Structural Investigation of Bacterial Pathogenesis and Host Immune Response

by

Jie (Jennifer) Cheng

Department of Biochemistry
Duke University

Date: 05/16/2023

Approved:

_____
Dr. Pei Zhou, Supervisor

_____
Dr. Seok-Yong Lee

_____
Dr. Meta Kuehn

_____
Dr. Alberto Bartesaghi

_____
Dr. Mario-Juan Borgnia

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Chemistry in the Graduate School
of Duke University

2023

ABSTRACT

Structural Investigation of Bacterial Pathogenesis and Host Immune Response

by

Jie (Jennifer) Cheng

Department of Biochemistry
Duke University

Date:_____05/16/2023_____
Approved:

_____
Dr. Pei Zhou, Supervisor

_____
Dr. Seok-Yong Lee

_____
Dr. Meta Kuehn

_____
Dr. Alberto Bartesaghi

_____
Dr. Mario-Juan Borgnia

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Chemistry in the Graduate School of
Duke University

2023

Abstract

Bacterial infection is a persistent threat to public health and agriculture. In nature, pathogenic bacteria utilize various mechanisms to invade the host and cause a variety of diseases in human, animals, and plants. Ongoing research is focused on revealing how several essential proteins manipulate the host immunity to cause disease. It's necessary to understand the host-bacterial interactions as it is also a potential strategic hotspot for developing the novel therapeutic interventions against infectious diseases.

Many animal and plant pathogenic bacteria utilize the type III secretion system to deliver effector proteins into host cells, modulate host cellular functions, and cause infectious diseases. The AvrE/DspE-family effectors are one of the widely conserved Type III effector proteins, which play a central role in the pathogenesis of diverse phytopathogenic bacteria.

In the past three decades, the study of the function of type III secreted effector proteins has been hampered by the lack of structural information. By means of cryo-electron microscopy, we report the structure and function of DspE from *Erwinia amylovora,* a representative of the AvrE/DspE-family type III effectors. Our structural information complemented the AlphaFold2.0 (AF2) predictions, and our *in vitro* activity assays indicate that the *Ea*DspE protein could function as a pore anchoring into cell membrane, mediating the interplay between the bacteria and the host. This offers an

alternative mechanism of action for the AvrE/DspE-family effectors in the development of the plant disease. Furthermore, we propose the AvrE/DspE-family effectors might act as a novel class of pathogenic channels creating osmotic/water potential perturbation during bacterial infection.

Due to the lack of mobile defender cells and an adaptive immune system, plants have relied exclusively on the innate immunity of individual cells and on systemic signals emanating from the infection sites. NONEXPRESSOR OF PATHOGENESIS-RELATED GENES 1 (NPR1) is a master immune regulator in plants. It orchestrates systemic acquired resistance (SAR) by activating PATHOGENESIS-RELATED (PR) genes in response to induction of salicylic acid (SA) during the plant response to pathogenic challenges. In contrast, NPR3 and NPR4, paralogs of NPR1, have been reported to serve as negative regulators of the SAR in *Arabidopsis*. Despite the important role of NPR proteins, the understanding of their regulatory mechanism is hindered by a lack of structural information. Here, we performed the X-crystallographic structural analysis of *At*NPR1, *At*NPR3 and its complex with the transcription factor TGA3. The disclosed high-resolution structures surprisingly uncover an unexpected palmitic acid at the interface of dimeric TGA3 NID. This new discovery of a fatty-acid as a ligand of TGA3 opens a new venue of research in the role of fatty-acid in plant immunity.

In summary, we demonstrate successful structural characterization and functional studies of proteins essential in bacterial-host interactions and plant immunity.

These studies have significant implications in understanding the fundamentals in pathogen-host interactions and in developing therapeutics to treat diseases both in plants and human.

Dedication

I dedicate my dissertation work to my dear family and many friends. A special feeling of gratitude to my loving parents and dear sister, Sanhong Cheng, Wenying Huang, and Chen Huang whose words of encouragement and push for tenacity ring in my ears. Thanks for making me see this adventure through to the end.

Contents

List of Tables

List of Figures

List of Abbreviations

3D – Three-dimensional

AF2 – AlphaFold 2.0

ANK – Ankyrin repeat

CBD – Chaperone-binding domain

CF – Carboxy-fluorescein

Cryo-EM – Cryo-electron microscopy

Cul3 – Cullin 3

DTT – Dithiothreitol

E.coli – *Escherichia coli*

E. amylovora – *Erwinia. amylovora*

EMSA – Electrophoresis mobility shift assay

HR – Hypersensitive response

LC-ESI/MS – Liquid chromatography-electrospray ionization-mass spectrometry

MBP – Maltose-binding protein

NID – NPR1-interacting domain

NMR – Nuclear magnetic resonance

NLS – Nuclear localization signal

PAMAM – Poly(amidoamine) dendrimer

PAMP – Pathogen-associated molecular pattern

PC - phosphatidylcholine

PE - phophatidylethanolamine

PFT – Pore-forming toxin

PI - phosphatidylinositol

P. syringae – *Pseudomonas syringae*

PR – Pathogenesis-related

RLK – Receptor-like kinase

SA – Salicylic acid

SAR – Systemic acquired resistance

SBD – SA-binding domain

T3E – Type III effector protein

T3SC – Type III secretion chaperone

TOPP – Type one protein phosphatase

Acknowledgements

I would like to thank my advisor, Dr. Pei Zhou, for his steady and detailed guidance throughout my long journal at Duke graduate school, his patience with me and his accessibility. As a scientific and life mentor, he has honed me as a true scientist with critical thinking and taught me to enjoy the ups and downs in the life with a positive attitude.

I would like to thank my lab mates, current and former, for making me feel at home and for making the lab a comfortable and cheerful place to work in. This includes Javaria Najeeb, Ben Fenton, Jinshi Zhao, Qinglin Wu, Carly Sciandra, and Skyler Cochrane. Additionally, I'd like to thank my classmates and friends: Aili Hao, You Wang, Wei Mao, Simiao Ren, Xiaoou Zhan, Shufan Xia, and Yongze Jia.

Furthermore, I would like to thank my family for supporting me overcome all the difficult times and their unconditional love, and especially my uncle, Jingbo Xu, a professor at Zhejiang University, who led me into the world of scientific research and unconditionally supported me in the past years.

Lastly, I would thank my committee members: Meta Kuehn, Seok-Yong Lee, Alberto Bartesaghi, and Mario-Juan Borgnia, for their valuable feedback and supports over the years.

Chapter 1. Introduction

In nature, plants are subject to various attacks from pathogens and microbes. These plant pathogens cause diverse plant disease and lead to major crop losses worldwide. Numerous devastating pandemics have been caused by plant pathogens through history, including the Great Famine during the 1840s, caused by *Phytophthora infestans* infection of potato crops throughout Europe which led to the starvation and deaths of millions of people.

Among diverse plant pathogens, the pathogenic bacteria have caused many serious diseases in vegetable crops and have been a persistent threat to the public health. For example, the fire blight is a common disease caused by the Gram-negative bacterium *Erwinia amylovora*, mostly affecting pears and apples. Specifically, in the Washington state, there have been annual outbreaks since 1991.

Pathogenic bacteria utilize various specialized mechanisms to invade the host plants and cause diseases. The process is a complex interplay between host plants and pathogens. In brief, plant pathogens first contact and pass through the physical barriers in the host, followed by penetrating and colonizing internal plant tissue, and then modulate host physiology to causes infection (1, 2). In turn, to respond, the plant host triggers its innate immune system to ward off the invaders.

As the plant diseases caused by bacterial pathogens have been a major threat to agriculture, which indirectly affecting human health, in the past decades, researchers

have been made extensive efforts to develop chemicals against plant diseases. In early 1950s, the discovery of streptomycin has brought success in controlling plant diseases and since then the streptomycin has become commonly used in agriculture in the U.S. Later in 1960s; however, the emergence of streptomycin resistant (Sm^R) plant pathogens, in particular, the Sm^R strains of *E. amylovora*, has impeded the plant diseases control and led to widespread economic losses (3). Given the devastating consequences, scientists are seeking for novel alternatives to combat bacteria-mediated diseases.

Understanding the biology of host-pathogen interactions and the process of disease development in plants is crucial for seeking for alternatives to bacterial antibiotics. Interference with bacterial virulence and/or enhancement of the host immune response could be a compelling approach for developing novel anti-virulence treatment or engineering disease resistant crops. Similar to human bacterial pathogens, phytopathogenic bacteria evade and suppress host innate immunity during infection, and manipulate host cell physiology to achieve microbial invasion (4). Many phytopathogenic bacteria use a variety of virulence strategies to cause disease in the host, including secretion of cell wall degrading enzymes that disassemble the plant cell wall, production of low-molecular-weight phytotoxins to kill or change signal transduction, and employment of specialized secretion systems to deliver effector proteins (5-7).

Bacterial pathogens have evolved a wide variety of highly specialized secretion systems to deliver an arsenal of virulent factors into the host to cause diseases (4, 8). In Gram-negative bacteria, so far, five double-membrane-spanning secretion machineries have been identified and are classified as type I secretion system (T1SS), T2SS, T3SS, T4SS and T6SS(8). Three of them, T3SS, T4SS, and T6SS, are involved in the translocation of virulence factors including small molecules, proteins, and DNA (9). In general, depending on the secretion machineries, these secreted virulence factors have three possible fates: remaining associated with the bacterial outer membrane, being released into the extracellular space, or being injected into the target host cell (8).

The T3SS are found in various Gram-negative bacterial genera such as *Salmonella*, *Pseudomonas*, and enteropathogenic *Escherichia coli* (8, 10). Made up of more than 20 proteins, it is considered as one of the most complex secretion systems known in bacteria, though its components are generally conserved(10). Functionally, the syringe-like T3SS delivers secreted bacterial Type III effector (T3E) proteins to the host cytoplasm in plants and then cause diseases (7, 11-13). These T3E proteins collectively alter the signal transduction to the benefit of the pathogen, suppress antimicrobial host defense and elicit the release of nutrients from host to enhance bacterial multiplication (13, 14).

Although the T3SS has been well characterized as a transfer apparatus for virulent effector proteins, the specific roles of T3SS effectors in regulating and

controlling the host invasion and the onset of the plant diseases have remained incomplete understood. Elucidation of how these effector proteins function in host cells is critical for understanding the mechanisms of bacterial virulence, which potentially could present a promising anti-virulence strategy for treatment of bacterial infection. In the past 20 years, efforts have been made to identify and characterize those pathogenic virulence factors and to reveal their mode of actions at the molecular level (15, 16). One well-studied example is the type III effector protein AvrPto from *Pseudomonas syringae*. In 2008, Xiang et al., reported that AvrPto blocks the innate immunity by targeting several receptor kinases including Pto in the plant host, and additional structural and biochemical analyses further confirmed that AvrPto serves as an inhibitor of the Pto kinases (14). Another well-known Type III effector protein is the widely conserved AvrE/DspE-family effectors, which play central roles in the pathogenesis of diverse phytopathogenic bacteria. Although biochemical studies have hinted its functions in the past decades, little is known about the molecular mechanism of this family protein.

In response to bacterial attacks, the host plants – due to a lack of mobile defender cells and the adaptive immune system in vertebrates - exclusively rely on the innate immunity of individual cells and systemic signals emanating from infection sites (17, 18). The co-evolution of diverse pathogens and plant hosts has shaped the plant host to develop a highly adaptable, two-set plant innate immune system (19). One set uses the pattern recognition receptors (PRRs) on the cell surface to mediate the recognition of

microbial- or pathogen-associated molecular patterns (MAMPS or PAMPs) to elicit pattern-triggered immunity (PTI). The second acts within the cell and is known as 'effector-triggered immunity' (ETI), specifically targeting those pathogens which evade PTI and infect their hosts through secreted virulence effectors.

However, stimulation of defense responses not only occurs at the local site of pathogen recognition but also in distal regions of the plant. The secondary immune response in distal tissues is triggered once the plant perceives pathogens at the infection sites and the local defense is activated. The process is known as systemic acquired resistance (SAR) (17, 20). SAR is an effective innate immune response that protects plants against a broad range of biotrophic pathogens, and can be induced by several signaling molecules, including the salicylic acid (SA). SA is a phytohormone and is well known as a master immune regulator critical in both local defenses and SAR. Once bacterial infection is detected, plants begin synthesizing and accumulating SA.

In the plant model system *Arabidopsis thaliana*, SA is perceived by two classes of protein receptors of the same family: NONEXPRESSOR OF PATHOGENESIS-RELATED GENES 1 (NPR1) and NPR1-LIKE PROTEIN 3 (NPR3)/NPR4. Although they share high similarity in sequences, NPR1 and NPR3/4 are proposed to have distinct functions (20, 21).

NPR1 orchestrates SAR by activating *PATHOGENESIS-RELATED (PR)* genes in response to induction of salicylic acid (SA) upon pathogen challenge. The *NPR1* gene

was first discovered and identified in early 1990s (22). Since then, the role of NPR1 in mediating plant responses against pathogen invasion has been extensively studied (23). Over the past two decades, NPR1 was established as a positive regulator of systemic acquired resistance (SAR) and for controlling the onset of SAR. It acts as the master regulator of the plant defense signaling network, mediating cross-talks between the SA and Jasmonic acid/ethylene (JA/ET) responses (23). During SAR induction, NPR1 is primarily translocated to the nucleus. And in the nucleus, it interacts with members of the TGA family of transcription factors to indirectly activates *PR* gene expression (24, 25).

The discovery and functional characterization of NPR1 protein has been a decades-long effort. Although a lot has been uncovered in recent years, the molecular basis, and the mechanism by which SA activates NPR1 is incompletely understood. In 2003, Mou et al., reported that NPR1 can be found as a large cytoplasmic oligomer in the non-stress condition, but during SA-induced SAR a reduction of the cytoplasmic oligomeric form is observed (23, 26). The redox state of NPR may contribute to the NPR1 monomerization, nuclear localization and *PR* gene expression(23). This discovery raises further questions how these different oligomeric states of NPR1, regulated through cellular redox changes, mediate plant immune responses upon pathogens attack.

In contrast to NPR1, NPR3 and NPR4, two paralogs of NPR1, have been reported to act oppositely and serve as negative regulators of the SAR in *Arabidopsi*s (27). Initially,

NPR3 and NPR4 were reported to function as adaptors of the Cullin 3 ubiquitin E3 ligase to mediate NPR1 degradation in an SA-regulated manner (21). However, in 2018, Ding et al., reported that NPR3/NPR4 function as transcriptional co-repressors and plays opposite roles to NPR1 in transcriptional regulation of SA-induced defense gene expression (20). The specific functional roles that NPR3/NPR4 are playing in plant immunity remain to be resolved.

Since NPR1 and NPR3/NPR4 both are proposed to be SA receptors but act oppositely, despite their high similarity in sequences, how they perceive SA and coordinate hormonal signaling at the molecular level have been a long-standing question. It is critical for us to obtain direct structural information to reveal the mystery. Uncovering their atomic structural information helps us better understand their mode of actions in the plant host immune responses.

Decades ago, X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy are two prevailing structural biology tools to obtain high-resolution atomic structures of biomolecules and to study protein dynamics. Thanks to recent technological advances in software and hardware, cryo-electron microscopy (cryo-EM) has gradually become another mainstream tool for structural investigation and analysis.

In the past 55 years, a series of critical discoveries and advances in hardware, software, and computational power was made to make it possible to visualize non-crystalline biomolecules at high-resolution. This also leads to a rapidly growing number

of published high-resolution structures and biomolecular assemblies challenging for either X-ray crystallography or NMR (28). In contrast to X-ray crystallography, which needs a large amount of sample and requires stable folding of protein molecules at certain size, cryo-EM requires less samples, allows visualization of larger biomolecules, and additionally preserves the sample at closer-to-native state during sample preparation. Similar to NMR, the cryo-EM permit the flexibility and could capture the dynamic information of intermolecular interactions. Cryo-EM continues to smash through barriers.

In 2017, three scientists who made major contributions to this scientific breakthrough: Jacques Dubochet, Joachim Frank, and Richard Henderson, were awarded the Nobel Prize in chemistry. This reflects the spectacular results that have been achieved in recent years with single-particle cryo-EM analysis in the field of structural biology (28). Supplemented with X-ray crystallography and NMR, the advances of cryo-EM have led to an exponential increase in available 3D structures of biomolecular and complex assemblies. As of April 2023, there are more than 25,000 of cryo-EM solved structures deposited in the Electron Microscopy Data Bank (EMDB).

Cryo-EM promotes our mechanistic understanding of proteins and protein-protein interactions essential in bacterial pathogenesis and host immune response via providing structural information.

For instance, the anthrax toxin is a virulence factor secreted by *Bacillus anthracis* and is known by causing high mortality in humans and other mammals (29). The toxin consists of three components, a membrane translocase PA, a cytotoxic substrates LF and edema factor (EF). Due to its suitability as a bioweapon, there is an urgent search for an antidote. Since 1990s, anthrax toxin components have been widely investigated by means of X-ray crystallography and NMR (30, 31), and the structures of each component were well determined (32). However, due to its complexity involving three protein components and its large size, it has been challenging to study the internal dynamics and to capture the overall assembly. For years, many research groups are trying to obtain detailed mechanistic understanding of the stepwise process of pore formation and subsequent translocation, hoping to find a means to mitigate or prevent anthrax disease (29, 33). With the recent advances in single-particle cryo-EM analysis, several cryo-EM-derived structures of toxin assembly and anthrax toxin translocation complex are able to provide deeper insights into the FL unfolding and refolding mechanism (29, 34).

The second case tells the story of how the cryo-EM-led mechanistic study of the pore-formation process of aerolysin, a bacterial β-pore-forming toxin (β-PFT), leads to a biotechnology invention. Aerolysin from bacteria *Aeromonas hydrophila*, is a well-known member of the β-PFT family (35, 36). It is a beta structure toxin that is produced as a soluble monomer but is capable of oligomerizing at the membrane of target host cells

forming pores, which may lead to osmotic lysis and cell death. Owing to its pathogenic role and its unique ability to exist both as soluble proteins and transmembrane complexes, there is a long-standing interest to fully reveal the molecular mechanism and dynamics of how aerolysin monomer oligomerizes, achieves membrane insertion, and further causes cell death. For decades, X-ray crystallographic analysis only reveals the soluble state of aerolysin, but not the membrane-bound state. The lack of the direct structure information on the process of pore-formation and membrane anchoring has hindered our mechanistic understanding of its sensing capability as a pore. Not until 2016, lacovache et al., reported several cryo-EM structures of its conformational intermediates and the final aerolysin pore, jointly providing structural insights in the conformational changes that allow pore formation, revealing a full picture of the process of aerolysin pore-forming (35). Furthermore, given these mechanistic insights, Cao et al. was able to use a range of computational and biophysical technique to precisely map the sensing spots of this toxin for ssDNA translocation and established the fundamentals for developing the aerolysin as a nano-sensor for single-molecule DNA sequencing analysis (37-39).

In the era of technological progress continues to grow exponentially, besides the advance in cryo-EM, the breakthrough in computational modeling, in particular, the release of AF2, accelerates our biomedical research in basic science. Protein structure predication with high accuracy has been one of the most challenging problems in

structural biology for decades. Scientist has long wondered how a protein constitutes parts - a string of different amino acids – map out the many twists and folds of its eventual shape and how the nature makes the rules. In 2021, *Nature* reported that the machine-learning based AF2.0 developed by DeepMind can predict 3D structures of around 100,000 unique proteins directly from amino acid sequences with atomic-level accuracy, which triggers another revolution in biology (40-42). AF2 represents a milestone advancement in protein structure prediction. Early attempts to use computers to predict protein structures in the 1980s and 1990s performed poorly. The advent of AF2 presents us an unprecedented opportunity to accelerate research in the fields of drug discovery, protein design, et al. So far, only about 200,000 proteins' structures have been determined by means of cryo-EM, X-ray crystallography and NMR, covering less than 0.1% of the protein universe. It is a long-term effort for us to fully reveal the protein atlas in the world. AF2 and its predicted models will enable researchers to accelerate and solve problems that are previously thought to be highly challenging.

Although the time is not long since the AF2 were released, there are already quite a few examples of applying AF2 in the fields of biomedical research, demonstrating the potential of AF2 in accelerating scientific research (40). For instance, Hutin et al. recently revealed a structure of the vaccinia virus DNA helicase, the helicase-primase D5, via combining cryo-EM and AF2 prediction (43). Jin et al. solved the structure of interleukin (IL) – 27 signal complex by using cryo-EM with the aid of

AF2 prediction, through which they revealed a new mechanism for the assembly and activation of IL-27 receptor recognition complex (44). These results proposed an integrated strategy of supplementing AF2 with mainstream structural biology tools.

Besides structure determination, AF2 prediction can also be used for designing reliable expression constructs of challenging protein targets. With a relatively reliable model, AF2 can assist in selecting low resolution cryo-EM classes for further refinement.

With the revolutionary advances in single-particle cryo-EM and AF2 modeling, supplemented with X-ray crystallography, we aim to reveal the molecular basis of essential protein components involved in bacterial pathogenesis and plant immune responses. In the past three decades, the mechanistic studies of the function of T3E proteins have been hampered by the lack of structure information. Here we report our structural and functional analysis of DspE from *Erwinia Amylovora,* a representative T3E from the AvrE superfamily. Our structural information complemented with the AF2 predictions indicating that the C-terminus of *Ea*DspE effector folds into a β-barrel structure resembling bacterial porins, and our *in vitro* biochemical assays suggested that the *Ea*DspE protein might functions as a pore transporting ions or chemicals as a virulence strategy, mediating the interplay between the bacteria and the host. This offers an alternative mechanism of action for the AvrE/DspE-family effectors in plant disease development.

We also performed the X-crystallographic structural analysis of *At*NPR1, *At*NPR3 and a transcription factor *At*TGA3. These high-resolution structures surprisingly uncover unexpected features of NPR1, NPR3, and a palmitic acid embedded within each unit of the dimeric TGA3 NID. This new discovery of a fatty-acid as a ligand of TGA3 opens a new research venue of investigating the regulatory/signaling role of fatty-acid cofactor in plant immunity.

Taken together, my thesis research has revealed a novel activity of T3E protein, *Ea*DspE, and its initial preliminary structure information, disclosed the high-resolution structures of key immune regulators, *At*NPR1 and *At*NPR3, and discovered a novel ligand of *At*TGA3. In the first study of bacterial effector protein, we show that DspE from *Erwinia amylovora* might function as a pathogenic pore as a virulence mechanism. In the second study of NPR protein and its interacting partners TGA3, we revealed insightful structural information of protein-protein interactions and its potential ligand using X-ray crystallography methods.

Chapter 2. Structural and functional studies of DspE revealing a new virulence strategy

Many bacterial pathogens utilize dedicated secretion systems to promote bacterial virulence in host cells, specifically, secreting virulence factors from the bacterial cytosol into host cells. In Gram-negative bacteria, the secretion systems are divided into more than five classes due to their distinct overall architectures, functions, and specificity (8, 9). One representative secretion apparatus is the T3SS, present in a large number of Gram-negative bacterial pathogens (9).

The T3SS is essential in promoting bacterial virulence and is known as a membrane-embedded 'injectisome' due to its 'needle and syringe'-like structure (45). Structurally, the T3SS can be broken down into three main components: a base complex spanning bacterial membranes, the needle component, and the translocon (Figure 1) (46). The ~800-Å-long T3SS needle complex connects the intracellular compartments of infection bacteria and hosts and has an inner hollow core permitting the passage of virulent effector proteins from bacteria to host cells. The translocation of Type III effectors (T3Es) into host cells is essential for the virulence of many phytopathogenic bacteria (9). Once these T3E proteins are delivered into the host cytoplasm, they modulate host cellular physiology, thwart host immune responses and promote bacterial infection.

As the inner diameter of the T3SS needle filament is only about 20 Å, too small to accommodate most large folded bacterial effectors, earlier it has been proposed that T3E may travel through T3SS in a largely unfolded manner. Not until the recent cryo-EM study of the structure of the T3SS by Radics et al., direct evidence has been provided (45). T3Es protein must be unfolded during transport through T3SS 'injectisome' (45).

In addition to the necessity for T3Es to unfold themselves, another key element for efficient transfer of T3Es in T3SS is the requirement of specialized chaperones (Figure 1) (47, 48). These customized cytoplasmic chaperones are known to distinctively interact with the amino termini of cognate secreted virulence factors/T3Es, and form highly specific complex before secreting from the bacterial cytoplasm. Such interactions at the N-termini of T3Es in the T3E-chaperone complexes may allow the T3Es to retain their enzymatic or effector activity while bound to the chaperones (49). As many crystal structures of chaperones bound with N-termini of T3Es have been revealed, for instance, the *Yersinia* SycE-YopE chaperone-effector complex, it has been suggested that the chaperones might help keep extended portions of the T3Es in a secretion-competent state suitable for transferring through the T3SS needle complex in an unfolded manner (48, 50, 51). However, the exact mechanism of the T3E-chaperone interactions needs to be determined.

