The Structural Basis of the Negative Regulation of Plant Immunity

Madeline Louise Gish

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Ning Zheng, Chair

Adam Steinbrenner

Stanley McKnight

Program Authorized to Offer Degree:

Pharmacology

University of Washington

Abstract

The Structural Basis of the Negative Regulation of Plant Immunity

Madeline Louise Gish

Chair of the Supervisory Committee
Ning Zheng
Pharmacology

Humans are dependent on plants for food, energy and materials, yet plants are constantly under threat from bacteria, viruses and insects. Plants defend themselves from pathogens primarily through the expression of specialized defense genes, initiated by the key defense hormone salicylic acid, which is structurally related to the drug aspirin. Defense gene transcription and translation is a massive metabolic commitment by the plant, requiring tight negative control for long-term plant health. The salicylic acid receptor, NPR1, interacts with transcription factors to initiate defense gene transcription through an unknown mechanism. The Nimin family of proteins, under the control of an SA responsive promoter, form a negative feedback loop by directly interacting with NPR1 to stop defense gene transcription. The mechanism of this interaction, although necessary for control of the plant immune response, is poorly defined. In this work, I present the first

biophysical characterization of NPR1's interaction with the Nimin family of proteins and report the x-ray crystal structure of NPR1 salicylic acid binding domain in complex with Nimin-1, showing separate Nimin and SA binding pockets. I then demonstrate that Nimin-1 proteins regulate NPR1 signaling by allosterically competing with salicylic acid to bind to NPR1's C-terminal domain, and further identify key residues in the allosteric pathway between the two pockets.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1: Historical Perspectives on Plant Immunity

Section 1.1: Defense of the thesis topic

"The history of man is the history of a hungry creature in search of food."

Van Loon, *The History of Mankind*[1]

Animal and plant health is interdependent yet mutually destructive, complementary at the metabolic level, fused by cannibalistic necessity. Nutrition is only the most obvious of these interdependencies. As long as plants have existed— suspended between earth and sky, quietly and faithfully performing the alchemical transmutation of the inanimate into the animate— parasites have siphoned this creative excess for their own ends. Plant materials are used directly for our food, medicine, construction, and energy even as we are indirectly sustained by the ecosystems that plants formed in the past and participate in presently. The glittering inanimate edifice of human technological civilization rests upon living, organic foundations— foundations which are presently susceptible to ecological destruction at a wide scale. Genetically homogenous commercial crops are optimized for production at the cost of resilience. Each year, almost half of crops are lost to pests and crop diseases,[2] despite making up 80% of world caloric intake.[3] As climate change rapidly alters seasonal weather patterns and the geographical distribution of plant pathogens, losses to disease will only increase with inevitable consequences for the most vulnerable of human lives.[4] Food insecurity began rising in 2020, and as of 2021, 149 million children were affected developmentally by malnutrition, worsening health outcomes in the afflicted communities.[5] The plant immune response is therefore of fundamental

importance to human flourishing. Careful management of the trade-offs between defense and yield is required for humans to continue our utilization of plants as a resource not only for food, but also as the source of novel compounds for medicines, construction materials, energy, and the maintenance of our ecosystem as a whole. This work investigates the regulation of plant defense at the molecular level, forming only the most recent iteration of a historical inquiry going back thousands of years.

Section 1.2: Early Plant Biologists

"All trees, it may be said, have worms, but some less, as fig and apple, some more, as pear. Speaking generally, those least liable to be worm-eaten are those which have a bitter acrid juice..."

Theophrastus, An Enquiry into Plants (325 BC)[6]

Plant biologists— or earlier, botanists— have long been concerned with crop diseases, beginning with fungal infection. Theophrastus described various diseases affecting trees and other crops in 325 BC, including wheat rust and mildew.[6] As early as the 1900s, it had been widely observed that plants could acquire an immunity of some kind, but that the character of this resistance was markedly different from the acquired immunity found in humans, being non-specific and granular.[7] Plants became resistant to future infection if they were capable of what was termed "masking" or "recovery" in which the plant showed initial symptoms of infection but later recovered, although the virus could remain latently present.[7] It was clear by the 1930s that plants did not rely on antibodies. Thung showed that the acquired immunity was non-specific, meaning that masked or recovered plants were resistant to a second, novel virus. Price's grafting experiments using resistant stock indicated that a circulating humoral substance was not the

cause of immunity: as previously infected, masking stock did not transmit its resistance to the scion.[8] Chester described this immunity as "closely bound up in the living cell"[7] and it was known, contemporaneously, that resistance had a significant genetic component. In 1905, Biffen had noted that susceptibility to rust differed by wheat strain, and that this resistance appeared to be recessive according to Mendelian genetics.[9] Following on this line of research, H. H. Flor first proposed the "gene-for-gene" hypothesis in 1944, observing that resistance depended upon the presence of "resistance genes" in the plant that matched an "avirulence gene" in the pathogen.[10] He was then able to show that resistance could be lost by mutation resulting from x-ray radiation.[10]

In 1961, A. F. Ross connected acquired and genetic resistance through trials with tobacco plants resistant to tobacco mosaic virus (TMV). This strain, Samsun NN, developed a pattern of necrotic lesions that correlated with an overall reduction in systemic, catastrophic infection, termed the "hypersensitive response". Samsun NN plants developed these lesions when inoculated with TMV, and crucially, during a secondary infection in the exposed plants, the lesions were significantly reduced. That is, the hypersensitive response conferred immunity that was both acquired and systemic— coining the term "systemic acquired resistance". Samsun NN was named for the resistance gene N, which Ross introgressed from a separate strain of tobacco, showing that this gene was sufficient for systemic acquired resistance.[11] The genetic nature of this resistance implies a gene product, and in 1970, Van Loon published a study showing that novel proteins were produced in Samsun NN in response to viral infection, having separated the soluble protein fraction by electrophoresis.[12] By comparing Samsun and Samsun NN, he revealed that four novel proteins were produced systemically in response to viral infection only in the resistant strain, and crucially, that the proteins accumulated proportionally to the severity

of the hypersensitive response, indicating a causal relationship to resistance.[13] Polyacrylic acid had, at that time, been shown to stimulate immune responses in humans, and in 1974, Gianinazzi and Kassanis showed that polyacrylic acid of low molecular weight could stimulate the production of virus-responsive genes in plants as well.[14] In the coming years, a variety of acidic compounds would be tested for their ability to induce an immune response in plants.

Section 1.3: Loss and Rediscovery of Salicylic Acid and its Role in Plant Immunity

"There is a bark of an English tree, which I have found by experience to be a powerful astringent, and very efficacious in curing agues and intermitting disorders."

Edward Stone, Letter to the Royal Society (1763 CE)[15]

Before modern synthetic chemistry, plants were the primary source of medicinal compounds for human disease. The oldest known drug, as described by the Ebers Papyrus in 1500 BC and then by Hippocrates in 400 BC, is a tea brewed from willow leaves, which was prescribed for pain relief.[16] Other drugs derived from plants are still in use today, such as digitalis, morphine, and caffeine. Endogenously, these drugs are poisons produced to deter herbivores, but have therapeutic effects in humans at the correct dosage.[17–19] Although willow tea was widely known as an analgesic and antipyretic among ancient physicians, this knowledge was lost to the west after the fall of the roman empire until rediscovered in translations of Arabic copies.

Although ignorant of this intellectual history at the time, Edward Stone rediscovered the effects of willow bark tea in 1763, administering it as a powder to those suffering from "the agues" or fever[15]. Bark from the cinchona tree, which we now know contains the bitter anti-

malarial compound quinine, was known to be effective against malarial symptoms at the time but was expensive to acquire in Britain and negatively associated with Catholicism.[20] Stone based his enquiry upon the widely accepted doctrine of signatures and the prevalence of malarial symptoms in wet swampy areas. He went searching for a particularly bitter tree that grew in wet conditions, and upon tasting bark from a young white willow, submitted it to various trials with great success.[15] Eventually the active component of this extract was isolated by Henri Leroux in 1826, then named "salicin" from *salix alba,* or white willow, by Johann Buchner in 1828.[16] Ten years later, Raffaele Piria prepared salicylic acid from salicin, and in 1897, Felix Hoffman acetylated salicylic acid to make the drug now known as aspirin.[16] After its commercialization by Bayer, aspirin became one of the most well-known and successful drugs ever developed.

Nearly a century later, in 1979, R. F. White applied an aspirin solution to tobacco plants, discovering that aspirin was sufficient to induce resistance to tobacco mosaic virus.[21] In 1990, Métraux et al. showed that endogenous salicylic acid rose in concentration at the onset of systemic acquired resistance in a cucumber.[22] In 1993, Gafney et al. used transgenic plants with a bacterial gene encoding a salicylate hydroxylase, NahG, to show that the failure to accumulate SA blocked PR gene expression and rendered plants susceptible to infection.[23] Although grafting and exudate experiments had indicated that it was not a signal that circulated systemically, these studies suggested that SA served an endogenous role as a defense hormone. The identity of the receptor for this signal was still unknown.[24,25] Large scale mutagenesis studies were conducted to find mutants deficient in SAR to deconvolute the pathway, discovering SA binding proteins, transport proteins, and enzymes. Isochorismate synthase 1 (ICS1), a key enzyme in the synthesis of SA from chorismate was found.[26] Another enzyme was discovered, PBS3, which conjugates isochorismate with glutamate. This intermediate spontaneously degrades into SA, or is converted

by an acyltransferase, EPS1.[27,28] All three enzymes, when mutated, produce plants highly susceptible to infection with limited ability to accumulate endogenous SA. PAL converts phenylalanine to SA through a trans-cinnamic acid intermediate which is itself converted to benzoic acid through AIM1, and finally, to SA through an unknown hydroxylase.[28] PAL and AIM1 mutants are also immune-impaired.

Chapter 2: Types of Plant Immunity

Section 2.1: Pathogen Triggered Immunity

Although the same kinds of pathogens (bacteria, viruses, and insects) attack both plants and animals, the plant immune response developed under very different evolutionary pressures. Because plants are sessile, they co-evolve with the pathogens in their environment and therefore developed an immune response that is non-adaptive, non-specific and primarily genetically encoded. This response occurs in several phases over time as the infection, and therefore the necessary defenses, escalates. As plants evolved more sophisticated surveillance and defense mechanisms, the pathogens that depend upon plant tissues for survival evolved ways of evading and exploiting these mechanisms in turn, all of which are deployed sequentially during infection.[29]

Plants first sense the presence of pathogens through cell-surface receptors that detect pathogen associated molecular patterns, or PAMPs. Before invading the cell itself, pathogens first enter the space between plant cells, the apoplast, through wounds, or small pores called stromata. During pathogen entry, bacteria, insects and viruses shed molecules like flagellin, chitin, or capsid proteins that bind to transmembrane receptors at the cell-surface. When the PAMP molecule binds to the extracellular ligand binding domain, it causes conformational changes in an intracellular kinase or tail domain.[29] This begins a signaling cascade that activates innate cellular immunity in the form of Ca^{2+} influx, callose deposition at the cell wall, the generation of reactive oxygen species, and the transcription of specialized defense genes.[30] This immune response resulting from PAMP detection is known as pathogen triggered immunity, or PTI. In many cases, PTI is effective in preventing further infection.

Section 2.2: Effector Triggered Immunity

In order to overcome PTI, pathogens evolved small signaling peptides known as effectors that inhibit the defense signaling cascade, or lower defenses by interacting directly with endogenous proteins.[29,31] For example, defense signaling closes leaf stromata during PTI. In order to gain entry to the leaf, *pseudomonas syringae* effector *AvrB* increases the opening of stromata by interacting with an endogenous protein RIN4.[31] Other pathogen effectors promote the degradation of key defense signals, or increase the transcription of favorable signaling proteins.[31,32] In response, plants evolved receptors to monitor the activity of these effectors through cytosolic nucleotide binding leucine rich repeat (NB-LRR) receptors.[29] These receptors both bind to effectors themselves or monitor endogenous proteins for damage associated with effector activity, known as the "guard hypothesis".[29] Effector perception results in the second phase of the immune response known as effector triggered immunity, ETI.

A hallmark characteristic of ETI is localized programmed cell death at the site of infection. This presents as a series of brown spots, small lesions, at the point of pathogen entry, called the "hypersensitive response".[7,33] Plants evolved this defense mechanism to evade biotrophic pathogens that require living tissue for pathogenesis, such as viruses that replicate within living cells. There are other pathogens, called necrotrophic, that require dead tissue. Some pathogens, like certain fungi, bridge these categories by forming long-term parasitic relationships with plants.[34] The hypersensitive response prevents systemic infection by sacrificing healthy cells surrounding the point of pathogen entry to quarantine biotrophs, preventing pathogen spread.

Section 2.3: Systemic Acquired Immunity

Although canonically presented as separate phases of the immune response, recent studies indicate that PTI and ETI are not fully independent. ETI intensifies PTI signaling but ETI alone is not sufficient for immunity.[29,31] Both PTI and ETI are local responses at the site of infection that also generate a systemic signal, N-hydroxypipecolic acid (NHP) that initiates pre-emptive defenses throughout the plant.[35] This means that while a plant may show signs of the hypersensitive response at the primary site of infection, when challenged after some time with the same pathogen elsewhere, the plant is unaffected and neither hypersensitive response nor infection occurs.[11] The plant is said to have acquired a resistance systemically that it did not display during the primary challenge called "systemic acquired resistance" or SAR.

Plant defense in all three responses, PTI, ETI and SAR, relies on the transcription of pathogenesis related (PR) genes. These genes encode small acidic or basic proteins with specialized defense roles. Acidic PR proteins are excreted into the apoplast where they interact with the invading pathogens while basic PR proteins are trafficked to vacuole.[36,37] Some known PR genes are PR-2, a β-1,3 glucanase and PR-3, PR-4, PR-8, and PR-11 which are chitinases.[36,37] These enzymes degrade bacterial or insect cell walls.[36] Others, like PR-1, have functions that are still un-unexplained. Although its mechanism is unknown, over-expressing PR-1 alone is enough to confer resistance to *P. Parasitica* indicating the sufficiency of this transcriptional response for defense in SAR.[38] During the defense response, PR genes are massively expressed, comprising 5-10% of total proteins in leaves.[39] While PTI and ETI involve other defense mechanisms in addition to PR genes, SAR is entirely driven by this transcriptional response, meaning that control of transcription is key to plant survival.

Chapter 3: Molecular Basis for Systemic Acquired Resistance

Section 3.1: Roles of Salicylic Acid in Plant Immunity

Following pathogen perception, a chain of signaling events is required to cause transcription of defense genes both locally, and systemically. PRR signaling generates reactive oxygen species, a transient influx of calcium, and transient activation of MAPK signaling. Calcium binding to transcription factors CBP60g (CALMODULIN BINDING PROTEIN 60g) and SARD1 (SYSTEMIC ACQUIRED RESISTANCE DEFICIENT 1) upregulates transcription of ICS1, key enzyme in salicylic acid synthesis.[30] The transcription of other enzymes key to the accumulation of salicylic acid is initiated through similar pathways, such as transporters. As SA accumulates, it binds to its receptors NPR1 and NPR4 in the nucleus, recruiting transcription machinery to PR gene promoters through an unknown mechanism.[29,40–42]

In addition to initiating PR gene expression, SA signaling also plays a key role in the hypersensitive response. Salicylic acid directly inhibits catalases required for maintaining the oxidation-reduction balance within the cell.[43,44] At high enough concentrations of salicylic acid the build-up of reactive oxygen species triggers apoptotic cell death. SA synthesis is controlled by NPR1-regulated TGA transcription factors so that SA participates in a positive feedback loop to quickly increase SA concentration in response to pathogenesis.[45,46] This causes necrosis that is spatially correlated to the severity of infection, preventing biotrophic pathogen spread in addition to initiating systemic PR gene expression.[47,48]

Section 3.2: Salicylic Acid Synthesis and Evolution

Salicylic acid is one of many plant phenols synthesized from chorismate.[28] Synthesis occurs through two known pathways, the isochorismate synthase 1 (ICS1) pathway and the

phenylalanine ammonia lyase (PAL) pathway.[28] The ICS1 pathway is conserved in the most recent common ancestor of land plants, but SA as a molecule is present even in green algae, and is thought to be synthesized through an ancient β-oxidation pathway that converts trans-cinnamic acid to benzoic acid.[49] In all land plants, salicylic acid is a defense signal that activates the transcription of PR genes. Endogenous SA concentration rises rapidly after pathogen perception and exogenous application of SA causes the transcription of marker PR genes even in the absence of infection.[22,25] Both the loss of SA synthesis or failure to accumulate SA compromises the plant defense response. avrPphB SUSCEPTIBLE3 (PBS3), ENHANCED PSEUDOMONAS SUCEPTIBILITY 1 (EPS1), and ENHANCED DISEASE SUSCEPTIBILTY 5 (EDS5) are enzymes involved in SA synthesis named for susceptibility phenotypes discovered in mutagenesis screens.[28,50] Transgenic *NahG* plants expressing the bacterial enzyme salicylate hydrolase cannot accumulate SA, do not express PR genes, and are therefore highly susceptible to infection.[23] PAL converts phenylalanine to SA through a trans-cinnamic acid intermediate which is itself converted to benzoic acid through AIM1, and finally, to SA through an unknown hydroxylase.[28] PAL and AIM1 mutants are also immune-impaired.

Section 3.3: Discovery of NPR Proteins and Proposed Mechanisms

The receptor linking SA to gene expression remained unknown, however, until a study in 1994 used a reporter gene with an SA inducible promoter to search for PR gene expression phenotypes.[51] This screen produced a mutant plant called “non-expressor of PR genes 1” or NPR1, that failed to express PR-genes in response to SA.[51] NPR1 was later found to encode a protein with ankyrin repeats.[52] NPR1 has been discovered over 20 times in genetic screens for SAR impairment as a variety of mutations can impair PR gene expression. Based on this genetic

evidence, the case for NPR1 being the receptor for SA would seem settled, but until very recently, there was mixed evidence that NPR1 even bound SA at all. When NPR1 was found to translocate from the cytosol to the nucleus during SAR, it was suggested that NPR1 indirectly sensed SA through redox-sensitive aggregation.[40,53] NPR1 would then monomerize during the oxidative burst associated with early PTI, catalyzed by thioredoxins.[54] It was later found that s-Nitrosylation, SUMOylation, and phosphorylation also regulate the nuclear localization of NPR1.[54] While some of, though not all, cysteines expected to underlie this reversible aggregation are conserved and confirmed to regulate translocation, the primacy of NPR1 directly binding SA in defense signaling was clearly supported by the genetic data. Most of point mutations affecting NPR1 mediated gene transcription were not expected to affect aggregation or translocation. Additionally, NPR1's homology to its paralog, NPR4 came to support NPR1's role as a direct receptor for SA.

