Title

NAIPs: Building an Innate Immune Barrier against Pathogenic Bacteria

Permalink

https://escholarship.org/uc/item/031643q1

Author

Kofoed, Eric Matthew

Publication Date

2012

Peer reviewed|Thesis/dissertation

NAIPs: Building an Innate Immune Barrier against Pathogenic Bacteria

By

Eric Matthew Kofoed

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cellular Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Russell E. Vance, Chair
Professor David Raulet
Professor Barbara Baker
Professor Qing Zhong

Spring 2012

NAIPs: Building an Innate Immune Barrier against Pathogenic Bacteria

©2012

By Eric Matthew Kofoed

Abstract

NAIPs: Building an Innate Immune Barrier against Pathogenic Bacteria

By Eric Matthew Kofoed

Doctor of Philosophy in Molecular and Cellular Biology

University of California, Berkeley

Russell E. Vance, Chair

The innate immune system of mammals encodes several families of immune detector proteins that monitor the cytosol for signs of pathogen invasion. One important but poorly understood family of cytosolic immunosurveillance proteins is the NLR (nucleotide binding domain, leucine-rich repeat containing) proteins. Work presented here demonstrates that one subfamily of NLRs, the NAIPs (NLR family, Apoptosis Inhibitory Proteins), are activated by specific interaction with bacterial ligands, such as flagellin. NAIP activation leads to assembly of a large multiprotein complex called the inflammasome, which initiates innate immune responses by activation of the Caspase-1 protease. NAIPs therefore appear to detect pathogen molecules via a simple and direct receptor-ligand mechanism. Interestingly, other NLR family members appear to detect pathogens indirectly, perhaps by responding to host cell 'stress' caused by the pathogen. Thus, the NLR family may have evolved surprisingly diverse mechanisms for detecting pathogens.

Susceptibility to the intracellular bacterial pathogen *Legionella pneumophila* was mapped to a discrete genetic locus in the mouse containing multiple tandem gene paralogues constituting the NAIP gene array over ten years ago, but functional understanding of the NAIP array, and in particular whether different genes in the array perform diverse or redundant functions in host defense remains enigmatic. Here we show that NAIP proteins monitor the cytoplasm of host innate immune cells for the presence of bacterial flagellin, and the inner-rod proteins of bacterial type III secretion systems. Specific interaction of NAIP5 and NAIP6 with bacterial flagellin, and NAIP2 with inner-rod proteins (from a variety of pathogenic bacteria) promotes inflammasome assembly and activation. The 'orphan' receptor NAIP1 appears to detect a narrow subset of inner-rod proteins, but the structural basis for this selectivity and its functional importance are unknown.

Finally, we have mapped the ligand-specificity determining region of the NAIPs to a ~200 amino acid region of the LRR domain that spans seven structural repeats that exhibit a high degree of polymorphism among the NAIP paralogues. Whether the region is sufficient for binding specific ligands is an open question. We have used structural modeling and polymorphism mapping of the NAIP LRR domains to identify putative ligand-binding residues that may be under selective pressure from bacterial pathogens. Thus, the NAIP gene array has evolved to recognize functionally constrained molecules of pathogenic bacteria, and represents an exquisite mechanism of host defense.

For

Farida Jhabvala Romero

Whose endless love and compassion have carried me through the academic jungle

&

Devika Kofoed

Who is flawless, an absolute joy, and a reminder that world is what you make of it

&

Maryanne and Russell Kofoed

Who have given me the gifts of wonder and the love of nature

Table of Contents

List of Figures

Acknowledgements

Chapter One:
Introduction to host recognition of pathogenic bacteria

*Significant portions of this chapter are scheduled for publication in the journal BioEssays (2012) in an article with the same title as this thesis, and co-authored by Eric M. Kofoed and Russell E. Vance. Permission to reprint this material is granted to contributors by Wiley-Blackwell publishers solely for academic purposes.

1.1. Pathogen recognition by germline-encoded receptors

Initial recognition of pathogens is the first critical step in immune defense and is the primary function of the innate immune system. In mammals, work over the last two decades has demonstrated that the primary molecular mechanism for recognition of pathogens is via germline-encoded receptors that recognize specific microbial molecules, such as flagellin or lipopolysaccharide, that are conserved among diverse pathogens. These molecules are often referred to as pathogen-associated molecular patterns, or PAMPs[1]. The most well-studied class of receptors that detect PAMPs are the transmembrane Toll-like receptors (TLRs). TLRs have now been shown to mediate recognition of a diversity of pathogen-derived PAMPs directly, or as part of a recognition complex[2-7]. Recognition of PAMPs by TLRs is mediated by a ligand-recognition domain that consists of a series of leucine-rich repeats (LRRs). LRRs are found in many proteins of immune and non-immune function, and structural studies of these proteins appear to suggest that the LRR domain is a highly evolvable and functionally flexible domain that can accommodate recognition of diverse ligands[8]. This may be critical in permitting the TLRs to maintain recognition of diverse and evolving microbial ligands.

The TLRs are type I transmembrane proteins with the LRR domain located at the N-terminus (co-translationally directed to the lumen of the ER). Thus, the topology of TLRs in the membrane implies that they recognize microbial ligands that are extracellular, phagosomal, or delivered to the lumen of intracellular compartments such as autophagosomes. TLRs are not, therefore, believed to be involved in the recognition of cytosolic PAMPs. This is not necessarily a major shortcoming of the TLR system, since many pathogens that reside primarily in the cytosol must nevertheless access or transit the extracellular space as part of their lifestyle. Even if a pathogen has the ability to spread directly from the cytosol of one host cell to another, the host has evolved mechanisms such as phagocytosis or autophagocytosis that can direct cytosolic material into compartments that are monitored by TLRs. In fact, few (if any) pathogens are able to completely avoid recognition by TLRs.

Still, it has become apparent that the TLR system for pathogen recognition is complemented by several families of innate immune pathogen-detector proteins that reside in the host cell cytosol[9]. We focus here on one particular family of cytosolic pathogen-detection proteins that are called the NLR (Nucleotide-binding domain (NBD), Leucine-rich Repeat (LRR)-containing) proteins. Although much less is known about NLRs than their TLR cousins, the NLR gene family appears to be a significant component of the germline-encoded immune detection system of mammals, with greater than 20 different NLR genes in both mice and humans. Some species, such as sea urchins, encode over 200 members of the NLR gene family[10]. NLRs are also central to the immune defense system of plants, where they have been studied for several years prior to the discovery of TLRs or NLRs in mammals[11,12].

Why have hosts evolved cytosolic pathogen-recognition proteins? This is an especially interesting question when it is considered, as mentioned above, that virtually all cytosolic pathogens are nevertheless detected by TLRs. What added benefit is conferred to the host by an elaborate system of cytosolic immunosurveillance? The answer to this question may be found in part by considering a second key question: how do cytosolic NLRs detect pathogens? Although TLRs and NLRs both contain LRR domains, the mechanism by which NLRs recognize pathogens remains unclear, and may not resemble the direct recognition of PAMPs mediated by TLRs. Below we consider how NLRs may provide unique information to the immune system not afforded by TLRs. It is also interesting to consider the kinds of immune responses orchestrated by NLRs. In many cases, these responses do not overlap with the responses initiated by TLRs. Elucidation of the distinct signaling outputs of NLRs may help explain the role of NLRs and cytosolic immunosurveillance in host immunity.

1.2. Identification of NAIPs as mediators of innate immunity

As is often the case, the discovery of NAIPs followed a circuitous path. The initial studies did not begin with the goal of understanding cytosolic immunosurveillance. Instead they began with a very different question: how is replication of the vacuolar pathogen, *Legionella pneumophila*, controlled by mouse macrophages? *L. pneumophila* is a gram-negative facultative intracellular bacterium that can cause a severe pneumonia in humans called Legionnaires' Disease[13]. The natural replicative niche for *L. pneumophila*, however, appears to be a variety of fresh water amoebae, and *L. pneumophila* does not appear to be transmitted from person to person[14]. Thus, *L. pneumophila* is often referred to as an accidental pathogen of humans. In both amoebae and humans, *L. pneumophila* appears to replicate within a specialized intracellular compartment called the Legionella Containing Vacuole (LCV). This compartment is required for *L. pneumophila* replication within host cells and its formation depends upon a specialized bacterial secretion system called the Dot/Icm system[15]. This system injects bacterial 'effectors' (enzymes) into the host cell cytosol that, among other things, orchestrate the creation of the LCV.

A key early observation was that *L. pneumophila* replication within mouse macrophages *ex vivo* varies with the genetic background of the mouse[16]. Macrophages from C57BL/6 mice restricted *L. pneumophila* replication, whereas those from the A/J strain were highly (~1000-fold more) permissive. Macrophages from other mouse strains, including 129, exhibit an intermediate and variable phenotype. The large phenotypic difference in *L. pneumophila* replication between B6 and A/J was mapped using classical genetics to a locus on mouse chromosome 13[17-19] that contains a tandemly repeated family of genes that were then called Neuronal Apoptosis Inhibitor Proteins (or NAIPs). Because the function of NAIP appears to not be restricted to neurons, and has mainly been studied in macrophages, a new nomenclature[20] has eliminated "Neuronal" from the name of NAIPs, which now stands for "NLR family, Apoptosis Inhibitory protein". (Unfortunately, there is also little evidence that NAIPs regulate apoptosis, but by now the NAIP name is engrained in the literature). Fine mapping and expression analysis identified the *Naip5* gene to be essential for the restriction of *L. pneumophila* replication in macrophages *ex vivo*, and further showed that functional transcripts of *Naip1*, *Naip2*, *Naip5* and *Naip6* exist in the C57BL/6 mouse strain[21-25]. *Naip5* has also been specifically deleted in mice of the C57BL/6

background and has been shown to be essential for restriction of *L. pneumophila* replication[26].

1.3. NAIP5 responds to flagellin and induces pyroptotic cell death

A key focus of research has been to understand how NAIP5 restricts bacterial replication. Several theories have been advanced, but this remains a controversial topic even now. One early report[27] correlated restriction with alterations in the early steps of formation of the LCV. Later it was suggested that restriction was "multifactorial", with a particularly important contribution being the induction of a rapid cell death in restrictive macrophages[28]. Since *L. pneumophila* depends upon an intracellular host cell niche in order to replicate, a proposed role for host cell death was particularly appealing since rapid host cell death would eliminate that niche and block replication. However, neither of these early studies was able to provide a molecular explanation of how NAIP5 connected either to cell death or altered LCV formation. A link between cell death and NAIP5 was, however, provided a couple of years later with the key observations of Dario Zamboni and Craig Roy[29], who showed that rapid cell death induced by *L. pneumophila* depends on the Caspase-1 protease, and that activation of Caspase-1 was regulated by NAIP5. Death induced by Caspase-1 involves overt cellular lysis and has been termed 'pyroptosis' ('fiery death') so as to distinguish it from the less inflammatory apoptotic death that is mediated by different Caspase family members[30,31]. Although several other labs have also reported NAIP5-dependent alterations in LCV maturation[32-34], it remains unclear whether or how NAIP5 regulates intracellular membrane trafficking events.

Because NAIP5 contains both an NBD and LRRs, the domain structure of NAIP5 places it within the NLR superfamily (Fig. 1). At the time that NAIP5 was identified, two members of the NLR superfamily, NOD1 and NOD2, were shown to be critical for immune responsiveness to fragments of bacterial peptidoglycan[35-38]. This established a precedent for the notion that NLRs might function as pathogen sensors – in essence, the cytosolic equivalents of the TLRs. However, there was little idea as to what the NAIPs or other orphan NLRs might recognize, if anything. It was therefore significant when two groups[39,40] reported that expression of flagellin by *L. pneumophila* was essential both for NAIP5-mediated restriction of *L. pneumophila* replication and for activation of Caspase-1-dependent cell death by *L. pneumophila*. The role of flagellin was not merely to promote bacterial motility, but appeared to be specifically required for NAIP5 activation, as non-motile but flagellin-expressing bacteria still induced NAIP5/Caspase-1-dependent pyroptosis and were still restricted. Moreover, in addition to flagellin, NAIP5- and Caspase-1-dependent cell death also required expression of the *L. pneumophila* Dot/Icm secretion system[40]. The most parsimonious explanation of the data was that flagellin was being delivered to the host cell cytosol via the Dot/Icm system where it activated NAIP5 and Caspase-1. However, direct evidence for the idea that NAIP5 was a flagellin-sensor was lacking in these early studies. Indeed, a lack of evidence for ligand recognition by NLRs is a chronic and ongoing issue for the field. Even now, direct uncontestable evidence that the NOD1/2 proteins recognize peptidoglycan fragments is lacking (but see [41,42]).

1.4. NLRC4 is an adaptor for NAIPs: Lessons from bacterial pathogens
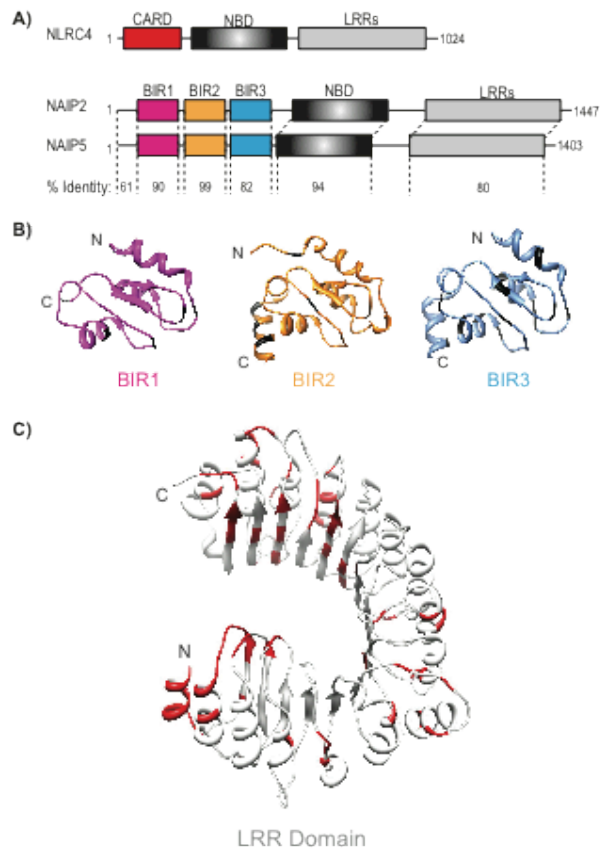
Figure 1.1. Domain organization and structural motifs of NLRC4 and NAIPs. **A:** Schematic representation of NLRC4, NAIP2, and NAIP5. Overall amino acid identity between NAIP2 and NAIP5 is 80%, whereas identity between NLRC4 and either NAIP is only ~15%. While the NBD and LRR folds are common to both NLRC4 and the NAIPs, they diverge greatly at the level of primary amino acid sequence. For example, even the minimal NBD of NLRC4 (152-248) and NAIP5 (453-551) are only 42% identical, whereas the NAIP2 and NAIP5 NBDs are 94% identical. Similarly, the predicted LRR domains of NAIP2 (1006-1447) and NAIP5 (961-1403) are 80% identical, whereas the LRR domain of NLRC4 (656-1024) and either NAIP share only 20% identity. Primary sequence identity between discrete domains of NAIP2 and NAIP5 are indicated below the alignment of NAIP2 and NAIP5. NH$_2$-terminus (1-57), BIR1 (58-129), BIR2 (157-229), BIR3 (276-347). **B:** Structural homology models of each NAIP BIR domain with polymorphisms between NAIP2 and NAIP5 mapped in black. The BIR1 domain was modeled after the crystal structure of bacculoviral iap repeat-containing protein 2 (cIAP2)(PDB ID#: c3t6pA)[43] covering amino acids 57-128 of NAIP2&5. The BIR2 domain was modeled after the structure of melanoma2 inhibitor of apoptosis (ml-iap) (PDB ID#: c1oy7c)[44] covering amino acids 156-252 of NAIP2&5. The BIR3 domain was modeled after the bacculoviral iap repeat-containing protein 1 (cIAP1)(PDB ID#:c2vm5A)[45] covering amino acids 275-372 of NAIP2&5. **C:** Structural homology model of the NAIP2 LRR domain with polymorphisms between NAIP2 and NAIP5 mapped in red. The LRR domain was modeled after the LRR of the plant auxin receptor TIR1 (PDB ID#: c2p1nE)[46] covering amino acids 1006-1441 of NAIP2, and 961-1396 of NAIP5. All homology models were generated using PHYRE, and polymorphisms between NAIP2 and NAIP5 were mapped onto the structures using UCSF Chimera.

A series of more recent studies have provided stronger evidence that NAIP5 is involved in recognition of flagellin. First, by direct expression of flagellin in macrophages from a mammalian promoter, or by transfection of purified flagellin into macrophages, it was demonstrated that flagellin itself was sufficient to activate NAIP5 and Caspase-1 in the absence of infection or other bacterially-derived stimuli[26,47,48]. In fact, expression of the C-terminal 35 amino acids of flagellin, fused to GFP as a carrier protein, is sufficient to activate NAIP5[26]. These 35 amino acids contribute to an alpha helical domain, called D0 (Fig. 2). Deletion or specific mutation of the C-terminus of flagellin abrogated its ability to activate host cell death, and this host cell death was abolished in $Naip5^{-/-}$ cells[26]. TLR5 is a distinct cell-surface receptor that also recognizes flagellin. The TLR5 ectodomain makes extensive direct interactions with the distinct D1 domain of flagellin[7,49,50]. Interestingly, the D0 and D1 domains of flagellin are highly conserved among different bacteria, and mutations in either of these regions of flagellin compromise bacterial motility[26,40,49]. Thus, NAIP5 is activated by a conserved and essential portion of the flagellin molecule, reminiscent of the conserved and essential molecular structures that serve as ligands for the TLRs.

The simple model that NAIP5 is a cytosolic flagellin sensor has been complicated by evidence indicating that NAIP5 does not act alone. In fact, a significant source of confusion for the field has been the observation that NAIP5-dependent restriction of *L. pneumophila* replication and flagellin-dependent activation of Caspase-1 requires a second member of the NLR superfamily, a protein now called NLRC4 (NLR family, CARD domain-containing 4) (formerly called IPAF (ICE-Protease Activating Factor)) [29,32]. As can be inferred from its name, NLRC4 contains an N-terminal Caspase Activation and Recruitment Domain (CARD), and this domain accounts for its ability to activate Caspase-1 (Fig.1A). NAIP5 lacks a CARD and thus its molecular role in Caspase-1 activation was not clear. Moreover, while *Naip5*-deficient mice retained the ability to respond to flagellin under certain circumstances, *Nlrc4*-deficient mice were defective in all responses to cytosolic flagellin[26,47,51]. For example, expression of full-length flagellin in the cytosol of macrophages could activate Caspase-1 in a manner dependent on *Nlrc4* but independent of *Naip5*[26]. Moreover, activation of Caspase-1 by bacterial pathogens such as *Salmonella typhimurium* or *Pseudomonas aeruginosa* depended entirely on *Nlrc4*, but only partially on *Naip5*[47,51-53]. For these reasons, the field assumed that *Nlrc4* was likely to be the true cytosolic sensor of flagellin, whereas the role of *Naip5* was unclear. It was proposed, for example, that *Naip5* might detect a ligand in *L. pneumophila* other than flagellin[54].

A significant study from Ed Miao, Alan Aderem and colleagues[48] provided a key insight that eventually led to a clearer understanding of the respective roles of NAIP5 and NLRC4. Surprisingly, it was found that NLRC4 could be activated by a second bacterial protein that is present in many bacterial pathogens, the inner rod protein of type III secretion systems (T3SS)(Fig. 2A). In *Salmonella*, the inner rod protein of the SPI-1 T3SS is called PrgJ. The two activators of NLRC4, PrgJ and flagellin, were suggested to share a superficial structural similarity, as C-terminal alpha-helical regions appeared to be recognized in both proteins (Fig. 2C)[48]. However, there is only low primary sequence homology between flagellin and PrgJ (Fig. 2B), raising the question of how these two proteins were specifically recognized. Interestingly, in subsequent work it was shown that PrgJ activates NLRC4 without a requirement for NAIP5, whereas NAIP5 seemed to be specifically required for the response to flagellin[52]. Since *L. pneumophila* lacks a T3SS and a PrgJ homolog, these results provide an explanation of why recognition of *L. pneumophila* is *Naip5*-dependent, whereas recognition of *Salmonella* and *Pseudomonas*,

which express both T3SSs and flagellin, is largely *Naip5*-independent. In retrospect, the functional specialization of NAIP5 should not have been a surprise. *Naip5* is one of four expressed *Naip* paralogues on Chromosome 13, yet susceptibility to *L. pneumophila* mapped to a single paralog, *Naip5*. Thus, it was clear early on that *Naip* paralogs are not completely redundant with each other. Nevertheless, it was the identification of PrgJ as an activator of NLRC4 that ultimately fostered the hypothesis that perhaps another NAIP paralogue might be responsible for detection of bacterial rod proteins such as PrgJ. What was needed was a biochemical system for analyzing NAIP function. I describe such a biochemical system in this dissertation and its use in identifying the function of NAIPs in NLRC4 inflammasome activation.

1.5. Role of inflammation in innate and adaptive immunity

1.5.1. Consequences of inflammasome activation

Inflammasome activity is required for host resistance to many infections, but has also been implicated in a number of inflammatory pathologies[55]. For example, Caspase1-deficient mice are highly susceptible to some bacterial pathogens and resistant to a variety of inflammatory diseases[55-58]. In fact, patients suffering from Crohn's disease, or several autoinflammatory syndromes have been found to contain polymorphic NLR alleles[38,59-61]. These observations raise two important questions: 1) why is inflammation important for host defense? and 2) why would evolution favor a defense mechanism so sensitive that it can inadvertently cause immunopathology?

1.5.2. Priming the adaptive immune response

Innate immunity is a feature of all metazoans and functions to identify invading microbes and mount a defense. In fact, many organisms lack an adaptive immune system and are thriving nevertheless. In contrast, mammals have developed another layer of immunity that functions in concert with the innate immune system, and it is clear that the 'two' immune systems are inexorably connected. For our purposes, we can think of the innate immune system as providing a rapid host defense, and adaptive immunity as a slower, but tailored response that generates functional memory of, and immunity to, previously encountered pathogens. Thus, the main functions of the innate immune system are to identify and attack pathogens, and to prime the adaptive immune response.

Adaptive immunity plays an important role in host defense, and is characterized by somatic recombination of immunoglobulin-family gene segments that generates combinatorial diversity of unique antigen receptors[62]. This process creates a dizzying repertoire of lymphocytes carrying unique B and T cell receptor specificities within a single organism. The diversity of receptors with unique specificity permits a random, but thorough coverage of antigens that the host may encounter over its lifetime. Surprisingly, the jawless vertebrates (agnathes: sea lamprey & hagfish) utilize a similar host defense strategy by developing somatic diversity of anticipatory lymphocytes by germline rearrangement. Interestingly, instead of shuffling immunoglobulin-family gene segments, agnathes shuffle LRR gene cassettes[63,64].

The activation of naive B and T cells depends on cells of the innate immune system that are specialized for antigen presentation. Dendritic cells (DCs) are 'professional' antigen presenting cells, and respond to TLR ligands by rapid upregulation of co-stimulatory molecules and the production of cytokines important for T cell

Figure 1.2. Characteristics of bacterial ligands detected by NAIP paralogues. **A:** Illustration of a bacterial flagellum (on left) and a type III secretion system (on right) oriented in a bacterial membrane. The NAIP ligands and their location in the respective apparatus are indicated by color: flagellin monomer (FlaA or FliC) in blue; inner rod protein (PrgJ) in green; needle protein (PrgI) in orange. Abbreviations: OM, outer membrane; PGN, peptidoglycan layer; IM, inner membrane. **B:** Pairwise amino acid sequence alignment of the C-terminal 35 amino acids of FlaA and PrgJ. Hydrophobic amino acids are highlighted in red. Overall amino acid identity between FlaA and PrgJ is low, but C-terminal hydrophobic patches in both proteins have been shown to be important for activation of the inflammasome. **C:** Structural homology model of *L. pneumophila* FlaA based on phase 1 flagellin from *S. typhimurium* (PDB ID#: d1ucua) covering amino acids 2-473 of FlaA, and a model of PrgJ from *S. typhimurium* based on the needle protein PrgI (PDB ID#: c2jowA)[65] covering amino acids 31-89 of PrgJ. Flagellin subdomains are labeled D0-D3. The enlarged portion of FlaA is the carboxy and amino terminal 35 amino acids of D0 that form a structure superficially similar to PrgJ.

Figure 1.3. TLRs and NLRs generate distinct responses to pathogens. The TLRs (e.g. 2 / 4 / 5) and NLRs (e.g. NAIPs / NLRC4) are two innate immune receptor classes that directly detect pathogen associated molecular patterns (PAMPs) and are essential for host defense. TLRs are localized to the plasma membrane and endosomal compartments, whereas NLRs are restricted to the cytoplasm. Both receptor classes share an LRR domain utilized for PAMP recognition, but the downstream signaling outputs vary significantly. For example, activation of TLR5 by flagellin stimulates NF-kB-dependent transcription, upregulation of antigen presentation, and the production of T-cell activating cytokines. In contrast, activation of the NAIP5-NLRC4 inflammasome by flagellin delivered to the host cytoplasm causes pyroptosis and the maturation of pro-inflammatory cytokines that results in an acute inflammatory response.

activation (Fig. 1.3.). Antigen presenting cells recognize microbial invasion, transport antigens derived from these microbes to the lymph nodes, and initiate adaptive immune responses by presenting pathogen-derived peptides on major histocompatibility (MHC) proteins.

1.5.3. NLRs in adaptive immune responses

While TLRs can stimulate antigen presentation by initiating NF-κB, AP-1, and IRF3 transcriptional programs, the role of cytoplasmic NLRs in adaptive immunity is much less clear. The most commonly used vaccine adjuvant (alum) activates the NLRP3/ASC inflammasome, suggesting that delivery of antigen in physical association with an inflammasome activator is sufficient to induce a robust antibody response[66].

However, the mechanism by which NLRP3 responds to alum and triggers humoral immunity remains controversial[67].

The bacterial cell wall fragments isoglutamyl-di-amino-pimelic acid (iE-DAP) and muramyl di-peptide (MDP) are detected by NOD1 and NOD2, respectively [36,37,68], and make important contributions to antigen-specific immunity against *Helicobacter pylori* and *Listeria monocytogenes*[69-72]. More importantly, MDP is the minimally active component of Freund's adjuvant, suggesting that in some cases NLR activation is sufficient for initiating adaptive immune responses[73]. However, NOD1/2 are best known for their involvement in upregulating chemokines that recruit neutrophils to the site of infection in response to many different bacterial pathogens[74-78].

Nevertheless, it is generally accepted that inflammation plays an important role in shaping the adaptive immune response, and is essential for transmigration of leukocytes into the site of infection from the blood stream. It should be emphasized here that neutrophil recruitment to wound sites is essential to stall microbial replication during the time (often 4-7 days) needed to generate adaptive immunity. In support of this notion, neutropenic individuals suffer from chronic and debilitating infections [79].

1.5.4. Pathogen evasion of host recognition

Both bacterial and viral pathogens have evolved ways to avoid recognition by the inflammasomes, and enhance their dissemination[80,81]. In fact, pathogens have developed IL-1 decoy receptors[82], dominant negative NLR homologues[83], and Caspase-1 inactivating proteins[84-87] to avoid the consequences of inflammation.

S.typhimurium evades detection by the NAIP5-NLRC4 inflammasome by turning off expression of flagellin during the systemic phase of infection[88]. However, if engineered to constitutively express flagellin, the pathogen is cleared by a pyroptosis-dependent mechanism that exposes bacteria for uptake and killing by neutrophils[81]. Furthermore, mice deficient in an enzyme involved in production of phagosome reactive oxygen species also failed to clear flagellin-expressing Salmonella *in vivo*. These data demonstrate a role for pyroptosis in exposing intracellular pathogens to neutrophil-mediated clearance *in vivo*, and support the concept that pathogens are evolving to avoid detection by inflammasomes. Intracellular bacterial pathogens are not frequently accessible to antibody or complement, which is why cellular immune responses frequently play important roles against intracellular pathogens[89].

1.5.5. Substrates of Caspase-1

The natural substrates of Caspase-1 include pro-IL-1β and pro-IL-18, but few other verified substrates have been described[90,91]. Both of these cytokines are important for immune responses to a variety of viral and bacterial infections[88,92,93]. However, the molecular targets that mediate pyroptosis downstream of Caspase-1 activation are poorly characterized. While a variety of global proteomic approaches have defined potential Caspase1 substrates, different groups using different experimental strategies have identified non-overlapping targets[94-96]. Nevertheless, identification of the Caspase-1 substrates involved in pyroptosis, and determining how substrates are selected remains an important question.

1.5.6. Role of IL-1β in host defense

Interleukin-1β is recognized as a potent pyrogen and a gatekeeper of inflammation in recognition of its commanding role in inducing fever and inflammation[97,98]. IL-1β is responsible for changes in local vasculature during acute infection, changes in immune cell function, and activation of the acute phase response.

Changes in vascular permeability around the site of infection permits an influx of cytotoxic leukocytes that are stimulated by IL-1β and IL-18 (and secondary cytokines) to destroy bacteria. IL-1β also directly stimulates production of factors that cause neutrophilia and the differentiation of monocytes into potent antigen presenting cells[98]. Finally, IL-1β and IL-6 activate the acute phase response by causing the liver to produce high levels of complement that mediates the systemic clearance of microbes from the circulation. Pyroptosis complements the function of IL-1β by exposing intracellular pathogens to extracellular defenses such as the complement system, antibodies, and cellular immunity while at the same time evicting the pathogen from its replicative niche.

1.6. Dissertation Overview

Inflammation shapes the host response to pathogen infection, but also underlies many pathological diseases. Phagocytic cells of the innate immune system serve as the front line of host defense against pathogen invasion, and are poised to respond using germline-encoded receptors that detect PAMPs. What makes an immune response 'appropriate' depends on the nature of the insult, but the steady-state exclusion of PAMPs from the host cytoplasm provides the basis for cytosolic immunosurveillance.

Here we show that the NAIP gene array has evolved to detect conserved and functionally constrained features of bacterial pathogens. The features targeted for recognition by the NAIPs cannot be easily mutated due to their critical roles in bacterial fitness (i.e. locomotion and virulence). I show that inner-rod proteins derived from bacterial type III secretion systems are specifically detected by NAIP1 and NAIP2, whereas flagellin is specifically detected by NAIP5 and NAIP6. NAIPs bind these bacterial protein ligands directly, recruit NLRC4, and form an inflammasome complex consisting of a NAIP protein, NLRC4, and a cognate bacterial ligand. Thus, recognition of bacterial ligands by NAIPs determines inflammasome specificity.

Here we map the region of NAIPs responsible for conferring ligand-specificity to the LRR domain between amino acids 822-1078. This region likely represents the ligand-interaction surface. The LRR domain is utilized for the recognition of foreign molecules by mammalian TLRs and NLRs, and also by plant NB-LRR resistance proteins[12,99,100]. Strikingly, the LRR is utilized as the basic unit of somatic recombination in the generation of lymphocyte diversity in jawless vertebrates[63,64]. Use of the LRR domain in pathogen recognition and host defense across kingdoms is a remarkable example of convergent evolution. In addition, the prevalence of non-synonymous substitutions among LRR-containing gene families indicates that these proteins are under diversifying selective pressure by pathogens. These observations suggest that the LRR itself may hold greater potential than we currently appreciate, both for understanding host-pathogen evolution, and in novel therapeutic applications.