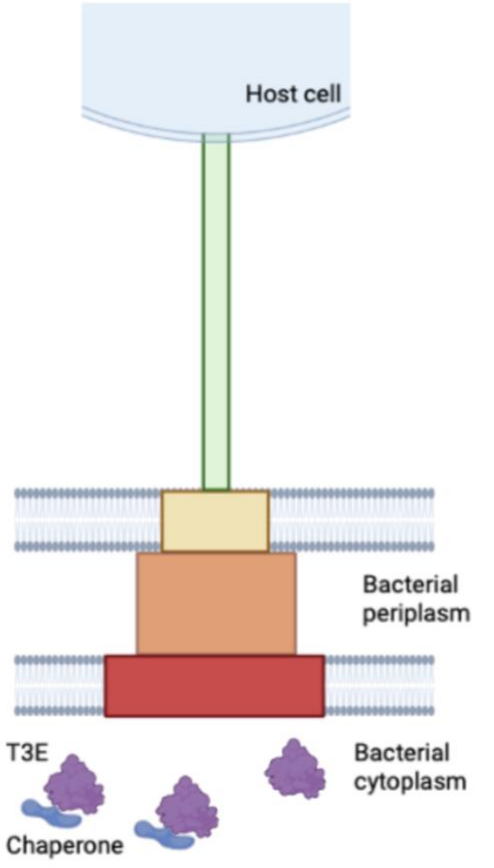


Figure 1 A schematic illustration of the type III secretion system from Gram-negative bacterial pathogen injecting virulence factors into host cells.

The green protrusion is the 'needle' filament. The virulence factors/effectors are bound to chaperones when they are waiting to be transported through the export machinery. Figure was made using BioRender.

Among numerous T3Es from Gram-negative bacteria, the AvrE superfamily effectors are broadly conservative and play crucial roles in bacterial pathogenesis in plants. As the *Pseudomonas syringe pv. tomato* (*Pst*) DC3000 is a well-characterized model pathogen and is widely used to study bacterial diseases in plants, the *Pst* AvrE effector

protein has been a most thoroughly characterized target to investigate the mode of action of T3Es belonging to the widely conserved AvrE/DspE superfamily (52).

Gene *avrE* was originally cloned from *Pseudomonas syringae pv. tomato* (*Pto*) PT23 in 1989 on the basis of its ability to confer avirulence to *P. syringae pv. glycinea* race 4 in soybean (53). Following Kobayashi et al., in 1995, Lorang and Keen characterized the *avrE* gene and found that *avrE* locates immediately adjacent to the *P. syringae pv. tomayo* gene cluster *hrp*, named for the "*h*ypersensitive *r*eaction" (HR) and "*p*athgenicity" (54). As *hrp* genes, present in all Gram-negative necrogenic plant pathogens, are known for controlling the ability of major groups of phytopathogenic bacteria to elicit the hypersensitive response (HR) in resistant plants and cause disease in susceptible plants (55), this discovery suggests that the *avrE* genes might be involved in suppressing plant immune response.

Another well-known phytopathogenic bacteria is the *Erwinia amylovora*, which causes fire blight in apple, pear, and other rosaceous plants (56). Similarly, prior genetic analyses identified a homologous and functionally similar *hrp*-linked pathogenicity locus in *E. amylovora,* known as "*d*isease-*sp*ecific" (*dsp*) region. This *dsp* region is located next to the *hrp* gene cluster in *E. amylovora* and is required for its pathogenicity but dispensable for HR elicitation (56). Interestingly, the *avrE* locus *in trans* could restore pathogenicity to *dspE* strains of *E. amylovora*, although restored stains were low in virulence (56).

Both *avrE* and *dspE* genes encode two representative T3E proteins of approximately 2000 amino acids, belonging to the AvrE superfamily. Members of the AvrE superfamily are known as vertically inherited core effectors and for contributing significantly to virulence by suppressing pathogen-associated molecular pattern (PAMP) - triggered immunity (57). They're also involved in inhibiting SA-mediated plant defenses, interfering with vesicular trafficking, and promoting bacterial growth *in planta*.

Despite their important roles in phytobacteria pathogenesis, the exact biochemical functions of AvrE/DspE-family effectors have been recalcitrant to mechanistic understanding for almost three decades. As elucidating how these T3Es function in the host cells is critical for us to understand the bacterial pathogenesis and disease developments in plants, we used a series of biochemical methods and structural biology tools to investigate their mode of actions in an attempt to reveal the virulence strategy of the AvrE/DspE-family T3Es.

2.1 AvrE/DspE-family protein sequence comparison

Members of the AvrE superfamily contribute significantly to virulence in host plants (57). Prior BlastP search revealed that AvrE shared 28% identity with the DspE protein required for *Erwinia amylovora* pathogenicity, and cross-complementation analysis confirmed conservation of function (58). Both encoding genes are located next to the T3SS cluster, and orthologues, also linked to the T3SS cluster, were later described in other plant-pathogenic bacteria: *Pantoea stewartia* ssp. *Stewarti*a (WtsE) (59), *Pantoea*

agglomerans pv. gysophilae (DspE/A) (60), and *Pectobacterium atrosepticum* and *Pectobacterium carotovorum* (DspE) (61).

Similar to AvrE from *Pst* DC3000, its homologs are essential for the pathogenicity of other four phytopathogenic bacteria: *E. amylovora*, *Pa. stewartii ssp. Stewartii*, *Pa. agglomerans pv. gypsophilae* and *Pe. Carotovorum*, respectively (58-61).

Functionally, these proteins may share certain similarities. In 2016, Xin et al reported that bacteria establish a humidity-dependent aqueous intercellular space (apoplast) during the infection of the phyllosphere, and several bacterial effectors are playing important roles in inducing such establishment of the aqueous apoplast (62). This symptom is known as 'water-soaking', the macroscopic hydration of the apoplast. This plant pathogen-induced water-soaking and necrosis allow bacteria to replicate and promote bacteria growth in plant host (62, 63).

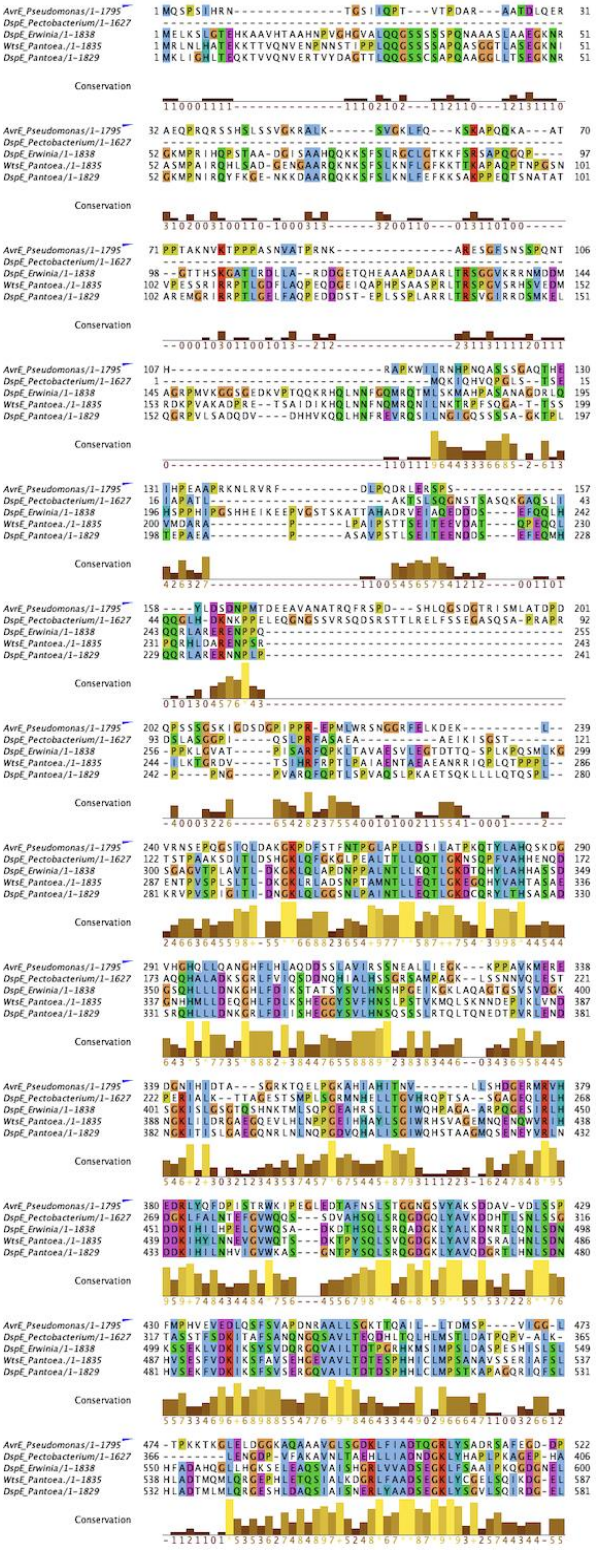
Following Xin et al., Gentzel et al. further reveals that the formation of a hydrated and nutritive apoplast in *Pantoea stewartii subsp. stewartii* (*Pnss*) is driven by WtsE, a T3E belonging to the AvrE/DspE-superfamily (64). WtsE perturbs host gene expression within phenylpropanoid pathway and leads to significant accumulation of the downstream metabolites (64). Bacteria *Pnss* relies on WtsE to survive and proliferate in its host.

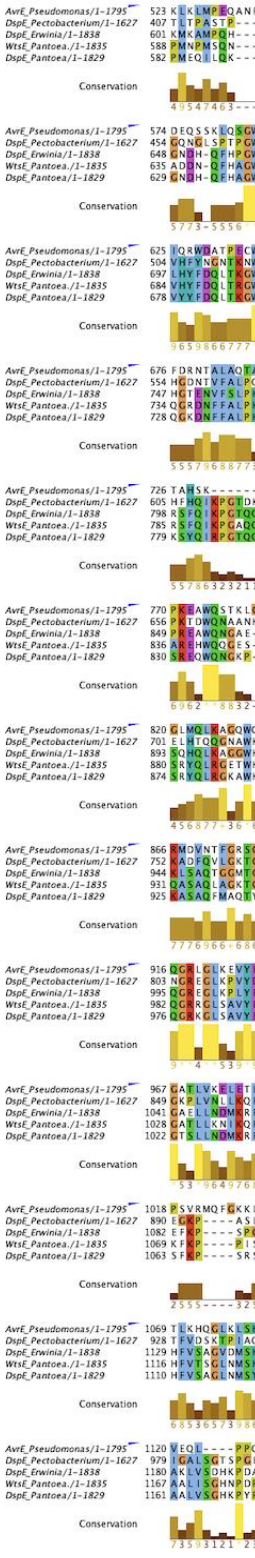
As above major virulence T3Es contribute to bacterial proliferation in hosts and induction of major disease symptoms including water-soaking (2, 57, 59, 61, 62, 65, 66).

First of all, we performed multiple sequence alignment (MSA) for comparison and analysis (Figure 1).

In the MSA, we found that starting from amino acid ~300, all five T3Es have higher consensus in sequences. The charged residues (Asp, Lys, Arg & Glu) and non-charged residues (Ala, Gly, Val & Leu) share high-level conservation at the C-terminus. Such a high conservation in sequences at the C-terminal half of AvrE/DspE-T3Es indicates the functional importance of this region.

Though the AvrE/DspE family effectors share high similarity in their sequences, currently no atomic structural information is available due to their extremely large size (approximately 200kD) and high toxicity to host plants. Thus, obtaining direct structural information of AvrE/DspE T3Es would be essential for us to elucidate how these effector proteins function in the host cell.


AvrE_Pseudomonas/1-1795
DspE_Pectobacterium/1-1627
DspE_Erwinia/1-1838
WtsE_Pantoea./1-1835
DspE_Pantoea/1-1829
Conservation


AvrE_Pseudomonas/1-1795
DspE_Pectobacterium/1-1627
DspE_Erwinia/1-1838
WtsE_Pantoea./1-1835
DspE_Pantoea/1-1829
Conservation

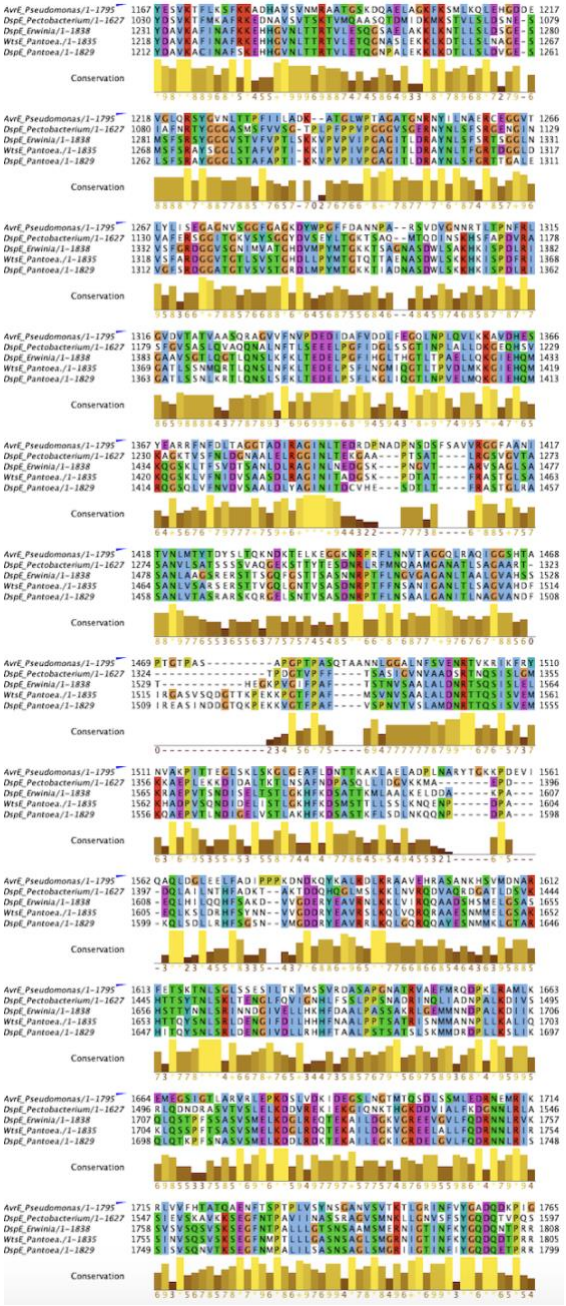


Figure 2: Multiple sequence alignment of amino acid sequences of five representative AvrE/DspE-family effectors.

The alignment includes sequences of *P. syringae pv. tomato* AvrE (*Pst* DC3000), *P. agglomerans* DspE , *E. amylovora* DspE,, *P. stewartii* WtsE and *P. carolovorum* DspE are aligned using Clustal Omega and Jalview (67-69).

2.2 AF2 prediction and cryo-EM imaging

Effectors of the AvrE superfamily are very large proteins containing approximately 2000 amino acids, without sequence similarity with proteins of known structures. Although their large size is consistent with a multifunctional activity, their biochemical function remains unknown (57).

Since no detailed structural information of any protein in the AvrE/DspE superfamily is currently available, it's challenging for us to perform homology modeling for structural prediction. Alternatively, we chose four representative protein sharing high sequence similarity from the AvrE/DspE-family and performed AF2 prediction (Figure 2&3). As the N-termini of these four AvrE/DspE-family proteins are intrinsically disordered, similar to other T3Es, here, we deleted the first N-terminal 300 amino acids in their AF2-predicted models for illustration and comparison (Figure 3).

The N-termini of T3Es from the AvrE/DspE family are known for interacting with type III secretion chaperones (T3SCs), required for the secretion or accumulation of T3Es inside the bacterial cytoplasm before injecting into the host (57). Two research groups, Triplett et al. and Oh et al., have mapped the region in *Ea*DspE responsible for such interactions and identified that the chaperone-binding domain (CBD) in *Ea*DspE is located within N-terminal 51-200 amino acids (70, 71). AF2-generated prediction of full-length DspE protein also indicate an intrinsically disordered polypeptide at its N-terminus, consistent with prior genetic mapping and mutational studies (Figure 4A).

The AF2-predicted models of DspE from *E. amylovora*, DspE from *P. carotovorum*, AvrE from *P. s.* pv *tomato* (*Pst*) DC3000 and WtsE from *P. stewartia* (Figure 2) all reveal an overall similar architecture resembling a mushroom, with a prominent central β-barrel forming the stem, which is surrounded by a globular N-terminal domain (*E. amylovora* DspE: K298-H672), a WD40 repeat domain (H673-P912), and two perpendicularly arranged helix bundles (E998-T1222 and P15569-S1648) on the top. The presence of WD40 repeat domain in the AF2-predicted model is consistent with prior bioinformatic analysis reported by Xin et al, that the N-terminal half of both *Pst*AvrE and *E. amylovora* DspE contains a conserved double β-propeller motif (52). The architectural similarity of these predicted structures seems remarkable considering divergent amino acid sequences among these proteins (e.g., ~30% between *Pst* DC3000 AvrE and *E. amylovora* DspE) (Figure 3).

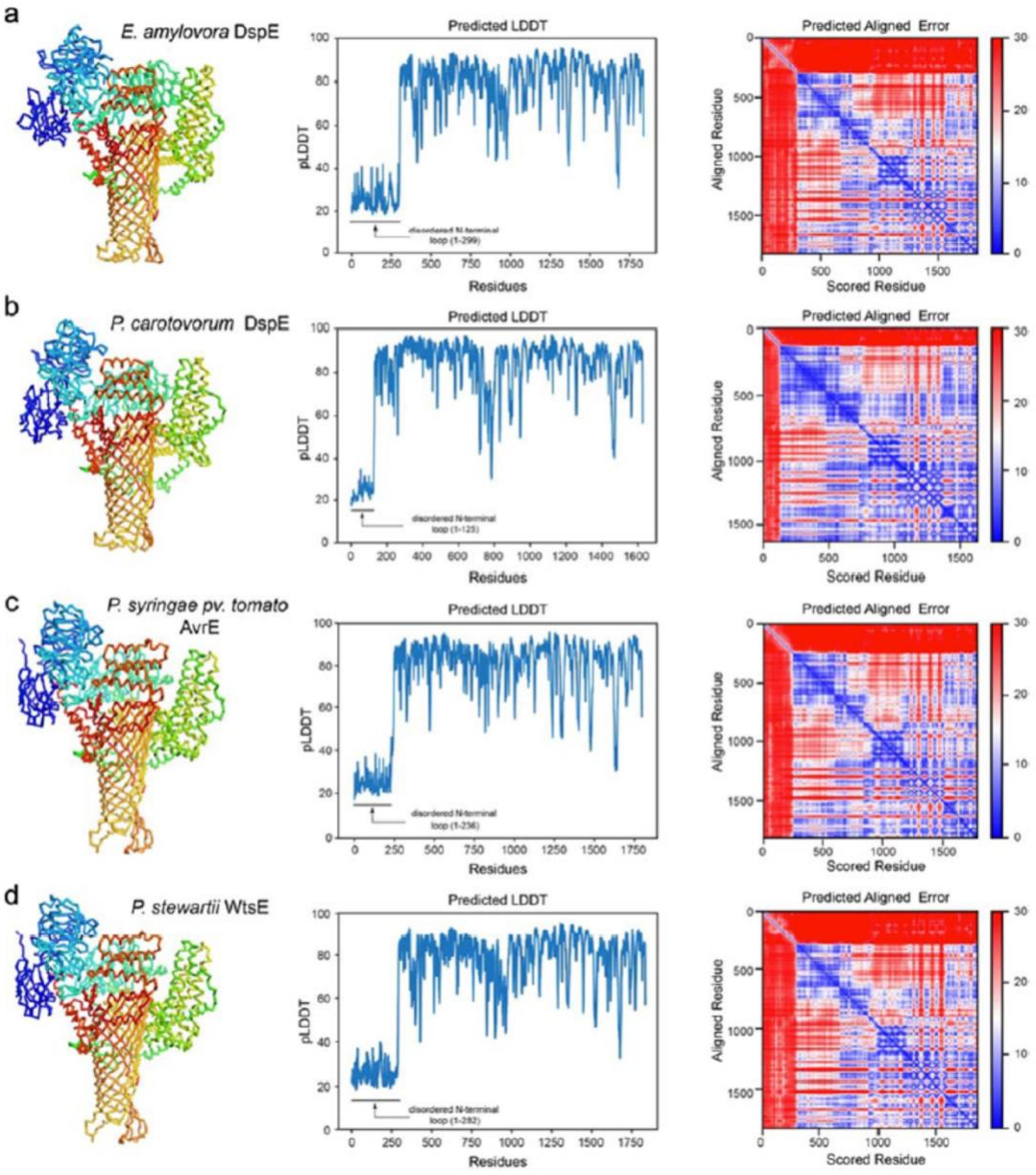


Figure 3 AF2-predicted models of major AvrE/DspE-family effector proteins.

The top-ranked models, predicted local distance difference test (pLDDT), and predicted aligned error (PAE) of *E. amylovora* DspE, *P. carotovorum* DspE, *P. syringae pv. tomato* AvrE, and *P. stewartia* WtsE are shown in panels a, b, c, and d, respectively. The models are predicted by AlphaFold2 using MMseqs2 (ColabFold). For each protein, the N-terminal disordered loop with low pLDDT scores is not shown, while the remaining residues are shown in Ca traces and colored in rainbow, with the N-terminal in blue and C-terminus in red.

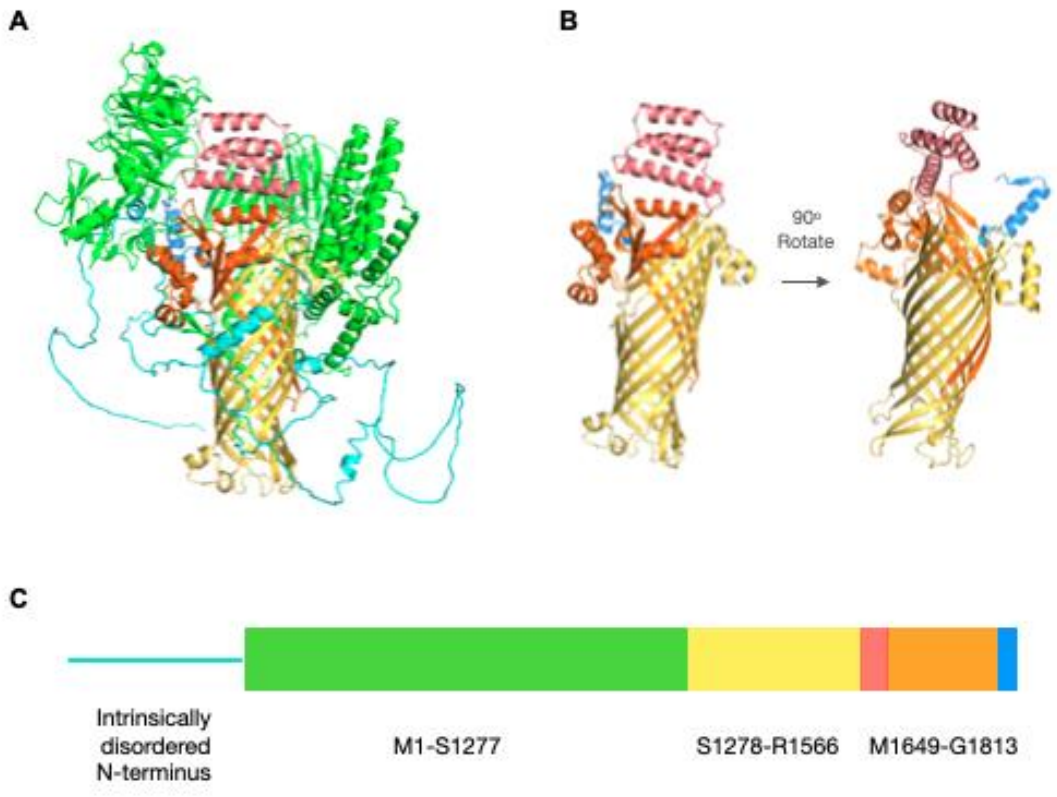


Figure 4 AF2-predicted model of full-length DspE, its c-terminal β-barrel portion, and a schematic representation.

(A) AF2 model of entire DspE from E. amylovora including the disordered N-terminus colored in cyan. (**B**) The AF2-predicted C-terminal β-barrel portion. Two parts consists of β-strands are colored in yellow and orange. (**C**) A schematic representation of the full length AF2-predicted DspE from *E. amylovora*.

Unexpectedly, AF2 predicts the C-terminus forming a bacterial porin-like β-barrel structure (Figure 4A). Specifically, take *E. amylovora* DspE as an example, the N-terminus of the AF2-predicted *Ea*DspE model consists of an intrinsically disordered loop (Figure 4A, colored in cyan), and several helix bundles and a β-propeller domain (Figure 4A, colored in green). The C-terminal half is organized into a β-barrel structure forming

a pore allowing chemicals/molecules passing through. A 4-stranded β sheets (Figure 4B, colored in orange) and a 10-stranded β sheets (Figure 4B, colored in yellow) are connected by a small helix bundle (Figure 4B, colored in light pink) and several short helix loops to form the β-barrel protruding out from the mushroom-like N-terminus.

This surprising prediction prompts us to utilize single-particle cryo-EM analysis to gain more atomic-level structural information of *Ea*DspE. Figure 5B presents 2D classes obtained from our initial cryo-EM data analysis. The predicted domain arrangement is supported by our cryo-EM imaging of *Ea*DspE, where the 2D class averages clearly reveal an overall similar top view as the AF2 model, with circularly arranged globular domains surrounding a central pore (Figure 5A&B).

Further examination of the predicted *E. amylovora* DspE 3D model reveals that surface β-barrel residues facing outside are enriched with hydrophobic residues, whereas inward-facing pore residues are predominantly hydrophilic amino acids. The length of the β-barrel stem covered by hydrophobic residues is estimated to be ~40 Å, roughly the thickness of a cellular membrane (Figure 5C). As AvrE has previously been reported to be membrane anchored (72), our analysis suggests that the protruding β-barrel stem of AvrE-family effectors likely inserts into the membrane and functions as a channel, similar to that of bacterial porins.

