. Chemical (i.e. the composition of the protein drop) and physical parameters (e.g. crystallisation temperature, rate of equilibrium or methodology to approach the equilibrium) can be combined with biochemical factors (e.g. posttranslational, chemical or genetic modification of the protein or controlled proteolysis of the sample) if a protein fails to crystallise.
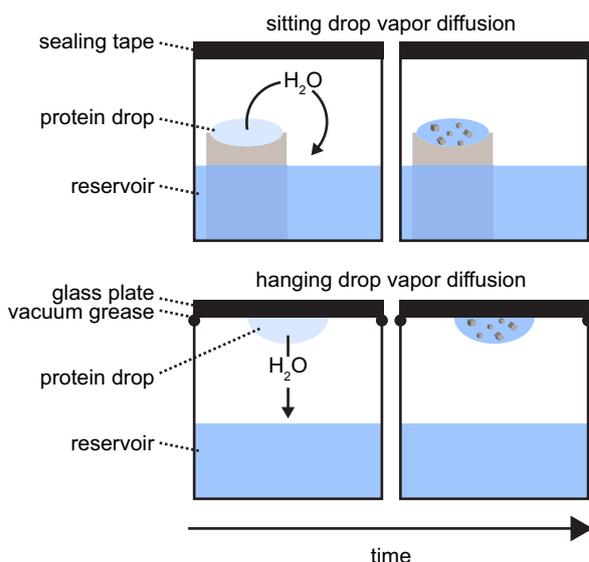
Figure 2.3 | Vapour diffusion technique. In both sitting- and hanging drop vapour diffusion, the protein solution is mixed in an arbitrarily ratio with the crystallant from the reservoir. Each condition is sealed and incubated at different temperatures to achieve diffusion of water through the gas phase out of the drop. This leads to the gradual saturation of the protein in solution and potentially to crystal growth.
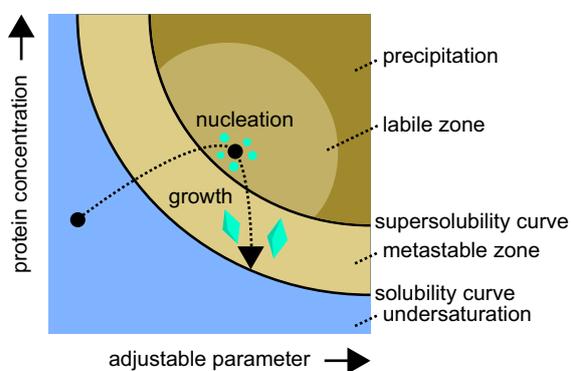
Figure 2.4 | The phase diagram of protein crystallisation. As described in the text protein crystallisation relies on supersaturation of the protein in solution. Through the combination of different parameters (see text) and a defined supersaturation of the protein, crystal nuclei appear in the labile zone of the phase diagram. As the crystals grow, the solution depletes in protein and the metastable zone is reached in which the crystals continue to grow while protein is still supersaturated. Figure adapted from [Chayen, 2004] .

For robotic sparse matrix sampling, protein solutions at varying concentrations (5-15 mg/ml) were mixed with crystallants from commercial screens that are listed in Table 2.4 using either a Matrix® Hydra® II eDrop (Thermo Scientific) or a mosquito® Crystal (TTP LabTech) robot and 96 well Intelli® (Art Robbins) or iQ (TTP LabTech) plates, respectively. This allowed to reproducibly pipet nanoliter volumes (Hydra® II: ≥ 500 nl per pipetting step; mosquito®: ≥ 100 nl per pipetting step) of protein and crystallant. Plates were kept at 4 °C or 20 °C for several months and inspected regularly under the microscope.

Initial hits were optimised by varying crystallant pH, crystallant concentration, growth temperature, drop sizes and through the addition of small molecule substances, primarily those of Additive and Detergent Screens 1-3 (Hampton Research).

Crystal Seeding As depicted in the phase diagram (Fig. 2.4), the conditions for the formation of protein crystal nuclei differ from those needed for crystal growth. To circumvent the necessity of spontaneous crystal nucleus formation, crystal seeds can be added to the protein drop in many ways. While even non-proteinous seeds potentially support nucleus formation and have been successfully used more or less systematically (synthetic silicates [Ino *et al.*, 2011] but also hairs from a 5-year-old horse [Thakur *et al.*, 2007] were deemed helpful), more common seeding techniques require that a protein was once crystallised. These crystals are then transferred into a new (equilibrated) protein drop either directly (macroseeding) or as submicroscopic seeds that have been obtained by crushing the initial crystal (microseeding) or by dislodging seeds from the initial crystal with an animal hair (e.g. cat whisker) that is then streaked through the new protein drop (for details on "streak seeding" see [Stura and Wilson, 1991]).

Microseeding and streak seeding were successfully used in this thesis. For microseeding experiments, a single crystal was transferred into a drop of reservoir or a stabilising solution (i.e. a solution in which the crystallant concentration is 1.05-1.2 $\times$ higher compared to the reservoir solution) using a nylon loop. This crystal was then crunched using a acupuncture needle and this seed stock was diluted with reservoir or stabilising solutions between 100- and 5000-fold (Fig. 2.5). New drops were set up as follows: 0.5 drop volume (DV) protein solution + 0.35-0.45 DV reservoir solution + 0.15-0.05 DV seed solution. Seeds were either added directly or 1 to 48 hours after protein and crystallant solutions had been mixed.

For streak seeding single crystals were briefly touched with human eyelashes, cat whiskers, hairs from a horse's tail or a young hedgehog and streaked three to four times through different (preequilibrated) crystallisation drops (Fig. 2.5).
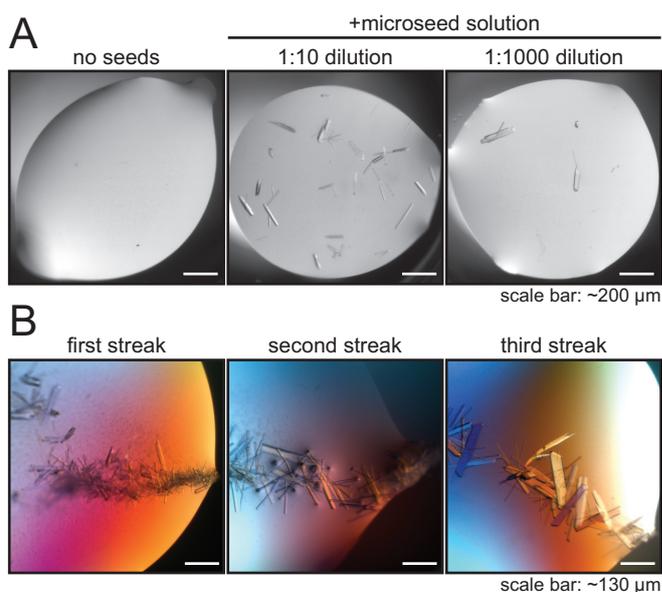

Figure 2.5 | Protein crystal seeding. (A) Microseed solutions were used in different dilutions to alter the outcome of crystallisation trials. (B) Streak seeds were analogously diluted by streaking the same seeding tool (e.g. a cat whisker) through multiple drops.

2.5.2. Cryoprotection

Macromolecular cryocrystallography as reviewed by Garman and Schneider [1997] became popular during the 1990s after recognising that protein crystals suffer significantly less from radiation damage if diffraction data are collected at low temperatures. As already described by Hope [1988], crystal need to be in a cryoprotecting solution before freezing them in liquid nitrogen or directly in a nitrogen stream to avoid "internal ice formation on cooling below the freezing point. While Hope transferred his protein crystals into a mixture of Paratone-N and mineral oil, cryoprotection can be achieved through the addition of alcohols (e.g. glycerol, methanol, 2-propanol, 2,3-butanediol), sugars (e.g. sucrose, xylitol, dextrose, trehalose), salts or glycols (e.g. polyethylenglycols of varying chain length or 2-methyl-2,4-pentanediol (MPD))[6]. Protein crystals may even directly grow in a cryoprotecting solution or, if one is less lucky, transfer into the cryoprotecting solution can be achieved through (1) direct transfer and soaking for arbitrary amount of time (depending on temperature, crystal size and nature of crystal solvent channels and the diffusion coefficient of the cryoprotectant) or (2) a stepwise increase of the cryoprotectant concentration to minimise the effect of an osmotic shock.

2.5.3. Heavy Atom Soaking

Heavy atom solutions were prepared by solving each compound listed in Table 2.17 directly in a solution suitable for cryoprotection of the protein at a stock concentration of approximately 30 mM. They were used after cryoprotection at a final concentration of 10 mM. Native crystals were soaked in the presence of heavy atoms for 1, 5 or 10 m.

2.5.4. Crystal Mounting

Following cryoprotection, crystals were picked up in CryoLoops™ with diameters between 0.05 and 0.4 mm (Hampton Research) installed in SPINE standard CryoCaps™ (Molecular Dimensions) and mounted on a goniometer head.

2.5.5. Diffraction Data Collection

Detailed descriptions of the techniques that underlie X-ray protein crystallography fill whole books [Blow, 2010; Rupp, 2009; Rhodes, 2006]. I used them to introduce some basic concepts in the following paragraphs.

Generation of X-rays X-ray structure solution from crystals uses that the atoms within a crystal diffract a beam of hard X-rays (usually with wavelengths below 2 Å) into specific directions.

The X-ray beam can be generated by an X-ray tube in which electrons are produced by heating up a cathode, commonly a tungsten filament. These electrons are then accelerated towards a metal (often copper or molybdenum) anode using an electric field of around 40 kV. These high energy electrons then displace electrons from energetically low-lying orbitals of the metal anode atoms upon collision. This, in turn, results in higher orbital electrons dropping down to

---

[6]Berejnov *et al.* [2006] lists these cryoprotectants and a critical concentration at which they should be used.

Table 2.17 | Heavy atom derivatives used for crystal soaking.

Platinum derivatives	MW (Da)	Useful? [Agniswamy et al., 2008]
1. Potassium tetrachloroplatinate(II)	415.11	++
2. Potassium tetranitroplatinate(II)	457.3	++
3. Potassium tetracyanoplatinate(II)	431.41	(+)
4. Potassium hexaiodoplatinate	1034.7	+
5. Diamminedichloroplatinum	300.06	
6. Dichloro(ethylenediamine)platinum	326.1	
7. Potassium tetrachloropalladate(II)	326.42	
8. Potassium hexachloroplatinate(IV)	486.01	
Gold derivatives		
9. Potassium tetrachloroaurate(III)	377.88	++
10. Gold(I)iodide	328.87	
11. Potassium dicyanoaurate(I)	288.1	++
Mercury derivatives		
12. Mersalyl salt	483.37	++
13. Mercury(II)acetate	318.7	++
14. Mercury(II)chloride	371.5	+
15. Ethylmercury(II)chloride	265.11	+
16. 2-(Ethylmercurimercapto)benzoic acid	404.81	
17. 4-Hydroxymercuribenzoic acid	360.7	
18. p-Aminophenylmercuric acetate	351.8	
19. Mercury dibromofluorescein	804.75	
20. Phenylmercuryacetate	336.74	
Lanthanoids		
21. Gadolinium(III)sulfate octahydrate	746.81	
22. Ytterbium(III)nitrate pentrahydrate	449.13	
23. Lanthanum(III)acetate hydrate	316.04	
24. Samarium(III)chloride hexahydrate	364.8	
Lead derivatives		
25. Trimethyllead acetate	310.35	
26. Triethyllead acetate	352.43	
Others		
27. Potassiumhexachloroosmate(IV)	481.12	+
28. Potassium osmate(VI) dihydrate	368.45	
29. Uranylnitrate hexahydrate	502.13	
30. Potassium hexachloroiridate(IV)	483.12	

fill the vacancy while loosing energy that they emit as X-ray photons with a wavelength that is characteristic for the anode metal type (characteristic X-rays). This mechanism is used by laboratory X-ray machines which are commonly referred to as home sources.

Alternatively, X-rays can be produced at a cyclic particle accelerator (synchrotron). Therein, electrons are constantly accelerated in a storage ring through bending magnets. As a by-product of magnetic bending, the electrons emit radiation that can be used to create a beam of monochromatic X-rays. Synchrotron radiation is characterised by a highly brilliant beam that is often tunable in wavelength through monochromatisation.

For diffraction data collection during this thesis, two home-source generators (that of a SuperNova™ (Agilent) or a MicroMax-007 HF generator (Rigaku)) were used. Diffraction data were additionally collected at the following synchrotron beamlines:

- X06SA (PXI), X06DA (PXIII); Swiss Light Source (Paul Scherrer Institute), Villingen
- BL14.1 and BL14.2; BESSY, Berlin

X-ray Diffraction by Protein Crystals   A protein crystal is made of molecules that are built up in a regular way so that equivalent molecules interact identically with their neighbours. As the molecules of protein crystals are irregularly shaped, they contain rather large solvent channels making them soft and fragile compared to inorganic crystals. The solvent content usually varies between 30 and 60 % [Matthews, 1968].

The molecules within the crystal reside on a crystal lattice and are related by symmetry operations. The smallest repeating unit that can generate the crystal with only translation operations is called the unit cell. Within each unit cell, several molecules can be related by other symmetry operations (rotations and reflections) and these form the asymmetric unit (see Fig. 2.6).

asymmetric unit
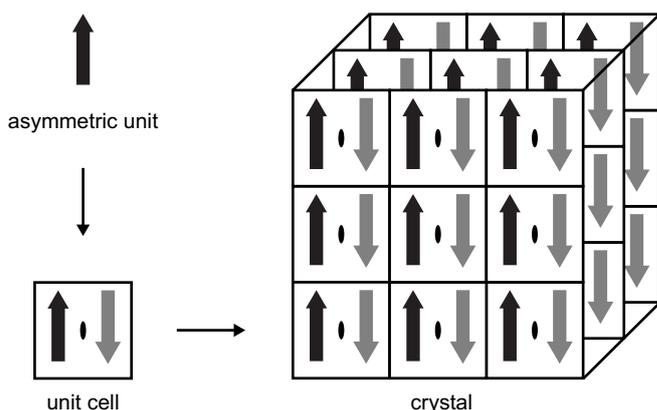
unit cell   crystal

Figure 2.6 | Building blocks of protein crystals. The protein crystal is formed by translation operations of the unit cell in all three dimensions. Within each unit cell, molecules of an asymmetric unit can be additionally described by rotations and reflections. Adapted from `http://www.rcsb.org/pdb/101/static101.do?p=education_discussion/Looking-at-Structures/bioassembly_tutorial.html` (request date: January, 2013)

The electrons of periodically arranged molecules within a crystal scatter the photons of an X-ray beam in random directions. The X-ray photons within the beam travel as waves that are characterised by a wavelength, an amplitude and a phase. While the scattered waves have the same wavelength as the incident beam (which depends on the photon source; see above), the amplitude and phase depend on the distribution of scattering objects within the crystal. After hitting the crystal, the randomly scattered waves will interfere with each other. While this

mostly leads to destructive interference in which the waves cancel each other out, there is also constructive interference occurring in which the waves' amplitudes reinforce each other (Fig. 2.7 A). The conditions required for constructive interference can be expressed through the Bragg's law [Bragg and Bragg, 1913].

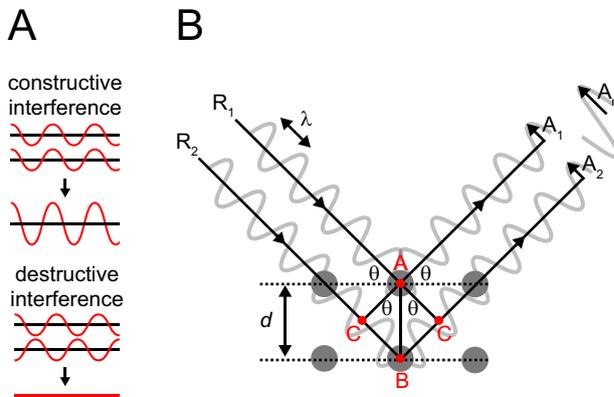**Figure 2.7 | Graphical representation of Bragg's law.** (A) Examples of constructive and destructive interference of two waves. (B) If the ray R_2 travels an additional distance that is an integral multiple of λ then constructive interference with ray R_1 occurs. See text for details.

If the grey dots in Figure 2.7 B represent crystal lattice points that lie on sets of planes separated by a distance d and the incident beam with a wavelength of λ impinges upon the planes at an angle θ and is reflected at the very same angle, then constructive interference occurs if the ray R_2[7] travels an additional distance of 2 BC as the scattered waves are then in phase: the single amplitudes A_1 and A_2 add to a reinforced amplitude A_r.

As $sin\,\theta = BC/AB$ and $BC = AB\,sin\,\theta = d\,sin\,\theta$, the Bragg's law is usually expressed as:

$$2d\,sin\,\theta = n\lambda$$

where n is an integer. So whenever the difference in path length ($2d\,sin\,\theta$) for waves reflected from successive planes is equal to an integral number of wavelengths λ, they emerge in phase with each other and interfere constructively.

The maximum possible resolution obtained in an X-ray diffraction experiment is then defined as:

$$d_{min} = \lambda/2\,sin\,\theta_{max}$$

Thus, reflections with large diffraction angles (or Bragg angles) θ contain most high-resolution information.

As the crystal, mounted on a goniometer head, is rotated within the beam, different layers of atoms satisfy Bragg's law and produce constructive interference. This results in diffraction peaks with an intensity related to the number and type of atoms in the diffracting layer.

From Diffraction to Structure The ultimate goal of an X-ray diffraction experiment is to obtain an electron density map of the asymmetric unit's contents. This map is then used to

---

[7]Pointed out by Rupp [2009], there are not two incoming X-rays involved (as schematically depicted in Figure 2.7) but it is the electromagnetic field vector of the same photon that excites all atoms within its coherence length.

automatically or manually build an atomic model of the contents. To calculate an electron density map in real space from a diffraction data set in reciprocal space, it requires a mathematical operation called Fourier transform. The recorded diffraction data set consists of numerous single reflections that represent the sum of the contributions of all scattering objects in the crystal's unit cell. The sum for each single reflection hkl is called the structure factor F_{hkl} where hkl describes the position of an individual reflection. The structure factor is directly used in a Fourier transform to calculate the electron density map and is expressed for each of the N atoms in the unit cell as:

$$F_{hkl} = \sum_{N} n_i f_i e^{2\pi(hx_i + ky_i + lz_i)} e^{-B_i sin^2\theta/\lambda^2}$$

where n_i is the occupancy factor for the ith atom, f_i is the atomic scattering factor, $2\pi(hx_i + ky_i + lz_i)$ is the phase factor and the last term is a description of thermal and crystal disorder known as the Debye-Waller factor (or B-factor).

The total occupancy factor n_i of atoms within the peptide chain is 1. However, amino acid side chains can adopt different conformations with partial occupancies below 1. The same can be true for any ligand or solvent molecule that is not homogeneously distributed among the crystal.

The atomic scattering factor f_i is the amplitude of a wave scattered by an atom relative to the amplitude of the wave scattered by an isolated electron and increases with the sum of scattering atoms inside the crystal unit cell and with their charge density. It further decreases with increasing Bragg angles at a fixed wavelength. The B-factor depends on the displacement of a crystal atom from its average position in the lattice[8] and high B-factors reduce the maximum possible resolution. The intensity of a diffracted beam is directly related to the amplitude of the structure factor, information on the structure factor phase, however, is lost and thus the phase factor $2\pi(hx_i + ky_i + lz_i)$ can not be directly estimated. This is known as the "phase problem".

2.5.6. De novo Phasing Based on Heavy Atoms

The Patterson function introduced by Patterson [1935] allows to determine the relative positions of atoms in a crystal unit cell without the knowledge of phases and is described as:

$$P(u, v, w) = 1/V \sum_{h} \sum_{k} \sum_{l} |F_{hkl}|^2 e^{-2\pi i(hu + kv + lw)}$$

where u,v,w are the coordinates of a point in the resulting Patterson map and $|F_{hkl}|$ is the structure factor amplitude that is directly measured in a diffraction experiment (as it is related to the intensity of a reflection: $|F_{hkl}| \propto \sqrt{I_{hkl}}$). As described above, the structure factor amplitude contains contributions of each scattering object in the unit cell. The determination of relative positions between scatterers thus becomes complicated and virtually impossible for too many scattering objects. While the Patterson function can be of direct use when solving the structures of small inorganic compounds, it is not directly applicable to proteins that regularly contain several thousand atoms within a single unit cell.

Nevertheless, the Patterson function is commonly used by protein crystallographers. The most popular phasing approach is that of a molecular replacement (this and other phasing techniques

---

[8]And is therefore often referred to as atomic displacement parameter (ADP)

are reviewed by Taylor [2010]). It requires the structure of a protein that is related to the protein in the crystal. This structure (the search model) is used to calculate its Patterson map. A Patterson map is also calculated from the structure factor amplitudes of the recorded reflections and these maps are then rotated and translated until the search model is correctly orientated relative to the new unit cell.

Yet, if no search model is known (usually a sequence identity of more than 25 % with an rmsd below 2 Å is needed), phases have to be determined experimentally. Again, the Patterson function is used, yet not to calculate the relative distances of protein atoms in the unit cell but rather the distances of heavy atoms that reside on fixed positions within the crystal unit cell. Given this limited number of scattering object, the Patterson function can be used for phasing. The introduction of <u>heavy</u> atoms is required as all scattering objects in the unit cell contribute to the intensity of a reflection and thus to the structure factor amplitude that we need to calculate a Patterson map. As described above, this amplitude is directly correlated with the atomic scattering factor f_i, that in turn depends on the charge density and therefore the number of electrons around atom i. Heavy atoms contain a large number of electrons (platinum for instance contains 13 times more electrons than carbon), and therefore contribute measurably to the structure factor amplitude.

Their precise contribution to the structure factor amplitude $|F_{hkl}|$ can be measured differently: in an isomorphous replacement experiment, data of two separate crystals are measured one of which is derivatised with a heavy atom. The structure factor amplitudes for the native and the derivatised crystals are then measured and the isomorphous difference between them can be used to estimate the heavy atom structure amplitude that is then used to generate a Patterson map. As the name suggests, this technique requires isomorphous crystals. Crick and Magdoff [1956] could show that already a 0.5 % change in the unit-cell dimensions leads to an average 15 % change of measured intensities rendering the determination of the heavy-atom structure factor amplitude impossible.

Alternatively, phases can be determined using a phenomenon called anomalous scattering that occurs because atoms scatter X-rays as a function of the incoming photons' wavelength. During normal scattering, an X-ray wave impinges on the atom and causes the electrons to oscillate with the same frequency as the incoming beam. The oscillating electrons, in turn, emit X-rays having the wavelength as the incident beam but a changed amplitude (as described above). If the energy of the X-ray beam, however, is close to the transition state that can bring the atom to an excited state (the absorption edge), anomalous scattering occurs in addition to normal scattering. This comes because a fraction of the X-ray beam is absorbed and re-emitted with a different phase, altering the structure factor by adding an anomalous and a dispersive component. These phase differences can be used to determine the position of anomalous scatterers within the asymmetric unit.

In a multi-wavelength anomalous dispersion (MAD) experiment, only one protein crystal is derivatised with a heavy atom. The absorption edges of commonly used heavy atoms are given in table 2.18. While these are close to wavelengths used in X-ray diffraction experiments (1 Å), carbon and other common elements of proteins have different edges and do not significantly contribute to anomalous scattering.

Table 2.18 | Absorption edges of different atoms (see `http://skuld.bmsc.washington.edu/scatter/AS_periodic.html` for more heavy atoms; request date: December, 2012).

Atom	K-edge	L I-edge	L II-edge	L III-edge
Carbon	43.7 Å			
Sulphur	5.02 Å			
Platinum	0.158 Å	0.893 Å	0.934 Å	1.072 Å
Gold	0.154 Å	0.864 Å	0.903 Å	1.040 Å
Mercury	0.149 Å	0.836 Å	0.873 Å	1.009 Å

Diffraction data of a derivatised crystal are collected at multiple wavelengths which leads to different contributions of the heavy atom to normal and anomalous scattering. Usually, wavelengths are chosen at the absorption peak (maximum anomalous scattering), at the point of inflection on the absorption curve (minimal dispersive scattering) and at a remote wavelength that is far away from the absorption edge.

2.5.7. Data Processing, Model Building and Refinement

Intensities measured in X-ray diffraction experiments were integrated and scaled using XDS [Kabsch, 2010] and Scala [Evans, 2006]. Scaled intensities were combined with phases that were experimentally determined as described above. The resulting structure factors were used for automated model building by the RESOLVE program [Terwilliger, 2003]. The final model served as a starting points for iterative cycles of manual model building in Coot [Emsley *et al.*, 2010] followed by structure refinement using the PHENIX software suite [Adams *et al.*, 2011]. The quality of the refined protein model is usually expressed by two R-factors, R_{free} and R_{work}. The R_{work} is defined as:

$$R_{work} = \frac{\sum ||F_{obs}| - |F_{calc}||}{\sum |F_{obs}|}$$

where $|F_{obs}|$ and $|F_{calc}|$ are the observed and calculated structure factor amplitudes. The resulting R_{work} value thus is a measure how well a simulated diffraction pattern that is calculated from the atomic model matches the diffraction pattern recorded in the actual experiment. While a perfect fit would produce an R_{work} of 0, a calculated diffraction pattern using randomly arranged atoms will produce an R_{work} of approximately 0.63. During the process of structure refinement, this value can be made artificially small through using too many refinement parameters (so-called overfitting). Brünger [1992] recognised this and proposed to use another R-factor, R_{free} in addition. Before the structure refinement starts, a certain percentage (usually about 1000 reflections are enough) are removed from the data that are used for refinement. The R_{free}-value is then calculated analogous to the R_{work} but expresses how well the diffraction pattern calculated from the refined model matches the diffraction data that were not used during refinement. Both

R-factors, R_{work} and R_{free} are dependent on the maximum resolution of the collected data and typically lie between 0.2 and 0.3.

2.5.8. Molecular Dynamics

Protein motions are essential for the function of most proteins. Molecular dynamics simulations allow to compute the movements of atoms and molecules in solution and have helped to understand small and large scale protein motions during enzyme catalysis, orchestrated motions within huge protein complexes and also mechanisms of protein folding [Karplus and Kuriyan, 2005]. Soon after Rahman and Stillinger [1974] had simulated the motions within liquid water, the first molecular dynamics simulation of a 58 amino acid protein was reported in 1977 by McCammon *et al.* [1977]. Although improved force fields that describe the energy of the protein as a function of its atomic coordinates and a rapid increase in computational power nowadays allow the simulation of larger proteins and even protein assemblies, large scale motions are comparatively slow [Zwier and Chong, 2010] and remain challenging to predict.

Preparation and simulation of proteins were performed according to a tutorial[9] with slight modifications. In brief, proteins were prepared from crystal structures (deletion of crystalline water molecules, alternative side-chain conformations) and handed over to the GROMACS (version 4.5.1) program [Van der Spoel *et al.*, 2005]. After putting the protein in a cubic water box (SPC/E model) and adding ions to neutralise the system's charge, an energy minimisation was performed in which the structure is relaxed. Subsequently, the system was equilibrated for each a 100 ps until a stable temperature and density were reached. The actual simulation was performed over 50 ns on the CHEOPS high-performance computing cluster (Regionales Rechenzentrum der Universität zu Köln)[10]

2.6. Yeast Methods

The yeast two-hybrid method allows to detect interactions between proteins. Established by Fields and Song [1989], nowadays several variants of this system coexist. The general principle, however, is always the same: for the easiest case of two proteins that shall be analysed for interactions, these protein are fused to each one domain of a transcription factor. The commonly used GAL4 transcription activator endogeneously exists in yeast cells and consists of a DNA binding domain (BD) and an activation domain (AD). While the domains are non-functional on their own, they are brought into close proximity if the fused proteins do physically interact with each other. This is sufficient to reconstitute the activity of the GAL4 activator and leads to transcription of a downstream reporter gene. We used the yeast ADE2 and HIS3 genes as reporters. These genes encode for proteins that are essential for the biosynthesis of the nucleobase adenine and the amino acid histidine, respectively. If those cells are plated on media that do lack adenine and/or histidine (both are essential for their survival), they can only grow if a functional GAL4 activator is reconstituted through the interaction of the proteins fused to the GAL4 activation and binding domain. This system is illustrated in Fig. 2.8.

---

[9]http://www.bevanlab.biochem.vt.edu/Pages/Personal/justin/gmx-tutorials/index.html(request date: April, 2010)

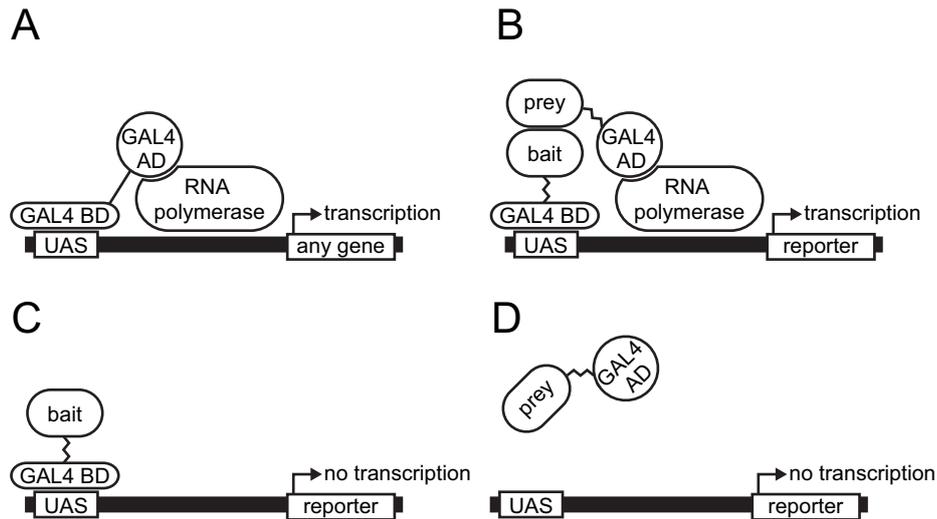[10]http://rrzk.uni-koeln.de/cheops.html (request date: December, 2012)

Figure 2.8 | **The yeast two-hybrid system.** (A) The GAL4 transcription activator is used naturally by yeast cells to control transcription of any gene. (B) The activation (AD) and binding domain (BD) of the GAL4 activator can be separately fused to two proteins (bait and prey) to test them for interaction. If the proteins physically interact, AD and BD domain are brought into close proximity and render the GAL4 activator functional. This can be used to selectively transcribe a reporter gene. Yet, if the bait and prey proteins do not interact, the activator remains non-functional and transcription of the reporter gene does not take place (C)/(D).

2.6.1. Transformation of Yeast Cells

Ito *et al.* [1983] found that yeast cells treated with alkali cations and polyethylene glycol of distinct chain length are able to pick up plasmid DNA from their environment. Based on this finding, Gietz and Schiestl [1991] proposed an optimised protocol for the transformation in which yeast cells are made competent by lithium acetate and polyethylene glycol and plasmids are transformed by heat-shocking in the presence of single-stranded carrier DNA.

For up to 40 parallel transformations, a 100 ml culture of yeast cells was grown over night at 30 °C in a shaker until the cells reached an OD_{600} of approximately 0.6. If cells grew to a higher optical density, the over-night culture was diluted to an OD_{600} of 0.2 in the morning and incubated shaking until they reached the desired density. Cells were then pelleted by 5 minutes of centrifugation at $1000 \times g$. The pellet was once resuspended in 20 to 30 ml sterile water and pelleted again by centrifugation as described. After resuspending the pellet in 1-2 ml of sterile 100 mM Lithium acetate in TE-buffer, the cells were incubated for 10 minutes at 30 °C. For a single transformation, 240 µl 50 % (w/v) PEG3350, 36 µl 1 M lithium acetate, 50 µl single-stranded DNA (previously boiled for 15 min at 94 °C and afterwards kept on ice) and 24 µl $10 \times$ TE-buffer were mixed with 50 µl of the competent cells and 0.4 µg of each plasmid to be transformed. After brief vortexing, this mixture was incubated for 30 minutes at 30 °C before heat-shocked for 20 minutes in a 42 °C waterbath. Transformed cells were subsequently spun down by centrifugation at $20.000 \times g$ for 30 seconds. The supernatant was discarded, the cells resuspended in 100 µl non-selective SD medium and spread out on (selective) SD-agar plates.

2.6.2. Selection of Co-Transformants

Gateway® cassettes were cloned into pGADT7 and pGBKT7 plasmids (see Clontech Matchmaker® Yeast Two-Hybrid System User Manual for details) to allow the creation of "Entry Clones" (see section 2.2.6). These plasmids carry the LEU2 and TRP1 gene, respectively, encoding for proteins that are essential for the biosynthesis of the essential amino acids leucine and tryptophan. Successful transformation of each plasmid into yeast cells can thus be checked by their ability to grow on SD-agar that lacks one or both of these amino acids.

2.6.3. Analysis of Protein-Protein Interactions

As described above, reporter genes encoding for proteins essential for bioynthesis of adenine and histidine (ADE2 and HIS3, respectively) were used to analyse the interaction between bait and prey proteins. Four single colonies of successfully co-transformed cells were streaked on SD-agar plates lacking leucine and tryptophan (-LW; control) and histidine (-LWH) and adenine (-LWHA). These plates were incubated at 30 °C and cell growth was documented after 2 and 5 days. As the GAL2-ADE2 reporter is more stringently regulated than GAL1-HIS3, cell growth on -LWH but not on -LWHA plates indicates residual interaction of the proteins.

2.6.4. Total Protein Extraction and Immunoblot Analysis

An apparent loss-of-interaction can be caused by the instability of one of the protein interaction partners in yeast cells. It is therefore critical to analyse protein accumulation by immunoblot detection of the protein in total protein extracts.

5 ml liquid YPD medium were inoculated with cells from SD (-LW) plates and incubated over night at 30 °C in a rotary shaker. On the next morning, the OD_{600} of each culture was measured and equivalent amounts of cells were harvested by centrifugation at $20.000 \times g$ for 10 minutes. Cell lysis was performed in three cycles of freezing in liquid nitrogen and thawing at room temperature. Each sample was mixed with 100 µl $2 \times$ SDS sample buffer, boiled for 15 minutes at 96 °C and the debris was spun down through centrifugation at $20.000 \times g$ (5 min). 10 µl of each sample were separated by SDS-PAGE and the proteins were detected by immunoblotting as described in section 2.4.3.

3. Results

"Every experiment proves something. If it doesn't prove what you wanted it to prove, it proves something else."

— Anonymous

Preamble

All protein sequences of EDS1, SAG101 and PAD4 were derived from the model plant *Arabidopsis thaliana* and represent the full-length variants EDS1 (i.e. amino acids 1-623)[11], SAG101 (1-537)[12] and PAD4 (1-541)[13] unless otherwise stated. According to the *Arabidopsis* Information Resource (TAIR) nomenclature, protein products of genes are written in capital letters without italics (e.g. EDS1), wildtype gene alleles in italic capital letters (e.g. *EDS1*) and mutant gene in lower case italic letters (*eds1*) whereas different alleles of the same gene are distinguished with a number following a hyphen (e.g. *eds1-2*).

As this thesis has been developed in a fruitful cooperation with the Max Planck Institute for Plant Breeding Research (MPIPZ, Köln), single experiments were not done by myself and are marked as such. Especially most *in planta* experiments were gratefully handed over to Johannes Stuttmann and Jaqueline Bautor. This is graphically illustrated in Figure 3.34 at the end of the results section.

3.1. Sequence Analysis: EDS1, SAG101 and PAD4 Exist in all Higher Plants and Share Lipase Homology

EDS1-related research has mainly focused the model plant *Arabidopsis*. The rapidly increasing number of published higher plant[14] genomes[15], however, allows to analyse if and by how far the EDS1 protein family is conserved among other plants.

Protein sequences were retrieved from plant databases[16] using *Arabidopsis* EDS1, SAG101 and PAD4 as query sequences in a tBLASTn search[17]. Putative orthologues were aligned and curated manually in several iterations. Sequences with an identity of more than 95 % were considered redundant and removed randomly using the ExPASy "Decrease redundancy" tool[18]. The final set was clustered within a phylogenetic tree using the Phylogeny.fr website [Dereeper *et al.*, 2008].

As shown in Appendix Figure A.1, EDS1, SAG101 and PAD4 fell into three clusters within this tree and allowed to distinguish between them even if the sequences had not been properly

---

[11]Ecotype Landsberg *erecta* and TAIR locus name AT3G48090

[12]Ecotype Columbia and TAIR locus name AT5G14930

[13]Ecotype Columbia and TAIR locus name AT3G52430

[14]Such that have vascular tissues necessary for the transport of fluids and nutrients

[15]Seven were available in 2008 and this number jumped to 36 by the end of 2012 (`http://en.wikipedia.org/wiki/List_of_sequenced_plant_genomes`; request date: January, 2013)

[16]phytozome.net, bioinformatics.psb.ugent.be/plaza, maizesequence.org, NCBI and solgenomics.net

[17]The tBLASTn search uses a protein sequence as input which is being compared to the six-frame translations of a nucleotide database. This is especially useful in coding regions of homologous proteins in unannotated sequences

[18]`http://web.expasy.org/decrease_redundancy/` (request date: December, 2012)

annotated. It was thus possible to analyse features of the proteins individually. Similar results were obtained when the sequences were aligned using T-Coffee [Notredame *et al.*, 2000] and used for tree-building by the SeaView program [Gouy *et al.*, 2010].

Interestingly, orthologues of EDS1 and PAD4 were present among genomes of mono- and dicotyledonous plants whereas SAG101 orthologues existed only in dicotyledonous species (Fig 3.1 A).

Genomes that contained orthologous sequences of EDS, PAD4 (and SAG101 in case of dicotyledonous species) were selected to analyse and compare sequence features of the EDS1 protein family in more detail. The organisation in two domains of which one is lipase-like and the other is unrelated to known proteins and therefore termed EP domain (as presented in Fig. 1.6) was found in all analysed sequences.

3.1.1. EDS1 and PAD4 Contain Conserved Triad Residues

In all orthologues of EDS1 and PAD4, three positions within the lipase-like domain are invariably conserved that correspond to a serine-aspartate-histidine (S-D-H) catalytic triad (Fig. 3.1 B).

A catalytic triad generally consists of a nucleuphilic, an acidic and a basic side chain that orchestrate the reactions of proteases, esterases, lipases and other enzymes [Dodson and Wlodawer, 1998]. Using the crystal structure of bovine chymotrypsin-A, Blow *et al.* [1969] proposed that these residues function as a "charge relay system" in which the proton of the catalytic serine is transferred along hydrogen bonds to histidine and finally aspartate rendering the deprotonated serine nucleophilic. The deprotonated serine is then able to perform a nuclophilic attack on the carbonyl C-atom of a substrate ester bond to start one of manifold reactions. Although it has been shown experimentally that the aspartate residue actually does not change its protonation state during the reaction [Bachovchin, 1985; Kossiakoff and Spencer, 1981], the acidic residue assists proton transfer from the serine to the histidine side-chain probably through stabilising the protonated histidine and fixing the configuration around the active site [Ishida and Kato, 2004]. A triad of serine, asparate and histidine is most commonly found and referred to as the "classical triad" − the concept, however, can also be realised with different residues and in rare cases a dyad is sufficient for catalysis [Rydel *et al.*, 2003; Dessen *et al.*, 1999; Ekici *et al.*, 2008].

Lipases and esterases that use a nucleophilic serine are often characterised by another sequence feature: the serine is embedded in a GxSxG-motif that structurally results in an exceptionally sharp turn in which the nucleophile can be easily approached by the substrate [Brenner, 1988; Nardini and Dijkstra, 1999]. This motif is as well strictly conserved among all EDS1 and PAD4 orthologues (Fig. 3.1 B).

