STRUCTURE AND FUNCTION OF A GUARDED ARABIDOPSIS IMMUNE SIGNALING UBIQUITIN LIGASE

Dissertation zur Erlangung der Würde des Doktors der Naturwissenschaften

der Fakultät Mathematik, Informatik und Naturwissenschaften,

Fachbereich Biologie

der Universität Hamburg

vorgelegt von

Jan Knop

aus Bad Segeberg

Hamburg, 2019

Vorsitzender der Prüfungskommission:	Jun.-Prof. Wim Walter, PhD
1. Gutacher:	Prof. Dr. Stefan Hoth
2. Gutachterin:	Prof. Dr. Julia Kehr
Datum der Disputation:	27.02.2020

“Alles keine leichten Tänze.”

Hildegard Knop

TABLE OF CONTENTS

ABBREVIATIONS

Å	Ångström
ABA	Abscisic acid
ABA1	ABA DEFICIENT 1
ACS2/6	1-AMINO-CYCLOPROPANE-1-CARBOXYLATE SYNTHASE 2/6
ADP	Adenosine diphosphate
AIF	ATBS1-INTERACTING FACTOR
AKS1	ARABIDOPSIS SKP1 HOMOLOGUE 1
AP2	APETALA 2
APC	ANAPHASE PROMOTING COMPLEX
Arg	Arginine
ARM	ARMADILLO
Asn	Asparagine
Asp	Aspartic acid
ATF2	THIOREDOXIN F2
ATP	Adenosine triphosphate
A. thaliana	*Arabidopsis thaliana*
ATMC1	ARABIDOPSIS THALIANA METACASPASE 1
ATR1	ARABIDOPSIS THALIANA RECOGNIZED 1
AU	Arbitrary units
AZF1	ZINC FINGER PROTEIN 1
BAK1	BRI-ASSOCIATED RECEPTOR KINASE 1
BAP1	BON ASSOCIATION PROTEIN 1
BEE2/3	BR ENHANCED EXPRESSION2/3
bHLH	Basic HELIX-LOOP-HELIX
BiFC	Bimolecular fluorescence complementation
BIK1	BOTRYTIS-INDUCED KINASE 1
BIM1	BES1-INTERACTING MYC-LIKE 1
BIR1/2/3	BAK-INTERACTING RECEPTOR-LIKE KINASE 1/2/3
BON1	BONZAI 1
bp	Base pairs
BR	Brassinosteroid
BTB	BRIC-A-BRAC-TRAMTRACK-BROAD
BZR1	BRASSINAZOLE-RESISTANT 1
C	Control
C_t	Cycle value
C2	PROTEIN KINASE C, CONSERVED REGION 2
Cas9	CRISPR associated protein 9
CC	Coiled-coil
CCS27/52	CELL CYCLE SWITCH PROTEIN 27/52
CD	Circular dichroism
CDC20	CELL DIVISION CYCLE 20
cDNA	Complementary DNA
CDPK	CALCIUM-DEPENDENT PROTEIN KINASES
CHIP	CARBOXYL TERMINUS OF HSP70-INTERACTING PROTEIN
CHS1	CHILLING SENSITIVE 1
CLR	CULLIN-RING ligase
CLSM	Confocal laser scanning microscopy
cm	Centimeter
CML41	CALMODULIN-LIKE 41

CNL	CC-NB-LRR
COI1	CORONATINE-INSENSITIVE 1
Co-IP	Co-immunoprecipitation
COP1	CONSTITUTIVE PHOTOMORPHOGENIC 1
CP	Core protease
CPK5	CALCIUM-DEPENDENT PROTEIN KINASE 5
CPR1	CONSITUTIVE EXPRESSER OF PR GENES 1
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CUL1/3/4	CULLIN 1/3/4
DAMP	Damage-associated molecular pattern
DDB	DNA DAMAGE-BINDING
DEG	Differential expressed gene
DLS	Dynamic light scattering
DNA	Deoxyribonucleic acid
DND	DEFENSE, NO DEATH 1
DUB	Deubiquitinating enzyme
dSTORM	Direct stochastical optical reconstruction microscopy
DUB	Deubiquitinating enzyme
DWD	DDB1 BINDING WD40
E	Amplification efficiency
E1	Ubiquitin-activating enzyme
E2	Ubiquitin-conjugating enzyme
E3	Ubiquitin ligase
E. coli	*Escherichia coli*
EDS1	ENHANCED DISEASE SUSCEPTIBILITY 1
EFR	EF-TU RECEPTOR
EREBP	ETHYLEN-RESPONSIVE ELEMENT BINDING FACTOR
ERF2	ETHYLENE RESPONSE FACTOR 2
ESI	Electrospray ionization
ET	Ethylene
ETI	Effector-triggered immunity
FC	Fold change
FLS2	FLAGELLIN-SENSITIVE 2
FMO1	FLAVIN-DEPENDENT MONOOXYGENASE 1
FPLC	Fast protein liquid chromatography
GA	Gibberellin
gDNA	Genomic DNA
GFP	Green fluorescent protein
GO	Gene ontology
GOI	Gene of interest
GST	Glutathione s-transferase
h	Hour
HBI1	HOMOLOG OF BEE2 INTERACTING WITH IBH1
HECT	HOMOLOGY TO E6-AP C-TERMINUS
HR	Hypersensitive response
IC	Internal calibrator
ICS1	ISOCHORISMATE SYNTHASE 1
IPTG	Isopropy-β-D-thiogalactopyranoside
JA	Jasmonic acid
JAZ	Jasmonate ZIM domain
kb	Kilobases
KD	Kinase domain
KEGG	*Kyoto Encyclopedia of Genes and Genomes*
LMB	Leptomycin B
LRR	Leucin-rich repeats

LYK5	LYSIN MOTIF RECEPTOR KINASE 5
Lys	Lysin
m	Marker
mA	Milliampere
MAPK	Mitogen-activated protein kinase
MAPKKK3/5	MAPK KINASE KINASE 3/5
MBP1	MYROSINASE-BINDING PROTEIN 1
min	Minute
MKK4/5	MAPK KINASE 4/5
ml	Milliliter
mm	Millimeter
MPK3/6	MITOGEN-ACTIVATED PROTEIN KINASE 3/6
MS	Mass spectrometry
MT	Mutant
MUB	Membrane-anchored ubiquitin-fold
MVB	Multi-vesicular body
MW	Molecular weight
MWCO	Molecular weight cut-off
N	Negative control
NAC	NAM, ATAF1,2, CUC2
NB	Nucleotide-binding
NCRQ	Normalized calibrated relative quantities
NDHA	NADH DEHYDROGENASE
NDR1	NON RACE-SPECIFIC DISEASE RESISTANCE 1
NGS	Next-generation sequencing
NHP	N-hydroxypipecolic acid
NLR	Nucleotide-binding leucine-rich repeat protein
NO	Nitric oxide
NRG1	N REQUIREMENT GENE 1
NSD	Normalized spatial discrepancy
ORA59	OCTADECANOID-RESPONSIVE ARABIDOPSIS AP2/ERF 59
OXI1	OXIDATIVE SIGNAL-INDUCIBLE 1
PAD4	PHYTOALEXIN DEFICIENT 4
PAGE	Polyacrylamide gel electrophoresis
PAMP	Pathogen-associated molecular pattern
PBL2	PBS1-LIKE 2
PBS1	AVRPPHB SUSCEPTIBLE 1
PCA	Principle component analysis
PCR	Polymerase chain reaction
PDB	*Protein data bank*
$PI(4,5)P_2$	Phosphatidylinositol 4,5-bisphosphat
PKC	PROTEIN KINASE C
PTI	Pattern-triggered immunity
Pto	*Pseudomonas syringae* pv. *tomato*
PR1	PATHOGENESES-RELATED 1
PRR	Pattern recognition receptor
PUB	Plant U-box type E3 ligase
qRT-PCR	Quantitative real-time PCR
R_g	Radius of gyration
RBOHD	RESPIRATORY BURST OXIDASE HOMOLOG D
REM19	REPRODUCTIVE MERISTIM 19
RIN4	RPM1-INTERACTING PROTEIN 4
RING	REALLY INTERESTING NEW GENE
PMR4	POWDERY MILDEW RESISTANT 4
RNA	Ribonucleic acid

RNA-seq	RNA sequencing
ROS	Reactive oxygen species
RP	Regulatory particle
RPM1	RESISTANCE TO PSEUDOMONAS SYRINGAE PV. MACULICOLA 1
RPP1	RECOGNITION OF PERONOSPORA PARASITICA 1
RPS2/4	RESISTANT TO PSEUDOMONAS SYRINGAE 2/4
RRS1	RESISTANT TO RALSTONIA SOLANACEARUM 1
RT	Room temperature
RT-PCR	Reverse transcription PCR
RTV1	RELATED TO VERNALIZATION 1
s	Second
SA	Salicylic acid
SAC51	SUPPRESSOR OF ACAULIS 51
SAG101	SENESCENCE-ASSOCIATED GENE 101
SAR	Systemic acquired resistance
SAUL1	SENESCENCE-ASSOCIATED E3 UBIQUITIN LIGASE 1
SAXS	Small-angle X-ray scattering
SCF	S PHASE KINASE-ASSOCIATED PROTEIN 1-CULLIN 1-F-BOX
SDS	Sodium dodecyl sulfate
SEC	Size-exclusion chromatography
Ser	Serine
SIRK	SENESCENCE-INDUCED RECEPTOR-LIKE KINASE
SNC1	SUPPRESSOR OF *npr1-1*, CONSTITUTIVE 1
SOC3	SUPRESSOR OF *chs1-2, 3*
SUMO	SMALL UBIQUITIN-LIKE MODIFIER
SZF1/3	SALT-INDUCIBLE ZINC FINGER1/3
TCP8	TCP DOMAIN PROTEIN 8
TF	Transcription factor
TGN	Trans-Golgi network
Thr	Threonine
TIR	TOLL-INTERLEUKIN-1 RECEPTOR
TNL	TIR-NB-LRR
TPR1	TOPLESS-RELATED 1
TRX5	THIOREDOXIN H-TYPE 5
TSS	Transcription start site
TTS	Transcription termination site
U	Units
UBC	Ubiquitin-conjugating enzyme
UBE2D/G	UBIQUITIN-CONJUGATING ENZYME E2 D/G
UBP26	UBIQUITIN BINDING PROTEIN 26
UND	U-box N-terminal domain
V	Volt
VWA	Von Willebrand factor, type A domain
vWF	Von Willebrand factor
WT	Wild type
Y2H	Yeast two-hybrid
YFP	Yellow fluorescent protein
ZAR1	ZYGOTIC ARREST 1
ZAT7	ZINC FINGER OF ARABIDOPSIS THALIANA 7
ZED1	HOPZ-ETI-DEFICIENT 1

1 INTRODUCTION

Although being toxic to many microbial organisms, life as it is known today would not be possible without oxygen. Green plants are the main producers of oxygen and provide the foundations for most life forms. In addition, they supply life's energy basis, because they are the major source of biomass. This can take place either directly in case of herbivores or indirectly in case of carnivorous organisms. Thus, understanding how plants work and interact in a currently changing environment is of high importance.

Due to the existence of a large variety of plant species experiments have to be performed with so-called model organisms, so that insights can be transferred to other species. In plant sciences, *Arabidopsis thaliana* (*A. thaliana*, or thale cress) is one of the organisms of choice. It was established as a model organism in 1995 and since the publication of its genome in 2000 many groundbreaking discoveries have been made using this organism [1]. These range from the molecular principles of photosynthesis to the complexities of plant hormone function and the elucidation of many molecular details of the two-sided plant immune system with its systemic signaling. Its significance is best illustrated by citations of *A. thaliana* papers more and more cited outside the *A. thaliana* community, thus highlighting the importance of *A. thaliana* as a model organism for plant research [2].

Physiological changes and interactions with other organisms are always based on molecular mechanisms, as these form the foundations of everything. Therefore, to gain insights into how plants work molecular details have to be analyzed by applying numerous disciplines such as biochemistry, genetics, structural biology and bioinformatics. This allows to understand what holds life together.

1.1 THE PLANT IMMUNE SYSTEM

Like every other organism, plants are exposed to a diverse range of external factors. Many factors are not beneficial to the organism. Thus, evolution lead to the development of different mechanisms to cope with these factors. In general, external factors can be divided into two categories: (i) abiotic factors such as the supply of nutrients and water or the amount or quality of light, and (ii) biotic factors that involve the interaction with organisms such as bacteria, nematodes or viruses. The latter can be subdivided into beneficial microbes or pathogens that can lead to disease and threaten plant survival. Coping with the latter is of high importance to the plant to sustain its fitness. Pathogens always try to evade plant defense mechanisms, and thus a constant arms race fighting for the upper hand exists between plants and pathogens. Accordingly, evolution resulted in a plant immune system that allows to deal with potential pathogenic organisms and to generally prevent disease [3,4].

The immune system of plants is quite unique. Unlike mammals, plants lack mobile defender cells and a somatic adaptive immune system [3]. Thus, the defense response is only based on the innate immunity of each singular plant cell and the systemic signals, which propagate the information to distant regions [5]. This is in particular remarkable, since plants have to cope with a very broad range of pathogens, which are quite diverse concerning their infection strategies. Pathogens can either proliferate in the apoplast or directly enter the cell. They are able to secret molecules antagonizing the immune system or invaginate haustoria to directly feed on cell compartments [3]. Nonetheless, the plant immune system may appear limited, but the plant is still able to defend itself against most of these different pathogens with great success. On a very basic level, the immune response can be divided into two different mechanisms.

(i) On the one hand pattern recognition receptors (PRRs) in the plasma membrane can recognize pathogen-associated molecular patterns (PAMPs) (Figure 1). To enable this recognition PRRs structurally consist of their intercellular (apoplastic)-facing recognition domain and cytosol-facing signal transmitting domain (Figure 1). Recognition can be achieved by multiple extracellular domains, for instance leucin-rich repeats (LRRs) [6]. PAMPs are thought to be common patterns, which can be found in a broad variety of pathogens. For instance, flg22 is an elicitor-active epitope of bacterial flagellin [6]. In addition, PRRs can respond to damage-associated molecular patterns (DAMPs) that are being released upon cell damage due to an infection [7]. Currently, only 16 PRRs are well described in *A. thaliana* highlighting the general spectrum of PAMP recognition [8].

Upon PAMP recognition receptor heterodimers, -trimers or -tetramers are formed, leading to an activation of the immune response (Figure 1) [7]. Some PRRs belong to the protein family of receptor kinases harboring an intracellular kinase domain. Thus, due to multimerization the kinase domains are phosphorylated and associated receptor like cytoplasmic kinases (RLCKs) are activated and dissociate to propagate the immune signal further downstream (Figure 1) [7]. As a result, the immune response is initiated (Figure 1). This pattern-triggered immunity (PTI) leads subsequently to an enhancement of basal immunity by activating local and systemic defense responses [8].

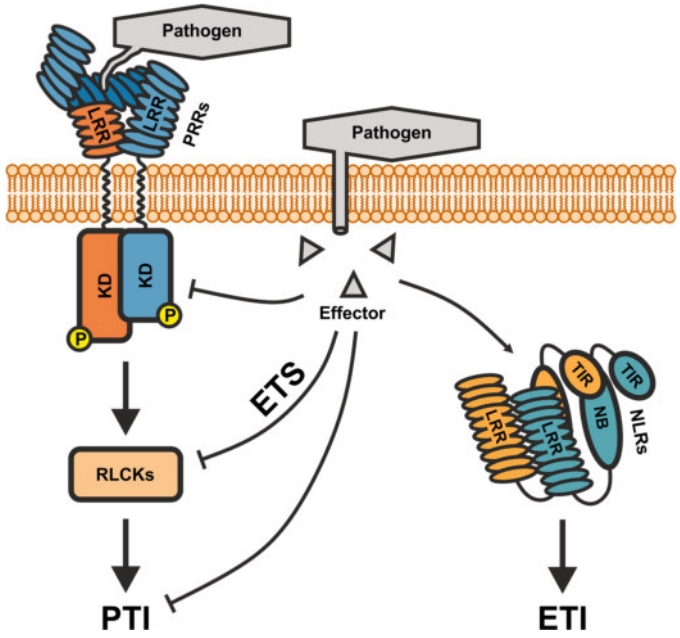


Figure 1: Schematic representation of the plant immune system. Upon PAMP recognition a PRR heterodimer is formed. The recognition domain consists in this case of LRRs. Upon activation RLCKs dissociate from the PRRs and induce the PTI. This can be targeted and blocked by effectors, which are secreted by pathogens into the cell. Furthermore, effectors and changes mediated by them can be recognized by NLRs which undergo a conformational change leading to the onset of ETI. In this example, NLRs contain N-terminal TIR domains.

(ii) On the other hand, plants harbor another class of receptors, because PRRs are thought to recognize only general patterns, but plants have to counteract a huge and diverse range of pathogens. Particularly, since certain pathogens can insert pathogenic virulence molecules, so-called effectors into host cells that interfere with PTI and reduce basal immunity resulting in effector-triggered susceptibility (ETS) [3]. These effectors and their inhibitory effects are recognized by this additional class of receptors. They are intracellular nucleotide-binding domain leucin-rich repeat proteins (NLRs) (Figure 1). There are 207 potential NLR family members in *A. thaliana* [9]. They contain a N-terminal dimerization domain, which can be either a coiled-coil (CC) motif or a domain containing a homology to TOLL-INTERLEUKIN-1 RECEPTOR (TIR) cytoplasmic domains. The central nucleotide-binding (NB) domain is followed by C-terminal LRRs. Thus, two functional classes of NLRs exist. On the one hand CC-NB-LRR (CNL) and on the other TIR-NB-LRR (TNL) NLRs. They are known to be functional as homo- or heterodimers (Figure 1) [10]. Heterodimers are thought to widen the detection spectrum. Such heterodimers are often formed between truncated NLRs that lack either the LRR or the N-terminal domain and these NLRs containing all domains [4,9,11].

Upon recognition, NLRs undergo a steric change from the closed state to an open adenosine triphosphate (ATP)-bound state. This renders new sites accessible to various signaling molecules and the immune response is propagated through the cell [12]. In contrast to PTI, this effector-triggered immunity (ETI) is often associated with localized cell lesions being caused by cell death [4]. In addition, ETI is thought to enhance basal defense responses and to be more prolonged and stronger compared to PTI [4]. Nonetheless, gaining more knowledge concerning PTI and ETI led to the assumption that both pathways are not as discrete as it has been hypothesized [13]. This assumption, how these signaling pathways actually work in detail and what components are involved will be discussed in the following chapters.

1.1.1 PATTERN-TRIGGERED IMMUNITY (PTI)

As mentioned above, PTI is characterized mainly by the recognition of PAMPs or DAMPs at the plasma membrane by PRRs. The hallmarks of PTI following activation of PRRs are Ca^{2+} signaling, production of reactive oxygen species (ROS), activation of calcium-dependent protein kinase (CDPK), activation of mitogen-activated protein kinase (MAPK) cascades, ethylene (ET), jasmonic acid (JA), and salicylic acid (SA) signaling, callose deposition and phytoalexin production [14]. PTI results in a basal local defense response, which can be propagated to distant cells to prime the defense systems for a possible attack. This long distance response is called systemic acquired resistance (SAR) [15].

PRR activation

PRRs can generally be divided into receptor kinases carrying an intracellular kinase domain and receptor proteins lacking the intracellular kinase domain. The kinase domain is essential for downstream signaling. Regulatory proteins lacking this domain, have been shown to dimerize with receptor kinases. The latter can further be subdivided by their different ectodomains for recognition of PAMPs/DAMPs. In addition to LRRs, they can contain a lysin-motif domain. For example, the lysin-motif domain receptor kinases CERK1 and LYK5 (LYSIN MOTIF RECEPTOR KINASE 5) form a heterodimer in *A. thaliana* for chitin recognition [7]. In addition, PRRs can consist of a lectin, an EGF-like or a lectin-like ectodomain [8]. In the absence of any PAMPs/DAMPs, PRRs reside in a resting state by keeping the kinase domain inactive. In case of the PRR FLS2 (FLAGELLIN-SENSITIVE 2) its binding to BAK1 (BRI1-ASSOCIATED RECEPTOR KINASE), which is mediated by binding to BIR2/3 (BAK1-INTERACTING RECEPTOR-LIKE KINASE 2 and 3), renders FLS2 inactive (Figure 2A) [16]. Mimicking PAMP recognition by application of the flg22 peptide, which is part of the bacterial flagellin, leads to the release of the FLS2/BAK1 heterodimer followed by autophosphorylation [7]. BIK1 (BOTRYTIS-INDUCED KINASE 1), a RLCK normally associating with FLS2, dissociates upon autophosphorylation to induce several downstream regulatory

processes (Figure 2B) [17]. The exact mechanism of FLS2 activation was determined by crystallography by analyzing the recombinantly expressed ectodomains of FLS2 and BAK1 with bound flg22 [18]. Thus, structural biology gave important insights into the exact mechanisms underlying the activation of the immune response.

Ca^{2+} and ROS signaling

Following PAMP recognition, a Ca^{2+} influx, resulting in membrane depolymerization and activation of multiple defense pathways, is subsequently initiated (Figure 2B). In case of flg22-perception this is mediated by the RLCK BIK1 right after receptor formation (Figure 2B) [19]. The PTI-associated Ca^{2+} influx is known to result in a singular peak in the cytosolic free Ca^{2+} concentration. The duration and intensity of this peak was shown to vary, depending on the kinds of PRRs being activated. Thus, a characteristic Ca^{2+} pattern has been proposed to modulate downstream signaling [20]. Immediately after the Ca^{2+} influx, an oxidative burst takes place in the apoplastic space. This is caused by a phosphorylation and subsequent activation of the intermembrane RBOHD (RESPIRATORY BURST OXIDASE HOMOLOG D) by BIK1 (Figure 2B) [21]. RBOHD is the main producer of ROS in plants. Although its phosphorylation being the main regulatory step RBOHD is known to be additionally activated by Ca^{2+} binding to its EF-hand motif resulting in a conformational change (Figure 2B). Thus, the activity of RBOHD is enhanced by Ca^{2+} influx as well [22]. Upon activation RBOHD generates superoxide (O_2^-) in the apoplast that is transformed into hydrogen peroxide (H_2O_2) due to a superoxide dismutase (Figure 2B) [23,24]. This results in an oxidative burst.

This regulation by ROS influences plant immunity in two different ways. On the one hand, ROS are highly reactive and thus directly toxic to pathogens by supposedly damaging their membranes (Figure 2B) [25,26]. On the other hand, ROS have been shown to be the main mediators propagating the defense signal from the local infection site to distant cells [27]. Consequently, a higher fitness in adjacent tissue is achieved and a potential secondary infection is inhibited or dampened [28]. At a distant cell or at the local infection site ROS can traverse membranes by being transported through aquaporins (Figure 2B) [29]. In the cytosol ROS directly activate kinases like OXI1 (OXIDATIVE SIGNAL-INDUCIBLE 1). These kinases initiate MAPK cascades resulting in an onset of downstream immune signals [26]. Additionally, Ca^{2+} channels in the plasma membrane and at chloroplasts are activated by ROS (Figure 2B) resulting in a Ca^{2+} wave expanding from the infection site to distal regions [26,30].

CDPK activation

Due to the Ca^{2+} peak in local or distant cells CDPKs are activated. The best described is CPK5 (CALCIUM-DEPENDENT PROTEIN KINASE 5), which was shown to be highly important for immune signal propagation [31]. Upon activation, CPK5 phosphorylates RBOHD, resulting in the activation of RBOHD and subsequent ROS production. Thus, CPK5 mediates the interplay between ROS and Ca^{2+}, resulting in SAR [32]. Nevertheless, additional CDPKs are activated by the Ca^{2+} influx as well. CPK4/5/6/11 are such CDPKs and have been reported to act as essential signaling hubs for PTI and PAMP-induced signal propagation (Figure 2B) [33]. Their function has been demonstrated to rely on phosphorylation and subsequent activation of downstream transcription factors (TF).

In case of CPK4/11 the TF WRKY28 is phosphorylated leading to its binding of the TF TCP8 (TCP DOMAIN PROTEIN 8). Both TF form a complex that binds to the promoter of *ICS1* (*Isochorismate Synthase 1*), resulting in elevated *ICS1* transcription levels (Figure 2B) [34–36]. ICS1 is active in the chloroplasts and is the main producer of salicylic acid (SA), which is the predominant regulatory hormone of the defense response against biotrophic pathogens [35]. Consequently, SA is produced and transported into the cytosol (Figure 2B).

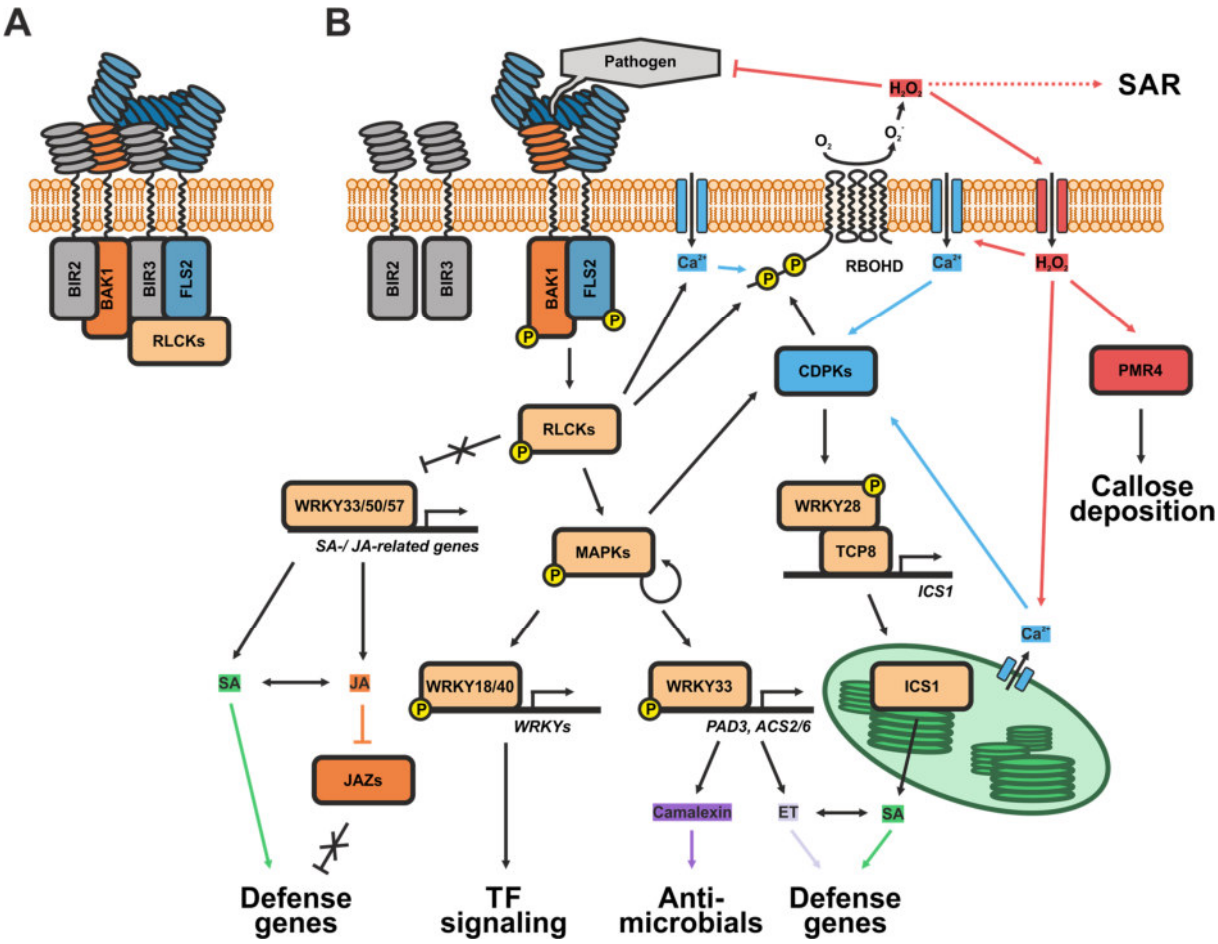


Figure 2: Pattern-triggered immunity model. (A) In the absence of pathogens PRRs reside in a heteromeric complex inhibiting their activation [7]. **(B)** Upon activation a heterodimer is formed, RLCKs (such as BIK1)dissociate and activate a Ca^{2+} influx [19]. Furthermore, the RLCKs and subsequent activated CDPKs (like CPK5) activate RBOHD leading to an oxidative burst [23,24]. CDPKs activate as well downstream WRKY transcription factors resulting in the increased production of SA and the translation of SA-dependent genes [34–36]. The oxidative burst also leads to the initiation of callose deposition (by activating PMR4) [37]. Further downstream a MAPK cascade is initiated, camalexin and ET are synthesized and further WRKYs are expressed leading to an onset of the defense response [38]. RLCK activation results as well in SA and JA synthesis and subsequent hormone-dependent gene expression [39].

RLCK signaling and MAPK cascades

In contrast to CDPKs, RLCKs are directly activated by phosphorylation by PRRs (Figure 2B). These kinases are involved both in positive and the negative regulation of the immune response. An example for such a dual function is the RLCK BIK1. On the one hand, BIK1 was shown to inhibit SA and JA synthesis in uninfected plants by repressing the TFs WRKY33/50/57 (Figure 2B) [39]. On the other hand, upon infection, BIK1 and other RLCKs phosphorylate and thus activate members of the MAPKKK (MAPK kinase kinase) family. This leads to the initiation of MAPK cascades and subsequent downstream signaling resulting in an elevated immune response.

In case of FLS2 activation, MAPKKK5 gets activated and phosphorylates several MKKs (MAPK kinases), resulting in their activation. These MKKs phosphorylate and thus activate MPKs (MITOGEN-ACTIVATED PROTEIN KINASEs) in a subsequent step, which further propagates the immune signal [17]. Thus, a MAPK cascade is initiated upon PAMP perception (Figure 2B). Which MPKs are exactly activated during PTI has been reported to depend on the PRRs, which initiate the immune response [7]. Downstream signaling of the MAPK cascade is known to rely on multiple mechanisms.

In one case, MPKs phosphorylate TFs and thus directly influence downstream expression of defense-related genes. The earliest ones that are phosphorylated and thus activated are WRKY18, WRKY33 and WRKY40 [40]. In particular WRKY18 and WRKY40 are known to bind to the promoters of further *WRKY* genes enabling their expression and therefore amplifying defense signaling [38]. In case of WRKY33 expression of *PAD3* (*Phytoalexin Deficient 3*) and *ACS2/6* (*1-Amino-Cyclopropane-1-Carboxylate Synthase 2* and *6*) genes is up-regulated (Figure 2B) [41,42]. Their expression has been reported to lead to camalexin and ET synthesis, since they are both involved in their synthesis [40,43]. In addition, MAPKs directly effect CDPKs activity and are thus also involved in Ca^{2+}-mediated signaling. MPK3/4/6 indirectly phosphorylate and subsequently activate CPK5/6 (Figure 2B) [44]. This could explain why CDPKs have been reported to be Ca^{2+}-independent in case of a strong MAPK cascade activation [45].

Hormone-mediated signaling

As already stated, the different pathways that are activated upon PAMP or DAMP perception result, among other mechanisms, in hormone synthesis. The main immune-related plant hormones are SA, JA and ET [3].

SA, which is synthesized by ICS1 in the chloroplast, is transferred into the cytosol. There, SA activates NPR1 (NONEXPRESSER OF PR GENES 1), which can interact with the TF family of TGAs (TGACG SEQUENCE-SPECIFIC BINDING PROTEINs) and bind to promoter sequences of SA-dependent genes (Figure 2B). Subsequently, defense related genes like *PR1* (*Pathogenesis-Related Gene 1*) and *WRKY70* are expressed [46]. SA has mainly been associated with defense responses against bacteria or biotrophic pathogens [47]. In contrast to SA, JA and ET have frequently been reported to be the main driving plant hormones in case of necrotrophic pathogens and to act antagonistically to SA [48]. Nonetheless, JA, ET and SA were all found to be important regulators of the early onset of PTI in case of flg22 treatment and other non-necrotrophic associated PAMPs. Accordingly, JA, ET and SA have been hypothesized to balance each other resulting in a more robust defense response [49]. Ja-dependent genes are in JA absence repressed by jasmonate ZIM domain (JAZ) proteins (Figure 2B), which bind to COI1 (CORONATINE-INSENSITIVE 1) and the key transcriptional activators of JA signaling. Upon JA-perception JAZ proteins are polyubiquitinated by COI1 and subsequently degraded, releasing transcriptional activators and initiating transcription of downstream genes [50].

Callose deposition and phytoalexin-mediated defense

As mentioned before, callose deposition in the cell wall is as well one of the hallmarks of PTI, providing a physical barrier. This barrier has been found to be an essential part of the defense response, especially during fungal infections, guaranteeing a successful protection against the pathogen [25,51]. Secondary cell wall formation has been reported to mainly rely on the callose synthase PMR4 (POWDERY MILDEW RESISTANT 4) [52,53]. Its function being highly dependent on RBOHD is thus directly controlled by ROS, highlighting again the high importance of ROS during the defense response (Figure 2B) [37]. Another dependency has been shown to be modulated by ethylene. Enzymes acting upstream of callose synthesis were shown to depend on ethylene production [54]. Interestingly, callose deposition has been reported to be highly PAMP-dependent in its strength [55].

Phytoalexins, like camalexin, act as antimicrobial substances, especially against necrotrophic pathogens. They are synthesized in a PAD3-dependent manner (Figure 2B) and lead to membrane disruption of the pathogens [56,57]. In summary, PRR activation induces a complex immune response demonstrating how variable PTI is regulated.

1.1.2 EFFECTOR-TRIGGERED IMMUNITY (ETI)

ETI is the counteraction against pathogenic effectors to prevent ETS. The E3 ubiquitin ligase AvrPtoB from *Pseudomonas syringae* (*P. syringae*) for instance is such an effector. AvrPtoB has been reported to inhibit the PRR FLS2 by polyubiquitination leading to degradation of FLS2 via the 26S proteasome [58,59]. In addition, effectors can act further downstream in the PTI-signaling cascade. AvrPphB from *P. syringae* has been shown to target the RLCKs PBS1 (AVRPPHB SUSCEPTIBLE 1) and BIK1 acting directly downstream of FLS2. In this case, AvrPphB acts as a protease that cleaves its targets [60]. Effectors are even known to target multiple compartments as in case of AvrPtoB, which in addition to FLS2 targets immunity-related parts of the vesicle transport machinery such as EXO70B1 [61]. To counteract these effector-mediated mechanisms ETI is thought to be a stronger and more prolonged kind of immune response, subsequently leading to the inhibition of pathogen growth [3]. This can be accompanied by cell death known in case of ETI as a hypersensitive response (HR) [4].

NLR activation

The first step in ETI is the activation of NLRs through changes in their conformation involving an exchange of bound adenosine diphosphate (ADP) for ATP. This change results in an accessible NB and C-terminal signaling domain. The dimerization of NLRs represents an essential step in this mechanism [4]. NLRs can act as homodimers, but heterodimers and -tetramers have been detected as well [62]. Even pairings with truncated NLRs, lacking the LRR domain, were reported. Truncated NLRs are thought to be dysfunctional, so the paired NLR is assumed to take over the role of the signal transducer [11]. The NLR activation is known to result from effector recognition. Effectors can be recognized by three distinct mechanisms.

(i) The first mechanism is a direct interaction of an effector such as ATR1 (*Arabidopsis thaliana* RECOGNIZED 1) by an NLR. ATR1 is an effector of the downey mildew *Hyaloperonospora arabidopsidis*. Upon secretion into the cell ATR1 is bound by the LRR domain of the NLR RPP1 (RECOGNITION OF *Peronospora parasitica* 1) (Figure 3A). This binding causes conformational changes of the RPP1 TIR domain, rendering the NB domain accessible, thus leading to homodimerization, ATP binding and therefore activation (Figure 3A) [63]. Because direct interactions between NLRs and effectors are thought to be an exception, two more common and similar mechanisms have been discovered.

(ii) The second mechanism follows the so-called guard model. Effectors are known to directly target components of the plant immune system to render the plant more susceptible. The target components are often guarded by NLRs and are therefore referred to as guardees. RPM1 (RESISTANCE TO *Pseudomonas syringae pv. maculicola* 1) is known to guard RIN4 (RPM1-INTERACTING PROTEIN 4), a plasma membrane localized negative regulator of PTI (Figure 3B). RIN4 is known to be targeted by AvrRpm1, an effector of *P. syringae*. An interaction between both proteins leads to phosphorylation of RIN4 resulting in its activation and thus suppression of PTI. This is recognized by RPM1, which is thereby activated (Figure 3B) [64,65]. Interestingly, RIN4 is guarded twofold. The interaction with the effector AvrRpt2 cleaving RIN4 is also recognized by the CNL RPS2 (RESISTANT TO *Pseudomonas* syringae 2). Guarding by more than one NLR has been observed for other guardees as well [62].

(iii) The third mechanism for perception of pathogenic effectors by NLRs follows the so-called decoy model. Not all proteins, which are known to be targeted by effectors, possess a function in the defense response. These proteins are thought to act as decoys and function as alarm signals [66]. A common example for an integrated decoy is the heterotetramer of RRS1 (RESISTANT TO *Ralstonia solanacearum* 1) and RPS4 (RESISTANT TO *Pseudomonas syringae* 4). RRS1 contains an additional C-terminal WRKY domain, which directly binds the effector AvrPS4 (Figure 3C). This domain only has the function to bind effectors. Binding leads to a conformational change of the TIR domain, activation of the heterotetramer and thus activation of ETI (Figure 3C) [67]. A direct

decoy model exists in case of the NLR ZAR1 (ZYGOTIC ARREST 1) that guards PBL2 (PBS1-LIKE 2), which is a pseudokinase without any apparent function [68]. Thus, PBL2 only acts as a decoy, too.

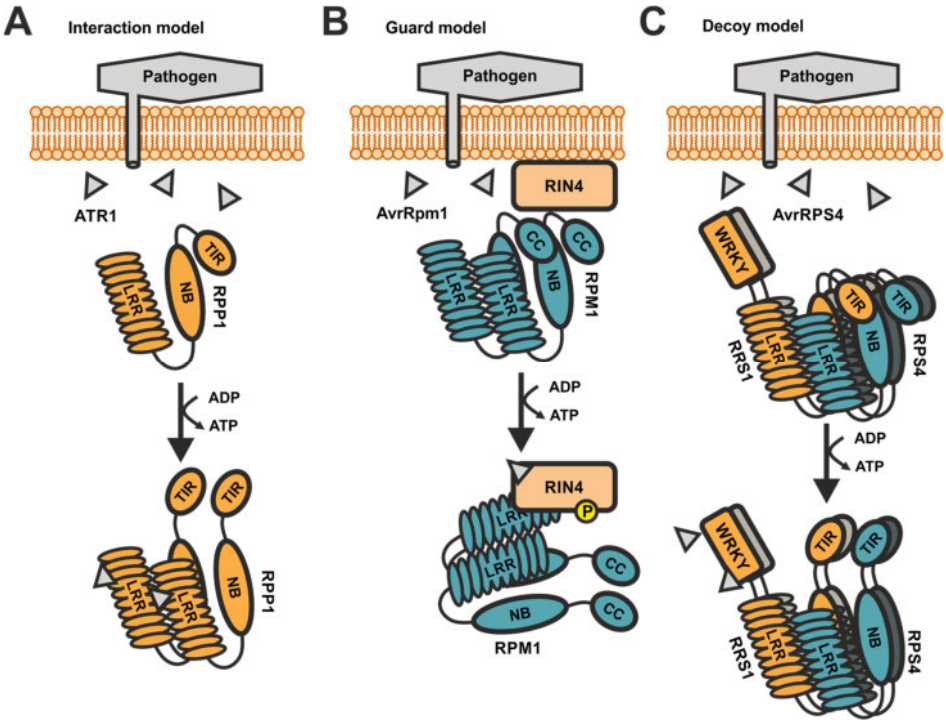


Figure 3: Different modes of NLR effector recognition. Different mechanisms exist, by which NLRs can recognize effectors. **(A)** RPP1 recognizes the effector ATR1 by direct binding at its LRR domains. Activation leads to homodimerization [63]. **(B)** RPM1 guards RIN4, which is bound to the plasma membrane, as a homodimer. AvrRpm1 binding to RIN4 results in phosphorylation of RIN4. This is recognized by RPM1, which gets activated [64,65]. **(C)** In case of RRS1, which forms a heterotetramer with RPS4, an additional WRKY domain functions as an integrated decoy. Upon AvrRPS4 binding to the WRKY domain the NLR complex is activated [67]. The decoy does not have to be a part of the NLR, but can be an otherwise non-functional protein.

Downstream signaling

Following NLR activation, an initial Ca^{2+} influx is necessary for ETI induction. In contrast to PRR-dependent immune signaling, the regulatory mechanisms that lead to a higher Ca^{2+} concentration in ETI are not known. Nonetheless, the main difference between PTI and ETI is a second Ca^{2+} peak. In contrast to PTI, the first Ca^{2+} peak in ETI, which occurs 8-12 min after induction, is followed by a second peak after 105-137 min. This second peak has been shown to be essential for ETI and subsequent HR [69].

After activation, the defense signal is transmitted through the cell by signaling cascades. Although a vast number of downstream regulators and mechanisms are known, the exact mechanisms connecting the activated NLR to these proteins is still elusive [62]. Nevertheless, regulations taking place further downstream have been studied intensively. Even though central hubs for downstream regulation differ depending on the N-terminal NLR domain, some common signaling pathways can be found for both classes. Upon activation, some NLRs localize to the nucleus. In the nucleus, these NLRs interact directly with TFs, leading to a change in gene expression (Figure 4) [4]. One of the most prominent examples is SNC1 (SUPPRESSOR OF *npr1-1*, CONSTITUTIVE 1), which upon activation binds to the corepressor TPR1 (TOPLESS-RELATED 1). As a result, TPR1 is impaired and its target genes *DND1/2* (*Defense, no Death 1* and *2*) are up-regulated and act as positive defense regulators [70]. Apart from these NLRs downstream signaling in ETI normally depends on the type of NLR. Whereas CNL signaling depends on NDR1 (NON RACE-SPECIFIC DISEASE RESISTANCE 1), TNL signaling requires the function of EDS1 (ENHANCED DISEASE

SUSCEPTIBILITY 1) [4]. Although CNLs and TNLs appear to rely on separate pathways, they share multiple downstream signaling components and are strongly intertwined.

CNL-initiated signaling

In case of CNL activation, NDR1 is the main regulatory hub. NDR1 is an integrin-like protein in the plasma membrane that anchors RIN4 to the plasma membrane (Figure 4). RIN4 has been shown to be guarded by the CNLs RPM1 and RPS2. Nonetheless, additional NDR1-dependent CNLs exist. Upon CNL activation, a Ca^{2+} influx into the cytosol is detectable leading to activation of several CDPKs. As a consequence of CDPK-activation several WRKY TFs are phosphorylated and transcription levels of defense associated genes like *ICS1* are elevated. Consequently, SA is generated [71]. These mechanisms are quite similar to those detected in PTI. Activation of the CDPKs leads as well to phosphorylation of RBOHD and ROS-production is initiated [72]. For RPM1-induced ETI a PTI-like phosphorylation of MPK3 and 6 was reported, being more prolonged compared to PTI during which MAPK-activation was found to be transient. This leads in case of CNL-dependent ETI to an independency of SA, because MAPKs are constantly active and can substitute SA-signaling pathways [45]. This cascade results in activation of WRKY33 and subsequent expression of *PAD3* and *ACS2/6*. Thus, camalexin and other defense-associated substances are produced (Figure 4). In summary, CNL-signaling utilizes pathways similar to those reported for PTI, which may differ only in duration and strength.

TNL-initiated signaling

In case of TNL-dependent ETI EDS1 is the main regulator. EDS1 is present in two different complexes, either in complex with PAD4 (PHYTOALEXIN DEFICIENT 4) or with SAG101 (SENESCENCE-ASSOCIATED GENE 101).

(i) The EDS1-PAD4 complex is known to be responsible for SA accumulation, activation of defense genes and production of N-hydroxypipecolic acid (NHP). SA synthesis and subsequent gene regulation is thought to take place due to induction of *ICS1* via a currently unknown mechanism (Figure 4) [73]. Production of SA subsequently leads to the initiation of SA-responsive genes by activation of NPR1 [74]. High SA levels as detected in case of ETI have as well been reported to initiate the expression of JA-dependent genes. This induced expression is mediated by NPR3/4. Both proteins are normally known to act as repressors of NPR1 under low SA concentrations and are inhibited upon elevated SA levels [46]. An additional function has been proposed as NPR3/4 were detected to bind and inhibit JAZ (JASMONATE ZIM DOMAIN) proteins at high SA concentrations. These are known to act as suppressors of JA signaling. Consequently JA signaling and synthesis were shown to be SA-dependent during ETI and are initiated when NPR1 is activated (Figure 4) [75]. However, a recent publication could demonstrate that not all defense genes, acting downstream of the EDS1-PAD4 complex, are SA-dependent [73]. An additional pathway, which was reported to be sufficient for a successful immune response, relies on the FMO1 (FLAVIN-DEPENDENT MONOOXYGENASE 1) [76]. Upon up-regulation of *FMO1* NHP is produced that acts as a positive regulator of the defense response. NHP accumulation has been reported to result in MAPK activation and therefore further initiation of defense genes (Figure 4) [77]. The EDS1-PAD4 complex has also been linked to MAPKs, because the TF WRKY33, a target of MAPKs, was reported to be activated in a PAD4-dependent manner [42]. This is noteworthy, since EDS1-dependent ETI has so far been associated with a lack of MAPK-regulation [73]. Thus, these pathways can confer immunity independently of prior SA, ROS and MAPK signaling, highlighting the diverse mechanisms leading to EDS1-dependent ETI [76].

(ii) As aforesaid, EDS1 also forms a complex with SAG101. This complex is an important downstream component for a different class of NLRs, the helper NLRs (hNLR). These are NLRs, which are not directly involved in the recognition of effectors, but which have been shown to be essential regulators of the immune response. In case of EDS1 and SAG101 this complex associates with NRG1 (N REQUIREMENT GENE 1) resulting in an onset of HR (Figure 4). Although cell death

was found to be initiated due to the EDS1-PAD4 complex-dependent pathways as well, the switch of EDS1 between PAD4 and SAG101/NRG1 is thought to control the amount of cell death observed during ETI [78]. Remarkably, hNLRs have been detected to function in other pathways as well. A prominent example is ZAR1. Its unique function was discovered when ZAR1 was investigated at the structural level using cryo-electron microscopy [79]. This unusual hNLR forms a pentamer upon activation that is localized at the plasma membrane (Figure 4). This may support for example Ca^{2+} fluxes and therefore the initiation of the defense response [79,80]. Such a mechanism may account for the activation of CDPKs, which were found to be regulated in an EDS1-dependent ETI [4]. Hence, TNL-mediated ETI, though differing from CNL-signaling, relies on similar mechanisms. This is highlighted by the CNL RPS2 whose signaling is mediated via EDS1 thus connecting both pathways directly [81]. In addition, CNL and TNL-dependent ETI have both been reported to lead to an elevated immune response and sometimes to HR. Structural investigations can therefore provide valuable new insights into the mechanisms or lead to the discovery of new functions of components of the plant immune system.

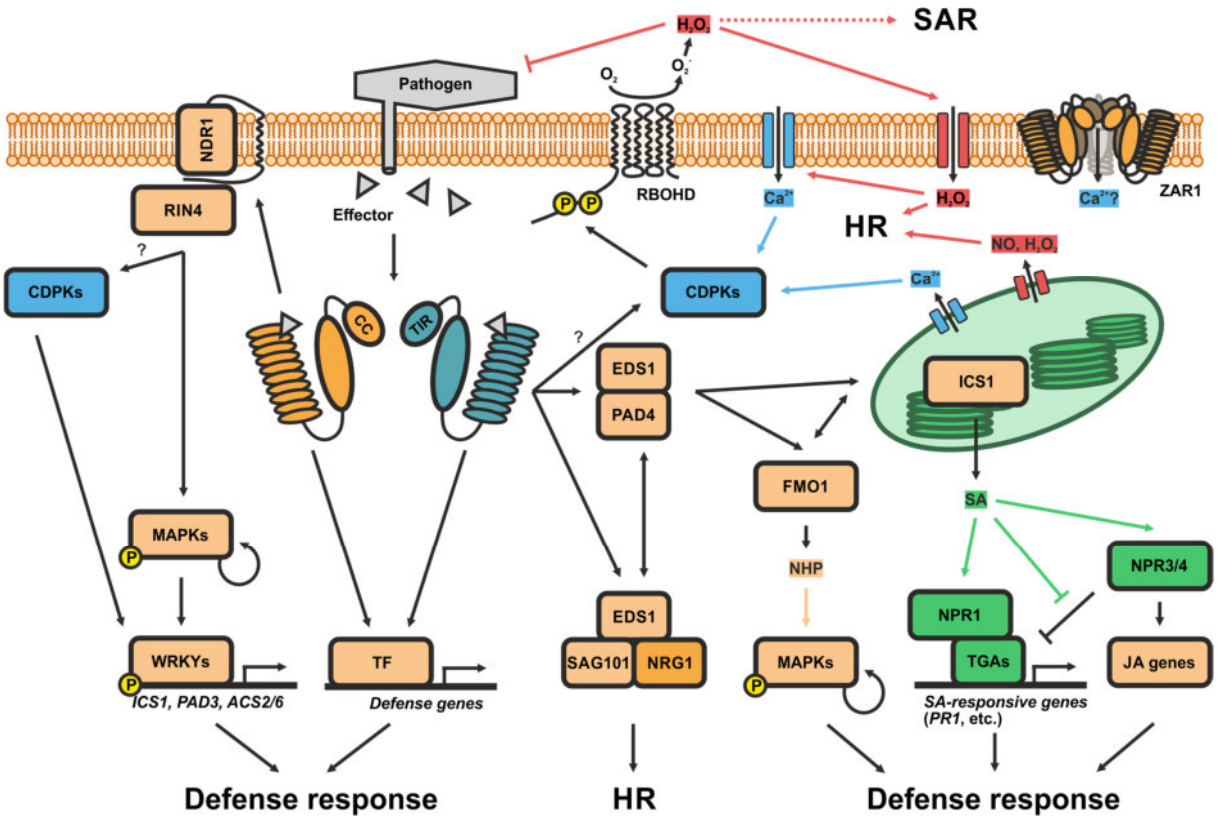


Figure 4: Effector-triggered immunity model. Upon effector perception different signaling pathways can take place. One possibility is a localization of the NLRs to the nucleus and an interaction with a TF leading to the expression of defense related genes [70]. CNLs are mostly dependent on NDR1. By NDR1, Ca^{2+} influx, signals to CPKs and MAPKs are mediated resulting in SA-, camalexin and ROS-production [45,71,72]. In case of TNLs the defense signal is propagated by EDS1. EDS1 can form two complexes EDS1/PAD4 and EDS1/SAG101/NRG1. EDS1/PAD4 results in expression of ICS1 and FMO1 both resulting in independent immunity. This is in case of ICS1 SA- and in case of FMO1 NHP-mediated [42,73,76,77]. SA leads as well, mediated by NRP3/4, to the induction of JA-dependent genes and JA production [75]. NPR3/4 normally acting as repressors of SA-dependent genes are inhibited upon increasing SA-levels [46]. The EDS1/SAG101/NRG1 branch is thought to mediate cell death and an interplay between both complexes regulates the intensity of observed HR [78]. HR is as well regulated by an interplay between NO and H_2O_2 [82]. Ca^{2+} influx being essential for ETI-regulation is hypothesized to be mediated partially by the hNLR ZAR1, which could form membrane pore complexes leading to ion leakage [80].

Hypersensitive response

Subsequent HR is a hallmark of ETI and is defined to be localized cell death, which occurs very rapidly at the point of infection [83]. This cell death is characterized by a shrinking of the plant protoplast which is accompanied by chromatin condensation, cytochrome c release, chloroplast disruption, ion leakage, vacuolization and finally lysis of the plasma membrane [84]. All these mechanisms antagonize further pathogenic spreading.

The regulation of HR has proven to be very complex, since accumulations of SA, ROS, nitric oxide (NO), activation of MAPK cascades and changing cellular Ca^{2+} levels have all been reported to be essential for HR [85]. Although SA is known to be necessary for the induction, the amount of SA is not linked to a specific outcome [86]. Nonetheless, several mutants exhibiting HR exist, which are SA-independent. In some cases, high SA concentrations even act antagonistically to HR [86,87]. Two additional signals, which have been shown to be indispensable for the initiation of HR, have been reported to be ROS and NO, because mutants lacking ROS-production are HR-deficient [85]. Comparable to PTI the main producer of ROS has been described to be RBOHD, thus generating an apoplastic oxidative burst [23]. In addition, chloroplasts and mitochondria are ROS sources by uncoupling photosynthesis and photorespiration leading to rising ROS levels [25]. NO is produced also in chloroplasts and in addition in mitochondria [82,88]. Downstream of ROS and NO production, multiple pathways are activated such as cytochrome c release, down-regulation of detoxifying enzymes and the activation of MAPK cascades [25,82]. The posttranslational modification S-nitrosylation that is NO-dependent has also been reported to affect downstream regulations. NPR1 as well as other proteins are known to be S-nitrosylated, thus being affected by redox-based NO-signaling [82]. In addition, NO and ROS act in a feed-back loop, which was reported to be necessary for HR initiation [82].

Further downstream, two pathways have been reported to be essential for HR-initiation during ETI. One relies on ATMC1 (*Arabidopsis thaliana* METACASPASE 1) as one of the main driving factors. ATMC1 is distantly related to the caspase family which is known to initiate cell death upon NLR activation in mammals. *A. thaliana atmc1* mutants have been shown to result in HR repression [83]. The other pathway resulting in HR is autophagy. Autophagy and ATMC1 seem to act in parallel, where autophagy could lead to vacuolar lysis due to an active overload [89]. In addition, both pathways are known to act antagonistically to cell death as they suppress age-dependent senescence [89]. Comparable to SA-regulation of HR, ATMC1 and autophagy seem to be precisely regulated and their function seems to depend on their cellular context.

It has been discussed, whether the observed HR is actually a part of the immune response or just a consequence of downstream signaling, in particular because an *atmc1* mutant led to a nearly suppressed HR but not to a measurable change in the efficiency of the defense response [83,85].

1.1.3 THE CLOSE CONNECTION BETWEEN ETI AND PTI

ETI and PTI immune responses, which rely on similar or almost identical downstream regulatory pathways, are very much alike. The main difference between both was hypothesized to be the strength and duration of the response [3]. Examples in this introduction underlined that such a clear differentiation between both answers is not possible. One example for this assumption is HR, which does not seem to be always required and may represent an overreaction [85]. In addition, exceptions of immune phenotypes are more frequently reported, such as PTI answers being more intense and thus ETI-like and ETI answers being less drastic and more PTI-like [13]. Both immune responses appear to be highly connected and members which have been thought to be exclusively assigned to one pathway are also found to be important for the other pathway [45,49,71]. Consequently, the only defining aspect of PTI and ETI is the recognition of the PAMPs or effectors and the subsequent activation of PRRs or NLRs [13].

1.2 UBIQUITINATION IN PLANTS

Posttranslational modifications are the main regulatory mechanisms, which define function, localization, interaction and fate of a protein. Although a broad range of posttranslational modifications exist, ubiquitination and the related SUMOylation (SMALL UBIQUITIN-LIKE MODIFIER) have been shown to be of particular interest [90]. Ubiquitin consists of 76 amino acids only and has a mass of 8.5 kDa. Remarkably, ubiquitin, and the related SUMO polypeptides are themselves potential targets of posttranslational modifications [91]. Ubiquitin contains seven lysines and one methionine, all of which can be ubiquitinated, enabling a plethora of ubiquitin-ubiquitin linkages. This results in a large complexity of ubiquitin derived modifications. These can range from monoubiquitination to polyubiquitination forming complex branched structures [92]. Whereas monoubiquitination has been reported to have a role in endocytosis and other regulatory mechanisms [93], polyubiquitination is mostly associated with proteasomal degradation [91].

Ubiquitination model

Proteins involved in ubiquitin-dependent posttranslational modifications can generally be divided into three classes: writers, readers and erasers. Writers are proteins which are involved in ubiquitination, readers interact with the ubiquitinated proteins and erasers are deubiquitinating enzymes (DUBs). In line with the four consecutive steps during the ubiquitination cycle, four different families of writer proteins exist. These are ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2), ubiquitin ligases (E3) (Figure 5) [94]. In addition, a fourth protein family has been discovered to be involved in ubiquitin chain elongation (E4). However, so far only one member of this family has been discovered in *A. thaliana* [95].

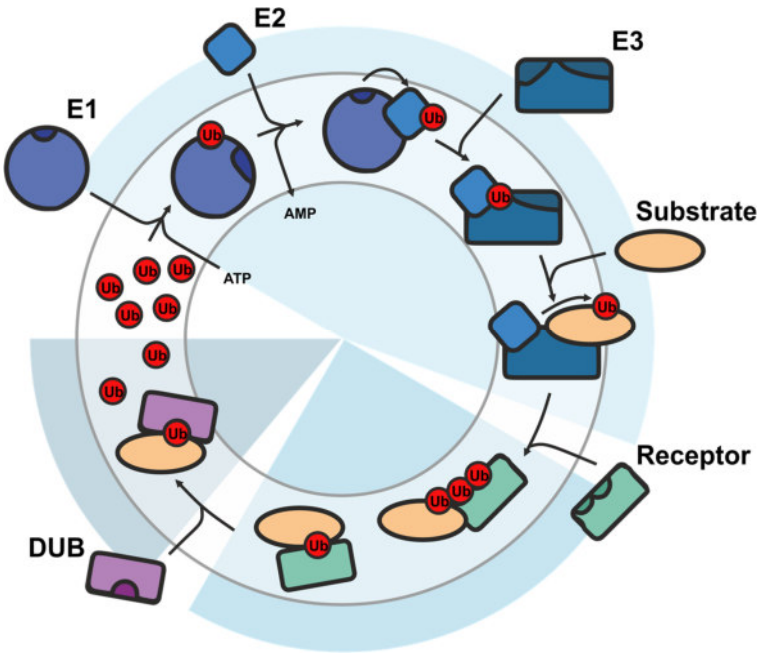


Figure 5: Ubiquitination model. In an initial step ubiquitin is bound to E1 enzymes thereby consuming ATP. Afterwards, ubiquitin is transferred to E2 enzymes. These associate with E3 enzymes and their corresponding substrates. In a subsequent reaction, substrates can either be mono- or polyubiquitinated. These posttranslational modifications are recognized by receptors. The mode of ubiquitination defines the fate of the substrate, which can be for instance delivery to vesicles or proteasomal degradation. In addition, ubiquitin molecules can be cleaved of the substrate by DUBs. Modified after Trujillo, M. (2017) [96].

The activation of ubiquitin involves the consumption of ATP when in the initial step an E1 binds ubiquitin via a thioesther linkage. The E1-ubiquitin linkage is formed between a conserved cysteine of the E1 and the terminal glycine of ubiquitin [97]. In addition, due to binding of ubiquitin, E1s undergo conformational changes resulting in an exposed E2-binding site, which is normally hidden [97]. Following the activation ubiquitin can be transferred to a cysteine of one of the 37 E2 enzymes in *A. thaliana* (Figure 5). The cysteine resides in a conserved region, which is surrounded by more variable domains [97]. In a next step, the E2-ubiquitin complex associates with an E3 (Figure 5). Thereby, surrounding variable regions in the E2 are thought to convey E3 specificity. Generally, about six amino acids from three different regions mediate E2-E3 interactions. These have been shown to be E3-dependent in case an E2 interacts with multiple E3s [97]. In addition to their ability to transfer ubiquitin to E3 enzymes, E2s are known to determine the linkage specificity and length of ubiquitin chains [97]. Accordingly, E2s and not E3s determine the fate of a ubiquitinated substrate. In a last step, the ubiquitin is transferred from the E2 to the substrate supported by an E3 enzyme. The transfer can either be direct or involves the transfer to the E3 first.

Ubiquitin linkage

The ubiquitin chains that are attached to target substrates by E3 ligases, may differ greatly in their structure. The most abundant chain type in plants is a ubiquitin linkage at Lys48 [94]. This linkage type, in case of a chain of at least four ubiquitins, is known to result in proteasomal degradation of the modified protein [98]. Nonetheless, ubiquitination types with differing linkage types are known, resulting as well in degradation. This is explained by the 'ubiquitination threshold' model. This model states that the linkage type is not the primary determinant but it is the number of ubiquitin proteins that matters [91].

Apart from proteasomal degradation, proteins are influenced by ubiquitin modifications in other ways. Polyubiquitination not always results in proteasomal degradation, such as in the case of Lys63-mediated ubiquitination. This ubiquitination was frequently observed in membrane-localized proteins. Additionally, trafficking-associated proteins were often polyubiquitinated by Lys63 linkage. Thus, this modification is hypothesized to function as a localization signal and to be involved in endocytosis [99,100]. Several additional ubiquitin linkage functions have been reported in mammals. For example Lys33 and Lys63 linkage have both been reported to result in endocytosis, thus being similar to Lys63 in plants [91]. Met1 polyubiquitination, which has not been reported in plants yet [101], was observed to act as a cofactor in the catalytic center of an oncogenic protein [91]. Therefore, a plethora of different ubiquitin functions could exist in plants as well. Nonetheless, the main function of polyubiquitination still seems to be proteasomal degradation, as this mechanism has also been observed for Lys11 and Lys29 mediated polyubiquitination [91].

Proteasomal degradation

Degradation is generally performed by the 26S proteasome, a 31 subunit-containing protein complex, in an ATP-dependent manner. These subunits are arranged into three complexes, the 20S core protease (CP) and two 19S regulatory particles (RPs). The RPs are known to restrict access to the CP so that only unfolded proteins can enter the internal protease chamber. Because the only function of the CP is protease-mediated degradation of the substrate, the RPs have several functions like substrate recognition, recycling of ubiquitin molecules, unfolding, transport of the substrate to the CP and subsequent release of the degraded substrate [94]. In addition, 26S proteasomes have been reported to associate with additional proteins, which could act as cofactors [102]. Consequently, different proteasome types would be assembled, which could have specific cellular localizations or targets. In the end degradation is facilitated in the CP by three proteases, which exist in a combination resulting in cleavage of nearly all peptide bonds [102]. Thus, only amino acids or small peptides remain, which are subsequently degraded. The ubiquitin chains removed by the RP are also split into single ubiquitins. This step is performed by DUBs [96].

Deubiquitination

Although 64 different DUBs are known in *A. thaliana* only few have been analyzed. Generally, DUBs are either cysteine or zinc metalloproteases and have mostly been reported to be involved in ubiquitin recycling after or prior to proteasomal degradation [94]. Until now the only DUB with known target in *A. thaliana* is UBP26 (UBIQUITIN BINDING PROTEIN 26). UBP26 is involved in ubiquitin removal of histones, which is important for heterochromatic histone methylation. Nevertheless, the importance of DUBs should not be underestimated, since mutants of different DUBs are characterized by strong developmental or immune-related miss regulations [103].

E3 enzymes

Specificity of the ubiquitin pathway depends on E3 ligases that bind the substrate protein and are required for ubiquitination. With more than 1400 members E3 enzymes are the largest protein family in *A. thaliana* involved in ubiquitination [94]. They can be divided into six families that can be categorized into two main types of E3s, depending on whether they contain a HECT (HOMOLOGY TO E6-AP C-TERMINUS) or a RING (REALLY INTERESTING NEW GENE) finger domain [94]. E3 ligases harboring a HECT domain are with seven members the smallest E3 family in *A. thaliana*. They are characterized by an additional binding site for ubiquitin in the HECT domain. This has been reported to be essential for HECT-mediated ubiquitination, because the ubiquitin binds this site transiently (Figure 6A) [104]. Although members of this family have at first only been associated with developmental processes such as trichome development, their involvement in the regulation of immune responses has recently been suggested. HECT E3 ligases appear to be indispensable for SA-dependent immune signaling and have been associated with the regulation of WRKYs potentially by targeting their repressors [105].

Members of the monomeric RING finger or the extremely similar U-box domain-containing family are the second largest E3 family with more than 400 members. This family differs from the HECT ligases by one significant property. Ubiquitin is not transferred to this kind of E3 ligase, but the ubiquitin is transferred directly from the bound E2 to the target substrate (Figure 6B). In RING and plant U-box type E3 ligases (PUBs) the required E3-E2 interactions are facilitated by a 40-70 amino acid long domain. RING ligases and PUBs are grouped together, as their E2-binding domains are structurally alike. In case of RING E3 ligases, E2 binding is mediated by two chelated zinc atoms. In contrast, PUBs lack these zinc atoms and mediate E2-E3 interaction via electrostatic interactions [94,104].

The remaining E3 ligases belong to the CULLIN-RING ligase (CLR) superfamily. These are multimeric E3 enzymes in which cullin or cullin-like proteins act as scaffolds for multiple associated adaptors. One of the associated proteins, which is recruited to the C-terminus of the cullin proteins, always harbors a RING finger domain. This associated protein is responsible for the E3-E2 interaction, which is mediated by the RING finger domain. In most cases this is RBX1 (RING-BOX 1). Three subgroups of the CLR family are known to associate with RBX1. These are the SCF (S PHASE KINASE-ASSOCIATED PROTEIN 1-CULLIN 1-F-BOX), BTB (BRIC-A-BRAC-TRAMTRACK-BROAD) and DDB (DNA DAMAGE-BINDING) complexes. They differ in the cullin member that serves as a scaffolding protein, being CUL1 (CULLIN 1), CUL3 and CUL4, respectively (Figure 6C,D,E) [104]. In addition, they contain different adapter proteins serving as substrate recognition modules.

Among these CLRs SCF ligases are especially interesting, because they are the largest family with more than 700 members and contain a F-box protein for substrate recognition (Figure 6C). Such proteins have been described to directly bind auxin and JA and therefore act in downstream hormone signaling [106]. The most common example is COI1 acting as a transcriptional regulator of JA by polyubiquitinating the JAZ repressors (see chapter 1.1.1) [50]. In addition to hormone recognition and signaling, SCF E3 ligase complexes are as well involved in the regulation of circadian rhythm, cell division, senescence and additional pathways [104].

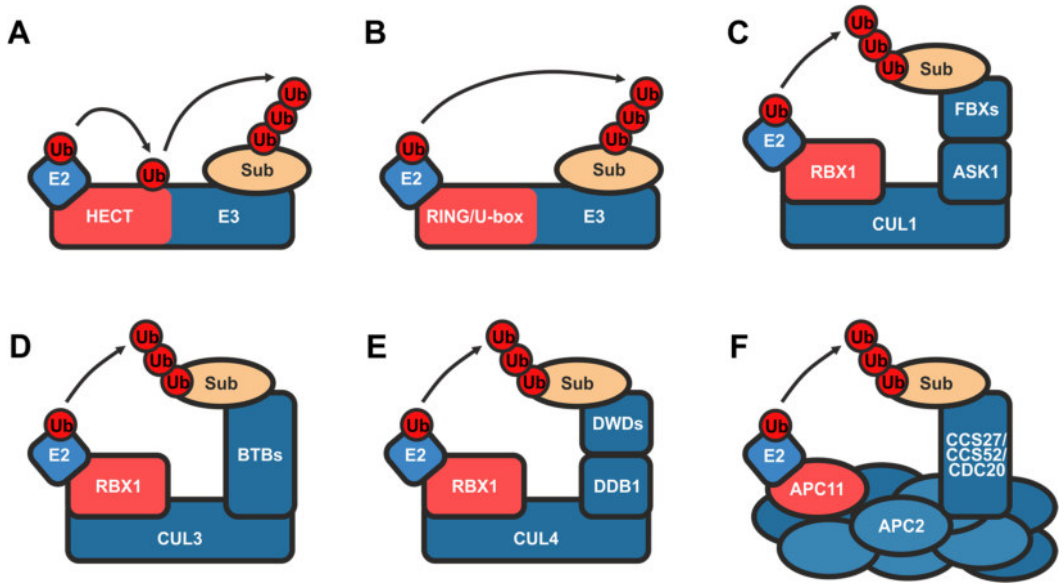


Figure 6: E3 ligase families in *A. thaliana*. E3 ligases consist of six distinct families, which can be grouped by their E2-interacting domain. **(A)** HECT E3 ligases are unique as they contain in their HECT domain an additional ubiquitin-binding site, which mediates the ubiquitination of the substrate (Sub). **(B)** RING/U-box E3 ligases contain either a RING finger or a U-box domain, which are structurally related. These E3 ligases can directly ubiquitinate a substrate. CRLs in contrast are multimeric RING E3 ligases, which mediate the ubiquitination by adaptor proteins. Scaffolding proteins are either cullin proteins or APC2. **(C)** SCF are E3 ligase complexes that contain CUL1 and RBX1 as a E2-interacting module. RBX1 harbors a RING finger domain to mediate the E3-E3 interaction. FBXs allow for substrate specificity, where ASK1 acts as a bridging protein. **(D)** BTB are E3 ligase complexes consisting of CUL3 that is associated with RBX1 and BTB proteins. BTB proteins are known to mediate the substrate specificity. **(E)** DDB E3 ligase complexes consist of CUL4 that is bound to RBX1 for E2 recognition. DWD proteins bind to specific target substrates and are associated with CUL4 via DDB1. **(F)** APC in contrast are large multimeric E3 ligase complexes that contain 11 or more subunits, which are related to cullin proteins or RBX1. E2 recognition is mediated by APC11. Three interchangeable recognition protein families (CCS27/CCS52/CDC20) are essential for substrate recognition [94]. Modified after Chen, L. *et al.* (2013) [104].

Whereas F-box containing proteins bind to CUL1 via the adaptor protein AKS1 (ARABIDOPSIS SKP1 HOMOLOGUE 1), BTB E3 complexes do not contain such a bridging factor. In the latter, recognition of the substrate is mediated by the BTB protein, which binds directly to CUL3 (Figure 6D). BTB proteins contain a conserved core structure, which is composed of five α-helices and three β-strands, surrounded by a variable region. This has been reported to mediate specific protein-protein interactions [104]. Due to this variability, BTB complexes have been reported to be involved in a broad range of interactions, specifically stress responses [104]. In this context, the afore mentioned NPR3/4 have been found to be BTB proteins with the additional ability to bind SA (see chapter 1.1.2) [87]. This has been demonstrated to result in an involvement in JA signaling as both proteins were found to target, upon SA binding, JAZ proteins for polyubiquitination and therefore degradation [75].

The third family, the DDB E3 ligase complexes contain a bridge protein similar to SCF complexes, which connects substrate recognition with the scaffolding protein CUL4. This DDB1 is connected to DWD (DDB1 BINDING WD40) proteins, which mediate substrate interaction (Figure 6E). DWD proteins have been reported to be essential for substrate recognition mediated by their protein-binding WD40 domain. In conclusion, all these CLRs are involved in different regulatory processes, while being relatively similar in their structure. All of them form a complex out of three to four components.

APC (ANAPHASE PROMOTING COMPLEX) E3 ligase complexes contrast strongly with these E3s. They form extremely large complexes consisting of at least of 11 subunits. Nonetheless, these subunits are related to the ones that mediate E3 function in SCF, BTB and DDB complexes. APC11

for instance contains a RING finger domain and fulfills therefore a similar function as RBX1. In addition, APC2 is functioning similar to CUL1/3/4 by acting as a central hub for all APC subunits (Figure 6F). Concerning their function APC complexes are known to be involved in cell cycle regulation and have been shown to be important for a vast range of developmental processes. In this context, they have been reported to ubiquitinate cyclins and thus label them for turnover. Similar to mammals, the proteins mediating substrate interaction belong to two protein families. These are CCS27/52 (CELL CYCLE SWITCH PROTEIN 27 and 52) and CDC20 (CELL DIVISION CYCLE 20) families that interact through specific amino acid motifs with the substrates [104,107]. Consequently, the great complexity of E3 ligases enables them to form a diverse range of E3-E2-substrate compositions. This results in a tremendously diverse range of signaling pathways that are regulated by these E3 ligases.

1.2.1 PLANT U-BOX TYPE E3 UBIQUITINATION LIGASES (PUBS)

Of all these E3 ligase families PUBs and RING enzymes are the only proteins that allow direct transfer of the ubiquitin from the E2 enzyme to the substrate without associating with adapters [94]. PUB E3 ligases, consisting of 64 members, have many divergent functions, which are based on their structural variability [108,109].

These functions are often associated with various stresses, because the *PUB* genes have been found to be differentially expressed upon various stress treatments [109]. In addition, *pub* mutants are often very sensitive to different stresses [96]. It has also been reported that the localization of PUBs changes upon external stimuli [110]. Single PUBs can even be involved in different regulatory processes based on their multiple interactions with different E2 enzymes [111].

This functional variability between different PUBs is based on structural variations, because they can contain additional protein domains. In addition to the eponymous U-box domain, which is important for E2-E3 interaction, PUBs contain other domains that affect their function such as LRRs, kinase domains or ARMADILLO (ARM) repeats [108]. Among these additional domains ARM repeats have been demonstrated to be present in most PUBs. They have been detected in 41 of 64 PUBs in *A. thaliana* [112]. These are mostly known to facilitate protein-protein interactions. This function has been frequently reported in case of PUBs mediating ligase-substrate interactions [96]. In addition, these ARM repeats have been shown to be quite versatile in their function, since an involvement in dimerization as well as plasma membrane localization has been reported [113,114].

Apart from the interaction of PUBs with multiple E2 enzymes or substrates, the function of PUBs can be directly influenced by regulating their activity. A very prominent one is PUB autoubiquitination resulting in their degradation. This is thought to be a regulatory mechanism to keep the levels of the respective PUBs low as long as they are not required [115]. An additional mechanism is the phosphorylation of PUBs [96]. For instance, PUB22 is phosphorylated by MPK3 and thus stabilized by inhibiting PUB22 oligomerization [116]. A similar mechanism has recently been reported for PUB25 and PUB26, which have been found to be targets of phosphorylation by CPK28 [117].

Consequently, oligomerization is one of the key mechanisms regulating PUB activity. The oligomerization status of PUB10 and PUB22 for example had a direct influence on their functionality [114,116]. This is highly interesting, since a change of oligomerization is very likely to be based on structural changes. Although multiple interaction partners such as E2 enzymes and substrates of PUBs have been identified, their structural properties are still largely unknown [96]. Until now, only structural data concerning their respective U-boxes exist. A complete structure that would help to understand how PUBs are regulated, function and oligomerize is still missing.

1.2.2 PUB-DEPENDENT REGULATION OF THE PLANT IMMUNE SYSTEM

Ubiquitination is one of the main regulatory mechanisms during plant immunity involving all E3 ubiquitin ligases families [118]. In addition to receptor internalization after activation, E3 ubiquitin ligases are also targeting other membrane-associated proteins to the vacuole resulting in a positive regulation of the immune response [119]. In addition, E3 ligases are involved in affecting the expression of defense-associated genes by targeting histones [120] or promoting the degradation of defense suppressors [121]. Nevertheless, E3 ubiquitin ligases act as both positive and negative regulators of the immune response [118]. In contrast to plant ligases, pathogenic effector proteins are known, which are E3 ubiquitin ligases that target components of the plant immune system and therefore repress the defense response [122]. Thus, E3 ubiquitin ligases act in a plethora of immune-associated pathways.

Among all involved E3 ligase families PUBs are particularly interesting. Despite being involved in a plethora of various abiotic stresses the main regulatory involvement of PUBs appears to be in plant immune responses [96]. This can already be observed by the fact that many PUB mutants possess an elevated immune response upon pathogen perception or have autoimmune like phenotypes mediated by systemic ETI responses. These phenotypes have been reported for *pub12/13*, *pub22/23/24* and *pub44* [96,123].

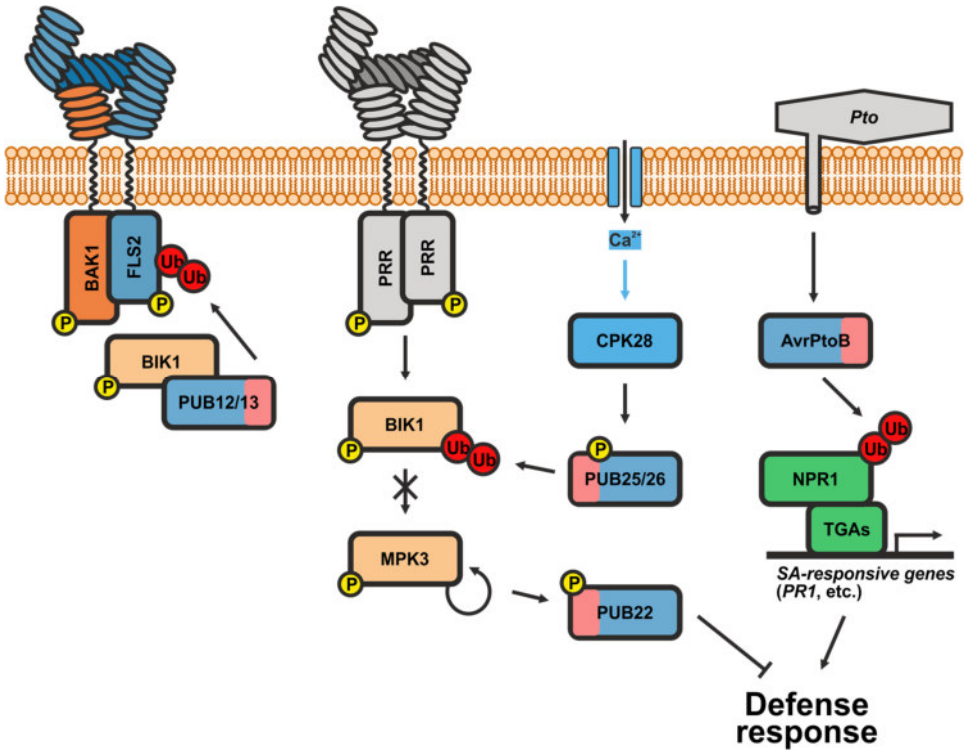


Figure 7: PUBs and their involvement in plant immunity. PUBs have been reported to have diverse roles during plant immunity. One major function is the negative regulation of PRR-mediated defense responses. This is mediated by PUB12/13 and PUB25/26, which polyubiquitinate either FLS2 or BIK1 upon immune activation [117,124]. Further downstream PUB22 was found to be phosphorylated and subsequently activated by MPK3 [125]. This results in a negative regulation of the defense response. In addition, bacterial effectors like AvrPtoB are as well U-box E3 ligases which target important regulators of the defense response like NPR1 [122]. The U-box is displayed in red.

For the paralogs PUB12/13 an involvement in the regulation of the PRR FLS2 has been demonstrated. Upon activation following pathogen perception, FLS2 is known to be internalized and subsequently degraded (Figure 7) [126]. Regulation mediated by endocytosis and subsequent degradation has been reported for multiple PRRs or RLCKs [126,127]. This mechanism is necessary,

especially during PTI, to avoid an unnecessary long activation of the PRRs that could result in an inadequately strong immune response. This could subsequently harm the plant cell, even potentially resulting in cell death. In case of PUB12/13 this negative regulation is mediated by binding to BIK1. This allows PUB12/13 to associate with FLS2 resulting in its ubiquitination (Figure 7) [124,128]. Remarkably PUB13 was recently found to be involved in the regulation of the PRR LYK5, the chitin receptor, as well[127]. These findings support the above-mentioned hypothesis (see chapter 1.2.1) that PUBs may have multiple substrates.

An additional RLCK-associated mechanism of PTI signal control has been shown for the closely related PUB25/26. However, this PUB pair acts further downstream in regulating the immune system. Due to Ca^{2+} influx, CPK28 is activated and has been reported to phosphorylate PUB25/26. This led to their activation and the subsequent polyubiquitination of their substrate BIK1, which then gets degraded (Figure 7) [117]. This inhibits or no longer activates downstream signals such as MAPK cascades.

Further downstream of BIK1 MPK3 has been reported to phosphorylate and therefore stabilize PUB22. Thus, PUB22, which normally autoubiquitinates to initiate its degradation, is directly controlled by an onset of PTI (Figure 7) [116]. When stabilized PUB22 targets a member of the exocyst complex, EXO70B2, which is known to be vital for early immune responses [125,129]. Accordingly, PUB22 functions as a negative regulator and thus counteracts a potential overmodulation of the plant immune response.

PUBs are not always acting as negative regulators, which has been demonstrated in tomato. StPUB17 was reported to re-localize into the nucleus upon pathogen perception and to act there as a positive regulator of PTI and even cell death [130]. This indicates the diversity of regulatory mechanisms employed by PUBs to control immune responses. In addition, the plant ubiquitin 26S proteasome pathway can be hijacked by pathogens. The bacterial effector AvrPtoB, which is secreted into the cell upon *Pto* infection, was described as a U-box containing E3 ligase. Upon secretion AvrPtoB targets the SA-dependent regulator NPR1 resulting in its polyubiquitination (Figure 7) [122]. Due to subsequent degradation of NPR1 SA-responsive genes are no longer expressed. In conclusion, PUBs and related proteins are key regulators of the plant immune response, whose functional and structural analyzes give valuable insights into the regulatory processes that balance the plant immune system.

1.3 SENESCENCE-ASSOCIATED E3 UBIQUITIN LIGASE 1 (SAUL1)

Together with its paralog PUB43, the PUB-ARM protein PUB44 or SAUL1 (SENESCENCE-ASSOCIATED UBIQUITIN LIGASE 1) was found to be unique in its structure in the PUB family. In contrast to other members, SAUL1 and PUB43 do not contain a UND domain. More importantly, they contain a larger number of 11 and 12 proposed ARM repeats, respectively, compared to all other members. These additional ARM repeats result in an elongated C-terminus in SAUL1 and PUB43 (Figure 8) [112].

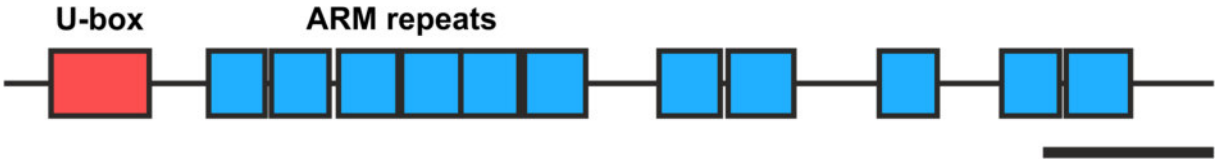


Figure 8: Schematic representation of SAUL1. Domain organization of SAUL1 where the U-box domain is depicted in red and the ARM repeat containing domains are pictured in blue. Scale bar represents 100 amino acids. Representation modified after Mudgil, Y. *et al.* (2004) [112].

First described in 2009 the *SAUL1* gene was found to be one of the few *PUB* members being differentially expressed upon ABA treatment [123,131]. Interestingly, the *saul1-1* mutant of this active E3 ligase appeared to be quite unique. Mutant *saul1-1* seedlings as well as older plants exhibit an immediate growth arrest and a systemic lesioning of all above-ground organs upon changes in the environment [123]. The characteristic phenotype could be triggered by a decrease in temperature from 25 °C to 24 °C or in an environment with a relative humidity below 80 % [132]. Until now, this phenotype is nearly unique in its simultaneous occurrence in all above-ground organs. Only the *chs1-2* gain of function mutant, an NLR, has been described to be similarly severe [133,134]. Due to the systematic lesions, reduced chlorophyll content, lower efficiency of the photosystem II and expression of senescence marker genes like *SAG12/13* and *SIRK* (*Senescence-Induced Receptor-Like Kinase*) *SAUL1* was first thought to be a suppressor of senescence [123]. Later on, a *PAD4*-dependency of the *saul1-1* phenotype was observed, because a *saul1-1 pad4-1* double mutant was able to revert the phenotype [135]. Therefore, the SA pathway was described to be essential.

Subsequent experiments could show that the actual phenotype observed in *saul1-1* is more likely related to immunity, because *PR* genes were found to be strongly up-regulated [132]. In addition, the phenotype with severe growth defects was similar to the ones observed in other autoimmune mutants like *snc1-1* or *chs1*-2 [133,136]. In addition, increased cell wall depositions containing callose and lignin were detected and pointed towards immune responses. Furthermore, a stronger resistance against various pathogens was reported. Thus, *SAUL1* was considered to be a negative regulator of autoimmunity [132]. This assumption was revised when the *saul1-1* phenotype could be suppressed by mutations in *SOC3* (*Suppressor of chs1-2, 3*) [137]. SOC3 is an NLR and had previously been reported to revert the highly similar autoimmune phenotype of *chs1-2* gain of function mutants [11]. Remarkably, the *chs1-2* phenotype was also inducible by low temperature [133]. SOC3 and CHS1 have been shown to interact and to form a heteromeric NLR complex [11]. Such complexes have also been observed for other NLR proteins, whose genes are in direct proximity on the genome [10]. Therefore, it has been hypothesized that SAUL1 is actually guarded by this heteromeric complex, because both NLRs have been found to act in close proximity to SAUL1 (Figure 9A,B) [137]. In conclusion, the *saul1-1* phenotype results most probably from the activation of ETI.

Intriguingly, not only the absence of *SAUL1* in *saul1-1* mutant appeared to be guarded. The overexpression of *SAUL1* in *35S::SAUL1* plants, which resulted in reduced growth and the expression of *PR* genes, was also reported to depend on the *SOC3* and *TN2* [137,138]. Like CHS1, *TN2* encodes a truncated NLR, which is lacking the LRR domain. The *TN2* gene is localized head-to-head with *SOC3* on the genome. In case of *35S::SAUL1* TN2 and SOC3 are thought to guard a

putative target substrate of SAUL1. When degraded due to high amounts of SAUL1, the absence of that substrate may lead to the activation of SOC3 and therefore to the onset of the ETI (Figure 9C) [138]. SAUL1 and its potential target substrate are therefore highly important for the immune response, since both are tightly guarded.