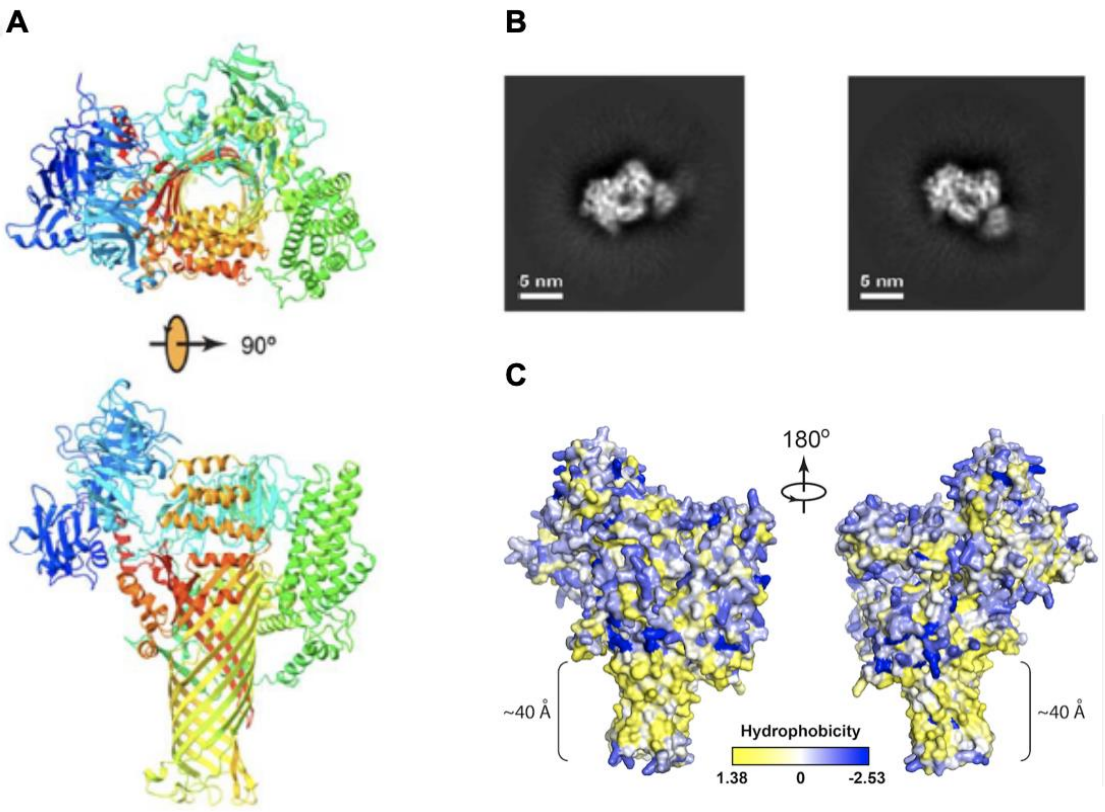


Figure 5 AF2-generated model, cryo-EM 2D classes and hydrophobicity scale.

(**A**) 3D model of *E. amylovora* DspE generated by AlphaFold2 using MMseqs2. DspE (residues 298-1838) is shown in the cartoon model in rainbow color, with N-terminus colored in blue and C-terminus colored in red. (**B**) Cryo-EM 2D class averages of DspE, revealing a circular arrangement of domains around a pore. (**C**) Hydrophobicity scale of DspE model, with hydrophobic residues colored in yellow and remaining residues colored in blue. The length of the β-barrel stem is labeled at 40 Å.

2.3 DspE anchors into liposome membrane

The predicted DspE pore have a diameter of 15-20 Å, significantly larger than the size of a water molecule or a simple ion. We hypothesized that, in addition to ions and water, the DspE pore may allow larger molecules to pass through. Thus, we conducted fluorescent dye permeability assays (Figure 6A) to test the membrane insertion

hypothesis and whether molecules smaller than the predicted pore size could pass through the membrane.

The liposomes made from soy extract lipids, comprising a mixture of phosphatidylcholine (PC), phosphatidylethanolamine (PE), and phosphatidylinositol (PI), mimic the natural cell membrane in the plant host. The unilamellar liposomes were prepared to encapsulate high concentration of carboxy-fluorescein (CF) inside, a small fluorescein molecule around 332 Da with an estimated maximum molecular diameter of 7 Å and can be self-quenched at high concentrations. Theoretically, only the breakdown of the CF-encapsulated liposomes or a pore on the liposome membrane allowing the pass through of small molecules leads to the increase of the fluorescence signals.

We did observe a significant gradual increase of fluorescence signal in the solution upon the addition of full-length DspE protein into the CF-loaded liposomes. A dosage-dependent response of DspE was also observed in the soy liposomes as DspE execute stronger effect on CF dye release with concentrations ranging from 0 uM to 0.5 uM (Figure 6B).

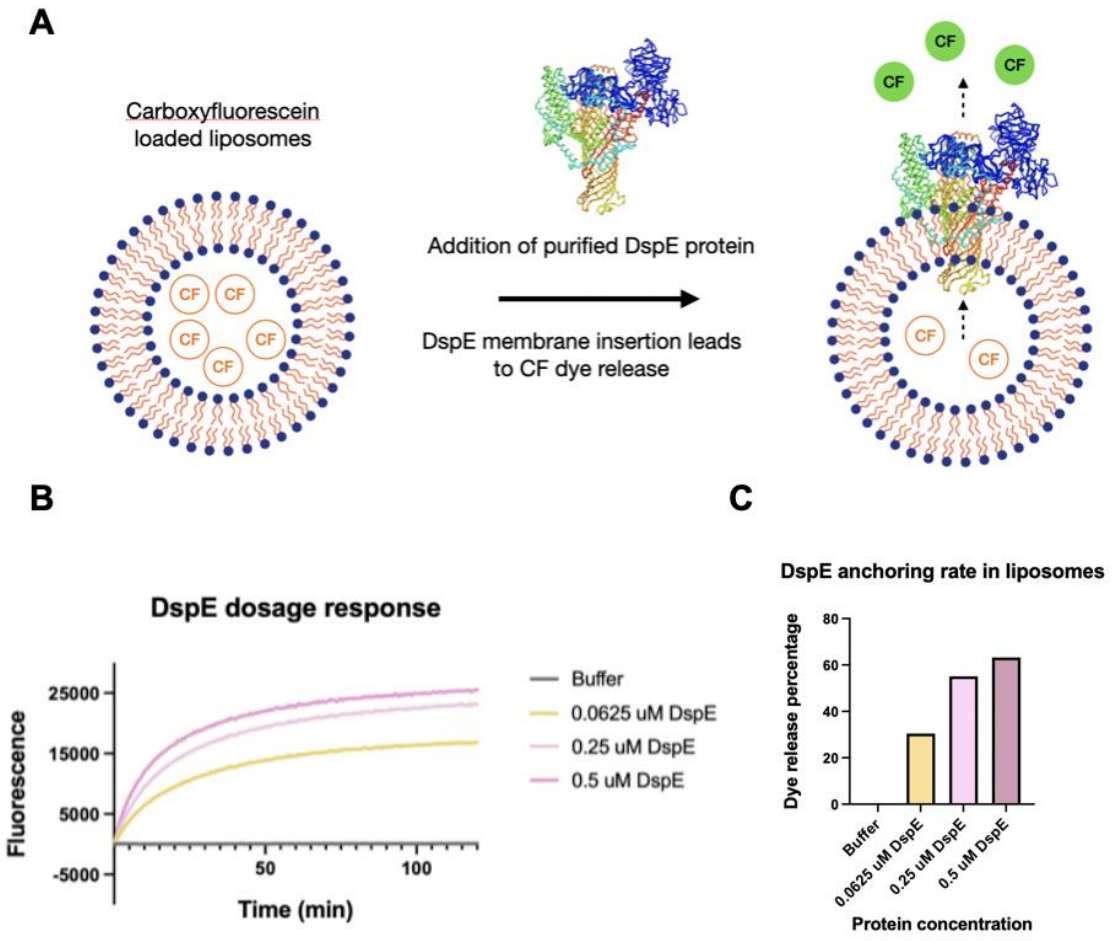


Figure 6 DspE causes dye release in carboxy-fluorescein (CF) dye permeability assay.

(**A**) Schematic illustration of DspE-dependent release of carboxy-fluorescein (CF) encapsulated within a soy liposome. (**B)** DspE dosage response in liposomes made from Soy extract (Avanti Polar). Curves in orange, grey, yellow, and blue indicating condition including DspE at 0.5, 0.25, 0.0625, and 0 uM concentrations. In this range, higher concentration of DspE added into the CF-loaded soy liposomes, higher fluorescence signals were detected. (**C**) DspE anchoring rate in soy liposomes. More than 60% fluorescence signals were detected once 0.5 uM DspE was added into the liposomes.

This liposome permeability assay confirms that the DspE pore alone is sufficient to allow the passage of small molecules such as fluorescein. Because at the end of assay, Triton was added to break down liposomes and to fully release the rest of CF dye, we

measured and calculated the percentage/ratio of the dye release caused by the addition of DspE protein. We found out that over 60% fluorescence signals were captured once 0.5 uM DspE was added (Figure 7A). Such a high percentage of dye release rate upon the addition of DspE indicates that DspE has a strong preference for membrane anchoring and insertion.

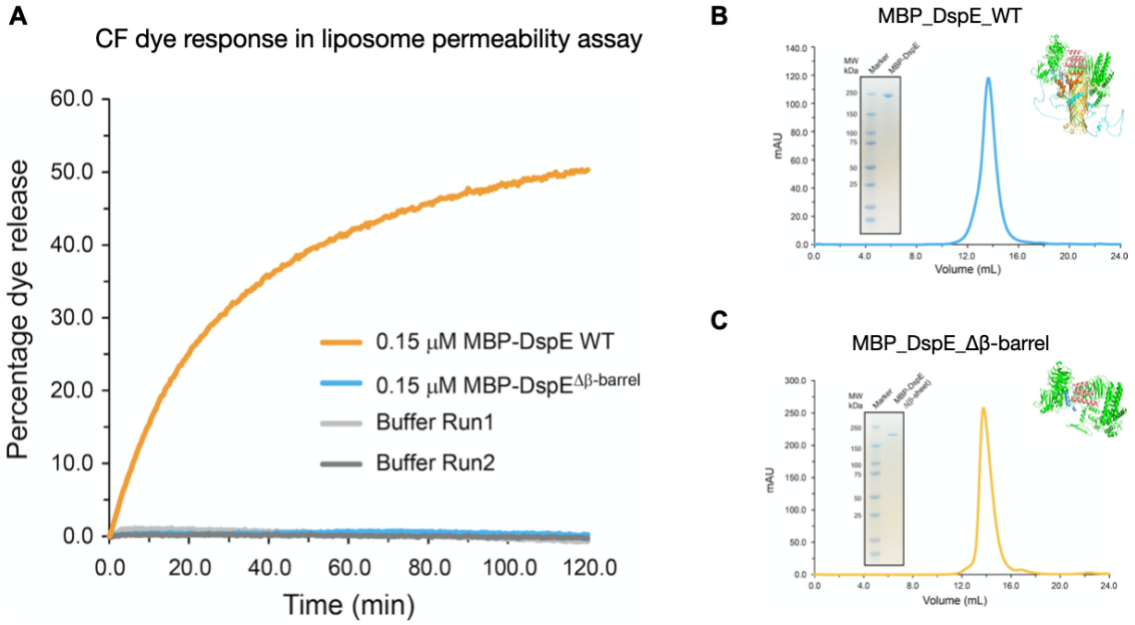


Figure 7 C-terminal β-barrel is essential for membrane insertion.

(**A**) CF dye release response in liposome permeability assay with 0.15 uM wild-type MBP_DspE and MBP_DspE $^{\Delta\beta\text{-barrel}}$, a β-barrel deletion construct. The addition of 0.15 uM MBP_DspE_WT causes over 50% dye release, however, no dye release response was observed upon the addition of MBP_DspE $^{\Delta\beta\text{-barrel}}$. (**B**) The gel filtration trace and SDS-PAGE gel of purified MBP_DspE_WT protein. (**C**) The gel filtration trace and SDS-PAGE gel of purified MBP_DspE $^{\Delta\beta\text{-barrel}}$, in which the main β-barrel portion is removed.

Based on the current AF2 model of DspE, we designed a truncation construct in which the C-terminal β-sheets in the β-barrel are mostly removed (Figure 7C) to

investigate the importance of DspE β-barrel in membrane insertion process. As Figure 7A presents, the addition of 0.15 uM full-length WT DspE causes nearly 50% of dye release, whereas no increase of fluorescence signal was observed upon the addition of 0.15 uM DspE$^{\Delta\beta\text{-barrel}}$ in the CF dye permeability assay. This evidence clearly demonstrates that the β-barrel is essential for DspE membrane anchoring and insertion.

In summary, all above indicate that DspE might insert a pore on the soy liposome permitting the CF molecules to pass through the pore and causing CF dye release in the liposome permeability assay. Furthermore, the integrity of the β-barrel at DspE C-terminus is indispensable for liposome membrane anchoring and insertion.

2.4 Mutational studies on DspE C-terminal β-barrel pore

As the C-terminal β-barrel is essential for DspE membrane anchoring and insertion, we further investigate the pore by mutational studies. According to the hydrophobicity analysis of AF2-predicted model (Figure 5C), in which the surface of β-barrel is enriched with hydrophobic residues, whereas the inside is full of hydrophilic residue, we identified several key residues in the transmembrane region which might be important for the pore function and performed a variety of activity assays including the oocyte assay and *in planta* water-soaking assay (Figure 8).

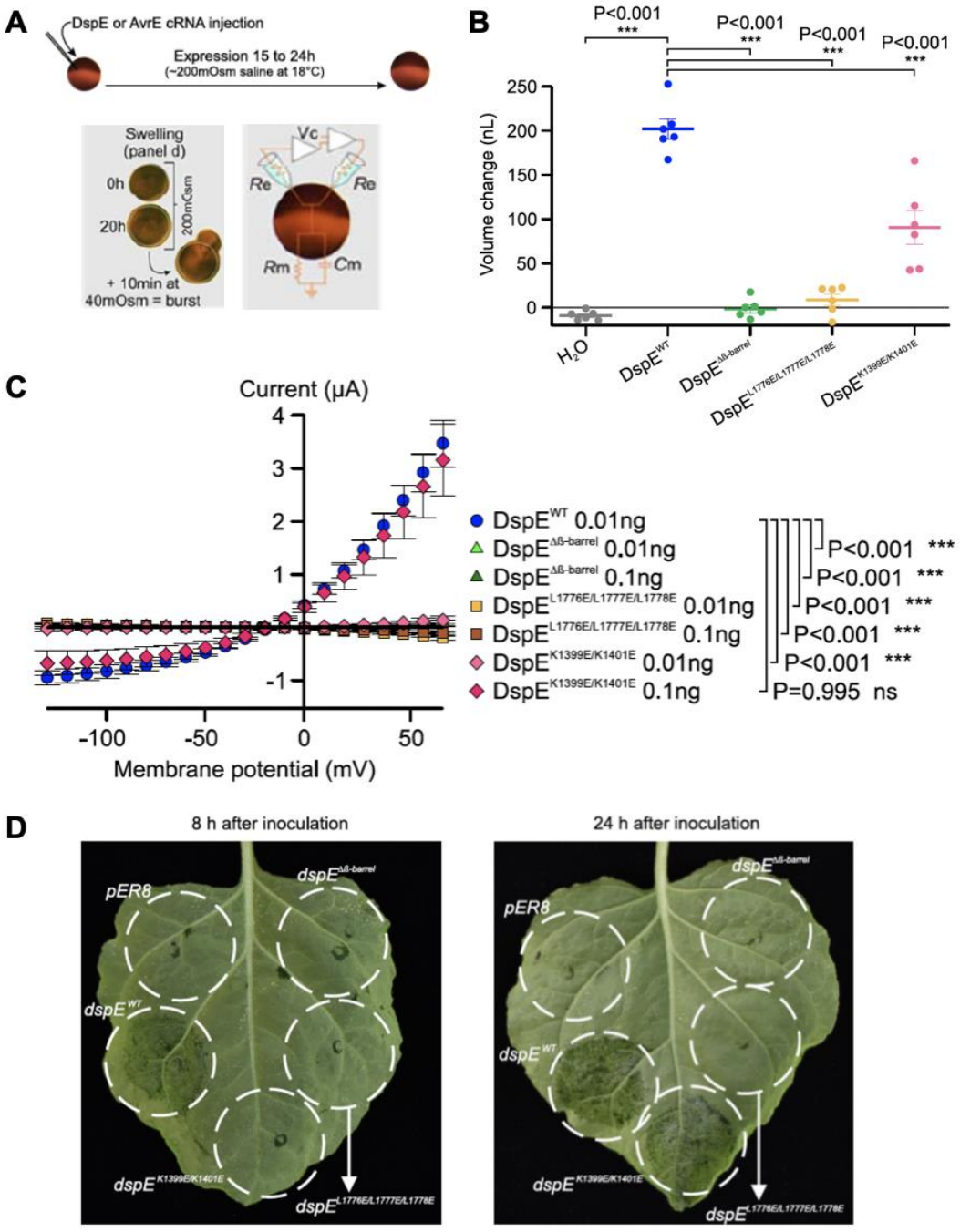


Figure 8 Oocyte and *in planta* assays of DspE mutants.

(**A**) Schematic representation of the two oocyte assays. Left: baseline and induced swelling/burst assay to test if water could flow across the oocyte membrane. Right: two-electrode voltage clamp

(TEVC, where Vc = voltage comman, Re = electrode resistance, Rm = membrane resistance, and Cm = membrane capacitance) to test if charged ions could flow across the oocyte membrane. (**B**) Quantitatively analysis of volume change of DspE mutants in oocyte swelling assay. Deletion of β-barrel abolishes the swelling symptom in oocytes, but both facing-in double mutant and facing out triple mutant show reduction of swelling symptom. (**C**) Ion currents captured in DspE mutants in the two-electrode voltage clamp recording experiments. (**D**) The ability of dspE mutants to induce water-soaking was assessed in N. benthamiana leaves. $1x10^8$ CFU/mL of Agrobacterium tumefaciens GV3101 containing pER8 empty vector, pER8-dspE or pER8-dspE mutants were syringe-inoculated into leaves of N. benthamiana and kept at 22°C for 24 hours before leaves were painted with 90 uM estradiol to induce DspE protein expression. Water soaking symptom was assessed at 8 h and 24 h post estradiol treatment.

As a negative control, $DspE^{\Delta\beta\text{-barrel}}$, in which the majority of β-barrel-forming sequences are deleted, failed to induce water-soaking/necrosis symptoms in *N. benthamiana* leaves, conduct ion currents, induce oocyte swelling or allow CF dye release in liposome permeability assay (Figure 7A & 8B-D). Then, we performed activity assays with a representative facing-out triple mutant (DspE_L1776E/L1777E/L1778E), in which three conserved hydrophobic out-facing residues (L1776, L1777, L1778) of the predicted transmembrane region is mutated. The facing-out triple mutant protein lost the ability to conduct ion currents in oocytes, to induce oocyte swelling and to cause water-soaking/necrosis symptoms in *N. benthamiana* leaves (Figure 8B-D).

Similarly, we mutated two in-facing residues in the β-barrel pore that is predicted to be important for DspE function. The double mutation at K^{1399} and K^{1401} partially abolish the ability of DspE to conduct ion currents in oocytes, induce oocyte swelling and cause water-soaking/necrosis symptoms in *N. benthamiana* leaves (Figure

8B-D). Future research is needed to comprehensively define essential in-facing residues that affect DspE function.

2.5 PAMAM G1 is an inhibitor of DspE

The discovery of DspE might act as a pore allowing flow of ions or molecules passing through offered us an opportunity to screen a series of compounds and to identify potential candidates that block DspE pore and inhibit its activity.

Early in 2010s, chemists became interested in using polyvalent scaffolds for rational pore-blocker design to inhibit channel/pore-forming bacterial toxins, secreted by pathogenic bacteria during invasion, to treat diseases. Since then, disruption of toxin activity as an anti-virulence strategy has been widely studied. And nowadays, there has been an assortment of inhibitors developed to specifically obstruct the activity of pore-forming toxins, in addition to vaccination and antibiotics that serve as a plausible treatment for most diseases caused bacterial infections (73).

There are two major types of bacterial pore-forming toxins, and one is membrane-perforating or pore-forming bacterial toxins (PFTs). The PFT is secreted as polypeptides or small proteins. Once accumulating on the membrane, they spontaneously form large, stable (last for hours) oligomeric transmembrane pore (36, 74). Another is the binary bacterial toxins, representatively, the binary anthrax toxin. These toxins are secreted as single-chain polypeptides consisted of at least two functionally distinct domains, the binding B domain that docks to specific surface

receptors and the active A domain that modified certain substrates in the cytosol (36, 74-76).

Recently, a variety of new synthetic multivalent pharmaceutical agents have been reported to be able to successfully neutralize several bacterial protein toxins (77). One unique pore-blocker is the dendrimer-based channel blockers. Dendrimers are rigid highly divaricate hyperbranched polymers, composed of regularly repeated branches originating from a central core. Due to their properties of molecular uniformity, narrow molecular weight distribution, defined size, rigidity, stability, and ability to form host-guest complexes, dendrimer has been wide used as antibacterial, antiviral, and antiparasitic agents (78).

In 2014, Föstner et al reported that poly(amidoamine) (PAMAM) dendrimers could be novel blockers for the pore-forming B components of the binary anthrax toxin (PA_{63}) and *Clostridium botulinum* C2 toxin (C2IIa), and suggested a potential role for PAMAM dendrimers as effective polyvalent channel-blocking inhibitors (75). This discovery led us to try the PAMAM blockers with DspE. PAMAM dendrimers are based on an ethylene diamine core and amidoamine repeat branching structure, which allows the variance in molecular diameters in different generations of PAMAM dendrimers ranging from 15 – 135 Å (Generation 1 – Generation 10). We first tested PAMAM dendrimers G0-G3 with diameters varying from 15 to 36 Å and then focused on PAMAM G1 (Figure 8C).

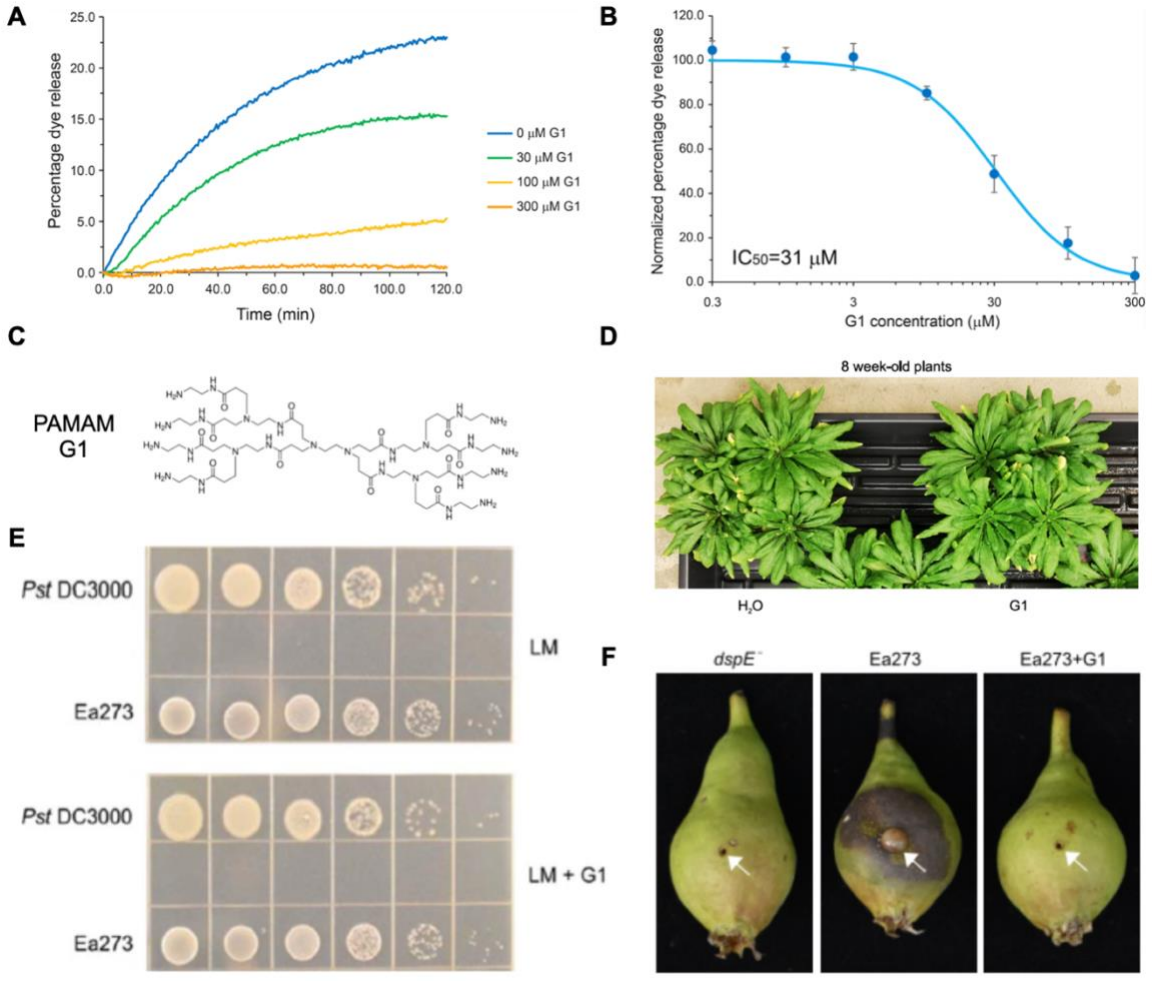


Figure 9 PAMAM G1 inhibits DspE.