In contrast, *Arabidopsis* SAG101 does not contain a single conserved residue of a S-D-H triad and although individual positions are conserved among SAG101 orthologues from other species, they never constitute a complete triad (Fig. 3.1 B). Paradoxically, SAG101 [He, 2002] and, more convincingly, PAD4 [Louis *et al.*, 2012] have been proposed to be active hydrolases and EDS1 is commonly annotated as such in widely used databases[19]. Still, their catalytic activity is debated [Wiermer *et al.*, 2005] and while many recent studies help to understand where the EDS1 protein family genetically acts [Heidrich *et al.*, 2011; Bhattacharjee *et al.*, 2011; Louis

---

[19]http://www.uniprot.org/; http://www.arabidopsis.org/ and others (request date: January, 2013)
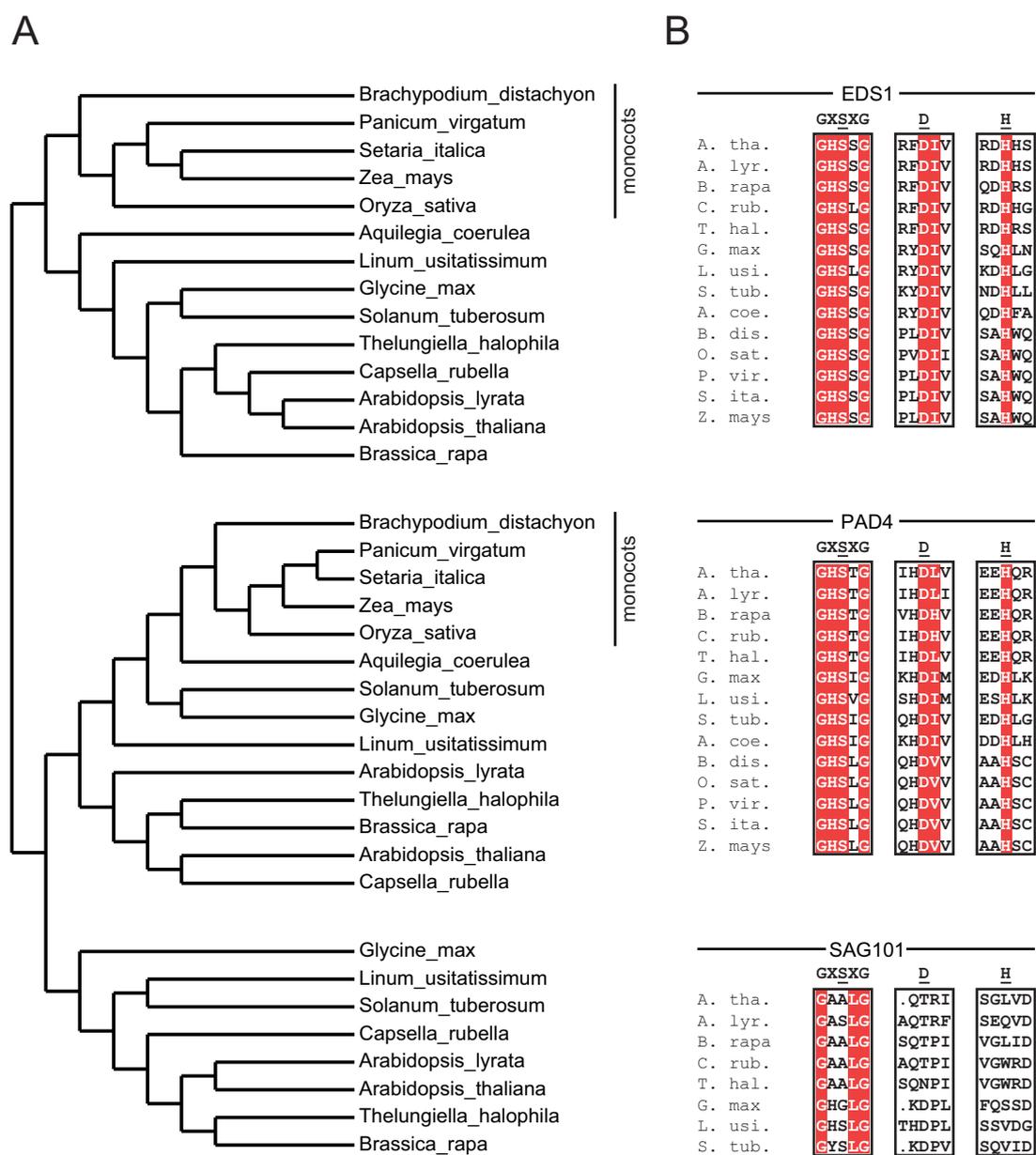
Figure 3.1 | **Sequence features of the EDS1 protein family.** Orthologues of *Arabidopsis* EDS1, PAD4 and SAG101 were retrieved from databases and clustered in a phylogenetic tree as described in the text. (A) shows a cladogram of a sequence subset that unambiguously fell into three clusters. Within each cluster, the sequences were aligned using T-Coffee [Notredame *et al.*, 2000]. The conservation of positions corresponding to the catalytic S-D-H of α/β hydrolases – including a lipase-typical GxSxG motif around the triad's serine – is shown in (B).

et al., 2012; Zhu *et al.*, 2011], the biochemical functions of EDS1, SAG101 and PAD4 remain elusive. Structural information can provide such insights and the following sections shall illustrate the progress towards my goal of structure elucidation.

3.2. The Production and Characterisation of Recombinant Proteins

Although technical progress in recent years has drastically reduced the amounts of protein needed for *in vitro* experiments, preparation on the scale of milligrammes remains a prerequisite, especially in protein crystallography for which proteins additionally have to be exceptionally pure. Even though Lohkamp and Dobritzsch [2008] were able to crystallise a protein (yet not the one that they actually looked for) from a 50 protein mixture and no predominant species, this remains an exception to the rule. In general, one should follow the advice of Alexander McPherson according to whom protein crystallisation relies on three important factors: "purity, purity and homogeneity" [McPherson, 2004].

For protein production, plants themselves have attracted attention as expression hosts with yields of up to 4 mg/g fresh biomass [Marillonnet *et al.*, 2005; Werner *et al.*, 2011]. Yet, this number is inferred from the green fluorescent protein (GFP) expression and differs significantly for other proteins. Additionally, purification from green plant compartments remains challenging and I therefore decided to produce EDS1-related proteins in a bacterial expression system (*E. coli*) that is flexible, easy to cultivate and allows for extremely high production levels of heterologous proteins (60 mg of green fluorescent protein per gramme dry weight were achieved in bioreactors [Aucoin *et al.*, 2006]).
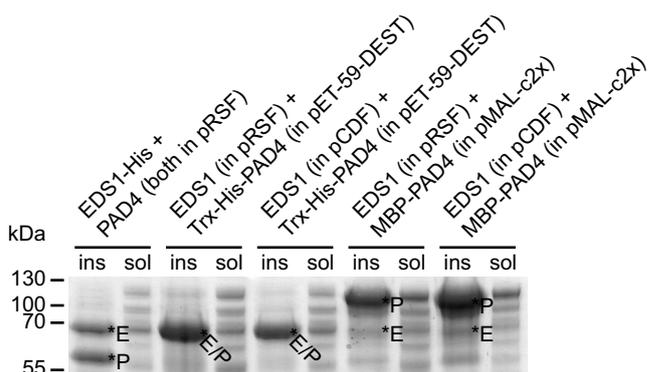
3.2.1. Protein Expression

Variants of EDS1, SAG101 and PAD4 were cloned into expression vectors by Steffen Rietz and J. Stuttmann (both MPIPZ, Köln) carrying different − mostly polyhistidine − or no additional affinity tags. I transformed these into an *E. coli* strain optimised for expression according to section 2.2.12. In addition to full-length proteins from different plant species, N- and C-terminal variants comprising the lipase-like and EP domain, respectively, were proposed on the basis of limited proteolysis experiments (Steffen Rietz) or structural information that I obtained during this thesis. I was able to express major soluble amounts of all proteins that are listed in Table 3.1.

Despite major effort, soluble PAD4 could not be produced recombinantly. Expression of different orthologues, co-expression with EDS1 as a binding partner (and reported stabiliser of PAD4 [Feys *et al.*, 2005, 2001; Rietz *et al.*, 2011]) and of PAD4-fusions with affinity-tags that potentially act as solubility enhancers [LaVallie *et al.*, 1993; Kapust and Waugh, 1999] did not yield noteworthy amounts of soluble protein. One of many expression tests performed according to section 2.3.1 is presented in Fig. 3.2. All protein variants of EDS1, SAG101 and PAD4 that were either not expressed or remained insoluble are listed in Table 3.2.

Table 3.1 | Variants that were expressible as soluble protein on the scale of milligrammes. Yield is given as the approximate amount of soluble protein in mg per litre *E. coli* culture

Variant	MW (kDa)	Used for	Yield (mg/l)
EDS1-His	72.5	Crystallisation, interaction studies, functional studies	8-12
EDS1L262P-His [Rietz *et al.*, 2011]	72.5	Crystallisation	8-12
EDS1 w/o tag	71.5	Crystallisation, interaction studies	8-12
EDS1-TEV-His	73.5	Crystallisation	8-12
Trx-His-EDS1	87	Crystallisation, interaction studies	3-4
EDS1$^{1-362}$-His (lipase-like domain)	41.5	Crystallisation	12-15
EDS1$^{1-384}$-His (lipase-like domain)	44.5	Crystallisation, interaction studies, functional studies	12-15
EDS1$^{1-384}$ w/o tag (lipase-like domain)	43.5	Interaction studies	12-15
EDS1$^{1-616}$-His	72	Crystallisation	7-10
EDS1-His from *Arabis alpina*	72	Crystallisation	3-4
His-SAG101	63	Crystallisation, interaction studies	6-9
SAG101 w/o tag	62	Interaction studies	6-9
SAG101$^{291-537}$-His (EP domain)	31	Interaction studies	>25

Figure 3.2 | **Co-expressions of PAD4 variants.** *Arabidopsis* PAD4 was expressed as a codon-optimised gene either untagged or fused to solubility-enhancing affinity tags (MBP: maltose binding protein; Trx: thioredoxin) in the presence of EDS1 using pRSF-Duet and pCDF-Duet expression vectors. Proteins were expressed over night at 12°C. Soluble (sol) and insoluble (ins) fractions were separated and analysed by SDS-PAGE. Minor amounts of soluble MBP-PAD4 aggregated during purification, probably due to improper folding. Asterisks mark the expected band positions of EDS1 (∗E) and PAD4 (∗P).

3.2.2. Protein Purification

Milligramme amounts of His-tagged EDS1 and SAG101 variants were prepared in two steps: each protein was purified separately through (1) immobilised metal-affinity chromatography (IMAC) and (2) gelfiltration as described in section 2.3.4. An exemplary IMAC purification of EDS1-His and gelfiltration results of frequently used proteins are presented in Fig. 3.3 A and B.

In gelfiltrations following IMAC purification, the majority of each protein eluted in monomeric form. However, especially full-length EDS1 and SAG101 repeatedly formed higher order molecular states that eluted in the unresolved void volume of the column and might be protein ag-

Table 3.2 | Protein variants that I was not able to express and purify due to low level expression or insolubility of the protein.

EDS1 variants	
EDS1$^{1\text{-}234}$-His	
EDS1$^{1\text{-}351}$-His	
EDS1$^{1\text{-}475}$-His	
EDS1$^{1\text{-}604}$-His	
EDS1$^{259\text{-}604}$-His	
EDS1$^{384\text{-}604}$-His	
EDS1$^{368\text{-}623}$-His	
EDS1$^{384\text{-}623}$-His	
His-EDS1 (including Gateway®-linker)	
*St*EDS1 (low soluble ammounts)	
PAD4 variants	
PAD4$^{1\text{-}310}$-His	
PAD4 (co-expressed with EDS1-His)	
GST-PAD4 (codon-optimised; including Gateway®-linker)	
His-PAD4 (codon-optimised; including Gateway®-linker)	
Trx-His-PAD4 (codon-optimised; including Gateway®-linker)	
MBP-PAD4 (codon-optimised)	
Trx-His-PAD4 (codon-optimised; including Gateway®-linker; co-expressed with EDS1)	
MBP-PAD4 (codon-optimised; co-expressed with EDS1)	
*Aa*PAD4-His	
*Aa*PAD4-His (co-expressed with *Aa*EDS1)	
*St*PAD4-His	
SAG101 variants	
SAG101$^{1\text{-}292}$-His	

gregates. Lipase-like N-terminal EDS1$^{1\text{-}384}$ additionally formed a minor state that is, judged by molecular weight size standards, a dimer (Fig. 3.3 B).

All proteins purified in this way were >95 % pure as analysed by SDS-PAGE (Fig. 3.4). For single crystallisation experiments, I additionally purified the proteins by anion exchange chromatography as described in section 2.3.4 directly preceding gelfiltration.

3.2.3. Protein Characterisation

As described in section 2.4.7, polydispersity of a sample (i.e. the amount of objects in solution that differ in size, shape or mass) can be detrimental to protein crystallisation. Polydispersity of selected EDS1 samples prepared for crystallisation was thus analysed directly after gelfiltration by means of dynamic light scattering. As Table 3.3 summarises, the samples did not exceed 30 % Pd and they were thus used for crystallisation without further manipulation. The estimation of molecular weight supported a predominantly monomeric EDS1 lipase-like domain (theoretical mass of 44.5 kDa) in solution, while the estimated weight of full-length EDS1 was in between that of a monomer (72.5 kDa) and a dimer (145 kDa).

Especially for truncated variants of heterologously expressed proteins, proper three-dimensional folding has to be confirmed before they can be used in further experiments. Consequently, I
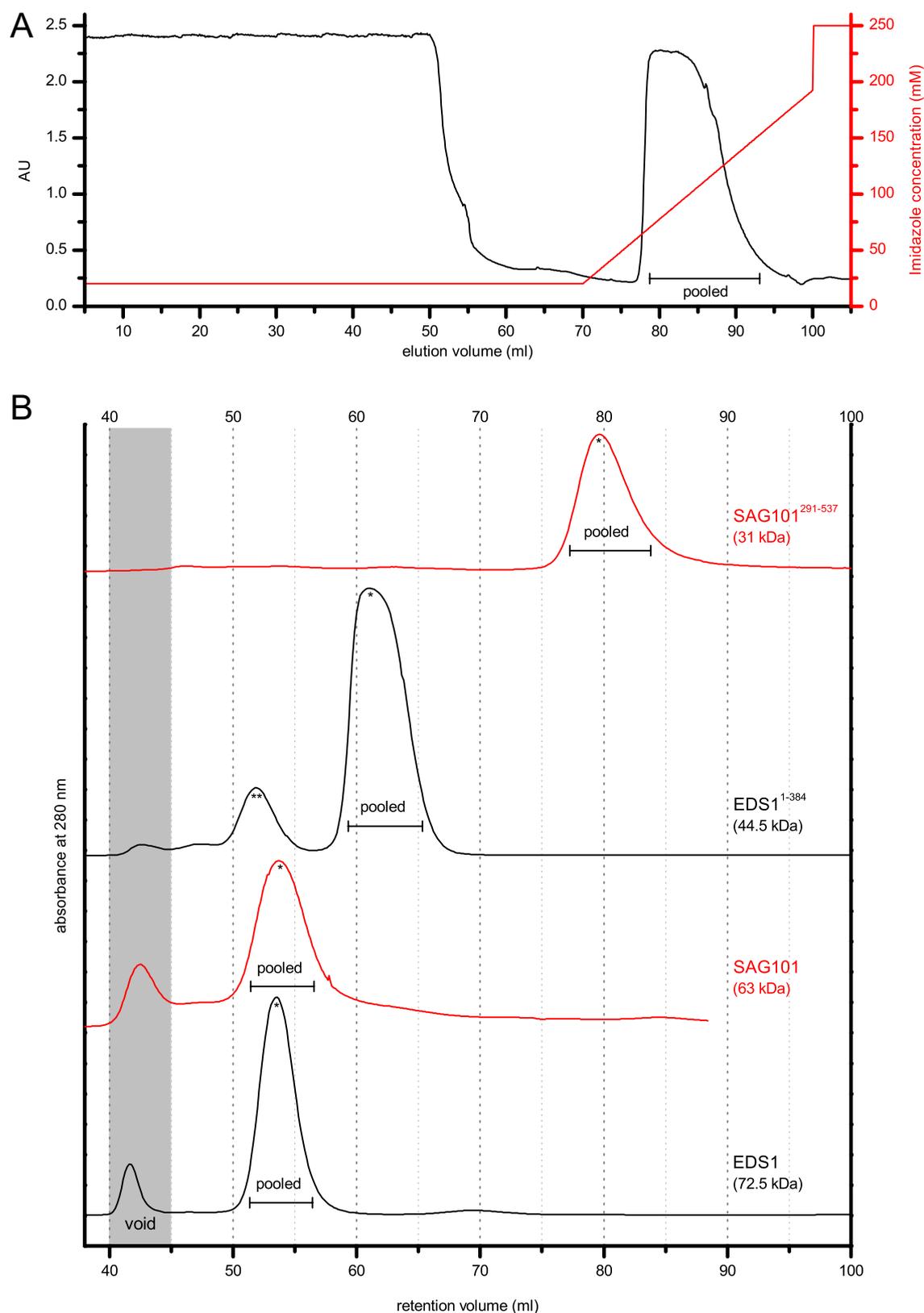
Figure 3.3 | Purification of isolated EDS1 and SAG101 variants. (A) Exemplaric chromatogram of a Ni-NTA purification of EDS1-His. Indicated fractions were pooled and used for further experiments. (B) Each of the presented proteins was applied to a preparative HiLoad 16/60 Superdex 75 column (GE Healthcare) directly after IMAC purification and eluted isocratically. The column's void volume – in which large molecular species are not separated – is indicated. Peaks corresponding to the monomeric form of each protein are marked with an asterisk whereas putative dimers of N-terminal EDS1[1-384] are indicated with two asterisks.
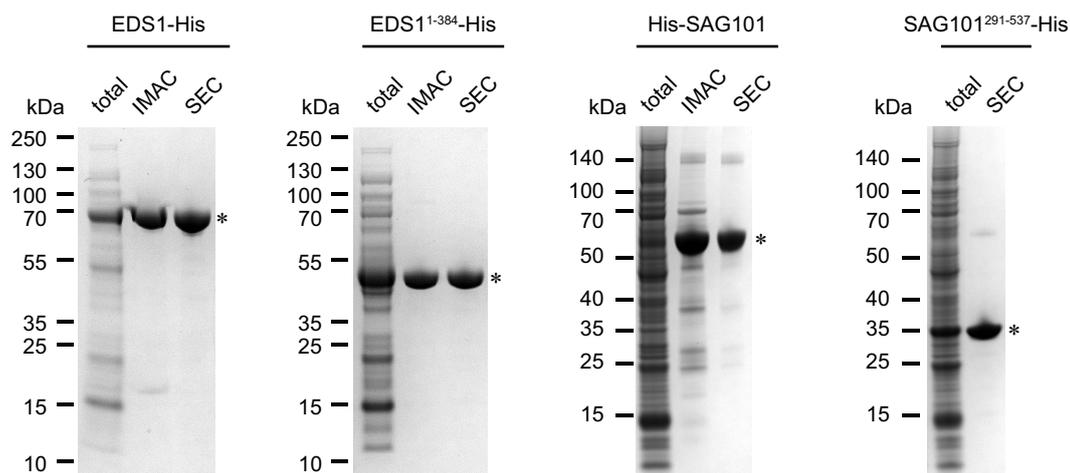
Figure 3.4 | **SDS-PAGE analysis of recombinantly produced proteins.** The total soluble fraction of *E. coli* cell lysates and samples from either both or the final purification step were separated by SDS-PAGE. Bands migrating at the expected size of each protein are indicated by an asterisk. IMAC − immobilised metal-affinity chromatography, SEC − size exclusion chromatography.

Table 3.3 | Polydispersity of protein samples used for crystallisation. Protein samples were analysed directly after gelfiltration and values for the normalised polydispersity (% Pd) and an estimate of the molecular weight were averaged from five consecutive measurements taking 10 acquisitions per measurement.

	EDS1	EDS1$^{1-384}$
Polydispersity (% Pd)	15.0	8.9
Molecular weight (kDa)	101.8	37.6

performed circular dichroism (CD) spectroscopy in the far UV range as described in section 2.4.5. Recorded CD spectra of EDS1$^{1-384}$ and SAG101$^{291-537}$ are shown in Fig. 3.5 A.

These fragments were additionally denatured through a temperature gradient between 10 and 80 °C. Unfolding of secondary structure elements was followed by monitoring the CD signal at a wavelength of 222 nm where especially α-helices contribute to a negative ellipticity value (Fig. 3.5 B and C). The results presented in Figure 3.5 suggested that both domains were intrinsically folded.

Thermofluor assays (section 2.4.6) allow the determination of a protein's melting temperature in the presence of different buffers, salts and additives. As thermostability of a protein is correlated with its ability to form crystals [Ericsson *et al.*, 2006], I performed Thermofluor assays prior to crystallisation setups.

In general, full-length EDS1, N-terminal EDS1 and SAG101 are most thermostable at pH values between 7.5 and 8.5 and less thermostable in the presence of increasing salt concentrations (Fig. 3.6 A and B). However, especially results obtained for full-length proteins have to be treated with caution as their melting curves occasionally showed two melting events, presumably resulting from separated unfolding of lipase-like and EP domain (see Appendix Fig. A.2).

The substances that reproducibly stabilised any of the tested proteins are listed in Table 3.4 for EDS1, SAG101 and EDS1$^{1-384}$. The most effective, trimethylamine N-oxide, is a small organic
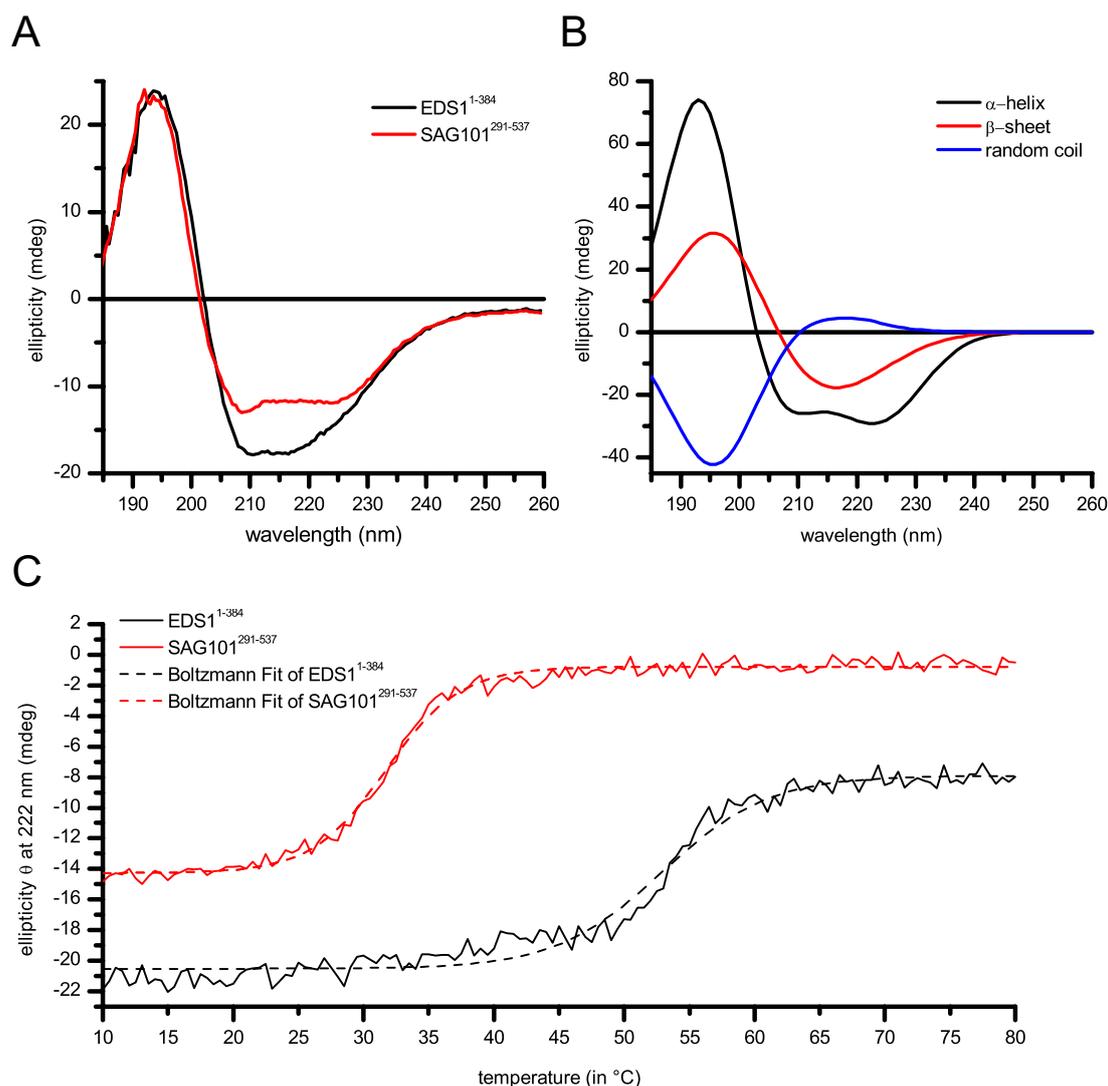
Figure 3.5 | CD spectroscopy of N-terminal EDS1 and C-terminal SAG101. (A) The CD spectra of $EDS1^{1-384}$ and $SAG101^{291-537}$ were taken at 10 °C and in 10 mM NaH_2PO_4, pH 7.5. (B) As a reference, idealised CD spectra of different secondary structure elements are shown in the same range of wavelengths between 185 and 260 nm. (C) The same samples were subsequently heated from 10 °C to 80 °C, following the ellipticity at a wavelength of 222 nm. Both curves were fitted to the Boltzmann equation that allows to derive a melting point, T_m.

osmolyte that naturally occurs in marine organisms where it protects proteins against elevated urea levels and deep sea pressure [Krywka *et al.*, 2008] and has been proposed to be used in protein crystallography by Bolen [2004] and others. I used it at concentrations between 0.5 and 5 % (w/v) in crystallisation screenings and optimisations after recognising the positive effect on protein thermal stability. Similar stabilising effects that promote crystal growth have also been described for glycerol Sousa [1995] and it was thus used at a final concentration of 20 % (v/v) in crystallisation trials. Bicine buffer at pH 8.0 was regularly chosen as gelfiltration buffer during purifications for the same reason.

Figure 3.7 summarises the melting points for EDS1, SAG101 and their domain variants as determined in either circular dichroism or Thermofluor assays. Compared with the corre-
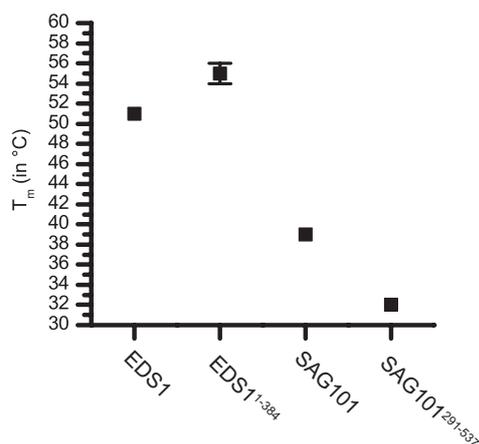
A

B

Figure 3.6 | Trends observed in Thermofluor experiments. Thermal stability of EDS1, the N-terminal EDS1 domain EDS1[1-384] and SAG101 in dependence of different buffer pH values (A) and salt concentration (B). T_m – protein melting temperature.

Table 3.4 | List of substances that reproducibly increased the thermostability in Thermofluor assays compared to a control (water).

Condition	Increase of T_m (in °C)
	EDS1
5 % (w/v) Trimethylamine N-oxide	3
50 mM Bicine, pH 8.0	1
20 % (v/v) Glycerol	1
	EDS1[1-384]
5 % (w/v) Trimethylamine N-oxide	2
50 mM Bicine, pH 8.0	1
20 % (v/v) Glycerol	1
	SAG101
5 % (w/v) Trimethylamine N-oxide	1

sponding full-length proteins, N-terminal EDS1[1-384] is more thermostable whereas C-terminal SAG101[291-537] with a T_m of 32 °C already unfolds at remarkably low temperatures.

Figure 3.7 | **Melting points of EDS1, SAG101 and two of their domains.** The melting temperatures were either determined by Thermofluor (EDS1 and SAG101), circular dichroism (SAG101$^{291\text{-}537}$) or both (EDS1$^{1\text{-}384}$).

3.2.4. Recombinant EDS1 Interacts With and Stabilises SAG101

As known from plants and yeast two-hybrid studies [Feys *et al.*, 2005], EDS1 interacts with the sequence-related protein SAG101. To prepare a recombinant complex of EDS1 and SAG101, I expressed both proteins separately and co-purified an excess of untagged EDS1 with an N-terminally His-tagged SAG101 as described in section 2.3.4 and [Wagner *et al.*, 2011] (Fig. 3.8 A and B). 40 mM imidazole, 50 mM NaCl, 0.5 mM EDTA, pH 7.4 was used as buffer during gelfiltration and in crystallisation trials. The homogeneous complex could be identified as a heterodimer by analytical gelfiltration and subsequent SDS-PAGE as shown in Figures 3.8 B and C. Approximately 8-12 mg of pure EDS1/SAG101 heterodimer were purified from two litres of *E. coli* cell culture (1 l per protein were unified). According to dynamic light scattering experiments directly following gelfiltration, the complex solution was more polydisperse (22.3 % Pd) compared to EDS1 and its lipase-like domain (see Table 3.3) while the molecular weight estimated from these experiments (146 kDa) further supported a heterodimer of actually 134.8 kDa in solution.

Notably, isolated SAG101 tended to precipitate in various buffer systems making *in vitro* experiments difficult and crystallisation virtually impossible. However, this tendency could be dramatically reduced in the presence of recombinant EDS1 as shown in Fig. 3.8 D.

3.3. Crystal Structure of a Functional EDS1/SAG101 Heterodimer

3.3.1. Appearance of EDS1/SAG101 Crystals

The mutual stabilisation of EDS1 and SAG101 (Fig. 3.8 D) may be the reason why of all proteins used in crystallisation trials (refer to Table 3.1) only the EDS1/SAG101 complex, regardless of a rather high polydispersity, yielded crystals suitable for X-ray structure determination. The first intergrown crystals appeared in sitting drop sparse-matrix screenings with 10 % (w/v) PEG4000 and 5 % (v/v) 2-propanol in 100 mM HEPES, pH 7.5 as a crystallant (Classic Screen 3, Jena Bioscience) but did not show any X-ray diffraction on a homesource (SuperNova™, Agilent) or a synchrotron X-ray beam (BESSY, Berlin). After optimisation of the crystallisation conditions, diffracting crystals grew in 5 % (w/v) PEG4000, 5 % (v/v) 2-propanol and 100 mM HEPES, pH 7.5 (Fig. 3.9 A and B). Through the use of crystal seeds, large single crystals grew reproducibly and diffracted to resolutions below 3 Å (Fig. 3.9 C and D).
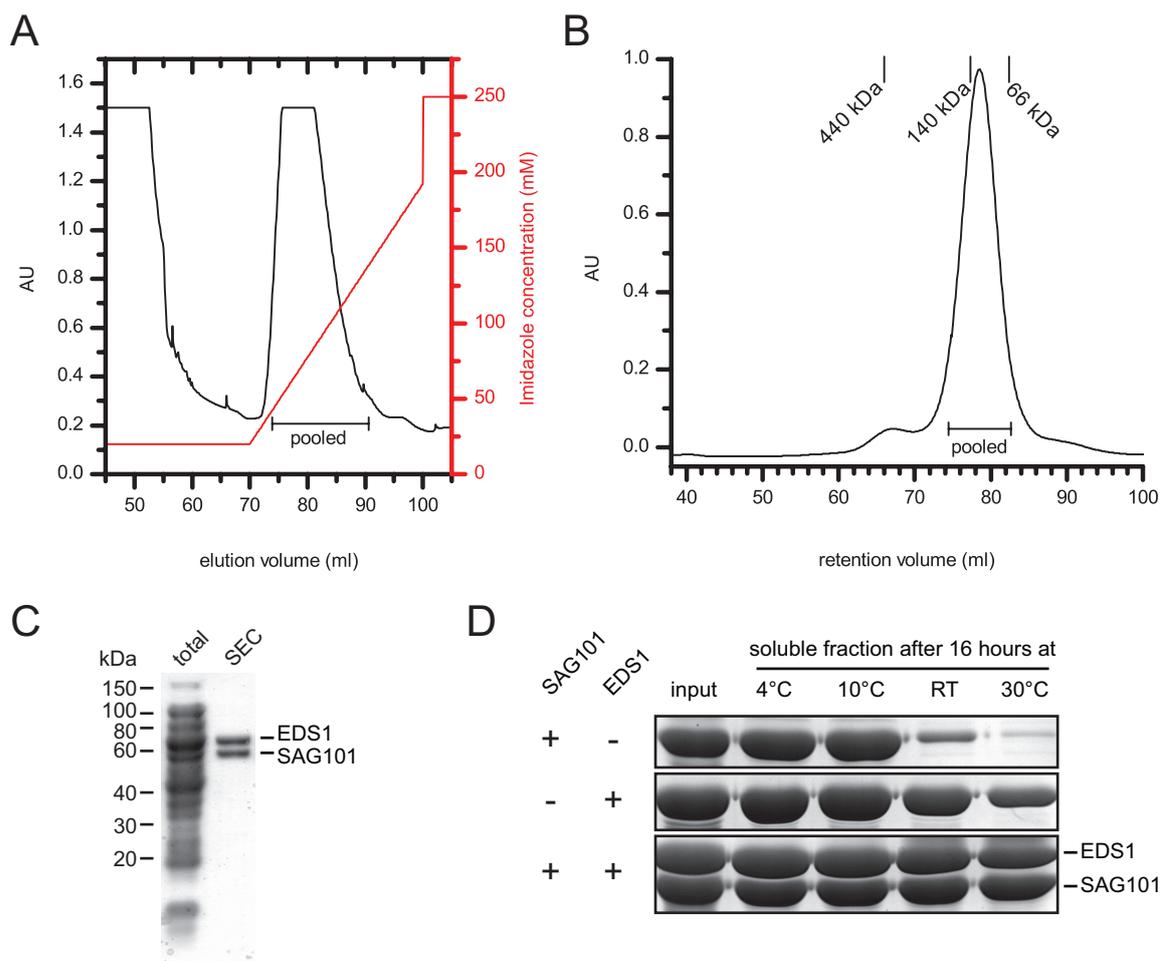
Figure 3.8 | Recombinant EDS1 and SAG101 form a stable heterodimer. (A) 50 ml of an *E. coli* lysate including His-tagged SAG101 and an excess of untagged EDS1 were applied onto a Ni-NTA column of appropriate size. After washing with at least 10 column volumes IMAC lysis buffer, the proteins were eluted using an imidazole gradient to a final concentration of 250 mM. Indicated fractions were pooled, concentrated and applied to a preparative HiLoad 16/60 Superdex 200 column. (B) The proteins were isocracally eluted at a flow-rate of 0.5 ml/min. Indicated fractions containing the EDS1/SAG101 complex were pooled and used for further experiments. The retention volumes of ferritin (440 kDa), lactate dehydrogenase (140 kDa) and bovine serum albumine (66 kDa) are indicated as molecular weight markers. (C) SDS-PAGE analysis of the total soluble fraction after lysing unified pellets of EDS1 and SAG101 and a purified sample after size-exclusion chromatography (SEC). (D) Recombinantly produced and purified EDS1 and SAG101 were left over night at different temperatures either as single protein (two top panels) or mixed in an equimolar ratio (lower panel). On the following day, the protein solutions were centrifuged at 20000 × g and 4 °C for 15 minutes to pellet precipitated protein. Equal amounts of the supernatant were then analysed on SDS-PAGE as shown.

Without using microcrystalline seeds, those crystals were hardly reproducible. Analysis of 288 equally pipetted crystallisation trials set up by a robot (mosquito® Crystal, TTP LabTech) revealed that crystals appeared in only about 5 % of all drops emphasising a statement from Chruszcz *et al.* [2008]: "We predict that, although the proliferation of the best experimental protocols and the development of new methodologies will decrease crystallographers' dependence on fortunate circumstances, luck will still play a significant role in the foreseeable future". Microseeding and streak-seeding (see section 2.5.1) helped to overcome this problem and this was
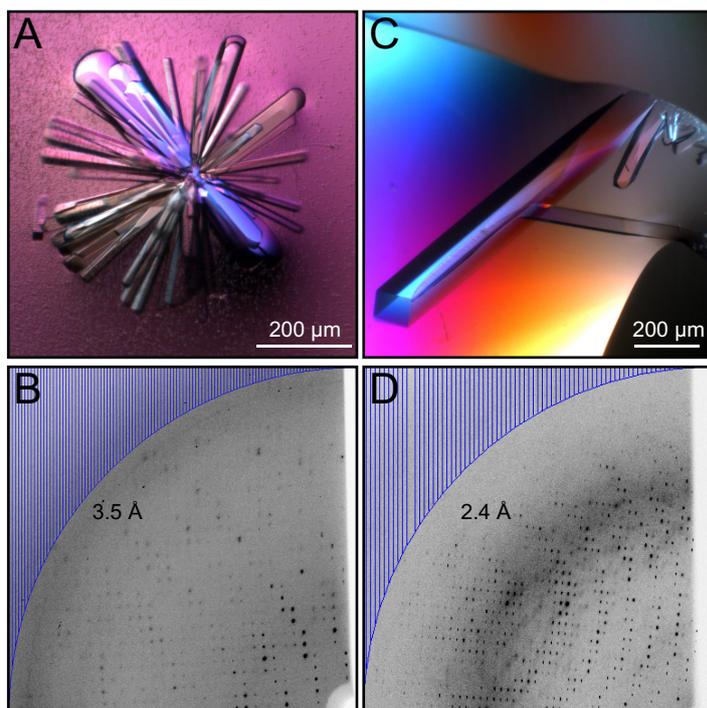
Figure 3.9 | EDS1/SAG101 protein crystals. (A) Intergrown crystals appeared in an optimised condition containing 5 % (w/v) PEG4000, 5 % (v/v) 2-propanol and 100 mM HEPES, pH 7.5. and diffracted to approximately 3.5 Å at a synchrotron beam (B). (C) Through the use of microseed dilutions, large single crystals could be obtained that show increased diffraction of up to 2.2 Å (D).

particularly important when I had to systematically test cryoconditions and heavy atom soaking using (and often destroying) native crystals.

3.3.2. A Gold Salt Allowed Experimental Phase Determination

Solving a protein structure *de novo* requires the experimental determination of crystallographic phases. One way to solve the so-called "phase problem" (see section 2.5.6) relies on the introduction of heavy atoms into the protein in its crystalline state (heavy-atom soaking). In a first experiment, native EDS1/SAG101 protein crystals were incubated in the presence of 30 different compounds at a final concentration of 10 mM (see section 2.5.3 for details). Such compounds that did not immediately lead to macroscopic crystal cracking were chosen for a second experiment in which native crystals were cryoprotected through stepwise transfers into a stabilising solution (100 mM HEPES, 12 % (w/v) PEG 4000, 5 % (v/v) 2-propanol, pH 7.5) containing increasing concentrations of glucose (up to 35 % (w/v) in steps of 5 % incubating for 2 min in between them). Heavy atom compounds solved in the final cryoprotectant (containing 35 % glucose) were added at final concentrations of 10 mM and crystals were incubated in their presence for 1, 5 and 10 minutes before frozen in liquid nitrogen and used for diffraction data collection at a tunable synchrotron X-ray beamline (BESSY, Berlin). Native protein crystals were cryoprotected using the same protocol without additional heavy-atom soaking.

Following a fluorescence scan, I used the CHOOCH program [Evans and Pettifer, 2001] to determine a peak, inflection and a remote wavelength suitable for a multiwavelength anomalous dispersion (MAD) experiment (see section 2.5.6) and collected data were analysed for the presence of an anomalous diffraction signal using XDS [Kabsch, 2010]. A crystal soaked for 10 minutes in the presence of potassium dicyanoaurate was found to be suitable for anomalous data collection. Statistics of the MAD experiment and a data set collected from a native crystal are given in Table 3.5.

The Auto-rickshaw pipeline [Panjikar *et al.*, 2005][20] was used to determine experimental phases. Auto-rickshaw uses several tools from the CCP4 software suite [CCP4, 1994] to prepare and scale the data sets and uses the SHELXC and SHELXD programs [Sheldrick, 2010] to solve the heavy-atom substructure. A total of 8 heavy atom binding sites (Fig. 3.10) were identified, refined and used for phase calculation by the CP3 program (CCP4). The programs DM (CCP4) and RESOLVE [Terwilliger, 2003] were then used for density modification calculations and an initial protein model was automatically build by Buccaneer (CCP4).

Table 3.5 | Data collection statistics for peak (pk), inflection (inf) and high energy remote (her) datasets used to determine experimental phases and a native dataset that was used for structure refinement. Values in parentheses are for highest-resolution shell.

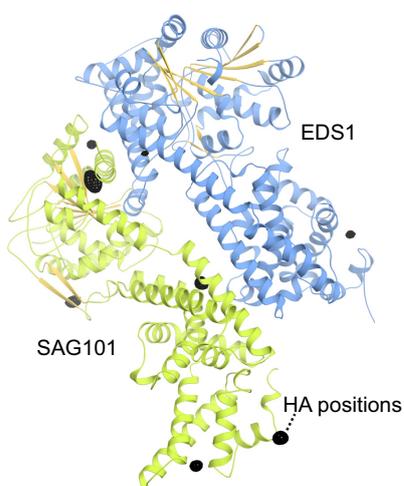|	pk	inf	her	native
X-ray source	BESSY MX14.1	BESSY MX14.1	BESSY MX14.1	BESSY MX14.1
X-ray detector	MAR225	MAR225	MAR225	MAR225
Wavelength (Å)	1.0399	1.0404	1.0313	0.9184
Spacegroup	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a,b,c (Å)	111.5, 112.5, 125.3	111.5, 112.5, 125.3	111.5, 112.5, 125.3	112.6, 113.6, 125.4
Oscillation angle (°)	0.5	0.5	0.5	0.5
Matthews coefficient (Å^3 Da^{-1})				2.32
Solvent content (%)				47
Molecules in ASU	1	1	1	1
Resolution, overall (Å)	35.00-2.80	35.00-2.80	35.00-2.80	35.00-2.21
Resolution, highest shell (Å)	2.87-2.80	2.87-2.80	2.87-2.80	2.25-2.21
Overall B-factor (Å^2)				43.71
Number of observations	276685 (13483)	269898 (13806)	280878 (14693)	613386 (19038)
Unique reflections	74252 (5129)	74668 (5398)	74531 (5183)	78935 (3608)
Completeness (%)	99.4 (93.1)	99.7 (97.5)	99.4 (94.0)	97.3(79.1)
Multiplicity	3.7 (2.6)	3.6 (2.6)	3.8 (2.8)	7.8 (5.3)
Mean I/σ(I)	14.3 (2.0)	17.6 (2.7)	20.8 (3.1)	21.6 (2.0)
R$_{meas}$ (see section A.3)	6.8 (54.8)	5.5 (40.9)	4.9 (38.9)	5.4 (97.2)

Figure 3.10 | **Location of anomalous signal.** An anomalous map calculated by PHENIX [Adams *et al.*, 2011] from the peak (pk) dataset using the final protein model as an additional input reveals the location of anomalous signal within the EDS1/SAG101 structure. Anomalous electron density shown in black was contoured at 7 standard deviation (σ) greater than the mean anomalous electron density of the unit cell and indicates the positions of heavy atoms (HA).

