Title
Playing with Fire: How NAIP Inflammasomes Sense and Respond to Bacterial Pathogens

Author
Tenthorey, Jeannette

Publication Date
2017

Peer reviewed|Thesis/dissertation

Playing with Fire:

How NAIP Inflammasomes Sense and Respond to Bacterial Pathogens

by

Jeannette L. Tenthorey

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Russell E. Vance, Chair
Professor Gregory M. Barton
Professor Andreas Martin
Professor Brian J. Staskawicz

Spring 2017

Playing with Fire: How NAIP Inflammasomes Sense and Respond to Bacterial Pathogens

By Jeannette L. Tenthorey

Abstract

Playing with Fire:
How NAIP Inflammasomes Sense and Respond to Bacterial Pathogens

by

Jeannette L. Tenthorey

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor Russell E. Vance, Chair

The innate immune system is responsible for initiating the host immune response to infection. The study of microbial virulence has uncovered numerous mechanisms for microbes to evade innate immune detection. In contrast, relatively little is understood about the strategies employed by the host to prevent microbes from evading innate immunity. The NAIP innate immune receptors provide an intriguing case study to investigate these strategies. In mice, *Naip* has undergone gene duplication and drift, recombination, and pseudogenization[1], all of which can be signatures of a co-evolutionary arms race with targeted pathogens[2]. This duplication and specialization has allowed NAIP paralogs to recognize several distinct bacterial proteins: NAIP5 binds bacterial flagellin (FlaA), and NAIP2 detects the inner rod protein (PrgJ) of the pathogen-associated type III secretion system (T3SS)[3,4].

I sought to address how gene duplication and drift enabled functional specialization by first defining which NAIP domains bind to bacterial ligands. I analyzed a panel of chimeric proteins, in which homologous domains of NAIP5 and NAIP2 were swapped, to determine which domains conferred the ability to recognize FlaA or PrgJ. A long-standing expectation in the field was that the auto-inhibitory C-terminal leucine-rich repeat (LRR) domain mediates ligand binding. Surprisingly, I found that the LRR was dispensable for ligand specificity. Instead, ligand recognition was mediated by several alpha-helical domains in the center of the protein. Strikingly, these domains are specifically evolving under positive selection, in which non-synonymous mutations are repeatedly selected to provide altered ligand binding surfaces. Separation of sensing and auto-inhibition functions into different domains may allow NAIPs to sample ligand recognition-altering mutations without disrupting steady-state auto-inhibition.

These data suggested that NAIPs are engaged in a co-evolutionary arms race with bacteria over innate immune detection. However, bacterial ligands can evolve much more rapidly than mammalian NAIPs. To determine how NAIPs can successfully

compete in such an arms race, I conducted alanine scanning screens of FlaA and PrgJ to comprehensively identify the ligand motifs recognized by NAIPs. Both NAIP5 and NAIP2 recognized multiple conserved surfaces, near the N- and C-termini, of their respective ligands. This multi-surface recognition strategy conferred NAIPs with robust detection of their bacterial ligands, as single point mutations in any recognition motif did not affect NAIP recognition. Rather, bacterial immune evasion required simultaneous mutation of multiple recognition motifs. However, highly mutated ligands that escaped immune detection also lost their native function, suggesting that multi-surface recognition serves to constrain bacterial immune evasion.

To verify these biochemical results, we have determined the cryo-EM structure of NAIP5 bound to FlaA, an event which triggers oligomerization with the adapter protein NLRC4 into a large (>1 MDa) signaling complex. The structure reveals direct contacts between the NAIP5 ligand recognition domains and both recognition surfaces of FlaA. The extensive and largely hydrophobic contacts between NAIP5 and FlaA are consistent with a lack of “hot spot” binding sites and likely contribute to the robust recognition of NAIP5 for FlaA single point mutants. Additionally, our structure reveals how binding to FlaA triggers a conformational change in NAIP5 to expose its oligomerization surface, allowing the recruitment of NLRC4. The polymerization of NLRC4[5] and subsequent recruitment of the signaling effector, CASPASE-1, illustrates the switch-like mechanism by which the detection of a single ligand monomer is amplified into oligomerization-induced signaling.

Collectively, this dissertation elucidates the biochemical mechanism of NAIP innate immune detection of bacterial ligands. Furthermore, it has provided surprising insights into strategies employed by innate immune receptors to compete with bacteria in an evolutionary arms race over host defense.

To Pat and Paul,
for teaching me to question,
even if it meant also questioning them;

To Luc,
for staying close to my heart
despite the miles that separate us;

To Thornton and Jesse,
who lovingly taught me how to build a community;

To Harvey, Stef, Megan, Amy and Ashley,
my adopted family;

And to Andy,
my heart and joy,
whose support and laughter made this possible,

I cannot thank you enough.

Table of Contents

List of Figures and Tables

Chapter One

Chapter Two

Chapter Three

Chapter Four

Chapter Five

Acknowledgments

‘Give credit where credit is due’ is a fine sentiment until you realize just how much credit there is to dole out. I suspect this acknowledgments section should be just as long as the rest of the dissertation, but as you will soon see that it isn’t, please consider this an abbreviated version of the thanks I owe.

~

The most important lesson I am taking away from my graduate career is the recognition of how much my thinking and development has been influenced by the incredible people around me. I could not have asked for a better group of scientists and friends to guide me.

The first thanks must of course go to Russell Vance, the ringleader of this crazy circus. I cannot overstate how much Russell has shaped the way I think about science. Luckily for me, he had the courage to try turning a dyed-in-the-wool biochemist into a passable, if not card-carrying, immunologist. I am a better, more rigorous, and more creative scientist for his guidance. What I find most impressive is that he accomplished that guidance while simultaneously giving me free rein to pursue my own ideas, even when we disagreed about the likely outcome. I’ll be the first to admit that I rarely win bets against Russell, scientific or otherwise (though I’ll note that I’m 50/50 in head-to-head bocce matches). For letting me take those bets, and for all of his patient mentorship, I will be forever grateful.

Joining Russell’s lab comes with the invaluable advantage of co-mentorship from Greg Barton. I have benefited tremendously from his input, not only in the challenge to think critically and deeply about my own work, but also in the breadth of ideas and questions to which his lab’s work has exposed me. Just as important, Greg has always been honest and open about what a career in academic science entails, which is, I suspect, a rather rare opportunity for a graduate student.

There are a number of other mentors, both more and less official, whom I must also sincerely thank: Brian Staskawicz, for teaching me the value of drawing lessons and parallels from other disciplines; Andy Martin, for keeping my dissertation grounded in solid biochemistry; Sarah Stanley, for illustrating that it’s possible to be critical and still find nuggets of truth after picking everything apart; Heran Darwin, for demonstrating how to be a bad-ass woman in science and have a ton of fun doing it; and Dan Portnoy, for teaching me how to study immunology through the lens of bacterial genetics.

Mentors aren’t always in positions of authority, and I’ve learned just as much over the years from members of the Vance and Barton labs and the broader MCB community. For their scientific input, and for making ‘work’ seem like play, I thank Bella Rauch,

Meghan Koch, Patrick Mitchell, Aaron Whiteley, Dara Burdette, Justin De Leon, Randilea Nichols, Jordan Price, Kevin Barry, Joe Chavarria-Smith, Andrew Sandstrom, Livy Majer, Steve Wilson, Jon Portman, Thomas Burke, Eric Kofoed, Jakob von Moltke, Daisy Ji, Katie Deets, Shally Margolis, April Price, Aloe Stanbery, Lieselotte Kreuk, Jacques Deguine, and Bo Liu.

I also thank the members of campus and the broader Bay Area scientific community for sharing their work and insight at collaborative meetings like the Microbial Pathogenesis P01, BAMPS, immunology RIP, and departmental retreats; with them I got the chance to try on my immunologist, microbiologist, and cell biologist hats. In addition, there are a number of collaborators and mentees who contributed to the work described here; their contributions are acknowledged specifically in the following chapters.

Of course, it takes more than scientific colleagues to navigate graduate school. I am forever indebted to so many people who provided a supportive community and reminded me to play hard when I wasn't working hard. Thornton Thompson has been, as he well knows, just "the best." He and Jesse Niebaum formed the bedrock of our community, establishing WNDC as well as our fledgling Burning Man camp. I learned so much from both of them about openness and sharing. Bella Rauch has been both a role model and a friend, and you can certainly blame her for the crazy hair colors. My incredible classmates, who navigated this process with me and found adventures along the way, especially Randilea Nichols, Brett Robison, Rachel Hood, Caleb Cassidy-Amstutz, Avi Samelson, Jonathan Braverman, and Andrew Birnberg. A lucky break on Craigslist landed me not just a room but also an adopted family: Harvey Brockman, Ashley Penna, Stef Woodward, Amy Radon, and Megan Hipsley. They welcomed me into their communities, brought me exploring with them, encouraged me through the ups and the downs, and generally made life richer.

Finally, and most importantly, I thank my (non-adopted) family. Pat, Paul, and Luc have put up with an awful lot of my stubbornness over the years, and returned it with love and encouragement. Their support gave me the courage to start a career in science, and their continued support, even when I keep failing to move closer to home, gives me the strength to continue. Andrew Buchanan has been my cheerleader-in-chief for nearly a decade now. His unfailing and enthusiastic encouragement, his jokes (both great and terrible) that make me laugh no matter what my mood, and his willingness to share the crazy adventure of life with me have made all of this not only possible but fun. All I can say, simply, is thank you.

Chapter One: An Introduction to Innate Immune Sensing of Cytosolic Pathogens

The conflict for limited resources has driven the diversification of species and their competition strategies, as described by Charles Darwin[6], including the origin of numerous microbial parasites whose invasion and exploitation of their hosts often triggers a disease state. In response to these pathogens, hosts have evolved an arsenal of immune defense systems. These systems are categorized as either 'innate' defense, meaning a set of genes that do not change within an individual host, or 'adaptive' immunity, meaning those defenses that undergo genetic alteration within an individual and adapt in response to specific infections experienced by that individual. Innate immunity is found across all kingdoms of life, from prokaryotes (e.g., restriction enzymes that digest foreign DNA)[7] to plants and mammals (discussed below). Adaptive immunity was once considered a unique feature of vertebrates[8]. However, this view was challenged by the discovery of the adaptive prokaryotic CRISPR system, which targets digestion of foreign DNA based on previously encountered viruses[9]. The diversification of immune defense is thus likely to be as widespread as that of the pathogens it is meant to combat.

1.1 A brief overview of innate immunity in mammals and plants

In mammals, which possess both innate and adaptive immune defenses, the innate immune system constitutes the first line of defense against invading pathogens. Innate immunity is triggered within minutes of invasion[10]. In a few described examples[11-13], innate defenses are themselves sufficient to fully clear the pathogen and prevent colonization, a necessary precursor for infection. The efficacy of innate immunity in preventing disease may be the norm rather than the exception, given that most research focuses not on the spectrum of *potential* pathogens[14] but on clinically relevant, disease-inducing pathogens that have evolved to evade innate immunity to some degree. In these cases, the innate immune system performs a second critical function, which is to initiate the adaptive immune response[15]. This function is accomplished through the secretion of chemokines and cytokines, which recruit and activate specialized immune cells, and the upregulation of co-stimulatory molecules that license adaptive immune activation[16]. Broadly, this licensing step is thought of as a way to inform the adaptive immune system, which cannot distinguish self from non-self antigens, that a *bona fide* infection is in progress[17].

In order to discriminate self from non-self in a genomic space-efficient manner, the innate immune system deploys a limited number of receptors that recognize patterns common to broad classes of invading microbes[17]. Thus innate receptors are often referred to as pattern recognition receptors (PRRs). In its first iteration, this idea was taken to mean that innate immune receptors would directly bind to conserved components of microbial pathogens, termed PAMPs (pathogen-associated molecular

patterns). Indeed, numerous examples of PRR-recognized PAMPs have been described over the last several decades. Recognized PAMPs include, but are by no means limited to: lipopolysaccharide (LPS)[18] and peptidoglycan (PGN)[19,20], essential components of the cell wall in Gram-negative or Gram-negative and Gram-positive bacteria, respectively; double-stranded RNA[21], a feature of some viral genomes; and β-glucans[22], major cell wall components in fungi. It is noteworthy that innate immune sensing of helminths (unlike bacteria, viruses, fungi, and possibly protozoa) does not appear to rely on recognition of broadly conserved PAMPs[23].

More recently, it was proposed that innate immune recognition of foreign microbes might include not only their physical components but also their virulence-associated activities. When shared between numerous microbial pathogens, these virulence-promoting activities might be considered 'patterns of pathogenesis'[24]. One such activity is the hijacking of the host cytoskeletal protein actin. Attaching-and-effacing bacteria like enterohemorrhagic *Escherichia coli* (EHEC) recruit actin to promote bacterial attachment and colonization[25]. Several bacteria that replicate in the host cytosol, such as *Listeria monocytogenes*[26], polymerize actin tails to power motility and cell-to-cell spread[27]. Intriguingly, an innate immune receptor was recently reported to detect bacterial modification of Rho GTPases[28], which regulate actin dynamics, and the same receptor was found to colocalize with *L. monocytogenes* actin tails[29]. Another possible pattern of pathogenesis is the proteolytic degradation of host restriction factors, and this protease activity can be sensed by the innate immune receptor NLRP1B[30-32].

Extending this principle more broadly, there is some evidence that the innate immune system might respond to a general state of cellular stress. For example, NOD1 and NOD2, PRRs described as sensors of bacterial PGN[33,34], were recently reported to mediate inflammation following induction of the endoplasmic reticulum (ER) stress response[35]. Numerous intracellular pathogens induce ER stress[36-38], although the ER stress response is not exclusively associated with infection. The innate immune receptor NLRP3 has also been linked to the ER stress response[39,40], as well as ionic imbalance following the disruption of host cell membranes by pore-forming toxins[41] and even osmotic stress[42]. There is also a suggestion that necrotic cell death may itself be immune-stimulatory[43,44], although this hypothesis is debated[45]. ER stress, ion fluxes, and necrosis can certainly be triggered during microbial pathogenesis, but they can also occur in the absence of infection. Thus, it is unclear whether and why sensing of these stresses *per se*, in the absence of other 'non-self' cues, should result in inflammation.

The detection of cellular disruption is reminiscent of 'effector-triggered immunity' in plants, in which the immune system is activated by the activity of pathogen effector molecules[41]. For example, some receptors have been hypothesized to act as 'guards' to monitor the integrity of other host defense mediators[46,47]. When pathogens manipulate the latter set of defenses in an effort to evade plant immunity, the guard receptors are activated to initiate alternate immune defenses. As a corollary strategy, plant receptors may incorporate a mimic of the pathogen-targeted defense protein, thereby serving as a

'decoy' for pathogen manipulation[48]. The decoy strategy has the advantage of separating the selective pressure to be targeted by pathogens (immune detection function) from the selective pressure to avoid pathogen manipulation of the original host response. Several examples of plant guard[49-52] and decoy[53,54] receptors have been described. Mammalian immune sensing of cytoskeletal manipulation or stress induction bears striking resemblance to the plant guard hypothesis, though the cytoskeleton and ER are not targeted by pathogens due to their role in immunity, and the protease-targeted domain of NLRP1B has been described as a decoy[55]. In addition, some plant immune receptors target conserved PAMPs, such as the bacterial motility-mediating protein flagellin[56,57]. Thus, there are remarkable parallels between plant and mammalian innate immunity. A notable difference is that plants seem to encode greater numbers of PRRs than mammals, possibly reflecting the lack of a plant adaptive immune system[58].

1.2 Cytosolic pattern recognition receptors

A criticism of the PAMP hypothesis is that conserved microbial features, like bacterial or fungal cell wall components, are often shared between pathogens and non-pathogens and thus do not allow the innate immune system to discriminate pathogens from non-invasive or commensal species[59]. The separation of innate immune sensing into extracellular and intracellular spaces confers the ability to distinguish cell-invasive pathogens from non-pathogens, which remain outside the cell[24]. It should be noted that phagosomes, which take up extracellular material including commensal microbes[60], are contiguous with the extracellular space. Phagosomes subsequently fuse with lysosomes to chemically and enzymatically degrade ingested material[61]. Many diverse pathogens access the host cytosol with the general aim of evading digestion, acquiring host nutrients, and avoiding extracellular immune defenses. Cytosolic access can include full escape from the endocytic/phagosomal compartment (e.g., viruses, bacteria like *Listeria*), limited permeabilization of the phagosomal membrane (e.g., *Legionella pneumophila*), or access from outside the cell via specialized secretion systems (e.g., EHEC) or pore-forming toxins (e.g., *Streptococcus pneumoniae*) (Figure 1.1).

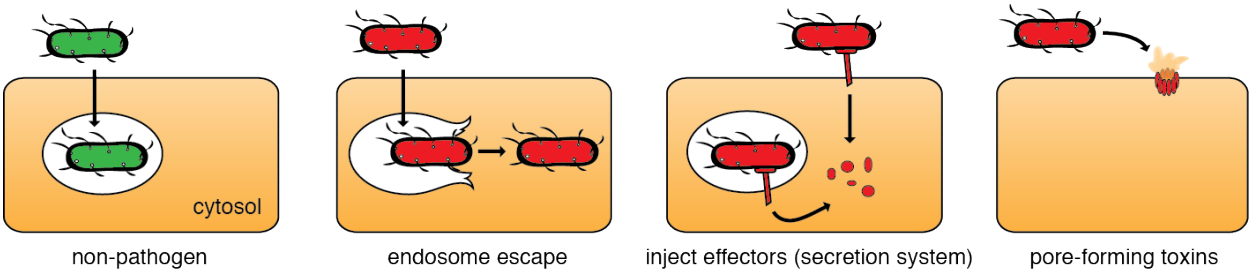


Figure 1.1. Pathogens can access the host cell cytosol

Phagocytosed non-pathogens are degraded within the phagolysosome, whereas pathogens can rupture the phagosome or prevent lysosomal fusion via effector proteins injected into the cytosol with specialized secretion systems. Injection of effectors can also prevent phagocytic uptake. Pore-forming toxins allow the release of cytosolic contents. In each case, the barrier separating the cytosol (orange) from the extracellular space is disrupted.

The first described PRRs were cell-surface receptors of PAMPs, but analogous intracellular PRRs that detect the same or at least similar PAMPs were discovered shortly thereafter (Table 1.1). The detection of the same PAMP by PRRs in two discrete locations often, though not universally, results in the induction of qualitatively different responses. For example, sensing of phagolysosomal DNA by TLR9 results in the NFκB-dependent transcription of numerous cytokines[62], or in some cell types the induction of type I interferons (IFN-I)[63]. In contrast, the detection of intracellular DNA by AIM2 results in a more rapid and severe response, including post-translational maturation of a limited set of cytokines and the induction of cell death[64,65], while the intracellular DNA sensor cGAS induces an IFN-I transcriptional response along with autophagy[66,67]. In principle, then, spatial cues can allow PRRs to tailor inflammatory responses to the degree of threat posed by encountered microbes.

Table 1.1. The same PAMPs can be detected extracellularly and intracellularly to induce distinct responses

PAMP	Extracellular PRR	Signaling outcome	Cytosolic PRR	Signaling outcome
DNA	TLR9[62] (endosome)	Transcription (NFκB targets[62], IFN-I[63])	AIM2[64,65]	Cell death, IL-1β maturation[57,58]
			cGAS[68]	Transcription (IFN-I[68]), autophagy[60]
dsRNA	TLR3[69] (endosome)	Transcription (NFκB targets, IFN-I)[69]	RIG-I[70], MDA5[21]	Transcription (IFN-I)[21]
LPS	TLR4[18] (cell surface)	Transcription (NFκB targets, IFN-I)[71]	CASPASE-11[72]	Cell death[73]
flagellin	TLR5[74] (cell surface)	Transcription (NFκB targets)[74]	NAIP5[75]	Cell death, IL-1β maturation[76]

At least some PAMPs that are sensed in the intracellular compartment are likely to have activated extracellular PRRs on their transit to the host cell cytosol. For example, bacterial pathogens that escape from the phagosome must first be taken up from the extracellular space within a phagosome, and their surface molecules should in principle be available for detection in both compartments. Thus, extracellular and intracellular PRRs are likely to act in concert to tailor immune responses. An excellent example of this synergy is the 'two-signal' model of activation[77] for inflammasomes such as NAIP5. Inflammasomes are intracellular PRRs that induce cell death as well the post-translational maturation of the cytokine IL-1β[78], but this cytokine requires prior transcriptional induction by cell-surface PRRs[77]. A single PAMP, detected by two PRRs, can thus induce three distinct responses: sensing of flagellin at the cell surface by TLR5 induces transcription of IL-1β and other cytokines, NAIP5 detection of cytosolic flagellin triggers cell death, and together TLR5 and NAIP5 can generate the mature form of the highly inflammatory[79] cytokine IL-1β.

1.3 Inflammasomes

Inflammasomes are particularly intriguing PRRs for a number of properties explored above: they are cytosolic PRRs that selectively respond to pathogens or patterns associated with pathogens to trigger non-transcriptional immune responses[80]. Inflammasomes were first described 15 years ago as cytosolic proteins that, upon stimulation, formed large oligomeric complexes to activate the protease CASPASE-1 (CASP1)[81]. CASP1 is activated by dimerization[82], mediated by this oligomeric complex, to cleave the immature forms of IL-1β (pro-IL-1β) and IL-18 (pro-IL-18) into signaling-competent cytokines[79]. Secreted IL-1β recruits neutrophils, a particularly phagocytic and microbicidal cell type, to the site of infection[83] and induces numerous inflammatory mediators[79]. IL-18 induces the expression of IFNγ[79], a cytokine that activates microbicidal effector functions in phagocytic cells and enhances the activity of natural killer (NK) cells[84]. Both cytokines also influence the development of the subsequent adaptive immune response[80].

CASP1 activation also results in a rapid, lytic form of cell death that is highly inflammatory, hence its name 'pyroptosis' (fiery cell death)[85]. Strikingly, pyroptosis occurs within minutes of detectable CASP1 activation[86]. Though pyroptosis is considered a characteristic output of inflammasome activation, neutrophils have been reported not to induce pyroptosis following CASP1 stimulation[87]. This finding could be explained by the lack of expression of GASDERMIN-D (GSDMD) in neutrophils, as GSDMD is required for pyroptosis[88,89]. Cleavage of GSDMD by CASP1 triggers GSDMD to form pores in the plasma membrane[90,91]. These pores then release cytosolic contents, including mature IL-1β and IL-18, into the extracellular milieu.

Paradoxically, the induction of cell death can be protective for the host[92,93]. Infected cells that activate the inflammasome have already failed to restrict the pathogen, as indicated by its invasion of the cytosolic compartment. By undergoing programmed suicide, the infected cell removes the replicative niche that is being exploited by the pathogen. Pyroptosis also exposes the pathogen to phagocytic cells, which have been recruited to the site of infection and activated by IL-1β and IL-18. Furthermore, exposed pathogens seem to be immobilized in the pyroptotic cell debris, preventing them from escaping recruited phagocytes[94]. Notably, CASP1 is not unique in its induction of pyroptosis, as the same form of cell death is observed upon activation of CASPASE-11[73]. Furthermore, in the absence of CASP1, at least some inflammasomes can activate an alternate form of cell death through CASPASE-8[95,96]. These data suggest that ablation of the intracellular replicative niche is an important immune defense.

There are several distinct inflammasomes that converge upon CASP1 activation but that respond to different infection-associated stimuli. As noted above, pyrin monitors the integrity of Rho GTPases, which regulate actin dynamics. Pyrin is activated when bacteria chemically modify Rho GTPases[28] to manipulate host actin, potentially allowing cells to sense this bacterial activity[97]. NLRP1 isoforms contain a relatively unstructured domain that serves as a decoy for pathogenic proteases, and proteolytic cleavage of

this domain activates NLRP1 inflammasomes[30,31,98]. Proteases are deployed by pathogens to interfere with immune defenses. For example, the *Bacillus anthracis* lethal factor protease degrades MAPK kinases to disrupt immune signaling cascades, and this activity is detected by NLRP1b[32,99]. The NLRP3 inflammasome is activated by an impressive variety of noxious stimuli, including pore-forming toxins[41], disruption of phagosomal membranes[100], disruption of mitochondrial membranes[101], and ER stress[40]. The mechanism by which NLRP3 responds to these diverse stimuli is unclear, although it has been suggested that at least some of these mechanisms converge on ionic imbalances like K^+ efflux[41].

In general, the pyrin, NLRP1, and NLRP3 inflammasomes seem to respond to 'patterns of pathogenesis'. By contrast, several inflammasomes detect specific microbial PAMPs in the host cytosol. AIM2 directly binds to DNA in the cytosol[102]. Given that host DNA is typically confined to the nucleus, cytosolic DNA is an indicator of viral infection[103] or the secretion or lytic release of DNA from cytosolic bacteria[104,105]. NLRP7, present in humans but not rodents, has been reported to detect bacterial lipoproteins in the cytosol[106]. NAIP inflammasomes, discussed further below, bind to conserved bacterial proteins like flagellin and structural components of the type III secretion system (T3SS), a virulence factor common to many Gram-negative bacterial pathogens[3,4].

Another interesting distinction between inflammasomes is the requirement for some NLRs to undergo transcriptional priming. Whereas NAIP, NLRP1, and pyrin inflammasomes are constitutively expressed and can trigger pyroptosis immediately after detecting their respective stimuli[107], NLRP3 and AIM2 are not normally expressed in most cell types. NLRP3 transcription requires NFκB-dependent signaling downstream of TLRs or inflammatory cytokines like IL-1β, whereas AIM2 is induced in response to IFN-I signaling[80]. The non-canonical inflammasome CASPASE-11 also requires transcriptional priming by either IFN-I or IFN-γ[108]. The requirement for inflammation-based priming of some inflammasomes may reflect the need to more strictly constrain their activation to *bona fide* infectious settings. Indeed, DNA and ionic fluxes are not strict indicators of the presence of non-self infections. This logic cannot explain the priming requirement of CASPASE-11, which responds to specifically to bacterial LPS.

In addition, inflammasomes may differ in their cell-type specificity. Most inflammasomes have been studied in macrophages, with less study of their role in other cell types. However, NAIP inflammasomes have been shown to be functional in neutrophils[87] and epithelial cells[109]. NLRP1 is highly expressed and functional in the skin[110]. The expression pattern of other NLRs also varies significantly by cell type. For example, pyrin is highly expressed in monocytes and conventional dendritic cells, whereas NLRP3 is highest in epidermal/Langerhans dendritic cells (ImmGen). It is not yet clear whether varying expression patterns affect inflammasome function in these different contexts.

Most inflammasomes share a core domain architecture consisting of an AAA+ family nucleotide-binding domain (NBD) and a leucine-rich repeat domain (LRR), hence their classification as NLR (NBD- and LRR-containing) proteins. Because AAA+ domains are common platforms for oligomerization[111], the NBD has long been assumed to mediate NLR inflammasome assembly, and this prediction has been borne out by recent structural work[5,112]. The LRR is required for auto-inhibition in the absence of an activating stimulus, as removal of this domain results in constitutive signaling in a number of NLRs[80]. Although not all inflammasomes contain an NLR core, they share one of two possible signaling domains capable of activating CASP1. Inflammasomes that contain a CARD (caspase activation and recruitment domain) are capable of directly interacting with CASP1[113], whereas those that contain a PYD (pyrin domain) indirectly recruit CASP1 through the PYD-CARD adaptor protein ASC[80]. An exception to this rule is the NAIP inflammasomes, which lack a CARD or PYD and require association with the CARD-containing NLRC4 to activate CASP1[3].

Strikingly, plant R (resistance) proteins that mediate cytosolic effector-triggered immunity share the NBD-LRR architecture, and they also induce rapid, local cell death upon pathogen invasion of the cytosol[47]. This 'hypersensitive response' isolates the pathogen and prevents its spread to the rest of the plant. Despite these similarities, plant and animal NLRs appear to have evolved independently[114]. Furthermore, R protein signaling domains and thus signaling cascades leading to cell death differ from those of animal NLRs[55]. Filamentous fungi also encode defense proteins that induce cell death in response to fusion and cytoplasmic mixing with genetically distinct, or non-self, strains[115]. The 'heterokaryon incompatibility' of these hybrid fusion cells generates a barrier of dead cells that limits invasion of the competing strain. Proteins that mediate heterokaryon incompatibility contain an AAA+ NBD but lack an LRR[116], and their signaling domains are also distinct from those of plants or animals[117]. Thus, while plants and fungi do not encode true inflammasomes, they both use AAA+ NBD proteins to mediate similar defense mechanisms in response to invasion by pathogens or non-self competitors.

1.4 NAIP inflammasomes

The NAIP inflammasomes were first discovered by mapping the genomic locus responsible for restricting the intracellular growth of the bacterium *L. pneumophila*[1]. In mice, the locus contains several tandemly arrayed paralogs, of which *Naip5* is required for *L. pneumophila* restriction[118]. Restriction also requires the bacterial protein flagellin[75,119], suggesting that NAIP5 is a cytosolic PRR for flagellin. Confusingly, another host NLR, NLRC4, was independently determined to be required for the response to cytosolic flagellin[120,121]. Adding further confusion, NLRC4 activation was also observed upon cytosolic introduction of an unrelated protein, the inner rod protein of the bacterial T3SS[122]. The confusion was resolved when it was determined that different NAIP paralogs directly bound to distinct bacterial ligands, triggering NAIP association with NLRC4[3,4]. NAIP proteins lack a CARD or PYD and instead contain three baculovirus inhibitor of apoptosis repeats (BIRs). The NAIP BIRs appear to be

required for NAIP function or at least stability[3], but they cannot mediate CASP1 recruitment. By contrast, NLRC4 contains a CARD but does not play a direct role in ligand sensing (Chapter 4). These findings explain the dual requirement for NAIPs and NLRC4 in cytosolic detection of bacterial proteins.

The mouse *Naip* locus has undergone several rounds of gene duplication and drift[1], pseudogenization of some *Naip* copies, and extensive recombination between tandemly arrayed paralogs[123]. This genomic volatility has allowed the functional specialization of NAIP proteins, despite their high sequence identity (≥ 80%). NAIP5 and NAIP6 are 94% identical at the amino acid level, and both bind to flagellin[3,4]. In particular, the flagellin D0 domain is both necessary and sufficient for NAIP5/6 detection[118,124]. In most contexts NAIP5 is the dominant sensor for flagellin[124], possibly reflecting a lower expression level or a weaker flagellin-binding affinity of NAIP6. NAIP1 and NAIP2 detect the needle[125,126] and inner rod proteins[3,4], respectively, of the T3SS. The ligand specificity of NAIPs was initially determined biochemically and has since been genetically confirmed[127,128].

Unlike rodents, primates encode only one full-length copy of NAIP. The human *NAIP* locus has undergone duplication, but duplicated *NAIPs* code for truncated proteins lacking some or all of the BIRs[129]. Human populations also exhibit distinct copy number variation with possible functional consequences[130]. Strikingly, the single human NAIP protein responds to flagellin and the T3SS rod and needle proteins ([131] and S. Shin, unpublished data). These data suggest that recognition of all three ligands is the ancestral phenotype of NAIP surveillance. Duplication and specialization of NAIP proteins in mice may have conferred the ability to evolve higher affinity or more flexible recognition of each independent ligand, although this idea has not been directly tested.

The three bacterial ligands recognized by NAIPs are unrelated but share several important characteristics. Each ligand homotypically polymerizes into a hollow tube, either the flagellum or the central channel of the T3SS, and monomers are secreted through the tube to assemble at the distal tip[132]. Given that NAIPs detect monomeric ligands[119], it is likely that some monomers fail to assemble, or are incorrectly secreted after capping of the distal tip, and are instead shed into the host cell cytosol. Flagellin can also be aberrantly secreted into the cytosol via the T3SS[133]. These mechanisms of delivery suggest that ligands are at low concentrations in the host cytosol, implying that NAIPs are highly sensitive to the translocation of just a few ligand monomers.

As monomeric proteins, both the flagellin D0 domain[134,135] and the T3SS rod protein[136] are largely unstructured. The T3SS needle protein is not unstructured as a monomer[137], but it is unfolded during secretion through the assembling needle apparatus[138]. Upon oligomerization both flagellin and the T3SS needle adopt a similar structure of two antiparallel alpha helices[139,140], and the T3SS rod protein is predicted to share this fold[122]. This shape may therefore be an important determinant for NAIP recognition of

different bacterial ligands. In addition, NAIP ligands share several hydrophobic amino acids in the C-terminal 5-10 amino acids that are critical for NAIP recognition[118,122,126].

1.5 Bacterial evasion of NAIP–NLRC4 inflammasomes

NAIP inflammasomes are sufficient to restrict "accidental" pathogens like *L. pneumophila*[118] or *Chromobacter violaceum*[12], environmental isolates that do not natively infect mammalian hosts. This restriction potential places a selective pressure on co-evolving pathogens to evade NAIP surveillance. Indeed, several instances of NAIP evasion have been reported. Virulent strains of *Listeria* repress the expression of flagellin at 37 ºC[141]. This repression is critical for escape from NAIP5 restriction, and forced expression of flagellin severely attenuates *L. monocytogenes*[93,142]. Flagellin repression is not without cost, as motility is critical for cell-to-cell spread and virulence[143]. Remarkably, *L. monocytogenes* has acquired the ability to co-opt host actin and polymerizes actin from one pole to power intracellular motility[26]. Numerous intracellular bacterial pathogens have acquired host actin-based motility through several distinct mechanisms[144], suggesting that it is a useful tool to avoid the NAIP immune barrier.

Salmonella typhimurium also represses flagellin expression once it invades host cells[121], and ectopic flagellin expression results in inflammasome-dependent clearance[92]. In addition, *S. typhimurium* has horizontally acquired two different T3SS, SPI1 and SPI2. While the SPI1 inner rod and needle proteins activate NAIP2 and NAIP1, respectively, the inner rod protein of SPI2 is highly divergent and is not recognized by NAIP2[122]. It has not yet been tested whether the SPI2 needle protein (SsaG) is recognized by NAIP1, although there is some evidence of SPI2-dependent inflammasome activation[145]. Flagellin repression is coordinated with the upregulation of SPI2 inside acidifying phagosomes[146]. The switch to SPI2 expression appears to allow systemic spread of *S. typhimurium*, as artificial expression of the SPI1 rod protein under a SPI2 promoter prevented this spread[122]. Interestingly, flagellin and SPI1 repression are less pronounced inside epithelial cells, the first cell type encountered by ingested *Salmonella*, than in macrophages underlying the intestinal epithelial layer[147]. This may reflect an orthogonal selective pressure on the bacterium to induce local inflammation in the gut, which generates a metabolic advantage for *S. typhimurium* to outcompete the intestinal microbiota in the lumen[148]. Nevertheless, the activation of NAIP inflammasomes in the intestinal epithelium limits bacterial colonization of host tissues[109].

Pathogen evasion of NAIP surveillance is not exclusively through repressed expression of NAIP ligands. For example, *Yersinia* species secrete several T3SS effectors that interfere with inflammasome recognition or signaling. YopK physically associates with the T3SS and prevents the activation of NAIP and NLRP3 inflammasomes, possibly through masking ligands or altering T3SS secretion[149]. YopM inhibits CASP1, though there is disagreement as to whether it serves as a competitive CASP1 substrate[150] or functions via other mechanisms[151,152]. NAIP inflammasomes can activate CASPASE-8 in the absence of CASP1[153], suggesting that CASP1 inhibition may not be a successful strategy for bacterial agonists of NAIP inflammasomes. Regardless, the myriad

strategies deployed to prevent NAIP activation indicate that inflammasome restriction is an important selective pressure driving the evolution of bacterial pathogens.