(**A**) G1-dependent dye release response in DspE liposome permeability assay. **(B**) IC50 curve of PAMAM G1. **(C**) Chemical structure of PAMAM G1 compound. **(D**) Picture of 8-week-old plants. Five-week-old Col-0 plants were sprayed with 10 uM PAMAM G1 every week until 12 weeks old. PAMAM G1 does not affect plant growth of *Arabidopsis* Col-0 **(E**) Bacteria growth assay. PAMAM G1 does not impact bacteria multiplication *in vitro* (in LM medium). Ea273 or *Pst* DC3000 cells in the logarithmic growth phase were spotted, at a 10-fold serial dilution (from left to right), on LB agar plates with or without 50 uM PAMAM G1. Plates were kept at 28°C for 2 days for colony visualization. Experiments were performed three times with similar results. **(F**) Effect of PAMAM G1 on bacterial infection. Left panel: a representative pear injected with *Erwinia dspE* knock-out strain. Middle panel: a representative pear injected with *Erwinia amylovora* WT strain. Right panel: a representative pear injected with *Erwinia amylovora* WT strain and G1 inhibitor. G1 effectively inhibits *Erwinia amylovora* multiplication in pear.

In the liposome dye release assay, we observed a reduction of fluorescence signal upon the addition of PAMAM G1 in the proteoliposomes (DspE + liposome), which implies that G1 inhibits DspE. Figure 9A presents that the G1 inhibition is dosage-dependent and the IC50 is calculated as 31 uM (Figure 9B). Next, we tested whether the G1 compound affect host plant growth or bacterial growth and conducted plant and bacterial growth assay (Figure 9D&E). Clearly, neither PAMAM G1 affect plant growth or impact bacteria multiplication *in vitro*. However, the pear infection assay suggests that PAMAM G1 can prevent *Erwinia amylovora* infection on pears (Figure 9D). All above indicates that PAMAM G1 is anti-virulent and could be a good candidate for developing anti-virulence strategies to combat bacteria-mediated disease.

2.6 Discussion

Since the initial report of AvrE in *P. syringae* almost three decades ago (54), understanding the functional roles of AvrE/DspE-family effectors in diverse pathogenic bacterial species has been a long-standing quest of research. To date, prior efforts have been largely focused on understanding the genetics and on mapping the interaction networks of these T3Es (6, 14, 66, 70, 79-82).

Recently, more details in understanding bacterial pathogenesis in plants have been revealed. In 2016, Xin et al. reported that phytobacteria establish an aqueous intercellular space (apoplast) in plants to achieve virulence and pathogenesis, and the authors advocated that the T3Es may play indispensable roles in creating the water-

soaking environment in the apoplast (62). Furthermore, in 2022, Gentzel et al. reported that water and metabolite accumulation were observed in the apoplast of maize leaves infected by *Pantoea stewartii subsp. stewartii* (*Pnss*), preceding the disruption of host-cell integrity, and suggested that the formation of a hydrated and nutritive apoplast is driven by an AvrE/DspE-family type effector, WtsE (64).

In brief, these phytopathogenic bacteria perturb their hosts to create an environment suitable for their proliferation, including the suppression of immune defense in the host and the promotion of water and nutrient availability in the apoplast. Corresponding T3Es might act as a key water/solute-permeable channel at the beginning of bacterial infection given its broad distribution and the above biochemical analyses.

Here, we employed biochemical methods and structural biology tools unravelling the function of a representative T3E, *Ea*DspE, in the AvrE superfamily. To our knowledge, this is the first time that bacterial virulence factors have been directly shown to act as pathogenic water/solute-permeable pores/channels as a virulence mechanism. Specifically, we identify the integrity of C-terminal β-barrel is indispensable for membrane anchoring and insertion function, combining the AF2-predicted model and data from various activity assays. It offers a novel mechanism of actions for AvrE/DspE-family effectors in creating a water/nutrient-rich apoplast and further causing host cell death.

In summary, we propose a new model elucidating the function of T3E protein DspE during the process of bacterial infection (Figure 10). In early infection, the T3E DspE is secreted from *Erwinia amylovora* and injected into the host cell via T3SS. Once delivered in the plant host cytoplasm, the AvrE/DspE-family effectors act primarily as a novel class of water/solute-permeable pores/channels dedicated to creating osmotic/water potential perturbation and a water/nutrient-rich apoplast where bacteria are ape to multiply and proliferate within the infected plant tissues.

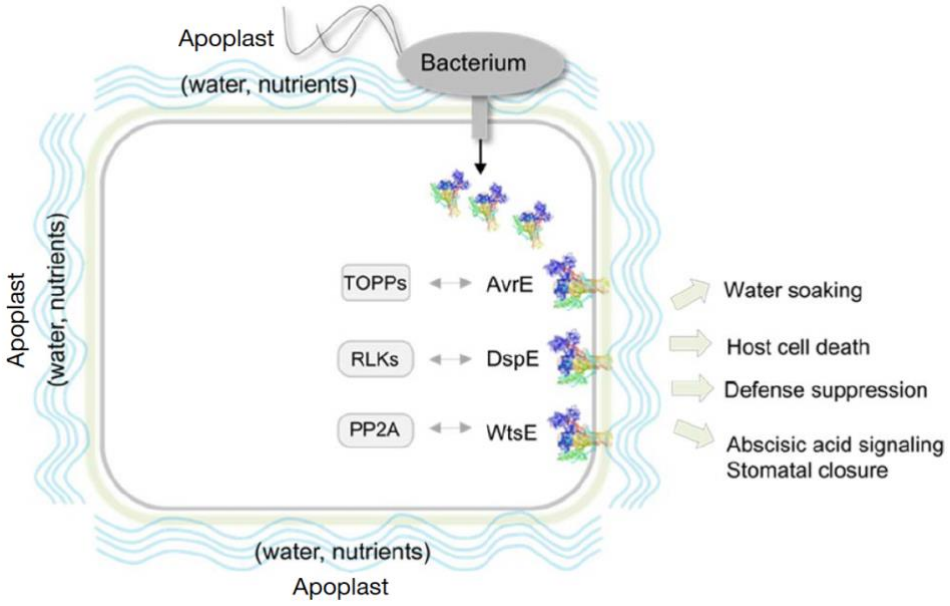


Figure 10 A proposed model for the mode of actions of AvrE/DspE-family effectors in host plants.

AvrE/DspE-family effectors act primarily as a novel class of water/solute permeable channels dedicated to creating osmotic/water potential perturbation and a water/nutrient-rich apoplast, in which bacteria multiply within the infected plant tissues. AvrE/DspE-family effectors can additionally engage host proteins, including plant protein phosphatase PP2A subunits, type one protein phosphatases (TOPPs) and receptor-like kinases (RLKs), possibly to modulate AvrE-family channel properties or to optimize the pathogenic outcomes of AvrE-family channel

activities, including water soaking, host cell death, defense suppression and /or abscisic acid signaling/stomatal closure (52, 57, 61, 62, 66, 72, 83-85).

Following current studies, in the future, we aim to continue capturing the native state of *Ea*DspE protein and to visualize the membrane insertion state of DspE via single-particle cryo-EM analysis. Additionally, it is our interest to design new pore-blockers and screen for novel anti-virulence compounds once atomic structural information is obtained. Beyond solving one of the longest-lasting puzzles in bacterial pathogenesis, discovery of AvrE/DspE-family channels may also lead to the development of pathogen-inspired molecular tools to study water/solute homeostasis in eukaryotic cells.

Moreover, recent studies in yeast indicated that certain T3Es activate protein phosphatase 2A and inhibit serine palmitoyl transferase, the first enzyme of the sphingolipid biosynthesis pathway (57), which could be alternative directions for future efforts. Investigating the interactions and dynamics between AvrE/DspE T3Es and their interacting partners could help us reveal the entire network in bacterial infection process and better understand how bacteria invade plants.

2.7 Methods

2.7.1 Cloning and Protein Purification

The *dspE* sequence was amplified using the genomic DNA of *Erwinia amylovora* strain Ea273 as the template and cloned into modified pET28a vector (Millipore Sigma)

as C-terminal fusion to maltose-binding protein (MBP) containing a His_8 tag, a preScission protease cleavage site (PPX), and a FLAG tag in the form of MBP-His_8-FLAG-DspE. Plasmid-transformed BL21(DE3) E. coli cells were grown in Luria-Bertani (LB) medium at 37 ^oC until OD_{600nm} reached 0.4-0.6 and were then induced with 0.1 mM IPTG and grown at 18 ^oC overnight. Harvested cell pellets were resuspended in a lysis buffer containing 20 mM HEPES (pH 7.5), 300 mM NaCl, and 2.5% glycerol, supplemented with cOmplete EDTA-free protease inhibitor tablet (Roche) and DNase I, and lysed French press. Following centrifugation at 20,000 rpm for 30min at 4 ^oC to remove cell debris, the fusion protein was purified using HisPur™ Ni^{2+} -NTA resin (Thermo Fisher Scientific). Following extensive washing in a buffer containing 20 mM HEPES (pH 7.5), 300 mM KCl, 2.5% glycerol, 1 mM ATP, 5 mM $MgCl_2$, and 30 mM imidazole, the fusion protein was eluted in a buffer containing 20 mM HEPES pH 7.5, 150 mM NaCl and 1 mM DTT at 4 ^oC. The protein fractions at the peak were aliquoted and flash-frozen at -80 ^oC for storage.

2.7.2 Cryo-EM sample preparation, data collection and processing

Cryo-EM grids were prepared using a Leica EM GP2 Automatic Plunge Freezer in a humidity-controlled chamber operated at 10 ^oC and 85-90% relative humidity. Homemade Gold Quantifoil® R1.2/1.3 300-mesh grids were glow-discharged using the PELCO easiGlow™ glow discharge cleaning system (TED PELLA, INC) before sample application. During sample freezing, a 3 μL sample of DspE (~1.1 mg/mL) was applied

to freshly glow-discharged grids and incubated on grids for 60 s before blotting with Whatman #1 filter paper (Whatman International Ltd) for 2.8 s. The grids were then immediately plunge-frozen in liquid ethane and stored in liquid nitrogen before data acquisition.

A total 7,810 movies were recorded on an FEI Titan Krios electron microscope (Thermo Fisher Scientific) operated at 300 kV equipped with a K3 direct electron detector (Gatan, Inc.) operated in the counting mode. Movies were collected at a nominal magnification of 81,000x using pixel size of 1.08 Å/pix with a defocus range from -2.4 μm to -0.8 μm using Latitude™ S (Version 3.51.379.0, Gatan, Inc.) automated image acquisition package. Each stack was exposed for 2.8 s with an exposure time of 0.047 s per frame, resulting in 60 frames per stack. The total dose was approximately 56.3 $e^-/Å^2$ for each stack.

Movie alignment and contrast transfer function (CTF) estimation were performed with the patch motion correction model and patch CTF estimation module in cryoSPARC (86). A total of 7,141 micrographs were selected from a total of 7,810 images based on the CTF fitting resolution using a cutoff value of 4.0 Å. A total of ~2.4 M particles were picked using pre-trained TOPAZ (87) models, of which ~167 K particles corresponding to MBP_His_Flag_DspE protein with high resolution features were selected to generate 2D class averages.

2.7.3 Preparation of unilamellar lipid vesicles

The soy extract lipids in chloroform were purchased from Avanti Polar Lipids and stored in glass vials (52). The lipids in chloroform were evaporated under a stream of nitrogen gas until it forms a thin lipid film and then dried in a vacuum desiccator chamber overnight. On the second day, the lipid film was dissolved in a suspension buffer (HBS: 20 mM HEPES, 300 mM NaCl, pH 8.0) containing 50 mM 5(6)-carboxy-fluorescein (Novabiochem ®). To solubilize lipids, the solution in the glass vial was sonicated for 15 min and incubated in a 37 °C water bath for at least an hour. Then lipids were subject to multiple freeze-thaw cycles. In each cycle, lipids were frozen in liquid nitrogen for 5 min and thawed in 37 °C water bath for 10 min. This freeze-thaw cycle was repeated 8 times to reduce the formation of multilamellar liposomes. To control the liposome diameter size, liposomes were extruded through a polycarbonate filter (200 nm, Whatman) 25 times using a mini extruder (Avanti Polar Lipids) with Hamilton glass syringes. Carboxy-fluorescein-loaded liposomes were purified by centrifugation at 41,000 rpm for 20 min in a Beckman TLA 100.3 rotor incorporating three sequential wash steps. After the final wash, carboxy-fluorescein-loaded soy liposomes were resuspended in HBS buffer to make the final liposome of 1 μg/μL (64, 88).

2.7.4 Liposome permeability assays

Release of the liposome contents was assessed using the self-quenching property and fluorescence of the carboxy-fluorescein dye. Permeability induced by protein DspE

was evaluated by incubating 0.5 μM protein with 90 μL carboxy-fluorescein-loaded liposomes (0.25 μg/μL). The total reaction volume was made up to 100 μL with HBS buffer. The fluorescence signal was measured for 2 hours once purified DspE protein was added into the liposome solution, afterwards 5% Triton X-100 (Sigma-Aldrich) was added to fully release the rest of dye and its readings were measured for 20 min. In PAMAM inhibitor tests, the buffer, DspE protein, and PAMAM G1 inhibitor were first mixed thoroughly with pipetting, then 90 μL carboxyfluorescein-loaded liposomes were added to make the total volume 100 μL. The following steps are the same.

The spectrofluorometric excitation and emission parameters were set at the wavelengths of 485 and 510 nm for carboxy-fluorescein molecules. The total fluorescence intensity was captured every 30s using a fluorescent plate reader (SpectraMax M3, Molecular Devices). The fluorescence signal was collected for 2 hours, afterwards, 5% Triton X-100 (2%; Sigma-Aldrich) was added and then fluorescence measurement was performed for 20 min.

2.7.5 Xenopus laevis oocyte preparation injection and expression of DspE and AvrE

Oocytes were purchased as ovary from Xenopus1 Co. (Dexter, MI). The ovary was treated with 0.55 mg/mL of collagenase B (0.191 U/mg) in calcium-free ND96 saline[38] for 20 min while on a naturing mixer at 21-22°C. Immediately after treatment, the enzymatic solution was rinsed off from the ovary with ND96 bath saline several

times and cell clusters were spread on several 70 mm plastic culture dishes with ND96 bath saline for temporary storage in an 18°C incubator. On the same day, the follicular cells and follicular membrane covering mature oocytes (stages IV and V) were manually peeled off with fine forceps and oocytes were kept in ND96 bath saline at 18°C until the time for cRNA injection. cRNA was mixed with diethyl pyrocarbonate (DEPC)-treated water to defined concentrations necessary to deliver desired amounts (ranging from 0.01 ng to 20 ng) of cRNA per oocyte when injecting a volume of 27.6 nL. Injection was performed with a nanoinjector (Nanoject II, Drummond Scientific Co. Broomall, PA) following the manufacturer's directions. DEPC-treated water was injected in control oocytes. Oocytes were kept in a 6-well plastic culture plate at 18°C to allow expression of proteins. Incubation solution was either control (ND96 bath saline), ND96 with inhibitor (PAMAM G0 or G1), ND96 with 0.0005% fluorescein or ND96 with 0.1% GFP protein.

2.7.6 Two-electrode voltage clamp recordings

Each oocyte was impaled by one voltage-sensing borosilicate microelectrode and one current-passing borosilicate microelectrode with resistance of 0.5±0.1 MOhm, while in an electrically grounded ND96 bath saline with or without PAMAM inhibitors. The glass microelectrodes were half-filled with 1.5% agar containing 3M KCl. The electrodes were connected to an oocyte clamp amplifier (OC-725C, Warner Instrument Corp. Hamden, CT) by chlorinated silver wires. The bath clamp headstage was connected to bath saline by two chlorinated silver wires inside a disposable polytetrafluoroethylene 18-

gauge tubing filled with 1.5% agar containing 3M KCl serving as agar-bridges. The oocyte clamp amplifier was connected to a computer by an analog-digital interface (Digidata 1440A, Molecular Devices, San Jose, CA). The command voltage protocols and data acquisition were performed in Clampex 10.7 (Molecular Devices, San Jose, CA). Oocytes were clamped to a desired potential using the fast clamp mode with maximum clamp gain and current gain set to 0.1 V/uA. Signal for both voltage and currents was recorded. Upon impaling an oocyte and before clamping it, both electrodes are capable to measuring the resting potential of that oocyte. Oocytes were clamped at their resting potential and test pulses of 100 ms toward more positive or more negative potentials in 10 mV increments were applied. The resultant current was recorded and analyzed. Current amplitude was determined at 2.5 ms to 5 ms of the test pulse, time at which there was the smallest or no overlaps with membrane capacitance or Ca-dependent chloride currents from endogenous oocyte channels[39]. Since resting potential values across individual oocytes varied (likely due to uncontrollable intrinsic differences in each oocyte, its size and in fine adjustment of electrode position and resistant), the voltage-current relationship data was fit to a quadratic polynomial regression (SigmaPlot 12.5 Systat Software Inc.) providing intermediate values and 95% confidence intervals. This also allowed currents elicited by the test potentials on control oocytes to be subtracted from the currents in treatment oocytes, so the resultant values represent only DspE- or AvrE-mediated current flowing across the membrane.

The two-electrode voltage clamp recordings were done by Dr. Felipe Andreazza, and the description is adapted from a manuscript under revision.

2.7.7 Oocyte swelling assay

Oocytes injected with 1 or 2 ng *dspE* or 20 ng *avrE* cRNA/oocyte and were imaged using Motic Images Plus 3.0 software connected to a Moticam X3 camera (Motic China Group Inc., China) on a SHR Plan Apo 1X WD:60 magnification lens of a stereoscope (Nikon SMZ18 Nikon Corp., Japan). At 0.01 ng *dspE* or 0.1 ng *avrE* cRNA/oocyte used for recording, no baseline oocyte swelling was observed. Baseline swelling was observed at 0.2 ng *dspE* or 2 ng *avrE* cRNA/oocyte. For induced oocyte swelling, starting oocyte images were recorded immediately after each injection, and then every 2 h to 4 h intervals for 24 h or until oocytes injected with *dspE* or *avrE* started to burst after the bath solution was adjusted to hypoosmotic. Oocytes were kept in bath saline of 200 mOsm with or without PAMAM inhibitors in the stereoscope room at 18-19°C for the entire period. Each picture depicting 5 oocytes (replicates) was analyzed with the Fiji software[40]. Data are presented as absolute volume at a given evaluation time or as change in volume in relation to the start point (immediately after cRNA injection). For osmotic swelling, oocytes expressing DspE or AvrE were transferred into a 5-fold diluted ND96 bath saline (40 mOsm) and were immediately imaged as described for baseline swelling once every 20 s for 10-20 min. Data are presented as change in volume in relation to the first picture in diluted saline. Pictures were also arranged in sequence to create time-lapse movies showing oocyte

swelling and bursting. One-way ANOVA with Tukey's test was used for multiple comparisons within a data set, with significance set to a P value < 0.05. For dataset with repeated measures over time, as in the osmotic swelling assay, a two-way repeated measure ANOVA with Dunnett's test was used instead, with significance also set to a P value < 0.05.

The oocyte assay was done by Dr. Felipe Andreazza. and the description is adapted from a manuscript under revision.

2.7.8 Bacterial media and plant growth assay

Bacterial strains used are wild-type *Pseudomonas syringae* strain DC3000 and its mutants: the ΔE mutant (89), ΔM mutant (89) and the ΔEM mutant (89) and wild-type *Erwinia amylovora* strain Ea273 and its *dspE* mutant (56). Bacteria were grown in low-salt Luria-Bertani (LB) medium at 28°C. Antibiotic ampicillin, kanamycin or rifampicin was added at 100, 50 or 100 μg/ml, respectively. *Arabidopsis thaliana* Col-0 plants were grown in Redi-Earth potting soil (Sun Gro Horticulture) in air-circulating growth chambers. Plants were grown under relative humidity set at 60%, temperature at 20°C, light intensity at 100 $\mu E \cdot m^2 \cdot s$ and photoperiod at an 8 h light-16 h dark cycle. Four- to five-week-old plants were used for bacterial disease assay. Immature pear fruits were gifts from George Sundin at Michigan State University.

The bacteria growth assay was done by Dr. Kinya Nomura. and the description is adapted from a manuscript under revision.

2.7.9 Transient expression of DspE in *Nicotiana benthamiana*

dspE and *dspE* mutant ORFs were PCR-amplified with *dspE* ORF forward primer (TTGGGCCCATGGAATTAAAATCACTGGGAACTG, underline indicates *Apa*I site) and *dspE* ORF reverse primer (TTTACTAGTTTAGCTCTTCATTTCCAGCCCTTCC, underline indicates *Spe*I site) and *pGH-dspE* or *pGH-dspE* mutant plasmids as template. PCR-amplified *dspE* and *dspE* mutants ORF (*Apa*I-*Spe*I fragment) were cloned into the binary vector *pER8* (https://pubmed.ncbi.nlm.nih.gov/11069700/) to create *pER-dspE* and *pER-dspE* mutant constructs. All constructs were transformed into *Agrobacterium tumefaciense* GV3101 for transient expression assay. $1x10^8$ CFU/mL of *A. tumefasciens* GV3101 containing *pER8* empty vector, *pER-dspE* or *pER-dspE* mutant were syringe-inoculated into leaves of *Nicotiana benthamiana* and kept at 22℃ for 24 h before leaves were painted with 90 μM estradiol. Eight hours later, leaf samples were collected for western blot. Water-soaking/and necrosis symptoms were recorded at 8 h and 24 h after estradiol treatment.

The *in planta* water-soaking assay was done by Dr. Kinya Nomura. and the description is adapted from a manuscript under revision.

2.7.10 Bacterial disease assay

Disease assays with immature pear fruits were performed as previously reported (56). Pears were surface-sterilized with 10% bleach for 5 min and rinsed in sterile water twice. Then, a small hole was made in the pear using a 200 μL tip. Ten μL of 10^3 CFU/mL

Ea273 or the *dspE* mutant was loaded into the hole. Inoculated pears were placed on a wet paper towel in a sterile box to maintain high humidity at 28℃ for 10 days. Disease assays with Arabidopsis plants were performed as follows. Arabidopsis plant leaves were infiltrated with *Pst* DC3000, ΔE, ΔM or ΔEM at 10^6 CFU/ml with a needle-less syringe. After visible water soaking disappeared (within 1 h), plants were kept under high humidity (~99%) at 23°C. Bacteria population in leaves was determined at day 3 post infiltration. Detached leaves were surface-sterilized in 75% ethanol for 30 s and rinsed in sterile water twice. Then, leaf discs (1 cm^2 in diameter) were punched out and ground in 100 μL sterile water. Ten μL of each 10-fold serial diluted leaf extract was plated on LB rifampicin and kept at 28℃ for 24 h. Colony forming units (CFUs) were counted under microscope before colonies started to coalesce and analyzed by GraphPad Prism software. Two-way ANOVA with Tukey's test was used for multiple comparisons within a data set, with significance set to a P value < 0.05. For inhibition assays, 50 nM PAMAM G1 was added to bacterial suspension and co-inoculated into plants.

The bacterial disease assay was done by Dr. Kinya Nomura. and the description is adapted from a manuscript under revision.

2.7.11 AvrE/DspE sequence alignments

Sequences of *E. amylovora* DspE, *P. carotovorum* DspE, *Pst* DC3000 AvrE, *P. agglomerans* DspE, and *P. stewartii* WtsE are aligned using Clustal Omega (67).

Chapter 3. Structural elucidation of the plant immune regulators NPR1 and NPR3

NPR1 is a master immune regulator in plants. It orchestrates SAR by activating *PR* genes in response to induction of SA during the plant response to pathogenic challenges. It was first discovered and identified in early 1990s (22), and since then, understanding the role of NPR1 and NPR1-SA interaction in mediating plant responses has been a long-standing research interest (22, 23).

In early 1990s, genetic studies showed that the *npr1* gene affected the *PR* gene expression in plants. Specifically, *Arabidopsis npr1* mutants have displayed increased disease susceptibility and a decrease in SAR-triggered *PR* gene expression, for instance, the *PR1* (22, 23). Importantly, complementing *npr1* mutants with wild-type NPR1 could restore resistance and *PR* gene expression (22).

Following the genetic studies, later functional examination mapped protein domains in NPR1 protein and identified several unique motifs including the ankyrin repeat (ANK) and the BTB/POZ domains. Both of them are known for facilitating versatile protein-protein interactions in a wide range of biological events (90). In 2000, Kinkema et al. characterized a bipartite nuclear localization sequence which allows for nuclear localization of NPR1 following SA induction (91). Meanwhile, a subset of the TGA family of bZIP transcriptional factors are found to be activated by NPR1 leading to expression of *PR* genes and SAR induction (25, 91, 92). So far, most of the key functional

motifs and interacting partners have been identified and characterized. And all these data suggest that NPR1 is a transcription cofactor responsible for effecting SA-dependent signaling (93).

In 2003, Mou et al. reported that NPR1 can be found as a large cytoplasmic oligomer in the non-stress condition, but a reduction of the cytoplasmic oligomeric form is observed during SA-induced SAR (23, 26). The redox state of NPR1 may contribute to the NPR1 monomerization, nuclear localization and *PR* gene expression (23). This discovery raises further questions of how these different oligomeric states of NPR1 in cells regulate through cellular redox changes and mediate plant immune responses. Due to the lack of structural information, the molecular basis, and the mechanism by which SA activates NPR1 have not been fully understood. Therefore, we resolved the structural details of truncated NPR1 and full-length NPR1 protein by X-ray crystallography and single-particle cryo-EM analysis.