---

[20]Accessible through `http://www.embl-hamburg.de/Auto-Rickshaw/` (request date: December, 2009)

The experimental phases determined by Auto-rickshaw were then combined with structure factor amplitudes collected from a native crystal that diffracted to a resolution of 2.21 Å and used for automated model building by RESOLVE. The final model was refined with the PHENIX software suite [Adams *et al.*, 2011] and iterative manual model building between refinements was carried out in Coot [Emsley *et al.*, 2010]. The process of model building is depicted in Figure 3.11. The coordinates and structure factors have been deposited in the Protein Data Bank under the code 4HVE.
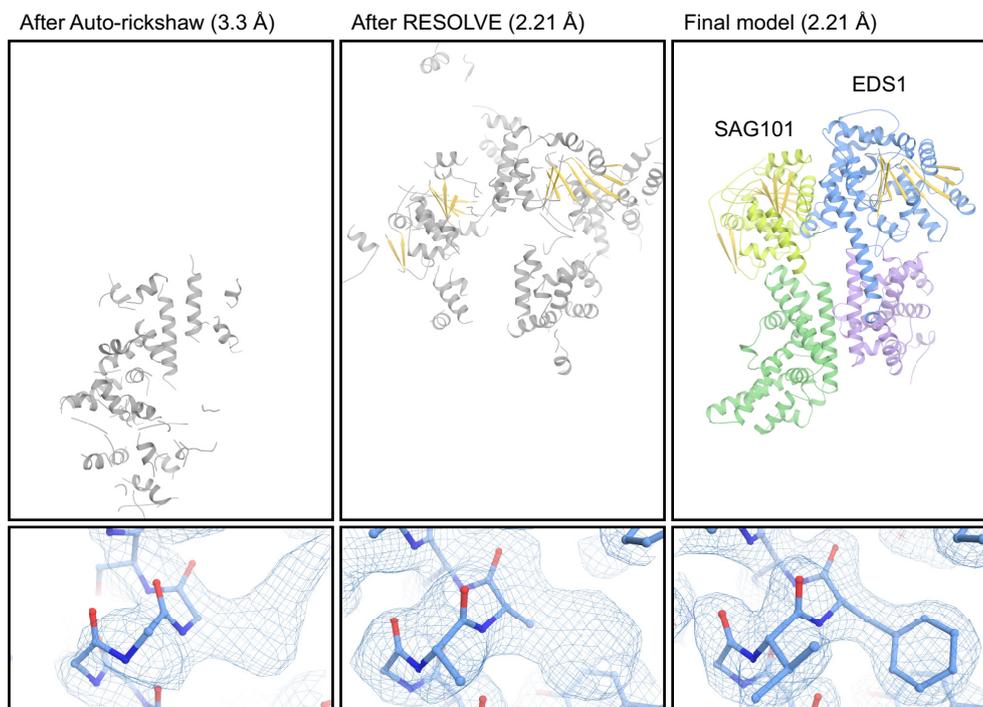
Figure 3.11 | The process of automated and manual model building. The EDS1/SAG101 complex model is depicted after automated model building through Auto-rickshaw using a resolution maximum of 3.3 Å, following automated model building by the RESOLVE program and after manual rebuilding that lead to the final model. The upper panel depicts parts of the model that were built while the lower panel shows the according electron density at each stage. Electron density was contoured at 1 standard deviation (σ) greater than the mean electron density of the unit cell. Automated tools often start by building the protein's backbone (polyglycines) and extend this to polyalanines or the correct side-chain group as the electron density improves.

3.3.3. Both Domains of EDS1 and SAG101 Participate in Heterodimer Formation

The final model contains all but 4 terminal EDS1 residues and all SAG101 residues with a segment missing between Proline 35 and Aspartate 52. As expected from sequence analysis, a central structural feature of EDS1 and SAG101 is their organisation in two distinct domains. As shown in Figure 3.12, both participate in the formation of a head-to-head dimer in which EDS1 and SAG101 interact through a large interface that will be further analysed in section 3.6.
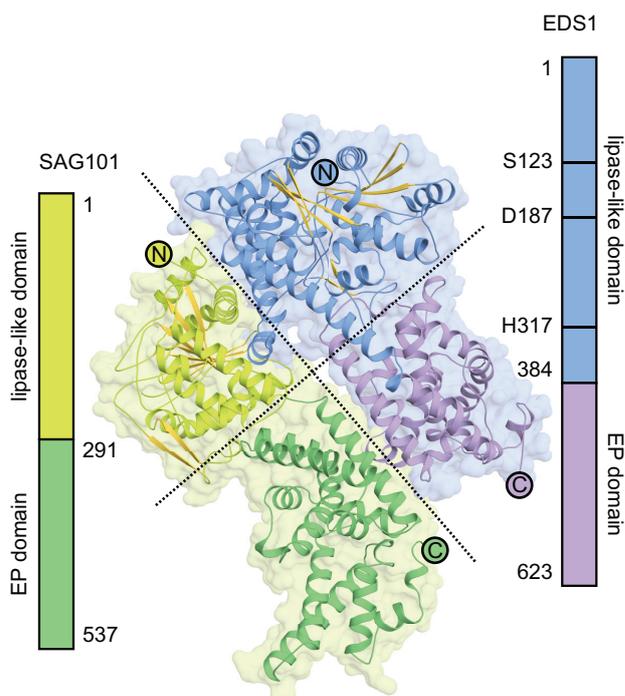
Figure 3.12 | Crystal structure of the EDS1/SAG101 heterodimer. EDS1 and SAG101 are each organised in two distinct domains. Both domains participate in complex formation and both N-terminal and C-terminal domains interact with each other burying a large surface area. While the lipase-like domains contain a core of mainly parallel β-strands (yellow) connected by α-helices, the C-terminal EP domains are exclusively helical. The positions of the conserved S-D-H triad within lipase-like EDS1 are shown schematically.

3.4. The Lipid Link

3.4.1. The EDS1 Lipase-Like Domain Is Closely Related to Hydrolases

As sequence analysis emphasised the relation of EDS1 (and less convincingly SAG101) to lipases, I wondered how closely this is at the structural level. The Dali server [Holm and Rosenström, 2010] allows to compare[21] a query protein structure to all structures deposited within the Protein Data Bank [Berman, 2000][22]. The five closest related structures to N-terminal EDS1 and SAG101 are listed in Table 3.6 and have a common lipase class 3 (triacylglycerol lipase) fold.

I chose a small and well-characterised triacylglycerol lipase from the fungus *Mucor miehei* (MML; PDB codes 3TGL and 4TGL) for a detailed structural comparison to the EDS1 and SAG101 lipase-like domains. For an individual description of secondary structure elements, I proposed a nomenclature that deviates from the canonical α/β hydrolase fold [Ollis *et al.*, 1992].

As shown in topology diagrams (Fig. 3.13), N-terminal EDS1 contains all but the first central β-strand of the MML with the most C-terminally β**10** being inverted and structural related to the MML N-terminus. This differently organised domain-terminus additionally involves α-helices α**K,L,M** and is necessary for the connection to the C-terminal EP domain. The two-domain organisation also results in an extended loop-out between β-strands b**1** and b**2** within EDS1 and SAG101 that contacts the EP domain and probably stabilises the orientation of both domains towards each other. Other eye-catching differences that will be discussed in further detail are: (1) the N-terminal α-helix α**A'** within SAG101 that occupies the position of EDS1 and MML

---

[21]The Dali-server calculates intramolecular distance matrices for each of the protein and aligns them pairwise resulting in a similarity score. The Z-score reported in Table 3.6 is a normalised similarity score and increases with higher similarity between two proteins. Z-scores below 2 indicate no structural relationship, while highly related proteins eventually produce Z-scores of above 16 (as stated in `http://ekhidna.biocenter.helsinki.fi/how/week2/day8/stralignment.pdf`; request date: March, 2013)

[22]Approximately 80.000 protein structures by the end of 2012

Table 3.6 | The five closest structural homologues of the N-terminal EDS1 or SAG101 domain as identified by the DALI-server. For each entry, the PDB entry code (PDB), the DALI Z-score (Z), the average distance between aligned atoms (rmsd), the sequence identity (id) and related entries in the Pfam database are given.

EDS1$^{1\text{-}384}$	PDB	Z	rmsd	id (%)	Pfam
SAG101$^{1\text{-}290}$		22.9	2.9 Å among 243 aligned residues	19	
Feruloyl esterase A	1uza	20.2	2.3 Å among 205 aligned residues	16	Lipase 3
Malassezia globosa LIP1 (SMG1)	3uue	19.5	2.7 Å among 215 aligned residues	16	Lipase 3
Thermomyces lanuginosa lipase (TLL)	1gt6	19.2	2.3 Å among 202 aligned residues	21	Lipase 3
Mucor miehei lipase (MML)	4tgl	18.9	2.6 Å among 208 aligned residues	21	Lipase 3
Gibberella zeae lipase (GZEL)	3ngm	18.4	3.3 Å among 210 aligned residues	17	Lipase 3
SAG101$^{1\text{-}290}$					
Feruloyl esterase A	1uza	19.2	2.6 Å among 209 aligned residues	14	Lipase 3
Malassezia globosa LIP1 (SMG1)	3uue	18.5	2.9 Å among 211 aligned residues	8	Lipase 3
Mucor miehei lipase (MML)	3tgl	18.2	2.8 Å among 206 aligned residues	16	Lipase 3
Rhizopus niveus lipase (RNL)	1lgy	18	2.8 Å among 203 aligned residues	16	Lipase 3
Thermomyces lanuginosa lipase (TLL)	1dt5	17.6	2.8 Å among 201 aligned residues	12	Lipase 3

β**10** and contributes decisively to the EDS1/SAG101 interaction interface,

(2) the α-helix a**1** that plays a crucial role in the regulation of MML and is degenerated to a 3_{10}-helix in EDS1 and SAG101 and, most prominent,

(3) an insertion into the canonical α/β hydrolase fold between EDS1 α**E** and β**8** that comprises 73 amino acids between residues I192 and Y267 folding into three α-helices, α**F,G,H**. This will be referred to as EDS1 "N-terminal insertion" throughout the rest of the text.

Notably, SAG101 misses this EDS1-like insertion between the aspartate and histidine triad positions. Moreover, the protein backbone surrounding the SAG101 aspartate-position is differently organised in SAG101 and EDS1 (Fig. 3.13 and 3.14). While this strengthens the notion that the SAG101 triad is generally unconserved, a detailed look at the EDS1 triad (S123, D187, H317) reveals that this is not only conserved sequentially but also structurally (Fig. 3.1 B and 3.14).

3.4.2. EDS1 Contains an Insertion Into the Canonical α/β Hydrolase Fold That is Conformationally Stabilised Through SAG101

The N-terminal EDS1 insertion introduced above is interesting in two ways: firstly, it covers the conserved triad residues and, secondly, it directly interacts with SAG101.
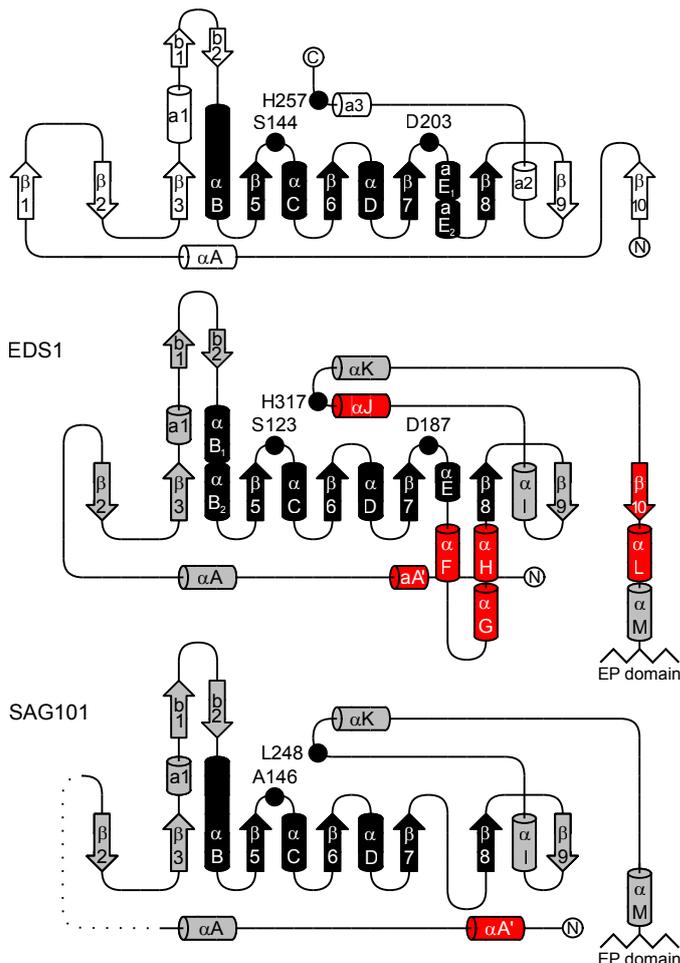
Mucor miehei lipase

Figure 3.13 | **Topology diagrams of the N-terminal EDS1 and SAG101 domains compared to an active lipase.** In the topologies of an active lipase from *Mucor miehei* (MML) and the lipase-like domains of EDS1 and SAG101 elements exhibiting the canonical fold of α/β hydrolases are in black while elements that are structurally conserved among EDS1 and SAG101 are in grey. Secondary structure elements that are differently organised in EDS1 and SAG101 are in red and contain the N-terminal EDS1 insertion (α**F,G,H**). A SAG101 region that could not been placed due to missing electron density is indicated as black dashes. The diagrams were created using TopDraw [Bond, 2003]
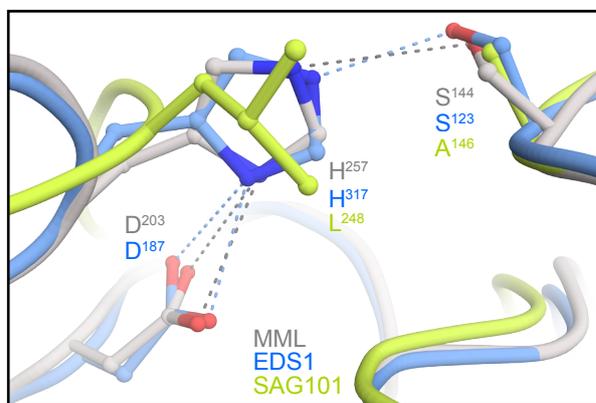
Figure 3.14 | **Detailed view on the EDS1 triad.** The S-D-H triad positions of EDS1 (blue), SAG101 (green) and *Mucor miehei* lipase (grey) were superimposed using least-square fitting and the residue numbers are given for all triad positions.

Intriguingly, the structure of the *Mucor miehei* lipase that I already used for structural comparison was initially crystallised as well in a state in which the triad was solvent-shielded [Brady *et al.*, 1990]. Way before that, Sarda and Desnuelle [1958] had observed that lipases are activated in the presence of an oil-water interface, a mechanism termed interfacial activation[23]. Brady *et al.* [1990] consequently speculated that a protein loop that they refer to as the "lid" covers the active site but can be "removed or displaced, possibly through interfacial activation".

---

[23]The authors find lipase activity and interface dimension to be as related as enzymatic activity and substrate concentration: "l'activité de la lipase dépend des dimensions de l'interface de l'émulsion comme l'activité d'un enzyme ordinaire dépend de la concentration de la solution du substrat."

They could confirm this by co-crystallising the same lipase with a small-molecule inhibitor that stabilises the enzyme's open conformation [Derewenda *et al.*, 1992]. The transition between open and closed state involves particularly hydrophobic side chains from α-helix a**1** (Fig. 3.13 and 3.15 A) that is able to bind a hydrophobic interface (for instance lipid or detergent micelles or a crystal contact), thereupon undergoes a conformational change and thus renders the active site accessible.

This regulatory MML helix a**1** is degenerated to a 3_{10}-helix in SAG101 and EDS1 that does not completely restrict solvent accessibility of the triad pocket (Fig. 3.15 B and C). Within EDS1, however, the hydrophobic side-chains F47 and F52 stabilise the position of helix α**F** (being part of the N-terminal insertion) that folds on top of the triad (Fig. 3.16 A).

Judging by crystallographic B-factors that are correlated with the structural flexibility of a protein region, α-helices α**F** and α**G** are rather rigid and only a small loop between them possesses flexibility (Fig. 3.16 B). In this state, which is not stabilised through crystal contacts, the triad would be hardly accessible for a small molecule. Yet, it is important to notice that helix α**H** and the preceding α**G** (both are part of the N-terminal EDS1 insertion) are in direct contact to SAG101 (Fig. 3.15 C). I therefore asked whether the N-terminal EDS1 insertion could – as a whole or partly – undergo conformational changes upon SAG101 dissociation.
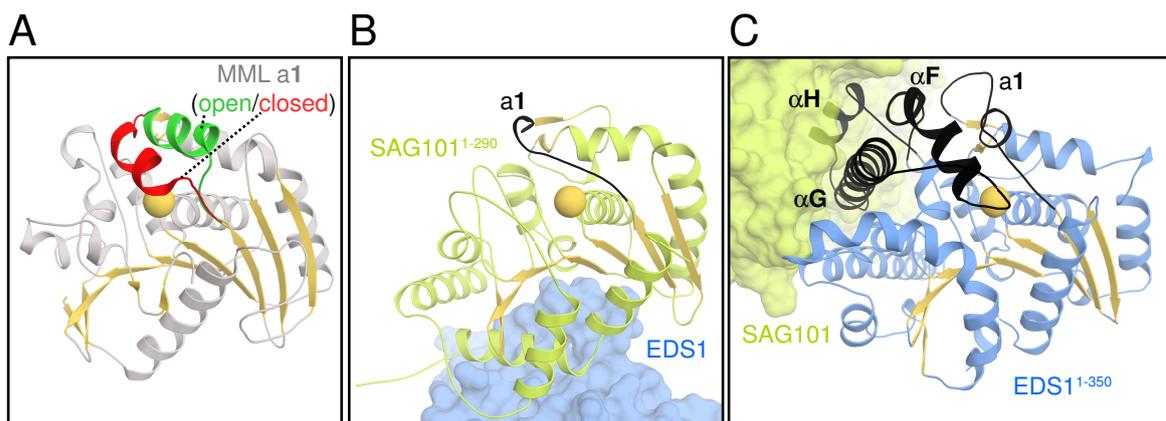

Figure 3.15 | The N-terminal domains of EDS1 and SAG101 adopt an α/β hydrolase fold. The active lipase from *Mucor miehei* (grey) was superimposed with either EDS1$^{1\text{-}384}$ (blue) or SAG101$^{1\text{-}290}$ (green) and the proteins are individually presented. (A) Within the MML, α-helix **a1** that regulates activity, as described in the text, is in green (open conformation) or red (closed conformation). (B) The 3_{10} helix **a1** of SAG101 is in black, bound EDS1 is shown as surface representation. (C) Within lipase-like EDS1, 3_{10} helix **a1** and the N-terminal insertion comprising helices α**F,G,H** are in black. The surface of bound SAG101 is shown. The MML active site and analogous positions within EDS1 and SAG101 are indicated as a yellow sphere.

Consequently, I computationally calculated the motions of either isolated EDS1 or EDS1 in complex with SAG101 in a water box as described in section 2.5.8 for a time frame of 50 ns. As shown in Figure 3.16 C, simulated motions surrounding the triad generally resemble what could have been inferred from crystallographic B-factors: while there is structural flexibility between α-helices α**F** and α**G**, the insertion adopts an overall stable conformation. Yet, if the motions of isolated EDS1 were simulated, the third helix within the N-terminal insertion (α**H**) became flexible as shown in Figure 3.17. This helix is characterised by hydrophobic side-chains

and will be important for the detailed discussion of EDS1/SAG101 interaction in section 3.6. Hence, a conformational change upon SAG101 dissociation in an aqueous environment was to be expected but could be accompanied by concerted molecular motions that take to long to be computationally modelled. I thus did not want to generally exclude the possibility that the EDS1 triad could be accessed by a small molecule.

Structural motions calculated by the NMSim tool in "large scale motions" mode [Krüger *et al.*, 2012][24] emphasised the flexibility of EDS1 helix αH upon SAG101 dissociation but did not point to associated motions of other structural elements (see Appendix Fig. A.3).
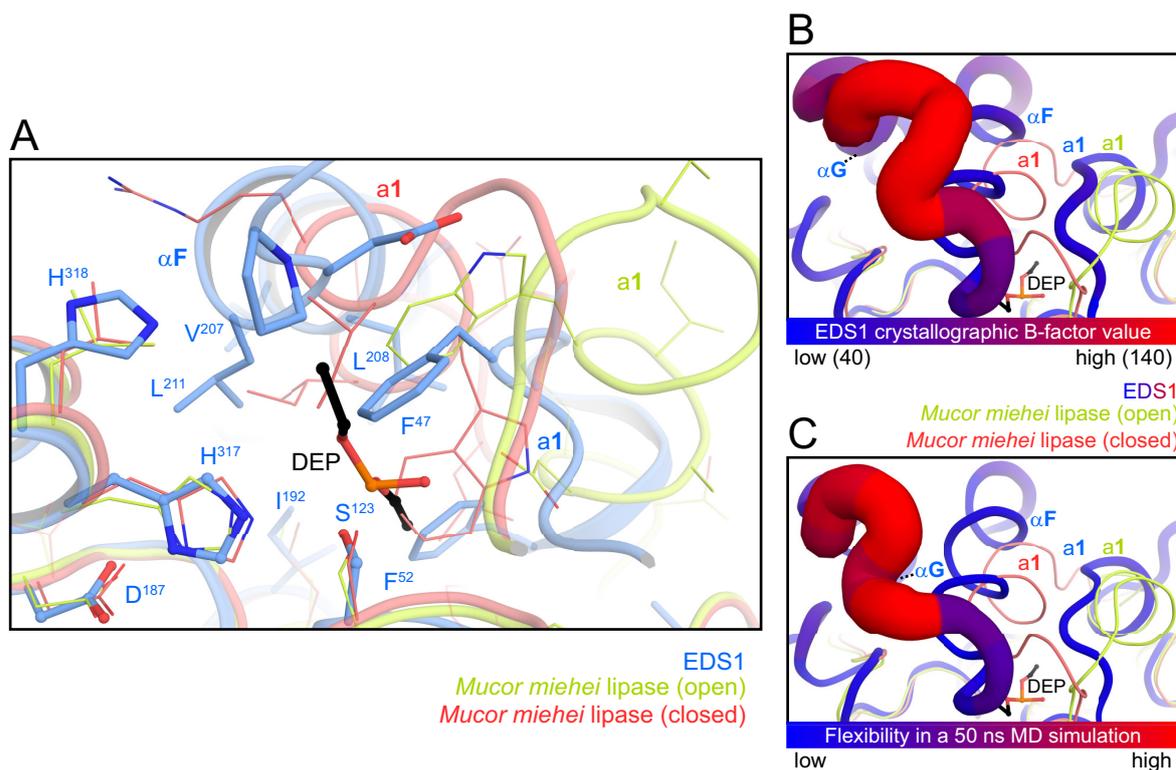
Figure 3.16 | The EDS1 triad might be inaccessible. (A) N-terminal EDS1 (blue) was superimposed with the *Mucor miehei* lipase (MML) in open (green) and closed (red) conformation. The lipase inhibitor DEP (diethylphosphonate) that stabilises the open conformation of MML is drawn to indicate where a ligand or substrate could bind. Within EDS1, the triad residues and such that effectively block solvent access to them are drawn as sticks. (B) is a schematic representation of (A) in which the EDS1 elements are coloured based on their crystallographic B-factors. High B-factor values (reddish) indicate high structural flexibility. (C) shows the flexibility of the same region as determined in a 50 ns molecular dynamics simulation of EDS1/SAG101.

Sequence analysis revealed that *Arabidopsis* PAD4 does not contain the EDS1-like N-terminal insertion and this, in combination with a conserved S-D-H triad among all PAD4 orthologues (Fig. 3.1 B), renders this protein probably the most obvious candidate for hydrolytic activity. It shall, however, be pointed out that actually most PAD4 orthologues <u>do</u> contain an N-terminal insertion that could as well block access to the triad (Fig. 3.18).

---

[24]http://cpclab.uni-duesseldorf.de/nmsim/ (request date: September, 2012)
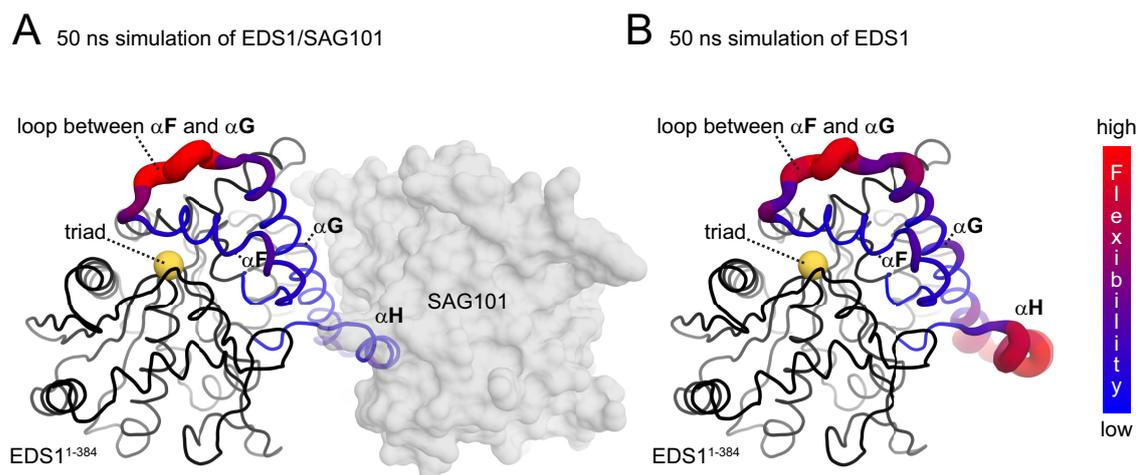
Figure 3.17 | **Structural flexibility of the EDS1 insertion upon SAG101 dissociation.** Atomic motions of EDS1/SAG101 (A) and EDS1 alone (B) in a waterbox were simulated over 50 ns. The N-terminal EDS1 insertion encompassing α-helices α**F,G,H** is coloured according to structural flexibility. The position where a small molecule ligand or a substrate could bind is indicated by a yellow sphere.

Figure 3.18 | The N-terminal insertion within EDS1 and PAD4 orthologues. Sequences of EDS1 and PAD4 orthologues between positions corresponding to the *Arabidopsis* EDS1 triad aspartate (D187) and Y267 were aligned by the MUSCLE tool [Edgar, 2004] and visualised using ESPript [Gouet *et al.*, 1999]. Protein sequences from *Brassicaceae* species are highlighted in red. Residues strictly conserved have a black background while residues well conserved according to Risler *et al.* [1988] are in bold letters and boxed.

3.4.3. EDS1 Does not Depend on Hydrolytic Activity in Central Immune Pathways

To test whether the triads of EDS1 and PAD4 were conserved throughout evolution due to a function in *Arabidopis* immune signalling, J. Stuttmann (MPIPZ, Köln) stably transformed an *eds1-2/pad4-1* double null mutant in accession Columbia (Col-0) with an EDS1 variant that misses all three triad residues (S123A+D187A+H317A) and a PAD4 variant in which the potentially catalytic serine S118 was exchanged to alanine and expressed both variants under their own promoters.

These plants were infected with an incompatible isolate (Emwa2) of the oomycete *Hyaloperonospora arabidopsis*. Incompatible *Hyaloperonospora* isolates trigger a TIR-NB-LRR-mediated hypersensitive response on wildtype *Arabidopsis* plants. As introduced, the hypersensitive response is a defencive cell death mechanism that effectively stops pathogen invasion. This cell death can be visualised by the diazo dye Trypan Blue that stains fungal structures (hyphae, oospores etc.) and dead but not vital plant cells. As controls, J. Stuttmann (MPIPZ, Köln) additionally infected plants that were transformed accordingly with wildtype variants of EDS1 and PAD4 and plants that did not express EDS1, PAD4 or both proteins. While these single and double null mutant plants failed to induce defencive cell death at the site of *Hyaloperonospora* infection, the triad mutant variants of EDS1 and PAD4 restored a hypersensitive response to wildtype levels indicating that catalytic activity is dispensable in this context (Fig. 3.19 A). Partial complementation of line #45_3 can be explained with a comparably low level of mutant protein (Fig. 3.19 C).

The same plants were further infected with the compatible *Hyaloperonospora* isolate Noco2. Compatible *Hyaloperonospora* isolates do not trigger localised cell death but are still restricted in growth by residual plant defence mechanisms – the basal resistance. We found that the triad mutants of EDS1 and PAD4 were fully functional in this context whereas plants missing EDS1, PAD4 or both proteins were significantly compromised in their ability to restrict the oomycete's growth (Fig. 3.19 B). For both experiments, J. Stuttmann (MPIPZ, Köln) selected plant lines that accumulated the EDS1 variant to wildtype amounts (Fig. 3.19 C).

3.4.4. Small Molecule Binding to the Triad Seems Improbable

As the EDS1 triad could be further conserved due to a function in binding instead of hydrolysing small molecules, I proposed structure-guided exchanges that further perturbed a putative binding pocket (Fig. 3.20 A). J. Stuttmann transformed these variants stably in *eds1-2* null mutant plants and infected these with the incompatible *Hyaloperonospora* isolate Cala2 that does not survive on wildtype Col-0 plants. While the oomycete was able to sporulate on *eds1-2* plants, all transformed EDS1 variants, including those with a completely artificial environment around the conserved triad, were fully functional in restricting pathogen growth (Fig. 3.20 B).

3.4.5. Recombinant EDS1 Does Not Hydrolyse Artificial Substrates

Given that EDS1 has been identified as a component of further signalling cascades including an immunity-unrelated response to photo-oxidative stress [Straus *et al.*, 2010; Rustérucci *et al.*,
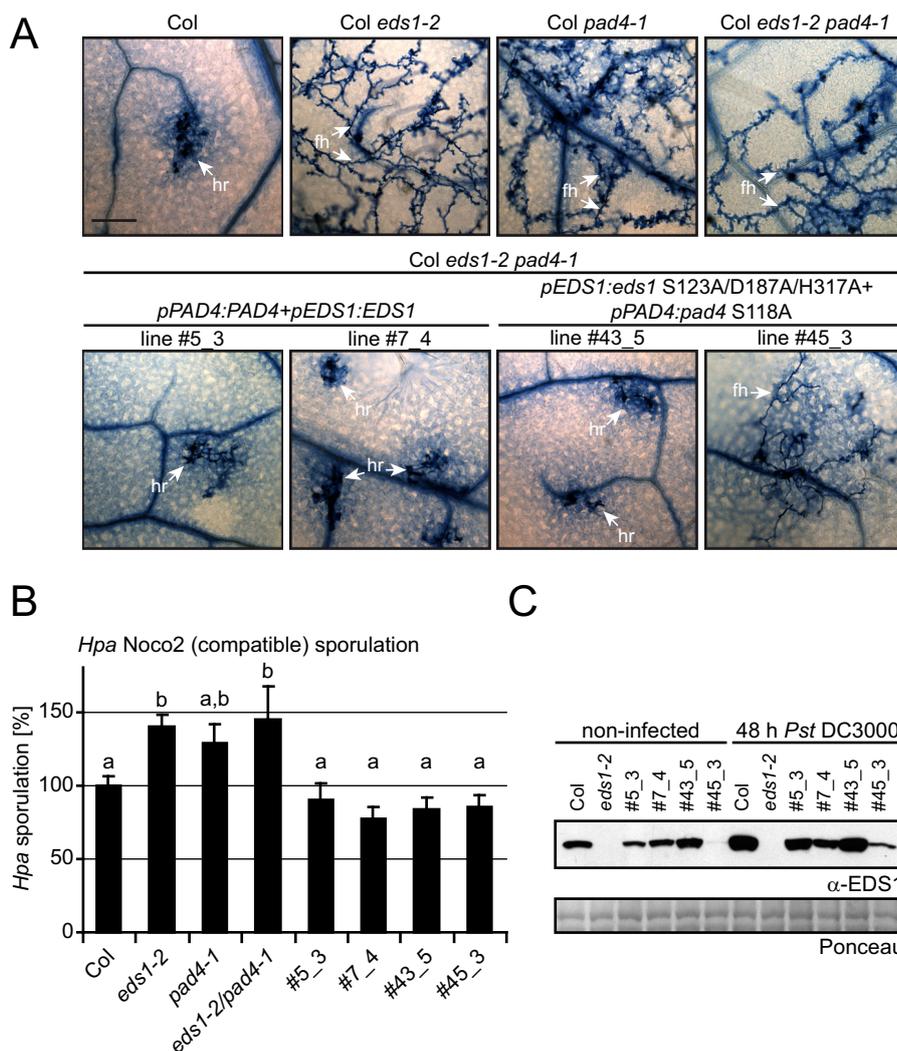
A

| Col | Col *eds1-2* | Col *pad4-1* | Col *eds1-2 pad4-1* |

Col *eds1-2 pad4-1*

| *pPAD4:PAD4+pEDS1:EDS1* | | *pEDS1:eds1* S123A/D187A/H317A+ *pPAD4:pad4* S118A | |
| line #5_3 | line #7_4 | line #43_5 | line #45_3 |

B

Hpa Noco2 (compatible) sporulation

C

Figure 3.19 | EDS1 and PAD4 are not acting as hydrolases in central immune pathways. (A) Control plants and homozygous *eds1-2 pad4-1* null mutant plants were transformed with PAD4, EDS1 or mutant variants carrying exchanges in residues of the catalytic triad under their own promoters (*pEDS1/pPAD4*). These were infected with the incompatible *Hyaloperonospora* isolate Emwa2. Infected leaves were stained with Trypan Blue 4 days after infection and analysed under a microscope. hr − hypersensitive response, fh − free hyphae. Scale bar= 200 μm. (B) The same lines as in (A) were infected with the compatible *Hyaloperonospora* isolate Noco2 and fungal growth was scored by spore counting. Data from four independent experiments, each containing ≥ 3 biological replicates, were normalised to Col= 100 % sporulation. Letters indicate statistically different classes. (C) Steady-state and pathogen-induced EDS1 protein accumulation in lines used in (A) and (B). Total protein extracts were prepared from three week-old plants that were either left uninfected or exposed for 48h to *Pseudomonas syringae* pv. tomato (*Pto*) DC3000 and subjected to SDS-PAGE followed by immunoblot analysis using α-EDS1. Ponceau staining of the membrane is shown as a loading control.
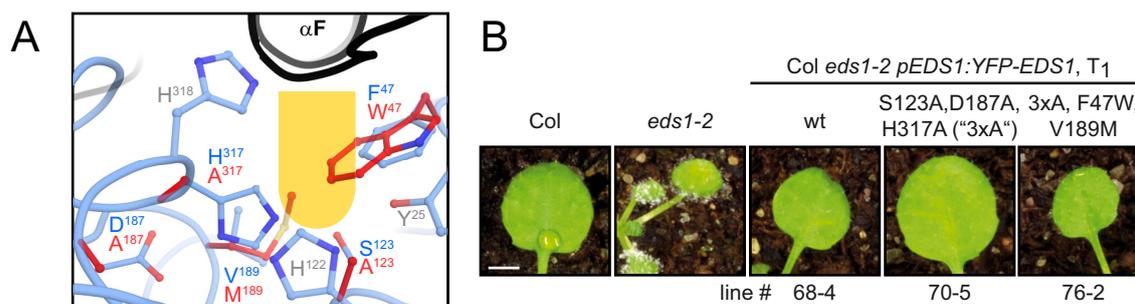
Figure 3.20 | The EDS1 triad does not bind a small molecule in TIR-NB-LRR-mediated resistance.
(A) shows a comparison of the putative ligand binding site in the EDS1-SAG101 structure (blue) and a model of the modified protein containing S123A, D187A, H317A, F47W and V189M mutations (red). Estimated space available for a ligand is indicated in yellow, residues potentially contacting a ligand and the αF helix covering the substrate binding site in the crystal structure are shown. (B) *eds1-2* null mutant plants with EDS1 variants expressed as C-terminal fusions to YFP under control of the EDS1 promoter (*pEDS1*) and control plants were infected with the incompatible *Hyaloperonospora* isolate Cala2 and macroscopic disease symptom formation was scored 7 days after infection. Scale bar = 2.5 mm. Expression of YFP-EDS1 fusion proteins and complementation in individual lines is shown in Appendix Figure A.4.

2001; Ochsenbein *et al.*, 2006], I analysed EDS1 hydrolytic activity *in vitro* by measuring the hydrolysation of artificial substrates. Recombinantly produced EDS1 and its lipase-like domain EDS1$^{1-384}$ were incubated separately in the presence of p-nitrophenyl esters of varying chain lengths while following the production of the photometrically active product p-nitrophenolate. While pNP-esters with shorter chains (pNP-acetate, -propionate, -butyrate) are preferentially hydrolysed by esterases, long-chain esters (pNP-octanoate, -myristate, -palmitate) are more typical lipase substrates. In two completely independent experiments, neither EDS1 nor its lipase-like domain hydrolysed any of the six substrates under the given conditions while the lipase of *Mucor miehei*, which I already used for structural comparison, does so (Fig. 3.21).

3.4.6. Lipase-Like EDS1 Is Not Sufficient for Proper Immune Signalling

These results show that EDS1 does neither bind nor hydrolyse a small molecule in central *Arabidopsis* immune pathways. Recombinant EDS1 was further not able to hydrolyse artificial esterase and lipase substrates.