1.6 NAIP inflammasomes and disease

The output of inflammasome activation is highly inflammatory and potentially dangerous. The danger is particularly pronounced for NAIP inflammasomes, which are constitutively expressed and poised to signal in many cell types. Systemic activation of the NAIP inflammasome in mice results in massive inflammation and ultimately death in less than an hour[107], and activation in select cell types also causes rapid pathology[153]. In human patients, gain-of-function point mutations in *NLRC4* are dominantly linked to a severe and sometimes lethal auto-inflammatory phenotype[154-158]. Strikingly, *in vitro* studies indicate that these mutations only mildly disrupt NLRC4 autoinhibition, suggesting that this disruption is amplified across many NLRC4-expressing cells to cause severe disease. The identified mutations all lie adjacent to the nucleotide-binding pocket of the NBD. Thus, these residues may contribute to coordinating ADP between several NLRC4 domains in order to maintain a closed, non-signaling conformation[159]. As yet there have been no identified auto-active alleles of *NAIP* in humans, but the above data hint that very few constitutively active *NAIP* or *NLRC4* genotypes are likely to be viable.

Recently several reports have linked NAIPs and/or NLRC4 to tumor progression. This connection is surprising, given the well-described role of NAIP inflammasomes in specific detection of bacterial pathogens. Furthermore, the findings of these studies are often directly conflicting. Most studies suggest that NAIP or NLRC4 are protective against tumor growth[160-162], although one report found that NLRC4 promotes obesity-associated breast cancer[163]. Surprisingly, one study found that the tumor-protective effect of NAIPs was independent of NLRC4 and CASP1[160], possibly implicating a role for the BIR domains in inhibiting tumor cell apoptosis. In another case, NLRC4-dependent protection against tumor growth was independent of CASP1[162]. Given that NAIPs are expressed in epithelial cells and are sensors for bacterial proteins, it is possible that altered composition of the microbiota may explain these conflicting results. Consistent with this hypothesis, most studies linking NAIPs or NLRC4 to cancer have induced colitis[160,161] or altered diet[163], plausibly leading to the outgrowth or translocation of potential pathobionts. Nevertheless, the idea that inflammasome-driven inflammation can alter tumor growth is certainly not without precedent[164].

1.7 Dissertation overview

Bacterial and host genetics have defined many of the molecular players in the detection of cytosolic pathogens. However, biochemical characterization of these interactions has been largely lacking. This is particularly true of the inflammasome family, whose members are large multi-domain proteins that are prone to overexpression-induced aggregation[165] and typically do not tolerate truncations. This dissertation aims to uncover the molecular workings of the NAIP inflammasomes, from ligand recognition to inflammasome assembly and regulation, with a view to understanding how these innate

immune receptors respond robustly to bacterial ligands while avoiding responses to self proteins.

In Chapter 2, I define the NAIP domains that mediate ligand specificity and, presumably, ligand binding. These domains are evolving under a positive selection regime, consistent with their direct interaction with rapidly evolving bacterial ligands. In addition, I demonstrate that each NAIP protomer must bind to its specific ligand to trigger inflammasome assembly, suggesting that NAIPs are highly specific to their cognate stimuli and cannot be activated as 'bystander' inflammatory mediators.

In Chapter 3, I explore mechanisms that NAIPs utilize to counteract rapid bacterial evolution of escape mutations. Specifically, NAIPs employ a 'multi-surface' recognition strategy that makes NAIP recognition robust to point mutations in bacterial ligands. To escape NAIP surveillance, ligands must acquire multiple or non-conservative mutations that also disrupt their function, thus providing a selective pressure against escape variants.

In Chapter 4, we determine the structure of the assembled NAIP5 inflammasome. The structure confirms and extends the findings of Chapters 2 and 3, defining a binding interface between several domains of NAIP5 and multiple surfaces of the flagellin D0 domain. This extensive interface likely contributes to the robustness of NAIP5 recognition of flagellin. From the structural data, we propose a model for NAIP5 activation in which flagellin binding sterically induces the rigid-body rotation of NAIP5 domains to expose the NBD oligomerization surface for recruitment of NLRC4.

In Chapter 5, I introduce tools to probe the regulation of NAIP inflammasome signaling. Using CRISPR targeting in mice, I generate several knock-in alleles of *Nlrc4* to test the role of NLRC4 phosphorylation or other potential regulators in the sensing of cytosolic bacterial pathogens. I also modify a method of cytosolic ligand delivery to allow for purification of endogenous NAIP inflammasomes with associated cofactors from mouse macrophages.

Finally, in Chapter 6, I review the findings of this dissertation. I discuss our insights into how NAIP PRRs maintain robust recognition of rapidly evolving bacterial pathogens without compromising the health of the host through aberrant self-recognition. I conclude with potential future directions and what, in general, biochemistry can contribute to the study of immunology.

Chapter Two: Molecular Basis for Specific Recognition of Bacterial Ligands by NAIP–NLRC4 Inflammasomes

2.1 Abstract

NLR (nucleotide-binding domain [NBD]- and leucine-rich repeat [LRR]-containing) proteins mediate innate immune sensing of pathogens in mammals and plants. How NLRs detect their cognate stimuli remains poorly understood. Here, we analyzed ligand recognition by NAIP (NLR Apoptosis Inhibitory Protein) inflammasomes. Mice express multiple highly related NAIP paralogs that recognize distinct bacterial proteins. We analyzed a panel of 43 chimeric NAIPs, allowing us to map the NAIP domain responsible for specific ligand detection. Surprisingly, ligand specificity was mediated not by the LRR domain, but by an internal region encompassing several NBD-associated α-helical domains. Interestingly, we find that the ligand specificity domain has evolved under positive selection in both rodents and primates. We further show that ligand binding is required for the subsequent co-oligomerization of NAIPs with the downstream signaling adaptor NLRC4 (NLR family, CARD-containing 4). These data provide a molecular basis for how NLRs detect ligands and assemble into inflammasomes.

2.2 Highlights

- The LRR domain of NAIPs is not utilized for discrimination of bacterial ligands.
- The NBD-associated helical domains of NAIPs dictate specific ligand recognition.
- The ligand specificity domain has evolved under positive selection.
- Ligand binding is required for NAIP assembly into inflammasomes.

2.3 Introduction

Many nucleotide-binding domain (NBD)- and leucine-rich repeat (LRR)-containing proteins (NLRs) function as innate immune sensors that monitor the cytosol for the presence of microbial products and other infection-associated stimuli[47,71,80,166]. Once activated, some NLRs assemble into high molecular weight complexes termed inflammasomes[81,166] that recruit and activate pro-inflammatory proteases such as CASPASE-1 (CASP1). CASP1 cleaves the pro-inflammatory cytokines IL-1β and IL-18 into their signaling-competent forms. CASP1 also initiates pyroptosis[85], a rapid, lytic form of cell death that releases pro-inflammatory molecules to trigger rapid and potent immune responses[107,167].

The NBD-LRR architecture is found in pathogen-sensing proteins in both mammals and plants[46,168], but remarkably little is known about how NLRs detect infectious stimuli and initiate signaling. The NBD of NLRs is classified as an AAA+ ATPase[169], a domain found in diverse proteins known to form homo- and hetero-oligomeric complexes[170]. The NBD is presumed to mediate assembly of NLR protomers into the active oligomerized inflammasome, analogous to the function of the NBD in assembly of the apoptosome[171].

The other domain that defines membership in the NLR superfamily, the LRR domain, is believed to have two distinct roles. The first is to function as an autoinhibitory domain, as truncation of this domain generally results in constitutive NLR activation[3,98,172,173]. The autoinhibitory function of the LRR domain is supported by the recently determined crystal structure of the monomeric/inactive form of NLRC4, in which the LRR domain curves back to occlude the NBD *in cis*[159]. In addition to its role in autoinhibition, the LRR domain has also been proposed to act as a 'sensor' that directly or indirectly detects ligands[170]. The ligand-binding function of the LRR domain is supported primarily by analogy to the well-established ligand-binding function of the LRRs in pathogen-sensing Toll-like receptors (TLRs) [174]. Association of ligands with the LRR is believed to disrupt autoinhibitory *cis* interactions between the LRR and the NBD, resulting in NBD-mediated oligomerization and inflammasome assembly[159,170,175]. However, direct evidence for ligand association with the LRR domain, or indeed any other domain, of mammalian NLRs is lacking.

In order to address the fundamental issue of how NLRs detect their specific ligands, we analyzed the ligand specificity of NAIP/NLRC4 inflammasomes. Mice express multiple NAIP paralogs, each of which recognizes a distinct bacterial ligand. Both NAIP5 and NAIP6 detect bacterial flagellin, whereas NAIP2 detects inner rod proteins of type III secretion systems[3,4]. Mouse NAIP1 and human NAIP respond to needle proteins of type III secretion systems[4,125,126]. Upon recognition of their ligands, NAIPs assemble with NLRC4 into an oligomerized inflammasome that contains both NLRs and the ligand[3]. The assembled inflammasome can then directly recruit and activate CASP1 via the NLRC4 CARD (caspase activation and recruitment domain) [80].

At present, the molecular basis for ligand recognition by NAIP/NLRC4 inflammasomes, or indeed by any mammalian NLR, remains unclear. It has not yet been possible to map the ligand recognition domain of mammalian NLRs by mutagenesis because mutations that disrupt NLR function may not specifically affect ligand binding but may instead disrupt the overall NLR fold or oligomerization competence[173]. We circumvented this difficulty by taking advantage of the fact that, although they recognize distinct bacterial ligands, the mouse NAIP paralogs share a high degree of amino acid identity and the same basic architecture (Figure 2.1A). Reasoning that chimeric NAIP proteins might retain their overall fold and function, we generated and analyzed a panel of 43 NAIP chimeras to identify the ligand recognition domain. To our surprise, we found that ligand specificity of NAIPs does not require the LRR domain, but instead depends on several largely alpha helical domains associated with the NBD. Furthermore, we observed an evolutionary signature of positive (diversifying) selection that maps to the ligand specificity domain, consistent with the direct association of these domains with rapidly evolving bacterial ligands[2,176]. We further demonstrate that ligand binding is required for NAIPs to co-assemble with NLRC4. These results suggest a new model for innate immune sensing by the NLR superfamily, and provide evidence for an ongoing molecular arms race between NAIPs and the pathogens they sense.

2.4 Results

In order to map the region(s) of NAIPs required for specific recognition of ligands, we generated chimeras between murine NAIP paralogs that respond to distinct ligands. We swapped homologous sequences between NAIP2 and either NAIP5 or NAIP6 at seven breakpoints (a-g) along the length of the protein (Figure 2.1A). These breakpoints were chosen because they are located within stretches of high sequence identity among the NAIP paralogs. We then assayed the ability of chimeric NAIPs to induce NLRC4 inflammasome assembly in response to *Legionella pneumophila* flagellin (FlaA) or *Salmonella typhimurium* type III secretion system inner rod protein (PrgJ). Assembly of the high molecular weight (~1MDa) inflammasome was monitored by native gel electrophoresis using an established reconstituted inflammasome assay[3]. Because this is the most sensitive assay for NAIP function thus far described, we reasoned it might allow us to detect NAIP regions that only weakly contribute to ligand recognition.

As previously observed[3], transient transfection of HEK293T cells with NLRC4 and NAIP2 yields an oligomerized inflammasome only in response to PrgJ, while NAIP5 and NAIP6 induce NLRC4 oligomerization only in response to FlaA (Figure 2.1 and 2.2). Thus, the parental NAIPs exhibit a high degree of selectivity for their ligands.

2.4.1 BIRs, NBD and LRR of NAIPs do not affect ligand specificity

We first analyzed a series of chimeras in which the N-terminal domains of NAIP5 or NAIP6 were fused to the C-terminal domains of NAIP2 (termed NAIP5.2 or NAIP6.2 chimeras; Figure 2.1B-E, 2.2). The NAIP5.2(a) chimera, containing the NAIP5 BIR (baculovirus inhibitor of apoptosis repeat) domains, still oligomerized normally in response to the NAIP2 ligand PrgJ (Figure 2.1B, C). The NAIP5.2(c) chimera, containing the BIRs, NBD, HD1 (helical domain 1), and part of the WHD (winged helix domain) of NAIP5, exhibited a slightly diminished response to PrgJ. The reduced response to PrgJ was more pronounced in the corresponding NAIP6.2(c) chimera (Figure 2.2), suggesting that a portion of the PrgJ-detecting element had been replaced in these chimeras. The NAIP5.2(d) and NAIP5.2(e) chimeras, with breakpoints in HD2 (helical domain 2), did not respond to either ligand (Figure 2.1B, C). Importantly, however, the NAIP5.2(f) and NAIP5.2(g) chimeras, which contained the NAIP5 N-terminal domains but the LRR domain of NAIP2, assembled inflammasomes only in response to the NAIP5 ligand FlaA. Both chimeras were also able to activate CASP1 to initiate cleavage of IL-1β (Figure 2.1E). In addition, the corresponding NAIP6.2(f) and NAIP6.2(g) chimeras recognized FlaA as well as wild-type NAIP6 (Figure 2.2), despite containing the LRR domain from NAIP2. These results demonstrate that the LRR domains of NAIP5 and NAIP6 are dispensable for the specific response to FlaA, while the BIR domains of NAIP2 are dispensable for detection of PrgJ.

We next generated a set of reciprocal NAIP2.5 chimeras, in which the N-terminal domains of NAIP2 were fused to the C-terminal domains of NAIP5 (Figure 2.3). Replacement of the NAIP5 BIRs and NBD with those of NAIP2 (NAIP2.5(b)) did not alter the extent or specificity of oligomerization in response to FlaA, indicating that these

domains were dispensable for detection of FlaA. However, the NAIP2.5(c) chimera, which additionally contains HD1 and the N-terminal half of the WHD from NAIP2, responded only weakly to FlaA, and notably, also responded weakly to PrgJ. This result suggests a contribution of HD1 and the WHD to ligand recognition. The partial response to each ligand could not be explained by decreased expression of NAIP2.5(c) (Figure 2.3C). Whereas NAIP2.5(d) did not respond to either ligand, the NAIP2.5(e) chimera detected PrgJ at least as well as wild-type NAIP2 (Figure 2.3B, C) and was capable of inducing IL-1β cleavage specifically in response to PrgJ (Figure 2.3D). These data indicate that NAIP2 domains C-terminal to HD2 are not required for the response of NAIP2 to PrgJ. The analogous set of NAIP2.6 chimeras confirmed that the LRR domain does not dictate the ligand specificity of NAIP2 (Figure 2.4).

2.4.2 Central NBD-associated domains dictate ligand specificity

The above results suggested that the central NBD-associated helical domains of NAIPs dictate ligand specificity. To confirm this, we analyzed 'double' chimeric NAIPs in which only the central domains were exchanged between paralogs. NAIP5.2.5(c-f) or NAIP6.2.6(c-f) chimeras, containing the C-terminal half of the WHD, HD2, and most of the subsequent unannotated domain from NAIP2, responded weakly or not at all to PrgJ (Figures 2.5, 2.6). However, additional inclusion of the HD1 and full WHD from NAIP2 allowed the NAIP5.2.5(b-f) chimera to oligomerize in response to PrgJ as well as wild-type NAIP2 (Figure 2.5). The same domains of NAIP2 within a NAIP6 context also conferred a specific, albeit reduced, response to PrgJ (Figure 2.6, NAIP6.2.6(b-f)). These results demonstrate that HD1, WHD, HD2 and part of the unannotated domain of NAIP2 dictate specific recognition of PrgJ.

We next tested whether the same domains also control the specificity of NAIP6 for FlaA (Figure 2.7). Analogous to NAIP2, the C-terminal half of the WHD, HD2 and unannotated domain of NAIP6 (as in chimeras NAIP2.6.2(c-f) and NAIP2.6.2(c-g)) were not sufficient to confer a robust response to FlaA. Surprisingly, these chimeras still responded partially to PrgJ, further emphasizing the importance of HD1 and the N-terminal half of the WHD from NAIP2 for PrgJ recognition. Extension of the NAIP6 region to include HD1 and the full WHD abrogated the response to PrgJ and conferred substantial response to FlaA (i.e., NAIP2.6.2(b-f) and NAIP2.6.2(b-g)). However, these chimeras supported only partial oligomerization in response to FlaA (Figure 2.7B) and exhibited background IL-1β processing even in the absence of ligand (Figure 2.7D). Furthermore, although NAIP5 chimeras with a single breakpoint were functional (Figure 2.1 and 2.3), we were unable to generate 'double' chimeras containing the NAIP5 central domains that were capable of responding to FlaA (Figure 2.8).

The ligand specificity and degree of response for the 43 chimeras tested are summarized in Figure 2.9. Collectively, our results demonstrate that the NBD-associated HD1, WHD, HD2, and unannotated domain of NAIPs are both necessary and sufficient to dictate ligand specificity. In contrast to our expectations and those of

others in the field (e.g.,[170]), our extensive analysis revealed no role for the annotated LRR domain in specific ligand recognition.

2.4.3 The NAIP ligand specificity domain has evolved under positive selection

We hypothesized that the NAIP ligand specificity domain might be subject to diversifying evolutionary selection, driven by the need to recognize relatively rapidly evolving bacterial ligands. We found evidence for extensive recombination between rodent *Naip* paralogs, with several recombination breakpoints occurring within the ligand specificity domain. While it is difficult to assess whether these recombinants have been adaptively selected for, recombinants that yield chimeric specificity domains could be a rapid means to evolve novel NAIP specificities. Indeed, recombination has contributed to differing specificities of at least one allelic series of plant NLRs[177]. Nonetheless, the rampant recombination precludes codon-based analyses of positive selection. Instead, we performed a pairwise sliding window analysis comparing the rate of synonymous changes (dS) to the rate of nonsynonymous changes (dN) between the mouse and rat *Naip2* genes to investigate whether the ligand specificity domain displays a signature of positive selection. This sliding-window approach is less susceptible to artifacts arising due to recombination. We found that much of *Naip2* is evolving with a dN/dS ratio of less than 1, indicative of purifying selection (Figure 2.10A). Strikingly, however, we found two peaks in which the dN/dS ratio was much greater than 1. Both of these peaks lie in the domains functionally implicated in ligand recognition, although the second peak falls within the unannotated domain that has no apparent role in PrgJ recognition by mouse NAIP2. The unannotated domain does appear to play a role in flagellin recognition by mouse NAIP5, and may also be important for recognition of ligand by rat NAIP2, thereby providing a possible explanation for the peak of elevated dN/dS ratio in this region. Overall, these data are consistent with the ligand specificity domain evolving under positive selection.

We next extended our evolutionary analysis to primate *NAIP* genes. Primate genomes encode only one intact *NAIP* ortholog, and unlike in rodents, we found no evidence for recombination. We therefore analyzed these genes by maximum likelihood methods for evidence of recurrent positive selection (Figure 2.10B). We found a statistically significant signature ($p < 0.01$) of positive selection, consistent with *NAIP* evolving adaptively in primates. Moreover, when we analyzed the ligand specificity domain alone, we found strong evidence for positive selection, as well as three individual codons that have evolved under recurrent positive selection. Importantly, removal of the specificity domain from our analysis of primate *NAIP* genes resulted in a loss of the signature of positive selection. Taken together, these data suggest that *NAIP* genes in both primate and rodent genomes have evolved under positive selection, specifically in the region of the protein that confers ligand specificity.

Localized positive selection is consistent with direct association of these domains with rapidly evolving bacterial ligands. Indeed, co-immunoprecipitation of chimeric NAIPs with either FlaA or PrgJ required the full ligand specificity domain from either NAIP6 or

NAIP2, respectively (Figure 2.11). In contrast, the C-terminal LRR domain and N-terminal BIRs and NBD of NAIP6 were dispensable for association with FlaA, and the NAIP2 LRR domain was dispensable for binding PrgJ. Thus, the ligand specificity domain also dictates the physical association of NAIPs with their cognate ligands.

2.4.4 Ligand binding is strictly required for NAIP oligomerization

Because our analyses mapped NAIP ligand recognition to NBD-associated domains, we investigated the effects of ligand binding on NAIP oligomerization. Consistent with a recent study[165], we found that NAIPs are monomeric in the absence of ligand, as evidenced by the inability of FLAG-tagged NAIP5 to co-immunoprecipitate HA-tagged NAIP5 when co-transfected into HEK293T cells (Figure 2.12A, lane 2). In the absence of NLRC4, FLAG-NAIP5 bound FlaA as previously reported[4] but remained unable to associate with HA-NAIP5 (Figure 2.12A, lane 4). However, binding FlaA did permit FLAG-NAIP5 to associate with co-expressed NLRC4, and within this assembled inflammasome, FLAG-NAIP5 was able to co-immunoprecipitate HA-NAIP5 (Figure 2.12A, lane 8). This association was not detected upon mixing pre-assembled FLAG-NAIP5 and HA-NAIP5 inflammasomes (Figure 2.12A, lane 10), suggesting that HA-NAIP5 and FLAG-NAIP5 co-immunoprecipitated within a single assembled inflammasome rather than through non-specific inter-oligomer association.

The above results imply that more than one NAIP protomer can be incorporated into a single inflammasome oligomer. To confirm this, we assessed the stoichiometry of NAIP and NLRC4 constituents within assembled inflammasomes. We labeled both NLRs with the same FLAG epitope, immunoprecipitated the complex via the 6myc-tagged ligand, and determined the ratio of NLRs by densitometry of an anti-FLAG immunoblot. We analyzed the stoichiometry of the NAIP2–NLRC4–PrgJ inflammasome because under our experimental conditions 6myc-PrgJ did not associate with either unassembled NAIP2 or NLRC4 (Figure 2.13A). Therefore, the observed ratio of 5 NLRC4 to 2 NAIP2 (Figure 2.13B, ratio = 2.5) reflects the average composition of assembled inflammasomes free of PrgJ-associated unassembled components. Inflammasome stoichiometry was insensitive to the relative expression levels of each NLR (Figure 2.13C). However, it remains possible that assembled inflammasomes are heterogeneous in composition, and thus the stoichiometry of individual oligomers could vary considerably from the observed 5:2 average. Indeed, we later determined that the apparent inclusion of multiple NAIP protomers in an inflammasome was an artifact of the lack of CASP1 co-expression, allowing the NLRC4 CARD domains to mediate homotypic interactions between inflammasome rings rather than with the CARD of CASP1. Instead, NAIP/NLRC4 inflammasomes are nucleated by a single ligand-bound NAIP (Chapter 4).

We took advantage of the artifact that inflammasomes apparently contained multiple NAIP protomers to determine under what conditions NAIP2 and NAIP5 could co-oligomerize. We found that a FLAG-NAIP5–NLRC4–FlaA inflammasome did not associate with co-expressed HA-NAIP2 in the absence of PrgJ, nor did FLAG-NAIP5

pull down an HA-NAIP2–NLRC4–PrgJ inflammasome in the absence of FlaA (Figure 2.12A, lanes 11-14). However, HA-NAIP2 was incorporated into a FLAG-NAIP5 inflammasome, at roughly the same efficiency as HA-NAIP5, when the NAIP2 cognate ligand PrgJ was present in addition to the NAIP5 ligand FlaA (Figure 2.12A, lane 16). Again, this association was not detected upon mixing pre-assembled FLAG-NAIP5 and HA-NAIP2 inflammasomes (Figure 2.12A, lane 18), suggesting that it was not due to non-specific inter-oligomer association. These data indicate that ligand binding is strictly required for assembly of NAIPs into inflammasomes. We confirmed this finding using the oligomerization assay to monitor inflammasome assembly (Figure 2.12C). Again, FLAG-NAIP2 and FLAG-NAIP6 were detectable in the oligomer only when provided with their cognate ligands, i.e., PrgJ and FlaA, respectively.

The model that ligand binding is required for assembly of each NAIP protomer into inflammasomes predicts a 1:1 ligand:NAIP stoichiometry. We therefore analyzed the ratio of GFP-NAIP and GFP-ligand in inflammasomes immunoprecipitated by FLAG-NLRC4, which does not associate with either NAIP or ligand individually (Figure 2.13E). Within assembled inflammasomes, both NAIP5 and NAIP2 were present at a 1:1 stoichiometry with their respective ligands (Figure 2.13F). These results are consistent with the model that cognate ligand is required for NAIP incorporation into NLRC4-containing inflammasomes.

2.5 Discussion

Despite their shared domain architecture, members of the NLR family respond to a surprisingly diverse set of ligands and agonists[71,80,166]. To gain insight into how the NBD-LRR platform can evolve such divergent specificities, we investigated the mechanism of ligand recognition by NAIP–NLRC4 inflammasomes. We took advantage of the existence of several highly related murine NAIPs with distinct ligand specificities. We mapped the NAIP specificity domain using a panel of 43 chimeric NAIPs, of which, remarkably, 31 (72%) retained at least some function.

The LRR domain has been identified as the specificity determinant in several plant NLR immune sensors[177-180]. Mammalian NLR specificity domains have not previously been conclusively mapped, although some evidence implicates the LRR in NOD2 and NLRP3[101,181]. However, our analysis of NAIP chimeras indicated that the annotated NAIP LRR domain was dispensable for ligand specificity. The SwissModel repository[182] predicts that the LRR domain extend N-terminally to NAIP2 residue 1069, a finding later confirmed by our structural analysis of NAIP5 (Chapter 4), but even this extended region had no effect on specificity (Figure 2.1, NAIP5.2(f); Figure 2.3, NAIP2.5(f)). It is possible that the unannotated region between HD2 and breakpoint “f” might adopt an LRR-like tertiary structure, but structural modeling predicts near-total alpha helicity (Figure 2.14B). Although we were unable to unambiguously model much of this region (Chapter 4), the unannotated domain of NAIP5 does not extend the LRR and consists of at least some ordered helices. Furthermore, recognition of PrgJ by NAIP2 does not require any domains C-terminal to HD2 (Figure 2.3), ruling out a role for LRRs in PrgJ specificity.

Taken together, these data indicate that the LRRs do not play a critical role in ligand recognition. Although we did not uncover a role for the NAIP LRR domain in dictating ligand specificity, our data do not undermine the previously established role of the LRR domain in NAIP autoinhibition[3]. Thus, taken together, our data demonstrate that ligand recognition and autoinhibition are mediated by separable domains.

Our data indicate that the LRR domains of NLRs are not necessarily pathogen-detection domains. Consistent with this possibility, the NLR family member NLRP1B can be activated upon direct cleavage by *Bacillus anthracis* lethal factor protease[32,98]. Importantly, cleavage of NLRP1B is sufficient for its activation, even by a heterologous protease, arguing against a requirement for specific recognition of lethal factor by the NLRP1B LRR domain[98]. Instead, the NLRP1B LRR is essential only for mediating NLR autoinhibition, analogous to the LRR of NAIPs. The mechanism of recognition by other NLRs remains undetermined and may involve the LRR domain. However, we propose that the total surface area of NLRs available for evolvable ligand recognition includes the NBD-associated domains in addition to the LRR. Recombination or mutation anywhere within this expanded region could result in novel ligand specificities. Indeed, our data suggest that the ligand specificity domain of NAIP5 is shifted C-terminally relative to the corresponding ligand specificity domain of NAIP2 (compare Figure 2.1, NAIP5.2(f) and Figure 2.3, NAIP2.5(e)). The existence of a large and evolvable ligand specificity domain may be critical for NAIP proteins to keep pace with rapidly evolving bacterial ligands.

Based on both functional and evolutionary analysis, we define the central NBD-associated domains as the NAIP ligand specificity domain (Figure 2.14A). However, these domains from NAIP5 were unable to mediate a response to FlaA (Figure 2.8), despite 'single' chimera analysis mapping NAIP5 recognition of FlaA to this region (Figures 2.1, 2.3, 2.9). The most likely explanation for the failure of NAIP2.5.2 chimeras to respond to FlaA is that chimerism at these breakpoints disturbed the tertiary structure enough to disrupt function without dramatically reducing expression levels (Figure 2.8B). The partial oligomerization and background IL-1β processing of NAIP2.6.2 chimeras (Figure 2.7) support this interpretation. The ligand recognition domains are thus unlikely to be entirely modular, but instead require placement within a compatible structural framework. Indeed, homology modeling of NAIP2 (Figure 2.14B) suggests that these largely alpha helical domains adopt a compact, closed structure with substantial inter-domain contacts. Chimerism at breakpoints that interfere with critical inter-domain contacts could therefore disrupt NAIP structure and function. This may explain why NAIP5.2.5(c-f) responds partially to PrgJ, whereas NAIP5.2.5(c-g), containing a larger NAIP2 fragment, does not (Figure 2.5).

The simplest interpretation of our data is that the 'specificity' domain we identify is also the site of ligand binding. This domain is both necessary and sufficient to dictate ligand specificity and confers the ability for chimeric NAIPs to physically associate with ligand (Figure 2.11). Moreover, we find evidence for positive selection acting on this domain

but not on the rest of the protein (Figure 2.10). This signature of positive selection is consistent with direct contact between this domain and rapidly evolving pathogen ligands, as has been observed in other host-pathogen interactions[2,176] and even some plant NLRs[183]. Furthermore, subsequent structural analysis confirmed that most contacts with flagellin are mediated by the NAIP5 'specificity domain' (Chapter 4).

We therefore propose a model in which NAIP activation is triggered by ligand binding to the NBD-associated helical domains (Figure 2.14C). Intriguingly, the NBD-associated HD2 of NLRC4 was recently shown to participate in autoinhibition[159]. If the HD2 of NAIPs plays a similar role, then ligand binding at HD2 could sterically preclude autoinhibition. Alternatively, ligand binding may allosterically activate NAIPs through rigid body rotation of the autoinhibitory LRR domain away from the NBD, reminiscent of the mechanism by which cytochrome C activates APAF-1 to trigger apoptosome assembly[184]. Finally, we cannot rule out the possibility that the LRR contributes to ligand binding, even though the LRR domain does not define PrgJ vs. FlaA specificity. For example, the LRR might recognize structural features common to both ligands. Indeed, a few residues of the NAIP5 LRR participate in flagellin binding (Figure 4.6). Because the NAIP2 and NAIP5 LRRs did not affect specific recognition of PrgJ or FlaA, we hypothesize that the analogous LRR residues of NAIP2 may contact similar motifs in PrgJ.

Our data indicate that NAIPs only assemble into inflammasomes in the presence of their cognate ligands (Figure 2.12). Strikingly, ligand-bound NAIPs are licensed to co-oligomerize with NLRC4 but remain unable to self-associate (Figure 2.12A). This finding was later explained by structural work that identified the oligomerization surfaces of NAIPs and NLRC4[5,112]. NAIPs contain a 'donor' oligomerization surface that is exposed by ligand binding (Chapter 5), which binds to the 'acceptor' surface of NLRC4. NLRC4 also contains a 'donor' surface to continue propagation of inflammasome assembly with additional NLRC4 protomers. However, NAIPs lack an 'acceptor' surface, thus explaining why further NAIP protomers are not recruited to the inflammasome. This propagation specificity is advantageous in recruiting only NLR protomers that contain the signaling CARD domain, which NAIPs lack. It remains unexplored whether other NLRs contain an 'acceptor' surface that would allow inclusion into the NAIP/NLRC4 inflammasome. It is tempting to speculate that, as a specialized sub-class of NLR that has segregated ligand sensing from signaling via a PYRIN or CARD domain, NAIPs are unique in this regard. Regardless, taken together, our data identify the NBD-associated helical domains as a highly evolvable surface that mediates pathogen recognition by a class of NLR innate immune sensors.

2.6 Methods

2.6.1 Expression constructs

All constructs except IL-1β (pSPORT, CMV promoter) were cloned into the MSCV2.2 retroviral vector, and expression was driven by the viral LTR. Murine NLRC4, NAIP2,

NAIP5, NAIP6, CASPASE-1, pro-IL-1ß, and N-terminally 6myc-tagged FlaA (*L. pneumophila*) and PrgJ (*S. typhimurium*) have been described [3], except that NAIP5 was modified to replace 3 nucleotides (3683-3685) missing relative to the reference sequence (NCBI refseq. NP_035000.2) using Quickchange PCR with forward (GGAAACGTCAGAAAAGTTTGCCCAGGCTCTGGGTTCTCTCAG) and reverse (CTGAGAGAACCCAGAGCCTGGGCAAACTTTTCTGACGTTTCC) primers.

NAIP chimeras were generated by splicing-by-overlap extension PCR as follows: the N-terminal segment was amplified with a 5' flanking BamHI site and a Kozak sequence (GCCACC) preceding the start codon using the forward primers for NAIP2 (TAAGGATCCGCCACCATGGCAGCCCAGGGAGAAG) or NAIP5/NAIP6 (TAAGGATCCGCCACCATGGCTGAGCATGGGGAG) and the reverse primers indicated in Table 2.1. The C-terminal segment was amplified with a 3' flanking NotI site using the reverse primers for NAIP2 (GTTGCGGCCGCTCACTTCTGAATGACAGGAGAG), NAIP5 (GTTGCGGCCGCTTACTCCAGGATAACAGGAGAG), or NAIP6 (GTTGCGGCCGCTTACTCCAGGACAACAGGAGAG) and the forward primers indicated in Table 2.1. Internal primers listed in Table 2.1 contained complementary overhanging sequence to allow for fusion of the two segments by PCR, using the external NAIP2, NAIP5, and NAIP6 forward and reverse primers as appropriate. Resulting chimeras were cloned into the BglII and NotI sites of MSCV2.2. To generate NAIP5.2(g), NAIP6.2(g), NAIP2.5(g), and NAIP2.6(g), the LRR was removed from each paralog by digestion with XhoI and NotI and ligated into the XhoI and NotI sites of the appropriate alternate paralog.

N-terminal epitope tags were added by conventional PCR using forward primers containing a 5' flanking BamHI site, Kozak sequence, and either 1xFLAG or 1xHA tag immediately after the start codon, as follows: FLAG-NLRC4 (AAAAGGATCCGCCACCATGGATTACAAGGACGACGATGACAAGAACTTTATAAGG AACAACAGACG), FLAG-NAIP5 and FLAG-NAIP6 (AAAAGGATCCGCCACCATGGATTACAAGGACGACGATGACAAGGCTGAGCATGGG GAGTCCTCCG), FLAG-NAIP2 (AAAAGGATCCGCCACCATGGATTACAAGGACGACGATGACAAGGCAGCCCAGGG AGAAGCCGTTGAGG), HA-NAIP5 (AAAAGGATCCGCCACCATGTACCCATACGATGTTCCAGATTACGCTGCTGAGCAT GGGGAGTCCTCCG), and HA-NAIP2 (AAAAGGATCCGCCACCATGTACCCATACGATGTTCCAGATTACGCTGCAGCCCAG GGAGAAGCCGTTGAGG). Reverse primers contained a 3' flanking NotI site, including the NAIP2, NAIP5 and NAIP6 reverse primers above and the NLRC4 reverse primer (TTTTGCGGCCGCTTAAGCAGTCACTAGTTTAAAGGTGCC). Tagged constructs were cloned into the BglII and NotI sites of MSCV2.2.