Chapter Two
Measuring the inflammasome

2.1. Introduction

Pioneering work from Jürg Tschopp's lab[102] has provided biochemical evidence that NLR-dependent activation of Caspase-1 occurs within a high molecular weight multiprotein complex that was termed the "inflammasome", by analogy to the "apoptosome"[103] that mediates activation of apoptotic caspases. Formation of the apoptosome depends upon oligomerization of a key scaffold protein called APAF-1 (Apoptosis Protease Activating Factor-1) that contains a NBD very similar to that found in NLRs. Indeed, one NLR family member, NLRP1, has been shown to assemble into pentameric and heptameric rings that exhibit striking structural similarity to the ring-like structures formed by APAF-1[104]. Oligomerization of APAF-1 is required for proteolytic activation of Caspase-9, which is believed to occur upon proximity induced dimerization[105]. Based largely on the model of the apoptosome, activation of Caspase-1 by NLRs is also believed to depend upon a ligand-induced oligomerization step, but it must be emphasized that the biochemical mechanism of NLR-mediated inflammasome assembly still requires rigorous biochemical investigation. In fact, the biochemical study of ligand-dependent NLR activation and inflammasome assembly has proven difficult for a variety of reasons, including low expression levels of key protein components, a lack of good antibodies for most NLRs, a predisposition for ligand-independent NLR aggregation, and the absence of specific known ligands. In light of these limitations, we turned to overexpression of known inflammasome components in 293T cells in an attempt to reconstitute a heterologous but more biochemically tractable NAIP5-NLRC4 inflammasome.

2.2. Reconstitution of NAIP5-NLRC4 inflammasome in a heterologous system

The NAIP inflammasomes present additional difficulties for *in vivo* and *ex vivo* study that include the expression of multiple NAIP proteins and their low expression levels. However, reconstituting the inflammasome in non-immune 293T cells, which do not express NLRC4, NAIPs, or Caspase-1 permits functional analysis of individual proteins. The crux of this approach is to identify a biological activity that can be quantified. One established assay for inflammasome activation is the 'retroviral lethality' assay that compares survival of macrophages after transduction with a vector expressing flagellin marked with green fluorescent protein (GFP), versus a GFP-marked vector alone[26]. If the protein in question activates the inflammasome it will cause pyroptosis and release of the GFP marker, whereas expression of the vector alone should not. We reasoned that expression of known components of the inflammasome in a heterologous system should be capable of recapitulating key aspects of flagellin-dependent activation.

A general caveat of reconstituted systems is that protein overexpression can lead to spontaneous non-specific (ligand-independent) activities. We were able to circumvent this problem by employing an MSCV2.2-IRES-GFP vector that expresses proteins at moderate levels, upstream of an internal ribosome entry site (IRES) and a GFP marker. 293T cells were transiently transfected with MSCV2.2-IRES-GFP encoding wild-type NLRC4, NAIP5 and Caspase-1. Importantly, the transfected cells did not exhibit significant spontaneous inflammasome activation, and instead, a majority of cells expressed GFP (Fig. 2.1). However, when flagellin (FlaA) from *L. pneumophila* was co-expressed with NLRC4/NAIP5/Caspase-1, we observed a significant loss of GFP^{high} cells and an increase in the number of dead (7AAD positive) cells (Fig. 2.1). This result was
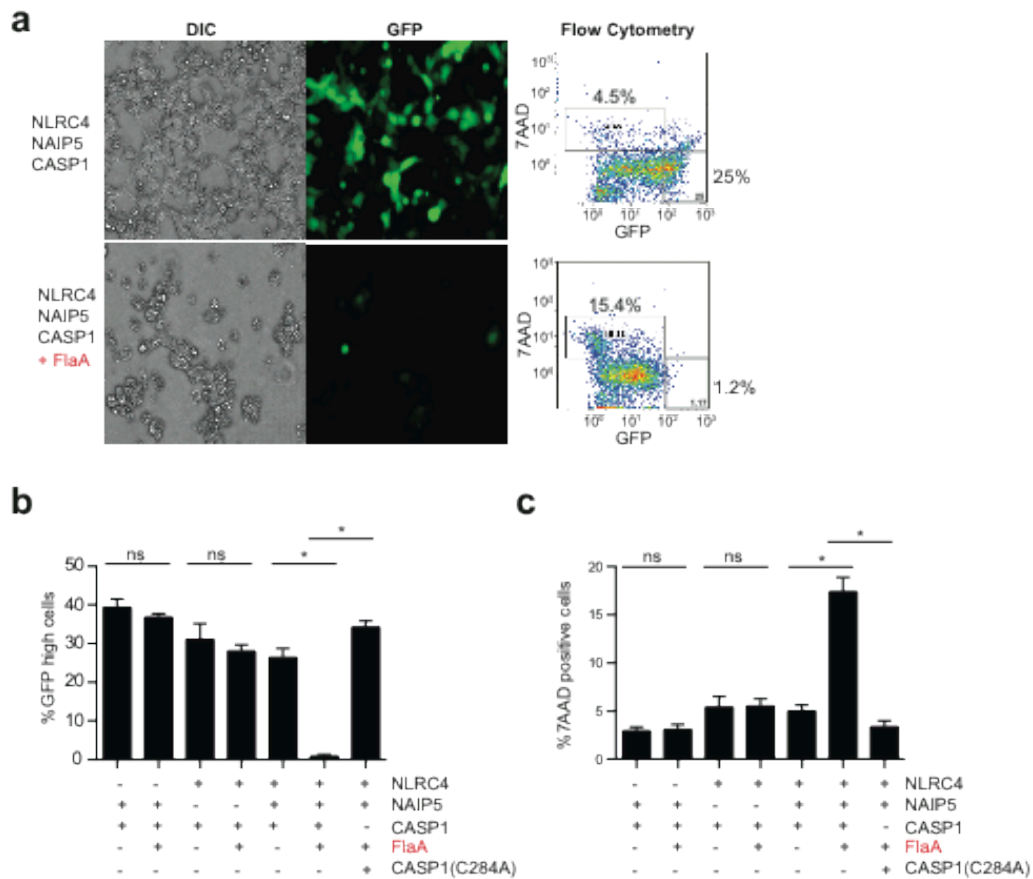
Figure 2.1. Reconstitution of the NAIP5-NLRC4 inflammasome in 293T cells. (**a**) MSCV2.2-IRES-GFP expression vectors encoding NLRC4, NAIP5, and Caspase-1 were transiently transfected into 293T cells, alone or in combination with MSCV2.2 encoding full-length flagellin from *L. pneumophila* (FlaA). Cells were imaged for differential interference contrast (DIC) and GFP fluorescence 48 hours later. Harvested cells were stained with 7AAD, a fluorescent dye that labels dead cells that have lost membrane integrity, and analyzed by flow cytometry for 7AAD and GFP. (**b**) GFP-high cells and (**c**) 7AAD positive cells were quantified as in **a**, but with specific expression vectors omitted from the transfection as indicated. Caspase-1 (C284A) is a catalytically dead mutant.

reminiscent of flagellin-dependent activation of the endogenous NAIP5/NLRC4 inflammasome in macrophages, which also results in a rapid Caspase-1-dependent cell death, loss of membrane integrity, and release of cytosolic contents and GFP[26,52]. Similar to the genetic requirement for *Nlrc4*, *Naip5* and *Casp1* in macrophages, we found that NAIP5, NLRC4, catalytically active Caspase-1 and FlaA are all required to trigger cell death and loss of membrane integrity/GFP in reconstituted 293T cells (Fig. 2.1).

Standard measures of inflammasome activation are the processing of pro-Caspase-1 and pro-IL-1β from inactive zymogens to biologically active proteins[106]. Importantly, the reconstituted NAIP5/NLRC4 inflammasome also recapitulated the ability of native inflammasomes to process Caspase-1 and IL-1β in response to cytosolic flagellin (Fig. 2.2).
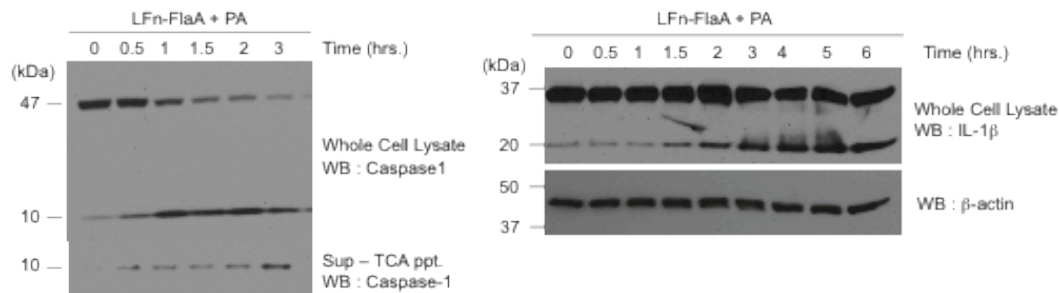
Figure 2.2. Reconstituted NAIP5-NLRC4 inflammasome processes Caspase-1 and IL-1β. (**a**) Transfection of 293T cells with NLRC4, NAIP5, and Caspase-1 was followed 18h later by delivery of purified LFn-FlaA [5μg/ml] fusion protein through a PA [1μg/ml] pore. In this system, PA alone or LFn-FlaA alone has no effect. Whole cell lysates and TCA precipitated cell supernatants were collected over time and analyzed by Western blotting for Caspase-1. (**b**) Processing of pro-IL-1β over a time-course of LFn-FlaA/PA treatment was monitored by Western blot of whole cell lysates from 293T cells transfected with murine pro-IL-1β, NAIP5, NLRC4, and Caspase-1.

The molecular basis of the protein-delivery method described above (Fig. 2.2) relies on the use of the bipartite lethal-factor (LF)/protective antigen (PA) toxin from *Bacillus anthracis* [107], and the protocol for delivery of purified LFn-flagellin through a PA pore has been described elsewhere[107,108]. One benefit of this delivery method is the opportunity to monitor kinetic events related to inflammasome activity following delivery of bacterial ligand in a bolus (Fig. 2.2). The activation of the NAIP5-NLRC4 inflammasome using this method of ligand-delivery also causes cell death of NAIP5/NLRC4/Caspase-1 expressing 293T cells as measured by loss of GFP[high] cells (Appendix Two).

2.3. Blue native gel electrophoresis used to monitor inflammasome assembly

Traditional methods of investigating inflammasome activity have focused on activation of the cysteine protease Caspase-1, and subsequent processing of the inflammatory cytokine IL-1β. These measurements report on the final output of

inflammasome activation, but do not address the mechanisms of inflammasome assembly and composition. It is believed that activated inflammasomes assemble into high molecular-weight multiprotein complexes[109], but despite considerable effort, this has not been demonstrated for the NLRC4 inflammasome. To visualize inflammasome assembly, 293T cells were transfected with NAIP5, NLRC4, and FlaA in various combinations, but Caspase-1 was omitted so that cell death and loss of cellular contents (and assembled inflammasomes) would not occur. Digitonin-solubilized cell lysates were resolved on Blue Native (BN) PAGE gels[110], and probed with a polyclonal antibody for NLRC4 (Fig. 2.3). A dramatic shift of NLRC4 from a monomer (~120kDa) to an oligomeric complex (~1000kDa) was seen in the presence of NAIP5 and FlaA. Complex formation was supported by either untagged FlaA, or a GFP-FlaA fusion protein, both of which were previously shown to activate NLRC4[26]. Importantly, assembly of the FlaA-inducible NLRC4 inflammasome was not observed in the absence of NAIP5 (Fig. 2.3A), indicating that a biochemical function of NAIP5 is to promote NLRC4 oligomerization. Accordingly, we observed that NAIP5 was also contained within the high molecular weight oligomeric complex (Fig. 2.3A). The association of NAIP5 and NLRC4 in the same complex was validated by co-immunoprecipitation studies (Fig. 2.3C). Samples resolved on SDS-PAGE showed that each inflammasome component was expressed evenly in the various transfectants, and thus the intense staining of the oligomerized form of NLRC4 is likely due to a greater avidity of the antibody for the oligomerized complex as opposed to the monomeric form of NLRC4.

Despite strong genetic evidence that NLR proteins, such as NAIP5 and NLRC4, function as microbial 'sensors', there is no biochemical evidence that NLRs interact directly with microbial ligands. In fact, some data on the NLRP3 inflammasome[111-113], as well as analyses of analogous proteins from plants[114], suggest that at least some NLR proteins recognize pathogens indirectly. Thus, the fundamental question of whether mammalian NLRs bind ligands remains controversial. In order to determine if the oligomerized NAIP5/NLRC4 complex also contains flagellin, we subjected samples separated in the first dimension by native PAGE to a second dimension of SDS-PAGE. To facilitate detection of flagellin, we used a 6x-Myc-tagged flagellin, which activates the inflammasome identically to native flagellin. Blots were probed for the presence of NLRC4, NAIP5, and 6x-Myc-FlaA. This approach revealed that FlaA was present in a high-molecular weight complex, along with NAIP5 and NLRC4 (Fig. 2.3B). Importantly, an additional control experiment using cells transfected only with NAIP5 and NLRC4 indicated the inflammasome complex does not form in the absence of FlaA (Appendix One). In addition, FlaA expressed alone was present in cell extracts only as a monomer (Appendix One). Reciprocal co-immunoprecipitations further corroborated that NLRC4, NAIP5, and FlaA are complexed with each other (Fig. 2.3C). Taken together, these observations provide the first evidence for a simple receptor-ligand model of NAIP5/NLRC4 activation by flagellin, analogous to extracellular PAMP detection by TLRs.

2.4. Targeted mutagenesis of NAIP5 and NLRC4 functional domains

In order to identify modular domains of both NAIP5 and NLRC4 involved in ligand recognition and inflammasome assembly we cloned various domain truncation and
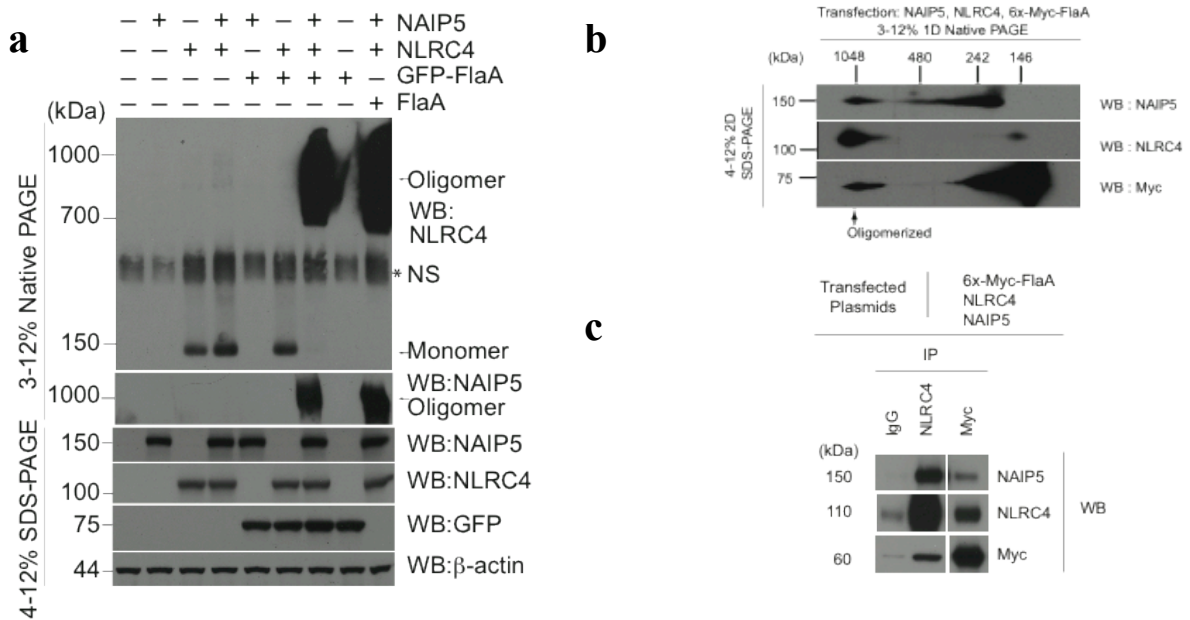
Figure 2.3. NAIP5 is required for formation of a hetero-oligomeric complex that contains NLRC4, NAIP5 and flagellin. (**a**) NAIP5, NLRC4, GFP-FlaA, and FlaA were expressed as indicated for 48 hours in 293T cells, followed by analysis by Blue Native-PAGE or SDS-PAGE, and western blotting. *NS, non-specific band. (**b**) 293T cells were transfected with NAIP5, NLRC4, and 6x-Myc-FlaA, and lysates were separated by a first dimension of Blue Native-PAGE followed by a second dimension of SDS-PAGE. Specific proteins were detected in both monomeric and oligomerized state by western blotting. (**c**) NLRC4, NAIP5, and flagellin (FlaA) can be co-immunoprecipitated. 293T cells were transfected with NAIP5, NLRC4, and 6x-Myc-FlaA for 48 hours, and digitonin cell extracts were immunoprecipitated with either control IgG, anti-NLRC4, or anti-Myc antibodies. Immunoprecipitates were separated by SDS-PAGE, and Western blotted for NAIP5, NLRC4, or Myc.

targeted deletion mutants for analysis using the assays described above. Consistent with the autoinhibitory function of LRRs in other NLRs, we found that NAIP5ΔLRR and NLRC4ΔLRR constitutively activated Caspase1-dependent cell death, independent of the presence of flagellin (Fig. 2.4A). Interestingly, NLRC4ΔLRR was able to activate Caspase-1 in the absence of NAIP5, whereas constitutively active NAIP5ΔLRR required wild-type NLRC4 in order to activate Caspase-1 (Fig. 2.4A & B; Appendix Two). This result suggests that NAIP5 functions upstream of NLRC4. Indeed, NAIP5ΔLRR was able to induce the oligomerization of wild-type NLRC4 (Fig. 2.4D), whereas the spontaneous oligomerization of NLRC4ΔLRR did not require NAIP5 (Fig. 2.4C). Spontaneous oligomerization of NLRC4ΔLRR did require the nucleotide binding domain (NBD) of NLRC4, as a K175R mutation previously shown to disrupt NBD function[115] abolished NLRC4ΔLRR auto-oligomerization (Fig. 2.4C). The ability of NAIP5 to induce oligomerization of NLRC4 in response to flagellin required both the NBD and N-terminal BIRs of NAIP5, but did not require the N-terminal CARD of NLRC4 (Fig. 2.4E & F), whereas functional Caspase-1 activation required all these domains (Fig. 2.5). Taken together, these data suggest a working model (Fig. 2.6) where NAIP5 is activated by flagellin and induces downstream NLRC4 oligomerization and Caspase-1 activation.
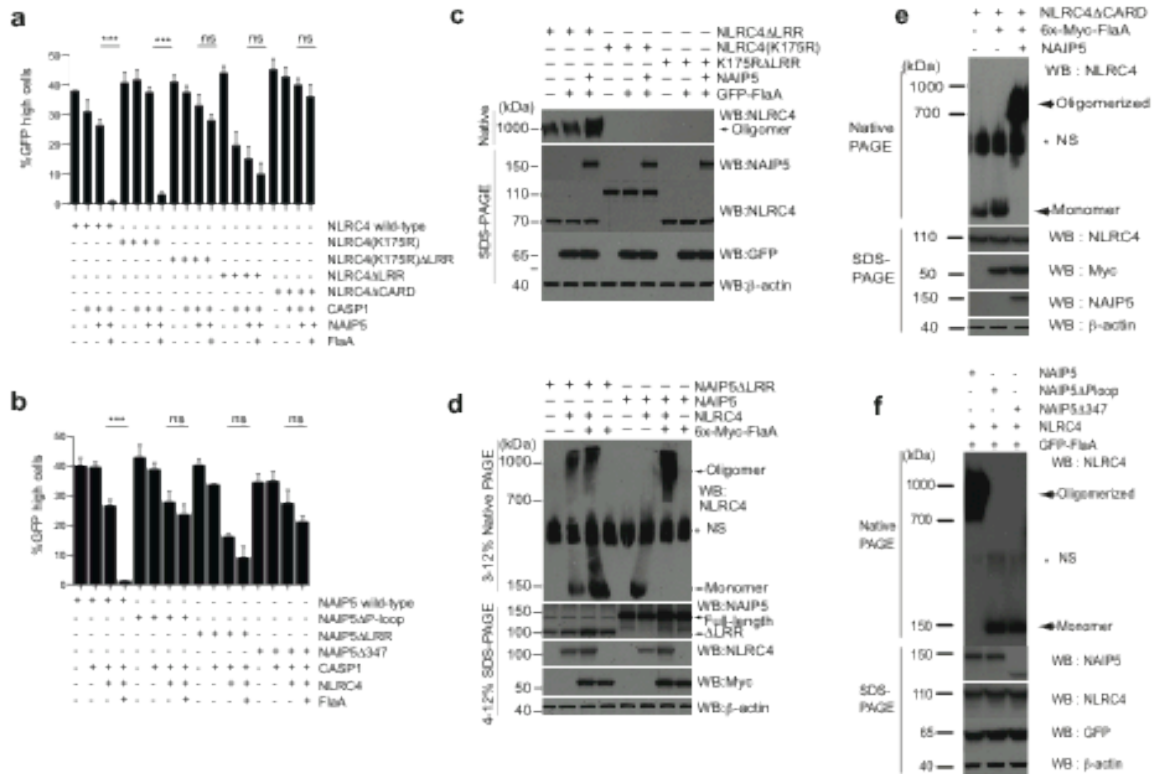
Figure 2.4. Functional domains involved in responsiveness to cytosolic flagellin & assembly of the inflammasome. (**a**) Wild-type NLRC4 or mutant constructs were transfected into 293T cells alone or with different combinations of Caspase-1, NAIP5, and FlaA MSCV2.2-IRES-GFP expression vectors. Forty-eight hours after transfection, cells were collected, stained with 7AAD, and analyzed by flow cytometry for GFP expression (**b**) Wild-type or mutant NAIP5 constructs were transfected alone or in combination with Caspase-1, NLRC4, and FlaA for 48 hours, followed by analysis for GFP expression. ns, not significant; ***, P<0.0001. (**c**) NLRC4ΔLRR, NLRC4(K175R), and NLRC4(K175R)ΔLRR were expressed alone or in combination with GFP-FlaA, or GFP-FlaA and NAIP5 in 293T cells, followed by analysis by Blue Native-PAGE or SDS-PAGE, and western blotting. *NS, non-specific band. (**d**) Expression constructs encoding NAIP5ΔLRR mutant and wild-type NAIP5 were transfected alone, or in different combinations with NLRC4 and 6x-Myc-FlaA, and analyzed as in **c**. (**e**) NLRC4ΔCARD was expressed alone, together with 6x-Myc-FlaA, or with both 6x-Myc-FlaA and NAIP5 for 48 hours, followed by sample preparation and analysis as in **c**. *NS denotes a non-specific band. (**f**) Wild-type NAIP5, NAIP5ΔPloop, and NAIP5Δ347 mutants were transfected into 293T cells in combination with NLRC4 and GFP-FlaA and analyzed as in **c**. The P loop is essential for the function of the nucleotide binding domain (NBD) and the Δ347 mutant truncates the N-terminal BIRs.

2.5. What is the function of the NLRC4 CARD domain?

The N-termini of NLRs have been proposed to determine interactions with downstream effector proteins, and therefore control the effector functions of inflammasomes. The CARD domain of NLRC4 has been shown to mediate homotypic interactions between the CARD domains of Caspase-1, ASC, and NLRC4 in overexpression systems[116,117]. We find that the CARD of NLRC4 is dispensable for
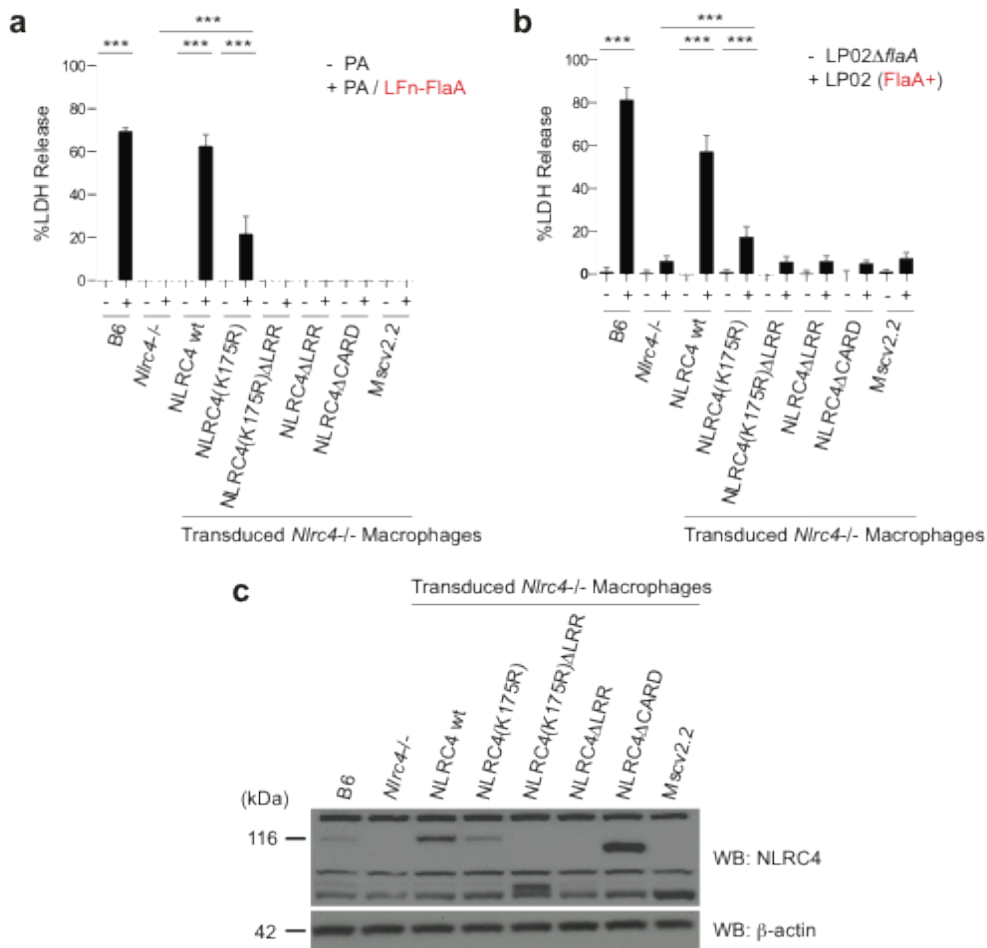
Figure 2.5. Complementation of immortalized *Nlrc4*$^{-/-}$ macrophages with wild-type and mutant *Nlrc4* alleles. (**a**) Immortalized *Nlrc4*$^{-/-}$ macrophages were stably transduced with retroviral vectors encoding wild-type, NLRC4(K175R), NLRC4(K175R)ΔLRR, NLRC4(ΔLRR), NLRC4(ΔCARD), or empty vector, and treated with PA alone or PA plus LFn-FlaA. Macrophage cell death was measured by LDH release. Immortalized B6 macrophages were used as a positive control. (**b**) Stably transduced macrophages were infected (MOI=2) with *Legionella pneumophila* (LP02) or flagellin-deficient *L. pneumophila* (LP02Δ*flaA*), and cell death was measured 4 hours later by quantifying LDH release. (**c**) Expression level of NLRC4 in B6, *Nlrc4*$^{-/-}$, and stably transduced macrophages, as measured by Western blot of whole cell lysates.

flagellin-dependent NAIP5-NLRC4 oligomerization (Fig. 2.4E), but absolutely required for Caspase-1 activation, cytokine processing, and pyroptosis (Fig. 2.5A & B; Appendix Two)[101,108]. This indicates that Caspase-1 recruitment to the NAIP-NLRC4 inflammasome requires the CARD of NLRC4. In contrast, ASC is required for cytokine processing in macrophages infected with various pathogenic bacteria, but is dispensable for pyroptosis[109].

2.6. What are the functions of the NAIP BIR domains?

A unique feature that distinguishes NAIPs from other NLRs is the presence of three N-terminal tandem baculovirus inhibitor of apoptosis repeat (BIR) domains, whose function remains elusive (Fig. 1.1A & B). The structurally conserved BIR repeat consists of a ~70 amino acid zinc finger-like motif that mediates protein-protein interactions in processes as diverse as mitosis, apoptosis, receptor signaling and ubiquitination[118-121]. Interestingly, BIR domains are complex protein-protein interaction motifs that contain multiple protein-interaction surfaces that can support dimerization[122] or interaction with a variety of protein structures such as coiled-coils[123], helical bundles[119], or NH$_2$-terminal tetrapeptide motifs[124]. BIR domain-containing proteins commonly regulate caspase activity, but intriguingly each individual BIR domain interacts with unique protein partners. For example, human XIAP contains three tandem BIR domains each with a unique function. The BIR1 domain is involved in dimerization[125], binding to TAB1[126], and initiating NF-kB activation; the BIR2 domain and its peptide linker directly inhibit Caspase -3[127] and Caspase-7[128,129]; whereas the BIR3 domain directly inhibits Caspase-9[130,131]. Interestingly, the BIR domains of human NAIP have been reported to bind TAB1 leading to MAP kinase (JNK) signaling that suppresses apoptosis independent of caspase inhibition[132,133]. However, these studies rely on overexpression of the NAIP BIR1-3 alone, and may not reflect the activity of the full-length proteins. In addition, crystal structures of human XIAP BIR2 reveal that the residues required for Caspase inhibition appear to be absent from the NAIP family paralogues[134,135]. Furthermore, BIR1 and BIR3 of hNAIP fail to bind NH$_2$-terminal tetrapeptides used to categorize BIR-containing family members[124]. The evolutionary divergence of NAIP BIR domains suggests novel interaction partners and functions have yet to be discovered.

Truncation of all three murine NAIP5 BIR domains results in an apparently non-functional protein that fails to assemble NAIP5-NLRC4 inflammasomes or activate Caspase-1 in response to cytoplasmic flagellin (Fig. 2.4B & F; Appendix Two)[101,108]. Individual NAIP5 BIR-deletions (NAIP5-ΔBIR1, -ΔBIR2, and -ΔBIR3) also fail to support NLRC4 oligomerization in response to FlaA (data not shown). Whether NAIP BIRs are required for proper protein folding, interaction with NLRC4, stabilization of ligand binding, or interaction with unknown effector proteins remains unanswered. It is likely that the BIR domains of the NAIPs interact with specific partners, and will be important going forward to clarify the role they play in regulating innate immunity.

2.7. What are the functions of LRR domains?

Leucine-rich repeats form binding surfaces for diverse molecular features including nucleic acids, peptidoglycan fragments, glycoproteins, lipo-proteins or proteins, and may also function in dimerization. Several TLR ectodomain-ligand structures have been solved, and demonstrate that ligand-interaction surfaces differ among TLRs. Ligand binding to LRRs is predicted to relieve auto-inhibition of the NBD, leading to conformational changes that permit nucleotide-biding, oligomerization, and recruitment of effector proteins[104]. NLRC4ΔLRR has been shown to spontaneously oligomerize and kill transfected cells (Fig. 2.4A & B; Fig. 2.5; Appendix Two)[101,117], similar to what is observed when the cytochrome-c-binding WD40 domain is deleted from Apaf-1[136]. In addition, truncation of the LRRs from NOD proteins leads to constitutive activation of an NF-kB promoter and a loss of ligand responsiveness[42,137]. Furthermore, work on plant R-proteins suggests that specificity for pathogen-derived

ligands map to narrow hypervariable regions of the LRR that are under intense diversifying selection[11,138-141]. Importantly, we found that the NAIP5ΔLRR mutant is capable of assembling a NAIP5ΔLRR/NLRC4 inflammasome independently of ligand, providing an epistasis argument for the LRR domain of NAIP5 binding flagellin and controlling inflammasome assembly[101,108]. Overall, the data support a model where the LRRs function in auto-inhibition of nucleotide binding, and perform a regulatory function important for preventing the consequences of unintentional inflammasome assembly.