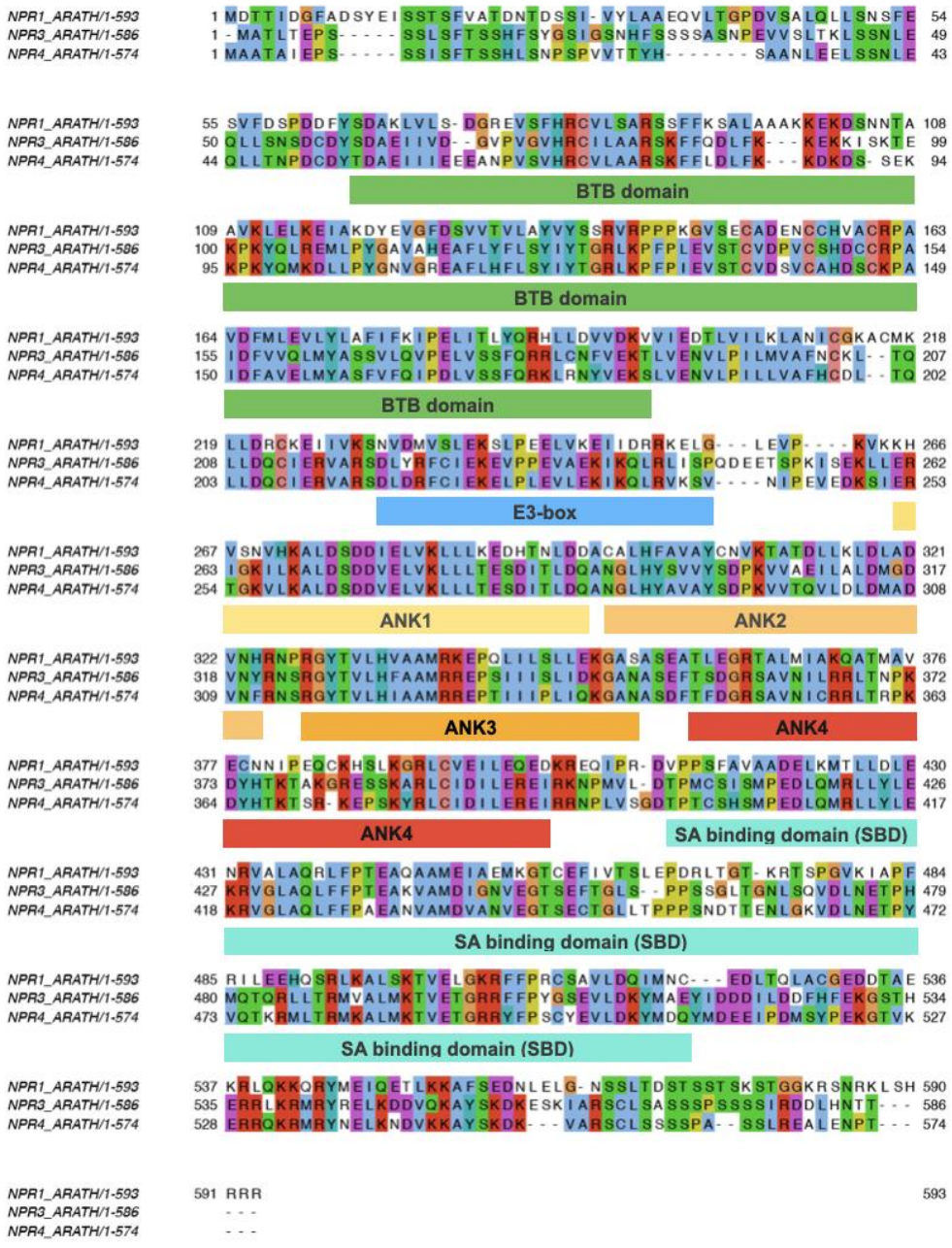


Figure 11 multi-sequence alignment of full-length *At*NPR1, *At*NPR3 and *At*NPR4.

The hydrophobic residues, hydrophilic residues, residues with polar uncharged side chains are highlighted in different colors. The colored boxes underneath the sequence mark specific

domains and motifs. All NPR protein contains a BTB domain and E3-box at the N-terminus, and 4 ankyrin repeats and a SA binding domain at the C-terminus.

Besides NPR1, numerous NPR1-like proteins, both putative and confirmed, have been identified in diverse plant species including *Oryza sativa* (rice) and *Nicotiana tabacum* (tobacco). In *Arabidopsis* only, five additional NPR1-like proteins have been reported: *At*NPR2, *At*NPR3, *At*NPR4, *At*NPR5, *At*NPR6 (94). Interestingly, all NPR proteins share a N-terminal BTB domain, a central ankyrin repeat (ANK) domain and a C-terminal (CT) domain with high conservation (23). Here, Figure 11 presents the sequence alignments of NPR1 and its two paralogs, NPR3 and NPR4, which share about 34.5% and 36% amino acid conservation with NPR1, respectively.

Despite the high levels of sequence and functional conservations among NPR proteins, the complexity and variability in functions also increase. In *Arabidopsis*, prior phylogenetic analysis indicate that *At*NPR1 and *At*NPR2 are involved with positive SAR regulation, whereas *At*NPR3 and *At*NPR4 with negative SAR regulation (95). In 2012, Fu et al. proposed that NPR3 and NPR4 primarily function as adaptor proteins for Cullin 3 (Cul3), an E3 ligase mediating NPR1 degradation (21). The authors demonstrated that NPR3 and NPR4 possess different affinities for the substrate SA, thus allowing them to effectively regulate NPR1-dependent gene expression through CUL3-mediated proteasome degradation of NPR1(21). On the contrary, Ding et al. suggested that the NPR3/NPR4 are negative regulators of SA in plant immunity, and instead act as transcriptional co-repressors (20). Particularly, NPR1 and NPR4 have opposite roles in

early defense gene expression in response to SA(20). These two opposite arguments raise our interests to obtain detailed structural information via X-ray crystallography to resolve the controversy.

3.1 Crystal structure of NPR1 reveals a zinc-finger motif for interacting with ANK and mediating NPR1 oligomerization

To determine the molecular details of the NPR1 protein, we used X-ray crystallography to visualize its detailed structural information. We first designed a NPR1(ΔSBD) (Thr39-Asp410) removing flexible regions for crystallography and obtained crystals through sitting-drop vapor diffusion. After extensive screening and optimization, a well-diffracting crystal was obtained and solved at 3.06 Å (Figure 12A).

The NPR1(ΔSBD) forms homodimer in the crystals and the overall architecture resembles a gliding bird (Figure 12B). The N-terminal NPR1 contains a BTB domain which consists of a three-stranded β-sheets (B1-B3) flanked by two short helices and three long helices (A1-A5; Figure 12D). Similar to other homo-dimeric BTB domains (90), the NPR1 BTB domain contains an additional dimerization helix at the far N-terminus earlier than the BTB domain (α0; Figure 12B&D) (90). Following the BTB domain, NPR1 contains a unique four-helix bundle (A6-A9, Figure 12D) that shares the helical pair repeat pattern with the BTB and carboxyterminal Kelch (BACK) domain (96, 97), but is much shorter than traditional BACK (4 versus 8 helices) and therefore is named as the BACK helix bundle (BHB) domain.

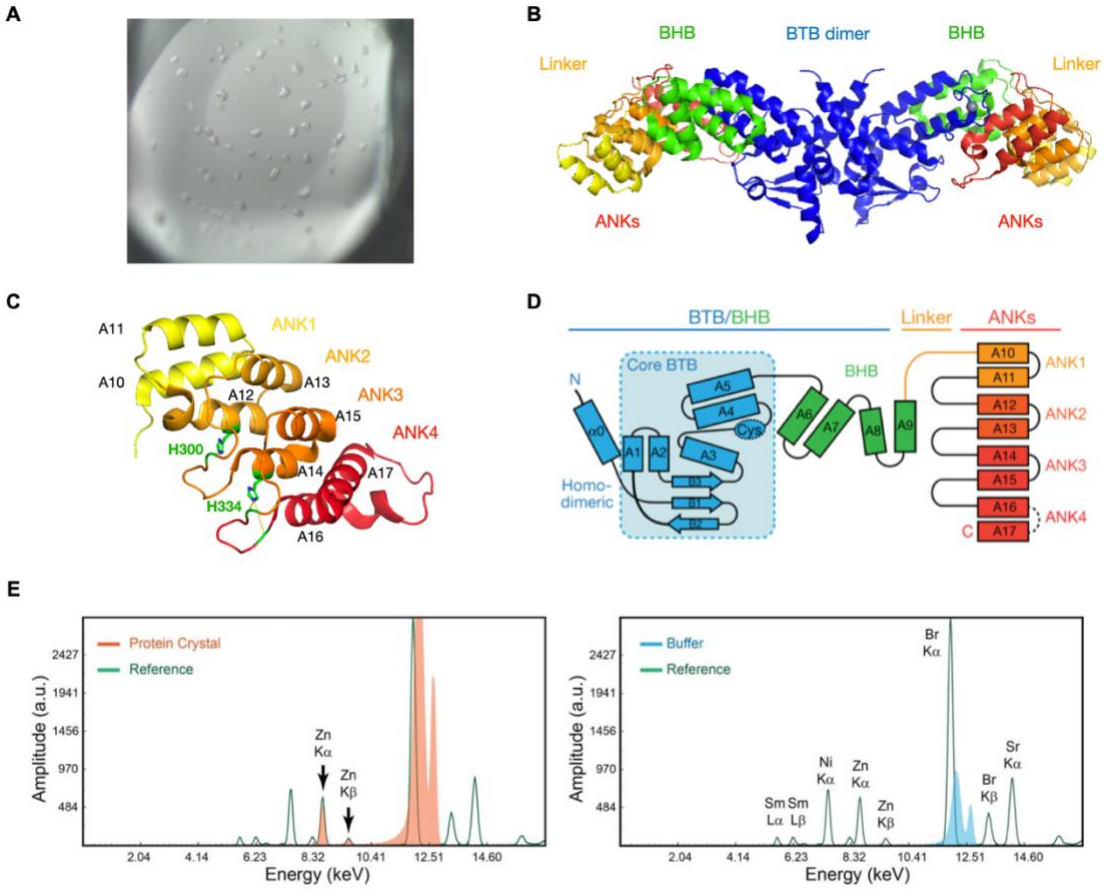


Figure 12 Analysis of NPR1(ΔSBD) crystal (PDB: 7mk3).

(**A**) Representative of NPR1(ΔSBD) crystals. (**B**) The NPR1 homodimer (residues 40-402) (**C**) Stabilization of NPR1 ANKs through His300/His334-mediated hydrogen-bond network. Bb, backbone amides. (**D**) Topology of NPR1 BTB, BHB and ANKs. Cys, cysteine cluster. (**E**) X-ray fluorescence scanning data revealed the presence of Zn^{2+} in NPR1(ΔSBD) crystals. Scanning results for the NPR1 protein crystal and buffers are shown in the left and right panels, respectively.

Right after the BHB, an extended linker loop connects the BHB to ankyrin repeats (ANKs) (Figure 12B&D). NPR1 contains three well-defined ANKs (98, 99) and a non-canonical fourth ANK with two α-helices linked by a 16-residue loop (Figure 12C). The last three ANK2-4 are connected by extended flexible loops in the length of ~15 amino

acids, whereas the first two ANK1-2 are joined by a short loop (Figure 12C). A distinguishing feature of ANKs is the presence of conserved histidine residues bridging neighboring repeats (99). Similarly, His300 of ANK2 and His334 of ANK3 form extensive hydrogen bonds with the neighboring charged residues on the inter-ANKs loops (Figure 12C). All three ANK1-3 consists of helix-turn-helix bundles with short turns consisting of 3 amino acids, however, the two helices in ANK4 is connected by a long turning loop made from 10 amino acids. As the NPR1 C-terminus is proposed to interact with SA for regulation and signaling purposes, the non-canonical ANK4 with an extended loop may provide greater structural flexibility for interacting with substrates. In brief, our crystal structure of NPR1 homodimer is consistent with the cryo-EM-derived NPR1 structure by Dr. Shivesh Kumar (100).

The X-ray fluorescence scanning result also reveals the presence of Zn^{2+} ions in the protein crystal, not in the buffer. Accordingly, analysis of the crystal structure disclosed a distinct C_2HC zinc finger within a functionally important cysteine cluster that is conserved in plant NPR proteins. The Zn^{2+} ion is coordinated by Cys150, Cys155, His157, and Cys160 in a consensus sequence of **C**-X-**D**-$X_{2/3}$-**C**-X-**H**-X_2-**C** located within the A3-A4 loop of the BTB domain (Figure 12A), whereas bold letters and Xs indicate conserved and non-conserved residues, respectively.

As part of its role in maintaining the structural integrity, the zinc finger bridges the BTB domain to ANK4 through a network of polar interactions (Figure 12A(insect)).

The carboxylate group of Asp152, the most conserved residue within the zinc finger besides the zinc-coordinating residues, form two hydrogen bonds with the side chains of Lys370 and Gln374 f ANK4. The zinc-finger – ANK4 interaction is also 'zip-locked' by backbone hydrogen bonds involving Glu149 and Ala151 from the zinc finger, and Lys370, Thr373 and Ala375 from ANK4. As Lys370 is located at the last turn of the ANK helix, the Ala151-Lys370 backbone hydrogen bond is fortified by the helix dipole effect (101).

Moreover, the formation of the zinc finger, with side chain of Cys155 and His157 pointing inwards to coordinate Zn^{2+}, pivots Cys156 directly into solvent (Figure 12A (insect)). Notably, we found that, in the crystal structure, the protein lattice is formed by head-to-head stacking of two NPR1 dimers, and the pair of Cys156 residues from each NPR1 dimer are juxtaposed against their counterparts in the tetramer (Figure 13B). Similar to the previous report of NPR1 BTB domain(102), we found that hydrogen peroxide treatment of the NPR1 BTB-BHB domain (Thr39-Lys262) readily yielded dimer, trimer and tetramer bands on the non-reducing SDS-PAGE (Figure 13C, courtesy of Dr. Qinglin Wu). The NPR1 C156A mutation strongly reduced dimer formation and eliminated oligomer bands (trimer and tetramer), supporting Cys156 as a redox sensor. Within the NPR1 dimer at the monomer-monomer interface lie two Cys83 residues also in proximity, and the C82A mutation reduced the formation of trimer and tetramer bands (Figure 13C), indicating its important role in the formation of higher oligomers.

By contrast, the C212V/C216A/C223 triple-mutant acts as a negative control and behaves similarly to the WT protein, suggesting that these non-zinc-chelating cysteines are not involved in the NPR1 oligomerization in vitro.

Besides the discovery of a zinc finger within the NPR1 crystal, the high-resolution crystal structure fills the gap on certain regions where the cryo-EM density did not provide sufficient resolution for confident model building.

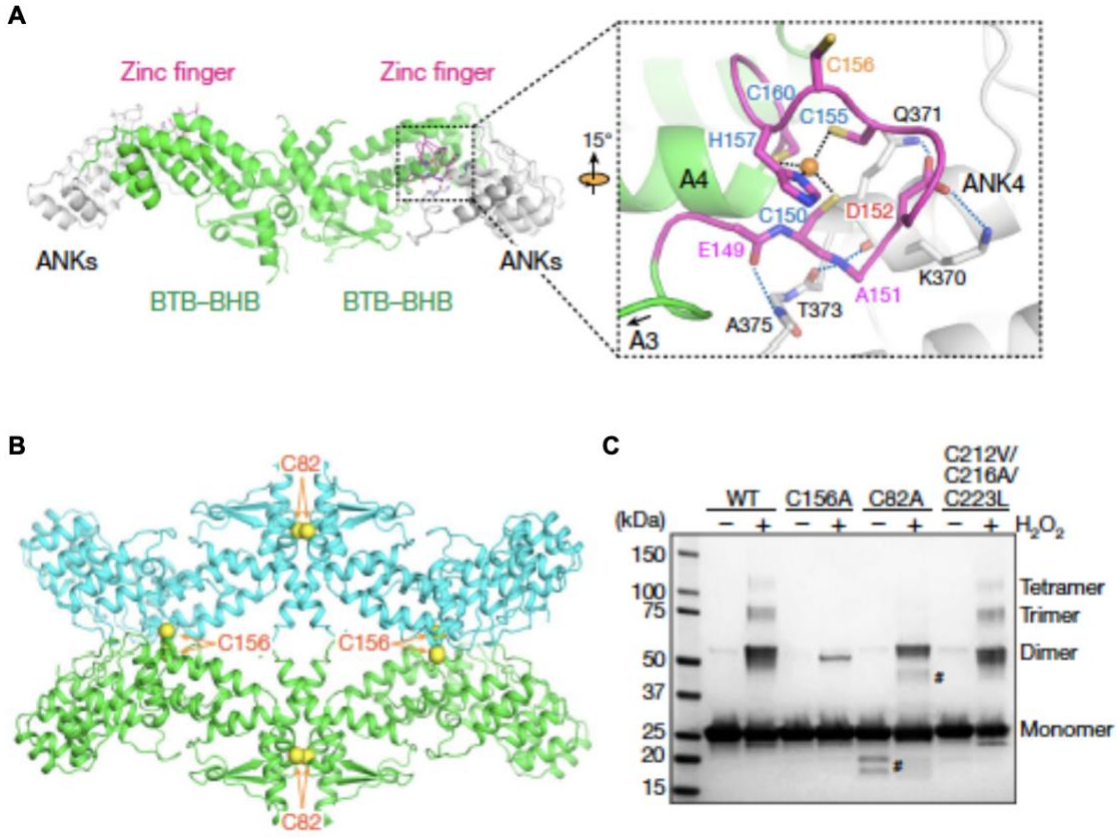


Figure 13 NPR1 contains a unique zinc finger motif and a redox sensor.

(**A)** NPR homodimer (residues 39-405) with the zinc finger detailed. The coordination of Zn^{2+} (coral sphere) by side chains of Zn^{2+} -chelating residues (shown in the stick model and labelled in blue) is indicated by black dashed lines. The hydrogen-bond network between the zinc finger and

ANK4 is indicated by blue dashed lines, with the pertinent side chain and backbone groups shown in the stick model. The side chain of Cys156 is shown in the stick model. (**B**) NPR1(ΔSBD) (Thr39-Asp410) tetramer in the protein crystal. (**C**) In vitro oligomerization of NPR1 (residues Thr39-Lys262), WT or the C156A, C82A and C212V/C216A/C223L mutants, indicated by hydrogen peroxide. The hash symbol (#) indicates impurity. The image is representative of two and three biological replicates, respectively (Courtesy of Dr. Qinglin Wu).

This work has been published and the description is adapted from our published manuscript (100).

3.2 Crystal structure of NPR3 reveals a zinc-finger stabilized by interactions between the BTB and the ANK4

Based on our prior efforts on structural studies of *At*NPR1, we designed a truncated *At*NPR3$_{S34\text{-}D406}$ construct in which the N-terminal and C-terminal flexible regions are removed, but the main BTB domain and the ankyrin repeats regions (ANKs) are intact (Figure 12). After extensive screening and optimization, we obtained well-diffracting crystals of *At*NPR3$_{S34\text{-}D406}$ using sitting-drop vapor diffusion method and determined its atomic structure at high resolution of 2.4 Å (Figure 14).

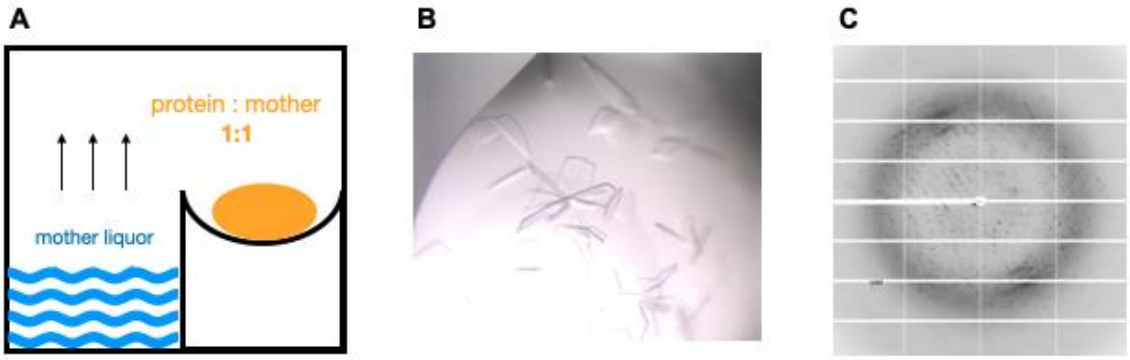


Figure 14 Diffraction quality protein crystals of *At*NPR3(S34D406) were obtained by sitting-drop vapor diffusion.

(**A**) Sitting-drop vapor diffusion illustration. (**B**) Representative *At*NPR3(S34D406) crystals. (**C**) Diffraction pattern of *At*NPR3(S34D406) crystals.

Table 1 X-ray data collection and refinement statistics of *Arabidopsis* NPR1(ΔSBD)

	NPR1ΔSBD (PDB 7MK3)
Data collection	
Space group	P 41 21 2
Cell dimensions	
a, *b*, *c* (Å)	78.98, 78.98, 395.37
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	48.62 – 3.06 (3.17 – 3.06) *
R_{merge}	0.127 (2.003)
R_{pim}	0.0202 (0.2969)
$CC_{1/2}$	0.998 (0.826)
$I / \sigma I$	19.97 (3.14)
Completeness (%)	99.84 (99.96)
Redundancy	42.7 (45.7)
Refinement	
Resolution (Å)	3.06
No. reflections	24,826 (2429)
R_{work} / R_{free}	0.2286 / 0.2628
No. atoms	5539
Protein	5472
Ligand/ion	67
Water	0
B-factors	113.76
Protein	113.78
Ligand/ion	112.03
Water	N/A
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.51

*values in parentheses are for highest-resolution shell.

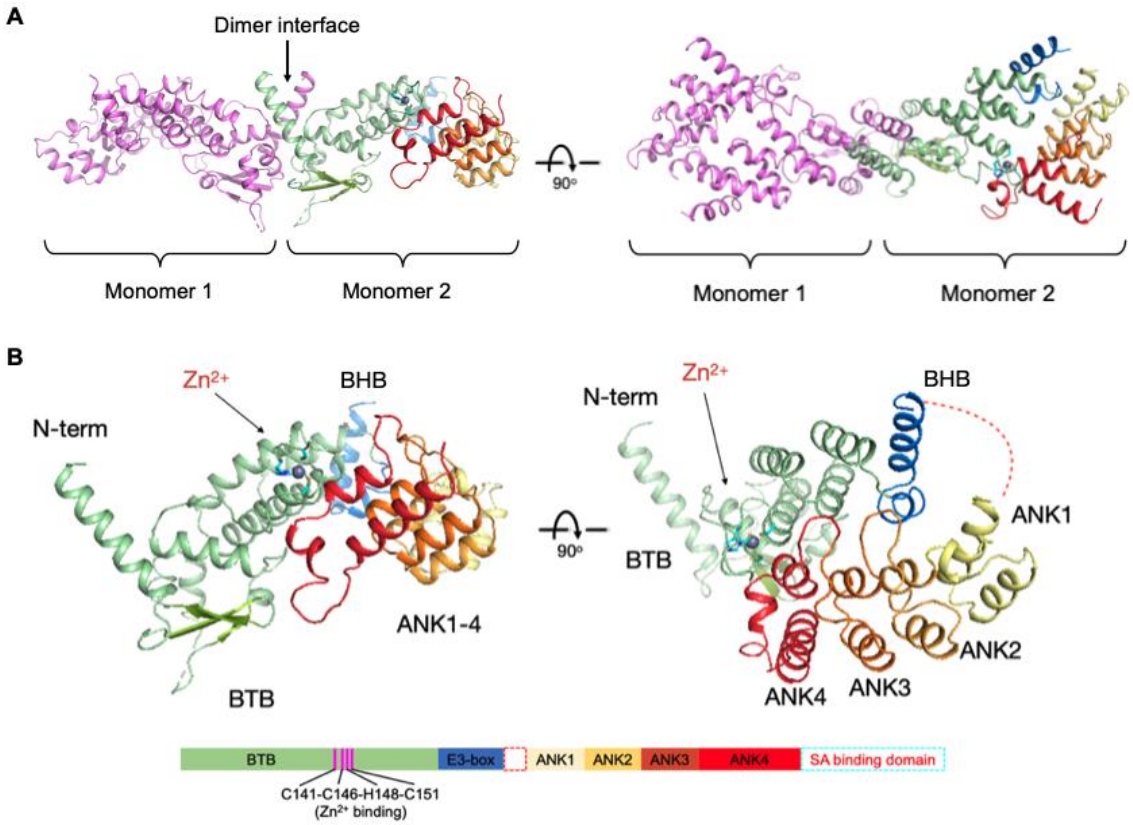


Figure 15 Overall architecture of *At*NPR3 homodimer and monomer.

(**A**) The overall architecture of AtNPR3 homodimer. (**B**) Left is a side view of NPR3 monomer including a BTB domain (green) with a Zinc binding motif, BHB, four ANK repeats. Right is a top view of the NPR3 monomer. The dashed red line represents a missing region of the e3-box and its linker to ANK1.