Constructs were fused with GFP by cloning in frame with an N-terminal GFP moiety in MSCV2.2 lacking an IRES-GFP. Coding sequences were amplified with flanking 5' NotI

(1 nt addition to keep in frame; adds a 3-Ala linker) and 3' SalI sites using the following primers: NAIP2 (Fwd: CAAGCGGCCGCAGCAGCCCAGGGAGAAGC; Rev: GGTGGTCGACTCACTTCTGAATGACAGGAGAG), NAIP5 (Fwd: CAAGCGGCCGCAGCTGAGCATGGGGAGTC; Rev: GGTGGTCGACTTACTCCAGGATAACAGGAGAG), PrgJ (Fwd: CCAGCGGCCGCATCGATTGCAACTATTGTCCCTG; Rev: GGTGGTCGACTCATGAGCGTAATAGCGTTTCAAC). Amplicons were cloned into the NotI and SalI sites downstream of GFP. GFP-FlaA has been previously described (Kofoed & Vance 2011).

All NAIP constructs were fully sequenced using the following primers: mscvF (AAGCCCTTTGTACACCCTAAGCC), mNAIP-F1 (GGGACACTGTGCAGTGTTT), mNAIP-F2 (GCCACATGAACTTGCCAGA), mNAIP-F3 (GTGTCCTCATGTGGGCAG), mNAIP-R4 (GTCCAGAAAACTCAATCTCTC), mscvR (CCTCACATTGCCAAAAGAC). NLRC4 constructs were sequenced using the following primers: mscvF, mNLRC4-F1 (CCTGCTTTTCTGAACTTCTACC), mNLRC4-F2 (CTGAGAAATCTGATGAAGACCC), mNLRC4-F3 (CCTTCGTAGAGTGTGGCATC), mscvR. All other constructs were sequenced using mscvF and mscvR.

Table 2.1. Primers used to construct NAIP chimeras

Chimera	N-terminal segment Rev primer	C-terminal segment Fwd primer
NAIP5.2(a), NAIP6.2(a), NAIP2.5(a), NAIP2.6(a)	TCCACCACTGTAGAATGAACTG CTGCTGCATCA	TGATGCAGCAGCAGTTCATTCT ACAGTGGTGGA
NAIP6.2(b), NAIP2.5(b), NAIP2.6(b)	AGACAGTATTATAGAAGGGAAA CTCTT	AAGAGTTTCCCTTCTATAATACT GTCT
NAIP5.2(c), NAIP6.2(c)	CATCTTCATCAACTCCTGCCTCT GCCAGGT	ACCTGGCAGAGGCAGGAGTTG ATGAAGATG
NAIP2.5(c)	CAAGAGGGTGGTAAGCTTTTCA TCTTCATCAACTCCTGCCTC	GAGGCAGGAGTTGATGAAGATG AAAAGCTTACCACCCTCTTG
NAIP2.6(c)	CTTTACATCTTCATCAACTCCTG CCTCTGCCAGGTTATCACTATT G	CAATAGTGATAACCTGGCAGAG GCAGGAGTTGATGAAGATGTAA AG
NAIP5.2(d), NAIP6.2(d), NAIP2.5(d), NAIP2.6(d)	TGGATGGAGCTTCATGTAATCC TCATTTTC	GAAAATGAGGATTACATGAAGC TCCATCCA
NAIP5.2(e), NAIP6.2(e)	CAGCAACTGTGTTGCTTTCAT	ATGAAAGCAACACAGTTGCTG
NAIP2.5(e)	GCAAAATAAATGGAGAGCATTC AGCAACTGTGTTGCTTTCATGA G	CTCATGAAAGCAACACAGTTGC TGAATGCTCTCCATTTATTTTGC
NAIP2.6(e)	GCAAAATAAATGGAGAGCATTC	CTCATGAAAGCAACACAGTTGC

	AGCAACTGTGTTGCTTTCATGAG	TGAATGCTCTCCATTTATTTTGC
NAIP5.2(f), NAIP6.2(f), NAIP2.5(f), NAIP2.6(f)	TAAACGGAACTCAATACTCTGTGAAGCTGA	TCAGCTTCACAGAGTATTGAGTTCCGTTTA
NAIP5.2(g), NAIP6.2(g), NAIP2.5(g), NAIP2.5(g)	[existing XhoI, NotI sites]	
NAIP5.2.5(c-f)	CTTCATCTTCATCAACTCCTGCCTCTGCCAGGTCATCACTATTG (template = NAIP5)	CAATAGTGATGACCTGGCAGAGGCAGGAGTTGATGAAGATGAAG (template = NAIP2.5(f))
NAIP5.2.5(c-g)	CTTCATCTTCATCAACTCCTGCCTCTGCCAGGTCATCACTATTG (template = NAIP5)	CAATAGTGATGACCTGGCAGAGGCAGGAGTTGATGAAGATGAAG (template = NAIP2.5(g))
NAIP5.2.5(b-f)	GTGAAAAAAGCTTCCGTAACACAGAGACAGTATTATAGAAGGGAAACTC (template = NAIP5)	GAGTTTCCCTTCTATAATACTGTCTCTGTGTTACGGAAGCTTTTTTCAC (template = NAIP2.5(f))
NAIP5.2.5(b-g)	GTGAAAAAAGCTTCCGTAACACAGAGACAGTATTATAGAAGGGAAACTC (template = NAIP5)	GAGTTTCCCTTCTATAATACTGTCTCTGTGTTACGGAAGCTTTTTTCAC (template = NAIP2.5(g))
NAIP6.2.6(c-f)	CTTCATCTTCATCAACTCCTGCCTCTGCCAGGTCATCACTATTG (template = NAIP6)	CAATAGTGATGACCTGGCAGAGGCAGGAGTTGATGAAGATGAAG (template = NAIP2.6(f))
NAIP6.2.6(c-g)	CTTCATCTTCATCAACTCCTGCCTCTGCCAGGTCATCACTATTG (template = NAIP6)	CAATAGTGATGACCTGGCAGAGGCAGGAGTTGATGAAGATGAAG (template = NAIP2.6(g))
NAIP6.2.6(b-f)	GTGAAAAAAGCTTCCGTAACACAGAGACAGTATTATAGAAGGGAAACTC (template = NAIP6)	GAGTTTCCCTTCTATAATACTGTCTCTGTGTTACGGAAGCTTTTTTCAC (template = NAIP2.6(f))
NAIP6.2.6(b-g)	GTGAAAAAAGCTTCCGTAACACAGAGACAGTATTATAGAAGGGAAACTC (template = NAIP6)	GAGTTTCCCTTCTATAATACTGTCTCTGTGTTACGGAAGCTTTTTTCAC (template = NAIP2.6(g))
NAIP2.6.2(c-f)	CTTTACATCTTCATCAACTCCTGCCTCTGCCAGGTTATCACTATTG (template = NAIP2)	CAATAGTGATAACCTGGCAGAGGCAGGAGTTGATGAAGATGTAAAG (template = NAIP6.2(f))
NAIP2.6.2(c-g)	CTTTACATCTTCATCAACTCCTGCCTCTGCCAGGTTATCACTATTG (template = NAIP2)	CAATAGTGATAACCTGGCAGAGGCAGGAGTTGATGAAGATGTAAAG (template = NAIP6.2(g))
NAIP2.6.2(b-f)	GTGAAAAAAACTTCCGTAACACAAAGACAGTATTATAGAAGGGAAACTC (template = NAIP2)	GAGTTTCCCTTCTATAATACTGTCTTTGTGTTACGGAAGTTTTTTTCAC (template = NAIP6.2(f))
NAIP2.6.2(b-g)	GTGAAAAAAACTTCCGTAACAC	GAGTTTCCCTTCTATAATACTGT

	AAAGACAGTATTATAGAAGGGAAACTC (template = NAIP2)	CTTTGTGTTACGGAAGTTTTTTTCAC (template = NAIP6.2(g))
NAIP2.5.2(c-f)	CAAGAGGGTGGTAAGCTTTTCATCTTCATCAACTCCTGCCTC (template = NAIP2)	GAGGCAGGAGTTGATGAAGATGAAAAGCTTACCACCCTCTTG (template = NAIP5.2(f))
NAIP2.5.2(c-g)	CAAGAGGGTGGTAAGCTTTTCATCTTCATCAACTCCTGCCTC (template = NAIP2)	GAGGCAGGAGTTGATGAAGATGAAAAGCTTACCACCCTCTTG (template = NAIP5.2(g))
NAIP2.5.2(b-f)	GTGAAAAAAACTTCCGTAATACAGAGACAGTATTATAGAAGGGAAACTC (template = NAIP2)	GAGTTTCCCTTCTATAATACTGTCTCTGTATTACGGAAGTTTTTTTCAC (template = NAIP5.2(f))
NAIP2.5.2(b-g)	GTGAAAAAAACTTCCGTAATACAGAGACAGTATTATAGAAGGGAAACTC (template = NAIP2)	GAGTTTCCCTTCTATAATACTGTCTCTGTATTACGGAAGTTTTTTTCAC (template = NAIP5.2(g))

2.6.2 Cell culture and transient transfection

HEK293T cells were grown in complete medium (DMEM, 10% FBS, 2mM L-glutamine, 100 U/mL penicillin, 100 µg/mL streptomycin). Cells were seeded into 6-well plates at a density of 8 x 10^5 cells per well or into 24-well plates at 1.5 x 10^5 cells per well, and transfected the following day using Lipofectamine 2000 (Invitrogen) according to manufacturer's instruction.

2.6.3 Reconstituted inflammasome and native PAGE

Inflammasome reconstitution was performed as previously described [3,185]. HEK293T cells were transfected in 6-well plates with 400 ng each of the indicated plasmids, harvested by trypsinization 48 hours after transfection, washed twice with cold PBS, and resuspended in cold native lysis buffer (50 mM BisTris, 50 mM NaCl, 10% w/v glycerol, 0.0001% Ponceau S, 1% digitonin, 2mM Na_3VO_4, 1 mM PMSF, 25 mM NaF, 1x Roche protease inhibitor cocktail [no EDTA], pH 7.2). Lysates were incubated for 30 min at 4ºC with rotation, and cell debris was then pelleted by centrifugation at 16,000 g for 30 min at 4ºC. Lysates were quantified for total protein by Bradford assay (Bio-Rad) to normalize gel loading, and then separated in parallel by blue native PAGE (3-12%) and SDS-PAGE (4-12%) using the Novex BisTris gel system according to manufacturer's instruction (Invitrogen). Native gels were soaked in 10% SDS for 10 min before transfer to Immobilon-FL PVDF membranes for immunoblotting.

2.6.4 Immunoprecipitation

Immunoprecipitation protocol was essentially as previously described [4]. HEK293T cells were transfected in 6-well plates with 400 ng each of the indicated plasmids, harvested by trypsinization after 48 hours, and lysed in cold IP buffer (50 mM Tris-HCl, 150 mM NaCl, 1% Triton X-100, 1x Roche protease inhibitor cocktail [no EDTA], pH 7.6). Where indicated, separately transfected HEK293T cells were mixed prior to lysis. Cell debris was pelleted by centrifugation at 16,000 g for 15 min at 4ºC, and lysates were pre-cleared with 30 µL of washed Protein G sepharose (GE Healthcare). Pre-cleared

lysates were divided equally by volume and immunoprecipitated with 1 µg of either specific antibody (anti-FLAG M2 [Sigma Aldrich], anti-c-myc [Clontech]) or normal mouse IgG (Santa Cruz Biotechnology). Immunoprecipitates were captured with Protein G sepharose, washed three times with IP buffer, eluted by boiling in SDS sample buffer, and separated by 4-12% SDS-PAGE for immunoblot analysis.

2.6.5 IL-1β processing

HEK293T cells were transfected in 24-well plates with 50 ng of pro-IL-1β and 160 ng each of the remaining plasmids. After 24 hours, media was removed and cells were lysed in plate in RIPA buffer (50 mM Tris, 150 mM NaCl, 1% NP-40, 0.5% sodium deoxycholate, 0.1% SDS, 1mM PMSF, 1x Roche protease inhibitor tablet [no EDTA], pH 8.0) for 20 min at 4ºC. Lysates were clarified by centrifugation at 16,000 g for 30 min at 4ºC, and supernatants (18% of total) were separated by 12% SDS-PAGE in MES buffer (Invitrogen) and immunoblotted for IL-1β.

2.6.6 Immunoblotting and densitometry

Proteins separated by either native or denaturing PAGE were transferred to Immobilon-FL PVDF membranes (Millipore). Membranes were blocked with Li-Cor Odyssey blocking buffer (denaturing gels) or 5% milk (native gels). Primary antibodies used were anti-NLRC4 (gift of S. Mariathasan and V. Dixit, Genentech); anti-NAIP5(961-978), which also detects NAIP6, and anti-NAIP2(33-46) [186]; anti-c-myc (9E10) (Clontech); anti-HA (3F10) (Roche); anti-FLAG (M2) (Sigma Aldrich); anti-mIL-1ß (R&D systems); and anti-β-actin (C4) (Santa Cruz). Secondary anti-mouse and anti-goat were conjugated to Alexa Fluor-680 (Invitrogen); anti-rabbit IgG was conjugated to Alexa Fluor-800 (Invitrogen) or HRP (GE Healthcare). Immunoblots were imaged using a Li-Cor fluorimeter, followed by conventional chemiluminescent immunoblotting for native gels. Native gel images shown are chemiluminescent immunoblots. Densitometry was performed on Li-Cor immunoblots using ImageJ. In at least one experiment, a FLAG-NLRC4 dilution series was used as a standard curve to ensure quantified bands of anti-FLAG immunoblots were in the linear range of signal.

2.6.7 Domain annotation and homology modeling

NAIP domains were identified by amino acid query using the NCBI Conserved Domain Database or by homology modeling NAIPs to the NLRC4 crystal structure [159] using the Phyre2 one-to-one threading tool [187]. A homology model of full-length NAIP2 was generated using the Phyre2 intensive modeling mode to allow multiple template modeling with the following templates: PDB 1I3O, RKXF, 1TFQ, 1PGV, 2A5Y, 3T6P.

2.6.8 Analysis of positive selection

Publicly available *NAIP* gene sequences were used for all analyses. Rodent *Naip* sequences (mouse *Naip1*, *2*, *5*, *6* and *7*; rat *Naip2* and *5*; and hamster *Naip*) were aligned and trimmed to remove all ambiguities and gaps. Similarly, *NAIP* orthologs from 8 primate species (human, chimpanzee, bonobo, gorilla, orangutan, rhesus macaque, baboon and African green monkey) were aligned and trimmed. Alignments were

analyzed by GARD in the HyPhy package [188] for evidence of recombination. Maximum likelihood analyses were performed on primate *NAIP* genes using PAML [189] to compare models that disallow (M7) or allow (M8) for codons to evolve under positive selection. Reported p-values compare the log likelihood values for each model using a chi-squared test with two degrees of freedom. Specific codons that evolved under recurrent positive selection with a posterior probability of >0.95 were identified using the Naive Empirical Bayes analysis within PAML. Sliding window analyses were performed using K-estimator [190].

2.7 Acknowledgments

This chapter was originally written as a peer-reviewed article, which was published in *Molecular Cell*[123]. Eric Kofoed initiated analysis of chimeric NAIP proteins. Matt Daugherty and Harmit Malik contributed positive selection analysis.

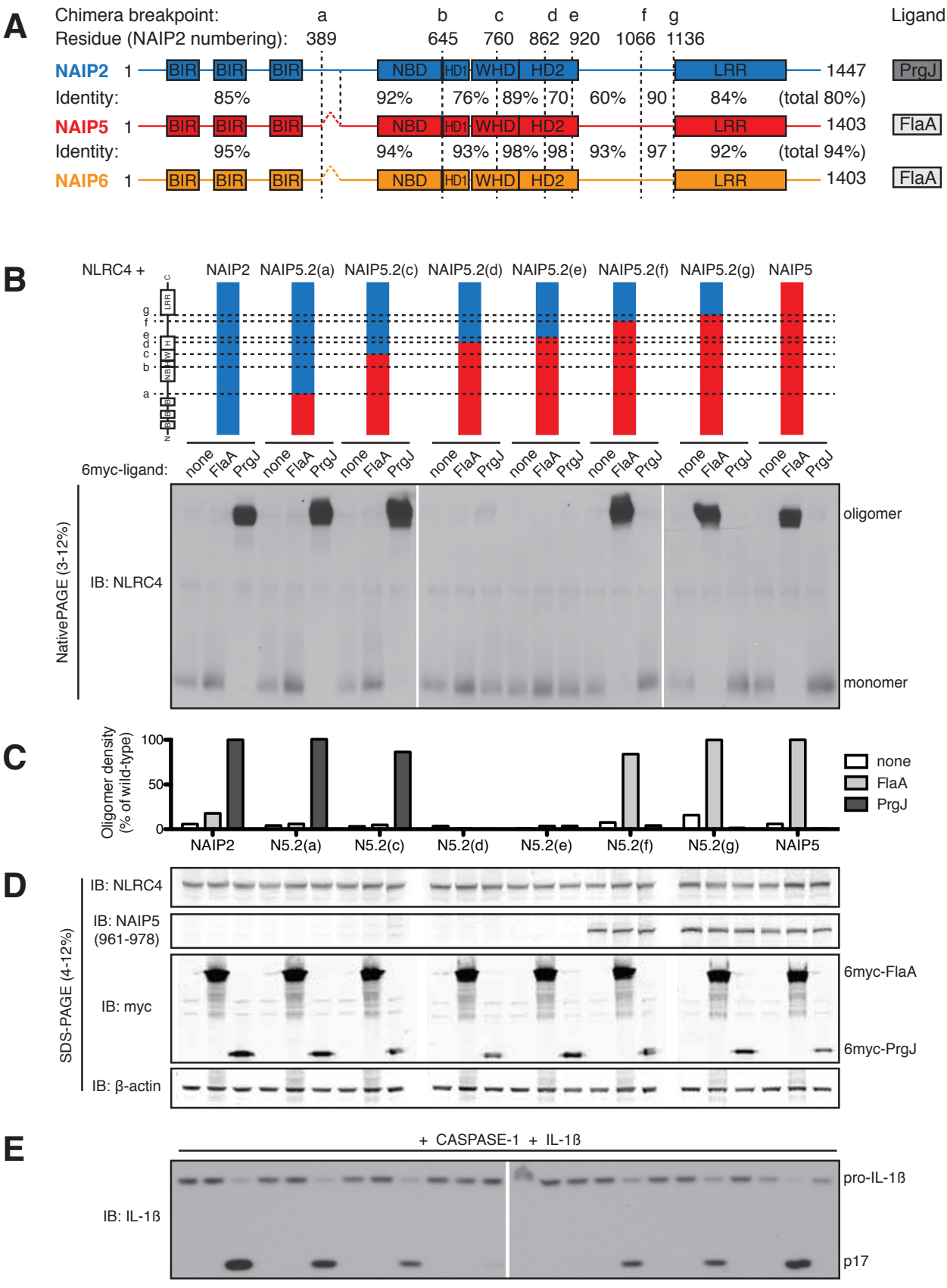

A
Chimera breakpoint:
Residue (NAIP2 numbering):
a 389
b 645
c 760
d 862
e 920
f 1066
g 1136
Ligand
NAIP2 1 BIR BIR BIR NBD HD1 WHD HD2 LRR 1447 PrgJ
Identity: 85% 92% 76% 89% 70 60% 90 84% (total 80%)
NAIP5 1 BIR BIR BIR NBD HD1 WHD HD2 LRR 1403 FlaA
Identity: 95% 94% 93% 98% 98 93% 97 92% (total 94%)
NAIP6 1 BIR BIR BIR NBD HD1 WHD HD2 LRR 1403 FlaA
B
NLRC4 +
NAIP2
NAIP5.2(a)
NAIP5.2(c)
NAIP5.2(d)
NAIP5.2(e)
NAIP5.2(f)
NAIP5.2(g)
NAIP5
6myc-ligand:
none FlaA PrgJ
NativePAGE (3-12%)
IB: NLRC4
oligomer
monomer
C
Oligomer density (% of wild-type)
100
50
0
NAIP2 N5.2(a) N5.2(c) N5.2(d) N5.2(e) N5.2(f) N5.2(g) NAIP5
none
FlaA
PrgJ
D
SDS-PAGE (4-12%)
IB: NLRC4
IB: NAIP5 (961-978)
IB: myc
IB: β-actin
6myc-FlaA
6myc-PrgJ
E
+ CASPASE-1 + IL-1ß
IB: IL-1ß
pro-IL-1ß
p17

Figure 2.1. NAIP5 LRR is dispensable for recognition of flagellin
(**A**) Schematic of NAIP predicted domains drawn to scale. Baculovirus inhibitor of apoptosis repeat (BIR) and leucine-rich repeat (LRR) domains were identified by amino acid query against the NCBI Conserved Domain Database. The nucleotide binding domain (NBD), helix domain 1 (HD1), winged helix domain (WHD), and helix domain 2 (HD2) were annotated by homology modeling to NLRC4 (PDB 4KXF). Chimera breakpoints are identified by lowercase letters (a-g) and by dashed lines. Amino acid identity between NAIP2 and NAIP5 or NAIP5 and NAIP6 is indicated for each segment and across total protein length. (**B**) Oligomerization assay to test the specificity of NAIP5.2 chimeras. HEK293T cells were transfected with NLRC4 and the indicated NAIP chimera and 6myc-tagged ligand. After 48 hours, cell lysates were harvested, normalized for total protein, subjected to blue native PAGE, and immunoblotted (IB) for NLRC4, as previously described [3]. Results shown are representative of at least 3 independent trials. See also Figure 2.2. (**C**) NAIP chimera responses to each ligand in (**B**) were quantified by densitometry of the oligomer species and normalized to wild-type NAIP2 or NAIP5 response to PrgJ or FlaA, respectively. (**D**) Lysates from (**B**) were subjected to denaturing SDS-PAGE and immunoblotted for NLRC4, NAIP5, myc, and β-actin to control for equal transfection and loading. (**E**) NAIP2.5 inflammasomes are functional and induce IL-1ß cleavage. HEK293T cells were transfected as in (**B**) but with the addition of CASP-1 and pro-IL-1ß. Cell lysates were harvested at 24 hours and immunoblotted for IL-1ß; full-length (pro) and cleaved (p17) forms are indicated. Results shown are representative of 2 independent trials.

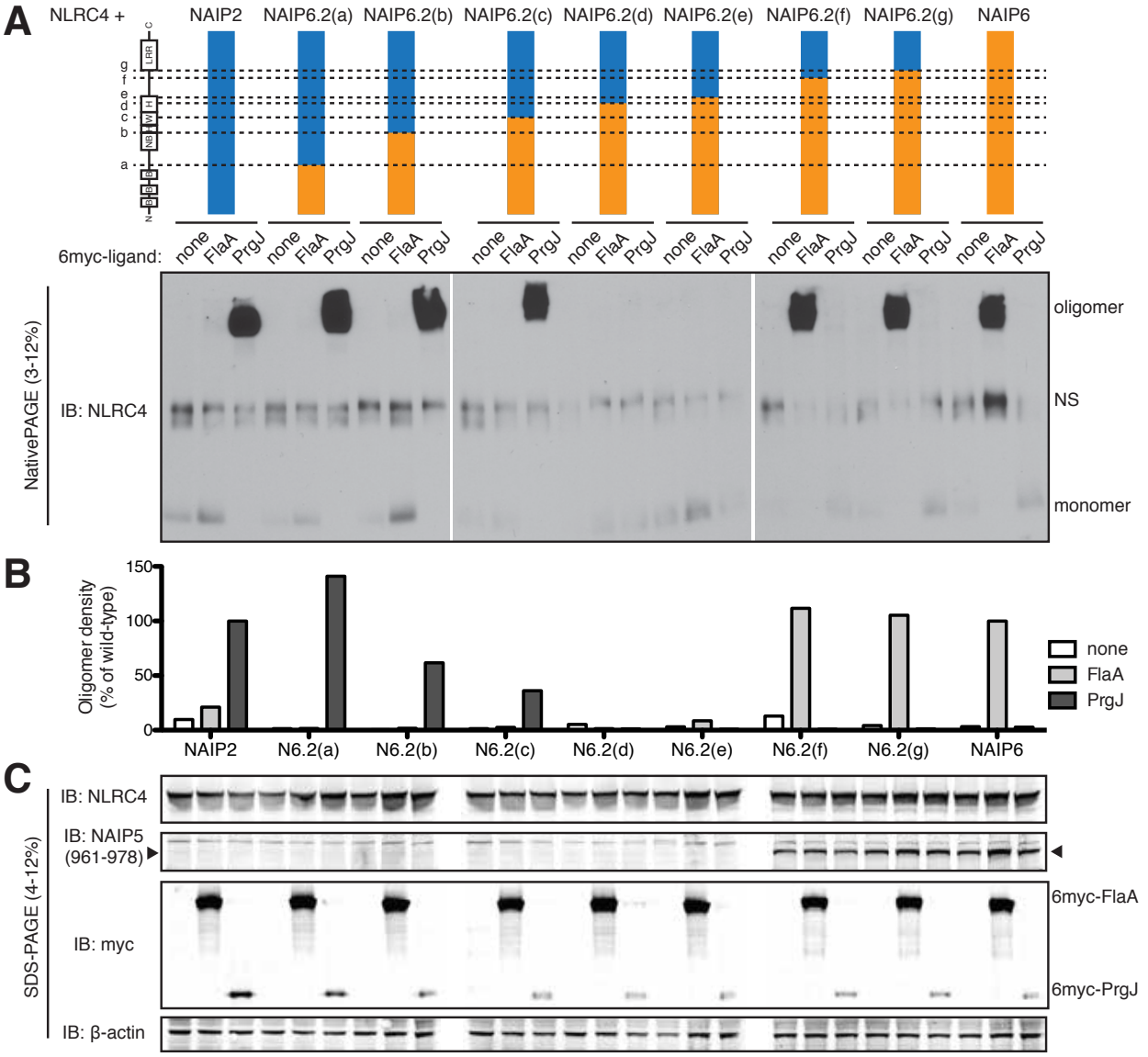


Figure 2.2. NAIP6 LRR is dispensable for recognition of flagellin
(**A**) Oligomerization assay to test the specificity of NAIP6.2 chimeras. HEK293T cells were transfected with NLRC4 and the indicated NAIP chimera and 6myc-tagged ligand. After 48 hours, cell lysates were harvested, normalized for total protein, subjected to blue native PAGE, and immunoblotted for NLRC4. NS, non-specific band. Results shown are representative of at least 3 independent trials. (**B**) NAIP chimera responses to each ligand in (**A**) were quantified by densitometry of the oligomer species and normalized to wild-type NAIP2 or NAIP6 response to PrgJ or FlaA, respectively. (**C**) Lysates from (**A**) were subjected to denaturing SDS-PAGE and immunoblotted for NLRC4, NAIP5, myc, and β-actin. The NAIP5 antibody also detects the same region of NAIP6; arrows indicate position of NAIP6 specific signal.

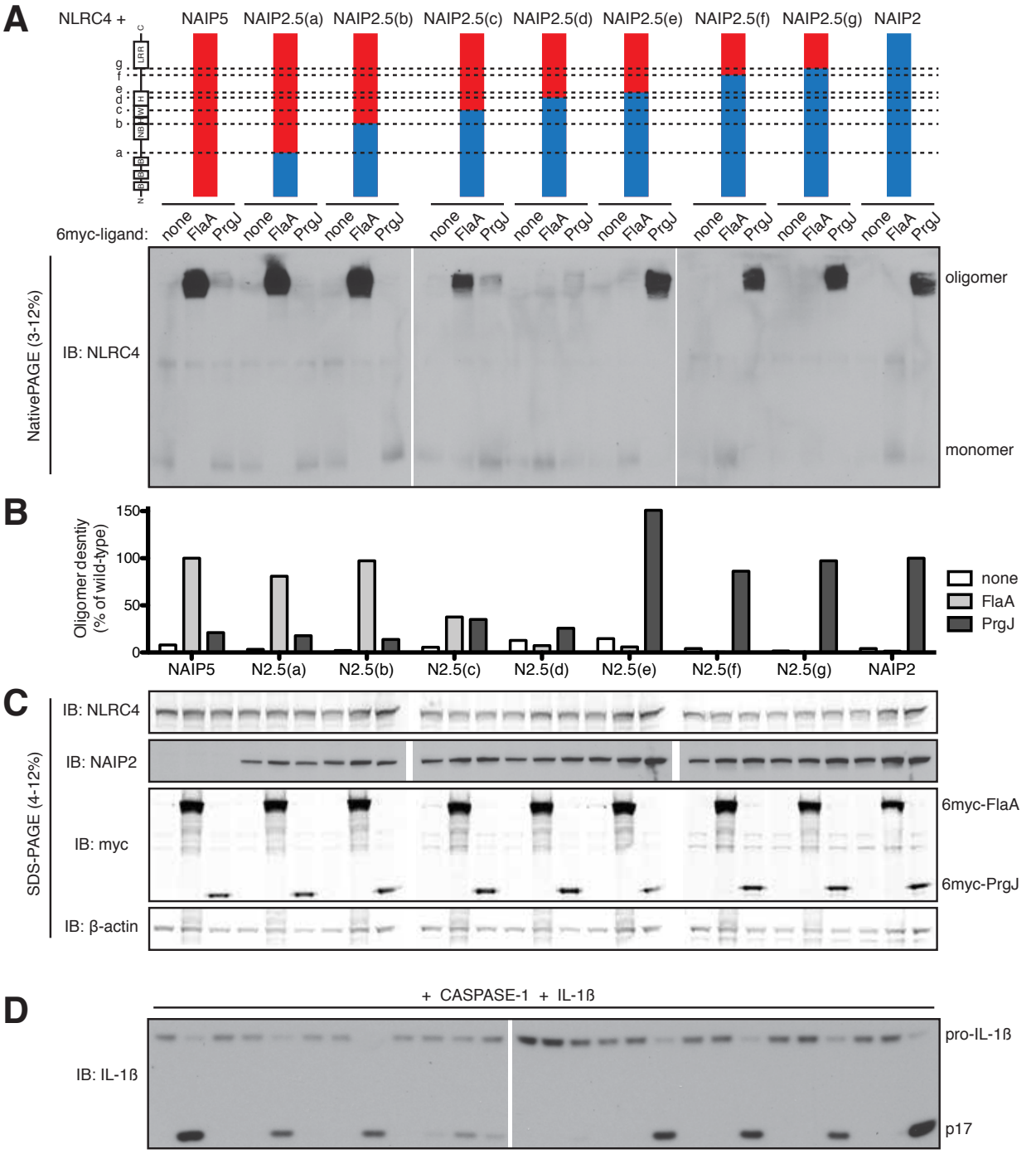


Figure 2.3. NAIP2 LRR is dispensable for recognition of PrgJ
(**A**) Oligomerization assay to test the specificity of NAIP2.5 chimeras. HEK293T cells were transfected with NLRC4 and the indicated NAIP chimera and 6myc-tagged ligand. Cell lysates were subjected to blue native PAGE as in Figure 2.1. Results shown are representative of at least 3 independent trials. See also Figure 2.4. (**B**) NAIP chimera responses to each ligand in (**A**) were quantified by densitometry of the oligomer species and normalized to wild-type NAIP2 or NAIP5 response to PrgJ or FlaA, respectively.

(**C**) Lysates from (**A**) were subjected to denaturing SDS-PAGE and immunoblotted for NLRC4, NAIP2, myc, and β-actin. (**D**) NAIP5.2 inflammasomes are functional. HEK293T cells were transfected as in (**A**) but with the addition of CASP-1 and pro-IL-1ß. Cell lysates were analyzed for IL-1ß cleavage after 24 hours. Results shown are representative of 2 independent trials.

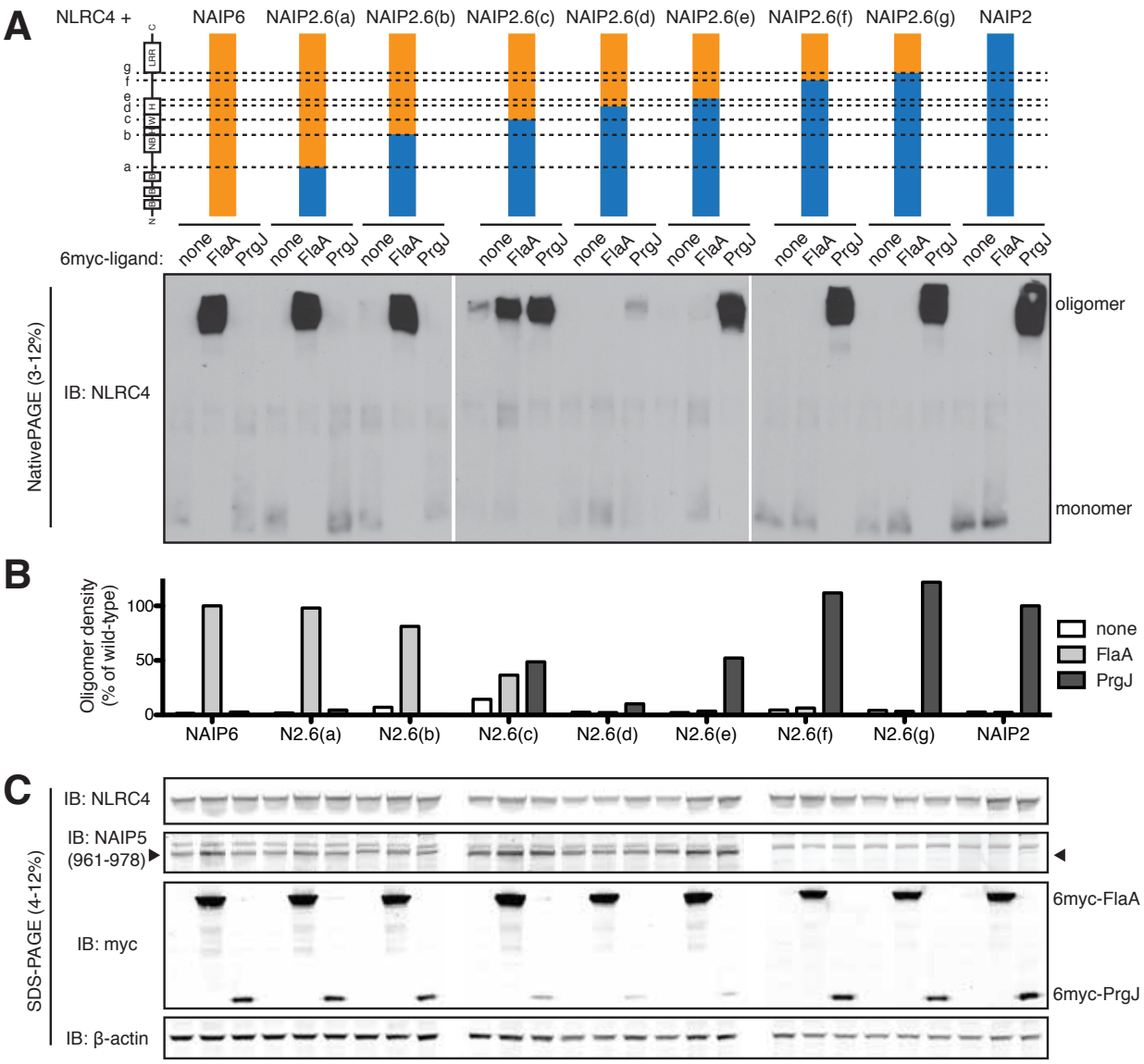


Figure 2.4. NAIP6 BIRs and NBD are dispensable for recognition of flagellin
(**A**) Oligomerization assay to test the specificity of NAIP2.6 chimeras. HEK293T cells were transfected with NLRC4 and the indicated NAIP chimera and 6myc-tagged ligand. Cell lysates were subjected to blue native PAGE as in Figure 2.2. Results shown are representative of at least 3 independent trials. (**B**) NAIP chimera responses to each ligand in (**A**) were quantified by densitometry of the oligomer species and normalized to wild-type NAIP2 or NAIP6 response to PrgJ or FlaA, respectively. (**C**) Lysates from (**A**) were subjected to denaturing SDS-PAGE and immunoblotted for NLRC4, NAIP5, myc, and β-actin. The NAIP5 antibody also detects the same region of NAIP6; arrows indicate position of NAIP6 specific signal.