In an additional experiment, an increased growth of *Pto* was detected when infecting the triple mutant *saul1-1 soc3-1 pub43-1* with *Pto* DC3000 hrcC-. *PUB43* was mutated as well, because it has been shown that this paralog is partially redundant to *SAUL1* [137]. NLRs are not activated upon an infection with *Pto* DC3000 hrcC-, since this bacterium is impaired in its type III secretion system. Therefore, ETI cannot be activated in plants, which are infected with *Pto* DC3000 hrcC-. Because of the elevated growth of *Pto* DC3000 hrcC- in the triple mutant *SAUL1* appeared to be a positive regulator of PTI being repressed in the triple mutant [137].

In conclusion, although SAUL1 is not directly involved in the regulation of ETI, it is highly associated with both immune pathways and may play an important role in the regulation of PTI. Its most remarkable hallmark is its simultaneous phenotype in all above-ground organs, resembling an onset of ETI, being nearly unique among the autoimmune mutants in *A. thaliana*. The possibility of initiating the phenotype by means of the change in temperature makes it possible to study the *saul1-1* mutant and the regulated genes very specifically. Therefore, the *saul1-1* mutant could be a valuable tool to study the signaling mechanisms of ETI, especially as ETI is still not yet fully understood in its complexity (see chapter 1.1.2).

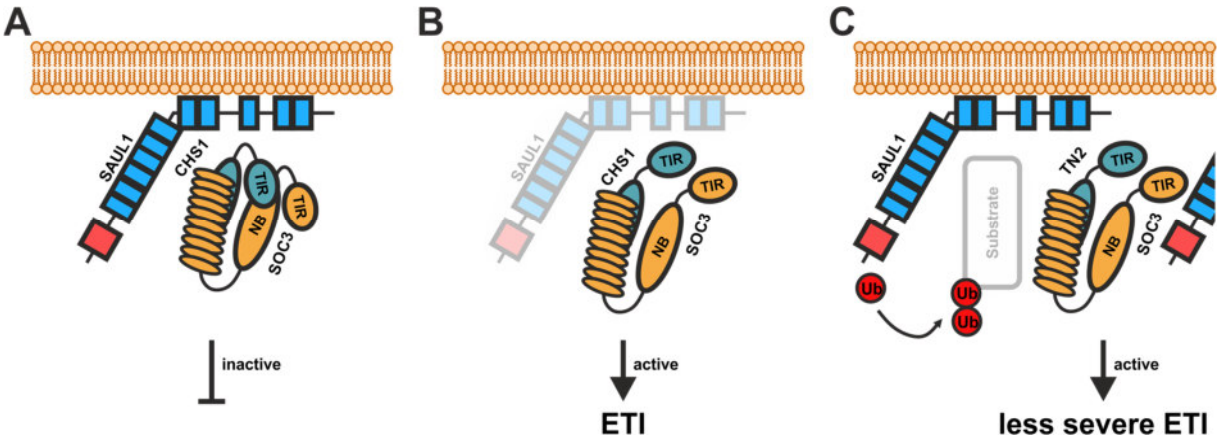


Figure 9: Working model for different *SAUL1* mutants. (A) In WT plants SAUL1 is guarded at the plasma membrane by the heteromeric NLR complex consisting of SOC3 and CHS1 or TN2. **(B)** In the *saul1-1* mutant, the absence of SAUL1 is recognized by the CHS1-SOC3 heterodimer. Thus, SOC3 gets activated and initiates the ETI response. A similar mechanism is thought to occur, if SAUL1 is not localized at the plasma membrane any more. For instance, due to a deletion of the predicted five C-terminally located ARM repeats. **(C)** In *35S::SAUL1* plants the degraded target substrate is thought to be guarded by the TN2-SOC3 heterodimer. Consequently, SOC3 gets activated and a less severe ETI response is initiated. Model modified after Tong, M. *et al.* (2017) [137] and Liang, W. *et al.* (2019) [138].

In addition to gene regulations in the *saul1-1* mutant, the potential function of SAUL1 was also investigated at the protein level. The SAUL1 protein was found to be uniquely localized at the plasma membrane, which was mediated by the predicted five C-terminally located ARM repeats 7-11 (Figure 9A). Upon deletion of these ARM repeats SAUL1 appeared to reside in the cytoplasm and the nucleus as well. A similar pattern was reported for its paralogs PUB43 and PUB42 and related PUB proteins in rice, poplar and moss [139]. Interestingly, a change of subcellular localization to the nucleus was as well reported upon ABA and 2,4D-treatment in cell culture [110]. Remarkably, complementation of *saul1-1* mutant plants with *35S::SAUL1ΔARM*$_{1\text{-}6}$, lacking the predicted C-terminally ARM repeats 7-11, did not result in a rescue of the *saul1-1* phenotype [113]. Therefore, the interaction between SAUL1 and the CHS1-SOC3 complex may be mediated by these ARM repeats. In addition, the SAUL1 localization could also be guarded. Since SAUL1 is located in most

cases at the plasma membrane, the still unknown potential target substrate of SAUL1 may also be located there (Figure 9C).

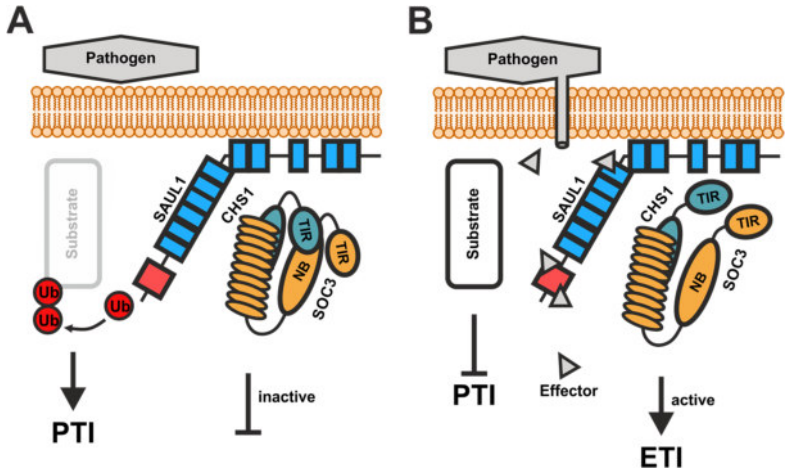


Figure 10: Model for the mode of action of SAUL1. (A) Upon infection with pathogen recognition at the plasma membrane SAUL1 is thought to ubiquitinate a specific target substrate. This protein being highly likely a negative regulator of the immune response is therefore deactivated or degraded. This leads to the onset of the PTI. **(B)** Upon infection with effector-inserting pathogens SAUL1 is hypothesized to be targeted by effectors, which could lead to its degradation, delocalization or deactivation. This is sensed by the CHS1-SOC heterodimer SOC3 gets activated and ETI is initiated. Model modified after Tong, M. *et al.* (2017) [137].

In addition, the localization of SAUL1 has recently been linked to the tethering of multi-vesicular bodies (MVBs) to the plasma membrane [140]. In this context, it has been shown SAUL1 is localized in patch-like structures at the plasma membrane [113]. The same pattern was observed for the SAUL1 homolog PUB43 and homologs from other plant species [139]. These patches are in case of SAUL1 regions where MVBs are tethered to the plasma membrane and SAUL1 could act as a linker between MVBs and the plasma membrane. In addition, the patches have been shown to be formed upon the induction of the immune response and the ARM repeats 7-11 were shown to be sufficient to induce patch-like distribution [140]. However, their exact function and how SAUL1 can act as a linker between MVBs and the plasma membrane is still unknown. In order to resolve this, structural investigations of SAUL1 might be helpful, as they would allow, among other things, to identify potential interaction sites.

In conclusion, the ARM repeats are of high importance for the function of SAUL1, because they could mediate protein-membrane and protein-protein interactions. How these repeating domains are actually organized and how a PUB like SAUL1 is structured, which contains so many ARM domains, is not understood. Another structural similarity to other PUBs was recently discovered. Experiments could show that SAUL1 does interact with itself [141] and could therefore form oligomers, as it has been reported for other PUBs [96]. Therefore, a structural analysis of SAUL1 could also give valuable insights into the organization and function of ARM-containing PUB ligases in general.

1.3.1 POTENTIAL SAUL1 INTERACTION PARTNER BON1

Since the exact function of SAUL1 is still unknown, its putative target substrate is of high interest. The determination of the SAUL1 substrate would allow to gain further knowledge, for example about the mechanisms how SAUL1 is linked to the positive regulation of PTI or the tethering of MVBs to the plasma membrane. Therefore, a potential substrate could be connected to the immune response as well and act as a regulator. Previous work identified BON1 (BONZAI 1) as a putative binding partner of SAUL1 [142]. BON1 was selected because the *bon1-1* mutant had a growth

arrest phenotype similar to *saul1-1*, which was also inducible by low temperatures [143]. It was also reported that BON1 is localized at the plasma membrane like SAUL1 [143]. BON1 itself belongs to the copine protein family. This protein family exists in a broad range of organisms and is characterized by its ability to bind Ca^{2+}-dependently to phospholipids or proteins [144]. The copine family was first described in humans in the form of synaptotagmin. This copine is involved in neurotransmitter transport [145] and Ca^{2+} binding and association to membranes and proteins are known to be mediated by the so-called C2 (C2 PROTEIN KINASE C, CONSERVED REGION 2) domains [146]. Because the *bon1-1* mutant exhibited a temperature-dependent growth arrest a connection to autoimmunity was assumed [143]. In addition, *BON1* gene expression was reported to be induced upon infection with various *Pto* strains and upon SA treatment [147]. Thus, BON1 is involved in the regulation of the immune response. Since infection with a *Pto* strain that is unable to induce ETI did not result in elevated *BON1* gene expression levels, BON1 acts as a regulator during ETI [147]. The *bon1-1* mutant phenotype is also associated with ETI, since double mutants with *snc1-11*, *pad4-1* and *eds1-1* led to WT-like plants [148]. Consequently, BON1 acts as a negative regulator of the TNL-dependent ETI. This was confirmed by overexpression of *CPR1* (*Constitutive Expresser of PR Genes 1*) in *bon1-1*, which led to the suppression of the *bon1-1* phenotype [149]. The suppression resulted from 26S-proteasome-mediated degradation of the TNL SNC1, which was targeted by the F-box protein CPR1 (Figure 11) [149]. A similar effect was observed as overexpression of BON1 and its interaction partner BAP1 (BON ASSICIATION PROTEIN 1) resulted in negative regulation of SNC1. This led to the suppression of effector-induced HR symptoms [150]. Therefore, BON1 is a negative regulator of the immune response.

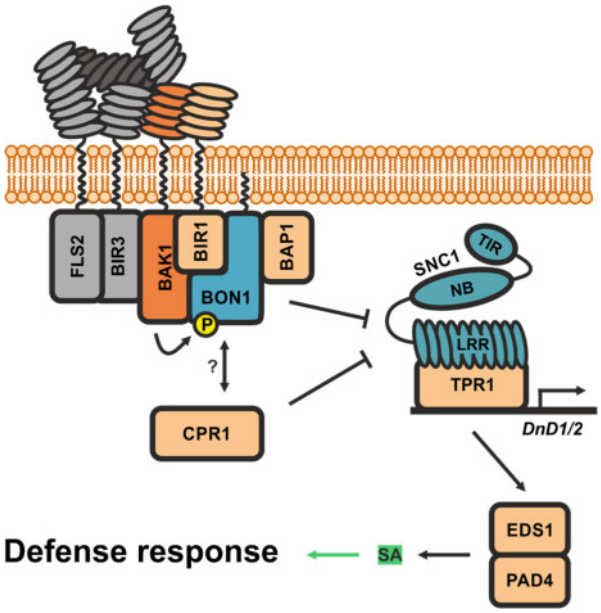


Figure 11: BON1 interaction model. BON1 is localized at the plasma membrane via a myristoylation site [151]. This copine is in a complex with BAP1, BAK1 and the negative regulator of FLS2/BAK1-complex formation BIR1 [150,152]. BON1 was found to be phosphorylated by BAK1, although the function and its role in the complex is still unknown. BON1 negatively regulates the TNL SNC1 in a potential interaction-dependent manner with the F-box protein CPR1. In a *bon1-1* mutant SNC1 accumulates and TNL-dependent ETI following regulation of EDS1/PAD4 and SA is initiated [148,149].

It was reported as well that BON1 is localized at the plasma membrane due to its myristoylation, which is also necessary for its function [151]. BON1 was found to be localized to the plasma membrane in a complex with BAK1 (BRI-ASSOCIATED RECEPTOR KINASE 1), BIR1 (BAK-INTERACTING RECEPTOR-LIKE KINASE 1) and BAP1 (Figure 11) [150,152]. BIR1 and BAK1 are regulators of FLS2 and may relate BON1 to the regulation of FLS2. BIR1 is thought to act as a negative regulator of FLS2 by preventing binding of BAK1 to FLS2 [7]. Thus, FLS2 remains inactive

when no PAMPs are present. Consequently, BON1 may act as a negative regulator of FLS2-BAK1 complex formation, especially as BON1 is phosphorylated by BAK1 [152]. Thus, BON1 could also be involved in the regulation of PTI.

Accordingly, investigating the interaction between SAUL1 and BON1 is of high interest. BON1 could be a target substrate of SAUL1, which is polyubiquitinated upon initiation of the immune response and subsequently degraded. By studying the structural mechanisms underlying the interaction between SAUL1 and BON1, the role of SAUL1 as a positive regulator of PTI can be explored. In addition, the interactions between a PUB and its substrate would be resolved for the first time.

2 AIM OF THE DOCTORAL THESIS

The *A. thaliana* E3 ubiquitin ligase SAUL1 has intriguing properties and is characterized by two hallmarks. (i) The *saul1-1* autoimmune mutant represents an optimal model system to study morphological changes and molecular mechanism of ETI responses, because autoimmunity and cell death occur in all above-ground organs. This is in contrast to all other autoimmune mutants (with the exception of *chs1-2* gain-of-function mutants) and allows to study molecular responses during ETI in very homogeneous plant material. In particular, the connection between ETI and PTI, which may be closer than previously known, could be investigated. To do so, a comprehensive next-generation sequencing (NGS) approach is used in this thesis. In addition, very early changes in gene expression will be detected by applying a time series experiment. These will show how the *saul1-1* phenotype develops and which genes drive regulations during the onset of ETI.

(ii) The SAUL1 protein is the first component of the ubiquitin pathway that has been shown to be (a) guarded by heteromeric NLR complexes. In addition, the (b) SAUL1 domain structure and its (c) localization at the plasma membrane are different from the other PUB-ARM protein family members. Therefore, by investigating SAUL1, the interaction with the NLRs and (d) its binding partner and potential substrate BON1 on a structural level, valuable insights will be given (a) how proteins are guarded by NLRs, (b) how PUB proteins are structurally arranged, (c) how SAUL1 interacts with the plasma membrane and (d) how PUBs interact with their target substrates in general. For this purpose, small-angle X-ray scattering (SAXS) and inline size-exclusion chromatography SAXS (SEC-SAXS) experiments will be performed with recombinant SAUL1 and its interaction partners. In this way, formed complexes can be analyzed separately and structural information can be obtained from proteins in a physiological-like environment.

Accordingly, SAUL1 will be analyzed in a comprehensive way to give new insights into its regulations, its structural organization and interactions with other proteins.

3 MATERIAL AND METHODS

3.1 BIOMOLECULAR METHODS

3.1.1 LIQUID AND SOLID MEDIA CULTURE

Media, which were used for bacterial growth during this thesis are listed below (Table 1). Antibiotics were used according to the manufacturer's specifications. All media were sterilized by autoclaving. For agar plates 0.8 % agar were added prior to autoclaving.

Table 1
Media used for bacterial growth

Medium	Components
LB	1 % Trypton, 0.5 % NaCl, 0.5 % Yeast extract
SOB	2 % Trypton, 0.05 % NaCl, 0.5 % Yeast extract, 2.5 mM KCl, 10 mM $MgCl_2$
2YT	1.6 % Tryptone, 0.5 % NaCl, 1 % Yeast extract

Liquid media cultures were inoculated with a sterile pipette tip and incubated at 37 °C and 180 rpm overnight using a *Multitron Standard* (Infors, Bottmingen, Switzerland). Liquid cultures with up to 5 ml of medium were incubated using a test tube. Larger volumes were incubated in Erlenmeyer flasks. In contrast, agar plates were inoculated by adding 150 µl of an *Escherichia coli* (*E. coli*) bacterial suspension onto the plate and evenly distributing the bacteria using a Drigalski spatula.

3.1.2 POLYMERASE CHAIN REACTION (PCR)

To amplify specific deoxyribonucleic acid (DNA) fragments different polymerase chain reaction (PCR) setups were performed. For colony PCR (see chapter 3.1.5), genotyping (see chapter 3.2.5) and RT-PCR (see chapter 3.2.8) the *DreamTaq™* polymerase (Thermo Fisher Scientific, Waltham, United States) was used by applying the respective conditions (Table 2). To amplify a specific DNA fragment for cloning (see chapter 3.1.8) the *Phusion* polymerase was used by applying the respective conditions (Table 3).

Table 2
(A) Composition of each PCR sample using the *DreamTaq™* polymerase

Component	Volume (µl)
DreamTaq™ buffer (10x)	2
dNTPs (10 µM)	0.4
Primer (0.1 µM)	1
Template DNA	1-2
DreamTaq™ polymerase (5 U μl^{-1})	0.2
ddH_2O	Add 20

(B) PCR program used for the *DreamTaq™* polymerase. Steps 2-4 were iterated 35 times.

Step	Temperature (°C)	Time (s)
1	95	30
2	95	30
3	Primer-dependent	30
4	75	60 kb^{-1}
5	75	300

Table 3
(A) Composition of each PCR sample using the *Phusion* polymerase

Component	Volume (µl)
Phusion HF buffer (5x)	4
dNTPs (10 µM)	0.4
Primer (0.1 µM)	1
Template DNA	1-2
Phusion polymerase (2 U μl^{-1})	0.2
DMSO	1.5
ddH_2O	Add 20

(B) PCR program used for the *Phusion* polymerase. Steps 2-4 were iterated 40 times.

Step	Temperature (°C)	Time (s)
1	98	30
2	98	10
3	Primer-dependent	30
4	72	30 kb^{-1}
5	72	600

3.1.3 *E.COLI* STRAINS AND GENERATION OF COMPETENT CELLS

E. coli strains

The following *E. coli* strains were used for cloning (Table 4). Additional strains for protein expression are listed below (Table 21).

Table 4
E. coli strains used for cloning

Strain	Genotype	Reference
DH5α	*E*. coli B. F- φ80*lacZ*ΔM15 Δ(*lacZYA*-argF)U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 λ- *thi*-1 *gyr*A96 *rel*A1	Hanahan, D. (1983) [153]
One Shot® TOP10	*E. coli* B. F- mcrA Δ(*mrr-hsd*RMS-*mcr*BC) φ80*lac*ZΔM15 Δ*lac*X74 *rec*A1 *ara*D139 Δ(*ara-leu*)7697 *gal*U *gal*K λ⁻ rpsL(Str^R) *end*A1 *nup*G	Thermo Fisher Scientific, Waltham, United States

Generation of chemical competent cells

Chemical competent *E. coli* cells of the corresponding cell lines were generated by inoculating 20 ml LB medium with the corresponding antibiotics with approximately 10 µl of a glycerol stock and incubating the culture at 37 °C and 180 rpm in a *Multitron Standard* (Infors, Bottmingen, Switzerland) incubator overnight. Afterwards 100 ml SOB medium with the corresponding antibiotics were inoculated with 1 ml of the overnight culture in incubated at 37 °C and 180 rpm until reaching an optical density at 600 nm ($OD_{600\ nm}$) of 0.5. In a next step the cells were pelletized for 10 min at 3220 x g and 4 °C. Afterwards the supernatant was discarded and the cells were resuspended using 40 ml sterile filtered TB buffer (10 mM PIPES pH 7.0, 55 mM $MnCl_2$, 15 mM $CaCl_2$, 250 mM KCl) and kept on ice for 10 min. The resuspended cells were pelletized again for 10 min at 3220 x g and 4 °C. In a last step cells were resuspended in 7.5 ml TB buffer with 600 µl DMSO. 50 µl aliquots were immediately frozen in liquid nitrogen and stored at -80 °C.

3.1.4 TRANSFORMATION OF BACTERIAL CELLS

Chemical competent *E. coli* cells were transformed by adding 10 ng plasmid or a respective sample to a thawed aliquot of chemical competent cells. Cells were incubated for 5 min on ice. In a next step cells were incubated for 45 s at 42 °C and immediately transferred on ice. After 5 min 250 µl LB or SOB medium were added and the cells were incubated for at least 30 min at 37 °C and 450 rpm using a *Thermomixer Compact* (Eppendorf, Hamburg, Germany). The cells were then spread out on agar plates containing the respective antibiotics (see chapter 3.1.1). Agar plates were incubated at 37 °C overnight or at 20 °C for 48-72 h in the dark.

3.1.5 COLONY PCR

To identify successfully transformed *E. coli* single colonies were extracted with a pipette tip. Bacteria were resuspended in 20 µl H_2O in a PCR reaction tube. 2 µl of the bacterial suspension were used as a template for colony PCR. The PCR was performed using the *DreamTaq™* polymerase (see chapter 3.1.2) and gene (Table 5) or the vector specific M13 rev (-29) or M13 uni (-21) primers (Eurofins Genomics, Anzinger, Germany). 10 µl of bacterial suspension were used for 5 ml liquid overnight cultures (see chapter 3.1.1).

3.1.6 GENERATION OF GLYCEROL STOCKS

Glycerol stocks of liquid culture were generated by mixing 500 µl of 50 % glycerol with 500 µl of bacterial cell suspension in a cryovial. The glycerol stock was vortexed and immediately stored at -80 °C.

3.1.7 PLASMID EXTRACTION FROM BACTERIA

To isolate plasmid DNA either the *QuickLyse Miniprep* (Qiagen, Venlo, Netherlands) for up to 2 ml of cell culture or the *NucleoBond® Xtra Midi* (Macherey-Nagel, Düren, Germany) kit for up to 200 ml of cell culture were used. Plasmid DNA was extracted according to the manufacturer's specifications and stored at -20 °C.

3.1.8 PLASMID CLONING

To generate new vectors harboring a specific gene, the desired sequence was amplified by PCR (see chapter 3.1.2) using the *Phusion* polymerase and gene specific primers (Table 5). As a template 2 µl of complementary DNA (cDNA) was used (see chapter 0).

Expression vectors

For cloning of expression vectors the primers contained additional restriction enzyme recognition sites (EcoRI and SalI) at the 5' end of the respective primers. After amplification PCR products were separated by agarose gel electrophoresis and extracted (see chapter 3.1.9). In a next step the PCR product was cloned into the pJet1.2/blunt vector using the *CloneJet* (Thermo Fisher Scientific, Waltham, United States) PCR cloning kit according to the manufacturer's specifications. Afterwards, the vector was transformed into competent DH5α *E. coli* cells (see chapter 3.1.3). Colonies were tested using colony PCR (see chapter 3.1.5).

In a next step the generated plasmid was extracted from liquid overnight culture (3.1.7) and the insertion was tested by Sanger sequencing (see chapter 3.1.10). In addition, the plasmid was used as a template for enzymatic digestion. Therefore, the *FastDigest®* (Thermo Fisher Scientific, Waltham, United States) restriction enzymes EcoRI and SalI were used according to the manufacturer's specifications. In addition, the pGEX-6P-1 (GE Healthcare, Chicago, United States) was digested as well. Both products were purified using gel electrophoresis and extracted (see chapter 3.1.9). The PCR product with the additional restriction sites and the digested were ligated in a subsequent step using the *T4* (Thermo Fisher Scientific, Waltham, United States) DNA ligase according to the manufacturer's specifications. The ligation product was transformed into *E. coli* cells (see chapter 3.1.3) and colonies were tested by colony PCR (see chapter 3.1.5). The sequence of the final expression vector was again tested by Sanger sequencing (see chapter 3.1.10). Finally, a glycerol stock of the successfully transformed *E. coli* cells was generated (see chapter 3.1.6) and the vector was stored at -20 °C

BiFC vectors

For constructs for bimolecular fluorescence complementation (BiFC) experiments primers contained additional recombination sites. These were necessary to use the 2in1 ratiometric BiFC system [154]. Cloning was performed as described in Grefen, C. & Blatt, M. R. (2012) [154] by using the *Gateway™* system (Thermo Fisher Scientific, Waltham, United States) according to the manufacturer's specifications. As intermediate vectors Gateway™ pDONR™ 221 006 and 007 (Thermo Fisher Scientific, Waltham, United States) were used. The destination vector was the BiFCt_2in1-CN vector (Table 6). Final constructs were verified by Sanger sequencing (see chapter 3.1.10).

Table 5
Primers used for plasmid cloning

ID	Gene	Sequence (5' to 3')	Reference
JK65	*CHS1*	GCGAATTCATGTCTACTTCTTATTC	This thesis
JK66	*CHS1*	GCGTCGACTTATCTTTGGGATGCTT	This thesis
JK86	*CPK5*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCA TGGGCAATTCTTGCCGTGG	This thesis
JK87	*CPK5*	GGGGACAACTTTGTATAGAAAAGTTGGGTGCT ACGCGTCTCTCATGCTAATG	This thesis
JK89	*UBC8*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCA TGGCTTCGAAACGGATCTTG	This thesis
JK90	*UBC8*	GGGGACAACTTTGTATAGAAAAGTTGGGTGTT AGCCCATGGCATACTTCTGAG	This thesis
JK92	*UBC10*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCA TGGCGTCGAAGCGGATCTTG	This thesis
JK93	*UBC10*	GGGGACAACTTTGTATAGAAAAGTTGGGTGTT AGCCCATGGCATACTTCTGAG	This thesis
JK95	*UBC28*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCA TGGCTTCGAAAAGGATCTTG	This thesis
JK96	*UBC28*	GGGGACAACTTTGTATAGAAAAGTTGGGTGTC ATCCCATTGCATATTTCTGAG	This thesis
JK98	*UBC29*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCA TGGCAACGAGACGGATATTG	This thesis
JK99	*UBC29*	GGGGACAACTTTGTATAGAAAAGTTGGGTGTT AAAACAAGGCATACTTCTGGG	This thesis
JK136	*UBC30*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCA TGGCGTCGAAAAGAATTAAC	This thesis
JK137	*UBC30*	GGGGACAACTTTGTATAGAAAAGTTGGGTGTC ATCCCATTGCATACTTCTGAGTCC	This thesis
JK139	*UBC37*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCA TGGCTCAAGCAGCGAGATTAAGCC	This thesis
JK140	*UBC37*	GGGGACAACTTTGTATAGAAAAGTTGGGTGTC AAGGTCTACATTTGAGGACTCTCC	This thesis
JK156	*UBC9*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCA TGTTTTTCTTCTTCTTAATTCATC	This thesis
JK157	*UBC9*	GGGGACAACTTTGTATAGAAAAGTTGGGTGTC AGCCCATGGCATACTTTTGGGTCCAG	This thesis
JK170	*UBC24*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCA TGGAAATGTCCCTTACTGACTC	This thesis
JK171	*UBC24*	GGGGACAACTTTGTATAGAAAAGTTGGGTGTT ATGATTCTGGTCCAATCTCTTG	This thesis
JK173	*UBC33*	GGGGACAAGTTTGTACAAAAAAGCAGGCTTAT GGCAGAAAAAGCTTGTAT	This thesis
JK174	*UBC33*	GGGGACAACTTTGTATAGAAAAGTTGGGTGTC ACAGCTGAAGCAAAGGCAAC	This thesis

Table 6
Vectors used for plasmid cloning

Vector	Reference
pJet1.2/blunt	Thermo Fisher Scientific, Waltham, United States
pGEX-6P-1	GE Healthcare, Chicago, United States
Gateway™ pDONR™ 221 006	Thermo Fisher Scientific, Waltham, United States
Gateway™ pDONR™ 221 007	Thermo Fisher Scientific, Waltham, United States
BiFCt_2in1-CN	Grefen and Blatt, 2012 [154]

3.1.9 AGAROSE GEL ELECTROPHORESIS AND DNA EXTRACTION

DNA fragments were separated according to their size using agarose gel electrophoresis. The agarose concentration was chosen based on the expected fragment size. Agarose was dissolved in 1x TAE buffer (40 mM Tris-HCl pH 8.5, 20 mM acetic acid, 5 mM EDTA) using a microwave. The agarose gel was poured into a casting stand and ethidium bromide was added to a concentration of 0.05 µg ml^{-1}. As a DNA ladder the *GeneRuler™ 1kb Plus* (Thermo Fisher Scientific, Waltham, United States) was used according to the manufacturer's specifications.

Separation of DNA fragments was performed by applying 120-140 V to the agarose gel for 30-60 min, depending on the size of the DNA fragments and the concentration of the gel. Fragments were documented using the *E-BOX VX2 gel documentation system* (VWR International, Darmstadt, Germany). If necessary, DNA was extracted using the *Hi Yield*® (Süd-Laborbedarf, Gauting, Germany) gel extraction kit according to the manufacturer's specifications.

3.1.10 SANGER SEQUENCING

To determine the correct sequence of a plasmid or the T-DNA insertion site in a plant, DNA was sequenced by Sanger sequencing. This was done by mixing the extracted plasmid (see chapter 3.1.7) or gDNA (see chapter 3.2.4) with the desired primers in a reaction tube according to the service provider's instructions. Sanger sequencing was performed by Eurofins Genomics (Anzinger, Germany).

3.2 PLANT HANDLING

3.2.1 PLANT MATERIAL

The following *A. thaliana* plants were used and generated in this thesis.

Table 7
Wild type *A. thaliana* plant and insertion lines

ID	ML-ID	Name	Stock-ID	Affected Gene	Reference
ED37	-	*Col-0*	-	-	Gerlach, E.-M. (2018) [155]

Table 7
continued

ID	ML-ID	Name	Stock-ID	Affected Gene	Reference
bTB1425	-	*saul1-1*	SALK-063974	*SAUL1*	Raab, S. *et al.* (2009) [123]
bTB1294	ML106	*at3g28580*	SALK_131046	*AT3G28580*	Meyer, S. (2018) [156]
bTB1412	-	*saul1-1 x at3g28580*	-	*SAUL1, AT3G28580*	Meyer, S. (2018) [156]
bTB1476	ML109	*trx5*	SALK_013684	*TRX5*	Meyer, S. (2018) [156]
bTB1386	-	*saul1-1 x trx5*	-	*SAUL1, TRX5*	Meyer, S. (2018) [156]
bTB1614	ML110	*at4g16260*	SALK_031479	*AT4G16260*	Marusoi, S. (2017) [157]
bTB1641	-	*saul1-1 x at4g16260*	-	*SAUL1, AT4G16260*	Marusoi, S. (2017) [157]
bTB1296	ML112	*at1g45145*	SALK_144259	*AT1G45145*	This thesis
bTB1305	-	*saul1-1 x at1g45145*	-	*SAUL1, AT1G45145*	This thesis
bTB1250	ML116	*wrky46*	SALK_134310	*WRKY46*	This thesis
bTB1387	-	*saul1-1 x wrky46*	-	*SAUL1, WRKY46*	This thesis

3.2.2 STERILIZATION, SOWING AND STRATIFICATION

All seeds were sterilized using chlorine gas. Accordingly, they were incubated for at least 2 h in an exicator in which a beaker with 25 ml sodium hypochlorite was placed, to which 2 ml of hydrochloric acid were added. Afterwards seeds were placed on soil containing two parts *Fruhstorfer Pikiersubstrat* (Wilsaflor, Neulehe, Germany), one part coarse-grained sand and one third part expanded clay *Fibobau* (Fibo ExClay, Lahmstedt, Germany). Prior to sowing, soil was treated for at least 2 h with a *Neudomück*® (Neudorff, Emmerthal, Germany) solution with a concentration of 2 ml l^{-1}. In a next step the soil was treated with a *Previcur*® *Energy* (Bayer CropScience, Monheim, Germany) solution with a concentration of 2.5 ml l^{-1}. After sowing, seeds were stratified by incubating the pots for 48 h at 4 °C in the dark.

3.2.3 GROWTH CONDITIONS

Depending on intended use, plants were grown using three different conditions. Plants, which were used for crossings and genotyping, were grown in climate chambers (Weiss Klimatechnik, Reiskirchen-Lindenstruth, Germany). As a more controlled environment was necessary for temperature-induction experiments, these plants were grown in *AR-36L3* growth cabinets (Percival Scientific, Perry, United States). In both cases plants were grown under long day conditions with 16 h light and 8 h dark. As for plants for protoplast generation, these were grown in *TS-110 GroBanks*® (CLF Plant Climatics, Wertingen, Germany) and under short day conditions with 8 h light and 16 h dark. In case of growth cabinets and grow banks, light intensity was set between 90 to 100 µmol m^{-2} s^{-1}.

3.2.4 gDNA EXTRACTION

For gDNA extraction from *A. thaliana* a small leaf was harvested and transferred into a 1.5 ml reaction tube. In a next step 200 µl lysis buffer (200 mM Tris-HCl pH 7.5-8, 250 mM NaCl, 25 mM

EDTA, 0.5 % SDS) were added and the leaf material was grounded using a battery drill with an attached pestle. Afterwards the sample was centrifuged for 5 min at 16873 x g and room temperature (RT). The supernatant was transferred into a new reaction tube containing 300 µl isopropyl alcohol and precipitated for at least 30 min at -20 °C. In a subsequent step the sample was centrifuged for 8 min at 16873 x g and RT and the supernatant was discarded afterwards. The pellet was washed using 1 ml 70 % ethanol and centrifuged again for 3 min at 16873 x g and RT. The supernatant was removed and the pellet dried. The dried pellet was resuspended in 40-60 µl H_2O by incubating the reaction tube for 10 min at 50 °C and 900 rpm using a *Thermomixer Compact* (Eppendorf, Hamburg, Germany). The extracted gDNA was stored at -20 °C.

3.2.5 GENOTYPING AND T-DNA INSERTION DETERMINATION

To analyze plants and determine if they are hetero- or homozygous for a mutation genotyping using PCR was performed. Therefore, a small leaf was harvested and gDNA extracted (see chapter 3.2.4). For genotyping 2 µl of gDNA were used. The PCR was performed using specific genotyping primers (Table 8A), which were designed flanking the T-DNA insertion using the *T-DNA Primer Design* tool (The Salk Institute for Biological Studies, La Jolla, United States) [158]. Primers were used in combinations to detect either the wild type or the mutant allele (Table 8B). The PCR was performed using the *DreamTaq™* polymerase (see chapter 3.1.2) and analyzed using gel electrophoresis (see chapter 3.1.9).

T-DNA insertion sites were determined by Sanger sequencing (see chapter 3.1.10) using extracted gDNA of the respective plants (see chapter 3.2.4) and gene and T-DNA insertion specific primers (Table 8C).

Table 8

(A) Primer used for genotyping and sequencing

ID	Gene	Sequence (5' to 3')	Reference
521	*T-DNA*	ATTTTGCCGATTTCGGAAC	Gerlach, E.-M. (2018) [155]
ED30	*SAUL1*	TTTCCCCATTCATGAGTGAAG	Gerlach, E.-M. (2018) [155]
ED29	*SAUL1*	TGAGGCCAATCAAATGATTTC	Gerlach, E.-M. (2018) [155]
694	*TRX5*	GATTTGATTACGTTATGGTTTAGGG	Meyer, S. (2018) [156]
695	*TRX5*	ATAGGGAAATGTTTGAATCTACAGG	Meyer, S. (2018) [156]
661	*AT5G52760*	CTCCACGTCACCCTAGACTTG	This thesis
662	*AT5G52760*	TTCTGGAGTAGTTTCTCGAGCC	This thesis
655	*AT3G28580*	GATCAGGCCTGTTGAGTTTTG	Meyer, S. (2018) [156]
656	*AT3G28580*	ATGCTCTCCGCTGTATTCATG	Meyer, S. (2018) [156]
664	*AT4G16260*	ACGTTGTTTACTCTGCCGTTG	Marusoi, S. (2017) [157]
665	*AT4G16260*	ACCTTAGGGCTTTGTTTGCTG	Marusoi, S. (2017) [157]
701	*WRKY46*	TCTGTCGATTCCAACAAAACC	This thesis
702	*WRKY46*	AAGCCAATTTTTATCCATCGG	This thesis

(B) Primer combinations used for genotyping

Gene	Primer 1	Primer 2	Product Size (bp)	Allele type
SAUL1	ED30	ED29	1166	WT
SAUL1	521	ED29	539-839	Mutant
TRX5	694	695	826	WT
TRX5	521	695	484-784	Mutant
AT5G52760	661	662	1233	WT

Table 8
(B) continued

Gene	Primer 1	Primer 2	Product Size (bp)	Allele type
AT5G52760	521	662	549-849	Mutant
AT3G28580	655	656	1129	WT
AT3G28580	521	656	565-865	Mutant
AT4G16260	664	665	1140	WT
AT4G16260	521	665	610-910	Mutant
WRKY46	701	702	1129	WT
WRKY46	521	702	575-875	Mutant

(C) Primer used for T-DNA localization

Gene	Primer 1	Primer 2
TRX5	521	695
AT5G52760	521	662
AT3G28580	521	656
AT4G16260	521	665
WRKY46	521	702

3.2.6 PROTOPLAST GENERATION AND TRANSFORMATION

To analyze protein-protein interactions and protein localization, *Arabidopsis thaliana* (*A. thaliana*) mesophyll protoplasts were generated. For that purpose, around 40 leaves of 4 to 6 week-old *Col-0* plants, which were grown as previously described (3.2.3), were cut into thin strips and placed into a petri dish filled 0.5x MCP buffer. Afterwards the buffer was replaced with an enzyme solution and lightproof wrapped with aluminum foil. The petri dish was incubated for 2 h at 26 °C at 60 rpm using a *Unimax 1010* (Heidolph Instruments, Schwabach, Germany). In a next step, protoplasts were detached using an incubation of 1 min at 26 °C at 80 rpm. Protoplasts were isolated by filtration using a nylon mesh with an opening size of 50 µm, which was moistened with MaMg buffer. Protoplasts were precipitated at 100 x g for 3 min at 23 °C with a very long deceleration. Afterwards the supernatant was discarded and protoplasts were cautiously resuspended using 20 ml MaMg buffer. In a next step, protoplasts were resuspended as described before. Protoplasts were cautiously resuspended using 1 ml MaMg buffer.

For each transformation 150 µl of protoplast suspension were transferred into 14 ml *#187261* reaction tubes (Greiner Bio-One, Kremsmünster, Austria) and 20 µg of the plasmid of interest were added (Table 10). Immediately thereafter 165 µl PEG-Ca buffer was added rotating the reaction tube cautiously. Reaction tubes were incubated for 30 min at RT in the dark. Afterwards 500 µl, 1 ml and 1.5 ml W5 buffer were added every 5 min rotating the reaction tube cautiously. Protoplasts were precipitated at 60 x g for 3 min at 23 °C with a very long deceleration. Afterwards the supernatant was discarded. Protoplasts were resuspended using 3 ml W5 buffer rotating the reaction tube cautiously. Protoplasts were precipitated at 60 x g for 3 min at 23 °C with a very long deceleration. In a final step, protoplasts were resuspended using 3 ml W5 buffer rotating the reaction tube cautiously and incubated at RT in the dark overnight. Afterwards protoplasts were used for microscopy.

Table 9
Buffers used for protoplast isolation and transformation

Buffer	Components	Comment
MCP	10 mM MES-HCl at a pH of 5.6, 500 mM Sorbitol, 1 mM CaCl	Autoclaved
Enzyme solution	0.25 % Macerozyme R-10, 1 % Cellulase	In MCP buffer
MaMg	5 mM MES-HCl at a pH of 5.6, 450 mM Sorbitol, 15 mM $MgCl_2$	Autoclaved
PEG-Ca	61.5 % w/w PEG 4000, 300 mM Mannitol, 150 mM $CaCl_2$	
W5	2 mM MES-NaOH at a pH of 5.7, 154 mM NaCl, 125 mM $CaCl_2$, 5 mM KCl, 5 mM Glucose	

Table 10
BiFC vectors transformed into protoplasts. Proteins were either C-terminally (Protein 1) or N-terminally (Protein 2) tagged. * In this case was SAUL1 N-terminally tagged.

ID	Name	Vector	Protein 1	Protein 2	Reference
bJK33	BiFC_CN_SAUL1_UBC8	BiFCt_2in1-CN	SAUL1	UBC8	This thesis
bJK57	BiFC_CN_SAUL1_UBC9	BiFCt_2in1-CN	SAUL1	UBC9	This thesis
bJK32	BiFC_CN_SAUL1_UBC10	BiFCt_2in1-CN	SAUL1	UBC10	This thesis
bJK46	BiFC_CN_SAUL1_UBC13	BiFCt_2in1-CN	SAUL1	UBC13	Meyer, S. (2018) [156]
bJK47	BiFC_CN_SAUL1_UBC17	BiFCt_2in1-CN	SAUL1	UBC17	Meyer, S. (2018) [156]
bJK38	BIFC_CN_SAUL1_UBC28	BiFCt_2in1-CN	SAUL1	UBC28	This thesis
bJK39	BiFC_CN_SAUL1_UBC29	BiFCt_2in1-CN	SAUL1	UBC29	This thesis
bJK37	BiFC_CN_SAUL1_UBC30	BiFCt_2in1-CN	SAUL1	UBC30	This thesis
gTL89	BiFC_CN_SAUL1_UBC33	BiFCt_2in1-CN	SAUL1	UBC33	This thesis
bJM352	BiFC_NN_SAUL1_UBC37	BiFCt_2in1-NN	SAUL1*	UBC37	Marusoi, S. (2017) [157]

3.2.7 RNA ISOLATION AND cDNA SYNTHESIS

To obtain RNA for transcriptional profiling three 12-day old seedlings, which were grown on soil, were harvested and frozen immediately in a 2 ml reaction tube with 2 steel balls in liquid nitrogen. Samples were stored at -80 °C.

For preparation, samples were mechanically shredded using a *TissueLyser* (Qiagen, Venlo, Netherlands) for 2 min at a frequency of 30 s^{-1}. Afterwards RNA was isolated using the *innuPREP Plant RNA Kit* (analytikjena, Jena, Germany) according to the manufacturer's specifications. RNA was stored at -80 °C. cDNA was synthesized using the *QuantiNova™* (Qiagen, Venlo, Netherlands) reverse transcription kit according to the manufacturer's specifications. cDNA was stored at -20 °C.

3.2.8 REVERSE TRANSCRIPTION PCR (RT-PCR)

To analyze the presence of specific gene transcripts in mutant plants and to test for a knock-out, reverse transcription PCR (RT-PCR) was performed. Therefore, RNA was isolated and cDNA was synthesized as described in chapter 3.2.7. Afterwards 1-2 µl of cDNA were used for RT-PCR. The RT-PCR was performed using specific RT-PCR primers (Table 11) and the *DreamTaq™* (Thermo Fisher Scientific, Waltham, United States) polymerase according to the manufacturer's

specifications. Primers were designed to detect full-length and partial cDNAs that are still expressed despite the T-DNA insertion. The RT-PCR was performed using the PCR program depicted in table 5.

Table 11
(A) Primers used for RT-PCR

ID	Gene	Sequence (5' to 3')	Reference
bTB149	*TRX5*	TTTGGAACGAGAAGGTCAAA	Meyer, S. (2018) [156]
bTB150	*TRX5*	ATGTTGCCTTCTTTCATGAA	Meyer, S. (2018) [156]
JK220	*AT5G52760*	ACGCTGTGTTGCAATTGAGT	This thesis
JK221	*AT5G52760*	CGTCTACCACTCTACAATTTCCG	This thesis
JK222	*AT5G52760*	TGTACCAGCTGTCGTGATGA	This thesis
bTB147	*AT3G28580*	ACCATTCCTTTATAGACTCTTCG	Meyer, S. (2018) [156]
bTB148	*AT3G28580*	TGTTCTTTTGCTCTATCGTCT	Meyer, S. (2018) [156]
777	*AT4G16260*	GCGGCAGAGAATGTGATCGG	Marusoi, S. (2017) [157]
779	*AT4G16260*	ATGACCACGTTATTCCTCCT	Marusoi, S. (2017) [157]
780	*AT4G16260*	GTCATCATCGGCGTTCCAAA	Marusoi, S. (2017) [157]
781	*AT4G16260*	CGTGGACCAAAACAAAGCCG	Marusoi, S. (2017) [157]
JK223	*WRKY46*	GGGAAAGAGCTTGCTAACCG	This thesis
JK224	*WRKY46*	CGACCACAACCAATCCTGTC	This thesis
JK225	*WRKY46*	TTCTCCAGCAGTGACCATCA	This thesis

(B) PCR conditions used for RT-PCR

Gene	Primer 1	Primer 2	Product Size (bp)	Position (from TSS)
TRX5	bTB149	bTB150	232	+205
AT5G52760	JK220	JK221	345	Full length
AT5G52760	JK221	JK222	212	+580
AT3G28580	bTB147	bTB148	413	+250
AT4G16260	779	777	1010	Full length
AT4G16260	779	780	120	+203
WRKY46	JK223	JK224	843	Full length
WRKY46	JK223	JK225	289	+244

3.2.9 QUANTITATIVE REAL-TIME PCR (qRT-PCR)

To determine the amount of specific gene transcripts in a sample a quantitative real-time PCR (qRT-PCR) was performed. All measurements were carried out using the *RT² SYBR® Green qPCR Mastermix* and the *Rotor-Gene® Q* (Qiagen, Venlo, Netherlands). Each cDNA sample was diluted 1:10 and 2 µl were used for each measurement (Table 12). 40 cycles of a specific PCR program were used to amplify the desired product (Table 13).

Genes were chosen as references according to an *in silico* analysis using *Genevestigator v6.01.05* [159] and the *RefGenes* tool. This allowed to determine genes which were thought not to be regulated by a temperature-shift as applied in the performed experiments. Primers for genes of interest (GOI) and reference genes (Table 14) were optimized to an annealing temperature of 60 °C using *Primer3web v4.1.0* [160]. Cycle (C_t) and amplification efficiency values (E) were obtained using the *Rotor-Gene® Q v2.3.1.49* (Qiagen, Venlo, Netherlands) software. Normalized calibrated relative quantity (NCRQ) values were calculated by using an internal calibrator (IC), consisting of a cDNA

mixture, for each run and normalizing these calibrated values using three reference primer sets. Calculations were done accordingly to Hellemans, J. *et al.* (2007) [161].

Table 12
Composition of each qRT-PCR sample

Component	Volume (µl)
RT² SYBR® Green qPCR Mastermix	5
Primer (each)	0.5
cDNA	2
H_2O	2

Table 13
PCR program used for the qRT-PCR. Steps 2 and 3 were iterated 40 times.

Step	Temperature (°C)	Time (s)
1	95	300
2	95	10
3	60	30
4	65-95	5 $°C^{-1}$

Table 14
Primers used for qRT-PCR.

ID	Gene	Sequence (5' to 3')	Usage
JK43	*WRKY53*	GCCTCTCTCTGGGCTTATTCT	GOI
JK44	*WRKY53*	CGAGAGTCATCATCGCCAAG	GOI
JK45	*ICS1*	TGGCAAGATCGCTGTTGAAT	GOI
JK46	*ICS1*	AGCCAACATTGAACTTCCACC	GOI
JK49	*RPL10*	TGGTCACCATGCTCAAGAGG	Reference gene
JK50	*RPL10*	CTTCGTGAAGCCCCATTTCC	Reference gene
JK55	*EER1*	AAGATGACCAAGATTCTGTGAGA	Reference gene
JK56	*EER1*	CAGGTAAAATGCGTGCAACAC	Reference gene
JK57	*VDAC1*	GCCTCCCTTACTGTGAACGA	Reference gene
JK58	*VDAC1*	GCTCACTTCGGCTCCAAC	Reference gene
JK176	*TRX5*	CATCATTGATCGTGTTGTCGG	GOI
JK177	*TRX5*	CAGAAGCTACAAGACCACCA	GOI
JK178	*AT1G52030*	GAGACAAGAAGCTCCACAAGA	GOI
JK179	*AT1G52030*	AAACTTCAGCGACGAAAGGT	GOI
JK180	*MBP1*	AGGAAGGATCTTTTCATGGTGTG	GOI
JK181	*MBP1*	AATAAGATACTCGGCTGGATGG	GOI
JK182	*THI2.1*	ACTCTGATACATGCCCTCGAG	GOI
JK183	*THI2.1*	CACCCTAACTTGCAGTGCTC	GOI
JK184	*ERF2*	ACTTGCTAGGAGGAGGAGGA	GOI
JK185	*ERF2*	AACCTCCCCAACTCTCTGTG	GOI
JK186	*AT5G52760*	AAAGCGCTTGTCACCGTTTC	GOI
JK187	*AT5G52760*	TGGTACATCAACTTCTCCAACT	GOI
JK188	*ABR*	GACAAGGCCAAGGATGCTG	GOI

Table 14
continued

ID	Gene	Sequence (5' to 3')	Usage
JK189	*ABR*	ACGACATCAGCAGCTCCTT	GOI
JK190	*AFL1*	CCACGAGAATTGTCCTAACTCAG	GOI
JK191	*AFL1*	CACACTTTCACCAACCTCCG	GOI
JK192	*AT3G28580*	ACTCTTCGGTCGCTTCTACC	GOI
JK193	*AT3G28580*	AGTAGGCTTCGCTTCTCTTGA	GOI
JK194	*ZAT7*	TGGTGGTCACATGAGGAGAC	GOI
JK195	*ZAT7*	TTTCTTCAAAGCCGTCACCG	GOI
JK196	*AT3G46370*	GTGATGAATTTTCCACAGTCTGA	GOI
JK197	*AT3G46370*	CCAGGAGGTTCTATTCAATCTGT	GOI
JK198	*CML41*	TCCGTCAAGTCTTCAGCCAT	GOI
JK199	*CML41*	ACCGACAGAGCCGAAGTAAT	GOI
JK200	*AT3G62460*	AGTTTCACTCCGCCGGTATC	GOI
JK201	*AT3G62460*	CCCATGCCAGTTAAAACAGC	GOI
JK202	*AT2G15020*	GCCTCTTCTGGTTGTGTGTG	GOI
JK203	*AT2G15020*	TGACATGTAAGCGCGTTCAC	GOI
JK204	*AT2G39310*	CAAAGACCTTGTACCCGGAG	GOI
JK205	*AT2G39310*	ATTCACCATCTTCAAGCACAAAC	GOI
JK206	*AT4G13575*	TAACCCGTAAAAGCCGCTTG	GOI
JK207	*AT4G13575*	ACGTTGTATTCTAAACCCATGCA	GOI
JK208	*AT4G16260*	TGGTGACGTTAGATGGTATATCG	GOI
JK209	*AT4G16260*	GGATAGATGTTGGCTAGTAACGC	GOI
JK210	*ORF149*	AGCCCCACGGACTTCATATT	GOI
JK211	*ORF149*	TTGTGCCGAGTCAAATCTGC	GOI
JK212	*NDHA*	TTGGGGATGGAATTTGTGGC	GOI
JK213	*NDHA*	ACGGTAACCTCTCGCATTCT	GOI

3.3 TRANSCRIPTOME ANALYSIS

3.3.1 RNA ISOLATION

RNA was isolated as described in chapter 3.2.7.

3.3.2 LIBRARY PREPARATION AND SEQUENCING

Library preparation and sequencing was conducted by the next-generation sequencing (NGS) technology platform in the Virus Genomics group of Adam Grundhoff (Heinrich Pette Institute, Hamburg, Germany).

In a first step RNA integrity was analyzed using the *RNA 6000 Nano Chip* on a *2100 Bioanalyzer* (Agilent, Santa Clara, United States). Next, mRNA was extracted using the *NEBNext® Poly(A) mRNA Magnetic Isolation module* (New England Biolabs, Ipswich, United States) and RNA-Seq libraries were generated using the *NEBNext® Ultra™ RNA Library Prep Kit for Illumina®* (New England Biolabs, Ipswich, United States) according to the manufacturer's specifications. Afterwards, concentrations were measured with a *Qubit® 2.0 Fluorometer* (Thermo Fisher Scientific, Waltham, United States) and fragment lengths distribution of the final libraries were analyzed using the *DNA*

High Sensitivity Chip on *2100 Bioanalyzer* (Agilent, Santa Clara, United States). In a last step all samples were normalized to 2 nM and pooled equimolarly.

The library pool was sequenced on a *HiSeq 2500* (Illumina, San Diego, United States) with 1 x 50 bp and 8.3 to 14 x 10^6 reads per sample.

3.3.3 MAPPING AND DIFFERENTIAL EXPRESSED GENE ANALYSIS

Quality control of raw reads was assessed using *MultiQC v1.5.* Afterwards reads were mapped against the TAIR10.34 database [162,163] and counted using *STAR v2.5.3a* [164,165]. Differential expressed gene analysis was performed using the *R v3.4.0* package *DESeq v1.32.0* [166]. Genes were defined as being differentially expressed having a Benjamini-Hochberg adjusted $p \leq 0.01$. Fold changes were always depicted as the binary logarithm. Z-score was defined as the deviation of the normalized read count (reads per million reads) of a sample from the mean normalized read count of all samples of one gene.

Table 15
Software used for RNA-seq read mapping and analysis

Usage	Software
Quality control	*MultiQC v1.5*
Read mapping	*STAR v2.5.3a* [164,165]
Differential expression analysis	*DESeq v1.32.0* [166] package in *R v3.4.0*

3.3.4 GENE ONTOLOGY AND EXPRESSION ANALYSIS

Gene ontology analysis was performed using *Cytoscape v3.4.0* [167] and the *ClueGo v2.3.5* [168] and *CluePedia v1.3.5* [168] apps. *Biological Process* terms of the *Gene ontology* (GO) [169,170] and pathways of the *Kyoto Encyclopedia of Genes and Genomes* (KEGG) [171–173] for *A. thaliana*, dated November 21, 2017, were used as data bases. Using these, GO tree intervals 3 to 8 were analyzed for enrichment. Terms were considered being enriched having a Bonferroni adjusted $p \leq 0.01$ and consisting initially of 1 % of the analyzed genes. As a next step GO terms were fused if they shared a common parent or child term. In a last step groups were clustered together if they shared at least 40 % of their genes based on their *kappa* score [168].

Expression clustering and heatmap analysis were carried out using the *heatmap.2* function of the *gplots v3.0.1 R v3.4.0*-package. Clusters were build using the *hclust* option.

Similar expression patterns were screened and data base analysis was performed using *Genevestigator v6.01.05* [159] with mRNA-seq data and applying the *Signature* or the *Perturbations* tool (Table 17).

Table 16
Software employed for gene ontology and expression analysis

Usage	Software
Gene ontology analysis	*Cytoscape v3.4.0* [167] using *ClueGo v2.3.5* [168] and *CluePedia v1.3.5* [168]
Heatmap analysis	*heatmap.2* of the *gplots v3.0.1 R v3.4.0* package
Expression patterns	*Genevestigator v6.01.05* [159] using the *Signature* tool

Table 17
Data sets used for cluster comparison and stress-induced expression change *Genevestigator* analysis

Usage	Condition	Database ID	Reference
Expression clustering	*cpr1-4 aba1-8*	GSE111238	Cai, Q. *et al.* (2018) [174]
	zed1-D	GSE95665	Wang, Z. *et al.* (2017) [175]
	Infection of Col-0 with *Pto* DC3000 / *Pto* AvrRpt2	GSE88798	Mine, A. *et al.* (2018) [176]
	Infection of *pad4-1* with *Pto* DC3000 / *Pto* AvrRpt2	GSE88798	Mine, A. *et al.* (2018) [176]
	flg22 treatment	GSE109150	Birkenbihl, R. P. *et al.* (2018) [38]
Genevestigator analysis	Drought stress	E-MTAB-3279	Clauw, P. *et al.* (2015) [177]
	Cold stress	GSE63406	Schlaen, R. G. *et al.* (2015) [178]
	Heat stress	GSE69077	Rawat, V. *et al.* (2015) [179]
	Phosphate stress	GSE74856	Liu, T. Y. *et al.* (2016) [180]
	Oxidative Stress	GSE65740	Xu, E. *et al.* (2015) [181]
	ABA treatment	GSE66737	Zhan, X. *et al.* (2015) [182]
	flg22 treatment	GSE63603	Li, B. *et al.* (2015) [183]
	Coronatine treatment	GSE64397	Liu, M.-J. *et al.* (2015) [184]
	Infection of Col-0 with *Pto* DC3000	E-MTAB-4450	Howard, B. E. *et al.* (2013) [185]
	Infection of Col-0 with *Pto* AvrRps4	E-MTAB-4450	Howard, B. E. *et al.* (2013) [185]
	NLR-dependent immunity	GSE39463	Maekawa, T. *et al.* (2012) [186]
	35S::*RPS4-HS* overexpression	GSE40329	Blanvillain-Baufumé, S. *et al.* (2012) [187]

3.3.5 PROMOTER ANALYSIS

To discover potential binding motifs of transcription factors in promoter regions (cis-elements), *MEME* v5.0.2 [188] and *TOMTOM v5.0.2* [189], which were implemented as a part of the *MEME suite v5.0.2* [190] were used. If not stated otherwise, sequences starting at the transcription start site (TSS) of a gene, with a length of 3 kbp, heading upstream, were used as a data basis. Detected motifs had to have a width between 5 and 50 bp.

Table 18
Software used for promoter analysis

Usage	Software
Motif detection	*MEME v5.0.2* [188]
Motif alignment	*TOMTOM v5.0.2* [189]

3.3.6 TRANSCRIPTION FACTOR ANALYSIS

To analyze which transcription factors (TF) were differentially expressed, the analyzed genes needed to be annotated. Therefore, a custom non-redundant TF database was generated, consisting of TF which were described in the TAIR gene family database [162,163], PlnTFDB v3.0 [191] and PlantTFDB v3.0 [192,193]. All databases were concatenated and family names harmonized, when possible.

Table 19
Databases used for the TF analysis

Database	Source	Accessed
TAIR gene family database [162,163]	https://www.arabidopsis.org/download/index-auto.jsp?dir=%2Fdownload_files%2FGenes%2FGene_families	2015-01-07
PlnTFDB v3.0 [191]	http://plntfdb.bio.uni-potsdam.de/v3.0/	2015-01-06
PlantTFDB v3.0 [192,193]	http://planttfdb.cbi.pku.edu.cn/index.php	2015-01-06

3.4 RECOMBINANT PROTEIN PURIFICATION

3.4.1 PROTEIN EXPRESSION

All proteins described and analyzed in this thesis were expressed in *E. coli* using the pGEX-6P-1 vector system. Constructs were either generated during this thesis as described in chapter 3.1.8 or were provided by Catharina Brieske or Marcel Bhattarai [142,194]. The following constructs were used (Table 20). The constructs were transformed (see chapter 3.1.4) into different *E. coli* strains for protein expression. Multiple strains were used during this thesis (Table 21).

Table 20
Expression constructs for recombinant protein expression

Name	Vector	Protein	Tag	Reference
pGEX6P1_SAUL1	pGEX-6P-1	SAUL1	GST	Brieske, 2017 [142]
pGEX6P1_BON1	pGEX-6P-1	BON1	GST	Bhattarai, 2017 [194]
pGEX6P1_CHS1	pGEX-6P-1	CHS1	GST	This thesis
pGEX6P1_AtKRP125b	pGEX-6P-1	AtKRP125b	GST	Brieske, 2017 [142]

Table 21
E. coli strains used for recombinant protein expression

Strain	Genotype	Reference
BL21-Gold (DE3)	*E. coli* B. F$^-$ *ompT hsdS*(r_{B^-}, m_{B^-}) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* Hte	Agilent, Santa Clara, United States
BL21 Star™ (DE3)	*E. coli* B. F$^-$ *ompT hsdS*(r_{B^-}, m_{B^-}) *galdcmrne*131 (DE3)	Thermo Fisher Scientific, Waltham, United States
BL21-CodonPlus (DE3) RIPL	*E. coli* B. F$^-$ *ompT hsdS*(r_{B^-}, m_{B^-}) *dcm*$^+$ Tet^r *gal* λ(DE3) *endA* The [*argU proL Cam*^{r}] [*argU ileY leuW Strep/Spec*^{r}]	Agilent, Santa Clara, United States
Origami™ 2 (DE3) pLysS	*E. coli* Δ(*ara*-leu)*7697 ΔlacX74 ΔphoA Pvull phoR araD139 ahpC gale galK rpsL* F'[*lac*$^+$ *lacI*^{q} *pro*] (DE3) gor522::*TN10 trxB* pLysS (Cam^R, Str^R ,Tet^R)	Merck, Darmstadt, Germany

The expression of the recombinant proteins was induced at a specific $OD_{600\,nm}$ with different concentrations of Isopropy-β-D-thiogalactopyranoside (IPTG). Therefore, cells were grown on LB agar plates to a high optical density. Afterwards, cells were resuspended using 2YT medium and

used as a starter culture for the liquid culture in Erlenmeyer flasks with baffles. One petri dish was used per liter of liquid medium. The cells were incubated at 37 C and 180 rpm using either a *MaxQ™ 5000* or a *MaxQ™ 6000* (Thermo Fisher Scientific, Waltham, United States) shaker up to a specific $OD_{600\ nm}$. After induction with IPTG, cells were incubated at 180 rpm at different temperatures for specific amounts of time and harvested afterwards (Table 22). For expression test 700 µl of cell culture were harvested in 1 h intervals.

Table 22
Expression conditions for different constructs

Protein	$OD_{600\ nm}$	IPTG (mM)	Temperature (°C)	Time (h)
GST-SAUL1	1	1	18	18
GST-BON1	0.6	0.5	18	22
GST-CHS1	1	1	18	2

3.4.2 PROTEIN SOLUBILITY SCREENING

Throughout this thesis different solubility screens were performed. As a starting point protein-expressing cells were harvested and aliquots of 12 ml were decanted in 15 ml reaction tubes. Afterwards cells were pelletized for 10 min at 3000 x g and 4 °C and stored at -20 °C. For cell lysis, the pellet was resuspended in a specific lysis buffers library [195], or in custom designed buffer palettes (Table 23-26) and transferred to 1.5 ml sonication-compatible reaction tubes. Subsequently cell walls were disrupted using 1 mg ml^{-1} lysozyme for 1 h at 4 °C and sonication. Sonication was performed using a *Bioruptor® Pico* (Diagenode, Seraing, Belgium) for 5 cycles with 30 s on and 30 s off. Cell debris was pelleted by centrifuging the cell lysate at 14000 x g at 4 °C for 30 min. Pellet and supernatant were analyzed by SDS-PAGE.

Table 23
Initial buffers, which were used for a GST-AtKRP125b solubility screen

No.	Buffer	pH	Ions	Detergent	Additives
1	100 mM Tris-HCl	7.0	200 mM NaCl	0.1 M Urea	1 mM DTT, 1 mM ATP
2	100 mM Tris-HCl	7.0	50 mM NaCl	0.1 M Urea	1 mM DTT, 1 mM ATP
3	100 mM Tris-HCl	7.0	50 mM NaCl	0.1 % Triton	1 mM DTT, 1 mM ATP
4	100 mM Tris-HCl	7.0	10 mM NaCl		1 mM DTT, 1 mM ATP
5	100 mM Tris-HCl	8.0	200 mM NaCl	0.1 M Urea	1 mM DTT, 1 mM ATP
6	100 mM Tris-HCl	8.0	50 mM NaCl	0.1 M Urea	1 mM DTT, 1 mM ATP
7	100 mM Tris-HCl	8.0	50 mM NaCl	0.1 %Triton	1 mM DTT, 1 mM ATP
8	100 mM Tris-HCl	8.0	10 mM NaCl		1 mM DTT, 1 mM ATP
9	50 mM Pipes	7.0	1 mM MgSO4		1 mM DTT, 1 mM ATP, 1 mM EGTA
10	50 mM Pipes	7.0	1 mM MgSO4	0.1 M Urea	1 mM DTT, 1 mM ATP, 1 mM EGTA
11	50 mM Pipes	7.0	1 mM MgSO4	0.1 %Triton	1 mM DTT, 1 mM ATP, 1 mM EGTA
12	10 mM HEPES	7.2	1 mM MgCl2 25 mM NaCl		1 mM DTT, 1 mM ATP, 1 mM EGTA
13	10 mM HEPES	7.2	1 mM MgCl2 25 mM NaCl	0.1 M Urea	1 mM DTT, 1 mM ATP, 1 mM EGTA

Table 23
continued

No.	Buffer	pH	Ions	Detergent	Additives
14	10 mM HEPES	7.2	1 mM MgCl2 25 mM NaCl	0.1 %Triton	1 mM DTT, 1 mM ATP, 1 mM EGTA
15	2x PBS				1 mM DTT, 1 mM ATP
16	2x PBS			0.1 %Triton	1 mM DTT, 1 mM ATP

Table 24
Different buffers, with different lysis conditions, which were used for a GST-AtKRP125b solubility screen

No.	Buffer	pH	Ions	Additives	Cell Lysis
1	100 mM Tris-HCl	8.0	10 mM NaCl	1 mM DTT, 1 mM ATP	1 mg ml^{-1} Lysozyme for 15 min
2	100 mM Tris-HCl	8.0	10 mM NaCl	1 mM DTT, 1 mM ATP	Sonication
3	100 mM Tris-HCl	8.0	10 mM NaCl	1 mM DTT, 1 mM ATP	1 mg ml^{-1} Lysozyme for 15 min, Sonication
4	100 mM Tris-HCl	8.0	10 mM NaCl	1 mM DTT, 1 mM ATP	1 mg ml^{-1} Lysozyme for 15 min, Sonication, 1 mM PMSF
5	100 mM Tris-HCl	8.0	10 mM NaCl	1 mM DTT, 1 mM ATP, 1 mM EGTA	1 mg ml^{-1} Lysozyme for 15 min, 1 mM PMSF
6	10 mM HEPES	7.2	1 mM $MgCl_2$ 25 mM NaCl	1 mM DTT, 1 mM ATP, 1 mM EGTA	1 mg ml^{-1} Lysozyme for 15 min
7	10 mM HEPES	7.2	1 mM $MgCl_2$ 25 mM NaCl	1 mM DTT, 1 mM ATP, 1 mM EGTA	Sonication
8	10 mM HEPES	7.2	1 mM $MgCl_2$ 25 mM NaCl	1 mM DTT, 1 mM ATP, 1 mM EGTA	1 mg ml^{-1} Lysozyme for 15 min
9	10 mM HEPES	7.2	1 mM $MgCl_2$ 25 mM NaCl	1 mM DTT, 1 mM ATP, 1 mM EGTA	1 mg ml^{-1} Lysozyme for 15 min, 1 mM PMSF

Table 25
Conditions, which were used to screen for a buffer in which GST-CHS1 appears to be soluble.

No.	Buffer	pH	Ions	Detergent	Additives
1	100 mM Tris-HCl	7.5	150 mM NaCl		1 mM DTT
2	100 mM Tris-HCl	8.0	150 mM NaCl		1 mM DTT
3	100 mM PIPES-NaOH	7.5	150 mM NaCl		1 mM DTT
4	100 mM Tris-HCl	7.5	150 mM NaCl		1 mM DTT, 5 % Glycerol
5	100 mM Tris-HCl	8.0	150 mM NaCl		1 mM DTT, 5 % Glycerol
6	100 mM PIPES-NaOH	7.0	150 mM NaCl		1 mM DTT, 5 % Glycerol
7	100 mM PIPES-NaOH	7.5	150 mM NaCl		1 mM DTT, 5 % Glycerol

Table 25
continued

No.	Buffer	pH	Ions	Detergent	Additives
8	100 mM Tris-HCl	7.5	150 mM KCl		1 mM DTT, 5 % Glycerol
9	100 mM Tris-HCl	8.0	150 mM KCl		1 mM DTT, 5 % Glycerol
10	100 mM PIPES-NaOH	7.0	150 mM KCl		1 mM DTT, 5 % Glycerol
11	100 mM PIPES-NaOH	7.5	150 mM KCl		1 mM DTT, 5 % Glycerol
12	100 mM Tris-HCl	7.5	150 mM LiCl		1 mM DTT, 5 % Glycerol
13	100 mM Tris-HCl	8.0	150 mM LiCl		1 mM DTT, 5 % Glycerol
14	100 mM PIPES-NaOH	7.0	150 mM LiCl		1 mM DTT, 5 % Glycerol
15	100 mM PIPES-NaOH	7.5	150 mM LiCl		1 mM DTT, 5 % Glycerol
16	100 mM Tris-HCl	7.5	150 mM NaCl	1 % Triton X-100	1 mM DTT, 5 % Glycerol
17	100 mM Tris-HCl	8.0	150 mM NaCOOH		1 mM DTT, 5 % Glycerol

Table 26
Different buffers which were used for a GST-BON1 solubility screen [194].

No.	Buffer	pH	Ions	Detergent	Additives
1	50 mM Tris-HCl	7.5	50 mM NaCl		5 mM EDTA
2	50 mM Tris-HCl	7.5	50 mM NaCl	500 mM Urea	5 mM EDTA
3					
4	50 mM NaOAc	5.0	50 mM NaCl		5 mM EDTA
5	50 mM Tris-HCl	7.0	50 mM NaCl		5 mM EDTA
6	50 mM Tris-HCl	9.0	50 mM NaCl		5 mM EDTA
7	50 mM Tris-HCl	7.5	100 mM NaCl		5 mM EDTA
8	50 mM Tris-HCl	7.5	200 mM NaCl		5 mM EDTA
9	50 mM Tris-HCl	7.5	50 mM KCl		5 mM EDTA
10	20 mM Tris-HCl	7.5	50 mM NaCl		2.5 % Glycerol
11	250 mM Tris-HCl	8.0	50 mM KCl	0.2 % Triton X-100	1 mM EDTA, 1 mM DTT, 1 mM AEBSF
12	10 mM Tris-HCl	7.4	150 mM NaCl		
13	50 mM Tris-HCl	7.5	100 mM NaCl		5 mM EDTA, 1 mM AEBSF
14	50 mM Tris-HCl	7.5	300 mM NaCl		5 mM EDTA, 0.1 mM AEBSF
15	50 mM Tris-HCl	7.5	200 mM NaCl	300 mM Urea	5 mM EDTA, 0.1 mM AEBSF
16	50 mM Tris-HCl	7.5	200 mM NaCl	100 mM Urea	5 mM EDTA, 0.1 mM AEBSF
17	100 mM Tris-HCl	7.5	200 mM NaCl		5 mM EDTA, 0.1 mM AEBSF
18	250 mM Tris-HCl	7.5	150 mM NaCl		5 mM EDTA, 0.1 mM AEBSF
19	50 mM Tris-HCl	7.5	200 mM NaCl	100 mM Urea	5 mM EDTA, 0.1 mM AEBSF
20	50 mM Tris-HCl	7.8	200 mM NaCl	300 mM Urea	5 mM EDTA, 0.1 mM AEBSF

Table 26
continued

No.	Buffer	pH	Ions	Detergent	Additives
21	100 mM Tris-HCl	7.8	200 mM NaCl	300 mM Urea	5 mM EDTA, 0.1 mM AEBSF
22	200 mM Tris-HCl	7.8	150 mM NaCl		5 mM EDTA, 0.1 mM AEBSF
23	20 mM Tris-HCl	7.8	100 mM NaCl		5 mM EDTA, 0.1 mM AEBSF
24	150 mM Tris-HCl	7.8	150 mM NaCl		5 mM EDTA, 0.1 mM AEBSF
25	20 mM HEPES	7.8	150 mM NaCl, 2 mM $MgCl_2$		0.1 mM AEBSF
26	50 mM HEPES	7.5	150 mM NaCl, 2 mM $MgCl_2$		0.1 mM AEBSF
27	50 mM HEPES	7.5	200 mM NaCl		5 mM EDTA
28	100 mM HEPES	7.5	200 mM NaCl		5 mM EDTA, 0.1 mM AEBSF
29	20 mM HEPES	7.8	150 mM NaCl, 2 mM $MgCl_2$		0.1 mM AEBSF
30	50 mM HEPES	7.8	20 mM NaCl		5 mM EDTA

3.4.3 PROTEIN EXTRACTION AND CLARIFICATION

To purify expressed proteins via affinity chromatography, the protein of interest needed to be extracted. Therefore, cells were pelletized for 10 min at 8000 x g and 4 °C and stored at -20 °C. For cell lysis, the pellet was resuspended in a specific lysis buffer (Table 27A) and cell walls were disrupted using lysozyme, sonication and/or pressure disruption (Table 27B). Sonication was performed using a *Branson Sonifier® S250A* with a *S450A Cell Disruptor* (Emerson, St. Louis, United States) with 50 % output for 5 cycles with 30 s on and 30 s off. For pressure disruption the high-pressure homogenizer *HPH* (IKA®-Werke, Staufen, Germany) was used applying 2000 bar. Cell debris was pelleted by centrifuging the cell lysate at 48000 x g at 4 °C for 30 min. To remove all remaining insoluble particles the supernatant was filtrated using a 0.45 µm syringe filter.

Table 27
(A) Lysis buffers for different recombinant proteins

Protein	Buffer	pH	Ions	Additives
GST-SAUL1	50 mM Tris-HCl	9.0	250 mM NaCl	0.1 mM AEBSF
GST-SAUL1	100 mM NaH_2PO_4	7.5	250 mM NaCl	0.1 mM AEBSF
GST-BON1	100 mM Tris-HCl	7.5	200 mM NaCl	0.1 mM AEBSF, 5 mM EDTA
GST-CHS1	100 mM Tris-HCl	8.0	150 mM NaCl	0.1 mM AEBSF, 1 mM DTT, 5 % Glycerol

(B) Lysis conditions for different recombinant proteins

Protein	1st Step	2nd Step	3rd Step
GST-SAUL1	1 mg ml^{-1} Lysozyme for 1 h	Sonication	
GST-BON1	1 mg ml^{-1} Lysozyme for 1 h	Sonication	Pressure disruption
GST-CHS1	1 mg ml^{-1} Lysozyme for 1 h	Sonication	

3.4.4 AFFINITY CHROMATOGRAPHY

As for the purification of the glutathione s-transferase (GST) coupled recombinant proteins manual or automated affinity chromatography using the fast protein liquid chromatography (FPLC) *ÄKTA™ pure 25L* system (GE Healthcare, Chicago, United States) was used.

In case of the manual affinity chromatography the cell lysate was applied directly to the different columns containing glutathione coupled sepharose (Table 28A) and incubating for at least 1 h at 4 °C, while being rotated. In case of the automated affinity chromatography a 50 ml superloop (GE Healthcare, Chicago, United States) with a flow rate between 0.05- and 0.15-ml min^{-1} was used (Table 28B). While performing manual affinity chromatography all steps were carried out using gravity flow. In case of automated affinity chromatography all other steps were carried out using a flow rate of 1 ml min^{-1}. Afterwards the column was washed with 3 column volumes (CV) of lysis buffer. Elution was performed using 3 CV lysis buffer containing 50 mM glutathione, collecting the eluate in 2 ml fractions. The fractions containing the recombinant protein were identified using the absorbance at 280 nm and pooled.

In a last step the column was washed with 5 CV lysis buffer, regenerated with 2 CV 6 M guanidine hypochloride, 5 CV PBS and equilibrated with 5 CV ddH_2O and 2 CV 20 % EtOH.

Table 28
(A) Columns and matrix used during manual affinity chromatography.

Column	Matrix	Matrix used (ml)	Manufacturer
Econo-Column® 1.0 x 10 cm	-	2	Bio-Rad, Hercules, United States
Econo-Column® 2.5 x 30 cm	-	10	Bio-Rad, Hercules, United States
Glutathione Sepharose™ 4B	Glutathione coupled sepharose	-	GE Healthcare, Chicago, United States

(B) Precast columns used during automated affinity chromatography.

Column	Matrix	Volume (ml)	Manufacturer
GSTrap™ FF	Glutathione coupled sepharose	5	GE Healthcare, Chicago, United States
GSTrap™ 4B	Glutathione coupled sepharose	5	GE Healthcare, Chicago, United States
GSTrap™ HP	Glutathione coupled sepharose	5	GE Healthcare, Chicago, United States

3.4.5 AFFINITY TAG REMOVAL

Afterwards the GST-tag was removed using the *PreScission™* protease (GE Healthcare, Chicago, United States), by adding two enzyme units per 100 µg of GST-fusion protein. Digestion was carried out incubating the mixture for at least 6 h at 4 °C.

3.4.6 SIZE-EXCLUSION CHROMATOGRAPHY (SEC)

For size-exclusion chromatography (SEC) the *ÄKTA™ pure 25L* system (GE Healthcare, Chicago, United States) was used with a *HiLoad™ 16/600 200pg* column (GE Healthcare, Chicago, United States). Sample loop size and flow rate were adapted to the different recombinant proteins purified in this thesis (Table 29). As for BON1, the pooled fractions of the affinity chromatography were concentrated to a volume of 2 ml using an *Amicon® Ultra-4* centrifugal filter unit (Merck Millipore, Burlington, United States) with a molecular weight cut-off (MWCO) of 10 kDa according to the manufacturer's specifications, prior to the SEC. The eluate was collected in 2 ml fractions.

Table 29
SEC conditions for different proteins

Protein	Sample loop	Flow rate
SAUL1	5 ml	1 ml min^{-1}
BON1	2 ml	0.1 ml min^{-1}

3.4.7 FINALIZATION

For protein analysis proteins were dialyzed, in case another buffer had to be used, using a *Spectra/Por®* dialyze membrane (Spectrum Chemical, New Jersey, United States) with a MWCO of 3.5 kDa. In case of a higher protein concentration was needed, protein solutions were concentrated using either a *Amicon® Ultra-4* centrifugal filter unit (Merck Millipore, Burlington, United States) with a MWCO of 10 kDa or *Vivaspin® 500 Centrifugal Concentrators* (Sartorius, Göttingen, Germany) with a MWCO of 10 kDA according to the manufacturer's specifications.

3.5 PROTEIN ANALYSIS

3.5.1 PROTEIN SEQUENCES

Amino acid sequences for the investigated recombinant proteins are listed in the supplement (Table S1). Due to the removal of the GST tag using the *PreScission™* protease eight additional amino acids are located at the N-terminus of each protein. Therefore, each amino acid position mentioned in this thesis does not correspond to the TAIR reference sequence database [162,163], but it is shifted by eight amino acids.

3.5.2 SODIUM DODECYL SULFATE-POLYACRYLAMIDE GEL ELECTROPHORESIS (SDS-PAGE)

To analyze proteins in general, sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was performed. Therefore, acrylamide gels were cast using the *Mini-Protean® Tetra Handcast System* (Bio-Rad, Hercules, United States). Spacer plates with 0.75 mm integrated spacers were used. All gels were run using a constant 40 mA per gel for at least 45 min. All gels were cast after the following recipes (Table 30).

Table 30
(A) SDS-PAGE stacking gel composition

Ingredient	Concentration
Tris-HCl pH 6.8	130 mM
Acryl/bisacrylamide mix (37.5:1)	5.1 %
SDS	0.1 %
APS	0.1 %
TEMED	0.001 %

(B) SDS-PAGE seperating gel composition

	Concentration		
Ingredient	10 %	12 %	15 %
Tris-HCl pH 8.8	382 mM	382 mM	382 mM
Acryl/bisacrylamide mix (37.5:1)	10 %	12 %	15 %
SDS	0.1 %	0.1 %	0.1 %
APS	0.1 %	0.1 %	0.1 %
TEMED	0.001 %	0.001 %	0.001 %

3.5.3 COLLOIDAL COOMASSIE G-250 STAINING

To detect proteins after SDS-PAGE, gels were stained using colloidal Coomassie G-250 staining [196]. Therefore, gels were washed three times for at least 10 min using ddH_2O. Afterwards gels were stained for at least 1 h. I necessary, gels were destained using destain or ddH_2O. Gels were stored in ddH_2O. Colloidal Coomassie G-250 stain and destain were prepared after the following recipe (Table 31) and protocol [196].

Table 31
Colloidal Coomassie G-250 stain and destain recipe

Ingredient	Concentration
Stain	
Coomassie Brilliant Blue G-250	0.02 %
Aluminium sulfate-(14-18)-hydrate	5 %
Ethanol	10 %
Phosphoric acid	2 %

Table 31
continued

Ingredient	Concentration
Destain	
Ethanol	10 %
Phosphoric acid	2 %

3.5.4 WESTERN BLOT ANALYSIS

To detect specific proteins after gel electrophoresis by SDS-PAGE (see chapter 3.5.2) semi-dry Western blotting was performed. Therefore, unstained SDS-gels were washed three times for 5 min with H_2O and equilibrated in transfer buffer. *Roti®-NC* nitrocellulose membrane (Carl Roth, Karlsruhe, Germany) and *Whatman™ 3MM Chr* chromatography paper (GE Healthcare, Chicago, United States) were as well equilibrated using transfer buffer. For the semi-dry Western blot the *FastBlot B43* (Biometra, Göttingen, Germany) was used with a sandwich containing two *Whatman™* papers, the nitrocellulose membrane, the SDS-gel and again two *Whatman™* paper. All air bubbles were removed and the transfer was performed using 60 mA per gel for 65 min.

In a next step the nitrocellulose membrane was washed for 5 min using TBS-T buffer and incubated overnight in blocking solution at 4 °C and continuous shaking. Afterwards the nitrocellulose membrane was washed twice for 5 min using TBS-T buffer and the primary antibody *Anti-Glutathione-S-Transferase IgG* (Sigma-Aldrich, St. Louis, United States) was added in a dilution of 1:5000 in 10 ml blocking solution and incubated at RT for 2 h an continuous shaking. After that the primary antibody was discarded and the nitrocellulose membrane was washed three times for 5 min using TBS-T buffer. The secondary antibody *Anti-Rabbit IgG HRP-conjugate* (Merck, Darmstadt, Germany) was then added in a dilution of 1:5000 in 10 ml blocking solution and incubated as well for 2 h at RT and continuous shaking. The secondary antibody was discarded as well and the nitrocellulose membrane washed three times for 5 min using TBS-T buffer. The detection of the luminescence signal was performed by incubating the nitrocellulose membrane with *Clarity™ Western ECL Substrate* (Bio-Rad, Hercules, United States) and using the *ChemiDoc™ Touch Imaging System* (Bio-Rad, Hercules, United States).

Table 32
Buffers used for Western blot

Buffer	Components
Transfer	50 mM Tris-HCl pH 8.3, 40 mM glycin, 20 % methanol
TBS-T	50 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.05 % Tween™ 20
Blocking solution	50 mM Tris-HCl pH 7.6, 150 mM NaCl, 0.05 % Tween™ 20, 5 % skim milk powder

3.5.5 PROTEIN CONCENTRATION DETERMINATION

All protein concentrations were determined measuring the absorbance at 280 nm. Each sample was normalized to the corresponding buffer. To correct for differences in tryptophane and tyrosine content, the corresponding extinction coefficients and the related molecular weight were considered for the calculations. The concentrations (c) in mg ml^{-1} were calculated according to the Beer-Lambert law as

$$c = \frac{Abs_{280\,nm}}{\varepsilon \times l} \times MW$$

where $Abs_{280\,nm}$ is the absorption at 280 nm in AU, ε is the extinction coefficient in M^{-1} cm^{-1}, l is the light path length in cm and MW is the molecular weight in Da.

3.5.6 MICROSCALE THERMOPHORESIS (MST)

To investigate protein interactions and determine the binding affinity MST experiments were carried out. Therefore, proteins of interest were labelled using the *Monolith NT™ Protein Labelling Kit Green* (NanoTemper Technologies, Munich, Germany) according to manufacturer's specifications. Measurements were performed using the *Monolith NT.115* (NanoTemper Technologies, Munich, Germany).

3.5.7 MASS SPECTROMETRY (MS)

To identify and analyze singular protein bands from SDS-PAGE experiments, mass spectrometry (MS) measurements were performed. These were carried out by the *Molecular Plant Genomics* research unit of Julia Kehr (Institute of Plant Science and Microbiology, Hamburg, Germany). All experiments were conducted and evaluated by either Patrizia Hanhart, Anna Ostendorp or Julia Kehr.

3.5.8 NATIVE MS

To analyze the interactions between two proteins under native conditions, native MS experiments were performed. These were planned and carried out in corporation with the *Dynamics of Viral Structures* research group of Charlotte Uetrecht (Heinrich Pette Institute, Hamburg, Germany). All experiments were conducted and evaluated by Julia Lockhauserbäumer.

For native MS proteins were used at concentrations of 5–11 µM. Purified proteins were buffer exchanged prior to MS analysis to 250 mM ammonium acetate at pH 9.0 or 7.5, via *Vivaspin® 500 Centrifugal Concentrators* (Sartorius, Göttingen, Germany) with a MWCO of 10 kDa at 15,000 × g. Native mass spectra were measured at 25 °C on a Q-ToF instrument (Waters, Milford, United States) modified for high mass experiments with a nano-electrospray ionization (ESI) source in positive ion mode [197]. Samples were directly infused from gold-coated electrospray capillaries without any accessory chromatographic separation. The voltages and pressures were optimized for noncovalent protein complexes [198].

The gas pressures were set to 10 mbar in the source region and 1.8 × 10^{-2} mbar argon in the collision cell. Mass spectra were recorded with applied voltages for the capillary, cone and collision cell of 1.25–1.35 kV, 150 V and 5–150 V, respectively, optimized for minimal complex dissociation. For the calibration of the raw data *MassLynx* (Waters, Milford, United States) and a

25 mg ml^{-1} cesium iodide spectrum from the same day was used. *MassLynx* was used to assign peak series to protein species and to determine the mass after minimal smoothing.

3.5.9 THERMOFLUOR ASSAY

To improve protein stability, the effects of different buffers on the protein of interest were compared. Therefore, a thermofluor assay was performed to determine the melting temperature (T_m) of the protein of interest. Thus, to the protein in its original buffer different buffers and *SYPRO™ Orange Protein Gel stain* (Thermo Fisher Scientific, Waltham, United States) were added. Afterwards a melting curve was recorded with the detected fluorescence corresponding to the amount of the protein being unfolded.

The analysis was performed after Boivin et al. [199], using the suggested buffers for global parameters. Measurements were performed using a *QuantStudio 6 Flex Real-Time PCR System* (Thermo Fisher Scientific, Waltham, United States). To calculate the melting temperature T_m of the protein of interest under different buffer conditions, the fluorescence of the melting curve was fitted against a modified Boltzmann equation [200], for which the relative fluorescence units (RFU) are defined as

$$RFU = \frac{RFU_{min} + (RFU_{max} - RFU_{min})}{(1 + e^{\frac{T_m - x}{m}})}$$

where RFU is the fluorescence in arbitrary units, RFU_{min} and RFU_{max} are the minimal and maximal fluorescence at low and high temperatures, T_m is the melting temperature of the protein in °C, x is the temperature in °C and m is the slope of the curve in °C^{-1}.