To examine the importance of lipase-like EDS1 in *Arabidopsis* resistance, J. Stuttmann (MPIPZ, Köln) transformed *eds1-2* null mutant plants with full-length EDS1 and the previously characterised variant EDS1$^{1-384}$ under a constitutive 35S promoter and infected these with an incompatible isolate of *Hyaloperonospora*, Cala2. While plants expressing full-length EDS1 were able to restrict *Hyaloperonospora* growth through the establishment of a hypersensitive response at the infection site, plants that did either miss EDS1 or expressed only its lipase-like domain were severely compromised in TIR-NB-LRR-mediated signalling (Fig. 3.22 A). It was ensured that several plant lines were tested and that these expressed the lipase-like domain to sufficient levels (Fig. 3.22 B). In localisation studies with YFP-fused EDS1$^{1-384}$, J. Stuttmann (MPIPZ, Köln) was additionally able to show that the inability to trigger an immune response was not due to mislocalisation of the protein domain (see Appendix Fig. A.5).
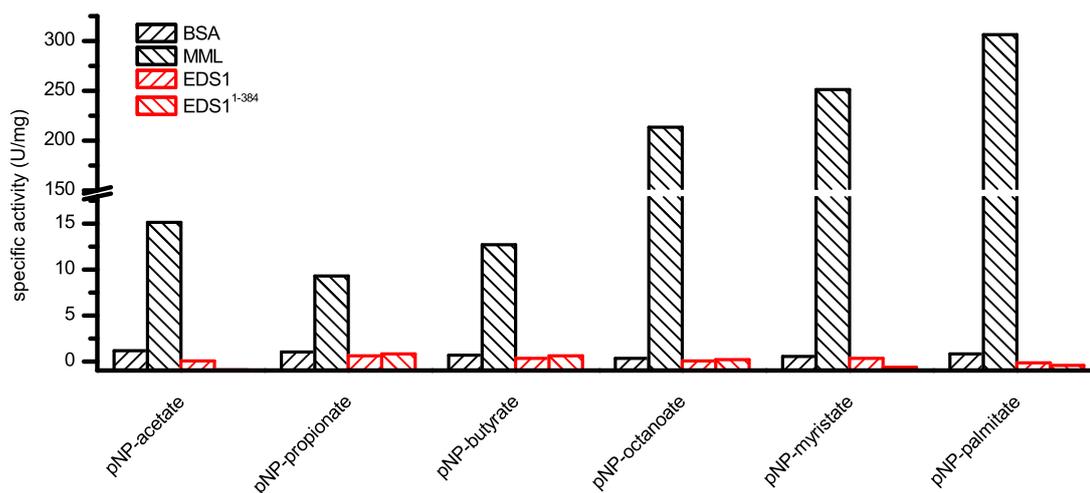
Figure 3.21 | Enzymatic activity of recombinant EDS1 on artificial p-nitrophenyl substrates. Recombinantly expressed EDS1 and its lipase-like domain EDS1$^{1-384}$ were assayed for hydrolytic activity as described in section 2.4.8. The specific activities of both, a positive control (MML − *Mucor miehei* lipase) and negative control (BSA − bovine serum albumin) are shown in the graph for different p-nitrophenyl (pNP) esters.
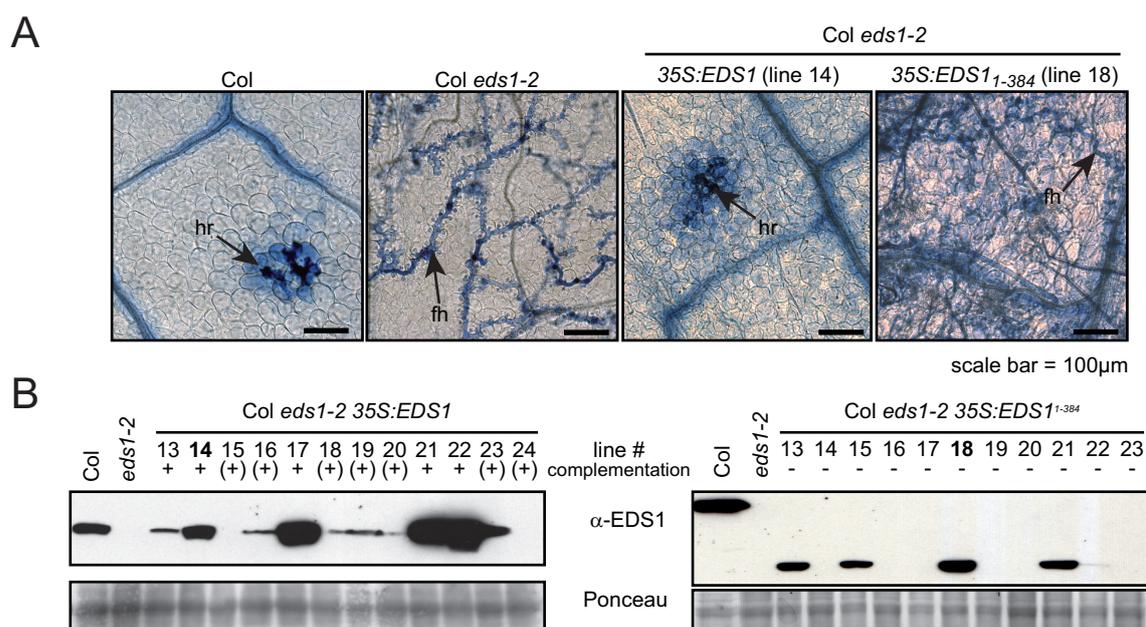

Figure 3.22 | The N-terminal EDS1 domain does not complement for loss of full-length EDS1. (A) Control plants and hemizygous T1 transformants expressing either full-length EDS1 or its N-terminal domain under control of the 35S promoter were infected with the incompatible pathogen *Hyaloperonospora* Cala2. Leaves were stained with Trypan Blue and documented 7 days after inoculation. Activation of an EDS1-mediated immune response leads to a hypersensitive response (hr) including cell death at the site of infection in wild type plants (Col) and plants expressing sufficient EDS1 under a 35S promoter instead of endogenous EDS1. Plants that either do not express EDS1 at all (Col *eds1-2*) or sufficient amounts of EDS1$^{1-384}$ under a 35S promoter show free hyphal growth (fh) as a result of impaired immune signalling. (B) Protein accumulation and complementation of individual T1 transformants shown in (A).

3.5. The C-terminal Domains of EDS1 and SAG101 Are Only Distantly Related to Known Proteins

The inability of the lipase-like EDS1 domain to complement for the loss of full-length EDS1 in plant immunity emphasised the importance of the C-terminal EP-domain. Sequence analysis had shown that this domain is present in all EDS1, PAD4 and SAG101 orthologues and has no other relatives on sequence level. To search for structural homologues, I compared EDS1[385-623] and SAG101[291-537] with all available protein structures in the PDB using the Dali algorithm as described in section 3.4.1 and Table 3.7 list the 5 closest homologues to both domains.

Table 3.7 | The five closest structural homologues of the C-terminal EDS1 or SAG101 domain as identified by the DALI-server. For each entry, the PDB entry code (PDB), the DALI Z-score (Z), the average distance between aligned atoms (rmsd), the sequence identity (id) and related entries in the Pfam database are given.

EDS1[385-623]	PDB	Z	rmsd	id (%)	Pfam
SAG101[291-537]		17.5	2.7 Å among 185 aligned residues	31	
Negative regulator of ofd1 (Nro1)	3qtm	5.5	4.2 Å among 145 aligned residues	8	n/a
Glutathione S-transferase	1yq1	5	3 Å among 92 aligned residues	8	Glutathione S-transferase domain
Nisin cyclase	2g02	4.8	4.3 Å among 99 aligned residues	4	Lanthionine synthetase C-like protein
Translocase of outer membrane 70 kDa subunit (Tom70)	2gw1	4.7	3.6 Å among 133 aligned residues	11	Tetratricopeptide repeat
26S proteasome regulatory subunit Rpn6	3txm	4.7	3.5 Å among 126 aligned residues	11	PCI domain
SAG101[291-537]					
Glutathione S-transferase	1ml6	7.6	3.3 Å among 102 aligned residues	8	Glutathione S-transferase domain (N- and C-terminal)
Syntaxin 6	1lvf	7	3.4 Å among 89 aligned residues	6	Syntaxin 6 domain
Mouse protein 25 alpha (MO25alpha)	2wtk	6.8	2.6 Å among 82 aligned residues	8	Mo25-like
Symmetrical sisters-1 (SYS-1)	3c2g	6.7	3.6 Å among 88 aligned residues	4	n/a
Target of Myb1 (Tom1)	1wrd	6.4	3.3 Å among 82 aligned residues	9	GAT domain

While the N-terminal domains of EDS1 and SAG101 are closely related to lipases (resulting in Dali Z-scores above 20), the list of homologues to the C-terminal domains consists of far less related proteins (Z-scores below 10)[25] that stem from all sorts of cellular functions and protein folds. Still, one common scheme was occurring: many of the homologous proteins were only

---

[25]As described before, Z-scores below 2 indicate no structural relationship, while highly related proteins produce Z-scores of 16 and above (source: `http://ekhidna.biocenter.helsinki.fi/how/week2/day8/stralignment.pdf`; request date: March, 2013)

functional in multi-protein complexes like the proteasome, the mitochondrial translocase complex or the TREX complex that acts in nuclear export of mRNA [Wu and Sha, 2006; Ellisdon *et al.*, 2012; Pathare *et al.*, 2012].

The EDS1 and SAG101 EP domain organisation and their structural relation to one of the more closely related homologues (Tom70 from the translocase complex) is depicted in Figure 3.23.
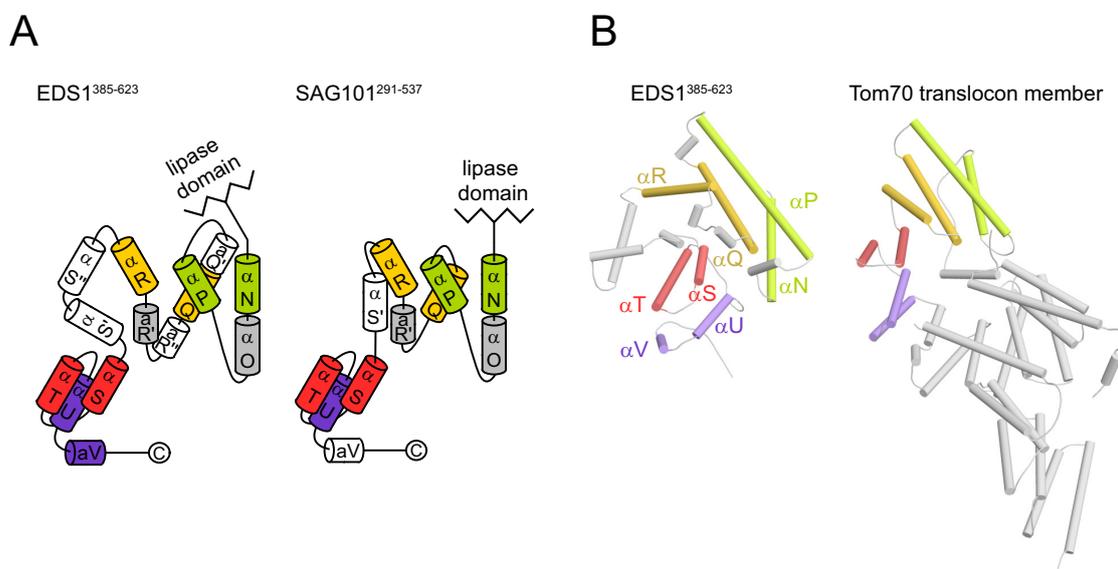

Figure 3.23 | EP domain organisation. (A) Topology diagrams of the EDS1 and SAG101 EP domains in which α-helices were named alphabetically by their order in the sequence following a nomenclature introduced in Figure 3.13. (B) highlights the structural relation between the EDS1 EP domain and one of its closest structural homologues, the transporter of outer membrane (TOM) protein Tom70. Helix tandems that belong to an tetratricopeptide repeat motif in Tom70 and their structural relatives in EDS1[385-623] and SAG101[291-537] are coloured accordingly in (A) and (B).

3.6. Interactions Within the EDS1 Protein Family

Protein-protein interactions of EDS1-related proteins have emerged as probably their central functional feature: not only does EDS1 interact with PAD4 and SAG101 but has also been reported to be in a physical complex with immune receptors [Bhattacharjee *et al.*, 2011; Bonardi and Dangl, 2012; Heidrich *et al.*, 2011] as introduced in section 1.3.5. Detailed molecular information on the formation of EDS1-containing complexes, however, is lacking. While I had already shown that full-length EDS1 and SAG101 do interact as recombinant proteins, I next asked which of the domains was able to bind the corresponding full-length partner. I expressed full-length variants of EDS1 and SAG101 as well as the EDS1 lipase-like domain (EDS1[1-384]) and the SAG101 EP domain (SAG101[291-537]) separately. *E. coli* cell suspensions containing different combinations of each one untagged and one His-tagged protein were then mixed, lysed and purified as described in sections 2.3.3 and 2.3.4. In these co-purification experiments, His-tagged SAG101 and an untagged variant of lipase-like EDS1[1-384] were able to interact as depicted in Figure 3.24. In contrast, the EP domain of SAG101 was not able to form a stable complex with full-length EDS1.
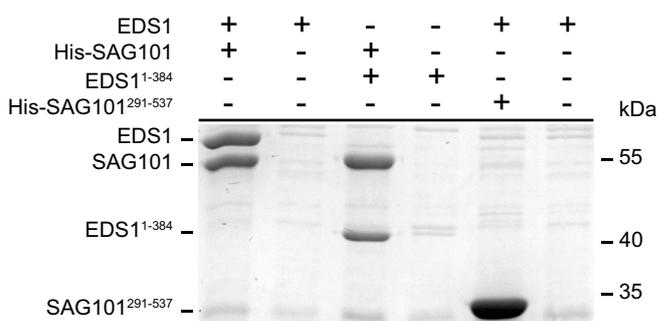
Figure 3.24 | Co-purifications of EDS1, SAG101 and two of their domains. For co-purifications, soluble cell lysates of His-tagged and an excess of untagged protein were mixed and applied to Ni-NTA matrix. After binding at 4 °C, the matrix was rigorously washed with at least 10 column volumes IMAC lysis buffer. His-tagged proteins were eluted by application of IMAC elution buffer and the elution fractions were separated by SDS-PAGE.

To further analyse the interactions within the EDS1 family, I made use of the yeast two-hybrid system introduced in section 2.6. Although it had been shown that EDS1 interacts with itself, SAG101 and PAD4 in such assays [Feys *et al.*, 2001, 2005; Rietz *et al.*, 2011], I had to re-establish the respective controls using Gateway-cloned genes in a GAL4-based yeast two-hybrid system (see sections 2.2.6 and 2.6 for further information). I was able to show the previously described interactions between combinations of full-length proteins (EDS1+EDS1, EDS1+SAG101 and EDS1+PAD4) and could additionally verify the interaction of N-terminal EDS1 with SAG101 3.25. As a result, I was able to further dissect the interactions between combinations of proteins.
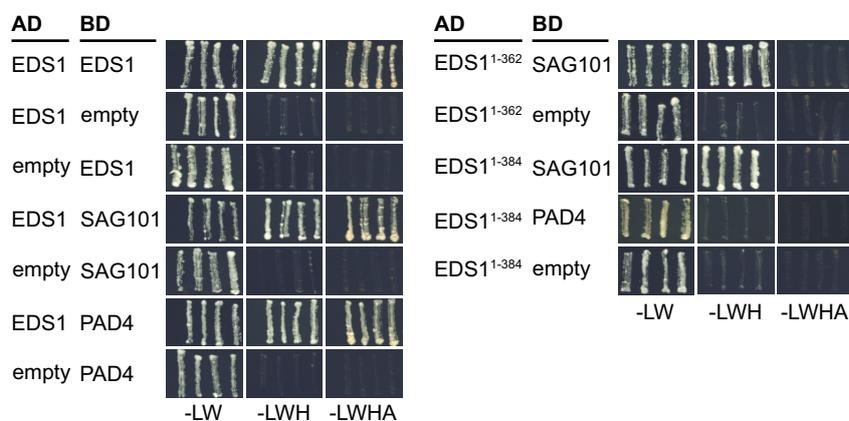
Figure 3.25 | Interactions within the EDS1 protein family in yeast two-hybrid assays. EDS1 or N-terminal domain variants (EDS1$^{1-362}$ and EDS1$^{1-384}$) were fused to the GAL4 activation domain (AD). In yeast two-hybrid assays, I analysed their interaction to EDS1, PAD4 and SAG101 fused to the GAL4 binding domain (BD). Successfully co-transformed cells that contain the AD- and BD-fusion proteins are able to grow on media lacking leucine and tryptophan (-LW). Cells only grow on -LWH and -LWHA plates (lacking histidine or additionally adenine) when the fused proteins do physically interact. As the GAL2-ADE2 reporter is more stringently regulated than GAL1-HIS3, cell growth on -LWH but not on -LWHA plates indicates a residual interaction.

3.6.1. Sequential Mutations in SAG101 Are Necessary to Suppress Interaction to EDS1

Manual selection of critical residues at the interface between EDS1[1-384] and SAG101 was assisted through the tools HotPoint[26] [Tuncbag *et al.*, 2010], HSPred[27] [Lise *et al.*, 2009] and KFC2[28] [Zhu and Mitchell, 2011; Darnell *et al.*, 2007] that predict so-called hot spots of protein interaction. Clackson and Wells [1995] had shown for a hormone-receptor that within a protein-protein interaction interface only a small and complementary set of hot spots residues contributes significantly to binding affinity. This was confirmed for larger sets of protein-protein interactions by Bogan and Thorn [1998] and others.

The hot spot of interaction between N-terminal EDS1 and SAG101, as predicted by the tools mentioned above, is made of a SAG101 hydrophobic pocket that accommodates EDS1 helix α**H** (Fig. 3.26 A and B). Within SAG101 and EDS1[1-384], single side chains at this interface were mutated to either alanine or the more polar serine. The effect of these exchanges on complex formation was evaluated in yeast two-hybrid assays using a GAL4 activation domain (AD) fusion of EDS1[1-384] and a binding domain (BD) fusion of SAG101. A comprehensive list of all tested variants is given in Appendix Table A.2.

Possibly due to an interaction that is weaker compared to the full-length protein complex, the interaction of EDS1[1-384] to SAG101 could only be shown on selective media lacking histidine (-LWH) but not additionally adenine (-LWHA) (Fig. 3.25). I therefore documented growth of yeast cells after two and five days incubation at 30 °C to qualitatively compare effects on interaction. Cells that were impaired in growth indicated weakened interaction between the tested variants and single mutations were sequentially combined to achieve an additive effect. As depicted in Fig. 3.26 C, only combinatorial exchanges within SAG101 were sufficient to significantly reduce the interaction to wildtype EDS1[1-384] whereas even single mutations in EDS1[1-384] lead to a significant loss-of-interaction to SAG101 (Fig. 3.26 D. and Appendix Table A.3).

Yet, when the resulting variants SAG101 **L**12A+**L**21A+**I**141A (**LLI**) and EDS1 **L**258A+**L**262A (**LL**) were probed against the respective binding partner in full length, interaction was at least partially restored, arguing for a contribution of the C-terminal interface to complex formation.

The C-terminal interface between EDS1 and SAG101 EP domains is a mixture of hydrophobic interactions, hydrogen bonding and salt bridges that are predominantly formed between two helices, EDS1 α**P** and SAG101 α**N** (Fig. 3.27 A and B). Several mutations within C-terminal SAG101 were tested in yeast two-hybrid experiments and an exchange of **Y**306 to alanine within the SAG101**LLI** variant reduced binding to full-length EDS1 (Fig. 3.27 C). The variant SAG101**LLIY** is severely compromised in its interaction to EDS1, yet a residual interaction is indicated by growth of the yeast cells on -LWH medium.

Contrary to SAG101, the combination of single exchanges within the N-terminal EDS1 interaction interface (resulting in a EDS1 **I**254A+**L**258A+**F**261A+**L**262A (**LLIF**) variant) turned out to be sufficient to achieve a similar effect (Fig. 3.27 D) and further mutations within the C-terminal interface did not reduce residual interaction that is indicated by the growth of cells

---

[26]http://prism.ccbb.ku.edu.tr/hotpoint/ (request date: April, 2011)
[27]http://bioinf.cs.ucl.ac.uk/hspred (request date: April, 2011)
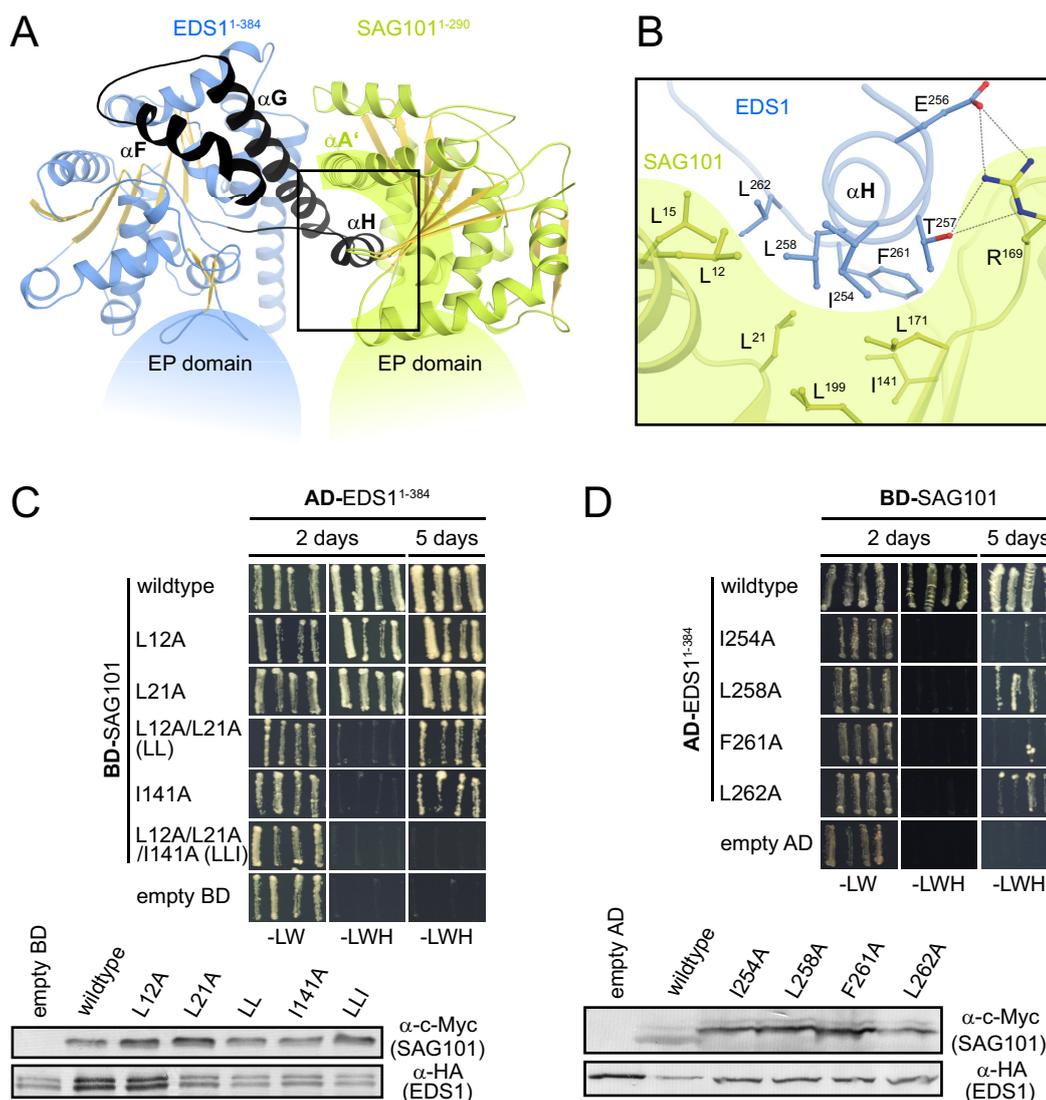[28]http://kfc.mitchell-lab.org/ (request date: April, 2011)

Figure 3.26 | **Interactions between N-terminal EDS1 and SAG101 in yeast two-hybrid assays.** (A) shows the anchoring point of the N-terminal interaction interface that consists of hydrophobic EDS1 helix αH that is accommodated by a schematically represented hydrophobic pocket within SAG101. EDS1 αH belongs to the N-terminal EDS1 insertion that is coloured black. Structural details of this interaction site is shown in (B) in which side chains that were exchanged according to the text are represented as sticks. (C) The effect of single and combinatorial exchanges within the hydrophobic pocket of SAG101 as analysed after 2 and 5 days of yeast cell growth on selective -LWH plates. Protein accumulation of AD- and BD-fusion proteins as recognised by α-HA and α-c-Myc antibodies, respectively, is shown below. (D) Accumulated protein amounts of EDS1[1-384] variants and their effect on SAG101 binding is shown according to (C).

on -LWH plates. The more comprehensive list of analysed variants is given in Appendix Tables A.4 and A.2.

To validate these effects *in planta*, J. Stuttmann (MPIPZ, Köln) infected tobacco (*Nicotiana benthamiana*) plants using the tobacco mosaic virus (TMV)-based magnICON system (Icon Genetics) that allows transient expression of the introduced genes. In co-purification experiments with StrepII-tagged SAG101 and untagged EDS1 variants that are schematically explained in Figure 3.28 A, he could verify that SAG101**LLIY** is indeed compromised in EDS1 interaction but also in protein stability (Fig. 3.28 B). In contrast, the stability of EDS1 variant **LLIF** is not
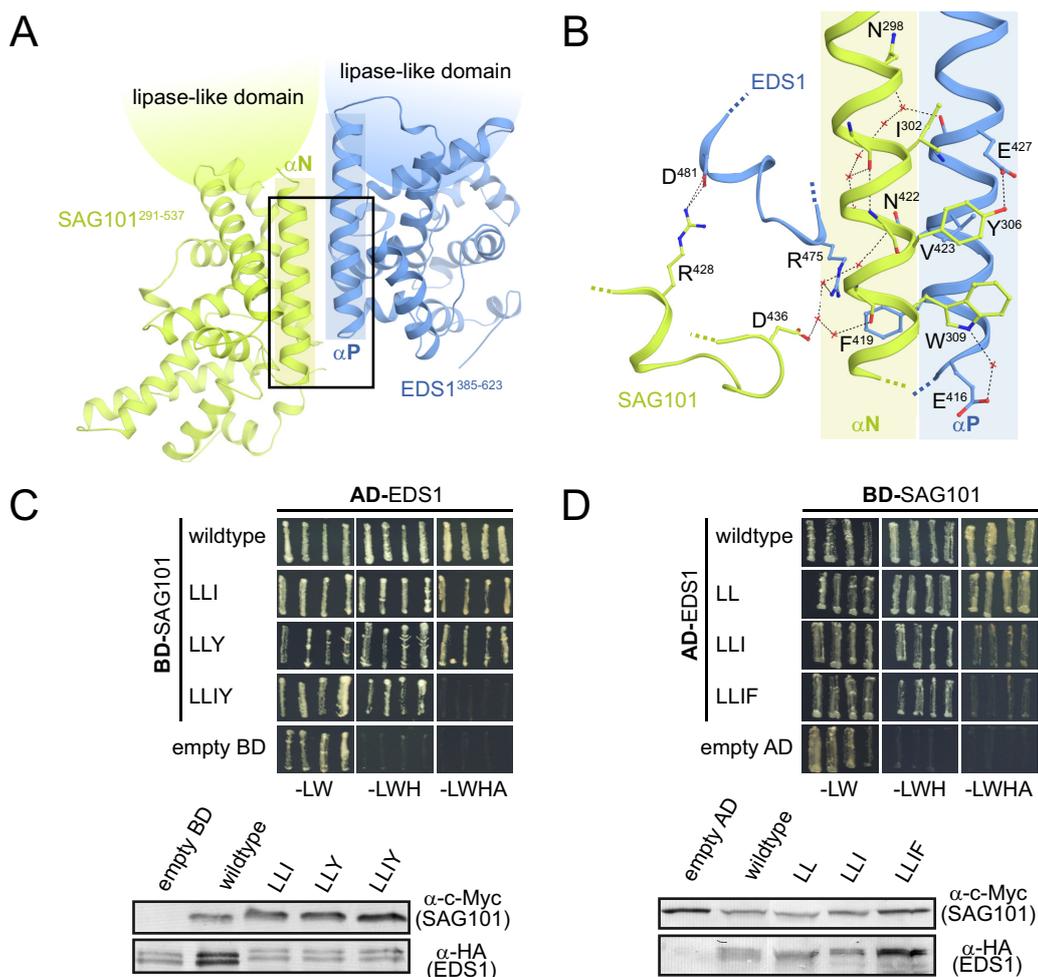
Figure 3.27 | Interactions between full-length EDS1 and SAG101 in yeast two-hybrid assays. (A) shows the interaction interface between the EDS1 and SAG101 EP domains. Two parallel helices, EDS1 αP and SAG101 αN, are highlighted that form most of the interactions. These consist of hydrophobic interactions, salt bridges and a network of hydrogens bonds as depicted in more detail in (B). (C) Yeast two-hybrid interactions between activation domain (AD) fusions of full-length EDS1 and binding domain (BD) fusions of different SAG101 variants (**LL**= **L**12A+**L**21A, **LLI**= **LL**+**I**141A, **LLIY**= **LLI**+**Y**306A). Pairs of plasmids were co-transformed into yeast cells, grown on plates selecting for co-transformation (-LW) and four single colonies were streaked onto plates selecting for weak (-LWH) and strong (-LWHA) protein interaction. Cell growth was monitored after 2 days at 30 °C. Accumulation of the protein fusions in yeast cells was assayed by total protein extraction, SDS-PAGE and subsequent immunoblot analysis using α-HA and α-c-Myc antibodies that recognise GAL4 activation- or binding domain, respectively. (D) Yeast two-hybrid interactions between activation domain (AD) fusions of different EDS1 variants (**LL**= **L**258A+**L**262A, **LLI**= **LL**+**I**254A, **LLIF**= **LLI**+**F**261A) and a binding domain (BD) fusion of full-length SAG101 that were analysed as described for (C).

affected and the exchanges result in a drastic reduction of interaction to SAG101 in plants (Fig. 3.28 C).
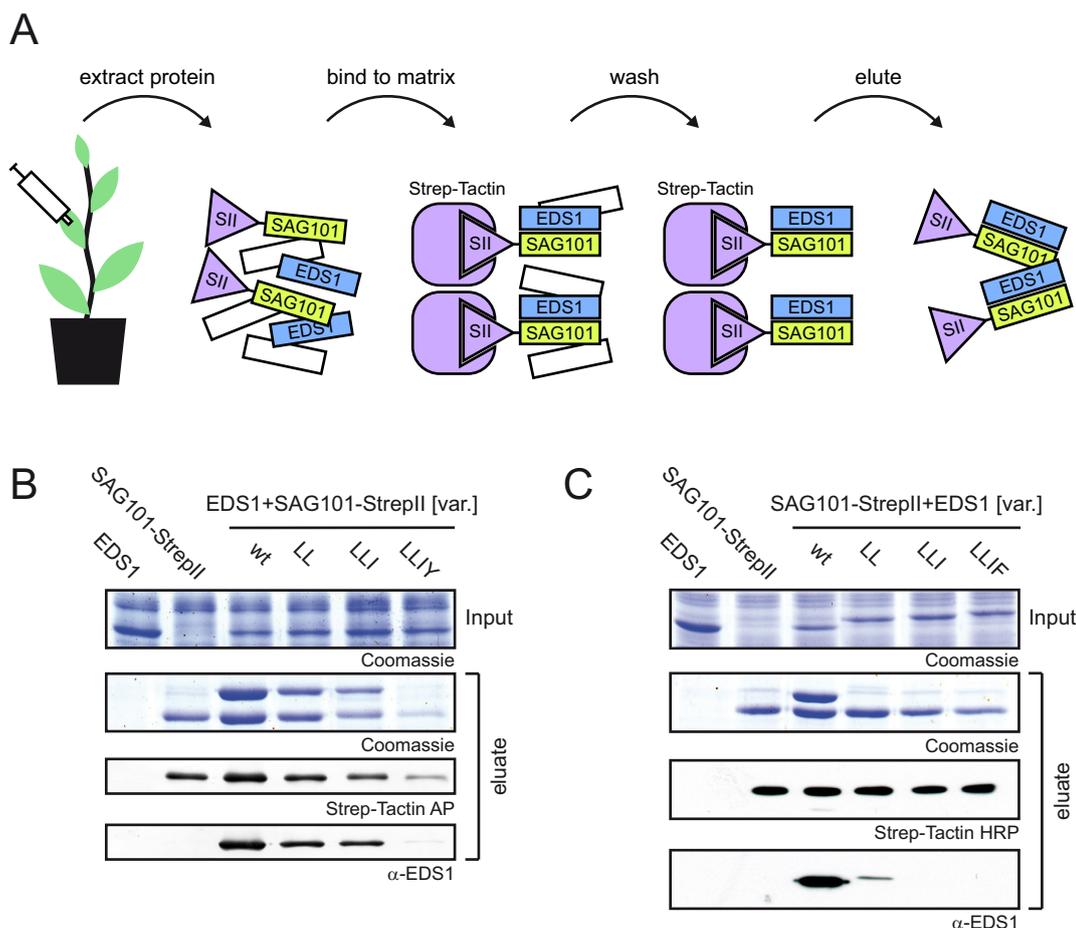
Figure 3.28 | Interactions between full-length EDS1 and SAG101 in tobacco. (A) Schematic representation of co-purification experiments from tobacco. According to section 2.3.6, total protein extracts of 2 g plant material were purified through the addition to Strep-Tactin matrix. Following rigorous washing, StrepII-tagged and co-purified proteins were eluted and separated by SDS-PAGE. (B) Co-purification of untagged EDS1 in wildtype form and variants of SAG101 fused to StrepII. Samples separated by SDS-PAGE were analysed by immunoblotting using Strep-Tactin conjugated with alkaline phosphatase (AP) or antibodies against EDS1 (α-EDS1). (C) Co-purification of untagged EDS1 variants (var.) and a StrepII-fusion of wildtype SAG101.

3.6.2. An EDS1 Variant Implies Similarities Between EDS1/SAG101 and EDS1/PAD4 Complex Formation

We analysed the same EDS1**LLIF** variant in yeast and tobacco for interaction to the well-characterised EDS1 binding partner PAD4 and surprisingly found that this interaction is similarly affected by the introduced mutations (Fig. 3.29 A and B). It was therefore tempting to speculate that SAG101 and PAD4 bind EDS1 in a similar manner. This finding and the notion that SAG101 and PAD4 are sequentially more closely related to each other than to EDS1 (Fig. 3.29 C) motivated us to build a homology model based on structural information I had obtained through EDS1/SAG101 complex crystals.
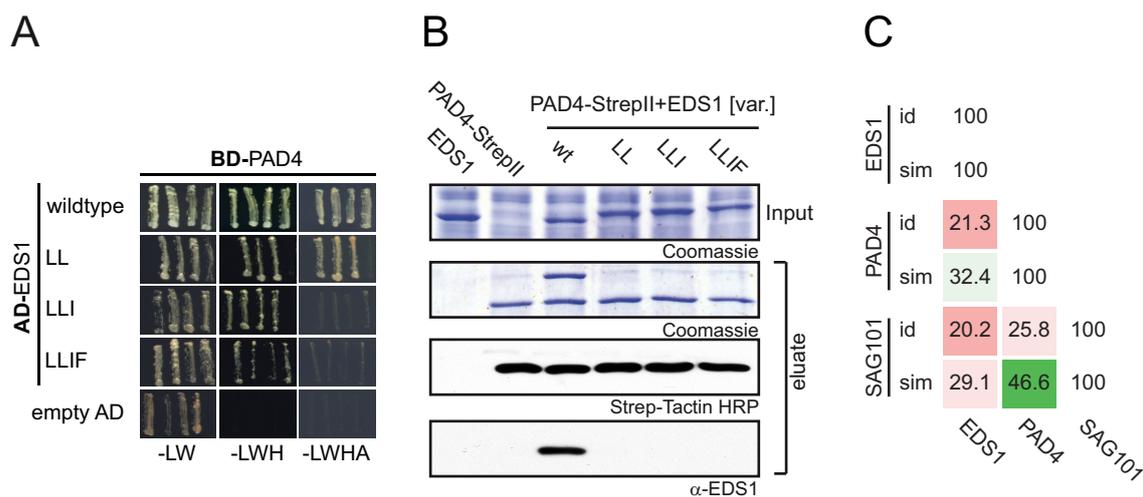
Figure 3.29 | Basis for a PAD4 homology model. (A) Yeast two-hybrid interactions between activation domain (AD) fusions of EDS1 variants (see text) and a binding domain (BD) fusion of full-length PAD4. Interaction was assayed as described for Figure 3.27 C. (B) Co-purification of untagged EDS1 variants shown in (A) and a StrepII-fusion of wildtype PAD4 that were transiently expressed in *Nicotiana benthamiana*. Input and elution fractions were separated by SDS-PAGE and analysed by immunoblotting using Strep-Tactin conjugated with horseradish peroxidase (HRP) or antibodies against EDS1 (α-EDS1). See Figure 3.28 for experimental details. (C) Sequence identities (id) and similarities (sim) between proteins of the EDS1 family as calculated by SIAS (`http://imed.med.ucm.es/Tools/sias.html`; request date: August, 2012).

Raphaël Guerois (CEA, Gif-sur-Yvette) used the EDS1/SAG101 crystal structure as a template and build a homology model of PAD4 in place of SAG101 using the Rosetta modeling suite [Kaufmann *et al.*, 2010]. Not surprisingly, the resulting complex model looks almost like the crystallised EDS1/SAG101 complex (Fig. 3.30 A). To validate this model, we analysed the putative interaction interface of EDS1 and PAD4 in yeast and tobacco. Interestingly, the predicted N-terminal interface between both proteins reveals that PAD4 could (similar to SAG101) possess a hydrophobic pocket for the accommodation of EDS1 helix αH (Fig. 3.30 B).

I chose critical residues within this pocket and J. Stuttmann (MPIPZ, Köln) mutated them into alanines. As EDS1[1-384] and PAD4 did not form a detectable protein complex in yeast two-hybrid assays (see Fig. 3.25), I directly analysed these variants using full-length EDS1 and PAD4. I found that three simultaneous exchanges within the hydrophobic pocket of PAD4 (PAD4**M**16A+**L**21A+**F**143A) are sufficient to significantly reduce the interaction to EDS1 (Fig. 3.30 C). This was verified by co-purification of transiently expressed EDS1 and PAD4 variants from tobacco plants (Fig. 3.30 D).

This does not only validate the general correctness of our model but additionally implies that EDS1 indeed binds SAG101 and PAD4 in a similar manner. All variants that were analysed in yeast two-hybrid assays are listed in Appendix Table A.5.
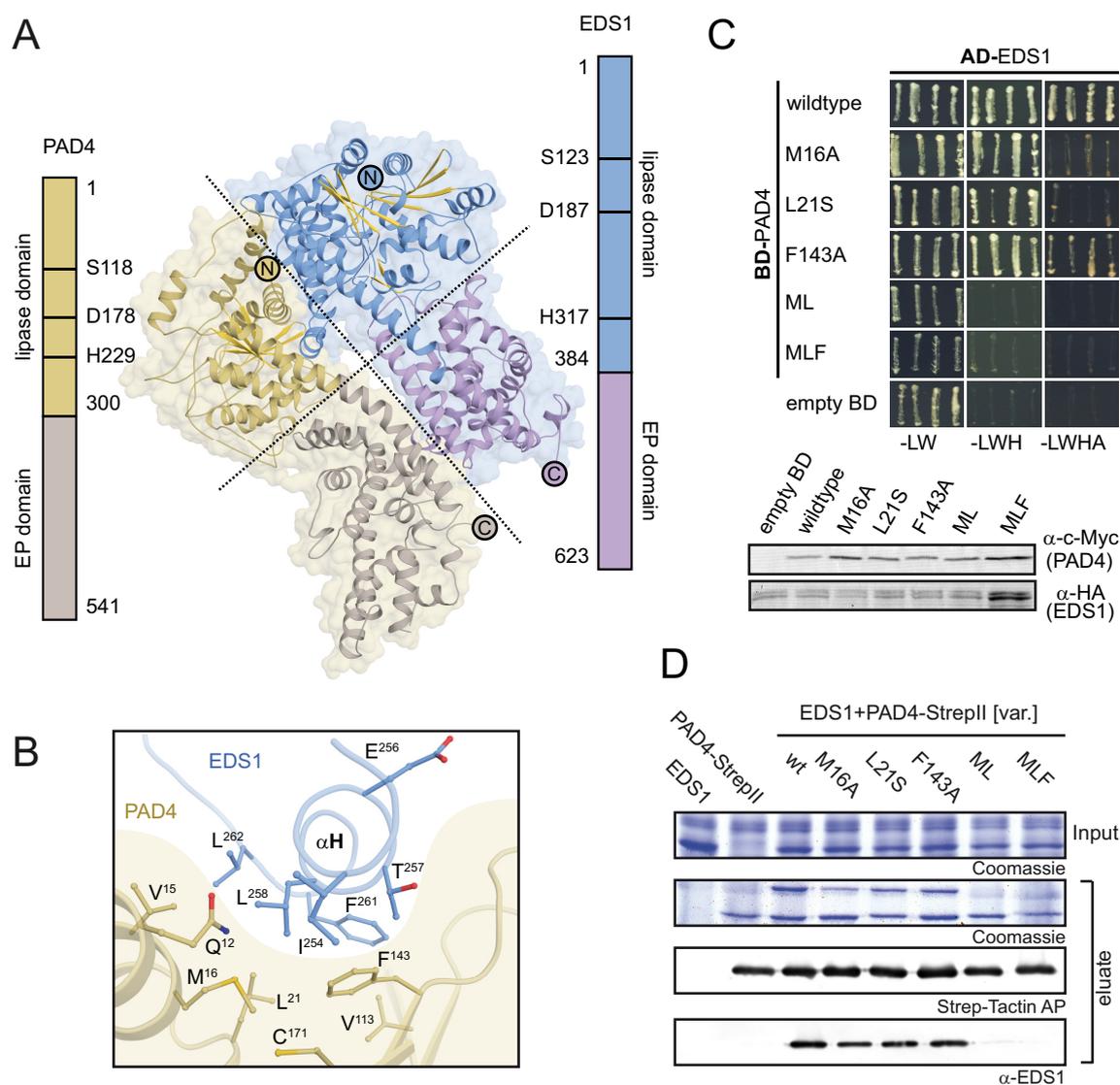
Figure 3.30 | Validation of the PAD4 homology model. (A) Structural overview of the EDS1/PAD4 heterodimer homology model. (B) Close-up view on the putative interaction site within the N-terminal interface of EDS1 and PAD4. The PAD4 hydrophobic pocket is schematically indicated. (C) Verification of the EDS1/PAD4 model by yeast two-hybrid assays. We designed variants of PAD4 that carry mutations within the hydrophobic pocket shown in (B) and tested these in yeast two-hybrid assays as described before. (D) Co-purification of untagged EDS1 and StrepII-fusions of PAD4 variants (var.) shown in (C) and transiently expressed in *Nicotiana benthamiana* as described for Figure 3.27.