NPR4 was discovered by its high sequence identity to NPR1, rather than a genetic screen for loss of function, meaning its role in SAR has been more disputed.[55] A very early paper proposed that NPR4 is required for defense but later studies indicated that NPR4 might play a repressive role.[55] *Npr3npr4* double knockout plants have elevated basal PR gene expression and an enhanced immune response for example.[56] The same group conducted a screen in a constitutively expressing cell-line *npr1 snc1-D* that had a severely stunted phenotype. A single point mutation, R419Q, in NPR4 restored the WT growth phenotype.[57] This mutation was later found to render NPR4 salicylic acid insensitive.[58] The same group also found that NPR3 and NPR4 repressed the transcription of a reporter gene in a GAL4 DNA binding domain based assay. From this, it was postulated that NPR4 represses gene expression and that this repression is lifted when it senses SA.[57] In 2020, the structure of the salicylic acid binding domain in the C-

terminus of NPR4 was solved by Wang et al. The architecture of the SA binding pocket in NPR4 is conserved in NPR1, and this group confirmed that NPR1 binds SA, although with lower occupancy than NPR4.[58]

Two opposing theories were proposed to explain NPR4 repression of NPR1-mediated PR gene expression. Because NPR proteins dimerize through BTB domains and this structure is also contained by several Cullin family E3 ligase substrate adapters, one group proposed that NPR4 functions as a substrate adapter for E3 ligase Cul3, dimerizing with and thus degrading NPR1 in the absence of SA. Once NPR4 binds SA, this interaction is lost, stabilizing NPR1 and allowing it to move into the nucleus and interact with TGAs as a positive regulator. This was based on co-IP data showing that NPR1 could pulldown Cul3, that this pull-down was reduced in *npr3npr4* lines, and that both proteasome inhibitors and SA could stabilize NPR1.[59] A contradictory theory was also proposed where both NPR1 and NPR4 interact with TGA transcription factors, one as a positive regulator, and one as a repressor. NPR4 contains a motif in its C-terminus similar to a well-known EAR motif that could potentially recruit a general transcription repressor, TOPLESS. In transcription reporter assays, fusing NPR4 to the DNA-binding domain repressed the reporter gene.[57]

Both theories have flaws. NPR4 is proposed to bind NPR1 monomers in the absence of SA despite NPR1 forming redox sensitive oligomers and only monomerizing in the presence of SA. While this could be explained by NPR4 serving a quality control function to police oligomerization, it does not explain why only NPR4 is expected to function as the substrate adapter for Cul3 even though NPR1 and NPR4 share a high sequence identity and contain conserved BTB domains. While in the alternate theory, both NPR1 and NPR4 act within the nucleus, which accounts for the sequence similarity between them, NPR4 is differentiated from

NPR1 by the putative EAR-like repressor motif which is also found in NPR1-like proteins in other, earlier plants meaning that NPR4 does not conserve this motif exclusively. Additionally, a family of proteins with this EAR motif had already been discovered and shown to reduce PR gene expression by interacting with NPR1 directly.[60]

Section 3.4: NPR Family and Domain Architecture

The SA receptor NPR1 is named for the "non-expression of pathogenesis related genes" phenotype when its function is lost.[51] There are six members of the NPR protein family. NPR1 was the first to be discovered, found in a genetic screen in 1994, while the other paralogs were identified by sequence similarity to NPR1. NPR1 and NPR2 are positive regulators of the plant immune response. Although NPR1 is expressed more highly, NPR2 can compensate for the loss of its role in transgenic plants. NPR3 and NPR4 are also redundant with each other, but unlike NPR1, they negatively regulate plant immunity. NPR4 is more highly expressed than NPR3, and thus more discussed in the literature. NPR5 and NPR6 are more widely known as BOP1 and BOP2 for "Blade-on-Petiole" because they regulate plant flowering, rather than PR gene expression.[61]

The NPR family of proteins share a common architecture made up of two conserved domains. They homodimerize through an N-terminal BTB-POZ (or Bric-a-Brac Tram-track Broad complex- POx virus and Zinc finger) domain, and interact with transcription factors through several ankyrin repeats.[41,52] The structure of this complex was recently solved and published in 2022, showing a dimer of NPR1 bridging two dimers of TGA3.[41] NPRs bind SA through a flexible C-terminal domain with a conserved LExRV motif. Wang et al. solved the structure of this "salicylic acid binding core" or "SBC" of NPR4 in 2020, showing that the

conserved arginine in this motif contacts SA within a deep pocket.[58] The domain is a small, highly labile bundle of alpha-helices connected to the N-terminal domain through a flexible linker. Both NPR1 and NPR4 conserve a key arginine residue required for salicylic acid binding.[58]

Section 3.5 TGA Transcription Factor Family and Architecture

Once in the nucleus NPR proteins interact with a family of transcription factors, the TGAs, to mediate the transcription of PR genes.[45,62] SA controlled PR genes contain tandem TGACG sites in the promoter region, recognized by TGACG-binding (TGA) transcription factors.[45,63–65] TGA2, TGA5, and TGA6 function redundantly to promote defense gene expression. The triple knock-out plant *tga2tga5tga6* recapitulates the *npr1*-like phenotype, lacking PR gene expression with greater susceptibility.[66] NPR1 is proposed to contain a C-terminal "transactivator" domain that strengthens TGA interaction with the PR gene promoter sequence, but this transactivation currently has no known structural basis.[41,67,68] While NPR1 must interact with general transcription factors to initiate the transcription of PR genes, how they are recruited and whether this interaction is SA dependent is not known.

The TGACG-binding (TGA) family of transcription factors interacts with NPRs to mediate PR gene transcription. There are ten TGAs in Arabidopsis, organized into five clades.[69,70] As basic region leucine zipper (bZIP) transcription factors, TGA dimers bind the TGACG motif in PR gene promoters through the leucine zipper of an N-terminal bZIP domain.[70,71] A C-terminal domain interacts with the ankyrin repeats of NPR proteins.[41]

.

Chapter 4: X-Ray Crystallography and Structural Perspectives

Section 4.1: History of X-ray Crystallography

"The secret of life lies in the structure of proteins, and there is only one way of solving it, and that is by x-ray crystallography."

— J. Desmond Bernal[72]

X-ray crystallography has its roots in two papers from 1912 and 1913 by father and son William Henry Bragg and William Lawrence Bragg. They mathematically described the first x-ray diffraction images produced by Max von Laue of copper sulfate pentahydrate. This was the genesis of Bragg's law ($n\lambda = 2dsin\theta$) relating the spacing of the crystal lattice with the angle of diffraction and the wavelength of the x-rays used. This equation is foundational to the practice of x-ray crystallography, earning W.L. Bragg a Nobel at the advanced age of twenty-five. While W. L. Bragg went on to characterize other salts and minerals during a long and successful career, in the next few decades the field progressed beyond simple inorganic salts without him, using x-ray diffraction to solve the structures of increasingly sophisticated macromolecules.[73] Today, protein crystallization and x-ray diffraction have become indispensable techniques in both basic and industrial research from structure-function studies to drug discovery.

X-rays are diffracted when the wavelength of the x-ray matches the inter-atomic spacing of the sample, scattering into many directions. Most of the x-rays are not scattered at all, and of the scattered rays, most interfere destructively, effectively erasing themselves. The scattered rays

that interfere constructively form the characteristic pattern of dots which are the fundamental data for structure determination. Because of the constructive interference of the scattering, the angles and intensity of the pattern are the sum of many coincidental waves, and therefore contain information about the electron density of the entire crystal.[73] The difficulty of macromolecular x-ray diffraction is that atoms, and therefore proteins, are mostly empty space, making diffraction infrequent. Thus, the diffraction power of a single protein is too weak to generate a signal sufficient for analysis. In order to collect a data set of reflections for structure determination, this signal must amplified. To do this, a situation must arise in which many copies of the target protein interact with the x-ray beam in nearly the exact same way, increasing the probability of diffraction. Protein crystallization is how this hypothetical situation is made real by the crystallographer.

Protein crystals are orderly arrangements of the asymmetric unit repeated until the crystal is visible to the human eye. Only if the packing of the asymmetric unit is rigid and uniform enough can the entire crystal substitute for the asymmetric unit, allowing for sufficient diffraction. The larger and more uniform the crystal, the stronger the signal and the higher the resolution of the resulting structure. Therefore, optimal crystallization of the protein target is required before a high resolution structure can be determined.

The first protein ever crystallized was hemoglobin, with reports of crystallites found in prepared blood and tissue samples appearing as early as 1840. Crystallization in its early application was a purification technique. Only in 1934 was the first x-ray photograph of protein crystal (pepsin) published, and the first protein structures solved in 1963 (hemoglobin) and 1965 (lysozyme), birthing the field of protein crystallography for structure determination. Early methods for crystallization would shock diligent crystallographers today, including the treatment

of crude extracts with alcohol, or even hot acetic acid.[74] Slowly, through the efforts of many dedicated scientists, perspectives on crystallization has now evolved beyond mere random chance, incorporating over time theories of purity, solubility, supersaturation, the use of specific salts, and the effect of temperature among others.[74] These advances have resulted in a process which is, today, only mostly random.

Section 4.2: Theory of Crystallization

Crystallization is, essentially, a controlled form of precipitation. The protein is encouraged by the correct environment to precipitate out of solution at a rate that allows each protein molecule to first access the conformation required for the asymmetric unit, and then to deposit molecule-by-molecule onto the growing crystal lattice.[75] The rate of deposit must exceed the rate of dissociation for the crystal to grow. If precipitation outpaces the on-rate of deposition onto the crystal face, the protein forms a crystalline precipitate or dust in which the individual crystals are too small to be harvested, but if it is too slow, the protein may remain perennially in phase-separated droplets without crystallizing.[75] Additionally, the off-rate of subunits that dissociate from the crystal must be slow enough to allow growth, but sufficient to reduce misalignment of the asymmetric unit, as internal flaws in the orientation of the crystal lattice also reduce resolution. Growth stops when deposition and dissociation reach equilibrium. The job of the crystallographer is to engineer the conditions that allow for crystal formation of sufficient size and quality for structure determination.[74]

The first condition for crystallization is supersaturation of the protein target. Any solvated substance precipitates when its concentration exceeds the capacity of the solvent to contain it, called “saturation”. A supersaturated solution is one where the solute has reached the

point of saturation, and yet remains dissolved. To supersaturate the protein sample, it is concentrated until near saturation, then evaporated slowly in the presence of crystallization reagents. In high-throughput hanging drop crystallization, a nanoliter amount of sample is suspended from a seal above a well filled with reagent. Once the seal is made, the drop evaporates into the space within the well. Now that the droplet is supersaturated, nucleation is more likely.[74]

Supersaturation is aided by the vapor diffusion of the crystallization reagents into the droplet. Because the volume of the reagent in the well is much greater than the small droplet, its vapor pressure is much higher, so after evaporation and condensation reach equilibrium, the well buffer will have entirely replaced the liquid within the protein droplet. This buffer includes precipitants that lower the concentration required for supersaturation. The choice of precipitant is essentially random, necessitating efficient trials of a wide variety of possible buffer conditions. Empirically formulated screens designed for use with 96 well microplates are commercially available, allowing many thousands of conditions to be assayed while conserving the protein sample. The Protein Data Bank requires crystallization information to be published with its deposited structure, and analysis of this dataset of successful crystallization conditions has identified certain trends. Post-hoc explanations for these trends now underlie the choice of matrixing in commercially available reagents.[74]

Salt, polymers, pH and temperature all play a role in lowering the solubility of the protein target. Salts in low concentration help to mediate like-charge repulsion of charged amino acids on the protein surface, thereby promoting solubility but when the salt concentration exceeds the capacity of the solvent (water) to dissolve both the salt and the protein, the salt is the more soluble of the two and forces the protein out of solution. Ammonium sulfate is one of the most

common and effective precipitating salts. Polymers, which are intrinsically disordered and therefore highly labile, also exclude the protein from solution by dominating the solvent space relative to the comparatively more rigid protein target. Protein solubility is also a function of pH and proteins are less soluble at their pI. Buffering at a pH above or below the pI alters the polarity of intermolecular attraction between asymmetric units, and may be optimized to promote crystal packing. Temperature is another key factor that regulates both solubility and the kinetic rates of deposition and dissolution. Storing micro-scale crystallization plates at 4ºC reduces the solvent capacity of the solution and slows the crystallization process, increasing crystal quality at the expense of the crystallographer's personal comfort— which in the mature stages of the doctoral process becomes increasingly negotiable.[74]

Although the precise condition that allows for crystallization of a specific target cannot be known in advance, the commercial availability of reagents and materials, the technological advances in automation allowing microscale and high throughput screening, as well as the wealth of knowledge accumulated since the 1930s means that crystallizing a novel target is easier and more organized now than at any time in history. Even in light of the increasing convenience of the crystallization process, it is still common for this aim to require several years of effort. This is because the most significant contributor to the crystallization of a protein is not the crystallization conditions at all, but rather the fitness of the protein itself. While hemoglobin almost spontaneously crystallizes, another protein may never crystallize at all, despite years of efforts. The main culprit in this paradox is the flexibility of the protein target. While all proteins "breathe" as they stochastically access different conformations, proteins with fewer available conformations are more rigid, and pack against themselves more easily, increasing the probability of crystallization. If the target protein has a wide variety of possible conformations—

If, for example, it contains an intrinsically disordered loop that can adopt of spread of positions— this reduces the probability that the protein will access the conformation required for deposition onto the crystal lattice. The presence of flexible and disordered regions may destroy the ability of the protein to crystallize entirely by obstructing key inter-subunit contacts that maintain the integrity of the crystal as a whole. These regions can be removed by the thoughtful crystallographer ahead of time through identification of the conserved “minimal” domain. Thus, the most important and influential stage in the crystallization process is, in fact, the beginning, when the construct is designed and optimized.

Chapter 5: Crystallization of NPR1-Nimin-1 Complex

Section 5.1: Crystallization of NPR1 Paralog NPR4

NPR proteins have two conserved domains, an N-terminal domain made up of BTB-POZ and several ankyrin repeats, and a C-terminal domain that binds salicylic acid. The crystal structure of a subdomain of the C-terminal domain, the salicylic acid binding core (SBC) of NPR4 was published in 2020 by Wang et al. This group identified the salicylic acid binding core within the larger C-terminal domain of NPR4 using limited trypsin digestion to determine the domain's boundaries, which correspond to residues aa373-516. Because this fragment was insoluble, it was purified under denaturing conditions and then refolded with SA. This protein was initially crystallized by hanging-drop vapor diffusion without any changes, but to improve the resolution, a short internal deletion was made, removing residues aa450-455, allowing a dataset with 2.28Å to be collected. NPR1 and NPR4 share 30% sequence identity, and NPR1 conserves the arginine required for binding SA. Wang et al. confirmed that NPR1 binds SA, but it appears to interact with SA differently than NPR4 does.[58] Because the structure of the C-terminal domain of NPR1 has not yet been solved, this has not yet been explained.

Section 5.2: Identification of the Minimal Domain for NPR1-Nimin-1 Complex

Minimal domains are identified by performing a conservation analysis, then confirmed by designing and testing increasingly radical domain boundaries. This minimal domain will also be the most rigid form of the target because it excludes flexible, un-conserved regions. In order to crystallize the NPR1-Nimin-1 complex, I first assessed the evolutionary conservation of NPR1.

A multiple sequence alignment (MSA) of NPR1 sequences of various plants from different evolutionary clades revealed that the conserved C-terminal domain of NPR1 extends beyond the region analogous to the SBC in NPR4. Using this MSA, I was able to assign conservation scores to each residue.[76] Then I predicted the structure of NPR1 in complex with Nimin-1 using AlphaFold, and mapped these conservation scores by residue, allowing me to evaluate the contribution of the conserved residues to the predicted structure of the protein.[76,77] The SBC is made up of four α-helices,[58] but the prediction indicated that the conserved residues form an extra α-helix in the C-terminal domain. The complex prediction indicated that the additional the C-terminal helices of NPR1 were sufficient to bind Nimin-1, and that the minimal domain for binding Nimin-1 includes two more α-helixes than shown in the solved structure of the NPR4 SBC. I confirmed the boundaries of this minimal domain by performing a pull-down with different domains of NPR1 and Nimin-1 (Fig. 1 and 2). The N-terminal domain, composed of a BTB domain and several ankyrin repeats (BTB-AR) is not required for interaction with Nimin-1, while reducing the additional helix by only 13 residues abolishes NPR1's interaction with Nimin-1 (Fig. 1). This novel domain from residues aa414-564 will be called the SBD, or salicylic acid binding domain.

Both NPR1 and NPR4 contain a loop of non-conserved residues between two helices within this domain. When crystallizing the SBC of NPR4, Wei et al. deleted residues aa450-455 to reduce flexibility. This deletion is analogous to aa466-470 in NPR1. I cloned, expressed and purified a minimal SBD with or without residues aa466-470, but was unable to crystallize either complex. As the NPR1-Nimin-1 complex is small, much of it is exposed to the solvent, and thus the failure of the complex to crystallize could indicate residual flexibility in the complex.

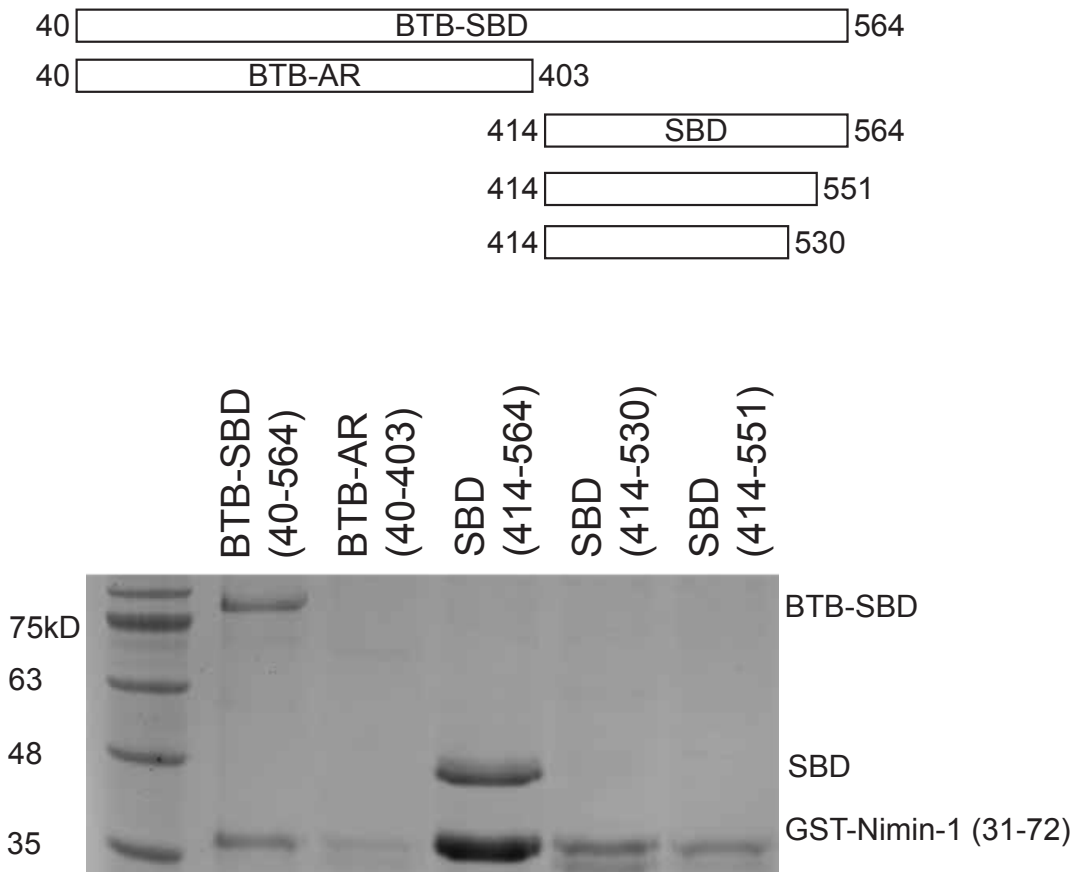


Figure 1: Diagram of NPR1 domain boundaries and minimal domain pulldown

NPR1 domain boundaries: NPR1 is made up of two broad domains, the N-terminal domain that contains a BTB domain and ankyrin repeats (BTB-AR), and a c-terminal domain that binds salicylic acid (SBD).

Pulldown demonstrating minimal domain for NPR1-Nimin-1 interaction: GST-Nimin-1 (31-72) pulling down his-msb-flag tagged NPR1 using glutathione agarose, eluted by glutathione.

Nimin-1 can pull-down the full-length NPR1, and the SBD, but has no interaction with the BTB-AR domain, or any truncation of the SBD.

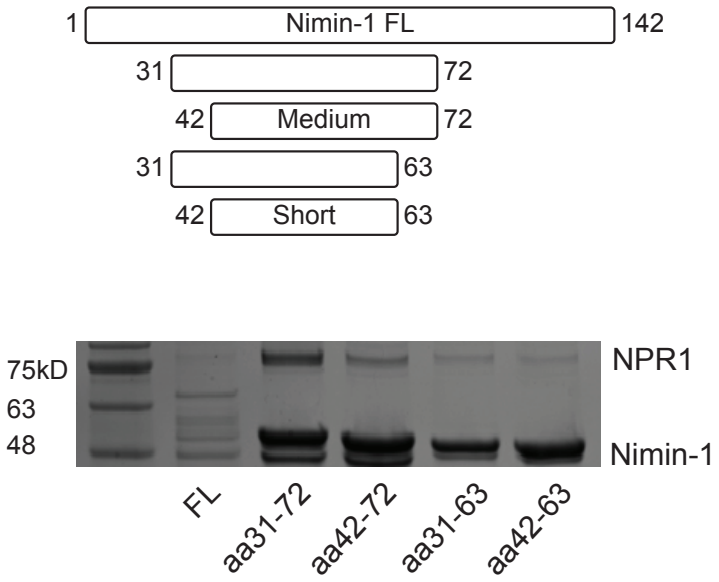


Figure 2: Diagram of Nimin-1 domain boundaries and minimal domain pulldown

Nimin-1 contains a highly conserved central region containing a DxFFK motif.

Pulldown demonstrating the minimal domain for interaction with NPR1, indicating that the central helix (aa31-72, and aa42-72) are sufficient for interaction with NPR1.

MBP-Nimin-1 domains pulling down His-msb-flag-NPR1 by maltose agarose, eluted by maltose.