Fits were performed using *OriginPro 2018G* (OriginLab Corporation, Northampton, United States) with the modified Boltzmann equation. To compare the effect of different buffer environments, the thermal shift (ΔT_m) was calculated being defined as

$$\Delta T_m = T_{m(POI)} - T_m$$

where $T_{m(POI)}$ is the melting temperature of the protein of interest in its original buffer.

3.6 STRUCTURAL PROTEIN ANALYSIS

3.6.1 PROTEIN FEATURE PREDICTION

To predict conserved protein domains *InterPro* was used (Table 33).

Table 33
Programs used for domain prediction

Usage	Software
Domain prediction	*InterPro* [201]

3.6.2 *IN SILICO* PROTEIN MODELLING

To obtain tertiary structures of proteins of interest, *in silico* models had to be generated. These were obtained by using the following homology-based programs, based on the amino acid sequence (Table 34).

Table 34
Programs used for homology-based structure modelling

Usage	Software
Hierarchical protein modelling	*I-TASSER* [202,203]
Markov model based modelling	*PHYRE v2.0* [204]

3.6.3 DYNAMIC LIGHT SCATTERING (DLS)

To analyze the distribution of radii, present in a solution, DLS experiments were carried out using the *SpectroSize™ 300* (Xtal Concepts, Hamburg, Germany). In advance protein solutions were centrifuged at 21000 xg for at least 10 min at 4 °C to precipitate large particles. Afterwards 20 µl of the measured protein were transferred to a Hellma™ Far UV quartz cuvette with a light path length of 10 mm. Scattering was measured twenty times for 20 s at 660 nm and 20 °C. The viscosity of the solution was adjusted in the program according to the components of the buffer.

3.6.4 CIRCULAR DICHROISM (CD) SPECTROSCOPY

To obtain insights into the secondary structure of proteins circular dichroism (CD) measurements were performed using the *J-815* spectrometer (JACO International, Tokyo, Japan). Due to the interference of chloride with CD spectroscopy proteins solutions were dialyzed using an identical buffer containing fluorine as a substitute. As an alternative, solutions were diluted 1:25 with ddH_2O to overcome the absorbance of chloride in the far UV spectrum [205]. The instrument was calibrated according to the manufacturer's specifications. The ellipticity of the sample was measured in a 1 mm quartz cuvette in a wavelength interval ranging from 185-260 nm. The baseline recorded for the corresponding buffer was subtracted. To compare different conditions the observed ellipticity was normalized. This molar ellipticity ([θ]) was calculated being defined as

$$[\theta] = \frac{m^\circ \times MW}{10 \times l \times c}$$

where m° is the ellipticity in millidegrees, *MW* is the molecular weight in Da, l is the light path length in cm and *c* is the protein concentration in mg ml^{-1}.

3.6.5 SMALL-ANGLE X-RAY SCATTERING (SAXS)

SAXS data was collected at beamline P12 operated by EMBL Hamburg at the PETRA III storage ring (DESY, Hamburg, Germany). Scattering data was collecting using a photon counting *Pilatus3 X 2M* pixel detector (Dectris, Baden-Daettwil, Switzerland) with a sample-detector distance of 3.1 m. A scattering vector q (where $q = 4\pi \sin\theta/\lambda$, 2θ is the scattering angle and λ is the radiation wavelength) range from 0.03 to 4.8 nm was recorded (Table 35A).

For batch measurements, monodisperse protein solutions in form of concentration series between 1 and 10 mg ml^{-1} were used. Scattering of protein solutions and corresponding buffers was detected successively 20 times for 45 ms. As a first data processing step all 20 measurements were normalized to H_2O and averaged. Afterwards the buffer scattering data was subtracted. All these steps were performed using the *SASFLOW* pipeline.

As a next step the radius of gyration (R_g) and the forward scattering intensity ($I(0)$) were calculated using the Guinier fit by applying *AUTORG*. In a next step all data points at very low angles, which were neglected for the Guinier fit, were trimmed. Using an indirect Fourier transformation, the P(r) function was calculated by *DATGNOM*. The maximum dimension (d_{max}) was defined manually and had to result in a smooth, not too elongated P(r) curve, with a reasonable fit to the experimental data. These programs were used as part of the *PRIMUSQT* package, which is part of the *ATSAS* suite.

To calculate theoretical radii of gyration (R_g) for proteins, Flory's equation was used, which is defined as

$$R_g = R_0 \times N^{\nu}$$

where R_0 is a constant, which depends on the type of protein, in Å, N is the number of amino acid residues and ν is a scaling factor [206].

To evaluate the fit, standardized residuals (Δ/σ) were plotted being defined as

$$\frac{\Delta}{\sigma} = \frac{I_{obs}(s) - I_{exp}(s)}{\sigma(I_{exp}(s))}$$

where $I_{obs}(s)$ is the detected scattering intensity, I_{exp}(s) is the expected intensity and $\sigma(I_{exp}(s))$ is the standard deviation of the expected intensities

In a next step the calculated radius distribution was used to generate *ab initio* models of the protein investigated using *DAMMIF* or *GASBOR*. *Ab initio* models were generated 20 times and compared using *DAMAVER*. The finals model was chosen based on the normalized spatial discrepancy (NSD).

To adjust the *in silico* models to the scattering data, different programs were applied. In case of low flexibility *SREFLEX* was used to refine the models. Afterwards models were evaluated by using *CRYSOL*. In case of higher predicted flexibility *EOM* was used. Therefore, the proteins of interest were divided into domains, which were predicted to be more ordered (Table S2, S3). Inter-domain regions were defined as flexible and allowed for adjustment of the model to the obtained scattering data.

Table 35
(A) SAXS data-collection parameters

Instrument	PETRA III (DESY, Hamburg, Germany), Beamline P12 [207]
Detector	Photon counting *Pilatus3 X 2M* pixel detector (253 x 288 mm^2) (Dectris, Baden-Daettwil, Switzerland)
Sample-detector-distance (m)	3.1
Wavelength (nm)	0.124
Focal spot (mm)	0.2 x 0.12

Table 35
(A) continued

s measurement range (nm^{-1})	0.03-4.8
Exposure time (ms)	45
s-axis calibration	Silver behenate
Sample temperature (°C)	20

(B) Software used for SAXS data processing

Usage	Software
Primary data reduction and processing	*SASFLOW v3* [208]
Program suite	*ATSAS v2.8.3* [209,210]
Graphical User Interface	*PRIMUSQT v2.8.3* [211]
Calculating the Guinier fit	*AUTORG v2.8.3* [209,210]
Calculating the P(*r*) function	*DATGNOM v2.8.3* [209,210]
Generating *ab initio* model (dummy atom model)	*DAMMIF* [212]
Generating *ab initio* model (chain-like dummy residues)	*GASBOR* [213,214]
Aligning and evaluation of *ab initio* models	*DAMAVER* [215]
Fit improvement of atomic models against scattering data	*SREFLEX* [216]
Fit improvement of flexible proteins against scattering data	*EOM v2.0* [217,218]
Evaluation of theoretical scattering curves of atomic models	*CRYSOL* [219]
Flexible refinement of atomic models	*SASREF* [214]
Generation of form factor files	*FFMAKER* [211]
Mixture analysis of polydisperse systems	*OLIGOMER* [211]

Prior to all measurements different concentrations of bovine serum albumin (BSA) were measured to test for technical errors. In addition, obtained values were used to calculate theoretical MW-values for the proteins of interest. The theoretical MW being defined as

$$MW_I = \frac{MW_S}{I(0)_S} \times I(0)_I$$

where MW_S is the molecular weight of BSA in Da, $I(0)_S$ is the forwards scattering intensity of BSA and $I(0)_I$ is the forwards scattering intensity of the protein of interest [220].

3.6.6 INLINE SIZE-EXCLUSION CHROMATOGRAPHY SAXS (SEC-SAXS)

Inline SEC-SAXS was performed using a *Superose™ 6 Increase 10/300 GL* (GE Healthcare, Chicago, United States) column with a flow rate of 0.5 ml min^{-1}. Continuous 1 s data-frame measurements of 50 µl of sample with 120 frames ml^{-1} were performed (Table 36). Data was pre-processed using *CHROMIXS* [221]. 25 frames from each sample were selected manually. 50 corresponding buffer

frames were selected automatically. Afterwards scattering data was processed as described in 3.6.5.

Table 36
SEC-SAXS parameters

HPLC	*1260 Infinity II Bio-Inert* sytem (Agilent, Santa Clara, United States)
SEC column	*Superose™ 6 Increase 10/300 GL* (GE Healthcare, Chicago, United States)
Injection volume (µl)	50
Flow rate (ml min^{-1})	0.5
Frame rate (frames ml^{-1})	120
Exposure time	Continuous1 s data-frame measurements

(B) Software used for SEC-SAXS data processing

Usage	Software
Data frame selection	*CHROMIXS v2.8.3* [221]

3.6.7 ELECTROSTATIC POTENTIAL ANALYSIS

To calculate the electrostatic potential of different structures, the *APBS-PDB2PQR* suite was used. In a first step the *pdb* files, which were analyzed, were prepared using *PDB2PQR* to account for missing information, which is needed to solve the equations of continuum electrostatics. Therefore, PDB2PQR was used with developer-recommended settings and a pH of 7. In a second step the obtained results were used to calculate the electrostatic potential of the solvent accessible surface. All calculations were performed using the web-based versions of the programs. Results were visualized using *PyMOL*.

Table 37
Software used for the calculations of electrostatic potentials of proteins

Usage	Software
File preparation	*PDB2PQR v2.1.1* [222]
Electrostatic potential calculation	*APBS v1.5* [223]

3.6.8 PROTEIN STRUCTURE VISUALIZATION

All structures were visualized using *PyMOL*. *Ab initio* models were depicted as spheres with a radius corresponding to the calculated one obtained from *DAMMIF* or *GASBOR*. Refined *in silico* structures were depicted in the cartoon style with flexible areas being visualized in the surface style. Proteins or domains were aligned using the align function of *PyMOL*. The solvent accessible surface was as well depicted in the surface style. Electrostatic potentials were illustrated using the *APBS Tools v2.1* plugin.

Table 38
Software used for protein structure visualization

Usage	Software
Visualization	*PyMOL v2.1.1* [224]
Electrostatic potential visualization	*APBS Tools v2.1* [223]

3.7 MICROSCOPY

To obtain high-resolution microscopy pictures, confocal laser scanning microscopy using the *Leica TCS SP8 Confocal Platform* (Leica Microsystems, Wetzlar, Germany), was performed. Emissions were recorded using photomultiplier (PMT) and *Leica HyD* detectors. Recording each picture, a pixel dwell time of 1.2 µs was used with a resolution of 1024 x 1024 pixels. In case of BiFC experiments pictures of transformed protoplasts (see chapter 3.2.6) were captured using the sequential scanning mode, separating the YFP and RFP channel. Settings for different stains and fluorophores are listed below (Table 39).

Table 39
(A) Excitation and emission parameters used for different fluorophores. Stains, fluorescing proteins and other signals are listed separately.

Fluorophores	Excitation (nm)	Emission (nm)	Detector
Fluorescing proteins			
Red fluorescent protein (RFP)	561	580-595	HyD
Yellow fluorescent protein (YFP)	514	521-535	HyD
Other signals			
Chlorophyll	514	700-730	PMT

(B) Sequential scanning settings

Fluorophores	Channel
Red fluorescent protein (RFP)	1
Yellow fluorescent protein (YFP)	2
Chlorophyll	2

4 RESULTS AND DISCUSSION

The SAUL1 system is a versatile model to investigate regulatory events in plant autoimmunity. On one hand, the homogeneous characteristic of autoimmunity in *saul1* mutant plants allows for accurate detection of downstream gene expression changes during autoimmune responses. On the other hand, guarding of SAUL1 homeostasis by different NLR heteromers renders the SAUL1 protein a highly interesting model for structural characterization.

4.1 GENE REGULATIONS IN THE *saul1-1* AUTOIMMUNE MUTANT

Whereas SAUL1 (SENESCENCE-ASSOCIATED E3 UBIQUITIN LIGASE 1) has originally been identified to be a negative regulator of senescence [114], it has been shown more recently that *saul1-1* mutants exhibit an autoimmune phenotype. The SAUL1 protein is guarded by the nucleotide-binding leucine-rich repeat (NLR) protein SOC3 (SUPPRESSOR OF CHS1-2 3) that forms a heterodimer with the truncated NLR CHS1 (CHILLING SENSITIVE 1) to trigger effector-triggered immunity (ETI) upon activation [137]. Upon a change from 25°C to 24°C or lower temperatures autoimmunity is initiated in *saul1-1* mutant plants. Our knowledge on regulatory events downstream of the activation of NLRs and NLR complexes is still limited. Because the autoimmune responses in *saul1-1* is initiated in all above-ground organs simultaneously, *saul1-1* autoimmunity represents an ideal model system to investigate gene regulations in ETI.

4.1.1 ANALYSIS OF MARKER GENE EXPRESSION AT EARLY POINTS IN TIME

Though the SOC3-dependent *saul1-1*-phenotype could be observed on the morphological level not before three days following the transfer from 25°C to 20 °C [135], gene expression changes were already resolved after 6 h [125]. It is known that gene regulations can be switched on or off almost immediately after applying an external stimulus [225]. Accordingly, it may well be that key genes are changed very early also in *saul1-1* autoimmunity upon the decrease in temperature, as after 6 h already hundreds of genes were changed in expression [125]. Therefore, it was of high interest to investigate, whether changes in the expression of genes encoding potential key regulators of autoimmune responses could be detected early after induction of ETI.

In a first step, the expression of distinct marker genes was analyzed by using quantitative real-time polymerase chain reaction (qRT-PCR), to test if any of them was differentially regulated before 6 h. For that purpose, plant samples were harvested as early as 30 min, 1 h and 3 h after the transfer of *saul1-1* and WT plants from 25°C to 20°C. Marker genes were selected that are known to be involved in the early regulations of the immune response or senescence. *ICS1* (*Isochorismate Synthase 1*) is one of the earliest differentially expressed genes during ETI and plant immunity in general. The *PAD4* (*PHYTOALEXIN DEFICIENT 4*) and *PR1* (*PATHOGENESIS-*

RELATED 1) genes are regulated slightly later on (see chapter 1.1.2) [76]. *WRKY53* was selected, because this transcription factor (TF) is one of the major regulators of plant senescence [226]. In the qRT-PCR experiments, *ICS1* was the only gene, which was differentially regulated after 1 h (Figure 12). Typical marker genes such as *PR1*, *PAD4* and *WRKY53* were not differentially expressed earlier than 24 h (Figure 12). In case of *WRKY53*, the late regulation may suggest a non-driving role of this TF. Although *WRKY53* has been shown to be an important TF in the defense against pathogens, it acts further downstream. This is due to the fact that its expression depends on salicylic acid (SA) previously synthesized by ICS1 [227,228].

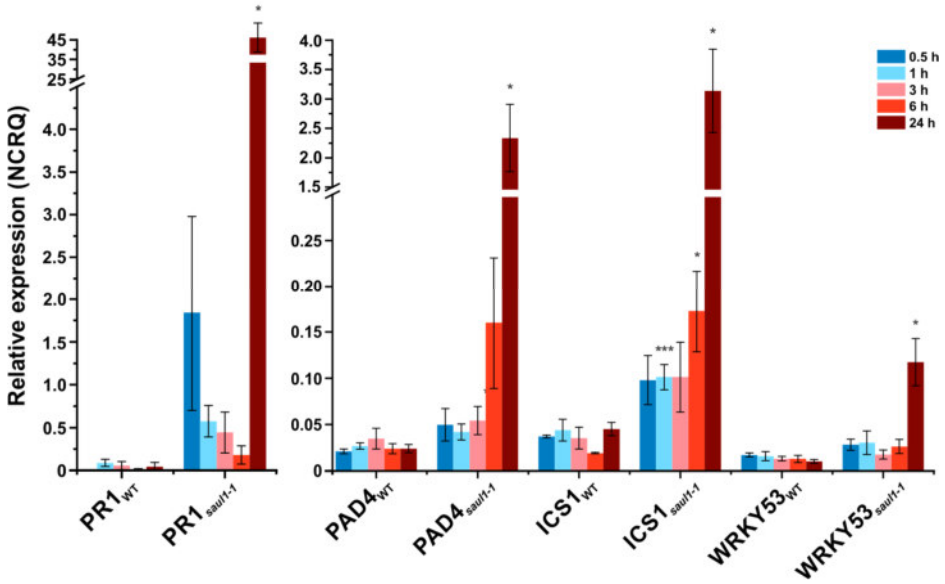


Figure 12: Marker gene expression in *saul1-1*. Depicted are the normalized calibrated relative quantities (NCRQ) of *PR1*, *PAD4*, *ICS1* and *WRKY53* expression in *saul1-1* and WT. RNA samples were collected before, 30 min, 1 h, 3 h, 6 h and 24 h after the plants were transferred from 25 °C to 20 °C. Expression changes in *saul1-1* were compared to WT values (N=3). Error bars represent ± SD. Significances are indicated by *, $p \leq 0.05$, ***, $p \leq 0.001$ (Student's t-test). Expression of *ADVAC*, *EER1* and *RPL10* served as control and calibration samples (see chapter 3.2.9).

The *ICS1* gene encodes the primary SA synthase during a pathogen infection [229]. Its early expression may hint to SA being one of the earliest regulators during the initiation of the *saul1-1* phenotype. Accordingly, an increase in the expression of *PR1*, which is one of the primary defense response marker genes and whose expression is SA-dependent, was detected later than the changes in *ICS1* expression. *PAD4*, a positive regulator of SA synthesis, was differentially expressed at later point in time. This is not surprising, as it has been reported that although *PAD4* gene expression is SA-dependent, the PAD4 protein acts upstream of SA accumulation [230]. Thus, SA synthesis appears to be an early step in *saul1-1* autoimmunity that triggers the expression of additional genes such as *PR1*, *PAD4* and *WRKY53* [135]. Taken together, these gene regulations pointed towards a defense response driven *saul1-1* phenotype and not a senescence-dependent one, although so far this conclusion was based on the analysis of only a few genes. Since it was shown that *ICS1* was expressed so early, gene regulations in *saul1-1* could already take place 1 h after the temperature shift. Therefore, in the following experiments time points from 1 h onwards were investigated.

4.1.2 ANALYSIS OF GENOME-WIDE EXPRESSION CHANGES AT EARLY POINTS IN TIME

In a microarray experiment 1043 genes were found to be differentially expressed as early as 6 h after the induction of the *saul1-1* phenotype [135]. This indicated that key changes in gene expression

may occur much earlier. In addition, this number of regulated genes was too large to identify key genes of autoimmunity easily. To analyze early gene expression changes and detect putative expression changes prior to the induction of SA synthesis, RNA-seq experiments were performed. For that purpose, plants were grown for 12 days at 25°C before growing the plants for additional five days at 20°C. Control plants were kept for the same time at 25°C. Since *ICS1* was detected to be differentially regulated already after 1 h (see chapter 4.1.1) samples were taken from 1 h onwards. To investigate gene expression changes on a kinetic level samples were collected 1 h, 2 h, 3 h. In addition, samples were taken 6 h after transferring *saul1-1* and WT plants from 25 °C to 20 °C in order to analyze subsequent expression changes and to identify similarities between the RNA-seq and the previously performed microarray experiment. Over time, visible phenotypical changes for *saul1-1* such as yellowing of leaves and reduced growth were detected after 24 h (Figure 13A, B). After RNA isolation from the respective *saul1-1* mutant and WT plants samples for t = 0 h, t = 1 h/ 2 h, t = 3 h and t = 6 h were sent for library preparation and sequencing to the *Virus Genomics Group* at the *Heinrich-Pette-Institute* (Hamburg, Germany).

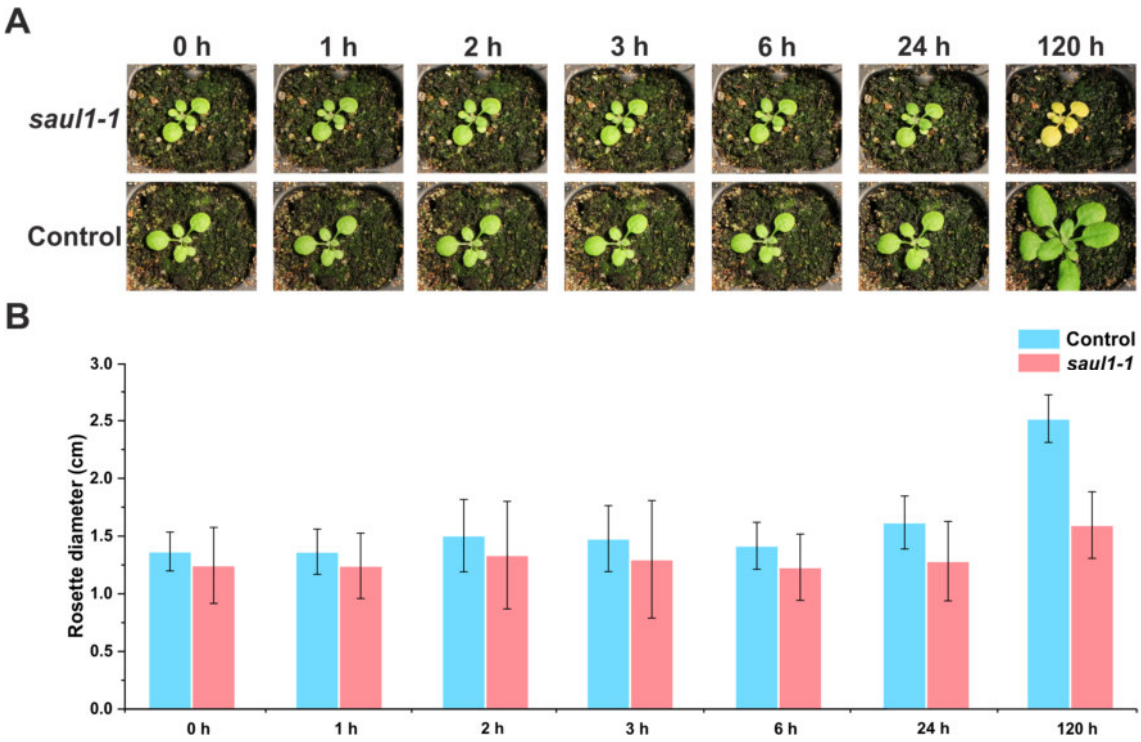


Figure 13: Phenotypical changes in the *saul1-1* mutant over time. (A) Control (WT) and *saul1-1* plants were grown on soil for 12 d at 25 °C and then shifted to 20 °C (0 h). **(B)** Rosette diameter of WT and *saul1-1* before (0 h) and after the temperature shift. Depicted is the median (N=4). Error bars are ± SD.

After sequencing, the quality of the data obtained was evaluated. To compare samples on a global scale a principle component analysis (PCA) of the mapped reads was performed (Figure 14). This revealed that control and *saul1-1* samples from the same time points were closely related, because almost all data points, each representing one biological replicate, were in close proximity. Note that one control sample at 0 h behaved differently compared to the other two biological replicates (Figure 14). This may have resulted in a lower number of detected differentially expressed genes (DEGs), because the analyzed group was more heterogenous. In general, it can be said that the *saul1-1* and control or wild type (WT) samples were initially expected to be more different. This disparity can be explained, because a PCA depicts global differences between samples. Further analysis showed that especially at early time points only few genes were differentially expressed (Figure 15A). Accordingly, all biological replicates were quite similar to each other. In addition, regarding a fold change (FC) being in most cases below 2, WT and *saul1-1* samples did not differ

on a global scale. The major influence in this experiment was the applied temperature shift. This can easily be observed, because the different time points differ more distinctly from each other than mutant and WT samples. Taken together, these and further quality control analyses using *FastQC* did not reveal any considerable data irregularities, which would prevent a further analysis of the RNA-seq data. The negligible degree of read duplications was considered to be not significant.

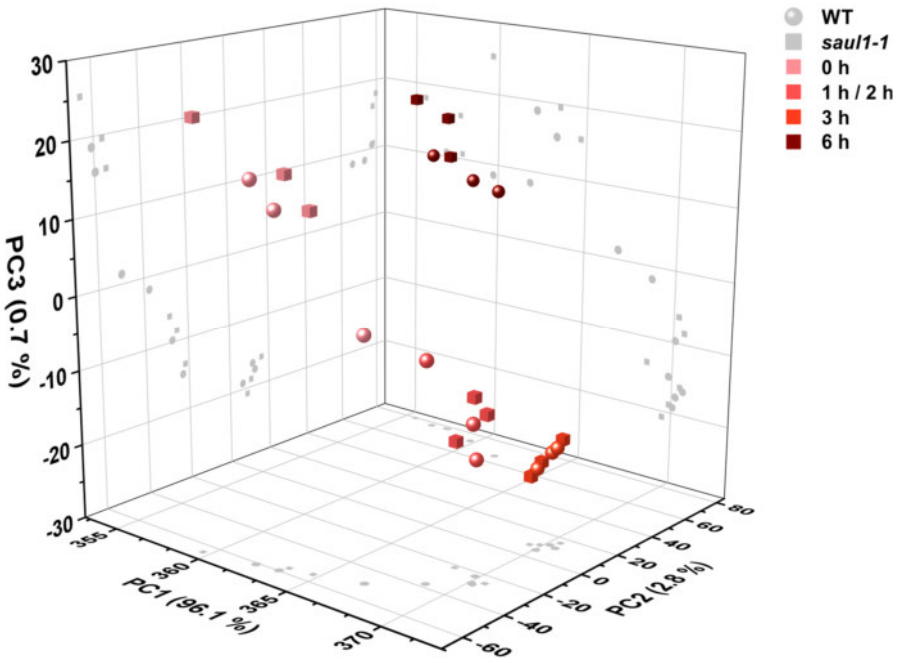


Figure 14: Three-dimensional principle-component analysis of all RNA-seq samples. WT samples are depicted as shells whereas *saul1-1* samples are depicted as cubes. Each time point consisted of three biological replicates.

The quality control allowed for analyzation of RNA-seq results with respect to DEGs. This analysis revealed that 19 genes were differentially expressed as early as 1 h and 2 h after the induction of the *saul1-1* autoimmune phenotype. Over time the amount of DEGs increased from 421 after 3 h up to 3431 genes after 6 h (Figure 15A). A Venn diagram showed that at 0 h most genes were quite distinct, because only 8 were shared with later time points (Figure 15B). The 6 h time point appeared to be most unique, since many genes were differentially expressed. In addition, changes were more prevalent compared to the earlier time points (Figure 22). In contrast, more than 95 % of the DEGs, which were detected after 3 h, were found as well after 6 h (Figure 15B). This indicates that all those gene regulations, which accumulate during *saul1-1* autoimmunity, appear to start on a small scale.

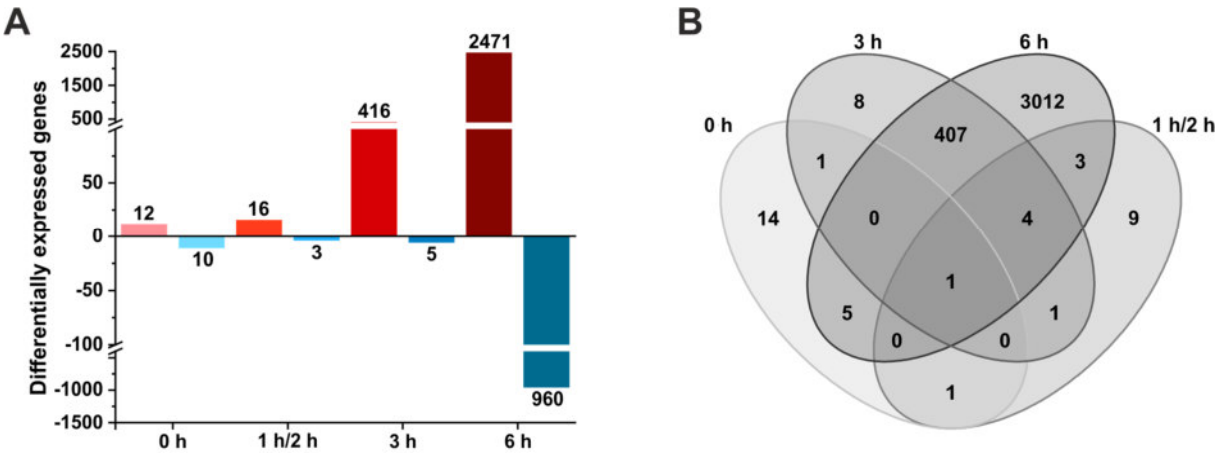


Figure 15: Differentially expressed genes in *saul1-1* at different time points. (A) Number of differentially expressed genes ($p \leq 0.01$) in *saul1-1* at different time points after transferring plants from 25 °C to 20 °C. Up-regulated (log_2(fold change) > 0) and down-regulated (log_2(fold change) < 0) genes are depicted in red and blue, respectively. **(B)** Intersection of all differentially expressed genes between different time points.

It is worth mentioning that the number of early DEGs could be underestimated. Merging two time points results in an averaging of mRNA levels. In case of genes, which may be opposingly regulated during 1 h and 2 h, this could result in a false negative detection, because the merged FC might be too low. For all 19 early regulated genes showing a moderate regulation with FC values less than 2 confirmation was required.

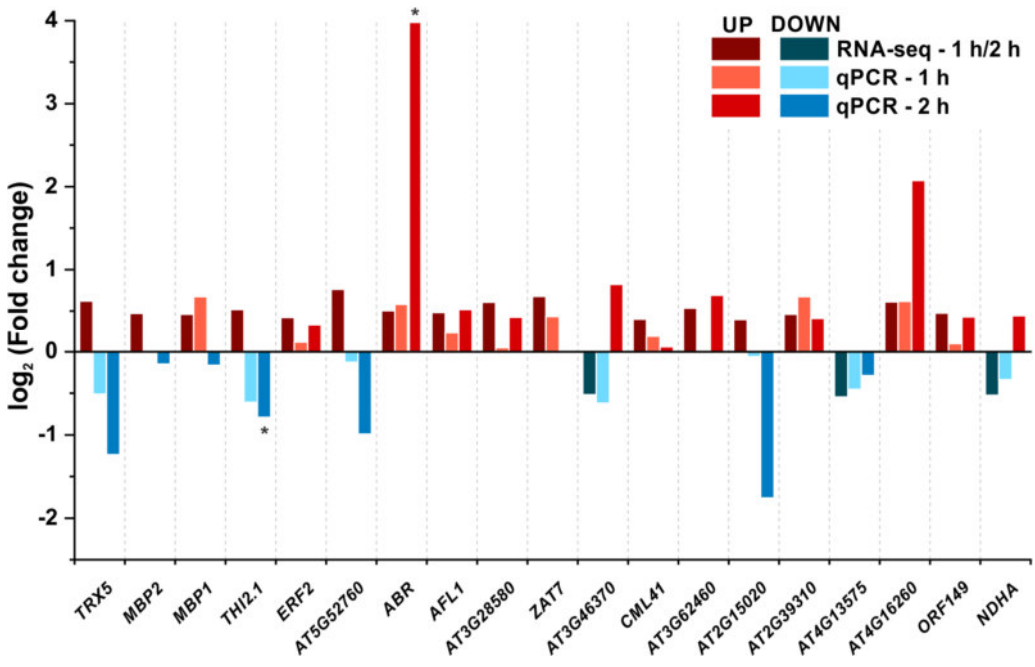


Figure 16: Conformation of early DEGs using qRT-PCR. Fold change (FC) values (log_2) of genes, which were detected to be up- or down-regulated in the RNA-seq experiment after 1 h/2 h. As a confirmation expression values were determined using qPCR and un-pooled time samples. FC values were obtained comparing *saul1-1* expression against WT-values (N=3). All RNA-seq values had an adjusted p value ≤ 0.01 (Benjamini-Hochberg). *, p value ≤ 0.05 (Student's t-test). Expression of *ADVAC*, *EER1* and *RPL10* served as control and calibration samples.

To verify the differential expression and to determine mRNA levels of these 19 DEGs in an independent experiment, a qRT-PCR analysis was performed (Figure 16). This revealed that most of the FC values detected were not significant, because the biological replicates were very heterogeneous. Nonetheless, clear tendencies towards an up- or down-regulation of those genes in *saul1-1* were observed. The qRT-PCR data showed that 14 out of 19 DEGs had FC values corresponding to those detected in the RNA-seq experiment for at least one time point (Figure 16). Three of these genes were found to be inversely regulated between t = 1 h and t = 2 h, namely *MBP1* (*Myrosinase-Binding Protein 1*), *AT3G46370* and *NDHA* (*NADH Dehydrogenase*). Nonetheless, all three genes were detected to be significantly differentially expressed after 3 h and 6 h. In addition, these genes were differentially expressed in the microarray experiment after 24 h and 48 h as well [135]. Although these genes were not clearly confirmed by qRT-PCR, they might be differentially regulated between 1 h and 2 h. In conclusion, those detected regulations were confirmed for a majority of the 19 DEGs. For this reason, these genes were further investigated to identify potential early regulators of the *saul1-1* phenotype (see chapter 4.1.5).

4.1.3 COMPARISON OF RNA-SEQ DATA TO PREVIOUS MICROARRAY DATA

In the following, the RNA-seq data were compared to gene expression data on *saul1-1*. These were obtained in previous microarray experiments, which detected differentially expressed genes in *saul1-1* at 6 h, 24 h and 48 h after induction of the autoimmune phenotype [135]. However, in contrast to the induction of autoimmunity in the RNA-seq experiments by decreased temperature, those plants were transferred from a photon flux density of 75 µmol m^{-2} s^{-1} to 20 µmol m^{-2} s^{-1} to change the light intensity [135]. It was later assumed that this led to a lowering of temperature and triggered the *saul1-1* phenotype (Hoth, S., personal communication). Nevertheless, both experiments can be compared to reveal similarities and discrepancies of DEGs.

By comparing the microarray with the RNA-seq data a 66 % (293 genes) intersection of the RNA-seq 6 h sample with the 6 h microarray sample was revealed (Figure 17A). When comparing additional points of time, the overlap between the 6 h RNA-seq sample and the 24 h microarray sample was found to be only around 40 %. Regardless, 92 % (2120 genes) of those genes, which were found to be differentially expressed after 6 h in the RNA-seq experiment, were differentially expressed in the microarray sample as well (Figure 17A). Many more genes were detected to be differentially expressed after 24 h in the microarray experiment compared to the 6 h time point in the RNA-seq experiment. This is not surprising, since the progression of the *saul1-1* phenotype results in more differentially expressed genes (Figure 15A) [135].

A more detailed comparison showed that both 6 h time points were strongly alike. Nevertheless, the RNA-seq data showed more genes (2266 in total) to be differentially expressed than the microarray data (438 in total) (Figure 17B). This can be explained by a higher sensitivity of the RNA-seq technique that results in more genes being detected as differentially expressed. Nevertheless, 293 genes (66.8 % of DEGs identified by the microarray experiment) were identified as up- or down-regulated by both techniques. Consequently, similar regulations take place in both experiments.

Comparing the latest time point of the RNA-seq experiment, which is 6 h, with the 24 h microarray data, an even more diverse pattern showed up. Here, 3050 genes were found to be uniquely differentially expressed in the microarray experiment. (Figure 17C). This is not surprising, since after 24 h more pathways in the *saul1-1* mutant should be regulated and the number detected DEGs increases over time (Figure 15A). A GO analysis of the DEGs, which were detected exclusively in the microarray experiment after 24 h, showed most of them being involved in *ubiquitin-dependent protein catabolic process*, *photosynthesis* and *plastid organization* (Figure 17D). These are closely connected to the *saul1-1* phenotype, since a reduction of photosynthesis, chlorophyll content and loss of chloroplast integrity has been observed in *saul1-1* [123,132]. The exclusively detected DEGs in the microarray experiment therefore correspond to the phenotypical changes that can be observed after 24 h at the earliest (Figure 13) [132]. Nevertheless, 2092 genes

matched as being up- or down-regulated by RNA-seq and the microarray experiment. Therefore, both experiments are very similar and the microarray data can be used to analyze later changes in gene expression.

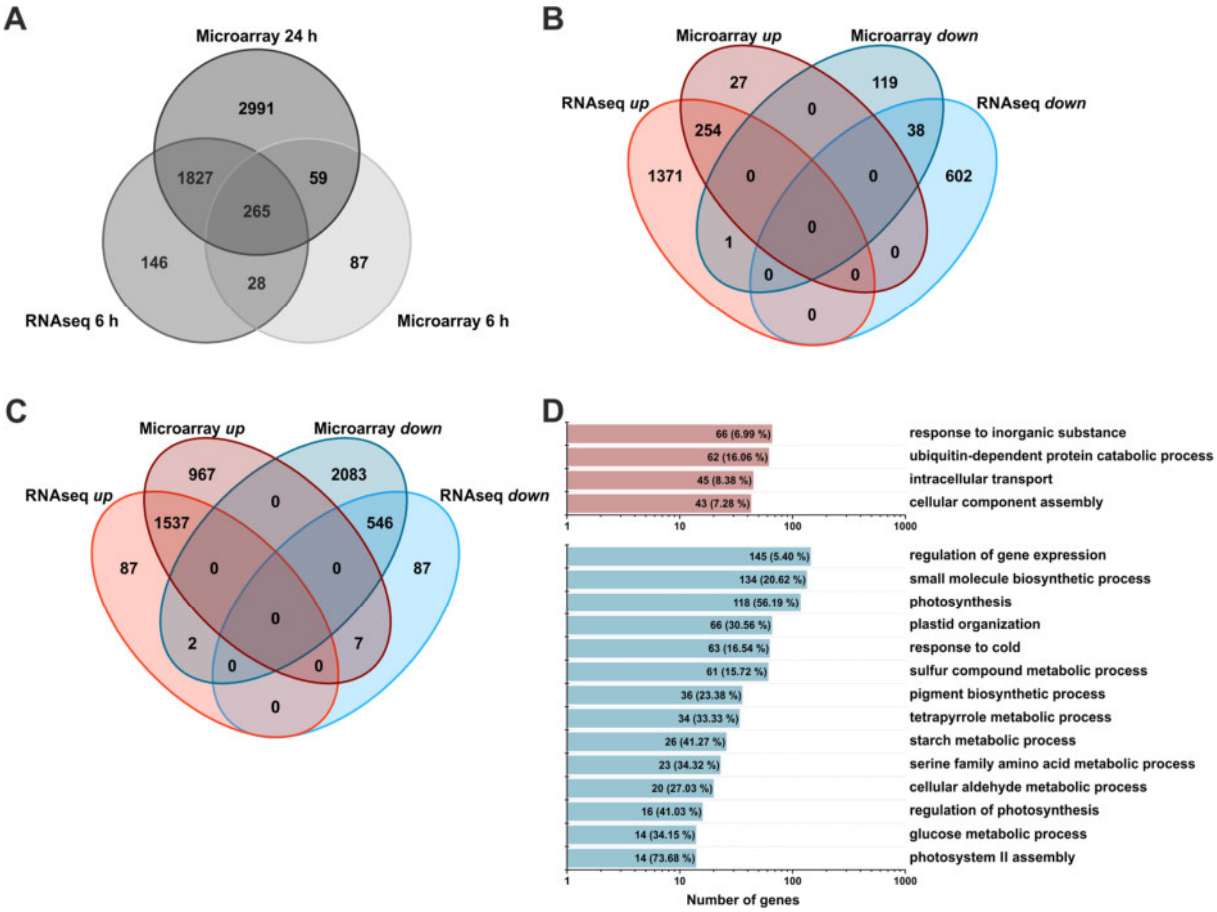


Figure 17: Comparison of RNA-seq data with a previously published microarray experiment [135]**. (A)** Venn diagram showing intersections of differentially expressed genes ($p \leq 0.01$). **(B)** Intersections of significantly up- (log_2(fold change) > 0) and down-regulated (log_2(fold change) < 0) genes, which were detected 6 h after transferring *saul1-1* plants. **(C)** Intersections of significantly up- and down-regulated genes after 6 h (RNA-seq) and after 24 h (microarray). **(D)** Most strongly enriched pathways of genes which were solely found to be regulated after 24 h in the microarray experiment. GO terms which were up-regulated are depicted in red and these which were down-regulated in blue. Percentage of genes associated with a specific term is given in brackets next to the gene count.

4.1.4 COMPARISON OF TRANSCRIPTOMIC CHANGES IN *saul1-1* TO EXPRESSION CHANGES IN ABIOTIC AND BIOTIC STRESS RESPONSES

As mentioned above, the senescence-associated phenotype of *saul1-1* mutants in lowered temperatures may be associated with global gene expression changes due to senescence. This assumption was based on observed lesioning and up-regulated marker genes [123]. However, early gene regulations were reminiscent of immune gene expressions [125]. *SOC3*-dependent autoimmunity pointed to an onset of ETI, resulting in the hypersensitive response in all above-ground organs [137]. Therefore, it can be hypothesized that *saul1-1* expression changes are highly similar to those in other autoimmune mutants or to those following the secretion of effectors into cells.

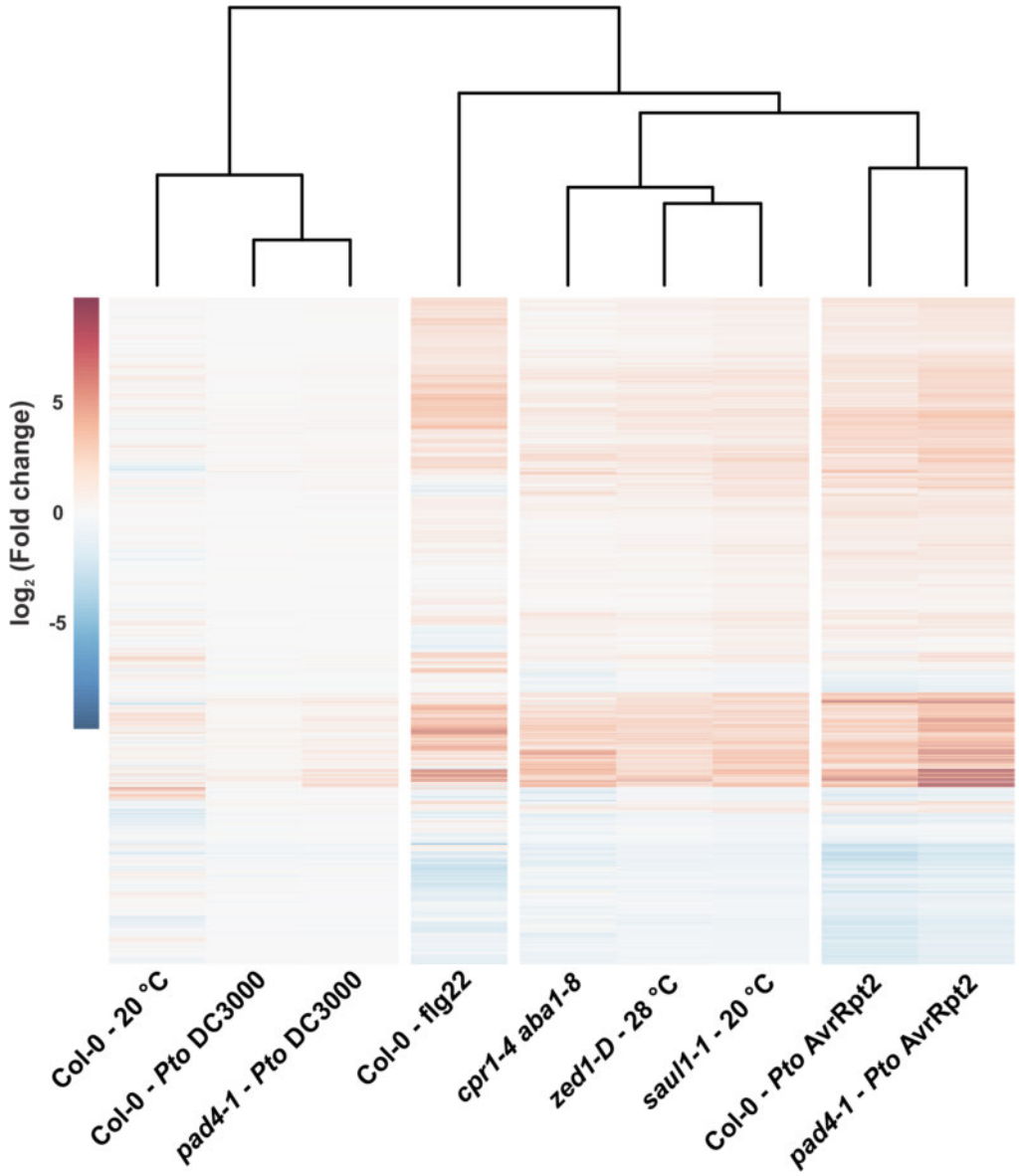


Figure 18: Comparison of gene expression changes of *saul1-1* with different autoimmune mutants and infection experiments. Cluster analysis of fold changes of DEGs in *saul1-1* 6 h after the temperature shift depicted as the binary logarithm. Fold changes were compared to gene expression changes of previously published stress experiments: infection (Col-0 – *Pto* DC3000, *pad4-1* - *Pto* DC3000, Col-0 – *Pto* AcrRpt2, *pad4-1* – *Pto* AvrRpt2) [176], flg22 treatment (Col-0 – *flg22*) [38], autoimmune mutants (*cpr1-4 aba1-8*, *zed1-D* – 28 °C) [174,175] and temperature shift (Col-0 – 20 °C) (Table 17).

To test this hypothesis, the expression pattern was compared to chosen published biotic and abiotic stress data sets in a cluster analysis. Gene expression data of experiments with the autoimmune mutants *cpr1-4 aba1-8* (*constitutive expresser of PR genes 1*; *ABA deficient 1*) and *zed1-D* (*HOPZ-ETI-deficient 1*) [174,175], an infection experiment of Col-0 and *pad4-1* plants with *Pseudomonas syringae* pv. *tomato* (*Pto*) DC3000 and *Pto* AvrRpt2 [176], and a treatment of Col-0 plants with flg22 peptide [38] were selected. In addition to the availability of the deposited data, the autoimmune mutants *cpr1-4 aba1-8* and *zed1-D* were chosen, because their phenotype is either very similar to the growth arrest phenotype observed in *saul1-1* or is temperature-dependent [174,175]. Infection experiments with *Pto* DC3000 and *Pto* containing the additional effector AvrRpt2 were selected to test for PTI and ETI-like inductions of the immune response. In addition, experiments were selected, which were performed in a *pad4-1* background using the

same pathogens, because the *saul1-1* phenotype has been reported to be *PAD4*-dependent [135]. In order to investigate similarities between regulations in the *saul1-1* mutant and PTI, which is dependent on the activation of pattern-recognition receptors (PRRs), an experiment of Col-0 plants treated with flg22 protein was selected.

The clustering analysis showed that all genes differentially regulated in *saul1-1* were not regulated in the cold stress and infection experiments with *Pto* DC3000 (Figure 18). In contrast, the overlap to gene regulations in plants, which were treated with flg22, was about 67 %. Thus, the gene expression changes in *saul1-1* appeared to be similar to those in pattern-triggered immunity (PTI) as expected. However, the gene-regulatory output was not identical, being in line with known differences between ETI and PTI. The overlap was higher when *saul1-1* data were compared to regulations in infection experiments with *Pto* expressing the effector AvrRpt2. Infections with these pathogens generally lead to a strong NLR activation and therefore a strong onset of ETI [231]. Hence, the high similarity supported the hypothesis that the *saul1-1* phenotype resembles autoimmune responses and therefore an initiation of ETI.

When comparing expression changes in Col-0 and *pad4-1* plants, which were both infected with *Pto* AvrRpt2, both are equally related to *saul1-1* (Figure 18). Only 5 % of the DEGs differed between both experiments. This is particularly remarkable, since PAD4 mediates SA-dependent downstream signaling during ETI [73]. One would expect that expression changes in *pad4-1*, a mutant that does not show any SA-dependent regulation, is more different from the expression changes in *saul1-1* observed after 6 h than Col-0. Interestingly, this does not seem to be the case. Thus, it can be concluded that at this point of time SA may no longer play the main role. This is similar to recent findings that could show that parts of ETI are SA-independent [76]. Nevertheless, the initiation of the *saul1-1* phenotype and therefore subsequent hypersensitive response (HR) has been described to be SA-dependent [132]. Although no complete rescue in a *saul1-1 pad4-1* double mutant was detected [135]. Thus, SA could act as an early initiator of the ETI phenotype in *saul1-1* and not be essential for downstream regulations in *saul1-1*.

The highest similarity of gene regulation was determined between *saul1-1* and both autoimmune mutants that are known to exhibit reduced growth phenotypes similar to *saul1-1* (Figure 18) [174,175]. In *zed1-D* plants, which showed the largest overlap in gene expression with *saul1-1*, a mutation in the *ZED1* gene (also called *HOPZ-ETI-DEFICIENT 1*) results in an Asn173Ser substitution in ZED1. The ZED1 protein is a pseudokinase that interacts with the NLR ZAR1 (ZYGOTIC ARREST 1) and is required for the defense response induction by the *Pto* effector HopZ1a [232]. In high temperature *zed1-D* exhibits a dwarf phenotype and high levels of defense associated genes. This was shown to depend on the *SNC1* gene (*Suppressor of NPR1-1, constitutive 1*), because ZED1 represses the accumulation of *SNC1* transcripts under high temperatures. In *zed1-D*, the Asn173Ser mutation results in accumulation of *SNC1* transcripts and thus in a dwarf phenotype and detectable HR [175]. A similar dependence on *SNC1* was described for the autoimmune phenotype of *cpr1-4 aba1-8* mutants. Accumulation of SNC1 was controlled by *CPR1*, which governs the abundance of miRNAs targeting *SNC1* transcripts [174]. The abscisic acid (ABA) deficiency of *aba1-8* mutants was related to an increase in nuclear localization of SNC1, which is required for its correct function [233].

Taken together, the cluster analyses confirmed the hypothesis that *saul1-1* is a bona fide autoimmune mutant. It is therefore an optimal model system to study the gene regulations during the ETI response. For future research it would also be highly interesting to investigate differences between *saul1-1* and the previously described autoimmune mutants, because it has been hypothesized that ETI pathways vary between different NLRs [4].

4.1.5 EXPLORING EARLY-REGULATED GENES IN *saul1-1*

The regulation of biological processes often involves a very early induction of key regulatory genes that may act high in the hierarchy of the signaling cascade [38]. Therefore, the molecular genetic analyses were at first focused on selected candidate genes among the 19 genes, which

were changed in their expression after 1 h and 2 h, to identify possible key regulators of ETI in *saul1-1* mutants. Their selection was dependent both on availability of mutants and on a previously described connection to plant immunity. The five genes, which were selected, were *TRX5* (*Thioredoxin H-Type 5*), *AT5G52760*, *AT3G28580*, *AT4G16260* and *WRKY46*.

The *TRX5* gene appeared to be highly interesting, because the encoded thioredoxin is involved in the SA-dependent monomerization and therefore activation of NPR1 (NONEXPRESSER OF PR GENES 1) [234]. NPR1 is the key regulator of SA-dependent defense downstream signaling, and accordingly *TRX5* could be one of the earliest regulators during ETI (see chapter 1.1.2) [74]. In case of *AT5G52760*, it has been reported that this gene is differentially expressed during flg22 treatment [183]. Thus, this gene may also function in the early regulation of ETI. *AT3G28580* is known to respond selectively to reactive oxygen species (ROS) [235,236] and to be regulated during the immune response [186,187]. Since the ROS burst is one of the first events in ETI (see chapter 1.1.2) [25], *AT3G28580* could be one of the earliest regulators. *AT4G16260* and *WRKY46* are both known to be involved in the immune response, although both have been associated with the basal immune response. *WRKY46* is known to interact with other *WRKY* TFs in the regulation of PTI [237]. AT4G16260 is a target of pathogenic effectors and acting as a positive regulator of plant immunity [238]. Since PTI and ETI are highly connected, both genes could as well be involved in the early regulation of ETI.

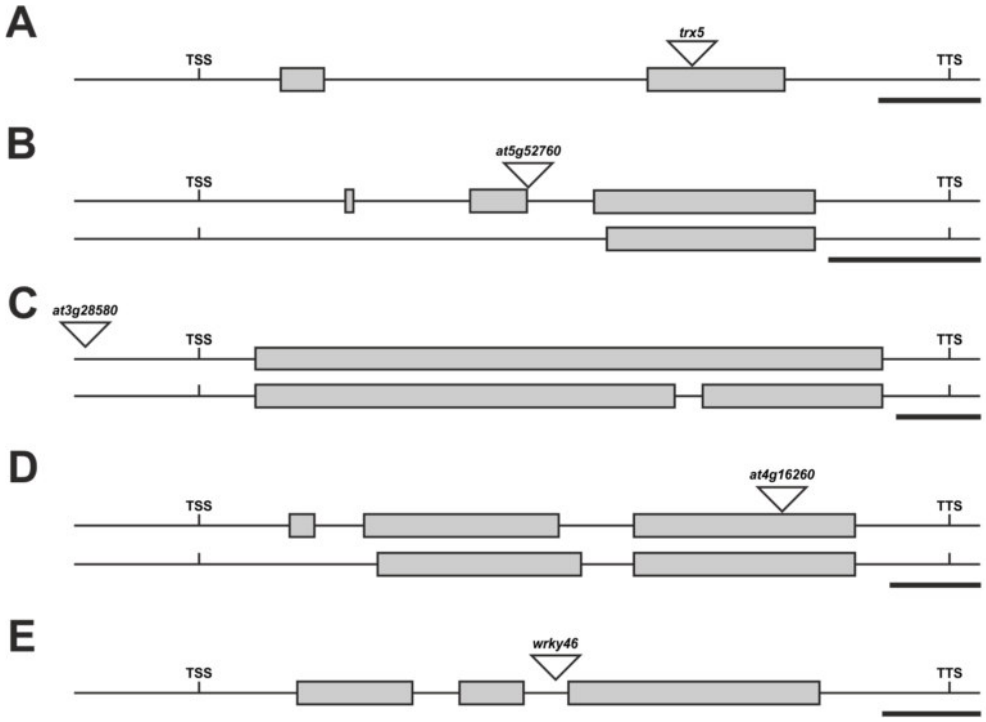


Figure 19: T-DNA insertion sites in different mutant lines of early regulated genes in *saul1-1* [156,157]: Schematic representation of five early regulated genes with the positions of the T-DNA insertions (indicated by a triangle) in the mutant lines. Exons are depicted as grey boxes. **(A)** *TRX5* with the T-DNA insertion being located +976 bp. **(B)** *AT5G52760* with the T-DNA insertion being located +436 bp. **(C)** *AT3G28580* with the T-DNA insertion being located -273 bp. **(D)** *AT4G16260* with the T-DNA insertion being located +1299 bp. **(E)** *WRKY46* with the T-DNA insertion being located +733 bp. Positions are in respect to the transcription start site (TSS). In case of different splice variants being known, both are depicted. TSS and transcription termination sites (TTS) are marked with a bar, respectively. Scale bars represent 200 bp.

To analyze their function and characterize their putative position in the signaling cascade, double mutants should be generated between *saul1-1* and the respective single mutants to screen for phenotypical changes compared to *saul1-1* single mutants. It was hypothesized that in case the early induction of a gene is indeed important for *saul1-1* autoimmunity, the phenotype may be

rescued or at least partially changed. The generation of three double mutants was performed in context of the bachelor theses of Sally Marusoi and Simon Peter Meyer [156,157]. Additional two mutant lines were generated by Teresa Wulf. All five mutant lines were genotyped, and T-DNA insertion sites were determined by sequencing (Figure 19, Table S4). Whereas the mutation was detected in the exon and intron of two lines, respectively, the insertion was detected to be located 273 bases upstream of the transcription start site (TSS) in the *at3g28580* mutant.

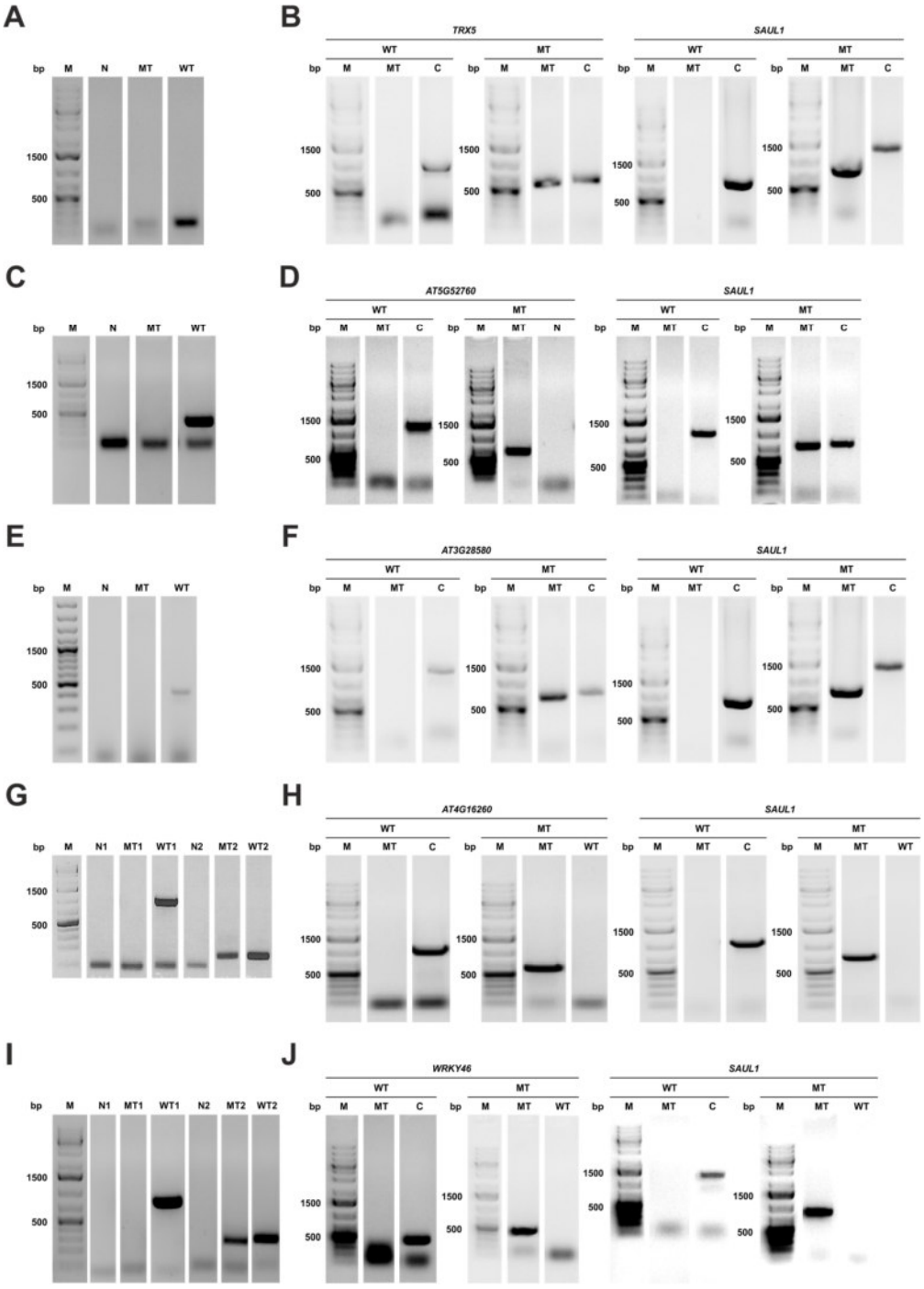


Figure 20: Genotyping and RT-PCR of double mutants of early regulated genes in *saul1-1* [156,157]**.** RT-PCR of single mutants with an expected product size of **(A)** *trx5* of 232 bp, **(C)** at*5g52760* of 345 bp, **(E)** *at3g28580* of 413 bp, **(G)** *at4g16260* of 1010 bp and 120 bp **(I)** *wrky46* of 843 bp and 289 bp. As a marker (M) the GeneRuler™ 1kb Plus DNA Ladder was used. Genotyping of double mutants of *saul1-1* with **(B)** *trx5*, **(D)** at*5g52760*, **(F)** *at3g28580*, **(H)** *at4g16260* and **(J)** *wrky46*. Depicted is for each gene the mutant (MT) and the wild type (WT) genotyping with the DNA ladder (M), the corresponding double mutant (MT) and the matching mutant or wild type control (C), negative control (N) or wild type sample (WT).

To investigate, whether these insertions indeed led to a loss of expression for the gene of interest (GOIs), transcript levels were monitored by reverse transcription PCR (RT-PCR) experiments. These could show that full-length transcripts of the GOI were not detectable in any of the mutant lines, because no corresponding band was observed in contrast to the WT samples (Figure 20A,C,E,G,I). In case of *at4g16260* and *wrky46* a partial and smaller transcript was detected. This corresponded in case of *at4g16260* to the first exon, as the T-DNA insertion was determined to reside in the first intron (Figure 19D). For *wrky46* the partial transcript corresponded to the first and second exons, in accordance with the T-DNA insertion localized in the second intron (Figure 19E). Although *AT4G16260* and *WRKY46* were still partially expressed, it was likely that such aberrant transcripts get degraded. This is either done by nonsense-mediated decay in case of a premature stop codon or by non-stop decay if the stop codon is missing due to the T-DNA insertion [239].

Consequently, all five lines were used to generate double mutants with *saul1-1*. Therefore, homozygous single mutants were used for crossings with *saul1-1*, and the F_1-generation was self-pollinated. Afterwards, the F_2-generation was genotyped and screened for double homozygous plants (Figure 20B, D, F, H, J). For each of the double mutants used, corresponding T-DNA insertion bands were detected for both affected genes (MT) and no bands in case of the control PCRs determining the WT gene status. Consequently, in all double mutants homozygous T-DNA insertions were detected, and thus the double mutants were used for further experiments.

It was assumed that these genes regulate the early onset of *saul1-1*autoimmunity and therefore hypothesized that in the double mutants the *saul1-1* phenotype could be reversed. Double mutants were grown together with *saul1-1* for 12 days at 25 °C and afterwards shifted to 20 °C. After five days, a point of time at which the autoimmune phenotype is clearly detectable in *saul1-1* plants (Figure 13A), all investigated double mutants displayed completely yellow leaves. In addition, a growth arrest, comparable to the one of *saul1-1* single mutant plants, was detected (Figure 21). All double mutants did not differ in their appearance from *saul1-1* plants and did thus not revert the phenotype.

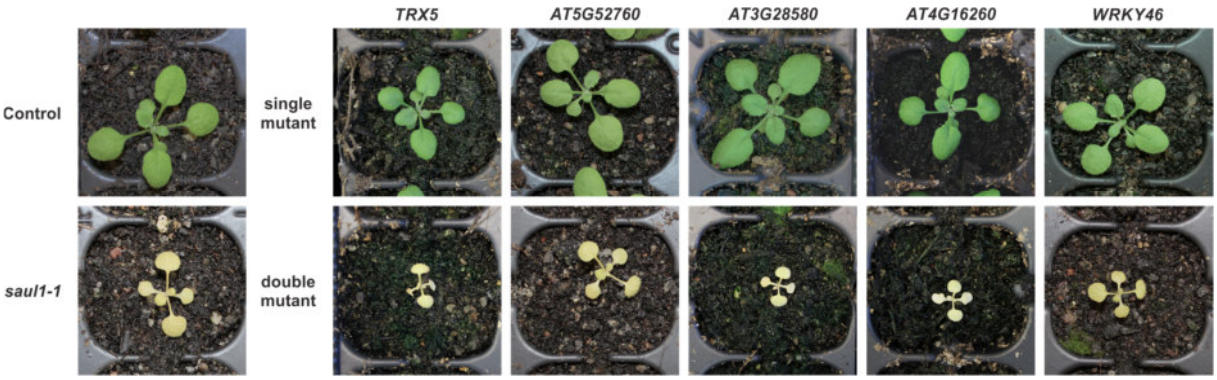


Figure 21: Phenotype of mutants of early regulated genes and double mutants with *saul1-1*. Control (WT), *saul1-1*, single mutants and double mutants, which were all crossed with *saul1-1*, were grown on soil for 12 days at 25 °C and shifted to 20 °C for 5 days.

In conclusion, the selected genes may not be essential or solely involved in the observed regulations of the ETI-like autoimmunity in *saul1-1*. Nevertheless, each of the investigated genes has two or more homologs in *A. thaliana* with a sequence identity of at least 50 % (Table S5). A sequence identity of 50 % and more is sufficient in a related protein to potentially result in a similar function [240]. Especially since *TRX5*, *AT5G52760* and *AT3G28580* have homologs with a sequence identity of more than 70 % that are highly likely to have a similar function (Table S5). Therefore, homologs of the investigated genes may take over their function, and a reversal of the *saul1-1* autoimmune phenotype would not occur. One particular example is *TRX3*, which is a homolog of *TRX5*, and is involved as well in the regulations of NPR1 oligomerization and therefore

the early onset of ETI [234]. In summary, it can be said that only mutants generated between *saul1-1* and all homologs may make it possible to identify potential early regulators of the ETI observed in *saul1-1*. Consequently, further experiments will focus on the generation of these multiple mutants to overcome redundancy.

In addition, there may be good candidates among the remaining 15 DEGs such as the two TF genes *ERF2* (*Ethylene Response Factor 2*) and *ZAT7* (*Zinc Finger of Arabidopsis thaliana 7*).Both genes are highly related to immune responses (see chapter 4.1.8) [241–243] and are thus good candidates to be involved in the regulation of the early regulatory mechanisms in *saul1-1* and possibly ETI. Their potential function will be discussed in following chapters. Although the mutant line for *ZAT7* is currently being investigated, no mutant lines were available for the *ERF2* gene. In the future, however, knock-out lines could be generated by using the Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated protein 9 (Cas9) approach. This method allows for specific generation of *erf2* mutants [244], without potential additional insertions that are found in around 13 % of T-DNA-derived mutant lines [245]. This would help to determine the function of *ERF2* in context of the ETI-based autoimmune phenotype in *saul1-1*.

4.1.6 EARLY EXPRESSION CHANGES IN *saul1-1*

With respect to the kinetics of expression changes in *saul1-1*, the identified DEGs can be clustered into seven sub clades (Figure 22). These consist of genes, which are most strongly regulated after 3 h (cluster III, V and VI), which are differentially expressed 6 h after the temperature shift (cluster I and VII) and which are more dynamic (Cluster II and IV). For clusters II and IV it was not possible to detect shared biological pathways, because those genes seemed to be very heterogenous. This was also the case for cluster VI that additional to the heterogeneity contained only a small number of regulated genes (95 DEGs). All other clusters will be described in more detail in the following.

Up-regulated genes - Clusters I and III

The genes up-regulated after 3 h (Cluster III) were strongly linked to the defense response, as directly indicated by a variety of GO terms including *defense response, innate immune responses or response to chitin* (Figure 22). The link to immune responses was also supported by many genes with the GO terms *signal transduction*, *response to oxygen-containing compound* and *response to acid chemical*. These GO terms are tightly associated with diverse plant hormone signaling pathways, namely jasmonic acid (JA), SA and ethylene (ET) pathways. Generally, JA and SA are important hormones in defense response. Whereas SA is tightly linked to signaling in response to biotrophic pathogens, JA is mostly linked to infections by necrotrophic ones [47]. It was shown, that plants, which were SA-non-accumulating and infected with the JA-inducing pathogen *Pto* DC3000 produced more JA than the non-infected plants [246]. These results and additional studies lead to the assumption that SA and JA act on a large scale antagonistically [247,248]. Mechanistically, SA mainly inhibits downstream signaling of the JA-mediated defense response and, to a certain degree, even its biosynthesis [249]. Whereas the accumulation of SA has been shown in *saul1-1* autoimmunity, JA levels were not studied [135]. The RNA-seq analyses suggested that in *saul1-1* autoimmunity both plant hormones were elevated. This may appear unusual, however, a production of both hormones has previously been observed in plants that were infected with a pathogen known to induce ETI [250,251]. Whereas an antagonistic interaction between SA and JA has generally been shown at high concentrations, a synergistic regulation was detected at very low hormone concentrations [252]. In addition, it was shown that SA may activate JA signaling by promoting degradation of JA transcriptional repressors and thus induce ETI [75]. Therefore, both hormones seem to be involved in the regulation of ETI. In *saul1-1* this synergistic interplay appeared to be shifted to a SA-dependent regulation after 6 h, because JA-associated pathways were not any more detectable in cluster I. This corresponded to the published SA dependence of the *saul1-1* autoimmune phenotype [135].

With regard to hormone signaling the up-regulation of multiple genes encoding members of the ERF transcription factor family suggested a participation of ET in *saul1* autoimmunity. ET is known to be associated with a broad spectrum of defense responses [253]. On one hand, ET is essential for the JA-dependent repression of SA in case of necrotrophic pathogens, mediated by the TF ORA59 (OCTADECANOID-RESPONSIVE ARABIDOPSIS AP2/ERF59) [249]. On the other hand, ET acts synergistically with SA in oil seed rape during an infection with a hemibiotrophic pathogen [254]. This was in line with the results on gene expression changes in *saul1-1*. The importance of ET-mediated regulations of plant defense were supported by recent findings showing that the type III effector HoPAF1 from *Pto* DC3000 is targeting the ET synthesis pathway and suppresses plant immunity [255]. With respect to the regulation of the ETI and HR it was shown, that ET levels are up-regulated at a very early time point, prior to the onset of the SA synthesis, and that this correlated with ROS detection [256]. Genes that are associated with ROS were detected in *saul1-1* as well, as part of the GO term *response to oxygen-containing compound* and *response to oxidative stress* (Cluster III). These GO terms were also detected in cluster I, although more genes were associated with them. Therefore, ROS-dependent gene regulation, which is initiated after 3 h, was clearly more predominant after 6 h and could regulate the progression of immune responses.

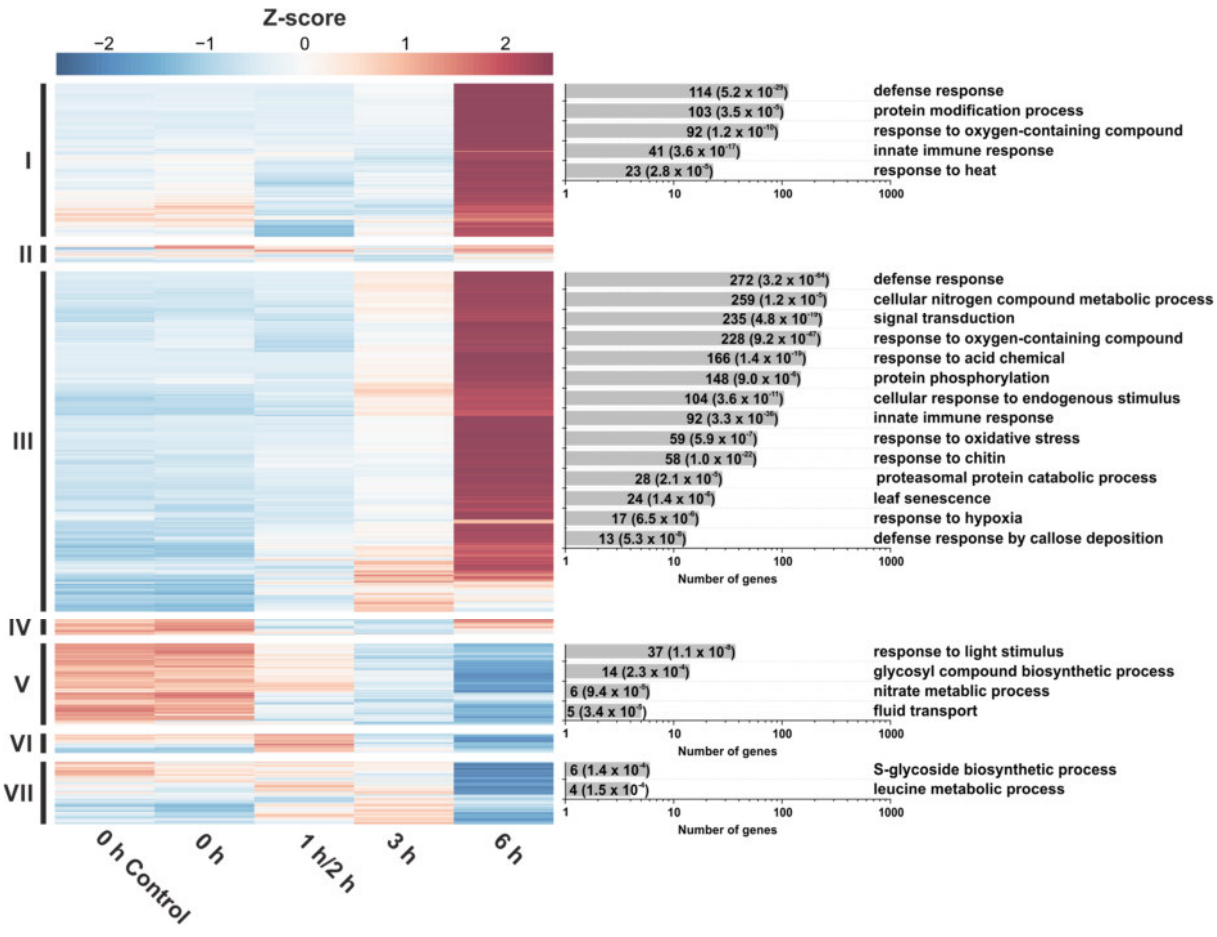


Figure 22: Transcriptional cluster analysis of early regulations in *saul1-1*. Heatmap of genes, which were differentially expressed ($p \leq 0.01$) after 0 h, 1 h/2 h, 3 h and 6 h in *saul1-1* in comparison to the segregating wild type. Expression values are depicted as normalized Z-scores. All seven expression clusters were used for further gene ontology analysis. Pathways which were significantly enriched ($p \leq 0.01$) are depicted on the right side of the heatmap. Adjusted p values of each term are written in brackets next to the gene count.

ABA responsive pathways were up-regulated in *saul1-1*, too (Cluster III). Despite its developmental role and association with abiotic stress, increased ABA levels are known to lead to a higher susceptibility to certain pathogens, for example *Pto* and *Blumeria graminis* f. sp. *hordei* [257,258]. These pathogens exploit this regulatory system by increasing the synthesis of ABA to dampen the SA and JA-ET answer [259]. This is due to increased ABA levels leading to a reduction of JA, ET [260] and SA, directly reducing the expression of ABA-synthesizing enzymes [261,262]. Therefore, this hormone is mostly described as a negative regulator of the immune response [263]. Contradictory to this, ABA also appears to play an important role in enhancing the defense response as well. In infection studies with necrotrophic pathogens such as *Alternaria brassicicola* higher ABA levels led to an increase in JA production and to a stronger resistance [264]. In conclusion, detecting ABA responsive genes in *saul1-1* plants could mean, that early regulations in ETI are ABA-dependent as well, as the ABA regulatory network is interwoven with nearly all defense pathways and might be a more subtle regulator of those [263,265]. In addition, ABA-dependent pathways were not detected in cluster I but only in cluster III. In conclusion, ABA-related pathways are not initiated after 6 h, but after 3 h. Therefore, ABA-dependent pathways are particularly involved in the early regulations of ETI. This seems to be quite reliable, because ABA has been linked to ETI before as well [266,267].

Besides the described hormonal regulations, the GO term *defense response by callose deposition* with several linked DEGs was detected in Cluster III. This was especially interesting, as a cell wall thickening due to lignin and callose deposition has been described as part of the *saul1-1* phenotype [132]. Those changes in the cell wall composition are known to highly increase disease resistance in plants by acting as a mechanical barrier [51,268]. The importance of this defense mechanism is illustrated by the fact, that the pathway is involved in cell wall thickening in *A. thaliana* and is suppressed by the type III effector AvrPto from *Pto* DC3000 [269]. The induction of cell wall thickening due to callose or lignin deposition is mostly described as PTI-related and takes place around 8 h after an infection or treatment with a PAMP, for example flg22 [55]. In *saul1-1* autoimmunity, the first morphological changes associated with cell wall thickening were detectable 2 d after the temperature shift [132], whereas the initial genetic regulations take place even as early as 3 h. Interestingly, after 6 h (Cluster I), no regulations of callose-associated genes were detectable. Apparently, the thickening of the cell wall was initiated at a very early stage. These results showed that cell wall thickening in *A. thaliana* occurs not only in association with PTI but also with ETI. This is supported by experiments in rice that revealed an induction of cell wall thickening due to an ectopic expression of the AvrXa27 effector from *Xanthomonas oryzae* pv. *oryzae* [270]. In an ETI context callose deposition and therefore thickening of cell walls could function as a physical barrier to prevent the spread of the pathogen and thus HR. This seems very reasonable, because during cell death a broad number of ROS, signaling compounds and vacuolar contents are released, which could drive adjacent cells into HR [85]. On a proteomic level it was shown that immune suppressing proteins are up-regulated in neighboring cells, which are thought to prevent an uncontrolled spread as well [271]. In conclusion, callose deposition in an ETI context seem to be quite reasonable. Future investigations on the *saul1-1* autoimmune model could give valuable insights into this aspect of plant defense.

Down-regulated genes - Clusters V and VII

The Clusters V and VII contained almost exclusively down-regulated genes. These were linked to the GO terms *glycosyl compound biosynthetic process* and *S-glycoside biosynthetic process*. This may indicate that glycosylation was negatively regulated during the onset of the *saul1-1* phenotype. By applying i*n silico* analyses using *Genevestigator* it was possible to detect that these genes were also down-regulated during NLR-mediated defense responses and upon flg22 treatment [183,185,186]. The down-regulation of those genes may thus generally be associated with immunity. However, glycosylation is mediated by a broad range of different pathways with diverging outcomes.

Remarkably, most detected genes are mainly involved in S- and not N-glycosylation. This is quite plausible, because N- in contrast to S-glycosylation is one of the major post-translational modifications, which is highly important for correct protein folding and trafficking. With respect

to the immune system an incorrect or missing N-glycosylation of the PRRs EFR (EF-TU RECEPTOR) and FLS2 (FLAGELLIN-SENSITIVE 2) has been shown to lead to non-functional receptors and retention in the endoplasmic reticulum [272]. Therefore, a down-regulation of N-glycosylation would be unexpected, since this posttranslational modification is vital for correct protein folding. In contrast, it is known that S-glycosylation is an important step in secondary metabolite synthesis [273]. Although those pathways, which are down-regulated in *saul1-1*, are quite diverse, most of them have one characteristic in common: they are highly connected to the interplay between JA and SA [274–276]. For instance, the down-regulated TF *MYB28* and *MYB29*, which promote glucosinolate biosynthesis, are known to be positively regulated by JA and suppressed by SA [276,277]. Additionally, repression of these TFs is known to lead to decreased growth [278]. This is especially interesting, because a characteristic of the *saul1-1* and other autoimmune-associated phenotypes is growth arrest [123,134]. Genes regulated by MYB28 or MYB29 comprise other TFs such as *WRKY18* and *WRKY40*, which in turn regulate members of the *CYP* (*Cytochrome P450*) family [276]. These, in particular *CYP79F2* and *CYP83A1*, were found to be down-regulated in *saul1-1* as well. Interestingly, it has been reported that a mutant of *CYP83A1* leads to a higher resistance against powdery mildew. This was due to an accumulation of camalexin, one of the major phytoalexins [43]. Phytoalexins are known to be expressed upon infection with necrotrophic pathogens and to be fungistatic [279]. In context of ETI, they have not been excessively studied, but are thought to be expressed to prevent a secondary infection [271]. In addition, the regulations in *cyp83a1* are linked to SA accumulation, because the increased resistance was found to be *EDS1*-, *PAD4*- and *NDR1* (NON RACE-SPECIFIC DISEASE RESISTANCE 1)-dependent [43]. Other genes involved in S-glycosylation, like *FMOGS-OX1* (*Flavin Monooxygenase GS-OX1*) and *FMOGS-OX3*, are known to be JA- and SA-responsive as well [275]. Therefore, those diverse pathways connected to S-glycosylation, could be part of downstream regulations reacting to elevated SA levels in *saul1-1* [135].

In summary, the gene expression changes 6 h and especially 3 h after the temperature shift in *saul1-1* appeared to be regulated in a quite complex manner involving an interplay of the hormones SA, JA, ET and ABA. These regulations seem to initiate all downstream regulations, since they were dominantly detected in cluster III and this cluster contained genes that were differentially expressed from 3 h onwards. This assumption is supported by the observation that after 6 h (Cluster I) fewer genes were differentially expressed for the first time. Consequently, the *saul1-1* autoimmunity was defined by a large number of genes, which initiates after 3 h pathways that result in the *saul1-1* phenotype. Therefore, clustering all DEGs in *saul1-1* for the different time points was extremely helpful to get first insights into genes with similar behavior and into activation of pathways. However, genes that were down- and up-regulated were separated during this analysis. This may result in an insufficient amount of enough genes being detected to significantly enrich specific GO terms. In conclusion, some signaling pathways involved in *saul1-1* may not be detected in GO analysis because they contain up- and down-regulated genes.

4.1.7 GLOBAL EXPRESSION CHANGES IN *saul1-1*

To address this potential problem just mentioned above, a global GO clustering was conducted covering all DEGs of all points of time (Figure 23). For each detected GO term, the composition of up- and down-regulated DEGs for the respective points in time was depicted as a pie chart. GO terms that belong to the same pathway were clustered into larger groups that were visualized by colored areas. GO terms were connected based on the genes shared between multiple GO terms, since some genes were described to be involved in multiple pathways. It should also be noted that some GO terms are more general than others. This is due to the fact that some pathways have been clustered on the basis of a common mechanism, which takes place in all analyzed GO terms.

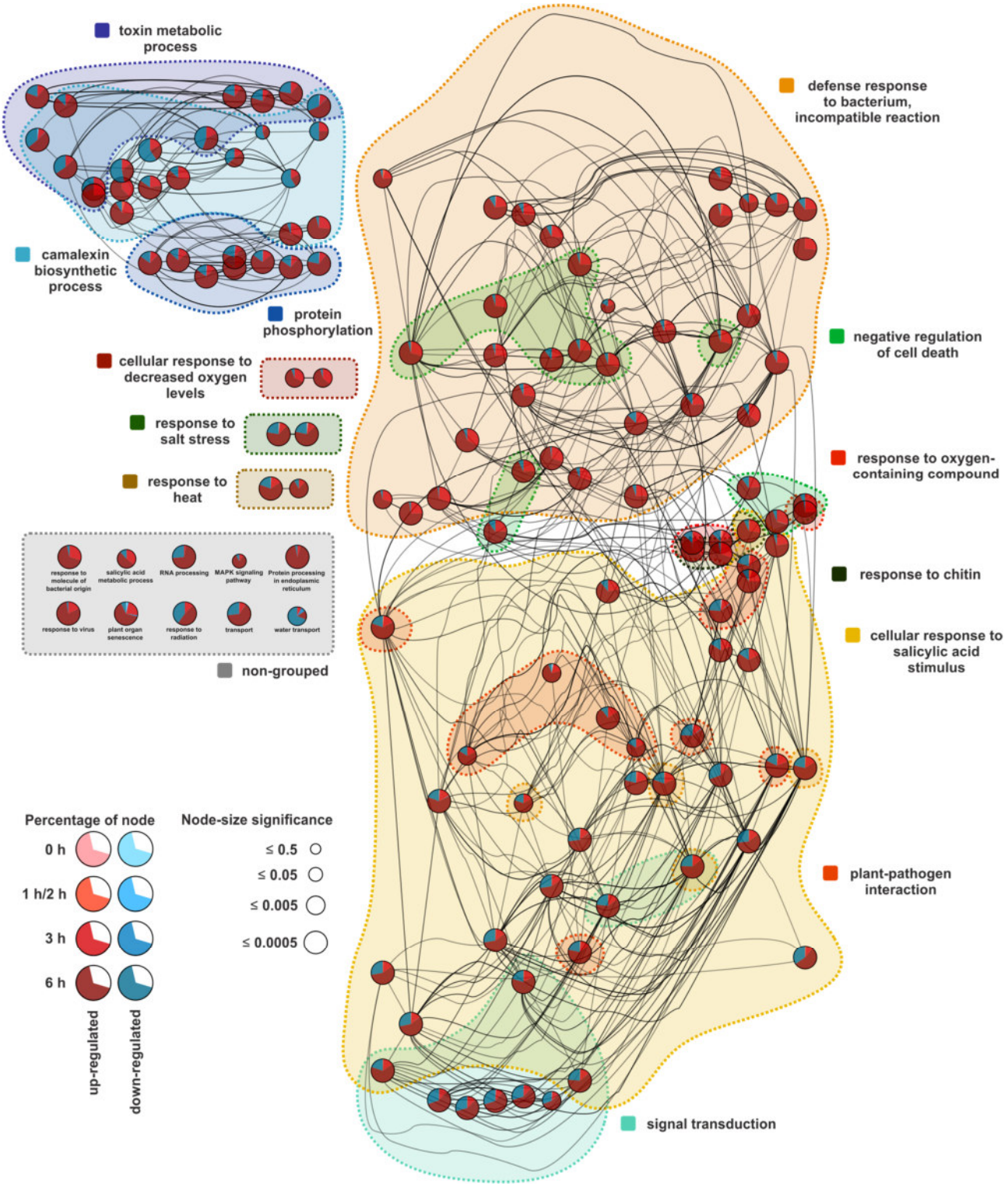


Figure 23: GO analysis of early regulated genes in *saul1-1*. Hierarchical clustering of GO terms in *saul1-1*. All GO term nodes are represented with the respective percentage of DEGs of a certain time point. Larger groups are defined by their color and the most significant enriched GO term. GO terms were determined as enriched, if the initial cluster consisted of at least 1 % of the respective DEGs and the final term had an adjusted p value ≤ 0.01. These were defined as connected, if their corresponding kappa score was higher or equal to 0.3. In addition, functional groups were clustered together if they shared 30 % or more of their genes. After all selection criteria were applied, 2183 (56.03 %) genes were associated to the pathways.