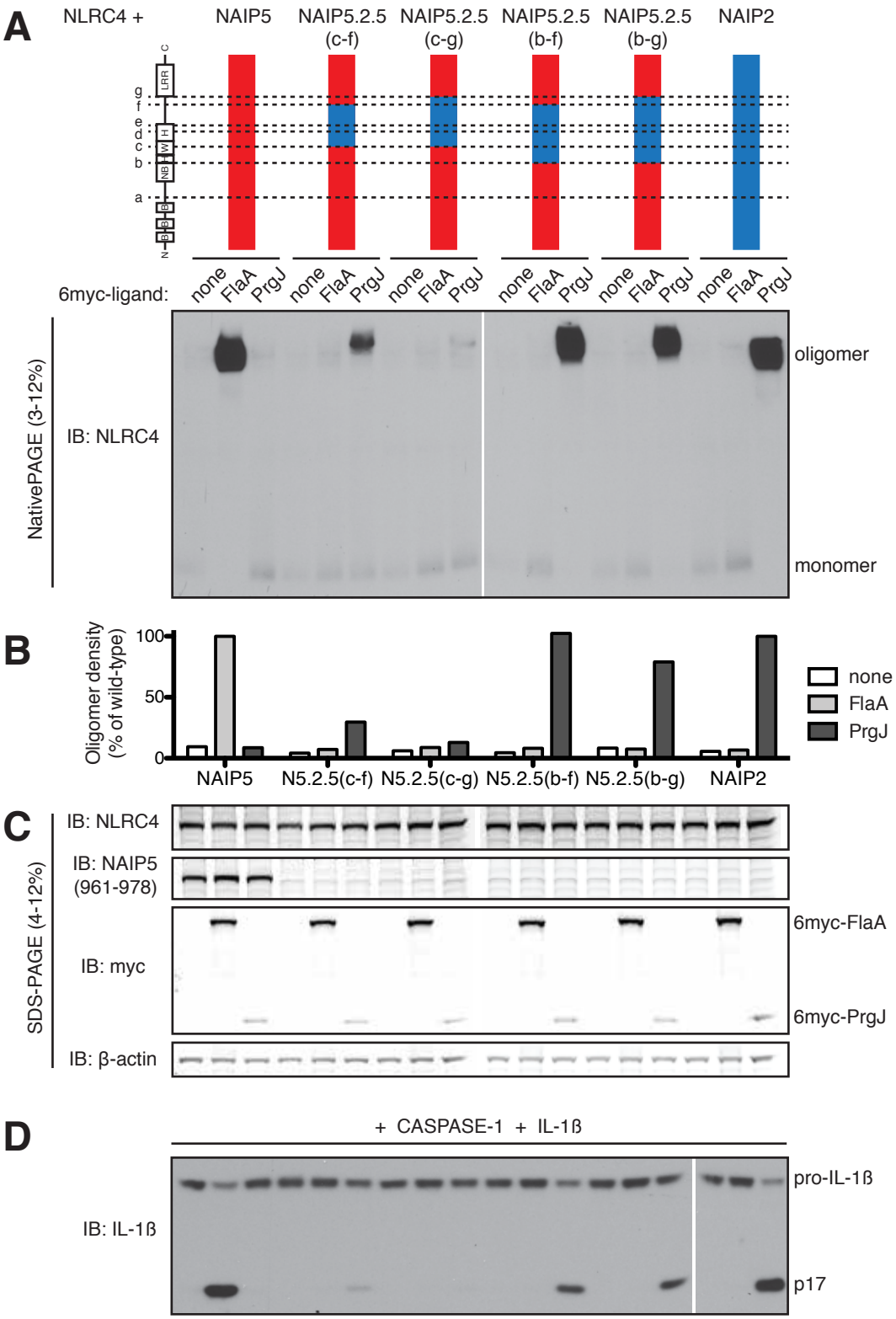


Figure 2.5. Central NBD-associated domains of NAIP2, including HD1, WHD, and HD2, are sufficient for recognition of PrgJ

(**A**) Oligomerization assay to test the specificity of NAIP5.2.5 chimeras. HEK293T cells were transfected with NLRC4 and the indicated NAIP chimera and 6myc-tagged ligand. Cell lysates were subjected to blue native PAGE as in Figure 2.1. Results shown are representative of at least 3 independent trials. See also Figure 2.6. (**B**) NAIP chimera responses to each ligand in (**A**) were quantified by densitometry of the oligomer species and normalized to wild-type NAIP2 or NAIP5 response to PrgJ or FlaA, respectively. (**C**) Lysates from (**A**) were subjected to denaturing SDS-PAGE and immunoblotted for NLRC4, NAIP5, myc, and β-actin. (**D**) NAIP5.2.5 inflammasomes are functional. HEK293T cells were transfected as in (**A**) but with the addition of CASP-1 and pro-IL-1ß. Cell lysates were analyzed for IL-1ß cleavage after 24 hours. Results shown are representative of 2 independent trials.

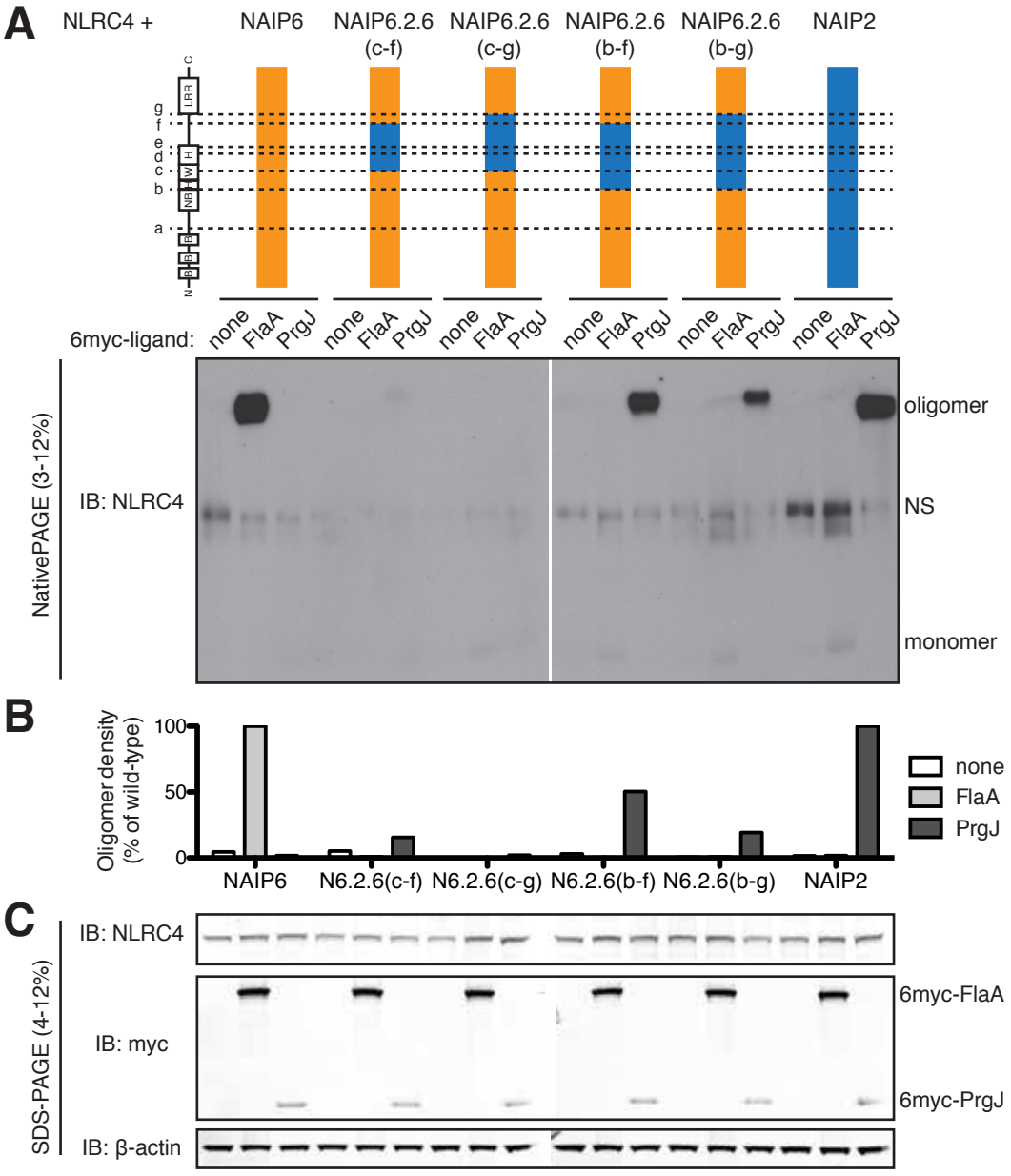


Figure 2.6. NAIP6.2.6 chimeras confirm that internal domains of NAIP2 mediate recognition of PrgJ

(**A**) Oligomerization assay to test the specificity of NAIP6.2.6 chimeras. HEK293T cells were transfected with NLRC4 and the indicated NAIP chimera and 6myc-tagged ligand. Cell lysates were subjected to blue native PAGE as in Figure 2.2. NS, non-specific band. Results shown are representative of at least 3 independent trials. (**B**) NAIP chimera responses to each ligand in (**A**) were quantified by densitometry of the oligomer species and normalized to wild-type NAIP2 or NAIP6 response to PrgJ or FlaA, respectively. (**C**) Lysates from (**A**) were subjected to denaturing SDS-PAGE and immunoblotted for NLRC4, myc, and β-actin.

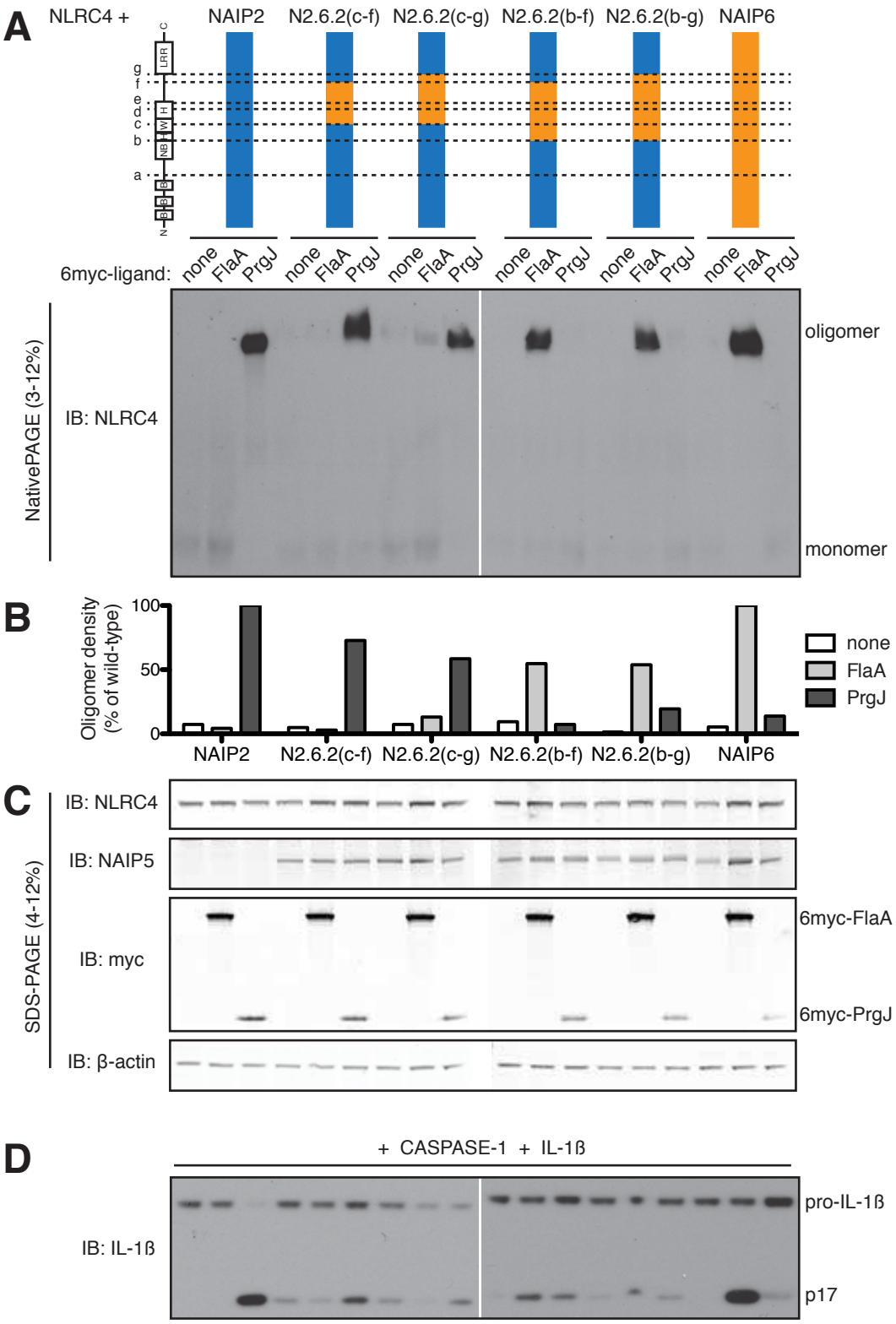


Figure 2.7. Central NBD-associated domains of NAIP6, including HD1, WHD and HD2, are sufficient for recognition of flagellin

(**A**) Oligomerization assay to test the specificity of NAIP2.6.2 chimeras. HEK293T cells were transfected with NLRC4 and the indicated NAIP chimera and 6myc-tagged ligand. Cell lysates were subjected to blue native PAGE as in Figure 2.1. See also Figure 2.8. (**B**) NAIP chimera responses to each ligand in (A) were quantified by densitometry of the oligomer species and normalized to wild-type NAIP2 or NAIP6 response to PrgJ or FlaA, respectively.(**C**) Lysates from (**A**) were subjected to denaturing SDS-PAGE and immunoblotted for NLRC4, NAIP5, myc, and β-actin. The NAIP5 antibody also detects NAIP6. (**D**) NAIP2.6.2 chimeras exhibit basal IL-1ß cleavage. HEK293T cells were transfected as in (**A**) but with the addition of CASP-1 and pro-IL-1ß. Cell lysates were analyzed for IL-1ß cleavage after 24 hours. All results shown are representative of 3 independent trials.

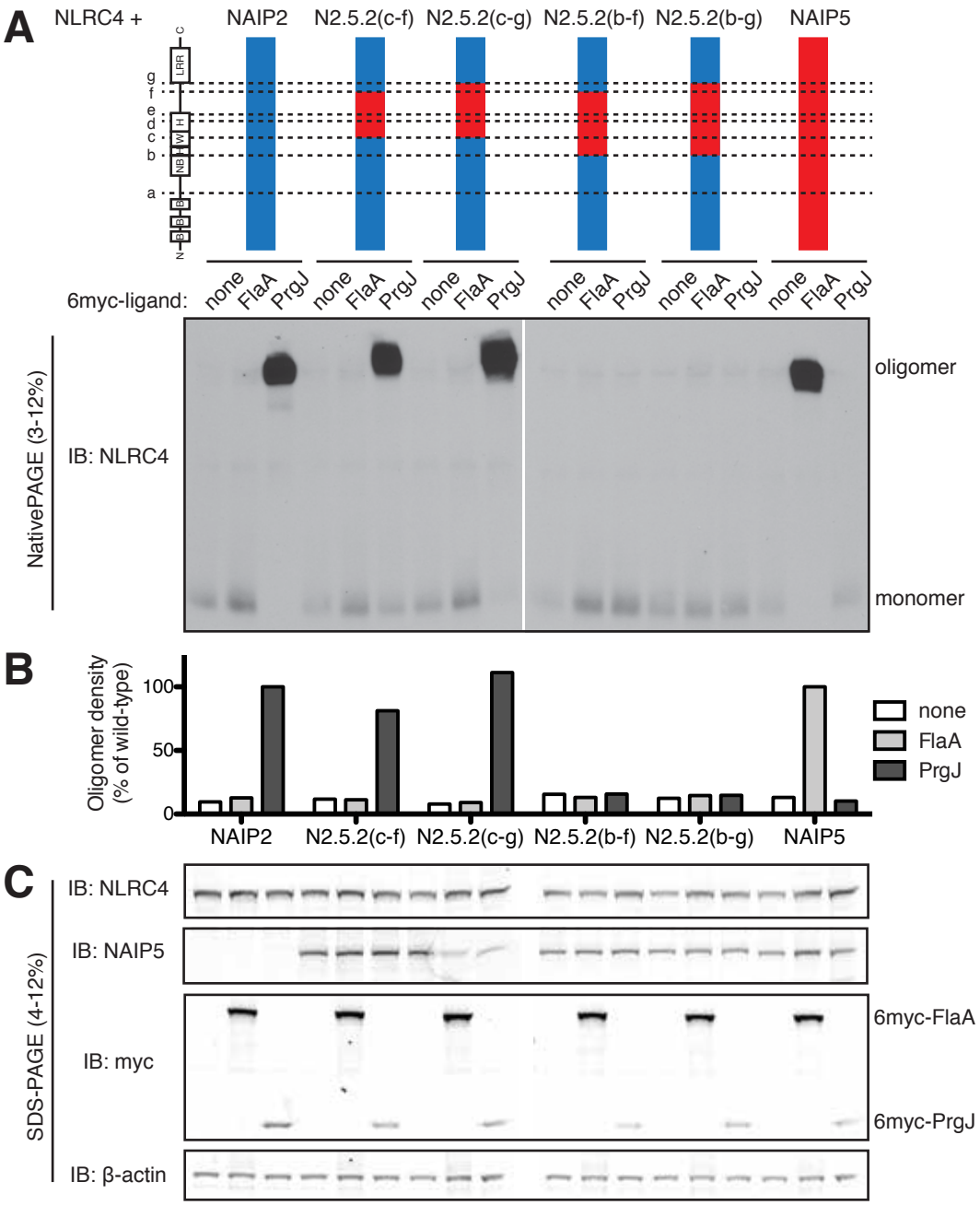


Figure 2.8. The HD1, WHD, HD2, and unannotated domains of NAIP5 cannot mediate specific recognition of flagellin

(**A**) Oligomerization assay to test the specificity of NAIP2.5.2 chimeras. HEK293T cells were transfected with NLRC4 and the indicated NAIP chimera and 6myc-tagged ligand. Cell lysates were subjected to blue native PAGE as in Figure 2.2. Results shown are representative of at least 3 independent trials. (**B**) NAIP chimera responses to each ligand in (**A**) were quantified by densitometry of the oligomer species and normalized to wild-type NAIP2 or NAIP5 response to PrgJ or FlaA, respectively. (**C**) Lysates from (**A**) were subjected to denaturing SDS-PAGE and immunoblotted for NLRC4, NAIP5, myc, and β-actin.

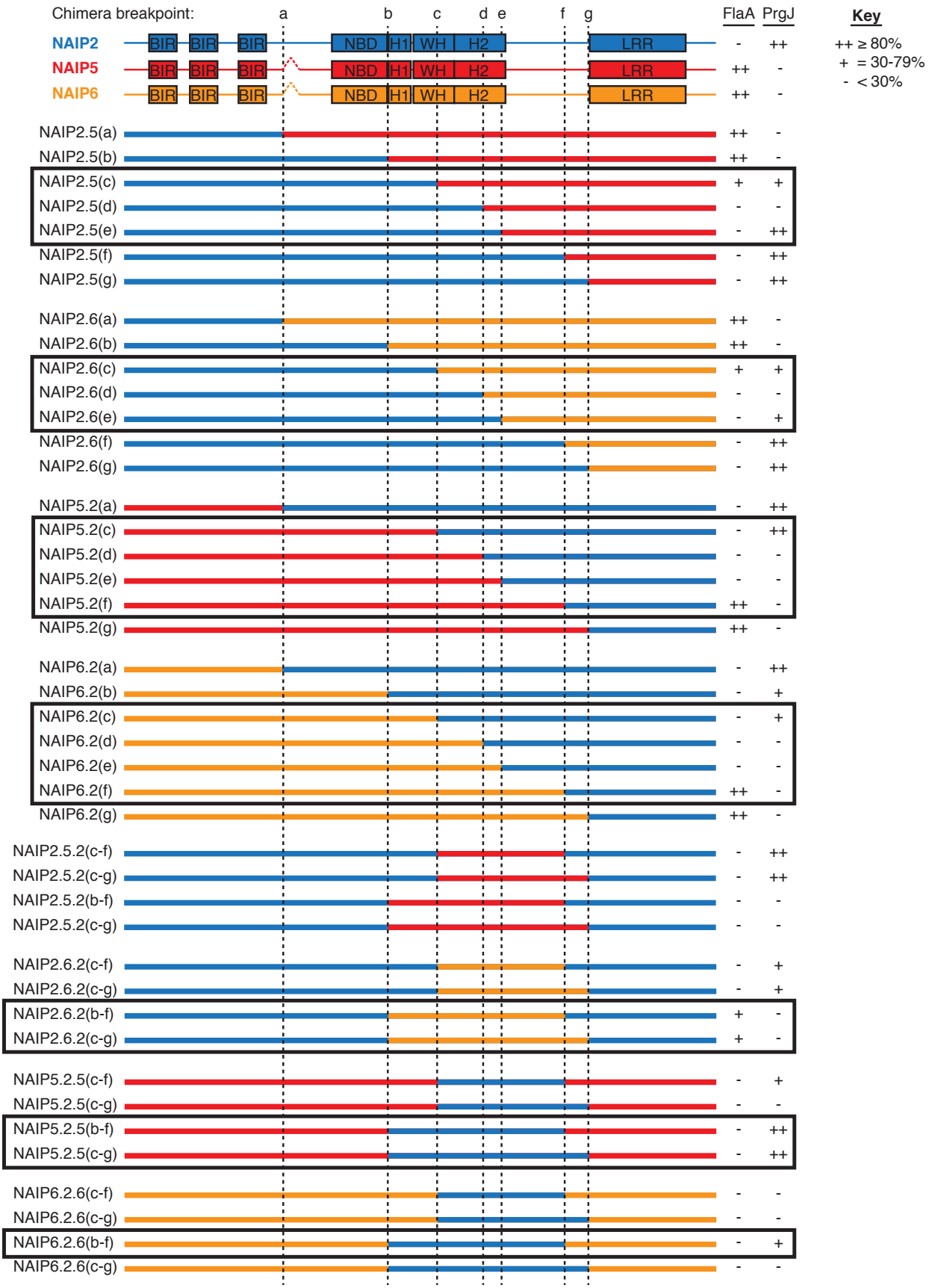

Chimera breakpoint:
a
b
c
d
e
f
g
FlaA
PrgJ
Key
++ ≥ 80%
+ = 30-79%
- < 30%
NAIP2
NAIP5
NAIP6
BIR
NBD
H1
WH
H2
LRR
NAIP2.5(a) ++ -
NAIP2.5(b) ++ -
NAIP2.5(c) + +
NAIP2.5(d) - -
NAIP2.5(e) - ++
NAIP2.5(f) - ++
NAIP2.5(g) - ++
NAIP2.6(a) ++ -
NAIP2.6(b) ++ -
NAIP2.6(c) + +
NAIP2.6(d) - -
NAIP2.6(e) - +
NAIP2.6(f) - ++
NAIP2.6(g) - ++
NAIP5.2(a) - ++
NAIP5.2(c) - ++
NAIP5.2(d) - -
NAIP5.2(e) - -
NAIP5.2(f) ++ -
NAIP5.2(g) ++ -
NAIP6.2(a) - ++
NAIP6.2(b) - +
NAIP6.2(c) - +
NAIP6.2(d) - -
NAIP6.2(e) - -
NAIP6.2(f) ++ -
NAIP6.2(g) ++ -
NAIP2.5.2(c-f) - ++
NAIP2.5.2(c-g) - ++
NAIP2.5.2(b-f) - -
NAIP2.5.2(c-g) - -
NAIP2.6.2(c-f) - +
NAIP2.6.2(c-g) - +
NAIP2.6.2(b-f) + -
NAIP2.6.2(c-g) + -
NAIP5.2.5(c-f) - +
NAIP5.2.5(c-g) - -
NAIP5.2.5(b-f) - ++
NAIP5.2.5(c-g) - ++
NAIP6.2.6(c-f) - -
NAIP6.2.6(c-g) - -
NAIP6.2.6(b-f) - +
NAIP6.2.6(c-g) - -

Figure 2.9. Summary of NAIP chimera results

Responses of NAIP chimeras to FlaA and PrgJ, as determined by densitometry of oligomer species on native gels (Figures 2.1-2.8), are categorized as full (++, ≥ 80% of wild-type), partial (+, 30-79% of wild-type), or negligible (–, 0-29% of wild-type). Boxed chimeras highlight the regions mediating specificity for ligands (i.e., where specificity switches from one ligand to the other). Chimeras are color-coded as follows: NAIP2 sequence (blue), NAIP5 sequence (red), NAIP6 sequence (orange).

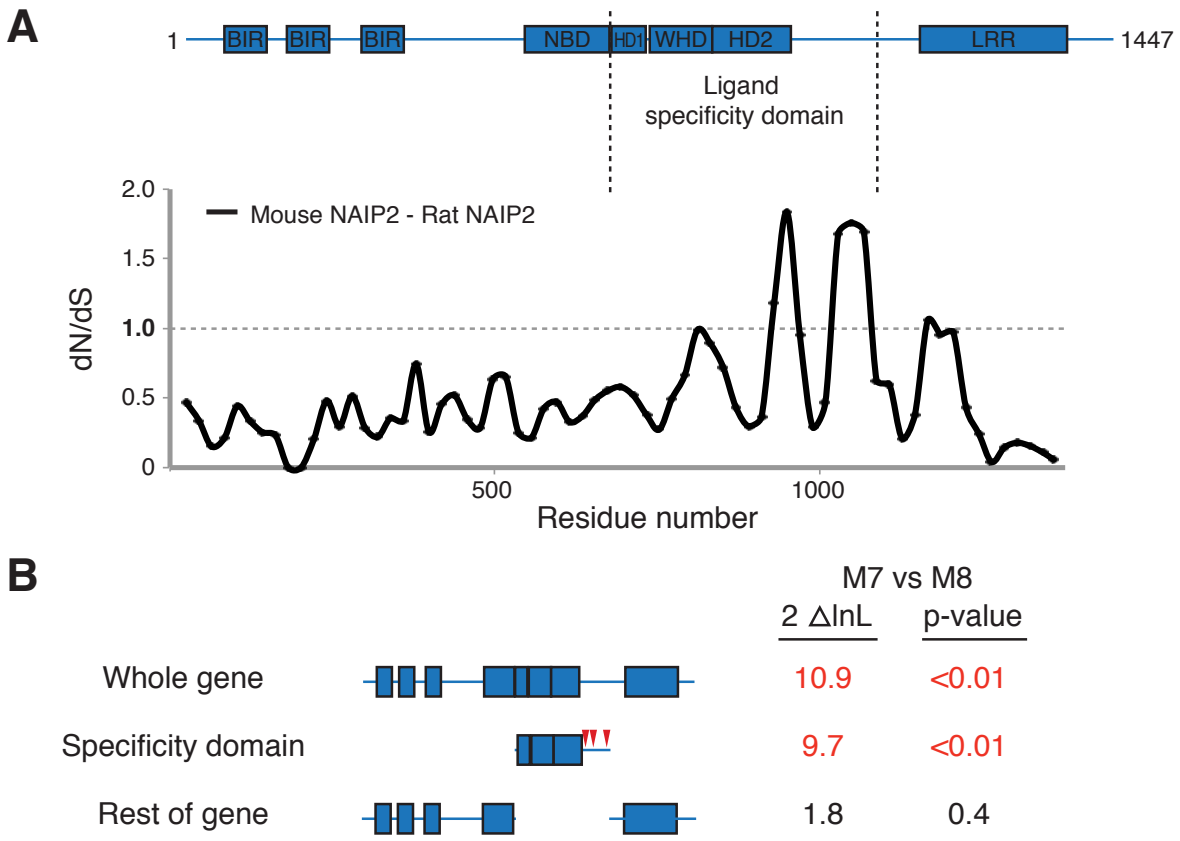


Figure 2.10. The ligand specificity domain has evolved under positive selection
(**A**) Sliding window comparison of the dN/dS ratio between mouse and rat *Naip2* genes. dN/dS ratios were calculated every 20 codons with a window size of 50 codons. Shown above is the domain structure of mouse NAIP2. (**B**) Results of PAML analyses on the entire primate *NAIP* gene, the ligand specificity domain alone, and the entire gene outside the ligand specificity domain. The left column shows two times the log likelihood difference between a model that allows (M8) or disallows (M7) positive selection. The right column shows the statistical significance of support for the gene, or domain, having evolved under positive selection. Values in red indicate strong support for positive selection. Red triangles indicate primate *NAIP* codons (corresponding to mouse NAIP2 residues 941, 965 and 1049) identified as having evolved under recurrent positive selection with a posterior probability of >0.95. See also Figure 2.11.

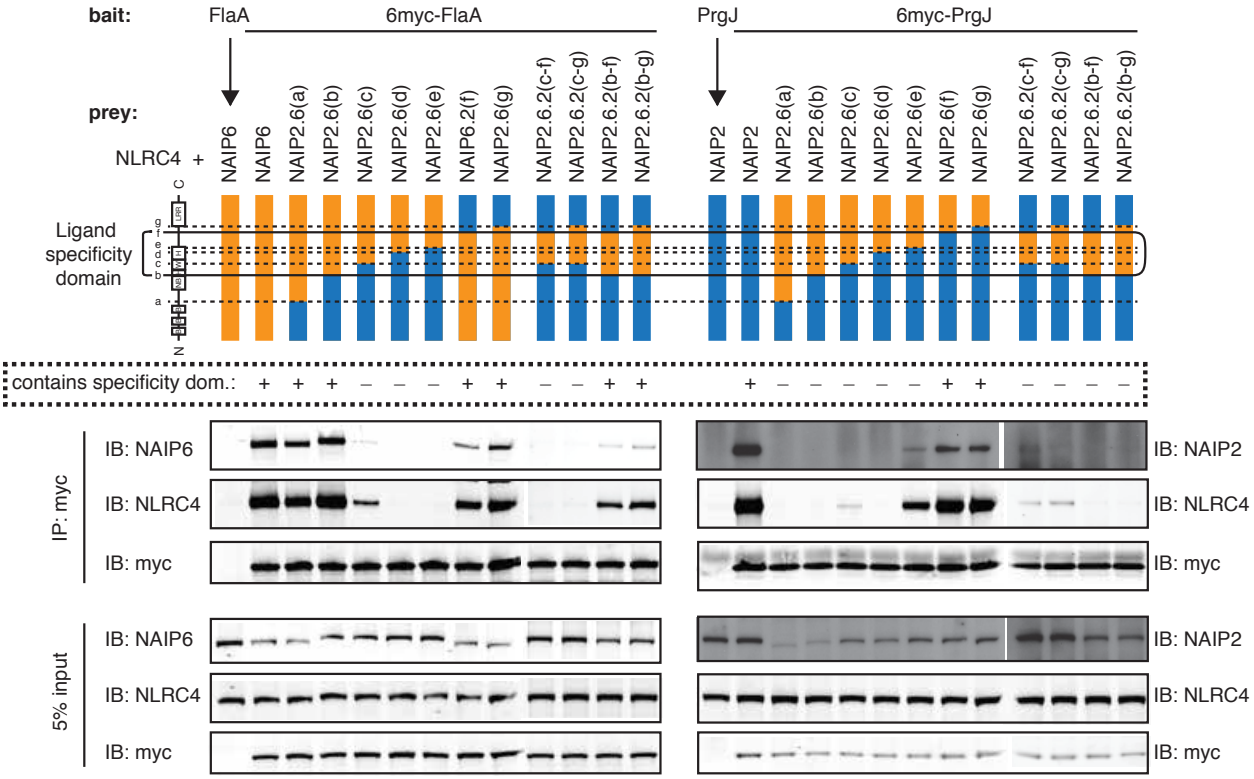


Figure 2.11. Ligand specificity domain confers the ability to bind ligand
HEK293T cells were transfected with the indicated NAIP chimera and ligand along with NLRC4. After 48 hours, cell lysates were harvested and subjected to immunoprecipitation (IP) with anti-myc antibody. Immunoprecipitates were separated by SDS-PAGE and immunoblotted for NAIP6 (left panel) or NAIP2 (right panel), followed by NLRC4 and myc. Prior to immunoprecipitation, 5% of lysate volume was removed as input, separated by SDS-PAGE, and immunoblotted in parallel with IP samples (lower panels). The boxed panel indicates whether the NAIP chimera contains the full ligand specificity domain (between breakpoints b and f).