Section 5.3: Assessing the Complex for Flexibility

To assess the NPR1-Nimin-1 complex for flexible or disordered regions that might interfere with crystal packing, I digested the complex with α-chymotrypsin, trypsin, elastase, papain, subtilisin, and Endoproteinase Glu-C (Fig. 3). Proteases only cleave the portions of the protein that are accessible, meaning that flexible, solvent exposed residues are more likely to be cut. The cleavage pattern then indicates stability, a technique Wang et al used to identify the SBC. Most of the proteases generated a spread of small fragments from NPR1-SBD. α-Chymotrypsin, however, cut the complex into two discrete bands, which suggested that it split the complex into yet smaller, more minimal sub-domains(Fig. 3). Hanging drop crystallization with *in situ* proteolysis using 1:4000 α-chymotrypsin yielded several small, un-replicable crystals (Fig. 4 and 6).

- 0.1M Sodium Hepes pH 7, 20% w/v PEG 8,000
- 0.1M Magnesium Acetate tetrahydrate, 0.1M MES pH 6.5, 10% w/v PEG 10,000
- 0.1M Sodium Hepes pH 7, 18% w/v PEG 12,000

To deconvolute the effect of cleavage by α-chymotrypsin, I identified the cut-site using N-terminal sequencing. α-Chymotrypsin cleaves the carboxyl side of the peptide bond of bulky, aromatic residues like phenylalanine, tyrosine, and tryptophan. Accordingly, N-terminal sequencing identified the residues at the N-terminus of each fragment, indicating that α-chymotrypsin cut the complex at the TEV cut site N-terminal to the SBD, and in the middle of a disordered loop. (Fig 5.) This loop could prevent crystal packing by disrupting the interface between

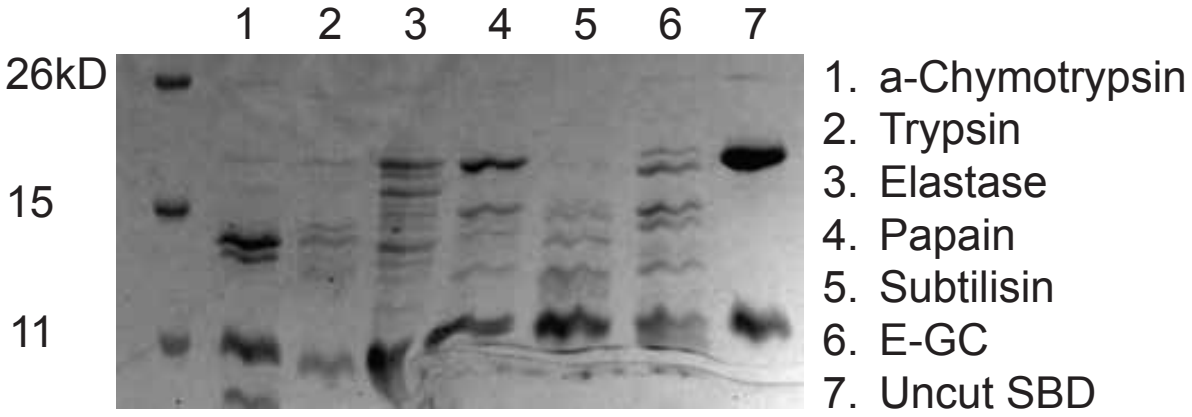


Figure 3: Panel of Proteases for assessing complex flexibility

Trial digestion of NPR1-SBD-Nimin-1 complex with different proteases, showing that α-Chymotrypsin cuts into two discrete bands while other proteases generate a spread of fragments.

9µg of complex incubated with 1:200 mass ratio of protease, using Hampton Reasearch HR-492 kit

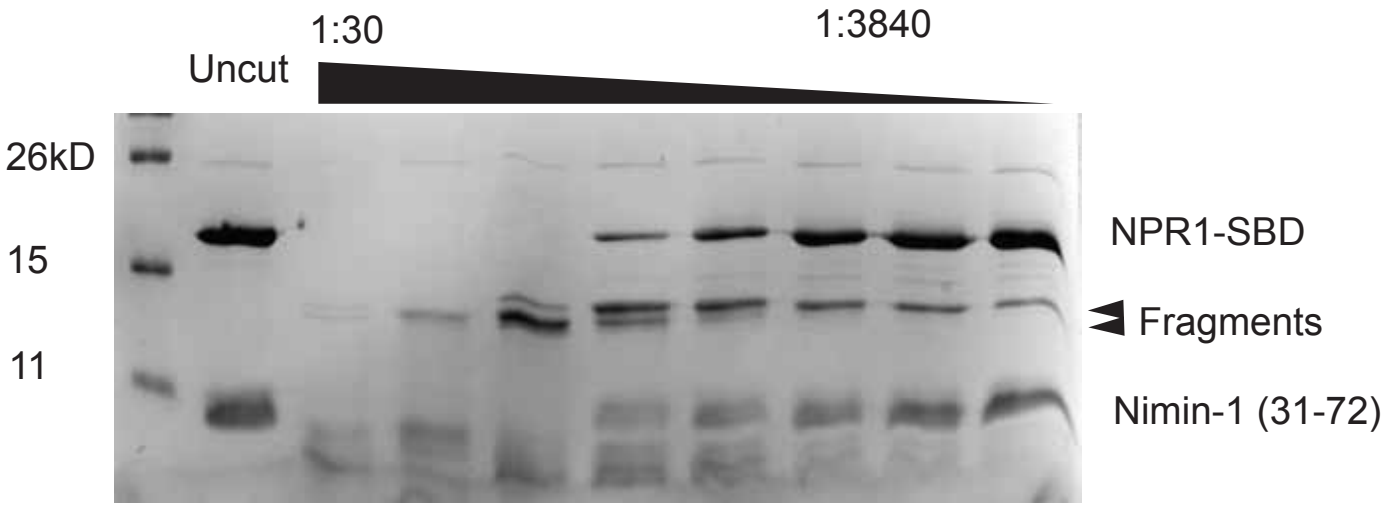
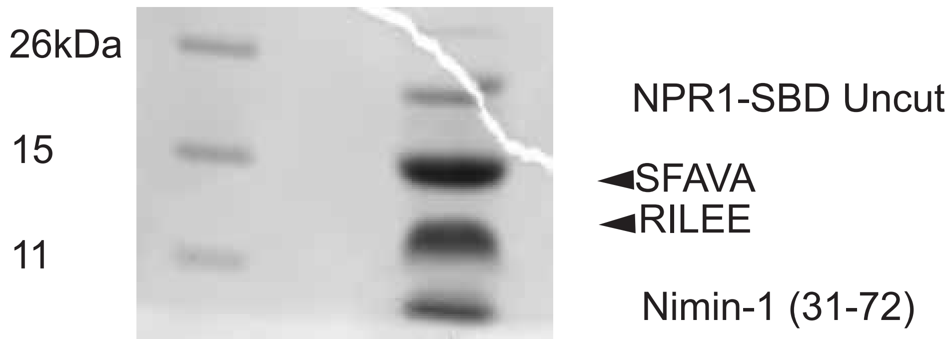


Figure 4: Effect of α-Chymotrypsin titration on complex cleavage

Serial dilution of α-Chymotrypsin cleaving the NPR1-SBD complex, showing that cleavage into two fragments is stable, optimized at 1:120 ratio at room temperature.

Serial dilution of 1:2, beginning at a 1:30 mass ratio of complex to protease.



NPR1	FAVAAD	ELKMTLLDLE	NRVALAQRLF	PTEAQAAMEI	AEMKGTCEFI	VTSLEPDRLT	GTKRTSPGVK	IAPFRILEEH
SBDΔ	FAVAAD	ELKMTLLDLE	NRVALAQRLF	PTEAQAAMEI	AEMKGTCEFI	VTSLE-----	GTKRTSPGVK	IAPFRILEEH
SBDΔ2	FAVAAD	ELKMTLLDLE	NRVALAQRLF	PTEAQAAMEI	AEMKGTCEFI	VT--------	--KRTSPGVK	IAPFRILEEH
SBDΔ3	FAVAAD	ELKMTLLDLE	NRVALAQRLF	PTEAQAAMEI	AEMKGTCEF-	----------	-----SPGVK	IAPFRILEEH
SBDΔ4	FAVAAD	ELKMTLLDLE	NRVALAQRLF	PTEAQAAMEI	AEMKGTCEF-	----------	-----GSG--	-SGFRILEEH

Figure 5: Diagram of loop deletion design

Map of the α-Chymotrypsin cleaved NPR1-SBD-Nimin-1 fragments sent for n-terminal sequencing.

α-Chymotrypsin cleaves at bulky hydrophobic residues, and cuts the NPR1-SBD at an exposed phenylalanine residue in a flexible loop, cutting the SBD in half. N-terminal sequence indicated by the arrows.

SBDΔ recapitulates the loop deletion by Wang, et al. in the NPR4-SBC.
SBDΔ2-4 make progressively radical deletions in the flexible loop between α2 and α3.
SBDΔ4 removes the entire loop and replaces it with five residues, GSGSG.

subunits. By cutting the loop, α-chymotrypsin digestion may have generated hyper-flexible fragments that could be excluded from the crystal contact site, allowing crystallization.

Based on this hypothesis, I designed three more increasingly radical loop deletions to reduce its interference: SBDΔ2 (aa462-471), SBDΔ3 (aa460-475), and SBDΔ4 (aa460-482, +GSGSG). In SBDΔ4, the entire loop was removed, and replaced with a flexible liner of five residues in order to minimize the loop as much as possible (Fig. 5). I cloned, expressed and purified each of these constructs with Nimin-1 (aa31-72). Only the most radical deletion SBDΔ4 with Nimin-1 α-helix (aa31-72) yielded repeated crystals (Fig. 5):

- Morpheus "G12" SBDΔL/S+ Nimin-1"Sh" and SBDΔ4L/S+Nimin-1 αH. 150:75nL, 74% humidity.

But the crystals diffracted poorly at 30Å. I titrated the individual components of the condition that yielded the first generation of crystals to optimize their behavior, but could not increase diffraction.

Section 5.4: Increasing The Diffraction of NPR1-Nimin-1 Crystals

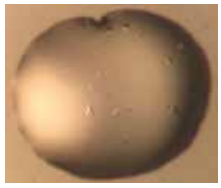
Poor diffraction is the result of insufficiently rigid crystal packing, such that the diffracted reflections are spread out, increasing noise. Crystal packing can be affected by sample purity, flexibility in the protein, choice of crystallization reagent, disturbances during crystal growth, crystal handling during harvesting or by the choice of cryoprotectants. The protein I prepared was free from obvious impurities, so to alter crystal packing and possibly allow the protein subunits to fit together more tightly, I chose to do biased and additive screens, I used the commercially available Hampton additive screen and the "Silver Bullets" screen at a ratio of 90μL Morpheus G12: 10μL of additive. I repeated this ratio with carboxylic acids and halogens

mixes as well. Although I saw that the additives altered crystal morphology, forming larger, flatter and more ragged crystals (Fig. 5). I did not observe an improvement in crystal diffraction.

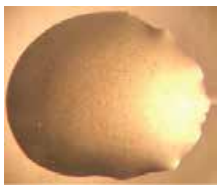
Crystals appear solid to the eye, but actually contain large channels filled with solvent. These channels maintain the crystallized protein in its hydrated form, preserving its semblance to native protein.[74] Although necessary, these large solvent filled channels can reduce rigidity. Dehydration by drawing the solvent out through osmotic pressure may cause the crystal to pack more tightly, increasing resolution. Unfortunately, attempting to dehydrate these crystals by immersion or vapor diffusion into more concentrated cryoprotectant did not improve diffraction.

Section 5.5: Crystallization of the complex via mMBP Fusion Technique

In x-ray crystallography, affinity tags are usually removed after the initial affinity purification step using an engineered cut site for targeted proteolysis. The affinity tag is then purified away before crystallization is attempted. A his-tag is often preferred because it is small enough that proteolysis is not required after affinity purification. For example, Wang et al. used a his-tagged construct of NPR4-SBC in order to refold the protein with SA under denatured conditions before crystallization.[58] Larger affinity tags, such as GST or MBP help solubilize proteins, and sometimes serve as folding chaperones for difficult or disordered proteins. Generally, larger affinity tags reduce the likelihood of crystallization by restricting the possible conformations of the target protein, possibly limiting crystal growth by excluding the conformation required for the asymmetric unit. This conformational heterogeneity can be leveraged in the case of difficult to crystallize small proteins or peptides, however. In this case, an easily crystallizable affinity tag may be fused to the protein target at its N-terminus such that as the affinity tag crystallizes, it incorporates the small protein. MBP or maltose binding protein,



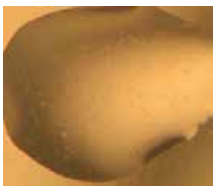
1:4000 α–Chymotrypsin
NPR1-SBDΔ + Nimin-1 (31-72)
Proplex F1
0.1M Sodium Hepes pH 7, 18% w/v PEG 12,000



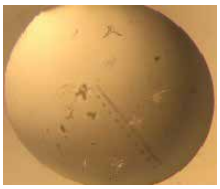
NPR1-SBDΔ
+Nimin-1 (31-72)

NPR1-SBDΔ4
+Nimin-1 (31-72)

12.5% w/v PEG 1000, 3350, MPD
0.03M Halides
0.1Mbicine/TrizmabasepH8.5



NPR1-SBDΔ + Nimin-1 (31-72)
12.5% w/v PEG v1000, 3350, MPD
0.02 Carboxylic Acids
0.1M bicine/Trizma base pH 8.5



Silver Bullets A12
0.25% w/v Benzidine
0.25% w/v Phenylglyoxal monohydrate
0.25% w/v Sulfaguanidine
0.25% w/v Sulfanilamide
0.02 M HEPES sodium pH 6.8

NPR1-SBDΔ4
+Nimin-1 (31-72)

Silver Bullets H2
0.2% w/v Ala-ala,Ala-gly,
Gly-asp, Gly-phe, Ser-Glu
0.02 M HEPES sodium pH 6.8 260

Figure 6: Comparison of crystal morphologies

Comparison of crystal morphology from different droplets to show the effect of the length of Nimin-1, the length of the loop in NPR1-SBD, and the crystallization condition. Silver Bullets additives affect the shape of the NPR1-SBDΔ4.

Crystallization screen, condition, and complex are annotated.

a well-known folding chaperone, may also contribute to the rigidity of the small protein by providing non-specific contacts between the tag and the target. The first x-ray crystal structures of a peptide fused to MBP were published in the early 2000s. Bokhove et al. made several modifications to MBP that increase its expression, reduced surface entropy, and increased the rigidity of its C-terminus.[78] Then in 2017, Jin et al. published several variants of fusion linkers that altered the crystallization of small protein constructs.[79]

I used this construct to make a fusion protein between mMBP and the NPR1-SBD. Using AlphaFold, I chose where to fuse the C-terminal helix of mMBP to the linker such that the linker and the N-terminal helix of NPR1-SBD formed one continuous helix with a minimum of bending (Fig. 7). I did this by discarding the un-conserved tail end of the mMBP gene, trimming mMBP until I felt the helix would be stably folded. Biased by the flexibility demonstrated by the NPR1-Nimin-1 complex in my previous crystallization trials, I chose to link the helices continuously to maintain the rigidity of the complex, rather than allow for random tag-complex packing through a flexible linker. Amino acids form a helical structure where each residue contributes 100° of rotation. Conceivably, the density of the small protein could sterically clash with the tag when fused at the wrong angle. Following the design laid out by Jin et al. I designed a series of linkers between mMBP and NPR1-SBD in addition to fusing their termini directly. These linkers were repeats of one to five alanine residues inserted between the two termini, intended to cover fully 2 rotations of the complex with respect to mMBP (Fig. 7). I cloned the fusion using overlap PCR to form a unified gene. I expressed and purified each of these constructs with Nimin-1 α-helix (aa31-72), Nimin-1 “med”, and Nimin-1 “sh” (Fig. 2). The shortest and longest linkers yielded crystals— that is, the direct fusion of the termini, and the termini spaced by five alanine residues.

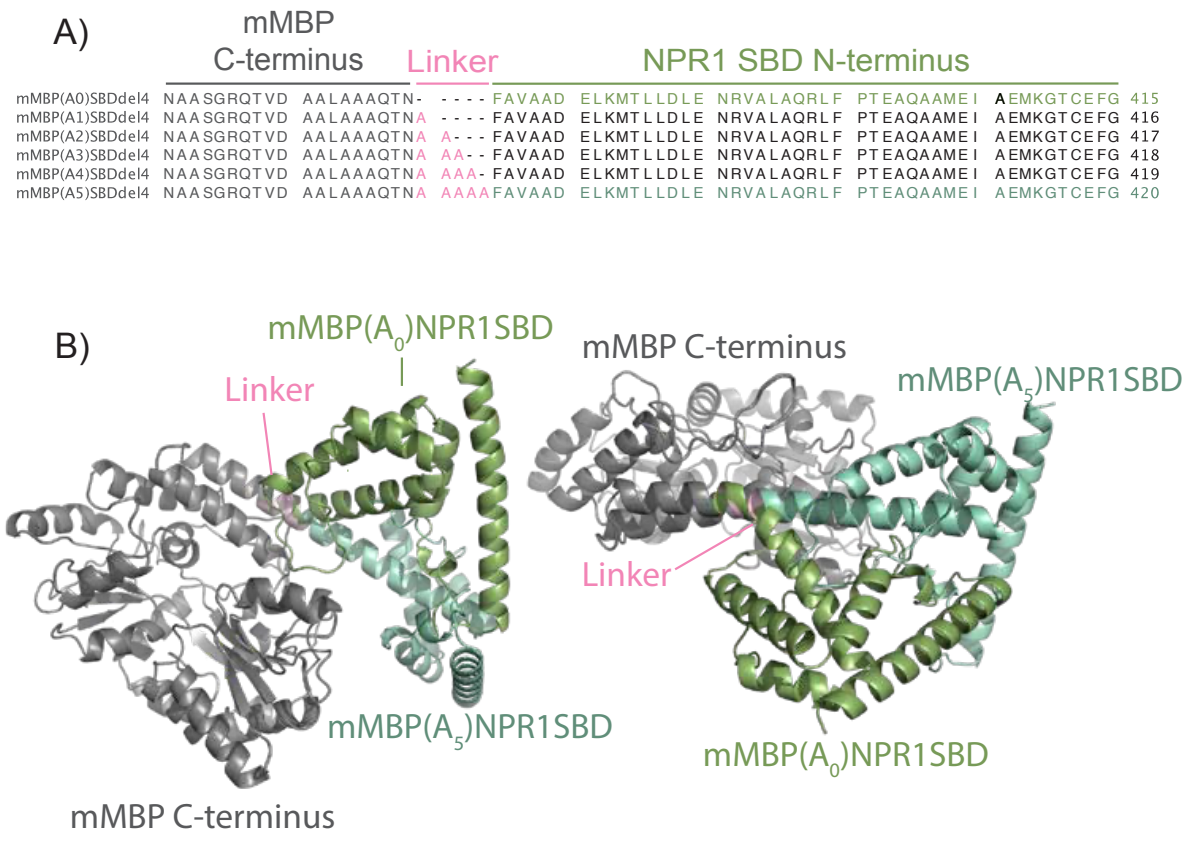


Figure 7: Design of mMBP Fusion constructs

A) Linkers of increasing length were added in between the c-terminal end of mMBP and the n-terminal end of NPR1-SBD. Each residue adds 100º of rotation.
B) Cartoon of AlphaFold prediction of the effect of linker length. mMBP shown in gray, linker residues in pink, shortest linker (A_0) in green, and longest linker (A_5) in blue. The alanine linkers form a continuous helix with the c-terminal helix of mMBP. Although the longest linker does not add great length to the fusion helix, the rotation is more favorable, and the helix is straight, but in the shortest linker is predicted to be bent, indicating possible steric clashes.