The clustering analysis identified the enrichment of four different groups of global GO terms. The two groups containing most GO terms were *defense response to bacterium, incompatible reaction*

and *cellular response to salicylic acid stimulus* (Figure 23, orange and yellow cluster). These pathways are highly connected to the ETI response. An incompatible reaction generally reflects a defense response against an infection with an avirulent pathogen, resulting in HR [280]. Accordingly, an incompatible defense reaction here reflects the *SOC3*-dependent autoimmune phenotype of *saul1-1* (orange cluster).

The *saul1-1* phenotype has been reported to depend on SA [135]. In line with this SA dependency the largest GO term cluster was determined to be *cellular response to salicylic acid stimulus* (yellow cluster). An additional smaller array, involving three different metabolic pathways, *toxin metabolic process*, *protein phosphorylation* and *camalexin biosynthetic process, respectively,* was detected as well (blue clusters).

Preceding cluster analysis already pointed towards an elevated camalexin synthesis in *saul1-1*, however, only down-regulated genes were detected in this analysis (see chapter 4.1.6). This global GO network in *saul1-1* autoimmunity identified additional up-regulated genes in this pathway. Among others, *CYP79B2* and *CYP79B3* were detected to be up-regulated after 3 h. Both are involved in the conversion of tryptophan to indole-3-acetaldoxime, a precursor of camalexin [281]. Acting further downstream *PAD3* gene expression was up-regulated after 6 h. PAD3 is a cytochrome P450 enzyme that catalyzes the final step during camalexin synthesis [281]. In conclusion, camalexin synthesis could be induced in *saul1-1* as early as 3 h. Camalexin, as well as other phytoalexins, are hypothesized to prevent a secondary infection with necrotrophic pathogens [282]. This hypothesis was supported by up-regulation of camalexin synthesis genes as shown in other ETI-inducing experiments [271]. An involvement of camalexin in the onset of the HR could take place as well, because higher concentrations were found to be cytotoxic and could therefore be the cause of subsequent HR [56]. This is consistent with previous reports that HR may not be a finely tuned defense pathway, but a consequence of downstream signaling during the immune response [85]. Thus, the defense against a secondary necrotrophic infection during ETI may be highly important, and first regulatory mechanisms can be detected a long time before the onset of HR.

This is underlined by findings showing that cell wall thickening, which could function as a physical barrier to necrotrophic pathogens, was detectable 3 d earlier in *saul1-1* than the lesioning of all aboveground organs. Accordingly, this analysis does not only confirm that *saul1*-1 exhibits an ETI-related phenotype, but gives insights into the early regulations and their complexity.

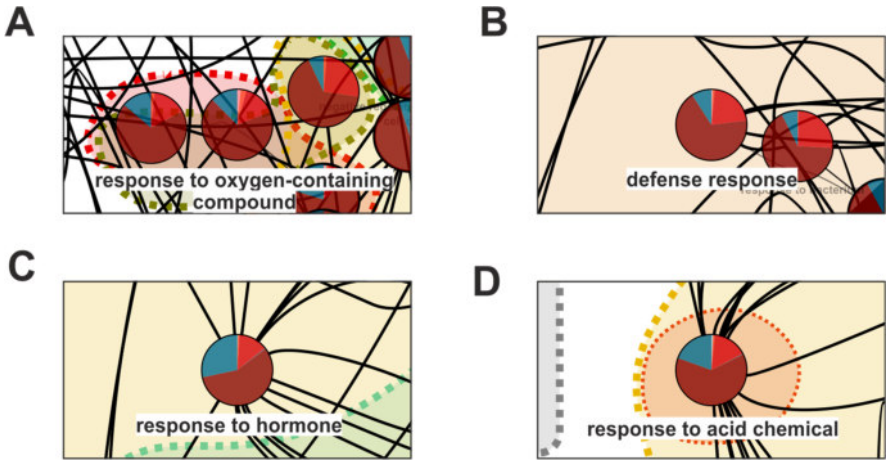


Figure 24: Excerpts of the GO analysis. Excerpts of GO terms, which contained several DEGs that were regulated after 1 h/2 h and 3 h. These are **(A)** *response to oxygen-containing compound* as part of the *response to oxygen-containing compound* cluster (red), which involved mostly DEGs being regulated after 1 h/2 h, **(B)** *defense response* part of the *defense response to bacterium* as part of the *incompatible reaction* cluster (bright orange), which involved mostly DEGs being regulated after 1 h/2 h, **(C)** *response to hormone* being part of the *cellular response to salicylic acid stimulus* cluster (yellow), which involved mostly DEGs being regulated after 1 h/2 h and **(D)** *response to acid chemical* being part of the *cellular response to salicylic acid stimulus* cluster (yellow), which involved mostly DEGs being regulated after 3 h.

This regulatory map was used thereafter to detect pathways containing DEGs, which were regulated early. These DEGs may function to trigger subsequent regulations (Figure 24). The detected GO terms, which contained at least three DEGs being regulated after 1 h/2 h, were *response to oxygen-containing compound*, *defense response*, *response to hormone*, *response to acid chemical* and *organonitrogen compound metabolic process*. The early regulated genes were *THI2.1* (*THIONIN2.1*), *ERF2* and *AT3G28580*. The comparison of these GO terms revealed that nearly all regulations were previously characterized as responses to signaling molecules, which are known to be essential immune regulators [283]. This was consistent with the analysis of Cluster III (Figure 22) pointing towards the involvement of the interplay between JA and SA in early regulations in *saul1-1* (see chapter 4.1.5). Increasing SA levels have been detected not before 12 h after the initiation of the *saul1-1* phenotype [135]. However, it has been proposed that such regulatory networks may depend on low levels of these hormones only [252]. The regulation of more than 50 SA and JA responsive genes after 3 h may suggest that a putative rise in SA levels was already present but below the detection limit in previous experiments [135]. Therefore, an interplay of JA, SA, ET and ABA appears to be necessary to successfully initiate ETI.

When comparing these GO terms, an additional similarity was detected. Nearly all GO terms contain the same early regulated genes (Figure 24). This could hint towards their importance in the regulation of the ETI phenotype, because they have been associated with multiple regulatory pathways.

The *THI2.1* gene encodes a thionin, which is expressed upon pathogen infection and has been used as a marker gene for increasing JA levels [252,284]. In addition, *THI2.1* is involved in the interplay between JA and SA at low concentrations. It was shown that *THI2.1* is detectable at low SA and JA concentrations and is down-regulated in response to higher amounts of SA [252]. This was also demonstrated in this RNA-seq experiment, since *THI2.1* was not differentially expressed after 3 h and 6 h, but only after 1 h/2 h. Accordingly, *THI2.*1 could be important for the early regulations in *saul1-1*. Nevertheless, it was not possible to confirm the regulation of *THI2.1* using qPCR (Figure 16). Thus, those results have to be interpreted carefully.

In contrast, the differential expression of *ERF2* was confirmed by qPCR (Figure 16). Interestingly, the *ERF2* gene was previously connected to defense response and has been reported to be ET-dependent [285]. In addition, the clustering analysis has already shown that other members of the *ERF* family were also up-regulated in *saul1-1* (Figure 22, Cluster III), emphasizing that this class of TFs may well be important for *saul1-1* autoimmunity. The putative function of ERF2 and other TFs will be discussed below (see chapter 4.1.8).

The *AT3G28580* gene encodes a putative AAA-ATPase. Whereas the protein function has not been shown yet, the *AT3G28580* promoter has been shown to specifically respond to singlet oxygen and has been used in promoter-reporter gene constructs to detect singlet oxygen levels [235,236]. Singlet oxygen and other ROS are rapid signaling molecules, which are of high importance during the defense response [32]. In addition, ROS contributes to HR in case of an incompatible defense response [85]. Hence, the up-regulation of *AT3G28580* may point towards the involvement of ROS in the early regulation of the ETI observed in *saul1-1*. *In silico* analysis support this hypothesis, because *AT3G28580* was shown to be up-regulated in plants overexpressing NLRs and in infection experiments [186,187]. In addition, further ROS responsive genes were detected after 3 h (89 genes) and after 6 h (320 genes), highlighting the importance of pathways, which are regulated by ROS.

The GO term analysis showed that already after 1 h/2 h different hormone-responsive pathways appeared to be active. This suggested that the initiation point of the ETI response could be a finely tuned interplay of small amounts of different plant hormones. These changes of hormonal levels could be very low and would therefore not be detectable with most techniques. Another problem with preceding infection experiments would be that they cannot be synchronized. This would result in the detection of averaged hormone levels. Thus smaller changes in hormone levels may be missed [286]. Consequently, *saul1-1* appears to be an ideal ETI model that for the first time suggests the presence of finely tuned hormone regulations resulting in the initiation of ETI.

4.1.8 TRANSCRIPTION FACTOR-MEDIATED REGULATION IN *SAUL1-1*

The regulation of complex biological processes often depends on the function of TFs. Many *TF* genes were found to be regulated early in *saul1-1*. Therefore, the enrichment of transcription factor families (TFF) was analyzed to investigate whether any TF families are involved in the regulation of the *saul1-1* phenotype. This revealed that most TFs were regulated not before 6 h and belong either to the *WRKY*, *NAC* (*NAM, ATAF1,2, CUC2*), *C3H*, *C2H2* or the *AP2/EREBP* (*APETALA 2/Ethylene-Responsive Element Binding Factor*) TF family (Figure 25). The largest TF family, whose members were differentially expressed, was the *WRKY* TF family. The *WRKY* TF members were generally up-regulated in *saul1-1*. In contrast, down-regulated TFs belonged to a large extent to the *bHLH* (*basic Helix-Loop-Helix*) family. Possible effects of these up- or down-regulated TF gene families are discussed below.

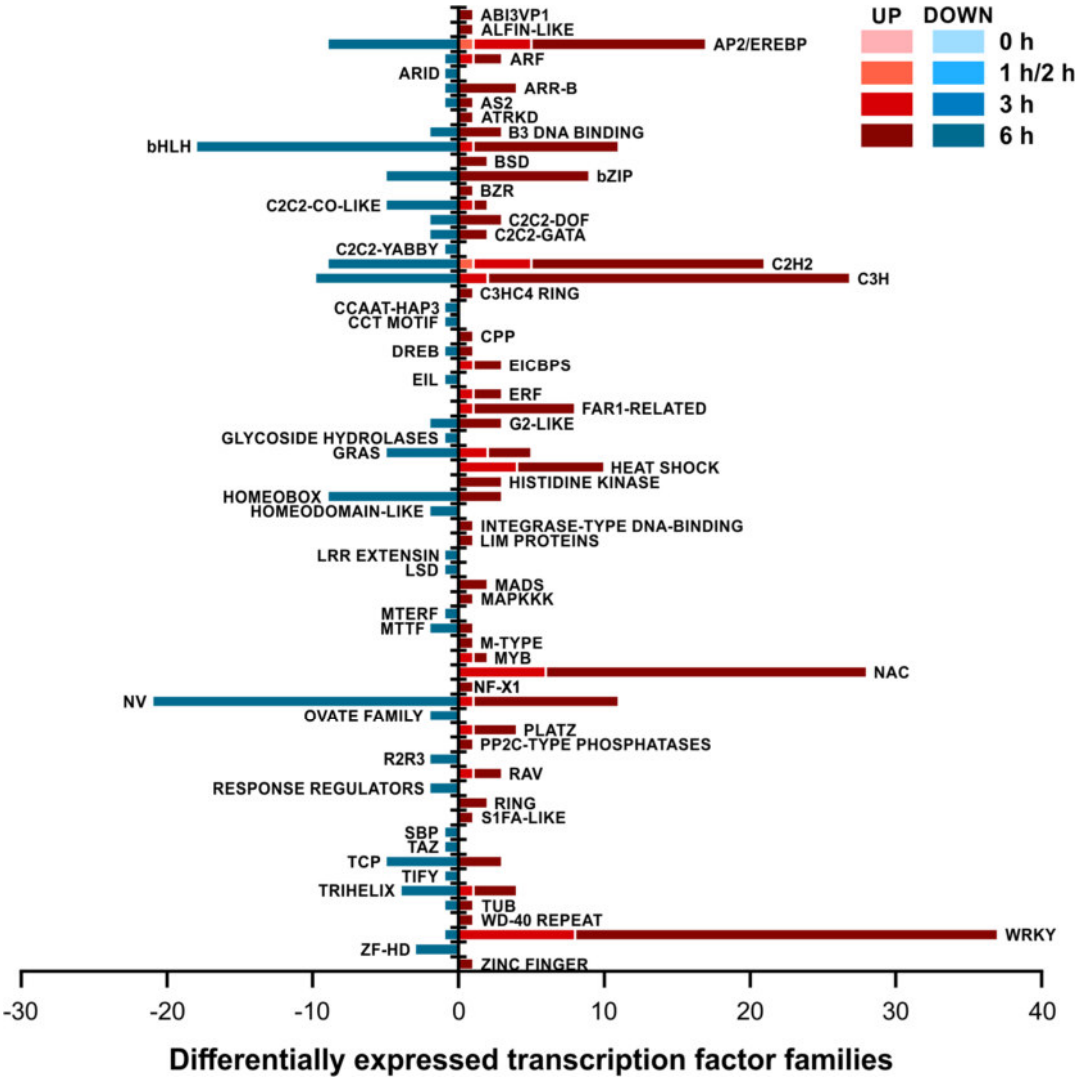


Figure 25: Differentially expressed transcription factor families. Histogram of transcription factor families and the number of corresponding genes, which were up- or down-regulated. Not all potential TFs were linked to a certain family and are therefore not assigned a value (NV).

WRKY family

Members of the *WRKY* gene family were named on the basis of their highly conserved WRKY domain, consisting of a WRKY peptide sequence and a zinc finger-like DNA-binding motif [287]. They are one of the 10 largest plant TF families and are known to be involved in a broad range of

regulatory processes [42,288]. Therefore, it is not possible to associate the whole family with a distinct process, because they are involved in various processes such as abiotic stress responses (*WRKY25*) [289], organ development (*WRKY44*) [290], senescence (*WRKY53*) [226] and biotic stress responses (*WRKY70*) [291]. Interestingly, neither of them can be associated with a single process, as most of them are known to be involved in multiple pathways [292]. The *WRKY* genes being regulated early during *saul1-1* autoimmunity may allow to decipher the interaction between different defense-associated hormone pathways. The up-regulated genes *WRKY18* and *WRKY60* are known to be negatively involved in ABA signaling upon elevation of SA levels [293]. Therefore, these up-regulated WRKYs may point towards rising SA levels between 3 h and 6 h. Other regulated WRKYs reflect this as well. *WRKY46*, which is known to be up-regulated upon SA was differentially expressed. The same was observed for the structurally related *WRKY70* and *WRKY54*. Particularly *WRKY70* and *WRKY54* are important regulators of SA-accumulation. Both repress SA synthesis, but enhance the induction of SA-responsive genes [294]. Recently both were strongly linked to JA-dependent defense responses as it was shown, that they repress the defense response against necrotrophic pathogens [295]. In addition *WRKY46* and *WRKY70* are known to be important in SA-dependent defense responses against *Pto* DC3000 [237,296]. As discussed before (see chapter 4.1.6 and 4.1.7) an early interplay between SA, JA and other plant hormones could be a characteristic of the *saul1-1* phenotype culminating in a SA-driven defense response. This is illustrated as well by the up-regulation of *WRKY50* and *WRKY51*, which were found to be essential for the SA-dependent repression of the JA-induced defense response [297]. The hypothesized interplay between JA and SA was supported by differential expression of *WRKY33*, which is known to be a major regulator of JA-dependent defense responses against necrotrophic pathogens. In addition WRKY33 suppresses SA-accumulation [298] and acts as a positive regulator of systemic acquired resistance (SAR) in a pipecolic acid-dependent manner [77].

NAC family

The second largest differentially expressed TF family are NAC TFs that are with over 100 members one of the largest plant TF families [299]. Although they are structurally related to each other, containing all a N-terminally conserved domain, they are involved in a broad range of processes. Some are involved in meristematic developmental processes (*CUC2*) [300], auxin signaling (*NAC1*) [301], abiotic stresses (*ANAC072*) [302] and the defense response (*NAC4*) [303]. The *NAC* TF that were regulated after 3 h were quite heterogenous as well. Three TFs, *NTL4 NTL6* and *ATF2* (*Thioredoxin F2*), are associated with plant immunity. *NTL4* and *NTL6* are both known to initiate immune responses upon abiotic stresses, although both respond quite differently. *NTL4* initiates ROS production in a positive feedback loop under heat stress by inducing the transcription of the *RBOH* (*Respiratory Burst Oxidase Homolog*) genes [304]. *NTL6* in contrast is responsible for the induction of several pathogenesis-related (PR) genes in response to cold stress [305]. In addition, overexpression of the wound-induced *ATF2* leads to a dwarfed phenotype that is very similar to autoimmune-related phenotypes [306]. However *ATF2* is slightly up-regulated upon infection with *Pto* DC3000 avrRps4 and acting mostly as a repressor of the defense response [185]. Thus, *ATF2* might be involved in different pathways, as it was shown that this TF can act as a transcriptional activator or repressor depending on the promoter sequence [307]. A performed *Genevestigator* analysis revealed that all detected NAC TFs are differentially expressed upon infection with *Pto* DC3000 AvrRps4 [185]. Therefore, their regulation could be typical in case of an ETI, although their specific function is not fully understood.

C3H family

The third largest TF family, which was found to be regulated after 3 h, was the *C3H* zinc finger TF family. C3H members contain a zinc binding motif that consists of three cysteines followed by one histidine [308]. Family members can contain up to six of these motifs and are widely distributed throughout the whole plant kingdom. In case of early regulated genes in *saul1-1* the closely related TFs *SZF1* (*Salt-inducible Zinc Finger 1*) and *SZF2* were found to be regulated. Both are known to be

salt stress responsively expressed and mediate the stress response [309]. Only *SZF1* has been linked to the immune response, since it is expressed upon herbivore infection and phosphorylated upon elevated Ca^{2+} levels [310]. Interestingly, *in silico Genevestigator* analyses showed, that these TFs are regulated in plants, which overexpress the NLR *RPS4* and were cultivated at 19 °C. These plants exhibited an autoimmune phenotype that is caused by the onset of ETI [187]. Therefore, these TFs are likely to be involved in the regulation of ETI, although their distinct function is still illusive.

bHLH family

A large group of down-regulated *TF* genes belonged to the *bHLH* family. Members of this family are characterized by their bHLH domain, which enables specific binding to DNA via the basic region. This domain is composed of two helices connected by a variable loop. The bHLH domain mediates protein-protein interactions, especially in homo- or heterodimer formation of bHLH TFs [311]. In *A. thaliana* TF family members are involved in various processes like trichome formation [311], vascular development [312] and JA signaling [313].

When taking a closer look at down-regulated *TF* genes, the involvement of gene products in growth regulation appeared to be a common feature. This is not very surprising since a dwarf phenotype is one of the hallmarks of *saul1-1* [123]. *HBI1* (*Homolog of BEE2 Interacting with IBH1*) is a prominent example, which is known to be involved in the growth-defense tradeoff, since *hbi1* mutants exhibit a growth arrest [314]. *HBI1* represses PAMP-induced growth arrest, acts as a negative regulator of immunity and promotes cell elongation. *HBI1* itself is a downstream regulator in brassinosteroid (BR) signaling and is one of the major mediators of BR-induced immunity suppression [314,315].

Table 40
FC values (log_2) of BR biosynthetic genes [316] in *saul1-1* after 3 h, 6 h (RNA-seq) and 6 h, 24 h, 48 h (Microarray [135]). Significant FC values (p value ≤ 0.01) are marked with an asterisk. Genes, which were not present on the ATH1 array are marked with *NA*.

	RNA-seq		Microarray		
Gene	3 h	6 h	6 h	24 h	48 h
DWF4	0.05	-0.19	-0.43	-0.76*	-1.09*
CPD	-0.03	-0.41*	-0.43	-1.08*	-2.25*
DET2	-0.09	-0.65*	-0.26	-0.61*	-0.85*
BR6ox1	0.01	1.07	*NA*	*NA*	*NA*
BR6ox2	-0.14	-0.53*	*NA*	*NA*	*NA*
CYP90C1	-0.02	0.09	*NA*	*NA*	*NA*
CYP90D1	-0.01	-0.11	*NA*	*NA*	*NA*
BAS1	-0.06	-0.62	-0.34	0.01	-0.89*
UGT73C5	0.20	0.28	*NA*	*NA*	*NA*
SMT2	0.08	0.09	-0.63*	0.05	-0.26
DWF1	0.00	-0.19	-0.19	-0.65*	-1.53*
DWF7	0.05	-0.15	0.05	-0.28*	-0.64*
DWF5	0.08	0.32	0.03	0.66*	0.76*

Interestingly, PTI and BR signaling are highly interconnected as they share *BAK1* (*BRI1-Associated Receptor Kinase 1*), which is involved in BR and flg22 perception. BAK1 dimerizes with the corresponding receptors BRI1 (BR INSENSITIVE 1) or FLS2 for BR or flg22 detection [317]. This interconnection can be observed in flg22-treated plants, which results in BR synthesis repression [316]. The same was observed in *saul1-1*. Three BR biosynthesis genes were found to be down-regulated in the RNA-seq experiment after 6 h (Table 40). Taking a look at a previously

performed microarray experiment [135] a similar pattern was detected during this thesis. A repression of BR biosynthesis-related genes was clearly observable (Table 40). Note that not all BR synthesis-involved genes were analyzed in the microarray data, since not all genes were represented on the ATH1 array.

This potential association of the dwarf phenotype in *saul1-1* with BR signaling was also demonstrated by down-regulation of other *bHLH* TFs, which are BR sensitive. *BEE2/3* (*BR Enhanced Expression 2* and *3*) and *BIM1* (*BES1-Interacting MYC-like 1*) have been described as BR inducible, positive regulators of BR signaling and promoters of plant growth [318]. They are also associated with a general growth-stress tradeoff [319]. In case of *BEE2*, which is a homolog of *HBI1*, over-expression led to the repression of defense associated genes and a higher susceptibility to pathogens [314]. Therefore, *BEE2* could act as well as a regulatory hub in BR signaling. All three TFs interact with negative regulators of BR signaling, the so-called *AIFs* (*ATBS1-Interacting Factors)*, and are therefore further connected [314,320]. Three of these *AIFs*, namely *AIF1*, *AIF2* and *AIF3*, were down-regulated in *saul1-1* as well. However, it might seem contradictory that these *AIFs*, which are negative regulators of BR signaling, were down-regulated. Nonetheless, the whole interplay between *HBI1* and *BEE2* and the *AIF* members is still not fully understood. A potential explanation could be a feedback loop involving negative regulators, which are under control of BR. This hypothesis seems to be plausible since previous experiments have shown that *AIF1* is up-regulated upon BR treatment and therefore BR-dependent [321]. Note that the experiment from which the discussed RNA-seq data were obtained was performed in roots. Thus, the regulation of *AIF* members may not be represented to the same extent in aboveground organs. In addition, two other TFs, which are closely related to the *AIF* family, were detected to be down-regulated. These were *SAC51* (*Suppressor of ACAULIS 51*) and *SACL1*. Both are involved in Xylem differentiation and hypocotyl growth [322]. They are negative regulators of xylem differentiation, which are up-regulated upon treatment with thermospermine. This is important, since plants, which lack thermospermine synthesis, exhibit a dwarf phenotype [323]. Because both TF are related to the *AIF* family they might be connected to BR signaling as well and exert similar functions [324].

Table 41
Differentially expressed TF families, involved genes and concerned regulatory pathways.

TF Family	Genes	Associated pathways	Time point
WRKY	*WRKY18*, *60*	Negative regulation of ABA signaling	3 h
	WRKY46, *54*, *70*	Negative regulation of SA synthesis, positive regulation of SA response	
	WRKY50, *51*	SA-dependent repression of JA	
NAC	*NTL4*, *6*	Induction of the immune response	3 h
	ATF2	Immune response repressor	
C3H	*SZF1*, *2*	Positive regulation of salt stress responses	3 h
AP2/EREBP	*ERF2*	Positive regulation of immune response, ethylene-dependent	1 h/2 h
	ERF1, *6*, *104*	Positive regulation of immune response against fungi, positive regulation of defensin expression	3 h
C2H2	*ZAT7*	Salt-stress related dwarf phenotype, induction of expression during infection with effector inserting pathogen	1 h/2 h
	AZF1	Negative regulation of auxin-dependent genes	3 h
bHLH	*HBI1*; *BEE2, 3*	Positive regulation of BR signaling	6 h

Table 41
continued

TF Family	Genes	Associated pathways	Time point
bHLH	*AIF1*, *2*, *3*	Negative regulation of BR signaling	6 h
	SAC51; *SACL1*	Positive regulation of hypocotyl growth	
	PIF1, *4*, *7*, *8*	Positive regulation of growth, GA-dependent, connected to BR	

Another class of bHLH TFs, which was found to be differentially regulated and is known to be involved in the growth-defense tradeoff, are the *PIFs* (*Phytochrome Interacting Factors).* They are described as interactors of phytochromes and are involved in various pathways [325]. In *saul1-1* *PIF1*, *PIF4*, *PIF7* and *PIF8* were found to be down-regulated. *PIF1* and *PIF4* are positive regulators of plant growth. In general, they are repressed by so-called DELLA proteins, which are degraded upon a gibberellin (GA) stimulus [326]. Recently, GA-mediated plant growth has been connected to BR signaling. It was shown that *PIF4* binds to the BR-activated TF *BZR1* (*Brassinazole-Resistant 1*) and both regulate synergistically multiple common target genes [327]. Thus, these results hint towards an interconnection between BR and GA signaling in promoting or repressing plant growth.

As a consequence, a negative regulation of BR signaling could be the driving factor of the growth reduction in *saul1-1*. This seems to be very likely, since a vast number of downstream *bHLH* TFs (especially the main regulatory TF *HBI1*) and multiple BR biosynthetic genes were down-regulated. In addition, *in silico* experiments could show, that all detected down-regulated bHLH members were found to be repressed upon flg22 treatment. An experiment to verify this potential connection would be a crossing between *saul1-1* and *35S::HBI1*. In case of growth reduction, which may rely on BR repression, the growth arrest in *saul1-1* could be reverted. Although it has to be mentioned that this double mutant would not uncouple growth and the immune response. An overexpression of *HBI1* would most likely lead to a suppression of the immune signaling [315].

AP2/EREBP family

Although the majority of TFs were up-regulated after 6 h (353 TFs), two TFs were already detected after 1 h/2 h. *ERF2* was one of them and belonged to the *AP2/EREBP* family. Members of this TF family are commonly known to be regulated due to abiotic stress [328]. However, *ERF2* belongs to the subgroup *ERF B3* that has been linked to ethylene-driven pathogen response [241]. *ERF2* has been described as a positive regulator of pathogen defense and JA-dependent responses, acting downstream of ethylene signaling [241,242]. Its binding motif is a GCC box that is present in the promoter region of various defense response associated genes [241]. A cis-element analysis performed during this thesis showed that the promotor regions of 46 and 122 DEGs after 3 h and 6 h respectively contained a GCC box. Nevertheless, an *in* silico analysis revealed that *ERF2* is not differentially expressed during PTI (Figure 26). However, *ERF2* could be more ETI-specific. In conclusion, *ERF2* might be one of the key regulators of the *saul1-1* phenotype [285,329].

In comparison, other *AP2/EREBP* members, such as *ERF1*, *ERF6* and *ERF104*, were up-regulated after 3 h. These TFs are known to be expressed upon nematode or fungi infection [330,331]. *ERF6* has been reported to be essential during the defense against nematodes and its overexpression results in a dwarf phenotype [332]. ERF6/104 are targets of MPK3/6 (MITOGEN-ACTIVATED PROTEIN KINASE 3 and 6) and are phosphorylated during infection or treatment with flg22, which results in their stabilization [331]. In addition, expression of *ERF6* is induced by MPK3/6 during PTI [333]. The function of *ERF6/104* might be different in *saul1-1*, since they are known to induce the expression of defensin genes [333], which was not observed in *saul1-1* as late as 6 h.

C2H2 family

Besides *ERF2*, *ZAT7* was as well up-regulated after 1 h/2 h. This TF is a member of the C2H2 family. C2H2 TFs are zinc finger proteins that are, similar to AP2/EREBP TFs, one of the largest TF families in *A. thaliana* [334]. Their members are mostly associated with abiotic stress responses [335,336], although recent findings point towards an additional role in pathogen defense [243,337].

Overexpression of *C2H2* TFs has been reported to result in a phenotype similar to autoimmune mutants [338]. *35S*::*SRG1* (*Senescence-related Gene 1*) and *35S*::*ATDOB9* (*DUF295 Organellar B 9*) mutants both display a dwarf phenotype. In case of *35S*::*SRG1* this is accompanied by an increased expression of *PR1* [337,339]. Overexpression of *ZAT7* led to a dwarf phenotype as well [340]. Gene expression analysis could show that *WRKY70*, which is one of the key marker TF in the defense response, was up-regulated [296,340]. In addition, *in silico* analyses performed during this thesis showed that expression of *ZAT7* is induced during different stress experiments. In particular, infection with avirulent *Pst* DC3000 AvrRps4 or treatment with flg22 led to an increased gene expression with a p value of 0.0101 (Figure 26). In summary, *ZAT7* may be a key regulator of the *saul1-1* phenotype.

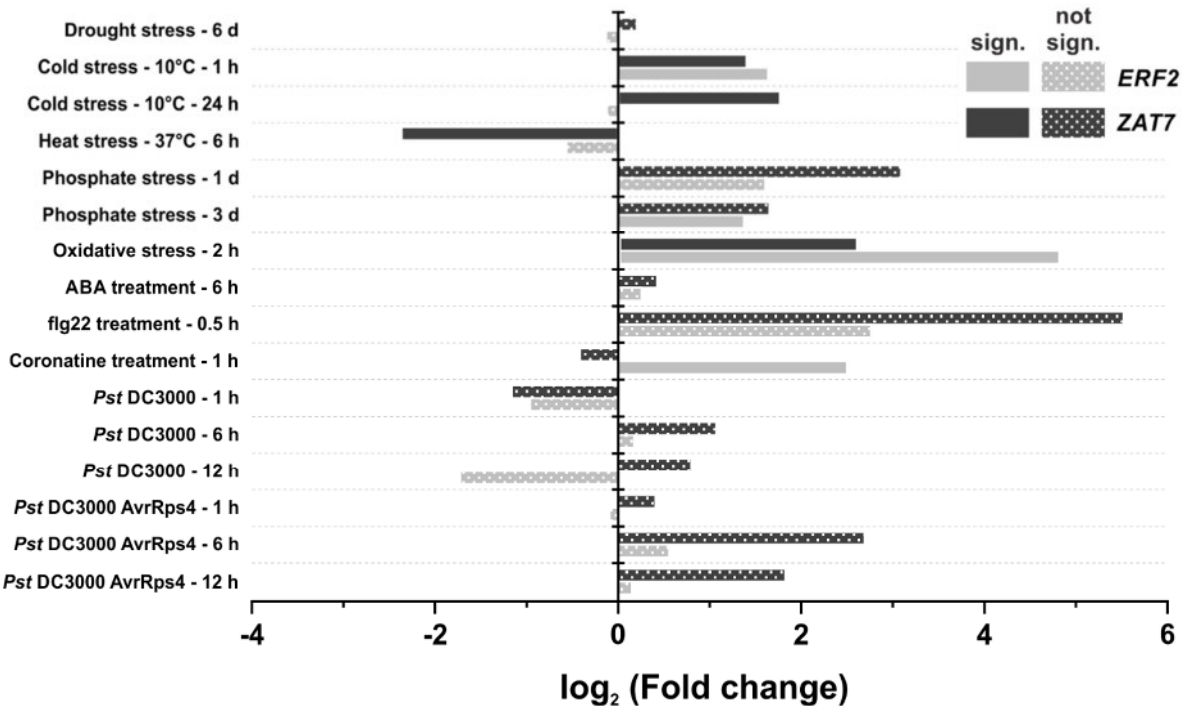


Figure 26: Database expression analysis of *ERF2* and *ZAT7*. Fold changes of *ERF2* and *ZAT7* under various stress conditions. Non-significant changes ($p > 0.01$) are depicted as dotted bars. Values were obtained using *Genevestigator*.

AZF1 (*Zinc Finger Protein 1*), which is another member of the *C2H2* family, was detected to be down-regulated. Overexpression of this TF led to a dwarf phenotype as well. *AZF1* is known to repress auxin-dependent genes. Therefore, the overexpression of *AZF1* leads to down-regulation of auxin-dependent genes [341]. These genes were down-regulated after 24 h in *saul1-1* [135].

In conclusion, this thesis is the first one that could show a connection between BR-dependent growth and ETI. All previous studies connected BR signaling only to PTI. In addition, TF analysis gave important insights into the regulations which underly the *saul1-1* phenotype. Such a strong connection between BR signaling and the dwarf phenotype, which has not been described for the ETI before, is worth mentioning and should be tested in the future. Furthermore, this analysis

revealed two potential regulators of the *SOC3*-dependent ETI phenotype of *saul1-1*, *ZAT7* and *ERF2*. These TFs could be driving factors in propagating the phenotype, since they are known to be involved in the defense response.

4.1.9 CIS-ELEMENT ANALYSIS IN *SAUL1-1*

If one or only a small number of specific TFs would be the main regulators of the *saul1-1* phenotype, a distinct DNA pattern may be enriched in the promoter regions of the DEGs. Accordingly, a cis-element analysis was performed. DNA motifs were discovered for all three tested points in time (Figure 27).

The detected motif in promoter regions of DEGs, which were regulated after 1 h/2 h had a length of 21 nucleotides. It consisted mostly of repeating adenines and guanines, with a slight prevalence to adenines (Figure 27A). In case of regulated DEGs after 6 h, a motif was detected containing a broad sequence of thymine nucleotides. Patches of two to three cytosines were detected between highly conserved thymines (Figure 27C). A similar motif was detected in the promoter regions of DEGs, which were found to be regulated after 3 h (Figure 27B). Guanines were detected with a lower frequency as well. Thus, this was determined to be strikingly similar to a motif which is associated with the TF RTV1 (*RELATED TO VERNALIZATION 1*) (Figure 27D) [342]. Although both motifs are clearly alike, a closer look into these findings revealed that RTV1 is neither regulated in *saul1-1* at any examined point of time, nor related to the defense response [343]. *RTV1* is known to be a major driver of the induction of flowering [343,344]. In conclusion, it is highly unlikely that the identified binding motif is related to the *saul1-1* phenotype. That this motif was detected nonetheless, might be caused by its quite long length and the rather unspecific sequence of thymines.

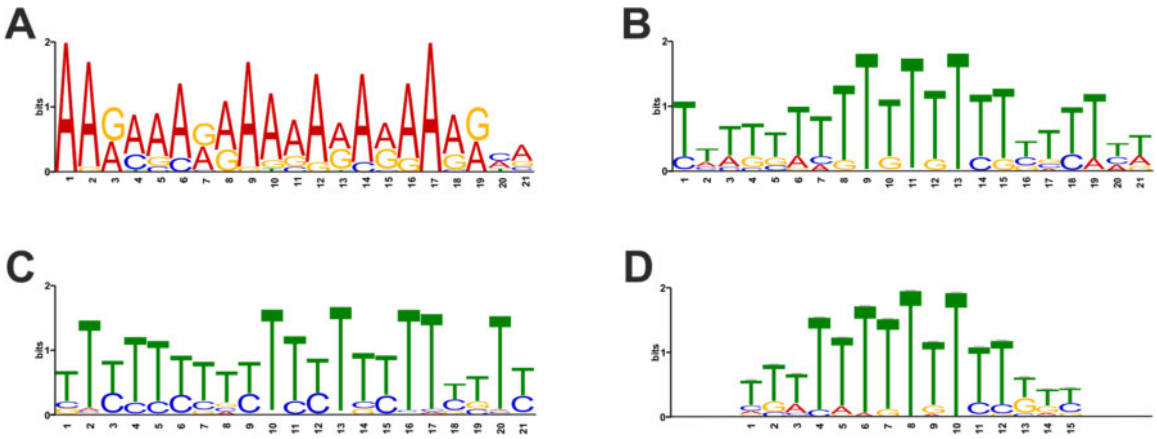


Figure 27: Cis-element analysis of early regulated genes in *saul1-1*. (A) Detected motif in the promoter region in 19 out of 19 regulated genes after 1 h/2 h with an E value of 5.9 x 10^{-6} **(B)** Detected motif with an E value of 3.7 x 10^{-18} in 407 out of 421 genes, which were differentially expressed 3 h after the temperature shift in *saul1-1.* **(C)** Motif, which was detected in the cis-region of 1104 out of 3431 differentially expressed genes after 6 h with an E value of 3.7 x 10^{-23}. **(D)** Binding motif of the TF RTV1 [342] being very similar (E value = 3.37 x 10^{-8}) to the one found in the cis-regions of genes, which were differentially expressed after (B) 3 h.

Therefore, the *saul1-1* phenotype is likely to be regulated by multiple TFs in a rather complex manner, since no binding motif of a single TF has been detected. This is consistent with previous findings that could show that ETI is regulated by multiple TFs, which are involved in a plethora of signaling pathways (see chapter 1.1.2) [4].

4.2 STRUCTURAL ANALYSES OF SAUL1

Structural properties of proteins are key characteristics of their function. The structure of SAUL1 may thus represent a key aspect to unravel its role and function in the cell. Structural analysis of a complete PUB protein is still missing. The analysis of such characteristics can be performed *in vivo* or *in vitro* experiments, respectively, both approaches having their advantages or drawbacks. The expression and purification of a recombinant protein may still be regarded as the gold standard, because vast and highly pure amounts of protein are required for most experiments in structural biology [345].

4.2.1 PURIFICATION OF SAUL1 BY USING AN AUTOMATED PIPELINE

To analyze its structural and putative binding properties, the SAUL1 protein had to be recombinantly expressed and purified. To conduct replicable experiments and guarantee an overall constant quality of purified protein, an automated purification approach using a fast protein liquid chromatography system had to be set up. Certain key features had been established by Haifa El Kilani using a manual protein purification protocol [346]. This had to be adapted to an automated purification pipeline using the *ÄKTA™* system for the affinity and size-exclusion chromatography.

For that purpose, SAUL1 was expressed as a GST-tagged protein in the pGEX-6p-1 vector. As no specific information was given in the thesis of Haifa El Kilani on the expression time after the induction of the construct with IPTG, a duration of 18 h was successfully tested and chosen to be a highly suitable time for SAUL1 expression. The analysis through SDS-PAGE revealed a distinct band at a size of approximately 110 kDa, corresponding to the GST-tagged SAUL1 (Figure 30A), which was confirmed by mass spectrometry (MS). The resuspension and lysis of the cell pellet varied strongly from the previously established protocol [346]. To increase the amount of solubilized protein, a lysozyme digestion step was added. In addition, the centrifugation was carried out at a higher speed to achieve a better separation of the soluble and insoluble fractions (see chapter 3.4.3). The lysis buffer with a pH of 9.0 was used in all steps of the purification (SAUL1 buffer).

To establish the purification of the GST-tagged SAUL1, different GST-purification columns and conditions were tested. These experiments revealed varying binding capabilities of all tested columns. In case of the *GSTrap™ FF* columns no absorbance signal could be detected during the elution (Figure 28A). SDS-PAGE analysis revealed that the protein had not bound to the matrix and was solely present in the flow-through (data not shown). *GSTrap™ 4B* columns, being the only column, whose matrix bound GST-SAUL1, were used for all purifications to follow. By testing varying flow rates during the loading of the column in the AC, differences in binding abilities were detected. In experiments with flow rates above 0.1 ml min^{-1} for loading the sample onto the column, no absorbance signal was detected during the elution (Figure 28B). Therefore, all SAUL1 purification experiments were carried out using the *GSTrap™ 4B* column at a flow rate of 0.1 ml min^{-1} during the loading process. By applying these settings, an absorbance peak in fractions 3 to 5 with an intensity of approximately 1400 mAU was detected, resembling GST-SAUL1 (Figure 28C).

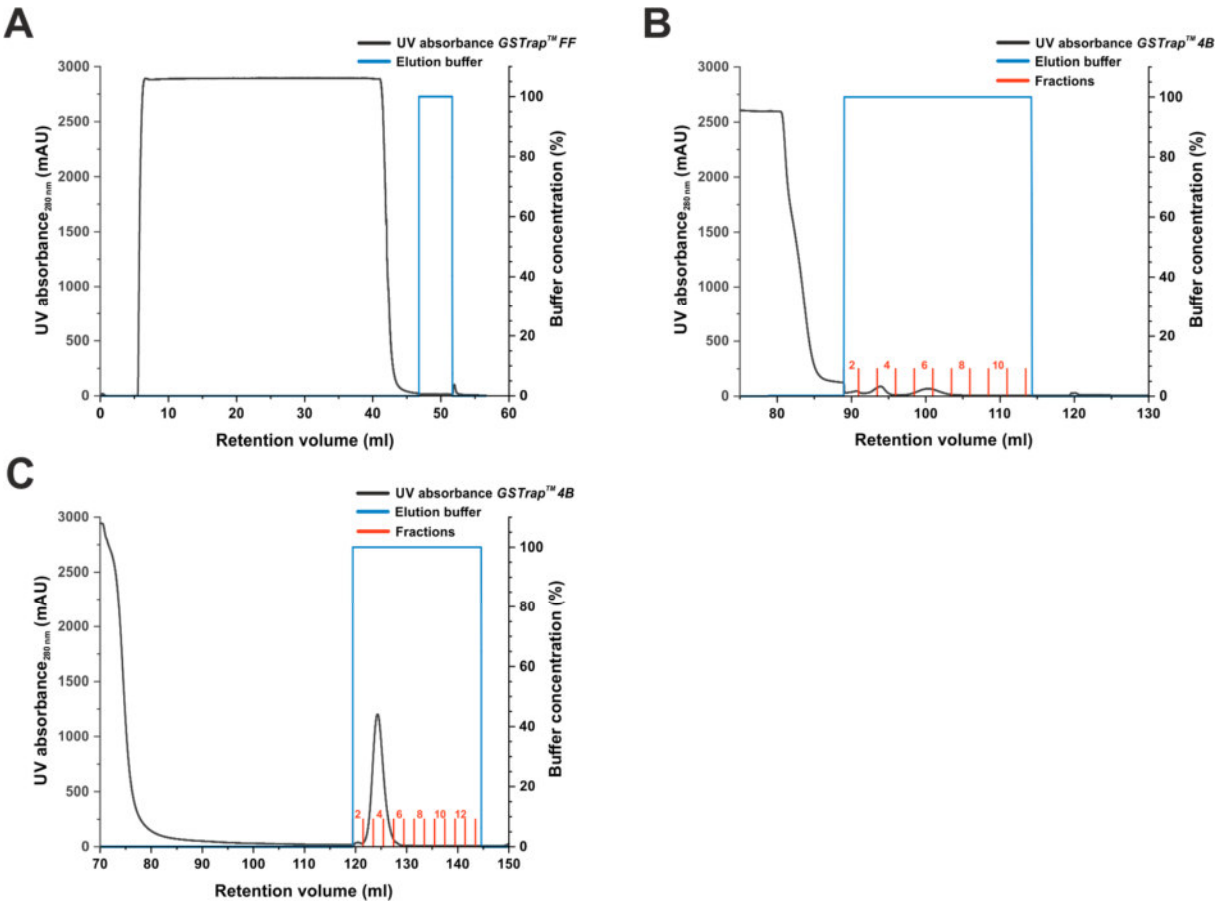


Figure 28: SAUL1 purification by affinity chromatography. Affinity chromatography of GST-SAUL1 using the *ÄKTA™ pure 25L* system and different chromatography columns and flow rates. The chromatography started with the incubation of the cell lysate on the column (absorbance higher than 2500 mAU) using a very slow flow rate. Afterwards the bound recombinant proteins were eluted using 50 mM glutathione in SAUL1buffer (blue curve). **(A)** Chromatogram of an affinity chromatography using a *GSTrap™ FF 5 ml* column. **(B)** Chromatogram of an affinity chromatography using a *GSTrap™ 4B 5 ml* column with a flow rate of 1 ml min^{-1}. **(C)** Chromatogram of an affinity chromatography using a *GSTrap™ 4B 5 ml* column with a flow rate of 0.1 ml min^{-1}.

Following the digestion of the GST-tag, SAUL1 was finally purified using the *HiLoad™ 16/600 200pg* column (Figure 29). SDS-PAGE experiments revealed that SAUL1 was present as a singular band in fractions 30-33 as a small left leaned peak with an absorption of about 60 mAU (Figure 30A). This low intensity results from the fact that SAUL1 contains only very few aromatic amino acids and has thus a low extinction coefficient. Therefore, the absorption corresponded to a protein concentration of nearly 0.2 mg ml^{-1}. The GST-tag was detected in fractions 38 to 43 with an absorption of about 250 mAU. A third peak was detected in fractions 53 to 57, which corresponded most likely to small peptides, as no signal was detected in the SDS-PAGE analysis (data not shown). Furthermore, this analysis revealed a clear separation of SAUL1 from the GST-tag resulting in 99 % pure protein (Figure 30B).

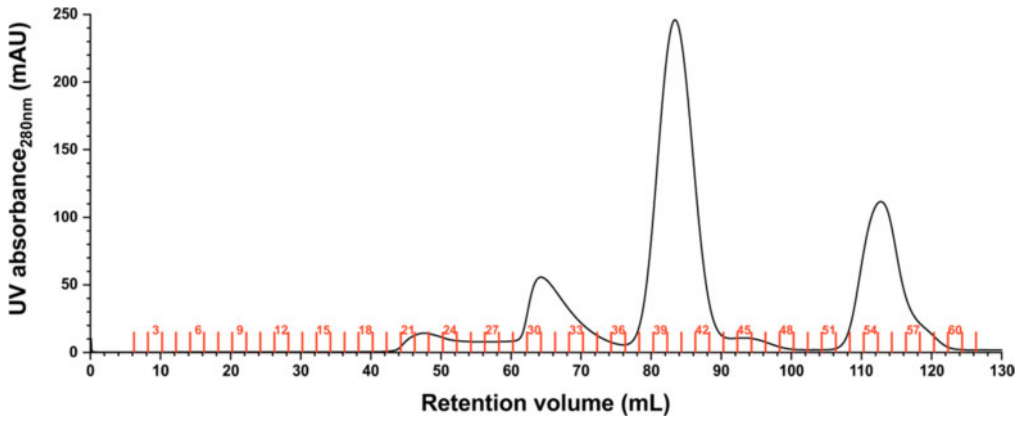


Figure 29: Separation of SAUL1 and the GST-tag using a size-exclusion chromatography. Size exclusion chromatography of digested SAUL1 using the *HiLoad™ 16/600 200pg* column and a sample volume of 5 ml. Peaks are detectable in fractions 21-23 (GST-SAUL1), 30-34 (SAUL1), 38-43 (GST-tag), 44-46 (nothing detectable by SDS-PAGE) and 53-58 (nothing detectable by SDS-PAGE).

To assess, whether the purified SAUL1 was, especially as it was purified in a rather basic buffer environment, folded in a native manner, circular dichroism (CD) spectroscopy measurements were performed. The sample was diluted to bypass high amounts of chloride ions in the SAUL1 buffer, which would otherwise interfere with the UV absorbance [205]. The resulting curve represented a secondary structure composition of 67.8 % α-helices, 14.5 % β-sheets, 6.5 % turns and 11.3 % random regions (Figure 31). In conclusion, SAUL1 appeared to be structurally ordered overall confirming the previous measurements [346]. Consequently, it was possible to purify native SAUL1 protein.

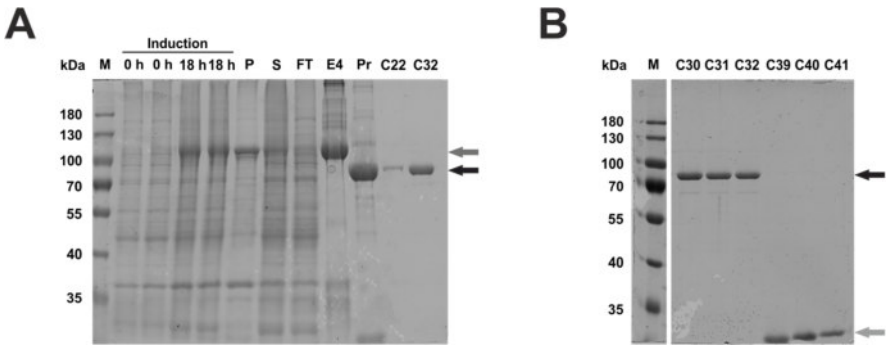
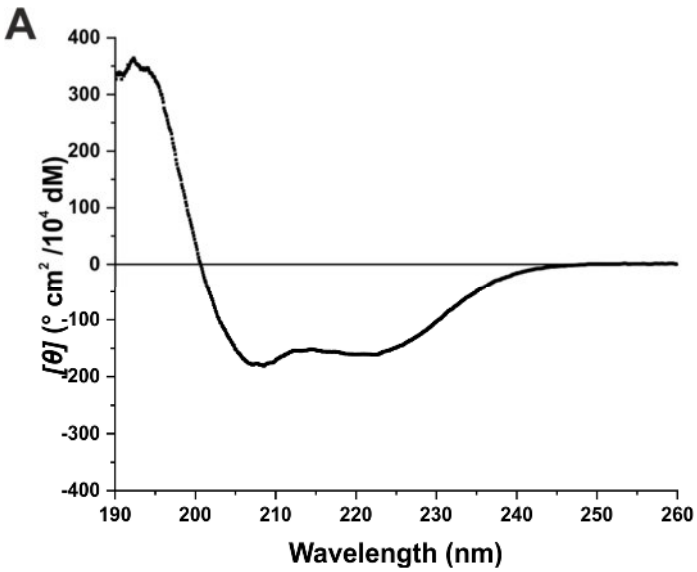


Figure 30: SDS-PAGE analysis of SAUL1 purification by affinity and size-exclusion chromatography. SDS-PAGE analysis of recombinant SAUL1 purification using a colloidal coomassie-stained 10 % SDS-PAGE gel. **(A)** Pellet (P), supernatant (S) and flow-through (FT). Elution (E4) was performed using the SAUL1 buffer containing 50 mM glutathione. GST-tag removal using the *PreScission™* protease (Pr) and size-exclusion chromatography fractions (C22, C32). **(B)** Size-exclusion chromatography fractions of the first (E30-33) and the second peak (E39-41). GST-SAUL1 has a theoretical size of 115.2 kDa (dark grey arrow), SAUL1 of 88.7 kDa (black arrow) and GST of 26.5 kDa (light grey arrow).



B

	Percentage
α-Helices	67.8 %
β-Sheets	14.5 %
Turns	6.5 %
Random	11.3 %

Figure 31: CD spectroscopy of recombinant of SAUL1. CD spectroscopy analysis of SAUL1, being diluted 1:25 in ddH_2O. **(A)** CD spectrum of SAUL1 depicted as molar ellipticity ($[\theta]$). **(B)** Secondary structure analysis of SAUL1. Analysis was done after Reed et al.[347].

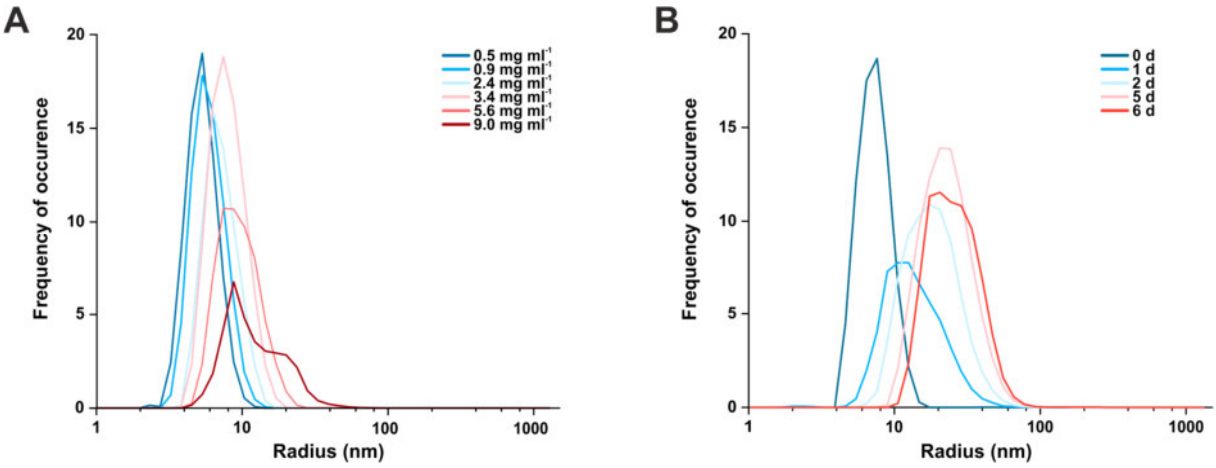


Figure 32: DLS Analysis of recombinant SAUL1. DLS measurements of purified SAUL1 in 50 mM Tris-HCl pH 9 and 250 mM NaCl. **(A)** DLS measurements on solutions of different concentration of SAUL1 from 0.5 mg ml^{-1} to 9.0 mg ml^{-1}. Detected radii shift to higher values and multiple radii with a protein concentration higher that 2.4 mg ml^{-1}. **(B)** Time-course measurements of purified SAUL1 with a concentration of 2.4 mg ml^{-1} from 0 d to 6 d. Measured radii shift to higher values already 1 d after the purification.

Experiments for the resolution of protein structures are highly sensitive towards the aggregation of the purified proteins [345]. Accordingly, the monodispersity of the protein solution was tested for quality control. Dynamic light scattering (DLS) was used to detect the radii present in the solution and thus to define the hydrodynamic radius distribution of SAUL1. The tested range of concentrations of SAUL1 solutions was from 0.5 to 9.0 mg ml^{-1}. The hydrodynamic radius increased to above 7.01±0.78 nm for concentrations higher than 2.4 mg ml^{-1} (Figure 32A). Broadly distributed radii of 9.1±1.73 nm could be measured for protein concentrations higher than 5 mg ml^{-1}. These results indicated that the more concentrated solutions were polydisperse and contained different multimers and aggregates. A monodisperse SAUL1 solution was observed with a concentration of less than or equal to 3 mg ml^{-1}. In addition to the concentration dependence, long-term stability is an important prerequisite for structural experiments. By analyzing the time-dependent stability in DLS experiments, a shift of the radius distribution as

early as 24 h after the SEC purification step was detected. The dispersion broadened from 7.01±0.78 nm to 12.37±1.89 nm (Figure 32B). This led to the conclusion, that SAUL1 is highly unstable over time. Therefore, recombinant SAUL1 should be used within 24 h to avoid problems resulting from aggregation.

4.2.2 PURIFICATION OF SAUL1 UNDER SEMI-PHYSIOLOGICAL CONDITIONS

The analysis of protein structures and of protein properties that allow for interactions with other molecules requires physiological or nearly physiological conditions [345]. Yet the established purification protocol for SAUL1 and the automated purification pipeline required a lysis buffer with a pH of 9.0. Such a high alkaline pH does not represent a physiological environment, which could be an issue regarding the structure and interaction properties of SAUL1. This might eventually cause difficulties when investigating for example the interaction between SAUL1 and BON1.

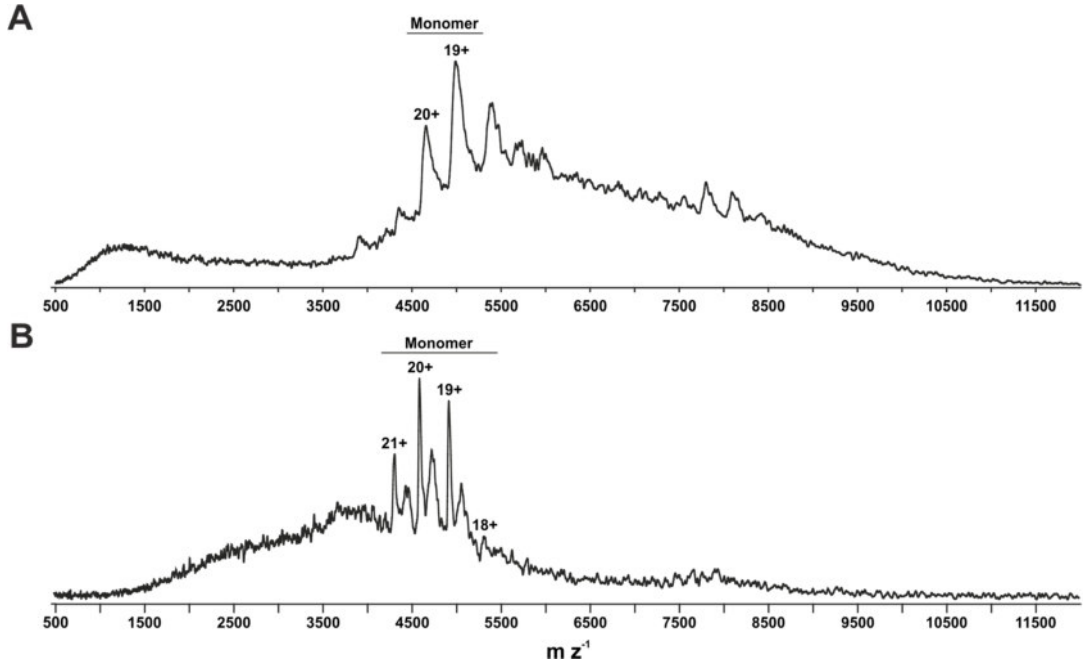


Figure 33: Native MS analysis of SAUL1. MS measurements of 11 µM SAUL1 in ammonium acetate at pH of 7.5. **(A)** SAUL1 at pH of 7.5 in 125 mM ammonium acetate. **(B)** SAUL1 at pH of 7.5 in 250 mM ammonium acetate. SAUL1 monomer with a molecular weight (MW) of 88.8 kDa was detected.

Nonetheless, SAUL1 appears to be dependent on a pH of 9.0. This has been investigated using native MS changing the pH and/or the ammonium acid concentration. In case of a pH of 9.0 and 250 mM ammonium acetate clear peaks corresponding to a SAUL1 monomer were detected (Figure 33A). In contrast, at a more physiological pH of 7.5 and a lower ionic strength of 125 mM ammonium acetate nearly no ions corresponding to SAUL1 were detectable. Thus, SAUL1 appeared to be extremely unstable at a pH of 7.5, potentially being unfolded, as it seemed to be hardly ionizable. An improvement was achieved by changing the concentration to 250 mM ammonium acetate stabilizing SAUL1, as more ions matching to the SAUL1 monomer were found (Figure 33B). Nevertheless, the detected ions were less defined, as SAUL1 still appeared to be very unstable under these conditions. Thus, SAUL1 appeared to be rather pH-dependent, at least under the chosen conditions.

To find a more suitable purification buffer for SAUL1, a thermofluor assay with the purified protein was performed. In this assay 96 different buffers [199] were compared in three independent

experiments. The T_m values, which were obtained from the slope (Figure 34A) of a fitted *Boltzmann* equation (see chapter 3.5.9) with a R^2 value higher than 0.95, were used to calculate the thermal shift (ΔT_m) of each condition. The T_m can be determined as well from the apex of the first derivative (Figure 34B).

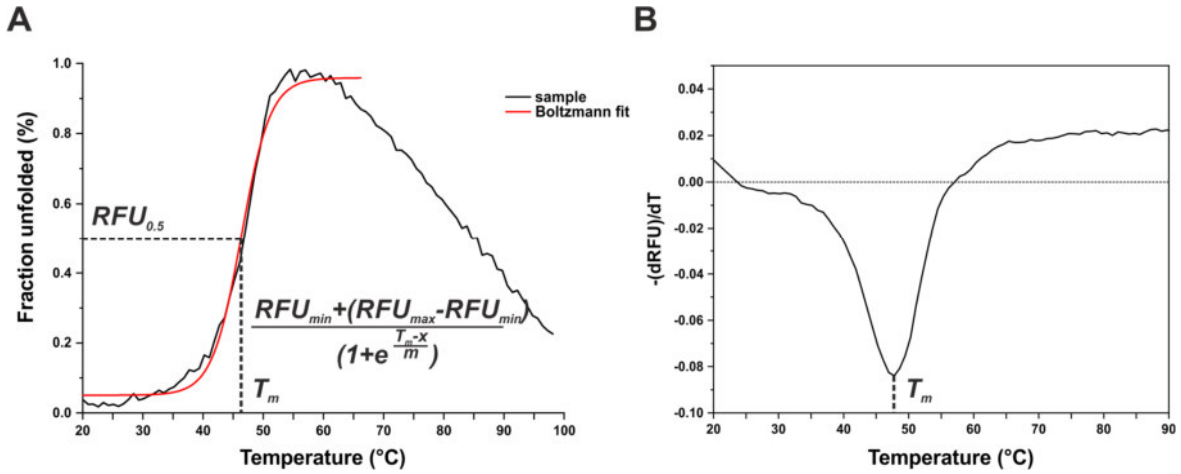


Figure 34: Typical melting curve analyzes of a protein using the thermofluor assay. (A) Melting curve of SAUL1 in 50 mM Tris-HCl pH 9.0 and 250 mM NaCl. To calculate the melting temperature (T_m) a Boltzmann equation [200] (red curve) was fitted to the melting curve. The T_m corresponds to the temperature at which 50% the protein is unfolded ($RFU_{0.5}$). **(B)** The T_m can as well be determined using the first derivative -dRFU of the melting curve.

As a broad spectrum of conditions was tested in the thermofluor assay (Figure 37), some buffers lead to a non-determinable T_m. In case of citric acid with pH 4.0 this was due to a high initial fluorescence (Figure 35A). This corresponded to aggregated protein being present in the solution. In addition, multiple apices were detected in case of KH_2PO_4 pH 7 (Figure 35B). A multimeric structure of SAUL1 could result in such a curve with multiple maxima.

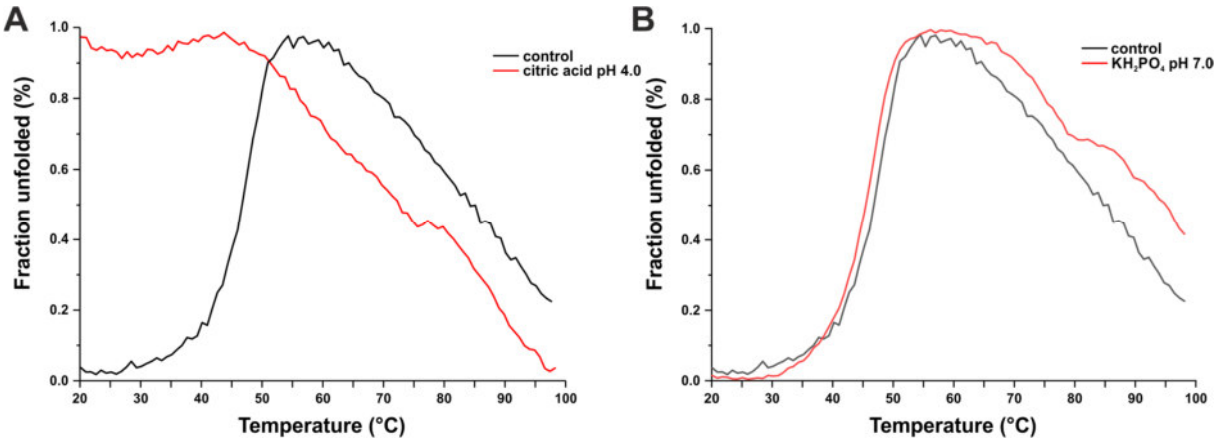


Figure 35: Melting curve comparison of different conditions for SAUL1. (A) Melting curve of SAUL1 in 100 mM citric acid pH 4 shows a high fluorescence background between 20 °C and 45 °C. **(B)** Melting curve of SAUL1 in 100 mM KH_2PO_4 pH 7 shows two maxima, which result out of two overlapping curves.

To resolve the stabilizing effect of all buffers tested, the thermal shift (ΔT_m) was determined in relation to the control sample. A high standard deviation, due to high versatility between the three independent experiments, lead in some cases to non-conclusive thermal shift values.

Nevertheless, most buffer environments lead to a decrease in T_m (Figure 37) and therefore to a destabilization of the recombinant SAUL1. Very alkaline pH levels above pH 8.2 (Figure 36A-B) lead to an increase of the T_m to 0.82 °C and therefore to a stabilization of SAUL1. Thus, a more acidic pH decreased the T_m by about 8.13 °C. Salt concentrations higher than 500 mM NaCl (Figure 36C-D) increased the T_m up to 2.92 °C, which corresponds to a stabilization of the protein. However, these buffer conditions were neglected because of their incompatibility with structural analysis methods, as discussed before.

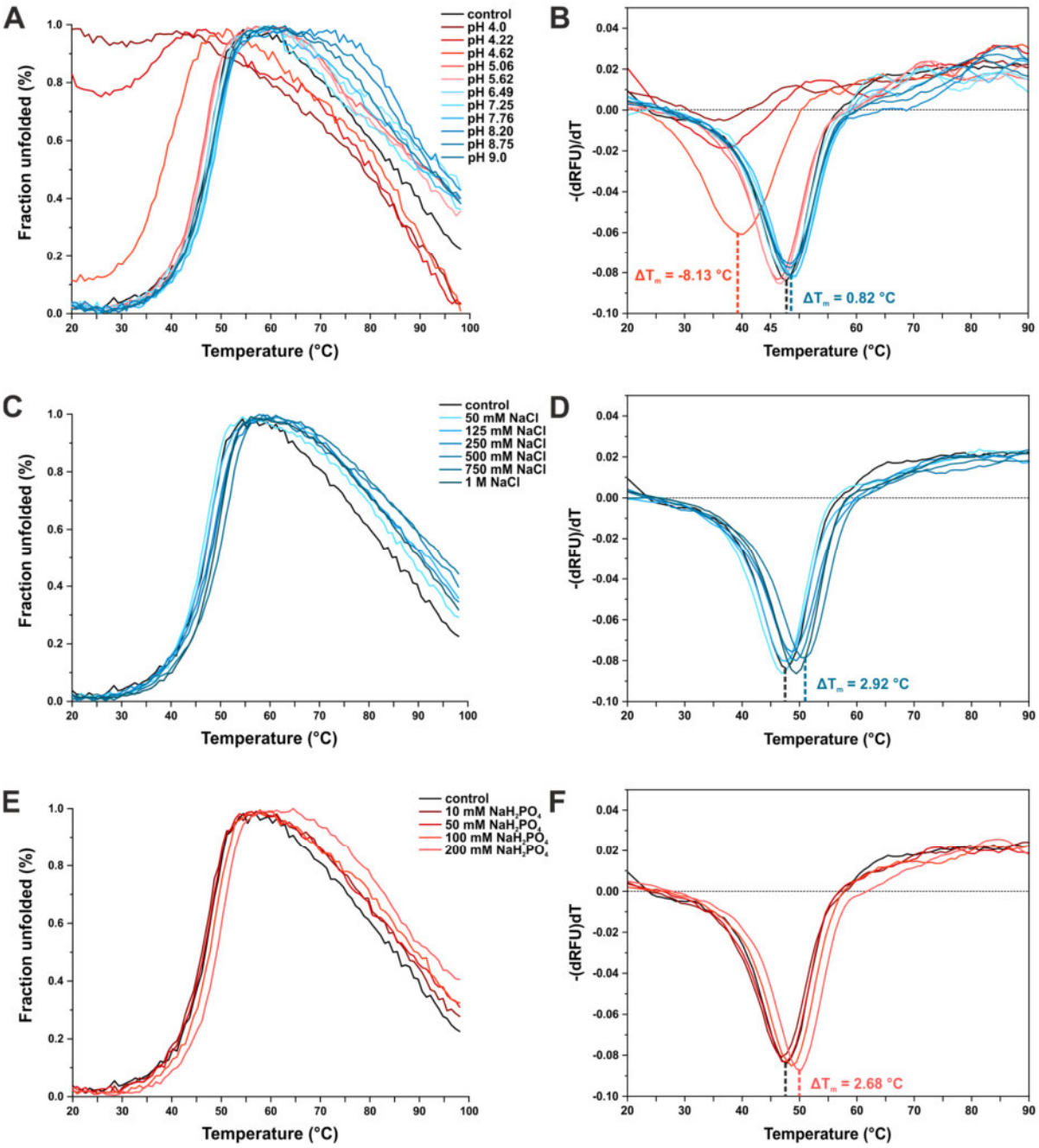


Figure 36: Thermal shift (ΔT_m) comparison of SAUL1 in different buffer conditions. (A) Melting curves of SAUL1 at different pH values. **(B)** First derivative of the melting curves with the thermal shift values (ΔT_m) in ratio to the control of pH 4.62 (-8.13 °C) and pH 9.0 (0.82 °C). **(C)** Melting curves of SAUL1 with different NaCl concentrations. **(D)** First derivative of the melting curves with the ΔT_m of 750 mM NaCl (2.93 °C). **(E)** Melting curves of SAUL1 in NaH_2PO_4 buffer with different molar concentrations. **(F)** First derivative of the melting curves with the highlighted ΔT_m of 200 mM NaH_2PO_4 (2.68 °C).

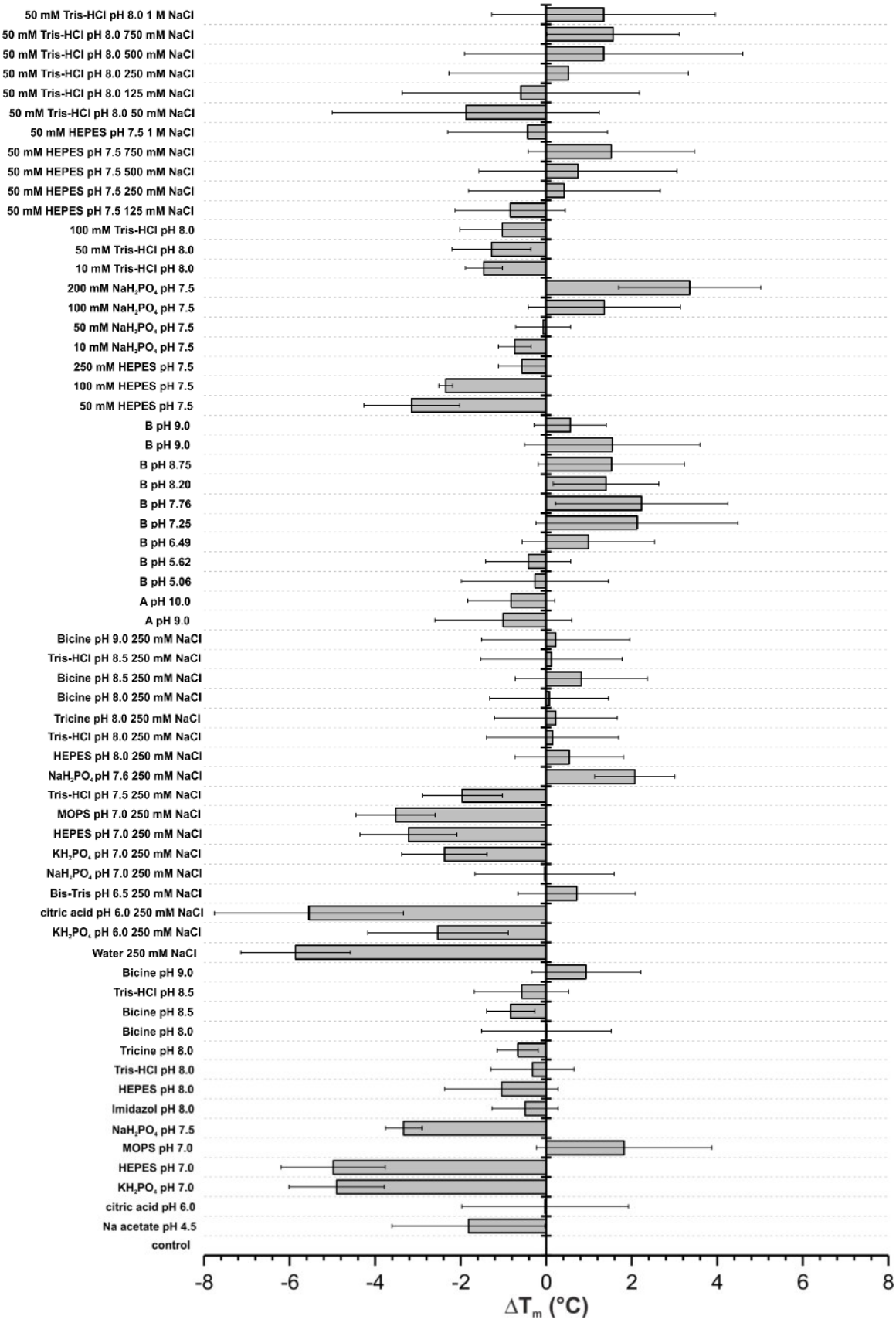


Figure 37: Thermal shift (ΔT_m) comparison of SAUL1 in different buffer conditions. The thermal shift is depicted as the median of three independent experiments. T_m values, which were obtained of a fitted *Boltzmann* curve (see chapter 3.5.9) with an R^2 smaller than 0.95 were neglected. All values are in proportion to the control condition (water) and error bars depict the standard deviation. Unless stated otherwise, buffers were used at a concentration of 100 mM.

An increase of the T_m was observed as well in conditions in buffers containing 100 mM NaH_2PO_4 or more (Figure 36E-F). An increase was also observed in samples containing in addition 250 mM NaCl (Figure 37). The buffer of choice for further purification experiments consisted of 100 mM NaH_2PO_4 and 250 mM NaCl.

Initially, a purification with SAUL1 following the established protocol (see chapter 4.2.1) with the identified buffer was carried out. In the next step, fractions of the purification steps were analyzed using SDS-PAGE. Bands corresponding to the size of the GST-tagged SAUL1 and tag-free SAUL1 were detectable in the pellet, in the supernatant, in the fractions following affinity elution and the fractions following size exclusion of the purification (Figure 38C). The corresponding band in the pellet fraction is clearly thinner, than the one observed in the supernatant. Therefore, SAUL1 seems to be highly soluble in the selected buffer. Following all purification steps, a peak with a height of 14.6 mAU was detected in the chromatogram of the SEC corresponding to SAUL1 in fraction 32 (Figure 38B). The SDS-PAGE analysis of this peak showed a singular defined band, which corresponded to the tag-free SAUL1 (Figure 38C).

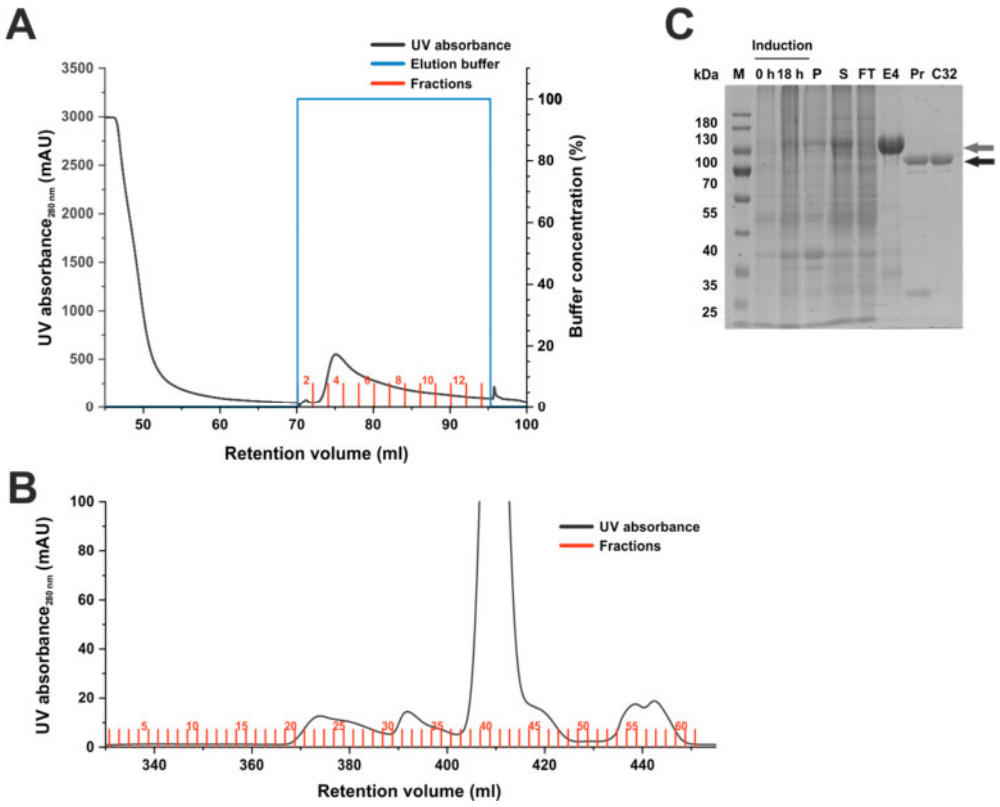


Figure 38: Purification of SAUL1 in semi-physiological buffer conditions. (A) Affinity chromatography of GST-SAUL1 using the *ÄKTA™ pure 25L* system. Chromatography starts with the incubation of the cell lysate on the column (absorbance higher than 2500 mAU) using a flow rate of 0.1 ml min^{-1}. Afterwards the bound recombinant proteins were eluted using 50 mM glutathione in 100 mM NaH_2PO_4 with a pH of 7.6 and 250 mM NaCl (blue curve). **(B)** Size exclusion chromatography of digested SAUL1 using the *HiLoad™ 16/600 200pg* column and a sample volume of 5 ml. **(C)** Analysis of the SAUL1 purifications using a colloidal coomassie-stained 10 % SDS-PAGE gels. Pellet (P), supernatant (S), flow-through (FT), elution (E4), GST-tag removal using the *PreScission™* protease (Pr) and size-exclusion chromatography fraction (C32). GST-SAUL1 has a theoretical size of 115.2 kDa (dark grey arrow), SAUL1 of 88.8 kDa (black arrow).

Consecutively, the time- and concentration-dependent stability of the purified protein was analyzed using DLS. In these experiments, a monodisperse solution of SAUL1 was detectable at concentrations less or equal than 2.8 mg ml^{-1} with a hydrodynamic radius between 5 and 8 nm. The time-course experiment revealed a shift of the hydrodynamic radius and the occurrence of multiple higher radii after 24 h (Figure 39). Thus, SAUL1 being purified in a buffer with a pH of

7.6 was stable for 24 h up to a concentration of 2.8 mg ml^{-1}. These results were comparable to the ones obtained for SAUL1 being purified at a pH of 9.0. In these experiments SAUL1 was stable up to 2.4 mg ml^{-1} for 24 h (Figure 32). Although no improvement due to the new buffer conditions were detected, it was possible to purify SAUL1 using buffer conditions with almost physiological pH values.

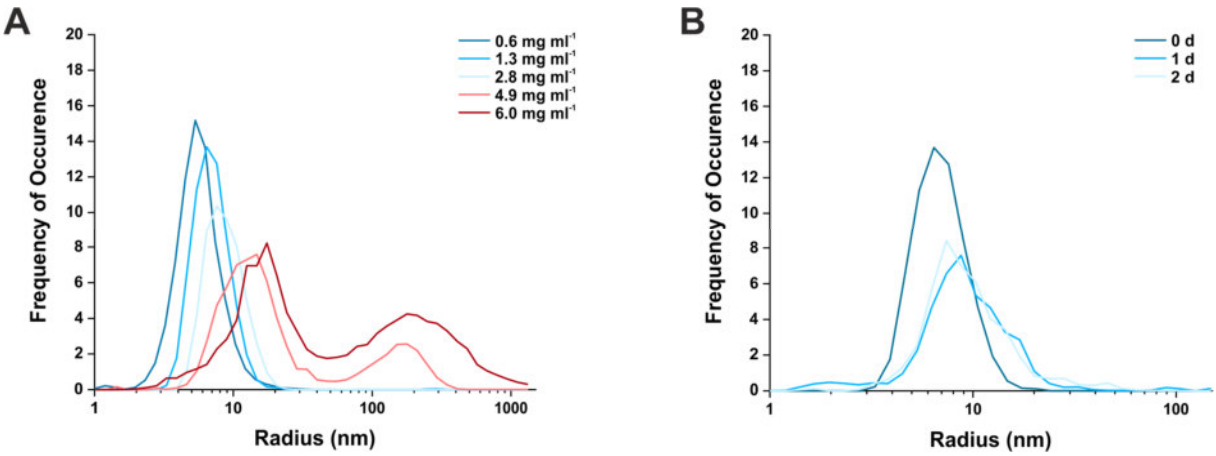


Figure 39: DLS measurements of SAUL1 being purified in a semi-physiological buffer. DLS measurements of SAUL1 purified in 100 mM NaH_2PO_4 pH 7.6 and 250 mM NaCl. **(A)** Measurements of different concentrations of SAUL1 from 0.6 mg ml^{-1} to 6.0 mg ml^{-1}. **(B)** Time-course measurements of purified SAUL1 with a concentration of 2.8 mg ml^{-1} over multiple days.

Despite these very promising results, it was not possible to use this buffer for further experiments for studying the SAUL1 structure or potential protein-protein interactions. As SAUL1 was observed to have a high tendency to form aggregates and to preserve its structural integrity, all purification steps were carried out at 4 °C. At this temperature the selected buffer was observed to crystalize in a concentration-dependent manner. This was very problematic, as crystals actually formed in the superloop, the column and the tubing used for automated affinity chromatography with the *ÄKTA™* system being placed in a refrigerator at 4 °C. Those crystals being formed resulted in an immediate increase in the internal pressure of the system, exceeding instrumental limits. Even at temperatures around 20 °C crystallization was observed over time. Further investigations determined, that the actual glass labware was the cause, as no crystallization was detected in new labware being made out of plastic. Thus, the cleaning procedure of the labware and maybe the used water might be the origin of the crystallization. Even small contaminations could have provided a seed for those crystals to be formed. Nonetheless, as large quantities of buffer were needed for the purification and thus glass labware had to be used, no feasible concentration of NaH_2PO_4 was found, which allowed for a potential purification of SAUL1, without the formation of crystals. Consequently, it was not possible to investigate SAUL1 under semi-physiological conditions.

4.2.3 STRUCTURAL ANALYSES OF SAUL1 AS A MONOMER

However, SAUL1 was analyzed using the pre-determined buffer with a pH of 9.0. Different methods can be applied to assess the structure of a protein. The gold standard here is X-ray crystallography. This allows for a structural resolution up to ≤0.7 Å [348]. Though, the most severe bottleneck in crystallography is always the generation of crystals and for some proteins it has never been possible to create these. Previous experiments with SAUL1 did not result in feasible crystals [346]. Thus, other methods for structural analysis had to be implemented. As it was possible to obtain a monodisperse solution of SAUL1, small-angle X-ray scattering (SAXS) was the method

of choice. Although only lower resolution data with a resolution up to 0.5 nm are collected [207], movement and flexibility can be taken into account [349]. Therefore, SAXS is an ideal technique to obtain structural information and accounting for movement and changes in formation of the protein of interest in solution.

It was known from earlier SAXS batch experiments, that SAUL1 is present in a concentration-dependent polydisperse system, with monomers, dimers and tetramers [346,350]. This was considered to be quite a hinderance, as the goal was to retrieve a monomeric structure with a higher resolution, which includes potential flexibility. To overcome this, in addition to normal batch measurements, all different oligomers would have to be measured separately. To achieve this goal, inline size-exclusion chromatography SAXS (SEC-SAXS) was used as an addition to the conventional batch method. This allowed a pre-separation of all molecules based on their size using an inline SEC separation with a continuous SAXS measurement [351,352].

For that purpose, SAUL1 was with respect to its instability (Figure 32B) purified on the same day of SEC-SAXS experiments. A SAUL1 solution with a concentration of 4.38 mg ml^{-1} was separated on a Superose™ 6 Increase 10/300 GL column. During the inline SEC one signal with an intensity of app. 300 AU (Figure 40A) and a corresponding absorbance of app. 25 mAU (Figure 40C) was detected. This resembled a protein concentration of 0.126 mg ml^{-1} (Table S6A). The identified signal was thought to consist of two overlapping peaks, as two potential maxima were detected (Figure 40B, arrows). In addition, DLS measurements of the SAUL1 solution displayed a slightly broadened radius distribution (Figure 40D). Consequently, 25 signal frames of the left part of the detected peak were used for subsequent data processing (Table S6A). Data analysis of the right shoulder of the signal revealed a scattering curve with a very low signal-to-noise ratio (Figure 41A). The corresponding dimensionless Kratky plot depicted a kind-of plateau, but was, at higher angles, not interpretable (Figure 41B). Due to the high noise, this second peak was not analyzed any further, as no reliable results could have been drawn from these scattering data. Thus, SAUL1 was present in two states, which are detected during the SEC, as these overlapping peaks, although only one was analyzed. The first state is the more prominent one with the higher absorbance. It can be hypothesized, that this is the monomer, as the previous experiments found this state to be the dominant one at low concentrations and especially as a very low concentration was measured [350].

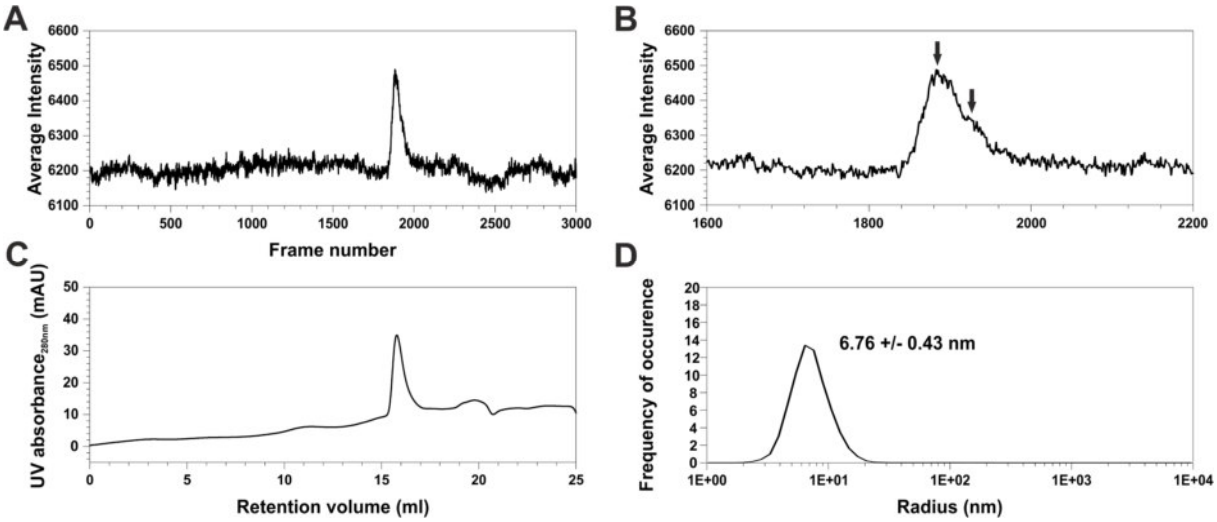


Figure 40: Chromatogram and radius distribution of SAUL1 in the SEC-SAXS experiment. (A) Scattering intensities of monodisperse SAUL1 with a loading concentration of 4.38 mg ml^{-1}. **(B)** Section of scattering intensities with a closer look on the SAUL1 peak highlighting (arrows) two potential maxima. **(C)** Chromatogram of the UV absorption at 280 nm. **(D)** DLS measurements of SAUL1 exhibiting one singular peak with a detected radius of 6.76 ± 0.43 nm.