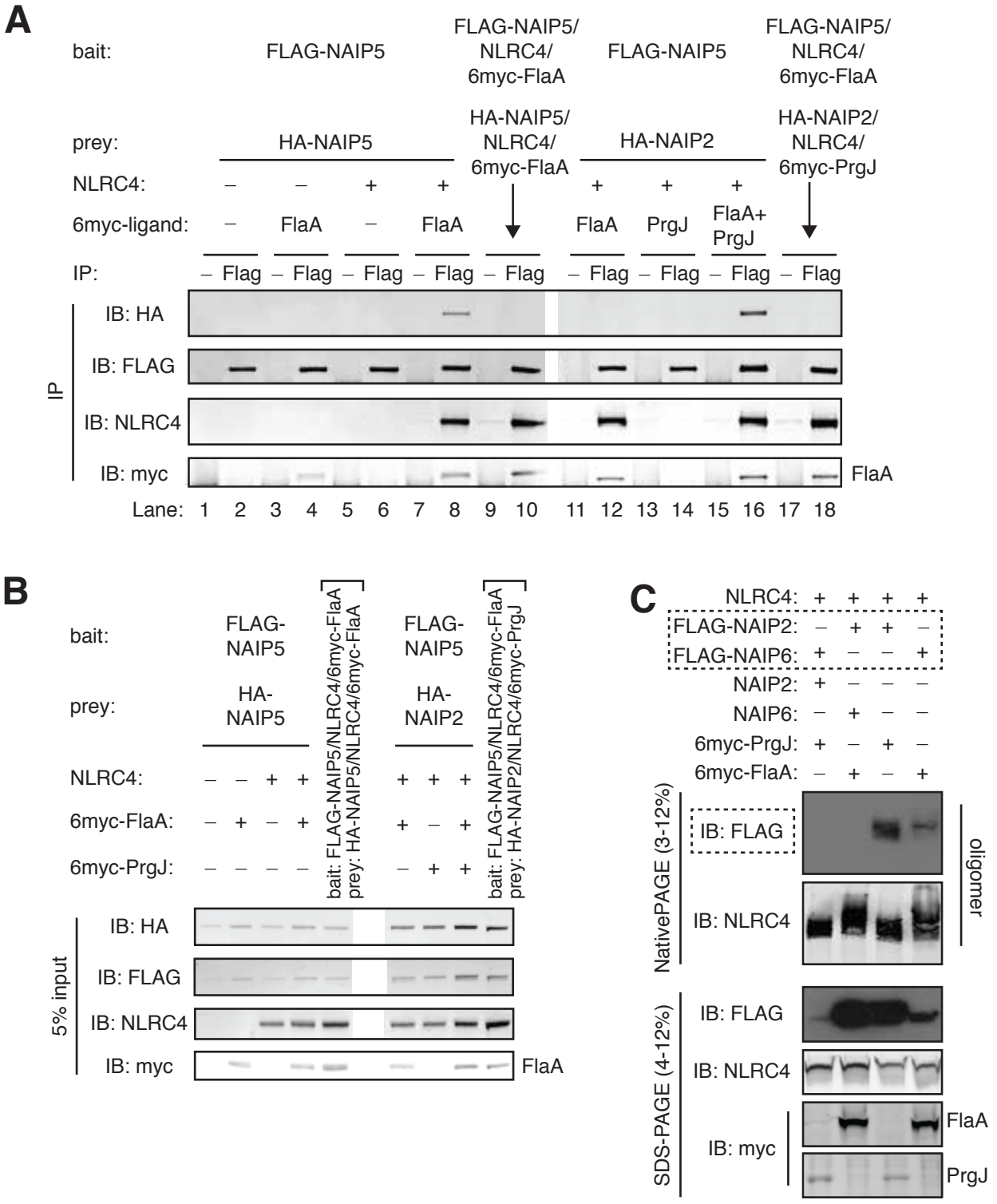


Figure 2.12. NAIPs require cognate ligand to assemble into an inflammasome
(**A**) HA-NAIP2 associates with FLAG-NAIP5 only when both cognate ligands are present. HEK293T cells were co-transfected with multiple NAIP constructs, distinguishable by FLAG or HA tags, and NLRC4 and 6myc-tagged ligands as indicated. After 48 hours, cell lysates were divided equally, and assembled inflammasomes were

immunoprecipitated with anti-FLAG antibody (Flag) or normal mouse IgG (–) as a negative control. To control for non-specific inter-oligomer association, separately assembled inflammasomes were mixed in lysate (denoted by arrows). Immunoprecipitates were separated by SDS-PAGE and immunoblotted for HA, FLAG, NLRC4 and myc. Results are representative of at least 4 independent trials. (**B**) Prior to immunoprecipitations in (**A**), 5% of lysate volume was removed as input, separated by SDS-PAGE, and immunoblotted in parallel with IP samples. (**C**) FLAG-NAIPs are detectable in assembled inflammasomes only when provided with their cognate ligand. HEK293T cells were co-transfected with NLRC4 and the indicated 6myc-tagged ligand and NAIP construct(s). After 48 hours, cell lysates were harvested, separated in parallel by blue native PAGE and SDS-PAGE, and immunoblotted for FLAG followed by NLRC4 and myc. See also Figure 2.13.

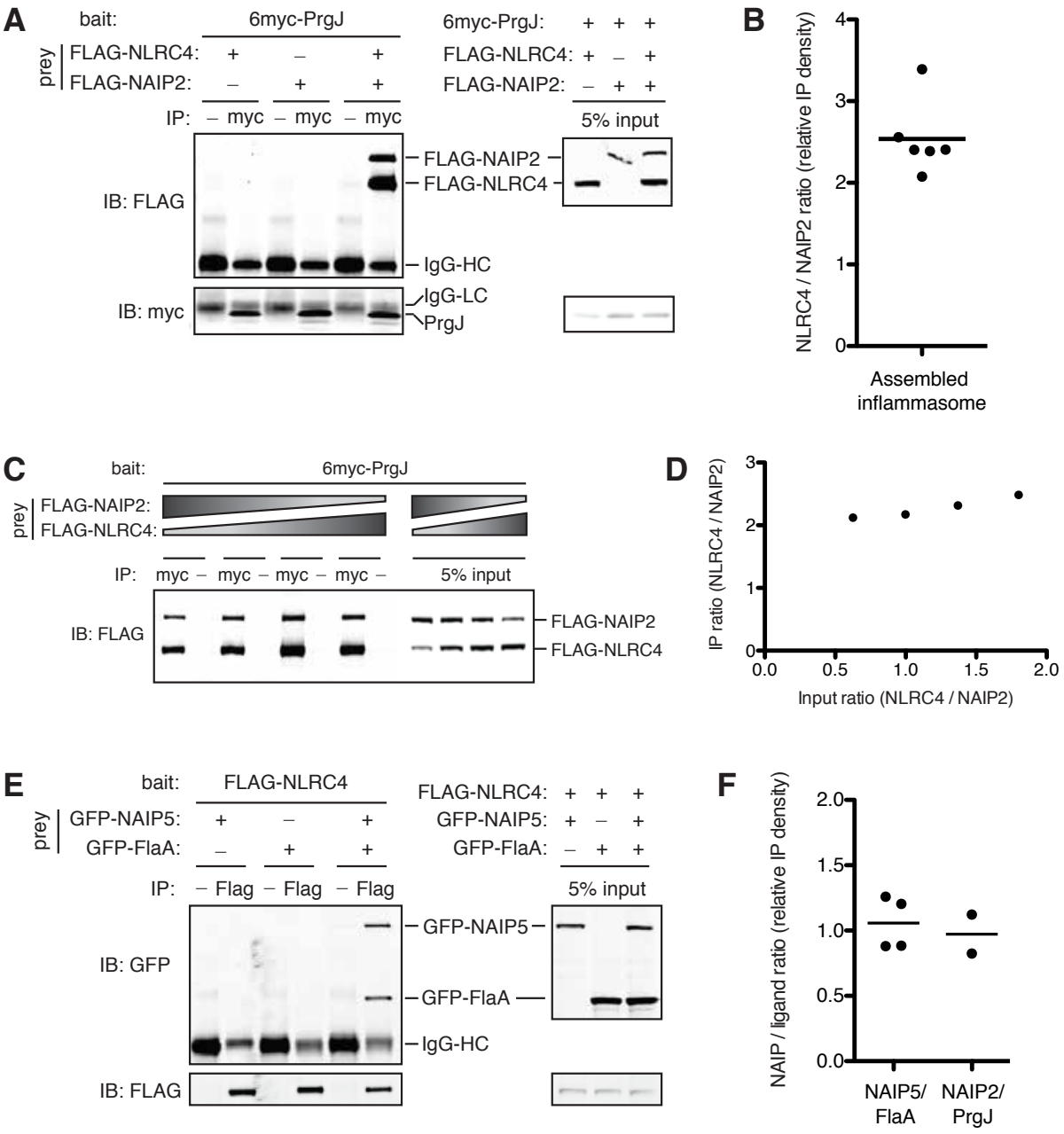


Figure 2.13. NAIP–NLRC4 inflammasome stoichiometry

(**A**) PrgJ associates only with assembled NAIP2 and NLRC4. HEK293T were transfected with 6myc-PrgJ and equal amounts of FLAG-NLRC4 and FLAG-NAIP2 as indicated. After 48 hours, cell lysates were divided equally and subjected to immunoprecipitation (IP) with anti-myc antibody (myc) or normal mouse IgG (–) as a negative control. Immunoprecipitates were separated by SDS-PAGE and immunoblotted for FLAG and myc. Prior to immunoprecipitation, 5% of lysate volume was removed as input, separated by SDS-PAGE, and immunoblotted in parallel with IP samples (right panel). (**B**) The ratio of NLRC4 to NAIP2 in assembled inflammasomes was calculated by densitometry of FLAG-NLRC4 and FLAG-NAIP2 bands immunoprecipitated by 6myc-PrgJ in (**A**, lane 6). Results and mean ratio are displayed from 6 independent trials.

(**C-D**) Inflammasome stoichiometry does not depend on relative NLR expression level. FLAG-NLRC4 and FLAG-NAIP2 plasmids were inversely titrated and transfected with 6myc-PrgJ into HEK293T cells. (**C**) Inflammasomes were immunoprecipitated as described in (**A**) and immunoblotted for FLAG. (**D**) Densitometry was performed as described in (**B**), and the ratio of NLRC4 to NAIP2 in IP lanes was plotted as a function of the ratio in the corresponding 5% input lanes. All results are representative of at least 3 independent trials. (**E**) NLRC4 associates only with assembled NAIP and ligand. HEK293T were transfected with FLAG-NLRC4 and equal amounts of GFP-NAIP5 and GFP-FlaA as indicated. Cell lysates were divided equally, immunoprecipitated with anti-FLAG antibody (FLAG) or normal mouse IgG (–), separated by SDS-PAGE, and immunoblotted for GFP and FLAG. Prior to IP, 5% of lysate was removed as input, separated by SDS-PAGE, and immunoblotted in parallel with IP samples (right panel). (**F**) NAIP and cognate ligand are stoichiometric within assembled inflammasomes. The ratio of NAIP to ligand was calculated by densitometry of GFP-NAIP5 and GFP-FlaA bands, or GFP-NAIP2 and GFP-PrgJ bands, immunoprecipitated by FLAG-NLRC4 as in (**E**, lane 6). Results and mean are displayed from 4 or 2 independent trials, respectively.

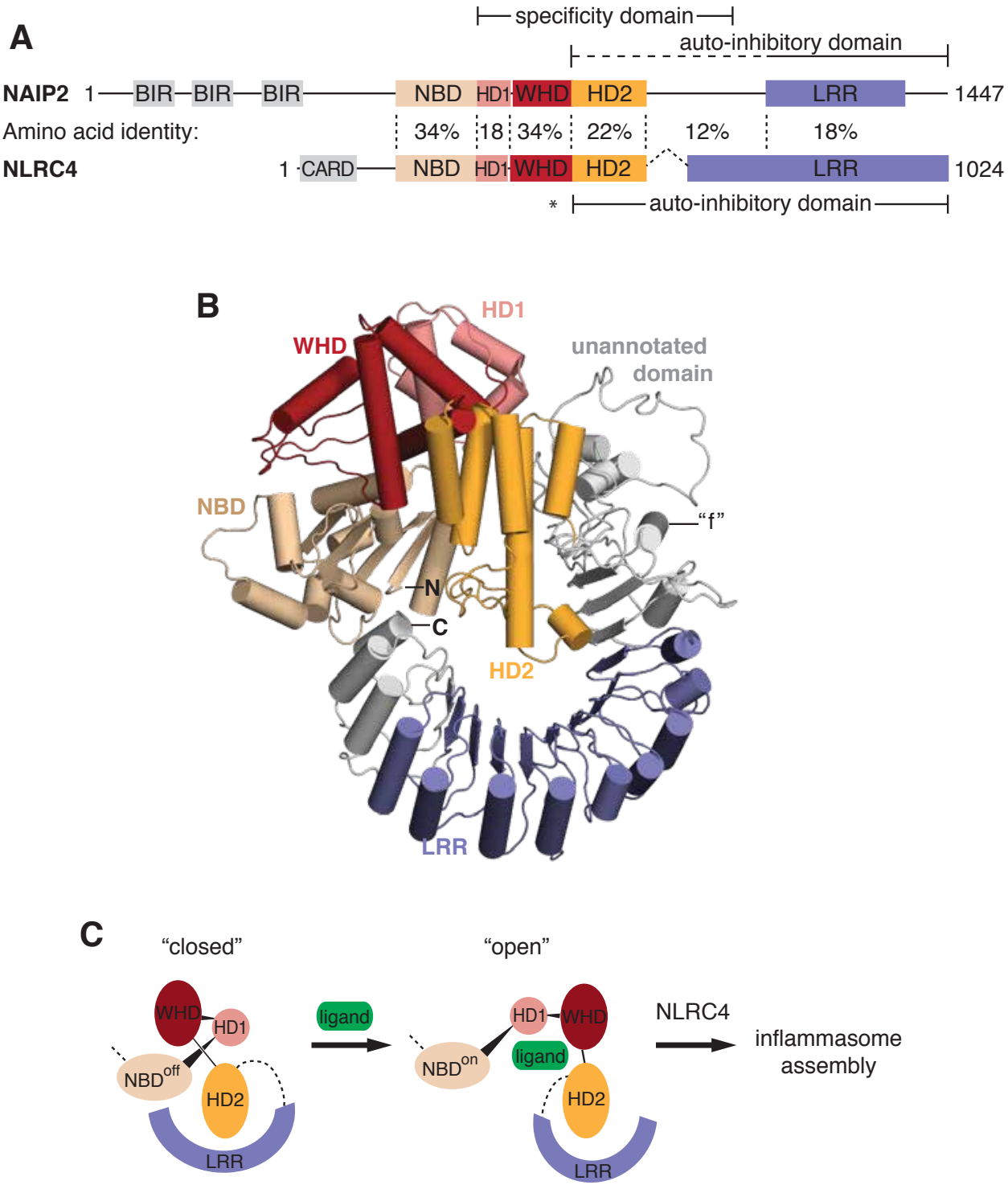


Figure 2.14. Model of NAIP autoinhibition relief by ligand binding
(**A**) Comparison of NLRC4 (PDB 4KXF) and predicted NAIP2 domain architecture. Solid lines indicate the NAIP ligand specificity domain identified in this study and autoinhibitory domains previously identified for NAIP[3] and NLRC4[159]. The dashed line denotes a potential extension of the autoinhibitory domain based on comparison with

NLRC4; the asterisk marks the NLRC4 autoinhibitory residue H443 that is not conserved in NAIPs. Amino acid identity between NAIP2 and NLRC4 is indicated for each segment. See also Figure 2.9. (**B**) Homology model of NAIP2 by multiple template threading using the Phyre2 server. NAIP2 predicted domains are colored as follows: NBD (tan), HD1 (pink), WHD (red), HD2 (orange), and LRR (blue). The N-terminal BIR domains of NAIP2 are not shown. Regions outside of the predicted domains are depicted in gray; predicted structure of the unannotated domain (especially residues 987-1040) is based on low-confidence *ab initio* modeling. (**C**) Model for NAIP activation by ligand binding. Regions of low-confidence structural modeling have been replaced with dashed lines. Ligand, depicted as a green oval, is predicted to bind primarily within the NAIP ligand specificity domain. Ligand binding may sterically occlude autoinhibitory interactions and/or allosterically induce rotation of autoinhibitory domains away from the NBD in a manner similar to apoptosome assembly[184]. Exposure of NBD oligomerization surfaces triggers assembly of the NAIP/NLRC4 inflammasome.

Chapter Three: A Multi-Surface Recognition Strategy Constrains Pathogen Evasion of Innate Immunity

3.1 Prefix

The above results indicate that NAIPs are engaged in an evolutionary arms race with bacterial pathogens for recognition of bacterial ligands. Indeed, there is evidence of positive selection and 'Red Queen' arms races across many, if not all, innate immune receptors from plants to insects to mammals[191-193]. However, it is difficult to understand how innate immune receptors can successfully compete in such a race. Compared to a mammalian host, microbes replicate extremely rapidly to massive population sizes. Thus, in the same time period, bacterial ligands will sample much more of the potential mutational landscape than the corresponding host receptor. As an additional challenge, an innate immune receptor must recognize the same ligand in many different microbial species, each of which is independently evolving and sampling different potential escape variants. A simple one-to-one mutational race is therefore unlikely to yield successful recognition by innate immune receptors over evolutionary time.

To increase their odds of success, pattern recognition receptors must employ additional strategies to constrain microbial evolution. An obvious candidate strategy is to detect conserved features that are difficult, if not impossible, for microbes to alter. For example, the phosphodiester backbone of DNA is universally conserved in all domains of life on Earth. However, that universality requires the innate immune system to employ additional filters to distinguish self from non-self DNA. These include the location of DNA in acidifying phagolysosomes[194,195] or the non-mitotic cytosol[196], or features like unmethylated CpG motifs[197]. Nevertheless, such filters are not perfect, and recognition of host DNA can trigger autoimmune disease[198].

The detection of bacterial proteins, by contrast, can be highly specific to avoid self-recognition but is subject to the challenge of mutable ligands. For example, point mutagenesis resulting in amino acid changes is a major driver in the evolution of bacterial antibiotic resistance[199] and can presumably allow escape of innate immune recognition as well. To minimize this path to escape, innate immune receptors typically bind to conserved sites of the bacterial protein that are critical for its function[118,200]. However, even constrained sites are not perfectly conserved across bacterial phyla, implying that they can tolerate some variation. Here we uncover an additional strategy of innate immune receptors to minimize the odds of bacterial immune evasion through point mutagenesis: the simultaneous recognition of multiple conserved motifs on bacterial proteins. We show that this strategy is widely employed by innate immune receptors of protein ligands, including in the plant and animal kingdoms, to constrain bacterial evasion of innate immune detection.

3.2 Abstract

To initiate responses to infection, the innate immune system deploys germline-encoded receptors that detect conserved pathogen-encoded molecules[17]. Although some innate

immune receptors recognize relatively invariant ligands, such as double-stranded DNA[62,64,65,68], other innate receptors engage more variable ligands. For example, NAIPs (NLR family, apoptosis inhibitory proteins) and TLR5 (Toll-like receptor 5) are receptors for bacterial proteins, including flagellin and structural components of type III secretion systems (T3SS)[3,4,74]. Given that large population sizes and short generation times provide pathogens with the capacity for rapid evolution to evade immune surveillance, it is unclear how germline-encoded innate immune receptors maintain recognition of variable pathogen ligands over evolutionary time. Here we provide experimental evidence that the innate immune system deploys a simple 'multi-surface' recognition strategy that constrains pathogen immune evasion. By employing systematic mutagenesis, we show that NAIP5 binds an N-terminal motif on flagellin, in addition to the previously described C-terminal recognition motif[118]. We found that the N-terminal motif ensures that NAIP5 recognition is robust to point mutations within the C-terminal motif and vice-versa. Simultaneous mutation of both motifs permitted evasion of NAIP5, but flagellin mutants that escaped NAIP5 recognition invariably lost bacterial motility. Similarly, NAIP2 and TLR5 limited pathogen immune evasion by recognition of multiple motifs on their cognate ligands. We propose that multi-surface recognition is a general strategy employed by innate immune receptors to recognize variable protein ligands and constrain pathogen immune evasion.

3.3 Results

Innate immunity restricts pathogen invasion[14] and promotes the induction of an adaptive immune response[15], thereby imparting selective pressure for pathogens to evade innate immune detection. We sought to understand how innate immune receptors such as the NAIP inflammasomes, which detect microbial proteins, are able to limit pathogen immune evasion. Murine NAIP2 detects the inner rod protein of the pathogen-associated T3SS, while the closely related paralogs NAIP5 and NAIP6 recognize bacterial flagellin[3,4]. Upon binding their ligands, NAIPs co-oligomerize with NLRC4 (NLR family, CARD containing protein 4) to induce a pro-inflammatory response that includes the release of the cytokines IL-1β and IL-18 and a lytic form of cell death termed pyroptosis[80].

3.3.1 NAIPs recognize multiple motifs on their cognate ligands

To investigate the strategies employed by NAIPs to maintain recognition of evolvable bacterial proteins, we first conducted a comprehensive analysis of the ligand motifs recognized by NAIPs. A short leucine-rich motif at the C-termini of flagellin and inner rod proteins was reported to be essential for NAIP recognition[118,122], but prior data[124,165] suggested that additional motifs might also exist. We thus performed alanine scanning mutagenesis across the length of PrgJ, the *Salmonella typhimurium* SPI1 T3SS inner rod protein, in three amino acid blocks. We assayed NAIP2 recognition of PrgJ mutants using a retroviral lethality assay[118] in which PrgJ cDNA was transduced into bone marrow-derived macrophages (BMM) using a retroviral vector marked with IRES-GFP. Whereas wild-type PrgJ robustly induced NLRC4-dependent pyroptosis, resulting in essentially no GFP^+ B6 BMM, several PrgJ mutants were wholly or partially incapable of

activating NAIP2, allowing recovery of GFP$^+$ cells (Figure 3.1a). As expected, mutations in the C-terminal seven amino acids of PrgJ (95-97A, 98-99A, 98-100A, and ΔC4) abolished NAIP2 recognition. Importantly, a second motif at residues 32-34 was also required for NAIP2 activation. In addition, the region spanning residues 65-88 consistently appeared to contribute weakly to NAIP2 recognition.

To ascertain whether recognition of multiple motifs is a general feature of NAIP ligand recognition, we also performed alanine scanning mutagenesis of the D0 domain of *Legionella pneumophila* flagellin (FlaA). NLRC4-dependent responses to FlaA in BMM are largely dependent on NAIP5[118] and partly on NAIP6[127]. As expected, mutations in the C-terminus of FlaA, including an L470A/L472A/L473A mutant previously reported to evade NAIP5 recognition[118], abrogated the activation of both NAIP5 and NAIP6 and permitted recovery of GFP$^+$ transductants (Figure 3.1c). In addition, motifs at residues 458-460 and 31-33 were partially required for NAIP5 and NAIP6 activation. The partial requirement for residues 31-33 of FlaA is consistent with reduced recognition of FlaA lacking its N-terminal 65 residues[124] and is particularly striking because of the positional correspondence to the NAIP2 recognition motif at residues 32-34 of PrgJ. The inability of these mutant ligands to activate their cognate NAIPs was not due to their lack of expression (Figure 3.1b, d). We conclude that NAIPs display a common strategy of recognizing multiple discrete motifs on their specific bacterial ligands.

3.3.2 NAIPs bind multiple surfaces on their cognate ligands

We next assessed whether the N- and C-terminal motifs each promote ligand binding to cognate NAIPs. We expressed the N- and C-terminal domains of each ligand as separate polypeptides fused to GFP. The C-terminal 35 amino acids of FlaA (FlaC35) co-immunoprecipitated with NAIP5 in a manner dependent on the C-terminal recognition motif L470/L472/L473 (Figure 3.2a). In contrast to a previous report using a stronger over-expression promoter[165], we did not observe NAIP5 binding to the N-terminal 65 amino acids of FlaA (FlaN65) unless FlaC35 was also bound to NAIP5. The binding of FlaN65 to NAIP5 was dependent on the N-terminal recognition motif R31/L32/S33. Mutations to disrupt a potential coiled-coil interaction between the two helices of the D0 domain[124,140] did not affect NAIP5 binding of FlaN65 (Figure 3.3). Furthermore, we found no evidence that FlaN65 is indirectly recruited to NAIP5 through a NAIP5-independent interaction with FlaC35 (Figure 3.2c). These data suggest that both the N- and C-terminal helices of flagellin are directly and cooperatively bound by NAIP5, even when expressed as separate polypeptides. Importantly, these data provide biochemical support for the structural model we have recently obtained of FlaA bound to the assembled NAIP5/NLRC4 inflammasome (Chapter 4).

A similar pattern was observed for the binding of FlaN65 and FlaC35 to NAIP6 (Figure 3.4). The N- and C-terminal regions of PrgJ were also capable of binding to NAIP2 when expressed as separate polypeptides fused to GFP. The C-terminal 38 amino acids of PrgJ (PrgJC38) bound to NAIP2 only when the last four amino acids were intact, whereas binding of the N-terminal 63 amino acids of PrgJ (PrgJN63) to NAIP2 required

both the N-terminal recognition motif L32/L33 and the binding of PrgJC38 (Figure 3.2b). NAIP2-independent binding of PrgJN63 to PrgJC38 was not responsible for the interaction of PrgJN63 with NAIP2 (Figure 3.2c).

The N- and C-terminal motifs we identified are some of the most conserved regions of PrgJ and FlaA (Figure 3.5) and are even similar between the otherwise relatively divergent rod and flagellin ligands (N: RLS in PrgJ, RLL in FlaA; C: VETLLRS in PrgJ, LSLLGR in FlaA). Despite this similarity, NAIP5 and NAIP6 did not bind to PrgJC38 nor to PrgJN63, even in the presence of FlaC35, while NAIP2 binding to PrgJC38 did not stimulate binding of FlaN65 (Figure 3.6). It is therefore likely that the specificity of NAIPs for their cognate ligands is contributed by additional variable regions outside the motifs. Collectively, these results suggest a direct and specific interaction between NAIPs and multiple surfaces on their cognate ligands.

3.3.3 Multi-surface recognition of flagellin constrains evasion of NAIP5

We reasoned that recognition of multiple ligand surfaces might confer NAIPs with an advantage in an evolutionary arms race with bacteria. If pathogen ligands were only recognized via a single surface, single point mutations that disrupt that recognition surface might be sufficient to allow pathogen escape from recognition. By contrast, these same point mutations might be tolerated in the context of a second intact binding interface, which could tether the ligand to NAIPs and reduce the likelihood of ligand-receptor dissociation (Figure 3.7). To test this hypothesis, we therefore assessed whether multi-surface recognition of ligands by NAIPs was insensitive to single point mutations. Consistent with the above model, mutation of any single FlaA residue within any of the NAIP5 recognition motifs had little to no effect on NAIP5-mediated induction of pyroptosis when FlaA was delivered retrovirally (Figure 3.8). Strikingly, FlaA molecules that harbored mutations in both the N- and C-terminal motifs were generally no longer recognized by NAIP5. NAIP5 binding to FlaC35, harboring only one binding interface, was also more sensitive to point mutations than a ligand with both binding surfaces (FlaN65-GFP-FlaC35) intact (Figure 3.9).

A major concern of the above studies is that overexpression of flagellin might mask a reduced affinity for NAIP5 and thus fail to reveal escape mutants. We therefore assessed the effects of individual and double point mutations in the endogenous copy of *flaA* as expressed from the *L. pneumophila* chromosome under its native promoter during bacterial infection of BMM at low multiplicity of infection (MOI). Importantly, we observed that single alanine substitutions in FlaA did not affect NAIP5 activation during *L. pneumophila* infection, as assessed by the induction of pyroptosis (Figure 3.10). In contrast, simultaneous point mutations in both the N- and C-terminal motifs more severely impacted NAIP5 recognition, particularly the combination of R31A and L470A. The robustness of NAIP5 recognition to single mutations as compared to combinatorial mutants was also observed in NAIP5-dependent restriction of *L. pneumophila* intracellular growth, a sensitive assay for NAIP5 recognition (Figure 3.10b). Intracellular growth of the R31A and L470A mutants was restricted in a NAIP5-dependent manner,

whereas the double R31A+L470A mutant evaded restriction and replicated almost as well as a FlaA-deficient strain.

The effectiveness of multi-surface recognition requires that mutations required to evade immune recognition are severe enough to also incur fitness costs to the pathogen. Critically, although most single point mutations did not affect flagellin function, we observed that the combinatorial mutations that evaded NAIP5 recognition also abrogated flagellar assembly and motility (Figure 3.10c). We also tested whether more disruptive mutations to L470 were sufficient to disrupt NAIP5 recognition. Several non-conservative mutations (L470P, L470N, L470R) reduced NAIP5-mediated pyroptosis in response to FlaA, but these mutations also disrupted motility (Figure 3.11). Taken together, these data suggest that single mutations in flagellin that do not affect its function are insufficient to escape NAIP5 detection, whereas more severe mutations that do escape recognition result in a fitness cost to the bacterium. We therefore propose that multi-surface recognition imposes a severe constraint on the mutational paths pathogens can utilize to escape immune recognition. Intriguingly, our cryo-EM structure of flagellin bound to NAIP5 provides insight into the cause of this evolutionary constraint (Chapter 4). The D0 domain of flagellin is largely disordered in its monomeric state[134,135] but adopts an organized structure in the polymerized flagellum that is critical for flagellum assembly[140,201]. A single protomer of flagellin binds to NAIP5, but the D0 domain nevertheless adopts an ordered conformation that is essentially superimposable on its conformation in the flagellar filament. Thus, we propose that mutations that disrupt the D0 conformation recognized by NAIP5 are also highly likely to disrupt the D0 conformation required to form the flagellar filament.

3.3.4 Multi-surface recognition constrains evasion of NAIP2 and TLR5

To ascertain whether multi-surface recognition also constrains pathogen escape of other innate immune sensors, we next asked whether multi-surface recognition of PrgJ by NAIP2 was similarly robust to single point mutations. Indeed, recognition of PrgJ was unaffected by single mutations in either recognition motif (Figure 3.12a, b). With one exception, single point mutations also did not affect the ability of PrgJ to assemble a functional T3SS capable of invading cells or secreting the SipA effector (Figure 3.12c, d). However, combined mutation of both motifs was sufficient for PrgJ to evade NAIP2 recognition. Many of these mutations also disrupted PrgJ function. Interestingly, unlike with FlaA, we were able to identify several immune escape double mutants of PrgJ that retain their native function, suggesting that the evolutionary constraint of multi-surface recognition is not always complete or that our functional assays were not sufficiently sensitive to detect modest loss of function. Nevertheless, taken together, our data indicate that multi-surface recognition of bacterial ligands by NAIPs limits the evolutionary paths available for bacteria to evade NAIP recognition while retaining ligand function.

Finally, we noted that several other innate immune receptors of protein ligands have been reported to recognize multiple surfaces on their ligands, including plant immune

receptors[57,202] as well as the cell-surface flagellin receptor, TLR5[200,203]. Importantly, single point mutations in flagellin almost always result in only modest decreases in TLR5 recognition, and the few point mutants that disrupt TLR5 recognition also abolish motility[200,204]. The binding of multiple flagellin surfaces has been proposed to contribute to the resiliency of TLR5 recognition[200], but this idea has not been directly tested. We observed that simultaneous point mutation of two TLR5 recognition motifs in *S. typhimurium* flagellin (FliC) dramatically reduced TLR5 activation by secreted FliC, whereas each single point mutation was largely tolerated by TLR5 (Figure 3.13a). Unlike single point mutations that were tolerated for flagellar motility, the combinatorial mutant that evaded TLR5 recognition was unable to assemble a functional flagellum (Figure 3.13b).

3.4 Discussion

It is a well-accepted principle — underlying the combinatorial use of antibiotics and antivirals[199] — that the simultaneous presence of dual selective pressures provides a greater constraint on the evolution of escape mutants than does each selective pressure individually. Our results provide evidence that the innate immune system exploits this general principle and employs a multi-surface recognition strategy to constrain bacterial immune escape. Instead, evasion of innate immunity generally involves more complex mechanisms, including the evolution of compensatory mutations to regain function of immune-evading mutants[205] or the acquisition of alternative virulence factors[122] such as intracellular actin- rather than flagellin-mediated motility[144]. The multi-surface recognition strategy we describe likely serves as a complement to additional strategies that constrain pathogen evolution. For example, targeting the most conserved features on microbial ligands helps to constrain mutagenic escape, although even constrained sites can tolerate some mutations (Figure 3.10c). Additionally, diversifying selection at the interface with microbial ligands, observed for both TLR5[206] and NAIPs[123], can counteract pathogen mutations that disrupt recognition[2]. Thus, we propose multi-surface recognition is one strategy in the arsenal deployed by hosts to counteract the intrinsic advantage held by large populations of rapidly evolving pathogens in their arms race with eukaryotic immune systems.

3.5 Methods

3.5.1 Mice

C57BL/6J (B6) mice were purchased from Jackson laboratories and grown at UC Berkeley. *Nlrc4*$^{-/-}$ mice on a B6 background were from S. Mariathasan and V. Dixit[207]. *Naip5*$^{-/-}$ mice on a B6 background have been described[118]. Animal experiments were approved by the UC Berkeley Animal Care and Use Committee.

3.5.2 Cell culture

HeLa, HEK293T, and GP2-293 (HEK293 expressing retroviral packaging genes *gag* and *pol* [Clontech]) cells were grown in DMEM. A CHO cell line stably expressing human TLR5 and an NFκB luciferase reporter[200] was grown in Ham's F-12 medium with

10 µg/mL blasticidin. BMM were differentiated in RPMI with 5% supernatant from stable macrophage colony stimulating factor (CSF) 3T3 transfectants. All media were supplemented with 10% FBS, 2 mM L-glutamine, 100 U/mL penicillin, and 100 µg/mL streptomycin. HeLa, HEK293T, GP2-293, and CHO cells were subcultured with trypsin digestion. BMM were harvested by chilling in cold PBS followed by gentle scraping.

3.5.3 Retroviral transduction

N-terminally 6myc-tagged *L. pneumophila* FlaA or *S. typhimurium* PrgJ were expressed from the viral LTR of the MSCV2.2 retroviral vector[3]. Triple-alanine mutants were synthesized in MSCV2.2 by Genewiz, whereas single-alanine mutants were generated by Quickchange PCR (Extended Data Table 1). GP2-293 were seeded in 6-well TC-treated plates at a density of 8 x 10^5 cells/well and transfected the following morning with 3.5 µg of MSCV2.2 derivatives and 0.5 µg of VSV-G expression plasmid using Lipofectamine2000 (Invitrogen). After 6 hr, media was replaced with 1mL of BMM media, and cells were incubated overnight at 32 ºC. Virus-containing supernatant was filtered (0.45 µm) and added to newly differentiating BMM, which had been harvested from bones the previous day and then collected in cold PBS, subjected to 2 min of ACK red blood cell lysis, and seeded in non-TC-treated 24-well plates at 2.5 x 10^5 cells/well that morning. Recipient cells were centrifuged (1200 x g, 90 min, 32 ºC) and incubated at 32 ºC overnight. Packaging cells were incubated with a fresh 1 mL of BMM media, and transduction was repeated the following day. After overnight incubation, recipient cells were supplemented with fresh BMM media and incubated 48 hr at 37 ºC. BMM were then harvested and analyzed by flow cytometry for GFP. Gates were set on BMM treated in parallel with media only.

3.5.4 Legionella strains

LP02 is a thymidine auxotroph derivative of *L. pneumophila* strain LP01, and the unmarked deletion of *flaA* in this background has been described[75]. To complement *ΔflaA*, the *flaA* open reading frame and its promoter were cloned into the pSR47S suicide vector; point mutations were then introduced by Quickchange PCR (Extended Data Table 1). These vectors were mated into LP02 via *pir*$^+$ DH5α, and single crossover events to introduce *flaA* onto the chromosome were selected with kanamycin. Strains were streaked for single colonies twice, and integration at the correct locus was confirmed by PCR and sequencing.

3.5.5 Salmonella strains

Flagellin-deficient *S. typhimurium* strains LT2*ΔfliCΔfljAB*, SL1344*ΔfliCΔfljAB* and the isogenic *ΔorgA* strain were a gift of G. Barton. Chromosomal *prgJ* was replaced with mutant alleles following a scarless genome editing protocol[208]. In brief, a mutation cassette containing homology arms of 200 bp 5' and 30 bp 3' of the desired point mutation, and a cassette containing 30 bp 5' and 200 bp 3' of the same point mutation were cloned on either side of the gentamicin resistance cassette in pT2SG using Gibson (NEB) cloning (Extended Data Table 1). SL1344*ΔfliCΔfljAB* was transformed with pSLTS, and λred recombinase expression was induced with L-arabinose. Bacteria were

made electrocompetent and transformed with linear mutation construct amplified out of pT2SG derivatives. Transformants were plated on gentamicin to select for chromosomal insertion of the mutation construct and streaked for single colonies. The mutation construct was transduced into pSLTS-bearing but non-λred-induced SL1344*ΔfliCΔfljAB* using the P22 phage and gentamicin selection. Lysogenic phage was counter-screened using green indicator plates. Mutation-cassette bearing colonies were plated onto anhydrotetracycline to induce I-SceI cleavage between the mutation cassettes. Outgrowing colonies were screened for repair of this lesion to a scarless point mutation via sequencing of the PrgJ locus. Strains were cured of pSLTS via culture at 37 ºC.