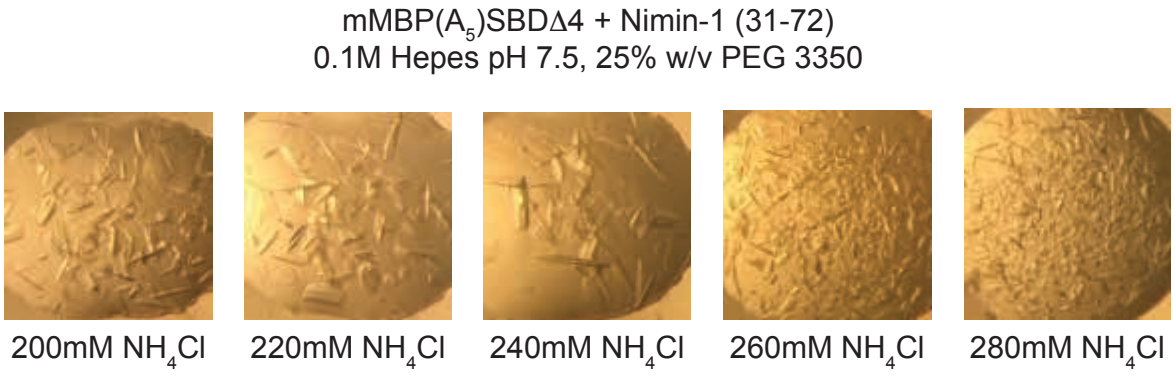


Figure 8: Effect of titration of ammonium chloride on crystal shape and nucleation

Effect of ammonium chloride salt titration on mMBP(A_5)NPR1-SBD + Nimin-1 (31-72) crystals, grown at 4ºC. Increasing the salt reduces initially increases the size of the crystals, making them more rounded. More than 240mM of the salt reduces crystal size, and increases crystal nucleation.

mMBP(A_0)NPR1SBDΔ4+Nimin-1 “Sh” 28mg/mL 150:75nL

- 0.2M Magnesium Chloride, 0.1M Sodium Cacodylate pH 6.5, 20% w/v PEG 1000

But these crystals were not replicable.

The mMBP(A_5)NPR1SBDΔ4+ Nimin-1 (aa31-72) yielded crystals in several conditions:

- 0.2M Ammonium Sulfate, 0.1M Hepes pH 7.5, 25% w/v PEG 3350
- 0.2M Sodium Chloride, 0.1M BIS-TRIS pH 6.5, 25% w/v PEG 3350

These crystals were replicable with diffraction of 6Å.

In order to improve diffraction, I optimized the crystal conditions by titrating the salt, buffer and precipitant. I also optimized the counter-ion by repeating the crystallization condition with a variety of ammonium and chloride salts (Fig. 8). Eventually, this optimization yielded a replicable crystal grown in:

- 0.2M Ammonium Chloride, 0.1M Hepes pH 7.5, 25% w/v PEG 3350

These crystals, when sent to the Gemini beam at the University of California, Berkeley, yielded a data set at 3.3Å, allowing me to solve the structure of NPR1-SBD with Nimin-1 (aa31-72) by molecular replacement with the AlphaFold prediction of the complex. (Fig. 9)

Section 5.6: Discussion and Conclusions

Here, I report the crystallization of the NPR1 C-terminus for the first time using a relatively recent, novel technique. The difficulty of crystallizing the NPR1 C-terminus is likely

due to the flexibility of this domain, as shown by Wang et al.'s hydrogen-deuterium exchange experiments. By minimizing the flexible loop, adding an additional binding partner, Nimin-1, and a crystallizing chaperone, mMBP, I achieved an SBD rigid enough to crystallize. The choice to use a rigid linker appears justified, but Jin et al. includes examples of flexible linkers that allow for target packing against mMBP in useful ways. It remains to be seen if such a technique could be useful for this crystallization system.

The increase in resolution of mMBP(A_5)NPR1-SBD-Nimin-1 (aa31-72) crystals may be attributed to the design of the Gemini beam for use with small crystals as well as the optimization of the crystallization conditions. The initial screen conditions contained ammonium sulfate, which dominates the Hoffmeister series of ions and readily precipitates proteins, as well as sodium chloride, which is much lower on the Hoffmeister series. The 3.3Å crystals grown in ammonium chloride, which is placed lower on the Hoffmeister series than ammonium sulfate but higher than sodium chloride, so perhaps it can be considered an average of the two conditions, optimizing the precipitation of the protein for higher crystal quality. For future structure studies, similar logic for optimization may be helpful.

Section 5.7: Method Details

5.7.1 Molecular Cloning

NPR1 domain inserts were made by PCR and cloned into pET vector using Ligation Independent Cloning (LIC) with N-terminally fused six Histidine (6xHis) tag, mysb tag (MSB) followed by a thrombin (Thr) cut site, DYKDDDDK tag (Flag) and tev cut-site (TEV): 6xHis-MSB-Thr-Flag-TEV-NPR1.

Nimin-1 domains were isolated by PCR, cloned into pET vector with N-terminally fused 6xHis tag, followed by a glutathione-S-transferase (GST) tag, and a TEV cut site using LIC: 6xHis-GST-TEV-Nimin-1.

Loop deletion inserts were ordered from Azenta/Genewiz, and cloned into pET vector using Ligation Independent Cloning (LIC) with N-terminally fused six Histidine (6xHis) tag, mysb tag (MSB) followed by a thrombin (Thr) cut site, DYKDDDDK tag (Flag) and tev cut-site (TEV): 6xHis-MSB-Thr-Flag-TEV-GSGS-SBD.

To generate mMBP fusions, mMBP and NPR1 SBD were fused by overlap PCR, then cloned into pAL vector using LIC with N-terminally fused 6xHis tag, maltose binding protein (MBP), TEV cut site: 6xHis-MBP-TEV-mMBP-SBD.

For DNA extraction, BL21(DE3) *E.Coli* was grown for 16h at 37ºC.

5.7.2: Minimal Domain Pull Down

0.03L cultures of BL21-DE3 *E. Coli* were transformed with His-msb-flag domain constructs and GST-Nimin-1 (aa31-72), grown in LB Miller Broth to OD_{600} 0.8-1.2 at 37ºC, then cooled to 16ºC and incubated for 16h after 500µM IPTG induction. The cells were harvested, resuspended, and lysed in lysis buffer (40mM Tris, pH 8.0, 200mM NaCl, 1mM TCEP) in the presence of protease inhibitors (2 µg/mL leupeptin, 2µg/mL pepstatin, 2µg/mL aprotinin, and 200µM PMSF) using sonication at 10% amplitude, 0.5 on/1s off for one minute. 0.6mg total protein from the soluble cell lysate was adjusted to 120µL and incubated with 30µL of prepared glutathione agarose resin (Thermofisher) for 16h at 4ºC, washed, and the protein was eluted with 10mM glutathione adjusted to pH 8.

5.7.3: Protein Purification

For protein expression, 2-6L of BL21(DE3) *E.Coli* were transformed with the DNA constructs, grown in LB Miller Broth to OD_{600} 0.8-1.2 at 37ºC, then cooled to 16ºC and incubated for 16h after 200µM IPTG induction.

The cells were harvested, resuspended, and lysed in lysis buffer (20mM Tris, pH 8.0, 200mM NaCl, 0.5mM TCEP, 20mM Imidazole) in the presence of protease inhibitors (1 µg/mL leupeptin, 1µg/mL pepstatin, 1µg/mL aprotinin, and 100µM PMSF) using a microfluidizer. The affinity-tagged proteins were isolated from the soluble cell lysate by Ni-NTA agarose (ThermoFisher). The affinity tags were removed by TEV cleavage overnight at 1:10 mass ratio. The complex was then purified by Mono Q chromatography (GE Healthcare). The NaCl eluates were further clarified by passing through Ni-NTA agarose (Thermofisher) and subjected to Superdex 75 (loop deletion complex) or Superdex200 (mMBP fusion) size exclusion chromatography (GE Healthcare). After collection, the SEC peaks were concentrated by ultrafiltration, and flash frozen in liquid nitrogen for further use.

5.7.4: Crystallization

NPR1-SBDΔ crystals were grown at 25ºC by hanging-drop vapor diffusion with 0.150µL of 16.9mg/mL complex mixed with 1:4000 mass ratio α-Chymotrpsin and 0.075µL of reservoir solution containing 0.1M Sodium Hepes pH 7, 20% w/v PEG 8,000; 0.1M Magnesium Acetate tetrahydrate, 0.1M MES pH 6.5, 10% w/v PEG 10,000; or 0.1M Sodium Hepes pH 7, 18% w/v PEG 12,000.

NPR1-SBDΔ+Nimin-1 (41-63) crystals were grown at 4ºC by hanging-drop vapor diffusion with 0.150µL of 20mg/mL complex mixed with 0.075µL of reservoir solution

containing 0.1M Carboxylic acids mix, 0.1M buffer system #3, pH 8.5, 37.5% precipitant mix #4 at 74% humidity; or 0.09M Halogens mix, 0.1M Buffer system #3, pH 8.5, 37.5% precipitant mix #4.

NPR1-SBDΔ4+Nimin-1 (aa31-72) crystals were grown at 4ºC by hanging-drop vapor diffusion with 0.150µL of 20mg/mL complex mixed with 0.075µL of reservoir solution containing 0.1M Carboxylic acids mix, 0.1M buffer system #3, pH 8.5, 37.5% precipitant mix #4 at 74% humidity. Crystals were harvested after two weeks. Cryoprotection was provided by the crystallization condition.

mMBP(A_5)NPR1-SBDΔ4 crystals were grown at 4ºC by hanging drop vapor diffusion with 0.225µL of 56mg/mL complex, supplemented to 10mM maltose, and mixed with 0.075µL of reservoir solution containing 0.2M Ammonium Sulfate, 0.1M Hepes pH 7.5, 25% w/v PEG 3350; 0.2M Sodium Chloride, 0.1M BIS-TRIS pH 6.5, 25% w/v PEG 3350' or 0.2M Ammonium Chloride, 0.1M Hepes pH 7.5, 25% w/v PEG 3350. The crystals were harvested after two weeks and cryoprotected by serial transfer to 27% sucrose with mother liquor.

Section 5.7.5: Biased Screens and Dehydration

Biased screens were set up with 90µL of 0.1M Carboxylic acids mix, 0.1M buffer system #3, pH 8.5, 37.5% precipitant mix #4 at 74% humidity with 10µL of other crystallization screens. Crystals were grown at 4ºC by hanging-drop vapor diffusion with 0.150µL of 20mg/mL complex mixed with 0.075µL of reservoir solution containing the biased condition.

Additive screen were set up with a ratio of 90µL of 0.1M Carboxylic acids mix, 0.1M buffer system #3, pH 8.5, 37.5% precipitant mix #4 at 74% humidity with 10µL of Silver Bullets and Additive Screen (Hampton Research). Crystals were grown at 4ºC by hanging-drop vapor

diffusion with 0.150µL of 20mg/mL complex mixed with 0.075µL of reservoir solution containing the additive.

Section 5.7.6: Protease digestion and N-terminal sequencing

The initial protease digestion was performed with 9µg of complex incubated with 1:200 mass ratio of proteases for 1.5h @ 4ºC using Hampton Reasearch HR-492 kit. 9 8µg samples of complex were incubated with α-Chymotrypsin (serially diluted at a ratio of 1:2, beginning at 1:30, and ending at 1:3,840) for 1h @ 25ºC.

To N-terminally sequence the cut site, a 34µg sample was digested by 1:4000 mass ratio α-chymotrypsin for 3.5h. The sample was run on a 9% SDS-PAGE gel, separated into an 18kD and 15kD band, then transferred to a PVDF membrane, stained with Coomassie, then cut and shipped to Tufts University.

Section 5.7.7: Data Collection and Structure Determination

After collecting a native dataset at Advanced Light Source Beamline 8.2.1 or 2.2.1 X-ray diffraction data were integrated and scaled with HKL2000 package (Otwinowski and Minor, 1997). The structure was solved by molecular replacement and phaser from the PHENIX suite of programs using molecular replacement with AlphaFold predicted complex. All structural models were built, refined and rebuilt using COOT (Emsley et al., 2010) and PHENIX (Adams et al., 2002). PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) was used to generate figures.

Data Collection	mMBP(A_5)NPR1-SBD Nimin-1 (31-72)
Space Group	C 2 2 21
Cell Dimensions	
a, b, c (Å)	118.483, 144.926, 145.329
α, β, γ (°)	90, 90, 90
Wavelength	
Resolution	3.54-3.33
R_{meas}; R_{pim}	0.761; 0.354
$I/\sigma I$	0.285
Completeness %	97.92 (90.81)
Redundancy	2.3 (2.1)
Refinement	
Resolution (Å)	3.33
No. reflections	16411 (2787)
R_{work}/R_{free}	0.2343 (0.3085) / 0.3084 (0.3798)
No. atoms	8467
Protein	8377
Ligand	90
Water	0
B-Factors	
protein	74.23
Ligand	65.8
Water	—
R.m.s. deviations	
Bond Lengths (Å)	0.004
Bond angles (°)	0.83
CC1/2	0.833

Figure 9: Crystallization table

Table with crystallographic information.

Chapter 6: Characterization of NPR1-Nimin-1 Interaction

Section 6.1: Negative Control of the Plant Immune Response

Because it is necessary for survival, the plant immune response allows trade-offs in the short-term that are deleterious if maintained indefinitely. PR gene transcription is a massive metabolic commitment by the plant, comprising 5-10% of leaf weight during the immune response.[39] Accordingly, plants that express PR genes constitutively are stunted, though the severity of this phenotype depends on the mutant background. An Arabidopsis mutant, *Snc2-1D npr1* shows severely reduced growth while being hyper-defended, but *cpr1, cpr5, and cpr6* plants only showed reduced fitness in the field, being unaffected under ideal laboratory conditions.[57,80] A transgenic plant encoding salicylate hydroxylase, *NahG*, had increased leaf biomass compared to the wild-type.[80] Later studies found that SA directly binds to a phosphatase that regulates auxin efflux, inhibiting growth.[81,82] Because defense and growth are mutually exclusive, attenuation of PR-gene transcription is necessary for long-term growth.

Plant hormone cross-talk partly controls salicylic acid signaling. During the hypersensitive response, the lesions represent concentric rings of salicylic acid and jasmonate signaling, in which the presence of jasmonate determines the boundaries of the necrotic lesions.[47] Jasmonate is a defense signal that regulates the plant's response to necrotrophs and broadly opposes salicylic acid signaling. By interacting with transcription factors, jasmonate represses ICS1 gene expression to suppresses salicylic acid synthesis.[34] In order to attenuate SAR more broadly, it has now been established that SA itself initiates a negative feedback loop that represses PR gene expression through a negative regulator that directly interacts with NPR1.

Section 6.2: Nimin-1 Mediated Negative Regulation of Plant Immune Response

Soon after the discovery of NPR1's role in initiating PR gene expression, a German group, Weigel et al., discovered a negative regulator of PR gene expression during a yeast-2-hybrid screen to find NPR1 interactors in 2001.[83] They identified three proteins with 12-17% sequence identity and named them "Nim-1 interactor" or "Nimin". These proteins are Nimin-1, Nimin-2, and Nimin-3. The same year, Ohta et al. described a conserved motif, called the ethylene amphipathic response or EAR motif, that recruits TOPLESS, a general transcription repressor involved in plant development.[84] This EAR motif L/FDLNL/F(x)P closely resembled the LDLNL motif found in the C-terminus of all three Nimin family members. By 2005, they had confirmed the role of these proteins as suppressors of PR gene expression.[60] Over-expressing wild type Nimin-1 reduced RNA transcripts of PR-1, PR-2, and PR-5 significantly. The Nimin-1 over-expressing line of Arabidopsis was also more susceptible to infection than wild-type, showing colony formation on par with the SAR impaired *Npr1-1*.[60] In 2013, Hermann et al. found that Nimin-1 and Nimin-2 genes are expressed in response to salicylic acid, neatly defining a negative feedback loop in which a set of SA responsive genes recruit transcription repressors to NPR1 regulated genes.[85] Despite persuasive genetic evidence indicating that NPRs are key proteins involved in plant immunity and years of research, basic questions about the mechanism of NPR regulation of the plant immune response lack conclusive data.

Section 6.2: Molecular Basis for Nimin-1 Regulation of NPR1-mediated Gene Transcription

The Nimin (NIM1 interacting) family of proteins interact with NPR1 to recruit a transcription repressor to NPR1 regulated gene promoters. Weigel et al. and Herman et al. admirably defined the Nimin-NPR1 interaction within their capabilities as plant geneticists. They

identified a key motif DXFFK, flanked by a common acidic and a basic region, as well as the C-terminal EAR motif. Deleting the DXFFK region abolished the interaction with NPR1.[60] Mutating the two phenylalanines in the DXFFK motif to serines also abolished NPR1 interaction, and when over-expressed, caused PR-genes to over-accumulate compared to the wild-type in which PR gene expression decreased over time.[60] Using the yeast-2-hybrid system, they attempted to identify the interacting domain of NPR1. They found Nimins interacted with wild type NPR1, but not the C150Y mutant *npr1-2,* which may be misfolded, nor a truncated form of NPR1 lacking the an ankyrin repeat and the C-terminus (Δ377-593 and Δ392-593).[60] This was the first indication that the C-terminal domain of NPR1 was sufficient to interact with Nimin-1 and Nimin-2. Yeast-2-hybrid is a less reliable method, however, because the selection pressure generates false positives, and because of potential false negatives due to the protein-protein interactions taking place outside of their canonical environments. Without structural data, this group could not rationally design their constructs, and accordingly, their data contains some anomalies.

Section 6.3: Crystal Structure of NPR1-Nimin-1 Complex

The asymmetric unit of the mMBP(A_5)NPR1-SBD is made up of two copies of mMBP(A_5)NPR1-SBD with two copies of Nimin-1. The mMBP proteins sandwich together at the core of the asymmetric unit with the two SBD domains participating in crystal contacts with SBD domains from other units. This is in line with the participation of Nimin-1 (aa31-72) playing a role in crystallization despite the resolvable density being much shorter. Although the complex was prepared with 10mM maltose, a maltetraose occupies the mMBP fusion protein

binding pocket, likely an endogenous metabolite included during protein folding and retained throughout purification and crystallization.

There is a clear distinction between mMBP and NPR1-SBD+Nimin-1, as predicted by AlphaFold. In fact, the AlphaFold prediction was suitable for molecular replacement. The two structures are quite close. The conserved, minimal helix of Nimin-1 fits into a groove formed by two alpha helices and the SBC of NPR1. (Fig. 10) The NPR1-SBD forms a hand that wraps around Nimin-1. Compared to the minimal domain required for binding salicylic acid in NPR4, the "SBC", the NPR1-SBD contains an extra two helices that form the clamp. Nimin-1 and NPR1 share a greasy, hydrophobic core at the center of this groove, where the two phenylalanine residues identified by Weigel et al. are inserted— explaining why substituting them with serine abolished binding. Within this groove, both NPR1 and Nimin-1 contribute hydrophobic residues that pack against each other like the teeth of a key within the tumblers of a lock. (Fig. 11) Nimin-1 contributes the bulky residues I48, F49, F50, and I53 while NPR1 surrounds these residues with smaller hydrophobic side chains, like A528. Flanking the greasy core are a series of oppositely charged residues that could participate in hydrogen bonds. Nimin-1's E42 and E43 coordinate with R538 and K542 from NPR1, while R60 and R63 coordinate with D428 and D420 from NPR1.

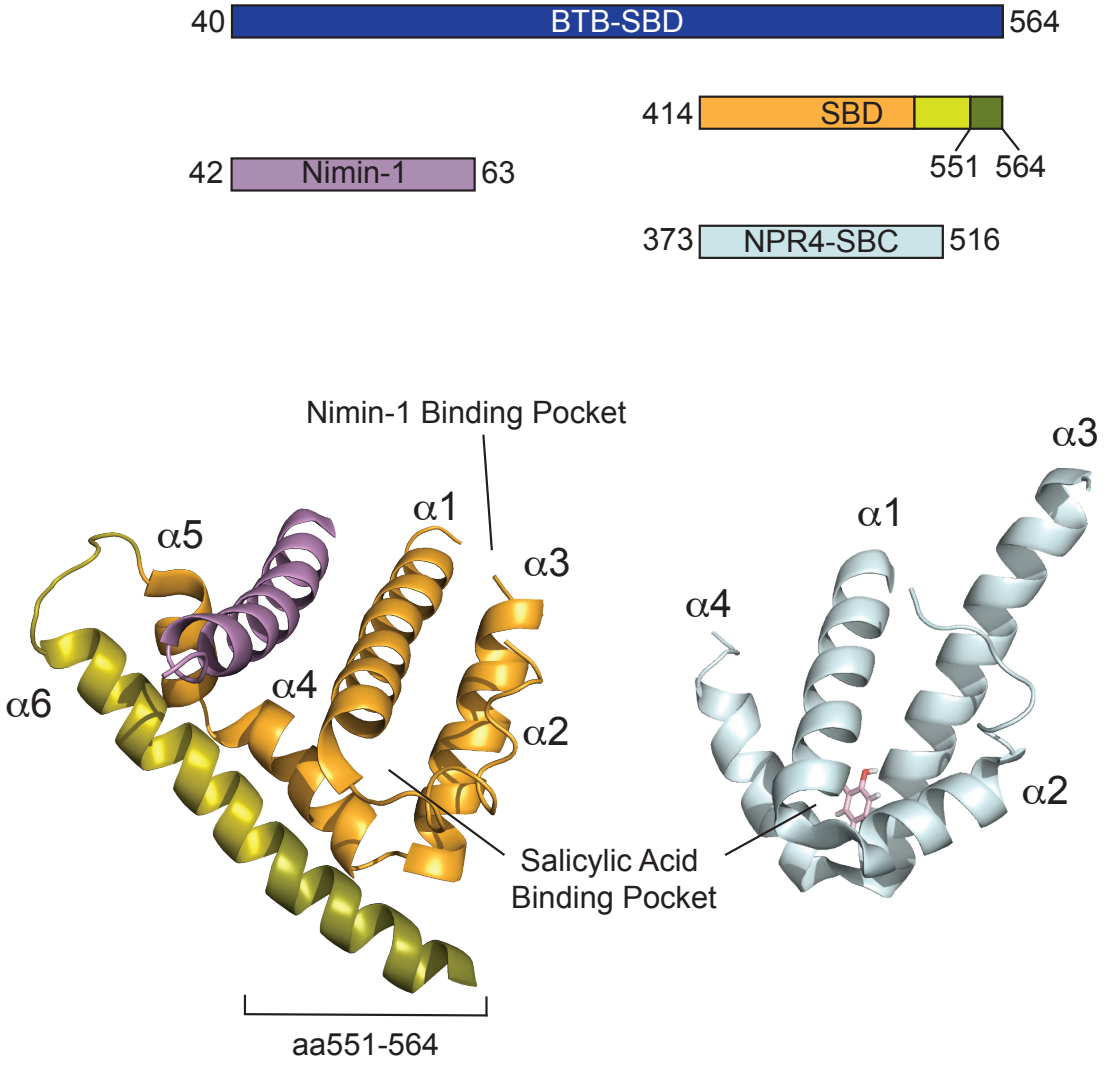


Figure 10: Schematic of NPR1 domains and comparison of NPR1-SBD structure with NPR4-SBC

Schematic of NPR1 domains. NPR1 BTB-SBD shown in blue, Nimin-1 shown in light purple, NPR1-SBD shown in a gradient of orange-lime-olive: aa414-530 shown in orange, aa530-551 shown in lime, and aa551-564 shown in olive. NPR4-SBC shown in light blue, with SA colored in pink. α-helices 1-6 labeled.