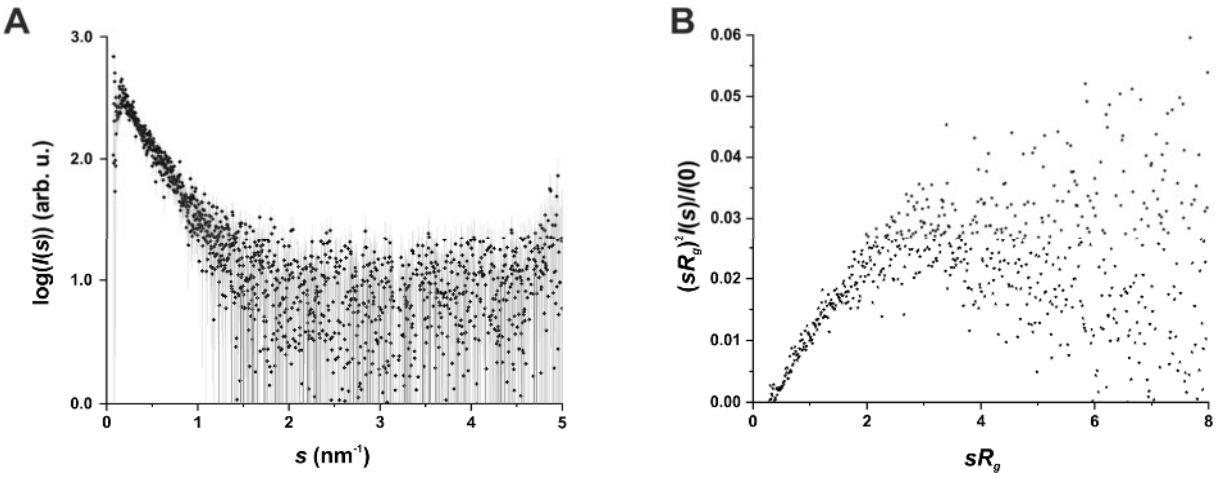


Figure 41: Scattering data of the second peak detected during the SEC-SAXS measurements. (A) Two-dimensional Log-linear plot of $I(s)$ versus s. **(B)** Dimensionless Kratky plot with the intensities normalized to the forward scattering intensity ($I(0)$) and the radius of gyration (R_g).

The log-linear plot of the scattering data of the first peak, which was thought to contain the SAUL1 monomer, showed a smooth decrease towards higher angles. This normally corresponds to a partially unfolded protein. As the decrease was not sigmoidal shaped, SAUL1 did not seem to be of globular shape, but maybe elongated and flexible (Figure 42A). In comparison to the batch measurements the detected intensities were significantly lower as well than the signal-to-noise ratio. The R_g based on the Guinier approximation was calculated to 4.64 ± 0.54 nm. Although the Guinier fit is in a small range (s_{min} = 0.005 nm^{-1} and sR_g = 1.3), the standardized residuals are mostly below 2 (Figure 42B). Therefore, the fit can be considered to be representative.

To assess the overall shape, the dimensionless Kratky plot was used. The intensities increased up to a maximum around 3.2. Towards higher angles the detected noise increased severely. Nonetheless, a tendency towards an intensity decrease in regions of higher angles was detected. Therefore, the curve was considered to be slightly bell-shaped (Figure 42C). As the peak was shifted towards higher angles and not around $\sqrt{3}$, as for globular proteins, and the curve seemed to decrease as well, SAUL1 appeared to be present as an elongated and flexible protein [349].

By calculating the P(r) function, a non-Gaussian-shaped curve was detected which had its peak around 3.15 nm. The frequency of radii higher than 6 nm decreased up to a d_{max} of 18.4 nm with a low slope. The fit of the distribution to the experimentally obtained data had an estimated quality of 0.77 and most standardized residuals were within a range of 2 (Figure 43 and Table S6B). Such a distribution would correlate with a rod-shaped elongated protein.

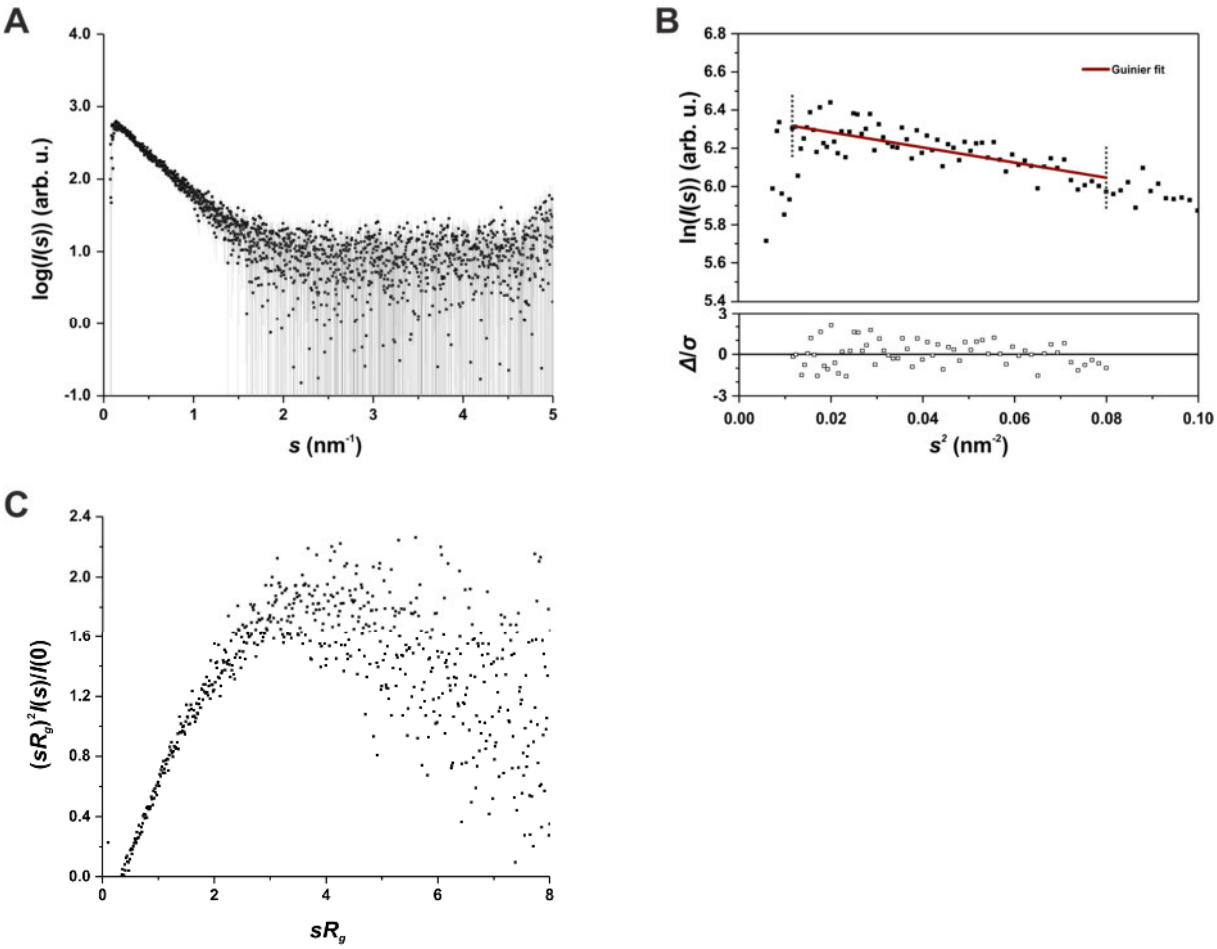


Figure 42: Scattering data of the first peak detected during the SEC-SAXS measurements of SAUL1. **(A)** Log-linear plot of $I(s)$ versus s. **(B)** Upper plot shows the Guinier fit (red) of the SAXS data. Dotted lines mark the fit range (s_{min} = 0.005 nm^{-1} and sR_g max = 1.3). Lower plot shows the standardized residual plot. **(C)** Dimensionless Kratky plot with the intensities normalized to the forward scattering intensity ($I(0)$) and the radius of gyration (R_g).

The dimensionless Kratky plot had a rather low signal-to-noise ratio. It was therefore not completely clear, whether SAUL1 was present in a folded state or not. Although, previous CD experiments pointed towards a structurally organized conformation of SAUL1, this had to be further investigated. The P(r)-derived R_g was compared to a theoretical one, using Flory's equation [206]. The P(r)-derived R_g for SAUL1 in this SEC-SAXS experiment was 5.14 nm (Table S6B). For chemically denatured proteins R_0 is 1.927 ± 0.27 and ν has a value of 0.598 ± 0.028 [353]. Therefore, SAUL1 would have, as a denatured protein, an app. R_g of 10.5 nm. For intrinsically disordered proteins (IDPs), for which R_0 is 2.54 ± 0.01 and ν is smaller with 0.522 ± 0.01 [354], SAUL1 would have a theoretical R_g of app. 8.35 nm. These calculated R_g values are substantial larger, than the derived one with 5.14 nm (Table S6B). Therefore, SAUL1 is present as an elongated protein, which is partially, but not completely intrinsically disordered and quite flexible.

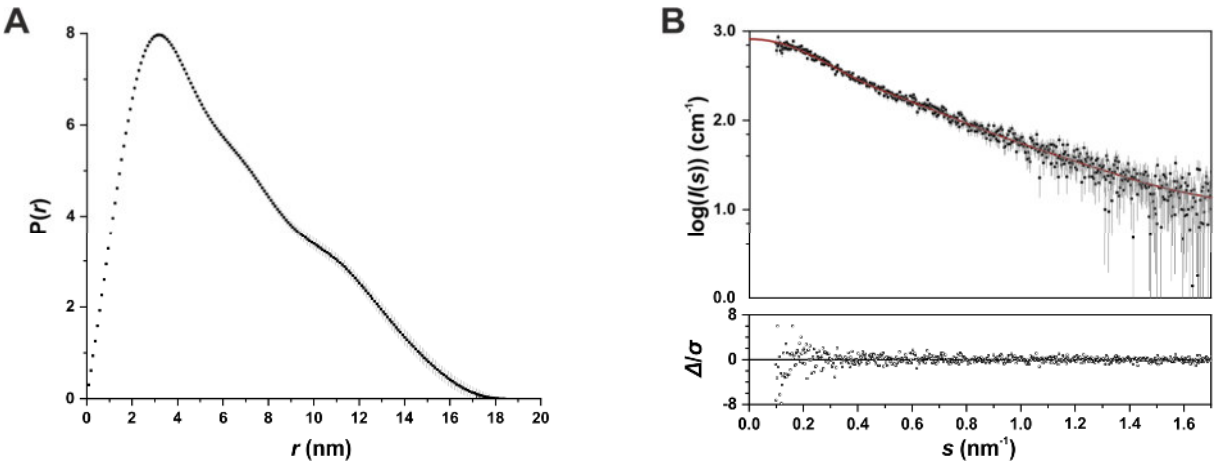


Figure 43: Radius distributions of SAUL1 in the SEC-SAXS measurement. (A) P(*r*) versus *r* profile of SAUL1. **(B)** Fit of the P(*r*) function (blue) to the experimental SAXS data. Lower plot depicts the standardized residual plot.

To further analyze the structure of this potential monomeric SAUL1, an *ab initio* model of SAUL1 was generated using *GASBOR*. A chain-like dummy residue model representing the overall shape was generated 20 times. In this model each bead represents a C-α atom. The final model was selected based in the normalized spatial discrepancy (NSD), which indicates which model is the most likely one, as it differs the least from all 20. Thus, a SAUL1 model with an NSD value of 1.498 was chosen (Figure 44A). To compare the *ab initio* model with the scattering data a theoretical scattering curve of this structure was calculated using *CRYSOL* (Figure 44B, black line) and compared to the original data. This comparison resulted in a low χ^2 value of 1.124, stating that the *ab initio* model represents the experimental data very well, as a perfect χ^2 value would be 1. With regard to the overall shape, SAUL1 is present as an elongated rod-like structure, which is slightly bended at one end. The whole rod has a mostly even diameter (Figure 44A). This is in high accordance to all previous calculations, like the distance distribution (Figure 43A).

In silico modeling was applied for further elucidation of the structural organization of SAUL1. In a first step a homology-based model using *I-TASSER* was generated. This model was fitted against the experimental data. By using *SREFLEX* to improve the fit of the model to the measured data, a χ^2 value of 1.190 was determined. This only partially corresponded to the calculated bead model and did not account for any flexibility (data not shown), which was observed in the experimental data (Figure 42C).

Thus, the ensemble optimization method (*EOM*) was used, which allows to take into account flexibility and therefore multiple dynamic states. *EOM* calculations with a total flexible SAUL1 lead to extremely elongated structures, which did not conform to the experimental data with a χ^2 value of 1.629 and a very high d_{max} of 380.17 nm (data not shown). As SAUL1 was characterized to be only partially flexible, some domains of the *in silico* structure had to be defined as structurally rigid. Thus, by using *I-TASSER* and *InterPro*, fixed domains were defined according to flexibility and domain predictions. As very high *B*-factor values were predicted, the N- and C-terminus were considered to be flexible (Figure 45A). Based on the *InterPro*-analysis, which predicted one U-box and eight potential ARM repeat domains (Figure 45C), rigid regions were defined. In general, rigid regions are connected via flexible linkers/loops [355], which would have higher *B*-Factor values. Thus, larger regions, based on the *InterPro* predictions, were determined, since these had lower *B*-Factor values. In contrast, some smaller regions were specified as linkers, based on a high *B*-Factor value (Figure 45B, Table S2A). Using this divided *I-TASSER* model, *EOM* calculations were carried out. These resulted in an *in silico* model, which contained five different states. Quality assessment using *CRYSOL* resulted in a χ^2 value of 1.045, which fitted very well to the experimental data (Figure 44D).

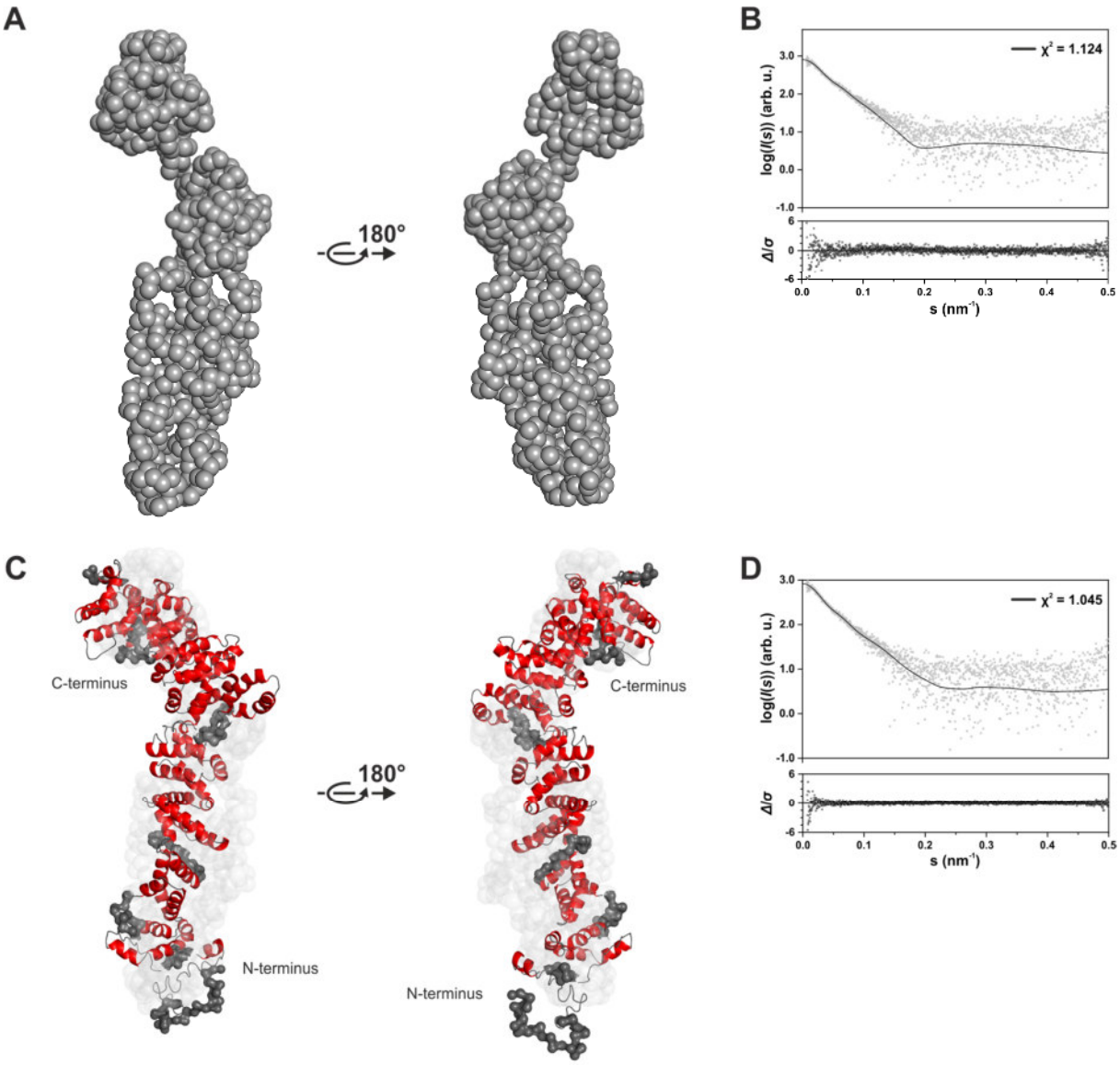


Figure 44: *Ab initio* and *in silico* structures of SAUL1 derived from SEC-SAXS measurements. (A) *Ab initio* bead model of SAUL1 with a resolution of 45 ± 3 Å and an NSD of 1.498. **(B)** Fit of the *ab initio* model to the raw scattering data with the standardized residual plot $\Delta/\sigma = (I_{exp}(s) - I_{mod}(s)) / \sigma(s)$. **(C)** *I-TASSER* derived *in silico* model, which was subdivided into different domains upon an *InterPro* analysis and refined using *EOM*. Depicted are α-helices (red), β-sheets (blue), loops (grey) and flexible regions (grey) **(D)** Fit of the *in silico* model to the raw scattering data with the standardized residual plot.

The *in silico* structure of the most abundant model consisted solely of α-helices and loops and fitted quite well to the rod-like *ab initio* model. Interestingly, the *in silico* model contained the same curvature, which was observed for the *ab initio* model. Around the U-box at the N-terminus the overall shape of the model was slightly expanded. As for the rest of the model most regions were quite condensed and α-helices formed two compact regions (Figure 44C). It is noteworthy that in particular the ARM repeat regions 3, 4 and 5, which were used for the *EOM* analysis, formed a rather compact cluster, although containing three flexible linkers (Figure 45B).

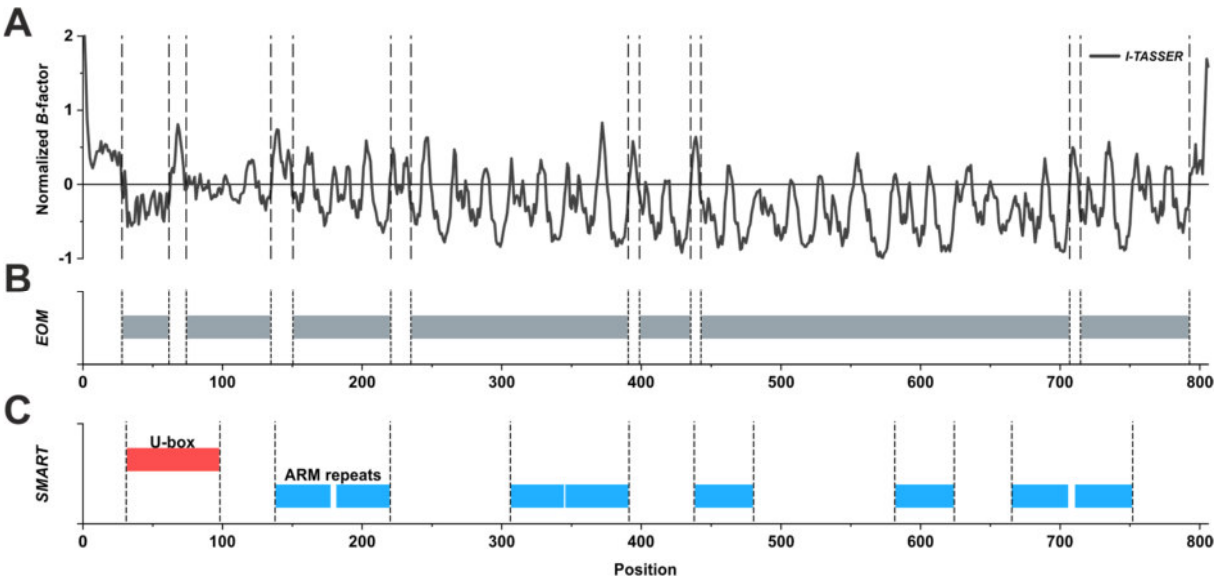


Figure 45: Flexibility and domain analysis of SAUL1. (A) The flexibility of SAUL1 was analyzed using *I-TASSER*. Depicted is the normalized B-factor. Values above zero indicate flexible residues, as values below zero suggest a more rigid structure. An *InterPro* analysis allowed for protein domain predictions using **(C)** *SMART* domain databases **(B)** Based on these and the flexibility values rigid regions were defined for the *EOM* analysis (Table S2A).

Five different models were calculated during the EOM analysis, corresponding to the experimental data. All in all, they were very similar to each other (Figure 46A). The models differed slightly in the orientation of the last ARM repeat containing regions and their compactness. Particular model 4 was more condensed. However, in all models the ARM repeat containing regions 3, 4 and 5 were quite compact and formed one part of the bend-divided rod-like structure (Figure 46A). As for the frequency of the *EOM* models, model 1 was the most abundant one with a percentage of 57 %. All other models were only present with a percentage of 11 % (Figure 46C). Therefore, SAUL1 seems to be present as an evenly formed rod-like structure, which consists mostly of α-helices.

This is in accordance with the SAXS data, which was previously obtained by Kilani and Kikhney. In these experiments batch SAXS measurements with SAUL1 at different concentrations were performed. These resulted in a polydisperse system containing a mixture of concentration-dependent multimers with a proportion of 57 % monomers at the lowest concentration [350]. The *I-TASSER* based monomer, which was fitted to the experimental data, was quite similar to the one obtained from the SEC-SAXS experiments in this thesis (Figure 47). However, the model was a bit more compact, which might have resulted from the fit of the *in silico* model to a polydisperse system and not accounting for predicted flexibility, as for the SEC-SAXS-derived model.

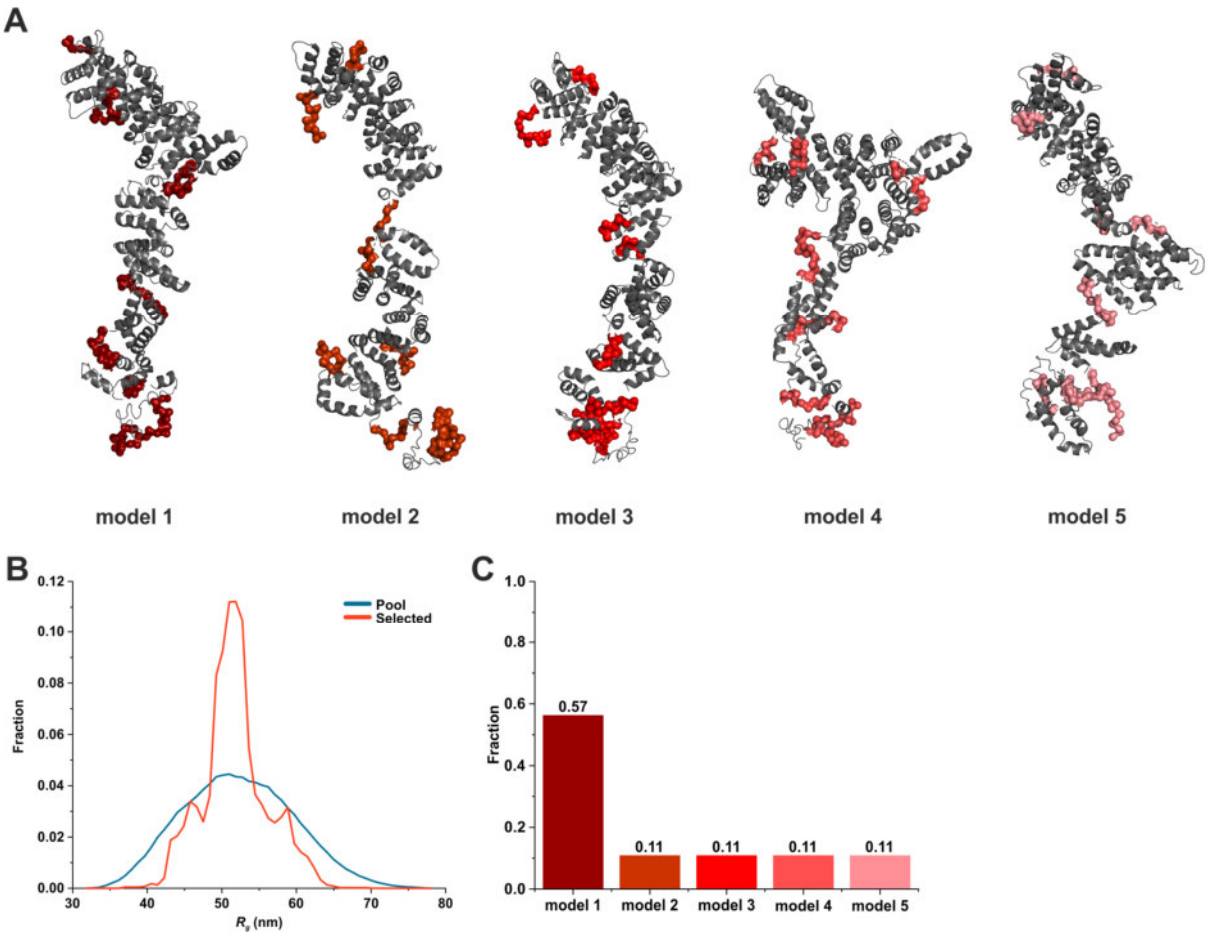


Figure 46: *EOM* analysis of SAUL1 in SEC-SAXS measurements. (A) Five different SAUL1 models were obtained by fitting an *in silico I-TASSER* model to the experimental SEC-SAXS data using *EOM*. **(B)** Fraction analysis of the R_g values of these models which were generated at the start of the *EOM* calculation (Pool) and those which fit best to the experimental data (Selected) **(C)** Fraction analysis of all five models.

When taking a closer look at the overall structure of this model, most predicted domains appeared to be correctly folded. Interestingly, although only eight ARM repeats were predicted (Figure 45C), more ARM repeats seem to be present. The refined *in silico* structure seems to consist solely of α-helices and loops (Figure 48A). Nevertheless, this model is only *in silico* derived and does not to represent every aspect of the correctly folded SAUL1. This is particularly problematic, since all *in silico* derived models are based on homologs of the protein of interest. The closest homolog of SAUL1, does only have a sequence identity of 16 % (β-catenin, *Danio rerio*) and does therefore represent SAUL1 only in a very limited way. Nonetheless, the observed secondary structure is in high accordance with former experimental data and does therefore represent SAUL1 very well. CD measurements could show, that SAUL1 consists of 67.8-75 % of α-helices and 6-14.5 % of β-sheets (Figure 31) [346]. This proportion corresponds to the refined *in silico* model. The proportion of β-sheets, which were missing in the model, could be localized in some regions, which were calculated now as loops or defined as completely flexible regions. To obtain, the complete atomic structure of SAUL1, experiments with a higher resolution, like X-ray crystallography, would be necessary.

Figure 47: SAXS structure of monomeric SAUL1. *In silico* structure of SAUL1, which was obtained in a batch SAXS experiment using different protein concentrations. Monomeric SAUL1 was detected in a ratio of 52:31:17 (Monomer:dimer:tetramer) measuring a concentration of 0.6 mg ml^{-1}. Depicted are α-helices (red), β-sheets (blue) and loops (grey) [350].

Although being divided in two parts via a flexible linker (Figure 45B), nearly all five different U-box models are folded in the same way. Consequently, in all models the two α-helices are ordered parallel to each other with loops and the flexible region in between them (Figure 48B-E). Thus, it seems as if there would be a need for the U-box to be folded in this way. In addition, there is a slight conformational similarity to the NMR-derived structure of the U-box of the E3 ubiquitin ligase E4B from *Mus musculus* (Figure 48F). Accordingly, the SAXS-derived U-box structure could come very close to the actual one.

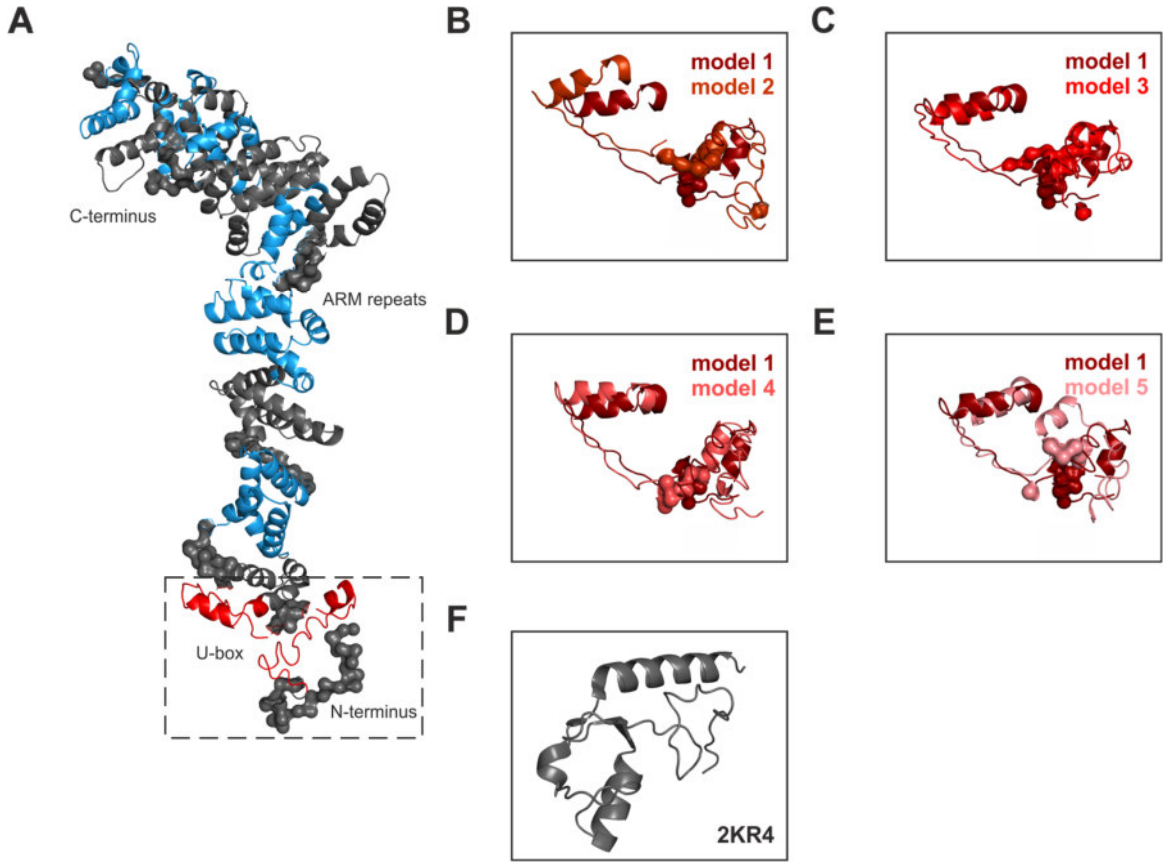


Figure 48: Domains of SAUL1. (A) *EOM* model 1 of SAUL1 with the U-box (red) and the ARM repeat domains (blue) from the *InterPro* predictions. Section of the U-box of model 1 aligned to **(B)** model2, **(C)** model 3, **(D)** model 4 and **(E)** model 5. **(F)** U-box of the E3 ubiquitin ligase E4B from *Mus musculus* (*PDB*-ID: 2KR4).

As already discussed, the *ab initio* and *in silico* structures of SAUL1 appeared to be bended, separating SAUL1 in two regions, resulting from the derived flexibility. Such an intrinsic flexibility has been described before for ARM repeats, in case of the adenomatous polyposis coli protein [356] and β-catenin from *H. sapiens* as well [357,358]. To investigate this flexibility and maybe determine the flexible regions, a feasible experiment would be limited proteolysis coupled with MS analysis. This would allow very easily to discover the flexible regions, which would result in a more precise and sophisticated structure analysis of SAUL1.

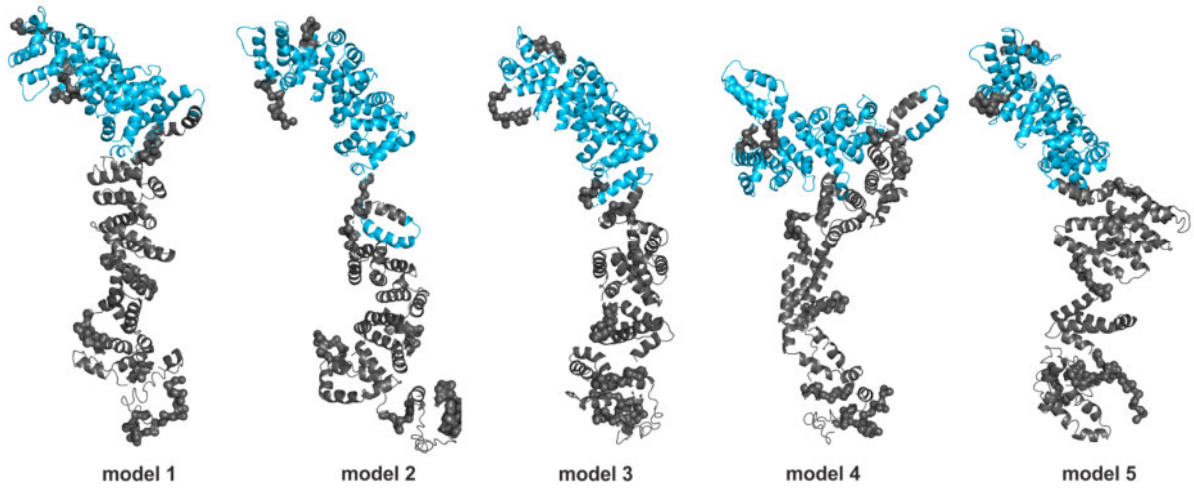


Figure 49: ARM repeats of SAUL1. Depicted are all five *EOM* models of SAUL1. The potential ARM repeats 7-11, which are responsible for the localization of SAUL1 at the plasma membrane [113], are colored in blue.

Since SAUL1 is separated into two parts, one part containing the U-box and the first ARM repeats and the other one consisting only of ARM repeats, one could speculate that these regions have different functions. Remarkably, these last ARM repeats are those, which have been shown to be responsible for the localization to the plasma membrane (Figure 49) [113]. Although ARM repeats mostly form important domains for protein-protein interactions [359], their role in protein-membrane interactions has been suggested. In case of the E3 ligase PUB13 it has been observed that it localizes specifically to phosphatidylinositol 4,5-bisphosphat ($PI(4,5)P_2$), which is a component of the plasma membrane [360]. Regions containing $PI(4,5)P_2$ are normally negatively charged and are known to regulate the surface charge of membranes, in concert with other phospholipids. Such regulations are crucial for some protein localizations [361] and ARM repeats contain positively charged groves, which are generally known to mediate interactions [356]. It has therefore been hypothesized that these regions could be the bridging factor to membranes [96]. Accordingly, SAUL1 may be localized to the plasma membrane due to a C-terminal positively charged region interacting with a negatively charged part of the plasma membrane.

To investigate, whether SAUL1 may contain such a positively charged grove, the electrostatic potential was analyzed by splitting the model 1 of SAUL1 (Figure 46A) into two domains based on their size. It has to be mentioned that only the structurally ordered regions were analyzed, because it was not possible to calculate the electrostatic potential of the flexible linkers. Thus, the actual electrostatic potential of SAUL1 might differ to some extent, also because no continuous solvent accessible surface could be determined. Nonetheless, it was possible to calculate the electrostatic potential and discover four distinct regions, which were characterized by their coherent charge (Figure 50). The most remarkable one is a positively charged grove (Figure 50, black arrow), which is composed of three arginines. Arg744, Arg745 and Arg783 are localized next to the C-terminus and are therefore on the far end of SAUL1. This is the only larger positively charged area on the surface of SAUL1. A smaller one was localized to the lower region of the upper ARM region (grey arrow) and might not actually exist in SAUL1, because in this part of the protein the domains would be connected and this region would be buried when both parts are joined. The

main positively charged grove was observed in three of the four remaining *EOM* models (model 3, 4 and 5) as well with a similar localization (data not shown). Consequently, the interaction of SAUL1 with the plasma membrane is likely to be mediated by this grove, especially as it is localized at the C-terminus. Previous experiments could show that even parts of this potential ARM 7-11 domain region, which are localized closer to the N-terminus, were required for the membrane localization [113]. However, it cannot be excluded that their deletion caused a structural change, which resulted in the disappearance of the positive grove. Another option would be, since it was not possible to calculate the electrostatic potential for all regions of the SAUL1 model, that additional positive stretches exist, which connect SAUL1 to the plasma membrane. These could be localized more to the N-terminal end of this 7-11 ARM domain region. However, in order to test whether this grove mediates the localization of SAUL1 to the plasma membrane, these arginines should be converted into alanines. Such a mutant SAUL1, which is tagged with green fluorescent protein (GFP), could be used in a localization experiments to test whether there is still a localization at the plasma membrane.

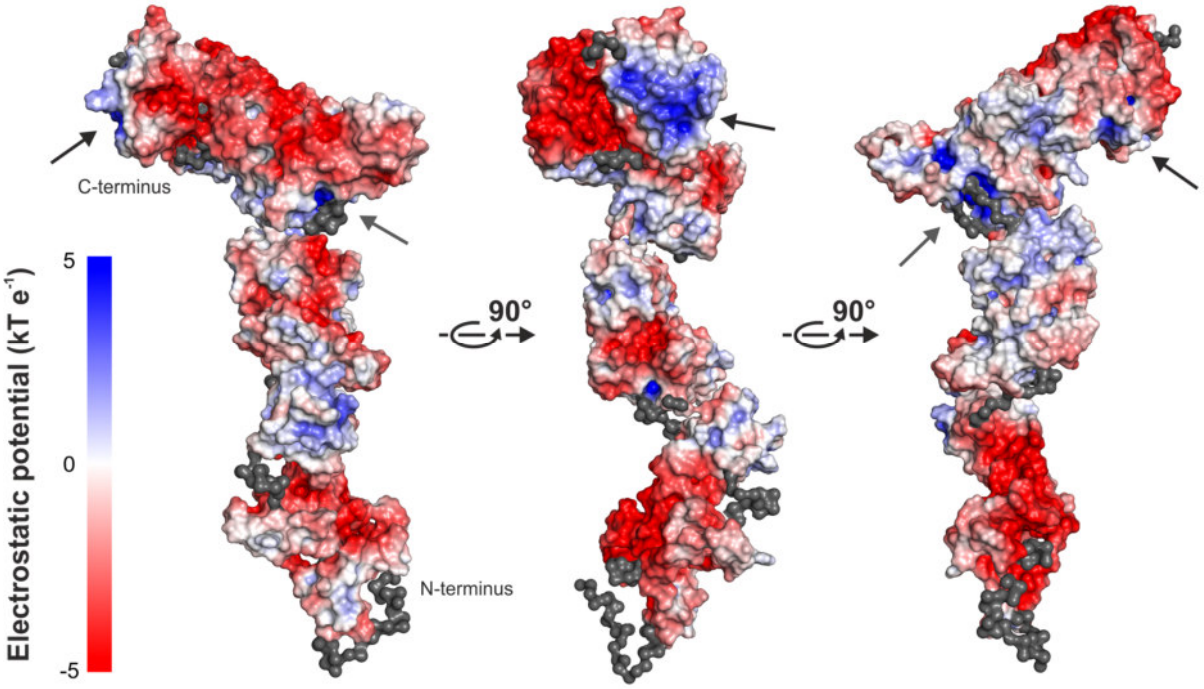


Figure 50: Electrostatic potential of SAUL1. Electrostatic potential of the solvent accessible surface of SAUL1 model 1 in a range of ± 5 kT e^{-1}. Calculations were performed by splitting the obtained SAUL1 model in two regions. One containing the potential ARM repeats 7-11 and the rest of the protein. Blue areas represent positively charged and red colored regions represent negatively charges areas. Marked (black arrow) is a positively charged grove next to the C-terminus in the area of the potential ARM repeats 7-11, as highlighted in Figure 49.

In addition to the positively charged grove, two large negatively charged regions were discovered. Whereas one was localized next to the N-terminus in the region of the U-box, the other one was found in the C-terminal ARM repeat region opposite of the positively charged grove. Charged regions are always of high interest as they are likely to mediate interactions. Particularly the discovery of such a region at the C-terminus, may point to a dual function of these ARM repeats at the C-terminal end of SAUL1. In addition to the association to the plasma membrane, it may thus allow for protein-protein interactions [359] with regard to the broad range of potential interaction partners of SAUL1 [138,142]. The second negatively charged region next to the N-terminus may support the interaction between SAUL1 and ubiquitin, because it has been shown previously that a negatively charged stretch is the interacting region between the U-box of such an E3 ligase and

ubiquitin [362]. Due to the proposed structure of SAUL1 it was not possible to take all amino acids into account, because some were located in the flexible linker regions. For this reason, a future model considering all amino acids could lead to a more specific electrostatic potential distribution. Additionally, such long negatively charged stretches could, potentially result in a repulsion between SAUL1 monomers, as described for the human E4B [363]. Since SAUL1 reportedly forms oligomers [141,346], higher oligomeric structures may therefore be less realistic. Another possibility would be that such negative stretches could be much smaller in a final model that takes all amino acids into account.

Taking together all these results, this work allowed for the first time to model SAUL1 and therefore an E3 ligase as a monomer and to obtain insights into its overall shape and structural order. SAUL1 flexibility was reliably confirmed and a more conserved region was detected that may mediate the protein-membrane interactions. The proposed structure of SAUL1 allows to narrow down the potential interaction site, being a positive grove at the C-terminus of SAUL1. In addition, it was possible to investigate the complexity of this region and thus to show that it may have different functions. To get further insights into the function of SAUL1 and to characterize its domains in more detail, additional experiments will be required. Those should focus on potential interaction sites, by determining the involved amino acids. It was not possible to crystalize SAUL1 by itself, possibly due to its polydispersity and flexibility. Future co-crystallization experiments may allow to determine the overall structure and potential interaction sites. Nonetheless, the performed experiments also allowed to model polydisperse and oligomeric systems, because the overall structure of SAUL1 has now been determined.

4.2.4 STRUCTURAL ANALYSIS OF SAUL1 AS A POLYDISPERSE SYSTEM

In batch measurements SAUL1 protein concentrations from 0.43 to 4.96 mg ml^{-1} were measured (Table S7A). No higher concentration was used for the series due to the tendency of SAUL1 to form aggregates (Figure 32A). A sample with a concentration of 3.3 mg ml^{-1} exhibited signs of aggregation in form of a steep intensity increase of the scattering data at very low angles and was therefore excluded from further analysis (data not shown). In a first step the scattering curves of the different concentrations were compared. These comparisons revealed a concentration dependency of the samples because the curves differed significantly from each other. The higher the concentration of the sample was, the steeper the slope, especially at lower angles. In addition, a bend was detected at higher concentrations (Figure 51A) and the detected $I(0)$ values in the Guinier plot differed. The $I(0)$ values were larger at higher concentrations, indicating that bigger particles were detected in the solution (Figure 51B). Furthermore, the calculated R_g values increased towards higher concentrations, from 6.27 ± 0.43 nm up to 10.64 ± 1.55 nm (Table S7B, C). and therefore, larger particles. The detected R_g values corresponded to molecule sizes, which were significantly larger than a monomer of SAUL1 with a R_g value of 4.64 ± 0.54 nm (Figure 51C). This detected concentration dependency is in accordance with previous DLS measurements (Figure 39A) and preceding native PAGE and SAXS experiments [346,350]. Thus, the scattering data seemed to originate of polydisperse systems containing multiple oligomers of SAUL1.

To compare the different shapes and compare these, a dimensionless Kratky plot was used that was independent of concentration- and size-effects. For all concentrations the plot showed a non-bell-shaped curve reaching a plateau around 3. With increasing concentration, the curves declined more delayed (Figure 51D). The overall shape was typical for elongated and partially flexible proteins, as globular ones tend to have a clear bell-shaped curve and a maximum around $\sqrt{3}$ [349].

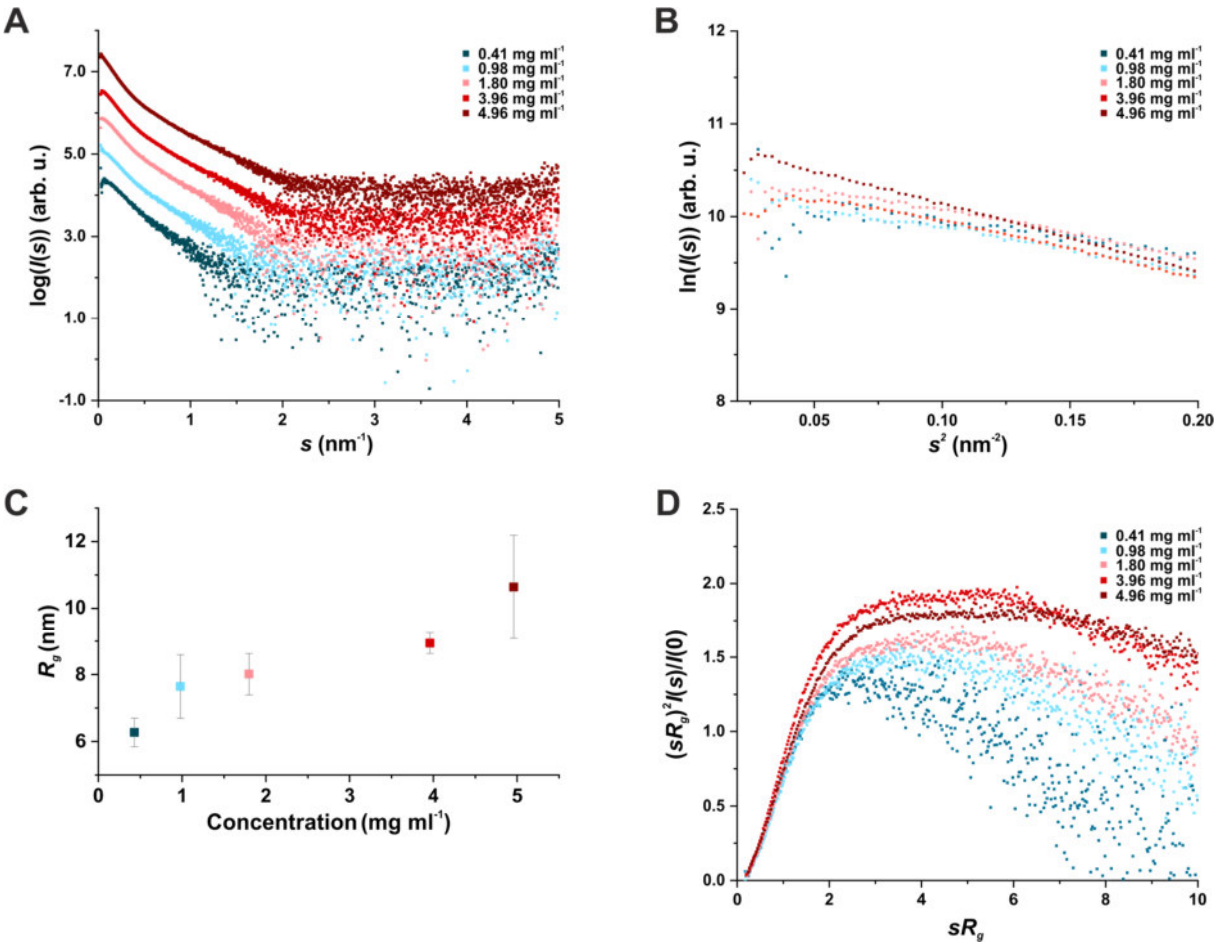


Figure 51: Scattering data of SAUL1 in a batch experiment. (A) Two-dimensional Log-linear plot of $I(s)$ versus s of the scattering data of different concentrations. Curves were plotted with an offset by being multiplied by a factor of 5 compared to the previous scattering curve. **(B)** Guinier plot of the SAXS data around $I(0)$. **(C)** Plot of the different Guinier-derived R_g values and standard deviations against their corresponding concentrations. **(D)** Dimensionless Kratky plot with the intensities normalized to the forward scattering intensity ($I(0)$) and the radius of gyration (R_g).

By using an indirect Fourier transformation of the scattering data (P(r)) distributions were calculated. All distributions represented a semi-bell-shaped left-shifted curve with a maximum increasing towards higher concentrations (Figure 52A). The d_{max} increased from 28.1 nm at 0.43 mg ml^{-1} to 45.7 nm at 4.96 mg ml^{-1} (Table S7B, C). All P(r) distributions represented the experimental data well, as the standardized residuals of the fit against the raw data were mostly below 2 and had a quality estimate values around 70 % (Figure 52B). By using the obtained $I(0)$ values and a measured BSA standard (data not shown) molecular weight values ranging from 176.6 kDa to 264.7 kDa were calculated (Table S7B, C). Those values did not represent SAUL1 as a monomer with 88.8 kDa or a defined multimer. Thus, these data also indicated that SAUL1 was present as a polydisperse system in the measured solutions.

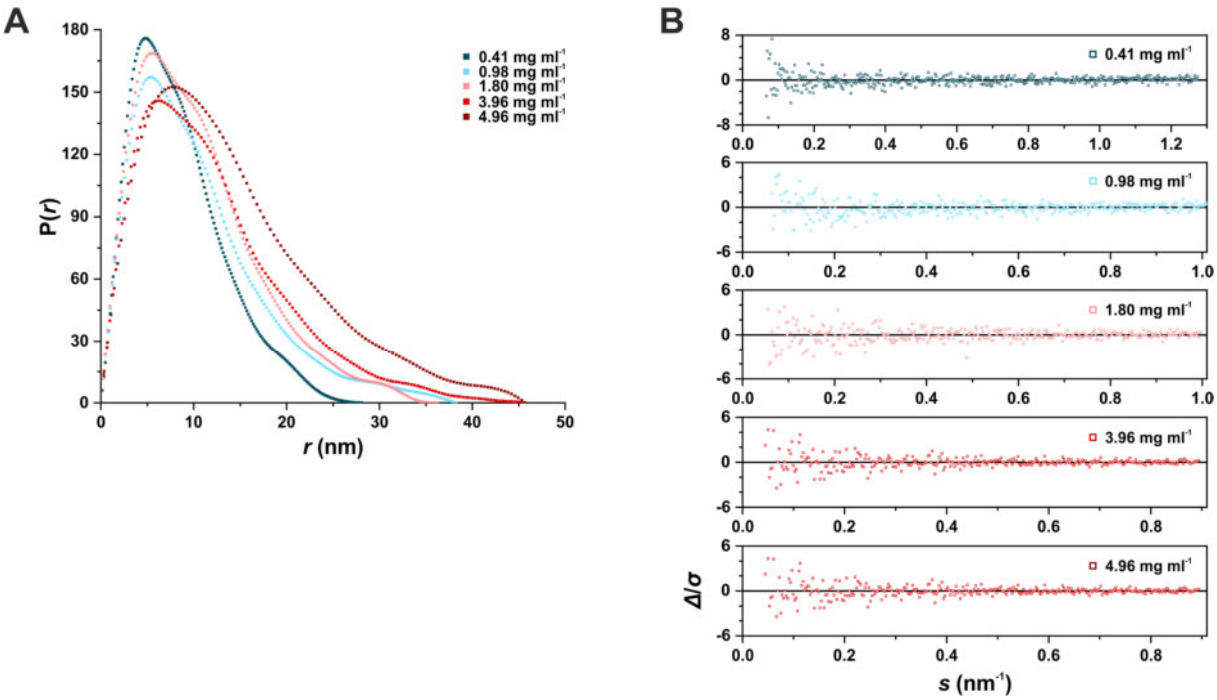


Figure 52: Radius distributions of SAUL1 in the batch measurement. (A) P(*r*) versus *r* profile of SAUL1 at different concentrations. **(B)** Plot of the different standardized residuals of the fits of the P(*r*) functions against the scattering data.

To determine the fraction of different oligomers *OLIGOMER* from the *ATSAS* package was used. Prior to this analysis, different multimeric models of SAUL1 had to be generated. This was done using *SASREF*, which enables a fit of oligomeric structures to experimental scattering data. As a monomeric model was obtained during the SEC-SAXS analysis, this one was thought to be the basis for the different multimers. Thus, the most frequent one, being model 1 (Figure 46C), was used. Models were calculated with respect to the different concentrations to achieve the most accurate fit of these structures. Because the radius distributions obtained at different concentrations suggested that very large structures were present in the polydisperse SAUL1 solutions, even octameric structures were generated to account for these findings. Although it was possible to generate all these models, using the *EOM*-derived monomer did not result in a good fit of the oligomers to the experimental data (data not shown).

To improve the quality in fitting the oligomers to the data, an alternative approach was chosen. Since the SAUL1 protein proved to be quite flexible (Figure 42C), the monomeric structure of the SEC-SAXS experiment might therefore not be present in solutions with higher concentrations. In these solutions interactions between the SAUL1 molecules might lead to conformational changes. To allow for potential flexibility and slight structural changes, SAUL1 multimers were modeled again with *SASREF*, but by using SAUL1 domains instead of a rigid monomeric structure. These domains were determined according to the *EOM* analysis, by looking at predicted protein flexibility (Table S2B). The *I-TASSER*-derived SAUL1 structure was divided into 10 domains (Table S8). These domains were used to generate multimeric structures based on the different protein concentrations. Although it was found that combinations of these structures fit better to the experimental data, the obtained fit was not sufficient (data not shown).

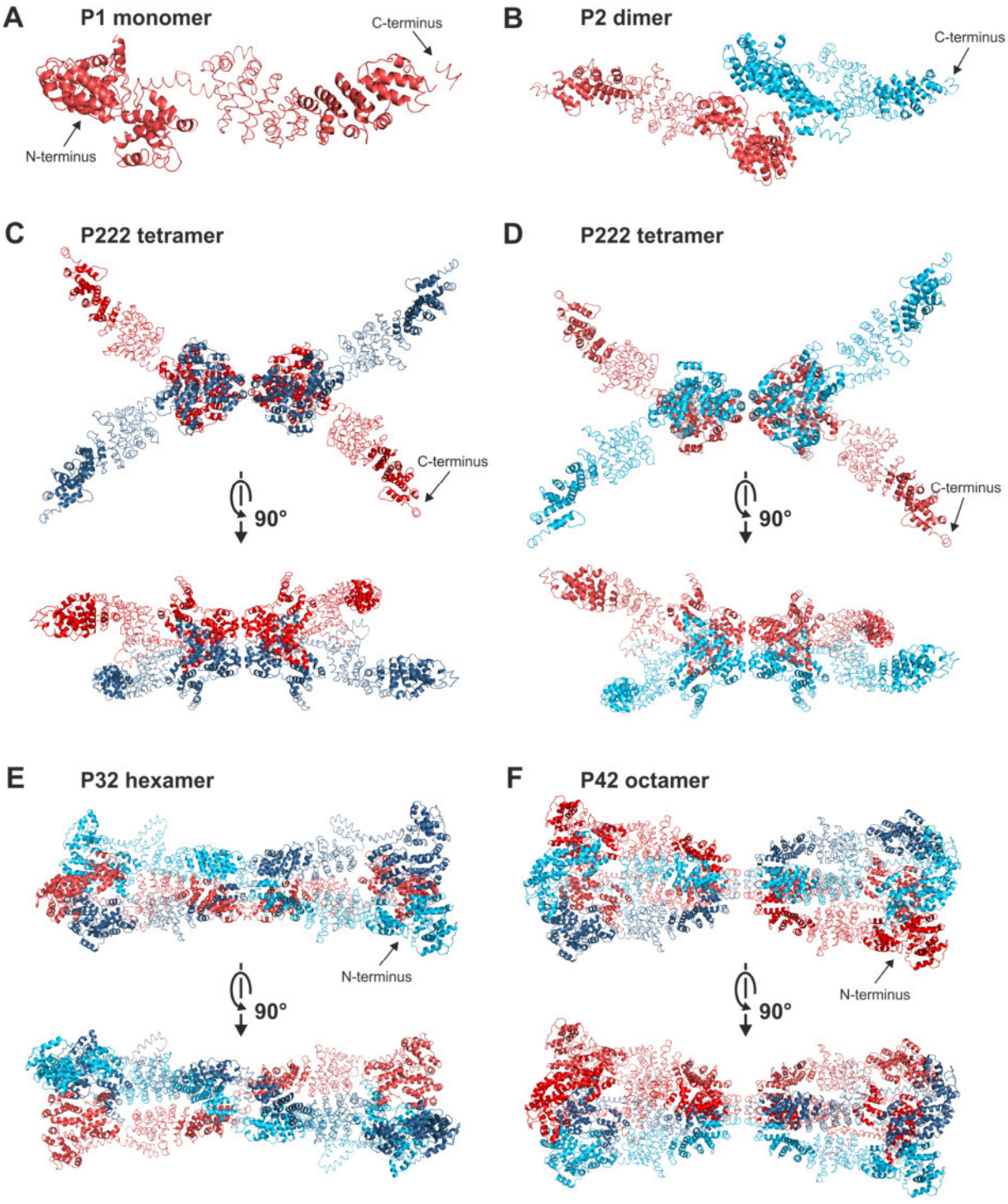


Figure 53: Oligomeric states of SAUL1. *SASREF* derived oligomeric models of SAUL1. Monomeric subunits are colored in one color. **(A)** Monomeric model, which was generated using SEC-SAXS experimental data. The model was built using *CORAL* and refined with *SREFLEX*. All oligomeric states are based on this model. **(B)** P2 dimer model fitted to a concentration of 0.43 mg ml^{-1}. P222 tetrameric model fitted to SAUL1 solutions with concentrations of **(C)** 0.98 and **(D)** 3.96 mg ml^{-1}. **(E)** P32 hexameric and **(F)** P42 octameric model, which were fitted to a SAUL1 concentration of 0.43 mg ml^{-1}.

In general, multimers are thought to consist of monomers, which are regularly arranged to form distinct space groups. Thus, the monomeric structure, which is the basis, is essential. To generate a most likely one, in a third approach, a monomeric model of SAUL1 was generated based on the 10 domains and fitted to the SEC-SAXS data using *SASREF*. In a second step the obtained model

was refined using *SREFLEX* allowing for a more probable structure. This monomer, having a χ^2 value of 1.084 was then used as a starting structure for *SASREF* calculations. Different multimeric structures were calculated, being based on the experimental data of the measured SAUL1 concentrations. Combinations of these multimers were used to calculate diverse mixtures, which were fitted against the experimental data using *OLIGOMER*.

This resulted in six different oligomers which are most likely present in the SAUL1 solutions. Those were the SEC-SAXS-based P1 monomer, a P2 dimer, which was derived from the SAUL1 solution containing 0.43 mg ml^{-1} protein, two P222 tetrameric structures, being based on 0.98 and 3.96 mg ml^{-1} and P32 hexameric as well as P42 octameric structures, which were calculated using a concentration of 0.43 mg ml^{-1} (Figure 53). Large oligomeric structures, like the P32 hexamer and P42 octamer were considered to be present due to the calculated P(*r*) functions, which pointed towards larger structures being present in solution (Figure 52A). In addition, previous direct stochastical optical reconstruction microscopy (dSTORM) experiments detected such structures *in planta*. In these experiments, SAUL1-GFP constructs were analyzed, detecting di-, tri-, tetra- and hexamers [142,346]. Although trimers were detected in dSTORM experiments, no fraction corresponding to these arrangements were detected in the analysis of the batch experiment.

When analyzing their oligomeric distributions, a concentration dependency was demonstrated. At low concentrations the SAUL1 dimer was found to be the dominant fraction (0.81 ± 0.02). The higher the concentration the lower was the abundancy. In contrast, the tetrameric structure was detected more frequently, changing from a fraction of 0.19 ± 0.02 at 0.43 mg ml^{-1} to 0.91 ± 0.01 at 4.96 mg ml^{-1} (Figure 54A, Table S9). Larger oligomeric structures were detected in samples containing 1.80 mg ml^{-1} SAUL1 and more. However, their fraction remained rather low, being always smaller than 0.10.

For the fit to the experimental data two different P222 tetramers were used. The one obtained from a SAUL1 concentration of 0.98 mg ml^{-1} (Figure 53C) fitted best to the SAXS measurements with a concentration of 0.43 and 0.98 mg ml^{-1}. The P222, which was based on the measurements at 3.96 mg ml^{-1} (Figure 53D), resulted in the best fit for 1.80, 3.96 and 4.96 mg ml^{-1}. In case of the batch measurements at 0.43, 0.98 and 3.96 mg ml^{-1} their fits to the original data were quite good with χ^2 values being mostly around 1.3 (Figure 54B, C, E). This was not the case for the measurements with 1.80 and 4.96 mg ml^{-1} of SAUL1. In case of the former no better fit by changing the oligomeric composition was obtained. A χ^2 value of 1.67 does not indicate a perfect fit. Therefore, aggregates may have been present in the solution, which would prevent determining an oligomeric composition. This represented the experimental data well, since SAUL1 is known to be prone to aggregation. Regarding the highest protein concentration, the goodness of fit was not as high as expected (χ^2 = 4.22). This may have resulted from aggregation in the sample since 4.96 mg ml^{-1} of protein is a SAUL1 concentration for which in which aggregation of SAUL1 has been documented (Figure 32A). Hence, this suboptimal fit may have been caused by this. This is supported by slight signs of aggregation in the form of a steep slope at small angles detected in the scattering plot (Figure 51A).

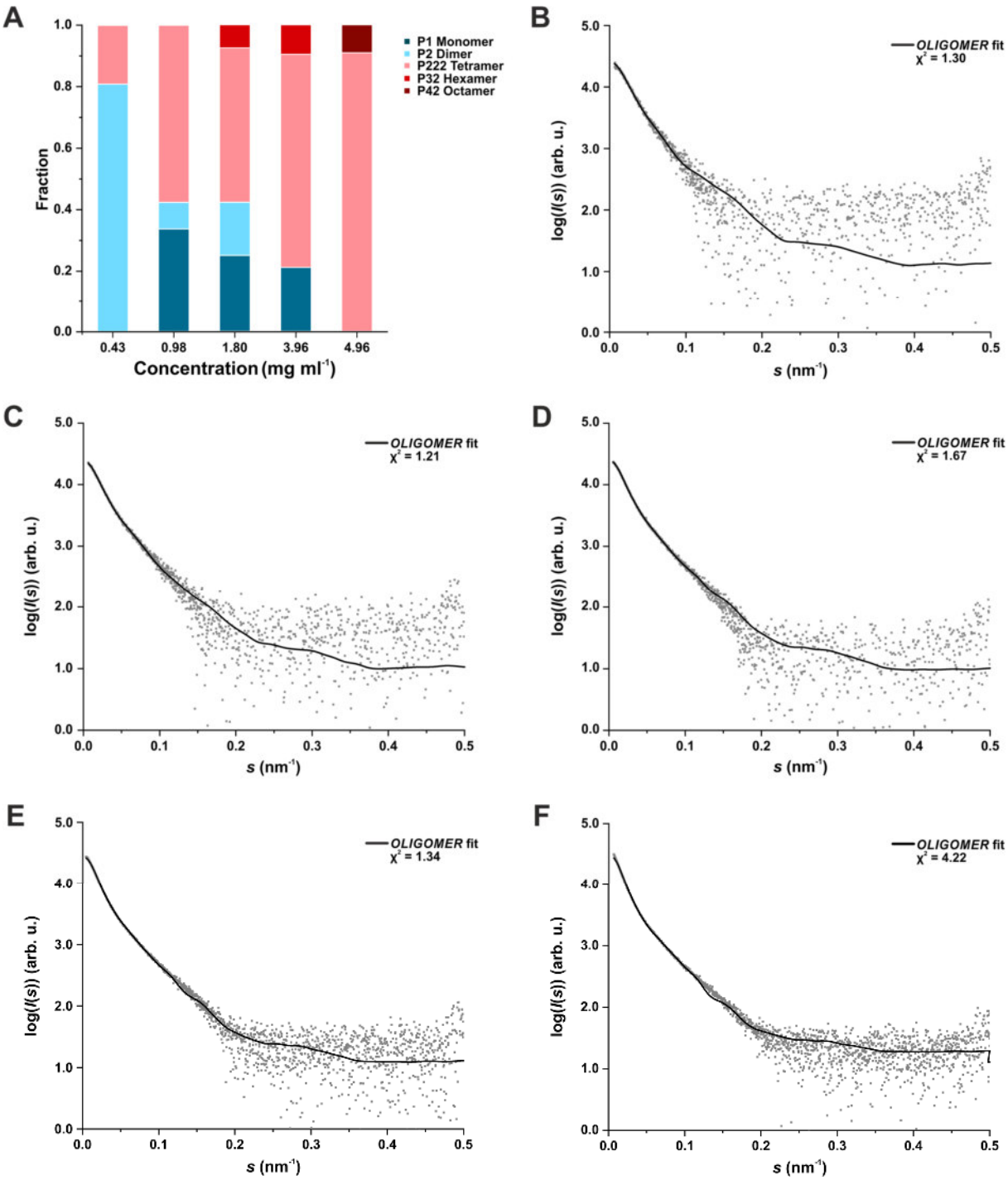


Figure 54: Oligomeric analysis of SAUL1. Analysis of polydisperse SAXS data of SAUL1 using *OLIGOMER*. **(A)** Fractions of different oligomeric states of SAUL1 at various protein concentrations. Fits of the *OLIGOMER*-derived fractions of oligomeric states of SAUL1to the experimentally-derived data in solutions with concentrations of **(B)** 0.43 mg ml^{-1}, **(C)** 0.98 mg ml^{-1}, **(D)** 1.80 mg ml^{-1}, **(E)** 3.96 mg ml^{-1} and **(F)** 4.96 mg ml^{-1}.

The presented results were compared to unpublished experimental data from Kilani and Kikhney, who measured and analyzed a SAUL1 concentration series ranging from 0.6 mg ml^{-1} up to 7.3 mg ml^{-1} [350]. During the analysis a concentration dependency was detected as well, as a P222 tetramer became the most frequent multimer at 7.3 mg ml^{-1}. Thus, both experiments seemed to be in accordance. However, no higher oligomeric structures were detected in the previous experiment. This is quite striking, as high concentrations (5.6 and 7.3 mg ml^{-1}) were measured

that have been shown to lead to larger particles and aggregation. In particular, DLS measurements which were repeated two times, clearly revealed that concentrations above 3 mg ml^{-1} contained aggregates, due to a strong increase in the detected radii and enlargement of their distribution (Figure 32A, 38A). Signs of aggregation were detected in the previous experiments in the Guinier plot of the scattering data for a concentration of 7.3 mg ml^{-1}, in form of a strongly decreasing slope at low angles [350]. As a result, it would be expected that larger R_g values would have been detected. In all measured concentrations, however, R_g values were rather small. In conclusion, the concentrations of SAUL1 in the previous experiment might have been overestimated. In spite of that, both experiments showed accordingly that SAUL1 is depending on its concentration mostly present as a P2 dimer or a P222 tetramer. This is in line with previous bimolecular fluorescence complementation (BiFC) and yeast two-hybrid (Y2H) experiments demonstrating that SAUL1 does interact with itself and can potentially form multimers [141].

Di- and tetramers of SAUL1 were found to be the most abundant oligomers even at low concentrations. Hence, these would be the most likely conformations of SAUL1 *in planta*. In particular because those were detected even at low concentrations. The electrostatic potentials calculated for the dimer (Figure 50) would allow for pairing of the first negatively charged stretch with a smaller positively charged grove to stabilize the dimer. In addition, SAUL1 was localized at the plasma membrane via its C-terminus [113]. Thus, oligomers which do not mask the C-terminus may be much more likely *in planta*. A steric hindrance for membrane association may exist in case of the modeled hexa- and octameric structures, because the C-termini were modeled to be on the inside of the oligomers (Figure 53E, F). In addition, a multimer formation via the C-terminus is less likely, because the negative stretches detected in the second ARM repeat containing region (Figure 50) would prevent such an interaction. Although a hexameric structure was detected in dSTORM experiments, this might have resulted out of too high concentrations of SAUL1, because the protein was expressed under the control of the very strong *35S*-promoter [142,346]. Furthermore, it was found that the detected P6 space group in the dSTORM experiment is less likely to fit to the experimental SAXS data than a P32 space group of a hexamer. It is therefore questionable whether such a high oligomer of SAUL1 is present under physiological conditions.

In support of the hypothesis that SAUL1 is mostly present as a di- or tetramer the interacting region between the SAUL1 monomers was modeled to be the U-box-containing region and parts of the proposed ARM domains of SAUL1 (Figure 53). These findings are in line with previous crystallization experiments analyzing the dimeric PUB CHIP (CARBOXYL TERMINUS OF HSP70-INTERACTING PROTEIN) revealing that its interaction is mediated via the U-box [364]. The same has been shown for PUB22 [116]. In addition, an important role of the ARM repeats for oligomerization in PUBs has been described as well. In case of PUB13 an intrinsic interaction of its ARM repeats has been reported [124], which could point towards an interacting mechanism. Thus, it is very likely that the self-interaction of SAUL1 also takes place via its U-box and the first ARM repeats, which has been proposed before for PUBs in general [96]. In order to investigate the oligomerization function of the U-box, a deletion construct could be expressed, purified and structurally investigated using DLS or SAXS. Data on PUB10 and PUB22 support the idea that the detected oligomerization of SAUL1 may be of physiological relevance. For both, PUB22 and PUB10, the activity depends on their oligomerization-state. The E3 ligase PUB10 was only active as a homodimer [114]. In contrast, a PUB22 homodimer lead to autoubiquitination keeping its abundance rather low, and only monomeric PUB22 bound to its targets [116]. Although an oligomerization dependent activity has not been shown for SAUL1 yet, it may well be regulated via this mechanism, in particular because multimers have been detected *in planta* [142,346]. Furthermore, SAUL1 oligomers formed already at very low concentrations, suggesting that they are present in the cell under physiological conditions.

4.2.5 RELEVANCE OF OLIGOMER FORMATION FOR THE SAUL1 FUNCTION

An interesting observation in the context of oligomer and aggregate formation is the occurrence of large patches at the plasma membrane of *35S::GFP-SAUL1*, *35S::SAUL1ΔARM$_{1-6}$-GFP* or *35S::YFP-SAUL1(ARM$_{7-11}$)* plants [112, 131]. Generally, patches have been observed at the plasma membrane either in case of N-terminally tagged SAUL1 or a construct lacking the U-box. Patch formation has also been described for additional SAUL1-type PUB-ARM E3 ligases from different organisms and its close homolog PUB43 pointing to a more general mechanism of this class of PUBs [139]. In contrast, SAUL1 being tagged at the C-terminus with GFP resulted always in a evenly distribution at the plasma membrane [113]. There are two different mechanisms that could potentially explain such a pattern.

(i) Firstly, the patterns could be the result of concentration-dependent oligomerization of SAUL1, because all experiments have been performed by using the very strong *35S*-promoter. SAUL1 tends to form oligomers and aggregates at higher concentrations over time and seems to be present as a monomer only at very low concentrations (see chapter 4.2.3). This hypothesis is also supported by the fact that the size of such patches has been described to depend on the expression level [140]. In this context, it has been shown that SAUL1 patch formation is immune response-dependent. This is particularly interesting because an *in silico Genevestigator* analysis could show that *SAUL1* gene expression is elevated upon infection (data not shown). Patch formation took place in tobacco plants transiently transformed with *35S::YFP-SAUL1*, mainly in cells infected with *Phytophtera capsisci* [140]. Therefore, *SAUL1* gene expression levels should be relatively stable. Nevertheless, the tobacco homolog of *SAUL1* could as well be induced upon the onset of immunity, which could then trigger SAUL1 oligomerization.

(ii) Nonetheless, the most striking observation concerning patch formation was that SAUL1 patches were detected in all plants expressing SAUL1 without the U-box (*35S::SAUL1ΔARM$_{1-6}$-GFP* and *35S::YFP-SAUL1(ARM$_{7-11}$)*) [112, 131]. In addition, plants with a potential hindered U-box, like *35S::GFP-SAUL1* did show in 25-76 % of all transformed protoplasts patch formation [113] (also, Lienemann, T., unpublished data). In contrast, C-terminally tagged SAUL1 (*35S::SAUL1-GFP*), in which the U-box was not hindered, did not show any patch formation [113]. Therefore, the U-box seems to be important for patch formation. This domain is highly likely involved in SAUL1 oligomerization (see chapter 4.2.4) [96]. Thus, absence or distortion of the U-box could result in hindered oligomerization and therefore in patch formation. This mechanism can be explained by the SAXS-derived structural information on SAUL1. Particularly those investigations, which focused on plasma membrane binding of SAUL1 (see chapter 4.2.3 and 4.2.4).

Plasma membrane binding of SAUL1 is mediated by the potential ARM repeats 7-11 and most likely by a positively charged grove next to the C-terminus (see chapter 4.2.3) [96]. This grove is hypothesized to bind to negatively charged phospholipids like PI(4,5)P_2, which are localized in membranes. Binding of SAUL1 to these components could result in patch formation, since these components are not distributed evenly [365]. If the positively charged grove is inaccessible, SAUL1 would therefore no longer be able to bind to specific regions in the plasma membrane. Experiments could show that all ARM repeats 7-11 are important for plasma membrane binding and not only the region containing the grove [113]. The positively charged grove could therefore only be necessary for specific interactions with negatively charged phospholipids. Thus, an inaccessible grove would result in an evenly distributed SAUL1 signal at the plasma membrane. The accessibility of the grove could be influenced by SAUL1 oligomerization (Figure 53). In higher oligomeric models for hexa- or octamers, the C-terminus of SAUL1 is completely hidden and is therefore very unlikely to specifically interact with negatively charged regions in the plasma membrane (Figure 53E, F). In addition, oligomerization may generally hinder accessibility and therefore prevent binding of the grove to the negatively charged phospholipids.

However, plasma membranes are not the only cellular compartments containing negatively charged phospholipids. SAUL1 is involved in tethering multi-vesicular bodies (MVBs) to the plasma membrane [140]. MVBs are known to contain in their membranes a high percentage of phospholipids like PI(4,5)P_2 [366]. Therefore, the putative ARM repeats 7-11 could mediate two

interactions: (i) interaction with the plasma membrane and (ii) interaction with MVBs. Unspecific interactions with the plasma membrane may be mediated by the whole ARM repeat domain and binding to MVBs by the positively charged grove at the C-terminus. Consequently, oligomerization of SAUL1 would interfere with tethering of MVBs to the plasma membrane by rendering the C-terminus inaccessible for $PI(4,5)P_2$ or similar phospholipids in MVBs. This would result in a continuous signal of SAUL1 at the plasma membrane.

The additionally observed patch formation during infection also fits into this model. Upon infection MVBs are known to accumulate in the cytosol, localize to the infection site and release intraluminal vesicles into the apoplast [366]. These vesicles have been shown to act positively on the immune response and phospholipids serve as signaling molecules to activate downstream defense signaling. The essential step in MVB release is the fusion with the plasma membrane, which may involve SAUL1 [140]. Thus, patch formation would result from the tethered MVBs.

Nevertheless, two questions remain. (i) First of all, which function does SAUL1 fulfill as an active E3 ubiquitin ligase at the plasma membrane? In context of the immune response the ubiquitination of a SAUL1 target substrate in combination with MVB tethering is particularly interesting. During the immune response PRRs are internalized to regulate downstream signaling [367]. Internalization can be mediated by monoubiquitinating the target substrate and has been reported, for example, for FLS2 monoubiquitinated by PUB12/13 [128]. An important compartment of receptor endocytosis are MVBs that can recycle internalized receptors or fuse with the vacuole, leading to receptor degradation [368]. Accordingly, SAUL1 could monoubiquitinate PRRs during the immune response, which are subsequently transported to MVBs.

(ii) And secondly, how does SAUL1 mediate tethering of MVBs, specifically in the case of infection? To address this, SAUL1 activity and regulation need to be discussed. SAUL1 is very likely not always present as an active E3 ligase. Its activity could be, similar to other PUB family members, regulated by its oligomeric state [114,116]. Thus, SAUL1 could be inactive as an oligomer, which is for example phosphorylated upon infection. This additional negative charge could lead to monomerization and SAUL1 could subsequently help tethering MVBs to the plasma membrane. This mechanism is actually very likely, since this type of regulation of PUB activity has been shown for PUB22 [116]. In conclusion, the proven oligomerization of SAUL1 could serve as a regulatory mechanism to prevent untimely tethering of MVBs to the plasma membrane.

In summary, this thesis allowed for a profound analysis of the different concentration-dependent oligomeric states of SAUL1 being based on a well-defined monomeric structure. These experiments gave insights into the interaction between the SAUL1 monomers forming di, tetra-, hexa- or octamers and showed that these are likely to be mediated via the U-box and the first ARM repeat containing region. In addition, SAUL1 activity could be regulated by oligomerization, which could regulate as well MVB tethering. Furthermore, it was possible to identify a potential dual function of the C-terminally ARM repeats that may mediate both plasma membrane and MVB interaction.

4.3 INTERACTIONS OF SAUL1

Another key to unravel the function of SAUL1 may be the identification of its target substrate or of other interaction partners. Since SAUL1 is linked to multiple proteins, which are involved in different regulatory pathways [142,369], SAUL1 could also be involved in multiple pathways. To understand how they may interact with SAUL1 and what influence they might have on the function of SAUL1, the potential interaction partners had to be analyzed. Therefore, some part of this thesis was focused on studying the interaction between SAUL1, CPK5, AtKRP125b, UBCs, SOC3, CHS1 and especially BON1.

4.3.1 INTERACTIONS OF SAUL1 WITH E2 UBIQUITIN-CONJUGATING ENZYMES

SAUL1 E3 ligase activity has previously been shown in *in vitro* experiments [123]. To ubiquitinate a target protein, SAUL1 has to transfer ubiquitin from a ubiquitin-conjugating enzyme (E2) to a target substrate. To investigate the function of SAUL1, knowledge about the interacting E2s will be extremely helpful. The kind of ubiquitination that ranges from mono- to polyubiquitination of the target is largely determined by the interacting E2 [370]. Therefore, the interaction between SAUL1 and different E2s was analyzed.

Due to the large number of E2s in *A. thaliana* (37) only selected ones were investigated [371]. Most promising candidate would be E2s that like SAUL1 are localized at the plasma membrane. In a previous study, all E2s were analyzed concerning their binding capability to membrane-anchored ubiquitin-fold (MUB) proteins [372]. These proteins are members of the ubiquitin-fold protein family whose members, although sharing a common β-grasp three-dimensional structure, have a wide variety of functions. The six members of the Arabidopsis MUB protein family are characterized by a C-terminal cysteine-containing sequence that is crucial for prenylation. The prenylation functions as a membrane anchor localizing the MUB proteins to the plasma membrane [373]. These proteins have been shown to be able to recruit specific E2s to the plasma membrane. In *A. thaliana* seven E2s (UBC8, UBC9, UBC10, UBC11, UBC28, UBC29 and UBC30) were found to bind selectively to MUB proteins. These were solely members of the E2 subgroup VI and were found to reside at the plasma membrane [372].Thus, these E2s were tested for a potential interaction with SAUL1, with the exception of UBC11 as it was not possible to generate a functional BiFC construct with this E2. In addition, E2s, which were not found to interact with MUB proteins, were analyzed as a control. These were chosen based on sequence similarity being either highly related to the subgroup VI or being more distantly related. The analysis was done by transforming BiFC constructs into *A. thaliana* mesophyll protoplasts.

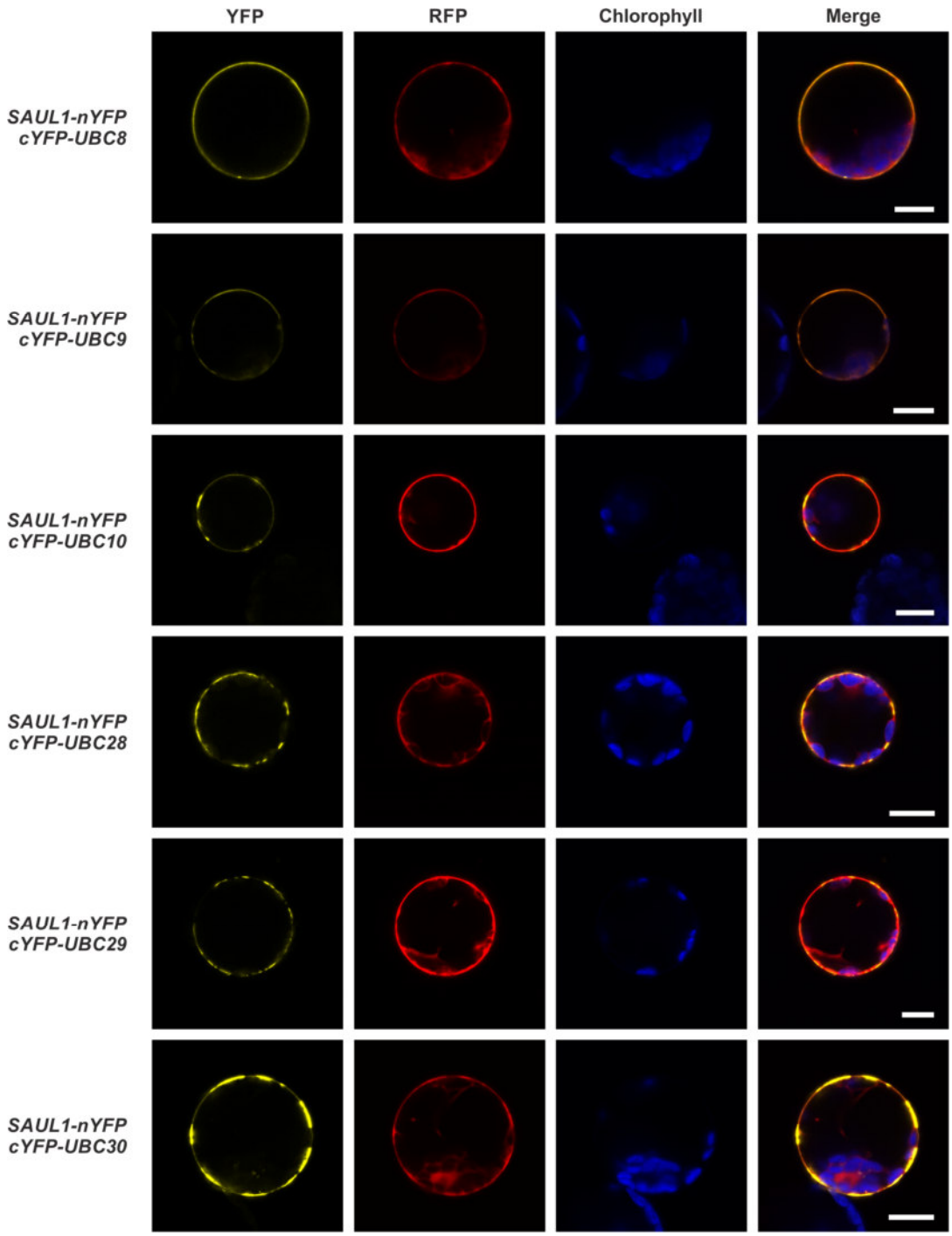


Figure 55: BiFC analysis of SAUL1 and E2s. Confocal laser scanning microscopy of transiently transformed mesophyll protoplasts. Protoplasts were transformed with BiFC constructs containing *SAUL1* being paired with *UBC8*, *UBC9*, *UBC10*, *UBC28*, *UBC29* and *UBC30*. *SAUL1* was always tagged C-terminally, while the different UBCs were tagged at the N-terminus. YFP, RFP and autofluorescence signals of chlorophyll were recorded. Scale bars represent 15 µm.

All plasma membrane-localized E2s of subgroup VI were found to interact with SAUL1, and the interaction was in all cases detected at the plasma membrane (Figure 55). Interestingly, the YFP-signal was found to differ in intensity between the E2s tested, which did seem to originate only in case of UBC9 from a lower transformation rate, as the control RFP-signal was proportionally weaker as well. In addition, a clustered YFP-signal was detected in case of an interaction with UBC10, UBC28, UBC29 and UBC30 that was reminiscent of patch formation (see chapter 4.2.5).

Surprisingly, an interaction between SAUL1 and E2s outside of subgroup VI, namely UBC13 and UBC17, was also detected at the plasma membrane. In case of UBC37 no YFP signal was detected (Figure 56), indicating that there is no interaction between SAUL1 and UBC37 when assuming that both proteins were expressed. In conclusion, SAUL1 showed interactions with many but not all tested E2s, in particular not only with those that were previously found to be localized at the plasma membrane.