2.8. Discussion

In this chapter we have outlined methodology for measuring the inflammasome, and applied these methods to address the basic questions: Do NLRs bind ligands directly?, and what controls inflammasome assembly? The answer is that flagellin specifically interacts with NAIP5 and NLRC4 in a megadalton inflammasome complex that can be visualized by blue native gel electrophoresis and by affinity co-immunoprecipitation. NAIP5 is absolutely required for assembly of the inflammasome in response to flagellin, and there appears to be direct interaction between flagellin and NAIP5.

Furthermore, mutational analysis of NLRC4 has shown that the CARD domain is dispensable for inflammasome assembly, but required for Caspase-1 activation and pyroptosis. Interestingly, the NLRC4 (K175R) mutant fails to assemble into the inflammasome, yet retains the functional ability (albeit reduced) to trigger pyroptosis both in the 293T cell assay and also in complemented macrophages. This raises the possibility that less stable complexes form between NAIP5 and NLRC4 (K175R), or that the NBD of NLRC4 is not required for interaction with NAIP5. In contrast, the NAIP5ΔPloop mutation is completely non-functional for inflammasome formation and pyroptosis. Together these results suggest that NAIP5 functions to bind ligand, whereas NLRC4 stabilizes the inflammasome complex and recruits Caspase-1.

Another mechanistic insight from these studies is that the LRR domains of both NAIP5 and NLRC4 are autoinhibitory. NLRC4ΔLRR spontaneously oligomerizes and induces pyroptosis, whereas NAIP5ΔLRR can spontaneously oligomerize with wild-type NLRC4 and stimulate pyroptosis. These results provide an epistatic argument for NAIP5 being upstream of NLRC4 in the recognition of flagellin, and also suggest that the LRR domain binds ligand directly and controls inflammasome assembly. The stoichiometry of the inflammasome components and the surface contacts between them remains speculative (Fig. 2.6). The precise structural detail of the inflammasome awaits crystallographic elucidation.

Here we have reconstituted the NAIP-NLRC4 inflammasome in a biochemically tractable heterologous system that recapitulates the genetic requirements of the native inflammasome. Four unique hallmarks of inflammasome activation can be easily monitored in this overexpression system: Inflammasome assembly (i.e. NLR oligomerization), Caspase-1 activation, IL-1β processing, and pyroptosis. The experimental approach outlined here is straightforward and can be used more broadly to investigate the assembly of multi-protein complexes.
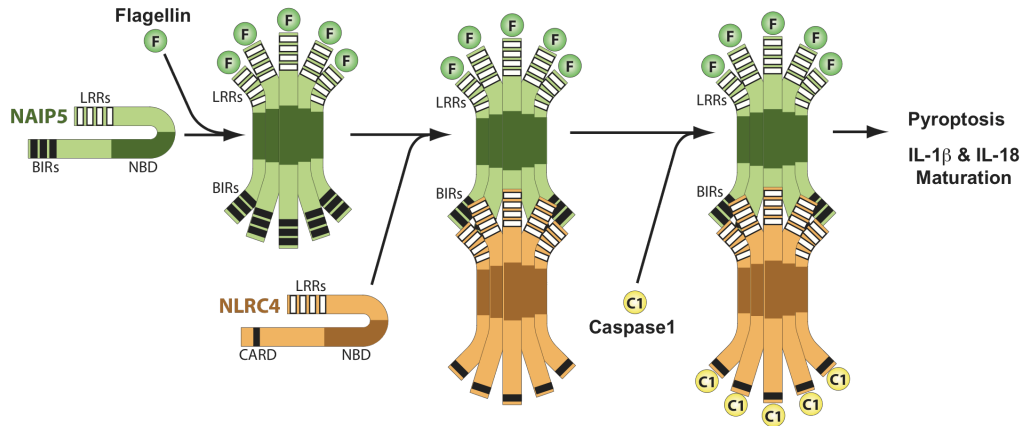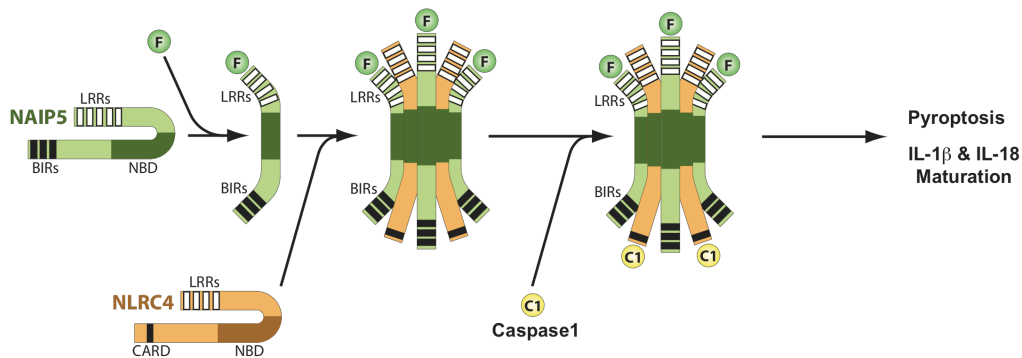
Figure 2.6. Speculative models of NAIP5-NLRC4 inflammasome assembly. In the absence of ligand, NAIP5 is maintained in an auto-inhibited state in the cytoplasm of innate immune cells. The LRR domain prevents spontaneous activation of NLRs in the absence of an infection and associated stimuli. The LRR domain of NAIP5 specifically binds bacterial flagellin in the host cytoplasm, autoinhibition is released, and the NAIP5-FlaA complex co-oligomerizes with NLRC4. The modular domains involved in each step of inflammasome assembly, and the stoichiometry of its components remains speculative. The stacked model **(A)** represents one explanation for why NAIP5Δ347 fails to assemble an inflammasome, and posits that the BIR domains are involved in binding NLRC4. Alternatively, truncation of the BIRs may have structural ramifications that prevent proper protein folding. The intercalated model **(B)** is favored by the assumption that the NBDs of NAIP5 and NLRC4 hetero-oligomerize in a manner conserved among AAA+ ATPase superfamily members.

Chapter Three:
NAIP paralogues determine inflammasome specificity

*Portions of this chapter are scheduled for publication in the journal BioEssays (2012), and permission to reproduce materials has been granted. Other portions of this chapter have been previously published in the journal Nature (2011)[101], and copyright of this material is retained by the co-authors Eric M. Kofoed and Russell E. Vance.

3.1. Introduction

L.pneumophila replication within mouse macrophages *ex vivo* was found to vary with genetic background[16], and was subsequently mapped to the Lgn1 interval on chromosome 13[18,19]. Fine mapping and expression analysis indentified the *Naip5* gene to be essential for the restriction of *L.pneumophila* replication in macrophages *ex vivo*, and that functional transcripts of *Naip1, Naip2, Naip5*, and *Naip6* exist in the C57Bl/6 mouse strain[23,24]. Because susceptibility mapped to a single paralogue, *Naip5*, it was considered unlikely that NAIPs are completely redundant, prompting the hypothesis that differential tissue/cell type expression and/or functional diversification of the NAIP locus[24,142]. The *Naip* locus is evolving rapidly[143], but whether this has functional consequences remains unknown.

During characterization of the FlaA-NAIP5-NLRC4 inflammasome described in *Chapter Two*, Ed Miao and colleagues reported host recognition of diverse bacterial pathogens via recognition of the inner rod component of type III secretion systems[48]. Macrophage response to inner rod components was completely dependent on NLRC4 and Caspase-1. These data fostered the hypothesis that other NAIP paralogues may be involved in detection of bacterial rod proteins, whereas NAIP5 was specifically required for detection of flagellin. Further support for this hypothesis was garnered by the observation that an inner rod protein from *S.typhimurium* (PrgJ) delivered to the cytoplasm of NAIP5-deficient macrophages caused pyroptosis and cytokine processing that required NLRC4 and Caspase-1[52]. NAIP5 was essential for flagellin-dependent NLRC4 activation, but secretion system rod components also operate through NLRC4, raising the question of how NLRC4 was capable of responding to multiple unique microbial ligands. Using the assays described above we tested each *Naip* paralogue for responsiveness to flagellin or PrgJ.

Ultimately, the *Naip* array provides examples of both functional diversity and redundancy. Moreover, we find that NAIP proteins are capable of recognizing broad classes of pathogenic bacteria. Thus, NLRC4 mainly serves an adaptor function for the NAIP family of proteins, important for recruitment of Caspase-1 into the inflammasome, but not for the specific detection of bacterial molecules. Therefore, the primary function of the NAIP-NLRC4 inflammasome is to form an innate immune barrier against broad classes of potentially virulent pathogens.

3.2. NAIP paralogues respond to bacterial protein ligands

Interestingly, and unlike flagellin, PrgJ could not induce formation of a high molecular weight complex containing NAIP5 and NLRC4 (Fig. 3.1A), and PrgJ expression did not result in death of 293T cells expressing NAIP5 and NLRC4 (Fig. 3.2)[101]. The lack of ability of a reconstituted NAIP5/NLRC4 inflammasome to respond to PrgJ is consistent with the observation that PrgJ activates NLRC4
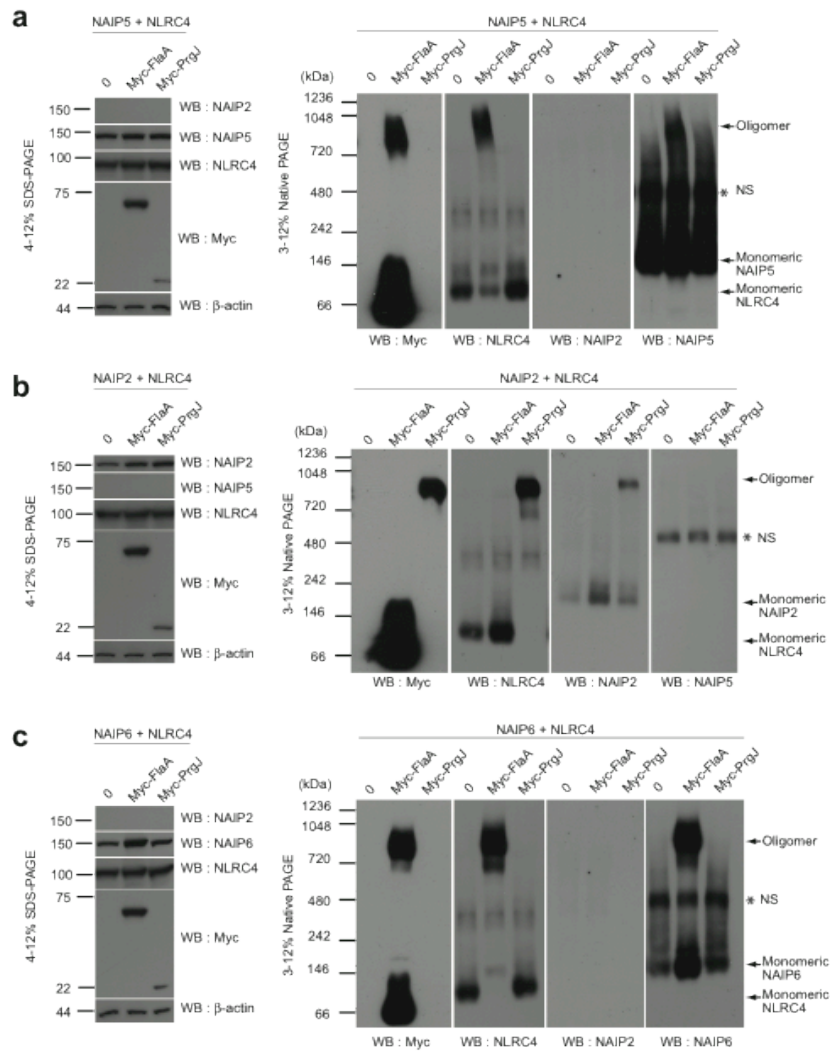
Figure 3.1. NAIP Paralogs Confer Specificity to the NLRC4 Inflammasome.
(**a**) 293T cells were co-transfected with wild-type NAIP5 and NLRC4, alone or in combination with 6x-Myc-FlaA or 6x-Myc-PrgJ followed by BN-PAGE 48 hours later. *NS, non-specific band. Whole cell lysates were also separated by conventional 4-12% SDS-PAGE to control for expression of each transfected gene construct (left panel). (**b**) 293T cells were transfected with wild-type NAIP2 and NLRC4 and analyzed as in **a**. (**c**) 293T cells were transfected with wild-type NAIP6 and NLRC4, and analyzed as in **a**.

independently of NAIP5 in macrophages[52], but raised the question of how PrgJ is recognized. One benefit of reconstituting NAIP5/NLRC4 inflammasome assembly in 293T cells is that it is technically straightforward to substitute components in order to identify factors that could mediate recognition of PrgJ. Cells expressing NAIP1 and NLRC4 did not respond to either flagellin or PrgJ (Fig. 3.3)[101,108], but interestingly, cells expressing NAIP2 and NLRC4 assembled a high molecular weight complex that contained NAIP2, NLRC4 and PrgJ (Fig. 3.1B)[101]. The additional expression of Caspase-1 resulted in death of the cells, indicating that the assembled NAIP2/NLRC4 inflammasome was functional (Fig. 3.2). NAIP2-NLRC4 did not respond to flagellin (Fig. 3.1B), whereas cells expressing NAIP6 and NLRC4 responded specifically to flagellin but not to PrgJ(Fig. 3.1C)[101,108]. The ability of NAIP6 to recognize flagellin
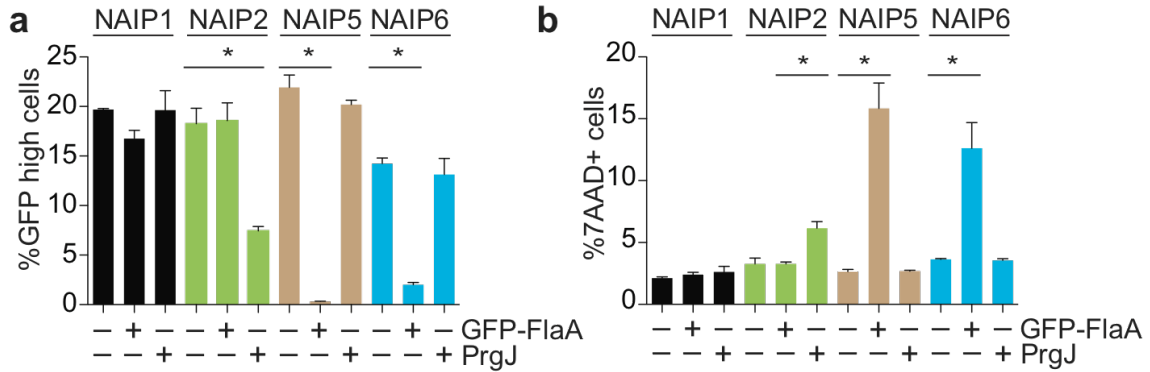
Figure 3.2. Reconstitution of the NAIP-NLRC4 inflammasomes in 293T cells. 293T cells were transfected with MSCV2.2-IRES-GFP expression vectors encoding NLRC4, Caspase-1, one of the NAIP paralogues, and either GFP-FlaA or PrgJ. Harvested cells were stained with 7AAD, a fluorescent dye that labels dead cells that have lost membrane integrity, and analyzed by flow cytometry for 7AAD and GFP. (**a**) GFP-high cells and (**b**) 7AAD positive cells were quantified as in Figure 2.1. Data shown are representative of at least three independent experiments. *, p<0.02 (Student's t test, two tailed); ns, not significant.
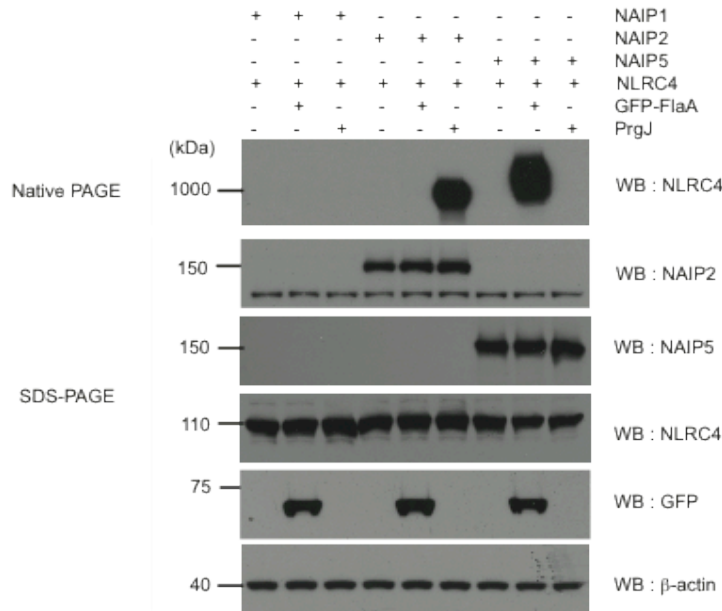
Figure 3.3. NAIP1 does not mediate NLRC4 inflammasome assembly in response to flagellin or PrgJ. 293T cells were co-transfected with vectors encoding NLRC4, NAIP1, NAIP2, NAIP5, GFP-FlaA, and/or PrgJ, followed by native gel electrophoresis and western blotting for oligomerized NLRC4 (top panel). Lysates were separated by SDS-PAGE and Western blotted for NAIP2, NAIP5, NLRC4, GFP, or β-actin to control for protein expression of each transfected inflammasome component (bottom panels). Data shown are representative of at least three independent experiments.

may account for the observation that *Naip5*$^{-/-}$ cells retain some responsiveness to flagellin[26].

Consistent with the above results, work by Feng Shao's group showed that NAIP5 and NAIP6 co-immunoprecipitated FlaA when expressed together in 293T cells, whereas NAIP1, NAIP2, and NLRC4 failed to pull down FlaA in the same assay[108]. A yeast two-hybrid interaction assay also confirmed that FlaA from *L. pneumophila* binds specifically to NAIP5 and NAIP6, but not NAIP1, NAIP2 or NLRC4[108]. In contrast, the *B. thailandensis* inner rod protein BsaK (a PrgJ homolog) interacted with NAIP2, but not NAIP1, NAIP5, NAIP6, or NLRC4[108]. Taken together these results provide multiple lines of evidence that NAIPs function to recognize specific bacterial proteins: NAIP5 and NAIP6 selectively recognize flagellin whereas NAIP2 selectively recognizes inner rod proteins such as PrgJ. Once activated, NAIPs appear to be critical for inducing assembly of an inflammasome that activates Caspase-1 (Fig. 5.1).

3.3. NAIP1 and NAIP2 recognize bacterial T3SS inner-rod proteins

To expand on the finding that NAIP2 specifically recognizes the type III secretion system inner-rod protein PrgJ found in *S.typhimurium,* we obtained a number of rod genes cloned into a pMXs-IRES-GFP retroviral vector (courtesy of Ed Miao) that we analyzed using the 293T NLRC4 oligomerization and cell death assays. Each inner-rod protein was derived from dedicated secretion systems important for pathogen virulence, and several were previously shown to be toxic to C57Bl/6 macrophages when expressed retrovirally[48]. Surprisingly, we found that NAIP1 appears to recognize a narrow subset of inner-rod proteins when compared to those detected by NAIP2 (Fig. 3.4). When NAIP1 was co-transfected with NLRC4 and either PscI (*P.aeruginosa*), YPO 0262 (*Y.pestis*), or BsaK (*B.pseudomallei*) (Fig. 3.4A), we observed the formation of high molecular weight inflammasomes that were also functional in the 293T cell death assay (Fig. 3.4B). NAIP1 failed to respond to flagellin (FlaA) and PrgJ, as shown above. In contrast, NAIP2 responds to the majority of inner rod proteins tested with the exception of SsaI (*S.typhimurium*)(Fig. 3.6C) and A1902 (*B.pseudomallei*)(Fig. 3.4C & D). The study of NAIP1 function is complicated by its redundancy with NAIP2, and by the fact that NAIP1 is not expressed in BMDMs. Instead NAIP1 appears to be highly expressed in resident peritoneal macrophages and the thymus (Fig. 3.5). Delivery of PrgJ, YPO0262, BsaK, and PscI to primary macrophages using engineered *Listeria* strains (see below) confirmed that these inner-rod proteins activated pyroptosis independent of *Naip5*, but nevertheless required *Nlrc4* (Appendix Five). These data have several important implications for how we understand NLR activation.

First, the finding that NAIP1 recognizes a small subset of rods that NAIP2 recognizes suggests that subtle amino acid changes in the ligand binding domains (i.e. the LRRs) must underlie these differences. Second, the inner-rod proteins examined are not highly similar in terms of overall amino acid identity (Appendix Fourteen), which raises the question of exactly what common features are detected. It has been proposed that the C-termini of both flagellin and inner-rod proteins are sufficient to bind the concave surface of the LRRs, because mutations in the C-terminus of FlaA[26], or truncation of the last seven amino acids of PrgJ render these proteins inactive[48]. Alternatively, there may be conserved structural features of the NAIP ligands that are sufficient to cause inflammasome assembly that are not apparent at the level of amino acid identity.

Finally, the magnitude of cell death in the 293T assay is not perfectly commensurate with the magnitude of oligomerization seen in the blue native assay (Fig.
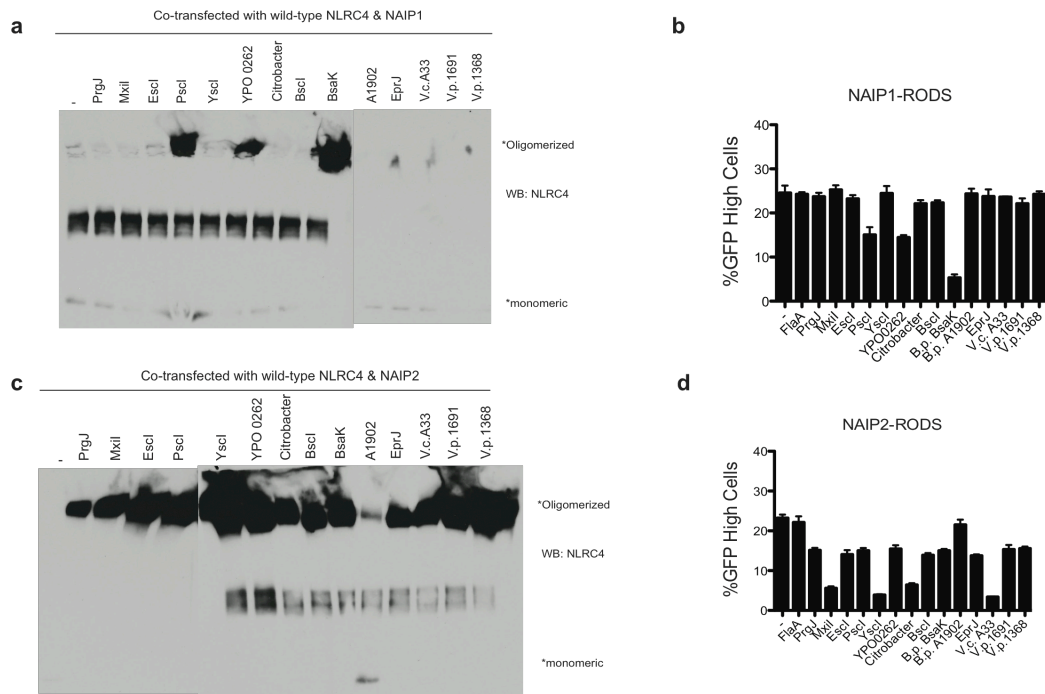
Figure 3.4. NAIP1 detects a subset of inner-rod proteins that are recognized by NAIP2.
(a) 293T cells were co-transfected with NLRC4, NAIP1, and either PrgJ, MxiI, EscI,
PscI, YscI, YPO0262, Citrobacter, BscI, BsaK, A1902, EprJ, V.c.A33, V.p.1691, or
V.p.1368 for 48hrs followed digitonin lysis, blue native gel electrophoresis and Western
blotting for NLRC4. Monomeric and oligomerized NLRC4 are denoted. **(b)** 293T cells
were transfected with MSCV2.2-IRES-GFP expression vectors encoding NLRC4,
Caspase-1, NAIP1 and either FlaA or inner rod proteins listed in **(a)**. Harvested cells
were analyzed by flow cytometry for GFP, and GFP-high cells were quantified as in
Figure 2.1. **(c,d)** 293T cells were treated as in **(a)** except NAIP2 was included in the
transfection. Results are representative of three independent experiments. *Some blots
show non-specific antibody staining in the middle of the native gels.
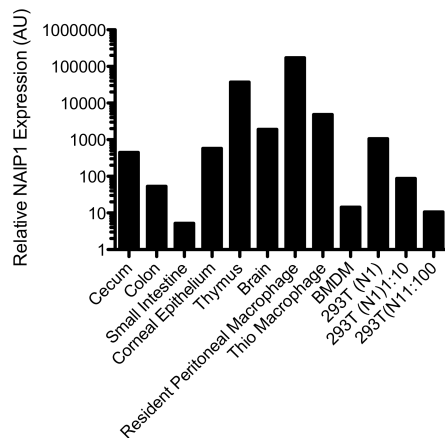
Figure 3.5. Tissue expression of NAIP1
determined by qPCR. Total mRNA was extracted
from flash frozen tissue samples using Trizol,
reverse transcribed and the expression of NAIP1
transcripts was assessed using TaqMan qPCR and
primers previously described[26]. 293T (N1)
dilution series are positive controls derived from
transient transfections. *Results require
verification as they are from a single experiment
performed using technical triplicates.

3.4C & D). If we assume that the ligands are structurally, and therefore qualitatively
different, then it is likely that unique ligands can generate unique conformations of the

inflammasome. We speculate that these qualitative differences may translate into quantitative differences in Caspase-1 recruitment or pyroptosis.

3.4. NAIP5 and NAIP6 recognize bacterial flagellin

NAIP5 and NAIP6 failed to recognize any T3SS inner-rod proteins that we examined, but respond robustly to flagellin in both 293T cell assays (Fig. 3.6). These data show the high degree of specificity that NAIP5 and NAIP6 have for flagellin. It will be useful to understand how this specificity is achieved on the molecular level, given the superficial structural similarity between the C-termini of flagellin and inner-rod proteins, despite low primary amino acid identity (Appendix Fourteen).
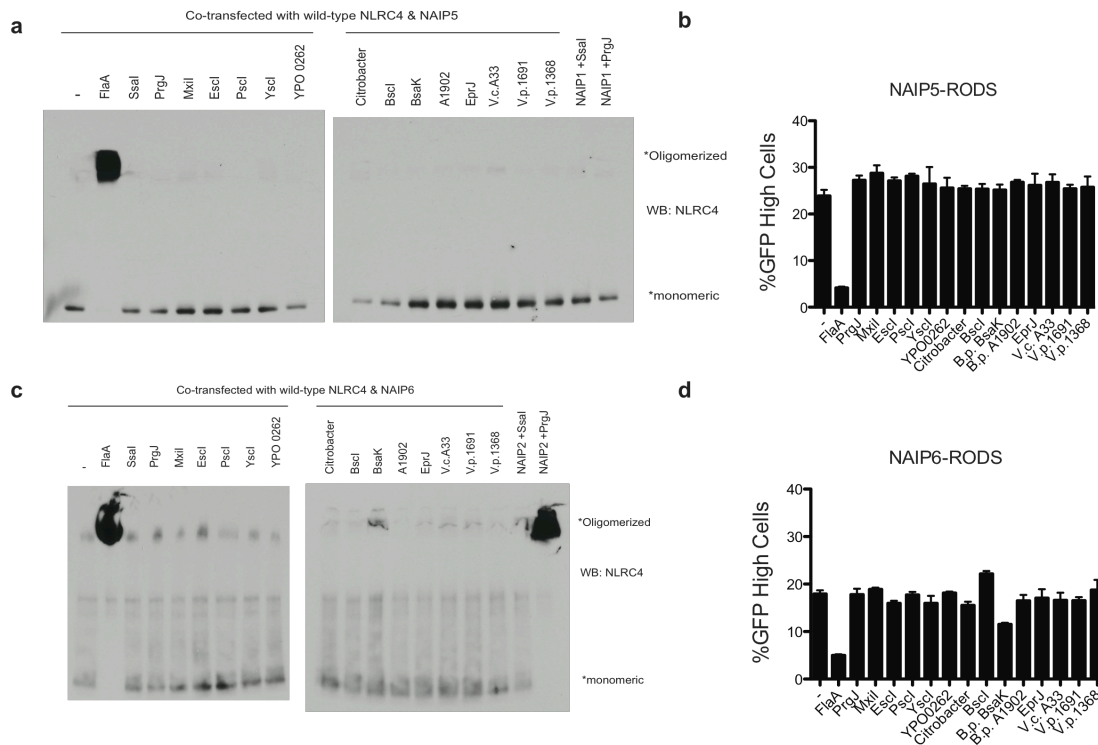
Figure 3.6. NAIP5 and NAIP6 do not detect inner-rod proteins recognized by NAIP2. **(a)** 293T cells were co-transfected with NLRC4, NAIP5, and either PrgJ, MxiI, EscI, PscI, YscI, YPO0262, Citrobacter, BscI, BsaK, A1902, EprJ, V.c.A33, V.p.1691, or V.p.1368 for 48hrs followed digitonin lysis, blue native gel electrophoresis and Western blotting for NLRC4. Monomeric and oligomerized NLRC4 are denoted. NAIP1 was also tested for response to SsaI and PrgJ on the same blot, and was negative. **(b)** 293T cells were transfected with MSCV2.2-IRES-GFP expression vectors encoding NLRC4, Caspase-1, NAIP5 and either FlaA or inner rod proteins listed in **(a)**. Harvested cells were analyzed by flow cytometry for GFP, and GFP-high cells were quantified as in Figure 2.1. **(c,d)** 293T cells were treated as in **(a)** except NAIP6 was included in the transfection. NAIP2 was tested for response to SsaI and PrgJ and responded only to PrgJ (as expected). Results are representative of three independent experiments. *Some blots show non-specific antibody staining in the middle of the native gels.