In crystals, *At*NPR3 forms a homodimer, in which the center was bridged by N-terminal BTB domain (Figure 15A). Consistent with sequence alignment, NPR3 contains N-terminal BTB domain (S60-T185) consisting of a three-stranded β-sheet and five helices and central ankyrin repeats (ANK1-4) (Figure 15B). This BTB domain is typically found in protein that contain only one or two other types of domain, such as BTB-zinc

finger (BTB-ZF), BTB-BACK-kelch (BBK), and BTB-BACK-PHR (BBP) families of proteins (90). Although the C-terminus of *At*NPR3 is truncated and the SA binding site of *At*NPR3 is invisible here. Similar to NPR1, we clearly observed a non-canonical fourth ANK identified in the crystal structure, which is distinct from the other three conventional ankyrin repeats (ANK1-3) (Figure 15). The two helices in the fourth ANK are connected through an extended loop of around 11 residues, and the second helix is dis-continued in the middle and is consisted of two smaller helices (Figure 16A).

Similar to NPR1, the Akyrin repeat 4 (ANK4) appears to form extensive interactions with the BTB domain in NPR3, specially surrounding the Zinc-binding region (Figure 16 B,C&D). Overall, the structure of NPR3 revealed that it appears to have two major lobes: the N-terminus containing the BTB_ZF domain and the BHB, and the C-terminus containing the 4 ankyrin repeats.

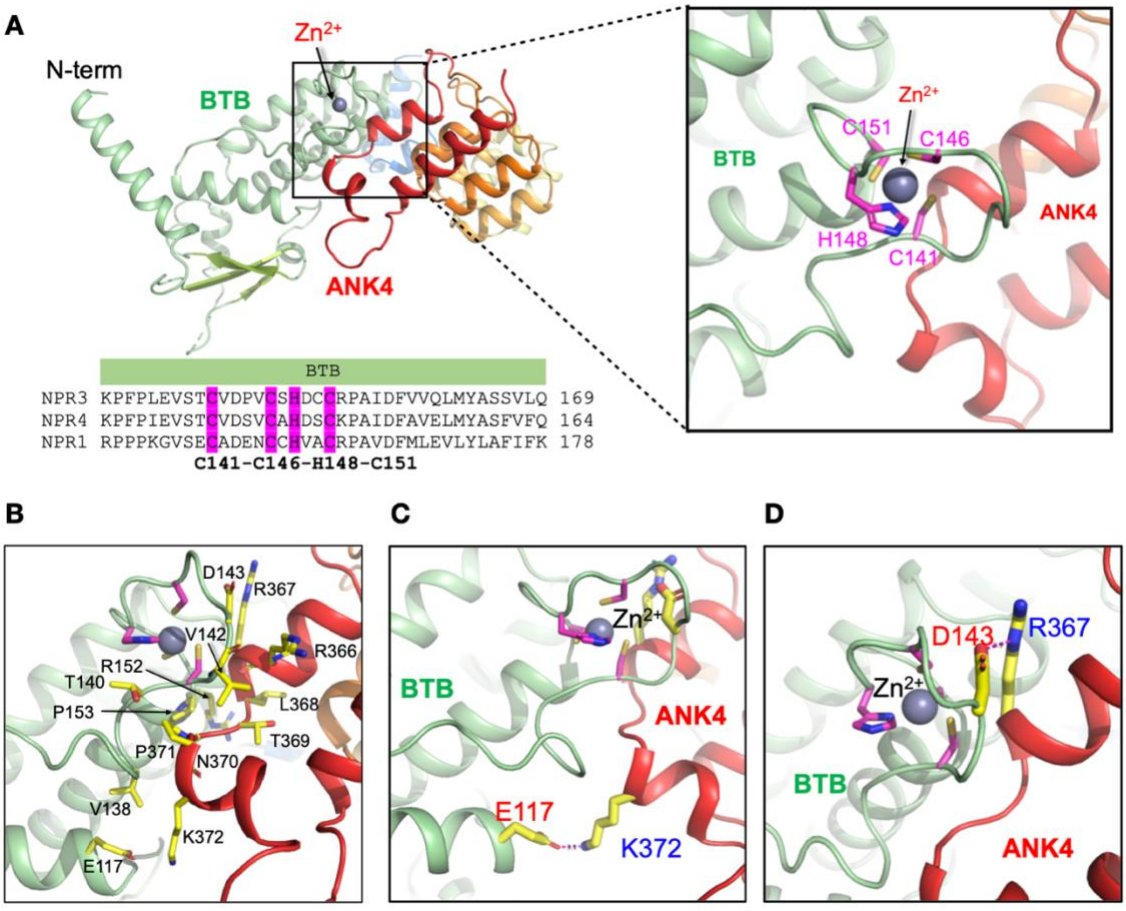


Figure 16 Close-up views of NPR3 zinc binding motif and inter-domain interactions.

(**A**) Zinc binding motif. Zinc ion is held by Cys141, Cys146, His148, and Cys151, which are highly conserved among NPR proteins. (**B**) The non-canonical ANK4 forms intensive interactions with the BTB_ZF domain (**C**) Key inter-domain interactions between ANK4 and the BTB domain.

3.3 Comparison of the crystal structures of AtNPR1 and AtNPR1

Structural comparison of the crystal structures of *At*NPR3 and *At*NPR1 suggests high resemblance as Figure 17 shows. Both *At*NPR3 and *At*NPR1 form bird-shaped homodimers bridged by N-terminal helices (Figure 17A). The zinc finger motif is present in both *At*NPR3 and *At*NPR1 monomers, highly conserved and stabilized by extensive

interactions between the BTB domain and the ANK4. Noted that in these crystal structures, the SA-binding domain (SBD) is absent. As Ding et al proposed that NPR1 and NPR3 might act in an opposite role (20), one as a transcriptional activator and another as a repressor, it remains to be established how the structural difference in NPR1 and NPR3 leads to their different functional roles. Possibilities are that their functional difference in regulating plant immune responses may be due to their distinct SA-binding domains or interactions with different binding partners.

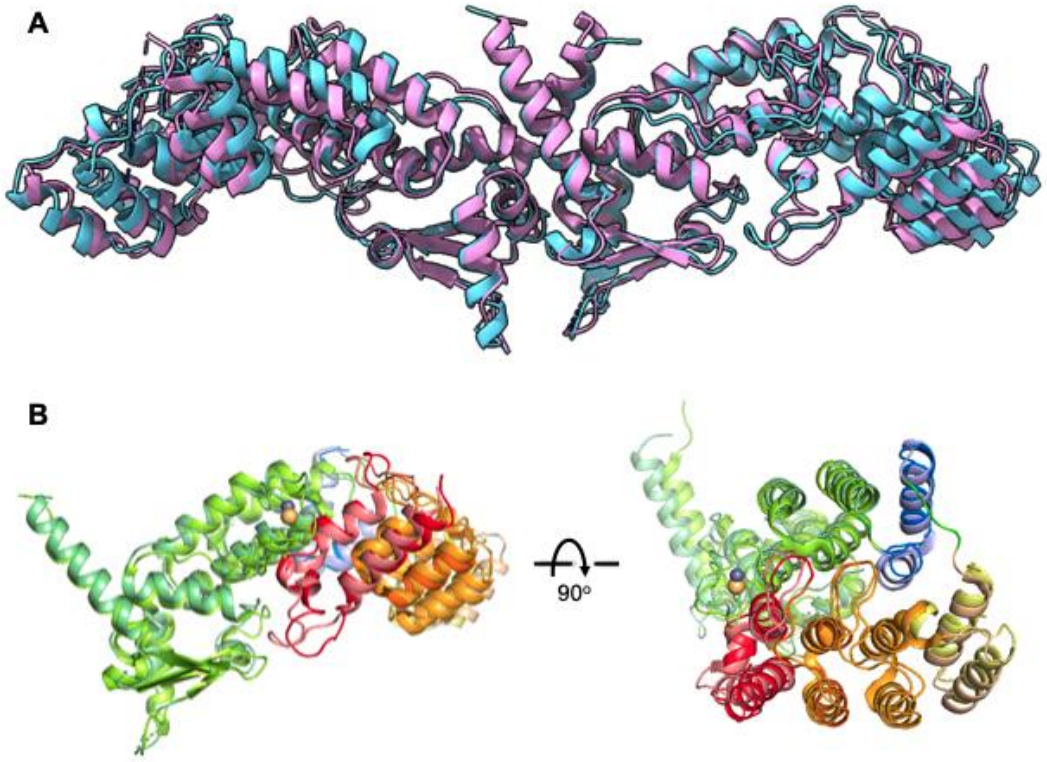


Figure 17 Superimposition of crystal structures of *At*NPR1 and *At*NPR3.

(A) Overlay of NPR1 homodimer (PDB: 7MK3) and NPR3 homodimer. Dimeric NPR1 is colored in blue, and dimeric NPR3 is colored in purple. (**B**) Overlay of NPR1 and NPR3

monomer from crystal structures. The BTB_ZF domain is colored in green (NPR1 in light green, NPR3 in dark green), the BHB is colored in blue (NPR3) and in purple (NPR1), and the ANKs are colored in dark orange or red (NPR3) and in light orange or red (NPR1).

3.4 Methods

3.4.1 Cloning and purification of NPR1(ΔSBD)

The *Arabidopsis* NPR1(ΔSBD) (Thr39–Asp410) gene was cloned into the pTrcHis vector (Thermo Fisher Scientific) with a His_{10}–MBP fusion protein containing a PreScission protease site following MBP and a C-terminal tandem Strep3 tag. Transformed BL21 (DE3) *E. coli* cells were grown in Luria-Bertani (LB) medium at 37 °C to an optical density between 0.4 - 0.6 at 600nm (OD_{600}), induced with 0.1mM IPTG at 18°C overnight. Cell pellet were harvested the second day, resuspended in the lysis buffer (25 mM HEPES pH 8.0 and 300 mM NaCl) with cOmplete protease inhibitor tablet (Sigma Aldrich) and DNase (Sigma Aldrich), and lysed by French press. After centrifugation at 20,000 rpm at 4 °C for 30 min, the supernatant was incubated with pre-equilibrated HisPur™ Ni-NTA resin (Thermo Fisher Scientific) for 20 min at 4 °C. After collecting the flow through, a gradient wash using lysing buffer with imidazole concentration ranging from 0mM to 20mM was conducted, followed by eluting protein with Elution buffer (25 mM pH 8.0, 300 mM NaCl, and 250 mM imidazole). Then PreScission protease was added for overnight cleavage in a Slide-A-Lyser dialysis cassette (Thermo Fisher Scientific) against lysis buffer to remove imidazole. After the removal of His_{10}–MBP tag by protease cleavage, the TGA3 NID protein was further

passed through the new batch of Ni-NTA resin and eluted from the column in the presence of 20 - 30 mM imidazole in the lysis buffer, whereas the His_{10}–MBP tag was retained on the Talon® cobalt resin (Takara Bio).

The NPR1(ΔSBD) was concentrated and further purified by size-exclusion chromatography (SEC) using the Superose 6 increase 10/300 column (Cytiva) pre-equilibrated with the buffer (20 mM HEPES pH 7.5, 150 mM NaCl, and 1 mM DTT). The SEC peak fractions were combined and concentrated up to 6 mg/ml for crystal set-up.

3.4.2 X-ray crystallography of NPR1(ΔSBD)

Protein crystals were grown using the sitting-drop vapor diffusion method at 20 °C. Each drop was prepared by mixing 1 μl of the protein solution at 6 mg ml^{-1} with 1 μl of the reservoir solution containing 0.1 M HEPES (pH 7.5), 19% PEG 3350, 0.25 M tri-sodium citrate. Diffraction-quality protein crystals were collected after 1 week and cryoprotected with the reservoir solution containing 25% glycerol.

X-ray fluorescence and diffraction datasets of NPR1(ΔSBD) were collected at the Northeastern Collaborative Access Team (NECAT) 24-ID-C beamline at the Advanced Photon Source at the Argonne National Laboratory. The X-ray diffraction data were processed using XDS (103).The phase information was obtained by molecular replacement with the PHASER module in the PHENIX suite (104) using the crystal coordinates of apo NPR3. Iterative model building and refinement were performed using COOT (105)and PHENIX (104). For NPR1(ΔSBD), the final model has 95.13%

residues in the favored region, 4.87% in the allowed region and 0.0% in the outlier region of the Ramachandran plot.

3.4.3 X-ray crystallography of NPR3

The *Arabidopsis* NPR3 (S34D406) gene was cloned into the pTrcHis vector (Thermo Fisher Scientific) with a His_{10}–GST fusion protein containing a PreScission protease site following MBP and a C-terminal tandem Flag tag. The expression of NPR3(S34D406) was performed using similar procedures to those described above for *E. coli*-expressed NPR1, and the *At*NPR3 (S34D406) was purified after GST, reverse GST, and reverse Ni-NTA affinity columns, followed by size exclusion chromatography containing SEC buffer (20mM HEPES pH 7.5, and 150 mM NaCl, 1mM DTT).

Protein crystals were grown using the sitting-drop vapor diffusion method at 20 °C. Each drop was prepared by mixing 1 μl of the protein solution at 6.6 mg ml^{-1} with 1 μl of the reservoir solution containing 0.1 M Bis-tris (pH 6.0), 18% PEG 3350, 0.2 M tri-sodium citrate. Diffraction-quality protein crystals were collected after 1 week and cryoprotected with the reservoir solution containing 25% glycerol.

X-ray fluorescence and diffraction datasets of NPR3 (S34D406)) were collected at the Northeastern Collaborative Access Team (NECAT) 24-ID-C beamline at the Advanced Photon Source at the Argonne National Laboratory. The X-ray diffraction data were processed using XDS (103).The phase information was obtained by molecular

replacement with the PHASER module in the PHENIX suite (104). Iterative model building and refinement were performed using COOT (105)and PHENIX (104).

Chapter 4. Structural elucidation of the NPR1-interacting domain of TGA3 reveals an unexpected fatty acid substrate

In early 2000s, another family of proteins, TGACG-binding (TGA) transcription factors, have raised wide attention as it was reported to play an important role in regulating defense-related genes through interacting with NPR1 directly (25, 106). Discovered in 1989, the TGA transcription factors were found to conservatively recognize and bind to DNAs containing the sequence of TGACG and they are among the first transcription factors that have ever studied in plants (107). Besides interacting with NPR1 to regulate plant immune defense, TGAs are also known to be essential regulators of various other cellular processes and involved in different hormonal pathways.

The TGA transcription factors are members of the basic leucine zipper (bZIP) superfamily, one of the most ancient and best characterized eukaryotic transcription factor families (108). As a member of bZIP transcriptional regulators, its genomic targets include the pathogenesis-related (PR), genes that are associated with plant defense (106). However, the role of individual TGA factors in the transcriptional control of PR defense genes has remained elusive as the expression of *PR* genes in *Arabidopsis* leaves is stimulated by SA through a stimulus pathway involving NPR1 (106). It was shown that the SA-induced expression of *PR-1* in Arabidopsis is mediated by *as-1*-type promoter *cis* elements (*as-1-like*) that is recognized by TGA factors, and such TGA - *as-1-like*

association is significantly enhanced by the presence of NPR1(109). Furthermore, the NPR1-TGA interaction seems to occur predominantly within the nucleus where NPR1 activates TGA factors by increasing their DNA binding affinity as evidences show improved TGA2-*as-1-like* complex formation in the presence of wild-type NPR1 (109).

There are two *as-1-lik* cis elements found within the PR1 promoter, a positive regulating element, *LS7*, and a negative regulating element, *LS5* (110). TGA3 was found to principally contribute to the *LS7* binding activity.

In terms of its structural topology, TGAs contains a highly conserved bZIP domain, which consists of an N-terminal basic region, responsible for DNA binding, and a leucine zipper, which mediates dimerization between bZIP domains (111). The bZIP domain determines DNA-binding specificity and further serves as a nuclear localization signal (NLS). Upon binding to target DNA sequences, TGAs forms either homo- or heterodimers for regulation, granting them broad variability in regulation of physiological responses (111). Surprisingly, though members of TGAs share the core DNA-binding sequence, the TGA paralogues exert various binding preferences (111).

To date, TGAs have been extensive studied using genetic tools, and there is limited knowledge on the relationship between its structure and function due to its complex involvement in various cellular processes and a lack of research supporting the proposed mechanisms. Since the discovery of TGAs 30 years ago, their three-dimensional (3D) structures have remained a mystery. Thus, understanding their

mechanism of actions at the molecular level, specifically, the functional role of the interaction between NPR1 and TGAs, will provide unparalleled insights into the function of TGAs in regulating plant immune responses.

As the recent breakthrough in computational modeling and prediction has led to the release of the AF2, we used AF2 for TGAs modeling. As shown in Figure 18, the AF2.0-generated predication model indicates an intrinsically disordered N-terminus, a bZIP domain, and a folded C-terminus, which are consistent with prior biochemical and bioinformatic analyses (Figure 18).

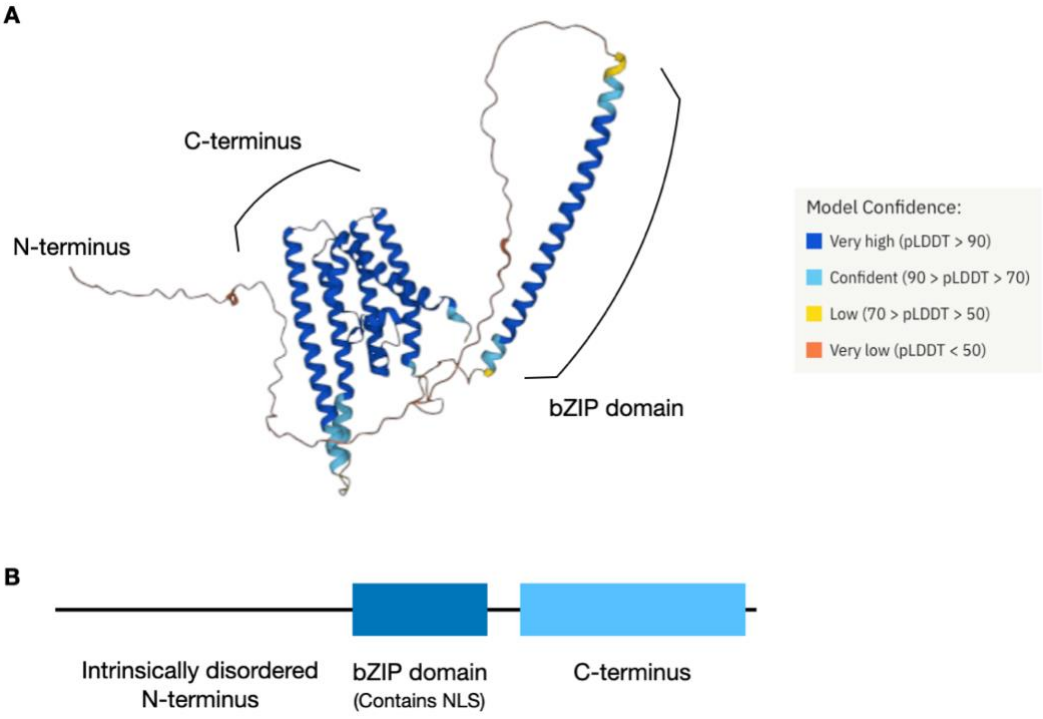


Figure 18 AlphaFold model of AtNPR3 and a schematic representation of TGA family of transcription factors.

(**A**) The AlphaFold generated 3D model of AtTGA3 (pLDDT, AlphaFold per-residue confidence score) and (**B**) A schematic representation of TGA domain organization, showing the intrinsically disordered N-terminus, the bZIP domain and the C-terminus. The nuclear localization signal (NLS) is included in the bZIP domain.

4.1 High resolution crystal structure of TGA3 NID reveals a dimerization state and a lipid substrate

In addition to NPR1, two of NPR1 paralogs in Arabidopsis, NPR3 and NPR4, have also been reported to interact with a number of TGA transcription factors including TGA3 in yeast two-hybrid assays (112). TGA3 is a transcriptional activator that binds specifically to the DNA sequence 5′-TGACG-3′, and *as-1*-like cis elements to mediate auxin- and SA-inducible transcription. Prior genetic studies has shown that among six

TGA family members, TGA3 has a particularly strong affinity for NPR1 (25). This observation led to our structural investigation of the interaction between TGA3 and NPR protein.

Assisted with AF2-predicted model and cryo-EM density map of $TGA3_2$-$NPR1_2$-$TGA3_2$ by Dr. Qinglin Wu, we designed a truncated NPR1-interacting domain (NID) of TGA3 (Ala160-Thr384) for X-ray crystallization. High diffraction quality crystals were obtained through sitting-drop vapor diffusion (Figure 19, Table 2).

Surprisingly, TGA3 NID forms dimers in the crystal (Figure 20 A&C). The TGA3 NID contains five prominent long helixes (α1, α2, α6, α7, and α8) and three short helices (α3, α4, and α5), forming an opening helix-bundle with one helix extruding out (α8) towards its dimer counterpart (Figure 20A). Interestingly, the three short helices (α3, α4, and α5) zigzag across the small helix bundle supporting the C-terminal helix (α8) and pushing it out. The NID dimer is formed through extensive interactions between the three short helixes (α3, α4, and α5) and between the tilted α8 helix in one monomer and the α6-α7 helices in the other monomer (Figure 20 C). Prior to our study, no one has reported that the NID of TGAs could form a dimer. Certainly, the AF2 model does not indicate that possibility.

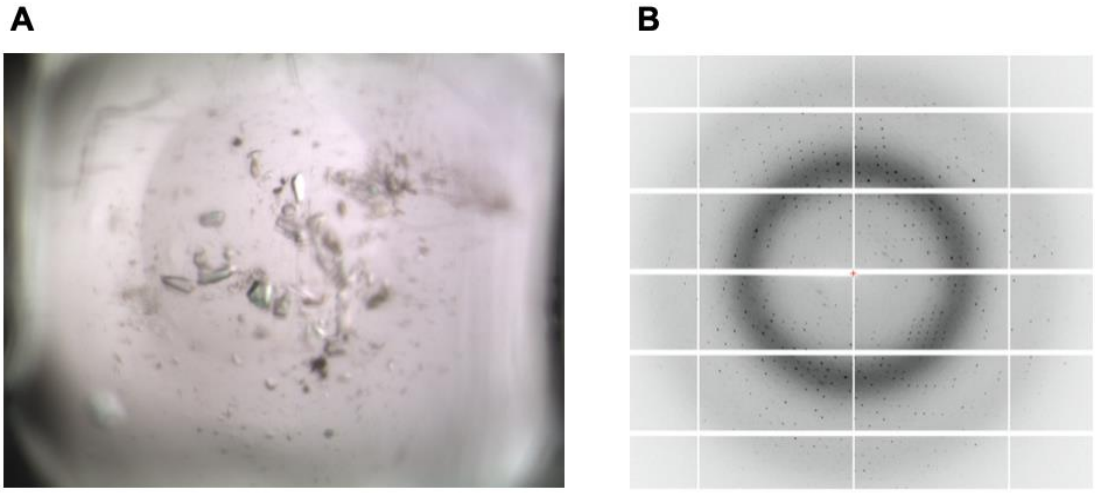


Figure 19 TGA3 NID protein crystals and the crystal diffraction pattern.

(**A**) TGA3 NID crystals. (**B**) The crystal diffraction pattern of a well-diffracted TGA3 NID protein crystal.

Table 2 X-ray data collection and refinement statistics of *Arabidopsis* TGA3 NID

	TGA3 NID (PDB 7TAE)
Data collection	
Space group	C 2 2 21
Cell dimensions	
a, *b*, *c* (Å)	68.39, 101.74, 83.95
α, β, γ (°)	90, 90, 90
Resolution (Å)	43.51 – 1.50 (1.55 – 1.50)
R_{merge}	0.04796 (1.002)
R_{pim}	0.0137 (0.2832)
$CC_{1/2}$	0.999 (0.947)
$I / \sigma I$	25.51 (3.19)
Completeness (%)	99.83 (99.33)
Redundancy	13.2 (13.1)
Refinement	
Resolution (Å)	1.50
No. reflections	622,057 (60,581)
R_{work} / R_{free}	0.1875 / 0.2089
No. atoms	1959
Protein	1757
Ligand/ion	30
Water	172
B-factors	34.72
Protein	33.49
Ligand/ion	44.48
Water	45.58
R.m.s. deviations	
Bond lengths (Å)	0.006
Bond angles (°)	0.90

*Values in parentheses are for highest-resolution shell.