Cartoon figure of the SBD in complex with Nimin-1 show α5 and α6 wrapping around Nimin-1, which takes the form of a long α-helix. NPR1-SBD in complex with Nimin-1 preserves the shape of the NPR4-SBC closely. NPR4-SBC and NPR1-SBD both conserve α1-4, with α5 and α6 of NPR1-SBD required for Nimin-1 binding. The structure indicates separate binding pockets.

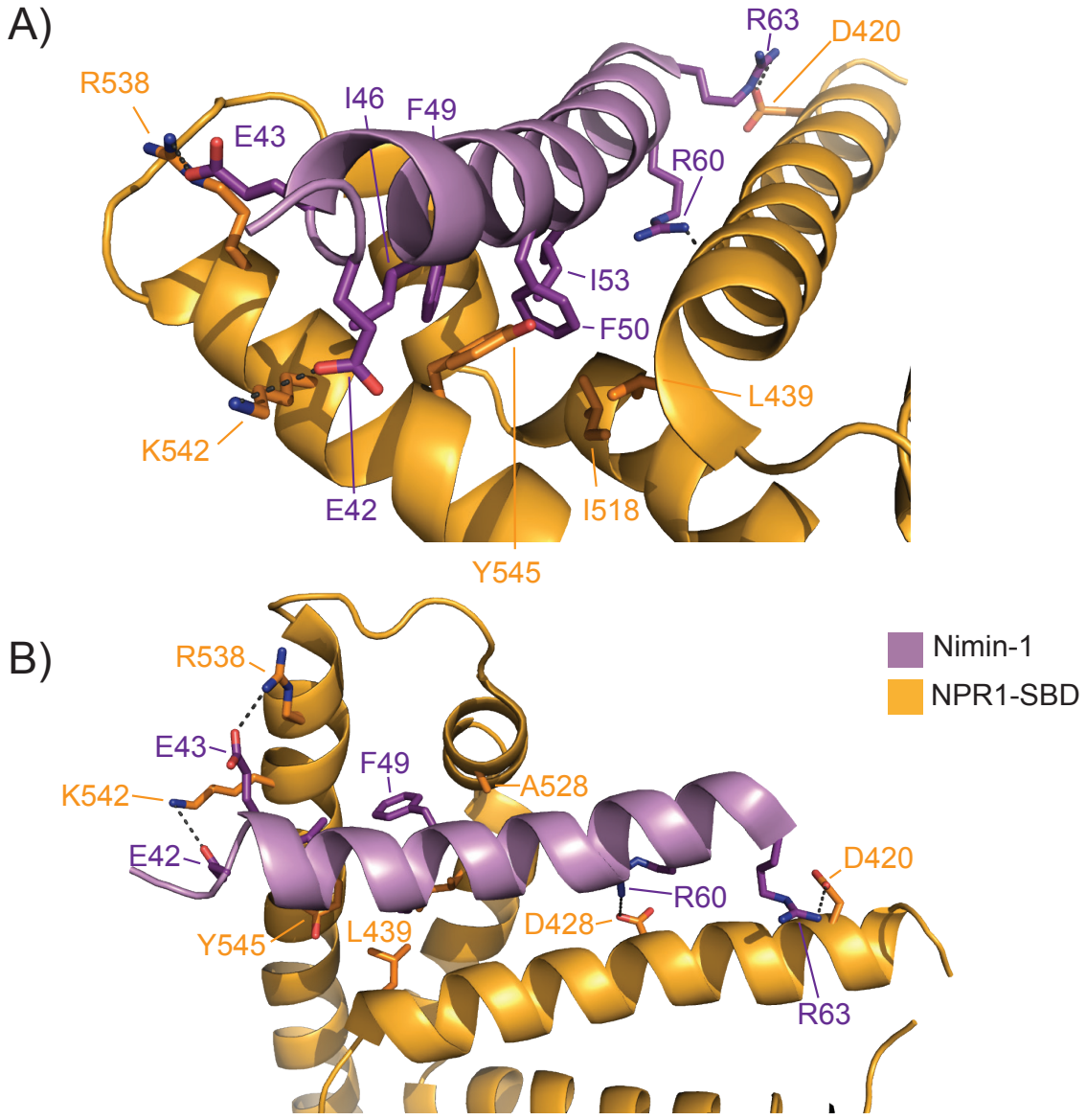


Figure 11: Cartoon schematic showing key side chains.

NPR1 colored in orange. Nimin-1 is colored in purple. NPR1 Side-chains are labeled in orange. Nimin-1 side chains are colored in purple.

A) View of Nimin-1 side chains. Nimin-1 projects several bulky hydrophobic residues into the cleft formed by α1, α4, α5, and α6. These are flanked by charged residues, n-terminally, by two glutamine side chains, and c-terminally, by two arginine residues that form contacts with oppositely charged NPR1 side chains.
B) View of NPR1 side chains. NPR1 contributes several smaller hydrophobic side chains, L439, I518, A528, and one bulky one, Y545. Additionally, A lysine, K542, arginine, R538, and two aspartic acid residues, D428 and D420, formed salt bridges with Nimin-1 residues.

Section 6.5: Contribution from Key Residues

To test the participation of key Nimin-1 residues to binding, I made E42A, E43A, I46A, F49A, F50A, I53A, R60A, and R63A mutants. This made the bulky residues smaller, and removed the charges from the glutamine and the arginine residues. Then, I verified the effect on binding by using GST-tagged Nimin-1 to pull-down wild-type NPR1. (Fig. 12) Interestingly, mutating the charged residues had little effect, but removing the bulky hydrophobic side-chains dramatically reduced binding. It is likely that these mutations would reflect changes in affinity, as the flanking residues in Nimin-2 may contribute to the difference in affinity between it and Nimin-1. To evaluate the contribution of NPR1 residues, I mutated the salt-bridge residues (D428A/K542A) to non-charged residues, and increased the size of the hydrophobic residues, I518Y and A528F. I then performed a pull-down with wild-type Nimin-1. (Fig. 12) Mutating both salt-bridges at the same time augmented the effect on binding. Increasing the size of the hydrophobic side-chains made the greatest difference, indicating that these side-chains clashed sterically with Nimin-1.

Section 6.6: Biophysical Characterization of Nimin Family affinity for NPR1

As discussed in the previous chapter, I identified the minimal interacting domains for NPR1 and Nimin-1. Using Amplified Luminescence Proximity Homogenous Assay (ALPHA), I determined the relative affinity of minimal Nimin-1 peptides for NPR1 by comparing the IC_{50}s of unlabeled peptides of varying lengths competing with biotinylated peptide. (Fig. 13 and 14) Similarly to the minimal domain data already shown, all peptides were able to compete, but the affinity reduced with size, such that the longest peptide competed better than the shortest. (Fig. 14) While the flanking residues outside of the core hydrophobic motif may not be required for

Nimin-1 binding, they may strengthen the interaction. While the shortest peptide had the lowest IC_{50}, it corresponds to the maximum resolvable density within the Nimin-1 binding groove in NPR1 even though the complex was prepared with a longer version (aa31-72). The shorter peptide would be expected to crystallize more readily, but in this case, the longer version may have stabilized the asymmetric unit in some way.

I also determined the relative affinity of Nimin-1, Nimin-2, and Nimin-3 using the same method. (Fig. 15) Nimin-2 had the highest IC_{50}, followed by Nimin-1, then Nimin-3. Nimin-3 diverges from the canonical NPR1 binding motif the most out of the Nimin family, substituting an LY pair for the double phenylalanine pair seen in Nimin-1 and Nimin-2. This likely reduces the affinity because leucine is smaller than phenylalanine, and tyrosine would introduce a polar group into the heart of the hydrophobic core. While Nimin-2 conserves the hydrophobic patch, the flanking residues differ significantly which could affect affinity, but the structure would need to solved to place sequence differences in context. For example, despite the different potencies the Nimin-2 peptide lacks one of the salt bridges NPR1 forms with Nimin-1 through R60.

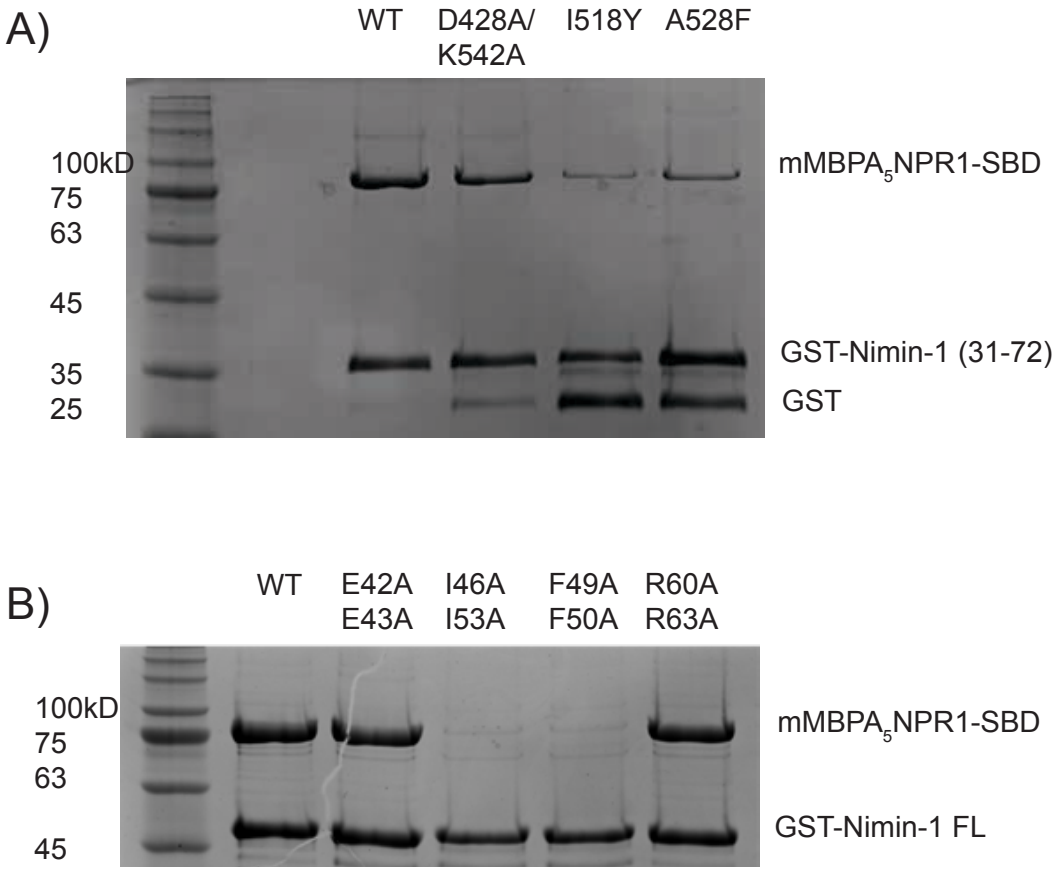


Figure 12: Pulldowns showing the impact of mutating key residues.

GST-Nimin-1 pulls down mMBP(A_5)NPR1-SBD construct using glutathione agarose, eluted with glutathione.

A) Mutating NPR1 residues to evaluate the impact on the NPR1-Nimin-1 interaction. D428A, K542A reduces interaction slightly, while I518Y and A528F abolishes the interaction completely, likely through the steric hindrance of bulky residues clashing with the Nimin-1 NPR1 binding motif. Without binding to NPR1, the short Nimin-1 helix becomes unstable, generating significant GST fragments.
B) Impact of mutating Nimin-1 on NPR1-Nimin-1 interaction. Mutating charged residues (E42A, E43A and R60A, R63A) has little effect, but reducing the size of the hydrophobic residues (I46A, I53A, and F49A, F50A) abrogates the interaction,

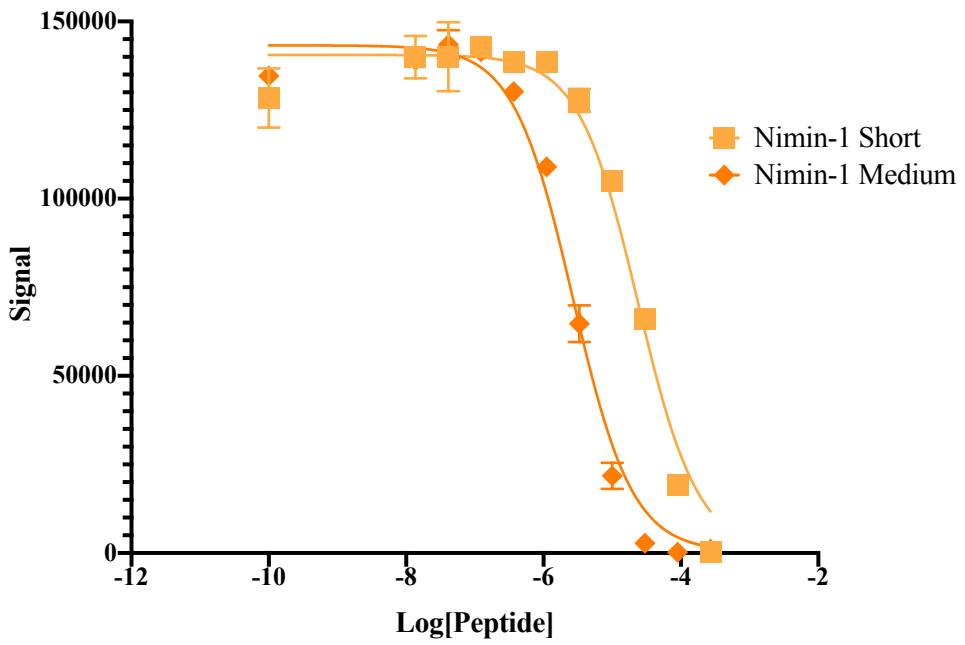
```
biotin-EDEEEEEEKKIDTFFKLIKHYQEARKRRREELA
     Medium: EEKKIDTFFKLIKHYQEARKRRREELA
      Short:   KKIDTFFKLIKHYQEARKRRR


Nimin-1: biotin-EDEEEEEEKKIDTFFKLIKHYQEARKRRREELA
Nimin-1:               EEKKIDTFFKLIKHYQEARKRRREELA
Nimin-2:          VVRTVTEEEVDEFFKILRRVHVATRTVAKVNG
Nimin-3:          EKEDDEEEKMEKLYTVLKNAREMRKYVNSSMEKK
```

Figure 13: Design of Nimin peptides for IC_{50} experiments

“Medium” (aa42-72) includes glutamine residues excluded by the “short” peptide (aa44-63).

Nimin-2 peptide preserves the core DxFFK motif, but omits two lysine and several arginine residues found in Nimin-1.Nimin-3 replaces the DxFFK motif with a smaller amino acid (L) and a more polar amino acid (Y).

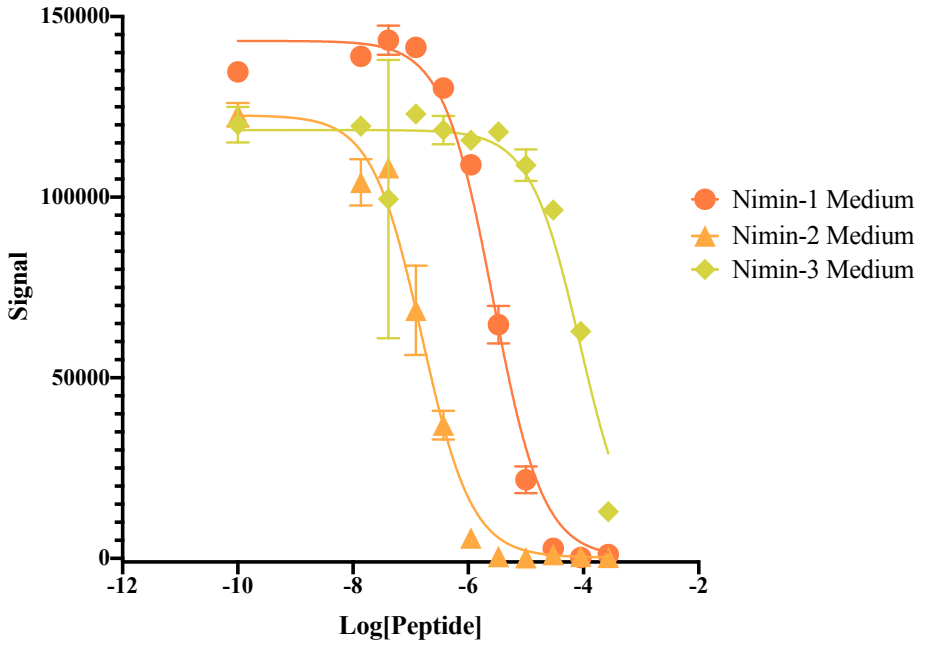


Peptide	IC_{50} (µM)
Medium	2.7
Short	24

Figure 14: Competition assay between short and medium Nimin-1 peptides

Alpha competition data showing the relative affinity of Nimin-1 "short" and Nimin-1 "medium". The longer peptide competes better than the shorter peptide, 2.7µM vs. 24µM.

MBP-NPR1 was mixed with synthetic biotinylated Nimin-1 peptide (Vivitide), MBP acceptor beads (Perkin Elmer) and Streptavidin donor beads (Perkin Elmer) to 0.45nM (MBP-NPR1), 37nM (biotinylated Nimin-1 peptide)



Peptide	IC_{50} (µM)
Nimin-1	0.16
Nimin-2	2.7
Nimin-3	87

Figure 15: Competition assay showing the relative affinity of Nimin-1, Nimin-2, and Nimin-3

Nimin-2 (0.16µM) competes with higher potency than Nimin-1 (2.7µM), or Nimin-3 (87µM).

MBP-NPR1 was mixed with synthetic biotinylated Nimin-1 peptide (Vivitide), MBP acceptor beads (Perkin Elmer) and Streptavidin donor beads (Perkin Elmer) to 0.45nM (MBP-NPR1), 37nM (biotinylated Nimin-1 peptide)
.

Section 6.6: Discussion

Nimin-1 regulation of PR gene expression is necessary for attenuation of SAR, an important role for long-term plant health. I report here the first crystal structure describing the molecular basis for Nimin-1's interaction with NPR1. I have identified key residues for binding, showing that hydrophobic contributions to a conserved greasy core shared between Nimin-1 and NPR1 flanked by salt bridges underlies their interaction. For the first time, I have quantified the relative affinity of different family members for NPR1. In light of these findings, work remains to be done to clarify the roles of the Nimin protein family. The different relative affinities may have a functional correlation to the differences in expression shown by Hermann et al.'s work. Nimin-1 and Nimin-2 are expressed in response to SA, but at different time scales. Perhaps their differing affinities for NPR1 play a role in fine-tuning regulation of PR gene expression at different stages of the immune response. While Nimin-3's affinity for NPR1 is much lower, it is expressed constitutively, and may interact with NPR4 instead. Solving the structures of other Nimin family members with both NPR1 and NPR4 SBDs and characterizing their interactions more fully may harmonize the data I have presented in this work.