3.5. Role of endogenous NAIP5 and NAIP2 in innate immunity

Genetic evidence indicates a requirement for NAIP5, NLRC4 and Caspase-1 in restricting the replication of *L.pneumophila* in *ex vivo* macrophages that is dependent on expression of bacterial flagellin[26,40]. NAIP5 deficiency renders macrophages unable to activate Caspase-1, undergo pyroptosis (Fig. 3.7D) or process pro-IL-1β after infection with *L.pneumophila*[26,40]. To determine the involvement of endogenous NAIP2 in recognition of bacterial ligands, we decided to take a short hairpin RNA knockdown approach to stably inhibit expression of NAIP2 in primary bone-marrow derived macrophages (BMDMs). N2shRNA#1 and #2 specifically reduced NAIP2 protein levels without targeting other NAIP paralogs, whereas empty vector, shRNA#3 or a scrambled control shRNA had little effect on NAIP2 protein levels (Fig. 3.7). Previously we
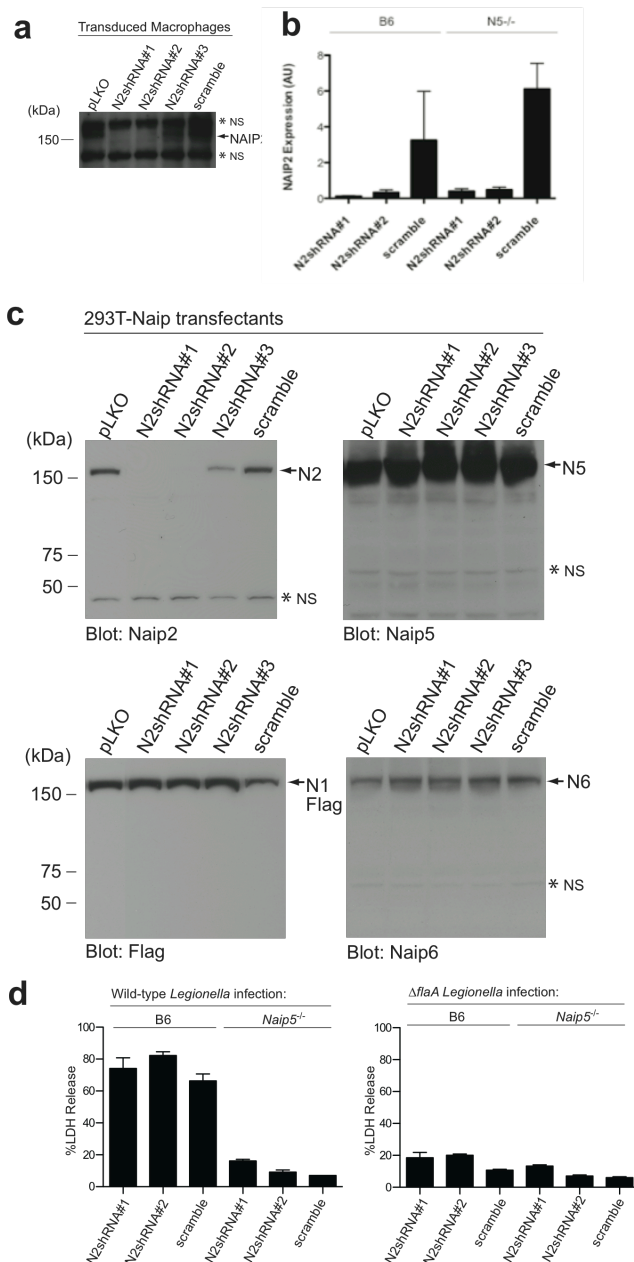
Figure 3.7. NAIP2 knockdown specifically affects NAIP2 expression and does not affect expression of other NAIP proteins. **(a)** Immortalized C57/Bl6 macrophages were stably knocked down for NAIP2 and cell lysates subjected to Western blot. **(b)** qPCR for NAIP2 transcript levels in primary NAIP2 knockdown BMDMs. **(c)** HEK293T cells were co-transfected with NAIP expression vectors and pLKO.1 control vector, the NAIP2 shRNA constructs (NAIP2 shRNA#1-#3), or the pLKO.1-scramble negative control for 48h, followed by western blotting for expression of NAIP proteins. No anti- NAIP1 antibody exists, so a Flag-tagged NAIP1 construct was employed. **(d)** NAIP2 knockdown cells were infected with wild-type or flagellin-deficient Legionella pneumophila (MOI=2) and LDH release was measured 4 hours after initial infection. Data shown +/- s.d. are representative of at least three independent experiments.

showed that endogenous NAIP5 deficiency does not alter responsiveness to PrgJ[52]. Here we show that endogenous NAIP2 deficiency has no effect on the flagellin-activated inflammasome that requires NAIP5, as shown by infection with *L.pneumophila* (Fig. 3.7D).

Macrophages expressing these shRNAs were then infected with flagellin-deficient *Listeria* strains that inducibly express PrgJ (*Listeria*-PrgJ) or flagellin (*Listeria*-FlaA)[52]. A *Listeria*-based system was chosen because it is an efficient means for delivering PrgJ to macrophages, and because it allows for controlled comparisons of PrgJ and FlaA within a single experimental system. Remarkably, knockdown of *Naip2* prevented pyroptosis and Caspase-1 activation by *Listeria*-PrgJ (Fig. 3.8A & C). By contrast, *Naip2* knockdown did not affect inflammasome activation by *Listeria*-FlaA (Fig. 3.8B & C) or *L. pneumophila*, which expresses flagellin but not PrgJ (Fig. 3.7D). Instead flagellin-dependent inflammasome activation depended on *Naip5*, as previously shown[26,29,39,40,52]. Inflammasome activation by wild-type *Salmonella*, which encodes both flagellin and PrgJ, was not significantly affected by *Naip2* knockdown (Fig. 3.8D & E). However, knockdown of *Naip2* in *Naip5*$^{-/-}$ macrophages significantly reduced or abolished inflammasome activation by wild-type *Salmonella* (Fig. 3.8D & E), indicating that both NAIP2 and NAIP5 recognize *Salmonella*. Interestingly, inflammasome activation by flagellin-deficient (FliC/FljB$^-$) *Salmonella*, which still express PrgJ, depended entirely on *Naip2* (Fig. 3.8D & E). Taken together, these data indicate that *Naip2* is specifically required for activation of the NLRC4 inflammasome by PrgJ, in contrast to *Naip5*, which appears to be specifically required for NLRC4 activation by flagellin (see Appendix Three).

This result demonstrated that *S.typhimurium* contains two unique protein ligands, flagellin and PrgJ, that are engaged by NAIP5 and NAIP2 respectively, that are both capable of activating an NLRC4 inflammasome. Furthermore, recently published data confirm that NAIP2 is capable of detecting the type III secretion system inner-rod protein BsaK from *B.thailandensis*, and inner-rods from a variety of other bacterial pathogens[108]. The broad spectrum of protection imparted by NAIP family paralogues highlight an important evolutionary strategy for detection of intracellular bacterial pathogens. Cytosolic surveillance for flagellin or T3SS inner-rod proteins (both conveyed through secretory apparatuses) suggests that the host is protecting itself from bacteria that translocate protein effectors using secretion systems.

3.6. Discussion

Despite their initial discovery as genes regulating intracellular replication of a single bacterial pathogen (*L. pneumophila*), it is now clear that NAIPs have broad roles in innate immunity, as the flagellin and inner rod ligands for NAIPs are broadly expressed by diverse bacterial species (Appendix Three). However, the above work has all been performed in the mouse. The *NAIP* gene locus in humans has not been thoroughly characterized and may be under copy number variation[144-147]. Nevertheless, humans appear to express at least one functional NAIP. This protein was shown to recognize neither flagellin nor the inner rod (PrgJ-like) proteins of T3SS[108]. Importantly, however, it instead appears that human NAIP can detect the needle subunit of several different T3SSs, including PscF from *P. aeruginosa*, PrgI from *S. typhimurium* (Fig. 2A), and MxiH from *S. flexneri*[108]. Although the sequence homology among flagellin, T3SS rods (e.g., PrgJ), and T3SS needles (e.g., PscF), is very limited, these NAIP ligands do all share a common structural function, which is essentially to form a hollow protein
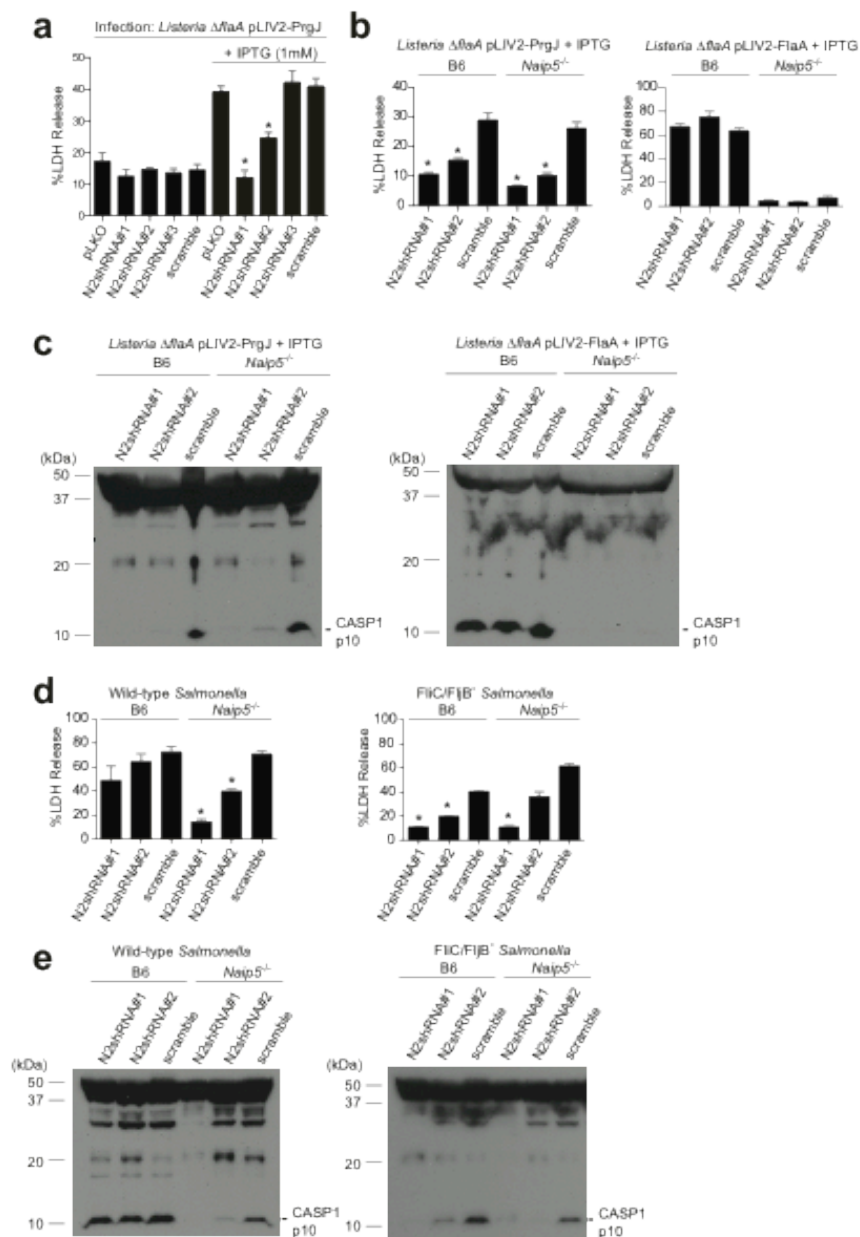
Figure 3.8. NAIP2 is required in macrophages for inflammasome activation in response to PrgJ. (**a-c**) Primary bone-marrow derived macrophages expressing shRNAs targeting NAIP2 (or controls) were infected with flagellin-deficient *Listeria monocytogenes* (MOI = 5) expressing a secreted ActA100-PrgJ (pPrgJ) or ActA100-FlaA (pFlaA) fusion protein under IPTG-inducible control. (**a, b**) Cell death (± s.d.) was measured in triplicate by LDH release 6 hours after infection, or (**c**) Active CASP1 (p10) was measured by western blotting of cell supernatants. (**d, e**) NAIP2 knockdown cells were infected with wildtype or flagellin-deficient (FliC/FljB⁻) *Salmonella* Typhimurium and inflammasome activation was measured by (**d**) LDH release (± s.d.) at 3h after infection or (**e**) CASP1 processing. Data shown are representative of two (c, e) or three (a, b, d) independent experiments. *, p<0.02 as compared to scramble (Student's t-test, two-tailed).

29

channel through which bacteria can secrete proteins. Interestingly, the data suggest that it is the monomeric form of flagellin (and likely, the monomeric form of the other NAIP ligands) that is detected[26,47]. Nevertheless, there may be something unique about the structure of proteins that polymerize to form a channel that is not found in self-proteins and that is therefore readily exploited for detection by the immune system. The structural basis by which NAIP ligands are recognized is an exciting avenue for future studies.

  A major caveat of the above biochemical studies is of course the fact that they have been performed in heterologous cell types and have employed overexpressed proteins. A general worry about overexpression is that it tends to amplify non-specific interactions. This worry is partially offset by the fact that the experiments described above contain ideal built-in specificity controls, e.g., NAIP5 responded to flagellin but not PrgJ, whereas NAIP2 responded to PrgJ but not flagellin. Worries about overexpression artifacts are also offset by genetic validation of the results in primary bone-marrow derived macrophages. *Naip5* knockout macrophages were selectively defective in responses to flagellin, but responded normally to PrgJ, whereas *Naip2* knockdown selectively affected responsiveness to PrgJ but not flagellin. Nevertheless, it will be important in future studies to develop methods to analyze the formation, specificity, and stoichiometry of native inflammasomes in a relevant cell type (e.g., macrophages). Finally, given the high degree of primary amino acid identity between the NAIP paralogues (Fig. 1 & Appendix Thirteen) and the panoply of bacterial ligands available, we believe the structural basis of ligand recognition will be an especially productive area of future research, and may reveal conserved mechanisms of NLR ligand recognition.

Chapter Four:
Mapping ligand specificity of the NAIP proteins

4.1. Introduction

The high degree of sequence similarity[23,148], and divergent ligand specificity[101,108] among NAIP paralogues presents the opportunity to map the surfaces on NAIP paralogues required for the interaction with cognate bacterial ligand. We chose to take a chimeric molecule approach to determine what NAIP domain/s are involved in ligand-binding and specificity. This approach has seen previous success in the plant NB-LRR literature[141], and also for some mammalian NLRs[109]. Inherent challenges of generating chimeric proteins are constraints on the choice of breakpoint, a need to avoid splitting functional modular domains, and generation of a functional and stable protein. Nevertheless, NAIPs exhibit significant primary sequence identity that makes the chimeric approach possible. The questions we are interested in answering using this approach are straightforward: 1) Can the minimal ligand specificity-determining region of the NAIPs be defined using a panel of chimeric molecules in the NLRC4 oligomerization assay? 2) Once the region of NAIPs that confers specificity is mapped, is expression of this minimal domain sufficient to interact with bacterial ligand? 3) Can targeted mutagenesis of the ligand-binding domain swap ligand specificity of the NAIP? The answer to these simple questions will greatly improve our mechanistic insight of how NLRs come to recognize specific bacterial ligands, and will shape the strategy for generating soluble NAIP ligand-binding domains for crystallographic analysis. The paucity of structural information available for mammalian NLRs is the major stumbling block in the inflammasome field. Our understanding of NLR regulation is based on analogy to other proteins (e.g. Apaf) that are not necessarily related functionally or structurally. Thus, understanding how NAIPs bind specific bacterial ligands will establish a mechanistic precedent for inflammasome-forming NLRs.

4.2. NAIP5-NAIP2 chimeras identify ligand specificity-determining region

In order to determine the region of NAIP5/6 that confers responsiveness to flagellin (e.g. FlaA) and the region of NAIP2 that confers responsiveness to inner-rod proteins (e.g. PrgJ), we made chimeric cDNAs that encode novel chimeric proteins by splicing of overlapping ends polymerase chain reaction[149]. We chose ten breakpoints located along the entire length of NAIP2, NAIP5 and NAIP6, and focused these breaks between known boundaries of modular domains (e.g. BIR1, BIR2, BIR3, NBD). Because the LRRs of NAIPs are potentially large (~800 amino acids) we chose breakpoints every 100 amino acids as long as identical stretches of primary sequence could be found. In fact, we found that modeling of the LRR domains using the prediction software PHYRE gave high confidence LRR structure between amino acids 1006-1447 for NAIP2, and between amino acids 961-1403 for NAIP5/6 (Fig. 1.1C). Yet, the region between the end of the NBD and the start of the LRR domain was uncharacterized both according to PHYRE and SMART domain prediction software. After mapping the ligand-specificity domain (see below) to this region, a rotation student in the lab, Jeannette Tenthorey, discovered that the region is also predicted to form LRRs (Fig. 4.4).
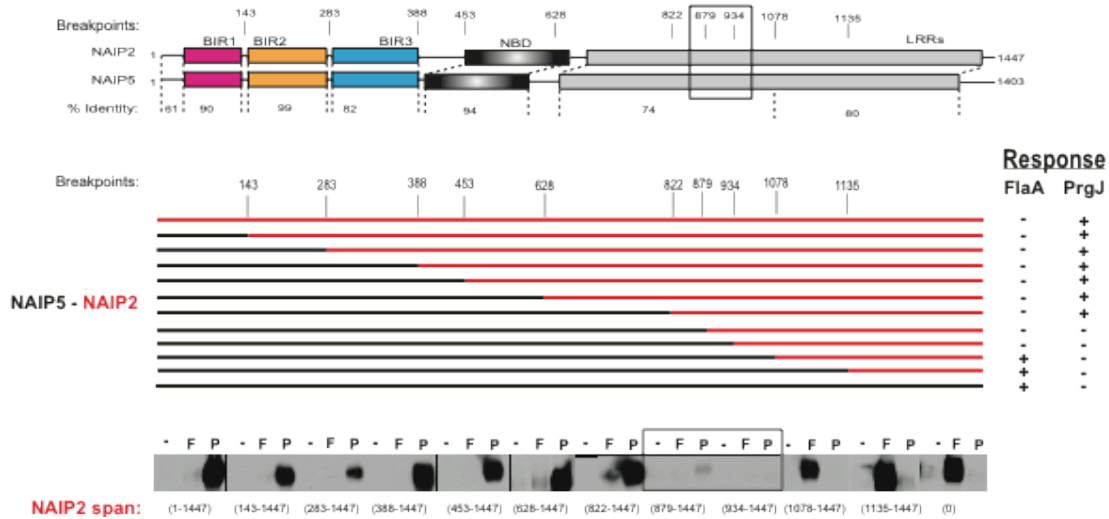
Figure 4.1. Novel chimeric NAIP5-NAIP2 molecules identify a narrow portion of the LRR domain (822-1078) responsible for conferring ligand-specificity to NAIPs. Ten different NAIP5-NAIP2 chimeric cDNAs were generated using splicing of overlap extension PCR and co-expressed as chimeric fusion proteins in 293T cells with NLRC4 alone (-), and in combination with 6x-Myc-FlaA (F) or 6x-Myc-PrgJ (P) and assayed for NLRC4 oligomerization using NativePAGE and Western blotting for NLRC4. Only the oligomerized NLRC4 (~1MDa) complex is shown for each chimera tested. Amino acids of NAIP2 for each chimera (i.e. - NAIP2 span) are listed below panels of the Western blot (NAIP2 contains a unique 44 amino acid linker between BIR3 and the NBD that is absent from NAIP5; therefore precise amino acids of NAIP5 in each chimera are determined by subtracting 44 from the indicated NAIP2 sequence (e.g. breakpoint 822 in NAIP2 will connect with amino acid 778 in NAIP5). Protein expression and loading controls can be found in Appendices Six through Eleven.

Thus, the NAIP LRRs are twice as large as originally predicted, highlighting a major caveat of relying on domain prediction algorithms.

NAIP5-NAIP2 chimeras show that ligand-specificity maps between amino acids 822-1078 of NAIP2, and 778-1034 of NAIP5/6. The data show that a protein that encodes NAIP5 from amino acid 0-778 still responds to PrgJ, but if it is NAIP5 from 0-1034 it responds to FlaA (Fig. 4.1). The two chimeras between these breakpoints (879 and 934) respond to neither ligand and are therefore non-functional for ligand recognition, but are nevertheless expressed to the same extent as other chimeras. These data clearly demonstrate that amino acids 822-1078[Naip2] & 778-1034[NAIP5] are necessary for determining ligand specificity. We are currently testing whether this region is sufficient for binding specific bacterial ligands. There may be other regions that contribute to ligand binding, but the region identified confers the dominant ability to distinguish between ligands. What features of bacterial ligands are recognized by NAIPs

Figure 4.2. Novel chimeric NAIP2-NAIP6 molecules identify a narrow portion of the LRR domain (657-1078) responsible for conferring ligand-specificity to NAIPs. Nine different NAIP2-NAIP6 chimeric cDNAs were generated using splicing of overlap extension PCR and co-expressed as chimeric fusion proteins in 293T cells with NLRC4 alone (-), and in combination with 6x-Myc-FlaA (F) or 6x-Myc-PrgJ (P) and assayed for NLRC4 oligomerization using NativePAGE and Western blotting for NLRC4. Only the oligomerized NLRC4 (~1MDa) complex is shown for each chimera tested. Expression and loading controls can be found in Appendices Six through Eleven.

is an important focus of future investigation.

4.3. NAIP2-NAIP6 chimeras confirm ligand specificity determining region

Similar to the approach used to characterize the NAIP5-NAIP2 chimeras, the 293T cell-based NLRC4 oligomerization assay was applied to NAIP2-NAIP6 chimeras, which were cloned using the same breakpoints described above (Fig. 4.1). The data confirm the region of ligand specificity maps to a region between amino acids 657 and 1078 of NAIP2. We did not obtain correct clones for breakpoints 822 or 934 for the NAIP2-NAIP6 chimeras, which is why the region identified is slightly larger. The NAIP2-NAIP6 chimera with breakpoint at amino acid 879 of NAIP2 responds to neither ligand, supporting the notion that this area is critical for a coherent and unified response to specific ligands. Overall, NAIP2-NAIP6 chimeras that are NAIP2 (1-657) respond to FlaA, but when NAIP2-NAIP6 chimeras are NAIP2 (1-1078) or more, they respond to PrgJ (Fig. 4.2). Lastly, it appears that there may be partial responsiveness to both ligands using NAIP2-NAIP6 chimeras NAIP2 (1-628) and NAIP2 (1-657), but this observation requires confirmation.

4.4. NAIP6-NAIP2 chimeras confirm ligand specificity determining region

Finally, we generated a complete panel of NAIP6-NAIP2 chimeras that confirm the ligand-specificity determining region to be located between NAIP2 amino acids 822-1078 (Fig. 4.3). Analysis and interpretation of these results are as described in sections 4.2 and 4.3 above. Note that we also generated NAIP2-NAIP5 chimeras, but these novel molecules were non-functional in our assays (data not shown). Why this is true for only this set of clones is open to speculation.
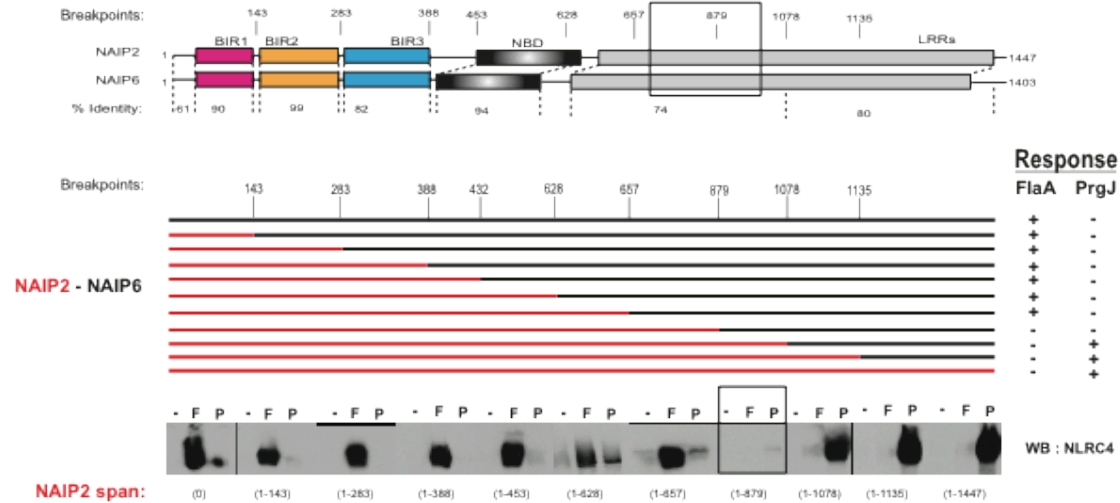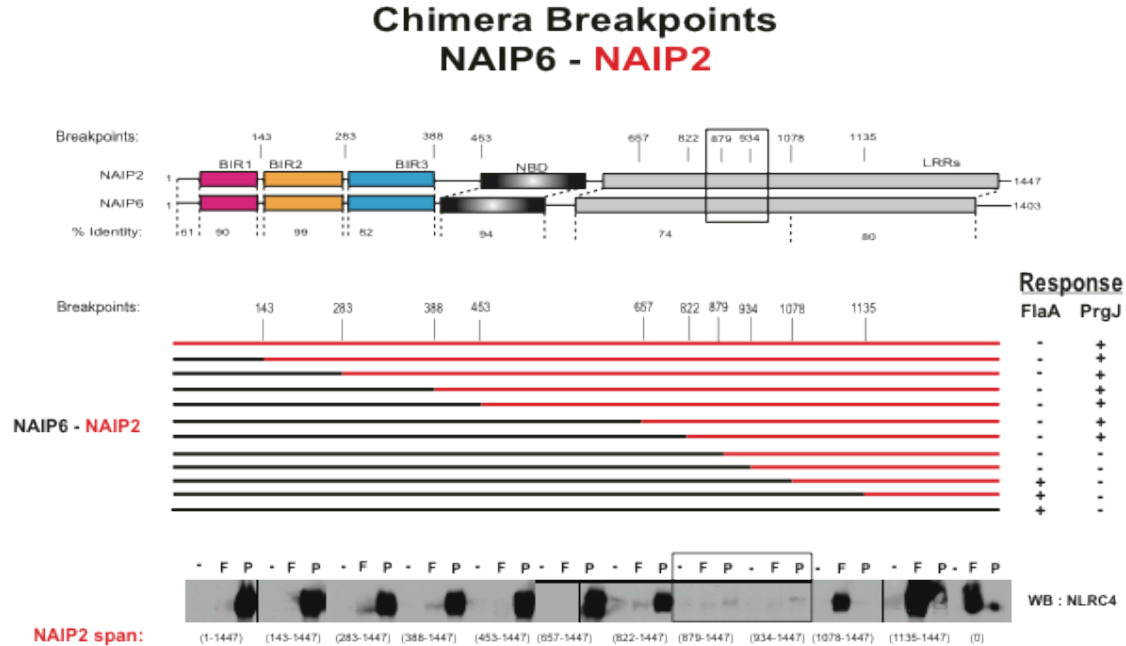

Figure 4.3. Novel chimeric NAIP6-NAIP2 molecules identify a narrow portion of the LRR domain (822-1078) responsible for conferring ligand-specificity to NAIPs. Ten different NAIP6-NAIP2 chimeric cDNAs were generated using splicing of overlap extension PCR and co-expressed as chimeric fusion proteins in 293T cells with NLRC4 alone (-), and in combination with 6x-Myc-FlaA (F) or 6x-Myc-PrgJ (P) and assayed for NLRC4 oligomerization using NativePAGE and Western blotting for NLRC4. Only the oligomerized NLRC4 (~1MDa) complex is shown for each chimera tested. Protein expression and loading controls can be found in Appendices Six through Eleven.

4.5. Putative NAIP-ligand interaction surface

Polymorphisms between NAIP paralogues cluster on the concave surface of the traditional solenoid shaped LRR (Fig. 4.4). Non-synonymous substitutions on the β-strands appear to be a common principle in the evolution of ligand recognition by LRR domains[63,64,150]. In fact, polymorphic residues between NAIP paralogues cluster to the β-strands all throughout the LRR, including the 3'-LRR region 1006-1447 of NAIP2 (Fig. 1.1C) that is not involved in specificity (see also Appendix Four). Nevertheless, the most important region for ligand binding (822-1078) has a greater number of
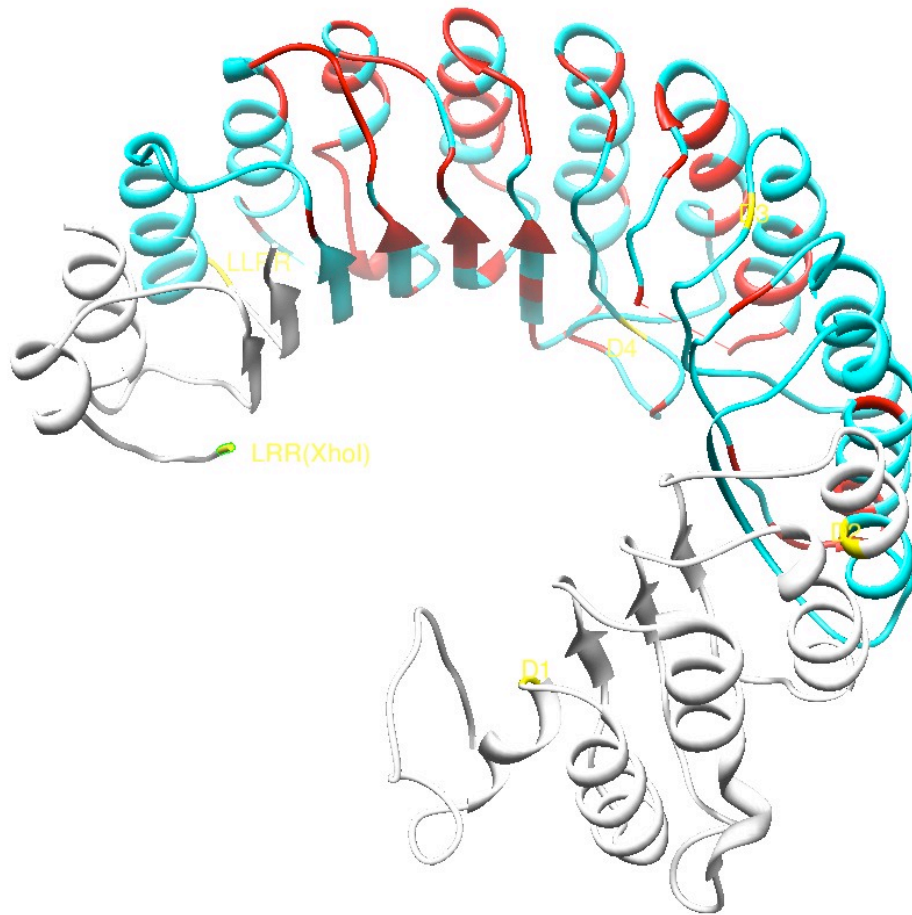
Figure 4.4. Structural model of the NAIP2 ligand-binding domain and polymorphisms between NAIP2 and NAIP5. The region that determines ligand-specificity is colored cyan, and polymorphisms between NAIP2 and NAIP5 are indicated in red. Yellow residues indicate the chimera breakpoints used to define the ligand specificity-determining region (822-1078^{NAIP2}). They correspond to amino acids of NAIP2: LRR (XhoI), 1135; LLRR, 1078; D4, 934; D3, 879; D2, 822; D1, 657. There is polymorphism in three adjacent β-strands on the concave surface and also their corresponding loops. Analogous features have been shown to be essential for contacting the D1 domain of flagellin in the TLR5 ectodomain/flagellin crystal structure, and indicate a likely site of interaction between flagellin and the NAIP5 LRR domain. Jeannette Tenthorey performed the modeling and polymorphism mapping.

polymorphisms (see Appendix Twelve).

Based on homology modeling using the PHYRE software we were able to generate a predicted NAIP2 LRR structure covering amino acids 657-1135 that encompasses the mapped ligand-specificity determining region of NAIPs (Fig. 4.4). The structure that is colored cyan denotes the experimentally determined minimal ligand-binding domain (Fig. 4.1-4.3). Polymorphisms between NAIP2 and NAIP5 are

highlighted in red on the ribbon diagram (Fig. 4.4). Polymorphic residues may contribute to ligand binding, but this hypothesis awaits experimental verification. These predicted structure models are important tools for the rational design of targeted mutagenesis to swap ligand specificity or to engineer gain or loss of function alleles. The polymorphism mapping also suggest that the loops above the variable β-strands may be important in ligand binding and specificity. Surprisingly, the structure of the TLR5 ectodomain LRR binding to D1 region of flagellin show that these loops are required for high-affinity interactions[7]. We predict that these variable and flexible loops mapped to the NAIP LRR model (Fig. 4.4) are critical for ligand binding and/or determining specificity.