3.5.6 Legionella infections

BMM were plated on TC-treated 24- or 96-well plates (for growth curves or cell death assays, respectively) at a density of 5 x 10^5 cells/mL in media lacking antibiotics. *L. pneumophila* was grown from 1 cm^2 patched colonies in BYE (supplemented with 40 mg/mL L-cysteine, 13.5 mg/mL ferric nitrate, and 10 mg/mL thymidine) shaking overnight at 37ºC in 2-fold serial dilutions. Stationary phase cultures ($3.5 < OD_{600} < 4.5$) were added to BMM at MOI = 3 (for cell death assay) or MOI = 0.01 (for growth curves), assuming 1 OD_{600} = 3.3 x 10^8 CFU/mL. Plates were centrifuged 10 min at 500 x g to normalize infection across motile and non-motile strains and then incubated at 37 ºC. Cell death was measured by release of lactate dehydrogenase[209] into culture supernatants after 4 hrs, normalized to BMM lysed for 15 min with 1% Triton-X-100. For growth curves, media was replaced after 1 hr with fresh media supplemented with thymidine (corresponding to 0 hr time point). BMM were lysed in water at the indicated times, and combined lysate and cell supernatant were spot-plated (20 µL) in 10-fold serial dilutions onto BYCE plates. CFU were counted from spots yielding 10-40 CFU.

3.5.7 Salmonella infections

BMM or HeLa cells were plated on TC-treated 96-well plates at 5 x 10^4 cells/well in media lacking antibiotics. Overnight shaking cultures of *S. typhimurium* were diluted 1:40 in fresh LB and grown an additional 3 hr shaking at 37 ºC to induce SPI1 expression. Cultures were added to BMM or HeLa at MOI = 100 (assuming 1 OD_{600} = 1.2 x 10^9 CFU/mL), and plates were centrifuged 10 min at 500 x g. For cell death assays, gentamicin (25 µg/mL) was added after 20 min of incubation at 37 ºC, and LDH release was measured after 1 hr. For cell invasion assays, after 10 min at 37 ºC HeLa cells were washed 3 times with warm PBS and incubated in gentamicin-supplemented media for 1 hr at 37 ºC to kill extracellular bacteria. Cells were then washed with PBS and lysed in 1% Triton. Lysates were spot-plated in 10-fold serial dilutions onto LB plates. CFU were normalized to input samples also spot-plated on LB.

3.5.8 Motility and SPI1 secretion assay

The motility of stationary phase *L. pneumophila* was assessed via microscopy by a blinded observer as described previously[75]. The motility of *S. typhimurium* was assessed by stabbing overnight cultures into 0.4% agar plates and incubating 8 hr at 37 ºC before measuring colony diameter. In parallel, bacterial cultures were washed in PBS

and then vortexed for 3 min to shear flagella. Bacteria were pelleted, and the supernatant was filtered (0.45 µm) and incubated with 10% TCA on ice for 1 hr. Precipitated protein was pelleted at 14,000 x g for 10 min at 4 ºC, washed with cold acetone, dried and resuspended in SDS sample buffer. To assess secretion of SPI1-dependent effectors, SPI-induced *S. typhimurium* culture supernatants were filtered and subjected to TCA precipitation. TCA precipitates were analyzed by SDS-PAGE and Coomassie stain.

3.5.9 Immunoprecipitation

Expression constructs in MSCV2.2 for mouse NLRC4, N-terminally FLAG-tagged NAIPs, and FlaN65-GFP and GFP-FlaC35 have been described previously[3,118,123,124]. PrgJN63 was fused 5' of GFP by SOE PCR with PrgJN63 and GFP fragments followed by conventional cloning into mscv2.2. PrgJC38 was cloned in frame downstream of mCherry in mscv2.2-HA-mCherry-MCS. Point mutations in FlaA or PrgJ were introduced by Quickchange PCR (Extended Data Table 1). HEK293T were seeded in 6-well TC-treated plates at 8 x 10^5 cells/well and transfected the following day using Lipofectamine2000 (Invitrogen) with 500 ng each of the indicated expression constructs, except GFP-tagged ligand domains (250 ng each). Cells were harvested 40-48 hr after transfection and lysed in chilled buffer (50 mM HEPES, 150 mM NaCl, 1% Triton, 10 mM KCl, 5 mM $MgCl_2$, 1x Roche EDTA-free protease inhibitor cocktail, pH 7.6). Lysates were cleared (16,000 x g, 15 min, 4 ºC), 2.5% of lysate was removed as input, and the remaining lysate was incubated 1-2 hr at 4 ºC with 30 µL of magnetic ProteinG beads (BioRad) coated with 1 µg of anti-FLAG antibody (M2, Sigma). Beads were washed 4 times with buffer and eluted with 1 x SDS sample buffer at 70 ºC. Eluates were analyzed by SDS-PAGE and immunoblotting with the following antibodies: anti-NLRC4 (gift of S. Mariathasan and V. Dixit, Genentech), anti-FLAG (M2, Sigma), anti-HA (3F10, Roche), anti-c-myc (9E10, Clontech), anti-GFP (JL8, Clontech), anti-β-actin (C4, Santa Cruz), and secondary anti-mouse or anti-rabbit antibodies conjugated to Alex Fluor-680 (Invitrogen). Immunoblots were analyzed on a Li-Cor imager.

3.5.10 TLR5 stimulation

FliC was cloned from pBBR1-MCS2[75] into pBBR1-MCS5, under an IPTG-inducible promoter, by excision at the MscI and XbaI sites. FliC point mutations were introduced via Quickchange PCR (Extended Data Table 1). FliC expression plasmids were electroporated into LT2*ΔfliCΔfljAB*. Transformants were cultured shaking overnight at 37 ºC in LB containing 1mM IPTG and diluted to 10^8 CFU/mL. Bacteria were pelleted, and culture supernatants were serially diluted and added (20 µL) to CHO hTLR5 reporter cells[200] (seeded the previous day in 96-well TC-treated plates at 2 x 10^4 cells/well) for 6 hr at 37 ºC. Cells were lysed in passive lysis buffer and analyzed for firefly luciferase expression (Promega). To verify that bacterial culture supernatants contained roughly equivalent amounts of FliC, total supernatant was TCA precipitated and analyzed by SDS-PAGE and Coomassie stain.

3.5.11 Alignments

Sequences were curated from UniProt, aligned using Clustal Omega with default settings, and analyzed for sequence conservation using JalView. FlaA and PrgJ sequences were queried using ConSurf with default settings and ConSurf-generated multiple sequences alignments. ConSurf mapped conservation scores onto FliC (PDB 1UCU) or a MODELLER-predicted structure of PrgJ.

Table 3.1. Primers used to generate constructs in Chapter 3

PRIMER	USE	SEQUENCE
Quickchange mutagenesis of FlaA		
oJT286	L12R (F)	CAACACTAATGTGGCGTCGCgCACAGCCCAACGTAATTTGG
oJT287	L12R (R)	CCAAATTACGTTGGGCTGTGcGCGACGCCACATTAGTGTTG
oJT288	L18R (F)	CGCTCACAGCCCAACGTAATcgGGGTGTTTCGGGCAACATG
oJT289	L18R (R)	CATGTTGCCCGAAACACCCcgATTACGTTGGGCTGTGAGCG
oJT290	M25R (F)	GGGTGTTTCGGGCAACATGAgGCAAACATCGATCCAGCG
oJT291	M25R (R)	CGCTGGATCGATGTTTGCcTCATGTTGCCCGAAACACCC
oJT292	L32R (F)	GCAAACATCGATCCAGCGTagATCATCGGGATTAAGGATTAACAGTG
oJT293	L32R (R)	CACTGTTAATCCTTAATCCCGATGATctACGCTGGATCGATGTTTGC
oJT331	I29A (F)	GGCAACATGATGCAAACATCGgcCCAGCGTTTATCATCGGGATTA
oJT332	I29A (R)	TAATCCCGATGATAAACGCTGGgcCGATGTTTGCATCATGTTGCC
oJT333	Q30A (F)	GCAACATGATGCAAACATCGATCgcGCGTTTATCATCGGGATTAAGG
oJT334	Q30A (R)	CCTTAATCCCGATGATAAACGCgcGATCGATGTTTGCATCATGTTGC
oJT335	R31A (F)	CATGATGCAAACATCGATCCAGgcTTTATCATCGGGATTAAGGATTAAC A
oJT336	R31A (R)	TGTTAATCCTTAATCCCGATGATAAAgcCTGGATCGATGTTTGCATCATG
oJT353	L32A (F)	GATGCAAACATCGATCCAGCGTgcATCATCGGGATTAAGGATTAACAG
oJT354	L32A (R)	CTGTTAATCCTTAATCCCGATGATgcACGCTGGATCGATGTTTGCATC
oJT337	S33A (F)	GCAAACATCGATCCAGCGTTTAgCATCGGGATTAAGGATTAACAGTG
oJT338	S33A (R)	CACTGTTAATCCTTAATCCCGATGcTAAACGCTGGATCGATGTTTGC
oJT339	M458A (F)	GCAACAAGCAGGTACAGCGgcGTTGGCACAAGCTAATAGCCTA
oJT340	M458A (R)	TAGGCTATTAGCTTGTGCCAACgcCGCTGTACCTGCTTGTTGC
oJT341	L459A (F)	CAACAAGCAGGTACAGCGATGgcGGCACAAGCTAATAGCCTACC
oJT342	L459A (R)	GGTAGGCTATTAGCTTGTGCCgcCATCGCTGTACCTGCTTGTTG
oJT343	L470A in mscv2.2 only (F)	CTAATAGCCTACCGCAATCTGTAgcATCTTTGTTAGGTCGATAGTCGAC
oJT344	L470A in mscv2.2 only (R)	GTCGACTATCGACCTAACAAAGATgcTACAGATTGCGGTAGGCTATTAG
oJT345	S471A in mscv2.2 only (F)	CTAATAGCCTACCGCAATCTGTATTAgCTTTGTTAGGTCGATAGTCGAC

oJT346	S471A in mscv2.2 only (R)	GTCGACTATCGACCTAACAAAGcTAATACAGATTGCGGTAGGCTATTAG
oJT347	L472A in mscv2.2 only (F)	GCCTACCGCAATCTGTATTATCTgcGTTAGGTCGATAGTCGACGG
oJT348	L472A in mscv2.2 only (R)	CCGTCGACTATCGACCTAACgcAGATAATACAGATTGCGGTAGGC
oJT349	L473A in mscv2.2 only (F)	CCTACCGCAATCTGTATTATCTTTGgcAGGTCGATAGTCGACGGTATC
oJT350	L473A in mscv2.2 only (R)	GATACCGTCGACTATCGACCTgcCAAAGATAATACAGATTGCGGTAGG
oJT351	R475A in mscv2.2 only (F)	CGCAATCTGTATTATCTTTGTTAGGTgcATAGTCGACGGTATCGATAAGC
oJT352	R475A in mscv2.2 only (R)	GCTTATCGATACCGTCGACTATgcACCTAACAAAGATAATACAGATTGCG
oJT357	L470A in pSR47S only (F)	CTAATAGCCTACCGCAATCTGTAgcATCTTTGTTAGGTCGATAGGCG
oJT358	L470A in pSR47S only (R)	CGCCTATCGACCTAACAAAGATgcTACAGATTGCGGTAGGCTATTAG
oJT359	S471A in pSR47S only (F)	CTAATAGCCTACCGCAATCTGTATTAgCTTTGTTAGGTCGATAGGCGG
oJT360	S471A in pSR47S only (R)	CCGCCTATCGACCTAACAAAGcTAATACAGATTGCGGTAGGCTATTAG
oJT361	L472A in pSR47S only (F)	GCCTACCGCAATCTGTATTATCTgcGTTAGGTCGATAGGCGGCC
oJT362	L472A in pSR47S only (R)	GGCCGCCTATCGACCTAACgcAGATAATACAGATTGCGGTAGGC
oJT363	L473A in pSR47S only (F)	CCTACCGCAATCTGTATTATCTTTGgcAGGTCGATAGGCGGCCG
oJT364	L473A in pSR47S only (R)	CGGCCGCCTATCGACCTgcCAAAGATAATACAGATTGCGGTAGG
oJT365	R475A in pSR47S only (F)	CGCAATCTGTATTATCTTTGTTAGGTgcATAGGCGGCCGCCACCG
oJT366	R475A in pSR47S only (R)	CGGTGGCGGCCGCCTATgcACCTAACAAAGATAATACAGATTGCG
oJT390	L470G in mscv2.2 only (F)	CTAATAGCCTACCGCAATCTGTAggATCTTTGTTAGGTCGATAGTCGAC
oJT391	L470G in mscv2.2 only (R)	GTCGACTATCGACCTAACAAAGATccTACAGATTGCGGTAGGCTATTA G
oJT392	L470S in mscv2.2 only (F)	CTAATAGCCTACCGCAATCTGTATcATCTTTGTTAGGTCGATAGTCGAC
oJT393	L470S in mscv2.2 only (R)	GTCGACTATCGACCTAACAAAGATgATACAGATTGCGGTAGGCTATTA G
oJT394	L470C in mscv2.2 only (F)	CTAATAGCCTACCGCAATCTGTATgcTCTTTGTTAGGTCGATAGTCGAC
oJT395	L470C in mscv2.2 only (R)	GTCGACTATCGACCTAACAAAGAgcATACAGATTGCGGTAGGCTATTAG
oJT396	L470P in mscv2.2 only (F)	CTAATAGCCTACCGCAATCTGTAccATCTTTGTTAGGTCGATAGTCGAC
oJT397	L470P in mscv2.2 only (R)	GTCGACTATCGACCTAACAAAGATggTACAGATTGCGGTAGGCTATTAG
oJT398	L470F in mscv2.2 only (F)	CTAATAGCCTACCGCAATCTGTATTcTCTTTGTTAGGTCGATAGTCGAC
oJT399	L470F in mscv2.2 only (R)	GTCGACTATCGACCTAACAAAGAgAATACAGATTGCGGTAGGCTATTAG
oJT400	L470D in mscv2.2 only (F)	CTAATAGCCTACCGCAATCTGTAgatTCTTTGTTAGGTCGATAGTCGAC
oJT401	L470D in mscv2.2 only (R)	GTCGACTATCGACCTAACAAAGAatcTACAGATTGCGGTAGGCTATTAG
oJT402	L470N in mscv2.2 only (F)	CTAATAGCCTACCGCAATCTGTAaatTCTTTGTTAGGTCGATAGTCGAC
oJT403	L470N in mscv2.2 only (R)	GTCGACTATCGACCTAACAAAGAattTACAGATTGCGGTAGGCTATTAG
oJT404	L470R in mscv2.2 only (F)	CTAATAGCCTACCGCAATCTGTAcggTCTTTGTTAGGTCGATAGTCGAC
oJT405	L470R in mscv2.2 only (R)	GTCGACTATCGACCTAACAAAGAccgTACAGATTGCGGTAGGCTATTAG

oJT426	L470G in pSR47S only (F)	CTAATAGCCTACCGCAATCTGTAggATCTTTGTTAGGTCGATAGGCG
oJT427	L470G in pSR47S only (R)	CGCCTATCGACCTAACAAAGATccTACAGATTGCGGTAGGCTATTAG
oJT428	L470P in pSR47S only (F)	CTAATAGCCTACCGCAATCTGTAccATCTTTGTTAGGTCGATAGGCG
oJT429	L470P in pSR47S only (R)	CGCCTATCGACCTAACAAAGATggTACAGATTGCGGTAGGCTATTAG
oJT430	L470D in pSR47S only (F)	CTAATAGCCTACCGCAATCTGTAgatTCTTTGTTAGGTCGATAGGCGGC
oJT431	L470D in pSR47S only (R)	GCCGCCTATCGACCTAACAAAGAatcTACAGATTGCGGTAGGCTATTAG
oJT432	L470N in pSR47S only (F)	CTAATAGCCTACCGCAATCTGTAaatTCTTTGTTAGGTCGATAGGCGGC
oJT433	L470N in pSR47S only (R)	GCCGCCTATCGACCTAACAAAGAattTACAGATTGCGGTAGGCTATTAG
oJT434	L470R in pSR47S only (F)	CTAATAGCCTACCGCAATCTGTAcggTCTTTGTTAGGTCGATAGGCGGC
oJT435	L470R in pSR47S only (R)	GCCGCCTATCGACCTAACAAAGAccgTACAGATTGCGGTAGGCTATTAG
oJT280	R31A/L32A/S33A (F)	CATGATGCAAACATCGATCCAGgcTgcAgCATCGGGATTAAGGATTAACAGTG
oJT281	R31A/L32A/S33A (R)	CACTGTTAATCCTTAATCCCGATGcTgcAgcCTGGATCGATGTTTGCATCATG
oJT184	L470/L472/L473A in mscv2.2 only (F)	GCTAATAGCCTACCGCAATCTGTAgcATCTgcGgcAGGTCGATAGTCGACGGTATCG
oJT185	L470/L472/L473A in mscv2.2 only (R)	CGATACCGTCGACTATCGACCTgcCgcAGATgcTACAGATTGCGGTAGGCTATTAGC
oJT375	L470A/L472A/L473A in pSR47S only (F)	GCTAATAGCCTACCGCAATCTGTAgcATCTgcGgcAGGTCGATAGGCGGCCGC
oJT376	L470A/L472A/L473A in pSR47S only (R)	GCGGCCGCCTATCGACCTgcCgcAGATgcTACAGATTGCGGTAGGCTATTAGC
Quickchange mutagenesis of PrgJ		
oEA01	R32A (F)	GACATTGTCTCGCTGGATGACgcGCTACTCCAGGCTTTTTCTGG
oEA02	R32A (R)	CCAGAAAAAGCCTGGAGTAGCgcGTCATCCAGCGAGACAATGTC
oEA03	L33A (F)	GTCTCGCTGGATGACCGGgcACTCCAGGCTTTTTCTGGTTCG
oEA04	L33A (R)	CGAACCAGAAAAAGCCTGGAGTgcCCGGTCATCCAGCGAGAC
oEA05	L34A (F)	CTCGCTGGATGACCGGCTAgcCCAGGCTTTTTCTGGTTCGG
oEA06	L34A (R)	CCGAACCAGAAAAAGCCTGGgcTAGCCGGTCATCCAGCGAG
oJT313	Q35A (F)	GCTGGATGACCGGCTACTCgcGGCTTTTTCTGGTTCGGCGAT
oJT314	Q35A (R)	ATCGCCGAACCAGAAAAAGCCgcGAGTAGCCGGTCATCCAGC
oJT315	F37A (F)	GACCGGCTACTCCAGGCTgcTTCTGGTTCGGCGATTGCCAC
oJT316	F37A (R)	GTGGCAATCGCCGAACCAGAAgcAGCCTGGAGTAGCCGGTC
oEA07	V95A (F)	CGTAAAGGAGTCGGGGCTGcTGAAACGCTATTACGCTCATGAG
oEA08	V95A (R)	CTCATGAGCGTAATAGCGTTTCAgCAGCCCCGACTCCTTTACG
oEA09	E96A (F)	GTAAAGGAGTCGGGGCTGTTGcAACGCTATTACGCTCATGAGTCG

oEA10	E96A (R)	CGACTCATGAGCGTAATAGCGTTgCAACAGCCCCGACTCCTTTAC
oEA11	T97A (F)	GTAAAGGAGTCGGGGCTGTTGAAgCGCTATTACGCTCATGAGTCG
oEA12	T97A (R)	CGACTCATGAGCGTAATAGCGcTTCAACAGCCCCGACTCCTTTAC
oEA13	L98A (F)	GAGTCGGGGCTGTTGAAACGgcATTACGCTCATGAGTCGACGG
oEA14	L98A (R)	CCGTCGACTCATGAGCGTAATgcCGTTTCAACAGCCCCGACTC
oEA15	L99A (F)	TCGGGGCTGTTGAAACGCTAgcACGCTCATGAGTCGACGGTATC
oEA16	L99A (R)	GATACCGTCGACTCATGAGCGTgcTAGCGTTTCAACAGCCCCGA
oEA17	R100A (F)	CGGGGCTGTTGAAACGCTATTAgcCTCATGAGTCGACGGTATCG
oEA18	R100A (R)	CGATACCGTCGACTCATGAGgcTAATAGCGTTTCAACAGCCCCG
oEA19	S101A (F)	GGCTGTTGAAACGCTATTACGCgCATGAGTCGACGGTATCGATAAG
oEA20	S101A (R)	CTTATCGATACCGTCGACTCATGcGCGTAATAGCGTTTCAACAGCC
oJT317	L32A/L33A (F)	GTCTCGCTGGATGACCGGgcAgcCCAGGCTTTTTCTGGTTCGGC
oJT318	L32A/L33A (R)	GCCGAACCAGAAAAAGCCTGGgcTgcCCGGTCATCCAGCGAGAC
Quickchange mutagenesis of FliC		
oJT450	L416A (F)	GCTGCAGAAAATTGATGCTGCTgcGGCACAGGTTGACACGTTACG
oJT451	L416 (R)	CGTAACGTGTCAACCTGTGCCgcAGCAGCATCAATTTTCTGCAGC
oJT452	L88A (F)	CGCTGAACGAAATCAACAACAACgcGCAGCGTGTGCGTGAACTG
oJT453	L88A (R)	CAGTTCACGCACACGCTGCgcGTTGTTGTTGATTTCGTTCAGCG
Quickchange modification of vectors		
oJT281	Add N-terminal HA tag to mCherry (F)	GGAATTAGATCCgccaccATGtacccatacgatgttccagattacgctGTGAGCAAGG GCGAGGAG
oJT282	Add N-terminal HA tag to mCherry (R)	CTCCTCGCCCTTGCTCACagcgtaatctggaacatcgtatgggtaCATggtggcGGAT CTAATTCC
oJT355	Add N-terminal FLAG tag to GFP (F)	CGGAATTAGATCCgccaccATGgattacaaggacgacgatgacaagGTGAGCAAGG GCGAGGAG
oJT356	Add N-terminal FLAG tag to GFP (R)	CTCCTCGCCCTTGCTCACcttgtcatcgtcgtccttgtaatcCATggtggcGGATCTAAT TCCG
Conventional cloning of FlaA into pSR47A		
oJT327	F primer (BamHI), 900 bp 5' of lpg1340 start codon	caaGGATCCccgctggcagtaaaacaaatgtatgg
oJT157	R primer (NotI), at lpg1340 stop codon	ggttgcggccgcCTATCGACCTAACAAAGATAATACAGATTGC
Generation of pT2SG		
oNL033.1	Amplify GentR cassette (F, pT2S 5' backbone overhang for Gibson reaction)	GAGGAAGAATTGTGAAACTATCACTAATGTTACGCAGCAGCAACGA

oNL034.2	Amplify GentR cassette (R, pT2S 3' backbone overhang for Gibson reaction)	TTACGCCCCGCCCTGCTTAGGTGGCGGTACTTGGGTC
oNL035	Amplify T2S backbone for Gibson assembly w/ GentR cassette (F)	TAGTGATAGTTTCACAATTCTTCCTCAG
oNL036	Amplify T2S backbone for Gibson assembly w/ GentR cassette (F)	GCAGGGCGGGGCGTAA
Gibson cloning of PrgJ mutation cassettes into pT2SG		
oNL017	Amplify 5' mutation cassettes for 33-37A from *S.t.* gDNA (F); HA3 5' overhang for Gibson assembly w/ pT2SG	aggcgtatcacgaggcccttgctctcggaatataacttgt
oNL021	Amplify 5' mutation cassette for L33A (R); SM5 3' overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatgaaccagaAAAAGCCTGGAGTgcCCGgtcatccagcgagacaatg
oNL023	Amplify 5' mutation cassette for L34A (R); SM5 3' overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatcgaaccagaAAAAGCCTGGgcTAGCCGgtcatccagcgagac
oNL025	Amplify 5' mutation cassette for Q35A (R); SM5 3' overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatcaatcgccgaaccagaAAAAGCCgcGAGTAGCCGgtcatccagcgag
oNL027	Amplify 5' mutation cassette for F37A (R); SM5 3' overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatggcaatcgccgaaccagaAgcAGCCTGGAGTAGCCGgtcatcc
oNL022	Amplify 3' mutation cassette for L33A (F); SM3 3' overhang for Gibson assembly w/ pT2SG	gcagggcggggcgtaacattgtctcgctggatgacCGGgcACTCCAGGCTTTTtctggttc
oNL024	Amplify 3' mutation cassette for L34A (F); SM3 3' overhang for Gibson assembly w/ pT2SG	gcagggcggggcgtaagtctcgctggatgacCGGCTAgcCCAGGCTTTTtctggttcg
oNL026	Amplify 3' mutation cassette for Q35A (F); SM3 3' overhang for Gibson assembly w/ pT2SG	gcagggcggggcgtaactcgctggatgacCGGCTACTCgcGGCTTTTtctggttcggcgattg
oNL028	Amplify 3' mutation cassette for F37A (F); SM3 3' overhang for Gibson assembly w/ pT2SG	gcagggcggggcgtaaggatgacCGGCTACTCCAGGCTgcTtctggttcggcgattgcc
oNL018	Amplify 3' mutation cassettes for 32-37A from *S.t.* gDNA (R); HA5 3' overhang for Gibson assembly w/ pT2SG	ctcacatgttctttcctgcgagtatatagatatcgacgaa

oNL001	Amplify 5’ mutation cassettes for 95A-ΔC4 from *S.t.* gDNA (F); HA3 5’ overhang for Gibson assembly w/ pT2SG	aggcgtatcacgaggcccttcggctgtggataaacagacg
oNL003	Amplify 5' mutation cassette for V95A (R); SM5 3’ overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatatcatgaGCGTAATAGCGTTTCAgCagccccgactcctttac
oNL005	Amplify 5' mutation cassette for E96A (R); SM5 3’ overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatcgaatcaTGAGCGTAATAGCGTTGCAACagccccgactcctta
oNL007	Amplify 5' mutation cassette for T97A (R); SM5 3’ overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatgaatcaTGAGCGTAATAGCGcTTCAACagccccgactcc
oNL009	Amplify 5' mutation cassette for L98A (R); SM5 3’ overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatcgacgaatcaTGAGCGTAATgcCGTTTCAACagccccgactc
oNL011	Amplify 5' mutation cassette for L99A (R); SM5 3’ overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatgatatcgacgaatcaTGAGCGTgcTAGCGTTTCAACagccccgac
oNL013	Amplify 5' mutation cassette for R100A (R); SM5 3’ overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatagatatcgacgaatcaTGAGgcTAATAGCGTTTCAACagcc
oNL015	Amplify 5' mutation cassette for S101A (R); SM5 3’ overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatgtatatagatatcgacgaatcaTGcGCGTAATAGCGTTTCAACagcc
oNL033	Amplify 5' mutation cassette for L98stop (R); SM5 3’ overhang for Gibson assembly w/ pT2SG	gcagggcggggcgtaagagtcggggctGTTGAAACGtgaTTACGCTCAtgattcgtcg
oNL004	Amplify 3' mutation cassette for V95A (F); SM3 3’ overhang for Gibson assembly w/ pT2SG	gcagggcggggcgtaagtaaaggagtcggggctGcTGAAACGCTATTACGCtcatgat
oNL006	Amplify 3' mutation cassette for E96A (F); SM3 3’ overhang for Gibson assembly w/ pT2SG	gcagggcggggcgtaataaaggagtcggggctGTTGCAACGCTATTACGCTCAtgattcg
oNL008	Amplify 3' mutation cassette for T97A (F); SM3 3’ overhang for Gibson assembly w/ pT2SG	gcagggcggggcgtaaggagtcggggctGTTGAAgCGCTATTACGCTCAtgattc
oNL010	Amplify 3' mutation cassette for L98A (F); SM3 3’ overhang for Gibson	gcagggcggggcgtaagagtcggggctGTTGAAACGgcATTACGCTCAtgattcgtcg

	assembly w/ pT2SG	
oNL012	Amplify 3' mutation cassette for L99A (F); SM3 3' overhang for Gibson assembly w/ pT2SG	gcagggcggggcgtaagtcggggctGTTGAAACGCTAgcACGCTCAtgattcgtcgatatc
oNL014	Amplify 3' mutation cassette for R100A (F); SM3 3' overhang for Gibson assembly w/ pT2SG	gcagggcggggcgtaaggctGTTGAAACGCTATTAgcCTCAtgattcgtcgatatct
oNL016	Amplify 3' mutation cassette for S101A (F); SM3 3' overhang for Gibson assembly w/ pT2SG	gcagggcggggcgtaaggctGTTGAAACGCTATTACGCgCAtgattcgtcgatatctatatac
oNL034	Amplify 3' mutation cassette for L98stop (F); SM3 3' overhang for Gibson assembly w/ pT2SG	accgctgccactcttgagatcgacgaatcaTGAGCGTAAtcaCGTTTCAACagccccgactc
oNL002	Amplify 3' mutation cassettes for 95A-ΔC4 from *S.t.* gDNA (R); HA5 3' overhang for Gibson assembly w/ pT2SG	ctcacatgttctttcctgcgggtaatgctatagcccaattttcc
SM5	Amplify pT2SG selection cassette for Gibson assembly (F)	ATCTCAAGAGTGGCAGC
SM3	Amplify pT2SG selection cassette for Gibson assembly (R)	TTACGCCCCGCCCTGC
HA5	Amplify pT2SG backbone for Gibson assembly (F)	CGCAGGAAAGAACATGTG
HA3	Amplify pT2SG backbone for Gibson assembly (R)	AAGGGCCTCGTGATACG
oNL031	Amplify linear mutation construct for 33-37A (F)	ggtggattatgtcgattgcaac
oNL032	Amplify linear mutation construct for 33-37A (R)	ctcttgcgaaatagccagctc
oNL029	Amplify linear mutation construct for 95A-ΔC4 (F)	ggctatttcgcaagagatgatttc
oNL030	Amplify linear mutation construct for 95A-ΔC4 (R)	gacctcattagcctgttcctg
Cloning FlaA or PrgJ domain fusions into mscv2.2		
oJT301	Clone PrgJN63 5' in-frame w/ GFP (F, BglII, Kozak); also external primer for SOE PCR	ccaagatcTgccaccATGTCGATTGCAACTATTGTCCC
oJT302	Clone PrgJN63 5' in frame w/ GFP (R, GFP 3' overhang for SOE PCR)	CCTCGCCCTTGCTCACcatATCCGTCACCAGATTAGGGTC

oJT303	Clone GFP 3’ in frame w/ PrgJN63 (F, PrgJN63 5’ overhang for SOE PCR)	GACCCTAATCTGGTGACGGATatgGTGAGCAAGGGCGAGG
oJT304	Clone GFP 3' in frame w/ PrgJN63 (R, NotI); also external primer for SOE PCR	gttgcggccgcTTACTTGTACAGCTCGTCCATGCC
oJT283	Clone full-length PrgJ 3’ in-frame w/ mCherry (F, BglII)	ccaAGATCTTCGATTGCAACTATTGTCCCTG
oJT284	Clone PrgJC38 3’ in-frame w/ mCherry (F, BglII)	ccaAGATCTCCTAAAGAGCTGGCTATTTCG
oJT113	Clone full-length or C38 of PrgJ 3’ in frame w/ mCherry (R, SalI)	GGTGgtcgacTCATGAGCGTAATAGCGTTTCAAC
oJT207	Clone PrgJC38ΔC4 3’ in frame w/ mCherry (R, SalI)	gtatgtcgaCTACGTTTCAACAGCCCCG
oJT308	Clone PrgJN63-GFP 5’ in frame w/ PrgJC38 (R, PrgJC38 3’ overhang for SOE PCR, use w/ oJT301)	GCGAAATAGCCAGCTCTTTAGGaccCTTGTACAGCTCGTCCATGCC
oJT309	Clone PrgJC38 3' in frame w/ PrgJN63-GFP (F, GFP 5’ overhang for SOE PCR)	GGCATGGACGAGCTGTACAAGggtCCTAAAGAGCTGGCTATTTCGC
oJT310	Clone PrgJC38 3' in frame w/ PrgJN63-GFP (R, NotI); also external primer for SOE PCR	ggttgcggccgcTCATGAGCGTAATAGCGTTTCAAC
oJT305	Clone FlaN65 5’ in frame w/ GFP-FlaC35 (F, BglII, Kozak); also external primer for SOE PCR	cgacagatctccaccATGGCTCAAGTAATCAACAC
oJT306	Clone FlaN65 5' in frame w/ GFP-FlaC35 (R, GFP 3’ overhang for SOE PCR)	CCTCGCCCTTGCTCACcatAACGGCTTGGTTCATCCCGC
oJT307	Clone GFP-FlaC35 3’ in frame w/ FlaN65 (F, FlaN65 5’ overhang for SOE PCR	GCGGGATGAACCAAGCCGTTatgGTGAGCAAGGGCGAGG
oJT157	Clone GFP-FlaC35 3' in frame w/ FlaN65 (R, NotI); also external primer for SOE PCR	ggttgcggccgcCTATCGACCTAACAAAGATAATACAGATTGC
Sequencing primers		
mscvF	Sequence mscv2.2 MCS (F)	AAGCCCTTTGTACACCCTAAGCC
mscvR	Sequence mscv2.2 MCS (R)	CCTCACATTGCCAAAAGAC
mscv-noIRES-R	Sequence MCS (R) in GFP/mCherry fusion mscv2.2	ATTTTATCGAATTCGATATCAAGCT
SR47-F	Sequence pSR47S MCS (F)	TGAACGGCAGGTATATGTG
pBBR1	Sequence pBBR1 MCS (F)	ctcagcttcctttcgggc

M13R	Sequence pSR47S, pBBR1 MCS (R)	AACAGCTATGACCATG
oJT386	Verify correct positional insertion of flaA (F); 910 bp 5' of lpg1340 codon (5' of targeting homology arm)	ccctcctgcatgcgagc
oJT387	Verify correct positional insertion of flaA (R); 160 bp 3' of lpg1340 start codon	CCCGCGAATTTGTGCAGTC
oJT388	Sequence *flaA* (F), 40 bp 5' of lpg1340 start codon	cccataccagggattcaggt
oJT389	Sequence *flaA* in single crossover *L. pneumophila* integrants (R, vector backbone 50 bp 3' of MCS)	cgaagtgcagttgttgctgc
oNL017	Sequence *prgJ* in *S. typhimurium* (F, 110 bp 5' of start codon, 5' of linear integration construct homology arm)	aggcgtatcacgaggcccttgctctcggaatataacttgt
oNL002	Sequence *prgJ* in *S. typhimurium* (R, 170 bp 3' of stop codon, 3' of linear integration construct homology arm)	ctcacatgttctttcctgcgggtaatgctatagcccaattttcc

3.6 Acknowledgments

I thank Elise Adamson and Ella Hartenian for their help with designing and testing alanine scanning libraries. I thank Nick Lind for his help constructing *Salmonella* strains. I thank Kelly Smith for the contribution of a CHO cell line stably expressing HsTLR5 and NFκB luciferase reporters, Greg Barton for *S. typhimurium* strains, Sky Brubaker and Denise Monack for the P22 transducing phage, and Heran Darwin for advice on analysis of flagellin expression.