3.7. EDS1 Interaction With Itself

I further tested whether the EDS1**LLIF** variant was still able to interact with itself in yeast and found that this was the case. This was to be expected as the monomers within homooligomers are generally symmetry-related to each other [Goodsell and Olson, 2000], a criterion that would not be fulfilled if EDS1 interacted with itself in a SAG101-like manner. As additional information on EDS1 homooligomerisation is contradictory – Feys *et al.* [2001] found in yeast two-hybrid assays that the EDS1 EP domain is sufficient for self-interaction while Zhu *et al.* [2011] were only able to see an interaction between EDS1 and the EDS1 lipase-like domain *in planta* – I can only speculate on how EDS1 homomers are formed. Although Feys *et al.* [2005] showed by

analytical size-exclusion chromatography of plant extracts that EDS1 monomers exist alongside homodimers, EDS1 is generally assumed to be homodimeric.

In contrast, full-length EDS1 produced in *E. coli* seemed to be predominantly monomeric as shown by size-exclusion chromatography (Fig. 3.31 A, B, C). I was further not able to produce a stable recombinant EDS1 homomer by chemical crosslinking with BS3 that irreversibly connects the primary amines of lysine residues or the N-termini of polypeptide chains (Fig. 3.31 D). Untagged EDS1 could further not be efficiently co-purified with a Trx-His-EDS1 fusion or a His-tagged variant of N-terminal EDS1[1-384] (Fig. 3.31 E). Surprisingly, EDS1[1-384] formed a minor dimeric fraction according to gelfiltration experiments depicted in Fig. 3.3 B.

While a prokaryotic protein expression system (e.g. *E. coli*) comes along with many benefits, including ease of cultivation and genetic manipulation, there is a major drawback: posttranslational modifications found naturally on many eukaryotic proteins are generally not applied to the recombinant protein. I wondered whether EDS1 carried such modifications that could be necessary for homodimerisation and tried to answer this question by purifying StrepII-tagged EDS1 from *Arabidopsis thaliana* as described in section 2.3.6. The purified protein solution was separated by SDS-PAGE (Fig. 3.32) and the band of EDS1 protein was cut from the gel and handed to Tobias Lamkemeyer (CECAD) for mass spectrometry analysis. Apart from a putative acetylation site on the EDS1 N-terminus that was already observed by Gobbato [2007], I did not find convincing hints pointing towards further modifications of EDS1 *in planta*.

3.7.1. Protein Crystals of EDS1 Might Reveal its Molecular State

In addition to the crystallisation and structure solution of the EDS1/SAG101 complex, I was able to produce protein crystals of His-tagged EDS1 without an additional binding partner. The protein was purified by IMAC as described in section 3.3 and was afterwards gelfiltrated on a HiLoad 16/60 Superdex 75 column (GE Healthcare) in 10 mM Bicine, 50 mM NaCl, 0.5 mM EDTA, pH 8. The peak fractions were pooled and concentrated while exchanging the protein buffer to 10 mM Bicine, 5 mM TCEP, pH 8. EDS1 at concentrations of 5, 10 and 15 mg/ml was used in sparse-matrix screening, mixing protein and crystallant in a 1:1 ratio. Initial crystals grew in the presence of 100 mM Tris (pH 8), 250 mM sodium tartrate and 10 % (w/v) PEG8000 as a crystallant (JBScreen Kinase screen, Jena Bioscience) and only at 4 °C but not at 20 °C (Fig. 3.33 A). Crystals grown in the presence of only 8 % (w/v) PEG8000 were considerably larger and used for collection of diffraction data (Fig. 3.33 B). They were cryoprotected by soaking them for 30 to 60 seconds in 100 mM Tris (pH 8), 250 mM sodium tartrate, 25 % (w/v) PEG8000 and 7.5 % (v/v) PEG400 and frozen in liquid nitrogen. While the crystals hardly showed X-ray diffraction on a home source, diffraction data to a maximum resolution of approximately 9 Å (Fig. 3.33 C) were collected at a synchrotron beam (Swiss Light Source, Villingen). Notably, these crystals degraded over time (Fig. 3.33 D), possibly due to the oligomerisation (aggregation?) of recombinant EDS1 over time that I described in Figure 3.31 C and that could deplete the crystal in protein.

Preliminary analysis according to Matthews [1968] suggested that there were multiple EDS1 molecules (at least two) within each asymmetric unit. However, data to this resolution are only
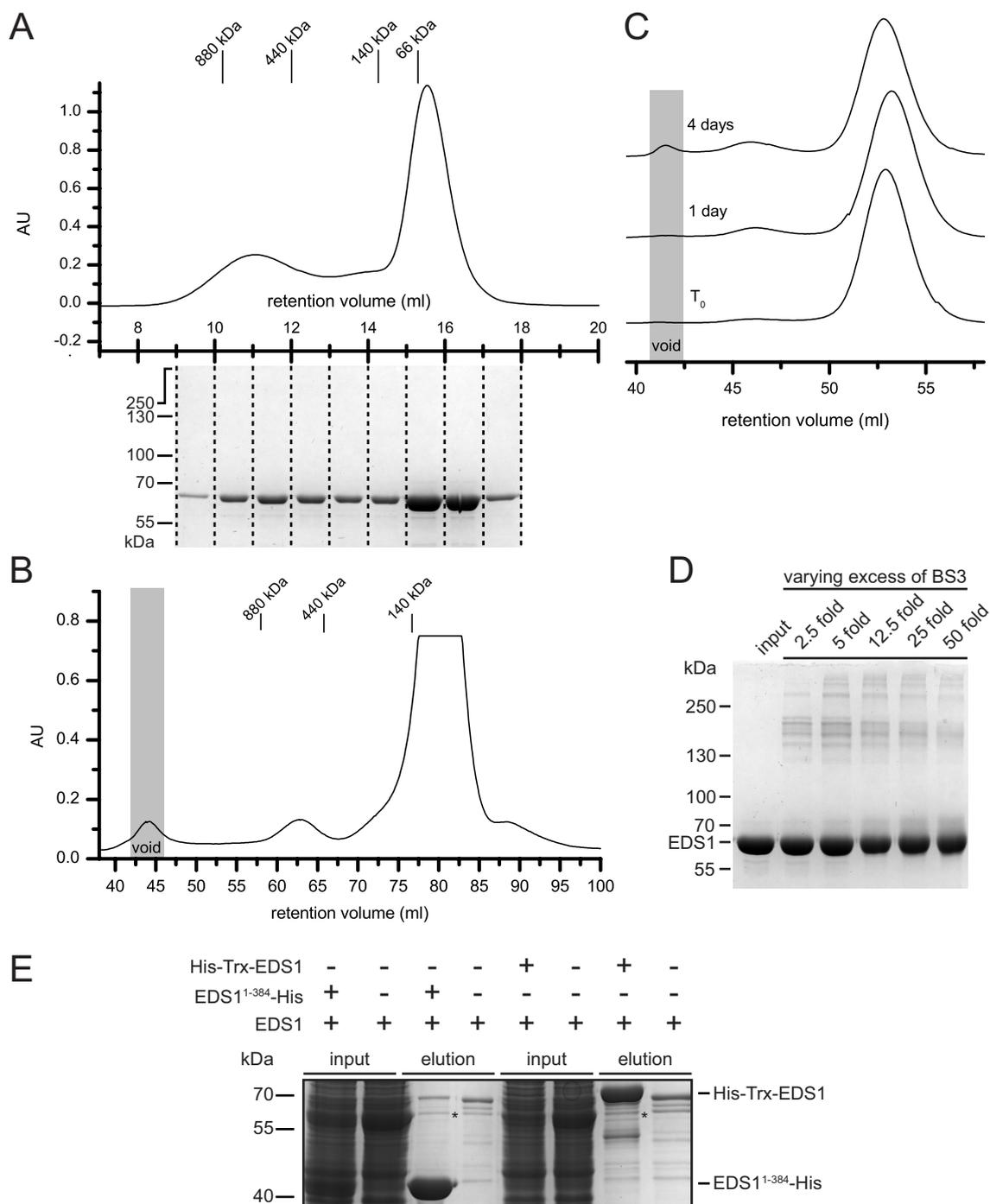
Figure 3.31 | EDS1 self-interaction analysis. (A) IMAC-purified EDS1 was analysed on a 10/300 GL Superdex 200 column (GE Healthcare) after letting it stand at 4 °C for 4 days. The elution profile and subsequent SDS-PAGE analysis of indicated fractions revealed that EDS1 exists predominantly as a monomer but apparently also as a higher order molecular complex. The retention volumes of ferritin (880 kDa, 440 kDa), lactate dehydrogenase (140 kDa) and bovine serum albumine (66 kDa) are indicated as molecular weight markers. (B) A similar distribution was found on a preparative HiLoad 16/60 Superdex 200 column (GE Healthcare). (C) This oligomerisation of EDS1 happens over time: while freshly gelfiltrated EDS1 exists almost exclusively as a monomer, a small fraction forms higher order molecular complexes after 1 or 4 days at 4 °C. (D) Chemical crosslinking with bissulfosuccinimidyl suberate (BS3) was performed for 2 h on ice using EDS1-His at a concentration of 1 mg/ml and different excesses of BS3 that potentially forms SDS-resistant amide bonds between proteins and allows analysis through SDS-PAGE. (E) For co-purifications, soluble *E. coli* lysates of His-tagged Trx-His-EDS1 and EDS1$^{1\text{-}384}$-His were mixed with an excess of untagged EDS1 and applied onto a Ni-NTA matrix. After binding at 4 °C, the matrix was rigorously washed with at least 10 column volumes IMAC lysis buffer and proteins were eluted by the application of IMAC elution buffer. Eluates were analysed by SDS-PAGE.
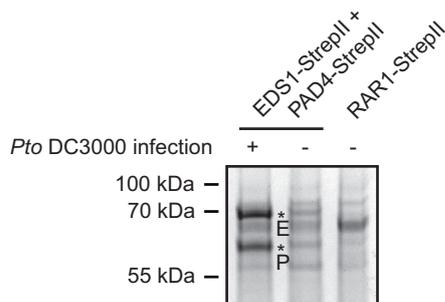
Figure 3.32 | EDS1 and PAD4 purified from Arabidopsis. EDS1 and PAD4 fused to an StrepII affinity tag were co-expressed under a constitutive 35S promoter in *Arabidopsis* and purified from leaves as described in section 2.3.6. Both proteins were constitutively under a 35S promoter and the plants were either left unchallenged or were infected with the bacterium *Pseudomonas syringae* pv. tomato (*Pst*) DC3000. As a control, an unrelated protein (RAR1) fused to a StrepII-tag was expressed accordingly. The expected position of EDS1 (∗E) and PAD4 (∗P) bands are indicated.
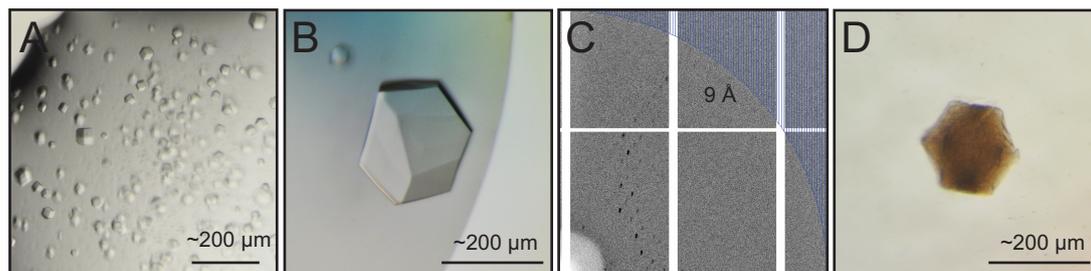
Figure 3.33 | Crystals of EDS1. (A) Crystals from the initial screen that grew within 6 days at 4 °C in 100 mM Tris pH 8.0, 250 mM sodium tartrate and 10 % (w/v) PEG8000. (B) Optimised crystals that appeared in the presence of only 8 % (w/v) PEG8000 leaving all other parameters unchanged. (C) A single diffraction image of one of the optimised crystals taken at a synchrotron beam (Swiss Light Source, Villingen). The maximum resolution of 9 Å is indicated. (D) Degraded crystals, approximately two weeks after first appearance.

hard to interpret and the crystals were predicted to be either twinned or characterised by a rather high crystallographic symmetry (a cubic spacegroup, probably P432).

As optimisation of diffraction by means of crystal annealing [Stevenson *et al.*, 2001], dehydration through transfer in dehydrating solutions [Heras and Martin, 2005], soaking in glutaraldehyde [Heras and Martin, 2005; Wine *et al.*, 2007] and micro- and streak-seeding [Bergfors, 2003] was unsuccessful, these crystals remain an outlook that might open the way to the structure of a further functional EDS1 state. Or, to put it in the words of Newman [2005], "There are a number of manipulations that are available to the experimentalist with poorly diffracting crystals. However, there are no systematic studies that have mapped particular problems to particular solutions and certainly no guarantees are issued with any of these methods."

Final Remark

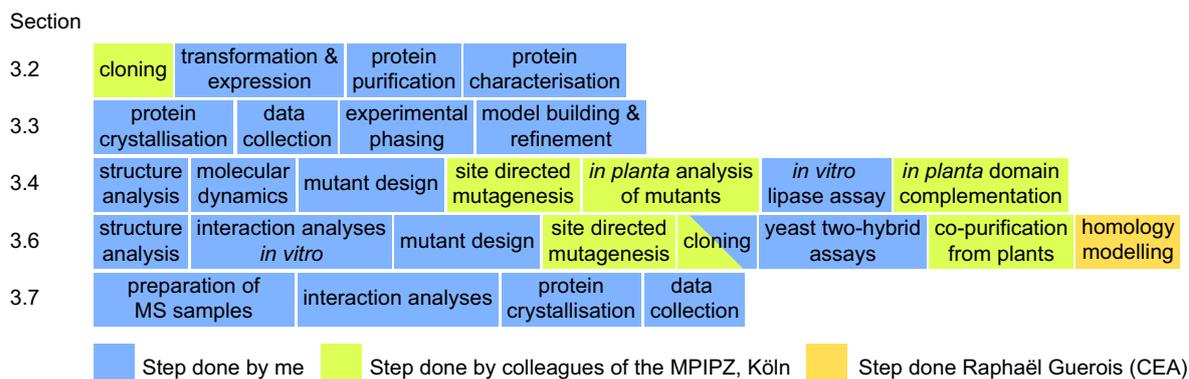The experimental contributions within each section are graphically summarised in Figure 3.34.

Figure 3.34 | Experimental contributions. This figure shall illustrate which experimental steps were performed by myself (blue) or by colleagues (mostly J. Stuttmann) at the cooperating Max-Planck-Institute for Plant Breeding Research (MPIPZ). Raphaël Guerois (CEA, Gif-sur-Yvette) contributed the PAD4 homolgy model.

4. Discussion

"The least initial deviation from the truth is multiplied later a thousandfold."

— Aristotle

As sessile organisms, plants must flexibly adapt to environmental stresses. To defend themselves against invading pathogens, they have evolved a multi-layered immune system that is brilliantly described by Jones and Dangl [2006]. While the first inducible layer recognises pathogens outside the plant cell through pathogen-associated molecular patterns (PAMPs) and triggers an immune response that is often sufficient to keep the plant healthy, successful pathogens have evolved effector molecules that suppress these defence mechanisms. Effector molecules are, in turn, recognised by intracellular plant immune receptors (the so-called R proteins) that activate a strong defencive response that is usually associated with localised cell death at the infection site and thus has to be tightly controlled.

R proteins of the TIR-NB-LRR type need the nucleo-cytoplasmatic plant protein EDS1 for successful signalling [Aarts *et al.*, 1998; Parker *et al.*, 1996]. EDS1 interacts with the sequence-related proteins PAD4 and SAG101 and they collectively form the EDS1 protein family that is defined by an N-terminal lipase-like domain and a C-terminal EP (EDS1-PAD4 defined) domain that has no obvious homologues at sequence level [Feys *et al.*, 2005]. The plant-specific EDS1 family has been shown to integrate biotic and abiotic stress signals as introduced in section 1.3.3. The molecular mechanism by which EDS1-requiring complexes transduce these signals are, however, poorly understood. In this thesis, I sought to elucidate the structure of EDS1 and its sequence-related binding partners PAD4 and SAG101 to provide a basis for their mechanistic description.

4.1. Defining Open Questions

By the beginning of my experimental work, virtually nothing was known about the molecular details that underlie EDS1 family protein functions. While it had been clear that N-terminal EDS1, PAD4 and, to a lesser extend, SAG101 were related to eukaryotic lipases at sequence level [Jirage *et al.*, 1999; Falk *et al.*, 1999; Feys *et al.*, 2005] (Fig. 1.6), hydrolytic activity had only been reported for SAG101 in an immunity-unrelated context [He, 2002]. The structural degree of lipase-relation was thus unclear. Further, it was unknown whether the C-terminal EP domain of all three proteins was not only sequentially but also structurally unique and to which extend both domains participated in the well-characterised interactions within the EDS1 protein family [Feys *et al.*, 2005; Rietz *et al.*, 2011].

Representing central components of plant pathogen resistance, the EDS1 protein family raised the interest of scientists worldwide, some of whom tried to create homology models of EDS1 and PAD4 utilising their relation to lipases to answer these questions [Singh and Shah, 2012].

The Limitations of Structure Prediction   Thanks to the growing knowledge of protein structures and sophisticated prediction methods, freely accessible web-tools have become popular

to predict the three-dimensional structures of proteins. Figure 4.1 shall illustrate that such predictions are not trivial, especially for multi-domain proteins.

Both the Phyre[2] [Kelley and Sternberg, 2009] and the I-TASSER [Zhang, 2008] web-tools are being used to predict several ten thousand structures per year[29]. Before solving the crystal structure of EDS1, I used these tools and will briefly compare the results to the experimentally determined structure. For prediction, I chose the primary sequences of either full-length EDS1 or that of EDS1[1-384] (encompassing the lipase-like domain) and EDS1[385-623] (EP domain) separately. The predicted models were superimposed on EDS1 extracted from the EDS1/SAG101 complex structure (Fig. 4.1).

Especially when the sequence of full-length EDS1 was used as query, the predicted structures differed significantly from crystallised EDS1 (upper panel of Fig. 4.1). Although separate predictions of the lipase-like and EP domain produced more reliable results (lower panel), the orientation of the domains towards each other and the position of functional features – for instance α-helix α**H** that is essential for SAG101 binding – would not have been answered by these models.
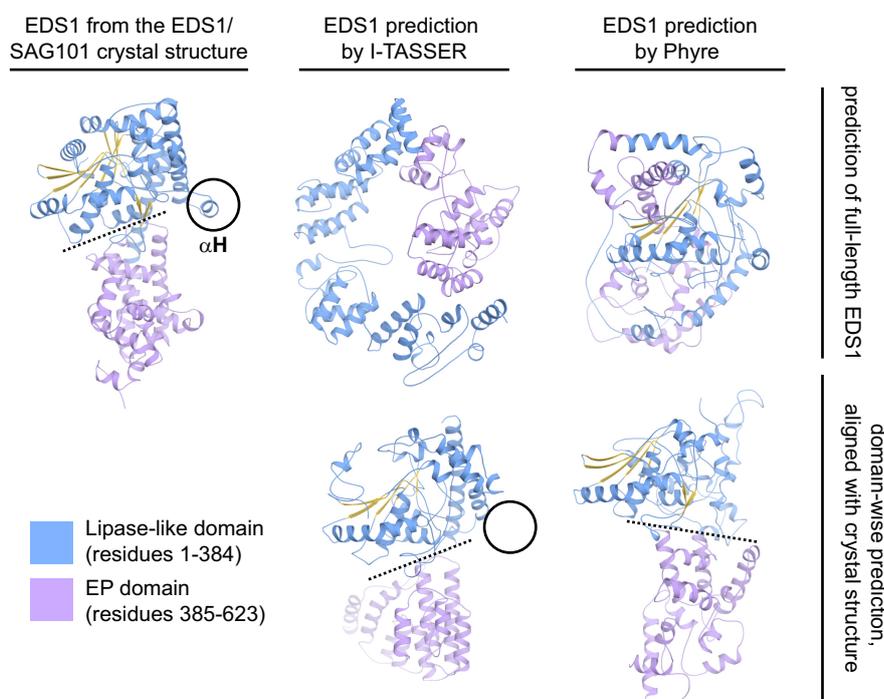
Figure 4.1 | EDS1 structure predictions. Primary sequences of EDS1 in full length or EDS1[1-384] and EDS1[385-623] (representing the lipase-like and EP domain, respectively) were used as sequence inputs for structure prediction by the Phyre[2] [Kelley and Sternberg, 2009] and the I-TASSER [Zhang, 2008] web-tools. The resulting models were superimposed on the experimentally determined EDS1 structure (extracted from the EDS1/SAG101 complex) shown on the upper left. The position of EDS1 helix α**H** is highlighted within the crystal structure to illustrate that even the most precise model (domain-wise prediction by I-TASSER) reveals significant differences to crystallised EDS1.

Structures that possess yet unknown folding aspects thus remain challenging to predict. When novel structures are solved experimentally and deposited in the freely accessible Protein Data

---

[29]see http://www.sbg.bio.ic.ac.uk/phyre2 (request date: February, 2013) and http://zhanglab.ccmb.med.umich.edu/I-TASSER/ (request date: February, 2013)

Bank (PDB) [Berman, 2000] they, in turn, help to improve the accuracy of structure prediction tools that rely on the availability of structural information. As will be discussed later, we used the novel EDS1/SAG101 structure to predict a homology model of EDS1 in complex with PAD4.

In the following pages, I will present functional implications of the obtained structures, beginning with a widely discussed aspect of EDS1 function, that of lipase homology.

4.2. Possible Roles of the Lipase-Like Domains

As expected from sequence analysis (section 3.1), I could show that N-terminal EDS1 and SAG101 are indeed closely related to lipases at the structural level (Table 3.6). Lipases belong to the α/β hydrolase fold family that was defined after Ollis *et al.* [1992] had recognised that enzymes of the same phylogenetic origin but diverse cellular functions share a closely related topology: a central arrangement of β-sheets that is surrounded by α-helices (Fig. 3.15). Together, they provide a skeleton for the correct position of critical catalytic residues – often a triad that I described in section 3.1.1.

While Ollis *et al.* [1992] introduced the canonical fold on the basis of five different enzymes, the family rapidly expanded [Nardini and Dijkstra, 1999; Carr and Ollis, 2009] and the ESTHER (ESTerases and α/β Hydrolase Enzymes and Relatives) database contains more than 30000 manually curated proteins as of today [Lenfant *et al.*, 2013]. Although most family members act as enzymes, numerous inactive α/β "hydrolases" have been characterised in recent years [Lenfant *et al.*, 2013].

I will introduce and discuss three models that could explain how the lipase-like domains of EDS1, SAG101 and PAD4 contribute to their biological function. Basically, I discriminate between the following modes of action: (1) hydrolytic activity in the synthesis or degradation of a small signalling molecule, (2) binding of a small molecule that is not hydrolysed but rather triggers a conformational change that is essential for downstream signalling and (3) a role that is independent of a small molecule and instead involves interactions between proteins.

4.2.1. The Lipase-Like Domains and Small Molecules: Implications From Plant Hormone-Binding α/β Hydrolases

The first two conceivable modes involve small molecules and are depicted in Figure 4.2 A and B. In the following paragraphs, I will present two proteins that act in either the hydrolytic synthesis or in non-hydrolytic binding of a plant hormone (Fig. 4.2 C and D) and highlight analogies and dissimilarities between them and the lipase-like domains of EDS1, SAG101 and PAD4.

The SABP2 Model   The 29 kDa protein SALICYLIC ACID BINDING PROTEIN2 (SABP2) was found in tobacco plants [Du and Klessig, 1997] and identified as an α/β hydrolase on sequence basis. Initially proposed as a receptor for salicylic acid, Kumar and Klessig [2003] could show that SABP2 hydrolyses artificial para-nitrophenyl (pNP) substrates using a serine-aspartate-histidine (S-D-H) triad (Fig. 4.3 A).

Using a crystal structure, Forouhar *et al.* [2005] finally revealed that SABP2 catalyses the reaction of methyl salicylate to salicylic acid, a plant hormone that I introduced as a central
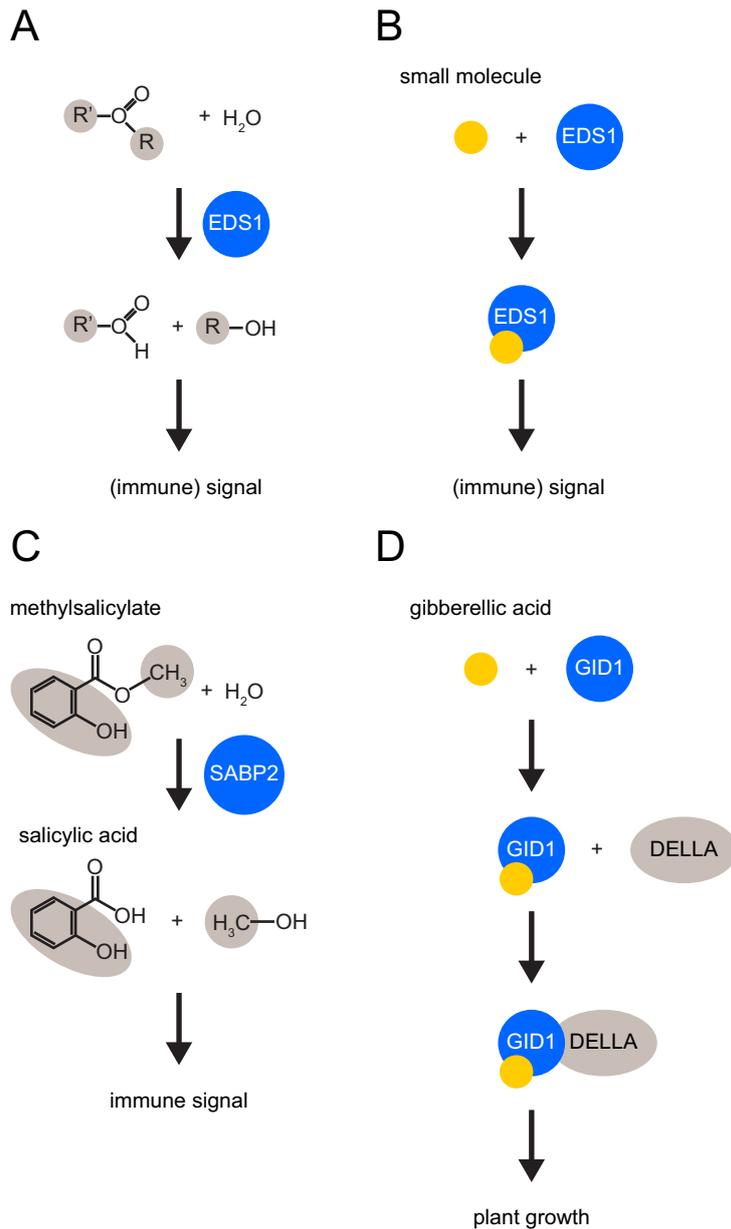
Figure 4.2 | Possible roles of the EDS1 lipase-like domain. (A) Hydrolysis of a small molecule is one possible mechanism through which the lipase-like domain of EDS1 (and PAD4) could participate enzymatically in the synthesis or degradation of a signalling component. (B) Instead, a non-hydrolysed small molecule could bind to lipase-like EDS1 (and PAD4 or SAG101) and induce conformational changes that are required for further signalling. The plant-specific proteins SABP2 (SALICYLIC ACID BINDING PROTEIN2) and GID1 (GIBBERELIC ACID INSENSITIVE DWARF1) are α/β hydrolase fold proteins that bind a plant hormone. While SABP2 generates the active hormone salicylic acid by hydrolysing bio-inactive methyl salicylate (C), GID1 binds gibberelic acid that is, however, not hydrolysed (D).
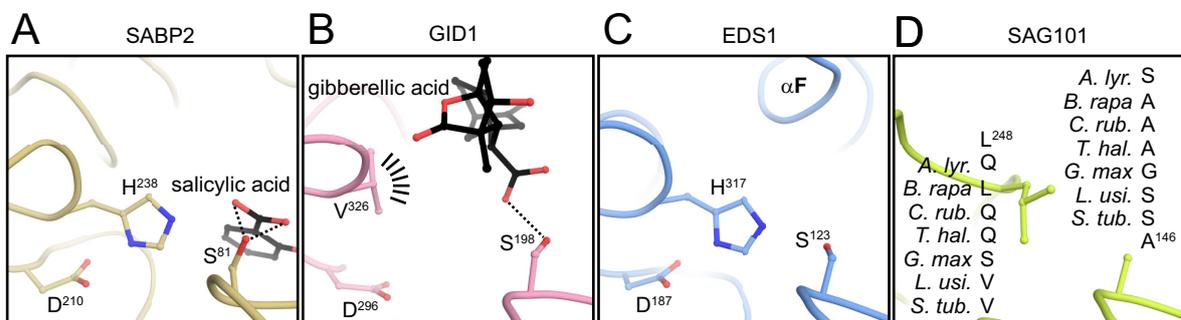
Figure 4.3 | The triads of discussed α/β hydrolase folded proteins. The triad positions of EDS1, SAG101 and of such proteins that are discussed in detail in the text (SABP2, GID1) were structurally superimposed but are presented separately. Bound ligands and substrates are represented as black sticks. The variability among SAG101 triad positions is emphasised in (D) by the presentation of equivalent positions in SAG101 orthologues from other plant species (see Fig. 3.1 for full names).

component of plant systemic acquired resistance (section 1.3.2). Methyl salicylate, in contrast, is biologically inactive and probably represents a transport form that needs SABP2-assisted conversion to salicylic acid [Park *et al.*, 2007]. Interestingly, SABP2 was crystallised bound to a salicylic acid molecule (Fig. 4.3 A) that is completely shielded from the solvent by a "cap domain" – the enzyme is therefore assumed to be inhibited by its reaction product [Forouhar *et al.*, 2005] (Fig. 4.4 A). While the authors could only speculate about an open-closed transition that allows for substrate binding and product release, the SA-bound and apo form of SABP2 look almost identical when crystallised and in solution (as revealed by nuclear magnetic resonance) [Forouhar *et al.*, 2005].
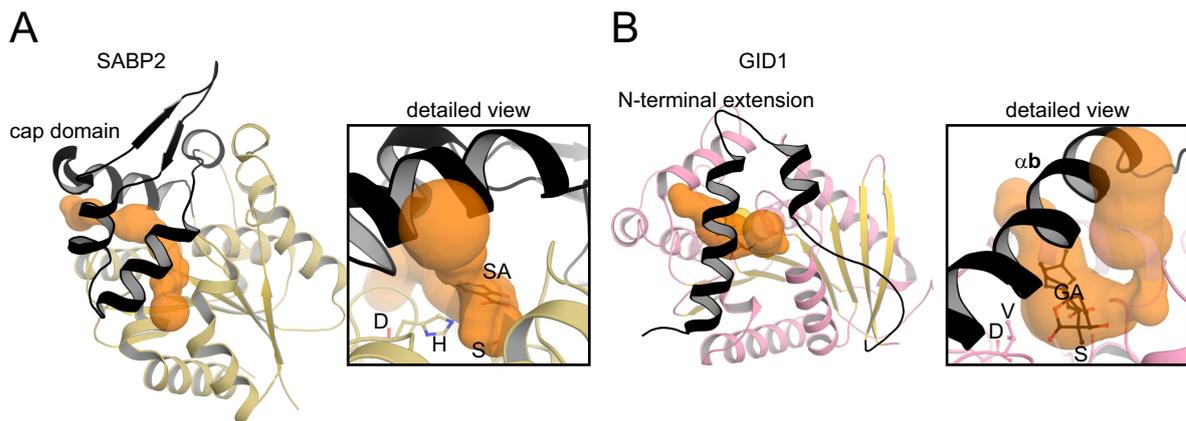
Figure 4.4 | Accessibility of the SABP2 and GID1 triads. Molecular tunnels leading from the surrounding solvent to the triad positions (indicated by a yellow sphere in the more schematic representation) were calculated using the CAVER tool [Petrek *et al.*, 2006; Chovancova *et al.*, 2012] within PyMOL [Schrödinger, LLC, 2010] and are coloured orange. The triad positions are further shown in more detail. (A) The cap domain of SABP2 that shields a salicylic acid (SA) molecule from the solvent, is in black. (B) The N-terminal extension of GID1 that folds on top of a gibberelic acid (GA) molecule, and subsequently binds to a transcription factor as described in the text, is in black.

The GID1 Model   Contrastingly, major conformational changes upon small molecule binding are essential for the function of another plant hormone-binding protein, GIBBERELIC ACID INSENSITIVE DWARF1 (GID1). GID1 participates in multiple plant growth responses including flowering, expansion of leaves and stem elongation [Sun, 2010].

Similar to SABP2, GID1 has been crystallised in the presence of a bound, solvent-shielded hormone, gibberelic acid [Shimada *et al.*, 2008]. Yet in contrast to SABP2, GID1 does not contain a S-D-H triad and is thus unable to hydrolyse gibberelic acid. Instead of the SABP2-like cap domain, GID1 contains an N-terminal extension to the canonical α/β hydrolase fold and Murase *et al.* [2008] could beautifully illustrate how this protein works: gibberelic acid binds to the non-conserved GID1 triad and triggers a conformational change within the N-terminal extension that stably folds onto the gibberelic acid binding pocket. This closed conformation (Fig. 4.4 B) allows binding of a transcriptional regulator that is subsequently marked for proteasomal degradation through the attachment of ubiquitin [Murase *et al.*, 2008].

Intriguingly, the triad's histidine within GID1 is replaced by a valine side chain that participates in gibberelic acid binding (Fig. 4.3 B). Shimada *et al.* [2008] could show that GID1 variants

in which either this particular valine-position or the triad's conserved serine were exchanged to alanine could not efficiently bind gibberelic acid.

Analogies Between SABP2, GID1 and Lipase-Like EDS1   N-terminal EDS1 possesses a conserved catalytic triad (Fig. 3.1 B and 4.3 C) and could thus act in the hydrolysis of a small signalling molecule. In the SAG101-bound state, however, this triad is efficiently shielded from the solvent through an insertion into the canonical α/β hydrolase fold (see Fig. 4.5 A and compare it to the closed state of an active lipase from *Mucor miehei* shown in Fig. 4.5 C).

Molecular dynamics simulations, as described in section 3.4.2, did not suggest that this insertion switches into a more open conformation spontaneously or upon SAG101 dissociation. These simulations, however, covered only a 50 ns which might be too short to allow for large scale molecular motions that regularly take microseconds to seconds [Henzler-Wildman and Kern, 2007].

To cover rather large motions, I used the NMSim webserver [Krüger *et al.*, 2012]. In a three-step approach, NMSim decomposes the protein into rigid clusters and flexible regions and in a second step, the rigid blocks are rotated and translated towards each other whereas the links between them are kept flexible. In a third step, the conformations of backbone and side chain atoms are re-adjusted and the computed structures are iteratively corrected regarding steric clashes and the violation of physico-chemical constraints [Ahmed *et al.*, 2011]. As shown in Appendix Figure A.3, large scale conformational changes that would expose the triad are not to be expected according to NMSim.

There were further no convincing experimental hints pointing towards an EDS1 enzymatic activity. Neither recombinant EDS1 in full length nor its lipase-like domain, EDS1[1-384], did hydrolyse pNP-esters analogous to SABP2 (Fig. 3.21). Similar assays are commonly used to identify and characterise hydrolytically active enzymes (see [Reiter *et al.*, 2000] and others).

We could additionally demonstrate that central immune functions of EDS1 do not depend on hydrolytic activity. The simultaneous exchange of all S-D-H triad residues within a EDS1 S123A/D187A/H317A mutant had no effect on the protein's function in *Arabidopsis* TIR-NB-LRR-mediated and basal resistance against the pathogenic oomycete *Hyaloperonospora* (Fig. 3.19).

Although the functionality of EDS1 S123A/D187A/H317A in immune signalling reasoned against a GID1-like function in binding (instead of hydrolysing) a small molecule, we further perturbed the triad's environment as described in section 3.4.4. In pathogen assays, even a quintuple EDS1 mutant variant was able to successfully fend of a *Hyaloperonospora* infection, ultimately arguing against small molecule binding to the EDS1 triad pocket as part of this immune response (Fig. 3.20).

Analogies Between SABP2, GID1 and Lipase-Like SAG101   *Arabidopsis* SAG101 does not contain residues of a S-D-H triad (Fig. 4.3 D) and thus can not be enzymatically active although otherwise reported by He [2002], who mass spectroscopically measured the release of oleic acid from triolein using SAG101 that was recombinantly produced in *E. coli.*
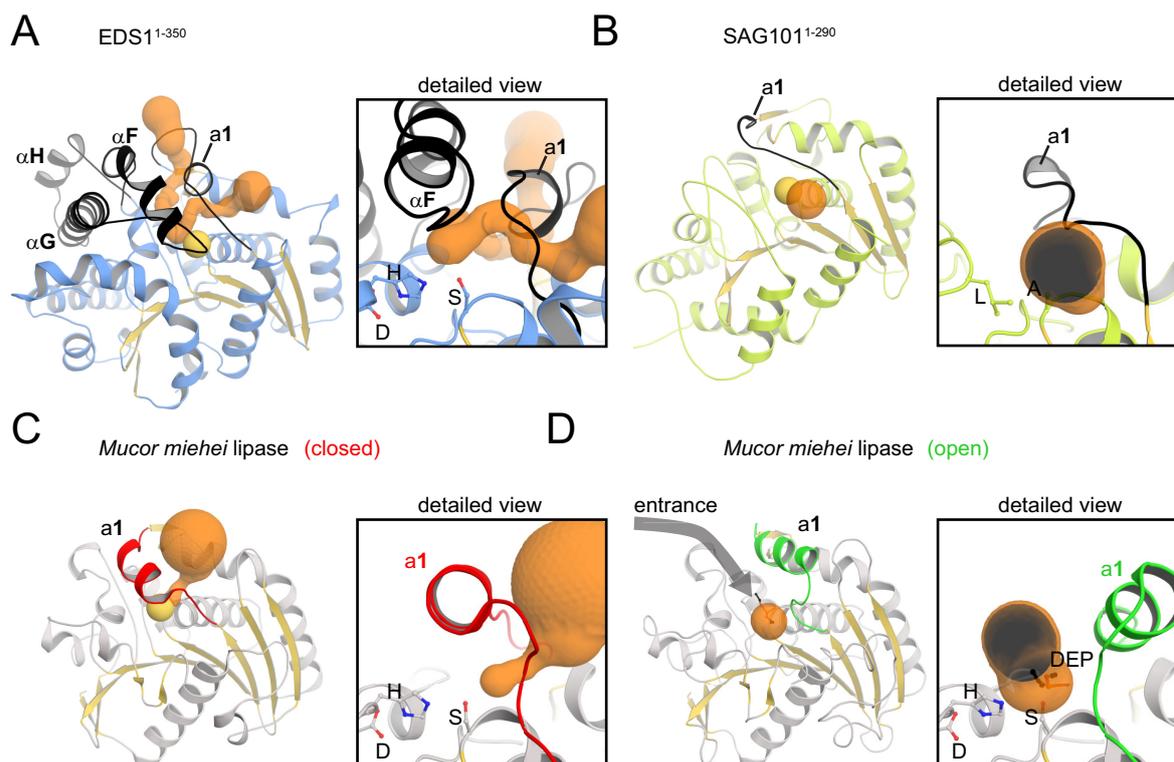
Figure 4.5 | Accessibility of the EDS1 and SAG101 triads. Molecular tunnels leading from the surrounding solvent to the triad positions (indicated by a yellow sphere in the more schematic representation) were calculated using the CAVER tool [Petrek *et al.*, 2006; Chovancova *et al.*, 2012] within PyMOL [Schrödinger, LLC, 2010] and are coloured orange. The triad positions are further shown in more detail. (A) The EDS1 N-terminal insertion as well as the 3_{10} helix a**1** – a remnant of the regulatory *Mucor miehei* lipase helix a**1** – are shown in black. (B) The SAG101 3_{10} helix a**1** is coloured black. (C)/(D) The closed and open conformation of the MML are shown in which α-helix a**1** is highlighted in red and green, respectively. Within (D), the lipase inhibitor DEP (diethylphosphonate), that stabilises the open conformation of MML, is shown as black sticks.

Not having an EDS1-like insertion that covers the former triad pocket (see Fig. 4.5 B and compare it to the open state of an active lipase from *Mucor miehei* shown in Fig. 4.5 D) SAG101 could, however, participate in ligand binding analogous to GID1.

Appealingly, I could not model SAG101 residues 35 to 53 within the crystal structure due to missing electron density that is a result of structural flexibility within this region. Lying close to the former triad pocket, binding of a small molecule could conformationally stabilise this region in order to form a protein binding platform comparable to the N-terminal GID1 extension.

This scenario becomes unlikely when SAG101 orthologue sequences are analysed: as shown in Figure 4.3 D, the SAG101 diad (the aspartate-position is not conserved at all) is highly variable in its composition. Specific recognition of a small molecule should, however, result in the conservation of such residues if they were to participate in ligand binding. Accordingly, in GID1 orthologues from various plant species, the histidine position is consistently replaced by either a valine or an isoleucine side chain, both being aliphatic and chemically related (sequence alignment not shown). They assist ligand binding by a hydrophobic interaction to giberellic acid (Fig. 4.3 B).

Further, the flexible region between *At*SAG101 Proline 35 and Aspartate 53 seems to represent an insertion that does not exist to the same extend in other SAG101 orthologues (see Appendix Fig. A.6). A functional importance is thus highly unlikely.