4.6. Discussion

The LRR domains of NAIPs bind directly to specific bacterial ligands. This result is not unexpected, and has been widely assumed to be the mechanism of ligand recognition among NLRs. Yet direct evidence for LRR interactions with specific stimuli are lacking. Here we have demonstrated using reciprocal chimeric molecules that the response to flagellin is conferred by the amino acids 778-1034 of NAIP5 and NAIP6, whereas the response to PrgJ is dictated by amino acids 822-1078 of NAIP2. This part of the LRRs is functionally the most important for ligand binding and specificity determination. Other regions may contribute to binding but are less critical. The NAIP LRR modeled here resembles the structure of the TLR5 LRR bound to flagellin[7] and highlights several important functional and structural properties of LRRs in general.

First, LRRs are simple β-strand-α-helix repetitive units that can be assembled into large solenoid-like scaffolds that typically have a hydrophobic parallel β-sheet lining the concave surface and α-helices lining the convex surface. The LRRs provide a framework upon which variable loops connect each β-strand to each α-helix, and the solvent-exposed side chains on the concave surface are evolutionarily malleable[140,150,151]. It has been proposed that variable loops of LRRs could function analogous to the complementarity-determining regions of immunoglobulin[150]. However, the hydrophobic concave surface is the most adaptable and variable in plant NBS-LRR proteins[100,138-140], and also in the VLRs of agnathes[64].

Second, the LRR domains are non-globular, and therefore have large solvent exposed surfaces that are more amenable to favorable binding interactions versus, for example, small and sterically occluded binding pockets common to steroid receptors (e.g. estrogen receptor)[152]. The LRR architecture can therefore accommodate very large protein ligands (e.g. flagellin, inner-rod proteins, ribonuclease[153]).

Finally, the LRR has been used to benefit the host in multiple instances of independent evolution for its ability to rapidly generate distinct specificities. The solvent-exposed concave surface provides a large area for protein-interaction whose affinity can be changed by amino acid substitutions[153]. The most stunning illustration of the versatility of this structural motif is that LRR diversification is the cornerstone of adaptive immunity in agnathes[64]. In addition, bio-informatic approaches have estimated the number of R-genes could account for 1% of the Arabidopsis genome[154]. Therefore, it is of general interest to many different scientific disciplines to understand the mechanisms by which specific host NB-LRR proteins distinguish between pathogens.

Chapter Five:
A paradigm of inflammasome activation

*Portions of this chapter are scheduled for publication in the journal BioEssays (2012), by the co-authors Eric M. Kofoed and Russell E. Vance, and permission to reproduce this material has been granted.

5.1. Review of findings

The NAIP gene array evolved to recognize functionally constrained molecules of pathogenic bacteria, and initiate innate immune responses that protect the host from potentially harmful infection. Expansion of the *Naip* gene array, and the proportion of non-synonymous substitutions among paralogues suggest that the locus is under selective pressure exerted by host-pathogen interactions. NAIPs have a critical role in cytoplasmic immunosurveillance because they directly bind bacterial proteins, assemble an inflammasome, and initiate immune responses beneficial to the host. We have found that NAIP5 and NAIP6 specifically detect flagellin, whereas NAIP2 specifically detects the inner-rod proteins from pathogen-associated type III secretion systems. NAIP1 is able to respond to a subset of inner rod proteins, indicating that at least partial functional redundancy exists between NAIP1 and NAIP2.

We have demonstrated that the bacterial ligands flagellin and PrgJ are in physical association with NAIPs and NLRC4 in a hetero-oligomeric inflammasome complex. The detection of flagellin by NAIP5 and NAIP6, and inner rod proteins by NAIP1 and NAIP2 have broad implications for host defense precisely because they are capable of 'sensing' infection from an expanding variety of pathogenic bacteria (e.g. *S.typhimurium*, *L.pneumophila, B.pseudomallei, Y.pestis, V.cholera, P.aeruginosa* etcetera). Thus, the NAIPs form an innate immune barrier against many pathogenic, or potentially pathogenic bacteria.

Finally, we have mapped the ligand-specificity determining region of the NAIPs to a small region of the LRR domain (amino acids 822-1078^{NAIP2}). This region is enriched for polymorphisms between NAIP5 and NAIP2, which strongly suggest that direct binding between the ligand and LRRs occurs at this location. We are currently testing whether these minimal LRRs are sufficient for ligand binding. Ultimately, we hope these studies will contribute to the rational design of soluble LRRs amenable to crystallographic analysis in association with cognate bacterial ligands.

5.2. A mechanism of inflammasome assembly

In other NLRs, the LRR domain is believed to function as an autoinhibitory domain. Autoinhibition is believed to be relieved upon recognition of a ligand by the LRR. The structural basis for autoinhibition by LRRs is poorly understood, but is believed to be due to the LRRs 'folding back' and occluding function of the NBD. In the NOD1 and NOD2 proteins, deletion of the C-terminal LRRs results in a constitutively active protein[137]. Fitting with this general mechanistic paradigm, truncated NAIP5ΔLRR and NLRC4ΔLRR proteins were also found to be constitutively active[101]. Importantly, NAIP5ΔLRR induced oligomerization of full-length NLRC4, and required NRLC4 in order to activate downstream Caspase-1. On the other hand, the

truncated NLRC4ΔLRR protein could oligomerize and activate Caspase-1 in the absence of NAIP5[101]. These data strongly suggest that NAIP5 functions upstream of NLRC4. The model that emerges is that NAIP5 is maintained by its LRRs in an inactive conformation, and that ligand interaction with the LRRs results in release of this autoinhibition, induction of NLRC4 oligomerization, and recruitment and activation of Caspase-1 (Figs. 2.6 & 5.1). It must be emphasized, however, that key aspects of this highly speculative model, including the autoinhibitory and ligand-binding functions of the LRRs, have yet to be established biochemically.
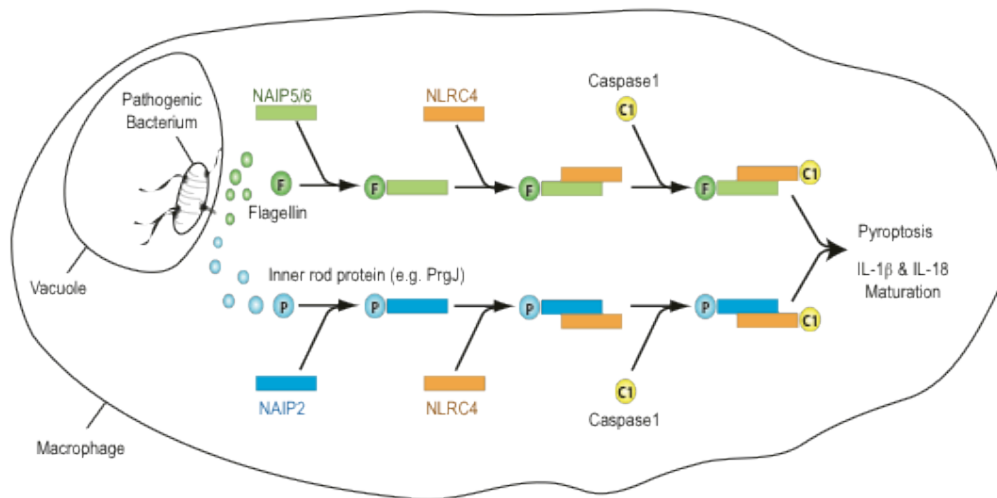
Figure 5.1. Model of NAIP inflammasome activation by specific bacterial ligands. Phagocytosed pathogenic bacteria (inadvertently) secrete proteins into the host cell cytoplasm where NAIP proteins respond by activation of the inflammasome. NAIP5 and NAIP6 specifically detect bacterial flagellin in the host cytoplasm, whereas NAIP2 mediates the detection of T3SS inner rod proteins (e.g. PrgJ from *Salmonella*). NLRC4 is recruited into the complex and serves as an adaptor for Caspase-1, that when activated leads to pyroptosis and cytokine maturation. Many aspects of this model have yet to be confirmed experimentally. For example, although NAIP proteins confer specificity for distinct bacterial ligands, NLRC4 may also contribute to binding. The stoichiometry and organization of NAIP/NLRC4 complexes also remain unknown. It also remains possible that additional host proteins are required.

5.3. Convergent evolution of the NB-LRR module across kingdoms of life

LRR domains are a structural motif found in all forms of life from viruses to mammals[8], are functionally versatile, and conformationally flexible. The convergent evolution of plant NBS-LRR and animal NBD-LRR containing proteins for host defense is remarkable. Plants and animals have converged on a solution to the problems pathogens pose by using common domain architecture for host defense: AAA+ ATPase oligomerization-domain connected to a LRR ligand-binding domain. This module organization combines the energetic favorability of oligomerization with an evolutionarily malleable LRR domain.

The LRR domain conforms to alternating β-strand/α-helix repeats that form a solenoid. LRR domains have been shown to bind directly to nucleic acid, glycoprotein, protein, and lipid chemical groups. This broad range of ligand binding capability is unusual for a single molecular fold. In contrast to the spatial constraints imposed by globular domains and the small surface area of ligand binding 'pockets', the LRR solenoid backbone forms a large surface-exposed ligand interaction scaffold that can accommodate large protein ligands.

The existence of NBD-LRR containing proteins in both plants and animals suggest that this combination of motifs has important mechanistic benefits under selection for host defense against rapidly evolving pathogens. In agnathes, the variable lymphocyte receptors (VLRs) are tethered to the cell surface by a GPI anchor, and this can be cleaved to generate soluble LRRs that can neutralize pathogens similar to the role of neutralizing antibodies in mammals. However, in plants and jawed vertebrates the genuine ability to create a diversity of ligand-binding specificity encoded in LRRs was combined with modules that confer oligomerization and downstream signaling capabilities (e.g. NBD, CARD, BIR & TIR domains). In plants and animals, the connection between pathogen recognition via LRRs and the NBD permits energetically favorable and rapid oligomeric assembly. We suggest that there is an evolutionary benefit to having a rapid and deliberate cell suicide program poised to respond to infection. The NBD-LRR combination yields proteins that display cooperative kinetics of oligomerization with a highly adaptable ligand binding surface. Therefore the response is highly specific, exhibits rapid kinetics, and has irreversible results (i.e. pyroptosis).

Finally, the fundamental mechanistic principles of gene cluster evolution holds great intrigue across scientific disciplines, yet remains poorly characterized. Both plant and animal NB-LRR gene arrays appear to be under strong selective pressure by pathogens. In plants, the abundance of non-synonymous substitutions in select regions of LRR domains indicates that diversifying selection is operating and maintaining haplotypic diversity[100,140]. In fact, differences in the susceptibility of mouse strains to *L.pneumophila* led to the discovery of the *Naip* gene array[18], whereas differences in susceptibility to *B.anthracis* led to the discovery of distinct NLR gene array (i.e. *Nalp1*)[155]. These observations suggest that pathogens drive the expansion and diversification of genes that are important in host defense across the kingdoms of life. However, many questions remain: how do the R-gene or NLR gene clusters arise? How is haplotypic diversity of gene arrays generated and maintained? Several theoretical mechanisms proposed rely on an ancestral gene-duplication event followed by unequal crossing over, meiotic instability, or gene conversion events[100,156,157]. The evolutionary pressure to maintain diversity at these loci must be immense considering the danger of generating deleterious alleles that may be constitutively active (e.g. *Nlrc4*ΔLRR) or hyper-sensitive (i.e. the 'hair-trigger'[156]).

5.4. What might NAIPs teach us about other NLR family members?

The human and mouse genomes each contain over 20 members of the NLR superfamily. The functions of most of these NLRs remain unknown and the mechanisms of their activation are even more mysterious. In the mouse, one relatively well-characterized NLR is NLRP3, which can be activated by structurally distinct stimuli including pore forming toxins such as nigericin, as well as crystalline material, such as

uric acid, alum, and silica[158]. Alum has recently been shown to alter lipid sorting and signal transduction cascades crucial for adjuvanticity, but these early events are upstream of NLRP3[67]. Since it is unlikely that NLRP3 directly binds this great diversity of ligands, it is widely assumed that NLRP3 stimuli induce some common downstream effects that result in indirect NLRP3 activation. Thus, the simple receptor-ligand model that governs NAIP/NLRC4 activation does not seem to apply to NLRP3. However, this remains an open question. Although several distinct mechanisms for NLRP3 activation have been proposed, no consensus has yet been reached in the field. One speculative possibility raised by work on the NAIP/NLRC4 inflammasome is that NLRP3, like NLRC4, may act downstream of other NLRs (or non-NLR proteins) that serve as direct upstream sensors. The concept of upstream 'helper' NLRs has also been recently proposed to apply to activation of certain plant NLR proteins[159,160].

A different NLR, NLRP1B, has been shown to be specifically required for Caspase-1 activation in response to anthrax lethal toxin[155]. Interestingly, lethal toxin is a protease and, moreover, mutation of the protease active site indicates that the protease activity of lethal toxin is required for activation of NLRP1B/Caspase-1[161,162]. Thus, again, a simple receptor-ligand model suggested by work on the NAIP/NLRC4 inflammasome also does not seem to apply to the case of mouse NLRP1B. Instead it seems likely that lethal toxin cleaves target proteins in the cell that indirectly results in NLRP1B activation. The human NLRP1 inflammasome has been reconstituted *in vitro* using purified NLRP1, and was shown to be activated by a fragment of bacterial peptidoglycan called muramyl dipeptide (MDP)[104]. This would seem to support a simple receptor-ligand model for human NLRP1 activation, but MDP has not been shown to directly bind NLRP1, and moreover, there is not strong genetic evidence supporting a role for NRLP1 in MDP detection. In fact, years of work studying NLRs in plants strongly suggest that NLRs can be regulated by surprisingly diverse mechanisms[114,160]. As evidence for such diversity in mammalian NLRs, a recent and surprising report found that the human NLRP1 protein contains an unusual ZU5-UPA domain that mediates autoproteolytic processing of NLRP1[163]. The function of such autoprocessing remains to be determined, but illustrates the broader notion that the mechanisms of NLR regulation should not be assumed to be highly conserved among family members. Instead, the continued development of rigorous biochemical systems to dissect activation of diverse NLR family members will be important.

5.5. Conclusion

In the ten years since inflammasomes were first described[102], the progress in our understanding has been remarkable. Nevertheless, much remains to be learned about the mechanisms of inflammasome activation. The NAIP-NLRC4 inflammasomes appear to function as receptors for specific bacterial molecules, but it remains unclear how NAIPs recognize and distinguish ligands, and what role NLRC4 plays in ligand recognition, if any. The role of the BIRs, a domain not found in other NLRs except for NAIPs, also remain to be deciphered.

The direct recognition of conserved bacterial ligands by NAIPs is reminiscent of PAMP recognition mediated by TLRs. Unlike TLRs, however, the activation of the NAIP-NLRC4 inflammasome is limited by its cytosolic localization. Indeed, NAIPs are only activated when specific bacterial ligands reach the host cell cytosol, and generally speaking, bacterial proteins do not reach the host cell cytosol unless the bacteria themselves enter the cytosol or the proteins are translocated by a bacterial secretion

system (though an interesting exception to this principle was recently described)[164]. Indeed, the bacterial pathogens that are known to robustly activate NAIP/NLRC4 all encode either virulence-associated type III or type IV secretion systems, and the function of these systems is generally required for bacterial activation of NAIP/NLRC4. Thus, the NAIP/NLRC4 system provides the host with the ability to respond specifically to the presence of bacterial pathogens, while ignoring harmless or commensal bacteria that lack such secretion systems. Fitting with its specialized role in recognizing pathogenic microbes, the NAIP/NLRC4 inflammasome induces unique responses (e.g., Caspase-1 activation, IL-1β and IL-18 processing, and pyroptosis) not induced by TLRs. These signaling outputs are highly inflammatory and pose a risk of self-damage to hosts. Thus, it may make sense that they are limited to situations in which hosts have encountered a *bona fide* pathogenic threat.

A question that is often raised is why bacterial pathogens cannot evade recognition by NAIP/NLRC4? The answer is that many bacterial pathogens likely *do* evade NAIP/NLRC4 recognition. *Salmonella* and *Listeria*, for example, tightly regulate flagellin expression and/or secretion, and tend not to express flagellin in hosts. In fact, enforced expression of flagellin by *Salmonella* and *Listeria* has a severely negative impact on bacterial virulence[81,165], and this negative impact can be ameliorated by elimination of host NLRC4. Thus, the primary function of the NAIP/NLRC4 inflammasome may not be to recognize highly host-adapted pathogens, but may rather be important to form an innate immune barrier against less well-adapted but still potentially virulent pathogens, such as *Legionella pneumophila*. An area to be considered in future work is whether the NAIP/NLRC4 inflammasomes might in fact play an important role in recognition of the host microbiota.

Materials and Methods

Cell Culture. HEK293T cells were grown in complete media (DMEM, 10%FBS, 100Units/ml Penicillin, 100μg/ml Streptomycin, 2mM L-glutamine). Bone marrow macrophages were differentiated from bone marrow-derived precursor cells using macrophage colony stimulating factor as described previously[26].

Transient Transfections. HEK293T cells were plated at 8×10^5 cells/well in 6-well plates, transfected the next day using Lipofectamine2000, and collected for flow cytometric analysis 48 hours later.

Measurement of Cell Death. Cell death of HEK293T cells was measured by flow cytometry measuring GFP and 7AAD fluorescence. Cells were stained with 7AAD (BD Pharmingen) according to the manufacturer's instructions. Death of immortalized macrophages in response to *L. pneumophila* and LFn-FlaA was measured by following release of the intracellular enzyme lactate dehydrogenase (LDH) as described previously[26]. Infection of immortalized macrophages with *L. pneumophila* was performed using a multiplicity of infection (MOI) of 2, and cell death was measured four hours later as described previously[26]. LFn-FlaA is a recombinant 6XHis-tagged fusion protein encoding the first (non-enyzmatically active) 263 amino acids of lethal factor from *Bacillus anthracis* fused to full length flagellin (FlaA) from *L. pneumophila* (J. von Moltke and R. Vance, unpublished). LFn-FlaA and *B. anthracis* protective antigen (PA) were purified from *E. coli* as described previously[166].   Endotoxin was removed from these proteins using Detoxi-Gel (Pierce, Rockford, IL) according to the manufacturer's protocol.

Western blotting and Native PAGE. Blue native gel electrophoresis was performed using the Bis-Tris NativePAGE system by Invitrogen according to the manufacturer's instructions. Briefly, HEK293T cells were transfected for 48 hours, washed twice with cold PBS, trypsinized for 3 minutes at 37°C, resuspended in complete media, and pelleted by centrifugation at 400x*g* for 5 minutes at 4°C. Cell pellets were washed twice with cold PBS, followed by resuspension in 1% digitonin native lysis buffer (50mM BisTris, 50mM NaCl, 10%w/v glycerol, 0.001% Ponceau S, 1% digitonin, 2mM Na₃VO₄, 1mM PMSF, 25mM NaF, 1x Roche protease inhibitor cocktail (no EDTA), pH 7.2). Cell lysates were triturated and incubated on ice 30 min, and then insoluble cell debris were pelleted by centrifugation at 400x*g* for 30min at 4°C. Lysates were quantified for total protein using the BCA protein assay (Pierce), equalized for total protein, and separated by NativePAGE using the Novex BisTris gel system according to manufacturer's instructions (Invitrogen). Native gels were soaked in 10% SDS for 5 minutes prior to transfer to PVDF membrane (Millipore) and conventional western blotting. Antibodies used were anti-mIL-1β (R&D systems), anti-CASPASE1 p10 (M20) (Santa Cruz), anti-NLRC4 (gift of S. Mariathasan and V. Dixit, Genentech), anti-NAIP5(961-978)[24], anti-c-myc (9E10)(Clontech), anti-β-actin (C4)(Santa Cruz), anti-GFP (JL-8) (Clontech), anti-NAIP2(33-46)[24], anti-rabbit IgG-HRP and anti-mouse IgG-HRP (GE Healthcare), anti-goat IgG-HRP (Santa Cruz). In some cases, native gels were subject to a second dimension of electrophoresis using SDS-PAGE. A 5.7cm slice (lane) of natively

resolved gel was placed in a dish containing 1x Laemmli sample buffer (50mM Tris-Cl (pH 6.8), 100mM DTT, 2% (w/v) SDS, 0.1% bromophenol blue, 10% (v/v) glycerol) for 10 min, microwaved on high for 20 seconds, and rocked for another 5 min prior to loading slice into the well of a precast 2D 4-12% SDS-PAGE gel (Invitrogen). For immunoprecipitations, digitonin cell extracts (100μg total protein) were pre-cleared with 25μl of washed Protein-G-sepharose (GE healthcare), and then cleared extracts were incubated with 1μg primary antibody (or isotype controls) overnight at 4°C, and complexes were captured the following day with 25μl of washed Protein-G-sepharose. Bound proteins were eluted by resuspension in Laemmli sample buffer, boiled for 5 minutes and separated by SDS-PAGE.

Expression Constructs. NAIP5 wild-type and mutant constructs were cloned into the MSCV2.2 retroviral vector containing an IRES-GFP downstream of the multiple cloning site. Expression is driven from the viral LTR and is considerably lower than that driven by the CMV promoter (data not shown). In general, wild-type and mutant ORFs were amplified between flanking BamHI and NotI sites, and a Kozak sequence (GCCACC) was engineered to proceed the start codon. The BamHI/NotI digested PCR product was cloned into complementary BglII/NotI digested MSCV2.2 vector. Wild-type NAIP5 (1-1402) and NAIP5ΔPloop(Δ464-487) were amplified from pcDNA3 using forward (AAAAGGATCCGCCACCATGGCTGAGCATGGGGAGTCCTCCG), and reverse (AAAAGCGGCCGCTTACTCCAGGATAACAGGAGAGAATGGGAC) primers. NAIP5Δ347(347-1402) was generated using the forward primer (AAAAG-GATCCGCCACCATGACCTTGAAGTCCTCTGCAGAAG) in combination with wild-type NAIP5 reverse primer. NAIP5ΔLRR(1-1039) was PCR cloned into BglII/PmeI sites of MSCV2.2 vector using the wild-type NAIP5 forward primer (see above), and reverse primer (GTTTAAAC<u>TCA</u>GCCACTGCTGTTGAATAAACG). Wild-type NLRC4(1-1024) was PCR amplified from pCDNA3 using forward (AAAAGGATCCGCCAC-CATGAACTTTATAAGGAACAACAGACG) and reverse (TTTTGCGGCCGCTTAAGCAGTCACTAGTTTAAAGGTGCC) primers, and ligated into BglII/NotI sites of MSCV2.2-IRES-GFP. NLRC4(K175R) was generated by site directed mutagenesis using the QuickChange protocol (Stratagene) using forward (GAGTCTGGCAAAGGGCGATCGACCCTGCTGCAG), and reverse (CTGCAG-CAGGGTCGATCGCCCTTTGCCAGACTC) primers. NLRC4ΔCARD(89-1024) was PCR amplified from pCDNA3 using forward primer (AAAAGGATCCGCCAC-CATGTCTTATCAGGTCACAGAAGAAGACC) paired with the wild-type reverse primer. NLRC4ΔLRR (1-656) was PCR amplified from pCDNA3 using the reverse primer (TTTTGCGGCCGCTTACTTCCAGTTGAAGAACAAAGACACAGC) in combination with the wild-type forward primer. Wild-type NAIP2 was PCR cloned from pBluescript(SK-) into the BglII/NotI sites of MSCV2.2-IRES-GFP using forward (AAAAGGATCCGCCACCATGGCAGCCCAGGGAGAAGCCGTTGAGG), and reverse (TTTTGCGGCCGCTCACTTCTGAATGACAGGAGAGAATGGCACTACCC) primers. Wild-type NAIP6 was PCR cloned from pBluescript(SK-) into the BglII/NotI sites of Mscv2.2-IRES-GFP using the wild-type NAIP5 forward primer (there is 100% conservation of the first 26 nucleotides among NAIP1, NAIP5, and NAIP6), and the reverse primer (TTTTGCGGCCGCTTACTCCAGGACAACAGGAGAGAACGGGAC). Wild-type NAIP1 was PCR cloned from pBluescript(SK-) into the BglII/NotI sites of

MSCV2.2-IRES-GFP using the wild-type NAIP5 forward primer, and the wild-type NAIP6 reverse primer (NAIP1 and NAIP6 share identical C-terminal nucleotides). MSCV2.2-CASPASE1, MSCV2.2-CASPASE1(C284A), and pSPORT-CASPASE11 were described previously[109].

We modified the MSCV2.2 vector to contain a 6x-Myc-tag in the MCS to facilitate generation of NH_2-terminal fusion proteins (6x-Myc-MSCV2.2-IRES-GFP). The 6x-Myc-tag from pCS-6x-Myc-SEC24 was PCR amplified using forward (AAAAAGATCTATCGATTTAAAGCTATG) and reverse (TTTTGCGGCCGCTGG-CCGGCCTGAATTCA) primers for insertion into the BglII site of MSCV2.2. 6x-Myc-FlaA was generated by PCR amplifying full-length FlaA (*L. pneumophila*) from MSCV2.2-FlaA [26] using forward (AAAAGCGGCCGCAG-CTCAAGTAATCAACACTAATGTGGC) and reverse (TTTTGTCGACTATC-GACCTAACAAAGATAATAC) primers and inserted into NotI/SalI sites. 6x-MYC-PrgJ was generated by amplifying PrgJ from MSCV2.2-PrgJ using forward (AAAAGCGGCCGCATCGATTGCAACTATTGTCCC) and reverse (TTTTGTCGACTCATGAGCGTAATAGCGTTTC) primers and insertion into NotI/SalI sites. All constructs were fully sequenced to confirm their identity. Sequencing primers used for NAIP5: MSCV2.2-F: (AAGCCCTTTGTACACCCTAAGCC), MSCV2.2-R: (CCTCACATTGCCAAAAGAC), NAIP5seq#1: (CAGCAAAAGCACTGAACGCC), NAIP5seq#2: (ATGAACAAATCCCTCGTAGC), NAIP5seq#3: (TCACTCCTAC-CCAAGTCCAC), NAIP5seq#4: (CTCAGACACA-CTTCACTAATGC), NAIP5seq#5: (TCCCTTAGTTCCATCACACC), NAIP5seq#6: (GACCCCTCTCTTTGTAGCAG), NAIP5seq#7: (GAGTTTCTTGCTGCCGTGAG), NAIP5seq#8: (TTAGAGGGTTGT-GGCTGGTGTC), NAIP5seq#9: (CTTCACAGAGTATTGAGTTCCG), NAIP5seq#10: (TTGAGTTTTCTGGACGATGC), NAIP5seq#11: (GGACAACTTGCCAAACCTAC). Sequencing primers for NAIP2: NAIP2seqF1: (TGGTGATGAGAAAGAGTCAC), NAIP2seqF2: (CTTCACAGAGTATTGAGTTCCG), NAIP2seqR1: (AGCAAATGG-TCAGTGCCGAG), NAIP2seqR2: (ACATACTGCTGCCACGAAG), NAIP2seqR3: (AATCCAGTGTTCTCCCTCG), MSCV2.2 -F and –R primers (see above). Sequencing primers for NAIP6: NAIP6seqF1: (CAGAAAGCCTGTTACTGTTGAG), NAIP6seqR1: (GATGGAACTAAG-GGAGAGGTAG), NAIP6seqR2: (TCTTGGTCTTCCTG-CCTATC), MSCV2.2-F, MSCV2.2-R, NAIP5seq#7 (see above).

Generation of chimeric NAIPs by SOE-ing PCR. We aligned primary nucleotide sequences of NAIP2, NAIP5, and NAIP6 to chose breakpoints based on pairwise sequence identity. We chose ten breakpoints along the length of the Naip cDNAs that would preserve modular domain boundaries in the case of BIRs and NBD, or divided the domain as evenly as possible. The SOE primers used were (using NAIP2 positioning):

PRIMER	nt position	5' to 3' sequence
N2BIR1:F	388-425	AAAGATGTTGGTAACATTGGCAAGTATGACATCCGGGT
N2BIR1:R		ACCCGGATGTCATACTTGCCAATGTTACCAACATCTTT
N2BIR2:F	769-819	CCTGGGTCAGAAGAGAATTACCTATGGTATCAGCTTACTGC
N2BIR2:R		GCAGTAAGCTGATACCATAGGTAATTCTCTTCTGACCCAGG
N2BIR3:F	1131-1163	TGATGCAGCAGCAGTTCATTCTACAGTGGTGGA
N2BIR3:R		TCCACCACTGTAGAATGAACTGCTGCTGCATCA

PRIMER	nt position	5' to 3' sequence
N2LinkerF	1322-1352	TTCAAGAGGCCAGGAGTCTGAGTGAGCAGCT
N2LinkerR		AGCTGCTCACTCAGACTCCTGGCCTCTTGAA
N2NBD1:F	1931-1956	AAGAGTTTCCCTTCTATAATACTGTCT
N2NBD1:R		AGACAGTATTATAGAAGGGAAACTCTT
N2NBD2:F	2264-2292	ACCTGGCAGAGGCAGGAGTTGATGAAGATG
N2NBD2:R		CATCTTCATCAACTCCTGCCTCTGCCAGGT
N2NBD3:F	2608-2637	GAAAATGAGGATTACATGAAGCTCCATCCA
N2NBD3:R		TGGATGGAGCTTCATGTAATCCTCATTTTC
N2NBD4:F	2747-2767	ATGAAAGCAACACAGTTGCTG
N2NBD4:R		CAGCAACTGTGTTGCTTTCAT

Generation of NAIP2 shRNA constructs. We used the lentiviral pLKO.1-TRC cloning vector (Addgene) to generate three vectors expressing three separate short hairpin RNAs targeting NAIP2 (NAIP2 shRNA#1-#3). NAIP2 shRNA#1 oligos: (CCGGGCCATTGCCTTTCAACCTATACTCGAGTATAGGTTGAAAGGCAATGGC TTTTTG) and (AATTCAAAAAGCCATTGCCTTTCAACCTATACTCGAGTATA-GGTTGAAAGGCAATGGC). NAIP2 shRNA#2 oligos (CCGGCCATCCAGAAA-CCTTGTTGTTCTCGAGAACAACAAGGTTTCTGGATGGTTTTTG) and (AATTCAAAAACCATCCAGAAACCTTGTTGTTCTCGAGAACAACAAGGTTTCT GGATGG). NAIP2 shRNA#3 oligos: (CCGGCTTTCAGTCTTGAAGAGACAACTC-GAGTTGTCTCTTCAAGACTGAAAGTTTTTG) and (AATTCAAAAACTTTC-AGTCTTGAAGAGACAACTCGAGTTGTCTCTTCAAGACTGAAAG). We included pLKO.1-TRC control vector (Addgene ID#10879) or pLKO.1 scramble (Addgene ID#1864) as negative controls.