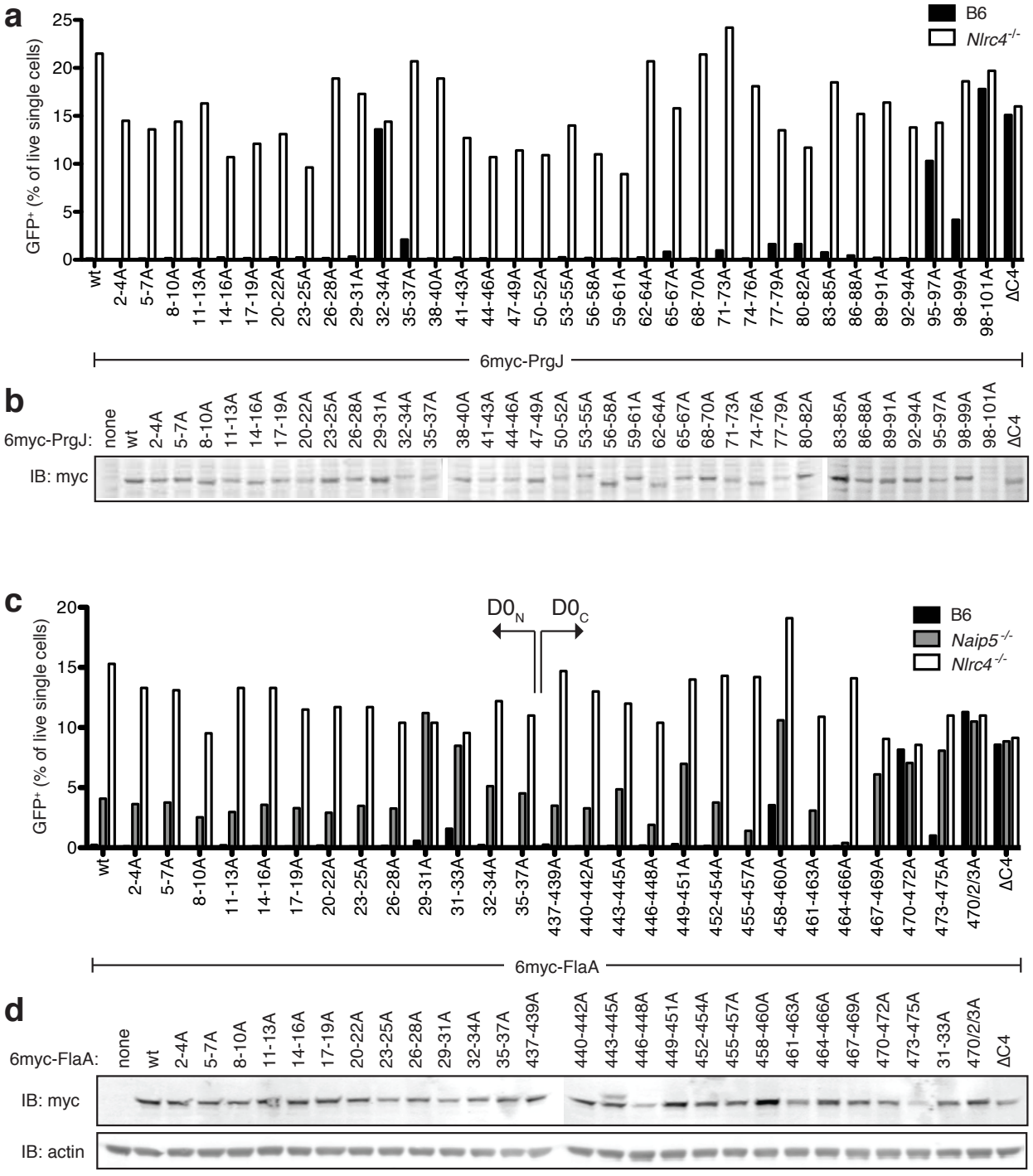


Figure 3.1. NAIP activation requires multiple ligand motifs

a, 6myc-tagged full-length *S. typhimurium* PrgJ, or variants with the indicated residues mutated to alanine (A), were transduced into BMM using a retrovirus marked with IRES-GFP. Transduction efficiency was assessed by flow cytometry for GFP expression at 4 days post transduction. Failure to transduce B6 BMM, as compared to *Nlrc4*$^{-/-}$ BMM, is indicative of NAIP2 activation. **b,** Constructs were transfected into HEK293T, and

lysates were probed for PrgJ expression by anti-myc immunoblot (IB). **c,** 6myc-tagged full-length *L. pneumophila* FlaA, or variants with the indicated residues mutated to alanine, were transduced into BMM as in **a**. *Naip5*$^{-/-}$ BMM responses to FlaA are NAIP6-dependent. **d,** Transduced *Nlrc4*$^{-/-}$ BMM lysates were probed for FlaA expression by anti-myc IB. Results representative of at least 2 independent experiments.

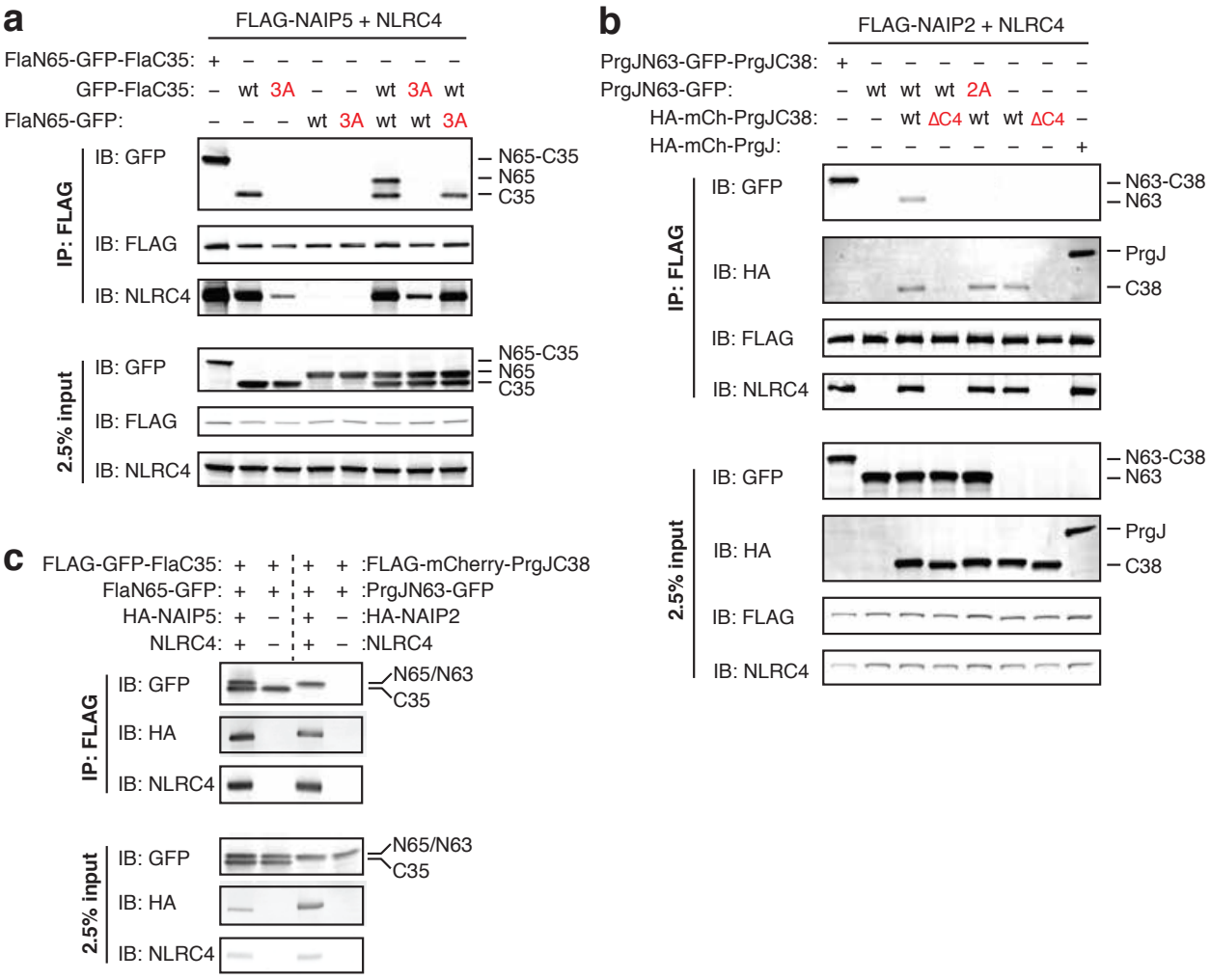


Figure 3.2. N- and C-terminal halves of ligands bind cooperatively to NAIPs
The indicated constructs were transfected into HEK293T, and lysates were subjected to anti-FLAG immunoprecipitation (IP). **a,** NAIP5 binds to the N- and C-terminal halves of the FlaA D0 domain. The 3A mutations in FlaN65 or FlaC35 are R31A/L32A/S33A and L470A/L472A/L473A, respectively. **b,** NAIP2 binds to the N- and C-terminal halves of PrgJ. The ΔC4 mutation of HA-mCherry-PrgJC38 removes the last 4 amino acids of PrgJ. PrgJN63-GFP-2A contains the mutations L33A and L34A. **c,** The N-terminal half of PrgJ or the FlaA D0 domain does not bind to the corresponding C-terminal half in the absence of NAIPs. Results representative of at least 3 independent experiments.

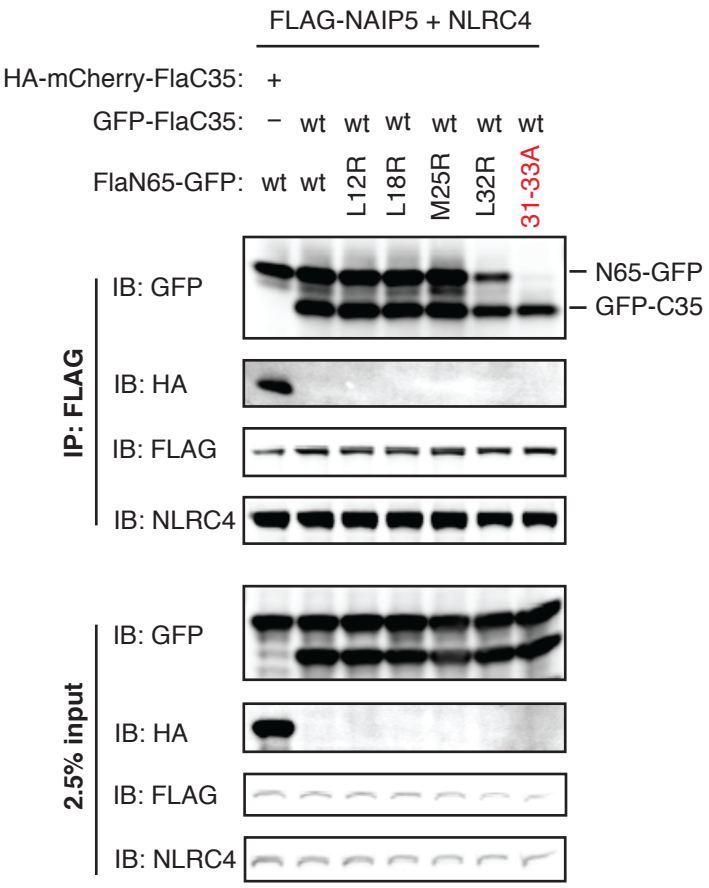


Figure 3.3. Binding of FlaN65 to NAIP5 is not mediated by interactions between FlaN65 and FlaC35

The indicated constructs were transfected into HEK293T and subjected to anti-FLAG IP as in Figure 2. Monomeric Cherry (mCherry) is incapable of mediating dimerization with GFP. The indicated leucine (L) to arginine (R) mutations are predicted to disrupt coiled-coil interactions between the helices of the D0 domain in the flagellar filament. The L32R mutation that reduces FlaN65 binding to NAIP5 is within the N-terminal NAIP5 recognition motif. Results representative of at least 3 independent experiments.

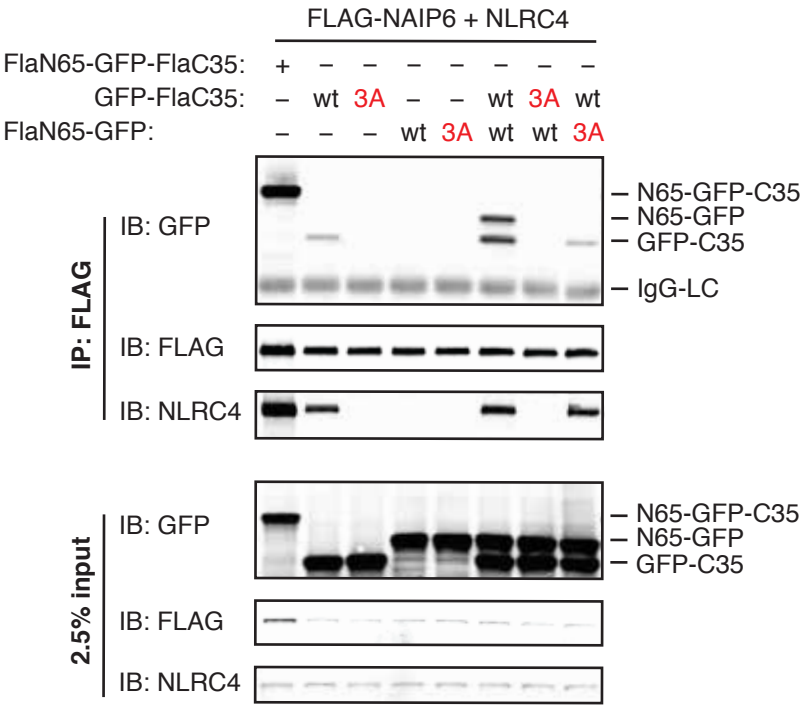


Figure 3.4. FlaA D0 halves bind cooperatively to NAIP6
The indicated constructs were transfected into HEK293T and subjected to anti-FLAG IP. The 3A mutations in FlaN65 or FlaC35 are R31A/L32A/S33A and L470A/L472A/L473A, respectively. Results representative of at least 3 independent experiments.

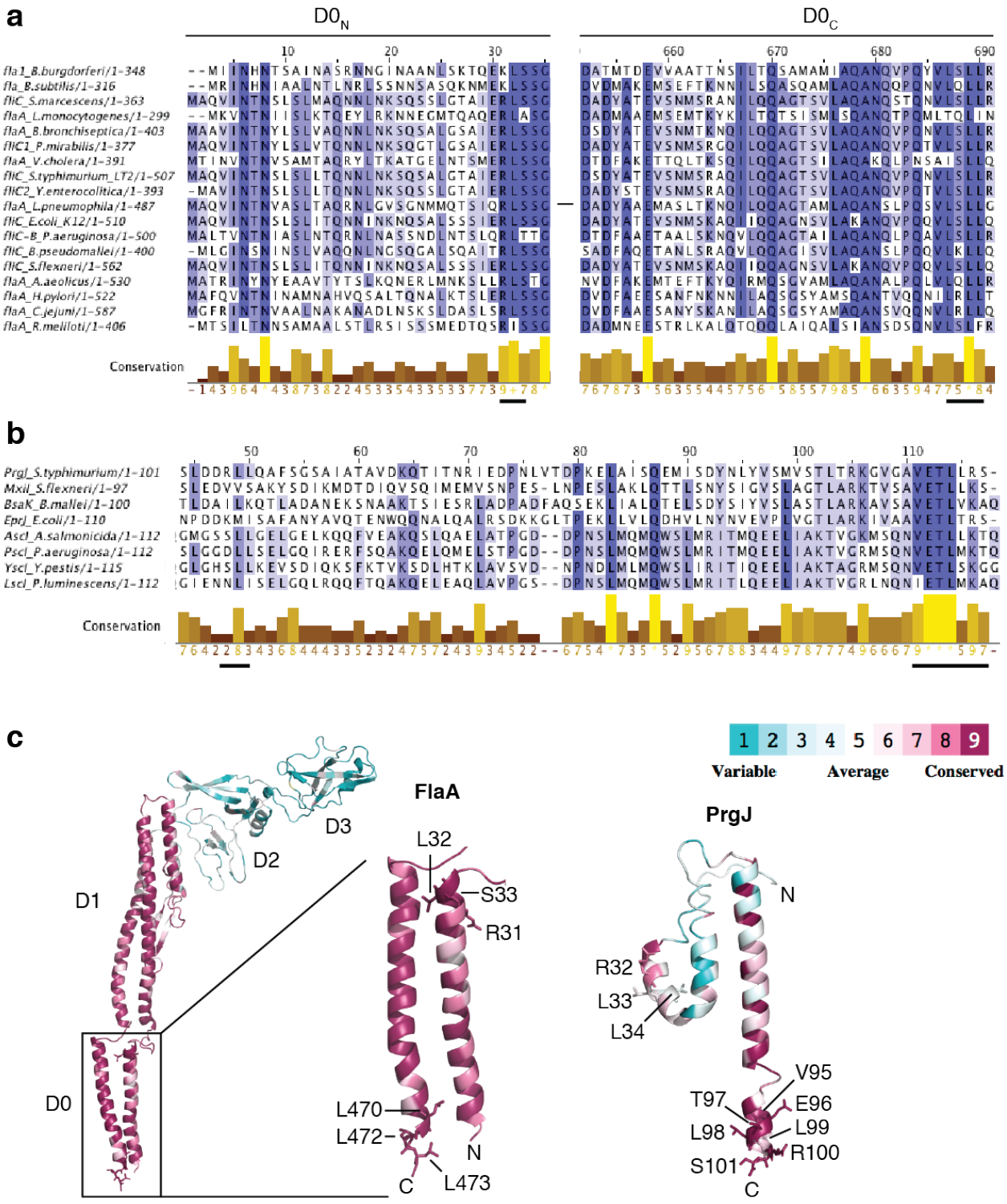


Figure 3.5. NAIP recognition motifs are conserved

a, Alignment of flagellins from *Borrelia burgdorferi*, *Bacillus subtilis*, *Serratia marcescens*, *Listeria monocytogenes*, *Bordetella bronciseptica*, *Proteus mirabilis*, *Vibrio cholerae*, *Salmonella typhimurium*, *Yersinia enterocolitica*, *Legionella pneumophila*, *Escherichia coli*, *Pseudomonas aeruginosa*, *Burkholderia pseudomallei*, *Shigella*

flexneri, *Aquifex aeolicus*, *Helicobacter pylori*, *Campylobacter jejuni*, and *Rhizobium meliloti*. **b,** Alignment of T3SS rod proteins from *Salmonella typhimurium*, *Shigella flexneri*, *Burkholderia pseudomallei*, *Escherichia coli*, *Aeromonas salmonicida*, *Pseudomonas aeruginosa*, *Yersinia pestis*, and *Photorhabdus luminescens.* NAIP recognition motifs are underlined. **c,** FlaA and PrgJ residue conservation scores mapped onto PDB 1UCU or a *de novo* structural model, respectively, using ConSurf.

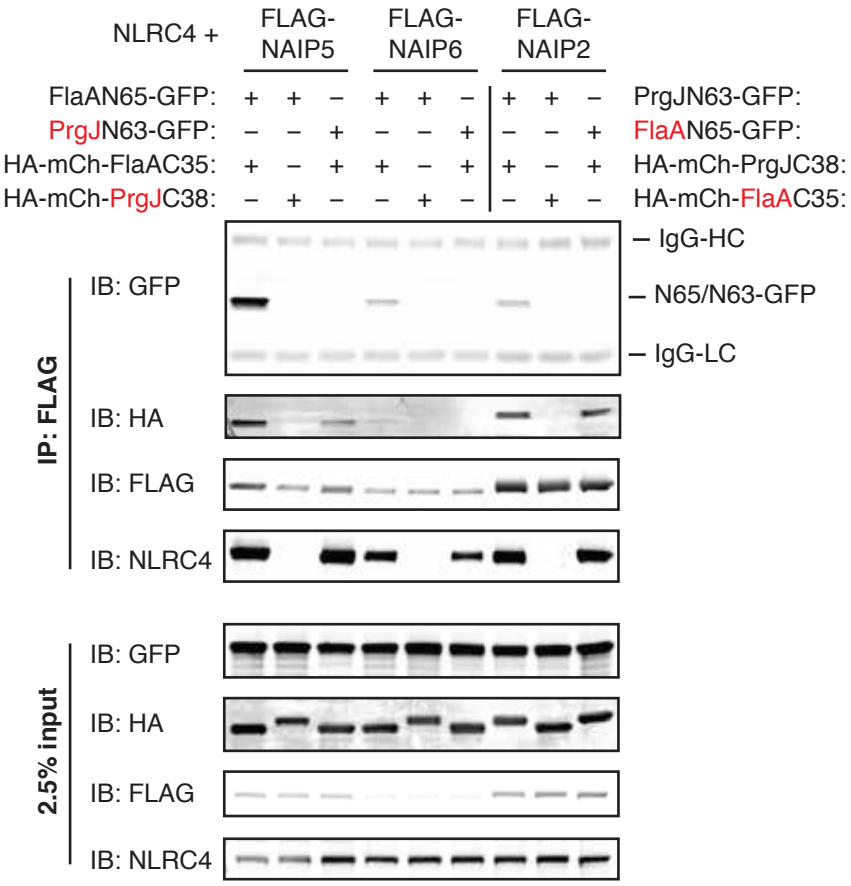


Figure 3.6. NAIPs retain specificity for both ligand halves
The indicated constructs were transfected into HEK293T and subjected to anti-FLAG IP. NAIPs do not bind the C-terminal half of a non-cognate ligand (indicated in red, IB: HA). Binding of the C-terminal half of the cognate ligand does not stimulate binding of the N-terminal half of a non-cognate ligand (indicated in red, IB: GFP). Results representative of at least 3 independent experiments.

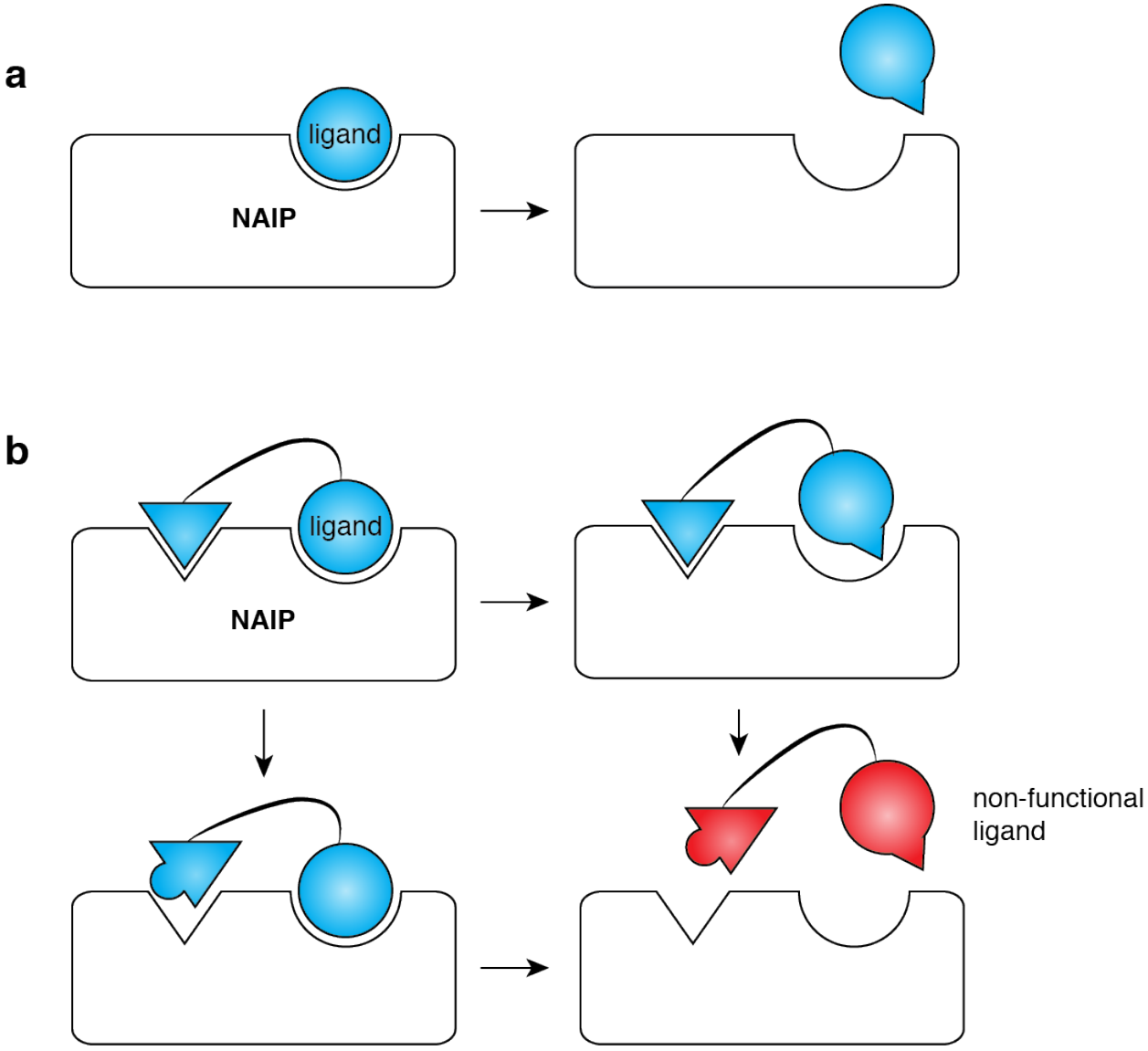


Figure 3.7. Model for the evolutionary advantage of multi-surface recognition of ligands

When only one ligand surface is recognized, a single point mutant that reduces binding affinity at that surface will concomitantly reduce immune recognition (left). However, a second intact site can compensate for affinity-reducing mutations in one binding interface through a tethering avidity model (right). In this case, multiple mutagenic steps are required to evade immune recognition but are more likely to impact ligand function.

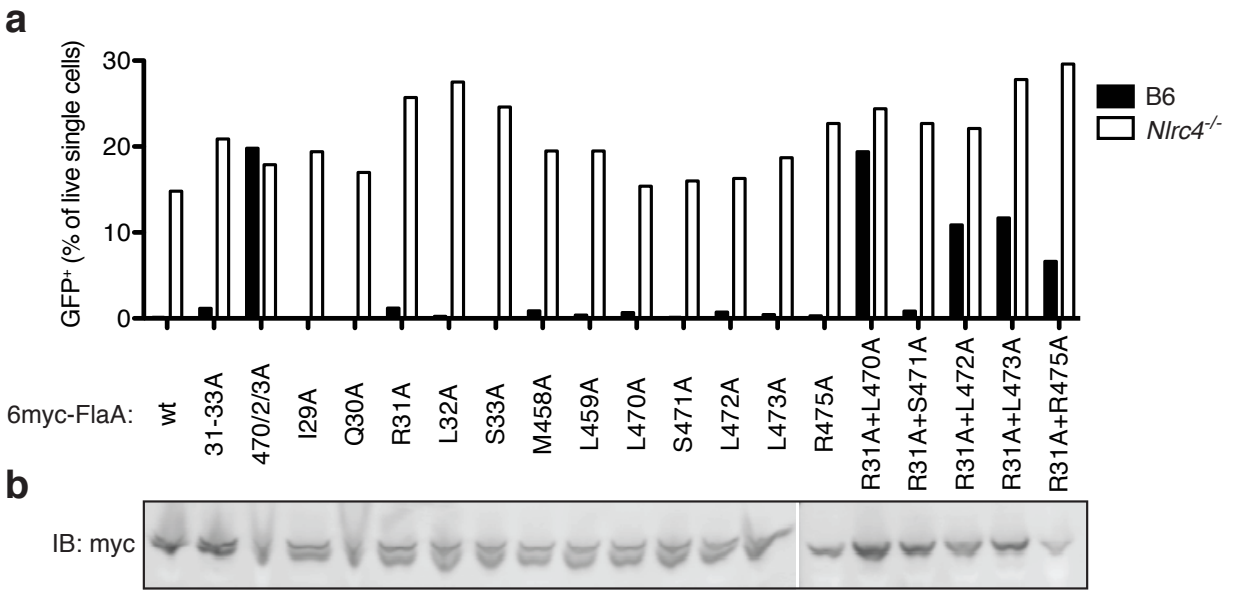


Figure 3.8. Retrovirally transduced flagellin requires multiple mutations to evade NAIP5 recognition

a, The indicated residues of *L. pneumophila* FlaA were mutated to alanine, and FlaA variants were retrovirally transduced into BMM as in Figure 1. **b,** FlaA mutants are expressed. Constructs were transfected into HEK293T, and lysates were probed for FlaA expression by anti-myc immunoblot (IB). Results representative of at least 2 independent experiments.

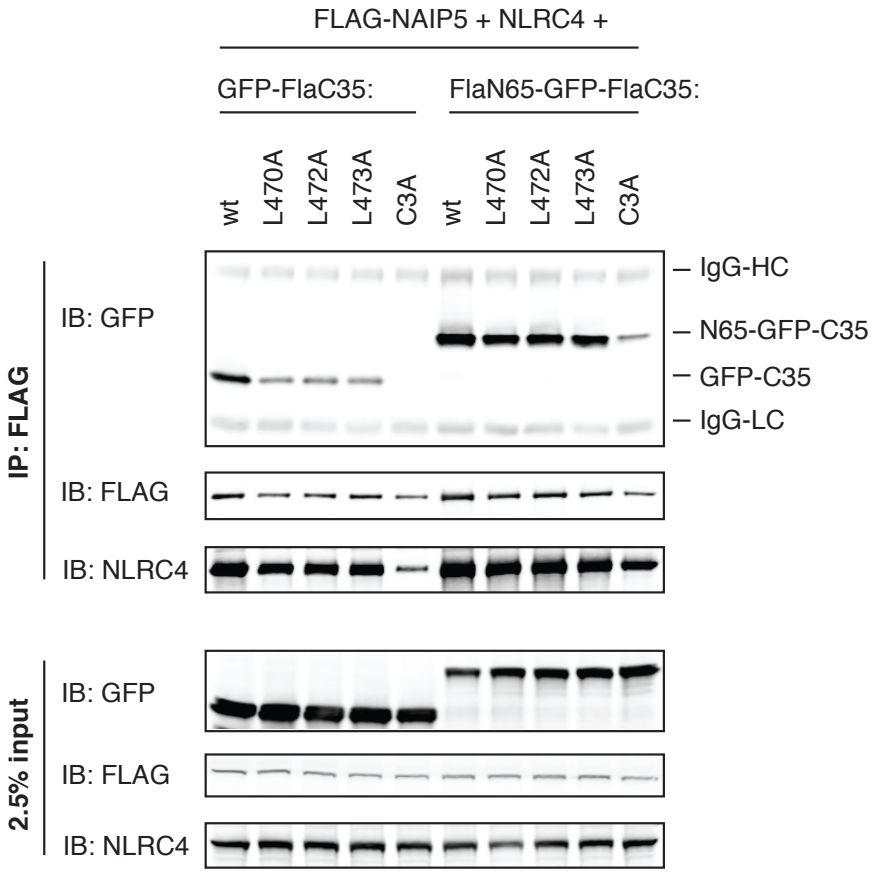


Figure 3.9. Point mutations disrupt NAIP5 binding to single-surface but not to a multi-surface ligand

The indicated constructs were transfected into HEK293T and subjected to anti-FLAG IP. Results are representative of at least 3 independent experiments.