Analogies Between SABP2, GID1 and Lipase-Like PAD4 While SAG101 seems to have lost an α/β hydrolase-related ability to bind or convert a small molecule, PAD4 has actually been proposed to be hydrolytically active in *Arabidopsis* resistance against insect feeding [Louis *et al.*, 2012]. The conservation of a S-D-H among PAD4 orthologues in mono- and dicotyledonous plant species (Fig. 3.1 B) generally supports this claim.

Louis *et al.* [2012] exchanged single amino acids from the PAD4 S-D-H triad and plants expressing these mutants were used to compare the feeding behaviour of the insects and their infestation rate. The authors found that an exchange of the catalytic serine resulted in plants that were less efficient in controlling insect infestation and in preventing insect feeding, arguing for a critical function of this residue. The same – although to a lesser extend – was observed for an exchange of the triad's aspartate. Paradoxically, an exchange of the histidine position had no effect on insect resistance. Although I described in section 3.1.1 that an incomplete triad can *per se* catalyse one of many α/β hydrolase reactions, any exchange within a conserved S-D-H triad should result in a measurable loss-of-function.

Still, following up on these results, we probed a PAD4 variant with an exchanged triad serine (S118A) in TIR-NB-LRR-mediated and basal resistance of *Arabidopsis* against *Hyaloperonospora* and found that this residue was dispensable (Fig. 3.19 A and B). PAD4 is thus not catalytically active in these immune pathways.

From a structural point of view, it is discussable whether PAD4 could actually hydrolyse a small molecule. As described in section 3.4.2, *Arabidopsis* PAD4 does not contain an EDS1-like insertion into the canonical α/β hydrolase fold. A sequence alignment spanning this region in various EDS1 and PAD4 orthologues (Fig. 3.18), however, reveals that non-existence of an insertion is only typical for PAD4 variants from *Brassicaceae* species. As we saw at the beginning of this discussion, predicted models of such uncommon insertions (using, for instance, our *Arabidopsis* PAD4 homology model) are not necessarily reliable. The question whether an insertion in non-*Brassicaceae* PAD4 orthologues shields the triad or not is, thus, left unanswered.

Could the Inactive Lipase-Like Domain Act as a Decoy? The enzymatically inactive lipase-like domain of SAG101 (and potentially EDS1 and/or PAD4) could act in concert with TIR-NB-LRR R proteins as a decoy in assisting pathogen effector recognition. As introduced in section 1.2.2, decoys are proteins that have no function in plant immunity but rather mimic the actual effector target(s). Thereby, they protect immune components from being modified and, together with R proteins, induce an effector-triggered immunity. In such a model, the N-terminal halves of EDS1 family proteins could act as molecular traps for effectors that are targeting active α/β hydrolases of the plant immune response.

Given that the EDS1 family participates in a broad range of plant signalling pathways and acts downstream of virtually all characterised TIR-NB-LRR R proteins [Aarts *et al.*, 1998; Parker *et al.*, 1996; Borhan *et al.*, 2004; Glazebrook *et al.*, 1997; Zhou, 1998], this scenario is highly unlikely.

4.2.2. The Lipase-Like Domains and Protein-Protein Interactions: Implications From an Inactive α/β "Hydrolase"

As described in section 3.4.1, numerous elements of lipase-like EDS1 and SAG101 are structurally differently organised compared to a structurally related lipase from the fungus *Mucor miehei* (MML) (Fig. 3.13), the most prominent difference being the N-terminal EDS1 insertion. Such insertions into the canonical α/β hydrolase fold are commonly found among lipases which use them as lids to regulate the activity of the enzyme through interfacial activation as described in section 3.4.2.

However, as shown above, also non-lipolytic enzymes like SABP2 can contain similar insertions (in this case the cap-domain) that are involved in the protein's regulation in a yet unknown manner. The N-terminal extension of GID1 does not only regulate small molecule access to the triad but additionally acts as a hub for protein-protein interactions upon binding of a non-hydrolysed ligand (gibberellic acid).

In another α/β hydrolase, the acetylcholinesterase (AChE), an insertion does not at all shield the active site from the solvent but forms a gorge towards a conserved S-E-H catalytic triad (Fig. 4.6 A and B). Faerman *et al.* [1996] proposed on the basis of molecular dynamics simulations that small motions along this Ω-loop allow rapid access of the substrate and release of products rendering the acetylcholinesterase one of the fastest existing enzymes.

The AChE is of interest for two reasons: first, a backdoor mechanism has been proposed as an alternative route towards or away from the catalytic triad (Fig. 4.6 A). This mechanism has been widely discussed since the first AChE crystal structure [Sussman *et al.*, 1991] to explain the astonishingly high turnover numbers that reach up to 10000 product molecules per second. Only recently, Sanson *et al.* [2011] used another crystal structure, in which AChE is inhibited by the snake-venom component aflatoxin, in combination with molecular dynamic simulations to show that the flexible side chains of a tyrosine and a tryptophan residue contribute to an efficient traffic of substrates and products.

While a related backdoor entrance does not necessarily exist in EDS1 and is actually not supported by my molecular dynamics simulations (section 3.4.2), this anecdote shall illustrate that even minor conformational changes (e.g. flexibility between EDS1 helices αF and αG as depicted in Fig. 3.16 B and C) may suffice to allow access of small molecules to the EDS1 triad.

The NL-1 Model Secondly, introduction of the AChE is important for the discussion of an AChE-related α/β hydrolase that does neither convert nor bind a small molecule and represents the third conceivable mode of action for the lipase-like domains of EDS1, SAG101 and PAD4. The mammalian protein NEUROLIGIN-1 (NL-1) exhibits a modified AChE Ω-loop that is referred to as Cys-loop and permits accessibility of the triad site (Fig. 4.6 C). Within the NL-1 triad, the serine position is further exchanged to a glycine residue (Fig. 4.6 D) and the combination of these two features ensures that NL-1 is enzymatically inactive.

Instead, NL-1 has become optimised for binding another protein, neurexin, using an interaction interface that resides away from the former active site [Araç *et al.*, 2007]. The NL-1/neurexin interaction is necessary for the formation of synapses in the mammalian brain.
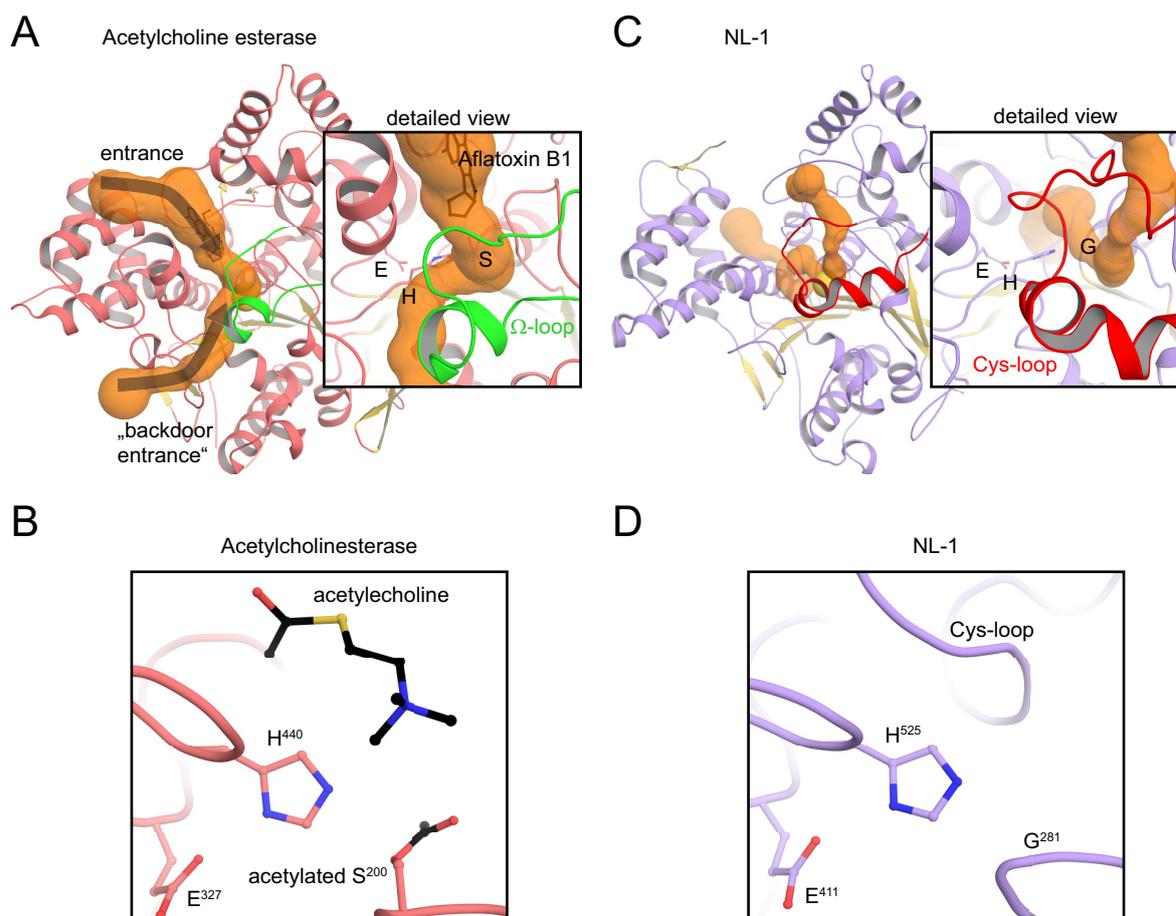
Figure 4.6 | Structural relation between NEUROLIGIN-1 and the acetylcholinesterase. Molecular tunnels leading from the surrounding solvent to the triad positions (indicated by a yellow sphere in the more schematic representation) were calculated using the CAVER tool [Petrek *et al.*, 2006; Chovancova *et al.*, 2012] within PyMOL [Schrödinger, LLC, 2010] and are coloured orange. The triad positions are further shown in more detail. (A) The acetylcholinesterase Ω-loop that guides substrates and products to and away from the active site is shown in green. The snake-venom inhibitor aflatoxin that blocks the enzyme's main entrance is shown as black sticks. The "backdoor entrance" proposed by Sanson *et al.* [2011] is indicated. (C) Structurally related to the Ω-loop, a so-called Cys-loop permits substrate entry to the (incomplete) triad of NEUROLIGIN-1 (NL-1). (B) The conserved S-E-H triad of AChE is shown and a bound acetylcholine molecule indicates where the substrate binds. (D) This substrate binding site is blocked by the Cys-loop in NL-1 that does moreover not contain a triad serine.

Modifications to the Canonical α/β Hydrolase Fold Within the EDS1 and SAG101 Lipase-Like Domains   Comparably, N-terminal EDS1 and SAG101 have become optimised for mutual binding. Two of the three α-helices that form the unique EDS1 N-terminal insertion are directly interacting with SAG101. Among the surface-exposed residues within this insertion, such amino acids that face SAG101 are highly conserved, emphasising their functional importance (Fig. 4.7 A).

SAG101 does, in turn, contain a hydrophobic pocket within its lipase-like domain that is formed by parallel β-strands within the protein's core and an α-helix, α**A'**, which does not exist in the *Mucor miehei* lipase (Fig. 3.13). This hydrophobic arrangement is untypical of lipases and accommodates helix α**H** from the N-terminal EDS1 insertion (Fig. 4.7 B).

97

The hydrophobic SAG101 pocket might be the reason why the protein precipitates in solution whenever EDS1 is not present (Fig. 3.8 D). Within the EDS1/SAG101 heterocomplex, the hydrophobic patch is shielded by EDS1 helix α**H** but inevitably becomes surface-exposed without EDS1 being present. This renders isolated SAG101 prone to protein aggregation through unspecific hydrophobic interactions. Isolated EDS1, in contrast, does not necessarily precipitate as the hydrophobic helix α**H** undergoes conformational changes upon SAG101 dissociation (Fig. 3.17 and A.3) that might result in a more stable conformation.
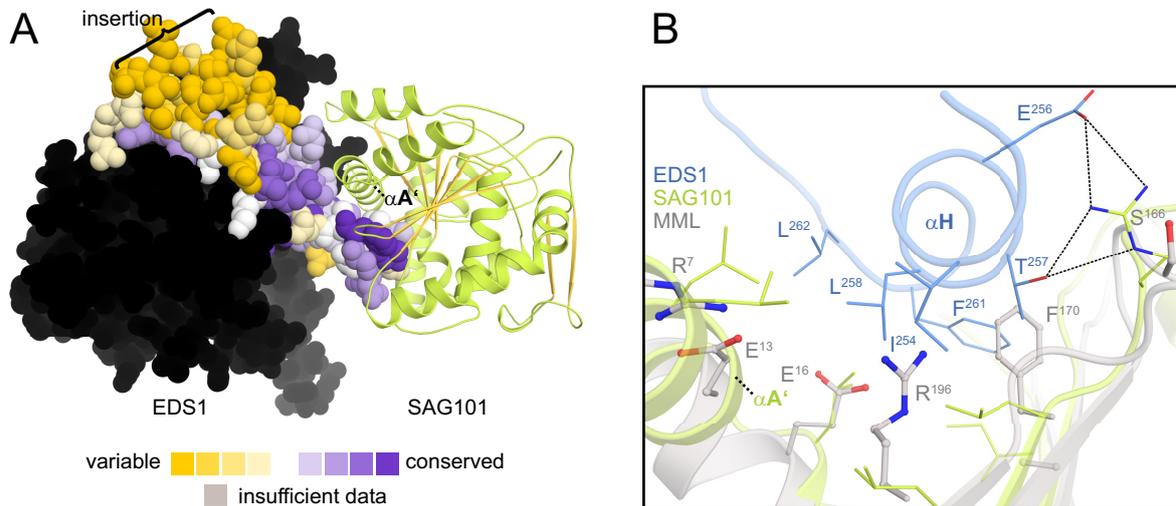
Figure 4.7 | **Adaptions of lipase-like EDS1 and SAG101.** (A) EDS1 is shown as spheres and the conservation score of amino acids among the N-terminal insertion is calculated based on 31 EDS1 orthologue sequences that clustered in a phylogenetic tree (see Appendix Fig. A.1) and was plotted on the structural surface using ConSurf [Ashkenazy *et al.*, 2010]. EDS1 residues that are not part of the insertion are in black. SAG101 is shown in green and the α-helix α**A'** that helps to form a hydrophobic pocket, as described in the text, is highlighted. (B) Comparison of the hydrophobic SAG101 (green) that accommodates EDS1 (blue) α-helix α**H** to the respective region in the *Mucor miehei* lipase (MML; grey) that is dominated by charge arginine and glutamine side chains.

Lipase-Like PAD4 Is Optimised to Bind EDS1 in a SAG101-Like Manner Zhu *et al.* [2011] and Rietz *et al.* [2011] proposed that SAG101 and PAD4 do not compete for EDS1 binding. Rietz *et al.* [2011] postulated this on the basis of an EDS1 variant that is compromised in interaction to PAD4 but not SAG101 arguing for differential modes of interaction. Zhu *et al.* [2011] conducted co-immunoprecipitation experiments with proteins expressed transiently in tobacco plants and found that the degree of complex formation between EDS1 and SAG101 or PAD4 was not altered in the absence or presence of the other EDS1 binding partner.

In the contrary, I found that exchanges within EDS1 helix α**H** equally affected the interaction to both SAG101 and PAD4 in yeast two-hybrid assays and co-purification experiments from tobacco plants as shown in Figures 3.27 D, 3.28 C and 3.29.

Based on this finding, Raphaël Guerois (CEA, Gif-sur-Yvette) built a homology model of PAD4 in complex with EDS1. This model suggested that PAD4 likewise possesses a hydrophobic pocket (Fig. 3.30 B) and I was able to show that exchanges within this pocket indeed significantly reduced the interaction to EDS1 arguing for the correctness of the EDS1/PAD4 heterodimer

model (Fig. 3.30 C and D). I thus propose that the lipase-like domains of SAG101 and PAD4 have likewise become evolutionary optimised for EDS1 binding and do so in a similar manner.

4.2.3. The Lipase-Like Domain Drives the Interactions Within the EDS1 Family

I showed that N-terminal EDS1 and SAG101 share features of the inactive α/β hydrolase NL-1: (1) SAG101 has a non-functional triad, (2) EDS1 has evolved a non-canonical insertion that restricts triad accessibility and (3) both proteins have co-evolved a key-lock-like protein-protein binding mechanism with EDS1 α-helix α**H** being the key that is inserted into a hydrophobic pocket of lipase-like SAG101 (and additionally PAD4). It might thus well be that N-terminal EDS1 and SAG101, albeit lipase-related, do (analogously to NL-1) exclusively function in protein-protein interactions within the EDS1 protein family (Fig. 4.8).
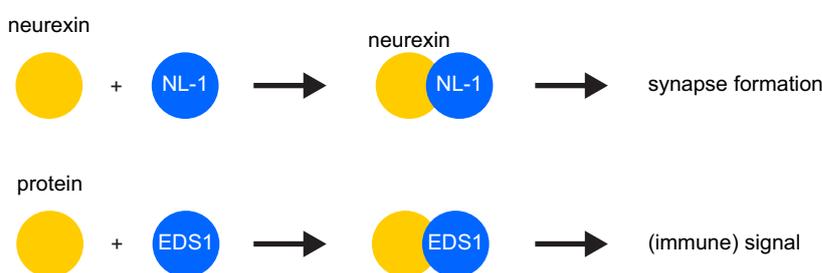
Figure 4.8 | Adaptions of lipase-like EDS1 and SAG101. As described in the text, NL-1 (NEUROLIGIN-1) possesses an α/β hydrolase fold but does not bind a small molecule. Instead, it interacts with another protein, neurexin, and the NL-1/neurexin complex is involved in the formation of synapses of the mammalian brain. As EDS1 and SAG101 share features of NL-1 (see text) and are optimised for mutual binding, an analogous role in protein-protein binding is conceivable for their lipase-like domains.

In line with this model, I found that the N-terminal lipase-like domains of EDS1, SAG101 and PAD4 are indeed driving the interactions within EDS1/SAG101 and EDS1/PAD4 hetero-complexes. Recombinantly produced lipase-like EDS1$^{1-384}$ was able to interact with full-length SAG101 while the C-terminal EP domain of SAG101 (SAG101$^{291-537}$) did not efficiently bind EDS1 (Fig. 3.24).

This is supported by computational analyses using the PPEpred [Su *et al.*, 2009][30] and PISA [Krissinel and Henrick, 2007][31] tools that predicted the lipase-like domain to contribute about two-third of the total EDS1/SAG101 complex binding affinity (Tab. 4.1).

Experimental support for the importance of the lipase-like domain for complex formation comes from *in planta* experiments in which Zhu *et al.* [2011] could show by bimolecular fluorescence complementation that SAG101 interacts with EDS1$^{1-350}$ but not EDS1$^{351-623}$.

Yeast two-hybrid assays that I used to characterise loss-of-interaction variants further empha-sised this. As presented in Figures 3.27, 3.29 and 3.30, four simultaneous amino acid exchanges at the protein-protein interaction interfaces between EDS1 and either SAG101 or PAD4 were

---

[30]PPEpred calculates a potential of mean force (PMF) that is based on the statistical analysis of protein-protein interactions within the Protein Data Bank (PDB) and thus is a knowledge-based potential.
[31]PISA calculates the solvation free energy gain upon formation of the interface considering interface size and composition and the number of salt bridges, H-bonds and disulfide bridges.

Table 4.1 | Computational predictions of EDS1/SAG101 affinities.

	PISA		
	EDS1 + SAG101	**EDS1$^{1\text{-}384}$ + SAG101**	**EDS1 + SAG101$^{291\text{-}537}$**
Affinity (kcal/mol)	-21.4	-16.3	-5.1
	PPEpred		
	EDS1 + SAG101	**EDS1$^{1\text{-}384}$ + SAG101**	**EDS1 + SAG101$^{291\text{-}537}$**
Affinity (kcal/mol)	-23.21	-15.59	-7.62

necessary and sufficient to drastically reduce the formation of EDS1 heterocomplexes. Notably, and with a single exception being the SAG101Y306A exchange, all these exchanges reside within the N-terminal interfaces between the proteins. Site-directed mutagenesis within the C-terminal interface, in contrast, had almost no effect on the interactions (see Appendix Tables A.2 and A.4).

The Conservation of EDS1 and PAD4 Triads   A conserved EDS1 S-D-H triad throughout higher plants, including mono- and dicotyledonous species that have diverged 140–150 million years ago [Chaw *et al.*, 2004], seems to be in disagreement with a functional model that does not involve small molecule hydrolysis. This conservation, however, could be a mere consequence of structural restraints as the triad is in direct contact with the N-terminal EDS1 insertion and might be necessary for the structural stability of this region (Fig. 3.16 A).

The PAD4 triad might be conserved for the same reason, especially within orthologs from non-*Brassicaceae* species that possess N-terminal insertions analogous to EDS1 (Fig. 3.18).

4.3. The EP Domain Contributes Additively to the Interactions Within the EDS1 Family

Although the interface between the lipase-like domains drives the interactions within the EDS1 protein family, I piled experimental evidence that the C-terminal interface between the EP domains additively contributes to efficient complex formation.

Exchanges that lead to a significant loss-of-interaction between recombinantly produced N-terminal EDS1$^{1\text{-}384}$ to SAG101 had no effect when the very same mutations were introduced into full-length EDS1 in an analogous experiment (see Appendix Fig. A.7).

The fact that an exchange within the C-terminal SAG101 interface (Y306A) was necessary to ultimately reduce the interaction between full-length EDS1 and SAG101 in yeast two-hybrid assays (Fig. 3.27 C) further supports a contribution of the EP domains to the EDS1/SAG101 heterodimer formation.

This contribution might explain why lipase-like EDS1$^{1\text{-}384}$ on its own is unable to complement for the loss of full-length EDS1 in *Arabidopsis* TIR-NB-LRR-mediated defence. When both proteins were expressed individually under a constitutive 35S promoter in a *eds1-2* mutant background, plants producing EDS1 in full length were able to initiate a wildtype-like hypersensitive

response upon infection with an incompatible *Hyalopoeronospora* isolate, Cala2. Contrastingly, plants producing only EDS1$^{1-384}$ were not able to halt pathogen infection (Fig. 3.22 A).

4.4. The Functional Importance of Interactions Within the EDS1 Family

4.4.1. EDS1 Heterocomplexes With SAG101 and PAD4 Seem Sufficient for TIR-NB-LRR Signalling

Analysis of available plant genomes, as described in section 3.1, revealed that EDS1 and PAD4 exist among all higher plant species while SAG101 is apparently restricted to dicotyledonous plants (Fig 3.1 A). EDS1 and PAD4 thus seem to represent the more ancestral signalling node while the appearance of SAG101 coincides with that of TIR-NB-LRR R proteins which are virtually non-existent in monocotyledonous species [Tarr and Alexander, 2009].

Yet somewhat paradoxically, SAG101 does not seem to be of central importance for TIR-NB-LRR-mediated resistance signalling: plants that lack SAG101 but still contain PAD4 are able to initiate a defencive hypersensitive response against biotrophic pathogens to the wildtype extend [Feys *et al.*, 2005].

J. Stuttmann (MPIPZ, Köln) could solidify the results of Feys *et al.* [2005] by inoculating wild-type *Arabidopsis* plants of accession Columbia as well as single (*eds1*, *pad4*, *sag101*), double (*pad4/sag101*) and triple (*eds1/pad4/sag101*) mutant plants with an incompatible *Hyaloperonospora* isolate, Cala2. While wildtype plants and such lacking SAG101 were fully resistant through a successful TIR-NB-LRR-mediated immune response, *pad4* mutant plants showed an increased susceptibility (Fig. 4.9). This susceptibility phenotype was dramatically enhanced in a *pad4/sag101* double null mutant due to the introduced redundancy of SAG101 and PAD4 in this type of resistance.

Strikingly, the same degree of susceptibility was observed in *eds1* null mutant plants or plants that lacked EDS1, PAD4 and SAG101, arguing for a model in which EDS1/PAD4 and EDS1/SAG101 complexes represent the functional components of TIR-NB-LRR-mediated immune signalling (Fig. 4.9).
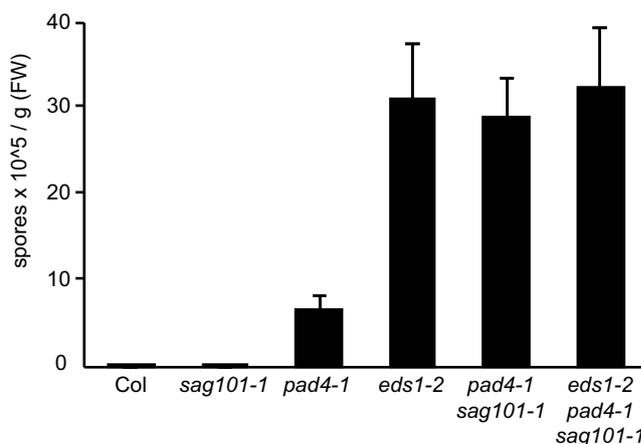

Figure 4.9 | **EDS1 only functions in complex within TIR-NB-LRR-mediated resistance.** *Arabidopsis* plants of Columbia (Col) ecotype were infected with the incompatible *Hyaloperonospora* isolate Cala2 that is restricted in growth through the establishment of a TIR-NB-LRR-mediated resistance. Accordingly, single (*sag101*, *pad4* and *eds1*), double (*pad4/sag101*) and triple (*eds1/pad4/sag101*) mutant plants were infected and susceptibility of these was measured in three different experiments by *Hyaloperonospora* spore counting 6 d after pathogen inoculation.

The EDS1 homooligomer does not seem to fulfil an additional function in this context. While experimental evidence points toward the ability of EDS1 to form such homooligomers (possibly homodimers) in plant cells [Feys *et al.*, 2005] and in yeast two-hybrid experiments [Feys *et al.*, 2005; Rietz *et al.*, 2011], its biological role remains elusive. I will thus focus my discussion on the EDS1/SAG101 and EDS1/PAD4 heterospecies for which I was able to obtain structural information. Protein crystals of isolated EDS1 that I presented in section 3.7.1 will potentially answer if and how EDS1 interacts with itself and what might be the functional consequences of that interaction.

4.4.2. Analysis of an EDS1 Loss-of-Interaction Variant in TIR-NB-LRR-Mediated Resistance

To analyse the functional importance of EDS1 family heterocomplexes in TIR-NB-LRR-triggered immunity, I designed a structure-guided EDS1 loss-of-interaction variant. EDS1**LLIF** carries four simultaneous alanine-exchanges of hydrophobic side chains that reside in EDS1 helix α**H**. As shown in yeast two-hybrid assays and by co-purification experiments from tobacco plants (Fig. 3.27 D, 3.28 C and 3.29), this variant is severely compromised in its interaction to both SAG101 and PAD4.

J. Stuttmann infected wildtype *Arabidopsis* plants and such that expressed EDS1 loss-of-interaction variants under the native EDS1 promoter with an incompatible isolate of *Hyaloperonospora* (Cala2). While wildtype EDS1 and EDS1 variants that retained interaction to SAG101 and PAD4 were able to induce a TIR-NB-LRR-mediated hypersensitive response, EDS1**LLIF** was unable to trigger localised cell death (Fig. 4.10). This strongly supports the biological significance of EDS1/SAG101 and EDS1/PAD4 heterocomplexes within this type of resistance.

Following the line of thought that these heterocomplexes represent the exclusive functional EDS1 states in TIR-NB-LRR-triggered resistance, I expected non-interacting EDS1**LLIF** to produce a phenotype that equals an *eds1* null mutant plant. A slightly less dramatic effect as seen in Fig. 4.10 could, however, be explained by residual complex formation.

Although co-purification experiments using transiently expressed proteins from tobacco plants (Fig. 3.28 B and 3.29 B) implied a full loss-of-interaction between EDS1**LLIF** and SAG101 or PAD4, yeast two-hybrid experiments argued for residual interaction as cells still grew on selective -LWH plates (Fig. 3.27 D and 3.29 A).

EDS1 variants in which the hydrophobic side-chains are not exchanged to neutral alanine side chains but rather polar or charged residues should be tested and could potentate the loss-of-interaction (and loss-of-function).

4.4.3. Functional Importance of Heterocomplexes Is Supported by Surface Conservation

The functional importance of intrafamilial interactions is further supported by a high degree of sequence conservation among residues that participate in EDS1 heterodimer formation. I combined sequence and structure information obtained during this thesis and used the ConSurf
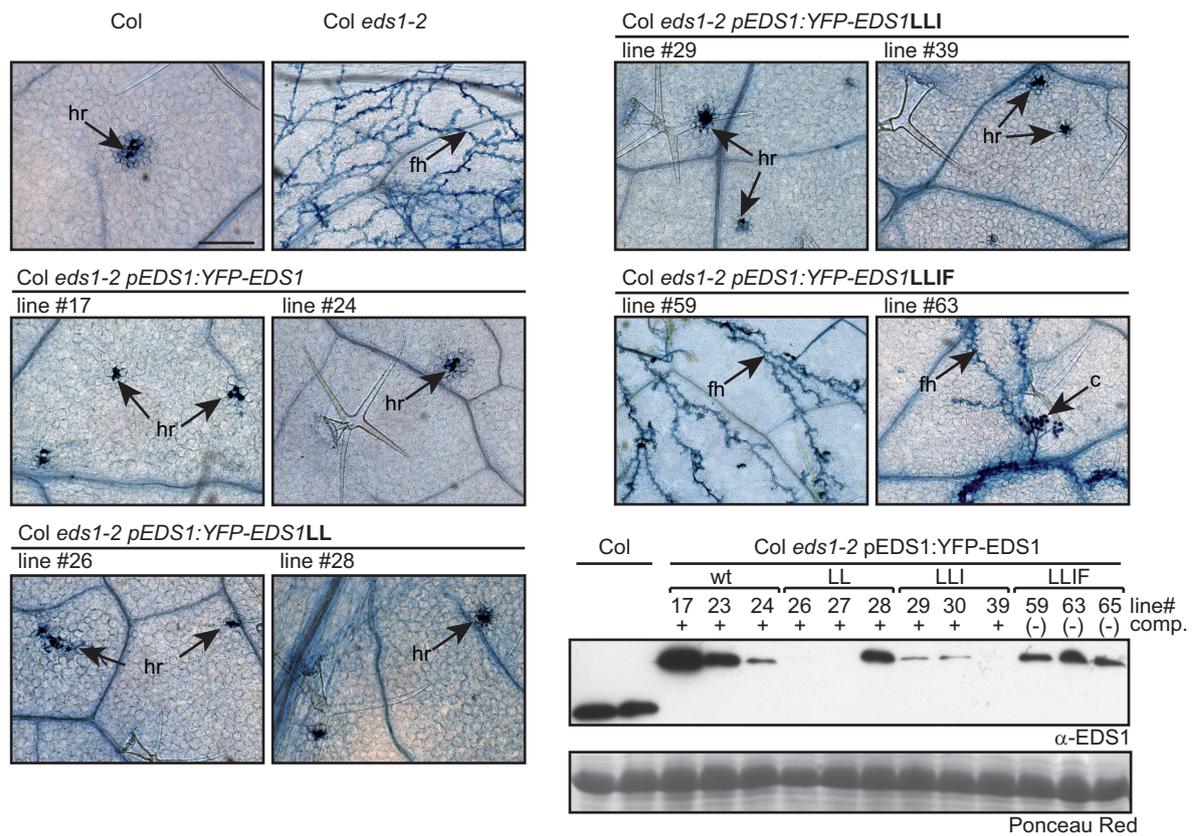
Figure 4.10 | Analysis of EDS1 loss-of-interaction variants in TIR-NB-LRR-mediated resistance.
Wildtype *Arabidopsis* plants of accession Columbia (Col), *eds1-2* null mutant plants and plants expressing different EDS1 variants under an EDS1 promoter (*pEDS1*) in a *eds1-2* null mutant background were infected with the incompatible *Hyaloperonospora* isolate Cala2. EDS1**LL** and EDS1**LLI** represent variants that were only partially compromised in their interaction to SAG101 and PAD4 while EDS1**LLIF** showed a severe loss-of-interaction (Fig. 3.27 D, 3.28 C and 3.29). The establishment of a hypersensitive response accompanied by localised cell death was visualised 6 d after infection by Trypan Blue staining. EDS1 protein accumulation of each tested line is shown in the lower right corner. hr – hypersensitive response, fh – free hyphae, c – conidiophores. Scale bar = 200 μm.

server [Ashkenazy *et al.*, 2010] to plot the degree of residue conservation onto the surfaces of EDS1, SAG101 and the PAD4 homology model.

As illustrated in Figure 4.11, residues at the heterodimer interfaces were significantly more conserved compared to the rest of each protein's surface. Such evolutionary conserved residues at protein surfaces have been shown to indicate functional importance and can be used to characterise and predict protein binding sites for other proteins [Ma *et al.*, 2003], peptides [Petsalaki *et al.*, 2009], DNA [Luscombe and Thornton, 2002], lipids [Adamian *et al.*, 2011] and a wide range of small molecule ligands [Huang and Schroeder, 2006].

The high degree of surface conservation at the interface of EDS1 to either PAD4 or SAG101 further allows to ask how a ternary EDS1/SAG101/PAD4 complex, proposed by Zhu *et al.* [2011], should form. Our experimentally verified EDS1/PAD4 homology model and presence of a single, highly conserved surface patch on EDS1 argue for a shared interface that is used for interaction to SAG101 <u>and</u> PAD4. A ternary complex would thus require the formation of an EDS1 homooligomer in which distinct EDS1 molecules either bind SAG101 or PAD4. Although

EDS1 homooligomerisation is inevitably different from heterodimerisation as the monomeric building blocks of homooligomers are generally symmetry related to each other [Goodsell and Olson, 2000], surface analysis does not reveal further patches of high sequence conservation that could be used as interfaces between EDS1 homomers (Fig. 4.11).
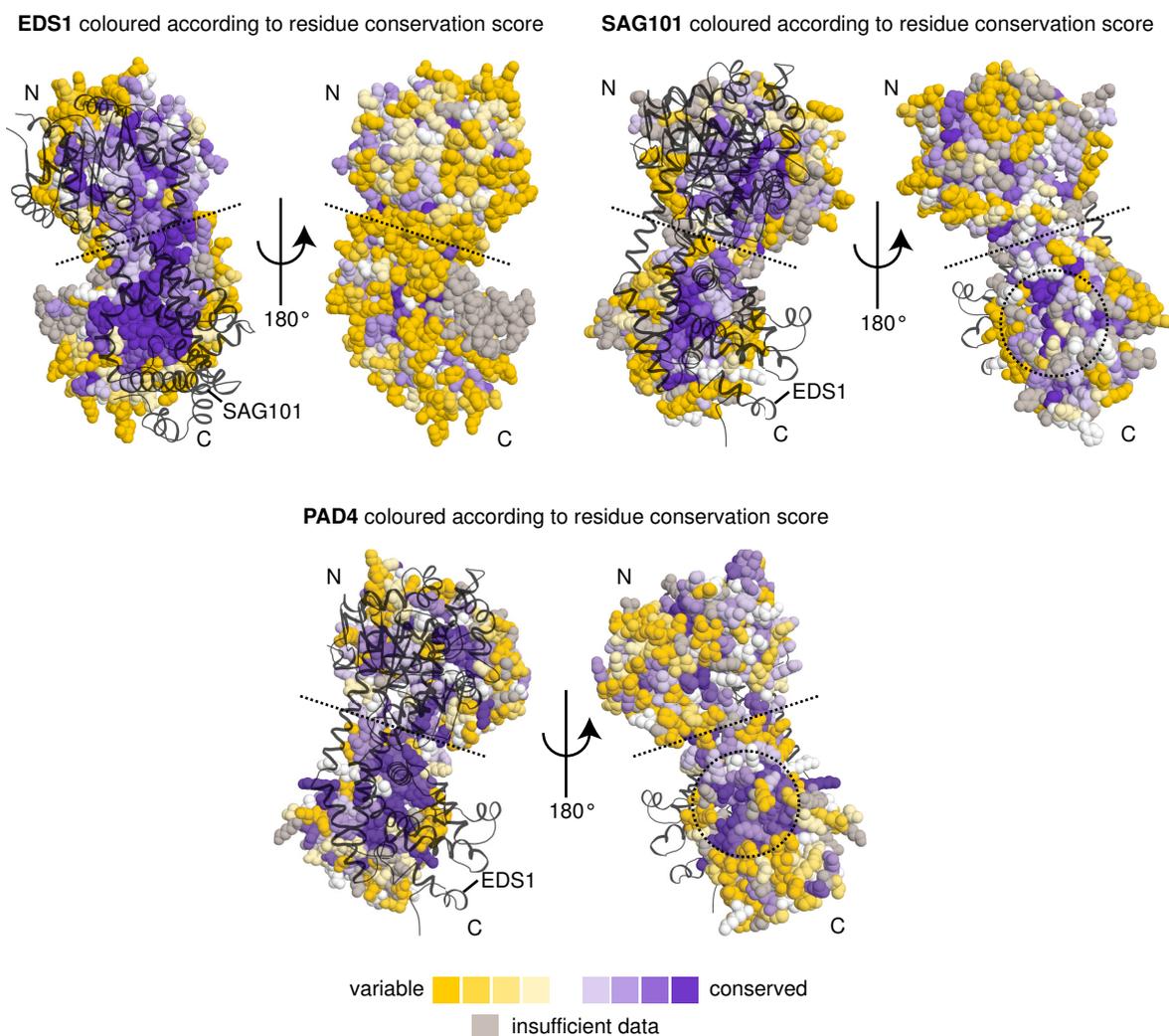

EDS1 coloured according to residue conservation score

SAG101 coloured according to residue conservation score

PAD4 coloured according to residue conservation score

variable   conserved

insufficient data

Figure 4.11 | **Surface conservation at the interaction interfaces of EDS1, SAG101 and PAD4.** The protein surfaces of EDS1 (top left) and SAG101 (top right) and PAD4 (bottom) are represented as spheres and are coloured according to the degree of residue conservation and polarity of interface residues. The conservation score is calculated based on orthologue sequences that clustered in a phylogenetic tree (see Appendix Fig. A.1) and was plotted on the structural surface using ConSurf [Ashkenazy *et al.*, 2010][32]. A dotted line indicates the boundary between N-terminal (N) lipase-like and C-terminal (C) EP domain. Dotted circles represent patches of conserved residues that will be further discussed in the text.

4.4.4. Is the EDS1/SAG101 Interaction Transient?

While the results presented above emphasise the importance of protein-protein interactions within the EDS1 family, their biochemical function remains elusive. As EDS1 has been proposed to exist in molecularly and spatially distinct complexes within the plant cell, transitions between

different states were and are assumed to be of central functional importance [Rietz *et al.*, 2011; García *et al.*, 2010].

The interface between EDS1 and SAG101, however, is highly untypical for a transient heterodimer. Both proteins interact through an interface of more than 4000 $\mathring{A}^2$ (see Tab. 4.2) that can be, according to a classification by Lo Conte *et al.* [1999], considered a "large interface" compared to "standard interfaces" between proteins that bury only an area of 1600 ($\pm$400) $\mathring{A}^2$.

Not only is the interface particularly large but also hydrophobic: Lo Conte *et al.* [1999] found that protein-protein interfaces on average do contain 56 % polar residues and thus do not significantly differ from surfaces of monomeric proteins (57 % non-polar residues [Miller *et al.*, 1987]). In contrast, 66 % of all residues within the EDS1/SAG101 interface are non-polar. Interfaces of that size and polarity are rather typical for homodimeric interactions: Nooren and Thornton [2003b] analysed homo- and heterodimeric proteins and found that obligate[33] homodimers have an average interface of 3818 $\mathring{A}^{2[34]}$ that consists of 64 % non-polar residues whereas the average interfaces of transient homodimers and heterodimers in general are much smaller and more polar (see Tab. 4.2).

Table 4.2 | Structural details of the EDS1/SAG101 complex interface compared to transient and non-transient homo- and heterodimers.

| | Data taken from [Nooren and Thornton, 2003b] | | | | |
| | Homodimers | | Heterodimers | | |
	Obligate	Transient	Non-transient	Transient	EDS1/SAG101
Area ($\mathring{A}^2$)	3818	1480	1788	2424	**4163**
Non-polar residues (%)	64	63	56	59	**66**

The predicted binding affinities of below -20 kcal/mol (Tab. 4.1) do further point towards a strong interaction between the two proteins. Nishi *et al.* [2011], for instance, used the PISA tool for a large scale analysis of protein-protein interactions that they classify as follows: weak transient (binding affinity $\sim$ 0 kcal/mol), strong transient (0 $\geq$ binding affinity $\geq$ -20 kcal/mol) and permanent (binding affinity < -20 kcal/mol).

A strong association between EDS1 and SAG101 is further supported by the fact that sequential mutations at the interface between them were necessary to reduce their interaction (Fig. 3.27 C and D). It is thus unlikely that EDS1 and SAG101 dissociate spontaneously – once formed, the complex will rather need an external trigger to fall apart.

In line with the importance of stable EDS1 heterocomplexes, the proteins do mutually stabilise each other *in vitro* and in plants. As discussed before, recombinant SAG101 tends to precipitate without the stabilising effect of EDS1 – probably due to the exposure of hydrophobic patches.