Knockdown of NAIP2 in primary macrophages. Bone marrow was collected from C57BL/6 mice on day 0, and plated into one 15cm plate in macrophage differentiation media (see above). Lentivirus encoding NAIP2 shRNAs were generated according to the Addgene protocol. On day 2, primary bone marrow macrophages were collected, red blood cells were lysed, and cells were plated at 1×10^6 cells/well in 6-well plates. Macrophages were spinfected with lentiviral particles at 32°C, 90 min., 1258xg, and placed in a 32C incubator for 48hours. On day 4, cells were collected in cold PBS, replated on 10cm plates containing fresh media containing puromycin [5μg/ml] for selection, and placed at 37°C. Puromycin containing media was replaced on day 6. On day 8, macrophages were collected, counted, and seeded at 1×10^5 cells/well in a 96-well plate, and infected with *Listeria* the following day. In some experiments, NAIP2 was knocked down in v-myc/v-raf immortalized bone marrow-derived macrophages that were previously generated by use of the J2 virus.

***Listeria monocytogenes* infection.** We generated flagellin-deficient strains of *Listeria monocytogenes* 10403S that express the secreted fusion protein ActAN100-PrgJ (L.m.ΔFlaA-ActAN100-PrgJ), under the control of an IPTG-inducible promoter, as

previously described[52]. Macrophages were spinfected in antibiotic free media at 400x*g* for 10 minutes at an MOI =5, with or without IPTG [1mM], and placed at 37°C for 30min. The media was then replaced with complete media containing gentamycin [10μg/ml] and IPTG, and supernatants were assayed for LDH assay 5.5 hours later.

***Salmonella enterica* serovar Typhimurium infection.** *Salmonella* strain LT2 or isogenic flagellin mutant (FliC/FljB[-]) was grown in 10mL LB standing cultures at 37°C overnight. The next morning, the cultures were diluted 1:100 in LB and grown for 4h (standing culture, 37°C). Bacteria were added to macrophages at an MOI of 10-30, followed by centrifugation at 400x*g* for 10 minutes. Gentamycin (25μg/ml) was added after 1h to kill remaining extracellular bacteria. Caspase-1 processing or LDH release was monitored as previously described[26]

References

1. **Janeway CA, Jr.** 1989. Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harb Symp Quant Biol* **54 Pt 1**:1-13.
2. **Jin MS, Kim SE, Heo JY, Lee ME, Kim HM and others.** 2007. Crystal structure of the TLR1-TLR2 heterodimer induced by binding of a tri-acylated lipopeptide. *Cell* **130**:1071-82.
3. **Kang JY, Nan X, Jin MS, Youn SJ, Ryu YH and others.** 2009. Recognition of lipopeptide patterns by Toll-like receptor 2-Toll-like receptor 6 heterodimer. *Immunity* **31**:873-84.
4. **Kim HM, Park BS, Kim JI, Kim SE, Lee J and others.** 2007. Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell* **130**:906-17.
5. **Liu L, Botos I, Wang Y, Leonard JN, Shiloach J and others.** 2008. Structural basis of toll-like receptor 3 signaling with double-stranded RNA. *Science* **320**:379-81.
6. **Park BS, Song DH, Kim HM, Choi BS, Lee H and others.** 2009. The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. *Nature* **458**:1191-5.
7. **Yoon SI, Kurnasov O, Natarajan V, Hong M, Gudkov AV and others.** 2012. Structural basis of TLR5-flagellin recognition and signaling. *Science* **335**:859-64.
8. **Bella J, Hindle KL, McEwan PA, Lovell SC.** 2008. The leucine-rich repeat structure. *Cell Mol Life Sci* **65**:2307-33.
9. **Takeuchi O, Akira S.** 2010. Pattern recognition receptors and inflammation. *Cell* **140**:805-20.
10. **Hibino T, Loza-Coll M, Messier C, Majeske AJ, Cohen AH and others.** 2006. The immune gene repertoire encoded in the purple sea urchin genome. *Dev Biol* **300**:349-65.
11. **Whitham S, Dinesh-Kumar SP, Choi D, Hehl R, Corr C and others.** 1994. The product of the tobacco mosaic virus resistance gene N: similarity to toll and the interleukin-1 receptor. *Cell* **78**:1101-15.
12. **Staskawicz BJ, Ausubel FM, Baker BJ, Ellis JG, Jones JD.** 1995. Molecular genetics of plant disease resistance. *Science* **268**:661-7.
13. **Stout JE, Yu VL.** 1997. Legionellosis. *N Engl J Med* **337**:682-7.
14. **Fields BS, Benson RF, Besser RE.** 2002. Legionella and Legionnaires' disease: 25 years of investigation. *Clin Microbiol Rev* **15**:506-26.
15. **Isberg RR, O'Connor TJ, Heidtman M.** 2009. The Legionella pneumophila replication vacuole: making a cosy niche inside host cells. *Nat Rev Microbiol* **7**:13-24.
16. **Yamamoto Y, Klein TW, Newton CA, Widen R, Friedman H.** 1988. Growth of Legionella pneumophila in thioglycolate-elicited peritoneal macrophages from A/J mice. *Infect Immun* **56**:370-5.
17. **Beckers MC, Yoshida S, Morgan K, Skamene E, Gros P.** 1995. Natural resistance to infection with Legionella pneumophila: chromosomal localization of the Lgn1 susceptibility gene. *Mamm Genome* **6**:540-5.

18.   **Dietrich WF, Damron DM, Isberg RR, Lander ES, Swanson MS.** 1995. Lgn1, a gene that determines susceptibility to Legionella pneumophila, maps to mouse chromosome 13. *Genomics* **26**:443-50.

19.   **Scharf JM, Damron D, Frisella A, Bruno S, Beggs AH and others.** 1996. The mouse region syntenic for human spinal muscular atrophy lies within the Lgn1 critical interval and contains multiple copies of Naip exon 5. *Genomics* **38**:405-17.

20.   **Ting JP, Lovering RC, Alnemri ES, Bertin J, Boss JM and others.** 2008. The NLR gene family: a standard nomenclature. *Immunity* **28**:285-7.

21.   **Diez E, Beckers MC, Ernst E, DiDonato CJ, Simard LR and others.** 1997. Genetic and physical mapping of the mouse host resistance locus Lgn1. *Mamm Genome* **8**:682-5.

22.   **Diez E, Lee SH, Gauthier S, Yaraghi Z, Tremblay M and others.** 2003. Birc1e is the gene within the Lgn1 locus associated with resistance to Legionella pneumophila. *Nat Genet* **33**:55-60.

23.   **Huang S, Scharf JM, Growney JD, Endrizzi MG, Dietrich WF.** 1999. The mouse Naip gene cluster on Chromosome 13 encodes several distinct functional transcripts. *Mamm Genome* **10**:1032-5.

24.   **Wright EK, Goodart SA, Growney JD, Hadinoto V, Endrizzi MG and others.** 2003. Naip5 affects host susceptibility to the intracellular pathogen Legionella pneumophila. *Curr Biol* **13**:27-36.

25.   **Diez E, Yaraghi Z, MacKenzie A, Gros P.** 2000. The neuronal apoptosis inhibitory protein (Naip) is expressed in macrophages and is modulated after phagocytosis and during intracellular infection with Legionella pneumophila. *J Immunol* **164**:1470-7.

26.   **Lightfield KL, Persson J, Brubaker SW, Witte CE, von Moltke J and others.** 2008. Critical function for Naip5 in inflammasome activation by a conserved carboxy-terminal domain of flagellin. *Nat Immunol* **9**:1171-8.

27.   **Watarai M, Derre I, Kirby J, Growney JD, Dietrich WF and others.** 2001. Legionella pneumophila is internalized by a macropinocytotic uptake pathway controlled by the Dot/Icm system and the mouse Lgn1 locus. *J Exp Med* **194**:1081-96.

28.   **Derre I, Isberg RR.** 2004. Macrophages from mice with the restrictive Lgn1 allele exhibit multifactorial resistance to Legionella pneumophila. *Infect Immun* **72**:6221-9.

29.   **Zamboni DS, Kobayashi KS, Kohlsdorf T, Ogura Y, Long EM and others.** 2006. The Birc1e cytosolic pattern-recognition receptor contributes to the detection and control of Legionella pneumophila infection. *Nat Immunol* **7**:318-25.

30.   **Fink SL, Cookson BT.** 2005. Apoptosis, pyroptosis, and necrosis: mechanistic description of dead and dying eukaryotic cells. *Infect Immun* **73**:1907-16.

31.   **Bergsbaken T, Fink SL, Cookson BT.** 2009. Pyroptosis: host cell death and inflammation. *Nat Rev Microbiol* **7**:99-109.

32.   **Amer A, Franchi L, Kanneganti TD, Body-Malapel M, Ozoren N and others.** 2006. Regulation of Legionella phagosome maturation and infection through flagellin and host IPAF. *J Biol Chem.*

33. **Amer AO, Swanson MS.** 2005. Autophagy is an immediate macrophage response to Legionella pneumophila. *Cell Microbiol* **7**:765-78.

34. **Fortier A, de Chastellier C, Balor S, Gros P.** 2006. Birc1e/Naip5 rapidly antagonizes modulation of phagosome maturation by Legionella pneumophila. *Cell Microbiol*.

35. **Chamaillard M, Hashimoto M, Horie Y, Masumoto J, Qiu S and others.** 2003. An essential role for NOD1 in host recognition of bacterial peptidoglycan containing diaminopimelic acid. *Nat Immunol* **4**:702-7.

36. **Girardin SE, Boneca IG, Carneiro LA, Antignac A, Jehanno M and others.** 2003. Nod1 detects a unique muropeptide from gram-negative bacterial peptidoglycan. *Science* **300**:1584-7.

37. **Girardin SE, Boneca IG, Viala J, Chamaillard M, Labigne A and others.** 2003. Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *J Biol Chem* **278**:8869-72.

38. **Inohara N, Ogura Y, Fontalba A, Gutierrez O, Pons F and others.** 2003. Host recognition of bacterial muramyl dipeptide mediated through NOD2. Implications for Crohn's disease. *J Biol Chem* **278**:5509-12.

39. **Molofsky AB, Byrne BG, Whitfield NN, Madigan CA, Fuse ET and others.** 2006. Cytosolic recognition of flagellin by mouse macrophages restricts Legionella pneumophila infection. *J Exp Med* **203**:1093-104.

40. **Ren T, Zamboni DS, Roy CR, Dietrich WF, Vance RE.** 2006. Flagellin-deficient Legionella mutants evade caspase-1- and Naip5-mediated macrophage immunity. *PLoS Pathog* **2**:e18.

41. **Askari N, Correa RG, Zhai D, Reed JC.** 2012. Expression, purification, and characterization of recombinant NOD1 (NLRC1): A NLR family member. *J Biotechnol* **157**:75-81.

42. **Laroui H, Yan Y, Narui Y, Ingersoll SA, Ayyadurai S and others.** 2011. L-Ala-gamma-D-Glu-meso-diaminopimelic acid (DAP) interacts directly with leucine-rich region domain of nucleotide-binding oligomerization domain 1, increasing phosphorylation activity of receptor-interacting serine/threonine-protein kinase 2 and its interaction with nucleotide-binding oligomerization domain 1. *J Biol Chem* **286**:31003-13.

43. **Dueber EC, Schoeffler AJ, Lingel A, Elliott JM, Fedorova AV and others.** 2011. Antagonists induce a conformational change in cIAP1 that promotes autoubiquitination. *Science* **334**:376-80.

44. **Franklin MC, Kadkhodayan S, Ackerly H, Alexandru D, Distefano MD and others.** 2003. Structure and function analysis of peptide antagonists of melanoma inhibitor of apoptosis (ML-IAP). *Biochemistry* **42**:8223-31.

45. **Cossu F, Milani M, Mastrangelo E, Vachette P, Servida F and others.** 2009. Structural basis for bivalent Smac-mimetics recognition in the IAP protein family. *J Mol Biol* **392**:630-44.

46. **Tan X, Calderon-Villalobos LI, Sharon M, Zheng C, Robinson CV and others.** 2007. Mechanism of auxin perception by the TIR1 ubiquitin ligase. *Nature* **446**:640-5.

47. **Miao EA, Alpuche-Aranda CM, Dors M, Clark AE, Bader MW and others.** 2006. Cytoplasmic flagellin activates caspase-1 and secretion of interleukin 1beta via Ipaf. *Nat Immunol* **7**:569-75.

48. **Miao EA, Mao DP, Yudkovsky N, Bonneau R, Lorang CG and others.** 2010. Innate immune detection of the type III secretion apparatus through the NLRC4 inflammasome. *Proc Natl Acad Sci U S A* **107**:3076-80.

49. **Smith KD, Andersen-Nissen E, Hayashi F, Strobe K, Bergman MA and others.** 2003. Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility. *Nat Immunol* **4**:1247-53.

50. **Andersen-Nissen E, Smith KD, Bonneau R, Strong RK, Aderem A.** 2007. A conserved surface on Toll-like receptor 5 recognizes bacterial flagellin. *J Exp Med* **204**:393-403.

51. **Franchi L, Amer A, Body-Malapel M, Kanneganti TD, Ozoren N and others.** 2006. Cytosolic flagellin requires Ipaf for activation of caspase-1 and interleukin 1beta in salmonella-infected macrophages. *Nat Immunol* **7**:576-82.

52. **Lightfield KL, Persson J, Trinidad NJ, Brubaker SW, Kofoed EM and others.** 2011. Differential requirements for NAIP5 in activation of the NLRC4 (IPAF) inflammasome. *Infect Immun*.

53. **Miao EA, Ernst RK, Dors M, Mao DP, Aderem A.** 2008. Pseudomonas aeruginosa activates caspase 1 through Ipaf. *Proc Natl Acad Sci U S A* **105**:2562-7.

54. **Lamkanfi M, Amer A, Kanneganti TD, Munoz-Planillo R, Chen G and others.** 2007. The Nod-Like Receptor Family Member Naip5/Birc1e Restricts Legionella pneumophila Growth Independently of Caspase-1 Activation. *J Immunol* **178**:8022-7.

55. **Siegmund B, Lehr HA, Fantuzzi G, Dinarello CA.** 2001. IL-1 beta -converting enzyme (caspase-1) in intestinal inflammation. *Proc Natl Acad Sci U S A* **98**:13249-54.

56. **Lara-Tejero M, Sutterwala FS, Ogura Y, Grant EP, Bertin J and others.** 2006. Role of the caspase-1 inflammasome in Salmonella typhimurium pathogenesis. *J Exp Med* **203**:1407-12.

57. **Li P, Allen H, Banerjee S, Franklin S, Herzog L and others.** 1995. Mice deficient in IL-1 beta-converting enzyme are defective in production of mature IL-1 beta and resistant to endotoxic shock. *Cell* **80**:401-11.

58. **Raupach B, Peuschel SK, Monack DM, Zychlinsky A.** 2006. Caspase-1-mediated activation of interleukin-1beta (IL-1beta) and IL-18 contributes to innate immune defenses against Salmonella enterica serovar Typhimurium infection. *Infect Immun* **74**:4922-6.

59. **Hugot JP, Chamaillard M, Zouali H, Lesage S, Cezard JP and others.** 2001. Association of NOD2 leucine-rich repeat variants with susceptibility to Crohn's disease. *Nature* **411**:599-603.

60. **Jeru I, Duquesnoy P, Fernandes-Alnemri T, Cochet E, Yu JW and others.** 2008. Mutations in NALP12 cause hereditary periodic fever syndromes. *Proc Natl Acad Sci U S A* **105**:1614-9.

61. **Ogura Y, Bonen DK, Inohara N, Nicolae DL, Chen FF and others.** 2001. A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. *Nature* **411**:603-6.

62. **Tonegawa S.** 1983. Somatic generation of antibody diversity. *Nature* **302**:575-81.

63. **Alder MN, Rogozin IB, Iyer LM, Glazko GV, Cooper MD and others.** 2005. Diversity and function of adaptive immune receptors in a jawless vertebrate. *Science* **310**:1970-3.

64. **Pancer Z, Amemiya CT, Ehrhardt GR, Ceitlin J, Gartland GL and others.** 2004. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature* **430**:174-80.

65. **Wang Y, Ouellette AN, Egan CW, Rathinavelan T, Im W and others.** 2007. Differences in the electrostatic surfaces of the type III secretion needle proteins PrgI, BsaL, and MxiH. *J Mol Biol* **371**:1304-14.

66. **Eisenbarth SC, Colegio OR, O'Connor W, Sutterwala FS, Flavell RA.** 2008. Crucial role for the Nalp3 inflammasome in the immunostimulatory properties of aluminium adjuvants. *Nature*.

67. **Flach TL, Ng G, Hari A, Desrosiers MD, Zhang P and others.** 2011. Alum interaction with dendritic cell membrane lipids is essential for its adjuvanticity. *Nat Med* **17**:479-87.

68. **Ogura Y, Inohara N, Benito A, Chen FF, Yamaoka S and others.** 2001. Nod2, a Nod1/Apaf-1 family member that is restricted to monocytes and activates NF-kappaB. *J Biol Chem* **276**:4812-8.

69. **Fritz JH, Le Bourhis L, Sellge G, Magalhaes JG, Fsihi H and others.** 2007. Nod1-mediated innate immune recognition of peptidoglycan contributes to the onset of adaptive immunity. *Immunity* **26**:445-59.

70. **Kobayashi KS, Chamaillard M, Ogura Y, Henegariu O, Inohara N and others.** 2005. Nod2-dependent regulation of innate and adaptive immunity in the intestinal tract. *Science* **307**:731-4.

71. **Viala J, Chaput C, Boneca IG, Cardona A, Girardin SE and others.** 2004. Nod1 responds to peptidoglycan delivered by the Helicobacter pylori cag pathogenicity island. *Nat Immunol* **5**:1166-74.

72. **Warren SE, Duong H, Mao DP, Armstrong A, Rajan J and others.** 2011. Generation of a Listeria vaccine strain by enhanced caspase-1 activation. *Eur J Immunol* **41**:1934-40.

73. **Lederer E.** 1986. New developments in the field of synthetic muramyl peptides, especially as adjuvants for synthetic vaccines. *Drugs Exp Clin Res* **12**:429-40.

74. **Berrington WR, Iyer R, Wells RD, Smith KD, Skerrett SJ and others.** 2010. NOD1 and NOD2 regulation of pulmonary innate immunity to Legionella pneumophila. *Eur J Immunol* **40**:3519-27.

75. **Hasegawa M, Yamazaki T, Kamada N, Tawaratsumida K, Kim YG and others.** 2011. Nucleotide-binding oligomerization domain 1 mediates recognition of Clostridium difficile and induces neutrophil recruitment and protection against the pathogen. *J Immunol* **186**:4872-80.

76. **Travassos LH, Girardin SE, Philpott DJ, Blanot D, Nahori MA and others.** 2004. Toll-like receptor 2-dependent bacterial sensing does not occur via peptidoglycan recognition. *EMBO Rep* **5**:1000-6.

77.   **Werts C, le Bourhis L, Liu J, Magalhaes JG, Carneiro LA and others.** 2007. Nod1 and Nod2 induce CCL5/RANTES through the NF-kappaB pathway. *Eur J Immunol* **37**:2499-508.
78.   **Travassos LH, Carneiro LA, Girardin SE, Boneca IG, Lemos R and others.** 2005. Nod1 participates in the innate immune response to Pseudomonas aeruginosa. *J Biol Chem* **280**:36714-8.
79.   **Rezaei N, Moazzami K, Aghamohammadi A, Klein C.** 2009. Neutropenia and primary immunodeficiency diseases. *Int Rev Immunol* **28**:335-66.
80.   **Lamkanfi M, Dixit VM.** 2011. Modulation of inflammasome pathways by bacterial and viral pathogens. *J Immunol* **187**:597-602.
81.   **Miao EA, Leaf IA, Treuting PM, Mao DP, Dors M and others.** 2010. Caspase-1-induced pyroptosis is an innate immune effector mechanism against intracellular bacteria. *Nat Immunol* **11**:1136-42.
82.   **Smith GL, Chan YS.** 1991. Two vaccinia virus proteins structurally related to the interleukin-1 receptor and the immunoglobulin superfamily. *J Gen Virol* **72 (Pt 3)**:511-8.
83.   **Gregory SM, Davis BK, West JA, Taxman DJ, Matsuzawa S and others.** 2011. Discovery of a viral NLR homolog that inhibits the inflammasome. *Science* **331**:330-4.
84.   **Ray CA, Black RA, Kronheim SR, Greenstreet TA, Sleath PR and others.** 1992. Viral inhibition of inflammation: cowpox virus encodes an inhibitor of the interleukin-1 beta converting enzyme. *Cell* **69**:597-604.
85.   **Komiyama T, Ray CA, Pickup DJ, Howard AD, Thornberry NA and others.** 1994. Inhibition of interleukin-1 beta converting enzyme by the cowpox virus serpin CrmA. An example of cross-class inhibition. *J Biol Chem* **269**:19331-7.
86.   **Schotte P, Denecker G, Van Den Broeke A, Vandenabeele P, Cornelis GR and others.** 2004. Targeting Rac1 by the Yersinia effector protein YopE inhibits caspase-1-mediated maturation and release of interleukin-1beta. *J Biol Chem* **279**:25134-42.
87.   **Master SS, Rampini SK, Davis AS, Keller C, Ehlers S and others.** 2008. Mycobacterium tuberculosis prevents inflammasome activation. *Cell Host Microbe* **3**:224-32.
88.   **Miao EA, Rajan JV, Aderem A.** 2011. Caspase-1-induced pyroptotic cell death. *Immunol Rev* **243**:206-14.
89.   **Kaufmann SH, Schaible UE.** 2005. Antigen presentation and recognition in bacterial infections. *Curr Opin Immunol* **17**:79-87.
90.   **Nett MA, Cerretti DP, Berson DR, Seavitt J, Gilbert DJ and others.** 1992. Molecular cloning of the murine IL-1 beta converting enzyme cDNA. *J Immunol* **149**:3254-9.
91.   **Thornberry NA, Bull HG, Calaycay JR, Chapman KT, Howard AD and others.** 1992. A novel heterodimeric cysteine protease is required for interleukin-1 beta processing in monocytes. *Nature* **356**:768-74.
92.   **Horvath GL, Schrum JE, De Nardo CM, Latz E.** 2011. Intracellular sensing of microbes and danger signals by the inflammasomes. *Immunol Rev* **243**:119-35.

93. **Ichinohe T, Lee HK, Ogura Y, Flavell R, Iwasaki A.** 2009. Inflammasome recognition of influenza virus is essential for adaptive immune responses. *J Exp Med* **206**:79-87.

94. **Shao W, Yeretssian G, Doiron K, Hussain SN, Saleh M.** 2007. The caspase-1 digestome identifies the glycolysis pathway as a target during infection and septic shock. *J Biol Chem* **282**:36321-9.

95. **Keller M, Ruegg A, Werner S, Beer HD.** 2008. Active caspase-1 is a regulator of unconventional protein secretion. *Cell* **132**:818-31.

96. **Agard NJ, Maltby D, Wells JA.** 2010. Inflammatory stimuli regulate caspase substrate profiles. *Mol Cell Proteomics* **9**:880-93.

97. **Dinarello CA.** 1984. Interleukin-1 and the pathogenesis of the acute-phase response. *N Engl J Med* **311**:1413-8.

98. **Dinarello CA.** 2011. A clinical perspective of IL-1beta as the gatekeeper of inflammation. *Eur J Immunol* **41**:1203-17.

99. **Bonardi V, Cherkis K, Nishimura MT, Dangl JL.** 2012. A new eye on NLR proteins: focused on clarity or diffused by complexity? *Curr Opin Immunol* **24**:41-50.

100. **Friedman AR, Baker BJ.** 2007. The evolution of resistance genes in multi-protein plant resistance systems. *Curr Opin Genet Dev* **17**:493-9.

101. **Kofoed EM, Vance RE.** 2011. Innate immune recognition of bacterial ligands by NAIPs determines inflammasome specificity. *Nature* **477**:592-5.

102. **Martinon F, Burns K, Tschopp J.** 2002. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. *Mol Cell* **10**:417-26.

103. **Zou H, Li Y, Liu X, Wang X.** 1999. An APAF-1.cytochrome c multimeric complex is a functional apoptosome that activates procaspase-9. *J Biol Chem* **274**:11549-56.

104. **Faustin B, Lartigue L, Bruey JM, Luciano F, Sergienko E and others.** 2007. Reconstituted NALP1 inflammasome reveals two-step mechanism of caspase-1 activation. *Mol Cell* **25**:713-24.

105. **Renatus M, Stennicke HR, Scott FL, Liddington RC, Salvesen GS.** 2001. Dimer formation drives the activation of the cell death protease caspase 9. *Proc Natl Acad Sci U S A* **98**:14250-5.

106. **Mariathasan S, Newton K, Monack DM, Vucic D, French DM and others.** 2004. Differential activation of the inflammasome by caspase-1 adaptors ASC and Ipaf. *Nature* **430**:213-8.

107. **Feld GK, Thoren KL, Kintzer AF, Sterling HJ, Tang, II and others.** 2010. Structural basis for the unfolding of anthrax lethal factor by protective antigen oligomers. *Nat Struct Mol Biol* **17**:1383-90.

108. **Zhao Y, Yang J, Shi J, Gong YN, Lu Q and others.** 2011. The NLRC4 inflammasome receptors for bacterial flagellin and type III secretion apparatus. *Nature* **477**:596-600.

109. **Broz P, von Moltke J, Jones JW, Vance RE, Monack DM.** 2010. Differential requirement for Caspase-1 autoproteolysis in pathogen-induced cell death and cytokine processing. *Cell Host Microbe* **8**:471-83.

110. **Schagger H, Cramer WA, von Jagow G.** 1994. Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal Biochem* **217**:220-30.

111. **Damiano JS, Oliveira V, Welsh K, Reed JC.** 2004. Heterotypic interactions among NACHT domains: implications for regulation of innate immune responses. *Biochem J* **381**:213-9.

112. **Hornung V, Latz E.** 2010. Critical functions of priming and lysosomal damage for NLRP3 activation. *Eur J Immunol* **40**:620-3.

113. **Zhou R, Tardivel A, Thorens B, Choi I, Tschopp J.** 2010. Thioredoxin-interacting protein links oxidative stress to inflammasome activation. *Nat Immunol* **11**:136-40.

114. **Chisholm ST, Coaker G, Day B, Staskawicz BJ.** 2006. Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* **124**:803-14.

115. **Lu C, Wang A, Wang L, Dorsch M, Ocain TD and others.** 2005. Nucleotide binding to CARD12 and its role in CARD12-mediated caspase-1 activation. *Biochem Biophys Res Commun* **331**:1114-9.

116. **Geddes BJ, Wang L, Huang WJ, Lavellee M, Manji GA and others.** 2001. Human CARD12 is a novel CED4/Apaf-1 family member that induces apoptosis. *Biochem Biophys Res Commun* **284**:77-82.

117. **Poyet JL, Srinivasula SM, Tnani M, Razmara M, Fernandes-Alnemri T and others.** 2001. Identification of Ipaf, a human caspase-1-activating protein related to Apaf-1. *J Biol Chem* **276**:28309-13.

118. **Crook NE, Clem RJ, Miller LK.** 1993. An apoptosis-inhibiting baculovirus gene with a zinc finger-like motif. *J Virol* **67**:2168-74.

119. **Jeyaprakash AA, Klein UR, Lindner D, Ebert J, Nigg EA and others.** 2007. Structure of a Survivin-Borealin-INCENP core complex reveals how chromosomal passengers travel together. *Cell* **131**:271-85.

120. **Rothe M, Pan MG, Henzel WJ, Ayres TM, Goeddel DV.** 1995. The TNFR2-TRAF signaling complex contains two novel proteins related to baculoviral inhibitor of apoptosis proteins. *Cell* **83**:1243-52.

121. **Srinivasula SM, Ashwell JD.** 2008. IAPs: what's in a name? *Mol Cell* **30**:123-35.

122. **Verdecia MA, Huang H, Dutil E, Kaiser DA, Hunter T and others.** 2000. Structure of the human anti-apoptotic protein survivin reveals a dimeric arrangement. *Nat Struct Biol* **7**:602-8.

123. **Mace PD, Smits C, Vaux DL, Silke J, Day CL.** 2010. Asymmetric recruitment of cIAPs by TRAF2. *J Mol Biol* **400**:8-15.

124. **Eckelman BP, Drag M, Snipas SJ, Salvesen GS.** 2008. The mechanism of peptide-binding specificity of IAP BIR domains. *Cell Death Differ* **15**:920-8.

125. **Lin SC, Huang Y, Lo YC, Lu M, Wu H.** 2007. Crystal structure of the BIR1 domain of XIAP in two crystal forms. *J Mol Biol* **372**:847-54.

126. **Lu M, Lin SC, Huang Y, Kang YJ, Rich R and others.** 2007. XIAP induces NF-kappaB activation via the BIR1/TAB1 interaction and BIR1 dimerization. *Mol Cell* **26**:689-702.

127. **Riedl SJ, Renatus M, Schwarzenbacher R, Zhou Q, Sun C and others.** 2001. Structural basis for the inhibition of caspase-3 by XIAP. *Cell* **104**:791-800.
128. **Chai J, Shiozaki E, Srinivasula SM, Wu Q, Datta P and others.** 2001. Structural basis of caspase-7 inhibition by XIAP. *Cell* **104**:769-80.
129. **Huang Y, Park YC, Rich RL, Segal D, Myszka DG and others.** 2001. Structural basis of caspase inhibition by XIAP: differential roles of the linker versus the BIR domain. *Cell* **104**:781-90.
130. **Srinivasula SM, Hegde R, Saleh A, Datta P, Shiozaki E and others.** 2001. A conserved XIAP-interaction motif in caspase-9 and Smac/DIABLO regulates caspase activity and apoptosis. *Nature* **410**:112-6.
131. **Shiozaki EN, Chai J, Rigotti DJ, Riedl SJ, Li P and others.** 2003. Mechanism of XIAP-mediated inhibition of caspase-9. *Mol Cell* **11**:519-27.
132. **Roy N, Deveraux QL, Takahashi R, Salvesen GS, Reed JC.** 1997. The c-IAP-1 and c-IAP-2 proteins are direct inhibitors of specific caspases. *EMBO J* **16**:6914-25.
133. **Sanna MG, da Silva Correia J, Ducrey O, Lee J, Nomoto K and others.** 2002. IAP suppression of apoptosis involves distinct mechanisms: the TAK1/JNK1 signaling cascade and caspase inhibition. *Mol Cell Biol* **22**:1754-66.
134. **Eckelman BP, Salvesen GS.** 2006. The human anti-apoptotic proteins cIAP1 and cIAP2 bind but do not inhibit caspases. *J Biol Chem* **281**:3254-60.
135. **Eckelman BP, Salvesen GS, Scott FL.** 2006. Human inhibitor of apoptosis proteins: why XIAP is the black sheep of the family. *EMBO Rep* **7**:988-94.
136. **Srinivasula SM, Ahmad M, Fernandes-Alnemri T, Alnemri ES.** 1998. Autoactivation of procaspase-9 by Apaf-1-mediated oligomerization. *Mol Cell* **1**:949-57.
137. **Tanabe T, Chamaillard M, Ogura Y, Zhu L, Qiu S and others.** 2004. Regulatory regions and critical residues of NOD2 involved in muramyl dipeptide recognition. *Embo J* **23**:1587-97.
138. **Botella MA, Parker JE, Frost LN, Bittner-Eddy PD, Beynon JL and others.** 1998. Three genes of the Arabidopsis RPP1 complex resistance locus recognize distinct Peronospora parasitica avirulence determinants. *Plant Cell* **10**:1847-60.
139. **McDowell JM, Dhandaydham M, Long TA, Aarts MG, Goff S and others.** 1998. Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the RPP8 locus of Arabidopsis. *Plant Cell* **10**:1861-74.
140. **Meyers BC, Shen KA, Rohani P, Gaut BS, Michelmore RW.** 1998. Receptor-like genes in the major resistance locus of lettuce are subject to divergent selection. *Plant Cell* **10**:1833-46.
141. **Rentel MC, Leonelli L, Dahlbeck D, Zhao B, Staskawicz BJ.** 2008. Recognition of the Hyaloperonospora parasitica effector ATR13 triggers resistance against oomycete, bacterial, and viral pathogens. *Proc Natl Acad Sci U S A* **105**:1091-6.
142. **Growney JD, Dietrich WF.** 2000. High-resolution genetic and physical map of the Lgn1 interval in C57BL/6J implicates Naip2 or Naip5 in Legionella pneumophila pathogenesis. *Genome Res* **10**:1158-71.