The fact that SAUL1 was interacting with multiple E2s is not very surprising, because this has been reported for other E3 ubiquitin ligases. In case of PUB22 all *A. thaliana* E2s were tested and the selective interaction with 11 E2s occurred across all subgroups and regardless of their proposed localization [111]. Thus, E2s do not seem to be restricted in their interaction with corresponding E3 ligases regardless of their localization or subgroup affiliation. Interestingly, among those interacting E2s were nearly all members of subgroup VI, which were not detected to interact at the plasma membrane with PUB22 [111]. Thus, the E2s do not seem to be confined in their localization, although the interaction with SAUL1 was found to reside at the plasma membrane.

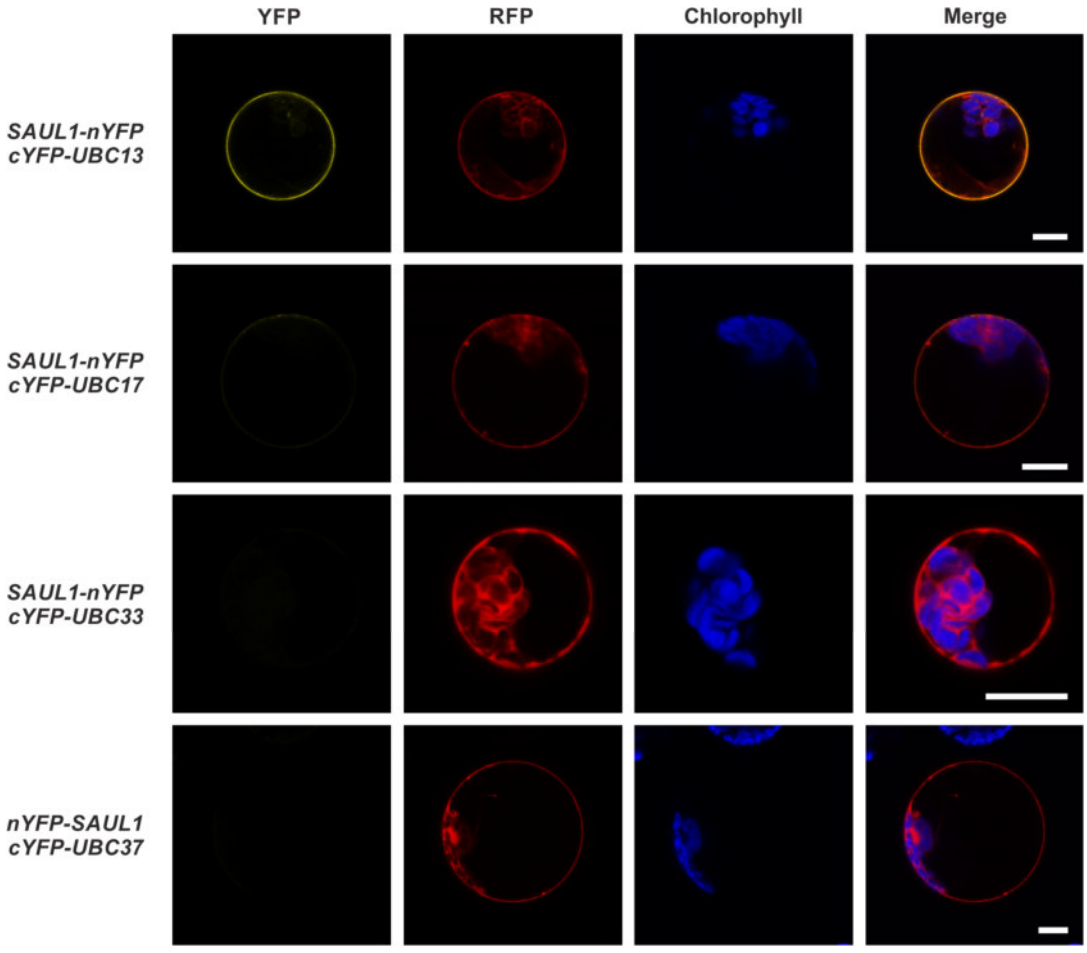


Figure 56: BiFC analysis of SAUL1 and E2s that have not been shown to be localized at the plasma membrane. Confocal laser scanning microscopy of transiently transformed mesophyll protoplasts. Protoplasts were transformed with a BiFC construct containing *SAUL1* being paired with *UBC13*, *UBC17*, *UBC33* and *UBC37*. *SAUL1* was always tagged C-terminally, while the different UBCs were tagged at the N-terminus. YFP, RFP serving as a transformation control and autofluorescence signals of chlorophyll were recorded. Scale bars represent 15 µm.

One important fact is that although target specificity is mediated via the E3 ligase, the actual ubiquitination specificity, determining poly- or monoubiquitination and the chain type, is mediated by the interacting E2 enzymes [370]. Thus, determining interacting E2s allows to specify

how a E3 ligase substrate is affected. In a previous study members of subgroup IV were found to be the most interactive ones [371]. Therefore, detected interactions and their suggested physiological relevant interactions might not be reliable. Even in the experiments on PUB22, in which all E2s were tested and specific interactions determined, the physiological role remained elusive [111].

Nevertheless, the patch-like interaction signal of SAUL1 with UBC10, UBC28, UBC29 and UBC30 was quite striking (Figure 55), because it resembled the patch formation, which had been described for SAUL1 [113,140]. Patch formation resulting from these interactions may still be artificial and does not necessarily have to originate from tethering of MVBs. Interestingly, SAUL1 was tagged at the C-terminus and for this orientation patch formation has never been described before [113]. However, such an interaction pattern was not observed in any other BiFC experiment that analyzed other binding partners of SAUL1 [141,142,369]. Therefore, the interaction between SAUL1 and the E2s could differ from all other tested binding partners resulting in patch formation. E2 enzymes bind to E3 ligases via the U-box. In case of SAUL1, it is assumed that the U-box also mediates oligomerization (see chapter 4.2.4). If an E2 binds to the U-box, oligomerization may no longer take place. This would result in an accessible C-terminus and tethering of MVBs being possible leading to patch formation. Therefore, a C-terminally tagged SAUL1 could as well result in patch formation. This has been shown for a construct that lacked the U-box and resulted in patch formation [113]. Consequently, only the E2s, which directly interact with the U-box of SAUL1, would result in patch formation during a BiFC experiment. For example, all other E2s could only interact with the ARM repeat domains of SAUL1. However, these interactions need to be further investigated to determine whether only some of the E2s bind to the U-box.

When taking a closer look at the potential physiological role, nearly all E2s interacting with SAUL1 (Figure 55) are closely related to the human UBE2D (UBIQUITIN-CONJUGATING ENZYME E2 D) family. Their sequence identity ranged from 60 to 80 % with a sequence coverage of about 99 % (Table S10). Accordingly, these E2s are highly alike. Furthermore, these are so-called minimal E2s that only contain an UBC domain. The human UBE2D family is known to mediate monoubiquitination of a substrate [374]. Similar findings in plants have been published recently, as the MUB-associated E2s were reported to be involved in monoubiquitination [375]. This monoubiquitination has been demonstrated to be mostly undirected, because these E2s were found to ubiquitinate any lysine without preference. This primary attached ubiquitin was reported to function as an initiation signal for other E2 enzymes. This is necessary since not all E2s are able to ubiquitinate a substrate without any primary attached ubiquitins. Thus, a so-called division of labor takes place, in which the UBE2D family has the role of priming E2s [374]. This may explain why a broad number of E3s is known to interact with these E2s, for example PUB22 [111]. Nevertheless, members of the UBE2D are also involved in receptor endocytosis in humans. They mediate Lys63-mediated polyubiquitination resulting in receptor internalization and subsequent endosomal degradation [376]. This could support the hypothesis, that SAUL1 is involved in immune-dependent endocytosis of PRRs, which are then transported to MVBs (see chapter 4.2.5).

Moreover, another E2, which was detected to interact with SAUL1, was UBC13. This E2 is not related to the UBE2D family, but to the UBE2G family (Table S10). In contrast to the former, this family is known to specifically mediate polyubiquitination at Lys48 [92]. This linkage type is known to result in proteasomal degradation [91], which is consistent with findings that SAUL1 is able to polyubiquitinate a target [123]. Therefore, a potential substrate of SAUL1 may be polyubiquitinated being mediated by UBC13. In future, it may be helpful to study the interaction of SAUL1 with all 37 Arabidopsis E2 enzymes with independent methods. This may allow to narrow down possible functions of SAUL1.

4.3.2 STRUCTURAL ANALYSIS OF ATKRP125B

Although the substrate target of SAUL1 is still unknown, SAUL1 has already been linked to positive regulation of PTI and tethering of MVBs to the plasma membrane [137,140]. To search for a target

substrate, tandem affinity purification experiments were performed and revealed an interaction with AtKRP125b, a family 5 kinesin [142]. This kinesin has been proposed to be involved in mitosis and was localized in the nucleus [369,377]. The interaction between SAUL1 and AtKRP125b was observed to take place in the nucleus as well, but also in the cytoplasm [142]. Although this seemed contradictory, because SAUL1 was predominantly localized at the plasma membrane. Nevertheless, localization experiments during this thesis in protoplasts using the native promoter of SAUL1 (*pSAUL1::SAUL1-GFP*) could show, that SAUL1 is transiently localized in the nucleus (data not shown). In addition, experiments using tobacco BY-2 cell culture cells showed that SAUL1 is localized in the nucleus 25 % of all transformed cells when treated with JA or auxin [110]. Therefore, an interaction may take place. However, the question about the physiological function of this interaction remains.

ATKRP125b is a member of the kinesin 5 family. In *A.* thaliana this family consists of four members [378]. First discovered in fungi [379], members have been reported to play an essential role in the spindle apparatus assembly and function [380]. A similar role has been reported in *A. thaliana* for AtKRP125c, which stabilizes the mitotic spindle [381]. However, SAUL1 has never been reported to be involved in the regulation of mitosis. Cells in the *saul1-1* mutant always divided regularly, in contrast to those observed in *rsw7* the *AtKRP125c* mutant [381]. However, another function of the kinesin 5 family was discovered in *Drosophila* S2 cells [382]. The family 5 kinesin Klp61F is involved in vesicle transport from the trans-Golgi network (TGN) to the cell surface. In the same study, a similar role was proposed for the human Eg5 [382]. Later on, an involvement of Eg5 in the transport of receptors to the cell surface of neurons was described as well [383]. Thus, this function seems to be conserved. Hence, members of the kinesin 5 family in *A. thaliana* could as well be part of the vesicle transport system.

This is highly remarkable, since SAUL1 is also connected to exocytosis. (i) On the one hand SAUL1 is involved in MVB tethering [140] that results in fusion of the MVBs with the plasma membrane and subsequent vesicle release [366]. (ii) On the other hand, SAUL1 has been linked to the exocyst complex. This complex, being first discovered in yeast consists of the proteins Sec3, Sec5, Sec6, Sec8, Sec10, Sec15, Exo70 and Exo84 [384]. The exocyst complex is responsible for the first contact of vesicles derived from the TGN to the plasma membrane and mediates the last step prior to exocytosis [385]. SAUL1 is linked to the exocyst, because SAUL1 overexpressing plants have been shown to be guarded by the truncated NLR TN2 [138]. TN2 itself is known to guard as well Exo70B1 [386]. In addition, Exo70 subunits bind selectively to negatively charged components of the plasma membrane like $PI(4,5)P_2$ [387,388]. SAUL1 is also likely binding to negatively charged phospholipids (see chapter 4.2.3).

Therefore, SAUL1 and Exo70B1 could reside in the same complex and SAUL1 may regulate potential components of exocytosis like AtKRP125b. A similar mechanism has also been reported for members of the PUB family. PUB22 is known to regulate Exo70B2 by polyubiquitination [129]. This results in degradation of Exo70B2, repression of exocytosis and subsequent suppression of the immune response. In contrast to PUB22, SAUL1 is a positive regulator of PTI and could therefore target suppressors of exocytosis for polyubiquitination. SAUL1 could also modify receptors or other proteins by monoubiquitination (see chapter 4.3.1), resulting in their internalization or changing their activity. Thus, investigating the interactions between SAUL1 and proteins associated with vesicle transport is of high interest. Particularly structural information on a binding partner like AtKRP125b will give valuable insights into their interaction and how both may mediate exocytosis.

Expression of AtKRP125b

To investigate the structure of AtKRP125b, the protein was recombinantly expressed. In a first step AtKRP125b was cloned into the pGEX-6P-1 vector [142]. To express GST-AtKRP125b, pGEX-6P-1-AtKRP125b was transformed into BL21 Star™ (DE3), BL21-CodonPlus (DE3) RIPL and Origami™ B (DE3) cells. Those cells were used in an expression test to analyze different temperature and induction conditions (Figure 57).

A band corresponding to GST-AtKRP125b was detected in BL21 Star™ (DE3) cells at 18 °C after 16 h and at 37 °C after 20.5 h (Figure 57A, B). Both conditions were induced with 1 mM IPTG at an $OD_{600\,nm}$ of 1. The expression test in BL21-CodonPlus (DE3) RIPL cells revealed only a slight induction after 22 to 24 h. This was independent of the incubation temperature and the induction conditions (Figure 57C, D). In Origami™ 2 (DE3) pLysS cells, which were induced with 1 mM IPTG at an $OD_{600\,nm}$ of 1, a band corresponding to GST-AtKRP125b with a size of 140.8 kDa was detected. The sample corresponded to cells that were incubated for 17 h at 18 °C (Figure 57E). Expression of AtKRP125b was verified in the corresponding bands using MS (data not shown).

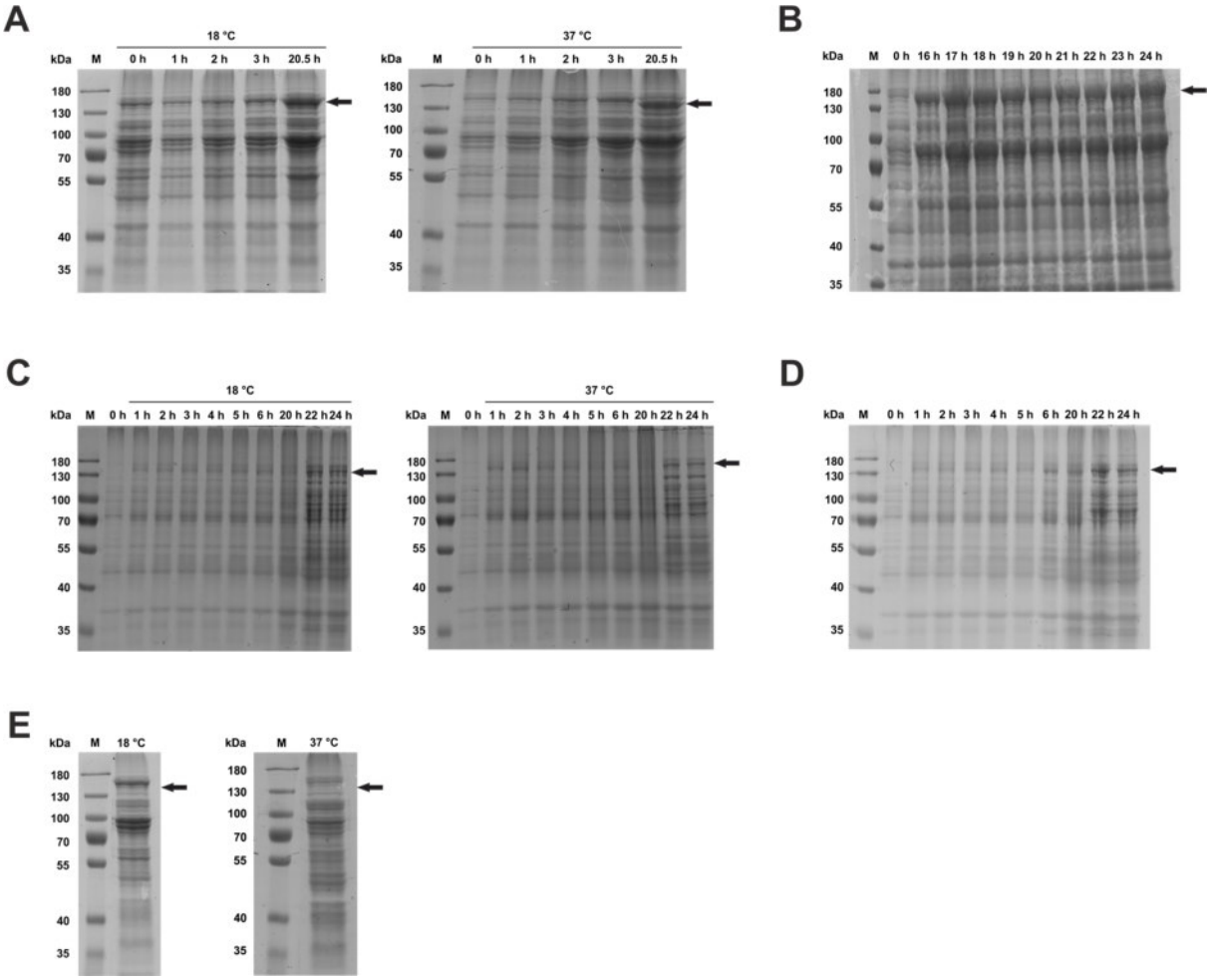


Figure 57: Expression test of AtKRP125b. Colloidal coomassie-stained 10 % SDS-PAGE gels. **(A)** Expression test of BL21 Star™ (DE3) cells containing the pGEX-6P-1-AtKRP125b plasmid. Expression was induced with 1 mM IPTG at an $OD_{600\,nm}$ of 1 **(B)** and a more detailed and prolonged time series of the same conditions. **(C)** Further expression test of BL21-CodonPlus (DE3) RIPL cells containing the pGEX-6P-1-AtKRP125b plasmid. In those cells the expression was induced with 0.5 mM IPTG at an $OD_{600\,nm}$ of 0.5. **(D)** Additional expression test at 18 °C with an induction at $OD_{600\,nm}$ of 1 and 1 mM IPTG. **(E)** Expression test of Origami™ 2 (DE3) pLysS cells containing the pGEX-6P-1-AtKRP125b plasmid. Expression was induced with 1 mM IPTG at an $OD_{600\,nm}$ of 1. Samples were taken after 17 h. Arrows depict the theoretical size of GST-AtKRP125b with 140.8 kDa.

Consequently, GST-ATKRP125b was expressed for all further experiments using BL21-Star™ (DE3) cells that were incubated for 16 h at 18 °C. Induction was carried out with 1 mM IPTG at $OD_{600\,nm}$ 1. In a next step, a solubility screen was performed to detect a lysis buffer (Table 23). A suitable buffer ensures that AtKRP125b is soluble and purifiable. A signal, corresponding to GST-AtKRP125b, was detected in supernatant samples 8 and 14 (Figure 58). Although, the detected signals were quite weak. Thus, different variants with varying lysis conditions were tested in an

additional screen to improve the solubility. This test showed no improvement over the previous one (Figure 59). Consequently, GST-AtKRP125b seemed to be approximately 50 % soluble in buffer conditions with 100 mM Tris-HCl pH 8.0, 10 mM NaCl, 1 mM DTT and 1 mM ATP.

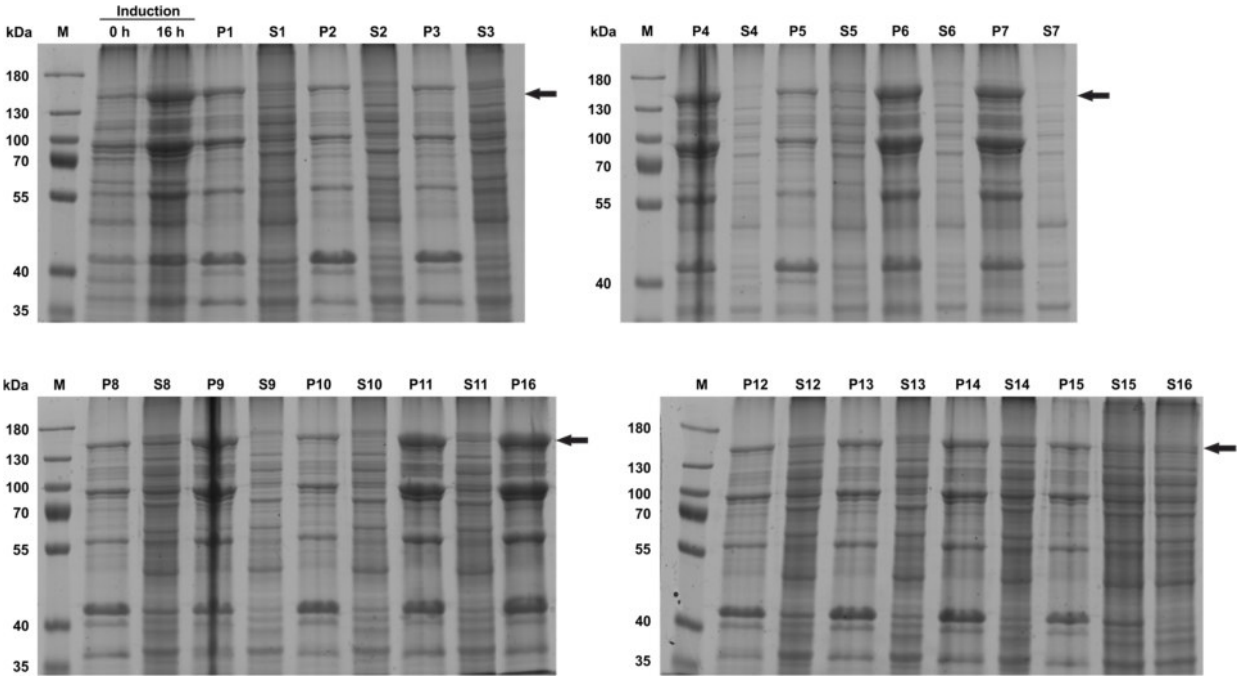


Figure 58: First solubility screen of GST-AtKRP125b. Analysis of a solubility screen of recombinantly expressed GST-AtKRP125b using colloidal coomassie-stained 10 % SDS-PAGE gels. Pellet (P) and supernatant fractions (S) of tested buffers (Table 23) are always next to each other. The expected size of GST-AtKRP125b is around 140.4 kDa (arrow).

Purification of AtKRP125b

Using this buffer an affinity chromatography was performed. Therefore, a self-packed Econo-Column® containing Glutathione Sepharose™ 4B beads was chosen. Afterwards all fractions were analyzed by SDS-PAGE. Signals equivalent to the size of the recombinant protein were detected in the supernatant, the pellet, the flow-through and the first wash fraction. No corresponding band was detected in any of the elution fractions (Figure 60). However, several additional bands corresponding to impurities were detected in all elution fractions. Consequently, it was not possible to purify GST-AtKRP125b using this method.

Bands resembling GST-AtKRp125b, which were detected in the flow-through and the washing fractions, corresponded in their intensity to the one detected in the supernatant fraction. Thus, the recombinant protein may not have bound to the matrix. In addition, the band of GST-AtKRP125b in the supernatant fraction was rather weak. Consequently, it would only have been possible to purify a very small amount of GST-AtKRP125b. Thus, the amount of GST-AtKRP125b in the elution fraction may have been too small to be detected by colloidal Coomassie staining. Most structural interactions studies rely on large quantities of protein [345]. Thus, a low yield of purified protein is very problematic. Reasons for a low binding affinity or low yield could be manifold. Aggregation or incorrect folding of GST-AtKRP125b could prevent a correct binding to the matrix. ATKRP125b is a rather large protein and could therefore mask interaction between GST and the matrix, if folded incorrectly. The size and complexity of the kinesin could as well be a problem in case of the expression, since incorrectly and aggregated proteins often result in so-called inclusion bodies, preventing a successful purification and rendering them insoluble [389].

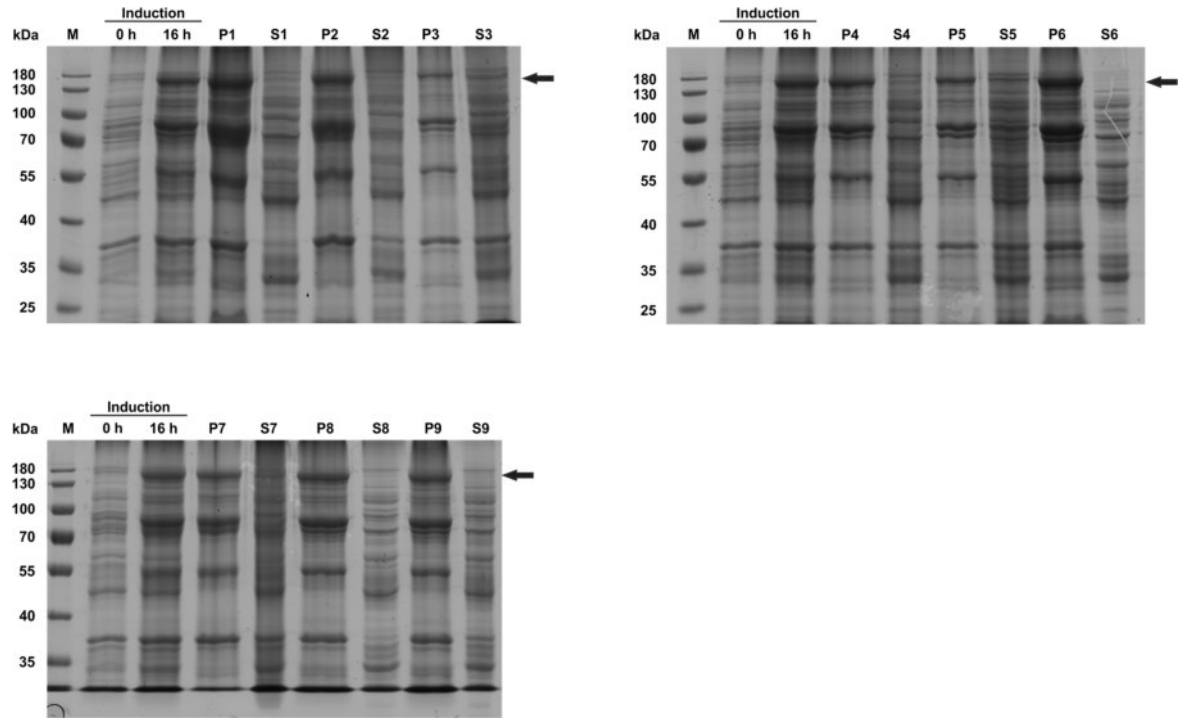


Figure 59: Second solubility screen of GST-AtKRP125b. Colloidal coomassie-stained 10 % SDS-PAGE gels of a solubility test of GST-AtKRP125b using nine different buffers (Table 24), which are variants of buffers 8 and 14 (Table 20). Pellet (P) and supernatant (S) samples of each buffer are always next to each other. Arrows depict the theoretical size of GST-AtKRP125b with 140.8 kDa.

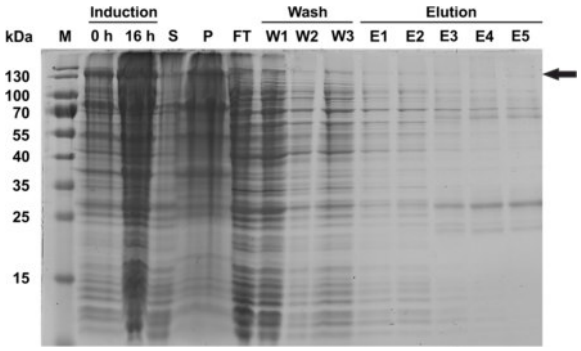


Figure 60: Affinity chromatography of GST-AtKRP125b. Analysis of the purification of recombinant GST-AtKPR125b using a colloidal coomassie-stained 10 % SDS-PAGE gel. Supernatant (S), pellet (P) and flow-through (FT). Beads were washed (W1-3) with 50 µM glutathione in buffer 3 (Table 24). Elution (E1-5) was performed with the same buffer containing 50 mM glutathione. GST-AtKRP125b has a theoretical size of 140.8 kDa (arrow).

For this reason, purifying GST-AtKRP125b in large quantities, did not seem to be feasible. Kinesins have always been reported to be not easy to purify. Their size and tertiary structure give rise to problems in bacterial expression systems [390]. These result often in missing modifications, leading to an incorrect folded protein or low yields. Consequently, a eukaryotic expression system, like baculovirus infected insect cells, would be more promising. This could result in a higher yield and properly folded proteins. In addition, impurities would be reduced. Establishing such a complex expression system is very time consuming. However, the main focus of the structural part of this work was the characterization of SAUL1 and its potential interaction partner BON1. Consequently, the purification of AtKRP125b was not pursued any further.

4.3.3 SAUL1 IN CONTEXT WITH CPK5

The predominant localization of SAUL1 is at the plasma membrane [110,113]. Hence, a potential interaction partner may reside there as well. Furthermore, SAUL1 is a positive regulator of PTI and is involved in MVB tethering and therefore exocytosis [137,140]. Therefore, an interaction partner could be involved in the same pathways. CPK5 (CACLIUM-DEPENDENT PROTEIN KINASE 5) acts as a calcium sensor and is known to also positively regulate plant immunity, by promoting cell-to-cell communication [31]. In addition, CPK5 was found to be located at the plasma membrane and interacts with a member of the plant exocyst complex [391]. Interestingly, the overexpression of *CPK5* results in an autoimmune phenotype that is dependent on the *TN2* gene [386]. The same dependency has been reported for overexpression of *SAUL1*, which is guarded by a heteromeric TN2-SOC3 complex, although not direct interaction takes place [138]. In addition, a direct interaction of TN2-CPK5 has been reported [386]. It was therefore speculated that SAUL1 and CPK5 act in the same protein complex. To investigate this potential interaction *in planta*, a BiFC experiment was performed. In this experiment a clear YFP-signal was indeed observed exclusively at the plasma membrane (Figure 61). This may indicate that both proteins interact directly or are located in the same protein complex at the plasma membrane. However, independent experiments such as Y2H or co-immunoprecipitation (Co-IP) are required for confirmation, because BiFC experiments are prone to false positive results [392].

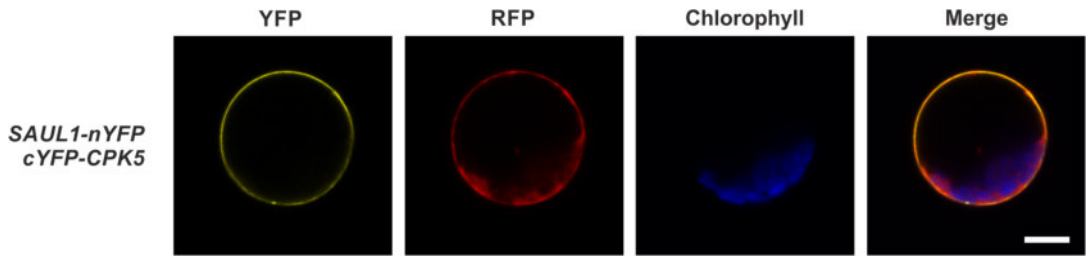


Figure 61: BiFC analysis of SAUL1 and CPK5. Confocal laser scanning microscopy of transiently transformed mesophyll protoplasts. Protoplasts were transformed with a BiFC construct containing *SAUL1* being paired with *CPK5*. *SAUL1* was tagged C-terminally, while *CPK5* was tagged at the N-terminus. YFP, RFP and autofluorescence signals of chlorophyll were recorded. Scale bars represent 15 µm.

Assuming that SAUL1 and CPK5 indeed interact, they may regulate each other in two different ways. SAUL1 may ubiquitinate CPK5 or CPK5 may phosphorylate SAUL1. Whereas there is no example for the ubiquitination of a CDPK, it has been shown that PUB25 and 26 are phosphorylated by CPK28 [117]. A phosphorylation of SAUL1 could suppress the tendency of SAUL1 to oligomerize and therefore regulate its activity (see chapter 4.2.5). Such a mechanism has been reported for PUB22, which is phosphorylated by MPK3 (MITOGEN-ACTIVATED PROTEIN KINASE 3) and subsequently activated due to the inhibition of its dimerization [116]. In addition, the assumption that SAUL1 and CPK5 interact supports another hypothesis, which has been discussed in this thesis (see chapter 4.2.5). SAUL1 has been connected to exocytosis by tethering MVBs to the plasma membrane and is hypothesized to reside in a complex which mediates immune-dependent exo- or endocytosis (see chapter 4.2.5 and 4.3.2) [140]. Remarkably, CPK5 is as well linked to exocytosis, because it has been reported that CPK5 phosphorylates Exo70B1, a member of the exocyst complex [386]. Thus, the activity of Exo70B1 is supposedly regulated by phosphorylation. Consequently, SAUL1 and CPK5 could be part of a complex, which is connected to the exocyst and regulate exocytosis during plant immunity. This is supported by findings, which could show that exocyst members are regulated by polyubiquitination and subsequent degradation [129].

The interaction of SAUL1 with CPK5 could therefore be the first step to explain how the activity of SAUL1 is regulated and should be investigated in subsequent experiments. Furthermore, this result strengthens the hypothesis that SAUL1 is involved in the regulation of exocytosis.

4.3.4 SAUL1 AND ITS REGULATORY INTERACTION PARTNERS CHS1 AND SOC3

The guard model is one of the hallmarks of plant immunity [393], and SAUL1 is guarded by the heteromeric NLR complex consisting of SOC3 and CHS1 [137,138]. The intense investigations on the activation of plant NLR proteins have been based on single-domain studies and homology models only [62,394]. By using cryo-electron microscopy, the structure of the CNL ZAR1 has been resolved very recently. These investigations revealed that ZAR1 may act as a pore in the plasma membrane during ETI [79]. Structural investigations therefore provided valuable insights into the mechanisms of ZAR1 in particular and hNLRs in general, which were not previously described. In addition, the structural context of an interaction between an NLR protein or complex and its guardee has never been successfully analyzed. Thus, understanding the interaction between SAUL1 and SOC3/CHS1 is of high interest [138,141]. The structural analysis would give valuable insights into the specific interactions of the heteromeric NLR complex with SAUL1. For gathering most detailed information, the expression in *Escherichia coli* (*E. coli*) and purification via GST-based affinity should be established. This would allow for the structural analysis of and binding studies on the recombinant proteins CHS1 and SOC3.

Expression and purification of CHS1 and SOC3

Both proteins were cloned into the expression vector pGEX-6P-1. This vector allows for an inducible expression of the respective protein using IPTG and a subsequent affinity chromatography via a GST-tag at the N-terminus. In case of CHS1 the construct was transformed into BL21 CodonPlus (DE3) RIPL cells, which were tested in an expression screen. When analyzing the results of the screen by using SDS-PAGE, a signal corresponding to the size of GST-CHS1 was detected in cultures, which were grown up to an $OD_{600\,nm}$ of 0.5 and induced with 0.5 mM IPTG (Figure 62A). The first apparent signal was found to arise between 2 and 3 h at 18 °C and as early as 1 h at 37 °C (Figure 62A). In an additional screen a higher $OD_{600\,nm}$ of 1 with an induction concentration of 1 mM IPTG was tested, which resulted in a strong GST-CHS1 signal, contrasting strongly with the background, after incubating the bacteria for 2 h at 18 °C (Figure 62B). The corresponding band was analyzed via MS, and the presence of GST-CHS1 in this band was confirmed (data not shown). Hence, for all further experiments these expression conditions were used.

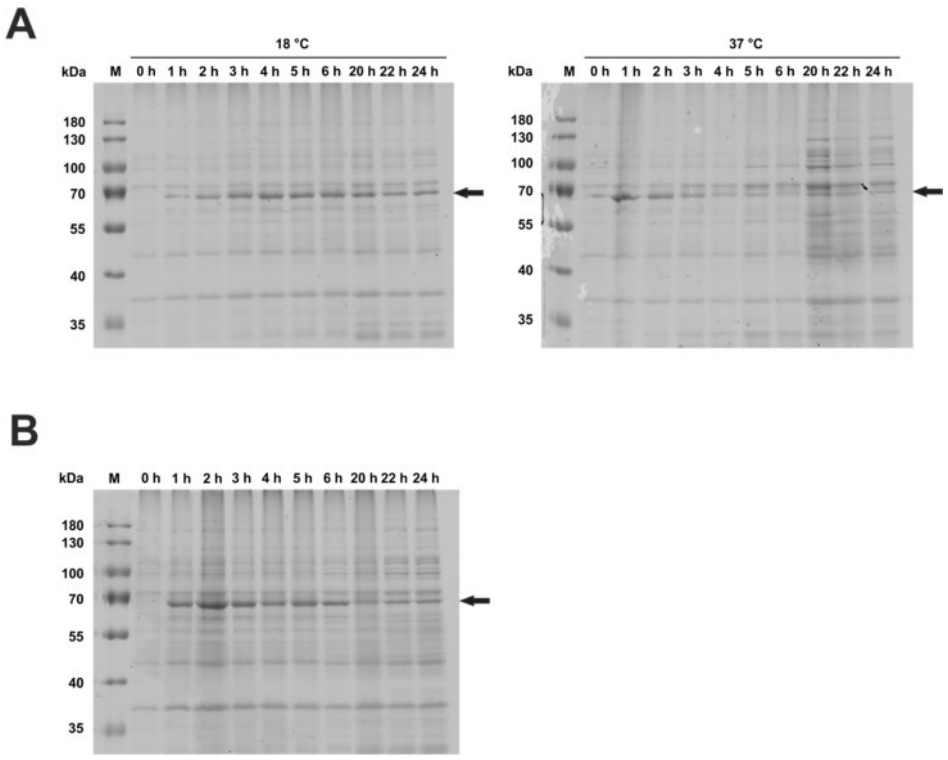


Figure 62: Expression analysis of GST-CHS1. SDS-PAGE analysis using colloidal coomassie-stained 10 % gels. **(A)** Expression test of proteins from BL21-CodonPlus (DE3) RIPL cells expressing pGEX-6P-1-CHS1, which were induced at $OD_{600\ nm}$ 0.5 with 0.5 mM IPTG and incubated at 18 °C and 37 °C. **(B)** BL21-CodonPlus (DE3) RIPL cells which were grown up to an $OD_{600\ nm}$ of 1 and induced with 1 mM IPTG and incubated at 18 °C. GST-CHS1 has a theoretical size of 74.4 kDa (arrows).

The generated construct for SOC3 was transformed into two different expression strains BL21-Gold (DE3) and BL21-CodonPlus (DE3) RIPL cells. Those experiments were conducted in context of the master thesis of Marcel Bhattarai [194]. During the expression screen the most promising results were obtained using BL21-Gold (DE3) cells, because no band in the correct size for GST-SOC3 was detected expressing the construct in BL21-CodonPlus (DE3) RIPL cells (data not shown) [194]. In case of an $OD_{600\ nm}$ of 0.6 and an induction with 1 mM IPTG, a band corresponding to the theoretical size of GST-SOC3 with 145.8 kDa was detected after 20 h incubation at 18 °C (Figure 63A). Nonetheless, the detected band did not stand out from the background bands corresponding to the basal expression of *E. coli* proteins. During the incubation at 37 °C no corresponding band was detected (Figure 63A). When testing a different $OD_{600\ nm}$ with a value of 1 and inducing the expression with 1 mM IPTG similar results were obtained. A weak band probably corresponding to GST-SOC3 was detected in case of an incubation at 18 °C and none in case of 37 °C (Figure 63B). Since the bands, which were proposed to correspond to GST-SOC3 did not stand out compared to the basal expression, GST-SOC3 was not expressed in those cells in a sufficient amount. Therefore, this approach was not feasible to be used for structural analysis. For this reason, subsequent experiments were focused on the expression and purification of CHS1.

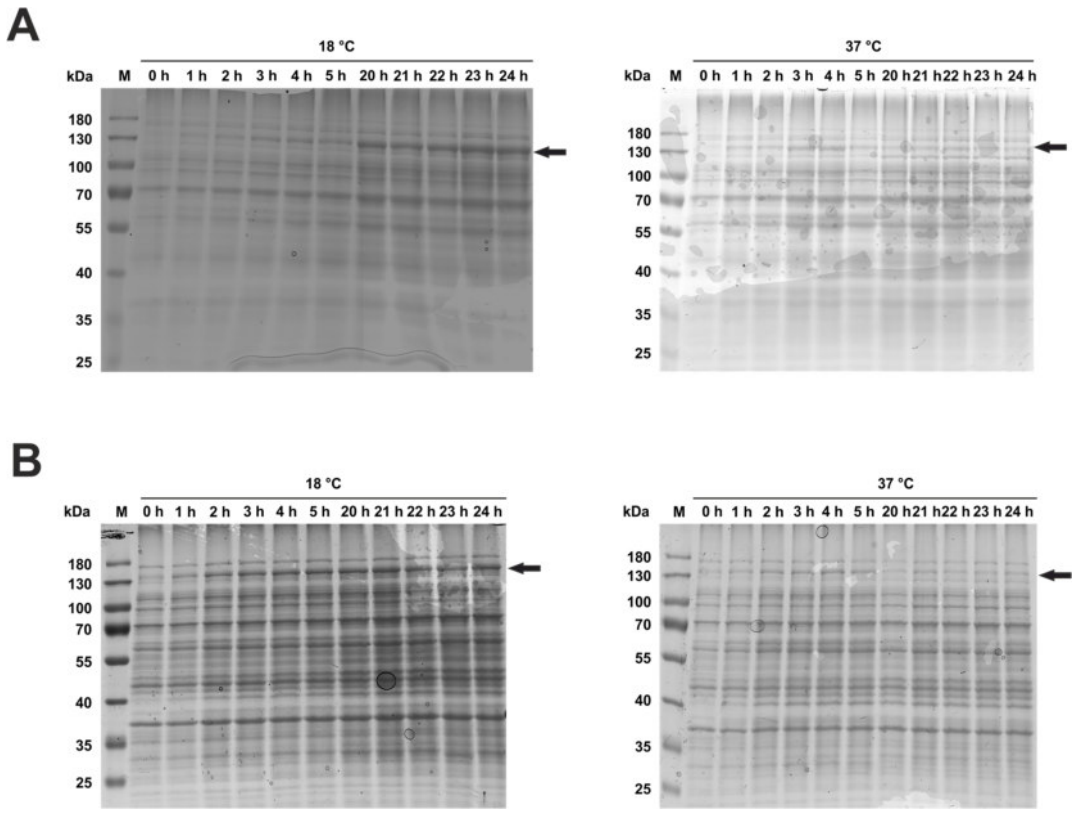


Figure 63: Expression analysis of GST-SOC3 [194]**.** SDS-PAGE analysis using colloidal coomassie-stained 10 % gels. Expression test of proteins from BL21-Gold (DE3) cells expressing pGEX-6P-1-SOC3, which were induced at **(A)** $OD_{600\,nm}$ 0.6 with 1 mM IPTG and incubated at 18 °C and 37 °C, or induced at **(B)** $OD_{600\,nm}$ of 1 and induced with 1 mM IPTG. GST-SOC3 has a theoretical size of 145.8 kDa (arrows).

In a next step, solubility assays were performed for CHS1. This allowed to test for a buffer with the best properties for purification of GST-CHS1. For that purpose, GST-CHS1 was expressed as stated before, and cells were pelleted and finally lysed using different buffers (Table 25). Soluble and insoluble fractions were separated by centrifugation. Afterwards supernatants and pellets were analyzed by SDS-PAGE. In many supernatant fractions, bands with a size of approximately 74.4 kDa, which would correspond to GST-CHS1, were detected. Thus, GST-CHS1 seemed to be soluble in multiple buffers (Figure 64A). To compare the amount and effectiveness of the lysis a Western blot of the supernatant fractions using a GST antibody was performed. This revealed, that the highest amount of GST-CHS1 was present in fractions 2 and 5 (Figure 64B). Since fraction 5 had the most distinct band, buffer 5 (Table 25) was used for all subsequent purification steps.

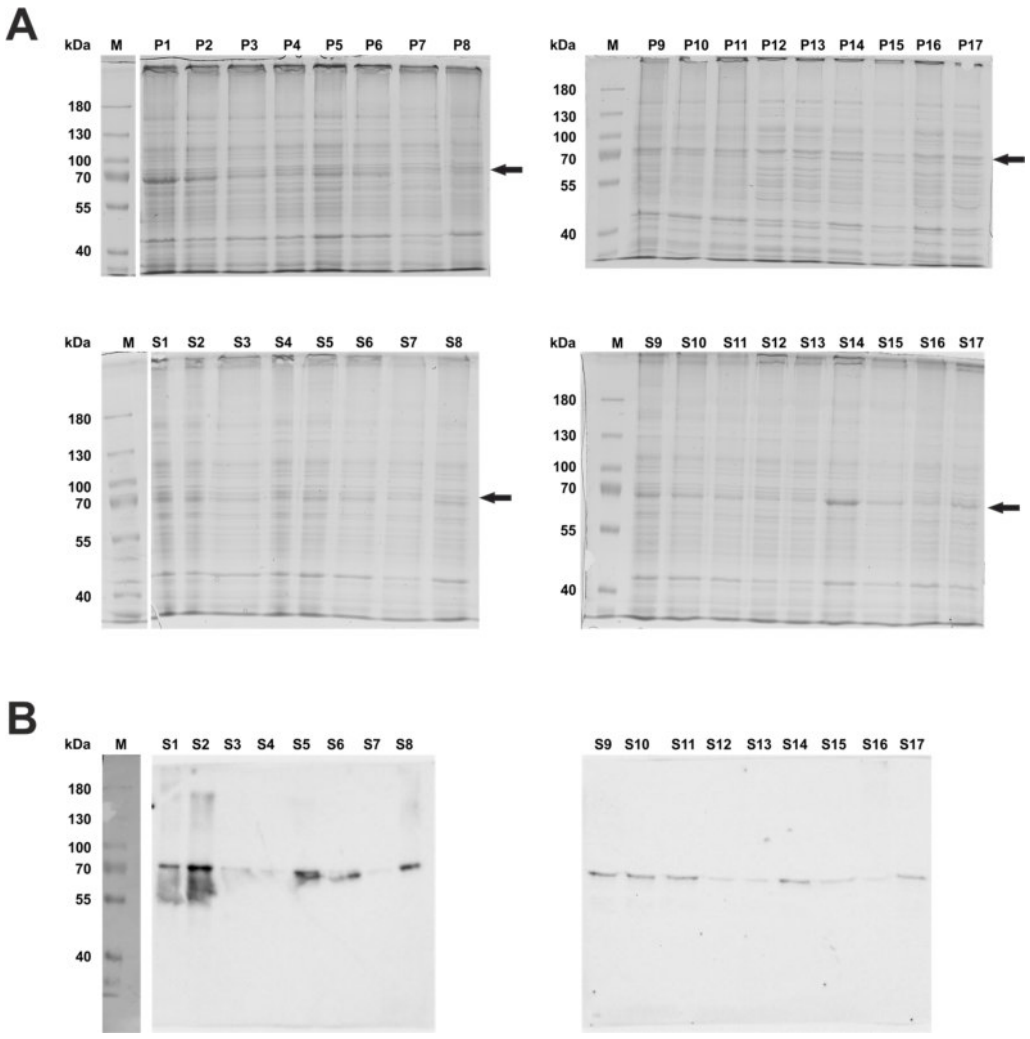


Figure 64: Solubility screen of GST-CHS1. Colloidal coomassie-stained 10 % SDS-PAGE gels. **(A)** Solubility Screen of GST-CHS1 expressed in BL21-CodonPlus (DE3) RIPL. Pellet (P) and supernatant (S) samples from 17 different buffers (Table 25) are depicted on different gels. Arrow depicts the theoretical size of 74.4 kDa of GST-CHS1. **(B)** Western blot of the supernatant fractions with *Anti-Glutathione-S-Transferase IgG* (Sigma-Aldrich, St. Louis, United States) used as a primary antibody.

In a next step, an affinity chromatography using Glutathione Sepharose™ 4B beads with an Econo-Column® was performed. A pellet from 200 ml cell culture was lysed in 5 ml of buffer 5 using lysozyme and sonication. The chromatography was executed using gravity flow, and all fractions were analyzed by SDS-PAGE. A clear band corresponding to GST-CHS1 was detected in the pellet fraction, with approximately 50 % of the protein remaining insoluble (Figure 65). Nonetheless, only a very faint band was detected in the flow-through and a bit stronger one in the first wash fraction. Bands corresponding to GST-CHS1 were detected in the elution fraction 2 to 5 (Figure 65). Additional bands were detected in all elution fractions, pointing towards the presence of some impurities. Accordingly, it was possible to purify GST-CHS1 via affinity chromatography. Nonetheless, as distinct bands for GST-CHS1 were detectable in the flow-through and wash fractions, only some parts of GST-CHS1seemed to have bound to the matrix and some parts were washed off, although no glutathione was used in the washing buffer. Thus, the binding affinity of GST-CHS1 was rather low. All fractions were pooled and the protein tag was cleaved off using the *PreScission™* protease.

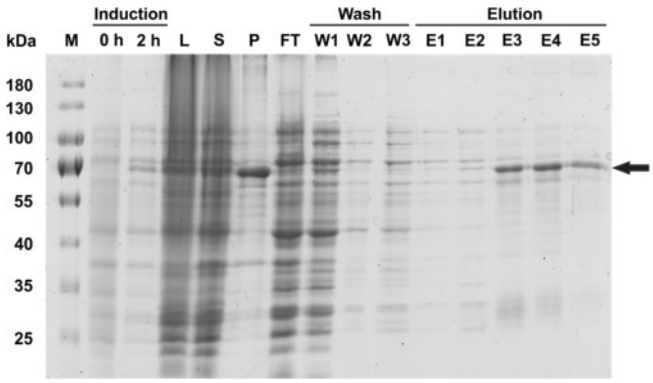
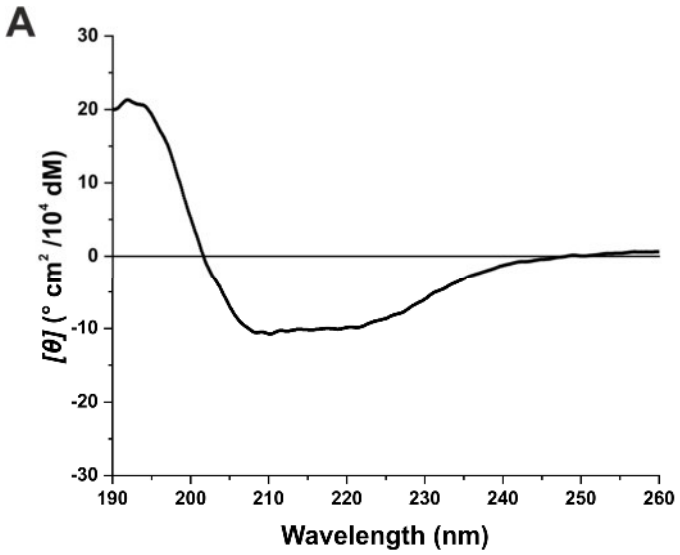


Figure 65: Affinity chromatography of GST-CHS1. Analysis of the purification of recombinant GST-CHS1 using a colloidal coomassie-stained 10 % SDS-PAGE gel. Lysate (L), supernatant (S), pellet (P) and flow-through (FT) samples. Beads were washed (W1-3) with buffer 5 (Table 25). Elution was performed with the same buffer containing 5 mM (E1, 2) and 50 mM glutathione (E3-5). GST-CHS1 has a theoretical size of 74.4 kDa (arrow).

Structural investigation of CHS1

To assess, whether the recombinant protein was correctly folded, CHS1 was thought to be measured via CD spectroscopy. In order to avoid interference of chloride ions with UV absorbance, the protein solution was dialyzed using a buffer containing NaFl. Afterwards the protein solution was concentrated up to 0.3 mg ml^{-1} in a volume of 200 µl. The obtained curve, depicting the normalized molar ellipticity was analyzed after Reed et al. (Figure 66A) [347]. The calculated secondary structure contained 34 % α-helices, 40 % β-sheets, 5 % turns, and 21 % of the protein were determined to be randomly organized (Figure 66B). Thus, CHS1 was present as a folded protein and not aggregated or intrinsically disordered. CHS1 as a truncated NLR protein, which is lacking the LRR domain, has been reported to consist of solely the N-terminal TOLL-INTERLEUKIN-1 RECEPTOR (TIR) and the C-terminal nucleotide-binding (NB) domain [395]. Therefore, as TIR and NB domains have been found to contain larger proportions of α-helices and β-sheets and less regions consisting of turns [79,396], the observed composition of CHS1 seems to be very likely.



B

	Percentage
α-Helices	34.0 %
β-Sheets	40.0 %
Turns	5.0 %
Random	21.0 %

Figure 66: CD spectroscopy analysis of CHS1. CD spectroscopy analysis of CHS1, being dialyzed against a CHS1-buffer 5 containing NaFl instead of NaCl. **(A)** CD spectrum of CHS1 depicted as molar ellipticity (*[θ]*). **(B)** Secondary structure analysis of CHS1. Analysis was done after Reed et al.[347].

Large scale purification of CHS1

Although the purification of recombinantly expressed CHS1 did start very promising, it was not possible to analyze CHS1 any further. Two major problems appeared to be the low concentration (Figure 62) and its insufficient solubility (Figure 64A). Only about 50 % of the GST-tagged protein appeared to be soluble, thus strongly reducing the amount of protein for purification. To overcome this problem extensive expression tests and solubility screens were performed. However, these did not result in a better yield (data not shown). GST-CHS1 seemed to be insoluble in most buffers. Despite those drawbacks, CHS1 should be used for structural analysis and therefore be purified using an automated approach with the *ÄKTA™* system for the affinity chromatography and SEC. This system would allow to overcome problems with the low yield, because larger amounts of cell lysate could be used, and the automated approach would help to reduce potential losses of protein. Therefore, an expression volume of 2 L was used for the automated affinity chromatography. When analyzing the chromatogram of the purification, only a very small peak with an absorption of 116 mAU corresponding to 0.21 mg ml^{-1} of protein was detected (Figure 67A). SDS-PAGE analysis revealed that nearly no GST-CHS1 had bound to the matrix, and a band corresponding to GST-CHS1 was mostly detectable in the flow-through fraction (data not shown). As a result, it was not possible to purify CHS1 using the automated approach. This could have been due to the fact that GST-CHS1 had not bound in a sufficient amount to the GSTrap™ 4B column. This may have resulted out of a combination of the previous observed low expression yield and low binding affinity. As a consequence, a large approach using a self-cast column was performed, because GST-CHS1 has been shown to bind partially to the Glutathione Sepharose™ 4B beads, which were used for the self-cast column (Figure 65). Surprisingly, this resulted in no detectable GST-CHS1 bands in the elution fractions (Figure 67B). Only in the flow-through and wash fractions corresponding bands were detected.

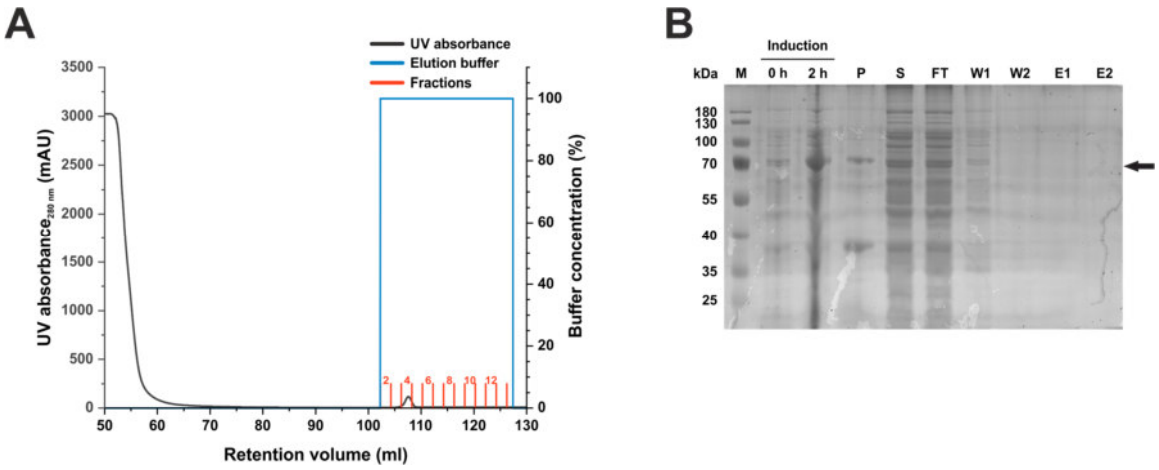


Figure 67: Automated affinity chromatography of GST-CHS1. (A) Affinity chromatography of GST-CHS1 using the *ÄKTA™ pure 25L* system. Chromatography starts with the incubation of the cell lysate on the column (absorbance higher than 2500 mAU) using a very slow flow rate. Afterwards the bound recombinant proteins were eluted using 50 mM glutathione in buffer 5 (blue curve, Table 25). **(B)** Analysis of the manual purification of recombinant GST-CHS1 using a colloidal coomassie-stained 10 % SDS-PAGE gel. Pellet (P), supernatant (S) and flow-through (FT) samples. Beads were washed (W1,2) with buffer 5 (Table 25). Elution was performed with the same buffer containing 50 mM glutathione (E1,2). GST-CHS1 has a theoretical size of 74.4 kDa (arrow).

Why the affinity of the matrix changed over time remains ambiguous. Particularly, since the beads were regularly cleaned using guanidine hydrochloride or NaOH according to the manufacturer's specifications. Thus, a reproducible affinity should have been given. Aggregation might have taken place, which could have resulted in a masked or altered GST being not able to bind any more to the affinity matrix. As a result, all undertaken purification experiments had problems with the

binding affinity, differing severely from the first purification. Consequently, it was not possible to purify recombinant CHS1 and analyze this NLR protein in the subsequent structural analysis.

As mentioned above, only one NLR protein, ZAR1, has ever been expressed as a full-length protein [79]. All other structural experiments used single domains of NLR proteins only [397]. The expression experiments on CHS1 showed that the most likely explanation for the absence of GST-CHS1 bands in the elution fractions was the low yield of soluble protein (Figure 62). Interestingly, it has been shown previously that the expression of NLR proteins in common expression strains turned out to be toxic for *E. coli* resulting in low expression even in case of a successful induction [398]. Thus, using a bacterial expression system does not seem to be a feasible approach to produce sufficient amounts of functional recombinant SOC3 and CHS1. Consequently, these NLRs are more likely to be successfully expressed using a eukaryotic expression system like Sf21 insect cells. In case of ZAR1 such approach led to considerable amounts of protein [79]. The change of the expression system may solve the solubility problem for GST-CHS1. Insoluble proteins often result from an incompatibility of the recombinant protein with the expression strain leading to aggregated proteins, which are stored in inclusion bodies. This would result in a huge amount of protein being insoluble [389]. Another very severe cause of observed insolubility could be that CHS1 and SOC3 were expressed on their own. NLR proteins are known to act as homo- or heterodimers and might stabilize themselves [10]. Without their interaction partners CHS1 and SOC3 could be prone to aggregation or misfolding, which would result in a low yield. Thus, a coexpression of both proteins might be a suitable option, especially, as in case of ZAR1 a coexpression with its interaction partner RKS1, resulted in a sufficient amount of recombinant protein [79]. Additionally, an exchange of the purification tag might be useful, as GST with its tendency to form dimers, could prevent a correct folding of CHS1 or SOC3 and could promote the formation of aggregates [399]. A suitable substitution would be a His-tag, combining a strong matrix-affinity with a small tag size. The tag size could be a great advantage, because such a small tag is less likely to interfere with the interaction of both NLR proteins. Another alternative purification tag could be MBP (MALTOSE BINDING PROTEIN), which can drastically improve the solubility of a target [400]. Accordingly, using such purification strategies seem to be very promising and could result in sufficient amounts of both NLR proteins, which can be used for further structural experiments.

4.3.5 INTERACTION BETWEEN SAUL1 AND ITS POTENTIAL BINDING PARTNER BON1

SAUL1 is an E3 ubiquitin ligase, which interacts with multiple E2 enzymes (see chapter 4.3.1) and has ligase activity [123]. Thus, it is most likely, that SAUL1 binds to a target to ubiquitinate it. Alike SAUL1 the putative target may have a function in the regulation of immune responses, eventually as a negative regulator of PTI at the plasma membrane that is marked for degradation by poly-ubiquitination through SAUL1 [137]. In *saul1* mutants this may result in the suppression of PTI, because degradation of the negative regulator is prevented. A suppression of PTI was indeed observed in the triple mutant *saul1-1 soc3-12 pub43-1*, which was infected with a *Pto* strain lacking the type III secretion system [137].

From BiFC experiments in *A. thaliana* protoplasts it was previously suggested that the negative regulator of plant immunity, BON1, interacts with SAUL1 [142,401]. Those results indicated that the interaction is mediated between the U-box and putative ARM repeats 1-6 of SAUL1 and the two protein kinase C conserved regions 2 (C2) of BON1 [142]. To confirm this interaction and in particular to receive further insights into the nature of the interaction between SAUL1 and BON1 further experiments were required. As function follows form, a lot of information on a protein and on interactions between proteins can be obtained from structural analyses. Hence, BON1 was thought to be analyzed structurally to obtain valuable data on its function and its interaction with SAUL1.

Expression and purification of recombinant BON1

As a first step, BON1 needed to be purified as a recombinant protein. This work and the initial structural measurements were performed in the context of the master thesis of Marcel Bhattarai [194]. As a suitable purification tag GST was chosen. A pGEX-6P-1-BON1 vector was generated and transformed into BL21 Star™ (DE3) cells, followed by expression analysis using different IPTG concentrations and temperatures (Figure 68). The most distinct band, corresponding to GST-BON1 with a size of 90.3 kDa, was observed after 22 h of incubation at 18 °C. Expression was induced with 0.5 mM of IPTG at an $OD_{600\ nm}$ of 0.6. The presence of BON1 in the corresponding band was confirmed by MS (data not shown). As a result, these conditions were applied for all further GST-BON1 expression experiments.

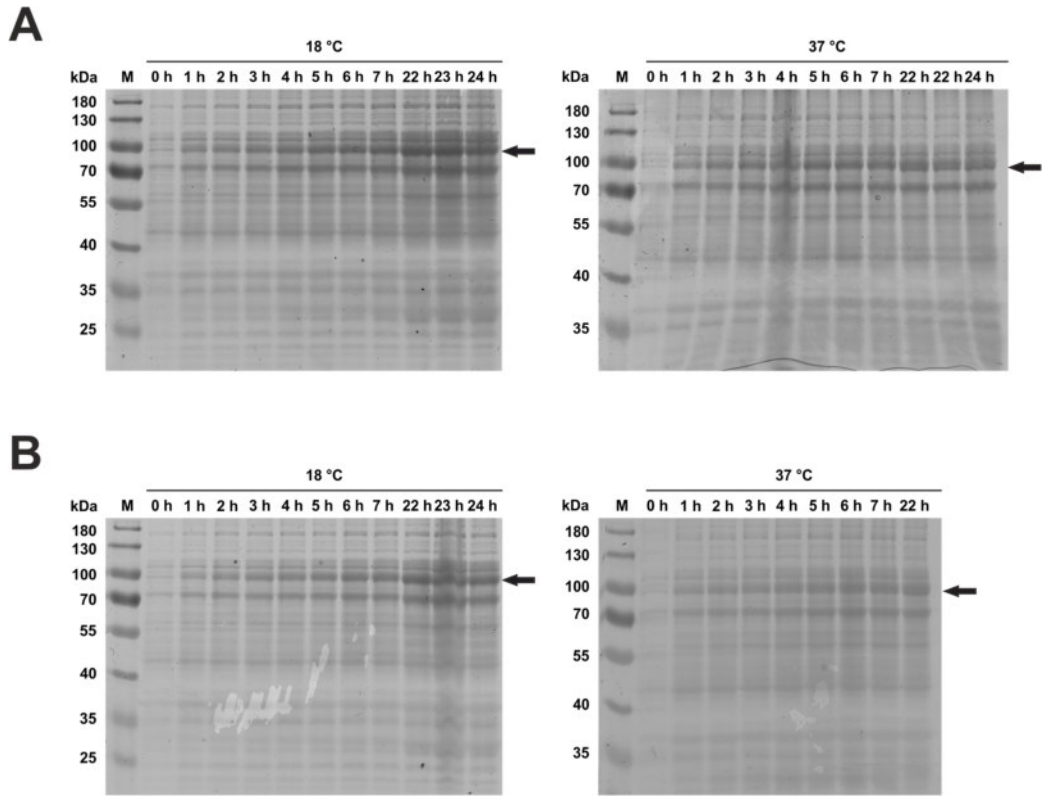


Figure 68: Expression analysis of GST-BON1 [194]**.** SDS-PAGE analysis using colloidal coomassie-stained 10 % gels. Expression test of proteins from BL21 Star™ (DE3) cells expressing pGEX-6P-1-BON1. Cells were induced at $OD_{600\ nm}$ 0.6 with **(A)** 0.5 mM IPTG and incubated at 18 °C and 37 °C **(B)** 1 mM IPTG and incubated at 18 °C and 37 °C. Arrows depict the theoretical size of 90.3 kDa of GST-BON1.

In a next step, a suitable lysis buffer for GST-BON1 had to be identified. For that purpose, a solubility assay was carried out employing a very broad range of buffers (Table 26). Multiple buffers were found, in which BON1 appeared to be soluble. Buffer systems containing only few ingredients, were favored for structural analysis, because ingredients may generally interfere with structural experiments, such as chloride ions with CD spectroscopy. For all further experiments, buffer 17 was selected on the basis of its simplicity consisting of 100 mM Tris-HCl at pH 7.5 with 200 mM NaCl and the additives 5 mM EDTA and 0.1 mM AEBSF (Table 26).

The following purification was performed using an automated approach with the *ÄKTA™* system for the affinity chromatography and SEC. By sonication 2 L expression culture were lysed in 50 ml of buffer 17. Afterwards, cellular debris was pelletized and the clarified lysate was used for the affinity chromatography. The lysate was applied to a *GSTrap™ 4B* column with a flow-rate between 0.05 and 0.1 ml min^{-1}. After washing the column and eluting any potentially bound proteins, a clear peak with an absorption maximum of 1332 mAU was detected in fractions 2 to 4 (Figure

70A). The analysis of these fractions by SDS-PAGE revealed a clear band corresponding to GST-BON1 with minor contaminations being present (Figure 70C). To remove the GST-tag by proteolytic digestion with the *PreScission™* protease, all three fractions were pooled and concentrated up to a volume of 2 ml. This reduction was necessary, because BON1 and GST appeared to be hardly separable, due to the tendency of GST to form dimers, which with a size of 52.8 kDa were very similar to the size of BON1 with 63.9 kDa. However, a separation of BON1 and GST was required for a better resolution in SEC. This was achieved by using a smaller loading volume and in addition a minimum flow-rate of 0.1 ml min^{-1}. Two separate peaks could be detected in the chromatogram, the first corresponding to BON1 and the second corresponding to GST (Figure 70B). An optimal separation of both proteins was not possible due to diffusion taking place at such low flow-rates. Very minor amounts of GST were detected in the fractions 24 and 25 corresponding to BON1. Nonetheless, using this approach BON1 was isolated to a purity of approximately up to 96 % and could therefore be used for all further structural and interaction experiments.

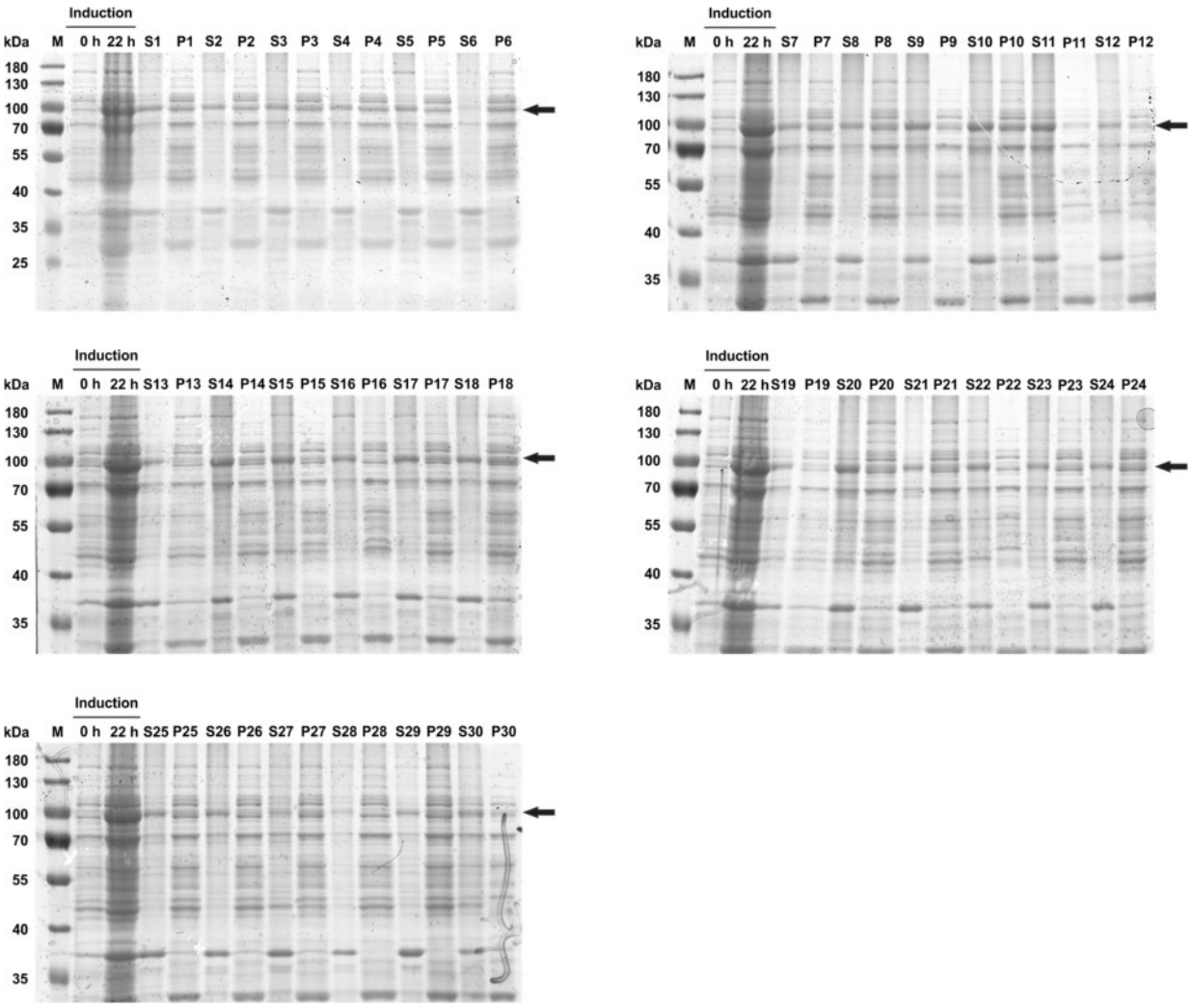


Figure 69: Solubility screen for GST-BON1 [194]**.** Colloidal coomassie-stained 10 % SDS-PAGE gels. Solubility Screen of GST-BON1 expressed in BL21 Star™ (DE3) cells. Supernatant (S) and pellet (P) samples from 30 different buffers (Table 26) are depicted on different gels. Arrows depict the theoretical size of GST-BON1 with 90.3 kDa.

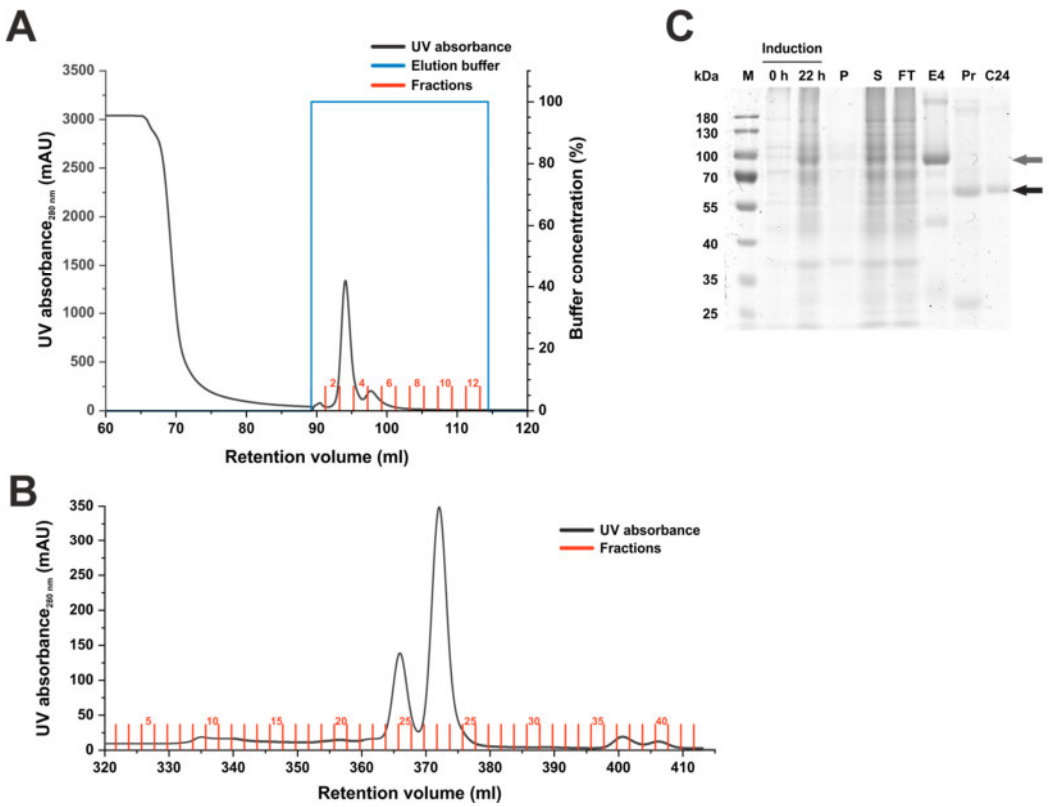


Figure 70: Purification of BON1. (A) Affinity chromatography of GST-BON1 using the *ÄKTA™ pure 25L* system. Chromatography started with the incubation of the cell lysate on the column (absorbance higher than 2500 mAU) using a very slow flow rate. Afterwards the bound recombinant proteins were eluted using 50 mM glutathione in buffer 17 (blue curve, Table 26). **(B)** Size exclusion chromatography of digested BON1 using the *HiLoad™ 16/600 200pg* column and a sample volume of 2 ml. **(C)** Analysis of the BON1 purifications using a colloidal coomassie-stained 10 % SDS-PAGE gels. Pellet (P), supernatant (S), flow-through (FT), elution (E4), GST-tag removal using the *PreScission™* protease (Pr) and size-exclusion chromatography fraction (C24). GST-BON1 has a theoretical size of 90.3 kDa (dark grey arrow), BON1 of 63.9 kDa (black arrow).

To confirm that BON1 was correctly folded and not aggregated, DLS and CD spectroscopy measurements were performed. DLS measurements revealed, that BON1 was stable for at least seven days, and no change in the distribution of radii was detected at a concentration of 5.2 mg ml^{-1} (Figure 71B). In addition, different concentrations were measured and revealed no change of the distribution of radii for concentrations up to 10.2 mg ml^{-1} (Figure 71A) [194].

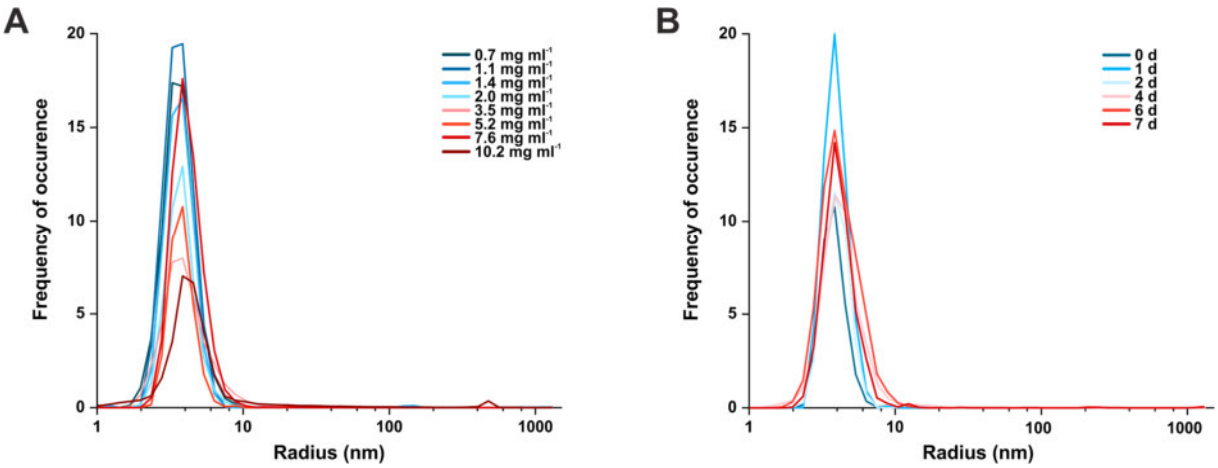
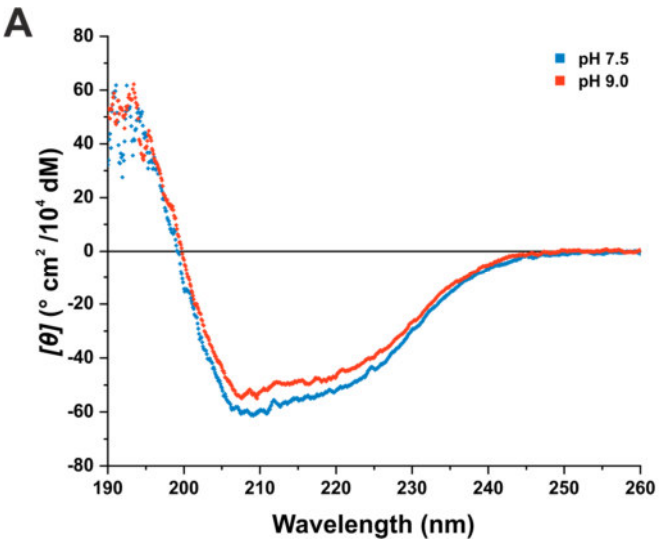


Figure 71: DLS Analysis of BON1 [194]**. (A)** Measurements of solutions with different concentrations of BON1 from 0.7 mg ml^{-1} to 10.2 mg ml^{-1}. **(B)** Time-course measurements of purified BON1 with a concentration of 5.2 mg ml^{-1} over multiple days.

CD spectroscopy was used to analyze whether BON1 was actually present in a folded state or as an unstructured protein. The obtained curve indicated that the signal-to-noise ratio was a bit low, because the protein solution was diluted 1:25 with ddH_2O, to avoid absorbance effects of chloride ions in the buffer (Figure 72A, blue curve). Nonetheless, it was possible to calculate the secondary structure, which was done after Yang et al. [402]. This resulted in a distribution of 18.7 % α-helices, 40.2 % β-sheets, 9.0 % turns and 32.1 % randomly structured regions (Figure 72B). BON1 has been reported to consist of two C2 domains, which are composed of β-strands, and one von Willebrand factor, type A domain (VWA) composed of α-helices surrounding the β-sheets. The arrangement of these three domains may be well represented by the determined distribution. The turn and random regions might be the linkers, which connect these domains. Apparently, BON1 was present in a native configuration, because such values correspond to a folded and not to an unstructured protein. Thus, BON1 was successfully purified, and the recombinant protein could be used for all further structural and interaction experiments.



B

pH	pH 7.5	pH 9.0
α-Helices	18.7 %	21.7 %
β-Sheets	40.2 %	37.2 %
Turns	9.0 %	9.1 %
Random	32.1 %	31.9 %

Figure 72: CD spectroscopy analysis of BON1. CD spectroscopy analysis of BON1, being diluted 1:25 in ddH_2O resulting in a concentration of 0.12 mg ml^{-1}. **(A)** CD spectra of BON1 in buffer environments with differing pH values depicted as molar ellipticity ($[\theta]$). **(B)** Secondary structure analysis of BON1 in buffers with different pH values. Analysis was done after Yang et al. [402].

Structural analysis of BON1

The structure of BON1 and its interaction with SAUL1 should be investigated by SAXS measurements. This allowed to account for potential flexibility and to investigate the interaction in a suitable buffer environment by using SEC-SAXS. The first step was to determine the singular BON1 structure. For that purpose, BON1 was measured in a SAXS batch experiment in a concentration range from 1.32 mg ml^{-1} to 7.5 mg ml^{-1}. The scattering curves of the different concentrations were compared to study the occurrence of concentration dependencies. Differences and changes in R_g values were not detected (data not shown), and the highest concentration with the highest signal-to-noise ratio, was used for all further analyses (Table S11A). However, small fractions of GST were detected in the BON1 fractions by SDS-PAGE, which were analyzed in the SAXS experiments [194]. As already described, this was due to the SEC that was performed beforehand (Figure 70). Nevertheless, no GST could be detected during the DLS (Figure 71). This was most likely due to its size, as GST itself, even as a dimer, is smaller than BON1. Since particles contribute to the detected scattering intensity with a magnitude which is proportional to the concentration and the volume squared, larger particles mask smaller ones. Therefore, a GST contamination would be impossible to detect by DLS. Nonetheless, despite this contamination the obtained data is still perfectly interpretable for the following reasons. On the one hand, the rule of scattering intensity being proportional to the volume of a molecule is applicable to SAXS measurements [345]. Hence, scattering of BON1, which was purified up to 96 % [194], is not affected by the small contamination of GST. On the other hand, multiple testing for a potential contamination was included during data analysis. This was done by fitting GST or mixtures of BON1 and GST to the experimental data. This testing did not show any indication for disturbances by contamination, because fitting of GST alone or a mixture of BON1 and GST to the scattering did not at all correspond to the experimental data (data not shown). In addition, experimental data obtained from batch measurements was very much alike to data obtained from SEC-SAXS measurements. In SEC-SAXS measurements only BON1 was measured due to dilution and separation effects. Since the data derived from SEC-SAXS and batch measurements are so similar, this means that the scattering data obtained from the batch measurements only correspond to BON1.

The log-linear plot of $I(s)$ versus s depicted a flattened sigmoidal-like decrease of $I(s)$ to higher angles. This would correspond to an overall globular shaped protein, which is just in some small parts unfolded or flexible (Figure 73A). The minor upturn at very low angles (s < 0.05 nm^{-1}) in the Guinier plot could arise from a slight aggregation (Figure 73B). Although no aggregation was detected in the DLS measurements, this could be the case nonetheless, because SAXS is much more sensitive than DLS (Figure 78C). The R_g was determined to 3.56 ± 0.37 nm using the Guinier approximation, which resulted in a reasonable fit ($\Delta/\sigma < 2$) of the linear range of the scattering data. Based on the Guinier fit, smaller angles were neglected due to reasons discussed before (Table S11C). To analyze the overall shape, a dimensionless Kratky plot was applied. This showed a right-shifted Gaussian-shaped curve with a maximum around 2.03 (Figure 73C). This correlated with a slightly globular protein, which was a bit flattened or elongated [349]. In comparison, total globular proteins would exhibit a maximum around $\sqrt{3}$. In addition, BON1 seemed to be quite rigid and to exhibit only small signs of flexibility, because the intensities flattened towards higher angles. Therefore, BON1 was present as a rigid globular, slightly flattened or elongated protein in the tested solutions.

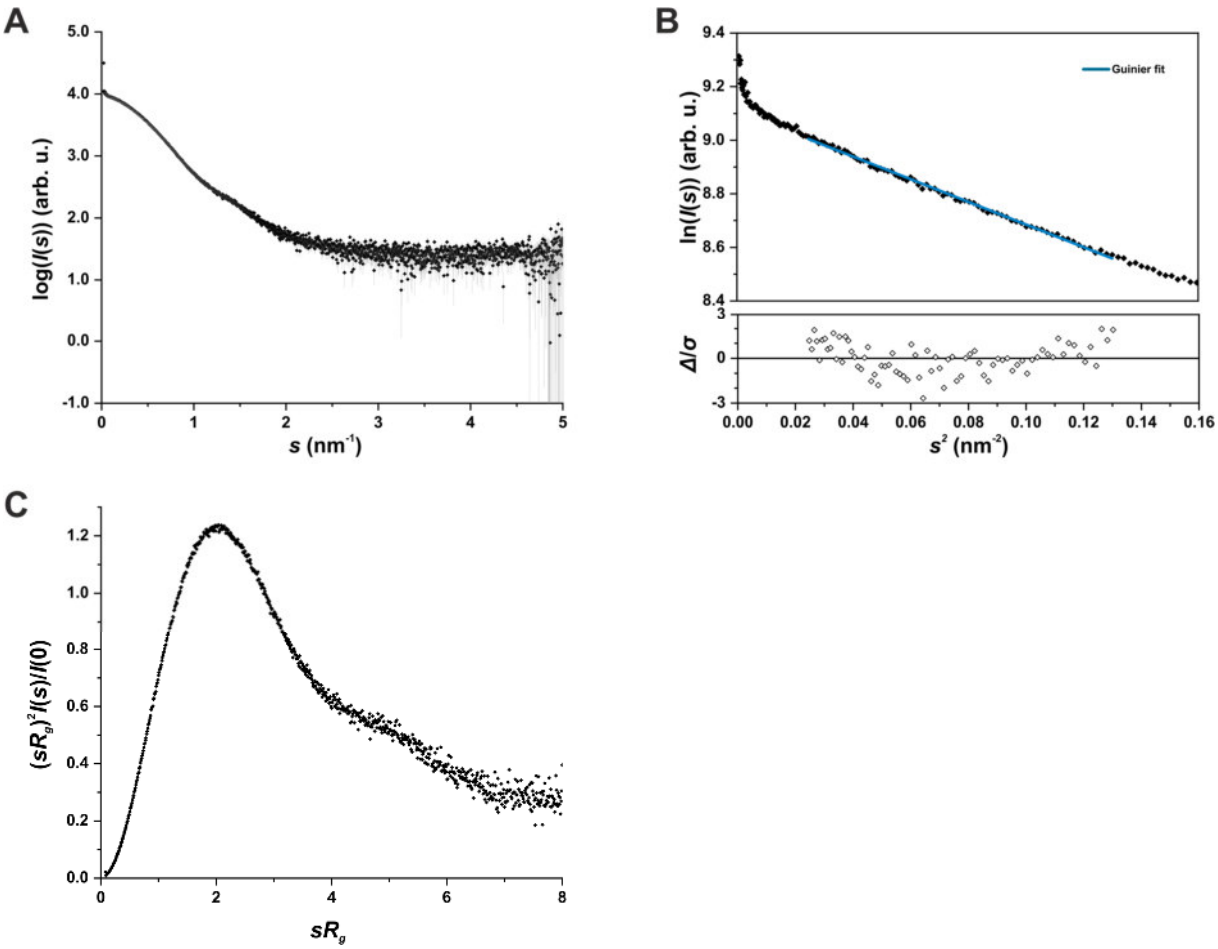


Figure 73: Scattering data of BON1 in a batch experiment. (A) Two-dimensional Log-linear plot of $I(s)$ versus s. **(B)** Upper plot shows the Guinier plot of the SAXS data. In blue the Guinier fit in a range of s_{min} = 0.1611 nm^{-1} and sR_g max = 1.28 (dotted lines) is depicted. Lower plot shows the standardized residual plot. **(C)** Dimensionless Kratky plot with the intensities normalized to the forward scattering intensity ($I(0)$) and the radius of gyration (R_g).