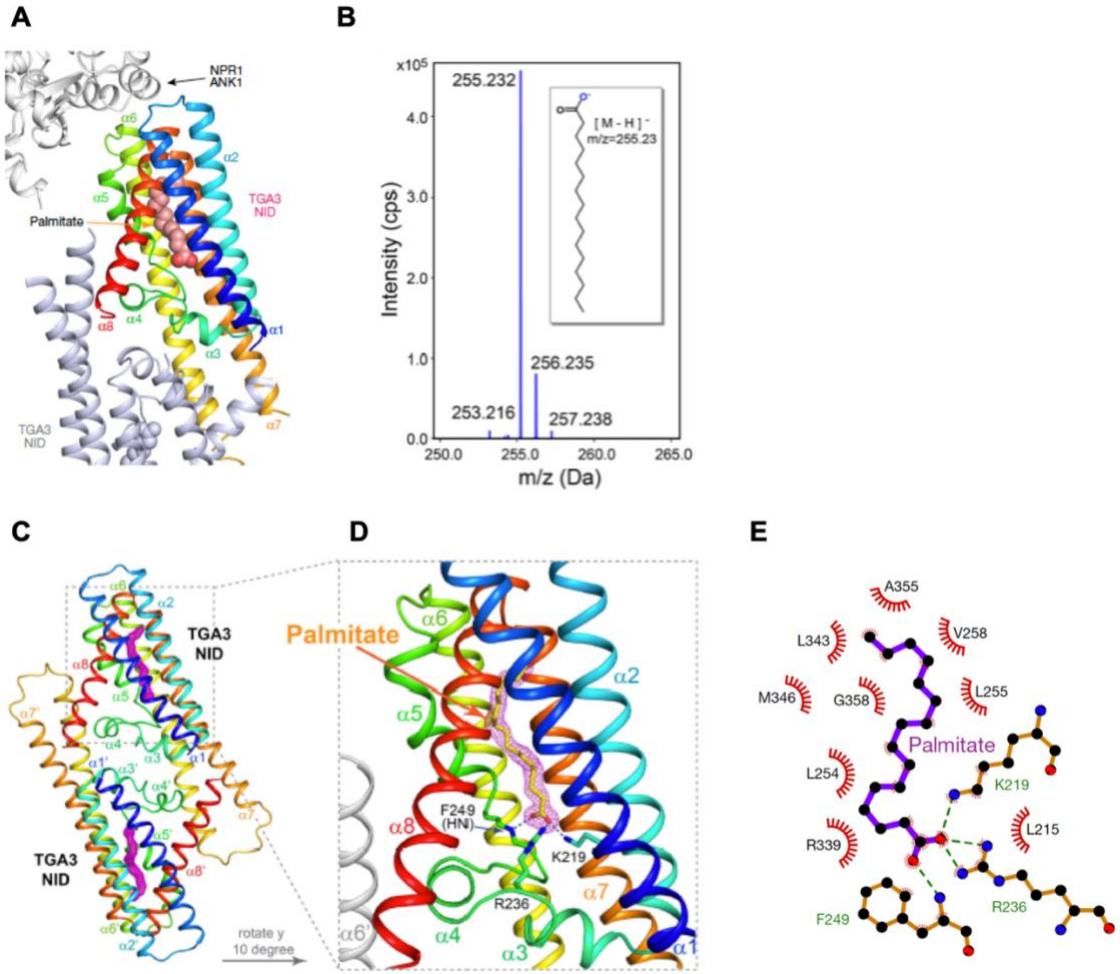


Figure 20 The palmitate containing TGA3 NID dimer and MS characterization of the palmitate ligand.

(**A**) The palmitate containing TGA3 NID dimer (colored rainbow and grey). (**B**) Mass spectrometry analysis of the fatty acid extracted from the protein sample, verifying the fatty acid as the palmitic acid (C16:0). (**C**) Crystal structure of the TGA3 NID dimer. The two TGA3-NID molecules are shown in the rainbow color, with N-terminus colored in blue and C-terminus colored in red. (**D**) A zoomed-in view of the location of the palmitate in monomer TGA3 NID. Polar interactions with the carboxylate group of the palmitates are indicated with dashed lines. For clarity, only one TGA3 molecules is colored in rainbow. Purple meshes in panels A and B represent 2mFo-DFc omit map of the palmitate plotted at the 1.0 σ level. (**E**) Palmitate recognition by the TGA3 NID in the 1.5-Å-resolution crystal structure of the TGA NID-palmitate complex. Polar interactions are denoted (dashed lines) and van der Waals contacts are shown (curved lines).

Moreover, the presence of homo-dimeric state of TGA3 in crystals assists in resolving the ambiguity in the cryo-EM density map of $TGA3_2$-$NPR1_2$-$TGA3_2$ complex and building the working model (Figure 21). With the atomic structural information provided by means of X-ray crystallography and single-particle cryo-EM analysis, we eventually finalize a working model, in which NPR1 forms a dimer with each ANK domain flanked by a TGA3 dimer (Figure 21A). Extensive interactions are observed between TGA3 NID and NPR1 (Figure 21B). Surface-exposed hydrophobic residues from Ankyrin repeat 1 (ANK1) of NPR1, such as Leu281 and Leu284, project onto a hydrophobic concave surface of one TGA3 subunit formed by the α5-α6 loop (Pro264 and Thr266) and α7-α8 loop (Thr351), whereas numerous hydrogen bonds and salt bridges are observed between ANK1 residues (Lys272, Ser276, Asp277, Asp278, Glu280, Glu288 and His290) and TGA3 NID residues from the α5-α6 loop (Glu263), α6 helix (Asp267), α8 helix (Thr352, Arg353, and Arg357). Based on these observations, mutating interfacial residues in NPR1 such as L281D and L284D was conducted and results in the abolishment of the NPR1-TGA3 interaction and SA-induced *PR1* gene expression (credits of Dr. Raul Zavaliev and Dr. Qinglin Wu).

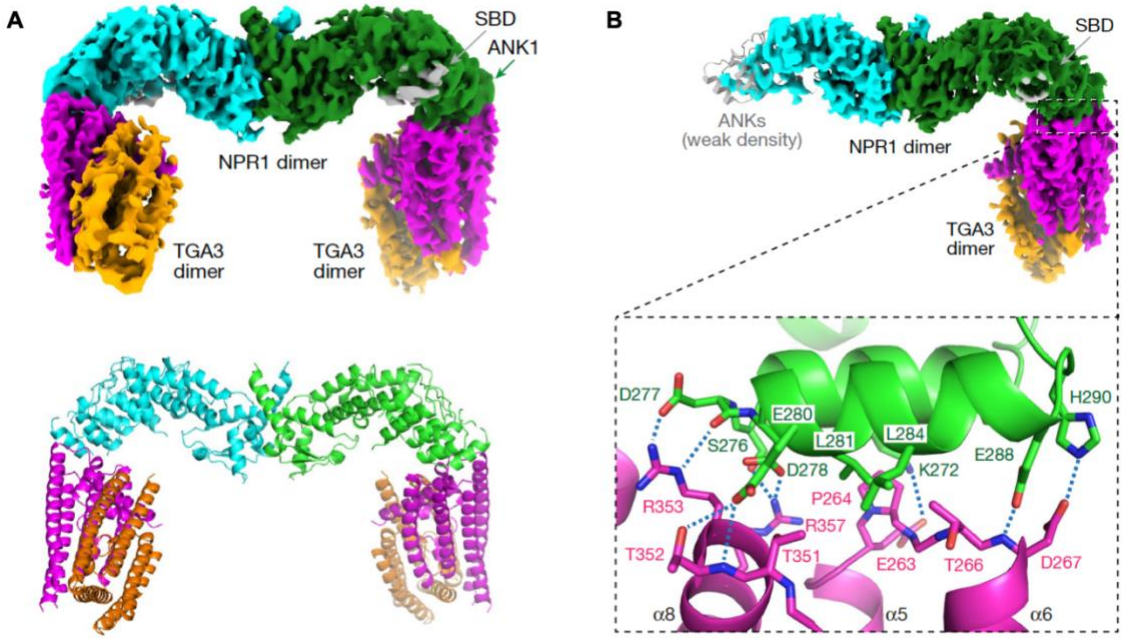


Figure 21 The cryo-EM structure of the NPR1-TGA3 complex.

(A) Cryo-EM density map and carton representation of the TGA_2-$NPR1_2$-$TGA3_2$ complex. (**B**) Cryo-EM density map of the $NPR1_2$-$TGA3_2$ assembly intermediate. The ANK region (grey) with weak density is indicated. The Zoom-in view shows the molecular interactions between the NPR1 ANK1 (green) and the TGA3 NID (purple). Polar interactions are indicated (dashed lines).

The disclosure of the structure of the $TGA3_2$-$NPR1_2$-$TGA3_2$ complex, with the dimeric NPR1 engaging two TGA3 dimers, leads us to further investigate its binding with DNA substrate, the *as-1* elements. We conducted the electrophoresis mobility shift assay (EMSA) using the PR1 gene promoter containing two *as-1* elements (LS5 and LS7), which are both reported to be required for NPR1-mediated transcription in a reporter assay (113) (Credits to Dr. Qinglin Wu). As figure 22A presents, NPR1 itself doesn't bind to DNA, but the transcription factor TGA3 binds to the *as-1* elements. In the presence of TGA3 and either LS5 or LS7, one or two upshifted bands of the promoter DNA

appeared. Upon addition of NPR1, it completely super-shifted the TGA3 double-occupancy band and depleted the single-occupancy band (Figure 22B). Moreover, mutating either of the two *as-1* elements in the DNA blocked the NPR1-mediated super-shifted band (Figure 22A).

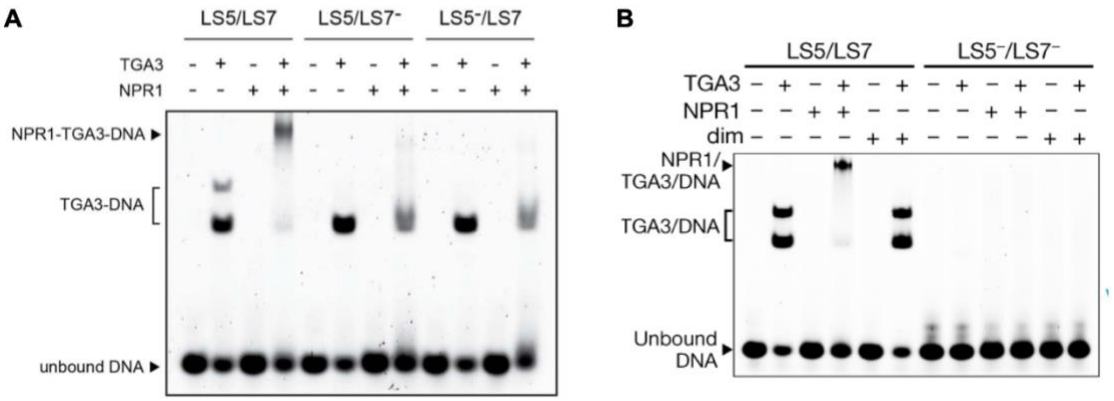


Figure 22 Electrophoresis mobility shift assay for TGA3 in complex with NPR1 or *npr1(dim)* with *as-1* elements.

(**A**) EMSA of NPR1 and TGA3 using 6-fluorescein-labelled DNA spanning the LS5-to-LS7 region of the *PR1* promoter containing two *as-1* elements (LS5/LS7) or a single *as-1* element in the LS5 region (LS5/LS7^{-}) or in the LS7 region (LS5^{-}/LS7). (**B**) The EMSA for TGA3 in complex with NPR1 or *npr1(dim)*. The *PR* promoter region with (LS5/LS7) or without (LS5^{--}/LS7^{--}) two *as-1* elements was used as the probe.

Besides the discovery of TGA3 NID forming a dimer, unexpectedly, a palmitic acid was found inside the TGA3 NID monomer, an observation confirmed by mass spectrometry (MS) analysis (Figure 20A&B). Within the TGA3 NID monomer, the palmitate is closely surrounded by two long helices (α2 and α7) and one short helix (α5). The extra space created by α8 helix extruding out and short helices (α3, α4, and α5) forming zigzag perfectly allows the 16:0 palmitate to perfectly fit in. In the TGA NID, the

palmitate has extensive Van der Waals contacts with the hydrophobic residues such as leucine and phenylalanine on the adjacent helices. Specifically, F249 and R236 from TGA3 NID form polar interactions with the carboxyl group and the hydroxyl group in the palmitate (Figure 20E). All those interactions protect the fatty acid from exposing to the outer hydrophilic environment.

This discovery of a fatty acid within TGA3 crystal opens a new venue for us to investigate the functional relation between palmitate and TGA transcriptional factors in plant defense. In fact, early in 2002, Maldonado et al. reported their discovery of a putative lipid transfer protein DIR1 involved in SAR in *Arabidopsis*. They found that defective in induced resistance 1-1 (*dir1-1*) exhibits wild-type local resistance to avirulent and virulent *Pseudomonas syringae,* but that PR gene expression is abolished in uninoculated distant leaves and *dir1-1* fails to develop SAR to virulent *Pseudomonas* or *Peronospora parasitica* (114). In their proposed model, DIR1 interacts with a lipid-derived molecule to promote long distance signaling. Since then, fatty acids have been suggested to link with plant immunity and were known for modulating signal transduction pathways by interacting with various transcription factors (115). This serendipity not only provides direct evidence of fatty-acid involvement in SAR but implicates the lipid metabolism in regulating plant immunity.

Following our structural study of TGA3, Shi et al. published their crystal structure of transcription factor TGA7 from *Arabidopsis* at 2.06 Å (116). Similarly, they

observed that TGA7 forms dimers in the crystals and found one molecule of palmitate in each TGA7 monomer, which supports our first report of TGA3 crystal structure forming a dimer with a palmitate substrate inside. The full length *At*TGA3 and *At*TGA7 show >70% identity in the sequence alignment (Figure 23A), and similarly the structural alignment presents high-level identity with only half of the helix(α6)-loop-helix(α7) diverging (Figure 23B). The corresponding sequences of the divergent loop region is S301 to A327 in *At*TGA3 or S286 to 311 in *At*TGA7, which indeed shares conservation in terms of sequence. The work on TGA7 by Shi et al. not only validates our new discovery that TGAs forms dimer but indicates that palmitate might act as an omnipresent ligand of TGA transcription factors. Although fatty acids have been reported to act as signaling molecules involved in SAR, the specific function of palmitate as a ligand of TGAs needs to be further investigated and established.

The work in this chapter has been published and the description is adapted from the published manuscript (100).

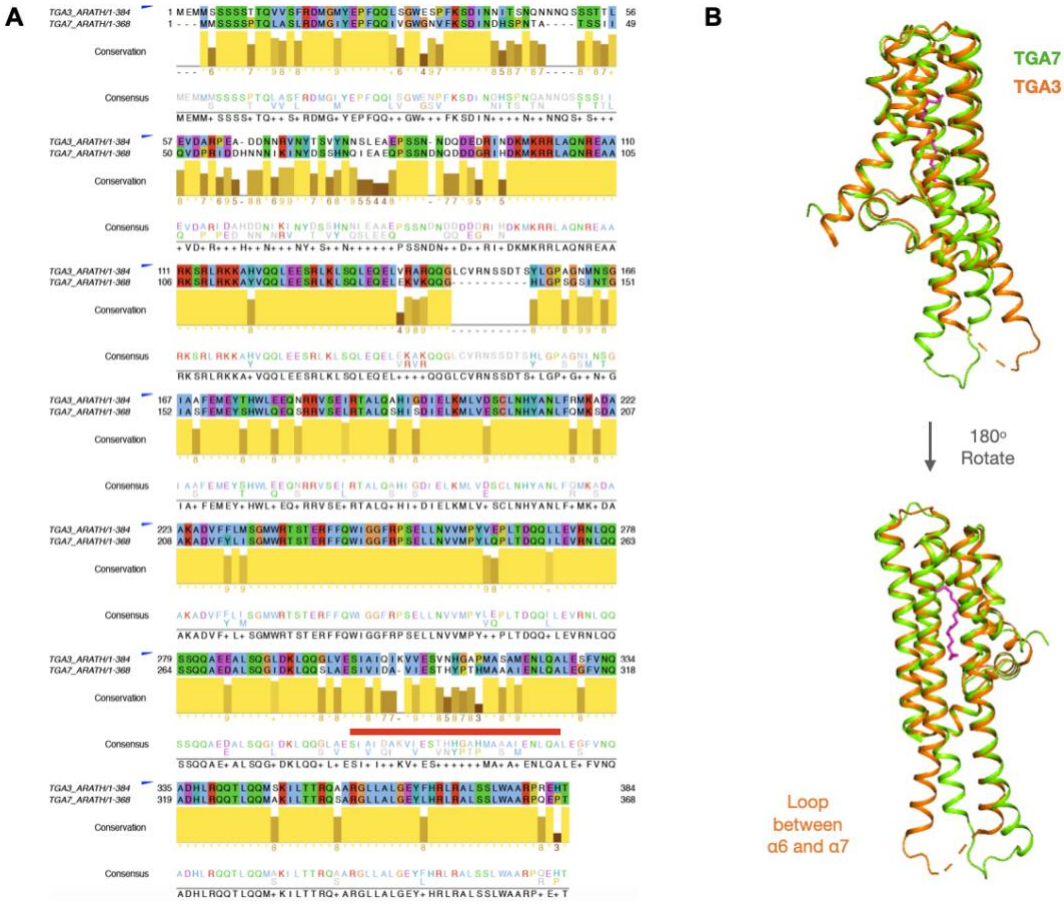


Figure 23 Sequence alignment of TGA3 and TGA3 from Arabidopsis and crystal structure comparison.

(**A**) sequence alignment of full length AtTGA3 and AtTGA7. The sequence region underscored in red indicates the flexible loop region between α6 and α7. (**B**) Structural alignment of crystal structures of TGA3 and TGA7 with a palmitate inside. The crystal structure of TGA7 is colored in green and the crystal structure of TGA3 is colored in orange. The alignment indicates of a divergence in the helix-loop-helix motif between helices α6 and α7.

4.2 Methods

4.2.1 TGA3_NID cloning and protein purification

The *Arabidopsis* TGA3 NID (Ala160–Thr384) gene was cloned into the pTrcHis vector (Thermo Fisher Scientific) with a His_{10}–MBP fusion protein containing a

PreScission protease site following MBP and a C-terminal tandem Flag–HA tag. Transformed BL21 (DE3) *E. coli* cells were grown in Luria-Bertani (LB) medium at 37 °C to an optical density between 0.4 - 0.6 at 600nm (OD_{600}), induced with 0.1mM IPTG at 18 °C overnight. Cell pellet were harvested the second day, resuspended in the lysis buffer (25mM HEPES pH8.0 and 300 mM NaCl) with cOmplete protease inhibitor tablet (Sigma Aldrich) and DNase (Sigma Aldrich), and lysed by French press. After centrifugation at 20,000 rpm at 4 °C for 30min, the supernatant was incubated with pre-equilibrated HisPur™ Ni-NTA resin (Thermo Fisher Scientific) for 20min at 4 °C. After collecting the flow through, a gradient wash using lysing buffer with imidazole concentration ranging from 0 mM to 20 mM was conducted, followed by eluting protein with Elution buffer (25 mM pH8.0, 300 mM NaCl, and 250 mM imidazole). Then PreScission protease was added for overnight cleavage in a Slide-A-Lyser dialysis cassette (Thermo Fisher Scientific) against lysis buffer to remove imidazole. After the removal of His_{10}–MBP tag by protease cleavage, the TGA3 NID protein was further passed through the new batch of Ni-NTA resin and eluted from the column in the presence of 10-20 mM imidazole in the lysis buffer, whereas the His_{10}–MBP tag was retained on the Talon cobalt resin (Takara Bio).

The TGA3_NID_FLAG_HA was concentrated and further purified by size-exclusion chromatography (SEC) using the Superose 6 increase 10/300 column (Cytiva) pre-equilibrated with the buffer (20 mM HEPES pH7.5, 150 mM NaCl, and 1 mM DTT).

The SEC peak fractions were combined and concentrated up to 9.6 mg/ml for crystal set-up.

4.2.2 X-ray Crystallography of TGA3_NID

The TGA3_NID crystals were obtained via sitting-drop vapor diffusion at 20 ^{o}C by mixing 9.6mg/ml purified TGA3_NID in 20 mM HEPES (pH7.5), 150 mM NaCl, 1 mM DTT with the mother liquor containing 0.1 M sodium acetate pH 5.0, and 20% (v/v) MPD at a 1:1 drop ratio. This crystal forming conditions were first screened using commercial JCSG general screening kits (Hampton Research). The high-quality diffracting crystal were obtained from JCSG I G2 condition after 1 week and harvested with cryo-protected solution containing 0.1 M sodium acetate pH 5.0, 20% (v/v) MPD, and 25% glycerol.

X-ray diffraction data were collected at the Northeastern Collaborative Access Team (NECAT) 24-ID-C beamline at the Advanced Photon Source at the Argonne National Laboratory. The X-ray diffraction data were processed using XDS (103). The final coordinates were built by Iterative model building using COOT(105) and refined using PHENIX (104). For the TGA3 NID, the final model has 99.01% residues in the favored region, 0.99% residues in the allowed region and 0.0% in the outlier region of the Ramachandran plot. The coordinates were deposited to the Protein Data Bank (PDB) with accession number of 7TAE.

4.2.3 Lipid extraction from the TGA3 NID and LC-ESI/MS analysis

The fatty acid was extracted from the purified TGA3 NID protein using the two-phase Bligh-Dyer method(117). The extract was analyzed using normal-phase liquid chromatography-electrospray ionization/MS (LC-ESI/MS) in negative mode using an Agilent 1200 Quaternary LC system coupled to a high-resolution TripleTOF5600 mass spectrometer (Sciex) as previously described (118). Data acquisition and analysis were performed using the Analyst TF1.5 software (Sciex)

The lipid extraction and LC-ESI/MS analysis were done by Dr. Jinshi Zhao, and the description is adapted from the published manuscript (100).

4.2.4 Electrophoresis mobility shift assay

Purified TGA3 (Asn87-Thr384) at a concentration of 0.6 μM and WT NPR1 or the npr1(dim) mutant at a concentration of 8 μM were used in the EMSA with the 6-fluorescein (6-FAM)-labelled oligonucleotide probes (200 nM) containing the LS5 and LS7 *as-1* cis elements of the SA-responsive *PR1* gene promoter (LS5/LS7, 6-FAM-C-gggCTATGACGTAAGTAAAATAGTGACGTAGAGAggg) or DNA sequences with the two *as-1* regions replaced with adenosines (LS5^{-}/LS7^{-}, 6-FAM-C-gggCTAAAAAAAAGTAAAATAGAAAAAAAAAGAggg), with the LS7 *as-1* element replaced with adenosines (LS5/LS7^{-}, 6-FAM-C-gggCTATGACGTAAGTAAAATAGAAAAAAAAAGAggg) or with the LS5 *as-1* element replaced with adenosines (LS5^{-}/LS7, 6-FAM-C-

gggCTAAAAAAAAAGTAAAATAGTGACGTAGAGAggg) as well as their complementary oligonucleotides (Integrated DNA Technologies). The uppercase bases indicate the LS5/LS7 sequence; the lowercase bases indicate overhangs outside the LS5/LS7 sequence; and bases replaced with adenosines are indicated with As in the small capital font. Pairs of complementary DNA strands were annealed before the binding reaction. For each binding reaction, the desired protein combination was mixed with the target DNA probe together with 30 ng μl^{-1} of poly(dI-dC) (non-specific competitor, Thermo Fisher Scientific) in a binding buffer containing 20 mM HEPES (pH 8.0), 300 mM NaCl, 10% glycerol, 1 mM DTT and 0.2 mM SA. The mixture was incubated at 4 ^{o}C for 30min and then run on a 3.8% native polyacrylamide gel in the TB buffer (100 mM Tris-borate pH 8.3, 10% glycerol). After electrophoresis, the gel was scanned with the Typhoon FLA 7000 imager (GE Healthcare)

The EMSA assay was done by Dr. Qinglin Wu, and the description is adapted from the published manuscript (100).

Chapter 5. Conclusion and future directions

Structural biochemistry has evolved exponentially due to the technical advances in tools, specifically, the single-particle cryo-EM analysis. So far, cryo-EM has resolved over four thousand structures at near-atomic resolution (2-4 Å) and become a mainstream structural biology tool complementing X-ray crystallography. By means of these two methodologies, we provide structural information important for understanding bacterial pathogenesis and host plant immune response.

In the first DspE project, we performed structural analysis and biochemical functional studies of the T3E DspE. The AF-predicted model hinted that the N-terminal half might perform regulatory or signaling purpose as it contains several functional motifs known for signal transduction and transcription regulation, whereas the C-terminal β-barrel structure, resembling bacterial porins, might function as a pore. Our biochemical functional assays further suggested the C-terminal β-barrel is essential for the pore activity and contributes to the water-soaking phenomenon observed during bacterial infection in plants. For years, the exact biochemical function of AvrE/DspE-family of enigmatic bacterial type III effectors have been recalcitrant to mechanistic understanding for more than three decades. Our structural investigation and functional studies first provide insightful information in revealing its mode of action and suggest that DspE is likely to function as a pathogenic water/solute-permeable pore/channel as a virulence strategy.

However, there are certain roadblocks that have hindered us from resolving the structure of DspE by single-particle cryo-EM and performing high-level structural analysis. For instance, the orientation bias in our cryo-EM sample has prevented us from obtaining diverse projections and side-view 2D classes for robust model building and refinement. Our current vision on how DspE is folded in the host cytoplasm and achieves membrane anchoring and penetration is limited by the data and statistics. Though data from our functional assays indicates that the C-terminal β-barrel might function as a pore penetrating into membrane, challenges arise while reconstituting the DspE in liposome for direct visualization via cryo-EM. Possibility is that DspE might need an assistant molecule or a protein partner to complete membrane anchoring or penetration.

Besides revealing the mode of action of DspE in the host cytoplasm, uncovering how DspE is folded and signaled for secretion in the bacteria would also benefit us in understanding bacterial pathogenesis. As a DspE-specific chaperone, DspF, has been reported to be necessary for DspE secretion and possibly for its translocation (70), disclosing the interaction and dynamics between DspE and DspF prior to secreting through the T3SS will be one of our interests in the future. Obtaining the mechanistic insights of their interaction and dynamics will assist us in revealing the mode of action of DspE during bacterial infection.

Moreover, AvrE-T3Es elicit cell death in both host and non-host plants independent of any known plant resistance protein, suggesting its involvement with the plant immune system (57). Prior studies in yeast have indicated that they activate protein phosphatase 2A (PP2A) and inhibit serine palmitoyl transferase, the first enzyme of the sphingolipid biosynthesis pathway (57, 84). Activation of PP2A by AvrE/DspE family T3Es would also reinforce PTI inhibition as PP2A is involved in the negative control of PTI *in planta* (112). Another future direction will be investigating the relationship between DspE with PP2A, and the functional relevance of their interaction in plant host immune response.