143. **Romanish MT, Lock WM, de Lagemaat LN, Dunn CA, Mager DL.** 2007. Repeated Recruitment of LTR Retrotransposons as Promoters by the Anti-Apoptotic Locus NAIP during Mammalian Evolution. *PLoS Genet* **3**:e10.

144. **Campbell L, Potter A, Ignatius J, Dubowitz V, Davies K.** 1997. Genomic variation and gene conversion in spinal muscular atrophy: implications for disease process and clinical phenotype. *Am J Hum Genet* **61**:40-50.

145. **Growney JD, Scharf JM, Kunkel LM, Dietrich WF.** 2000. Evolutionary divergence of the mouse and human Lgn1/SMA repeat structures. *Genomics* **64**:62-81.

146. **Rajcan-Separovic E, Mahadevan MS, Lefebvre C, Besner-Johnston A, Ikeda JE and others.** 1996. FISH detection of chromosome polymorphism and deletions in the spinal muscular atrophy (SMA) region of 5q13. *Cytogenet Cell Genet* **75**:243-7.

147. **Romanish MT, Nakamura H, Lai CB, Wang Y, Mager DL.** 2009. A novel protein isoform of the multicopy human NAIP gene derives from intragenic Alu SINE promoters. *PLoS ONE* **4**:e5761.

148. **Endrizzi MG, Hadinoto V, Growney JD, Miller W, Dietrich WF.** 2000. Genomic sequence analysis of the mouse Naip gene array. *Genome Res* **10**:1095-102.

149. **Heckman KL, Pease LR.** 2007. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc* **2**:924-32.

150. **Bell JK, Mullen GE, Leifer CA, Mazzoni A, Davies DR and others.** 2003. Leucine-rich repeats and pathogen recognition in Toll-like receptors. *Trends Immunol* **24**:528-33.

151. **Baker B, Zambryski P, Staskawicz B, Dinesh-Kumar SP.** 1997. Signaling in plant-microbe interactions. *Science* **276**:726-33.

152. **Kofoed EM, Guerbadot M, Schaufele F.** 2010. Structure, affinity, and availability of estrogen receptor complexes in the cellular environment. *J Biol Chem* **285**:2428-37.

153. **Kobe B, Deisenhofer J.** 1995. A structural basis of the interactions between leucine-rich repeats and protein ligands. *Nature* **374**:183-6.

154. **Ellis J, Dodds P, Pryor T.** 2000. Structure, function and evolution of plant disease resistance genes. *Curr Opin Plant Biol* **3**:278-84.

155. **Boyden ED, Dietrich WF.** 2006. Nalp1b controls mouse macrophage susceptibility to anthrax lethal toxin. *Nat Genet* **38**:240-4.

156. **Hammond-Kosack KE, Jones JD.** 1997. Plant Disease Resistance Genes. *Annu Rev Plant Physiol Plant Mol Biol* **48**:575-607.

157. **Parniske M, Hammond-Kosack KE, Golstein C, Thomas CM, Jones DA and others.** 1997. Novel disease resistance specificities result from sequence exchange between tandemly repeated genes at the Cf-4/9 locus of tomato. *Cell* **91**:821-32.

158. **Schroder K, Tschopp J.** 2010. The inflammasomes. *Cell* **140**:821-32.

159. **Bonardi V, Tang S, Stallmann A, Roberts M, Cherkis K and others.** 2011. Expanded functions for a family of plant intracellular immune receptors beyond specific recognition of pathogen effectors. *Proc Natl Acad Sci U S A* **108**:16463-8.

160.	**Eitas TK, Dangl JL.** 2010. NB-LRR proteins: pairs, pieces, perception, partners, and pathways. *Curr Opin Plant Biol* **13**:472-7.

161.	**Fink SL, Bergsbaken T, Cookson BT.** 2008. Anthrax lethal toxin and Salmonella elicit the common cell death pathway of caspase-1-dependent pyroptosis via distinct mechanisms. *Proc Natl Acad Sci U S A* **105**:4312-7.

162.	**Klimpel KR, Arora N, Leppla SH.** 1994. Anthrax toxin lethal factor contains a zinc metalloprotease consensus sequence which is required for lethal toxin activity. *Mol Microbiol* **13**:1093-100.

163.	**D'Osualdo A, Weichenberger CX, Wagner RN, Godzik A, Wooley J and others.** 2011. CARD8 and NLRP1 undergo autoproteolytic processing through a ZU5-like domain. *PLoS One* **6**:e27396.

164.	**Kupz A, Guarda G, Gebhardt T, Sander LE, Short KR and others.** 2012. NLRC4 inflammasomes in dendritic cells regulate noncognate effector function by memory CD8(+) T cells. *Nat Immunol* **13**:162-9.

165.	**Sauer JD, Pereyre S, Archer KA, Burke TP, Hanson B and others.** 2011. Listeria monocytogenes engineered to activate the Nlrc4 inflammasome are severely attenuated and are poor inducers of protective immunity. *Proc Natl Acad Sci U S A* **108**:12419-24.

166.	**Krantz BA, Melnyk RA, Zhang S, Juris SJ, Lacy DB and others.** 2005. A phenylalanine clamp catalyzes protein translocation through the anthrax toxin pore. *Science* **309**:777-81.

Appendix One. Formation of the NLRC4 inflammasome depends on flagellin. (**A**) NAIP5 and NLRC4 were expressed in 293T cells for 48 hours (without flagellin), followed by digitonin native cell extract preparation, and separation by BN-PAGE. Lanes were then excised from the first dimension gel and run on an SDS-PAGE second dimension. Specific proteins were detected in both monomeric and oligomerized state by western blotting. (**B**) NAIP5-Flag was expressed in 293T cells in the absence of NLRC4 and with or without GFP-flagellin, and samples were as in **a**, but with western blotting to detect NAIP5-Flag. (**C**) As in **A**, but with expression of 6x-Myc-FlaA in place of NLRC4 and NAIP5.

A)
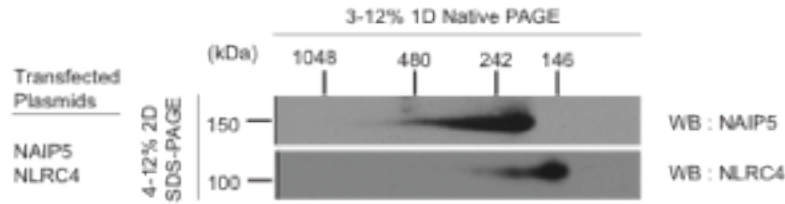
B)
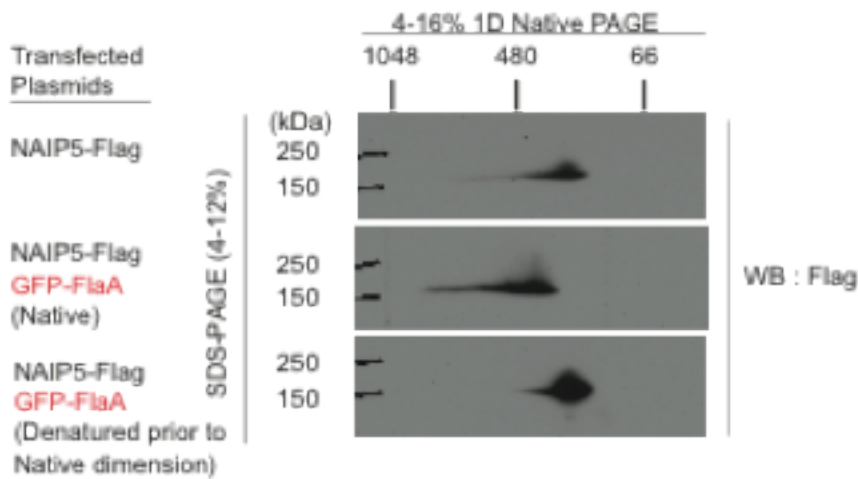
C)

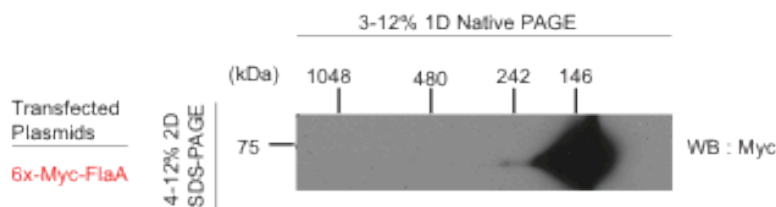**Appendix Two.** The response of wild-type and mutant NLRs to cytoplasmic flagellin in the reconstituted HEK293T system. Cells were transiently transfected with the indicated combination of plasmids for 36 hrs., followed by cytosolic delivery of FlaA[5μg/ml] or vehicle alone for 12 hrs. Cells were analyzed for GFP expression by flow cytometry. These data are a representative of two independent experiments.
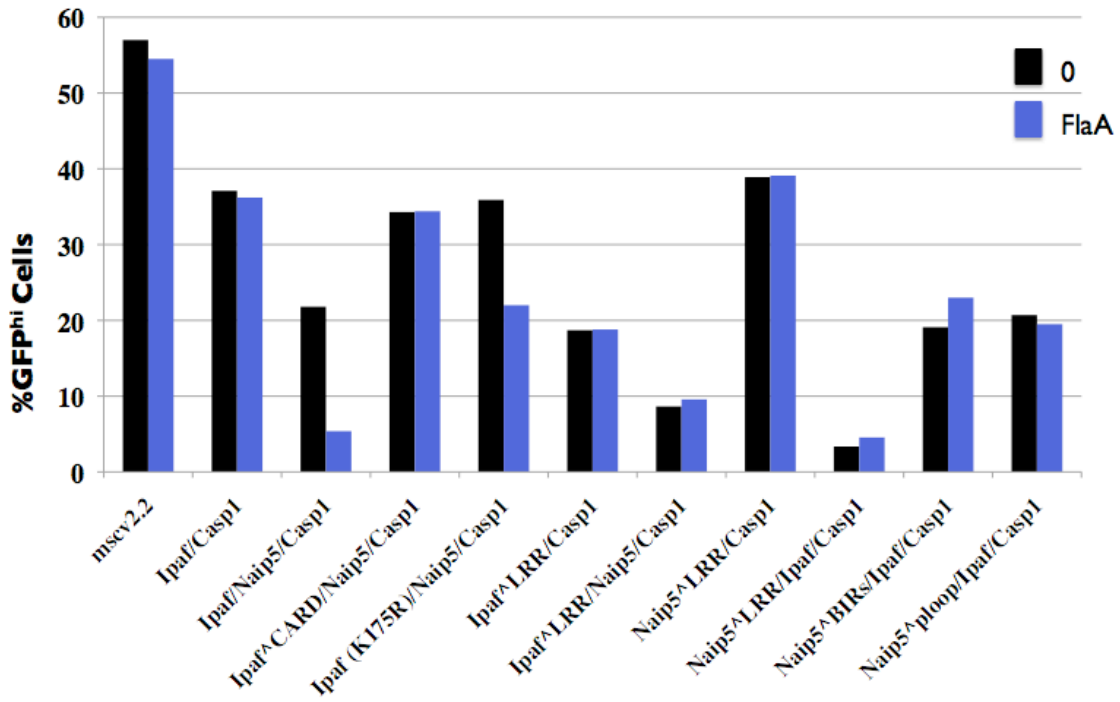
Appendix Three. Table showing the ability of inner-rod proteins to cause retroviral lethality in C57Bl/6 macrophages, and induce NLRC4 oligomerization in the 293T blue native assay. The matrix lists the name of the inner rod protein, the bacteria they are derived from, whether they are lethal when delivered into macrophage cytoplasm (column R.L.), and their ability to activate various NAIP-NLRC4 inflammasomes (columns: Naip-1, -2, -5, and -6). ND = not determined; - means it fails, + means it works (kills cells and/or activates inflammasome oligomerization).

Inner-rod protein	Bacterial Genera & Species	R.L.*	Naip 1	Naip 2	Naip 5	Naip 6
Prg J	Salmonella typhimurium (SPI I)	+	-	+	-	-
EscI	Escheria coli (enteropathogenic; LEE)	+	-	+	-	-
MxiI	Shigella flexneri	+	-	+	-	-
SsaI	Salmonella typhimurium (SPI 2)	-	-	-	-	-
A33	Vibrio cholera	ND	-	+	-	-
VP1691	Vibrio parahemolyticus	+	-	+	-	-
VP1368	Vibrio parahemolyticus	+	-	+	-	-
PscI	Pseudomonas aeruginosa	+	+	+	-	-
YscI	Yersinia pestis	ND	-	+	-	-
YPO0262	Yersinia pestis	ND	+	+	-	-
Citrobacter	Citrobacter rodentium	ND	-	+	-	-
HrpB	Pseudomonas syringae	ND	-	-	-	-
BscI	Bordetella pertussis	+	-	+	-	-
BsaK	Burkholderia pseudomallei; Rod1	+	+	+	-	-
A1902	Burkholderia pseudomallei; Rod2	-	-	-	-	-
EprJ	Escheria coli (enterohemorrhagic)	ND	-	+	-	-

Flagellin	Bacterial Genera & Species	R.L.	Naip 1	Naip 2	Naip 5	Naip 6
FlaA	Legionella pneumophila	+	-	-	+	+
FliC	Salmonella typhimurium	+	-	-	+	+
FliC	Pseudomonas aeruginosa	+	-	-	+	+

*Retroviral lethality assays for all rod proteins (except PrgJ) were conducted by Ed Miao[48].

Appendix Four. Comparison of the NAIP paralogues by pairwise sequence alignments, and mapping NAIP polymorphisms onto the predicted LRR structure. Red boxes indicate where β-strands are located on primary sequence. A) Polymorphisms between NAIP2 and NAIP5 mapped onto model; primary alignments of all paralogues (amino acids 1230-1448^{NAIP2}). B) Comparison of polymorphic residues between NAIP1 and NAIP2 on left, and NAIP2 and NAIP5 on right (see Fig.1C legend for PHYRE modeling data).

A)

B)

NAIP1 & NAIP2   NAIP2 & NAIP5
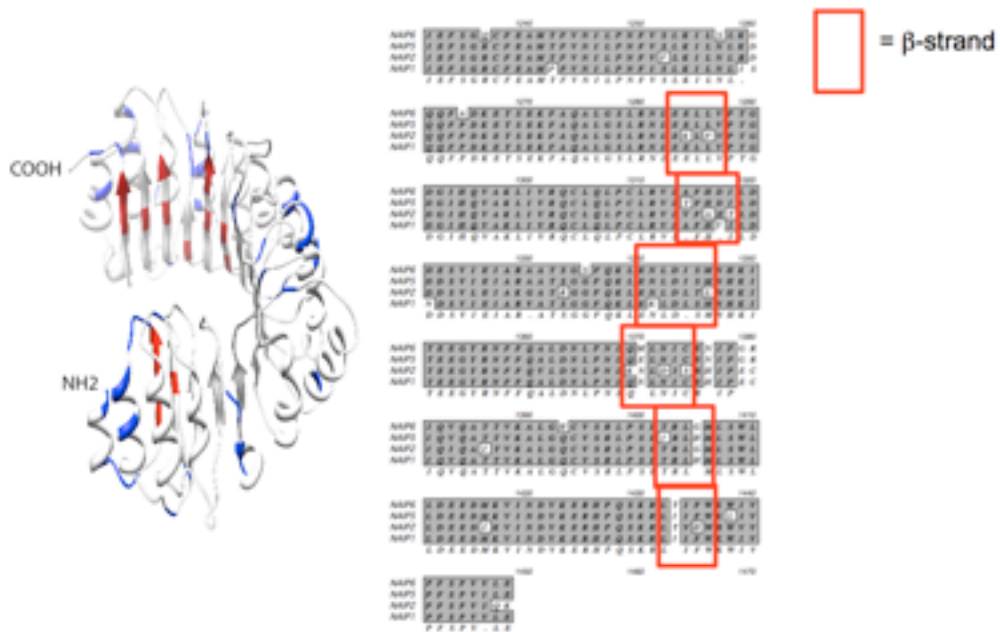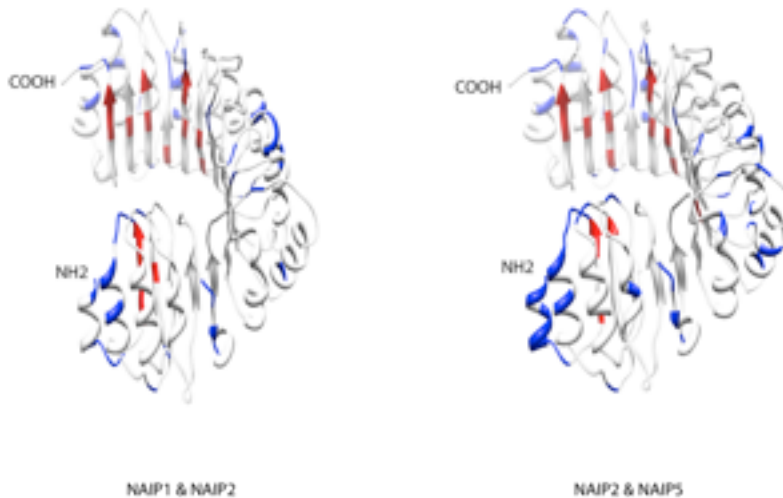
Appendix Five. Listeria strains engineered to inducibly secrete ActA-Inner-rod fusion proteins during infection are tested for the ability to induce pyroptosis in primary macrophages. All strains were generated by single copy-recombination and are under control of an IPTG inducible promoter, and are on the L.m.ΔFlaA background. Rods that were found to activate both NAIP1 and NAIP2 in the 293T assay were tested. A) Primary BMDMs were infected with the various L.m. strains (MOI=5) and LDH release measured 5.5 hours later. B) infections and assays were as in (A), except the macrophages were Naip5$^{-/-}$. C) as in (A) and (B) except macrophage genotype was Nlrc4$^{-/-}$ (a.k.a. Ipaf$^{-/-}$). Results are avg. and std. dev. of six technical replicates from a single experiment. Listeria strains also need genotype verification before conclusions can be drawn.

A) B6

B) NAIP5-/-

C) IPAF-/-
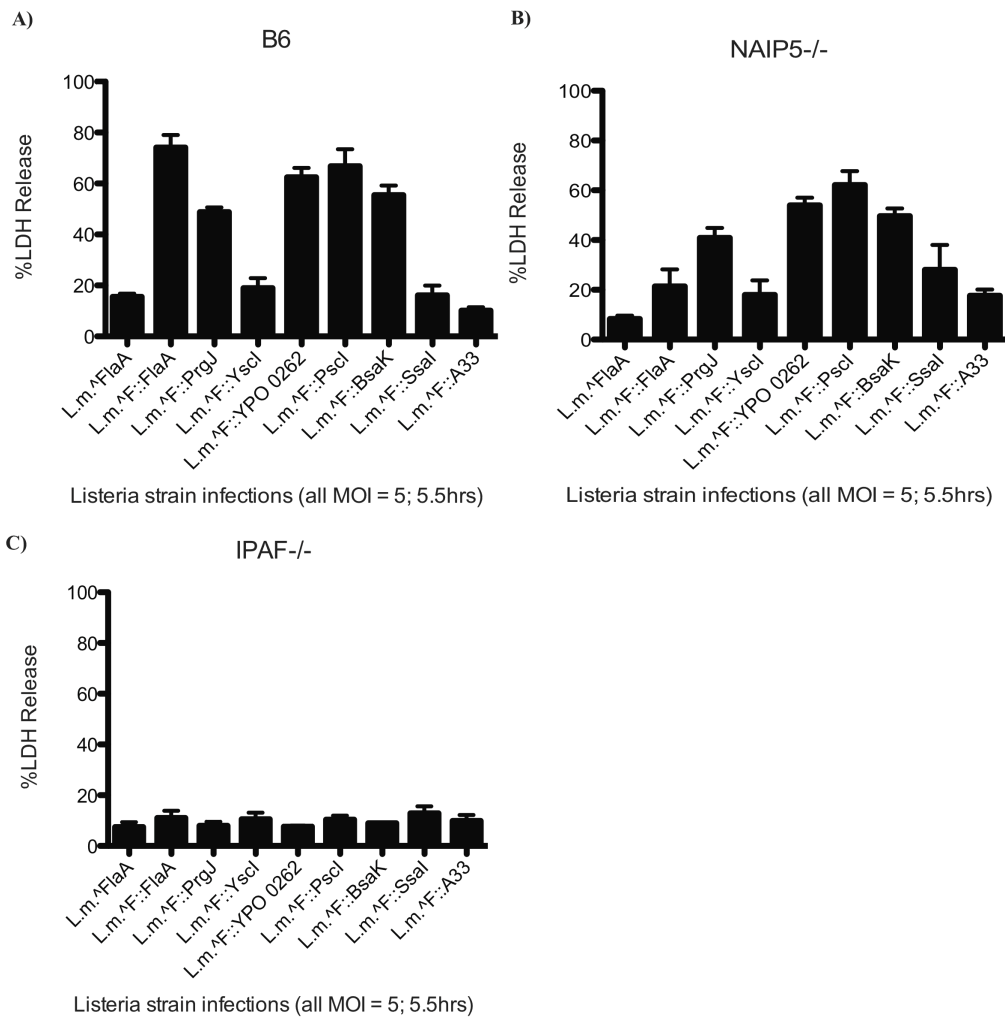
Appendix Six. Chimera oligomerization and loading controls (part A). Weak staining for 6x-Myc-PrgJ in the second set is an artifact of the G250 dye clogging the second PVDF membrane in a transfer sandwich, and therefore causing 'blow-though' of small MW proteins (This is true for Appendices Six-Eleven). The NAIP5 antibody recognizes amino acids 961-978 of NAIP5 and NAIP6 but does not cross-react with NAIP2. Therefore, absence of staining with the NAIP5 antibody indicates that the protein is NAIP2 for that region (in the LRRs). Staining indicates the chimera is either NAIP5 or NAIP6 in this region (this is true for Appendices Six-Nine, because Ten and Eleven lack accompanying NAIP5 blots).
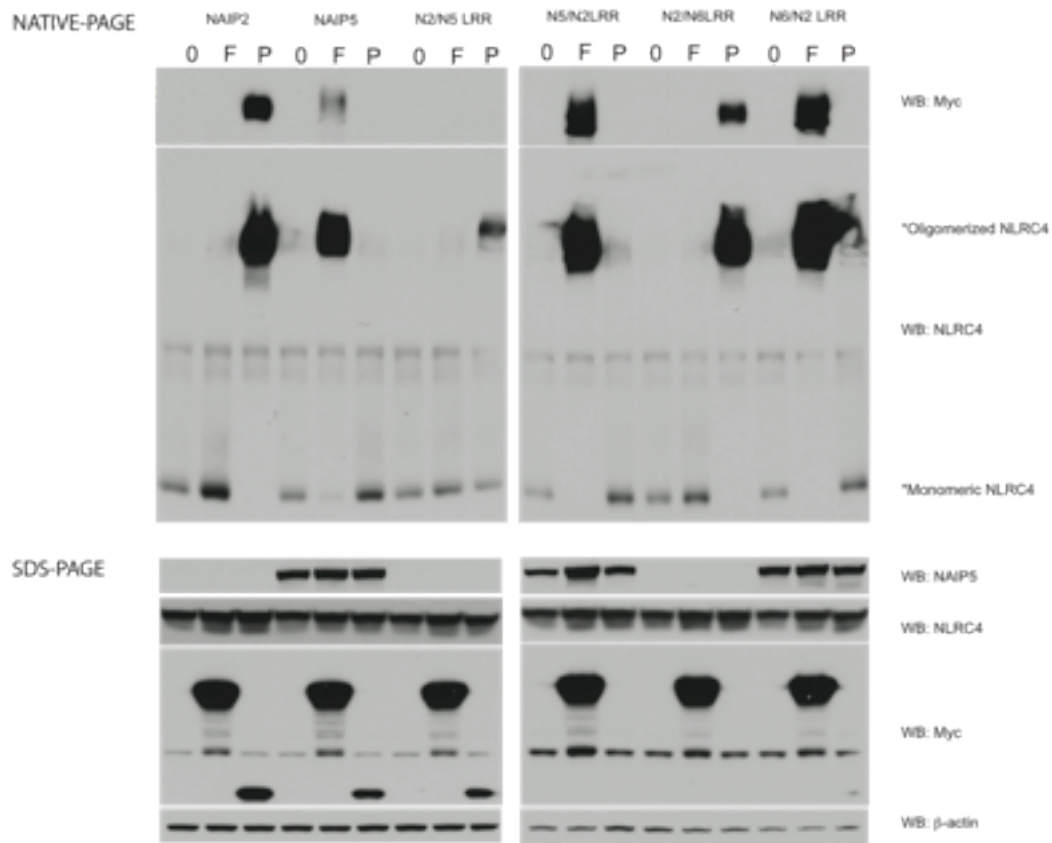
Appendix Seven. Chimera oligomerization and loading controls (part B). Weak staining for 6x-Myc-PrgJ in the second set is an artifact of the G250 dye clogging the second PVDF membrane in a transfer sandwich, and therefore causing 'blow-though' of small MW proteins.

Appendix Eight. Chimera oligomerization and loading controls (part C). Weak staining for 6x-Myc-PrgJ in the second set is an artifact of the G250 dye clogging the second PVDF membrane in a transfer sandwich, and therefore causing 'blow-though' of small MW proteins.
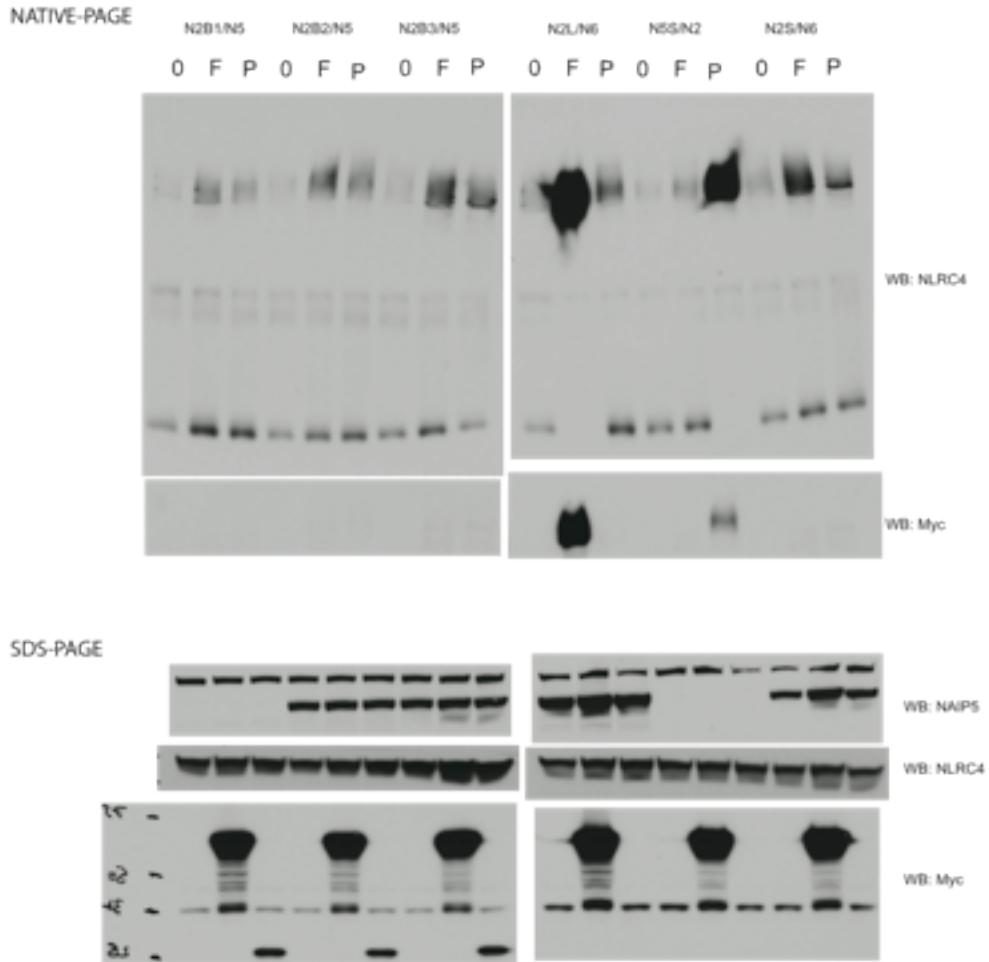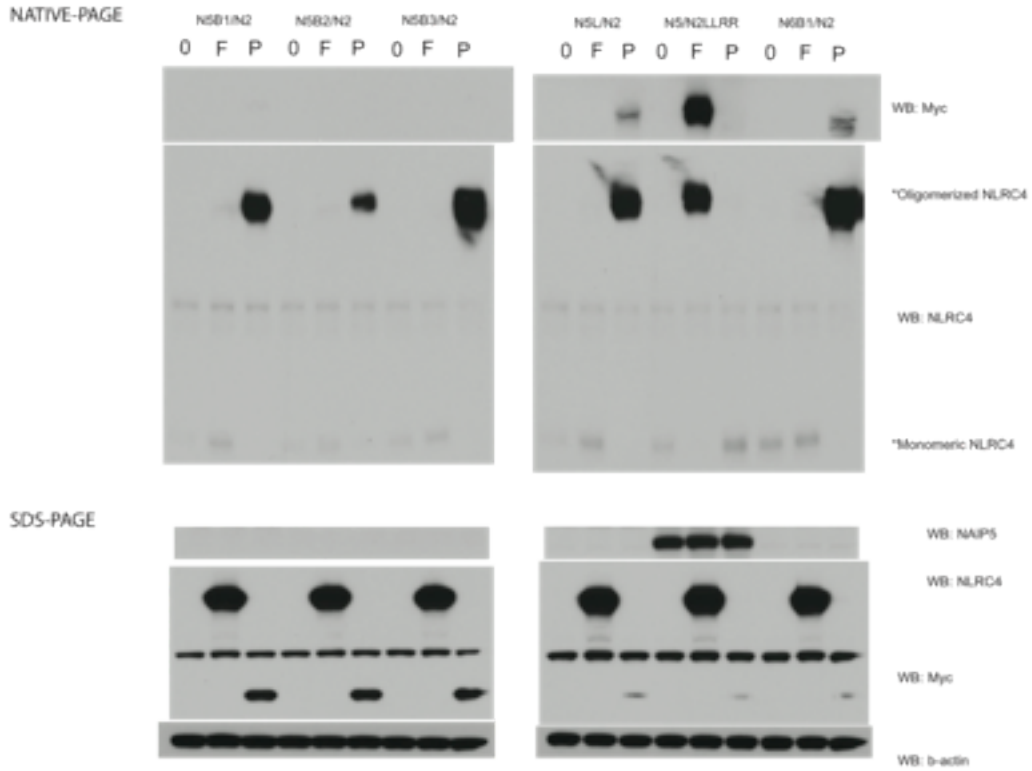
Appendix Nine. Chimera oligomerization and loading controls (part D). Weak staining for 6x-Myc-PrgJ in the second set is an artifact of the G250 dye clogging the second PVDF membrane in a transfer sandwich, and therefore causing 'blow-though' of small MW proteins.
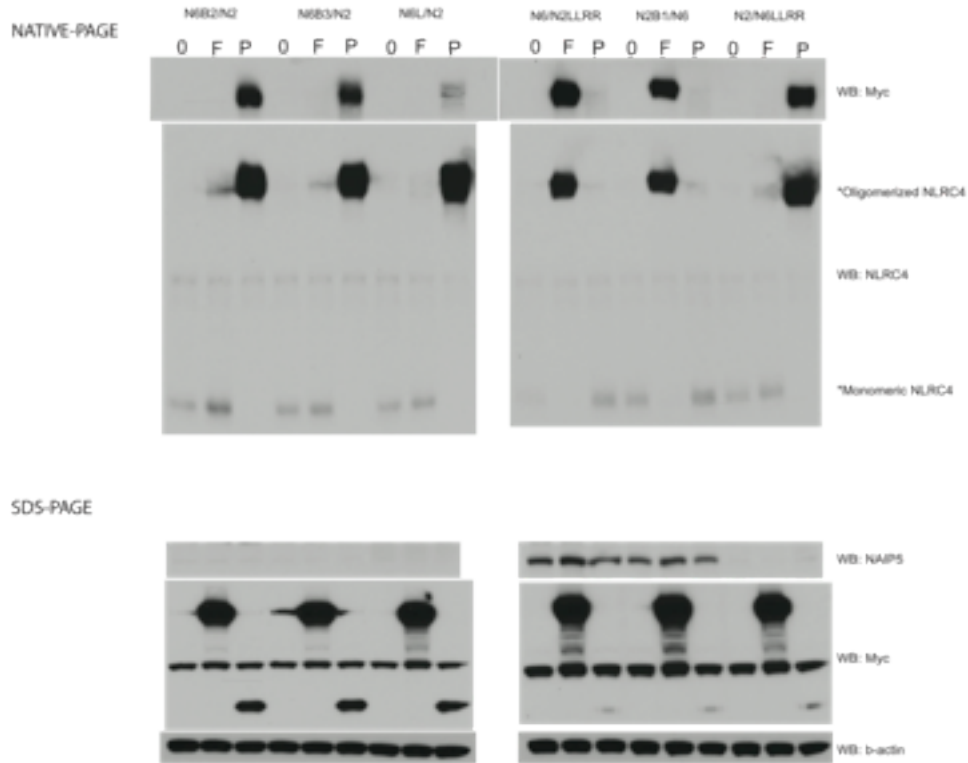
Appendix Ten. Chimera oligomerization and loading controls (part E)