The distance distribution was calculated by indirect Fourier transformation and showed a nearly Gaussian-shape with a peak at 3.49 nm. In addition, a small proportion of longer distances between 8 nm and the maximum dimension (d_{max}) of 14.7 nm was detected (Figure 74A). This corresponded well to a mostly globular protein, which exhibits a small elongated protrusion. The fit to the experimental data had for most part standardized residuals in a range of 2. The fit differed more from the data at low angles (Figure 74B). In summary, the fit does represent the scattering data very well. The molecular weight of BON1 was with the help of a BSA standard determined to 65.9 kDa (Table S11C), which was with a factor of 1.03 very close to the predicted molecular weight of 63.9 kDa. Taken together, the experimentally data indicated that BON1 is present in solution as a monomer.

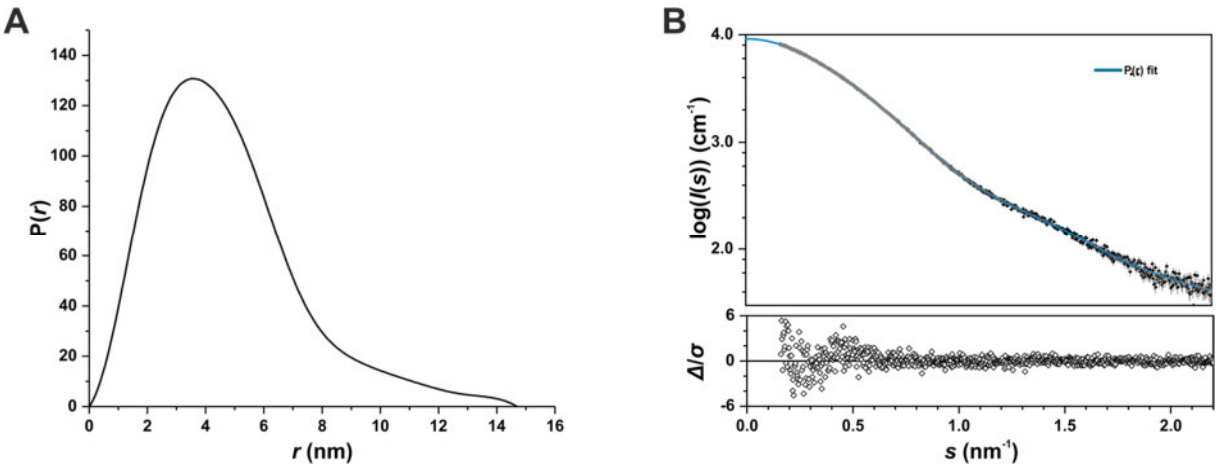


Figure 74: Distance distribution of BON1 in the SAXS batch experiment. (A) P(*r*) versus *r* profile of BON1. **(B)** Fit of the P(*r*) function (blue) to the experimental SAXS data. Lower plot depicts the standardized residual plot.

In a next step, an *ab initio* model of BON1 was derived. The *DAMMIF* program was used and resulted in a dummy atom model. This model represented the theoretical volume of BON1 and was generated 20 times. The model with the lowest NSD value of 0.854 was selected (Figure 75A). The standard deviation of the 20 NSDs was 0.121, which means that all models did not differ severely from each other and are therefore very probable. Thus, the calculated dummy atom model is most likely representing BON1. In a next step, the *ab initio* model was compared to the experimental data. Therefore, a theoretical scattering curve was calculated for the model using *CRYSOL* (Figure 75B, black line), which was compared to the scattering data. The fit of the model described the data well with standardized residuals mostly below 2 and a low χ^2 value with 1.094. The model demonstrated bowl- or hand-like shape, with a deepening in the middle. On one end was a small bulky protrusion, which was connected to the main structure via a small linker (Figure 75A).

To gain a more detailed view on the structure of BON1 an *in silico* model was calculated. A homology-based model was generated using *I-TASSER*. At the same time flexibility and domains were analyzed, and potential conserved regions in BON1 were elucidated. The BON1 protein contains a VWA and two C2 domains [143], which were found to be strongly conserved in the *InterPro* analysis (Figure 76C, Table S3B). Surprisingly, the *I-TASSER*-derived structure did not exhibit these domains and did not correspond to the CD spectroscopy analysis (Figure 72B), as the model consisted purely of α-helices. In addition, a *SREFLEX*-refined model did not represent the scattering data very well with a χ^2 value of 13.61 and did not fit to the *ab initio* model, because the detected protrusion was not present.

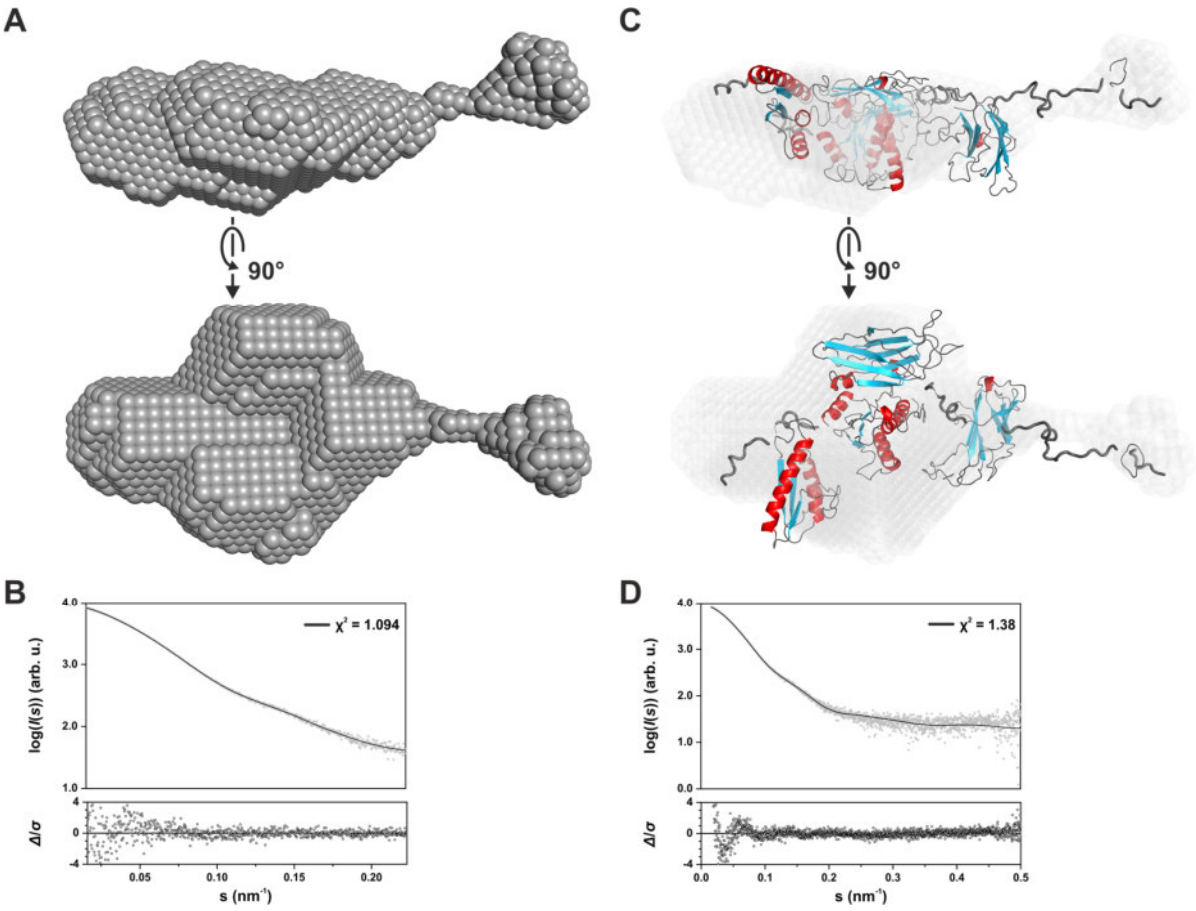


Figure 75: *Ab initio* and *in silico* structures of BON1 derived from batch SAXS measurements. (A) *Ab initio* bead model of BON1 with a resolution of 43 ± 3 Å and an NSD of 0.854. **(B)** Fit of the *ab initio* model to the raw scattering data with the standardized residual plot. **(C)** *ITASSA* derived in *silico model*, which was refined using *CORAL*. α-Helices are depicted in red and β-sheets in blue. **(D)** Fit of the *in silico* model to the raw scattering data with the standardized residual plot.

It was decided not to perform *EOM* analysis, because due to the dimensionless Kratky plot BON1 had a low flexibility. In view of three conserved and well-defined domains in BON1, a domain-based modelling approach was chosen. The BON1 protein was divided into two parts, one harboring the two C2 domains (amino acids 1-302) and one harboring the VWA (amino acids 321-586) domain. Both parts were used for homology-based modeling using *I-TASSER*. The obtained models did not directly fit to the scattering data (data not shown) and were thus divided into seven smaller domains with defined linker regions (Figure 76B, Table S3A). Those linkers were determined based on an elevated B-factor from the flexibility analysis, because this would point towards potential flexible regions in a protein, which might connect rigid domains. The parts were then used in a *CORAL* analysis, which allowed for multi-domain modeling. The obtained model had a χ^2 value of 1.38, representing a reasonable fit with mostly low standardized residuals (Figure 75D).

When analyzing the *in silico* structure itself, three domains connected by longer linkers were detected. They formed a more condensed region with the first β-sheet containing domain closer to the protrusion of the *ab initio* model. Whereas the β-sheet containing regions formed one side of the bowl-like structure, consisting of six β-sheets each, the other part was formed by the α-helices containing domain. The latter was composed of six α-helices, which surrounded six β-sheets in a Rossmann fold manner. This domain seemed to be divided into two regions, because one part of the α-helices was located next to the C-terminus and the other one more in the center. The protrusion itself was formed by the N-terminus. This was not unexpected, because the first amino acids 1-7 and 17-37 of BON1 were predicted to be quite flexible (Figure 76A). By and large,

this refined *in silico* structure fitted reliably to the experimental data and corresponded very well to the preceding CD spectroscopy experiments with 17.9 % α-helices, 22.1 % β-sheets and 59.8 % regions with turns or randomly organized amino acids (Figure 72B). Therefore, the derived model may represent the BON1 structure with a very high reliability.

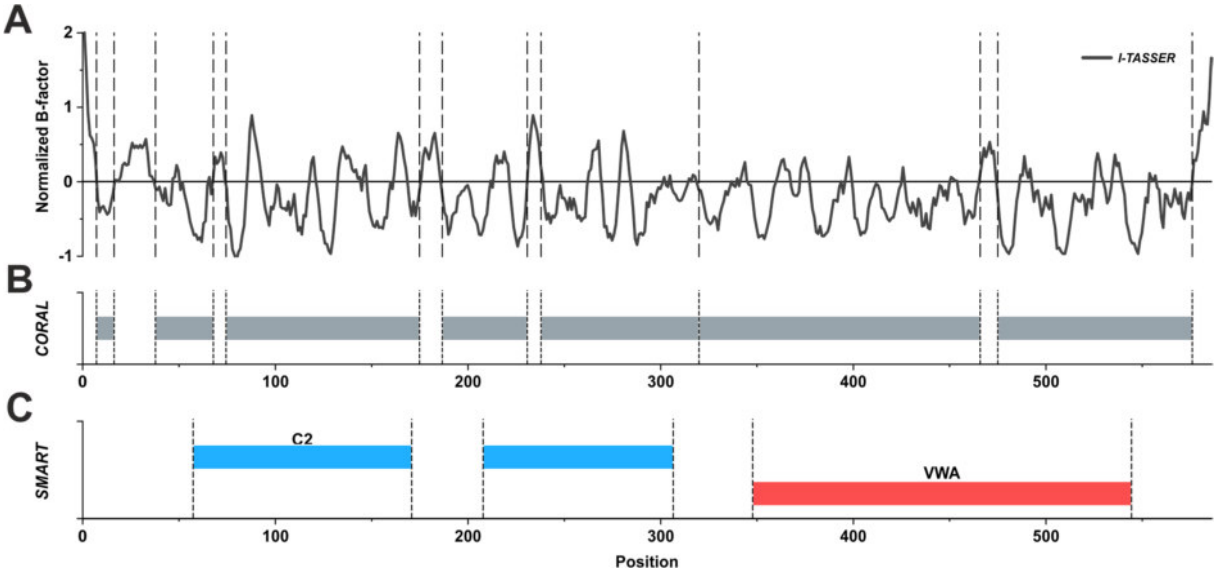


Figure 76 Flexibility and domain analysis of BON1. (A) The flexibility of BON1 was analyzed using *I-TASSER*. Depicted is the normalized B-factor. Values above zero indicate flexible residues and values below zero suggest a more rigid structure. An *InterPro* analysis allowed for protein domain predictions using **(C)** *SMART* domain databases **(B)** Based on these and the flexibility values rigid regions were defined for the *CORAL* analysis (Table S3A).

The regions that appeared to be structurally ordered reflected the predicted domain distributions. When analyzing the C2 domains it became apparent, that these are highly structurally conserved. They could be aligned to the crystal structure of the C2 domain of the protein kinase C (PKC) from *Rattus norvegicus* (Figure 77B). In this alignment, only slight changes in the linker regions in BON1 were visible, and the overall structure was nearly the same. As for the VWA, this domain unexpectedly appeared to be pulled into two parts, which were not forming one compact domain (Figure 77C). This resulted from the *CORAL* modelling that had considered one linker region in between this domain to be flexible. This was necessary to achieve a reasonable fit, because BON1 appeared to be not too condensed and a bit more flatly shaped. Nonetheless, this might not correspond to the real structure, because the sequence of the VWA domain was shown to be highly conserved. Thus, other linker regions might be more flexible or elongated in BON1, which would result in a more probable structure of BON1. However, a structure that fitted the data more efficiently could not be generated based on the SAXS data. In addition, the alignment of the second part of BON1 that was modeled by *I-TASSER* to the crystal structure of the A3 domain of the von human Willebrand factor indicated high accordance (Figure 77D). Thus, all three predicted domains seemed to be represented in the BON1 model and the overall flattened shape might be caused by linker regions and turns in between. This would mean that in a native environment with a pH of 7.5 the BON1 domains are not too densely packed and are lining up like pearls on a string.

The *ab initio* structure indicated that the VWA domain was opposite to the two C2 domains. They seemed to form a pocket that was also slightly visible in the dummy atom model (Figure 75A) and would allow interaction partners or ions to enter. In line with such a structure the VWA is known to bind divalent metal ions when involved in protein-protein interactions, which are mediated via a so-called MIDAS motif [144]. This motif normally appears in the reduced form Asp-x-Thr-x-Ser and is sufficient for ion binding [144,403]. Interactions between BON1 and BIR1 or between BON1 and BAK1 might indeed be mediated via the VWA [152].

Interestingly, the putative interaction between BON1 and SAUL1 has been reported to be VWA-independent. BiFC experiments could show that the N-terminal half, which carries the C2 domains, is essential and sufficient for this interaction [142]. This is not totally unexpected, because C2 domains are also well-known to be involved in protein-protein interactions. The most prominent example is the one of the human synaptotagmin I, a Ca^{2+}-dependent sensor in neurotransmitter release, which contains two C2 domains [145]. The C2A domain is known to interact with SNARE proteins in a Ca^{2+}-dependent manner [404]. This is especially interesting, because the C2B domain of BON1 shares a sequence similarity of 50.5 % with the C2A of synaptotagmin I. It may be hypothesized that the C2B domain of BON1 is interacting with SAUL1 in a potential Ca^{2+}-dependent manner. In addition, the C2B domain of BON1 was found to be highly related, with a sequence identity of 40 %, to the C2B domain of human copine 3, indicating nearly structural consistency. This C2B domain has been reported to be involved in Ca^{2+}-dependent membrane translocation [405]. Therefore, the C2B domain of BON1 might also be involved in membrane binding. Nonetheless, C2 domains can act in a bipartite way, as discovered for the C2A domain of synaptotagmin I, which does as well interact with lipid bilayers and is involved in protein-protein interactions [404,406].

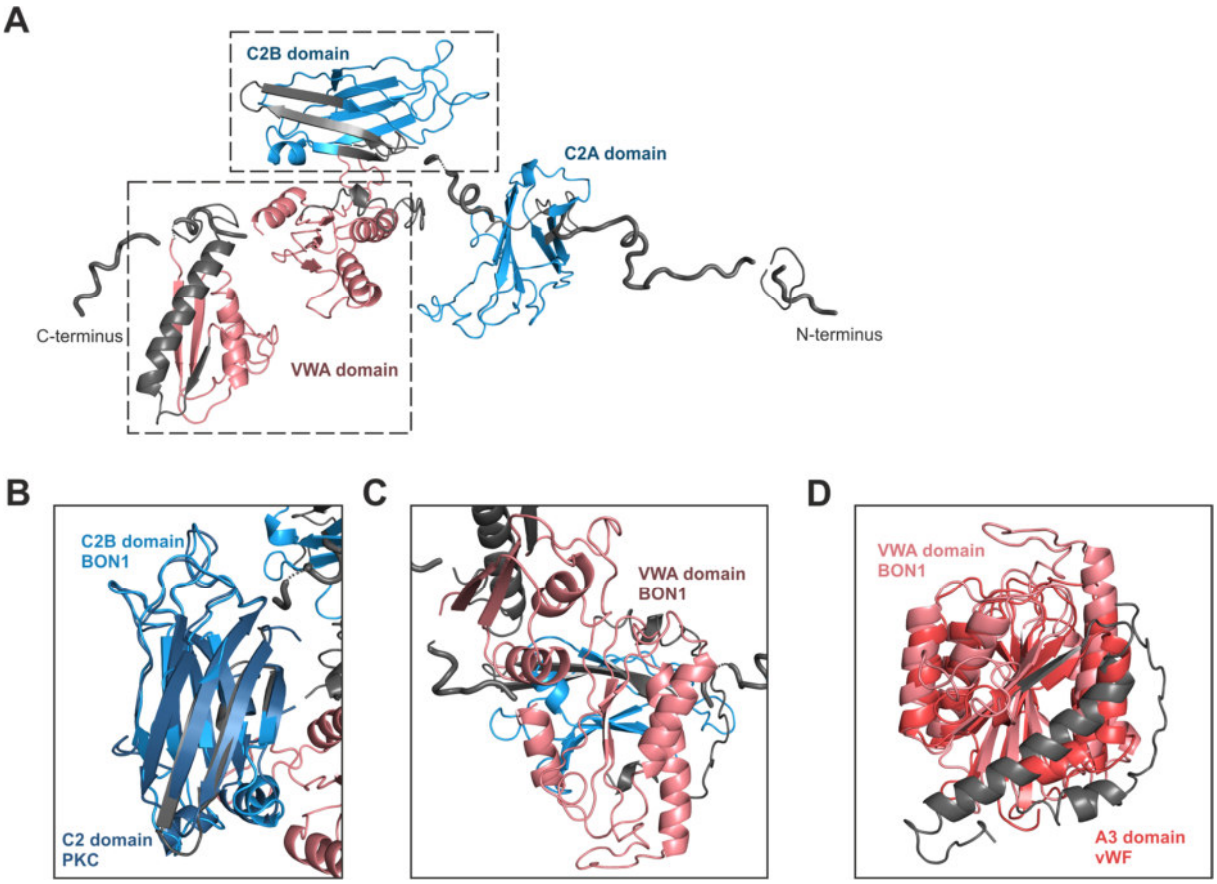


Figure 77:Domains of BON1 in SAXS batch experiment. (A) *In silico* structure of BON1 based on *CORAL*. Predicted and highlighted C2 (blue) and VWA domains (red). **(B)** Section of the C2B domain. Aligned is the crystal structure of the C2 domain from PKC (PROTEIN KINASE C) from *Rattus norvegicus* (*PDB*-ID: 1A25). **(C)** Section of the VWA domain. **(D)** *I-TASSER*-predicted VWA domain of BON1 (bright red) aligned with the A3 domain of the von Willebrand factor (vWF) from *Homo sapiens* (*PDB*-ID: 1AO3)

The C2A domain of BON1 shares a high similarity with the C2B domain of Munc13, with a sequence similarity of 40.1 %. Munc13 is involved in neurotransmitter release like synaptotagmin I. The C2B domain of Munc13 showed Ca^{2+}-dependent binding of phospholipids, thus localizing Munc13 to the plasma membrane [407]. Accordingly, the C2A domain that is located

on the outside in the determined structure (Figure 77A) may be important for the localization of BON1 at the plasma membrane [143]. This Ca^{2+}-dependent plasma membrane association through C2 domains appears to be a general feature of copines [405,408]. These interactions are generally mediated specifically via long linker regions between the β-sheets, comparable to those identified in the BON1 model [409,410]. Nevertheless, BON1 is thought to be localized to the plasma membrane primarily via a myristoylation at G2 [151]. However, it may well be that the C2A domain supports the interaction or serves as an additional protein-protein interaction site at the plasma membrane. Accordingly, mutations in the linker regions changed the function of BON1 [151]. Hence, the orientation and overall shape of the BON1 C2A domain seems to be quite plausible. Nonetheless, more movement might take place, because proteins are in general not rigid structures and SAXS gives only a broad overview of these.

Structural analysis of BON1 in the SAUL1 buffer

In addition to resolving the SAXS structure of BON1, this work was aiming at the characterization of its possible interaction with SAUL1. For that purpose, a mixture of both proteins should be measured by applying SEC-SAXS. Due to the instability of SAUL1, these measurements were thought to be realized using the SAUL1 purification buffer, which had a pH of 9.0. As a change in pH from 7.5 to 9.0 is rather severe, potential structural changes needed to be examined. Therefore, the same concentration of 4.93 mg ml^{-1}, which was thought to be analyzed in the interaction study, was measured beforehand in an individual SEC-SAXS experiment. The solution was proven to be monodisperse, as indicated by a singular and quite narrow peak in the performed DLS experiment (Figure 78C). During the chromatography, a singular peak with an intensity of about 1000 AU (Figure 78A) with a corresponding absorbance of about 200 mAU (Figure 78B) was detected. In the subsequent analysis, 25 frames of the detected signal and 50 buffer frames were used (Table S11B).

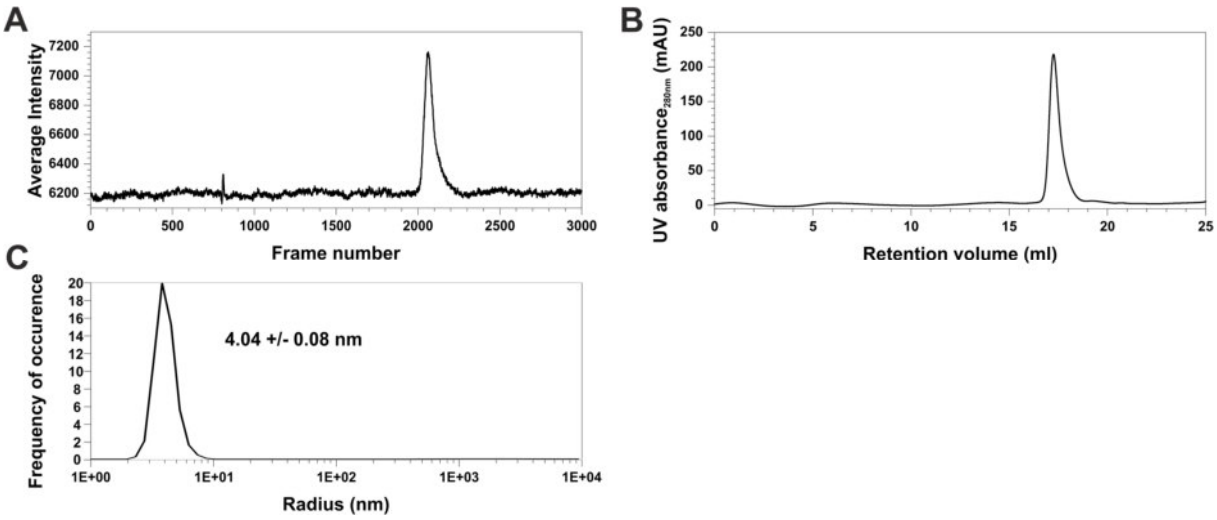


Figure 78: Chromatogram and radius distribution of BON1 in the SEC-SAXS experiment. (A) Scattering intensities of monodisperse BON1 with a loading concentration of 4.93 mg ml^{-1}. **(B)** Chromatogram of the UV absorption at 280 nm. **(C)** DLS measurements of BON1 exhibiting one singular peak with a detected radius of 4.04 ± 0.08 nm.

The log-linear plot of the intensity and scattering angle revealed a sigmoidal decreasing curve with a moderate signal-to-noise ratio. This shape would correspond to a globular, mainly folded protein. At low angles (< 0.05 nm^{-1}) an intensity increase was detected (Figure 79A). This could eventually have resulted from a slight aggregation in the sample. The R_g was determined to a size of 2.85 ± 0.21 nm by Guinier approximation (Table S11C). This was significantly smaller than the

one, which was calculated in the batch measurement. The calculated $I(0)$ had a value of 1858.84 ± 4.84. The fit underlying these values was accurate, because a broad linear range (between 0.0118 nm^{-1} and 0.206 nm^{-1}) was detected and the standardized residuals were mostly in a range of 2 (Figure 79B). In the dimensionless Kratky plot a bell-shaped curve with a maximum around 1.7 and a height of approximately 1.1 was detected. The intensities declined evenly towards higher angles (Figure 79C). Compared to the curve determined from the batch measurement (Figure 73C), this curve was shifted to the left. Such a plot is very typical for globular and condensed proteins since these exhibit a maximum around $\sqrt{3}$ [349]. Thus, BON1 appears to be more condensed in the SEC-SAXS experiments at pH 9.0 and seems to have lost some of its flexibility.

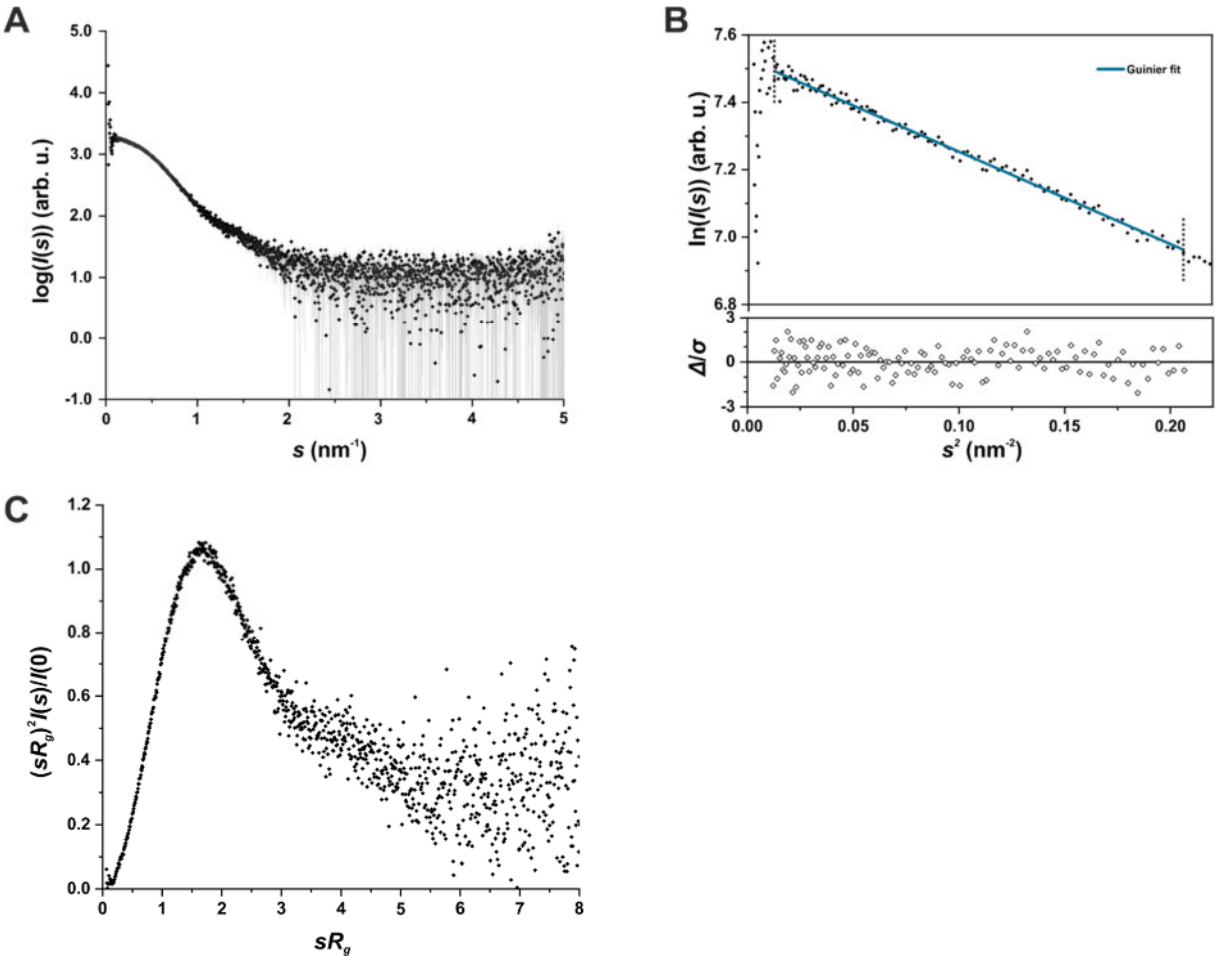


Figure 79: Scattering data of BON1 in the SEC-SAXS experiment. (A) Two-dimensional Log-linear plot of $I(s)$ versus s. **(B)** Upper plot shows the Guinier plot of the SAXS data. In blue the Guinier fit in a range of s_{min} = 0.1611 nm^{-1} and sR_g max = 1.28 (dotted lines) is depicted. Lower plot shows the standardized residual plot. **(C)** Dimensionless Kratky plot with the intensities normalized to the forward scattering intensity ($I(0)$) and the radius of gyration (R_g).

When analyzing the distance distribution from the SEC-SAXS data, a slightly left-shifted Gaussian-like curve with very small errors was detected. The peak of the curve was around 3.29 nm and declined smoothly to a maximum dimension of 8.6 nm (Figure 80A). This $P(r)$ represented the experimental data very accurately, because most standardized residuals were smaller than two, and especially in the low angular range the fit was considerably good (Figure 80B). In addition, the estimated quality was around 0.96 (Table S11C). Thus, BON1 is present in the SEC-SAXS experiment in a nearly globular, maybe slightly flattened, shape. In comparison to the batch measurements, BON1 seems to be more condensed, as a potential protrusion was not detectable

any more. Thus, structural pH-dependent alterations did take place and hindered the flexibility of BON1.

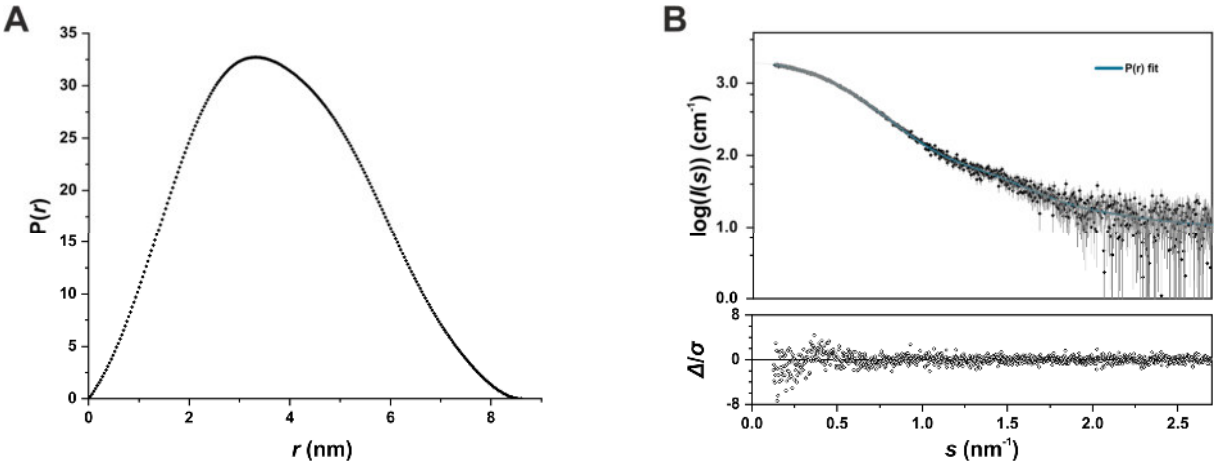


Figure 80: Distance distribution of BON1 in the SEC-SAXS experiment. (A) P(*r*) versus *r* profile of BON1. **(B)** Fit of the P(*r*) function (blue) to the experimental SAXS data. Lower plot depicts the standardized residual plot.

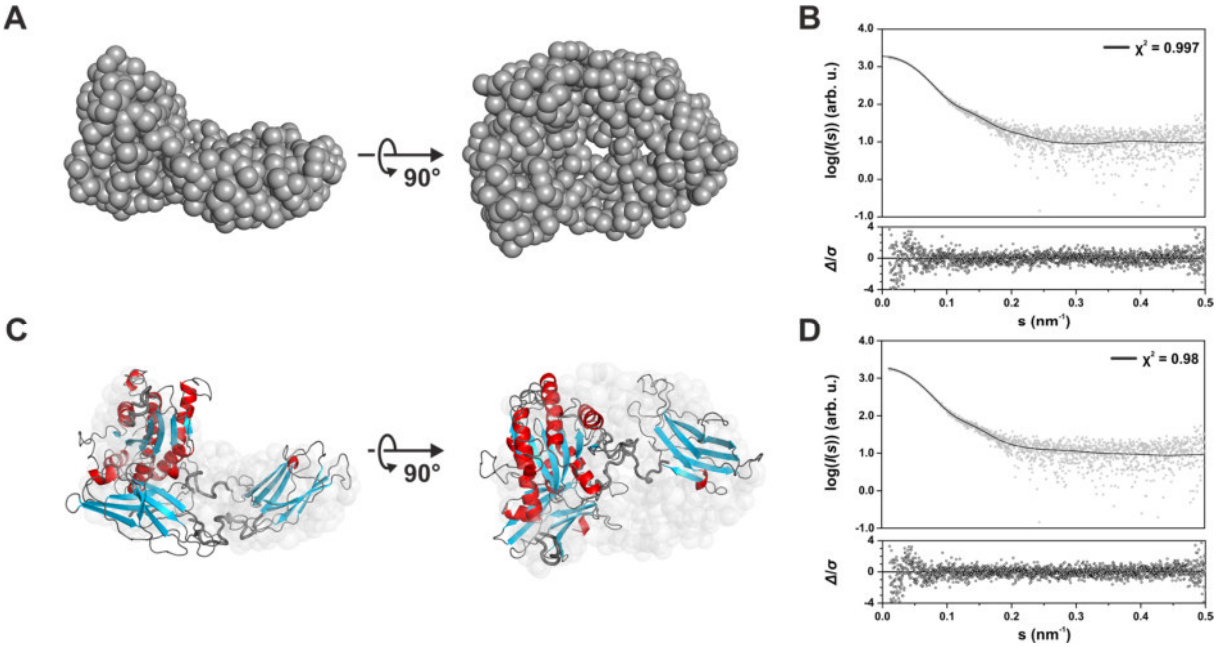


Figure 81: *Ab initio* and *in silico* structures of BON1 derived from SEC-SAXS measurements. (A) *Ab initio* bead model of BON1 generated by *GASBOR* with a resolution of 35 ± 3 Å and an NSD of 0.968. **(B)** Fit of the *ab initio* model to the raw scattering data with the standardized residual plot. **(C)** *ITASSA* derived *in silico* model, which was refined using *CORAL*. **(D)** Fit of the *in silico* model to the raw scattering data with the standardized residual plot.

For a more detailed picture the *ab initio* model, using *GASBOR*, was generated 20 times and the model with the lowest NSD value was selected. This was an NSD value of 0.968 (Figure 81A). Therefore, the model was chosen, which was the most similar one to all other models. The standard deviation of the 20 NSDs was 0.046, which suggested that all models did not differ from each other and therefore, the model calculated is highly likely. The model generated, was a chain-like dummy residue model, where each bead represented a C-α atom and therefore depicts the overall shape of the protein. In a next step, the *ab initio* model was compared to the experimental

data. A theoretical scattering curve was calculated for the model using *CRYSOL* (Figure 81B, black line), which was compared to the scattering data. Standardized residuals were mostly below 2 and χ^2 value with 0.997 was rather low, indicating a very good fit of the model to the data. The model itself, was like the batch-derived model, bowl- or hand-like shaped, with a deepening in the middle. Although, as detected before in the P(*r*) function, the protrusion was missing, leading to a more condensed structure of BON1.

For better comparison of the batch- and SEC-SAXS-derived structures, the *in silico* structure in the SEC-SAXS experiment was generated relying on the same programs as in the batch experiment. Therefore, *CORAL* was used again to allow for multi-domain-based modeling, and the previously determined BON1 domains were selected as a basis for the calculations. As a result, an *in silico* model was obtained, with a χ^2 value of 0.98, resulting from a nearly perfect fit with very low standardized residuals (Figure 81D). In this structure, the three predicted domains, being connected via linker regions, were detected as well (Figure 81C). As observed before in *ab initio* structure, the overall shape was much more condensed. This resulted in the VWA not being pulled apart, but being more similar to the crystal structure of the A3 domain of the von Willebrand factor from *Homo sapiens* (*PDB*-ID: 1AO3) (Figure 77D, 81).

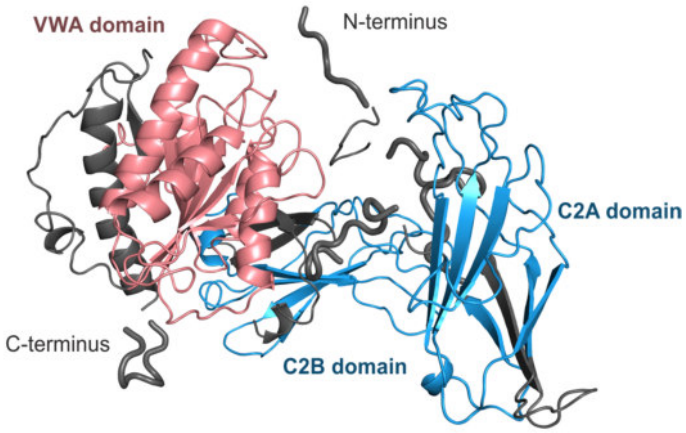


Figure 82: Domains of BON1 in the SEC-SAXS experiment. Depicted is the *CORAL*-derived *in silico* structure of BON1, fitted to the SEC-SAXS data. The predicted C2 domains are highlighted in blue, as in contrast the VWA domain is highlighted in red.

Consequently, BON1 is more compact in the SAUL1 buffer, with an N-terminus that is more rigid than in its own buffer. This restriction of the BON1 movement in the SAUL1 buffer could result in domains no longer being accessible. Therefore, a possible interaction between BON1 and SAUL1 during the SEC-SAXS experiment could be prevented. Nevertheless, this experiment gives new insights into the structure which could not have been obtained from the batch measurements. As stated before, BON1 appears to be bowl-shaped in both experiments. The SEC-SAXS measurements now allowed to predict more accurately, how this bowl-like structure is actually formed. The upper side of the "bowl" is formed by the VWA and the C2B domain, whereas the C2A domain is located on the opposite side (Figure 82). Consequently, BON1 could be structured in the same way under physiological-like conditions, although BON1 is likely to be more flexible and less condensed.

4.3.6 INTERACTIONS OF SAUL1 AND BON1

MST analysis of SAUL1-BON1 interactions

On the basis of successful purification of recombinant SAUL1 and BON1 as recombinant proteins (Figure 30, 70) [194,346] their interaction was thought to be investigated *in vitro*. In a first attempt, microscale thermophoresis (MST) ought to be used. This technique would allow for a determination of binding kinetics and to investigate the SAUL1-BON1 interaction in more detail. In a first experiment, SAUL1 was chosen for the required tagging with a fluorescent dye. In the following MST measurements, no fluorescence was detected, despite of protein being present. It was shown that labeling of SAUL1 with the fluorescing dye was not possible. The buffer exchange that was necessary to label proteins of interest could have led to an aggregation of SAUL1, resulting in a low labeling efficiency. It was shown in preceding DLS experiments that SAUL1 was rather unstable and tended to form aggregates (Figure 32). Alternatively, BON1 was labeled and a corresponding fluorophore signal was detected. To determine the labelling efficiency, a concentration curve of unbound dye and labelled BON1 was measured. This resulted in a comparatively very low signal for BON1 (Figure 83). As an efficiency of 1:1 of dye to BON1 would have been preferable, the detected fluorescence should have been as high as a similar concentration of the dye. The calculated efficiency was 41:1 and therefore too low. Accordingly, investigating the interaction between SAUL1 and BON1 was not pursued using MST as a technique of choice.

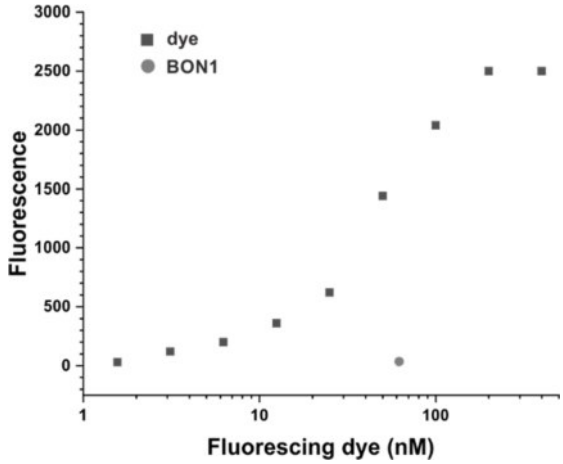


Figure 83: Microscale thermophoresis test of labelled BON1. 62 nM BON1 (circle) was labelled using the *Monolith NT™ Protein Labelling Kit Green*. Its fluorescence intensity was compared to a dye concentration series (squares).

Native MS analysis of SAUL1-BON1 interactions

In a next experiment, the BON1-SAUL1 interaction was investigated by native MS. This technique allowed to determine, with a very high specificity, whether any BON1-SAUL1 complexes were present under native conditions using ammonium acetate. Non-covalent oligomeric structures would stay intact during the used nano-electrospray ionization (ESI) [411]. At first, the recombinant SAUL1 and BON1 proteins were measured separately. In case of SAUL1, ions corresponding to a monomer with 88.8 kDa and a dimer with 177.6 kDa were detected (Figure 84A). Ions corresponding to the dimer were detected with a significantly lower intensity. Thus, SAUL1 seemed to be present in this experiment mostly as a monomer. The ratio between the monomer and the dimer is quite noteworthy, because SAUL1 has been reported to be present as a concentration-dependent oligomer [346,350]. These findings are consistent with previous SEC-SAXS measurements. During both experiments a low SAUL1 protein concentration (less than 1 mg ml^{-1}) was measured where the monomer was the dominant oligomer.

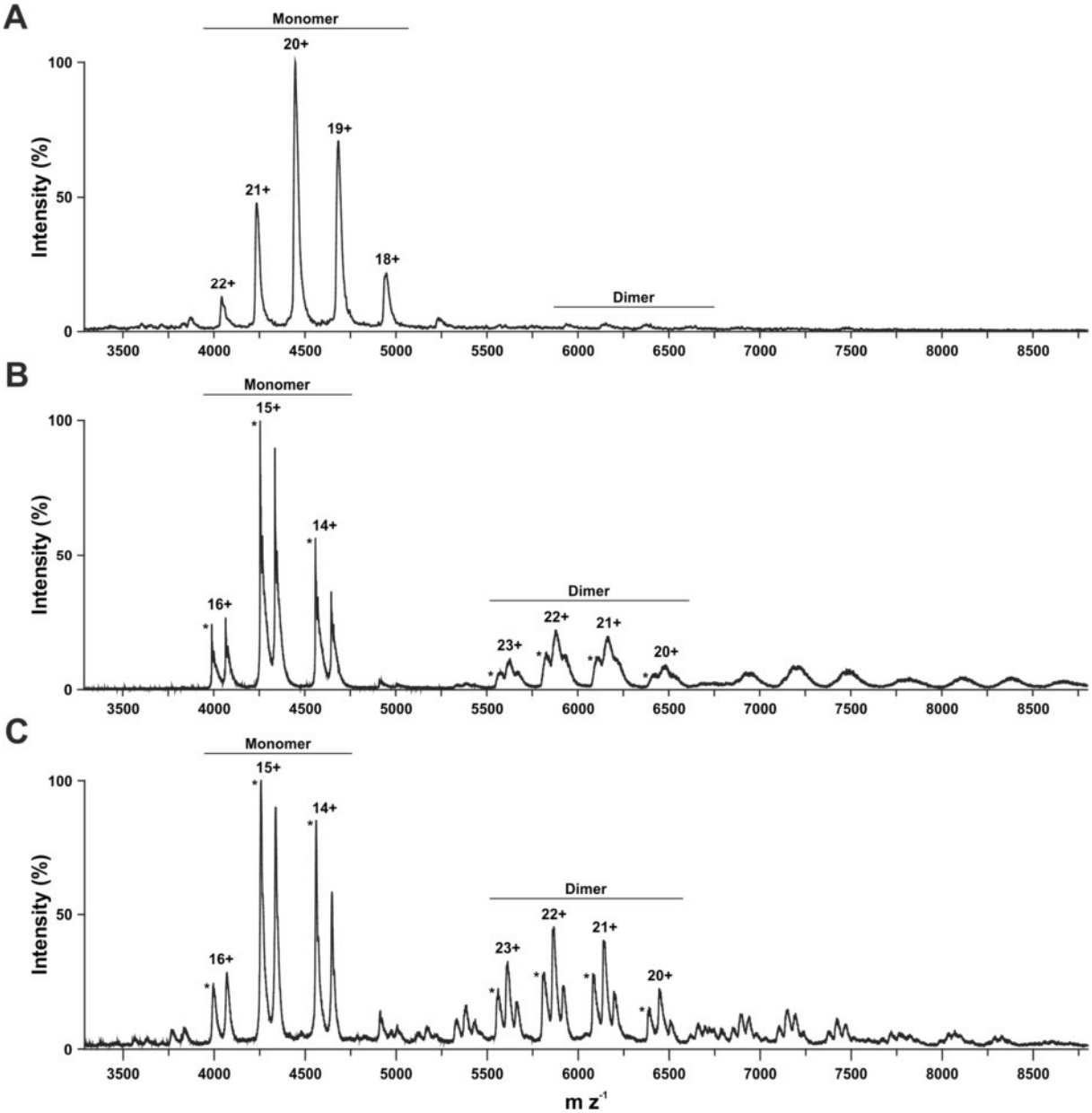


Figure 84: Native MS analysis of SAUL1 and BON1. Native MS measurements of proteins in 250 mM ammonium acetate. **(A)** 11 µM of SAUL1 at pH of 9.0 with a detected monomeric MW of 88.8 kDa, **(B)** 5 µM of BON1 at pH of 7.5 with a detected monomeric MW of 63.9 kDa (asterisk), monomeric MW of 65.2 kDa and dimeric MW of 127.8 kDa (asterisk) and **(C)** SAUL1 and BON1 in a ratio of 11 µM to 5 µM at a pH of 9.0. Detected were a monomeric MW of 63.9 kDa (asterisk), monomeric MW of 65.2 kDa and dimeric MW of 127.8 kDa (asterisk) corresponding to BON1.

In native MS, ions matching a monomeric and dimeric structure were detected for BON1. In contrast to SAUL1, two ion distributions corresponding to two different masses were detected (Figure 84B). In theory, monomeric BON1 has a molecular weight (MW) of 63.9 kDa, which was detected in the spectrogram (Figure 84B, asterisk). The corresponding ions for the dimer with a mass of 127.8 kDa were detected as well. Interestingly, the additional signals, which were detected for a monomeric BON1, corresponded to a mass of 65.2 kDa, thus deviating with 1.3 kDa from the theoretical size. Consequently, two different forms of BON1 were detected. This phenomenon, has been described before [411], and it was not surprising regarding the bowl-like shape of BON1 (Figure 75). BON1 may trap molecules in the detected cavity, which would result in a higher detected MW. What the 1.3 kDa actually correspond to, can only be speculated as no MS/MS data were available. Remnants of the purification could have been attached to the cavity

of BON1 and may therefore have resulted in the higher MW. Nonetheless, BON1 was detected in two oligomeric states, which is remarkable because previous SAXS experiments could only detect BON1 as a monomer. Due to the effect of larger particles contributing to the detected scattering intensity with a magnitude being proportional to the concentration and the volume squared [345], a dimer would have been detected in the SAXS measurements. Thus, BON1 was for the first time detected as a dimer in a ratio of 1:5 in respect to the monomer. This was highly interesting, because the potential Ca^{2+}-dependent oligomerization of plant copines has been observed before [412]. Thus, one could speculate, that BON1 might be present as an oligomer *in planta*, which was not detected in SAXS experiments due to the absence of Ca^{2+} ions. In particular, 5 mM EDTA, which was used for the BON1 buffer, could have prevented any Ca^{2+} binding and subsequent effect.

During her master thesis Mareike Schmidt, who continued the work of Marcel Bhattharai, found first evidence for a divalent ion-dependent oligomerization. She used a buffer for protein purification containing divalent ions ($MgSO_4$) that resulted in DLS experiments in a mean radius of 5.03 ± 0.12 nm. This was significantly larger than in a purification, in which only monovalent ions (NaCl) were used (mean radius of 3.61 ± 0.23 nm) (Figure 85). Both measurements were performed with a similar protein concentration. An effect of the higher pH can be excluded, because BON1 had been purified before by using the SAUL1-buffer at a pH of 9.0 resulting in no change of the radius distribution (data not shown). Thus, BON1 seems to oligomerize in the presence of divalent ions, as it has been described before for human copines [144,412].

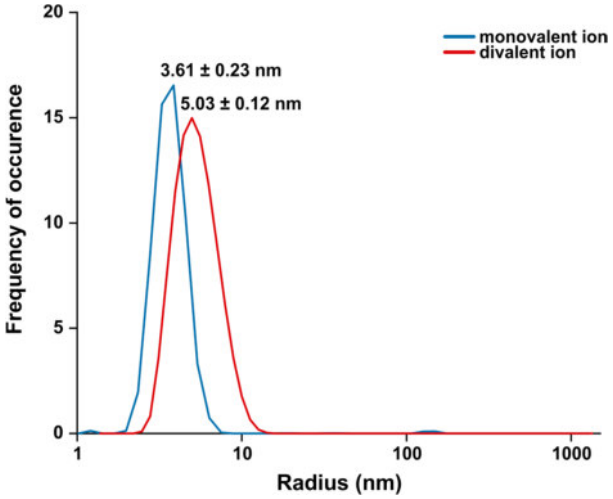


Figure 85: DLS analysis of the influence of divalent ions on BON1. DLS analysis of BON1 being purified using a buffer containing monovalent ions (100 mM Tris-HCl, 200 mM NaCl, 5 mM EDTA, pH 7.5) at a concentration of 1.3 mg ml^{-1} or divalent ions (100 mM triethanolamine, 50 mM $(NH_4)_2SO_4$, 10 mM $MgSO_4$, pH 8.5) at a concentration of 1.7 mg ml^{-1}.

Following the separate analysis of SAUL1 and BON1, they were analyzed together in a molar ratio of 11:5. This ratio was chosen on the basis of previous DLS experiments that pointed towards a complex most visible around a molar ratio of SAUL1:BON1 around 2 to 1 (data not shown). When analyzing SAUL1 and BON1 together, no ions corresponding to a potential interacting complex were detected (Figure 84C). Detected ions matched the BON1 monomers and dimers. Surprisingly, no ions corresponding to a SAUL1 mono- or dimer were found. This might have been caused by the overall shape of both proteins, as the ionization is based on the binding of ammonium ions to basic side groups on the surface of the proteins [411]. In comparison to the more elongated SAUL1 (Figure 44), BON1 is rather globular, especially at a pH of 9.0 (Figure 75, 80). Thus, BON1 has a smaller surface, leading to an easier and faster ionization and therefore earlier detection [411]. As a result, the BON1 peaks could have superimposed the SAUL1 ones. In additional experiments this would have to be investigated using MS/MS, to identify SAUL1 in the BON1 peaks.

In summary, no interaction between BON1 and SAUL1 at a pH of 9.0 and under these buffer conditions was demonstrated. How the pH and the buffer environment may potentially hinder the interaction will be discussed later, as similar issues are proposed to have taken place during the SEC-SAXS experiments. Still, the buffer composition might be an essential key to investigate the interaction, as it has been shown, that Ca^{2+} could be an essential factor mediating the formation of the complex [413]. Nonetheless, this experiment gave valuable insights into the potential oligomerization of BON1, which needs to be investigated in future and into the multimerization of SAUL1 as well, thus highlighting their very interesting structural complexity. Accordingly, native MS experiments should definitely be considered for future experiments, because they could give valuable insights into the formation of non-covalent structures. Specifically, the proposed concentration dependencies of SAUL1 could be studied very easily.

SEC-SAXS analysis of SAUL1-BON1 interactions

SEC-SAXS measurements were performed to investigate potential interactions of BON1 and SAUL1 on a structural level. With SEC-SAXS it is possible to analyze different complexes separately, as the SEC was upstream of the synchrotron beam. Therefore, SAUL1 and BON1 were mixed in a molar ratio of 2:1 and loaded onto the SEC column. During SEC-SAXS, two peaks were detected in the chromatogram (Figure 86A, B). The retention volume of those peaks was compared to singular measurements of SAUL1 (Figure 40A) and BON1 (Figure 78A). In an overlay both detected peaks corresponded either to the singular SAUL1, or the singular BON1 measurement (Figure 86C). No third peak was detected.

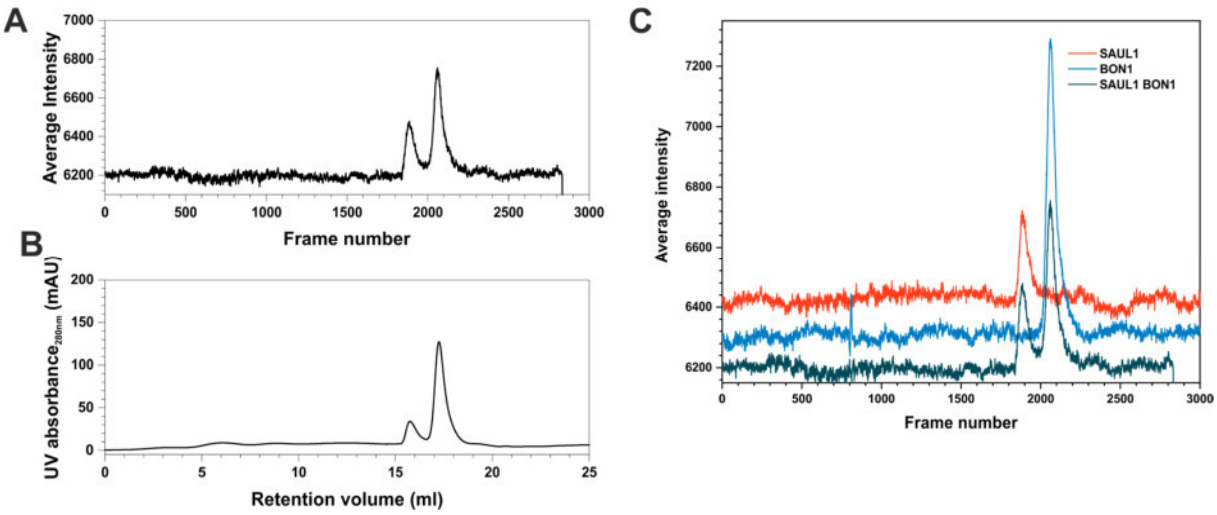


Figure 86: No peak corresponding to an interaction between SAUL1 and BON1 was detected in the chromatograms of the SEC-SAXS experiments. Scattering intensities and chromatograms of SEC-SAXS experiments using a *Superose™ 6 Increase 10/300 GL* column. **(A)** Scattering intensities of monodisperse SAUL1 and BON1 in a molar ratio of 1:2. **(B)** Chromatogram of the UV absorption at 280 nm. **(C)** Overlay of scattering intensities of singular SEC-SAXS measurements of SAUL1 (red), BON1 (blue) and the mixture of both (turquoise). Curves are slightly shifted on the intensity axes to achieve a better comparison.

Therefore, no complex corresponding to an interaction between SAUL1 and BON1 was detected. Consequently, SAUL1 and BON1 were found not to interact during SEC-SAXS measurements. Nevertheless, both scattering peaks, which correspond to the singular SAUL1 and BON1, were analyzed. Thus, effects on the structure of SAUL1 and BON1 by the presence of the respective other were investigated. The information obtained was compared with the individual measurements of SAUL1 and BON1.

For SAUL1 a slight steeper slope decrease was detected in the log-linear plot of the scattering intensities for the combined measurements. In addition, decreasing intensity values were detected at very low angles (Figure 87A). Such an effect is known to result from inter-particle rejections, which may originate from brief interactions with BON1 or a lower concentration of SAUL1 in the sample. The lower concentration is due to the fact that SAUL1 and BON1 were mixed in an unchanged sample volume. For BON1 no changes were detected for the scattering plot (Figure 87B).

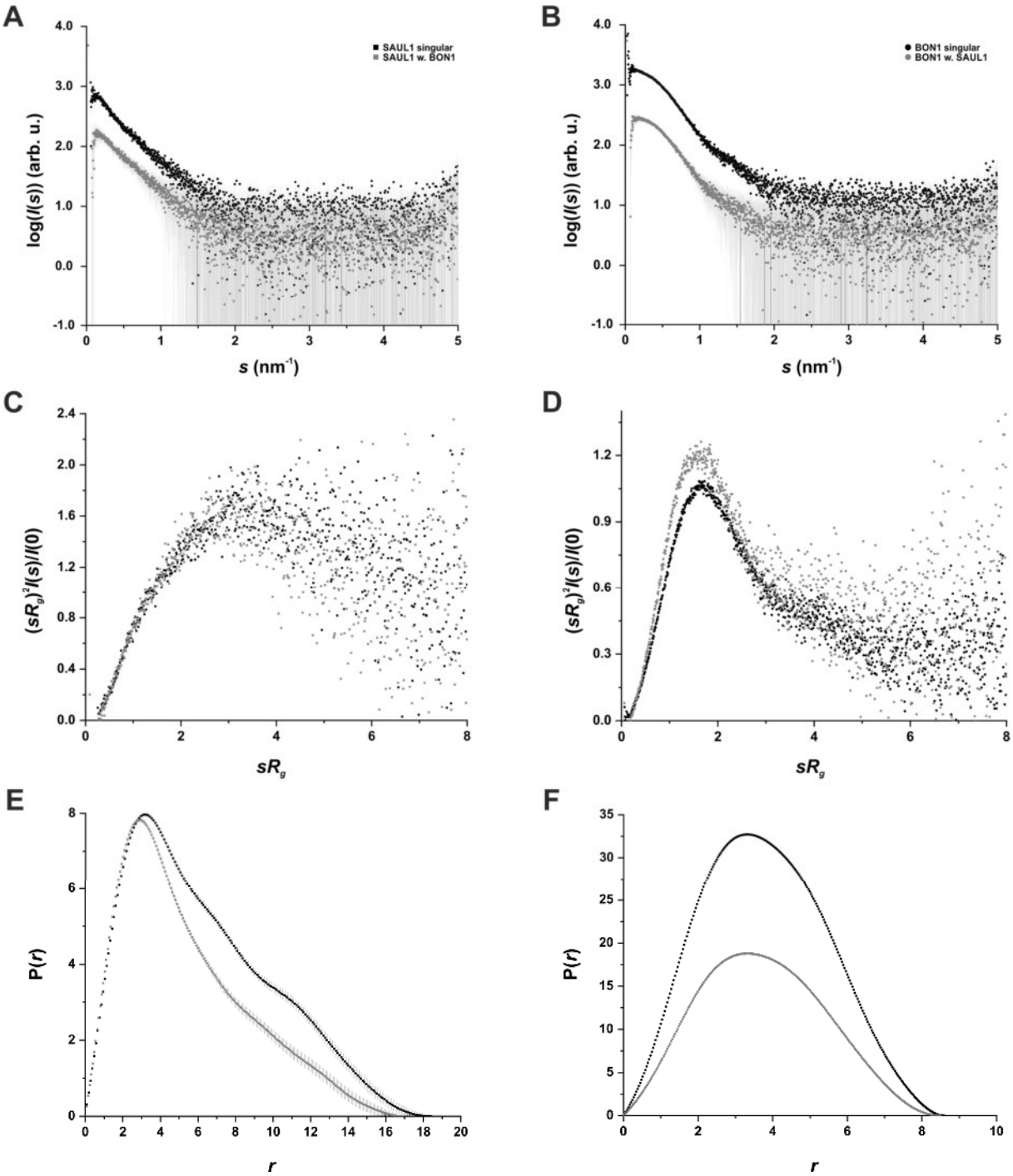


Figure 87: Combined SAUL1 and BON1 SEC-SAXS measurements using a molar ratio of 2:1. The combined SEC-SAXS measurement (grey) is always compared to the singular SEC-SAXS measurement (black). Log-linear plot of $I(0)$ and s for **(A)** SAUL1 and **(B)** BON1. Curves are slightly shifted on the intensity axes for a better comparison. Dimensionless Kratky plot for **(C)** SAUL1 and **(D)** BON1. Distance distribution of **(E)** SAUL1 and **(F)** BON1.

In a next step a dimensionless Kratky plot was used to analyze potential changes of the shape of SAUL1 and BON1 in the different measurements. In case of SAUL1 no change of the plot was detected compared to the individual SAUL1 measurement (Figure 87C). Consequently, the shape of SAUL1 was not changed in the presence of BON1. Thus, changes in the scattering curve seem to have originated out of a lower concentration of SAUL1, because the dimensionless Kratky plot is concentration-independent. A similar behavior was detected for the distance distribution. Although the curve is slightly changed, the d_{max} remains almost the same (Figure 87E). Thus, SAUL1 seems to be ordered in the same way compared to the monomer measurement. This was confirmed with an *ab initio* model, which was generated with *GASBOR*. The *ab initio* models for the singular SAUL1 and the combined measurement were highly alike (data not shown). For BON1 the dimensionless Kratky plot had a similar shape, even though the maximum was slightly altered (Figure 87D). Because the maximum was still around $\sqrt{3}$, which is an indication for globular proteins, BON1 still remained mostly globular. Analyzing the distance distribution similar results were obtained. The curve for BON1 in the combined measurement was still Gaussian-shaped, indicating a globular shaped BON1 (Figure 87F). Nevertheless, less medial radii were detected. As a consequence, BON1 could be more disk-shaped than BON1 in the singular SEC-SAXS measurement. However, no severe structural changes were detected for BON1 in the combined measurement. SAUL1 and BON1 therefore have no influence on the structure of the other.

Nevertheless, these findings are contradictory, since BON1 was found to interact with SAUL1 in two separate experiments [142] (also, Lienemann, T., unpublished data). There are three possible reasons why interaction may not have been observed in the SEC-SAXS measurements. (i) SAUL1 and BON1 could interact indirectly by binding both to a mediating protein. This seems to be quite likely, because SAUL1 interacts with multiple proteins. These could all be part of a larger complex that regulates exo- or endocytosis during immunity (see chapter 4.2.5 and 4.3.3) [142].

(ii) The second potential obstacle to SAUL1-BON1 interaction could have been the buffer environment. Due to the instability of SAUL1 its lysis buffer with a pH of 9.0 was used. This basic pH did hinder BON1 flexibility and resulted in steric changes (see chapter 4.3.5 and figure 81). Particularly, since the two C2 domains of BON1 are hypothesized to mediate the interaction and even their conformation in respect to each other might be important for the interaction. In case of binding of a streptococcal protein G to human immunoglobulin G both C2 domains were involved in protein-protein binding [414]. A change in the orientation of the C2 domains was observed in the SEC-SAXS measurement of BON1 at a pH of 9.0 (Figure 81). As a consequence, this could result in an impaired interaction between SAUL1 and BON1.

(iii) The third reason that no interaction between SAUL1 and BON1 was detected could have been the absence of Ca^{2+}. Most C2 domains that are known are Ca^{2+}-dependent [413]. Similar to other findings the C2 domains BON1 may even need a specific Ca^{2+} concentration to bind to SAUL1 [404]. Comparable to other copines, a potential oligomerization of BON1, which was dependent on divalent ions, was shown during this thesis (Figure 85). Thus, BON1 might bind to Ca^{2+}. Accordingly, protein-protein interaction might also be dependent on Ca^{2+}.

To overcome these hinderances the interaction might be investigated in a more suitable buffer containing a considerable amount of Ca^{2+}. In addition, an MST experiment could be performed using cell extracts, to provide a more *in vivo*-like environment and therefore, accounting for a lack of Ca^{2+} and/or additional interaction partners. To avoid the low dye binding efficiency, plants which express SAUL1 or BON1 fused to a fluorophore, could be used. As a result, this experiment did not confirm the direct interaction between SAUL1 and BON1. However, the experiment provided important insights into the interdependencies that could underlie the interaction and laid the foundation for future research.

5 ABSTRACT

The *Arabidopsis thaliana* E3 ubiquitin ligase SAUL1 (SENESCENCE-ASSOCIATED E3 UBIQUITIN LIGASE 1), which acts as a positive regulator during pattern-triggered immunity (PTI), is characterized by two intriguing hallmarks.

(i) On the one hand, SAUL1 is guarded by two heteromeric nucleotide-binding leucine-rich repeat protein (NLR) complexes. These initiate an inducible effector-triggered immunity (ETI) in the *saul1-1* mutant. Although ETI is known now for more than 15 years its early regulations and the demarcation from PTI are still not fully understood. Therefore, the *saul1-1* phenotype was used to investigate gene regulations during ETI by transcriptomics. These analyses revealed that observed gene regulations were highly similar to other autoimmune mutants and *saul1-1* is therefore an ideal model to study ETI. By accessing global changes an early interplay of multiple plant hormones, like salicylic acid, jasmonic acid, ethylene and abscisic acid, was identified, which is likely to initiate the onset of ETI. In addition, repression of brassinosteroid (BR) signaling may be a potential reason of the autoimmune-related growth arrest in *saul1-1*. Before, BR signaling, as well as synthesis of camalexin and JA signaling have been associated exclusively with PTI. In conclusion, it was possible to show that ETI and PTI are not as distinct as originally thought. In addition, 19 very early differentially expressed genes were identified during the first two hours of the onset of the *saul1-1* phenotype. Five of them were analyzed concerning their potential driving role during ETI and two transcription factors, *ERF2* (*Ethylene Response Factor 2*) and *ZAT7* (*Zinc Finger of Arabidopsis thaliana 7*), were identified that pose ideal candidates being initial regulators of ETI.

(ii) On the other hand, SAUL1 is characterized by its plasma membrane localization and its domain structure which differ from other plant U-box type E3 ligases (PUBs). Structural experiments allowed to determine for the first time the structure of a PUB. This revealed that SAUL1 is most likely present as a U-box-mediated concentration-dependent oligomer, which is mostly present as a dimer and tetramer. Therefore, SAUL1 activity may be controlled by oligomerization rendering the U-box inaccessible in a multimeric state. Analyzing the domain organization, it was possible to identify a positively charged stretch next to the C-terminus that is likely to bind to negatively charged phospholipids. Therefore, and due to the predicted flexibility of the C-terminal armadillo repeats, SAUL1 may bind to multi-vesicular bodies (MVBs) and to the plasma membrane at the same time. This could result in the observed patches at the plasma membrane, originating from MVB tethering, which would be regulated by oligomerization as well. In addition, the low-resolution structure of the SAUL1 binding partner BON1, a plant copine, was as well determined. Although no direct interaction between SAUL1 and BON1 could be observed, the performed experiments may suggest that Ca^{2+} ions may mediate a SAUL1-BON1 interaction.

In conclusion, this thesis provides new insights into regulations during ETI, the structural organization of SAUL1, the concentration-dependent oligomerization of SAUL1, its membrane binding capacities and the interaction between SAUL1 and BON1.

6 LITERATURE

1. Koornneef, M. & Meinke, D. The development of Arabidopsis as a model plant. *Plant J.* **61**, 909–921 (2010).
2. Provart, N. J. *et al.* 50 years of Arabidopsis research: highlights and future directions. *New Phytol.* **209**, 921–944 (2016).
3. Jones, J. D. G. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).
4. Cui, H., Tsuda, K. & Parker, J. E. Effector-Triggered Immunity: From Pathogen Perception to Robust Defense. *Annu. Rev. Plant Biol.* **66**, 487–511 (2015).
5. Stael, S. *et al.* Plant innate immunity - sunny side up? *Trends Plant Sci.* **20**, 3–11 (2015).
6. Chinchilla, D., Bauer, Z., Regenass, M., Boller, T. & Felix, G. The Arabidopsis Receptor Kinase FLS2 Binds flg22 and Determines the Specificity of Flagellin Perception. *Plant Cell* **18**, 465–476 (2006).
7. Wan, W. L., Fröhlich, K., Pruitt, R. N., Nürnberger, T. & Zhang, L. Plant cell surface immune receptor complex signaling. *Curr. Opin. Plant Biol.* **50**, 18–28 (2019).
8. Boutrot, F. & Zipfel, C. Function, Discovery, and Exploitation of Plant Pattern Recognition Receptors for Broad-Spectrum Disease Resistance. *Annu. Rev. Phytopathol.* **55**, 257–286 (2017).
9. Meyers, B. C., Kozik, A., Griego, A., Kuang, H. & Michelmore, R. W. Genome-Wide analysis of NBS-LRR-Encoding Genes in Arabidopsis. *Plant Cell* **15**, 809–834 (2003).
10. Nishimura, M. T. & Dangl, J. L. Paired Plant Immune Receptors. *Science* **344**, 267–268 (2014).
11. Zhang, Y. *et al.* Temperature-dependent autoimmunity mediated by chs1 requires its neighboring TNL gene SOC3. *New Phytol.* **213**, 1330–1345 (2017).
12. Qi, D. & Innes, R. W. Recent advances in plant NLR structure, function, localization, and signaling. *Front. Immunol.* **4**, 1–10 (2013).
13. Thomma, B. P. H. J., Nürnberger, T. & Joosten, M. H. A. J. Of PAMPs and Effectors: The Blurred PTI-ETI Dichotomy. *Plant Cell* **23**, 4–15 (2011).
14. Li, B., Meng, X., Shan, L. & He, P. Transcriptional Regulation of Pattern-Triggered Immunity in Plants. *Cell Host Microbe* **19**, 641–650 (2016).
15. Ryals, J., Uknes, S. & Ward, E. Systemic Acquired Resistance. *Plant Physiol.* **104**, 1109–1112 (1994).
16. Imkampe, J. *et al.* The Arabidopsis Leucine-Rich Repeat Receptor Kinase BIR3 Negatively Regulates BAK1 Receptor Complex Formation and Stabilizes BAK1. *Plant Cell* **29**, 2285–

2303 (2017).

17. Couto, D. & Zipfel, C. Regulation of pattern recognition receptor signalling in plants. *Nat. Rev. Immunol.* **16**, 537–552 (2016).

18. Sun, Y. *et al.* Structural Basis for flg22-Induced Activation of the Arabidopsis FLS2-BAK1 Immune Complex. *Science* **342**, 624–629 (2013).

19. Li, L. *et al.* The FLS2-associated kinase BIK1 directly phosphorylates the NADPH oxidase RbohD to control plant immunity. *Cell Host Microbe* **15**, 329–338 (2014).

20. Wan, W. L. *et al.* Comparing Arabidopsis receptor kinase and receptor protein-mediated immune signaling reveals BIK1-dependent differences. *New Phytol.* **221**, 2080–2095 (2019).

21. Kadota, Y. *et al.* Direct Regulation of the NADPH Oxidase RBOHD by the PRR-Associated Kinase BIK1 during Plant Immunity. *Mol. Cell* **54**, 43–55 (2014).

22. Ogasawara, Y. *et al.* Synergistic Activation of the Arabidopsis NADPH Oxidase AtrbohD by Ca2+ and Phosphorylation. *J. Biol. Chem.* **283**, 8885–8892 (2008).

23. Torres, M. A., Dangl, J. L. & Jones, J. D. G. Arabidopsis gp91phox homologues Atrbohd and Atrbohf are required for accumulation of reactive oxygen intermediates in the plant defense response. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 517–522 (2002).

24. Torres, M. A. & Dangl, J. L. Functions of the respiratory burst oxidase in biotic interactions, abiotic stress and development. *Curr. Opin. Plant Biol.* **8**, 397–403 (2005).

25. Torres, M. A. ROS in biotic interactions. *Physiol. Plant.* **138**, 414–429 (2010).

26. Marcec, M. J., Gilroy, S., Poovaiah, B. W. & Tanaka, K. Mutual interplay of Ca2+ and ROS signaling in plant immune response. *Plant Sci.* **283**, 343–354 (2019).

27. Waszczak, C., Carmody, M. & Kangasjärvi, J. Reactive Oxygen Species in Plant Signaling. *Annu. Rev. Plant Biol.* **69**, 209–236 (2018).

28. Miller, G. *et al.* The Plant NADPH Oxidase RBOHD Mediates Rapid Systemic Signaling in Response to Diverse Stimuli. *Sci. Signal.* **2**, 1–11 (2009).

29. Tian, S. *et al.* Plant Aquaporin AtPIP1;4 Links Apoplastic H2O2 Induction to Disease Immunity Pathways. *Plant Physiol.* **171**, 1635–1650 (2016).

30. Pel, Z. M. *et al.* Calcium channels activated by hydrogen peroxide mediate abscisic acid signalling in guard cells. *Nature* **406**, 731–734 (2000).

31. Dubiella, U. *et al.* Calcium-dependent protein kinase/NADPH oxidase activation circuit is required for rapid defense signal propagation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 8744–8749 (2013).

32. Choi, W. G. *et al.* Orchestrating rapid long-distance signaling in plants with Ca2+, ROS and electrical signals. *Plant J.* **90**, 698–707 (2017).

33. Boudsocq, M. *et al.* Differential innate immune signalling via Ca2+ sensor protein kinases. *Nature* **464**, 418–422 (2010).

34. Gao, X. & He, P. Nuclear dynamics of Arabidopsis calcium-dependent protein kinases in effector-triggered immunity. *Plant Signal. Behav.* **8**, (2013).

35. Seyfferth, C. & Tsuda, K. Salicylic acid signal transduction: the initiation of biosynthesis, perception and transcriptional reprogramming. *Front. Plant Sci.* **5**, 1–10 (2014).

36. Wang, X. *et al.* TCP transcription factors are critical for the coordinated regulation of ISOCHORISMATE SYNTHASE 1 expression in Arabidopsis thaliana. *Plant J.* **82**, 151–162 (2015).

37. Zhang, J. *et al.* A Pseudomonas syringae Effector Inactivates MAPKs to Suppress PAMP-

Induced Immunity in Plants. *Cell Host Microbe* **1**, 175–185 (2007).

38. Birkenbihl, R. P. *et al.* Principles and characteristics of the Arabidopsis WRKY regulatory network during early MAMP-triggered immunity. *Plant J.* **96**, 487–502 (2018).

39. Lal, N. K. *et al.* The Receptor-like Cytoplasmic Kinase BIK1 Localizes to the Nucleus and Regulates Defense Hormone Expression during Plant Innate Immunity. *Cell Host Microbe* **23**, 485-497.e5 (2018).

40. Mao, G. *et al.* Phosphorylation of a WRKY transcription factor by two pathogen-responsive MAPKs drives phytoalexin biosynthesis in Arabidopsis. *Plant Cell* **23**, 1639–1653 (2011).

41. Petersen, K., Fiil, B. K., Mundy, J. & Petersen, M. Downstream targets of WRKY33. *Plant Signal. Behav.* **3**, 1033–1034 (2008).

42. Phukan, U. J., Jeena, G. S. & Shukla, R. K. WRKY Transcription Factors: Molecular Regulation and Stress Responses in Plants. *Front. Plant Sci.* **7**, 1–14 (2016).

43. Liu, S., Bartnikas, L. M., Volko, S. M., Ausubel, F. M. & Tang, D. Mutation of the Glucosinolate Biosynthesis Enzyme Cytochrome P450 83A1 Monooxygenase Increases Camalexin Accumulation and Powdery Mildew Resistance. *Front. Plant Sci.* **7**, 1–13 (2016).

44. Rayapuram, N. *et al.* Quantitative Phosphoproteomic Analysis Reveals Shared and Specific Targets of Arabidopsis Mitogen-Activated Protein Kinases (MAPKs) MPK3, MPK4, and MPK6. *Mol. Cell. Proteomics* **17**, 61–80 (2018).

45. Tsuda, K. *et al.* Dual Regulation of Gene Expression Mediated by Extended MAPK Activation and Salicylic Acid Contributes to Robust Innate Immunity in Arabidopsis thaliana. *PLoS Genet.* **9**, (2013).

46. Ding, Y. *et al.* Opposite Roles of Salicylic Acid Receptors NPR1 and NPR3/NPR4 in Transcriptional Regulation of Plant Immunity. *Cell* **173**, 1454-1467.e10 (2018).

47. Pieterse, C. M. J., Van der Does, D., Zamioudis, C., Leon-Reyes, A. & Van Wees, S. C. M. Hormonal Modulation of Plant Immunity. *Annu. Rev. Cell Dev. Biol.* **28**, 489–521 (2012).

48. Glazebrook, J. CONTRASTING MECHANISMS OF DEFENSE AGAINST BIOTROPHIC AND NECROTROPHIC PATHOGENS. *Annu. Rev. Phytopathol.* **43**, 205–227 (2005).

49. Tsuda, K., Sato, M., Stoddard, T., Glazebrook, J. & Katagiri, F. Network Properties of Robust Immunity in Plants. *PLoS Genet.* **5**, 1–16 (2009).

50. Bigeard, J., Colcombet, J. & Hirt, H. Signaling Mechanisms in Pattern-Triggered Immunity (PTI). *Mol. Plant* **8**, 521–539 (2015).

51. Ellinger, D. *et al.* Elevated Early Callose Deposition Results in Complete Penetration Resistance to Powdery Mildew in Arabidopsis. *Plant Physiol.* **161**, 1433–1444 (2013).

52. Nishimura, M. T. *et al.* Loss of a Callose Synthase Results in Salicylic Acid–Dependent Disease Resistance. *Science* **301**, 969–972 (2003).

53. Blümke, A., Somerville, S. C. & Voigt, C. A. Transient expression of the Arabidopsis thaliana callose synthase PMR4 increases penetration resistance to powdery mildew in barley. *Adv. Biosci. Biotechnol.* **04**, 810–813 (2013).

54. Gigolashvili, T. *et al.* The transcription factor HIG1/MYB51 regulates indolic glucosinolate biosynthesis in Arabidopsis thaliana. *Plant J.* **50**, 886–901 (2007).

55. Luna, E. *et al.* Callose Deposition: A Multifaceted Plant Defense Response. *Mol. Plant Microbe Interact.* **24**, 183–193 (2011).

56. Rogers, E. E., Glazebrook, J. & Ausubel, F. M. Mode of Action of the Arabidopsis thaliana Phytoalexin Camalexin and Its Role in Arabidopsis-Pathogen Interactions. *Mol. Plant Microbe Interact.* **9**, 748–757 (1996).

57. Stotz, H. U. *et al.* Role of camalexin, indole glucosinolates, and side chain modification of

glucosinolate-derived isothiocyanates in defense of Arabidopsis against Sclerotinia sclerotiorum. *Plant J.* **67**, 81–93 (2011).

58. Xiang, T. *et al.* Pseudomonas syringae Effector AvrPto Blocks Innate Immunity by Targeting Receptor Kinases. *Curr. Biol.* **18**, 74–80 (2008).

59. Göhre, V. *et al.* Plant Pattern-Recognition Receptor FLS2 Is Directed for Degradation by the Bacterial Ubiquitin Ligase AvrPtoB. *Curr. Biol.* **18**, 1824–1832 (2008).

60. Zhang, J. *et al.* Receptor-like Cytoplasmic Kinases Integrate Signaling from Multiple Plant Immune Receptors and Are Targeted by a Pseudomonas syringae Effector. *Cell Host Microbe* **7**, 290–301 (2010).

61. Wang, W., Liu, N., Gao, C., Rui, L. & Tang, D. The Pseudomonas Syringae Effector AvrPtoB Associates With and Ubiquitinates Arabidopsis Exocyst Subunit EXO70B1. *Front. Plant Sci.* **10**, 1–16 (2019).

62. Cesari, S. Multiple strategies for pathogen perception by plant immune receptors. *New Phytol.* **219**, 17–24 (2018).

63. Schreiber, K. J., Bentham, A., Williams, S. J., Kobe, B. & Staskawicz, B. J. Multiple Domain Associations within the Arabidopsis Immune Receptor RPP1 Regulate the Activation of Programmed Cell Death. *PLoS Pathog.* **12**, 1–26 (2016).

64. El Kasmi, F. *et al.* Signaling from the plasma-membrane localized plant immune receptor RPM1 requires self-association of the full-length protein. *Proc. Natl. Acad. Sci. U. S. A.* **114**, E7385–E7394 (2017).

65. Toruño, T. Y., Shen, M., Coaker, G. & Mackey, D. Regulated Disorder: Posttranslational Modifications Control the RIN4 Plant Immune Signaling Hub. *Mol. Plant Microbe Interact.* **32**, 56–64 (2019).

66. Moffett, P. Using Decoys to Detect Pathogens: An Integrated Approach. *Trends Plant Sci.* **21**, 369–370 (2016).

67. Cesari, S., Bernoux, M., Moncuquet, P., Kroj, T. & Dodds, P. N. A novel conserved mechanism for plant NLR protein pairs: The 'integrated decoy' hypothesis. *Front. Plant Sci.* **5**, 1–10 (2014).

68. Wang, G. *et al.* The Decoy Substrate of a Pathogen Effector and a Pseudokinase Specify Pathogen-Induced Modified-Self Recognition and Immunity in Plants. *Cell Host Microbe* **18**, 285–295 (2015).

69. Grant, M. *et al.* The RPM1 plant disease resistance gene facilitates a rapid and sustained increase in cytosolic calcium that is necessary for the oxidative burst and hypersensitive cell death. *Plant J.* **23**, 441–450 (2000).

70. Zhu, Z. *et al.* Arabidopsis resistance protein SNC1 activates immune responses through association with a transcriptional corepressor. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 13960–13965 (2010).

71. Genot, B. *et al.* Constitutively Active Arabidopsis MAP Kinase 3 Triggers Defense Responses Involving Salicylic Acid and SUMM2 Resistance Protein. *Plant Physiol.* **174**, 1238–1249 (2017).

72. Gao, X. *et al.* Bifurcation of Arabidopsis NLR Immune Signaling via Ca2+-Dependent Protein Kinases. *PLoS Pathog.* **9**, (2013).

73. Cui, H. *et al.* A core function of EDS1 with PAD4 is to protect the salicylic acid defense sector in Arabidopsis immunity. *New Phytol.* **213**, 1802–1817 (2017).

74. Wu, Y. *et al.* The Arabidopsis NPR1 Protein Is a Receptor for the Plant Defense Hormone Salicylic Acid. *Cell Rep.* **1**, 639–647 (2012).

75. Liu, L. *et al.* Salicylic acid receptors activate jasmonic acid signalling through a non-

canonical pathway to promote effector-triggered immunity. *Nat. Commun.* **7**, 1–10 (2016).

76. Joglekar, S. *et al.* Chemical Activation of EDS1/PAD4 Signaling Leading to Pathogen Resistance in Arabidopsis. *Plant Cell Physiol.* **59**, 1592–1607 (2018).

77. Wang, Y. *et al.* A MPK3/6-WRKY33-ALD1-Pipecolic Acid Regulatory Loop Contributes to Systemic Acquired Resistance. *Plant Cell* **30**, 2480–2494 (2018).

78. Lapin, D. *et al.* A Coevolved EDS1-SAG101-NRG1 Module Mediates Cell Death Signaling by TIR-domain Immune Receptors. *Plant Cell* (2019). doi:10.1105/tpc.19.00118

79. Wang, J. *et al.* Reconstitution and structure of a plant NLR resistosome conferring immunity. *Science* **364**, 1–11 (2019).

80. Jubic, L. M., Saile, S., Furzer, O. J., El Kasmi, F. & Dangl, J. L. Help wanted: helper NLRs and plant immune responses. *Curr. Opin. Plant Biol.* **50**, 82–94 (2019).

81. Bhandari, D. D. *et al.* An EDS1 heterodimer signalling surface enforces timely reprogramming of immunity genes in Arabidopsis. *Nat. Commun.* **10**, (2019).

82. Wang, Y., Loake, G. J. & Chu, C. Cross-talk of nitric oxide and reactive oxygen species in plant programed cell death. *Front. Plant Sci.* **4**, 1–7 (2013).

83. Coll, N. S. *et al.* Arabidopsis Type I Metacaspases Control Cell Death. *Science* **330**, 1393–1397 (2010).

84. Mur, L. A. J., Kenton, P., Lloyd, A. J., Ougham, H. & Prats, E. The hypersensitive response; the centenary is upon us but how much do we know? *J. Exp. Bot.* **59**, 501–520 (2008).