---

[33] The constituents (protomers, monomers) of an obligate complex are unstable on their own *in vivo* whereas those of a non-obligate complex can exist on their own [Nooren and Thornton, 2003a]

[34] The therein cited interface area was doubled as it is differently calculated. While Lo Conte *et al.* [1999] define the interface area as the sum of the solvent-accessible surface areas (ASA) of the isolated components less that of the complex, Nooren and Thornton [2003a] define it as the sum of the surface areas of the two individual proteins minus the surface area of the protein complex divided by two. I decided to choose the former definition.

Accordingly, the protein levels of SAG101 (and PAD4) in plant cells were dramatically reduced in *eds1* null mutant plants [Feys *et al.*, 2005].

4.4.5. SAG101 Might Redirect EDS1 to the Nucleus

The tight association of SAG101 to EDS1 might be necessary for proper localisation of EDS1 during an immune response. García *et al.* [2010] found that triggering of TIR-NB-LRR R proteins through an infection with *Pseudomonas syringae* pv. tomato (*Pto*) DC3000 expressing the AvrRps4 effector leads to an increase of nuclear EDS1 levels 3 hours after infection. As higher nuclear EDS1 accumulation is not reflected by an increased *EDS1* gene expression, this is likely due to a post-transcriptional mechanism. Nuclear accumulation of EDS1 was further shown to precede or coincide with the expression of EDS1-dependent defence genes arguing for a role of nuclear EDS1 in transcriptional reprogramming [García *et al.*, 2010]. Accordingly, EDS1 that is artificially kept in the cytoplasm is substantially compromised in pathogen-induced expression changes [García *et al.*, 2010].

Interestingly, Zhu *et al.* [2011] found by co-expressing pairs of EDS1, SAG101 and PAD4, fused to fluorescent proteins that EDS1 preferentially localises to the nucleus when SAG101 but not PAD4 is co-expressed with EDS1. When the authors co-expressed all three proteins, some portion of EDS1 remained cytosolic. PAD4 could, however, not alter the localisation of EDS1 if it was expressed 24 or 48 h after EDS1 and SAG101. SAG101 might thus have a function in retaining EDS1 inside the nucleus. This could be achieved through the tight association of exclusively nuclear SAG101 (Fig. 1.5) to EDS1 for that I have gathered structural and experimental evidence.

The dispensability of SAG101 in TIR-NB-LRR-mediated resistance of *Arabidopsis* against *Hyaloperonospora* infection (Fig. 4.9) appears to be inconsistent with such a function. Even in the absence of SAG101, EDS1, however, remains nuclear-cytoplasmatic [Zhu *et al.*, 2011] and low levels of nuclear EDS1 could be sufficient for successful TIR-NB-LRR signalling in this pathogen-host interaction. In another immune response or in a situation where the plant has to integrate biotic and abiotic stress signals simultaneously, a tighter regulation of EDS1 accumulation through SAG101 might be necessary.

When García *et al.* [2010] artificially removed EDS1 from the nucleus through fusion of a functional NES (nuclear export signal) sequence to EDS1, the respective mutant plants were compromised in TIR-NB-LRR-mediated as well as basal resistance against incompatible and compatible pathogens (*Hyaloperonospora* and *Pseudomonas syringae*). Notably, the plants retained a substantial degree of resistance and this could imply that even a minor portion of EDS1, which is not efficiently removed from the nucleus, is sufficient for residual immune signalling using these systems [García, 2009].

4.4.6. Loss-of-Interaction Variants Will Help to Discriminate EDS1 Family Functions

Recent studies suggested that the EDS1 protein family members EDS1, PAD4 and SAG101 exist in distinct molecular assemblies with individual functions in plant signalling pathways [Louis *et al.*,

2012; Zhu *et al.*, 2011; Rietz *et al.*, 2011; Kim *et al.*, 2012]. This is generally supported by a phylogenetic analysis described in section 3.1. EDS1, PAD4 and SAG101 orthologue sequences form three distinct clusters within a phylogenetic tree (Fig. A.1), arguing for differential functions of SAG101- and PAD4-containing complexes despite the functional redundancy in *Arabidopsis* TIR-NB-LRR-mediated defence [Feys *et al.*, 2005].

This is in accord with the observation that SAG101 and PAD4 do fulfil non-redundant roles in *Arabdiopsis* defence against the plant-pathogenic turnip crinkle virus: Zhu *et al.* [2011] could show that plants lacking SAG101 are equally susceptible to viral infection compared with plants that do either lack EDS1 or PAD4, indicating that PAD4 can not even partially compensate for the loss of SAG101. The authors further described an individual role of PAD4 in salicylic acid-dependent signalling using the same pathogen-host system, based on the finding that an exogenous treatment of plants with a salicylic acid analog restored resistance in *pad4* but neither *eds1* nor *sag101* null mutant plants [Zhu *et al.*, 2011].

As introduced before, PAD4 acts independent of EDS1 (and SAG101) in another type of *Arabidopsis* resistance against the phloem-feeding insect green peach aphid [Pegadaraju *et al.*, 2007, 2005; Louis *et al.*, 2012]. Hence, PAD4 and SAG101 are able to fulfil individual roles in plant defence signalling and are not generally redundant.

Despite the stabilising effect of EDS1 on PAD4 and SAG101 within presumably tightly bound heterodimers (see section 4.4.4), at least PAD4 seems able to exist and function without EDS1 since the its ability to confer resistance against green peach aphid was not altered in plants lacking EDS1 [Pegadaraju *et al.*, 2007].

The EDS1, PAD4 and SAG101 loss-of-interaction variants that I was able to characterise will be useful to further dissect the functions of protein complexes and isolated components of the EDS1 protein family in these and probably other signalling pathways.

The EDS1L262P variant that is compromised in interaction to PAD4 [Rietz *et al.*, 2011] was, for instance, used by Kim *et al.* [2012] to show that EDS1, PAD4 and also EDS1L262P (i.e. EDS1 not being in complex with PAD4) do interact with the TIR-NB-LRR protein VICTR to function in root growth arrest of *Arabidopsis* plants. This suggests that the EDS1 helix α**H**, in which Leucine 262 resides, is involved in the interactions of EDS1 to SAG101 and PAD4 (as comprehensively shown in this thesis) but not necessarily participates in interactions to R proteins of the TIR-NB-LRR type.

4.5. The SAG101 and PAD4 EP Domains Could Mediate Protein-Protein Interactions Outside the EDS1 Family

With the lipase-like domains being optimised for interactions within the EDS1 family, the EP domain could emerge as a docking platform for proteins outside the family. Strikingly, the EP domains of EDS1 and SAG101 are, yet distantly, related to proteins that fulfil a function in large protein assemblies (Table 3.7 and Fig. 3.23).

Structurally, these proteins are (at least partially) made of helical repeat motifs like tetratricopeptide repeats (TPR) or armadillo motifs [Groves and Barford, 1999]. Such helical repeats possess an extended surface area and have a modular character: helical tandems can be struc-

turally deformed and modified in their length and composition of surface residues while retaining an overall stable fold. These tandems can act in a combinatorial manner to mediate interaction with multiple proteins and thus their interfaces "provide not only a scaffolding function that allows the assembly of multiprotein complexes, but also directly modulate the conformation and biological behaviour of the proteins with which they interact" [Groves and Barford, 1999].

Interestingly, and in contrast to EDS1, SAG101 and PAD4 contain surface-exposed patches of highly conserved amino acids within parts of the EP domain that are not covered by EDS1 in the heterodimers (see dotted circles in Fig. 4.11 that represent an area of ~ 700 Å^2). The EP domains are thus candidates for rather transient interactions to TIR-NB-LRR R proteins, effector proteins and other emerging EDS1 interactors.

Although TIR-NB-LRR R proteins have long been known to require the EDS1 family for downstream signalling, it is unclear how a signal is transduced between them. In recent studies, EDS1 has been shown by co-immunoprecipitation and bimolecular fluorescence complementation to exist in a physical complex with TIR-NB-LRR proteins (RPS4, RPS6) and additionally effector proteins (AvrRps4 and HopA1 that are recognised by RPS4 and RPS6, respectively) [Heidrich et al., 2011; Bhattacharjee et al., 2011].

Intriguingly, RPS4, in turn, cooperates genetically with RRS1 (RESISTANCE TO R. SOLANA-CEARUM1) in plant immunity and both genes are encoded in a head-to-head configuration, separated by only 264 base pairs, in the Arabidopis genome [Narusaka et al., 2009]. RRS1 is an atypical NB-LRR protein that contains a C-terminal WRKY domain [Deslandes et al., 2002]. The WRKY domain is able to bind DNA and the defining feature of WRKY transcription factors that modulate numerous plant processes [Rushton et al., 2010]. Although both proteins have yet not been shown to physically interact, EDS1 could act in concert with RPS4 and RRS1 as an adaptor protein that connects recognition of an effector (e.g. AvrRps4) to transcriptional reprogramming during an immune response [Bonardi and Dangl, 2012].

Beyond effector and NB-LRR proteins, EDS1 has been shown to physically interact with another protein, SRFR1 (SUPPRESSOR OF rps4-RLD1). SRFR1 is a negative regulator of effector-triggered immunity and Bhattacharjee et al. [2011] found that it exists in a complex together with EDS1 and the TIR-NB-LRR R proteins RPS4 and RPS6. As the interaction of SRFR1 to EDS1 was significantly reduced in the presence of the corresponding effector proteins AvrRps4 and HopA1 (recognised by RPS4 and RPS6, respectively), the authors propose that SRFR1 could set a threshold at which the R proteins are activated. Maybe coincidental, SRFR1 contains α-helical tetratricopeptide repeats [Kwon et al., 2009].

4.6. Conclusions and Future Perspectives

Structural information that I obtained during this thesis will help to explain the biochemical functions of the immune-regulatory EDS1 plant protein family.

Despite being sequentially and structurally related to lipases, this is not the central key to their understanding. We found that neither EDS1 nor PAD4 need critical residues from a putative catalytic S-D-H triad in Arabidopis TIR-NB-LRR-mediated and basal resistance against biotrophic pathogens. SAG101 misses critical lipase features and does definitely not act in the hydrolysis of small molecules.

Experimental and structural evidence further suggests that lipase-like EDS1 could likewise have lost its enzymatic activity in general as the protein failed to hydrolyse typical hydrolase substrates and the triad seems to be effectively shielded from the solvent.

While I can not ultimately rule out that either N-terminal EDS1 or PAD4 retained enzymatic function in other signalling pathways, the lipase-like domains of EDS1, PAD4 and SAG101 have co-evolved a function as the anchoring point for protein-protein interactions within the EDS1 family.

I analysed these interactions *in vitro* and *in vivo* and proposed structure-guided loss-of-interaction variants of EDS1, SAG101 and PAD4. These were verified and preliminarily analysed *in planta* by J. Stuttmann (MPIPZ, Köln) and provide an excellent basis to further separate between biological functions of EDS1-requiring complexes and the single components in plant signalling.

They have already helped to emphasise the importance of protein-protein interactions within the EDS1 family in TIR-NB-LRR-mediated immune signalling. Supported by genetic studies, our analysis suggests that EDS1 therein functions solely in complex with SAG101 and PAD4. Exchanges of hydrophobic residues to polar or charged side chains at the interface between the proteins might suppress residual complex formation and should be used in a more comprehensive analysis.

Instead of being transient, the large interaction interfaces within EDS1 heterodimers suggest surprisingly stable complexes. A tight association between exclusively nuclear SAG101 and EDS1 could be required for EDS1 accumulation in the nucleus that is required for transcriptional reprogramming during an immune response [García *et al.*, 2010]. Transition between molecular EDS1 states should be analysed, for instance, by competition assays using EDS1, SAG101 and PAD4 variants that are expressed *in planta* and fused to different affinity tags.

Beyond the scope of *Arabidopsis* resistance against biotrophic pathogens that is discussed in this thesis, the EDS1 protein family participates in the regulation of cell death upon photo-oxidative stress [Mateo *et al.*, 2004; Rustérucci *et al.*, 2001; Straus *et al.*, 2010], insect resistance [Pegadaraju *et al.*, 2007, 2005; Louis *et al.*, 2012] and root growth arrest [Kim *et al.*, 2012]. The herein characterised mutant variants, and the presented structural information in general, will help to further understand the underlying signalling events.

The Elucidation of Further EDS1 Structures For a comprehensive structure-function relationship, further EDS1 states should be structurally elucidated. An EDS1 structure without binding partner could, for instance, answer open questions concerning its ability to form homomers.

Additionally, EDS1 can almost be assumed to adopt diverse conformations in different molecular assemblies. Lo Conte *et al.* [1999] and Nooren and Thornton [2003b] consistently found, when analysing larger sets of protein-protein interactions, that the assembly or disassembly of complexes with large interaction interfaces ($>1000\,\text{Å}^2$) is regularly associated with a conformational change. Thus, each functional EDS1 state adds a piece to the puzzle that connects EDS1 structure to function.

Without putting too much interpretation into these results, the NMSim webserver [Krüger *et al.*, 2012] introduced before gives an idea how such changes may look like (Fig. 4.12).
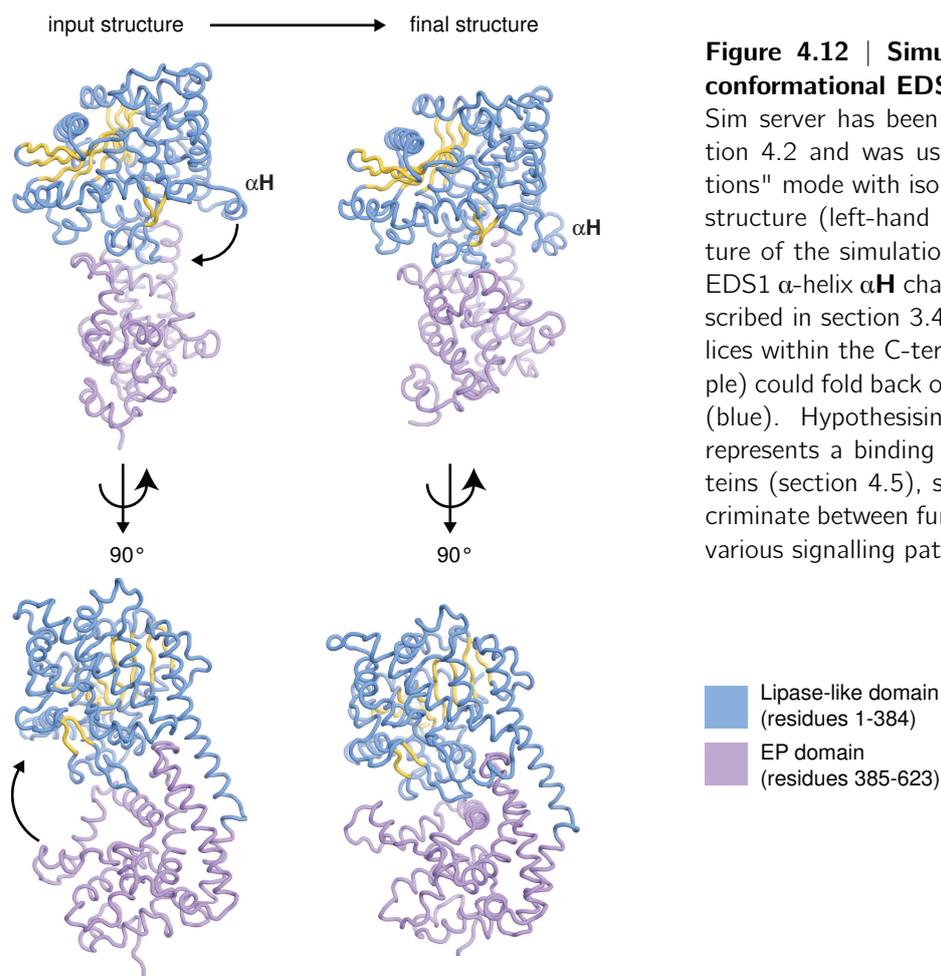
Figure 4.12 | Simulation of large scale conformational EDS1 changes. The NM-Sim server has been described before (section 4.2 and was used in "large scale motions" mode with isolated EDS1 as an input structure (left-hand side). The final structure of the simulation reveals that not only EDS1 α-helix αH changes its position, as described in section 3.4.2, but additionally helices within the C-terminal EP domain (purple) could fold back on the lipase-like domain (blue). Hypothesising that the EP domain represents a binding platform to other proteins (section 4.5), such motions could discriminate between functional EDS1 states in various signalling pathways.

Lipase-like domain (residues 1-384)

EP domain (residues 385-623)

A further focus should lie on the preparation of PAD4. Although we used a PAD4 homology model to characterise its interaction to EDS1, we are still unable to answer molecular details. It would be interesting to see whether PAD4 contains a solvent-accessible triad that acts in signalling molecule hydrolysis. As attempts to produce it recombinantly in *E. coli* failed, eukaryotic expression systems (e.g. yeast, insect cell culture or tobacco plants) should be considered.

In this context, the herein presented PAD4 homology model could and should be used to derive domain fragments or introduce chemical modifications that potentially stabilise the protein.

Once a soluble variant is obtained, the homology model could assist structure-guided mutagenesis of entropic surface residues that can be detrimental to crystallisation.

These approaches can be successfully combined as shown by Longenecker *et al.* [2001]. The authors produced a domain fragment of a RhoGTPase nucleotide exchange factor (RhoGEF) and exchanged lysine and glutamine residues that are statistically probable surface residues to alanines. By replacing polar, flexible side chains that formed clusters in the primary sequence, they created a binding region that allowed new intermolecular contacts and promoted protein crystallisation.

This and numerous other approaches applicable to non-crystallising proteins including, (1) hydrophobic to hydrophilic mutations on the protein surface, (2) deletion of potentially flexible regions and (3) the use protein-fusions to enhance solubility and crystallisability are review by Dale *et al.* [2003]. The authors further suggest to consider using orthologous proteins from

different species that, in case of PAD4, should include such containing an N-terminal insertion (Fig. 3.18).

A. Appendix

A.1. Appendix Tables

Table A.1 | Thermofluor screen composition table.

A1	50 mM Sodium acetate, pH 4.5	E1	150 mM Sodium bromide
A2	50 mM Sodium citrate, pH 5.0	E2	150 mM Sodium fluoride
A3	50 mM Sodium citrate, pH 5.5	E3	150 mM Sodium thiocyanate
A4	50 mM Bis-Tris, pH 6.0	E4	1 mM EDTA
A5	50 mM MES, pH 6.5	E5	0.2 % (w/v) Tween-80
A6	50 mM Sodium phosphate, pH 7.0	E6	0.2 % (v/v) Triton X-100
A7	50 mM Potassium phosphate, pH 7.0	E7	150 mM Glycine, pH 7.5
A8	50 mM HEPES, pH 7.0	E8	50 mM Glycylglycylglycine, pH 7.5
A9	50 mM Tris, pH 7.5	E9	500 mM Arginine, pH 7.5
A10	50 mM HEPES, pH 7.5	E10	5 % (v/v) Glycerol
A11	50 mM TEA, pH 7.5	E11	10 % (v/v) Glycerol
A12	50 mM Bicine, pH 8.0	E12	20 % (v/v) Glycerol
B1	50 mM Tris, pH 8.0	F1	5 % (w/v) D-Glucose
B2	50 mM HEPES, pH 8.0	F2	5 % (w/v) D-Fructose
B3	50 mM Imidazole, pH 8.0	F3	5% (w/v) Sucrose
B4	50 mM Tris, pH 8.5	F4	5 % (w/v) D-D-sorbitol
B5	50 mM CHES, pH 9.0	F5	5 % (w/v) Trehalose dihydrate
B6	50 mM CAPS, pH 10.0	F6	10 (v/v) Ethylenglycol
B7	A2 +50 mM NaCl	F7	1 % (v/v) PEG400
B8	A2 +150 mM NaCl	F8	1 % (w/v) PEG4000
B9	A2 +500 mM NaCl	F9	1 % (w/v) PEG8000
B10	A4 +50 mM NaCl	F10	10 mM urea
B11	A4 +150 mM NaCl	F11	10 mM GuHCl
B12	A4 +500 mM NaCl	F12	3 % (v/v) DMSO
C1	A10 +50 mM NaCl	G1	5 % (v/v) MPD
C2	A10 +150 mM NaCl	G2	5 % (w/v) Trimethyl-N-oxide
C3	A10 +500 mM NaCl	G3	1 mM DTT
C4	B4 +50 mM NaCl	G4	5 mM DTT
C5	B4 +150 mM NaCl	G5	1 mM β-ME
C6	B4 +500 mM NaCl	G6	5 mM β-ME
C7	150 mM NaCl	G7	1 mM TCEP pH 7.5
C8	150 mM KCl	G8	5 mM TCEP pH 7.5
C9	150 mM $CaCl_2$	G9	5 % (w/v) DDAO
C10	150 mM $MgCl_2$	G10	10 mM CHAPS
C11	150 mM LiCl	G11	Water
C12	10 mM NH_4Cl	G12	5 mM NAD
D1	10 mM $ZnCl_2$	H1	5 mM NADH
D2	10 mM $NiCl_2$	H2	5 mM ATP
D3	10 mM $CuCl_2$	H3	5 mM ADP
D4	10 mM CsCl	H4	5 mM AMP
D5	10 mM $MnSO_4$	H5	10 mM cholic acid
D6	150 mM Sodium citrate	H6	5 mM Salicylic acid in 10 mM HEPES, pH 7.5
D7	150 mM Sodium acetate	H7	5 mM 2,3 DHBA in 10 mM HEPES, pH 7.5
D8	150 mM Sodium sulfate	H8	5 mM 2,5 DHBA in 10 mM HEPES, pH 7.5
D9	150 mM Sodium dihydrogenphosphate, pH 7.0	H9	4 % (v/v) Ethanol
D10	150 mM Sodium nitrate	H10	4 % (v/v) Acetonitrile
D11	150 mM Sodium tartrate	H11	4 % (v/v) 2-Propanol
D12	150 mM Sodium iodide	H12	4 % (v/v) Methanol

Table A.2 | Effects of mutations within SAG101 as tested in the yeast two-hybrid assay.

Mutation in BD-SAG101	Effect against AD-EDS1[1-384]	Effect against AD-EDS1
L12A	no effect	no effect
L12S	no effect	no effect
L15A	protein instability	protein instability
L21A	no effect	no effect
L21S	partial loss-of-interaction	no effect
I141A	partial loss-of-interaction	no effect
R169E	protein instability	protein instability
L171A	no effect	no effect
L199A	no effect	no effect
L12A/L21A	partial loss-of-interaction	no effect
L12S/L21S	protein instability	protein instability
L12A/L21A/I141A	loss-of-interaction	no effect
R169E/L171A	protein instability	protein instability
N298A	not tested	no effect
I302A	not tested	no effect
Y306A	not tested	no effect
W309A	not tested	no effect
L12A/L21A/N298A	not tested	no effect
L12A/L21A/I302A	not tested	no effect
L12A/L21A/Y306A	not tested	no effect
L12A/L21A/W309A	not tested	protein instability
L12A/L21A/R428E	not tested	no effect
L12A/L21A/I141A/N298A	not tested	no effect
L12A/L21A/I141A/I302A	not tested	protein instability
L12A/L21A/I141A/Y306A	not tested	loss-of-interaction
L12A/L21A/I141A/R428E	not tested	no effect

Table A.3 | Effects of mutations within EDS1[1-384] as tested in the yeast two-hybrid assay.

Mutation in AD-EDS1[1-384]	Effect against BD-SAG101
R234E	partial loss-of-interaction
C245A	no effect
I254A	loss-of-interaction
E256K	no effect
T257A	no effect
L258A	loss-of-interaction
F261A	loss-of-interaction
L262A	loss-of-interaction
W306A	no effect
R353E	no effect
Y357A	no effect

Table A.4 | Effects of mutations within EDS1 as tested in the yeast two-hybrid assay.

Mutation in AD-EDS1	Effect against BD-SAG101
I254A	no effect
L258A	no effect
F261A	no effect
L262A	no effect
L258A/L262A	no effect
I254A/L258A/L262A	partial loss-of-interaction
I254A/L258A/F261A/L262A	loss-of-interaction
I254A/L258A/F261A/L262A/E416K	no additional loss-of-interaction
I254A/L258A/F261A/L262A/F419A	no additional loss-of-interaction
I254A/L258A/F261A/L262A/N422A	no additional loss-of-interaction
I254A/L258A/F261A/L262A/V423A	no additional loss-of-interaction
I254A/L258A/F261A/L262A/V423S	no additional loss-of-interaction
I254A/L258A/F261A/L262A/E427A	no additional loss-of-interaction
I254A/L258A/F261A/L262A/R475E	no additional loss-of-interaction
I254A/L258A/F261A/L262A/D481K	no additional loss-of-interaction
I254A/L258A/F261A/L262A/V423S/E427A	no additional loss-of-interaction

Table A.5 | Effects of mutations within PAD4 as tested in the yeast two-hybrid assay.

Mutation in BD-PAD4	Effect against AD-EDS1
L11A	no effect
L11S	no effect
Q12E	no effect
V15A	no effect
V15S	no effect
M16A	partial loss-of-interaction
L21A	no effect
L21S	partial loss-of-interaction
F143A	partial loss-of-interaction
R247E	no effect
L315S	no effect
W319S	no effect
M16A/L21A	partial loss-of-interaction
M16A/L21S	loss-of-interaction
M16A/L21A/F143A	partial loss-of-interaction
M16A/L21S/F143A	loss-of-interaction

A.2. Appendix Figures
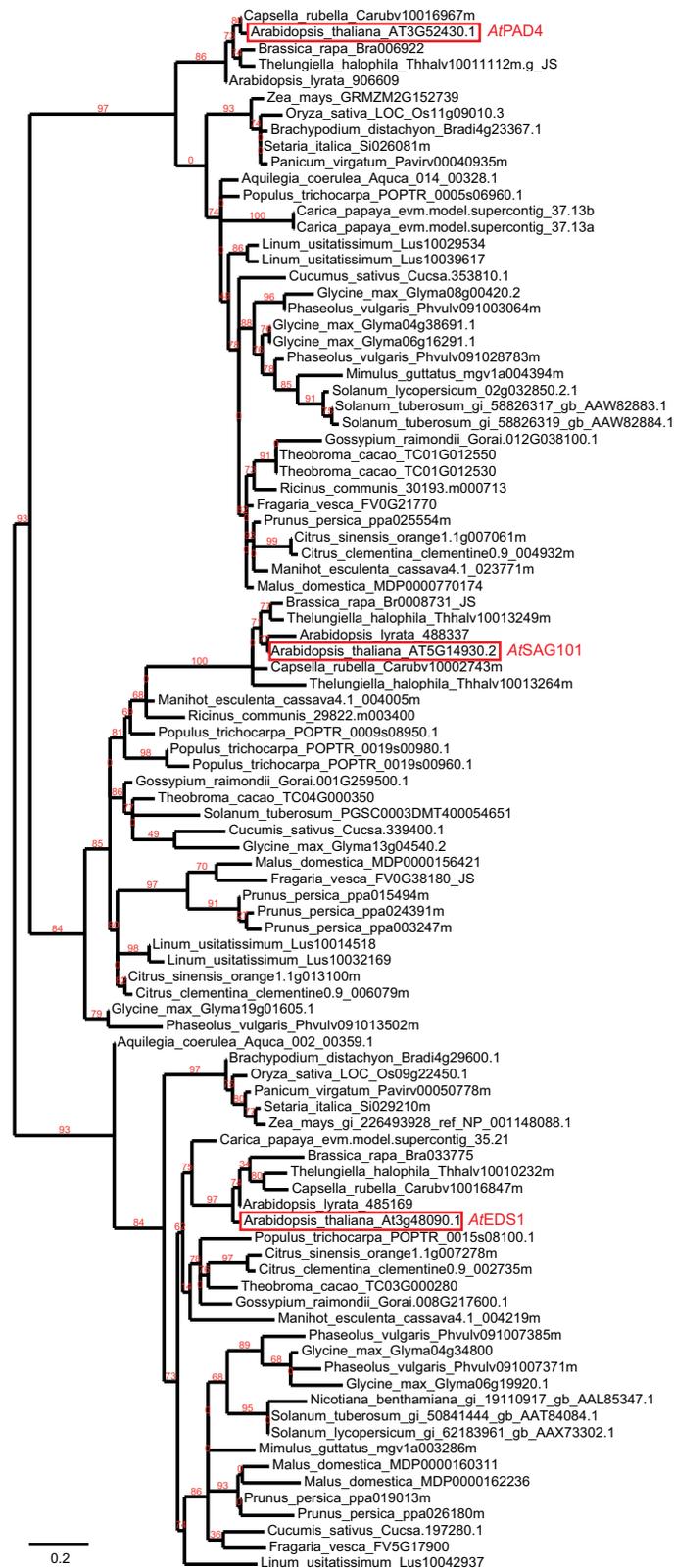
Figure A.1 | **Phylogenetic tree of the EDS1 protein family.** The sequences were retrieved and clustered in a phylogenetic tree as described in section 3.1. Annotated sequences of *Arabidopsis thaliana* EDS1, PAD4 and SAG101 are marked within each cluster. Bootstrap values are given in percentages for each branch. The scale bar equals a distance of 20 changes per 100 amino acid positions.
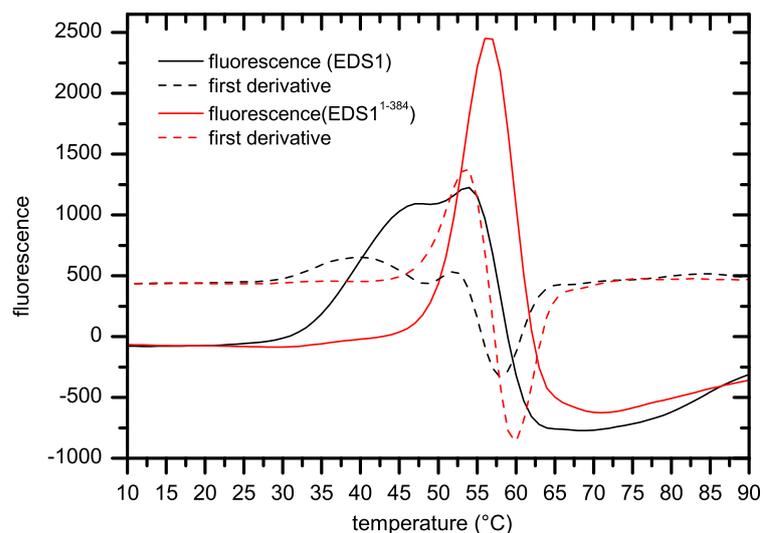
Figure A.2 | Thermufluor melting curves. EDS1 and EDS1[1-384] were thermally unfolded in the same sodium phosphate buffer, pH 7.5 and in presence of SYPRO® Orange dye. Unfolding events lead to an increase of fluorescence as the dye emits of a distinct wavelength upon binding to hydrophobic patches that become accessible upon protein denaturation. The maximum value of the first derivative curve equals the protein's melting point T_m.
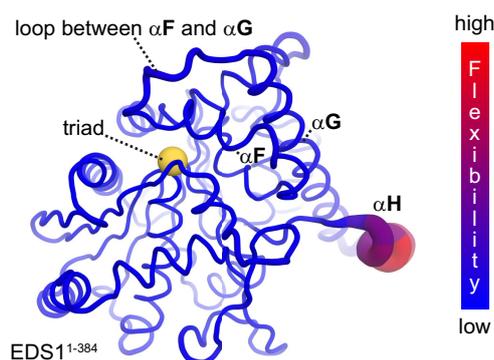
Figure A.3 | Flexibility of EDS1 helix αH. Full length EDS1 without bound SAG101 was used as input structure for the calculation of "large scale motions" using the NMSim server [Krüger *et al.*, 2012] (refer to section 4.2.1 for details). Here, the lipase-like EDS1 domain is coloured according to structural flexibility. Analogous to Figure 3.17, helices αF,G,H within the N-terminal EDS1 insertion are marked. The triad position, where a small molecule ligand or a substrate could bind, is indicated by a yellow sphere.
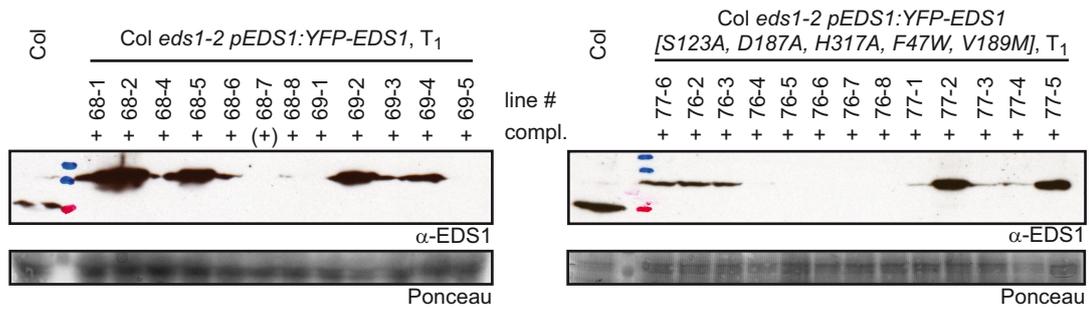
Figure A.4 | The EDS1 triad does not bind a small molecule in TIR-NB-LRR-mediated resistance.
Complementation of lines that express the depicted amounts of protein was tested as described for
Figure 3.20. All protein variants were expressed under the native EDS1 promoter (*pEDS1*)
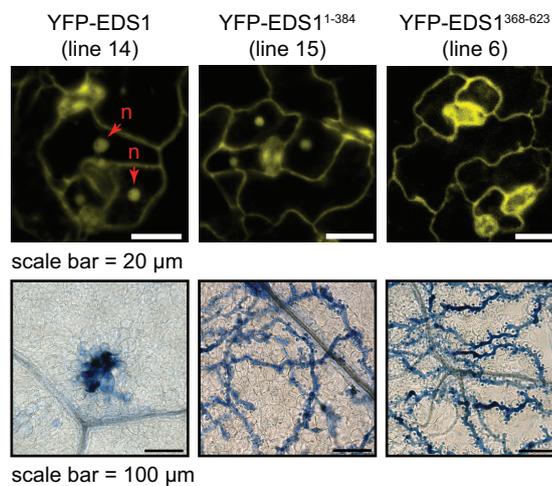
Figure A.5 | The N-terminal EDS1 domain does not complement for the loss of full length EDS1.
Upper panel: localisation of the indicated protein fusions upon 35S-driven expression in transgenic *Arabidopsis eds1-2* plants (T2). Full length EDS1 and the N-terminal domain localise to plant nuclei and the cytoplasm. The C-terminal domain, in contrast, never was visible in nuclei. Guard cell autofluorescence appears very bright due to generally lower expression/fluorescence of the C-terminal construct. Lower panel: infection phenotypes of the same lines upon infection with the incompatible *Hyaloperonospora* isolate Cala2. Infected leaves were stained with Trypan Blue 7 days after inoculation and analysed under a microscope.

```
                                          ↓Pro35                      ↓Asp53
Arabidopsis_thaliana     ILEIHNPP.YSNHDPGLQVSKKKKDSGLEFQ.I
Populus_trichocarpa      IKDLHAET.NPNQQ...........LSVRHK.L
Capsella_rubella         LEEER.........................M
Eutrema_parvulum         ISEIHELS.YPNKD...........SALGIN.I
Arabis_alpina            ISEIHAST.YPNQD...........SALGVK.I
Arabidopsis_lyrata       ISEIHKSP.YSNQD...........SGLGIK.I
Solanum_lycopersicum     ISKLQKQNLTLLEDPSLP.......FSVKNQIY
Fragaria_vesca           IEKQKQIN..PNAEPCL............FN.E
Ricinus_communis         IYTLQSEA.DPNQQ.......QRQSLSLRWK.F
Vitis_vinifera           ILELQRDQ.........................
Medicago_truncatula      SFKAHNSENEPY.............ISELFRIS
```

Figure A.6 | The flexible SAG101 loop. Sequence alignment of SAG101 orthologs from different plant species was carried out using SeaView [Gouy *et al.*, 2010] and presented using ESPript [Gouet *et al.*, 1999]. The positions of *At*SAG101 Proline 35 and Aspartate 53, which represent the borders of a loop that could not been modelled in the EDS1/SAG101 crystal structure due to high structural flexibility, are indicated.
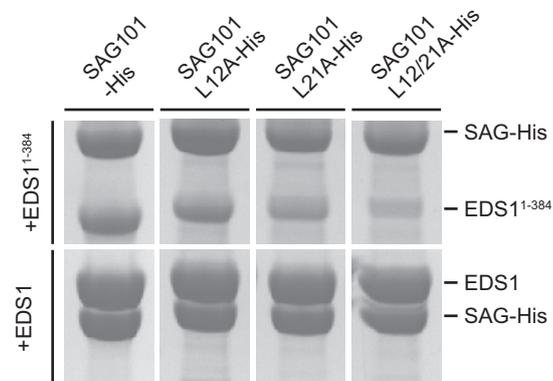
Figure A.7 | Full length and lipase-like EDS1 co-purifications with SAG101 variants. EDS1, EDS1[1-384] and SAG101 loss-of-interaction variants were expressed separately in *E. coli*. *E. coli* cells that expressed each of the SAG101 variant were lysed together with cells expressing an excess of either untagged EDS1 or EDS1[1-384]. His-tagged SAG101 (and interacting EDS1 or EDS1[1-384], respectively) was purified by IMAC and the elution fractions were separated by SDS-PAGE. Bands corresponding to each protein are indicated on the right.

A.3. Definition of Crystallographic R-Factors

The R_{meas} value was defined by Diederichs and Karplus [1997] and should be used as a measure of X-ray diffraction data quality instead of R_{sym} and R_{merge} that both depend on the multiplicity of collected data. It is defined as:

$$R_{meas} = \frac{\sum\limits_{h} \sqrt{\frac{n_h}{n_h - 1}} \sum\limits_{i}^{n_h} |\hat{I}_h - I_{h,i}|}{\sum\limits_{h} \sum\limits_{i}^{n_h} I_{h,i}}$$

where $\hat{I}_h = \frac{1}{n_h} \sum\limits_{i}^{n_h} I_{h,i}$ (the mean intensity for all observations i of reflection h) and n_h is its multiplicity (the total number of observations).

List of Figures

List of Tables

References

Aarts N, Metz M, Holub E, Staskawicz BJ, Daniels MJ, Parker JE (1998). Different requirements for EDS1 and NDR1 by disease resistance genes define at least two R gene-mediated signaling pathways in Arabidopsis. *Proc Natl Acad Sci U S A* **95**:10306–11.

Adamian L, Naveed H, Liang J (2011). Lipid-binding surfaces of membrane proteins: evidence from evolutionary and structural analysis. *Biochim Biophys Acta* **1808**:1092–1102.

Adams PD, Afonine PV, Bunkóczi G, Chen VB, Echols N, Headd JJ, Hung LW, Jain S, Kapral GJ, Grosse Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner RD, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH (2011). The Phenix software for automated determination of macromolecular structures. *Methods* **55**:94–106.

Agniswamy J, Joyce MG, Hammer CH, Sun PD (2008). Towards a rational approach for heavy-atom derivative screening in protein crystallography. *Acta Crystallogr D Biol Crystallogr* **64**:354–67.

Ahmed A, Rippmann F, Barnickel G, Gohlke H (2011). A normal mode-based geometric simulation approach for exploring biologically relevant conformational transitions in proteins. *J Chem Inf Model* **51**:1604–22.

Altenbach D, Robatzek S (2007). Pattern recognition receptors: from the cell surface to intracellular dynamics. *Mol Plant Microbe Interact* **20**:1031–9.

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

Araç D, Boucard AA, Ozkan E, Strop P, Newell E, Südhof TC, Brunger AT (2007). Structures of neuroligin-1 and the neuroligin-1/neurexin-1 beta complex reveal specific protein-protein and protein-Ca2+ interactions. *Neuron* **56**:992–1003.

Ashkenazy H, Erez E, Martz E, Pupko T, Ben-Tal N (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res* **38**:529–33.

Aucoin MG, McMurray-Beaulieu V, Poulin F, Boivin EB, Chen J, Ardelean FM, Cloutier M, Choi YJ, Miguez CB, Jolicoeur M (2006). Identifying conditions for inducible protein production in E. coli: combining a fed-batch and multiple induction approach. *Microb Cell Fact* **5**:27.

Ausubel FM (2005). Are innate immune signaling pathways in plants and animals conserved? *Nat Immunol* **6**:973–9.

Bachovchin WW (1985). Confirmation of the assignment of the low-field proton resonance of serine proteases by using specifically nitrogen-15 labeled enzyme. *Proc Natl Acad Sci U S A* **82**:7948–51.

Baltrus DA, Nishimura MT, Romanchuk A, Chang JH, Mukhtar MS, Cherkis K, Roach J, Grant SR, Jones CD, Dangl JL (2011). Dynamic evolution of pathogenicity revealed by sequencing and comparative genomics of 19 Pseudomonas syringae isolates. *PLoS Pathog* **7**:e1002132.

Benbrook CM (2012). Impacts of genetically engineered crops on pesticide use in the U.S. – the first sixteen years. *Env Sci Europe* **24**:24.