NATIVE-PAGE

| N2L/N5 | N2L/N6 | N2/N5LLRR | N2B3/N6 | N6 |
| 0 F P | 0 F P | 0 F P | 0 F P | 0 F P |

WB: Myc

*Oligomerized NLRC4

WB: NLRC4

*Monomeric NLRC4

SDS-PAGE

WB: NAIP5

WB: NLRC4

WB: Myc

WB: b-actin
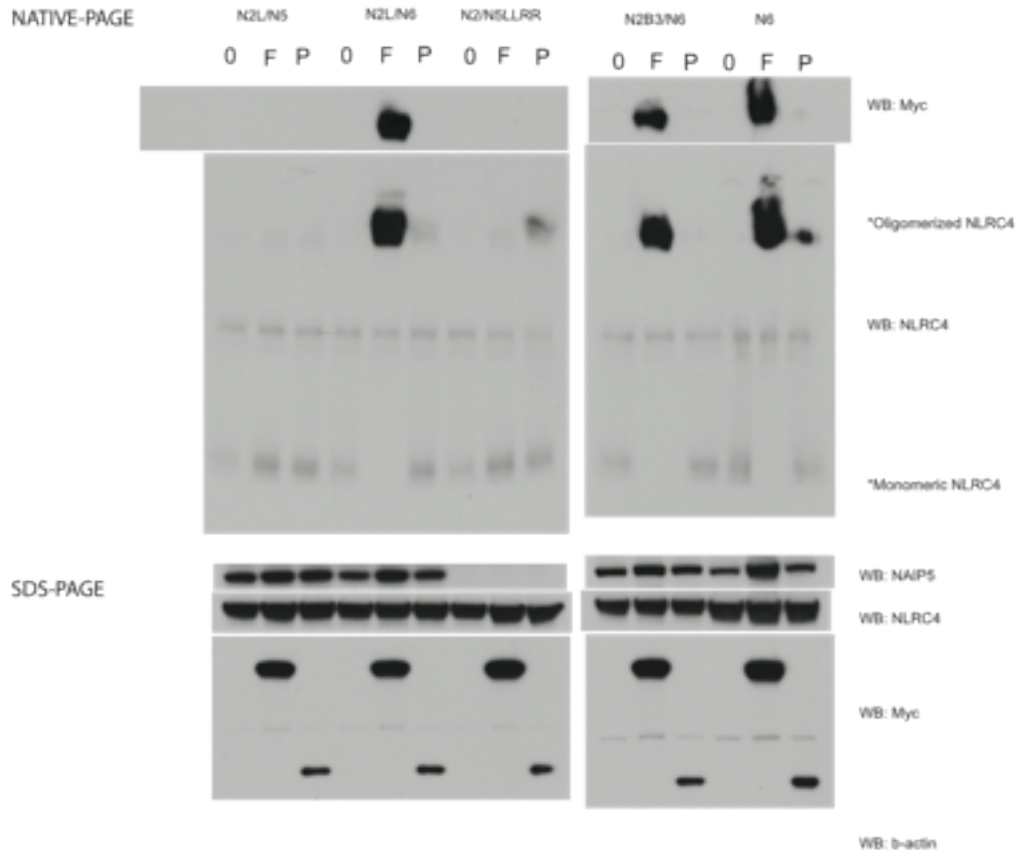
Appendix Eleven. Chimera oligomerization (part F). Western blot for NLRC4 oligomerization after blue native PAGE.
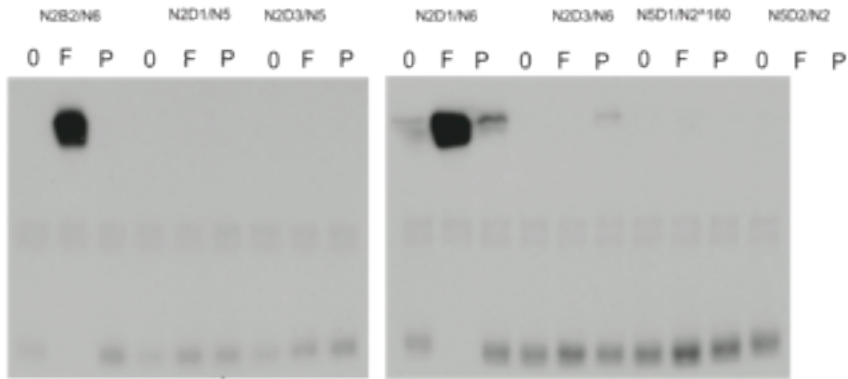
Appendix Twelve. Chimera oligomerization (part G). Western blot for NLRC4 following blue native PAGE.
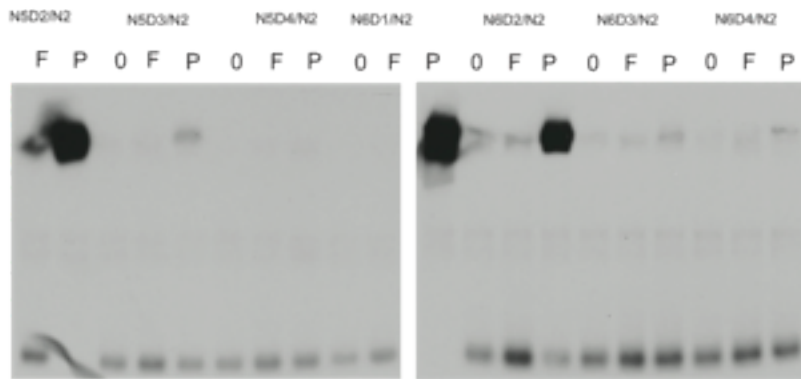
Appendix Thirteen. Multiple sequence alignment of C57Bl/6 NAIP1, NAIP2, NAIP5 and NAIP6. * = conserved residue; chemically similar amino acid substitutions = (.); non-similar substitution = (˙). Red highlights the minimal (822-1078) ligand specificity determining region of the NAIP LRRs.

```
NAIP1     1 MAEHGESSEDRISEIDYEFLPELSALLGVDAVQLAKSQEEEEHKERMKMKKGFNSQMRSE   60
NAIP2     1 MAAQGEAVEEIICEFDDDLVSELSTLLRVDALSVLKRQQEEDHKTRMKMKKGFNSQMRSE   60
NAIP5     1 MAEHGESSEDRISEIDYEFLPELSALLGVDAFQVAKSQEEEEHKERMKMKKGFNSQMRSE   60
NAIP6     1 MAEHGESSEDRISEIDYEFLAELSARFGMNLVQLAKSQEEEDHKERMKMKKGFNSQMRSE   60
            **  .**. *. *.* *  . . ***.     .      . * *.**.** **************

NAIP1    61 AKRLKTFETYDTFRSWTPQEMAAAGFYHTGVKLGVQCFCCSLILFGNSLRKLPIERHKKL  120
NAIP2    61 AKRLKTFETYDKFRSWTPQEMAAAGFYHTGVKLGVQCFCCSLILFSTRLRKLPIENHKKL  120
NAIP5    61 AKRLKTFETYDTFRSWTPQEMAAAGFYHTGVRLGVQCFCCSLILFGNSLRKLPIERHKKL  120
NAIP6    61 AKRLKTFESYDTFRSWTPQEMAAAGFYHTGVKLGVQCFCCSLILFGNSLRKLPIERHKKL  120
            ********.** ******************.*************  . ******* ****

NAIP1   121 RPECEFLQGKDVGNIGKYDIRVKSPEKMLRGGKARYHEEEARLESFEDWPFYAHGTSPRV  180
NAIP2   121 RPECEFLLGKDVGNIGKYDIRVKSPEKMLRGDKARYHEEEARLESFEDWPFYAHGTSPRV  180
NAIP5   121 RPECEFLQGKDVGNIGKYDIRVKRPEKMLRGGKARYHEEEARLESFEDWPFYAHGTSPRV  180
NAIP6   121 RPECEFLQGKDVGNIGKYDIRVKSPEKMLRGGKARYHEEEARLESFEDWPFYAHGTSPRA  180
            ******* *************** ******* **************************

NAIP1   181 LSAAGFVFTGKRDTVQCFSCGGSLGNWEEGDDPWKEHAKWFPKCEFLQSKKSSEEIAQYI  240
NAIP2   181 LSAAGFVFTGKRDTVQCFSCGGCLGNWEEGDDPWKEHAKWFPKCEFLQSKKSPEEITQYV  240
NAIP5   181 LSAAGFVFTGKRDTVQCFSCGGSLGNWEEGDDPWKEHAKWFPKCEFLQSKKSSEEIAQYI  240
NAIP6   181 LSAAGFVFTGKRDTVQCFSCGGSLGNWEEGDDPWKEHAKWFPKCEFLQSKKSSEEIAQYI  240
            *********************.**************************** ***.**.

NAIP1   241 QGYEGFVHVTGEHFVNSWVRRELPMVSAYCNDSVFANEELRMDTFKDWPHESPVAVDALV  300
NAIP2   241 QSYEGFLHVTGEHFVNSWVRRELPMVSAYCNDSVFANEELRMDTFKDWPHESPGAVEALV  300
NAIP5   241 QSYEGFVHVTGEHFVKSWVRRELPMVSAYCNDSVFANEELRMDMFKDWPQESPVGVEALV  300
NAIP6   241 QDYEGFVHVTGEHFVKSWVRRELPMVSAYCNDSVFTNEELRMDMFKDWPQESPVGFEALV  300
            * ****.******** *****************.******* *****.***    .***

NAIP1   301 RAGLFYTGKKGIVQCFSCGGCMEKCTEGDDPIQEHNKFFPNCVFLQTPKSSAEVIPALQS  360
NAIP2   301 KAGLFYTGKRDIVQCFSCGGCMEKWAEGDNPIEDHTKFFPNCVFLQTLKSSAEVIPALQS  360
NAIP5   301 RAGFFYTGKKDIVRCFSCGGCLEKWAEGDDPMEDHIKFFPECVFLQTLKSSAEVIPTLQS  360
NAIP6   301 RAGFFYTGKKDIVRCFSCGGCLEKWAEGDDPMEDHIKFFPECVFLQTLKSSAEVIPTLQS  360
            .** *****. **.*******.** .*** *...* **** ****** ********.***

NAIP1   361 HCALPEAMETTSESNHDDPAAVHSTVVG--------------------------------  388
NAIP2   361 HCALPEAMETTSESNHDDAAAVHSTVVDVSPSEAQELEPASSLVSVLCRDQDHSEAQGRG  420
NAIP5   361 QYALPEATETTRESNHGDAAAVHSTVVD--------------------------------  388
NAIP6   361 QYALPEATETTRESNHDDAAAVHSTVVD--------------------------------  388
            . ***** *** **** * ********

NAIP1   389 ------------LGRSEAQWFQEARSLSEQLRDNYTKATFRHMNLPEVCSSLGTDHLIGC  436
NAIP2   421 CASSGTYLPSTDLGQSEAQWLQEARSLSEQLRDTYTKATFRHMNLPEVYSSLGTDHLLSC  480
NAIP5   389 ------------LGRSEAQWFQEARSLSEQLRDNYTKATFRHMNLPEVCSSLGTDHLLSC  436
NAIP6   389 ------------LGRSEAQWFQEARSLSEQLRDTYTKTSFCHMNLPEVCSSLGTDHLLGC  436
                        **.***** *************.***..* ******* ********. *

NAIP1   437 DVSIISKHISQPVQGALTIPEVFSNLSSVMCVEGETGSGKTTFLKRIAFLWASGCCPLLY  496
NAIP2   481 DVSIISKHISQPVQGSLTIPEVFSNLNSVMCVEGEAGSGKTTFLKRIAFLWASGCCPLLY  540
NAIP5   437 DVSIISKHISQPVQEALTIPEVFSNLNSVMCVEGETGSGKTTFLKRIAFLWASGCCPLLY  496
NAIP6   437 DVSIISKHVSQPVQGALTIPEVFSNLSSVMCVEGEAGSGKTTFLKRIAFLWASGCCPLLY  496
            ********.***** .********** *******.***********************

NAIP1   497 RFQLVFYLSLSSITPDQGLANIICAQLLGAGGCISEVCLSSIIQQLQHQVLFLLDDYSGL  556
NAIP2   541 RFQLVFYLSLSSITPGQELAKIICAQLLGAGGCISEVCLSSIIQQLQHQVLFLLDDYSGL  600
```

```
NAIP5   497 RFQLVFYLSLSSITPDQGLANIICAQLLGAGGCISEVCLSSSIQQLQHQVLFLLDDYSGL  556
NAIP6   497 RFQLVFYLSLSSITPDQGLANIICTQLLGAGGCISEVCLSSSIQQLQHQVLFLLDDYSGL  556
            ************** * ** *** .*************** ******************

NAIP1   557 ASLPQALHTLITKNYLSRTCLLIAVHTNRVRGIRSYLDTSLEIKEFPLSNTVYILKKFFS  616
NAIP2   601 ASLPQALHTLITKNYLSRTCLLIAVHTNKVRGIRPYLDTSLEIKEFPFYNTVSVLRKLFS  660
NAIP5   557 ASLPQALHTLITKNYLSRTCLLIAVHTNRVRDIRLYLGTSLEIQEFPFYNTVSVLRKFFS  616
NAIP6   557 ASLPQALHTLITKNYLFRTCLLIAVHTNRVRDIRPYLGTSLEIQEFPFYNTVFVLRKFFS  616
            **************** *********** .** ** ** *****.***   *** .*.* **

NAIP1   617 HNIKRLLEFMVYFGQNEDLQGIHKTPLFVAAVCTDWFENPSDQ-PFQDMALFKSYMQYLS  675
NAIP2   661 HDIMRVRKFINYFGFHEELQGIHKTPLFVAAVCTDWFKNPSDQ-PFQDVALFKAYMQYLS  719
NAIP5   617 HDIICVEKLIIYFIDNKDLQGVYKTPLFVAAVCTDWIQNASAQDKFQDVTLFQSYMQYLS  676
NAIP6   617 HDIICVEKLIIYFSENKDLQGVYKTPLFVAAVCNDWNQNASAQDDFQDVTLFHSYMQYLS  676
            * *  .   . **   . .***. **********.**   * * *   ***..**..******

NAIP1   676 LKHKGAAKPLQATVSSCGQLALTGLFSSCFEFNSDDLAEAGVDEDEELTTCLMSKFTAQR  735
NAIP2   720 LKHKGAAKPLQATVSSCGQLALTGLFSSCFEFNSDNLAEAGVDEDEELTTCLMSKFTAQR  779
NAIP5   677 LKYKATAEPLQATVSSCGQLALTGLFSSCFEFNSDDLAEAGVDEDEKLTTLLMSKFTAQR  736
NAIP6   677 LKYKATAESLQATVSSCGQLALTGLFSSCFEFNSDDLAEAGVDEDVKLTTFLMSKFTAQR  736
            ** * .*  ****************************** *********  *** *********

NAIP1   736 LRPVYRFLGPLFQEFLAAMRLTELLSSDRQEDQDLGLYYLRQINSPLKALTTYNNFLKYV  795
NAIP2   780 LRPVYRFLGPLFQEFLAAVRLTELLSSDRQEDQDLGLYYLRQINSPLKAMSIYHTFLKYV  839
NAIP5   737 LRPVYRFLGPLFQEFLAAVRLTELLSSDRQEDQDLGLYYLRQIDSPLKAINSFNIFLYYV  796
NAIP6   737 LRPVYRFLGPLFQEFLAAVRLTELLSSDRQEDQDLGLYYLRQIDSPLKAINSFNIFLYYV  796
            ****************** .*********************** *****.   .. ** **

NAIP1   796 FSHPSSKAGPTVVSHLLHLVDETELLENTYKNEDYVNHPPGTSRIMKGLKELWLLSPEYY  855
NAIP2   840 SSHPSSKAAPTVVSHLLQLVDEKESLENMSENEDYMKLHPEALLWIECLRGLWQLSPESF  899
NAIP5   797 SSHSSSKAAPTVVSHLLQLVDEKESLENMSENEDYMKLHPQTFLWFQFVRGLWLVSPESS  856
NAIP6   797 SSHSSSKAAPTVVSHLLQLVDEKESLENMSENEDYMKLHPQTFLWFQFVRGLWLVSPESF  856
             ** **** ********.**** * ***    ****.    * .     .. ** .***

NAIP1   856 SSFVSEHLLRIALNFAYESNTVAECSPFILQFLRGRTLALKVLNLQYFRDHPESLLLVKS  915
NAIP2   900 SLFISENLLRICLNFAHESNTVAACSPVILQFLRGRTLDLKVLSLQYFWDHPETLLLLKS  959
NAIP5   857 SSFVSEHLLRLALIFAYESNTVAECSPFILQFLRGKTLALRVLNLQYFRDHPESLLLLRS  916
NAIP6   857 SSFVSEHLLRLALIFAYESNTVAECSPFILQFLRGRTLALRVLNLEYFWDHPESLLLLRS  916
            * *.**.***..* ** ****** *** *******.** *.** *.** ****.***..*

NAIP1   916 LEVSINGNKVPKVVDYSVMEKSFETLQPPTIDQDYASAFEQMKEHEKNLSENEETIKSIK  975
NAIP2   960 IKISLNGNNWVQRIDFSLIEKSFEKVQPPTIDQDYAIAFQPINEVQKNLSEKKHIIKKYE 1019
NAIP5   917 LKVSINGNKMSSYVDYS-FKTYFENLQPPAIDEEYTSAFEHISEWRRNFAQDEEIIKNYE  975
NAIP6   917 LKVSINGNKMSSYVDYS-FKTYFENLQPPAINEEYTSAFEHVSEWRRNFAQDEEIIKNYE  975
            . .*.***     .*.*     ** .***.* ..*. **.  *  .* ..     **

NAIP1   976 NIFPLQPPKISSGYWKLSPKPCKIPRLEVGVTNMGPADQALLQVLMEVFSASQSIEFRLS 1035
NAIP2  1020 DMKHQIPLNISTGYWKLSPKPYKIPKLEVQVTNTGPADQALLQVLMEVFSASQSIEFRLS 1079
NAIP5   976 NIRPRALPDISEGYWKLSPKPCKIPKLEVQVNNTDAADQALLQVLMEVFSASQSIEFRLF 1035
NAIP6   976 NIWPRALPDISEGYWNLSPKPCKIPKLEVQVNNMGPADQALLQVLMEVFSASQSIEFHLF 1035
             .          ** *** ***** ***.*** *.*   ********************.*

NAIP1  1036 DSSGFLESIRPALELSKASVTKCSMSRLELSRAEQELLLTLPALQSLEVSETNQLPDQLF 1095
NAIP2  1080 DSSGFLESIRPALELSKASVTKCSMSRLELSREDQKLLLTLPTLQSLEVSETNQLPDQLF 1139
NAIP5  1036 NSSGFLESICPALELSKASVTKCSMSRLELSRAEQELLLTLPALQSLEVSETNQLPEQLF 1095
NAIP6  1036 NSSGFLESIRPALELSKASVTKCSMSRLELSRAEQELLLTLPALQSLEVSETNQLPDQLF 1095
            ******** ********************** .* ****** .************* ***

NAIP1  1096 HNLHKFLGLKELCVRLDGKPDVLSVLPGEFPNLLHMEKLSIRTSMESDLSKLVKLIQNSP 1155
NAIP2  1140 HNLHKFLGLKELCVRLDSKPDVLSVLPGEFPNLHHMEKLSIRTSTESDLSKLVKLIQNSP 1199
NAIP5  1096 HNLHKFLGLKELCVRLDGKPNVLSVLPREFPNLLHMEKLSIQTSTESDLSKLVKFIQNFP 1155
NAIP6  1096 HNLHKFLGLKELCVRLDGKPDVLSVLPEEFLNLHHMEKLSIRTSTESDLSKLVKFIQNFP 1155
            ***************** ** ****** ** ** *******.** ********* *** *
```

```
NAIP1  1156 NLHVFHLKCDFLSNCDSLMAVLASCKKLREIEFSGRCFEAMPFVNILPNFISLKILNLIS 1215
NAIP2  1200 NLHVFHLKCNFLSNCEPLMTVLASCKKLREIEFSGRCFEAMTFVNILPNFVFLKILNLRD 1259
NAIP5  1156 NLHVFHLKCDFLSNCESLMAVLASCKKLREIEFSGRCFEAMTFVNILPNFVSLKILNLKD 1215
NAIP6  1156 NLHVFHLKCDFLSNCESLMTALASCKKLREIEFSGQCFEAMTFVNILPNFVSLKILSLKG 1215
            ********* *****. **. **************.***** ********. **** *

NAIP1  1216 QQFPDKETSEKFAQALGSLRNLEELLVPTGDGIHQVAKLIVRQCLQLPCLRVLAFHYILD 1275
NAIP2  1260 QQFPDKETSEKFAQALGSLRNLEKLFVPTGDGIHQVAKLIVRQCLQLPCLRVLVFAETLD 1319
NAIP5  1216 QQFPDKETSEKFAQALGSLRNLEELLVPTGDGIHQVAKLIVRQCLQLPCLRVLTFHDILD 1275
NAIP6  1216 QQFADKETSEKFAQALGSLRNLEELLVPTGDGIHQVAKLIVRQCLQLPCLRVLAFHDILD 1275
            *** ****************** * ************************** *    **

NAIP1  1276 NDSVIEIARVATSGGFQKLEKLDLSMNHKITEEGYRNFFQALDNLPNLQNLNICRHIPEC 1335
NAIP2  1320 DDSVLEIAKGATRGGFQKLENLDLTLNHKITEEGYRNFFQVLDNLPNLKNLDISRHIPEC 1379
NAIP5  1276 DDSVIEIARAATSGGFQKLENLDISMNHKITEEGYRNFFQALDNLPNLQELNICRNIPGR 1335
NAIP6  1276 DESVIEIARAATSGSFQKLENLDISMNHKITEEGYRNFFQALDNLPNLQMLNICRNIPGR 1335
             .**.***. ** * ***** **...************** *******. * *.*.**

NAIP1  1336 IQVQATTVKALGQCVSRLPSLTRLHMLSWLLDEEDMKVINDVKERHPQSKRLIIFWKWIV 1395
NAIP2  1380 IQIQAITVKALGQCVSRLPSLTRLGMLSWLLDEEDIKVINDVKERHPQSKRLTVHWRWVV 1439
NAIP5  1336 IQVQATTVKALGQCVSRLPSLIRLHMLSWLLDEEDMKVINDVKERHPQSKRLIIFWKLIV 1395
NAIP6  1336 IQVQATTVKALGHCVSRLPSLTRLGMLSWLLDEEDMKVINDVKERHPQSKRLTIFWKWIV 1395
            **.** ******.******** ** **********.************** . *. .*

NAIP1  1396 PFSPVVLE 1403
NAIP2  1440 PFSPVIQK 1447
NAIP5  1396 PFSPVILE 1403
NAIP6  1396 PFSPVVLE 1403
            *****.
```

Appendix Fourteen. Multiple sequence alignment of inner-rod proteins. Each rod protein shown here was tested in the 293T NLRC4 oligomerization assay in Chapter Three. Rod name in left column, followed by bacterial genus and species in parentheses, followed by multiple sequence alignments. There is not a single absolutely conserved amino acid among these inner rod proteins (using this ClustalW alignment).

```
ROD        Species                        Alignment
PrgJ    (S.t.)    1 ------------------------MSIATIVPENAVIGQAVNIRSMETDIVSLDDRLL  34
BscI    (B.p.)    1 -----MNLDLTAINAVQERLLAGSFDMPRSPAMADQARFELALGEMPGASAPNGAIALAP  55
Citro   (C.r.)    1 -------------------MNQSEKINIQLNNTSQTNFVDEYTPLAPTPSAAGAAQFLD  40
BsaK    (B.p.)    1 --------------------------MNITNPHAVPALPSLSEIESPERPATLDAILK  32
A1902   (B.p.)    1 --------------MNTPLSASDLSRALDVAFSDAAAAPAGAGAPAAAPVPAEVAGRFQ  45
A33     (V.c)     1 -------------------------------MFD---LGIQGIDVTQPSVME  18
EprJ    {EHEC}    1 -------------------MSVSNMPPIDRAEQSTAHEIQQAKVIDLNDRVLNLDNPDD  40
EscI    (EPEC)    1 MDALCYCLSHEKRLTVNMNNINQSENINIQLNKASQTNVVDEHIPLASTPSAAGAAQFLD  60
MxiI    {S.f.}    1 --------------------------MNYIYPVNQVDIIKASDFQSQEISS-LEDVVS  31
PscI    (P.a.)    1 --------------------------MDISRMGAQAQITSLEELSGGPAGAAHVAEFE  32
SsaI    (S.t.)    1 ----------------------------------------MSVVPVSTQSYVKSSAE  17
Vp1368  {V.p.}    1 ---------------------------------------MFDNGILGLQGVDVTQPSVME  21
Vp1691  {V.p.}    1 --------------------------MINTQYTEVMQTNLEKLQDAQVEDGLSERFE  31
YPO0262 (Y.p.)    1 ---------------------------------------MNMSSAQAVAGLFQATE  17
YscI    (Y.p.)    1 ----------------------MPNIEIAQA-DEVIITTLEELGPAEPTTDQIMRFD  34


PrgJ    (S.t.)     56 VALDEPLGRRILGQLRG-------------GLADVAGKWRAVQTGLAEVSQAPTVVGMLD 102
BscI    (B.p.)     41 QLLPKTAGVSSPERVLIE----------EIKKKHLATMDTNRRFDALS-AGELSIEELRI  89
Citro   (C.r.)     33 QTLADAN------------------------EKSNAAKTSIESRLADPVDFAQSEKLIA  67
BsaK    (B.p.)     46 ALMSNAVPTPPLAHAQETSAIAKLVETSDAEIRKVLDNVEYLNLHANEMTMNQMFAASLQ 105
A1902   (B.p.)     19 KTYNS----------------------------VGESFEQILATIESIESSGGMSAVTALR  51
A33     (V.c)      41 KMISAFAN--------------------YAVQTENWQQNALQALTSDKKGLTPEKLLV  78
EprJ    {EHEC}     61 QLLPKTAGVSSPEQVLIE----------EIKKRHLATMNTDLSFDALS-AGGLSPEDVLT 109
EscI    (EPEC)     32 AKYSD------------------------IKMDTDIQVSQIMEMVSNPESLNPESLAK  65
MxiI    {S.f.}     33 RAMGG---AGSLGGDLLS----------ELGQIRERFSQAKQELQMELSTPGDDPNSLMQ  79
PscI    (P.a.)     18 PSQEQ--------------------------INFFEQLLKDEASTSNASALLPQVMLT  49
SsaI    (S.t.)     22 KTYNN------------------------VGESFEQILATIDSIG-DGMSAVDALR  52
Vp1368  {V.p.}     32 QAMSIPEGNNGLEGGLLE----------NISELKNTIDNAKSSLQDSMKMVGDDPAQLLQ  81
Vp1691  {V.p.}     18 PPVGN-----------------------TQQVAKFTQLMAQPASTELMMTPESLLM  50
YPO0262 (Y.p.)     35 AAMSED--TQGLGHSLLK----------EVSDIQKSFKTVKSDLHTKLAVSVDNPNDLML  82
YscI    (Y.p.)     35 QAFSG-----------------------SAIATAVDKQTITNRIEDPNLVTDPKELAI  69


PrgJ    (S.t.)     70  SQEMISDYNLYVSMVSTLTRKGVGAVETLLRS- 101
BscI    (B.p.)     103 LQARLLQASVEYELVGKAIGRATQNVDTLARMS 135
Citro   (C.r.)     90  LQKKVLNANVNVDVLSKFASLLSTSLTKLISQQ 122
BsaK    (B.p.)     68  LQTELSDYSIYVSLASTLARKAVSAVETLVKAQ 100
A1902   (B.p.)     106 ASAEATAMQIDMQAKMGVVTSTKDAIGSLMKNQ 138
A33     (V.c)      52  LTQDVFHYAIYQETVTKIAVKAANAINDVAKAQ  84
EprJ    {EHEC}     79  LQDHVLNYNVEVSLVGTLARKIVAAVETLTRS- 110
EscI    (EPEC)     110 LQKNVLNANVNVDVVSKLASLLSTSVTKLVSMQ 142
MxiI    {S.f.}     66  LQTTLSNYSIGVSLAGTLARKTVSAVETLLKS-  97
PscI    (P.a.)     80  MQWSLMRITMQEELIAKTVGRMSQNVETLMKTQ 112
SsaI    (S.t.)     50  RQMDYMQLTVGVDYLARISGAASQALNKLDNMA  82
Vp1368  {V.p.}     53  LQQEVFHYSIYQETVTKIASKAATAVNEVMKAQ  85
Vp1691  {V.p.}     82  MQWALTRITFQEELIAKTVGKTTQNVETLLKAQ 114
YPO0262 (Y.p.)     51  QQAEWMHASLAIDLTAKVAGVMGQNINKLVNMQ  83
YscI    (Y.p.)     83  MQWSLIRITIQEELIAKTAGRMSQNVETLSKGG 115
```