Section 6.7: Method Details

6.7.1: Data Collection and Structure Determination

After collecting a native dataset at Advanced Light Source Beamline 8.2.1 or 2.2.1 X-ray diffraction data were integrated and scaled with HKL2000 package (Otwinowski and Minor, 1997). The structure was solved by molecular replacement and phaser from the PHENIX suite of programs using molecular replacement with AlphaFold predicted complex. All structural models were built, refined and rebuilt using COOT (Emsley et al., 2010) and PHENIX (Adams et al.,

2002). PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC) was used to generate figures.

Section 6.7.2: Molecular Cloning

Nimin-1 mutations were introduced by site-directed mutagenesis using the NEBuilder HiFi DNA assembly kit, and cloned into pAL vector using Ligation independent cloning (LIC) with N-terminally fused glutathione-s-transferase (GST) tag, and a TEV cut site to generate: 6xHis-GST-TEV-Nimin-1.

To make mMBP fusion mutants, NPR1-SBD mutant inserts: D428A/K542A, I518Y, A528F were ordered from Azenta/Genewiz, then fused to mMBP by overlap PCR then cloned into pET vector using Ligation Independent Cloning (LIC) with N-terminally fused six Histidine (6xHis) tag, mysb tag (MSB) followed by a thrombin (Thr) cut site, DYKDDDDK tag (Flag) and tev cut-site to generate: (6xHis-MSB-Thr-Flag-TEV-mMBPSBD.

For DNA extraction, BL21(DE3) *E.Coli* was grown for 16h at 37ºC.

Section 6.7.3: Pull Down Assays

0.03L cultures of BL21-DE3 *E. Coli* were transformed with his-msb-flag mutant constructs and/or GST-Nimin-1 (aa31-72) mutant constructs, grown in LB Miller Broth to OD_{600} 0.8-1.2 at 37ºC, then cooled to 16ºC and incubated for 16h after 500µM IPTG induction. The cells were harvested, resuspended, and lysed in lysis buffer (40mM Tris, pH 8.0, 200mM NaCl, 1mM TCEP) in the presence of protease inhibitors (2 µg/mL leupeptin, 2µg/mL pepstatin,

2µg/mL aprotinin, and 200µM PMSF) by sonication at 10% amplitude, 0.5 on/1s off for one minute. 0.6mg total protein from the soluble cell lysate was adjusted to 120µL and incubated with 30µL of prepared glutathione agarose resin (Thermofisher) for 16h at 4ºC, washed, and the protein was eluted with 10mM glutathione adjusted to pH 8.

Section 6.7.4: Amplified Luminescence Homogenous Assay

AlphaScreen Assays for determining and measuring protein-protein interacts were performed using EnSpire reader (PerkinElmer).

MBP-NPR1 was mixed with synthetic biotinylated Nimin-1 peptide (Vivitide), MBP acceptor beads (Perkin Elmer) and Streptavidin donor beads (Perkin Elmer) to 0.45nM (MBP-NPR1), 37nM (Nimin-1, Nimin-2, and Nimin-3 peptide) and 5µg/mL (beads IC_{50} was calculated using nonlinear regression: log(inhibitor) vs. response—variable slope (4 parameters) constrained to the averaged signal for 4 single protein well controls. Assay was performed in 20mM Hepes pH 7.5, 200mM NaCl, 0.1% BSA, 0.02%Tween20, 0.5mM TCEP.

Chapter 7: Identifying an Allosteric Pathway Between the Salicylic Acid and Nimin-1 Binding Pockets

Section 7.1: Mechanism of SA binding by NPR Proteins

The C-terminal domain of NPR proteins binds salicylic acid. The structure solved by Wang et al shows SA buried deeply within a pocket formed by a four-helix bundle of four α-helices, αSC1-4. Between αSC2 and αSC3 there is a flexible loop, the site of the short internal deletion Δ450-455. This loop contains the proposed EAR-like motif proposed to underlie NPR4's negative regulation of PR gene expression. It is also completely absent from the electron density of the structure. The four helices form two V-shapes, like nested nutcrackers where SA binds at the space between the hinges of the nutcrackers. In this pocket, a strictly conserved arginine, R419, forms two hydrogen bonds to the carboxylic acid group of salicylic acid, and when this residue is mutated, SA binding is abolished. NPR1 possesses an analogous arginine residue, and the group was able to confirm that both NPR1 and NPR4 can bind to SA, but it appears that they interact with SA differently. The group showed that a very small fraction of NPR1 was actually occupied by SA in their sample, 0.02% compared to 8% of the NPR4 sample but this difference had yet to be explained.[58]

Although it is known that the SBC of NPR1 and NPR4 interacts with SA, how this influences the plant immune response mechanistically is poorly understood. The lack of a solvent accessible channel in the pocket indicates that the apo form has a radically different conformation, one that would allow for diffusion of SA into the pocket. Whatever interactions the C-terminal domain engages in prior to SA binding must be disrupted once the C-terminus assumes an SA bound conformation. In the case of NPR4, this is thought to lift its repression of

either NPR1's function, or gene transcription directly.[57] In light of the SBC structure, several mutations in NPR1 found in SAR impaired plants map to key SA binding residues, indicating that NPR1 sensing SA is important for its function in SAR. It may be that SA-bound NPR1 is able to interact with general transcription factors to induce gene expression. This interaction would then be disrupted by Nimin binding to NPR1 during the negative feedback loop controlling SAR. While this hypothesis conveniently explains Nimin-1's repression of NPR1, the mechanism of Nimin-1 competition with salicylic acid is not known.

Section 7.2: SA Competes with Nimin-1 to bind NPR1

As discussed previously, the C-terminal domain of NPR1 is sufficient to bind Nimin-1, corroborated by the AlphaFold prediction of the complex. Because both bind to the same domain even though SA binding to NPR1 is required for PR gene expression while Nimin-1 binding represses PR gene expression, I hypothesized that the SA and Nimin-1 might compete to bind NPR1, providing a mechanistic explanation for negative regulation of PR gene expression by Nimin-1. To test this, I performed a pull-down assay between full-length NPR1 and Nimin-1 (aa31-72) with or without SA, showing that SA disrupts NPR1 binding to Nimin-1. (Fig. 16)

Next, I quantified this interaction by performing a competition ALPHA assay with NPR1 and Nimin-1 immobilized on the beads while titrating increasing concentration of SA, finding an IC_{50} of 7μM. (Fig. 17)

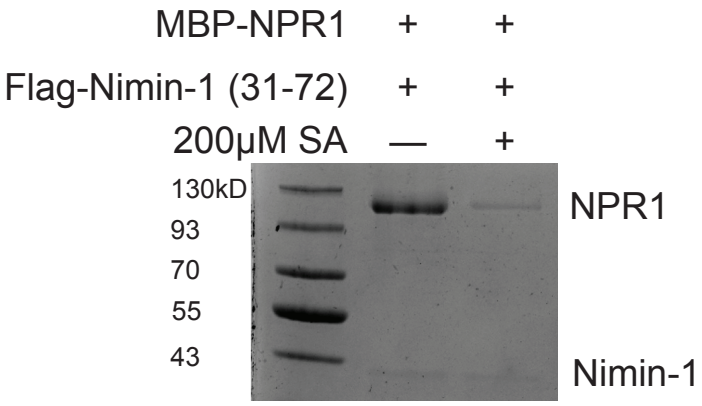
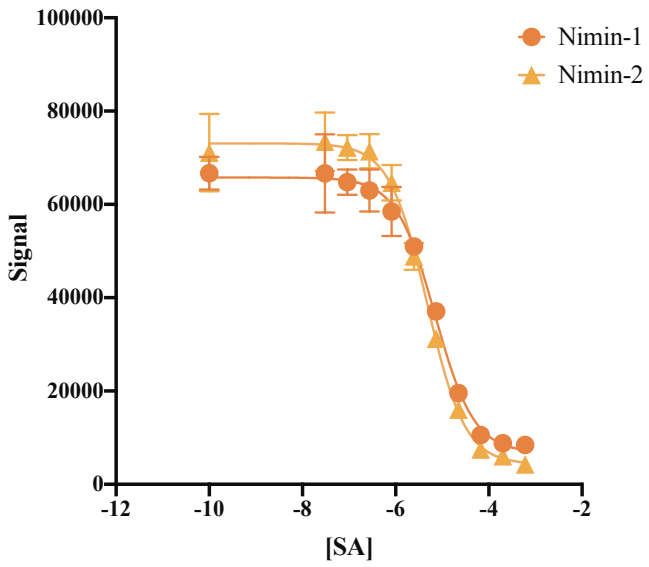


Figure 16: Pulldown showing SA and Nimin-1 competition

His-msb-flag-Nimin-1 (31-72) pulling down MBP-NPR1 (FL) using flag resin, eluted with flag peptide.

Nimin-1 pulls down NPR1 except in the presence of 200M SA



Biotinylated Peptide	SA IC_{50} (μM)
Nimin-1	7μM
Nimin-2	4.8μM

Figure 17: Competition between salicylic acid and biotinylated Nimin-1 and Nimin-2 peptide

Alpha competition assay titrating salicylic acid against biotinylated Nimin-1 and Nimin-2 peptide indicating potencies of 7μM and 4.8μM.

MBP-NPR1 was mixed with synthetic biotinylated Nimin-1 and Nimin-2 peptide (Vivitide), MBP acceptor beads (Perkin Elmer) and Streptavidin donor beads (Perkin Elmer) to 0.45nM (MBP-NPR1), 37nM (biotinylated Nimin-1 peptide).

Section 7.3: Crystal Structure of the NPR1 Salicylic Acid Binding Domain (SBD)

By crystallizing and solving the structure of the NPR1-Nimin-1 complex, I visualized the C-terminus of NPR1 for the first time, allowing comparison to the SA-bound NPR4-SBC structure solved by Wang et al. Interestingly, the structure shows clearly that Nimin-1 and SA do not bind to the same pocket, being separated by at least 13Å. The presence of separate pockets explains the difference between the experimentally determined boundaries of the SBC, and the bioinformatically determined domain boundaries of the SBD as SA would not stabilize the helices required for binding Nimin-1. Separate pockets also indicates that Nimin-1 competes with SA indirectly, implying an allosteric pathway between the pockets.

Unexpectedly, however, the structure of the SBC is preserved within the Nimin-1 bound SBD almost unchanged, despite evident changes in NPR1-mediated function in the plant as a result of Nimin-1 binding. There are few places of difference to identify the allosteric pathway between the Nimin-1 and SA binding pockets. One place of interaction between the the extra helices of the SBD and the SBC is at the C-terminus of α6: α5 and the N-terminal end of α6 wrap around Nimin-1 while the C-terminus of α6 packs against α4 at the base of the salicylic acid binding pocket. (Fig. 18) Although the AlphaFold prediction of the complex and the solved structure align closely with an RMSD of 0.819, one place of difference between the structure and the prediction is this C-terminal end of α6. In the prediction, this helix faces away from the SBC, but in the solved structure, a conserved hydrophobic patch packs against the base of the salicylic-acid binding pocket. (Fig. 18) Here, a proline, P340, is deflected downwards in the Nimin-1 bound structure, packing against a phenylalanine, F556. Another phenylalanine, F399 at the base of the SA binding pocket faces outward in the SA bound structure, but tucks inward in the Nimin-1 bound structure. Both P340 and F399 appear to interact with F556 which is flanked by

L552 and

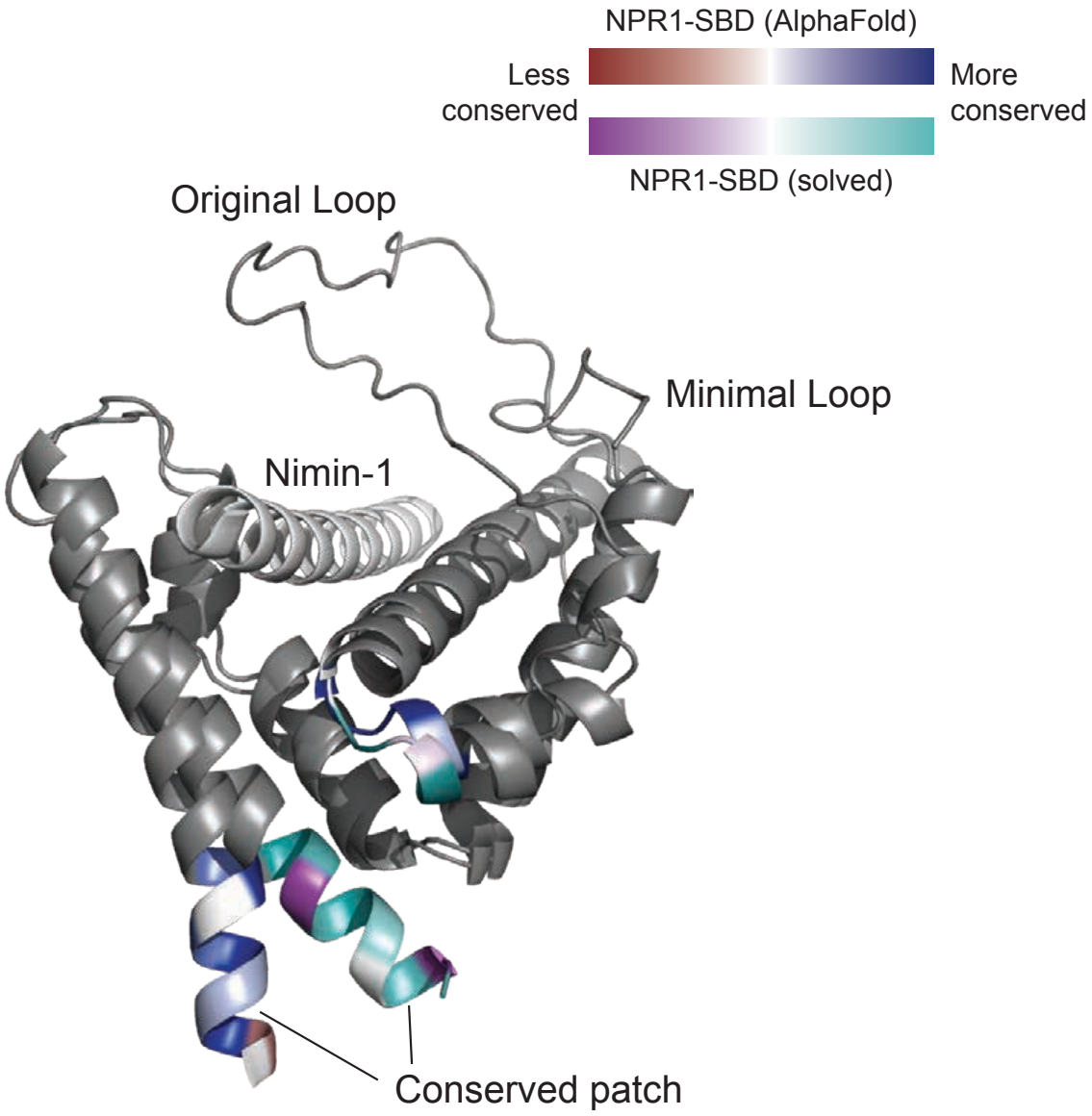


Figure 18: Alignment showing the packing of a conserved patch against the salicylic acid binding pocket

Alignment of AlphaFold prediction with solved structure, colored by conservation score (computed by Consurf). A section of α6 and a loop at the base of the SA binding pocket are both conserved, but they are not predicted to interact by AlphaFold, although they pack against each other in the solved structure.

Prediction and solved strucutre colored in gray, and red to dark blue spectrum. Solved SBD colored in dark gray, and magenta to cyan spectrum. Nimin-1 colored in white.

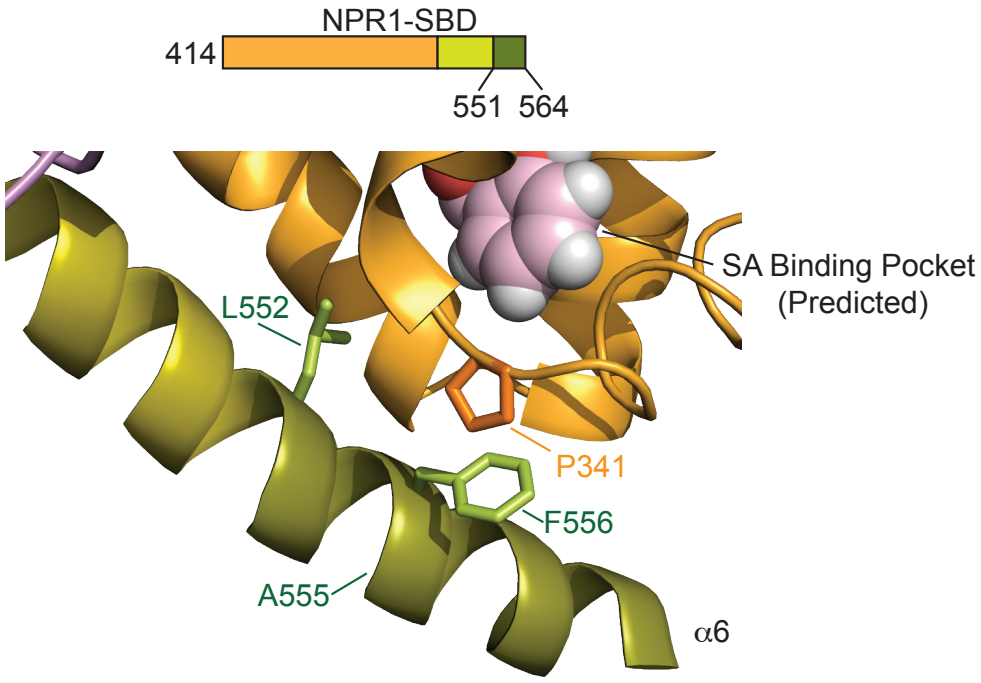


Figure 19: Diagram of potential allosteric pathway

NPR1 shown in an orange-lime-olive spectrum: aa414-530 colored in orange, aa530-551 colored in lime, aa551-564 colored in olive. Predicted placement of SA based on alignment to NPR4-SBC shown in pink. Side chain (P340) in loop between α1 and α2 colored orange. Side chains in α6 colored in olive (L552, A555, and F556). F556 appears to pack against a proline at the base of salicylic acid binding pocket, potentially stabilized by L552 and A555.

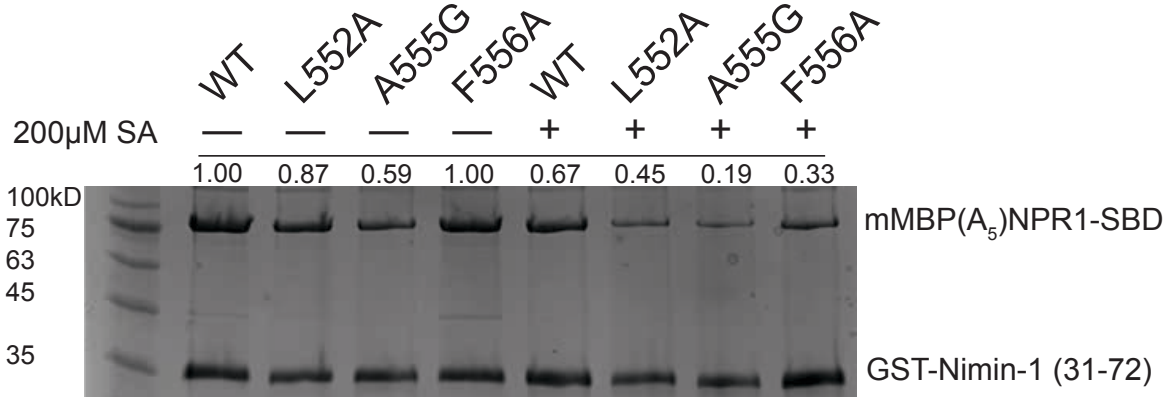
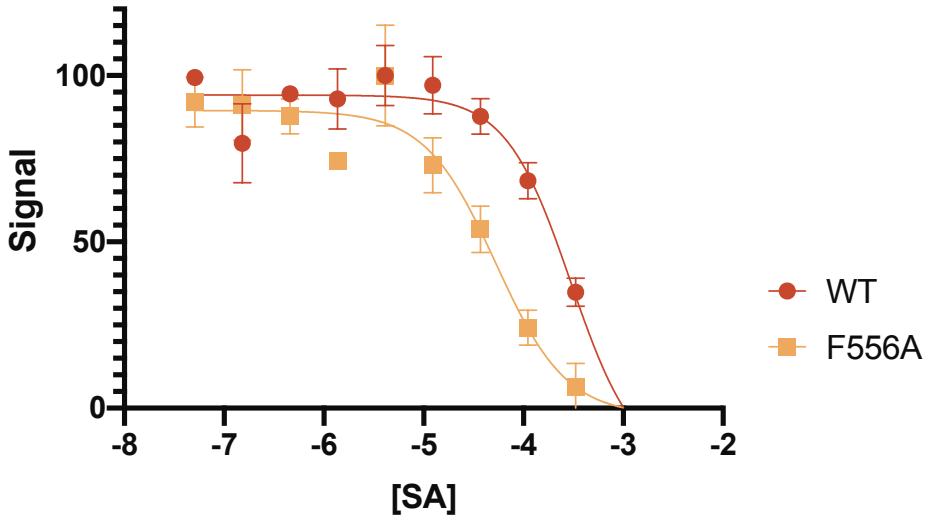


Figure 20: Pulldown evaluating mutations in a potential allosteric pathway

Pulldown showing the effect of mutating potential residues in the allosteric pathway on SA competition with Nimin-1.