85. Coll, N. S., Epple, P. & Dangl, J. L. Programmed cell death in the plant immune system. *Cell Death Differ.* **18**, 1247–1256 (2011).

86. Radojičić, A., Li, X. & Zhang, Y. Salicylic Acid: A Double-Edged Sword for Programed Cell Death in Plants. *Front. Plant Sci.* **9**, 1–5 (2018).

87. Fu, Z. Q. *et al.* NPR3 and NPR4 are receptors for the immune signal salicylic acid in plants. *Nature* **486**, 228–232 (2012).

88. Tewari, R. K., Prommer, J. & Watanabe, M. Endogenous nitric oxide generation in protoplast chloroplasts. *Plant Cell Rep.* **32**, 31–44 (2013).

89. Coll, N. S. *et al.* The plant metacaspase AtMC1 in pathogen-triggered programmed cell death and aging: Functional linkage with autophagy. *Cell Death Differ.* **21**, 1399–1408 (2014).

90. Friso, G. & Van Wijk, K. J. Posttranslational Protein Modifications in Plant Metabolism. *Plant Physiol.* **169**, 1469–1487 (2015).

91. Swatek, K. N. & Komander, D. Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016).

92. Kulathu, Y. & Komander, D. Atypical ubiquitylation - the unexplored world of polyubiquitin beyond Lys48 and Lys63 linkages. *Nat. Rev. Mol. Cell Biol.* **13**, 508–523 (2012).

93. Barberon, M. *et al.* Monoubiquitin-dependent endocytosis of the IRON-REGULATED TRANSPORTER 1 (IRT1) transporter controls iron uptake in plants. *Proc. Natl. Acad. Sci. U. S. A.* **108**, E450–E458 (2011).

94. Vierstra, R. D. The ubiquitin - 26S proteasome system at the nexus of plant biology. *Nat. Rev. Mol. Cell Biol.* **10**, 385–397 (2009).

95. Huang, Y. *et al.* An E4 ligase facilitates polyubiquitination of plant immune receptor resistance proteins in Arabidopsis. *Plant Cell* **26**, 485–496 (2014).

96. Trujillo, M. News from the PUB: plant U-box type E3 ubiquitin ligases. *J. Exp. Bot.* **69**, 371–384 (2017).

97. Ye, Y. & Rape, M. Building ubiquitin chains: E2 enzymes at work. *Nat. Rev. Mol. Cell Biol.* **10**, 755–764 (2009).

98. Hershko, A. & Chiechanover, A. THE UBIQUITIN SYSTEM. *Annu. Rev. Biochem.* **67**, 425–479 (1998).

99. Johnson, A. & Vert, G. Unraveling K63 Polyubiquitination Networks by Sensor-Based Proteomics. *Plant Physiol.* **171**, 1808–1820 (2016).

100. Romero-Barrios, N. & Vert, G. Proteasome-independent functions of lysine-63 polyubiquitination in plants. *New Phytol.* **217**, 995–1011 (2018).

101. Walsh, C. K. & Sadanandom, A. Ubiquitin chain topology in plant cell signaling: A new facet to an evergreen story. *Front. Plant Sci.* **5**, 1–6 (2014).

102. Sadanandom, A., Bailey, M., Ewan, R., Lee, J. & Nelis, S. The ubiquitin–proteasome system: central modifier of plant signalling. *New Phytol.* **196**, 13–28 (2012).

103. Isono, E. & Nagel, M. K. Deubiquitylating enzymes and their emerging role in plant biology. *Front. Plant Sci.* **5**, 1–6 (2014).

104. Chen, L. & Hellmann, H. Plant E3 Ligases: Flexible Enzymes in a Sessile World. *Mol. Plant* **6**, 1388–1404 (2013).

105. Furniss, J. J. *et al.* Proteasome-associated HECT-type ubiquitin ligase activity is required for plant immunity. *PLoS Pathog.* **14**, 1–22 (2018).

106. Santner, A., Calderon-Villalobos, L. I. A. & Estelle, M. Plant hormones are versatile chemical regulators of plant growth. *Nat. Chem. Biol.* **5**, 301–307 (2009).

107. Fülöp, K. *et al.* Arabidopsis Anaphase-Promoting Complexes: Multiple Activators and Wide Range of Substrates Might Keep APC Perpetually Busy. *Cell Cycle* **4**, 1084–1092 (2005).

108. Azevedo, C., Santos-Rosa, M. J. & Shirasu, K. The U-box protein family in plants. *Trends Plant Sci.* **6**, 354–358 (2001).

109. Yee, D. & Goring, D. R. The diversity of plant U-box E3 ubiquitin ligases: from upstream activators to downstream target substrates. *J. Exp. Bot.* **60**, 1109–1121 (2009).

110. Salt, J. N., Yoshioka, K., Moeder, W. & Goring, D. R. Altered Germination and Subcellular Localization Patterns for PUB44 / SAUL1 in Response to Stress and Phytohormone Treatments. *PLoS ONE* **6**, e21321 (2011).

111. Turek, I., Tischer, N., Lassig, R. & Trujillo, M. Multi-tiered pairing selectivity between E2 ubiquitin-conjugating enzymes and E3 ligases. *J. Biol. Chem.* **293**, 16324–16336 (2018).

112. Mudgil, Y., Shiu, S.-H., Stone, S. L., Salt, J. N. & Goring, D. R. A Large Complement of the Predicted Arabidopsis ARM Repeat Proteins Are Members of the U-Box E3 Ubiquitin Ligase Family. *Plant Physiol.* **134**, 59–66 (2004).

113. Drechsel, G. *et al.* C-terminal armadillo repeats are essential and sufficient for association of the plant U-box armadillo E3 ubiquitin ligase SAUL1 with the plasma membrane. *J. Exp. Bot.* **62**, 775–785 (2010).

114. Jung, C. *et al.* PLANT U-BOX PROTEIN10 regulates MYC2 stability in Arabidopsis. *Plant Cell* **27**, 2016–2031 (2015).

115. Lee, D., Kim, M. & Cho, K. H. A design principle underlying the paradoxical roles of E3 ubiquitin ligases. *Sci. Rep.* **4**, 1–13 (2014).

116. Furlan, G. *et al.* Changes in PUB22 ubiquitination modes triggered by MITOGEN-ACTIVATED PROTEIN KINASE3 dampen the immune response. *Plant Cell* **29**, 726–745 (2017).

117. Wang, J. *et al.* A Regulatory Module Controlling Homeostasis of a Plant Immune Kinase. *Mol. Cell* **69**, 493-504.e6 (2018).

118. Zhou, B. & Zeng, L. Conventional and unconventional ubiquitination in plant immunity. *Mol. Plant Pathol.* **18**, 1313–1330 (2017).

119. Gu, Y. & Innes, R. W. The KEEP ON GOING Protein of Arabidopsis Recruits the ENHANCED DISEASE RESISTANCE1 Protein to Trans-Golgi Network/Early Endosome Vesicles. *Plant Physiol.* **155**, 1827–1838 (2011).

120. Zou, B., Yang, D. L., Shi, Z., Dong, H. & Hua, J. Monoubiquitination of Histone 2B at the Disease Resistance Gene Locus Regulates Its Expression and Impacts Immune Responses in Arabidopsis. *Plant Physiol.* **165**, 309–318 (2014).

121. Zhang, L., Du, L., Shen, C., Yang, Y. & Poovaiah, B. W. Regulation of plant immunity through ubiquitin-mediated modulation of Ca2+-calmodulin-AtSR1/CAMTA3 signaling. *Plant J.* **78**, 269–281 (2014).

122. Chen, H. *et al.* A Bacterial Type III Effector Targets the Master Regulator of Salicylic Acid Signaling, NPR1, to Subvert Plant Immunity. *Cell Host Microbe* **22**, 777-788.e7 (2017).

123. Raab, S. *et al.* Identification of a novel E3 ubiquitin ligase that is required for suppression of premature senescence in Arabidopsis. *Plant J.* **3**, 39–51 (2009).

124. Zhou, J. *et al.* The dominant negative ARM domain uncovers multiple functions of PUB13 in Arabidopsis immunity, flowering, and senescence. *J. Exp. Bot.* **66**, 3353–3366 (2015).

125. Trujillo, M., Ichimura, K., Casais, C. & Shirasu, K. Negative Regulation of PAMP-Triggered Immunity by an E3 Ubiquitin Ligase Triplet in Arabidopsis. *Curr. Biol.* **18**, 1396–1401 (2008).

126. Mbengue, M. *et al.* Clathrin-dependent endocytosis is required for immunity mediated by pattern recognition receptor kinases. *Proc. Natl. Acad. Sci. U. S. A.* **113**, 11034–11039 (2016).

127. Liao, D. *et al.* Arabidopsis E3 ubiquitin ligase PLANT U-BOX13 (PUB13) regulates chitin receptor LYSIN MOTIF RECEPTOR KINASE5 (LYK5) protein abundance. *New Phytol.* **214**, 1646–1656 (2017).

128. Lu, D. *et al.* Direct Ubiquitination of Pattern Recognition Receptor FLS2 Attenuates Plant Innate Immunity. *Science* **332**, 1439–1442 (2011).

129. Stegmann, M. *et al.* The Ubiquitin Ligase PUB22 Targets a Subunit of the Exocyst Complex Required for PAMP-Triggered Responses in Arabidopsis. *Plant Cell* **24**, 4703–4716 (2012).

130. He, Q. *et al.* U-box E3 ubiquitin ligase PUB17 acts in the nucleus to promote specific immune pathways triggered by Phytophthora infestans. *J. Exp. Bot.* **66**, 3189–3199 (2015).

131. Hoth, S. *et al.* Genome-wide gene expression profiling in Arabidopsis thaliana reveals new targets of abscisic acid and largerly impaired gene regulation in the abi1-1 mutant. *J. Cell Sci.* **115**, 4891–4900 (2002).

132. Disch, E. *et al.* Membrane-Associated Ubiquitin Ligase SAUL1 Suppresses Temperature- and Humidity-Dependent Autoimmunity in Arabidopsis. *Mol. Plant Microbe Interact.* **29**, 69–80 (2016).

133. Wang, Y., Zhang, Y., Wang, Z., Zhang, X. & Yang, S. A missense mutation in CHS1, a TIR-NB protein, induces chilling sensitivity in Arabidopsis. *Plant J.* **75**, 553–565 (2013).

134. Zbierzak, A. M. *et al.* A TIR-NBS protein encoded by Arabidopsis Chilling Sensitive 1 (CHS1) limits chloroplast damage and cell death at low temperature. *Plant J.* **75**, 539–552 (2013).

135. Vogelmann, K. *et al.* Early Senescence and Cell Death in Arabidopsis saul1 Mutants Involves the PAD4 -Dependent Salicylic Acid Pathway. *Plant Physiol.* **159**, 1477–1487 (2012).

136. Li, X., Clarke, J. D., Zhang, Y. & Dong, X. Activation of an EDS1-Mediated R -Gene Pathway in the snc1 Mutant Leads to Constitutive, NPR1-Independent Pathogen Resistance. *Mol. Plant Microbe Interact.* **14**, 1131–1139 (2001).

137. Tong, M. *et al.* E3 ligase SAUL1 serves as a positive regulator of PAMP-triggered immunity and its homeostasis is monitored by immune receptor SOC3. *New Phytol.* **215**, 1516–1532

(2017).

138. Liang, W., van Wersch, S., Tong, M. & Li, X. TIR-NB-LRR immune receptor SOC3 pairs with truncated TIR-NB protein CHS1 or TN2 to monitor the homeostasis of E3 ligase SAUL1. *New Phytol.* **221**, 2054–2066 (2019).

139. Vogelmann, K. *et al.* Plasma membrane-association of SAUL1-type plant U-box armadillo repeat proteins is conserved in land plants. *Front. Plant Sci.* **5**, 1–8 (2014).

140. Tao, K., Waletich, J. R., Wise, H., Arredondo, F. & Tyler, B. M. Tethering of Multi-Vesicular Bodies and the Tonoplast to the Plasma Membrane in Plants. *Front. Plant Sci.* **10**, 1–23 (2019).

141. Kotur, T. Die Funktion der Gene Suppressor of saul1-1 1 und SNOWY COTYLEDON4 in Arabidopsis thaliana. (Universität Hamburg, 2016).

142. Brieske, C. Proteinbiochemische und molekularbiologische Analysen pflanzlicher PUB-ARM E3 Ubiquitinligasen. (Universität Hamburg, 2017).

143. Hua, J., Grisafi, P., Cheng, S. H. & Fink, G. R. Plant growth homeostasis is controlled by the arabidopsis BON1 and BAP1 genes. *Genes Dev.* **15**, 2263–2272 (2001).

144. Tomsig, J. L. & Creutz, C. E. Biochemical Characterization of Copine: A Ubiquitous Ca2+-Dependent, Phospholipid-Binding Protein. *Biochemistry* **39**, 16163–16175 (2000).

145. Brose, N., Petrenko, A., Südhof, T. C. & Jahn, R. Synaptotagmin: A Calcium Sensor on the Synaptic Vesicle Surface. *Science* **256**, 1021–1025 (1992).

146. Shao, X. *et al.* Synaptotagmin–Syntaxin Interaction: The C2 Domain as a Ca2+-Dependent Electrostatic Switch. *Neuron* **18**, 133–142 (1997).

147. Jambunathan, N. & McNellis, T. W. Regulation of Arabidopsis COPINE 1 Gene Expression in Response to Pathogens and Abiotic Stimuli. *Plant Physiol.* **132**, 1370–1381 (2003).

148. Yang, S. & Hua, J. A haplotype-specific Resistance gene regulated by BONZAI1 mediates temperaturea-dependent growth control in Arabidopsis. *Plant Cell* **16**, 1060–1071 (2004).

149. Gou, M. *et al.* The F-box protein CPR1/CPR30 negatively regulates R protein SNC1 accumulation. *Plant J.* **69**, 411–420 (2012).

150. Yang, H., Yang, S., Li, Y. & Hua, J. The Arabidopsis BAP1 and BAP2 Genes Are General Inhibitors of Programmed Cell Death. *Plant Physiol.* **145**, 135–146 (2007).

151. Li, Y., Gou, M., Sun, Q. & Hua, J. Requirement of Calcium Binding, Myristoylation, and Protein-Protein Interaction for the Copine BON1 Function in Arabidopsis. *J. Biol. Chem.* **285**, 29884–29891 (2010).

152. Wang, Z., Meng, P., Zhang, X., Ren, D. & Yang, S. BON1 interacts with the protein kinases BIR1 and BAK1 in modulation of temperature-dependent plant growth and cell death in Arabidopsis. *Plant J.* **67**, 1081–1093 (2011).

153. Hanahan, D. Studies on Transformation of Escherichia coli with Plasmids. *J. Mol. Biol.* **0**, 557–580 (1983).

154. Grefen, C. & Blatt, M. R. A 2in1 cloning system enables ratiometric bimolecular fluorescence complementation (rBiFC). *BioTechniques* **53**, 311–314 (2012).

155. Gerlach, E.-M. Physiological , molecular and ultrastructural properties define saul1 as an autoimmune mutant. (Universität Hamburg, 2018).

156. Meyer, S. P. Analysis of putative up- and downstream components of SAUL1 signaling in immune responses. (Universität Hamburg, 2018).

157. Marusoi, S. In search of new components of SAUL1 function and autoimmunity. (Universität Hamburg, 2017). doi:10.1192/bjp.112.483.211-a

158. The Salk Institute for Biological Studies. The Salk Institute Genomic Analysis Laboritory (SIGnAL). (2005). Available at: http://signal.salk.edu/tdnaprimers.2.html.

159. Hruz, T. *et al.* Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes. *Adv. Bioinformatics* **2008**, 1–5 (2008).

160. Untergasser, A. *et al.* Primer3-new capabilities and interfaces. *Nucleic Acids Res.* **40**, 1–12 (2012).

161. Hellemans, J., Mortier, G., De Paepe, A., Speleman, F. & Vandesompele, J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biol.* **8**, (2007).

162. Berardini, T. Z. *et al.* The Arabidopsis Information resource: Making and Mining the 'Gold Standard' Annotated Reference Plant Genome. *Genesis* **53**, 474–485 (2015).

163. The Arabidopsis Information Resource (TAIR). Available at: https://www.arabidopsis.org/download/index-auto.jsp?dir=%2Fdownload_files%2FGenes%2FTAIR10_genome_release. (Accessed: 17th March 2017)

164. Dobin, A. *et al.* STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

165. Dobin, A. & Gingeras, T. R. Mapping RNA-seq Reads with STAR. *Curr. Protoc. Bioinformatics* **51**, 11.14.1-11.14.19 (2015).

166. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

167. Shannon, P. *et al.* Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* 2498–2504 (2003). doi:10.1101/gr.1239303.metabolite

168. Bindea, G. *et al.* ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics* **25**, 1091–1093 (2009).

169. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000).

170. Carbon, S. *et al.* Expansion of the Gene Ontology knowledgebase and resources. *Nucleic Acids Res.* **45**, D331–D338 (2017).

171. Kanehisa, M. & Goto, S. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

172. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462 (2016).

173. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).

174. Cai, Q. *et al.* The disease resistance protein SNC1 represses the biogenesis of microRNAs and phased siRNAs. *Nat. Commun.* **9**, 1–14 (2018).

175. Wang, Z. *et al.* Arabidopsis ZED1-related kinases mediate the temperature-sensitive intersection of immune response and growth homeostasis. *New Phytol.* **215**, 711–724 (2017).

176. Mine, A. *et al.* The Defense Phytohormone Signaling Network Enables Rapid, High-amplitude Transcriptional Reprogramming During Effector-Triggered Immunity. *Plant Cell* **30**, tpc.00970.2017 (2018).

177. Clauw, P. *et al.* Leaf Responses to Mild Drought Stress in Natural Variants of Arabidopsis. *Plant Physiol.* **167**, 800–816 (2015).

178. Schlaen, R. G. *et al.* The spliceosome assembly factor GEMIN2 attenuates the effects of

temperature on alternative splicing and circadian rhythms. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 9382–9387 (2015).

179. Rawat, V. *et al.* Improving the Annotation of Arabidopsis lyrata Using RNA-Seq Data. *PLoS ONE* **10**, 1–12 (2015).

180. Liu, T. Y. *et al.* Identification of plant vacuolar transporters mediating phosphate storage. *Nat. Commun.* **11095**, 1–11 (2016).

181. Xu, E., Vaahtera, L. & Brosché, M. Roles of Defense Hormones in the Regulation of Ozone-Induced Changes in Gene Expression and Cell Death. *Mol. Plant* **8**, 1776–1794 (2015).

182. Zhan, X. *et al.* An Arabidopsis PWI and RRM motif-containing protein is critical for pre-mRNA splicing and ABA responses. *Nat. Commun.* **8139**, 1–12 (2015).

183. Li, B. *et al.* Phosphorylation of Trihelix Transcriptional Repressor ASR3 by MAP KINASE4 Negatively Regulates Arabidopsis Immunity. *Plant Cell* **27**, 839–856 (2015).

184. Liu, M.-J. *et al.* Determinants of nucleosome positioning and their influence on plant gene expression. *Genome Res.* **25**, 1182–1195 (2015).

185. Howard, B. E. *et al.* High-Throughput RNA Sequencing of Pseudomonas-Infected Arabidopsis Reveals Hidden Transcriptome Complexity and Novel Splice Variants. *PLoS ONE* **8**, e74183 (2013).

186. Maekawa, T., Kracher, B., Vernaldi, S., Ver Loren van Themaat, E. & Schulze-Lefert, P. Conservation of NLR-triggered immunity across plant lineages. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 20119–20123 (2012).

187. Blanvillain-Baufumé, S., Parker, J. E., Huettel, B., Srinivasan, A. & Jiminez-Gomez, J. Investigating putative RPS4-chromatin associations during Arabidopsis plant disease resistance, using a temperature-inducible system which resembles pathogen-infection. *Gene Expr. Omnibus* **GSE40216**, (2012).

188. Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Second Int. Conf. Intell. Syst. Mol. Biol.* **1**, 28–36 (1994).

189. Gupta, S., Stamatoyannopoulos, J. A., Bailey, T. L. & Noble, W. S. Quantifying similarity between motifs. *Genome Biol.* **8**, (2007).

190. Bailey, T. L. *et al.* MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.* **37**, 202–208 (2009).

191. Pérez-Rodríguez, P. *et al.* PlnTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res.* **38**, 822–827 (2010).

192. Jin, J., Zhang, H., Kong, L., Gao, G. & Luo, J. PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. *Nucleic Acids Res.* **42**, 1182–1187 (2014).

193. Jin, J. *et al.* An Arabidopsis Transcriptional Regulatory Map Reveals Distinct Functional and Evolutionary Features of Novel Transcription Factors. *Mol. Biol. Evol.* **32**, 1767–1773 (2015).

194. Bhattarai, M. Untersuchungen zu Proteininteraktoren des positiven Immunregulators SAUL1. (Universität Hamburg, 2017).

195. Lindwall, G., Chau, M.-F., Gardner, S. R. & Kohlstaedt, L. A. A sparse matrix approach to the solubilization of overexpressed proteins. *Protein Eng. Des. Sel.* **13**, 67–71 (2000).

196. Dyballa, N. & Metzger, S. Fast and Sensitive Colloidal Coomassie G-250 Staining for Proteins in Polyacrylamide Gels. *J. Vis. Exp.* 2–5 (2009). doi:10.3791/1431

197. van den Heuvel, R. H. H. *et al.* Improving the performance of a quadrupole time-of-flight instrument for macromolecular mass spectrometry. *Anal. Chem.* **78**, 7473–7483 (2006).

198. Tahallah, N., Pinkse, M., Maier, C. S. & Heck, A. J. R. The effect of the source pressure on the

abundance of ions of noncovalent protein assemblies in an electrospray ionization orthogonal time-of-flight instrument. *Rapid Commun. Mass Spectrom.* **15**, 596–601 (2001).

199. Boivin, S., Kozak, S. & Meijers, R. Optimization of protein purification and characterization using Thermofluor screens. *Protein Expr. Purif.* **91**, 192–206 (2013).

200. Huynh, K. & Partch, C. L. Analysis of Protein Stability and Ligand Interactions by Thermal Shift Assay. *Curr. Protoc. Protein Sci.* **79**, 28.9.1-28.9.14 (2015).

201. Mitchell, A. L. *et al.* InterPro in 2019: improving coverage, classification and access to protein sequence annotations. *Nucleic Acids Res.* **47**, D351–D360 (2019).

202. Yang, J. & Zhang, Y. I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res.* **43**, W174–W181 (2015).

203. Zhang, C., Freddolino, P. L. & Zhang, Y. COFACTOR: Improved protein function prediction by combining structure, sequence and protein-protein interaction information. *Nucleic Acids Res.* **45**, W291–W299 (2017).

204. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. E. J. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).

205. Kelly, S. M., Jess, T. J. & Price, N. C. How to study proteins by circular dichroism. *Biochim. Biophys. Acta - Proteins Proteomics* **1751**, 119–139 (2005).

206. Flory, P. J. *Principles of polymer chemistry*. (Cornell University Press, 1953).

207. Blanchet, C. E. *et al.* Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J. Appl. Crystallogr.* **48**, 431–443 (2015).

208. Franke, D., Kikhney, A. G. & Svergun, D. I. Automated acquisition and analysis of small angle X-ray scattering data. *Nucl. Instrum. Methods Phys. Res. A* **689**, 52–59 (2012).

209. Petoukhov, M. V., Konarev, P. V., Kikhney, A. G. & Svergun, D. I. ATSAS 2.1 - Towards automated and web-supported small-angle scattering data analysis. *J. Appl. Crystallogr.* **40**, 223–228 (2007).

210. Franke, D. *et al.* ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225 (2017).

211. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: a Windows PC-based system for small- angle scattering data analysis. *J. Appl. Crystallogr.* **36**, 1277–1282 (2003).

212. Franke, D. & Svergun, D. I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **42**, 342–346 (2009).

213. Svergun, D. I., Petoukhov, M. V. & Koch, M. H. J. Determination of Domain Structure of Proteins from X-Ray Solution Scattering. *Biophys. J.* **80**, 2946–53 (2001).

214. Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).

215. Volkov, V. V & Svergun, D. I. Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* **36**, 860–864 (2003).

216. Panjkovich, A. & Svergun, D. I. Deciphering conformational transitions of proteins by small angle X-ray scattering and normal mode analysis. *Phys. Chem. Chem. Phys.* **18**, 5707–5719 (2016).

217. Bernadó, P., Mylonas, E., Petoukhov, M. V., Blackledge, M. & Svergun, D. I. Structural Characterization of Flexible Proteins Using Small-Angle X-ray Scattering. *J. Am. Chem. Soc.* **129**, 5656–5664 (2007).

218. Tria, G., Mertens, H. D. T., Kachala, M. & Svergun, D. I. Advanced ensemble modelling of

flexible macromolecules using X-ray solution scattering. *IUCrJ* **2**, 207–217 (2015).

219. Svergun, D. I., Barberato, C. & Koch, M. H. J. CRYSOL– a Program to Evaluate X-ray Solution Scattering of Biological Macromolecules from Atomic Coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).

220. Mylonas, E. & Svergun, D. I. Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J. Appl. Crystallogr.* **40**, 245–249 (2007).

221. Panjkovich, A. & Svergun, D. I. CHROMIXS: automatic and interactive analysis of chromatography-coupled small-angle X-ray scattering data. *Bioinformatics* **34**, 1944–1946 (2018).

222. Dolinsky, T. J., Nielsen, J. E., McCammon, J. A. & Baker, N. A. PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.* **32**, 665–667 (2004).

223. Jurrus, E. *et al.* Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**, 112–128 (2018).

224. Schrödinger LCC. The PyMOL Molecular Graphics System, Version 2.1. (2018).

225. Zipfel, C. *et al.* Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts Agrobacterium-Mediated Transformation. *Cell* **125**, 749–760 (2006).

226. Zentgraf, U., Laun, T. & Miao, Y. The complex regulation of WRKY53 during leaf senescence of Arabidopsis thaliana. *Eur. J. Cell Biol.* **89**, 133–137 (2010).

227. Yu, D., Chen, C. & Chen, Z. Evidence for an Important Role of WRKY DNA Binding Proteins in the Regulation of NPR1 Gene Expression. *Plant Cell* **13**, 1527–1539 (2001).

228. Murray, S. L., Ingle, R. A., Petersen, L. N. & Denby, K. J. Basal Resistance Against Pseudomonas syringae in Arabidopsis Involves WRKY53 and a Protein with Homology to a Nematode Resistance Protein. *Mol. Plant Microbe Interact.* **20**, 1431–1438 (2007).

229. Wildermuth, M. C., Dewdney, J., Wu, G. & Ausubel, F. M. Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature* **414**, 562–565 (2001).

230. Feys, B. J. & Parker, J. E. Interplay of signaling pathways in plant disease resistance. *Trends Genet.* **16**, 449–455 (2000).

231. Kim, M. G., Geng, X., Lee, S. Y. & Mackey, D. The Pseudomonas syringae type III effector AvrRpm1 induces significant defenses by activating the Arabidopsis nucleotide-binding leucine-rich repeat protein RPS2. *Plant J.* **57**, 645–653 (2009).

232. Lewis, J. D. *et al.* The Arabidopsis ZED1 pseudokinase is required for ZAR1-mediated immunity induced by the Pseudomonas syringae type III effector HopZ1a. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 18722–18727 (2013).

233. Mang, H.-G. *et al.* Abscisic Acid Deficiency Antagonizes High-Temperature Inhibition of Disease Resistance through Enhancing Nuclear Accumulation of Resistance Proteins SNC1 and RPS4 in Arabidopsis. *Plant Cell* **24**, 1271–1284 (2012).

234. Tada, Y. *et al.* Plant Immunity Requires Conformational Changes of NPR1 via S-Nitrosylation and Thioredoxins. *Science* **321**, 952–956 (2008).

235. Baruah, A., Šimková, K., Apel, K. & Laloi, C. Arabidopsis mutants reveal multiple singlet oxygen signaling pathways involved in stress response and development. *Plant Mol. Biol.* **70**, 547–563 (2009).

236. Šimková, K. *et al.* Integration of stress-related and reactive oxygen species-mediated signals by Topoisomerase VI in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 16360–16365 (2012).

237. Hu, Y., Dong, Q. & Yu, D. Arabidopsis WRKY46 coordinates with WRKY70 and WRKY53 in

basal resistance against pathogen Pseudomonas syringae. *Plant Sci.* **185–186**, 288–297 (2012).

238. Hamamouch, N. *et al.* The interaction of the novel 30C02 cyst nematode effector protein with a plant b-1,3-endoglucanase may suppress host defence to promote parasitism. *J. Exp. Bot.* **63**, 3683–3696 (2012).

239. Zhang, X. & Guo, H. mRNA decay in plants: both quantity and quality matter. *Curr. Opin. Plant Biol.* **35**, 138–144 (2017).

240. Addou, S., Rentzsch, R., Lee, D. & Orengo, C. A. Domain-Based and Family-Specific Sequence Identity Thresholds Increase the Levels of Reliable Protein Function Transfer. *J. Mol. Biol.* **387**, 416–430 (2009).

241. Fujimoto, S. Y., Ohta, M., Usui, A., Shinshi, H. & Ohme-Takagi, M. Arabidopsis Ethylene-Responsive Element Binding Factors Act as Transcriptional Activators or Repressors of GCC Box-Mediated Gene Expression. *Plant Cell* **12**, 393 (2000).

242. Brown, R. L., Kazan, K., McGrath, K. C., Maclean, D. J. & Manners, J. M. A Role for the GCC-Box in Jasmonate-Mediated Activation of the PDF1.2 Gene in Arabidopsis. *Plant Physiol.* **132**, 1020–1032 (2003).

243. Mitsuya, Y. *et al.* Spermine signaling plays a significant role in the defense response of Arabidopsis thaliana to cucumber mosaic virus. *J. Plant Physiol.* **166**, 626–643 (2009).

244. Feng, Z. *et al.* Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 4632–4637 (2014).

245. O'Malley, R. C., Barragan, C. C. & Ecker, J. R. A User's Guide to the Arabidopsis T-DNA Insertion Mutant Collections. in *Plant Functional Genomics: Methods and Protocols* **1284**, 323–342 (2015).

246. Spoel, S. H. *et al.* NPR1 Modulates Cross-Talk between Salicylate- and Jasmonate-Dependent Defense Pathways through a Novel Function in the Cytosol Steven. *Plant Cell* **15**, 760–770 (2003).

247. Doares, S. H., Narvaez-Vasquez, J., Conconi, A. & Ryan, C. A. Salicylic Acid Inhibits Synthesis of Proteinase Inhibitors in Tomato Leaves Induced by Systemin and Jasmonic Acid. *Plant Physiol.* **108**, 1741–1746 (1995).

248. Spoel, S. H., Johnson, J. S. & Dong, X. Regulation of tradeoffs between plant defenses against pathogens with different lifestyles. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 18842–18847 (2007).

249. Leon-Reyes, A. *et al.* Ethylene Signaling Renders the Jasmonate Response of Arabidopsis Insensitive to Future Suppression by Salicylic Acid. *Mol. Plant Microbe Interact.* **23**, 187–197 (2010).

250. de Ilarduya, O. M., Xie, Q. & Kaloshian, I. Aphid-Induced Defense Responses in Mi-1-Mediated Compatible and Incompatible Tomato Interactions. *Mol. Plant Microbe Interact.* **16**, 699–708 (2003).

251. Tamaoki, D. *et al.* Jasmonic acid and salicylic acid activate a common defense system in rice. *Plant Signal. Behav.* **8**, (2013).

252. Mur, L. A. J., Kenton, P., Atzorn, R., Miersch, O. & Wasternack, C. The Outcomes of Concentration-Specific Interactions between Salicylate and Jasmonate Signaling Include Synergy, Antagonism, and Oxidative Stress Leading to Cell Death. *Plant Physiol.* **140**, 249–262 (2006).

253. Broekgaarden, C., Caarls, L., Vos, I. A., Pieterse, C. M. J. & Van Wees, S. C. M. Ethylene: Traffic Controller on Hormonal Crossroads to Defense. *Plant Physiol.* **169**, 2371–2379 (2015).

254. Šašek, V. *et al.* Recognition of Avirulence Gene AvrLm1 from Hemibiotrophic Ascomycete

Leptosphaeria maculans Triggers Salicylic Acid and Ethylene Signaling in Brassica napus. *Mol. Plant Microbe Interact.* **25**, 1238–1250 (2012).

255. Washington, E. J. *et al.* Pseudomonas syringae type III effector HopAF1 suppresses plant immunity by targeting methionine recycling to block ethylene induction. *Proc. Natl. Acad. Sci. U. S. A.* **113**, E3577–E3586 (2016).

256. Mur, L. A. J. *et al.* Biphasic ethylene production during the hypersensitive response in Arabidopsis. *Plant Signal. Behav.* **4**, 610–613 (2009).

257. Fan, J., Hill, L., Crooks, C., Doerner, P. & Lamb, C. Abscisic Acid Has a Key Role in Modulating Diverse Plant-Pathogen Interactions. *Plant Physiol.* **150**, 1750–1761 (2009).

258. Jensen, M. K. *et al.* Transcriptional regulation by an NAC (NAM-ATAF1,2-CUC2) transcription factor attenuates ABA signalling for efficient basal defence towards Blumeria graminis f. sp. hordei in Arabidopsis. *Plant J.* **56**, 867–880 (2008).

259. De Torres Zabala, M., Bennett, M. H., Truman, W. H. & Grant, M. R. Antagonism between salicylic and abscisic acid reflects early host-pathogen conflict and moulds plant defence responses. *Plant J.* **59**, 375–386 (2009).

260. Anderson, J. P. Antagonistic Interaction between Abscisic Acid and Jasmonate-Ethylene Signaling Pathways Modulates Defense Gene Expression and Disease Resistance in Arabidopsis. *Plant Cell* **16**, 3460–3479 (2004).

261. Yasuda, M. *et al.* Antagonistic Interaction between Systemic Acquired Resistance and the Abscisic Acid-Mediated Abiotic Stress Response in Arabidopsis. *Plant Cell* **20**, 1678–1692 (2008).

262. Jiang, C.-J. *et al.* Abscisic Acid Interacts Antagonistically with Salicylic Acid Signaling Pathway in Rice–Magnaporthe grisea Interaction. *Mol. Plant Microbe Interact.* **23**, 791–798 (2010).

263. Lievens, L., Pollier, J., Goossens, A., Beyaert, R. & Staal, J. Abscisic Acid as Pathogen Effector and Immune Regulator. *Front. Plant Sci.* **8**, 1–15 (2017).

264. Ton, J. & Mauch-Mani, B. β-amino-butyric acid-induced resistance against necrotrophic pathogens is based on ABA-dependent priming for callose. *Plant J.* **38**, 119–130 (2004).

265. Mauch-Mani, B. & Mauch, F. The role of abscisic acid in plant-pathogen interactions. *Curr. Opin. Plant Biol.* **8**, 409–414 (2005).

266. Cao, F. Y. *et al.* Arabidopsis ETHYLENE RESPONSE FACTOR 8 (ERF8) has dual functions in ABA signaling and immunity. *BMC Plant Biol.* **18**, 1–16 (2018).

267. Mohr, P. G. & Cahill, D. M. Suppression by ABA of salicylic acid and lignin accumulation and the expression of multiple genes, in Arabidopsis infected with Pseudomonas syringae pv. tomato. *Funct. Integr. Genomics* **7**, 181–191 (2007).

268. Hernandez-Blanco, C. *et al.* Impairment of Cellulose Synthases Required for Arabidopsis Secondary Cell Wall Formation Enhances Disease Resistance. *Plant Cell* **19**, 890–903 (2007).

269. Hauck, P., Thilmony, R. & He, S. Y. A Pseudomonas syringae type III effector suppresses cell wall-based extracellular defense in susceptible Arabidopsis plants. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 8577–8582 (2003).

270. Gu, K. *et al.* R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435**, 1122–1125 (2005).

271. Elmore, J. M., Liu, J., Smith, B., Phinney, B. & Coaker, G. Quantitative Proteomics Reveals Dynamic Changes in the Plasma Membrane During Arabidopsis Immune Signaling. *Mol. Cell. Proteomics* **11**, (2012).

272. Häweker, H. *et al.* Pattern Recognition Receptors Require N-Glycosylation to Mediate Plant

Immunity. *J. Biol. Chem.* **285**, 4629–4636 (2010).

273. Le Roy, J., Huss, B., Creach, A., Hawkins, S. & Neutelings, G. Glycosylation Is a Major Regulator of Phenylpropanoid Availability and Biological Activity in Plants. *Front. Plant Sci.* **7**, (2016).

274. Mikkelsen, M. D. *et al.* Modulation of CYP79 Genes and Glucosinolate Profiles in Arabidopsis by Defense Signaling Pathways. *Plant Physiol.* **131**, 298–308 (2003).

275. Kong, W. *et al.* Two Novel Flavin-Containing Monooxygenases Involved in Biosynthesis of Aliphatic Glucosinolates. *Front. Plant Sci.* **7**, 1–9 (2016).

276. Zhang, X. *et al.* The Transcription Factor MYB29 Is a Regulator of ALTERNATIVE OXIDASE1a. *Plant Physiol.* **173**, 1824–1843 (2017).

277. Beekweelder, J. *et al.* The Impact of the Absence of Aliphatic Glucosinolates on Insect Herbivory in Arabidopsis. *PLoS ONE* **3**, (2008).

278. Francisco, M. *et al.* The Defense Metabolite, Allyl Glucosinolate, Modulates Arabidopsis thaliana Biomass Dependent upon the Endogenous Glucosinolate Pathway. *Front. Plant Sci.* **7**, 1–14 (2016).

279. Hammerschmidt, R. PHYTOALEXINS: What Have We Learned After 60 Years? *Annu. Rev. Phytopathol.* **37**, 285–306 (1999).

280. Tao, Y. *et al.* Quantitative Nature of Arabidopsis Responses during Compatible and Incompatible Interactions with the Bacterial Pathogen Pseudomonas syringae. *Plant Cell* **15**, 317–330 (2003).

281. Schuhegger, R. *et al.* CYP71B15 (PAD3) Catalyzes the Final Step in Camalexin Biosynthesis. *Plant Physiol.* **141**, 1248–1254 (2006).

282. Ahuja, I., Kissen, R. & Bones, A. M. Phytoalexins in defense against pathogens. *Trends Plant Sci.* **17**, 73–90 (2012).

283. Birkenbihl, R. P., Liu, S. & Somssich, I. E. Transcriptional events defining plant immune responses. *Curr. Opin. Plant Biol.* **38**, 1–9 (2017).

284. Vignutelli, A., Wasternack, C., Apel, K. & Bohlmann, H. Systemic and local induction of an Arabidopsis thionin gene by wounding and pathogens. *Plant J.* **14**, 285–295 (1998).

285. McGrath, K. C. *et al.* Repressor- and Activator-Type Ethylene Response Factors Functioning in Jasmonate Signaling and Disease Resistance Identified via a Genome-Wide Screen of Arabidopsis Transcription Factor Gene Expression. *Plant Physiol.* **139**, 949–959 (2005).

286. Pastor, V. *et al.* Detection, characterization and quantification of salicylic acid conjugates in plant extracts by ESI tandem mass spectrometric techniques. *Plant Physiol. Biochem.* **53**, 19–26 (2012).

287. Rushton, P. J. *et al.* Interaction of elicitor-induced DNA-binding proteins with elicitor response elements in the promoters of parsley PR1 genes. *EMBO J.* **15**, 5690–5700 (2018).

288. Rushton, P. J., Somssich, I. E., Ringler, P. & Shen, Q. J. WRKY transcription factors. *Trends Plant Sci.* **15**, 247–258 (2010).

289. Jiang, Y. & Deyholos, M. K. Functional characterization of Arabidopsis NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. *Plant Mol. Biol.* **69**, 91–105 (2009).

290. Johnson, C. S., Kolevski, B. & Smyth, D. R. TRANSPARENT TESTA GLABRA2 , a Trichome and Seed Coat Development Gene of Arabidopsis, Encodes a WRKY Transcription Factor. *Plant Cell* **14**, 1359–1375 (2002).

291. Ren, C. M. *et al.* Transcription Factor WRKY70 Displays Important but No Indispensable Roles in Jasmonate and Salicylic Acid Signaling. *J. Integr. Plant Biol.* **50**, 630–637 (2008).

292. Pandey, S. P. & Somssich, I. E. The Role of WRKY Transcription Factors in Plant Immunity. *Plant Physiol.* **150**, 1648–1655 (2009).

293. Yan, L. *et al.* Auto- and Cross-repression of Three Arabidopsis WRKY Transcription Factors WRKY18, WRKY40, and WRKY60 Negatively Involved in ABA Signaling. *J. Plant Growth Regul.* **32**, 399–416 (2013).

294. Wang, D., Amornsiripanitch, N. & Dong, X. A Genomic Approach to Identify Regulatory Nodes in the Transcriptional Network of Systemic Acquired Resistance in Plants. *PLoS Pathog.* **2**, 1042–1050 (2006).

295. Li, J., Zhong, R. & Palva, E. T. WRKY70 and its homolog WRKY54 negatively modulate the cell wall-associated defenses to necrotrophic pathogens in Arabidopsis. *PLoS ONE* **12**, 1–22 (2017).

296. Li, J., Brader, G. & Palva, E. T. The WRKY70 Transcription Factor: A Node of Convergence for Jasmonate-Mediated and Salicylate-Mediated Signals in Plant Defense. *Plant Cell* **16**, 319–331 (2004).

297. Gao, Q.-M., Venugopal, S., Navarre, D. & Kachroo, A. Low Oleic Acid-Derived Repression of Jasmonic Acid-Inducible Defense Responses Requires the WRKY50 and WRKY51 Proteins. *Plant Physiol.* **155**, 464–476 (2010).

298. Liu, S., Kracher, B., Ziegler, J., Birkenbihl, R. P. & Somssich, I. E. Negative regulation of ABA Signaling By WRKY33 is critical for Arabidopsis immunity towards Botrytis cinerea 2100. *eLife* **4**, 4–27 (2015).

299. Olsen, A. N., Ernst, H. A., Leggio, L. Lo & Skriver, K. NAC transcription factors: structurally distinct, functionally diverse. *Trends Plant Sci.* **10**, 79–87 (2005).

300. Aida, M., Ishida, T., Fukaki, H., Fujisawa, H. & Tasaka, M. Genes Involved in Organ Separation in Arabidopsis: An Analysis of the cup-shaped cotyledon Mutant. *Plant Cell* **9**, 841–857 (1997).

301. Xie, Q., Frugis, G., Colgan, D. & Chua, N.-H. Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development. *Genes Dev.* **14**, 3024–3036 (2000).

302. Fujita, M. *et al.* A dehydration-induced NAC protein, RD26, is involved in a novel ABA-dependent stress-signaling pathway. *Plant J.* **39**, 863–876 (2004).

303. Lee, M. H., Jeon, H. S., Kim, H. G. & Park, O. K. An Arabidopsis NAC transcription factor NAC4 promotes pathogen-induced cell death under negative regulation by microRNA164. *New Phytol.* **214**, 343–360 (2017).

304. Lee, S. *et al.* The Arabidopsis NAC transcription factor NTL4 participates in a positive feedback loop that induces programmed cell death under heat stress conditions. *Plant Sci.* **227**, 76–83 (2014).

305. Seo, P. J. *et al.* Cold activation of a plasma membrane-tethered NAC transcription factor induces a pathogen resistance response in Arabidopsis. *Plant J.* **61**, 661–671 (2010).

306. Delessert, C. *et al.* The transcription factor ATAF2 represses the expression of pathogenesis-related genes in Arabidopsis. *Plant J.* **43**, 745–757 (2005).

307. Nagahage, I. S. *et al.* An NAC domain transcription factor ATAF2 acts as transcriptional activator or repressor dependent on promoter context. *Plant Biotechnol.* **35**, 285–289 (2018).

308. Bogamuwa, S. P. & Jang, J. C. Tandem CCCH Zinc Finger Proteins in Plant Growth, Development and Stress Response. *Plant Cell Physiol.* **55**, 1367–1375 (2014).

309. Li, H. *et al.* The CCCH-Type Zinc Finger Proteins AtSZF1 and AtSZF2 Regulate Salt Stress Responses in Arabidopsis. *Plant Cell Physiol.* **48**, 1148–1158 (2007).

310. Kanchiswamy, C. N. *et al.* Regulation of Arabidopsis defense responses against Spodoptera littoralis by CPK-mediated calcium signaling. *BMC Plant Biol.* **10**, 97 (2010).

311. Feller, A., MacHemer, K., Braun, E. L. & Grotewold, E. Evolutionary and comparative analysis of MYB and bHLH plant transcription factors. *Plant J.* **66**, 94–116 (2011).

312. Ohashi-Ito, K. & Fukuda, H. Functional mechanism of bHLH complexes during early vascular development. *Curr. Opin. Plant Biol.* **33**, 42–47 (2016).

313. Goossens, J., Mertens, J. & Goossens, A. Role and functioning of bHLH transcription factors in jasmonate signalling. *J. Exp. Bot.* **68**, 1333–1347 (2017).

314. Malinovsky, F. G. *et al.* Antagonistic Regulation of Growth and Immunity by the Arabidopsis Basic Helix-Loop-Helix Transcription Factor HOMOLOG OF BRASSINOSTEROID ENHANCED EXPRESSION2 INTERACTING WITH INCREASED LEAF INCLINATION1 BINDING bHLH1. *Plant Physiol.* **164**, 1443–1455 (2014).

315. Fan, M. *et al.* The bHLH Transcription Factor HBI1 Mediates the Trade-Off between Growth and Pathogen-Associated Molecular Pattern-Triggered Immunity in Arabidopsis. *Plant Cell* **26**, 828–841 (2014).

316. Jiménez-Góngora, T., Kim, S.-K., Lozano-Durán, R. & Zipfel, C. Flg22-Triggered Immunity Negatively Regulates Key BR Biosynthetic Genes. *Front. Plant Sci.* **6**, 1–7 (2015).

317. Lozano-Durán, R. & Zipfel, C. Trade-off between growth and immunity: role of brassinosteroids. *Trends Plant Sci.* **20**, 12–19 (2015).

318. Cifuentes-Esquivel, N. *et al.* The bHLH proteins BEE and BIM positively modulate the shade avoidance syndrome in Arabidopsis seedlings. *Plant J.* **75**, 989–1002 (2013).

319. Moreno, J. E., Moreno-Piovano, G. & Chan, R. L. The antagonistic basic helix-loop-helix partners BEE and IBH1 contribute to control plant tolerance to abiotic stress. *Plant Sci.* **271**, 143–150 (2018).

320. Wang, H. *et al.* Regulation of Arabidopsis Brassinosteroid Signaling by Atypical Basic Helix-Loop-Helix Proteins. *Plant Cell* **21**, 3781–3791 (2009).

321. Vragović, K. *et al.* Translatome analyses capture of opposing tissue-specific brassinosteroid signals orchestrating root meristem differentiation. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 923–928 (2015).

322. Imai, A. *et al.* The dwarf phenotype of the Arabidopsis acl5 mutant is suppressed by a mutation in an upstream ORF of a bHLH gene. *Development* **133**, 3575–3585 (2006).

323. Takano, A., Kakehi, J. I. & Takahashi, T. Thermospermine is Not a Minor Polyamine in the Plant Kingdom. *Plant Cell Physiol.* **53**, 606–616 (2012).

324. Cai, Q. *et al.* The SAC51 Family Plays a Central Role in Thermospermine Responses in Arabidopsis. *Plant Cell Physiol.* **57**, 1583–1592 (2016).

325. Castillon, A., Shen, H. & Huq, E. Phytochrome Interacting Factors: central players in phytochrome-mediated light signaling networks. *Trends Plant Sci.* **12**, 514–521 (2007).

326. Hauvermale, A. L., Ariizumi, T. & Steber, C. M. Gibberellin Signaling: A Theme and Variations on DELLA Repression. *Plant Physiol.* **160**, 83–92 (2012).

327. Oh, E., Zhu, J. Y. & Wang, Z. Y. Interaction between BZR1 and PIF4 integrates brassinosteroid and environmental responses. *Nat. Cell Biol.* **14**, 802–809 (2012).

328. Dietz, K. J., Vogel, M. O. & Viehhauser, A. AP2/EREBP transcription factors are part of gene regulatory networks and integrate metabolic, hormonal and environmental signals in stress acclimation and retrograde signalling. *Protoplasma* **245**, 3–14 (2010).

329. Pré, M. *et al.* The AP2/ERF Domain Transcription Factor ORA59 Integrates Jasmonic Acid and Ethylene Signals in Plant Defense. *Plant Physiol.* **147**, 1347–1357 (2008).

330. Warmerdam, S. *et al.* Mediator of tolerance to abiotic stress ERF6 regulates susceptibility of Arabidopsis to Meloidogyne incognita. *Mol. Plant Pathol.* **20**, 137–152 (2019).

331. Bethke, G. *et al.* Flg22 regulates the release of an ethylene response factor substrate from MAP kinase 6 in Arabidopsis thaliana via ethylene signaling. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 8067–8072 (2009).

332. Dubois, M. *et al.* The ETHYLENE RESPONSE FACTORs ERF6 and ERF11 Antagonistically Regulate Mannitol-Induced Growth Inhibition in Arabidopsis. *Plant Physiol.* **169**, 166–179 (2015).

333. Meng, X. *et al.* Phosphorylation of an ERF Transcription Factor by Arabidopsis MPK3/MPK6 Regulates Plant Defense Gene Induction and Fungal Resistance. *Plant Cell* **25**, 1126–1142 (2013).

334. Englbrecht, C. C., Schoof, H. & Böhm, S. Conservation, diversification and expansion of C2H2 zinc finger proteins in the Arabidopsis thaliana genome. *BMC Genom.* **5**, 1–17 (2004).

335. Ciftci-Yilmaz, S. & Mittler, R. The zinc finger network of plants. *Cell. Mol. Life Sci.* **65**, 1150–1160 (2008).

336. Kiełbowicz-Matuk, A. Involvement of plant C2H2-type zinc finger transcription factors in stress responses. *Plant Sci.* **185–186**, 78–85 (2012).

337. Cui, B. *et al.* S-nitrosylation of the zinc finger protein SRG1 regulates plant immunity. *Nat. Commun.* **9**, 4226 (2018).

338. van Wersch, R., Li, X. & Zhang, Y. Mighty Dwarfs: Arabidopsis Autoimmune Mutants and Their Usages in Genetic Dissection of Plant Immunity. *Front. Plant Sci.* **7**, 1–8 (2016).

339. Sendon, P. M. *et al.* Activation of C2H2-type Zinc Finger Genes Induces Dwarfism in Arabidopsis thaliana. *J. Korean Soc. Appl. Biol. Chem.* **57**, 35–41 (2014).

340. Ciftci-Yilmaz, S. *et al.* The EAR-motif of the Cys2/His2-type Zinc Finger Protein Zat7 Plays a Key Role in the Defense Response of Arabidopsis to Salinity Stress. *J. Biol. Chem.* **282**, 9260–9268 (2007).

341. Kodaira, K.-S. *et al.* Arabidopsis Cys2/His2 Zinc-Finger Proteins AZF1 and AZF2 Negatively Regulate Abscisic Acid-Repressive and Auxin-Inducible Genes under Abiotic Stress Conditions. *Plant Physiol.* **157**, 742–756 (2011).

342. O'Malley, R. C. *et al.* Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **166**, 1598 (2016).

343. Heo, J. B., Sung, S. & Assmann, S. M. Ca2ɦ -dependent GTPase, Extra-large G Protein 2 (XLG2), Promotes Activation of DNA-binding Protein Related to Vernalization 1 (RTV1), Leading to Activation of Floral Integrator Genes and Early Flowering in Arabidopsis. *J. Biol. Chem.* **287**, 8242–8253 (2012).

344. Mantegazza, O. *et al.* Analysis of the arabidopsis REM gene family predicts functions during flower development. *Ann. Bot.* **114**, 1507–1515 (2014).

345. Jeffries, C. M. *et al.* Preparing monodisperse macromolecular samples for successful biological small-angle X-ray and neutron-scattering experiments. *Nat. Protoc.* **11**, 2122–2153 (2016).

346. Kilani, H. El. Structural Characterization of Cell Wall and Plasma Membrane Proteins of Arabidopsis thaliana. (Universität Hamburg, 2016).

347. Reed, J. & Reed, T. A. A set of constructed type spectra for the practical estimation of peptide secondary structure from circular dichroism. *Anal. Biochem.* **254**, 36–40 (1997).

348. Blakeley, M. P., Hasnain, S. S. & Antonyuk, S. V. Sub-atomic resolution X-ray crystallography and neutron crystallography: promise, challenges and potential. *IUCrJ* **2**, 464–474 (2015).

349. Kikhney, A. G. & Svergun, D. I. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. *FEBS Lett.* **589**, 2570–2577 (2015).

350. Kilani, H. El & Kikhney, A. Senescence-associated ubiquitin ligase 1 (SAUL1). *SASBDB* (2017). Available at: https://www.sasbdb.org/project/315/vat47980g3/. (Accessed: 23rd July 2019)

351. Mathew, E., Mirza, A. & Menhart, N. Liquid-chromatography-coupled SAXS for accurate sizing of aggregating proteins. *J. Synchrotron Radiat.* **11**, 314–318 (2004).

352. Graewert, M. A. *et al.* Automated Pipeline for Purification, Biophysical and X-Ray Analysis of Biomacromolecular Solutions. *Sci. Rep.* **5**, 1–8 (2015).

353. Kohn, J. E. *et al.* Random-coil behavior and the dimensions of chemically unfolded proteins. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 12491–12496 (2004).

354. Bernadó, P. & Svergun, D. I. Structural analysis of intrinsically disordered proteins by small-angle X-ray scattering. *Mol. Omics* **8**, 151–167 (2012).

355. Papaleo, E. *et al.* The Role of Protein Loops and Linkers in Conformational Dynamics and Allostery. *Chem. Rev.* **116**, 6391–6423 (2016).

356. Zhang, Z. *et al.* Crystal structure of the armadillo repeat domain of adenomatous polyposis coli which reveals its inherent flexibility. *Biochem. Biophys. Res. Commun.* **412**, 732–736 (2011).

357. Huber, A. H., Nelson, W. J. & Weis, W. I. Three-Dimensional Structure of the Armadillo Repeat Region of β-Catenin. *Cell* **90**, 871–882 (1997).

358. Ritco-Vonsovici, M., Ababou, A. & Horton, M. Molecular plasticity of β-catenin: New insights from single-molecule measurements and MD simulation. *Protein Sci.* **16**, 1984–1998 (2007).

359. Samuel, M. A., Salt, J. N., Shiu, S. H. & Goring, D. R. Multifunctional Arm Repeat Domains in Plants. *Int. Rev. Cytol.* **253**, 1–26 (2006).

360. Antignani, V. *et al.* Recruitment of PLANT U-BOX13 and the PI4Kβ1/β2 phosphatidylinositol-4 kinases by the small GTPase RabA4B plays important roles during salicylic acid-mediated plant defense signaling in arabidopsis. *Plant Cell* **27**, 243–261 (2015).

361. Yeung, T. *et al.* Membrane Phosphatidylserine Regulates Surface Charge and Protein Localization. *Science* **319**, 210–213 (2008).

362. Pruneda, J. N. *et al.* Structure of an E3:E2~Ub Complex Reveals an Allosteric Mechanism Shared among RING/U-box Ligases. *Mol. Cell* **47**, 933–942 (2012).

363. Benirschke, R. C. *et al.* Molecular basis for the association of human E4B U box ubiquitin ligase with E2-conjugating enzymes UbcH5c and Ubc4. *Structure* **18**, 955–965 (2010).

364. Zhang, M. *et al.* Chaperoned Ubiquitylation—Crystal Structures of the CHIP U Box E3 Ubiquitin Ligase and a CHIP-Ubc13-Uev1a Complex. *Mol. Cell* **20**, 525–538 (2005).

365. Simon, M. L. A. *et al.* A multi-colour/multi-affinity marker set to visualize phosphoinositide dynamics in Arabidopsis. *Plant J.* **77**, 322–337 (2014).

366. Rutter, B. D. & Innes, R. W. Extracellular vesicles as key mediators of plant–microbe interactions. *Curr. Opin. Plant Biol.* **44**, 16–22 (2018).

367. Smith, J. M., Salamango, D. J., Leslie, M. E., Collins, C. A. & Heese, A. Sensitivity to Flg22 Is Modulated by Ligand-Induced Degradation and de Novo Synthesis of the Endogenous Flagellin-Receptor FLAGELLIN-SENSING2. *Plant Physiol.* **164**, 440–454 (2014).

368. Irani, N. G. & Russinova, E. Receptor endocytosis and signaling in plants. *Curr. Opin. Plant Biol.* **12**, 653–659 (2009).

369. Strauß, T. Characterization of a SAUL1-interacting kinesin-5. (Universität Hamburg, 2016).

370. Christensen, D. E., Brzovic, P. S. & Klevit, R. E. E2-BRCA1 RING interactions dictate synthesis of mono- or specific polyubiquitin chain linkages. *Nat. Struct. Mol. Biol.* **14**, 941–948 (2007).

371. Kraft, E. *et al.* Genome Analysis and Functional Characterization of the E2 and RING-Type E3 Ligase Ubiquitination Enzymes of Arabidopsis. *Plant Physiol.* **139**, 1597–1611 (2016).

372. Dowil, R. T., Lu, X., Saracco, S. A., Vierstra, R. D. & Downes, B. P. Arabidopsis Membrane-anchored Ubiquitin-fold (MUB) Proteins Localize a Specific Subset of Ubiquitin-conjugating (E2) Enzymes to the Plasma Membrane. *J. Biol. Chem.* **286**, 14913–14921 (2011).

373. Downes, B. P., Saracco, S. A., Sang, S. L., Crowell, D. N. & Vierstra, R. D. MUBs, a Family of Ubiquitin-fold Proteins That Are Plasma Membrane-anchored by Prenylation. *J. Biol. Chem.* **281**, 27145–27157 (2006).

374. Stewart, M. D., Ritterhoff, T., Klevit, R. E. & Brzovic, P. S. E2 enzymes: more than just middle men. *Cell Res.* **26**, 423–440 (2016).

375. Lu, X. *et al.* A MUB E2 structure reveals E1 selectivity between cognate ubiquitin E2s in eukaryotes. *Nat. Commun.* **7**, 1–11 (2016).

376. Vina-Vilaseca, A. & Sorkin, A. Lysine 63-linked Polyubiquitination of the Dopamine Transporter Requires WW3 and WW4 Domains of Nedd4-2 and UBE2D Ubiquitin-conjugating Enzymes. *J. Biol. Chem.* **285**, 7645–7656 (2010).

377. Lee, Y.-R. J. & Liu, B. Cytoskeletal Motors in Arabidopsis. Sixty-One Kinesins and Seventeen Myosins. *Plant Physiol.* **136**, 3877–83 (2004).

378. Reddy, A. S. N. & Day, I. S. Kinesins in the Arabidopsis genome: A comparative analysis among eukaryotes. *BMC Genom.* **2**, (2001).

379. Enos, A. P. & Morris, N. R. Mutation of a Gene That Encodes a Kinesin-like Protein Blocks Nuclear Division in A. nidulans. *Cell* **60**, 1019–1027 (1990).

380. Ferenz, N. P., Gable, A. & Wadsworth, P. Mitotic functions of kinesin-5. *Semin. Cell Dev. Biol.* **21**, 255–259 (2010).

381. Bannigan, A. *et al.* A conserved role for kinesin-5 in plant mitosis. *J. Cell Sci.* **120**, 2819–2827 (2007).

382. Wakana, Y. *et al.* Kinesin-5/Eg5 is important for transport of CARTS from the trans-Golgi network to the cell surface. *J. Cell Biol.* **202**, 241–250 (2013).

383. Ari, C., Borysov, S. I., Wu, J., Padmanabhan, J. & Potter, H. Alzheimer amyloid beta inhibition of Eg5/kinesin 5 reduces neurotrophin and/or transmitter receptor function. *Neurobiol. Aging* **35**, 1839–1849 (2014).

384. Elias, M. *et al.* The exocyst complex in plants. *Cell Biol. Int.* **27**, 199–201 (2003).

385. Saeed, B., Brillada, C. & Trujillo, M. Dissecting the plant exocyst. (2019).

386. Liu, N. *et al.* CALCIUM-DEPENDENT PROTEIN KINASE5 Associates with the Truncated NLR Protein TIR-NBS2 to Contribute to exo70B1-Mediated Immunity. *Plant Cell* **29**, 746–759 (2017).

387. He, B., Xi, F., Zhang, X., Zhang, J. & Guo, W. Exo70 interacts with phospholipids and mediates the targeting of the exocyst to the plasma membrane. *EMBO J.* **26**, 4053–4065 (2007).

388. Bloch, D. *et al.* Exocyst SEC3 and Phosphoinositides Define Sites of Exocytosis in Pollen Tube Initiation and Growth. *Plant Physiol.* **172**, 980–1002 (2016).

389. Basu, A., Li, X. & Leong, S. S. J. Refolding of proteins from inclusion bodies: rational design and recipes. *Appl. Microbiol. Biotechnol.* **92**, 241–251 (2011).

390. Korten, T., Chaudhuri, S., Tavkin, E., Braun, M. & Diez, S. Kinesin-1 Expressed in Insect Cells

Improves Microtubule in Vitro Gliding Performance, Long-Term Stability and Guiding Efficiency in Nanostructures. *IEEE Trans. Nanobioscience* **15**, 62–69 (2016).

391. Lu, S. X. & Hrabak, E. M. The myristoylated amino-terminus of an Arabidopsis calcium-dependent protein kinase mediates plasma membrane localization. *Plant Mol. Biol.* **82**, 267–278 (2013).

392. Horstman, A., Tonaco, I. A. N., Boutilier, K. & Immink, R. G. H. A Cautionary Note on the Use of Split-YFP/BiFC in Plant Protein-Protein Interaction Studies. *Int. J. Mol. Sci.* **15**, 9628–9643 (2014).

393. Dangl, J. L. & Jones, J. D. G. Plant pathogens and integrated defence responses to infection. *Nature* **411**, 826–833 (2001).

394. Maekawa, T., Kufer, T. A. & Schulze-Lefert, P. NLR functions in plant and animal immune systems: so far and yet so close. *Nat. Immunol.* **12**, 818–826 (2011).

395. Meyers, B. C., Morgante, M. & Michelmore, R. W. TIR-X and TIR-NBS proteins: Two new families related to disease resistance TIR-NBS-LRR proteins encoded in Arabidopsis and other plant genomes. *Plant J.* **32**, 77–92 (2002).

396. Hyun, K. gi, Lee, Y., Yoon, J., Yi, H. & Song, J. J. Crystal structure of Arabidopsis thaliana SNC1 TIR domain. *Biochem. Biophys. Res. Commun.* **481**, 146–152 (2016).

397. Burdett, H., Kobe, B. & Anderson, P. A. Animal NLRs continue to inform plant NLR structure and function. *Arch. Biochem. Biophys.* **670**, 58–68 (2019).

398. Yang, Y., Wu, X., Xuan, H. & Gao, Z. Functional analysis of plant NB-LRR gene L3 by using E. coli. *Biochem. Biophys. Res. Commun.* **478**, 1569–1574 (2016).

399. Niedziela-Majka, A., Rymarczyk, G., Kochman, M. & Ozyhar, A. Pure, Bacterially Expressed DNA-Binding Domains of the Functional Ecdysteroid Receptor Capable of Interacting Synergistically with the hsp 27 20-Hydroxyecdysone Response Element. *Protein Expr. Purif.* **14**, 208–220 (1998).

400. Waugh, D. S. Making the most of affinity tags. *Trends Biotechnol.* **23**, 316–320 (2005).

401. Li, Y., Pennington, B. O. & Hua, J. Multiple R-Like Genes Are Negatively Regulated by BON1 and BON3 in Arabidopsis. *Mol. Plant Microbe Interact.* **22**, 840–848 (2009).

402. Yang, J. T., Wu, C.-S. C. & Martinez, H. M. Calculation of Protein Conformation from Circular Dichroism. *Methods Enzymol.* **130**, 208–269 (1986).

403. Whittaker, C. A. & Hynes, R. O. Distribution and Evolution of von Willebrand/Integrin A Domains: Widely Dispersed Domains with Roles in Cell Adhesion and Elsewhere. *Mol. Biol. Cell* **13**, 3396–3387 (2002).

404. Sugita, S. & Südhof, T. C. Specificity of Ca2+-Dependent Protein Interactions Mediated by the C2A Domains of Synaptotagmins. *Biochemistry* **39**, 2940–2949 (2000).

405. Perestenko, P. V. *et al.* Copines-1, -2, -3, -6 and -7 show different calcium-dependent intracellular membrane translocation and targeting. *FEBS J.* **277**, 5174–5189 (2010).

406. Zhang, X., Rizo, J. & Südhof, T. C. Mechanism of Phospholipid Binding by the C2A-Domain of Synaptotagmin I. *Biochemistry* **37**, 12395–12402 (1998).

407. Shin, O. H. *et al.* Munc13 C2B domain is an activity-dependent Ca 2+ regulator of synaptic exocytosis. *Nat. Struct. Mol. Biol.* **17**, 280–288 (2010).

408. Perestenko, P., Watanabe, M., Beusnard-Bee, T., Guna, P. & McIlhinney, J. The second C2-domain of copine-2, copine-6 and copine-7 is responsible for their calcium-dependent membrane association. *FEBS J.* **282**, 3722–3736 (2015).

409. Gerber, S. H., Rizo, J. & Südhof, T. C. The Top Loops of the C2 Domains from Synaptotagmin and Phospholipase A2 Control Functional Specificity. *J. Biol. Chem.* **276**, 32288–32292

(2001).

410. Creutz, C. E. *et al.* The Copines, a Novel Class of C2 Domain-containing, Calcium- dependent, Phospholipid-binding Proteins Conserved from Paramecium to Humans. *J. Biol. Chem.* **273**, 1393–1402 (1998).

411. Heck, A. J. R. & van den Heuvel, R. H. H. INVESTIGATION OF INTACT PROTEIN COMPLEXES BY MASS SPECTROMETRY. *Mass Spectrom. Rev.* **23**, 368–389 (2004).

412. Creutz, C. E. & Edwardson, J. M. Organization and synergistic binding of copine I and annexin A1 on supported lipid bilayers observed by atomic force microscopy. *Biochim. Biophys. Acta - Biomembr.* **1788**, 1950–1961 (2009).

413. Rizo, J. & Südhof, T. C. C2-domains, Structure and Function of a Universal Ca2+-binding Domain. *J. Biol. Chem.* **273**, 15879–15882 (1998).

414. Sauer-Eriksson, A. E., Kleywegt, G. J., Uhlén, M. & Jones, T. A. Crystal structure of the C2 fragment of streptococcal protein G in complex with the Fc domain of human IgG Structure of the Fc domain of IgG DNA constructions and protein preparation. *Structure* **3**, 265–278 (1995).

7 SUPPLEMENT

Table S1
Amino acid sequences of recombinant proteins investigated during this thesis

Protein	Amino acid sequence
SAUL1	GPLGSPEFMVGSSDGDQSDDSSHFERGVDHIYEAFICPLTKEVMHDPVTLENGRTFE REAIEKWFKECRDSGRPPSCPLTSQELTSTDVSASIALRNTIEEWRSRNDAAKLDIA RQSLFLGNAETDILQALMHVRQICRTIRSNRHGVRNSQLIHMIIDMLKSTSHRVRYK ALQTLQVVVEGDDESKAIVAEGDTVRTLVKFLSHEPSKGREAAVSLLFELSKSEALC EKIGSIHGALILLVGLTSSNSENVSIVEKADRTLENMERSEEIVRQMASYGRLQPLLG KLLEGSPETKLSMASFLGELPLNNDVKVLVAQTVGSSLVDLMRSGDMPQREAALKA LNKISSFEGSAKVLISKGILPPLIKDLFYVGPNNLPIRLKEVSATILANIVNIGYDFDKA TLVSENRVENLLHLISNTGPAIQCKLLEVLVGLTSCPKTVPKVVYAIKTSGAIISLVQFI EVRENDDLRLASIKLLHNLSPFMSEELAKALCGTAGQLGSLVAIISEKTPITEEQAAA AGLLAELPDRDLGLTQEMLEVGAFEKIISKVFGIRQGDIKGMRFVNPFLEGLVRILAR ITFVFNKEARAINFCREHDVASLFLHLLQSNGQDNIQMVSAMALENLSLESIKLTRM PDPPPVNYCGSIFSCVRKPHVVNGLCKIHQGICSLRETFCLVEGGAVEKLVALLDHEN VKVVEAALAALSSLLEDGLDVEKGVKILDEADGIRHILNVLRENRTERLTRRAVWM VERILRIEDIAREVAEEQSLSAALVDAFQNADFRTRQIAENALKHIDKIPNFSSIFPNI A
CHS1	GPLGSPEFMSTSYSFLLAGRELDVFLSFSGKIALDVDFGYDLSRNGIKAFKSESWKES SFKPIDLRTLEALTESKVAVVMTSDEEVSSVGFLEELIVIIEFQEKRSLTVIPVFLTKH PLDVEKVSQIFPERAKIWRTAIAKLDNIAAQYSFSRNLAVMHGTHRIKQIADDIRLM FLSSASSDFKGLAGMDRHMKALYALLALESDEKVRTIGIWGSSGVGKTTLARYTYAE ISVKFQAHVFLENVENMKEMLLPSENFEGEDLRSVNHEMNEMAEAKQKHRKVLLI ADGVNNIEQGKWIAENANWFAPGSRVILITQEKSLLVQSGVNHVYEVGSLRYDEAL QLFSRFAFKQPYPSPDFERLSVRAVQLAGFLPVTIRLFGSFLTGRDKEEWEATLLKL NAKQGKDIKEVWKIMEALEDKDIVEASQR
BON1	GPLGSPEFMGNCCSDVASGAGATAGVGGSGSSAALGATNDALDYYLKSKGFNGLFS QIELSFSASNLRDRDVLSKSDPMVVVYQKEKDATLSEVFRSEVVLNSLAPKWIKKFI VAYHFETVQTLVFRVYDVDTKFQNSREEMLKLDEQQFLGEATCALSEIITKSTRTST LELKRKDGFAPQAQPHHGKLIIHAEESLASKISTEIVFRCSNLESKDLFSKSDPFLVVS KIVEHGTPIPVSKTEVRKNDLNPIWKPVFLSVQQVGSKDSPVIIECSDFNSNGKHSLI GKVQKSLSDLEKLHLAGQGINFSLPTGAGQNKVLKSQLFVDKFTETVHHTFLEYLAS GFELNFMVAIDFTASNGNPRLPDSLHYIDPSGRLNAYQRAIMDVGEVLQFYDSDKRF PAWGFGARPIDAPVSHCFNLNGSSSYSEVDGIQGIMTSYTSALFNVSLAGPTLFGPVI NAAAMIASASLAQGSRKYYVLLIITDGVITDLQETKDALVSASDLPLSILIVGVGGADF KEMEILDADKGERLESSSGRLASRDIVQFVALRDVQYGEISVVQALLAELPSQFLTY MRIRNMKPIPP

Table S2
(A) Protein domains of SAUL1, which used for *EOM* modelling. Start and end position are in relation to the length of the SAUL1 amino acid sequence used in the SAXS experiments.

Domain	Start	End
1	28	62
2	73	134
3	150	219
4	234	390
5	398	435
6	442	706
7	714	792

(B) Protein domains of SAUL1, which were detected using *InterPro*.

Database	Name	ID	Start	End	E-value
SMART	Ubox	SM00504	31	99	10^{-26}
SMART	ARM	SM00185	138	178	68
SMART	ARM	SM00185	180	220	17
SMART	ARM	SM00185	306	345	300
SMART	ARM	SM00185	346	391	200
SMART	ARM	SM00185	438	480	58
SMART	ARM	SM00185	582	624	100
SMART	ARM	SM00185	665	706	0.3
SMART	ARM	SM00185	710	752	120

Table S3
(A) Protein domains of BON1, which used for *CORAL* modelling. Start and end position are in relation to the length of the BON1 amino acid sequence used in the SAXS experiments.

Domain	Start	End
1	8	16
2	38	68
3	74	175
4	187	231
5	238	320
6	321	466
7	475	576

(B) Protein domains of BON1, which were detected using *InterPro*.

Database	Name	ID	Start	End	E-value
SMART	C2	SM00239	57	170	2.2×10^{-5}
SMART	C2	SM00239	207	306	5.1×10^{-9}
SMART	VWA	SM00327	347	544	3.1×10^{-10}

Table 4
T-DNA insertion sites in different mutants in respect to the transcription start site [156,157]

Mutant	Site
trx5	+976 bp
at5g52760	+436 bp
at3g28580	-273 bp
At4g16260	+1299 bp
wrky46	+733 bp

Table S5
Early differential expressed genes and their homologs in *A. thaliana* with sequence identity of at least 50 % according to a BLASTp search.

Gene	Homolog	Seq. identity (%)
TRX5	*TRX1*	63
	TRX2	54
	TRX3	74
	TRX4	60
AT5G52760	*AT5G52750*	72
	AT5G52770	62
AT3G28580	*AATP1*	77
	AT3G28510	60
	AT3G28520	53
	AT3G28540	58
	AT4G05340	51
AT4G16260	*BLG2*	50
	AT3G57240	55
	AT3G57270	54
WRKY46	*WRKY41*	57
	WRKY54	53
	WRKY67	50

Table S6
(A) Sample details of SAUL1 in SEC-SAXS experiments

Loading concentration (mg ml^{-1})	4.38
Buffer condition	50 mM Tris-HCl pH 9.0 250 mM NaCl
Selected sample frames	1873-1897
Selected buffer frames	1426-1476
Sample concentration of integrated curve (mg ml^{-1})	0.126

Table S6
(B) Structural parameters auf SAUL1 in SEC-SAXS experiments

	SEC
Guinier analysis	
$I(0)$	791.87 ± 14.53
R_g (nm)	4.64 ± 0.54
s_{min} (nm^{-1})	0.005
sR_g max	1.3
M from $I(0)$ (ratio to predicted)	-
P(r) analysis	
$I(0)$	815.5
R_g (nm)	5.14
d_{max} (nm)	18.4
s range (nm^{-1})	0.097-1.731
Quality estimate	0.77
M (Da) from $I(0)$ (ratio to predicted)	-
Porod Volume (nm^3)	116.82

Table S7
(A) Sample details of SAUL1 used in SAXS batch experiments

Molecular weight (Da)	88791.41
Extinction coefficient (M^{-1} cm^{-1})	26930
Concentration series (mg ml^{-1})	0.43; 0.98; 1.8; 3.3; 3.96; 4.96
Buffer condition	50 mM Tris-HCl pH 9.0 250 mM NaCl
Selected concentration (mg ml^{-1})	3.96

(B) Structural parameters of SAUL1 in SAXS batch experiments

Concentration (mg ml^{-1})	0.43	0.98	1.80
Guinier analysis			
$I(0)$	23802.6 ± 256.22	25804.3 ± 199.47	29578.4 ± 130.88
R_g (nm)	6.27 ± 0.43	7.64 ± 0.95	8.01 ± 0.62
s_{min} (nm^{-1})	0.004	0.041	0.003
sR_g max	1.3	1.27	1.22
M from $I(0)$ (ratio to predicted)	171236 (1.92)	185636 (2.09)	212787 (2.39)
P(r) analysis			
$I(0)$	24550	27000	30280
R_g (nm)	6.9	8.84	8.74
d_{max} (nm)	28.1	38.3	36.3
s range (nm^{-1})	0.07-1.27	0.06-1.04	0.05-0.99
Quality estimate (%)	0.68	0.68	0.73
M (Da) from $I(0)$ (ratio to predicted)	176613 (1.99)	194238 (2.18)	217834 (2.45)
Porod Volume (nm^3)	442.36	503.17	536.7

Table S7

(C) Structural parameters of SAUL1 in SAXS batch experiments		
Concentration (mg ml^{-1})	3.96	4.96
Guinier analysis		
$I(0)$	24074.6 ± 103.8	35713.7 ± 271.01
R_g (nm)	8.94 ± 0.31	10.64 ± 1.55
s_{min} (nm^{-1})	0.002	0.004
sR_g max	1.28	1.27
M from $I(0)$ (ratio to predicted)	173029 (1.95)	256924 (2.89)
P(r) analysis		
$I(0)$	29810	36800
R_g (nm)	10.03	11.78
d_{max} (nm)	45.6	45.7
s range (nm^{-1})	0.045-0.894	0.06-0.75
Quality estimate (%)	0.69	0.65
M (Da) from $I(0)$ (ratio to predicted)	214251 (2.41)	264739 (2.98)
Porod Volume (nm^3)	726.64	1160

Table S8
Protein domains of SAUL1, which used for *SASREF* modelling.

Domain	Start	End
1	1	27
2	28	67
3	68	138
4	139	245
5	246	371
6	372	393
7	394	438
8	439	734
9	735	796
10	797	806

Table S9
OLIGOMER-derived fractions of SAUL1 multimers

Oligomer	0.43 mg ml^{-1}	0.98 mg ml^{-1}	1.80 mg ml^{-1}	3.96 mg ml^{-1}	4.96 mg ml^{-1}
P1 monomer		0.34 ± 0.02	0.25 ± 0.02	0.21 ± 0.01	
P2 dimer	0.81 ± 0.02	0.09 ± 0.03	0.17 ± 0.02		
P222 tetramer (0.98 mg ml^{-1})	0.19 ± 0.02	0.57 ± 0.01			
P222 tetramer (3.96 mg ml^{-1})			0.50 ± 0.01	0.69 ± 0.01	0.91 ± 0.01
P32 hexamer			0.08 ± 0.01	0.10 ± 0.01	
P42 octamer					0.09 ± 0.01

Table S10
Homologs of E2s from *A. thaliana* in *Homo* sapiens with sequence identity and sequence coverage according to a BLASTp search.

Gene	Homolog	Seq. identity (%)	Seq. coverage (%)
UBC8	*UBE2D4*	79	99
UBC9	*UBE2D4*	80	82
UBC10	*UBE2D4*	99	79
UBC13	*UBE2G1*	63	96
UBC17	*UBE2W*	46	78
UBC28	*UBE2D4*	61	99
UBC29	*UBE2D4*	76	99
UBC30	*UBE2D4*	81	99
UBC33	*UBE2J2*	52	72
UBC37	*UBE2T*	53	38

Table S11
(A) Sample details of BON1 in batch experiments

Molecular weight (Da)	63904.74
Extinction coefficient (M^{-1} cm^{-1})	38850
Concentration series (mg ml^{-1})	1.32; 1.98; 3.4; 4.92; 6.2; 7.5
Buffer condition	100 mM Tris-HCl pH 7.5 200 mM NaCl 5 mM EDTA
Selected concentration (mg ml^{-1})	7.5

(B) Sample details of BON1 in SEC-SAXS experiments

Loading concentration (mg ml^{-1})	4.93
Buffer condition	50 mM Tris-HCl pH 9.0 250 mM NaCl
Selected sample frames	2051-2075
Selected buffer frames	501-551
Sample concentration of integrated curve (mg ml^{-1})	0.485

(C) Structural parameters auf BON1 in batch and SEC-SAXS experiments

	Batch	SEC
Guinier analysis		
$I(0)$	11693.00 ± 23.38	1858.84 ± 4.84
R_g (nm)	3.56 ± 0.37	2.85 ± 0.21
s_{min} (nm^{-1})	0.1611	0.0118
sR_g max	1.28	1.29
M from $I(0)$ (ratio to predicted)	84121 (1.33)	-
P(r) analysis		
$I(0)$	9149	1889
R_g (nm)	3.75	2.92
d_{max} (nm)	14.7	8.6
s range (nm^{-1})	0.16-2.23	0.13-2.80
Quality estimate (%)	0.73	0.96
M (Da) from $I(0)$ (ratio to predicted)	65822 (1.03)	-

Table S11
(C) continued

	Batch	SEC
P(*r*) analysis		
Porod Volume (nm^3)	126.45/129.08	94.26

8 EIDESSTATTLICHE VERSICHERUNG

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Hamburg, den

Jan Knop

9 ACKNOWLEDGEMENTS

Das war es. Die Arbeit ist zu Papier gebracht. Deswegen möchte ich mich bei vielen Menschen bedanken, ohne die diese Arbeit nicht möglich gewesen wäre.

Zuerst gilt mein Dank Stefan Hoth, in dessen Arbeitsgruppe diese Arbeit entstanden ist. Stefan, ich möchte Dir ganz herzlich für unsere Zusammenarbeit, die vielen Gespräche, während der wir auch mal unterschiedlicher Meinung waren, die immerwährende Unterstützung und natürlich auch für Deine geduldigen Korrekturen dieser Arbeit danken.

Dann gilt mein Dank Julia Kehr. Bei Dir Julia möchte ich mich für eine tolle Zeit während der DELIGRAH und Deine stets offene Bürotür bedanken. Und natürlich dafür, dass Du ohne zu zögern die Rolle der zweiten Gutachterin dieser Arbeit übernommen hast.

Ich möchte Andrew Torda ganz herzlich dafür danken, dass er mich in den letzten Jahren immer wieder unterstützt hat. Andrew, ich bin Dir sehr dankbar, dass ich bei Dir meine Bachelorarbeit machen durfte und bei Dir jederzeit mit einer Bitte für ein Referenzschreiben oder ein Gutachten vorbeikommen konnte.

Was die praktische Arbeit angeht, so wäre vieles ohne Kooperationen gar nicht möglich gewesen. Deswegen möchte ich mich besonders bei Sven Falke, Julia Lockhauserbäumer, Anna und Steffen Ostendorp, Melissa Graewert und Al Kikhney bedanken. Danke, dass Ihr mich bei Strukturmessungen, nativen MS-Untersuchungen, MALDI-TOF MS, Thermofluor Analysen sowie der Auswertung der SAXS-Daten unterstützt habt.

In den fast fünf Jahren Arbeit bleibt es nicht aus, dass ich etliche Studenten betreut habe, die somit auch an dieser Arbeit mitgewirkt haben. Hierbei haben mich besonders Mareike Schmidt, Simon Meyer, Sally Marusoi, Senem Tuncer und Chi Yan Lam unterstützt. Besonders hervorheben möchte ich die Arbeit von Marcel Bhattarai. Du hast immer unermüdlich gearbeitet und zusammen haben wir es geschafft, BON1 voranzutreiben. Das größte Dankeschön gilt an dieser Stelle aber Jenny Deke. Danke, dass Du mich nicht nur während meiner ersten Schritte im Bereich der Proteinanalytik unterstützt hast, sondern seitdem eine ganz tolle Freundin geworden bist.

Ohne die vielen netten Kollegen in der Molekularen Pflanzenphysiologie wäre diese Arbeit sicher nicht so gut gelaufen. Ich danke Euch allen sehr für die letzten fünf Jahre. Besonders Tim Lienemann, Teresa Wulf und Judith Mehrmann, möchte ich für Ihre Unterstützung danken. Ich danke auch Lisa Amelung, Monique Liebers und Wim Walter für eine tolle Zeit, gemeinsames Haare raufen und viele tolle Gespräche, die mir noch lange im Gedächtnis bleiben werden. Es war sehr schön mit Euch zusammen zu arbeiten. Ganz besonders danke ich Catharina Krings für Salmiak im richtigen Augenblick, mutmachende Dialoge und den von ihr sorgfältig ausgesuchten Eppiständer. Du hast mir diese Zeit wirklich zu einer denkwürdigen gemacht.

Aber natürlich auch von außerhalb der Universität haben einige Menschen ihren Teil zu dieser Arbeit beigetragen. Felix Ballhausen, Dir danke ich für zahllose Gespräche abends und sich gegenseitiges Leidklagen. Ohne Dich bei Steam zu sehen, wäre mancher Abend des Schreibens sicher um einiges ärmer gewesen. Des Weiteren danke ich Florian Sieck, René Mosenthien und

Flemming Feß dafür, dass Ihr immer ein offenes Ohr hattet, für grandiose Rollenspielrunden und füw denkwüwdige Gespwäche, ohne die ich diese Zeit nicht so gut überstanden hätte.

Ich möchte mich sehr bei Wiesia und Stephan Hanhart bedanken, die mich so liebevoll in ihre Familie aufgenommen haben. Bei Euch hatte ich immer einen Ruhepol und konnte von der Arbeit abschalten. Und wenn das bedeutet hat, ein Beet umzugraben, hüfthoch im Teich zu stehen oder zusammen zu kochen.

Von Herzen möchte ich mich bei meinen Eltern Hildburg und Gerd Knop bedanken, die immer für mich da waren, an mich geglaubt und mich jederzeit unterstützt haben. Ich bin Euch wahnsinnig dankbar, dass ihr euren Großen so sehr unterstützt habt und darauf vertraut habt, dass ich meinen Weg finde. Sogar „ohne Kopf".

Mein größter Dank gilt Patrizia Hanhart. Ohne Dich wäre diese Arbeit wahrscheinlich nie fertig geworden und Du hast mir in jeder Hinsicht bei dieser geholfen. Ich danke Dir für deine ganzen tollen Einfälle, dein Verständnis für meine Arbeitswut und Arbeitszeiten, deine Gelassenheit, wenn ich nur mal kurz eine Meinung von dir haben wollte, und dafür, dass Du in den letzten Jahren so sehr für mich da warst.

Danke.

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

**Fakultät für
Mathematik, Informatik
und Naturwissenschaften**

PhD Committee

Prof. Dr. Andrew Torda
Zentrum für Bioinformatik
Bundesstr. 43
20146 Hamburg

Tel. +49 40 42838-7331
Fax +49 40 42838-7332
torda@zbh.uni-hamburg.de
www.zbh.uni-hamburg.de/torda

18.12.19

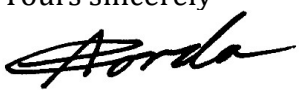
Re: Jan Knop, "Structur and function ... ligase"

Dear Committee Members,

This thesis is a delight to read. The English is far better the baseline for scientific texts. The grammar is correct and even the sentence construction is English. The word choice is English and not the usual literal translation.

I annoint this thesis with my english-mother-tongue blessing.

Yours sincerely



Andrew Torda