In the second story of elucidating key regulators in plant immunity, we successfully revealed the structural information of *At*NPR1, *At*NPR3, and *At*TGA3 at atomic levels. Consistent with sequence alignment, *At*NPR1(ΔSBD) and *At*NPR3(ΔSBD) show high resemblance in crystal structures. Unexpectedly, NPR3 forms dimers in crystals, which has never been reported priorly. Certainly, the functional significance of NPR3 dimer needs to be established. Our current understanding of the opposite function roles of NPR1 and NPR3 are playing are also limited due to the missing part of SA-binding domain in current crystals.

Our structural analysis of TGA3 opens a new venue for our future work as in crystals we found that TGA3 forms a dimer and contains a novel ligand in each monomer. Prior to our study, the dimerization of TGA3 has never been reported.

Importantly, it helps us resolve the cryo-EM density map of $TGA3_2$-$NPR1_2$-$TGA3_2$ complex and build a final model (Figure 21). In addition, the discovery of palmitate as a TGA3 ligand provides direct evidence supporting the involvement of fatty acids in regulating plant immune responses. Future work could be focused on investigating the specific role of palmitate while interacting with TGA3 in signal transduction. All these set foundations for future mechanistic studies to understand the dynamics of protein-protein interactions in regulating plant immune response.

References

1. Melotto M, Panchal S, Roy D. Plant innate immunity against human bacterial pathogens. Front Microbiol. 2014;5:411.

2. Nomura K, Debroy S, Lee YH, Pumplin N, Jones J, He SY. A bacterial virulence protein suppresses host innate immunity to cause plant disease. Science. 2006;313(5784):220-3.

3. McManus PS, Stockwell VO, Sundin GW, Jones AL. Antibiotic use in plant agriculture. Annu Rev Phytopathol. 2002;40:443-65.

4. Melotto M, Kunkel BN. Virulence Strategies of Plant Pathogenic Bacteria. The Prokaryotes2013. p. 61-82.

5. Lindeberg M, Cunnac S, Collmer A. Pseudomonas syringae type III effector repertoires: last words in endless arguments. Trends Microbiol. 2012;20(4):199-208.

6. Block A, Alfano JR. Plant targets for Pseudomonas syringae type III effectors: virulence targets or guarded decoys? Curr Opin Microbiol. 2011;14(1):39-46.

7. Rasko DA, Sperandio V. Anti-virulence strategies to combat bacteria-mediated disease. Nat Rev Drug Discov. 2010;9(2):117-28.

8. Costa TR, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M, et al. Secretion systems in Gram-negative bacteria: structural and mechanistic insights. Nat Rev Microbiol. 2015;13(6):343-59.

9. Green ER, Mecsas J. Bacterial Secretion Systems: An Overview. Microbiol Spectr. 2016;4(1).

10. Galan JE, Wolf-Watz H. Protein delivery into eukaryotic cells by type III secretion machines. Nature. 2006;444(7119):567-73.

11. Angot A, Peeters N, Lechner E, Vailleau F, Baud C, Gentzbittel L, et al. Ralstonia solanacearum requires F-box-like domain-containing type III effectors to promote disease on several host plants. Proc Natl Acad Sci U S A. 2006;103(39):14620-5.

12. Bocsanczy AM, Nissinen RM, Oh CS, Beer SV. HrpN of Erwinia amylovora functions in the translocation of DspA/E into plant cells. Mol Plant Pathol. 2008;9(4):425-34.

13. Buttner D. Behind the lines-actions of bacterial type III effector proteins in plant cells. FEMS Microbiol Rev. 2016;40(6):894-937.

14. Xiang T, Zong N, Zou Y, Wu Y, Zhang J, Xing W, et al. Pseudomonas syringae effector AvrPto blocks innate immunity by targeting receptor kinases. Curr Biol. 2008;18(1):74-80.

15. Koornneef M, Meinke D. The development of Arabidopsis as a model plant. Plant J. 2010;61(6):909-21.

16. Nishimura MT, Dangl JL. Arabidopsis and the plant immune system. Plant J. 2010;61(6):1053-66.

17. Jones JD, Dangl JL. The plant immune system. Nature. 2006;444(7117):323-9.

18. Chisholm ST, Coaker G, Day B, Staskawicz BJ. Host-microbe interactions: shaping the evolution of the plant immune response. Cell. 2006;124(4):803-14.

19. Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant-pathogen interactions. Nat Rev Genet. 2010;11(8):539-48.

20. Ding Y, Sun T, Ao K, Peng Y, Zhang Y, Li X, et al. Opposite Roles of Salicylic Acid Receptors NPR1 and NPR3/NPR4 in Transcriptional Regulation of Plant Immunity. Cell. 2018;173(6):1454-67 e15.

21. Fu ZQ, Yan S, Saleh A, Wang W, Ruble J, Oka N, et al. NPR3 and NPR4 are receptors for the immune signal salicylic acid in plants. Nature. 2012;486(7402):228-32.

22. Cao H, Bowling SA, Gordon AS, Dong X. Characterization of an Arabidopsis Mutant That is Nonresponsive to inducers of systemic acquired resistance. Plant Cell. 1994;6(11):1583-92.

23. Backer R, Naidoo S, van den Berg N. The NONEXPRESSOR OF PATHOGENESIS-RELATED GENES 1 (NPR1) and Related Family: Mechanistic Insights in Plant Disease Resistance. Front Plant Sci. 2019;10:102.

24. Chen J, Mohan R, Zhang Y, Li M, Chen H, Palmer IA, et al. NPR1 Promotes Its Own and Target Gene Expression in Plant Defense by Recruiting CDK8. Plant Physiol. 2019;181(1):289-304.

25. Zhou JM, Trifa Y, Silva H, Pontier D, Lam E, Shah J, et al. NPR1 differentially interacts with members of the TGA/OBF family of transcription factors that bind an element of the PR-1 gene required for induction by salicylic acid. Mol Plant Microbe Interact. 2000;13(2):191-202.

26. Mou Z, Fan W, Dong X. Inducers of plant systemic acquired resistance regulate NPR1 function through redox changes. Cell. 2003;113(7):935-44.

27. Zhou JM, Zhang Y. Plant Immunity: Danger Perception and Signaling. Cell. 2020;181(5):978-89.

28. Renaud JP, Chari A, Ciferri C, Liu WT, Remigy HW, Stark H, et al. Cryo-EM in drug discovery: achievements, limitations and prospects. Nat Rev Drug Discov. 2018;17(7):471-92.

29. Antoni C, Quentin D, Lang AE, Aktories K, Gatsogiannis C, Raunser S. Cryo-EM structure of the fully-loaded asymmetric anthrax lethal toxin in its heptameric pre-pore state. PLoS Pathog. 2020;16(8):e1008530.

30. Rajapaksha M, Eichler JF, Hajduch J, Anderson DE, Kirk KL, Bann JG. Monitoring anthrax toxin receptor dissociation from the protective antigen by NMR. Protein Sci. 2009;18(1):17-23.

31. Vourtsis DJ, Chasapis CT, Pairas G, Bentrop D, Spyroulias GA. NMR conformational properties of an Anthrax Lethal Factor domain studied by multiple amino acid-selective labeling. Biochem Biophys Res Commun. 2014;450(1):335-40.

32. Pannifer AD, Wong TY, Schwarzenbacher R, Renatus M, Petosa C, Bienkowska J, et al. Crystal structure of the anthrax lethal factor. Nature. 2001;414(6860):229-33.

33. Machen AJ, Akkaladevi N, Trecazzi C, O'Neil PT, Mukherjee S, Qi Y, et al. Asymmetric Cryo-EM Structure of Anthrax Toxin Protective Antigen Pore with Lethal Factor N-Terminal Domain. Toxins (Basel). 2017;9(10).

34. Machen AJ, Fisher MT, Freudenthal BD. Anthrax Toxin Translocation Complex Reveals insights into the Lethal Factor Unfolding and Refolding Mechanism. bioRxiv. 2020.

35. Iacovache I, De Carlo S, Cirauqui N, Dal Peraro M, van der Goot FG, Zuber B. Cryo-EM structure of aerolysin variants reveals a novel protein fold and the pore-formation process. Nat Commun. 2016;7:12062.

36. Dal Peraro M, van der Goot FG. Pore-forming toxins: ancient, but never really out of fashion. Nat Rev Microbiol. 2016;14(2):77-92.

37. Cao C, Li MY, Cirauqui N, Wang YQ, Dal Peraro M, Tian H, et al. Mapping the sensing spots of aerolysin for single oligonucleotides analysis. Nat Commun. 2018;9(1):2823.

38. Cao C, Cirauqui N, Marcaida MJ, Buglakova E, Duperrex A, Radenovic A, et al. Single-molecule sensing of peptides and nucleic acids by engineered aerolysin nanopores. Nat Commun. 2019;10(1):4918.

39. Cao C, Ying YL, Hu ZL, Liao DF, Tian H, Long YT. Discrimination of oligonucleotides of different lengths with a wild-type aerolysin nanopore. Nat Nanotechnol. 2016;11(8):713-8.

40. Yang Z, Zeng X, Zhao Y, Chen R. AlphaFold2 and its applications in the fields of biology and medicine. Signal Transduct Target Ther. 2023;8(1):115.

41. Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, et al. Protein structure prediction using multiple deep neural networks in the 13th Critical Assessment of Protein Structure Prediction (CASP13). Proteins. 2019;87(12):1141-8.

42. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583-9.

43. Hutin S, Ling WL, Tarbouriech N, Schoehn G, Grimm C, Fischer U, et al. The Vaccinia Virus DNA Helicase Structure from Combined Single-Particle Cryo-Electron Microscopy and AlphaFold2 Prediction. Viruses. 2022;14(10).

44. Jin Y, Fyfe PK, Gardner S, Wilmes S, Bubeck D, Moraga I. Structural insights into the assembly and activation of the IL-27 signaling complex. EMBO Rep. 2022;23(10):e55450.

45. Radics J, Konigsmaier L, Marlovits TC. Structure of a pathogenic type 3 secretion system in action. Nat Struct Mol Biol. 2014;21(1):82-7.

46. Abrusci P, McDowell MA, Lea SM, Johnson S. Building a secreting nanomachine: a structural overview of the T3SS. Curr Opin Struct Biol. 2014;25(100):111-7.

47. Akeda Y, Galan JE. Chaperone release and unfolding of substrates in type III secretion. Nature. 2005;437(7060):911-5.

48. Stebbin C, Galán J. Priming virulence factors for delivery into the host. Nat Rev Mol Cell Biol. 2003;4:738-44.

49. Page AL, Parsot C. Chaperones of the type III secretion pathway: jacks of all trades. Mol Microbiol. 2002;46(1):1-11.

50. Birtalan S, Phillips R, Ghosh P. Three-dimensional secretion signals in chaperone-effector complexes of bacterial pathogens. Molecular Cell. 2002;9:971-80.

51. Stebbin C, Galán J. Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion. Nature. 2001;414:77-81.

52. Xin X-F, Kinch L, Cai B, Paasch BC, Kvitko B, Grishin NV, et al. *Pseudomonas syringae* effector AvrE associates with plant membrane nanodomains and binds phosphatidylinositides *in vitro*. bioRxiv. 2021.

53. Kobayashi DY, Tamaki SJ, Keen NT. Cloned avirulence genes from the tomato pathogen *Pseudomonas syringae pv. tomato* confer cultivar specificity on soybean. Proc Natl Acad Sci U S A. 1989;86.

54. Lorang JM, Keen NT. Characterization of avrE from Pseudomonas syringae pv. tomato: A hrp-Linked Avirulence Locus Consisting of at Least Two Transcripitonal Units. 1995. 1995;8(1):49-57.

55. Roine E, Wei W, Yuan J, Lassila LE, Kalkkinen E, Romantschuk M, et al. Hrp pilus: An hrp-dependent bacterial suface appendage produced by Pseudomonas syringae pv. tomato DC3000. Proc Natl Acad Sci U S A. 1997;94(7):3459-64.

56. Bogdanove AJ, Kim JF, Wei Z, Kolchinsky P, Charkowski AO, Conlin AK, et al. Homology and functional similarity of an hrp-linked pathogenicity locus, dspEF, of Erwinia amylovora and the avirulence locus avrE of Pseudomonas syringae pathovar tomato. Proc Natl Acad Sci U S A. 1998;95(3):1325-30.

57. Degrave A, Siamer S, Boureau T, Barny MA. The AvrE superfamily: ancestral type III effectors involved in suppression of pathogen-associated molecular pattern-triggered immunity. Mol Plant Pathol. 2015;16(8):899-905.

58. Gaudriault S, Malandrin L, Paulin JP, Barny MA. DspA, an essential pathogenicity factor of Erwinia amylovora showing homology with AvrE of Pseudomonas syringae, is secreted via the Hrp secretion pathway in a DspB-dependent way. Mol Microbiol. 1997;26(5):1057-69.

59. Frederick RD, Ahmand MD, Majerczak DR, Arroyo-Rodríguez AS, Manulis S, and Coplin DL. Genetic Organizaition of the *Pantoea stewartii* subsp. *stewartii hrp* Gene Cluster and Sequence Analysis of the *hrpA*, *hrpC*, *hrpN*, and *wtsE* Operons. MPMI. 2001;14:1213-22.

60. Mor H, Manulis S, Zuck M, Nizan R, Coplin DL, Barash I. Genetic Organization of the hrp Gene Cluster and dspAE/BF Operon in *Erwinia herbicola* pv. *gypsophilae*. MPMI. 2007;14:431-6.

61. Kim HS, Thammarat P, Lommel SA, Hogan CS, Charkowski AO. *Pectobacterium carotovorum* Elicits Plant Cell Death with DspE/F but the *P. carotovorum* DspE Does Not Suppress Callose or Induce Expression of Plant Genes Early in Plant-Microbe Interactions. MPMI. 2011;24:773-86.

62. Xin XF, Nomura K, Aung K, Velasquez AC, Yao J, Boutrot F, et al. Bacteria establish an aqueous living space in plants crucial for virulence. Nature. 2016;539(7630):524-9.

63. Potnis N, Colee J, Jones JB, Barak JD. Plant pathogen-induced water-soaking promotes Salmonella enterica growth on tomato leaves. Appl Environ Microbiol. 2015;81(23):8126-34.

64. Gentzel I, Giese L, Ekanayake G, Mikhail K, Zhao W, Cocuron JC, et al. Dynamic nutrient acquisition from a hydrated apoplast supports biotrophic proliferation of a bacterial pathogen of maize. Cell Host Microbe. 2022;30(4):502-17 e4.

65. Ham JH, Majerczak DR, Nomura K, Mecey C, Uribe F, He SY, et al. Multiple activities of the plant pathogen type III effector proteins WtsE and AvrE require WxxxE motifs. Mol Plant Microbe Interact. 2009;22(6):703-12.

66. Ham JH, Majerczak D, Ewert S, Sreerekha MV, Mackey D, Coplin D. WtsE, an AvrE-family type III effector protein of Pantoea stewartii subsp. stewartii, causes cell death in non-host plants. Mol Plant Pathol. 2008;9(5):633-43.

67. Sievers F, Higgins DG. Clustal Omega for making accurate alignments of many protein sequences. Protein Sci. 2018;27(1):135-45.

68. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.

69. Waterhouse AM, Procter JB, Martin DM, Clamp M, Barton GJ. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics. 2009;25(9):1189-91.

70. Oh CS, Carpenter SCD, Hayes ML, Beer SV. Secretion and translocation signals and DspB/F-binding domains in the type III effector DspA/E of Erwinia amylovora. Microbiology (Reading). 2010;156(Pt 4):1211-20.

71. Triplett LR, Melotto M, Sundin GW. Functional analysis of the N terminus of the Erwinia amylovora secreted effector DspA/E reveals features required for secretion, translocation, and binding to the chaperone DspB/F. MPMI. 2009;22(10):1282-92.

72. Xin XF, Nomura K, Ding X, Chen X, Wang K, Aung K, et al. Pseudomonas syringae Effector Avirulence Protein E Localizes to the Host Plasma Membrane and Down-Regulates the Expression of the NONRACE-SPECIFIC DISEASE RESISTANCE1/HARPIN-INDUCED1-LIKE13 Gene Required for Antibacterial Immunity in Arabidopsis. Plant Physiol. 2015;169(1):793-802.

73. Omersa N, Podobnik M, Anderluh G. Inhibition of Pore-Forming Proteins. Toxins (Basel). 2019;11(9).

74. Geny B, Popoff MR. Bacterial protein toxins and lipids: pore formation or toxin entry into cells. Biol Cell. 2006;98(11):667-78.

75. Forstner P, Bayer F, Kalu N, Felsen S, Fortsch C, Aloufi A, et al. Cationic PAMAM dendrimers as pore-blocking binary toxin inhibitors. Biomacromolecules. 2014;15(7):2461-74.

76. Yamini G, Nestorovich EM. Multivalent Inhibitors of Channel-Forming Bacterial Toxins. Curr Top Microbiol Immunol. 2017;406:199-227.

77. Branson TR, Turnbull WB. Bacterial toxin inhibitors based on multivalent scaffolds. Chem Soc Rev. 2013;42(11):4613-22.

78. Brett. H, W. ME. Dendrimers at Work. Science. 2006;313:5789.

79. Collmer A, Badel JL, Charkowski OA, Deng WL, D.F. F, A. RR, et al. *Pseudomonas syringae* Hrp type III secretion system and effector proteins. Proc Natl Acad Sci U S A. 2000;97(8770-8777).

80. Axtell MJ, Chisholm ST, Dahlbeck D, Staskawicz BJ. Genetic and molecular evidence that the Pseudomonas syringae type III effector protein AvrRpt2 is a cysteine protease. Mol Microbiol. 2003;49(6):1537-46.

81. Cunnac S, Lindeberg M, Collmer A. Pseudomonas syringae type III secretion system effectors: repertoires in search of functions. Curr Opin Microbiol. 2009;12(1):53-60.

82. Gangadharan A, Sreerekha MV, Whitehill J, Ham JH, Mackey D. The Pseudomonas syringae pv. tomato type III effector HopM1 suppresses Arabidopsis defenses independent of suppressing salicylic acid signaling and of targeting AtMIN7. PLoS One. 2013;8(12):e82032.

83. Degrave A, Fagard M, Perino C, M.N. B, S. G, S. L, et al. Erwinia amylovora type three-secreted proteins trigger cell death and defense responses in Arabidopsis thaliana. MPMI. 2008;21:1076-86.

84. Jin L, Ham JH, Hage R, Zhao W, Soto-Hernandez J, Lee SY, et al. Direct and Indirect Targeting of PP2A by Conserved Bacterial Type-III Effector Proteins. PLoS Pathog. 2016;12(5):e1005609.

85. Roussin-Leveillee C, Lajeunesse G, St-Amand M, Veerapen VP, Silva-Martins G, Nomura K, et al. Evolutionarily conserved bacterial effectors hijack abscisic acid signaling to induce an aqueous environment in the apoplast. Cell Host Microbe. 2022;30(4):489-501 e4.

86. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017;14(3):290-6.

87. Bepler T, Morin A, Rapp M, Brasch J, Shapiro L, Noble AJ, et al. Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. Nat Methods. 2019;16(11):1153-60.

88. StGelais C, Tuthill TJ, Clarke DS, Rowlands DJ, Harris M, Griffin S. Inhibition of hepatitis C virus p7 membrane channels in a liposome-based assay system. Antiviral Res. 2007;76(1):48-58.

89. Badel JL, Shimizu R, Oh H-S, Collmer A. A *Pseudomonas syringae pv. tomato avrE1/hopM1* mutant is severely reduced in growth and lesion formation in tomato. MPMI. 2006;19(2):99-111.

90. Stogios PJ, Downs GS, Jauhal JJ, Nandra SK, Prive GG. Sequence and structural analysis of BTB domain proteins. Genome Biol. 2005;6(10):R82.

91. Kinkema M, Fan W, Dong X. Nuclear localization of NPR1 is required for activation of *PR* gene expression. Plant Cell. 2000;12:2339-50.

92. Zhang Y, Fan W, Kinkema M, Li X, Dong X. Interaction of NPR1 with basic leucine zipper protein transcription factors that bind sequences required for salicylic acid induction of the PR-1 gene. Proc Natl Acad Sci U S A. 1999;96:6523-8.

93. 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. 2006;2(11):e123.

94. Liu G, Holub EB, Alonso JM, Ecker JR, Fobert PR. An Arabidopsis NPR1-like gene, NPR4, is required for disease resistance. Plant J. 2005;41(2):304-18.

95. Cao H, Li X, Dong X. Generation of broad-spectrum disease resistance by overexpression of an essential regulatory gene in systemic acquired resistance. Proc Natl Acad Sci U S A. 1998;95:6531-6.

96. Stogios PJ, Prive GG. The BACK domain in BTB-kelch proteins. Trends Biochem Sci. 2004;29(12):634-7.

97. Canning P, Cooper CDO, Krojer T, Murray JW, Pike ACW, Chaikuad A, et al. Structural basis for Cul3 protein assembly with the BTB-Kelch family of E3 ubiquitin ligases. J Biol Chem. 2013;288(11):7803-14.

98. Sedgwick SG, Smerdon SJ. The ankyrin repeat: a diversity of interactions on a common structural framework. Trends Biochem Sci. 1999;24(8):311-6.

99. Li J, Mahajan A, Tsai M. Ankyrin repeat: a unique motif mediating protein-protein interactions. Biochemistry. 2006;45:15168-78.

100. Kumar S, Zavaliev R, Wu Q, Zhou Y, Cheng J, Dillard L, et al. Structural basis of NPR1 in activating plant immunity. Nature. 2022;605(7910):561-6.

101. Hol WG. Effects of the α-helix dipole upon the functioning and structure of proteins and peptides. Adv Biophys. 1985;19:133-65.

102. Tada Y, Spoel SH, Pajerowska-Mukhtar K, Mou Z, Song J, Wang C, et al. Plant immunity requires conformational charges of NPR1 via S-nitrosylation and thioredoxins. Science. 2008;321:952-6.

103. Kabsch W. Xds. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 2):125-32.

104. Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, et al. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr D Struct Biol. 2018;74(Pt 6):531-44.

105. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60(Pt 12 Pt 1):2126-32.

106. Johnson C, Boden E, Arias J. Salicylic acid and NPR1 induce the recruitment of trans-activating TGA factors to a defense gene promoter in Arabidopsis. Plant Cell. 2003;15(8):1846-58.

107. Fumiaki. K, Eric. L, Chua N-H. Two tobacco DNA-binding proteins with homology to the nuclear factor CREB. 1989. 1989;340:727-30.

108. Jindrich K, Degnan BM. The diversification of the basic leucine zipper family in eukaryotes correlates with the evolution of multicellularity. BMC Evol Biol. 2016;16:28.

109. Després C, Delong C, Glaze S, Liu E, Fobert PR. The Arabidopsis NPR1/NIM1 Protein Enhances the DNA Binding Activity of a Subgroup of the TGA family of bZIP Transcription Factors. The Plant Cell. 2000;12:279-90.

110. Lebel E, Heifetz P, Thorne L, Uknes S, Ryals J, Ward E. Functional analysis of regulatory sequences controlling PR-1 gene expression in Arabidopsis. Plant J. 1998;16(2):223-33.

111. Tomaz S, Gruden K, Coll A. TGA transcription factors-Structural characteristics as basis for functional variability. Front Plant Sci. 2022;13:935819.

112. Zhang Y, Cheng YT, Qu N, Zhao Q, Bi D, Li X. Negative regulation of defense responses in Arabidopsis by two NPR1 paralogs. Plant J. 2006;48(5):647-56.

113. Pape S, Thurow C, Gatz C. The Arabidopsis PR-1 promoter contains multiple integration sites for the coactivator NPR1 and the repressor SNI1. Plant Physiol. 2010;154(4):1805-18.

114. Maldonade AM, Doerner P, Dixon RA, Lamb CJ, Cameron RK. A putative lipid transfer protein involved in systemic resistance signaling in Arabidopsis. Nature. 2002;419(6905):399-403.

115. Kachroo A, Kachroo P. Fatty Acid-derived signals in plant defense. Annu Rev Phytopathol. 2009;47:153-76.

116. Shi X, Che Z, Xu G, Ming Z. Crystal structure of transcription factor TGA7 from Arabidopsis. Biochem Biophys Res Commun. 2022;637:322-30.

117. Bligh EG, Dyer WJ. A rapid method of total lipid extraction and purification. Canadian Journal of Biochemistry and Physiology. 1959;37.

118. Li C, Tan BK, Zhao J, Guan Z. In Vivo and in Vitro Synthesis of Phosphatidylglycerol by an Escherichia coli Cardiolipin Synthase. J Biol Chem. 2016;291(48):25144-53.

Biography

I attended Hangzhou No.2 high school till 2011 and Zhejiang University for 2 years after the college examination. Then I transferred to Bryn Mawr College to complete my undergraduate degree (BA in Chemistry) in May 2016. I matriculated at the Duke University Department of Biochemistry in August 2016. I have been a graduate research assistant in the laboratory of Dr. Pei Zhou since May 2017. A list of my publications is as follows:

1. Nomura K[†], Andreazza F[†], **Cheng J**[†] et al. Bacterial Pathogens Deliver water/solute-permeable channels as a virulence strategy. *Nature*. (co-first-authored, under review)

2. Kumar S[†], Zavaliev R[†], Wu Q[†], Y Zhou, **Cheng J**, Dillard L, Withers J, Guan Z, Borgnia M, Bartesaghi A, Dong X, and Zhou P. Structural Basis of NPR1 Activation in Plant Immunity. *Nature* **605**, 561-566 (2022) (100)