GST-Nimin-1 (31-72) pulls down mMBP(A_5)NPR1-SBD using glutathione resin, eluted by glutathione, with or without 200μM SA. Quantification of mMBP(A_5)NPR1-SBD normalized to GST-Nimin-1 band (ImageJ) placed above the lane.

L552A and A555G both destabilize binding to Nimin-1, commensurate with making α6 more flexible. F556A has no effect on binding to Nimin-1, but in the 200μM SA lane, appears to



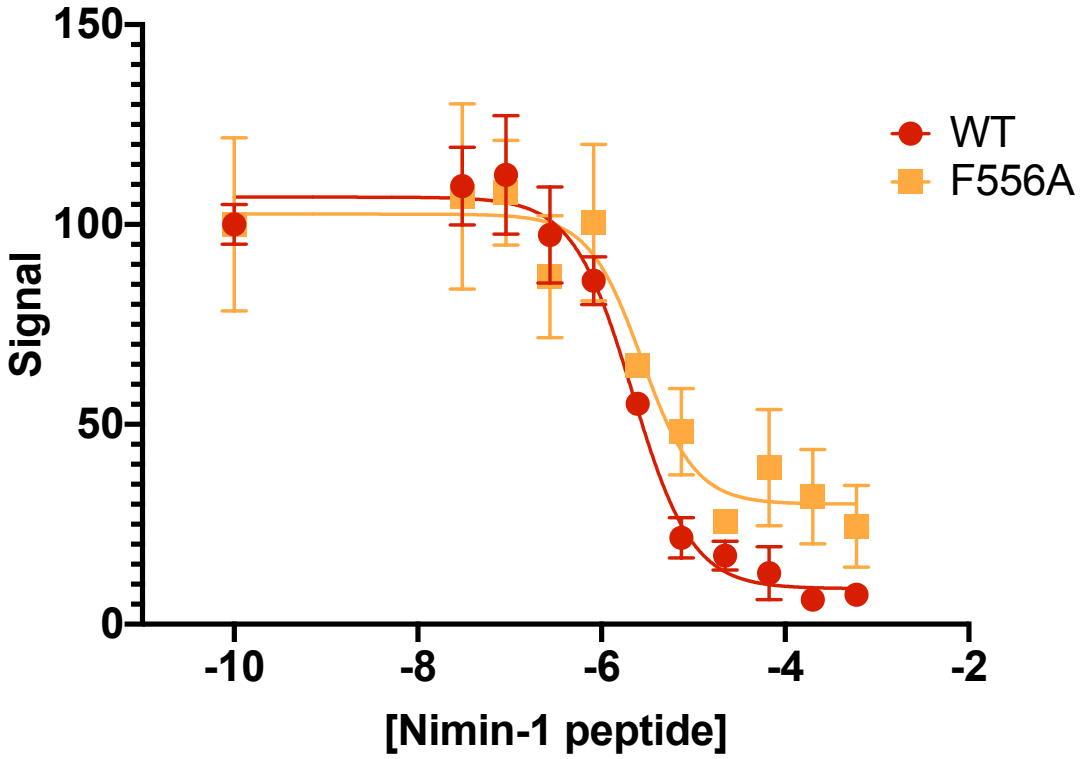
mMBP(A_5)NPR1-SBD	SA IC_{50} (µM)
WT	297.9
F556A	0.5

Figure 21: Competition assay showing effect of F556A mutation of potency of SA

Alpha assay titrating SA against GST-Nimin-1 (31-72) and mMBP(A_5)NPR1-SBD immobilized on anti-MBP and anti-GST beads, normalized to 100% of signal.

Mutating F556A increases the potency of SA against the NPR1-Nimin-1 complex by two orders of magnitude.

mMBP(A_5)NPR1-SBD (WT or F556A) co-expressed with GST-Nimin-1 (aa31-72) were mixed with MBP acceptor beads and GST donor beads to a final concentration of 9nM (complex) and 5µg/mL (beads).



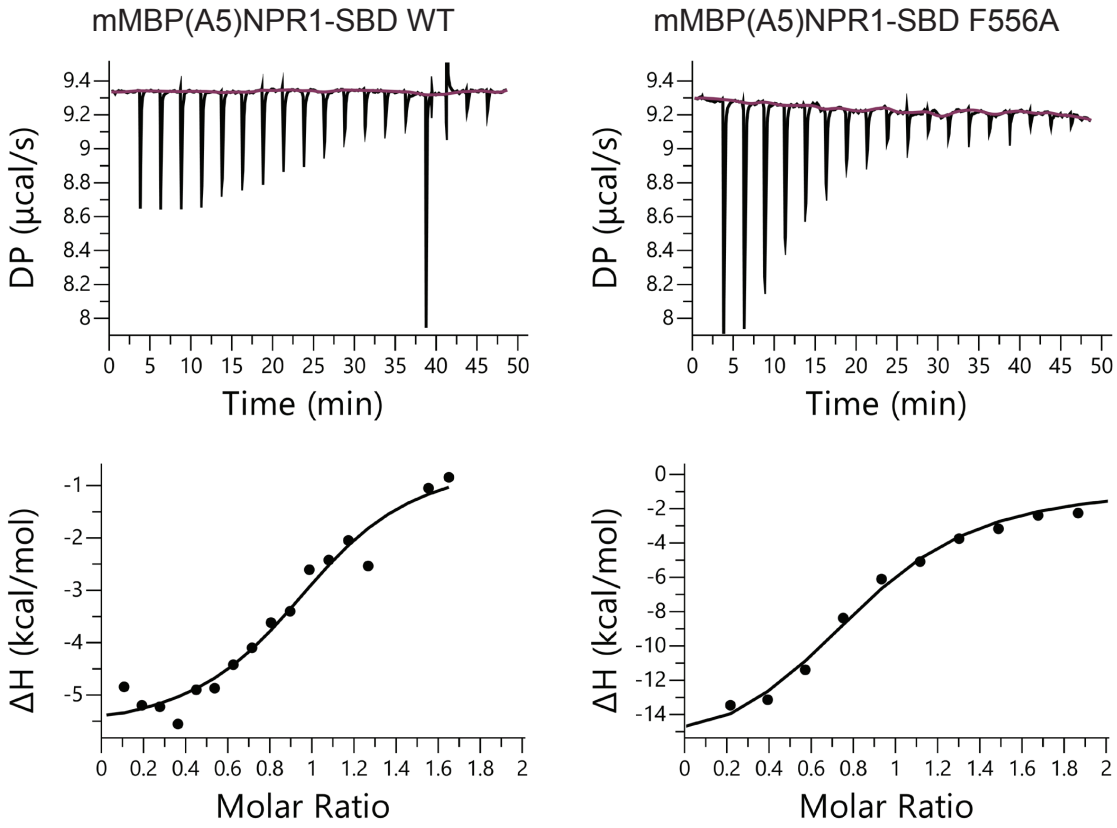
mMBP(A_5)NPR1-SBD	Nimin-1 Peptide IC_{50} (µM)
WT	2.2
F556A	2.7

Figure 22: Competition assay showing the effect of F556A mutation on affinity for Nimin-1

Alpha competition assay titrating biotinylated Nimin-1 peptide against GST-Nimin-1 (31-72) and mMBP(A_5)NPR1-SBD immobilized on anti-MBP and anti-GST beads, normalized to 100% of signal.

F556A mutation does not appear to affect the potency of Nimin-1 peptide greatly.

mMBP(A_5)NPR1-SBD (WT or F556A) co-expressed with GST-Nimin-1 (aa31-72) were mixed with MBP acceptor beads and GST donor beads to a final concentration of 9nM (complex) and 5µg/mL (beads).



mMBP(A_5)NPR1-SBD	SA K_D (µM)
WT	5.2 ± 3.1
F556A	4.9± 1.5

Figure 23: Effect of F556A mutation on NPR1-SBD affinity for SA

Isothermal Titration Calorimetry assay evaluating the affinity of NPR1-SBD for salicylic acid, showing that F556A mutation and WT NPR1-SBD have similar affinities for salicylic acid.

mMBP(A_5)NPR1-SBD WT (43.2µM) and mMBP(A_5)NPR1-SBD F556A (32.6µM) were mixed with 18 4s 2µL injections of 600µM SA spaced by 150s.

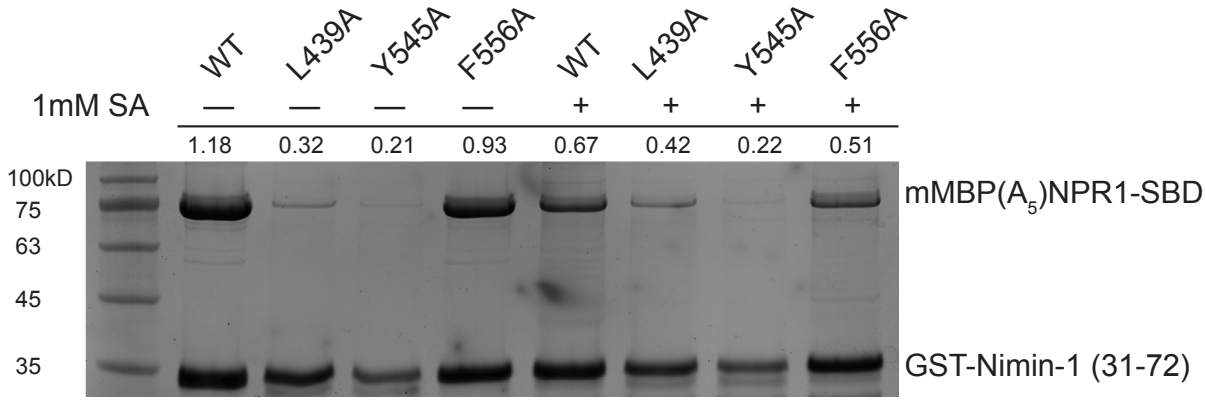


Figure 24: Pulldown showing effect of mutations in alternate allosteric pathway

Pulldown of mMBP(A_5)NPR1-SBD by GST-Nimin-1 (31-72), showing the effect of potential residues in an alternate allosteric pathway. Quantification of mMBP(A_5)NPR1-SBD band normalized to GST-Nimin-1 band placed above the lane.

Mutating L439A and Y545A significantly affect affinity for Nimin-1 binding. The residual level of binding, however, appears unaffected by salicylic acid.

Pulldown by GST-Nimin-1 using glutathione agarose resin, eluted by glutathione.

A555, possibly providing the allosteric pathway between the two pockets. (Fig. 19) Because altering P340 and F399 could affect affinity for SA, I chose to mutate F556, and L552 to alanine to reduce α6 interaction with P340. I also mutated A555 to glycine to perhaps destabilize the rigidity of the alpha-helix and weaken the allosteric competition.

Section 7.3: Evaluating contribution of point mutations on the allosteric pathway

To test the effect of L552A, A555G, and F556A mutation of SA competition, I performed a pull-down assay using the wild-type Nimin-1 to pull-down WT or mutant NPR1. (Fig. 20) Using ImageJ, I quantified the results, finding that both L552A and A555G reduce NPR1 affinity for Nimin-1. Far from reducing the competition by SA, however, the F556A mutation counter-intuitively appeared to increase its potency. To follow up on these results, I tried to quantify this difference in potency using an ALPHA competition assay with SA as the titrant, and found that F556A mutation does increase the potency of SA competition. (Fig. 21) I also verified that the F556A mutation does not affect NPR1-SBD affinity for Nimin-1 using an ALPHA competition assay with unlabeled peptide, showing similar IC_{50} values (Fig. 22). While SA binding to the SBC does not stabilize α5 and α6, Nimin-1 binding to the SBD stabilizes the entire domain, perhaps locking the SBC in a conformation without solvent access. F556A is able to influence this conformation, potentially by allowing the SA binding pocket to flex so that SA can access the pocket.

Concerned that this mutation could affect the affinity of NPR1 for SA, I measured the K_d using isothermal titration calorimetry, or ITC. (Fig. 23) Titrating SA into F556A and WT samples showed similar affinity for SA, meaning that the F556A mutation solely affects the allosteric pathway.

between the pockets. While these results may seem counter-intuitive, this method of evaluating the effect of the mutation on competition is backwards to the competition observed *in planta.* During the plant immune response, NPR1 binds to SA first, before Nimin-1 competition disrupts SA binding after it is expressed in response to SA. In these IC_{50} experiments, NPR1 binds Nimin-1 first. Conceivably, the F556A mutation could in actuality decrease the potency of Nimin-1, were the converse experiment performed.

So far, I have only identified residues contributing to the allosteric pathway in α6 near the SA binding pocket. Because the allosteric pathway involves both pockets, I theorized that L439 and Y545, residues that pack against the Nimin-1 binding groove, might also participate. I mutated these to alanine residues and performed a pull-down with WT Nimin-1 with or without SA. (Fig. 24) I found that mutating both residues significantly affected the affinity for Nimin-1, abolishing Nimin-1 binding in Y545A, and nearly abolishing it with regard to L439A. Without SA, the L439A mutant still binds a residual level of Nimin-1, and retains this binding in the presence of 1mM SA, indicating that mutating this residue not only alters the affinity for Nimin-1 but also disrupts competition by SA.

Section 7.4 Discussion and Conclusions

Here, I describe the structure of the NPR1-SBC for the first time, and interrogate the effect of SA on the NPR1-SBD-Nimin-1 complex. I show novel data indicating that Nimin-1 and SA compete with each other to bind NPR1 and that this competition is allosteric. I further identify key residues participating in this pathway, and show that one mutation, F556A, affects the potency of SA but not its binding affinity. I also identified L439 which both reduces the affinity for Nimin-1 and abolishes competition with SA. There may be a two pronged pathway,

one at the shared helix between the SA binding pocket and the Nimin-1 binding pocket, and one through α6. While the shared helix affects the affinity for Nimin-1, the pathway through α6 does not.

Section 7.5: Method Details

7.5.1: Molecular Cloning

To make mMBP fusion mutants, NPR1-SBD mutant inserts: L552A, A555G, F556A, L439A, Y454A were ordered from Azenta/Genewiz, then fused to mMBP by overlap PCR then cloned into pET vector using Ligation Independent Cloning (LIC) with N-terminally fused six Histidine (6xHis) tag, mysb tag (MSB) followed by a thrombin (Thr) cut site, DYKDDDDK tag (Flag) and tev cut-site to generate: (6xHis-MSB-Thr-Flag-TEV-mMBPSBD.

For DNA extraction, BL21(DE3) *E.Coli* was grown for 16h at 37ºC.

Synthetic peptides were ordered from Vivitide.

7.5.2: Conservation Analysis

Conservation analysis was performed using Consurf server maintained by the Edmond J. Safra Center for Bioinformatics at Tel Aviv University using the default parameters: Neighbor Joining with ML distance for the phylogenetic progress; Bayesian calculation with best fit model of substitution for the conservation scores; The MSA was built using MAFFT, and homologes collected from UNIREF90 using HMMER search algorithm with an E-value cut off of 0.0001.

7.5.3: Pulldown Assays

0.03L cultures of BL21-DE3 *E. Coli* were transformed with his-msb-flag mutant constructs and/or GST-Nimin-1 (aa31-72) mutant constructs, grown in LB Miller Broth to OD_{600} 0.8-1.2 at 37ºC, then cooled to 16ºC and incubated for 16h after 500µM IPTG induction. The cells were harvested, resuspended, and lysed in lysis buffer (40mM Tris, pH 8.0, 200mM NaCl, 1mM TCEP) in the presence of protease inhibitors (2 µg/mL leupeptin, 2µg/mL pepstatin, 2µg/mL aprotinin, and 200µM PMSF) by sonication at 10% amplitude, 0.5 on/1s off for one minute. 0.6mg total protein from the soluble cell lysate was adjusted to 120µL and incubated with 30µL of prepared glutathione agarose resin (Thermofisher) with or without 0.2mM SA for 16h at 4ºC, washed, then the protein was eluted with 10mM glutathione adjusted to pH 8.

Quantification was performed in ImageJ using a single gel.

7.5.4: Amplified Homogenous Luminescence Proximity Assay

AlphaScreen Assays for determining and measuring protein-protein interacts were performed using EnSpire reader (PerkinElmer).

MBP-NPR1 was mixed with synthetic biotinylated Nimin-1 and Nimin-2 peptides (Vivitide), MBP acceptor beads (Perkin Elmer) and Streptavidin donor beads (Perkin Elmer) to 0.45nM (MBP-NPR1), 37nM (peptide) and 5µg/mL (beads), titrating salicylic acid from 600µM to 30.48nM via 1:3 dilution. IC_{50} was calculated using nonlinear regression: log(inhibitor) vs. response—variable slope (4 parameters) constrained to the averaged signal for 4 single protein well controls. Assay was performed in 20mM Hepes pH 7.5, 200mM NaCl, 0.1% BSA, 0.02%Tween20, 0.5mM TCEP.

mMBP(A_5)NPR1-SBD (WT or F556A) co-expressed with GST-Nimin-1 (aa31-72) were mixed with MBP acceptor beads and GST donor beads to a final concentration of 9nM (complex) and 5µg/mL (beads), titrating SA from 1mM to 152nM via 1:3 titration, or synthetic

biotinylated Nimin-1 peptide (Vivitide) from 60μM to 15.2nM via 1:3 dilution. Values were normalized to 100% of signal, and IC_{50} was calculated using nonlinear curve regression log(inhibitor)vs. response—variable slope (4 parameters). Assay was performed in 50mM Hepes pH 7.5, 200mM NaCl, 0.1%BSA, 0.5mM TCEP, 0.02%Tween.

7.6.4: Isothermal Titration Calorimetry

Affinity for SA was determined by isothermal titration calorimetry using Malvern Pananalytical Microcal PEAQ-ITC. 150μL samples of mMBP(A_5)NPR1-SBD WT (43.2μM) and mMBP(A_5)NPR1-SBD F556A (32.6μM) in 20mM Hepes pH7.5, 200mM NaCl, 0.5mM TCEP were mixed with 18 4s 2μL injections of 600μM SA spaced by 150s.

Chapter 8: Summary and Future Directions

Section 8.1: Summary of Results

The plant immune response relies heavily on the transcription of pathogenesis related genes, initiated by perception of the key defense signal salicylic acid by its receptor, NPR1. NPR1 interacts with transcription factors in the nucleus to initiate gene expression through an unknown mechanism. The widescale expression of PR genes is a metabolic burden, however, leading to various fitness costs if improperly regulated. PR gene expression is controlled by Nimin-1, which is under the control of an NPR1 regulated promoter and negatively regulates PR gene expression through a poorly understood mechanism.

In this work, I have described the interaction between NPR1 and Nimin-1 in much greater detail. I quantified the relative affinity of Nimin family members for NPR1, identified the minimal interacting domains of each protein, as well as the major contributing residues. I purified, crystallized, and solved the structure of the NPR1-Nimin-1 complex using a relatively novel technique, mMBP fusion, revealing the mechanism of the NPR1-Nimin-1 interaction. Additionally, I showed that Nimin-1 and SA compete to bind NPR1, providing a mechanistic explanation for Nimin-1 regulation of PR gene transcription. The NPR1-Nimin-1 structure indicates that this competition is indirect, with Nimin-1 and SA binding to separate pockets, meaning that Nimin-1 allosterically inhibits the SA-bound conformation. I then identified two possible allosteric pathways between the Nimin-1 and SA binding pockets, and identified a key residue, F556, that solely affects the potency of SA. Taken together, these results reveal the mechanism of Nimin-1 regulation of NPR1 in the plant defense response.