Berejnov V, Husseini N, Alsaied O, Thorne R (2006). Effects of cryoprotectant concentration and cooling rate on vitrification of aqueous solutions. *J Appl Crystallogr* **39**:244–51.

Bergfors T (2003). Seeds to crystals. *J Struct Biol* **142**:66–76.

Berman HM (2000). The Protein Data Bank. *Nucleic Acids Res* **28**:235–42.

Bernoux M, Ve T, Williams S, Warren C, Hatters D, Valkov E, Zhang X, Ellis JG, Kobe B, Dodds PN (2011). Structural and functional analysis of a plant resistance protein TIR domain reveals interfaces for self-association, signaling, and autoregulation. *Cell Host Microbe* **9**:200–11.

Bhattacharjee S, Halane MK, Kim SH, Gassmann W (2011). Pathogen effectors target Arabidopsis EDS1 and alter its interactions with immune regulators. *Science* **334**:1405–8.

Blow D (2010). *Outline Of Crystallography For Biologists*. Oxford University Press, U.S.A., 1 ed.

Blow DM, Birktoft JJ, Hartley BS (1969). Role of a buried acid group in the mechanism of action of chymotrypsin. *Nature* **221**:337–40.

Bogan AA, Thorn KS (1998). Anatomy of hot spots in protein interfaces. *J Mol Biol* **280**:1–9.

Bolen DW (2004). Effects of naturally occurring osmolytes on protein stability and solubility: issues important in protein crystallization. *Methods* **34**:312–22.

Bonardi V, Dangl JL (2012). How complex are intracellular immune receptor signaling complexes? *Front Plant Sci* **3**:237.

Bond CS (2003). TopDraw: a sketchpad for protein structure topology cartoons. *Bioinformatics* **19**:311–2.

Borhan MH, Holub EB, Beynon JL, Rozwadowski K, Rimmer SR (2004). The arabidopsis TIR-NB-LRR gene RAC1 confers resistance to Albugo candida (white rust) and is dependent on EDS1 but not PAD4. *Mol Plant Microbe Interact* **17**:711–9.

Bradford MM (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**:248–54.

Brady L, Brzozowski AM, Derewenda ZS, Dodson E, Dodson G, Tolley S, Turkenburg JP, Christiansen L, Huge-Jensen B, Norskov L (1990). A serine protease triad forms the catalytic centre of a triacylglycerol lipase. *Nature* **343**:767–70.

Bragg WH, Bragg WL (1913). The Reflection of X-rays by Crystals. *Proc R Soc Lond. Series A* **88**:428–38.

Brenner S (1988). The molecular evolution of genes and proteins: a tale of two serines. *Nature* **334**:528–30.

Brünger AT (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. *Nature* **355**:472–5.

Carr PD, Ollis DL (2009). Alpha/beta hydrolase fold: an update. *Protein Pept Lett* **16**:1137–48.

CCP4 (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* **50**:760–3.

Chandra-Shekara AC, Navarre D, Kachroo A, Kang HG, Klessig D, Kachroo P (2004). Signaling requirements and role of salicylic acid in HRT- and rrt-mediated resistance to turnip crinkle virus in Arabidopsis. *Plant J* **40**:647–59.

Chaw SM, Chang CC, Chen HL, Li WH (2004). Dating the monocot-dicot divergence and the origin of core eudicots using whole chloroplast genomes. *J Mol Evol* **58**:424–41.

Chayen NE (2004). Turning protein crystallisation from an art into a science. *Curr Opin Struct Biol* **14**:577–83.

Chinchilla D, Shan L, He P, de Vries S, Kemmerling B (2009). One for all: the receptor-associated kinase BAK1. *Trends Plant Sci* **14**:535–41.

Chothia C, Lesk AM (1986). The relation between the divergence of sequence and structure in proteins. *EMBO J* **5**:823–6.

Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, Gora A, Sustr V, Klvana M, Medek P, Biedermannova L, Sochor J, Damborsky J (2012). CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput Biol* **8**:e1002708.

Chruszcz M, Wlodawer A, Minor W (2008). Determination of protein structures–a series of fortunate events. *Biophys J* **95**:1–9.

Clackson T, Wells JA (1995). A hot spot of binding energy in a hormone-receptor interface. *Science* **267**:383–6.

Coates ME, Beynon JL (2010). Hyaloperonospora Arabidopsidis as a pathogen model. *Annu Rev Phytopathol* **48**:329–45.

Collier SM, Moffett P (2009). NB-LRRs work a "bait and switch" on pathogens. *Trends Plant Sci* **14**:521–9.

Crick FHC, Magdoff BS (1956). The theory of the method of isomorphous replacement for protein crystals. I. *Acta Crystallogr* **9**:901–8.

Dagert M, Ehrlich SD (1979). Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells. *Gene* **6**:23–8.

Dale GE, Oefner C, D'Arcy A (2003). The protein as a variable in protein crystallization. *J Struct Biol* **142**:88–97.

Darnell SJ, Page D, Mitchell JC (2007). An automated decision-tree approach to predicting protein interaction hot spots. *Proteins* **68**:813–23.

de St Groth S, Webster RG, Datyner A (1963). Two new staining procedures for quantitative estimation of proteins on electrophoretic strips. *Biochim Biophys Acta* **71**:377–91.

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M, Claverie JM, Gascuel O (2008). Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**:465–9.

Derewenda U, Brzozowski AM, Lawson DM, Derewenda ZS (1992). Catalysis at the interface: the anatomy of a conformational change in a triglyceride lipase. *Biochemistry* **31**:1532–41.

Deslandes L, Olivier J, Theulieres F, Hirsch J, Feng DX, Bittner-Eddy P, Beynon J, Marco Y (2002). Resistance to Ralstonia solanacearum in Arabidopsis thaliana is conferred by the recessive RRS1-R gene, a member of a novel family of resistance genes. *Proc Natl Acad Sci U S A* **99**:2404–9.

Dessen A, Tang J, Schmidt H, Stahl M, Clark JD, Seehra J, Somers WS (1999). Crystal Structure of Human Cytosolic Phospholipase A2 Reveals a Novel Topology and Catalytic Mechanism. *Cell* **97**:349–60.

Diederichs K, Karplus PA (1997). Improved R-factors for diffraction data analysis in macro-molecular crystallography. *Nat Struct Biol* **4**:269–75.

Dodds PN, Lawrence GJ, Catanzariti AM, Teh T, Wang CIA, Ayliffe MA, Kobe B, Ellis JG (2006). Direct protein interaction underlies gene-for-gene specificity and coevolution of the flax resistance genes and flax rust avirulence genes. *Proc Natl Acad Sci U S A* **103**:8888–93.

Dodson G, Wlodawer A (1998). Catalytic triads and their relatives. *Trends Biochem Sci* **23**:347–52.

Doye JPK, Louis AA, Vendruscolo M (2004). Inhibition of protein crystallization by evolutionary negative design. *Phys Biol* **1**:9–13.

Du H, Klessig DF (1997). Identification of a Soluble, High-Affinity Salicylic Acid-Binding Protein in Tobacco. *Plant Physiol* **113**:1319–27.

Edgar RC (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**:113.

Ekici OD, Paetzel M, Dalbey RE (2008). Unconventional serine proteases: variations on the catalytic Ser/His/Asp triad configuration. *Protein Sci* **17**:2023–37.

Ellisdon AM, Dimitrova L, Hurt E, Stewart M (2012). Structural basis for the assembly and nucleic acid binding of the TREX-2 transcription-export complex. *Nat Struct Mol Biol* **19**:328–36.

Emsley P, Lohkamp B, Scott WG, Cowtan K (2010). Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* **66**:486–501.

Ericsson UB, Hallberg BM, Detitta GT, Dekker N, Nordlund P (2006). Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal Biochem* **357**:289–98.

Evans G, Pettifer RF (2001). CHOOCH : a program for deriving anomalous-scattering factors from X-ray fluorescence spectra. *J Appl Crystallogr* **34**:82–6.

Evans P (2006). Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* **62**:72–82.

Faerman C, Ripoll D, Bon S, Le Feuvre Y, Morel N, Massoulié J, Sussman JL, Silman I (1996). Site-directed mutants designed to test back-door hypotheses of acetylcholinesterase function. *FEBS Lett* **386**:65–71.

Fairbanks G, Steck TL, Wallach DF (1971). Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. *Biochemistry* **10**:2606–17.

Falk A, Feys BJ, Frost LN, Jones JD, Daniels MJ, Parker JE (1999). EDS1, an essential component of R gene-mediated disease resistance in Arabidopsis has homology to eukaryotic lipases. *Proc Natl Acad Sci U S A* **96**:3292–7.

Ferré-D'Amaré AR, Burley SK (1994). Use of dynamic light scattering to assess crystallizability of macromolecules and macromolecular assemblies. *Structure* **2**:357–9.

Feys BJ, Moisan LJ, Newman MA, Parker JE (2001). Direct interaction between the Arabidopsis disease resistance signaling proteins, EDS1 and PAD4. *EMBO J* **20**:5400–11.

Feys BJ, Wiermer M, Bhat RA, Moisan LJ, Medina-Escobar N, Neu C, Cabral A, Parker JE (2005). Arabidopsis SENESCENCE-ASSOCIATED GENE101 stabilizes and signals within an ENHANCED DISEASE SUSCEPTIBILITY1 complex in plant innate immunity. *Plant Cell* **17**:2601–13.

Fields S, Song O (1989). A novel genetic system to detect protein-protein interactions. *Nature* **340**:245–6.

Forouhar F, Yang Y, Kumar D, Chen Y, Fridman E, Park SW, Chiang Y, Acton TB, Montelione GT, Pichersky E, Klessig DF, Tong L (2005). Structural and biochemical studies identify tobacco SABP2 as a methyl salicylate esterase and implicate it in plant innate immunity. *Proc Natl Acad Sci U S A* **102**:1773–8.

García AV (2009). *Intracellular dynamics of Arabidopsis EDS1 immune regulatory complexes*. Dissertation, Universität zu Köln.

García AV, Blanvillain-Baufumé S, Huibers RP, Wiermer M, Li G, Gobbato E, Rietz S, Parker JE (2010). Balanced nuclear and cytoplasmic activities of EDS1 are required for a complete plant innate immune response. *PLoS Pathog* **6**:e1000970.

Garman E, Schneider T (1997). Macromolecular Cryocrystallography. *J Appl Crystallogr* **30**:211–37.

Gassmann AJ, Petzold-Maxwell JL, Keweshan RS, Dunbar MW (2011). Field-evolved resistance to Bt maize by western corn rootworm. *PLoS One* **6**:e22629.

Gietz RD, Schiestl RH (1991). Applications of high efficiency lithium acetate transformation of intact yeast cells using single-stranded nucleic acids as carrier. *Yeast* **7**:253–63.

Glazebrook J, Rogers EE, Ausubel FM (1996). Isolation of Arabidopsis mutants with enhanced disease susceptibility by direct screening. *Genetics* **143**:973–82.

Glazebrook J, Zook M, Mert F, Kagan I, Rogers EE, Crute IR, Holub EB, Hammerschmidt R, Ausubel FM (1997). Phytoalexin-deficient mutants of Arabidopsis reveal that PAD4 encodes a regulatory factor and that four PAD genes contribute to downy mildew resistance. *Genetics* **146**:381–92.

Gobbato E (2007). *Analysis of mechanisms underlying EDS1-PAD4 cooperation in Arabidopsis immune signaling*. Dissertation, Universität zu Köln.

Gòmez-Gòmez L, Boller T (2000). FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Mol Cell* **5**:1003–11.

Goodsell DS, Olson AJ (2000). Structural symmetry and protein function. *Annu Rev Biophys Biomol Struct* **29**:105–53.

Gouet P, Courcelle E, Stuart DI, Métoz F (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**:305–8.

Gouhier-Darimont C, Schmiesing A, Bonnet C, Lassueur S, Reymond P (2013). Signalling of Arabidopsis thaliana response to Pieris brassicae eggs shares similarities with PAMP-triggered immunity. *J Exp Bot* **64**:665–74.

Gouy M, Guindon S, Gascuel O (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**:221–4.

Greenfield NJ (2006a). Using circular dichroism collected as a function of temperature to determine the thermodynamics of protein unfolding and binding interactions. *Nat Protoc* **1**:2527–35.

Greenfield NJ (2006b). Using circular dichroism spectra to estimate protein secondary structure. *Nat Protoc* **1**:2876–90.

Grimsley GR, Pace CN (2004). Spectrophotometric determination of protein concentration. *Curr Protoc Protein Sci* **Chapter 3**:Unit 3.1.

Groves MR, Barford D (1999). Topological characteristics of helical repeat proteins. *Curr Opin Struct Biol* **9**:383–9.

Hartley JL, Temple GF, Brasch MA (2000). DNA cloning using in vitro site-specific recombination. *Genome Res* **10**:1788–95.

He Y (2002). A Gene Encoding an Acyl Hydrolase Is Involved in Leaf Senescence in Arabidopsis. *Plant Cell Online* **14**:805–15.

Heidrich K, Wirthmueller L, Tasset C, Pouzet C, Deslandes L, Parker JE (2011). Arabidopsis EDS1 connects pathogen effector recognition to cell compartment-specific immune responses. *Science* **334**:1401–4.

Henzler-Wildman K, Kern D (2007). Dynamic personalities of proteins. *Nature* **450**:964–72.

Heras Bn, Martin JL (2005). Post-crystallization treatments for improving diffraction quality of protein crystals. *Acta Crystallogr D Biol Crystallogr* **61**:1173–80.

Holm L, Rosenström P (2010). Dali server: conservation mapping in 3D. *Nucleic Acids Res* **38**:545–9.

Hope H (1988). Cryocrystallography of biological macromolecules: a generally applicable method. *Acta Crystallogr B* **44 (Pt 1)**:22–6.

Huang B, Schroeder M (2006). LIGSITEcsc: predicting ligand binding sites using the Connolly surface and degree of conservation. *BMC Struct Biol* **6**:19.

Illergård K, Ardell DH, Elofsson A (2009). Structure is three to ten times more conserved than sequence–a study of structural response in protein cores. *Proteins* **77**:499–508.

Ino K, Udagawa I, Iwabata K, Takakusagi Y, Kubota M, Kurosaka K, Arai K, Seki Y, Nogawa M, Tsunoda T, Mizukami F, Taguchi H, Sakaguchi K (2011). Heterogeneous nucleation of protein crystals on fluorinated layered silicate. *PLoS One* **6**:e22582.

Ishida T, Kato S (2004). Role of Asp102 in the catalytic relay system of serine proteases: a theoretical study. *J Am Chem Soc* **126**:7111–8.

Ito H, Fukuda Y, Murata K, Kimura A (1983). Transformation of intact yeast cells treated with alkali cations. *J Bacteriol* **153**:163–8.

Jancarik J, Kim SH (1991). Sparse matrix sampling: a screening method for crystallization of proteins. *J Appl Crystallogr* **24**:409–11.

Jirage D, Tootle TL, Reuber TL, Frost LN, Feys BJ, Parker JE, Ausubel FM, Glazebrook J (1999). Arabidopsis thaliana PAD4 encodes a lipase-like gene that is important for salicylic acid signaling. *Proc Natl Acad Sci U S A* **96**:13583–8.

Jones JDG, Dangl JL (2006). The plant immune system. *Nature* **444**:323–9.

Kabsch W (2010). XDS. *Acta Crystallogr D Biol Crystallogr* **66**:125–32.

Kang C, Kim HJ, Kang D, Jung DY, Suh M (2003). Highly sensitive and simple fluorescence staining of proteins in sodium dodecyl sulfate-polyacrylamide-based gels by using hydrophobic tail-mediated enhancement of fluorescein luminescence. *Electrophoresis* **24**:3297–3304.

Kapust RB, Waugh DS (1999). Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein Sci* **8**:1668–74.

Karplus M, Kuriyan J (2005). Molecular dynamics and protein function. *Proc Natl Acad Sci U S A* **102**:6679–85.

Kaufmann KW, Lemmon GH, Deluca SL, Sheehan JH, Meiler J (2010). Practically useful: what the Rosetta protein modeling suite can do for you. *Biochemistry* **49**:2987–98.

Kelley LA, Sternberg MJE (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nat Protoc* **4**:363–71.

Kelly SM, Jess TJ, Price NC (2005). How to study proteins by circular dichroism. *Biochim Biophys Acta* **1751**:119–39.

Kim TH, Kunz HH, Bhattacharjee S, Hauser F, Park J, Engineer C, Liu A, Ha T, Parker JE, Gassmann W, Schroeder JI (2012). Natural Variation in Small Molecule-Induced TIR-NB-LRR Signaling Induces Root Growth Arrest via EDS1- and PAD4-Complexed R Protein VICTR in Arabidopsis. *Plant Cell* **24**:5177–92.

Koopmann R, Cupelli K, Redecke L, Nass K, Deponte DP, White TA, Stellato F, Rehders D, Liang M, Andreasson J, Aquila A, Bajt S, Barthelmess M, Barty A, Bogan MJ, Bostedt C, Boutet S, Bozek JD, Caleman C, Coppola N, Davidsson J, Doak RB, Ekeberg T, Epp SW, Erk B, Fleckenstein H, Foucar L, Graafsma H, Gumprecht L, Hajdu J, Hampton CY, Hartmann A,

Hartmann R, Hauser G, Hirsemann H, Holl P, Hunter MS, Kassemeyer S, Kirian RA, Lomb L, Maia FRNC, Kimmel N, Martin AV, Messerschmidt M, Reich C, Rolles D, Rudek B, Rudenko A, Schlichting I, Schulz J, Seibert MM, Shoeman RL, Sierra RG, Soltau H, Stern S, Strüder L, Timneanu N, Ullrich J, Wang X, Weidenspointner G, Weierstall U, Williams GJ, Wunderer CB, Fromme P, Spence JCH, Stehle T, Chapman HN, Betzel C, Duszenko M (2012). In vivo protein crystallization opens new routes in structural biology. *Nat Methods* **9**:259–62.

Kossiakoff AA, Spencer SA (1981). Direct determination of the protonation states of aspartic acid-102 and histidine-57 in the tetrahedral intermediate of the serine proteases: neutron structure of trypsin. *Biochemistry* **20**:6462–74.

Krissinel E, Henrick K (2007). Inference of macromolecular assemblies from crystalline state. *J Mol Biol* **372**:774–97.

Krüger DM, Ahmed A, Gohlke H (2012). NMSim web server: integrated approach for normal mode-based geometric simulations of biologically relevant conformational transitions in proteins. *Nucleic Acids Res* **40**:310–6.

Krywka C, Sternemann C, Paulus M, Tolan M, Royer C, Winter R (2008). Effect of osmolytes on pressure-induced unfolding of proteins: a high-pressure SAXS study. *ChemPhysChem* **9**:2809–15.

Kumar D, Klessig DF (2003). High-affinity salicylic acid-binding protein 2 is required for plant innate immunity and has salicylic acid-stimulated lipase activity. *Proc Natl Acad Sci U S A* **100**:16101–6.

Kwon SI, Kim SH, Bhattacharjee S, Noh JJ, Gassmann W (2009). SRFR1, a suppressor of effector-triggered immunity, encodes a conserved tetratricopeptide repeat protein with similarity to transcriptional repressors. *Plant J* **57**:109–19.

Laemmli UK (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**:680–5.

LaVallie ER, DiBlasio EA, Kovacic S, Grant KL, Schendel PF, McCoy JM (1993). A thioredoxin gene fusion expression system that circumvents inclusion body formation in the E. coli cytoplasm. *Nat Biotechnol* **11**:187–93.

Lenfant N, Hotelier T, Velluet E, Bourne Y, Marchot P, Chatonnet A (2013). ESTHER, the database of the α/β-hydrolase fold superfamily of proteins: tools to explore diversity of functions. *Nucleic Acids Res* **41**:423–9.

Lesley SA, Kuhn P, Godzik A, Deacon AM, Mathews I, Kreusch A, Spraggon G, Klock HE, McMullan D, Shin T, Vincent J, Robb A, Brinen LS, Miller MD, McPhillips TM, Miller MA, Scheibe D, Canaves JM, Guda C, Jaroszewski L, Selby TL, Elsliger MA, Wooley J, Taylor SS, Hodgson KO, Wilson IA, Schultz PG, Stevens RC (2002). Structural genomics of the Thermotoga maritima proteome implemented in a high-throughput structure determination pipeline. *Proc Natl Acad Sci U S A* **99**:11664–9.

Lise S, Archambeau C, Pontil M, Jones DT (2009). Prediction of hot spot residues at protein-protein interfaces by combining machine learning and energy-based methods. *BMC Bioinformatics* **10**:365.

Little D, Gouhier-Darimont C, Bruessow F, Reymond P (2007). Oviposition by pierid butterflies triggers defense responses in Arabidopsis. *Plant Physiol* **143**:784–800.

Lo Conte L, Chothia C, Janin J (1999). The atomic structure of protein-protein recognition sites. *J Mol Biol* **285**:2177–98.

Lohkamp B, Dobritzsch D (2008). A mixture of fortunes: the curious determination of the structure of Escherichia coli BL21 Gab protein. *Acta Crystallogr D Biol Crystallogr* **64**:407–15.

Longenecker KL, Lewis ME, Chikumi H, Gutkind JS, Derewenda ZS (2001). Structure of the RGS-like domain from PDZ-RhoGEF: linking heterotrimeric g protein-coupled signaling to Rho GTPases. *Structure* **9**:559–69.

Louis J, Gobbato E, Mondal HA, Feys BJ, Parker JE, Shah J (2012). Discrimination of Arabidopsis PAD4 activities in defense against green peach aphid and pathogens. *Plant Physiol* **158**:1860–72.

Luscombe NM, Thornton JM (2002). Protein-DNA interactions: amino acid conservation and the effects of mutations on binding specificity. *J Mol Biol* **320**:991–1009.

Ma B, Elkayam T, Wolfson H, Nussinov R (2003). Protein-protein interactions: structurally conserved residues distinguish between binding sites and exposed protein surfaces. *Proc Natl Acad Sci U S A* **100**:5772–7.

Maekawa T, Cheng W, Spiridon LN, Töller A, Lukasik E, Saijo Y, Liu P, Shen QH, Micluta MA, Somssich IE, Takken FLW, Petrescu AJ, Chai J, Schulze-Lefert P (2011). Coiled-coil domain-dependent homodimerization of intracellular barley immune receptors defines a minimal functional module for triggering cell death. *Cell Host Microbe* **9**:187–99.

Malthus T (1798). *An Essay on the Principle of Population*. J. Johnson, in St. Paul's Churchyard.

Marillonnet S, Thoeringer C, Kandzia R, Klimyuk V, Gleba Y (2005). Systemic Agrobacterium tumefaciens-mediated transfection of viral replicons for efficient transient expression in plants. *Nat Biotechnol* **23**:718–23.

Mateo A, Mühlenbock P, Rustérucci C, Chang CCC, Miszalski Z, Karpinska B, Parker JE, Mullineaux PM, Karpinski S (2004). LESION SIMULATING DISEASE 1 is required for acclimation to conditions that promote excess excitation energy. *Plant Physiol* **136**:2818–30.

Matthews BW (1968). Solvent content of protein crystals. *J Mol Biol* **33**:491–7.

McCammon JA, Gelin BR, Karplus M (1977). Dynamics of folded proteins. *Nature* **267**:585–90.

McDowell JM, Simon SA (2006). Recent insights into R gene evolution. *Mol Plant Pathol* **7**:437–48.

McPherson A (2004). Introduction to protein crystallization. *Methods* **34**:254–65.

Métraux JP, Signer H, Ryals J, Ward E, Wyss-Benz M, Gaudin J, Raschdorf K, Schmid E, Blum W, Inverardi B (1990). Increase in salicylic Acid at the onset of systemic acquired resistance in cucumber. *Science* **250**:1004–6.

Meyers BC, Kozik A, Griego A, Kuang H, Michelmore RW (2003). Genome-wide analysis of NBS-LRR-encoding genes in Arabidopsis. *Plant Cell* **15**:809–34.

Miller S, Janin J, Lesk AM, Chothia C (1987). Interior and surface of monomeric proteins. *J Mol Biol* **196**:641–56.

Mishina TE, Zeier J (2007). Pathogen-associated molecular pattern recognition rather than development of tissue necrosis contributes to bacterial induction of systemic acquired resistance in Arabidopsis. *Plant J* **50**:500–13.

Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H (1986). Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol* **51 Pt 1**:263–73.

Murase K, Hirano Y, Sun Tp, Hakoshima T (2008). Gibberellin-induced DELLA recognition by the gibberellin receptor GID1. *Nature* **456**:459–63.

Naito K, Taguchi F, Suzuki T, Inagaki Y, Toyoda K, Shiraishi T, Ichinose Y (2008). Amino acid sequence of bacterial microbe-associated molecular pattern flg22 is required for virulence. *Mol Plant Microbe Interact* **21**:1165–74.

Nardini M, Dijkstra BW (1999). Alpha/beta hydrolase fold enzymes: the family keeps growing. *Curr Opin Struct Biol* **9**:732–7.

Narusaka M, Shirasu K, Noutoshi Y, Kubo Y, Shiraishi T, Iwabuchi M, Narusaka Y (2009). RRS1 and RPS4 provide a dual Resistance-gene system against fungal and bacterial pathogens. *Plant J* **60**:218–26.

Newman J (2005). A review of techniques for maximizing diffraction from a protein crystal in stilla. *Acta Crystallogr D Biol Crystallogr* **62**:27–31.

Nicaise V, Roux M, Zipfel C (2009). Recent advances in PAMP-triggered immunity against bacteria: pattern recognition receptors watch over and raise the alarm. *Plant Physiol* **150**:1638–47.

Nishi H, Hashimoto K, Panchenko AR (2011). Phosphorylation in protein-protein binding: effect on stability and function. *Structure* **19**:1807–15.

Nooren IMA, Thornton JM (2003a). Diversity of protein-protein interactions. *EMBO J* **22**:3486–92.

Nooren IMA, Thornton JM (2003b). Structural characterisation and functional significance of transient protein-protein interactions. *J Mol Biol* **325**:991–1018.

Notredame C, Higgins DG, Heringa J (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J Mol Biol* **302**:205–17.

Ntoukakis V, Balmuth AL, Mucyn TS, Gutierrez JR, Jones AME, Rathjen JP (2013). The tomato prf complex is a molecular trap for bacterial effectors based on pto transphosphorylation. *PLoS Pathog* **9**:e1003123.

Nürnberger T, Lipka V (2005). Non-host resistance in plants: new insights into an old phenomenon. *Mol Plant Pathol* **6**:335–45.

Ochsenbein C, Przybyla D, Danon A, Landgraf F, Göbel C, Imboden A, Feussner I, Apel K (2006). The role of EDS1 (enhanced disease susceptibility) during singlet oxygen-mediated stress responses of Arabidopsis. *Plant J* **47**:445–56.

Oerke EC (2006). Crop losses to pests. *J Agric Sci* **144**:31–43.

Ollis DL, Cheah E, Cygler M, Dijkstra B, Frolow F, Franken SM, Harel M, Remington SJ, Silman I, Schrag J (1992). The alpha/beta hydrolase fold. *Protein Eng* **5**:197–211.

Pace CN, Vajdos F, Fee L, Grimsley G, Gray T (1995). How to measure and predict the molar absorption coefficient of a protein. *Protein Sci* **4**:2411–23.

Panjikar S, Parthasarathy V, Lamzin VS, Weiss MS, Tucker PA (2005). Auto-rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr D Biol Crystallogr* **61**:449–57.

Pantoliano MW, Petrella EC, Kwasnoski JD, Lobanov VS, Myslik J, Graf E, Carver T, Asel E, Springer BA, Lane P, Salemme FR (2001). High-density miniaturized thermal shift assays as a general strategy for drug discovery. *J Biomol Screen* **6**:429–40.

Park SW, Kaimoyo E, Kumar D, Mosher S, Klessig DF (2007). Methyl salicylate is a critical mobile signal for plant systemic acquired resistance. *Science* **318**:113–6.

Parker JE, Holub EB, Frost LN, Falk A, Gunn ND, Daniels MJ (1996). Characterization of eds1, a mutation in Arabidopsis suppressing resistance to Peronospora parasitica specified by several different RPP genes. *Plant Cell* **8**:2033–46.

Pathare GR, Nagy I, Bohn S, Unverdorben P, Hubert A, Körner R, Nickell S, Lasker K, Sali A, Tamura T, Nishioka T, Förster F, Baumeister W, Bracher A (2012). The proteasomal subunit Rpn6 is a molecular clamp holding the core and regulatory subcomplexes together. *Proc Natl Acad Sci U S A* **109**:149–54.

Patterson AL (1935). A direct method for the determination of the components of interatomic distances in crystals. *Z Kristallogr* **90**:517–42.

Peart JR, Cook G, Feys BJ, Parker JE, Baulcombe DC (2002). An EDS1 orthologue is required for N-mediated resistance against tobacco mosaic virus. *Plant J* **29**:569–79.

Pegadaraju V, Knepper C, Reese J, Shah J (2005). Premature leaf senescence modulated by the Arabidopsis PHYTOALEXIN DEFICIENT4 gene is associated with defense against the phloem-feeding green peach aphid. *Plant Physiol* **139**:1927–34.

Pegadaraju V, Louis J, Singh V, Reese JC, Bautor J, Feys BJ, Cook G, Parker JE, Shah J (2007). Phloem-based resistance to green peach aphid is controlled by Arabidopsis PHYTOALEXIN DEFICIENT4 without its signaling partner ENHANCED DISEASE SUSCEPTIBILITY1. *Plant J* **52**:332–41.

Petrek M, Otyepka M, Banás P, Kosinová P, Koca J, Damborský J (2006). CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics* **7**:316.

Petsalaki E, Stark A, García-Urdiales E, Russell RB (2009). Accurate prediction of peptide binding sites on protein surfaces. *PLoS Comput Biol* **5**:e1000335.

Porath J, Carlsson J, Olsson I, Belfrage G (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* **258**:598–9.

Price WN, Chen Y, Handelman SK, Neely H, Manor P, Karlin R, Nair R, Liu J, Baran M, Everett J, Tong SN, Forouhar F, Swaminathan SS, Acton T, Xiao R, Luft JR, Lauricella A, DeTitta GT, Rost B, Montelione GT, Hunt JF (2009). Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. *Nat Biotechnol* **27**:51–7.

Rahman A, Stillinger F (1974). Propagation of sound in water. A molecular-dynamics study. *Physical Review A* **10**:368–78.

Reiter B, Glieder A, Talker D, Schwab H (2000). Cloning and characterization of EstC from Burkholderia gladioli, a novel-type esterase related to plant enzymes. *Appl Microbiol Biotechnol* **54**:778–85.

Rhodes G (2006). *Crystallography Made Crystal Clear. A Guide for Users of Macromolecular Models.* Academic Press, 3 ed.

Rietz S, Stamm A, Malonek S, Wagner S, Becker D, Medina-Escobar N, Vlot AC, Feys BJ, Niefind K, Parker JE (2011). Different roles of Enhanced Disease Susceptibility1 (EDS1) bound to and dissociated from Phytoalexin Deficient4 (PAD4) in Arabidopsis immunity. *New Phytol* **191**:107–19.

Risler JL, Delorme MO, Delacroix H, Henaut A (1988). Amino acid substitutions in structurally related proteins. A pattern recognition approach. Determination of a new and efficient scoring matrix. *J Mol Biol* **204**:1019–29.

Robatzek S, Wirthmueller L (2012). Mapping FLS2 function to structure: LRRs, kinase and its working bits. *Protoplasma* .

Rupp B (2009). *Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology*. Taylor & Francis Ltd., 1 ed.

Rushton PJ, Somssich IE, Ringler P, Shen QJ (2010). WRKY transcription factors. *Trends Plant Sci* **15**:247–58.

Rustérucci C, Aviv DH, Holt BF, Dangl JL, Parker JE (2001). The disease resistance signaling components EDS1 and PAD4 are essential regulators of the cell death pathway controlled by LSD1 in Arabidopsis. *Plant Cell* **13**:2211–24.

Rydel TJ, Williams JM, Krieger E, Moshiri F, Stallings WC, Brown SM, Pershing JC, Purcell JP, Alibhai MF (2003). The crystal structure, mutagenesis, and activity studies reveal that patatin is a lipid acyl hydrolase with a Ser-Asp catalytic dyad. *Biochemistry* **42**:6696–708.

Sanson B, Colletier JP, Xu Y, Lang PT, Jiang H, Silman I, Sussman JL, Weik M (2011). Backdoor opening mechanism in acetylcholinesterase based on X-ray crystallography and molecular dynamics simulations. *Protein Sci* **20**:1114–8.

Sarda L, Desnuelle P (1958). [Actions of pancreatic lipase on esters in emulsions]. *Biochim Biophys Acta* **30**:513–21.

Schrödinger, LLC (2010). The PyMOL Molecular Graphics System, Version 1.4.1.

Seeholzer S, Tsuchimatsu T, Jordan T, Bieri S, Pajonk S, Yang W, Jahoor A, Shimizu KK, Keller B, Schulze-Lefert P (2010). Diversity at the Mla powdery mildew resistance locus from cultivated barley reveals sites of positive selection. *Mol Plant Microbe Interact* **23**:497–509.

Sheldrick GM (2010). Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr* **66**:479–85.

Shimada A, Ueguchi-Tanaka M, Nakatsu T, Nakajima M, Naoe Y, Ohmiya H, Kato H, Matsuoka M (2008). Structural basis for gibberellin recognition by its receptor GID1. *Nature* **456**:520–3.

Singh I, Shah K (2012). In silico study of interaction between rice proteins enhanced disease susceptibility 1 and phytoalexin deficient 4, the regulators of salicylic acid signalling pathway. *J Biosci* **37**:563–71.

Sousa R (1995). Use of glycerol, polyols and other protein structure stabilizing agents in protein crystallization. *Acta Crystallogr D Biol Crystallogr* **51**:271–7.

Stevenson CE, Mayer S, Delarbre L, Lawson D (2001). Crystal annealing - nothing to lose. *J Cryst Growth* **232**:629–37.

Straus MR, Rietz S, Ver Loren van Themaat E, Bartsch M, Parker JE (2010). Salicylic acid antagonism of EDS1-driven cell death is important for immune and oxidative stress responses in Arabidopsis. *Plant J* **62**:628–40.

Stura EA, Wilson IA (1991). Applications of the streak seeding technique in protein crystallization. *J Cryst Growth* **110**:270–82.

Su Y, Zhou A, Xia X, Li W, Sun Z (2009). Quantitative prediction of protein-protein binding affinity with a potential of mean force considering volume correction. *Protein Sci* **18**:2550–8.

Sun Tp (2010). Gibberellin-GID1-DELLA: a pivotal regulatory module for plant growth and development. *Plant Physiol* **154**:567–70.

Sussman JL, Harel M, Frolow F, Oefner C, Goldman A, Toker L, Silman I (1991). Atomic structure of acetylcholinesterase from Torpedo californica: a prototypic acetylcholine-binding protein. *Science* **253**:872–9.

Tabashnik BE, Carrière Y, Dennehy TJ, Morin S, Sisterson MS, Roush RT, Shelton AM, Zhao JZ (2003). Insect resistance to transgenic Bt crops: lessons from the laboratory and field. *J Econ Entomol* **96**:1031–8.

Takken FLW, Goverse A (2012). How to build a pathogen detector: structural basis of NB-LRR function. *Curr Opin Plant Biol* **15**:375–84.

Tarr DEK, Alexander HM (2009). TIR-NBS-LRR genes are rare in monocots: evidence from diverse monocot orders. *BMC Research Notes* **2**:197.

Taylor GL (2010). Introduction to phasing. *Acta Crystallogr D Biol Crystallogr* **66**:325–38.

Terwilliger TC (2003). Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr D Biol Crystallogr* **59**:38–44.

Thakur AS, Robin G, Guncar G, Saunders NFW, Newman J, Martin JL, Kobe B (2007). Improved success of sparse matrix protein crystallization screening with heterogeneous nucleating agents. *PLoS One* **2**:e1091.

Towbin H, Staehelin T, Gordon J (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**:4350–4.

Tsuda K, Katagiri F (2010). Comparing signaling mechanisms engaged in pattern-triggered and effector-triggered immunity. *Curr Opin Plant Biol* **13**:459–65.

Tuncbag N, Keskin O, Gursoy A (2010). HotPoint: hot spot prediction server for protein interfaces. *Nucleic Acids Res* **38**:402–6.

Turvey SE, Broide DH (2010). Innate immunity. *J Allergy Clin Immunol* **125**:24–32.

Van der Biezen EA, Jones JD (1998). Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochem Sci* **23**:454–6.

Van der Hoorn RAL, Kamoun S (2008). From Guard to Decoy: a new model for perception of plant pathogen effectors. *Plant Cell* **20**:2009–17.

Van der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005). GROMACS: fast, flexible, and free. *J Comput Chem* **26**:1701–18.

VanEtten HD, Mansfield JW, Bailey JA, Farmer EE (1994). Two Classes of Plant Antibiotics: Phytoalexins versus "Phytoanticipins". *Plant Cell* **6**:1191–2.

Vedadi M, Niesen FH, Allali-Hassani A, Fedorov OY, Finerty Jr PJ, Wasney GA, Yeung R, Arrowsmith C, Ball LJ, Berglund H, Hui R, Marsden BD, Nordlund P, Sundstrom M, Weigelt J, Edwards AM (2006). Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc Natl Acad Sci U S A* **103**:15835–40.

Vera A, González-Montalbán N, Arís A, Villaverde A (2007). The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures. *Biotechnol Bioeng* **96**:1101–6.

Vlot AC, Dempsey DA, Klessig DF (2009). Salicylic Acid, a multifaceted hormone to combat disease. *Annu Rev Phytopathol* **47**:177–206.

Wagner S, Rietz S, Parker JE, Niefind K (2011). Crystallization and preliminary crystallographic analysis of Arabidopsis thaliana EDS1, a key component of plant immunity, in complex with its signalling partner SAG101. *Acta Crystallograph Sect F Struct Biol Cryst Commun* **67**:245–8.

Werner S, Breus O, Symonenko Y, Marillonnet S, Gleba Y (2011). High-level recombinant protein expression in transgenic plants by using a double-inducible viral vector. *Proc Natl Acad Sci U S A* **108**:14061–6.

Wiermer M (2005). *Molecular and spatial characterisation of Arabidopsis EDS1 defence regulatory complexes*. Dissertation, Universität zu Köln.

Wiermer M, Feys BJ, Parker JE (2005). Plant immunity: the EDS1 regulatory node. *Curr Opin Plant Biol* **8**:383–9.

Wine Y, Cohen-Hadar N, Freeman A, Frolow F (2007). Elucidation of the mechanism and end products of glutaraldehyde crosslinking reaction by X-ray structure analysis. *Biotechnol Bioeng* **98**:711–8.

Witte CP, Noël LD, Gielbert J, Parker JE, Romeis T (2004). Rapid one-step protein purification from plant material using the eight-amino acid StrepII epitope. *Plant Mol Biol* **55**:135–47.

Wu Y, Sha B (2006). Crystal structure of yeast mitochondrial outer membrane translocon member Tom70p. *Nat Struct Mol Biol* **13**:589–93.

Xiang T, Zong N, Zhang J, Chen J, Chen M, Zhou JM (2011). BAK1 is not a target of the Pseudomonas syringae effector AvrPto. *Mol Plant Microbe Interact* **24**:100–7.

Xiang T, Zong N, Zou Y, Wu Y, Zhang J, Xing W, Li Y, Tang X, Zhu L, Chai J, Zhou JM (2008). Pseudomonas syringae effector AvrPto blocks innate immunity by targeting receptor kinases. *Curr Biol* **18**:74–80.

Xiao S, Calis O, Patrick E, Zhang G, Charoenwattana P, Muskett P, Parker JE, Turner JG (2005). The atypical resistance gene, RPW8, recruits components of basal defence for powdery mildew resistance in Arabidopsis. *Plant J* **42**:95–110.

Xing W, Zou Y, Liu Q, Liu J, Luo X, Huang Q, Chen S, Zhu L, Bi R, Hao Q, Wu JW, Zhou JM, Chai J (2007). The structural basis for activation of plant immunity by bacterial effector protein AvrPto. *Nature* **449**:243–7.

Zhang Y (2008). I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**:40.

Zhou N (1998). PAD4 Functions Upstream from Salicylic Acid to Control Defense Responses in Arabidopsis. *Plant Cell Online* **10**:1021–30.

Zhu S, Jeong RD, Venugopal SC, Lapchyk L, Navarre D, Kachroo A, Kachroo P (2011). SAG101 forms a ternary complex with EDS1 and PAD4 and is required for resistance signaling against turnip crinkle virus. *PLoS Pathog* **7**:e1002318.

Zhu X, Mitchell JC (2011). KFC2: a knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features. *Proteins* **79**:2671–83.

Zulauf M, D'Arcy A (1992). Light scattering of proteins as a criterion for crystallization. *J Cryst Growth* **122**:102–6.

Zwier MC, Chong LT (2010). Reaching biological timescales with all-atom molecular dynamics simulations. *Curr Opin Pharmacol* **10**:745–52.

Affidavit/Erklärung

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

Köln, 25.04.2013
. . . .
Ort, Datum   Stephan Wagner