Section 8.2: Nimin-1 as a potential regulator of NPR1 in transgenic crops

Although we rely on plants and plant materials for food, energy and construction, 40% of crops are lost each year to pathogens.[2] Current strategies involve the use of expensive pesticides, which are not available to all farmers, and which can engender resistance if not used in sufficient quantities. Over-expression of both NPR1 and selected PR genes has been explored as a transgenic strategy for developing highly resistant crops. Over-expression of specific PR genes can confer high degrees of resistance to some pathogens, but not all, depending on the anti-fungal or anti-viral activity of the specific gene.[39,86] Over-expression of NPR1 leads to a variety of fitness costs, including reduced yield due to high basal SA levels or constitutive PR gene expression. Careful dosage or control of NPR1 expression has been shown to mediate these fitness costs.[86] Regulation by Nimin-1 may also be leveraged to mediate the downsides associated with NPR1 expression, particularly the constitutive expression of PR genes. If Nimin-1 can be over-expressed without affecting the timing of its expression, then crops may be more resistant to infection without sacrificing long term growth. In the future, more studies on the regulation of Nimin-1 expression in response to SA, or tuning Nimin-1 affinity for NPR1 to increase its potency against SA could alleviate the burden of NPR1 over-expression in transgenic crops.

Section 8.3: NPR1-SBD as a model for novel aspirin binding proteins

Low dose aspirin is now used widely as a prophylactic anti-thrombotic agent to prevent heart attacks and stroke. Longitudinal studies of low dose aspirin patients are widely available, and some effects are not fully explained by the covalent acetylation of COX1/2, aspirin's known mechanism of action.[16] These findings, along with the observation that aspirin quickly metabolizes to salicylic acid in the body, suggests that there may be unknown salicylic acid

targets. Early in NPR research, NPR1 was compared to Iκβ and NFκβ due to the similarity in their ankyrin repeat domains with one paper showing that aspirin can inhibit NFκβ activation, but only at millimolar concentrations.[87,88] Structural studies of the SA bound NPR1 and NPR4 may inform the search for these novel salicylic acid binding targets.

Section 8.4: Repurposing of the NPR1-Nimin-1-SA System

The regulation of the NPR1 SBD by both a small molecule and a small peptide is unique. Outside of agricultural or human pharmacological applications, this system may be leveraged as a novel affinity tag system, as a FRET system regulated by SA, or in other contexts where controlling protein interactions using small molecules is useful. The NPR1-SBD is a small domain, only fifteen kilodaltons. Although it is difficult to express by itself, when fused to mMBP, its behavior improves significantly, becoming much more stable. The small helix of Nimin-1 that binds NPR1 can be both synthetically derived or genetically encoded, and SA is a cheap, robust and widely available chemical, likely to be well-tolerated even in animal models. Although Nimin-1 affinity for NPR1 is in the μM range, targeted improvements to the NPR1-Nimin-1 interface may increase this affinity to the range of other common affinity tags. The potential of this system for general laboratory repurposing is exciting, and its full applications remain to be seen.

Section 8.4: Final Thoughts

Pharmacology and pharmacy have their etymological roots in the pre-Greek word *pharmakon*, from which overlapping terms for medicine, poison, and potion were derived.

Herbal medicinal compounds precisely embody this tripartite nature: the molecules that begin as endogenous poisons developed against insects become effective therapies in humans only at the correct dose, a magic which has only become science in recent history. Witchcraft in the western tradition, however, carries a double legacy, not only of poisons and potions, but also revolt against a supposed divine order. Pharmacology today lives out this rebellion against nature in its rejection of the destiny of the human body to age, sicken, and decay. We have drafted plants into this revolution not only as the source of medicinal compounds, but for food, energy, and shelter. NPRs as a subject of enquiry lies uniquely at this intersection of medicine and agriculture. Human domestication of crops is a Faustian bargain: to the degree that we rely upon plants' abundance, we share in their vulnerability to disease. I hope that the small contributions I have added to the body of work addressing this problem can be of service to our rebellion, this fragile and temporary profusion of life.

Works Cited

1. Van Loon, H. W. *The Story of Mankind*. (Boni and Liveright, 1921).
2. Savary, S. *et al.* The global burden of pathogens and pests on major food crops. *Nat. Ecol. Evol.* **3**, 430–439 (2019).
3. D'Odorico, P., Carr, J. A., Laio, F., Ridolfi, L. & Vandoni, S. Feeding humanity through global food trade. *Earths Future* **2**, 458–469 (2014).
4. Singh, B. K. *et al.* Climate change impacts on plant pathogens, food security and paths forward. *Nat. Rev. Microbiol.* **21**, 640–656 (2023).
5. Solving Child Hunger. *Action Against Hunger* https://www.actionagainsthunger.org/the-hunger-crisis/world-hunger-facts/who-does-hunger-affect/child-hunger/.
6. href= & href=. Theophrastus, Enquiry Into Plants. *https://archive.org/stream/enquiryintoplant02theouoft/enquiryintoplant02theouoft_djvu.txt* https://topostext.org/work/242 (300AD).
7. Chester, K. S. The Problem of Acquired Physiological Immunity in Plants.
8. Price, W. C. Acquired Immunity from Plant Virus Diseases. *Q. Rev. Biol.* **15**, 338–361 (1940).
9. Biffen, R. H. Mendel's Laws of Inheritance and Wheat Breeding. *J. Agric. Sci.* **1**, 4–48 (1905).
10. Flor, H. H. Current Status of the Gene-For-Gene Concept. *Annu. Rev. Phytopathol.* **9**, 275–296 (1971).
11. Ross, A. F. Systemic acquired resistance induced by localized virus infections in plants. *Virology* **14**, 340–358 (1961).
12. Loon, L. C. V. Induced resistance in plants and the role of pathogenesis-related proteins.

13. Loon, V. Polyacrylamide Disc Electrophoresis of the Soluble Leaf Proteins from Nicotiana tabacum var. ‘Samsun’ and ‘Samsun NN’ IV. Similarity of Qualitative Changes of Specific Proteins after Infection with Different Viruses and their Relationship to Acquired Resistance.
14. Kassanis, B., Gianinazzi, S. & White, R. F. A Possible Explanation of the Resistance of Virus-infected Tobacco Plants to Second Infection. *J. Gen. Virol.* **23**, 11–16 (1974).
15. Stone, E. An Account of the Success of the Bark of the Willow in the Cure of Agues. (1763).
16. Montinari, M. R., Minelli, S. & De Caterina, R. The first 3500 years of aspirin history from its roots – A concise summary. *Vascul. Pharmacol.* **113**, 1–8 (2019).
17. Howe, G. A. & Jander, G. Plant Immunity to Insect Herbivores. *Annu. Rev. Plant Biol.* **59**, 41–66 (2008).
18. Kytidou, K., Artola, M., Overkleeft, H. S. & Aerts, J. M. F. G. Plant Glycosides and Glycosidases: A Treasure-Trove for Therapeutics. *Front. Plant Sci.* **11**, (2020).
19. Kim, Y.-S., Choi, Y.-E. & Sano, H. Plant vaccination: Stimulation of defense system by caffeine production in planta. *Plant Signal. Behav.* **5**, 489–493 (2010).
20. Wood, J. N. From plant extract to molecular panacea: a commentary on Stone (1763) ‘An account of the success of the bark of the willow in the cure of the agues’. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20140317 (2015).
21. White, R. F. Acetylsalicylic acid (aspirin) induces resistance to tobacco mosaic virus in tobacco. *Virology* **99**, 410–412 (1979).
22. Métraux, J. P. *et al.* Increase in Salicylic Acid at the Onset of Systemic Acquired Resistance in Cucumber. *Science* **250**, 1004–1006 (1990).
23. Gaffney, T. *et al.* Requirement of Salicylic Acid for the Induction of Systemic Acquired Resistance. *Science* **261**, 754–756 (1993).

24. Rasmussen, J. B., Hammerschmidt, R. & Zook, M. N. Systemic Induction of Salicylic Acid Accumulation in Cucumber after Inoculation with *Pseudomonas syringae* pv *syringae*. *Plant Physiol.* **97**, 1342–1347 (1991).
25. Malamy, J., Carr, J. P., Klessig, D. F. & Raskin, I. Salicylic Acid: A Likely Endogenous Signal in the Resistance Response of Tobacco to Viral Infection. *Science* **250**, 1002–1004 (1990).
26. Wildermuth, M. C., Dewdney, J., Wu, G. & Ausubel, F. M. Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature* **414**, 562–565 (2001).
27. Chen, H. *et al.* Connecting the Dots: A New and Complete Salicylic Acid Biosynthesis Pathway. *Mol. Plant* **12**, 1539–1541 (2019).
28. Lefevere, H., Bauters, L. & Gheysen, G. Salicylic Acid Biosynthesis in Plants. *Front. Plant Sci.* **11**, 338 (2020).
29. Jones, J. D. G. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).
30. DeFalco, T. A. & Zipfel, C. Molecular mechanisms of early plant pattern-triggered immune signaling. *Mol. Cell* **81**, 3449–3467 (2021).
31. Cui, H., Tsuda, K. & Parker, J. E. Effector-Triggered Immunity: From Pathogen Perception to Robust Defense. *Annu. Rev. Plant Biol.* **66**, 487–511 (2015).
32. Zhang, S., Li, C., Si, J., Han, Z. & Chen, D. Action Mechanisms of Effectors in Plant-Pathogen Interaction. *Int. J. Mol. Sci.* **23**, 6758 (2022).
33. Balint-Kurti, P. The plant hypersensitive response: concepts, control and consequences. *Mol. Plant Pathol.* **20**, 1163–1178 (2019).
34. Glazebrook, J. Contrasting Mechanisms of Defense Against Biotrophic and Necrotrophic Pathogens. *Annu. Rev. Phytopathol.* **43**, 205–227 (2005).

35. Li, Q. *et al.* N-hydroxypipecolic acid triggers systemic acquired resistance through extracellular NAD(P). *Nat. Commun.* **14**, 6848 (2023).
36. Carr, J. P. & Klessig, D. F. The Pathogenesis-Related Proteins of Plants. in *Genetic Engineering* (ed. Setlow, J. K.) 65–109 (Springer US, Boston, MA, 1989). doi:10.1007/978-1-4615-7084-4_5.
37. Van Loon, L. C. & Van Strien, E. A. The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins. *Physiol. Mol. Plant Pathol.* **55**, 85–97 (1999).
38. Alexander, D. *et al.* Increased tolerance to two oomycete pathogens in transgenic tobacco expressing pathogenesis-related protein la. *Plant Biol.*
39. Ali, S. *et al.* Pathogenesis-related proteins and peptides as promising tools for engineering plants with multiple stress tolerance. *Microbiol. Res.* **212–213**, 29–37 (2018).
40. Kinkema, M., Fan, W. & Dong, X. Nuclear Localization of NPR1 Is Required for Activation of PR Gene Expression.
41. Kumar, S. *et al.* Structural basis of NPR1 in activating plant immunity. *Nature* **605**, 561–566 (2022).
42. Wu, Y. *et al.* The Arabidopsis NPR1 Protein Is a Receptor for the Plant Defense Hormone Salicylic Acid. *Cell Rep.* **1**, 639–647 (2012).
43. Durner, J. & Klessig, D. F. Salicylic Acid Is a Modulator of Tobacco and Mammalian Catalases. *J. Biol. Chem.* **271**, 28492–28501 (1996).
44. Morel, J.-B. & Dangl, J. L. The hypersensitive response and the induction of cell death in plants. *Cell Death Differ.* **4**, 671–683 (1997).

45. 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* **15**, 1846–1858 (2003).
46. Sun, T. *et al.* TGACG-BINDING FACTOR 1 (TGA1) and TGA4 regulate salicylic acid and pipecolic acid biosynthesis by modulating the expression of SYSTEMIC ACQUIRED RESISTANCE DEFICIENT 1 (SARD1) and CALMODULIN-BINDING PROTEIN 60g (CBP60g). *New Phytol.* **217**, 344–354 (2018).
47. Betsuyaku, S. *et al.* Salicylic Acid and Jasmonic Acid Pathways are Activated in Spatially Different Domains Around the Infection Site During Effector-Triggered Immunity in Arabidopsis thaliana. *Plant Cell Physiol.* **59**, 8–16 (2018).
48. Dorey, S. *et al.* Spatial and Temporal Induction of Cell Death, Defense Genes, and Accumulation of Salicylic Acid in Tobacco Leaves Reacting Hypersensitively to a Fungal Glycoprotein Elicitor. *Mol. Plant-Microbe Interactions®* **10**, 646–655 (1997).
49. Jia, X. *et al.* The origin and evolution of salicylic acid signaling and biosynthesis in plants. *Mol. Plant* **16**, 245–259 (2023).
50. Torrens-Spence, M. P. *et al.* PBS3 and EPS1 Complete Salicylic Acid Biosynthesis from Isochorismate in *Arabidopsis*. *Mol. Plant* **12**, 1577–1586 (2019).
51. Cao, H., Bowling, S. A., Gordon, S. & Dong, X. Characterization of an Arabidopsis Mutatn That Is Nonresponsive to Inducers of Systemic Acquired Resistance. *Plant Cell* **6**, 1583–1592 (1994).
52. Cao, H., Glazebrook, J., Clarke, J. D., Volko, S. & Dong, X. The Arabidopsis NPR1 Gene That Controls Systemic Acquired Resistance Encodes a Novel Protein Containing Ankyrin Repeats. *Cell* **88**, 57–63 (1997).

53. Rochon, A., Boyle, P., Wignes, T., Fobert, P. R. & Després, C. The Coactivator Function of *Arabidopsis* NPR1 Requires the Core of Its BTB/POZ Domain and the Oxidation of C-Terminal Cysteines. *Plant Cell* **18**, 3670–3685 (2007).

54. Tada, Y. *et al.* Plant Immunity Requires Conformational Charges of NPR1 via S-Nitrosylation and Thioredoxins. *Science* **321**, 952–956 (2008).

55. Liu, G., Holub, E. B., Alonso, J. M., Ecker, J. R. & Fobert, P. R. An Arabidopsis *NPR1* -like gene, *NPR4* , is required for disease resistance. *Plant J.* **41**, 304–318 (2005).

56. Zhang, Y. *et al.* Negative regulation of defense responses in Arabidopsis by two *NPR1* paralogs. *Plant J.* **48**, 647–656 (2006).

57. Ding, Y. *et al.* Opposite Roles of Salicylic Acid Receptors NPR1 and NPR3/NPR4 in Transcriptional Regulation of Plant Immunity. *Cell* **173**, 1–14 (2018).

58. Wang, W. *et al.* Structural basis of salicylic acid perception by Arabidopsis NPR proteins. *Nature* **586**, 311–316 (2020).

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

60. Weigel, R. R., Pfitzner, U. M. & Gatz, C. Interaction of NIMIN1 with NPR1 Modulates *PR* Gene Expression in Arabidopsis. *Plant Cell* **17**, 1279–1291 (2005).

61. Zhou, P., Zavaliev, R., Xiang, Y. & Dong, X. Seeing is believing: Understanding functions of NPR1 and its paralogs in plant immunity through cellular and structural analyses. *Curr. Opin. Plant Biol.* **73**, 102352 (2023).

62. Boyle, P. *et al.* The BTB/POZ Domain of the *Arabidopsis* Disease Resistance Protein NPR1 Interacts with the Repression Domain of TGA2 to Negate Its Function. *Plant Cell* **21**, 3700–3713 (2009).

63. Garretón, V., Carpinelli, J., Jordana, X. & Holuigue, L. The *as-1* Promoter Element Is an Oxidative Stress-Responsive Element and Salicylic Acid Activates It via Oxidative Species. *Plant Physiol.* **130**, 1516–1526 (2002).

64. Pape, S., Thurow, C. & Gatz, C. Exchanging the *as - 1* -like element of the *PR-1* promoter by the *as-1* element of the CaMV *35S* promoter abolishes salicylic acid responsiveness and regulation by NPR1 and SNI1. *Plant Signal. Behav.* **5**, 1669–1671 (2010).

65. Niggeweg, R., Thurow, C., Kegler, C. & Gatz, C. Tobacco Transcription Factor TGA2.2 Is the Main Component of as-1-binding Factor ASF-1 and Is Involved in Salicylic Acid- and Auxin-inducible Expression of as-1-containing Target Promoters. *J. Biol. Chem.* **275**, 19897–19905 (2000).

66. Zhang, Y., Tessaro, M. J., Lassner, M. & Li, X. Knockout Analysis of Arabidopsis Transcription Factors *TGA2* , *TGA5* , and *TGA6* Reveals Their Redundant and Essential Roles in Systemic Acquired Resistance. *Plant Cell* **15**, 2647–2653 (2003).

67. Després, C., DeLong, C., Glaze, S., Liu, E. & Fobert, P. R. The Arabidopsis NPR1/NIM1 Protein Enhances the DNA Binding Activity of a Subgroup of the TGA Family of bZIP Transcription Factors.

68. Després, C. *et al.* The Arabidopsis NPR1 Disease Resistance Protein Is a Novel Cofactor That Confers Redox Regulation of DNA Binding Activity to the Basic Domain/Leucine Zipper Transcription Factor TGA1. *Plant Cell* **15**, 2181–2191 (2003).

69. Gatz, C. From Pioneers to Team Players: TGA Transcription Factors Provide a Molecular Link Between Different Stress Pathways. *Mol. Plant-Microbe Interactions®* **26**, 151–159 (2013).

70. Tomaž, Š., Gruden, K. & Coll, A. TGA transcription factors—Structural characteristics as basis for functional variability. *Front. Plant Sci.* **13**, 935819 (2022).
71. Miller, M. The Importance of Being Flexible: The Case of Basic Region Leucine Zipper Transcriptional Regulators. (2009).
72. Perutz, M. *Science Is Not a Quiet Life: Unravelling the Atomic Mechanism of Haemoglobin*. (World Scientific Publishing Co. Pte. Ltd., P O box 128, Farrer Road, Singapore 912805, 1997).
73. Jaskolski, M., Dauter, Z. & Wlodawer, A. A brief history of macromolecular crystallography, illustrated by a family tree and its N obel fruits. *FEBS J.* **281**, 3985–4009 (2014).
74. McPherson, A. & Gavira, J. A. Introduction to protein crystallization. *Acta Crystallogr. Sect. F Struct. Biol. Commun.* **70**, 2–20 (2014).
75. Schmit, J. D. & Dill, K. Growth Rates of Protein Crystals. *J. Am. Chem. Soc.* **134**, 3934–3937 (2012).
76. Ashkenazy, H. *et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).
77. Landau, M. *et al.* ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res.* **33**, W299–W302 (2005).
78. Bokhove, M. *et al.* Easy mammalian expression and crystallography of maltose-binding protein-fused human proteins. *J. Struct. Biol.* **194**, 1–7 (2016).
79. Jin, T. *et al.* Design of an expression system to enhance MBP-mediated crystallization. *Sci. Rep.* **7**, 40991 (2017).

80. Heidel, A. J., Clarke, J. D., Antonovics, J. & Dong, X. Fitness Costs of Mutations Affecting the Systemic Acquired Resistance Pathway in Arabidopsis thaliana. *Genetics* **168**, 2197–2206 (2004).

81. Li, X. *et al.* The Systemic Acquired Resistance Regulator OsNPR1 Attenuates Growth by Repressing Auxin Signaling through Promoting IAA-Amido Synthase Expression. *Plant Physiol.* **172**, 546–558 (2016).

82. Navarrete, F. *et al.* TOPLESS promotes plant immunity by repressing auxin signaling and is targeted by the fungal effector Naked1. *Plant Commun.* **3**, 100269 (2022).

83. Weigel, R. R., Bäuscher, C., Pfitzner, A. J. P. & Pfitzner, U. M. NIMIN-1, NIMIN-2 and NIMIN-3, members of a novel family of proteins from Arabidopsis that interact with NPR1/NIM1, a key regulator of systemic acquired resistance in plants.

84. Ohta, M., Matsui, K., Hiratsu, K., Shinshi, H. & Ohme-Takagi, M. Repression Domains of Class II ERF Transcriptional Repressors Share an Essential Motif for Active Repression.

85. Hermann, M. *et al.* The Arabidopsis NIMIN proteins affect NPR1 differentially. *Front. Plant Sci.* **4**, (2013).

86. Silva, K. J. P., Mahna, N., Mou, Z. & Folta, K. M. NPR1 as a transgenic crop protection strategy in horticultural species. *Hortic. Res.* **5**, 1–11 (2018).

87. Ryals, J. & Johnson, J. The ArabidopsisNlMl Protein Shows Homology to the Mammalian Transcription Factor lnhibitor IKB.

88. Kopp, E. & Ghosh, S. Inhibition of NF-κB by Sodium Salicylate and Aspirin. *Science* **265**, 956–959 (1994).