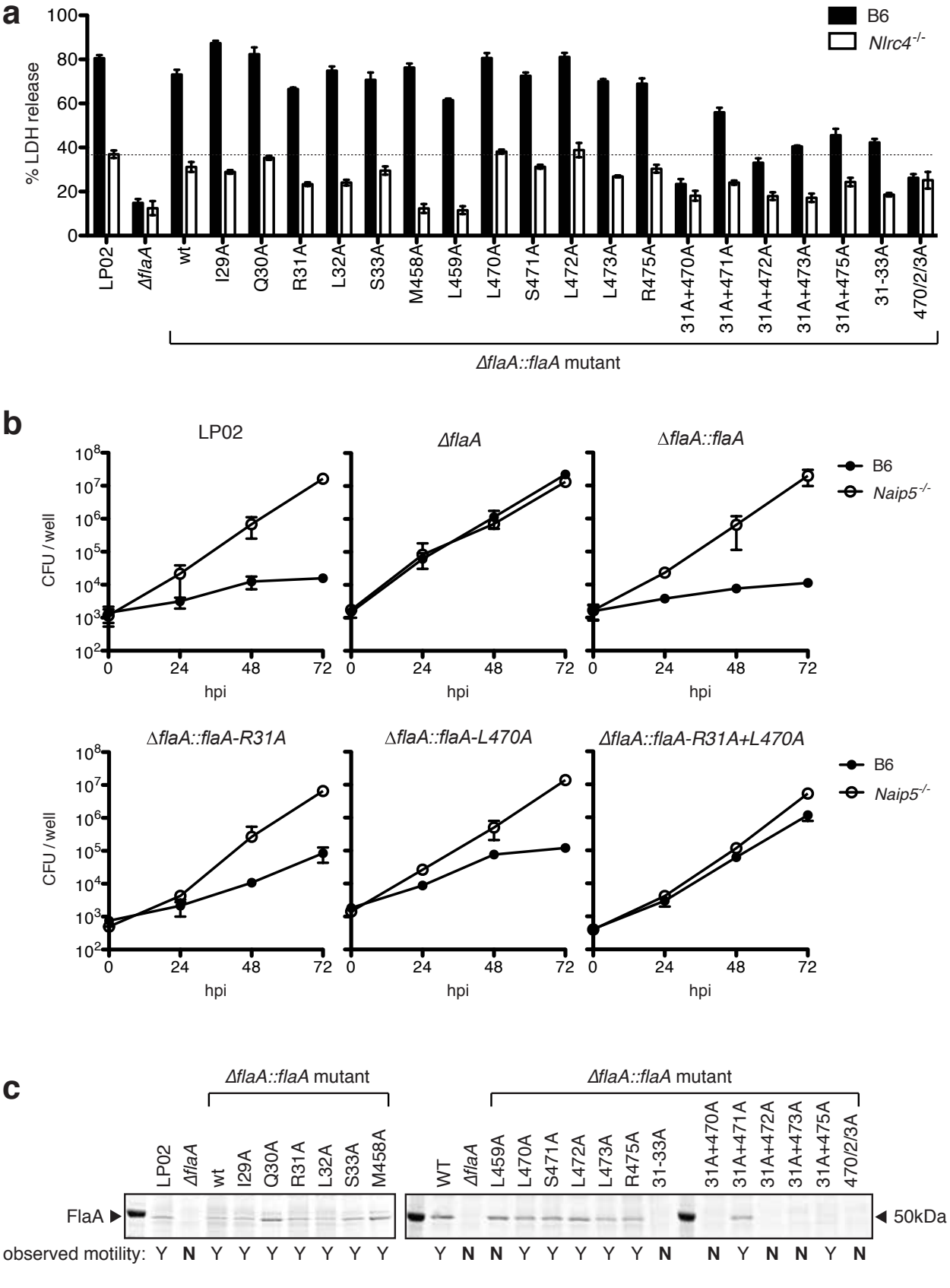
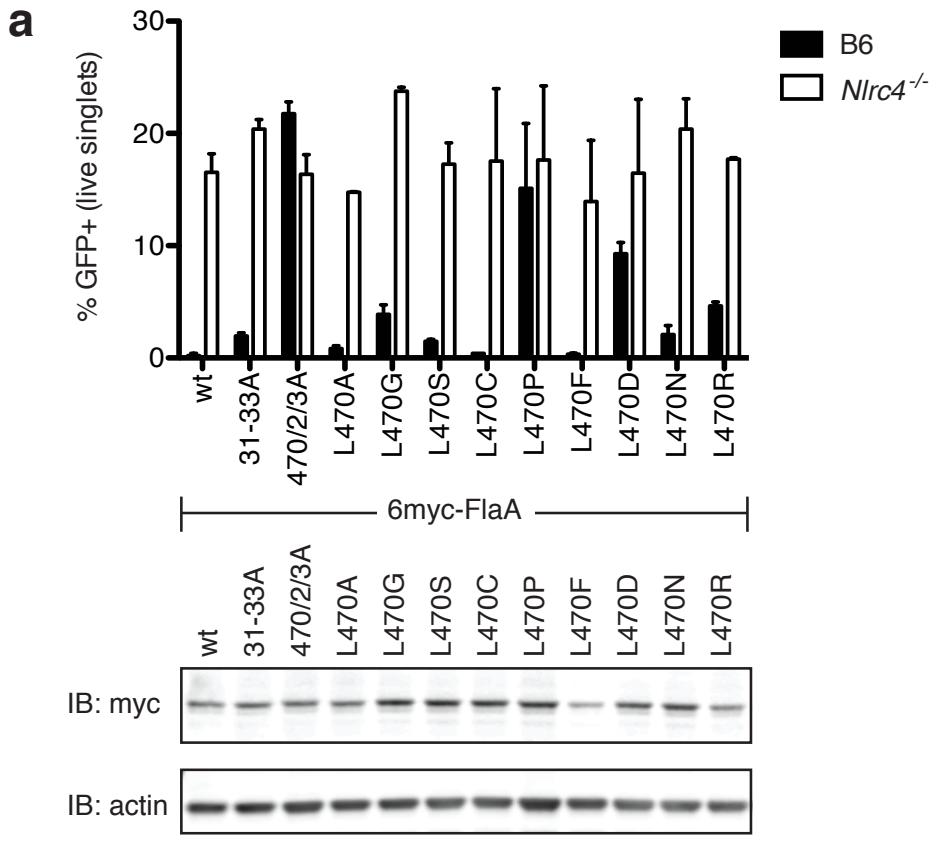

a
% LDH release
B6
Nlrc4-/-
LP02
ΔflaA
wt
I29A
Q30A
R31A
L32A
S33A
M458A
L459A
L470A
S471A
L472A
L473A
R475A
31A+470A
31A+471A
31A+472A
31A+473A
31A+475A
31-33A
470/2/3A
ΔflaA::flaA mutant
b
LP02
ΔflaA
ΔflaA::flaA
ΔflaA::flaA-R31A
ΔflaA::flaA-L470A
ΔflaA::flaA-R31A+L470A
CFU / well
hpi
B6
Naip5-/-
c
ΔflaA::flaA mutant
ΔflaA::flaA mutant
LP02
ΔflaA
wt
I29A
Q30A
R31A
L32A
S33A
M458A
WT
ΔflaA
L459A
L470A
S471A
L472A
L473A
R475A
31-33A
31A+470A
31A+471A
31A+472A
31A+473A
31A+475A
470/2/3A
FlaA
50kDa
observed motility: Y N Y Y Y Y Y Y Y
Y N N Y Y Y Y Y N N Y N N Y N

Figure 3.10. Simultaneous mutation of multiple NAIP5 recognition motifs is required to evade NAIP5 recognition but disrupts flagellar motility
The indicated mutations were introduced at the endogenous FlaA locus of *L. pneumophila* strain LP02. **a,** BMM were infected with *L. pneumophila* strains at MOI = 3, and cell death was measured by lactate dehydrogenase (LDH) release at 4 hrs post-infection. **b,** NAIP5- and FlaA-dependent restriction of *L. pneumophila* replication in BMM. BMM were infected at MOI = 0.01, and colony-forming units (CFU) were measured at the indicated time points. **c,** *L. pneumophila* were classified as motile (“Y”) or non-motile (“N”) based on the observation of swimming “runs.” Bacteria were vortexed to dissociate cell-surface flagella, and supernatants were analyzed by Coomassie stain. Results representative of at least 3 independent experiments. Error bars in **a, b** indicate standard error for 3 biological replicates.


a
% GFP+ (live singlets)
30
20
10
0
B6
Nlrc4 -/-
wt
31-33A
470/2/3A
L470A
L470G
L470S
L470C
L470P
L470F
L470D
L470N
L470R
6myc-FlaA
IB: myc
IB: actin

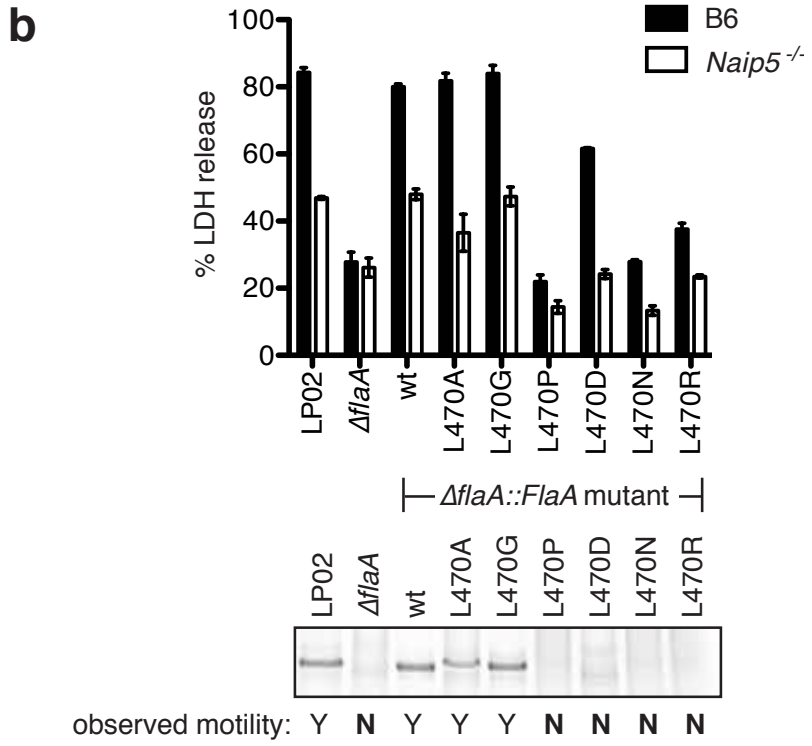

b
% LDH release
100
80
60
40
20
0
B6
Naip5 -/-
LP02
ΔflaA
wt
L470A
L470G
L470P
L470D
L470N
L470R
ΔflaA::FlaA mutant
observed motility: Y N Y Y Y N N N N

Figure 3.11. Non-conservative single point mutations in flagellin can disrupt NAIP5 recognition but also disrupt motility

a, The indicated FlaA variants were retrovirally transduced into BMM as in Figure 1. Transduced *Nlrc4*$^{-/-}$ BMM lysates were probed for FlaA expression by anti-myc IB. **b,** The indicated mutations were introduced at the endogenous FlaA locus of *L. pneumophila* strain LP02. BMM were infected at MOI = 3, and cell death was measured by lactate dehydrogenase (LDH) release at 4 hrs post-infection. Motility and the presence of cell-surface flagella were assessed as in Figure 3. Results representative of at least 3 independent experiments. Error bars indicate standard error for 2 (**a**) or 3 (**b**) biological replicates.

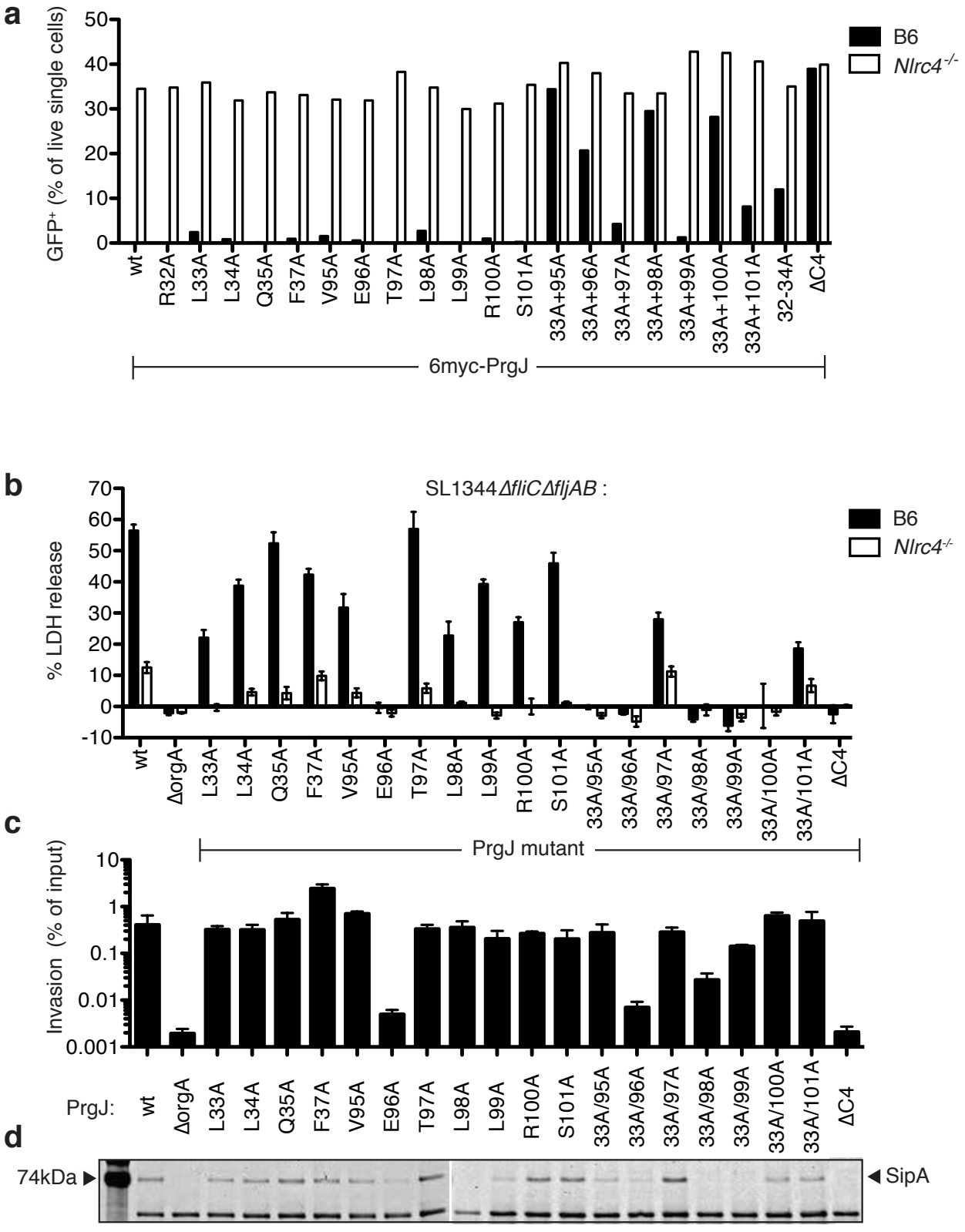


Figure 3.12. Multiple mutations in PrgJ are required to evade NAIP2 recognition but are more likely to disrupt T3SS function

a, The indicated PrgJ variants were retrovirally transduced into BMM as in Figure 1. **b-d,** The chromosomal copy of PrgJ in *S. typhimurium* strain SL1344*ΔfliCΔfljAB* was replaced with the indicated PrgJ variants. The *ΔorgA* strain serves as a negative control for SPI1 function. Bacteria were grown under SPI1-inducing conditions. **b,** BMM were infected with the indicated *S. typhimurium* strains at MOI = 100. LDH release was measured at 1 hr post-infection. In the *ΔfliCΔfljAB* background, NLRC4-dependent cell death is NAIP2- and PrgJ-dependent[122,127]. The E96A mutation does not induce cell death because bacteria fail to invade cells (see **c**). **c,** HeLa cells were infected at MOI = 100. Cells were treated with gentamicin to kill extracellular bacteria, then lysed and plated to determine invasion efficiency (intracellular CFU normalized to infection input CFU). **d,** Culture supernatants were TCA-precipitated and analyzed by Coomassie stain for the secretion of SPI1-dependent bacterial effectors such as SipA. Results representative of at least 3 independent experiments. Error bars in (**b, c**) indicate standard error for 3 biological replicates.

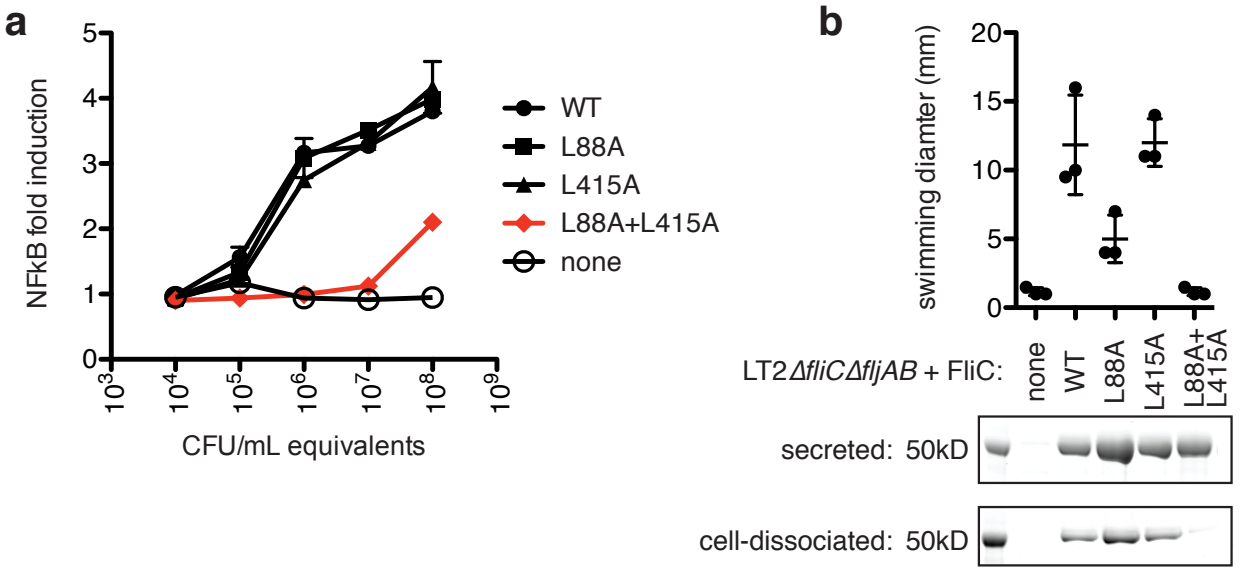


Figure 3.13. Mutation of both TLR5 recognition motifs enhances flagellin evasion of TLR5 recognition at the cost of flagellar motility

S. typhimurium strain LT2*ΔfliCΔfljAB* was transformed with an expression vector encoding wild-type FliC or the indicated variants. **a,** Overnight culture supernatants were incubated 6 hr with CHO cells expressing HsTLR5 and an NFκB luciferase reporter. Reporter cells were analyzed for luciferase expression. **b,** Diameter of colonies incubated on 0.4% agarose plates for 8 hr. Culture supernatants and the supernatants of vortexed bacteria (as in Fig 3c) were analyzed for the presence of secreted or cell-dissociated flagellin, respectively. Results representative of at least 3 independent experiments. Error bars indicate standard error for 4 (**a**) or 3 (**b**) biological replicates.

Chapter Four: Structural Basis for Activation of the NAIP5–NLRC4 Inflammasome by Flagellin

4.1 Prefix

Biochemical studies have made significant progress in uncovering the mechanism of NAIP and NLRC4 cooperation in sensing of cytosolic bacterial pathogens. NAIPs directly bind to bacterial ligands and subsequently recruit NLRC4[3,4,123]. Several alpha-helical domains adjacent to the NBD are likely to mediate binding (Chapter 2) to multiple surfaces of bacterial ligands (Chapter 3). While NAIPs are highly specific for an individual ligand, NLRC4 is involved in the response to all known NAIP ligands, implying either a minimal or non-existent role in direct ligand binding. However, the NLRC4 CARD domain is necessary for CASP1 recruitment and activation[3].

In this chapter I undertake structural studies of the NAIP inflammasomes in the pursuit of two major goals. First, I sought to obtain structural validation of the above biochemical results. More importantly, I sought to obtain structural insights into how the detection of bacterial ligands promotes the activation and assembly of the inflammasome, which is difficult to ascertain from biochemical studies alone. Below, we determine why ligand binding results in a massive rearrangement of NAIP domains that disrupts numerous inter-domain contacts keeping unliganded NAIP in an autoinhibited state[159]. This conformational rearrangement exposes the oligomerization surface of NAIP to recruit and activate NLRC4. Unlike NAIP, NLRC4 oligomerization is self-propagating[5], resulting in the rapid, switch-like assembly of inflammasomes.

Initial efforts to purify monomeric NAIPs in their apo or ligand-bound state proved largely unsuccessful. These attempts included transient expression of several NAIP paralogs in *E. coli*, Sf9 and Hi5 insect cell lines, and HEK293T cells; the stable introduction of NAIP5 into HEK293T to avoid heterogeneity of overexpression; co-expression of NAIP5 or NAIP6 with flagellin to stabilize the NAIP; fusion of flagellin or its C-terminal 35 residues to the C-terminus of NAIP5 to decrease flagellin dissociation; and a variety of LRR truncations that, rather than increasing protein yield, led to NAIP6 destabilization. In most cases, the protein yield was too low for crystallographic analysis, and when large amounts of NAIP were purified from insect cells, size exclusion chromatography indicated heterogeneous aggregation of NAIPs (data not shown).

However, transient expression of NAIP5 and flagellin with NLRC4 in HEK293T yielded assembled inflammasome particles of sufficient yield and size for analysis by electron microscopy. The initial heterogeneity of these particles was largely due to CARD-mediated stacking between inflammasome particles. This stacking was disrupted by several point mutations in a homotypic interaction surface[210] of the NLRC4 CARD, as well as an adjacent GFP fusion that provided steric hindrance to stacking. With these more homogeneous particles in hand, we investigated the structure of assembled NAIP5 inflammasomes in the absence and presence of the downstream CASP1 signaling effector.

In the course of this work, several other groups also resolved the structure of a CARD-deficient NLRC4 in the inflammasome[5,112]. However, radial averaging of the NAIP and NLRC4 constituents prevented the resolution of NAIP and ligand within the inflammasome, and the CARD deletion prohibited insight into CASP1 recruitment. In contrast to these reports, we resolve the NAIP5–flagellin and NAIP5–NLRC4 interface, and resolution of the NLRC4–CASP1 interface is in progress. However, additional insight into the mechanism of NLRC4 propagation following activation of the first NLRC4 protomer by flagellin-bound NAIP5 is drawn from their work. Collectively, the structures of NAIP inflammasomes illustrate the switch-like mechanism by which the detection of a single ligand monomer is amplified into oligomerization-induced signaling via the CASP1 enzyme.

4.2 Abstract

The innate immune system specifically detects and responds to diverse pathogen-encoded molecules[17,71]. Members of the Nucleotide-binding domain, Leucine-rich Repeat-containing (NLR) superfamily function as cytosolic pathogen sensors in plants and animals[55]. NAIP5 (NLR family, apoptosis inhibitory protein 5) is an NLR that binds the bacterial protein flagellin and then co-oligomerizes with NLRC4 (NLR family, CARD [Caspase Activation and Recruitment Domain]-containing 4) into a large multi-protein complex called an inflammasome[3-5,112]. Assembled inflammasomes initiate potent immune responses via the recruitment and activation of the Caspase-1 protease[166,211,212]. However, the mechanism by which NAIP5 binds flagellin to nucleate the inflammasome remains unclear. Here, we report the structure of the complete ~1.4 MDa flagellin-NAIP5-NLRC4 inflammasome complex, revealing the interface between NAIP5 and flagellin, and between NAIP5 and NLRC4. In the assembled complex, a single flagellin molecule is recognized solely by a single NAIP5 protomer, whereas NLRC4 does not contribute to flagellin recognition. The two helices of the flagellin D0 domain contact an extended surface formed from six distinct NAIP5 domains, apparently prying NAIP5 into an open and active conformation that assembles with NLRC4. Further NLRC4 oligomerization results in formation of an inner ring of CARD domains that recruit Caspase-1. This first structure of a full-length NLR in complex with a ligand provides key insights into the molecular mechanisms by which the immune system senses pathogens and initiates a protective cellular response.

4.3 Introduction

NAIPs are cytosolic innate immune receptors that detect the intracellular presence of conserved structural components of cell-invasive bacteria. Mouse NAIP5 binds to the terminal D0 domain of bacterial flagellin, whereas NAIP1 and NAIP2 detect the needle and inner rod proteins of bacterial type III secretion systems[3,4,125,126]. Ligand binding causes NAIPs to co-oligomerize with NLRC4, forming a high molecular weight inflammasome complex[3,5,112]. Inflammasome oligomerization is believed to mediate signaling by multimerizing the CARDs of NLRC4 to form a platform that recruits and activates the Caspase-1 protease[82]. Active Caspase-1 processes and mediates secretion of pro-inflammatory interleukin-1β and -18 and triggers a lytic form of cell

death called pyroptosis[78]. NAIP5 and NLRC4 are ~15% identical at the amino acid level, and share a similar overall domain architecture, found in all mammalian NLRs, which includes a conserved nucleotide binding domain (NBD), helical domain 1 (HD1), winged helix domain (WHD), helical domain 2 (HD2) and leucine-rich repeat (LRR) domain (Figure 4.1a). However, NAIP and NLRC4 encode different N-terminal domains: NLRC4 includes an N-terminal CARD, which recruits Caspase-1, whereas NAIP5 includes three N-terminal Baculovirus Inhibitor-of-apoptosis Repeat (BIR) domains (Figure 4.1a), the function of which remains unknown. No NAIP structure has been reported, and the existing NLRC4 structures[5,112,159] all lack the N-terminal CARD. Several prior cryo-electron microscopy (cryo-EM) studies revealed that the assembled NAIP/NLRC4 inflammasome adopts an overall ring-like structure[5,112,165]. However, high-resolution reconstructions[5,112] were obtained by applying a symmetry constraint that assumed the NAIP and NLRC4 protomers to be identical. Thus, these reconstructions failed to reveal the NAIP or its bound ligand.

4.4 Results

4.4.1 Structure of the NAIP5–NLRC4 inflammasome

In order to define the molecular mechanism driving the assembly of a functional inflammasome, we used cryo-EM to visualize the structure of the NAIP5–NLRC4 inflammasome bound to flagellin (Figure 4.1b). Two-dimensional image analysis showed the presence of inflammasome particles of different sizes (Figure 4.2). In addition, inflammasomes with full-length NLRC4 tended to form loosely associated stacks, likely due to the propensity of CARDs to associate with one another (Figure 4.3). The stacking interactions were conformationally heterogeneous and severely limited resolution. Rather than removing the CARD of NLRC4 to reduce ring stacking[5,112], we used NLRC4 with an N-terminal GFP fusion, which provided enough steric hindrance to prevent most ring stacking. The additional introduction of two point mutations (F79A and D83A) in a homotypic interaction surface of the CARD[213] eliminated the remaining ring stacking (Figure 4.3). Although the resolution of the CARDs was very low due to flexibility, this strategy allowed us to visualize the CARDs at the center of the complex (Figure 4.4). The CARDs, which mediate NLRC4–Caspase-1 interactions, are partially occluded in the stacked rings, making it unlikely that the stacked rings represent a biologically relevant form of the complex.

In order to visualize the NAIP5 subunit and its unique contacts with flagellin and NLRC4, we did not apply symmetry during our image analysis, in contrast to previous studies[5,112]. The complexes invariably appear to contain a single NAIP5 subunit bound to a single flagellin monomer, providing direct evidence in support of previous suggestions that the oligomer is nucleated by a single NAIP[5,112,165], which then associates with a variable number of NLRC4 subunits (Figure 4.2). To obtain the best possible resolution for NAIP5 and its critical interactions, we carried out focused refinement that concentrated on a region of the inflammasome containing the unique NAIP5 and two of the adjacent NLRC4 subunits (Figure 4.1c). This image analysis strategy (see Methods and Figure

4.2) yielded an improved reconstruction of the selected region with a final overall resolution of 7.2 Å, though the resolution of much of the structure was 6 Å. The combination of the available atomic structures, homology modeling, and advanced fitting techniques[214-216] (see Materials and Methods) enabled the near-atomic characterization of NAIP5 and its binding interfaces with NLRC4 and flagellin for the first time.

4.4.2 NAIP5 activation of NLRC4

The interface between NAIP5 and NLRC4 is well resolved (Figure 4.5). The NAIP5 oligomerization donor surface, contributed to by the WHD and NBD, is similar overall to that of NLRC4, and is composed primarily of basic and hydrophobic residues, as previously predicted[5,112]. The acceptor surface of NLRC4 that contacts NAIP5 consists of complementary acidic and hydrophobic residues. Interestingly, the conformations of the two adjacent NLRC4 molecules are identical even though symmetry was not imposed. Thus, the acceptor surface of NLRC4 is the same regardless of whether it contacts a NAIP5 or NLRC4 donor surface. Using a co-immunoprecipitation (IP) assay, we found that NLRC4 acceptor surface residues I124 and D125 are required for NLRC4 to bind to NAIP5 (Figure 4.5c), analagous to their role in binding a neighboring NLRC4[5]. In contrast, D123 appears to be less important for NLRC4 to bind NAIP5 (Figure 4.5c) than to NLRC4[5]. In addition, we found that NLRC4 residues N116, Y118, L120, I127, and M349 (which were not previously tested for their role in NLRC4 homotypic interactions[5,112]) were important for interaction with NAIP5. Taken together, we conclude that the structural homology between NAIP5 and NLRC4 is high in the regions that mediate oligomerization, although there are some subtle changes, particularly in the positioning of the acceptor surface loop (Figure 4.6a).

In contrast to prior averaged reconstructions of the NAIP–NLRC4 inflammasome, or the structures of related complexes such as apoptosomes[217], we find that the NAIP5–NLRC4 inflammasome does not form closed, planar rings (Figures 4.1, 4.2). The lack of ring closure is consistent with the prior proposal that although a donor surface on NAIP5 can recruit and activate the cognate acceptor surface on NLRC4, the donor surface on NLRC4 can only interact with additional NLRC4 protomers, and cannot close the ring by interacting with an acceptor surface of NAIP5[5,112].

4.4.3 NAIP5 diverges from NLRC4 to bind flagellin

Some regions of NAIP5 are distinct from the corresponding areas of NLRC4 (Figure 4.6b), and several of these differences have clear functional consequences. For example, NAIP5 contains a large insertion in the N-terminal region of the LRRs, which we term the ID (inserted domain, residues 922-984) (Figures 4.1a, 4.6b). The lack of structures homologous to this otherwise not clearly resolved region precluded modeling except for a single helix (964-976) that appears to contact both helices of the flagellin D0 domain (Figure 4.7a, b). To validate this interaction, we used a co-IP assay[4] to assess the effect of NAIP5 mutations on flagellin binding. In order to separately evaluate the effect of NAIP5 mutations on binding of the $D0_N$ and $D0_C$ flagellin helices, we expressed $D0_N$ and $D0_C$ as separate polypeptides fused to GFP. We and others have

previously found that this split flagellin ligand is still able to activate NAIP5[118,124,165]. As confirmation, we also tested binding of full-length flagellin. The binding assays show that several residues in the ID, especially S952/F954 and Y974/E975, play a role in binding to both the $D0_N$ and $D0_C$ helices (Figure 4.7c, Figure 4.8).

Another important difference between NAIP5 and NLRC4 is seen in HD2, where the phosphorylation loop of NLRC4[159,218] is replaced with two helices in NAIP5 (Figure 4.6b). One of these helices (residues H835-L850) makes numerous contacts (via Q837, L840, F844, G847, L848) with the flagellin C-terminal helix ($D0_C$, Figure 4.7a, b). These contacts are primarily hydrophobic, and point mutations at these sites reduce the binding of $D0_C$ in our IP assay (Figure 4.7c). Because association of $D0_N$ with NAIP5 depends on the binding of $D0_C$ (Chapter 3), these mutations also reduce $D0_N$ binding. HD2 also has hydrophobic residues (W841 and F844) that appear to contact flagellin and, when mutated, disproportionately disrupt $D0_N$ binding (Figure 4.7c). This effect may be mediated indirectly through repositioning the modeled ID helix that contacts $D0_N$. In addition, HD1 contains several residues (I626, I627) that are important for binding to $D0_C$ (Figure 4.7c, d). Interestingly, the inserted helices in HD2 (837-847) and the ID (922-984), as well as HD1 residues contacting flagellin, are poorly conserved between NAIP2 and NAIP5 (Figure 4.9). This divergence and the direct contacts with flagellin likely explain why these regions appear to play a role in conferring ligand specificity across NAIPs[123].

The final major difference between NAIP5 and NLRC4 is a second insertion in the NAIP5 LRR that takes the form of an additional helix and beta strand after the sixth leucine rich repeat (residues 1102-1138, Figure 4.6b). This insertion is of unclear functional significance as it does not appear to contact flagellin (Figure 4.7a).

Given the differences between NAIP5 and NLRC4, it is not surprising that our model of NLRC4 differs slightly from previous structures, since those studies averaged NAIP5 with NLRC4 by imposing symmetry during the cryo-EM reconstruction[5,112]. The largest discrepancy between our NLRC4 model and prior models is, as could be expected, in the phosphorylation loop (where NAIP5 differs significantly from NLRC4), with smaller deviations in the structures seen in other loops throughout (Figure 4.10).

4.4.4 Additional flagellin contacts outside of the NAIP specificity region

In addition to HD1, HD2, and the ID, all of which lie in the specificity-determining region of NAIPs[123], our NAIP5–flagellin structure unexpectedly reveals at least three additional regions that contact flagellin. First, a small portion of the LRR domain contains residues (R1330, H1360, and S1363) that contact the $D0_C$ helix of flagellin (Figure 4.7e). Binding assays suggest these residues contribute modestly to flagellin binding (Figure 4.7c). Second, our modeling of the NAIP5 BIR domains revealed that residue S108 of BIR1 is in direct contact with flagellin's $D0_C$ helix (Figure 4.7d) and is required for binding of this helix (Figure 4.7c). This result reveals the first specific function attributable to the NAIP BIR domains. Lastly, our NAIP5 model also contains a helix, just N-terminal to BIR1,

with residues (Q33 and V34) that contact one or both D0 helices of flagellin (Figure 4.7d). Mutation of these sites also modestly reduced flagellin binding (Figure 4.7c). Taken together, these results unexpectedly reveal that the binding pocket for flagellin is not formed by a single NAIP5 domain but instead involves residues contributed by six different NAIP5 domains, namely BIR1, HD1, HD2, ID, LRR, and the N-terminal helix (Figure 4.7). We propose that the extended recognition surface results in broadly distributed binding energy, which might make pathogen escape from detection through single point mutations in flagellin unlikely (see Chapter 3). Mutations in NAIP5 had a greater effect on binding of the flagellin $D0_N$ or $D0_C$ domains expressed as separate polypeptides, as compared to the binding of full-length flagellin (Figure 4.7c), consistent with our hypothesis that flagellin is bound via multiple contacts with the $D0_N$ and $D0_C$ helices.

When flagellin is a monomer in solution, the D0 domain is thought to be at least partially disordered[134,135], but this domain becomes highly ordered when it forms the core of the flagellar filament[140,201]. Interestingly, the D0 domain of the flagellin monomer bound to NAIP5 adopts a conformation that is strikingly similar to its structure in the flagellar filament (Figure 4.11a). Thus, mutations that disrupt this conformation in order to escape NAIP5 recognition would also likely disrupt the ability of the D0 domain to support flagellar filament formation (Figure 4.11b, c), a conclusion that is supported by companion functional studies (Chapter 3).

4.4.5 Model of inflammasome assembly

Our structural model provides insight into the initiating events of inflammasome assembly and activation (Figure 4.12a-c). Prior studies[5,112] showed that the activated form of NLRC4 found in the mature inflammasome undergoes a rigid body rotation between domains relative to the autoinhibited state[159], with the hinge located between HD1 and the WHD. This conformational change unfurls NLRC4 and positions its donor surface to activate and recruit the next incoming NLRC4 (Figure 4.12b). Our NAIP5 structure now shows how the acceptor surface of NLRC4 is also receptive to the donor surface of flagellin-activated NAIP5. Activated NAIP5 may trap an existing conformation of NLRC4 or promote the adoption of this structure after an initial, partial contact between NLRC4 and NAIP5. Importantly, we find no role for NLRC4 in directly binding flagellin (Figure 4.1, 4.7), indicating that NLRC4 activation is solely mediated by interaction with either flagellin-activated NAIP5 or other activated NLRC4 protomers.

We propose that NAIP5 also undergoes a rigid-body rotation hinged around the WHD when it is activated by flagellin binding (Figure 4.12a). This conclusion is supported by the high degree of structural homology between NAIP5 and NLRC4 in the NBD, HD1, and WHD domains, which are involved in the hinging and unfurling motion, in contrast to the HD2 and ID of NAIP5, which have diverged from NLRC4 to form the primary binding interface for the flagellin D0 domain (Figure 4.6, 4.7). If we make the parsimonious assumption that the inactive form of NAIP5 is similar to that of NLRC4 (Figure 4.12a, b), then it is apparent that crucial parts of the flagellin-binding surface, particularly those in

HD2, would be sterically occluded by the NBD and LRR (Figure 4.12a). Flagellin binding would necessarily displace the occluding NBD and LRR from HD2, freeing the NBD to serve as a donor surface for NLRC4 and allowing inflammasome assembly—and the ensuing inflammatory response—to begin (Figure 4.12c). Although NLRs in animals and plants are responsive to diverse stimuli that are likely detected via distinct mechanisms, we propose that the hinged opening process may represent a common feature of NLRs that confer them switch-like behaviors in diverse signaling contexts.

4.4.6 Purification of CASP1-associated inflammasomes

Consistent with the idea that head-to-head stacked inflammasomes are not the signaling-relevant form of inflammasome, the F79A and D83A mutations in the NLRC4 CARD disrupted both stacking and CASP1 activation (Figure 4.13a). These data indicate that the NLRC4 CARD interface that recruits CASP1 mediates a stacking artifact when inflammasomes are purified in the absence of CASP1. The lack of CASP1 activation by GFP-NLRC4-F79A/D83A precluded our use of this construct in structural studies assessing CASP1 recruitment. Instead, we co-expressed GFP-NLRC4 with a catalytically dead C2824A CASP1 to avoid cell toxicity associated with CASP1 activation[3]. The residual aggregation of GFP-NLRC4 inflammasomes was disrupted when we purified inflammasomes using a FLAG tag on CASP1-C284A rather than on NAIP5, as indicated by a rightward shift in the size exclusion chromatography profile (Figure 4.13b).

Using this strategy, we are currently analyzing CASP1-associated inflammasomes. The 2D class averages of these particles indicate that inflammasomes are single, unstacked discs, and CASP1 is associated with one face of this disc (Figure 4.14). This face contains the NLRC4 CARD domains, located below the ring in particles purified without CASP1 (Figure 4.1b). Interestingly, CASP1 appears to be initiating a spiral filament extending from the NLRC4 CARDs. Although further refinement will be necessary to draw firm conclusions, this pattern appears to be similar to the CASP1 filamentation observed upon overexpression of the CASP1 CARD domain in isolation[219] or nucleated with ASC filaments[220]. These data suggest that interactions between CASP1 CARDs can mediate filamentation regardless of the presence or absence of the adaptor ASC, provided that the nucleating inflammasome contains a CARD for direct CASP1 recruitment[113]. The formation of an extending CASP1 filament likely serves to further amplify inflammasome signaling by activating more CASP1 protomers than can be directly recruited to a single inflammasome particle with 8-10 NLRC4 CARD domains. Thus in addition to the amplification provided by NLRC4 self-propagation, CASP1 participates in a prion-like feed-forward signal expansion. Together, these mechanisms ensure that NAIP detection of a single bacterial ligand monomer in the host cell cytosol can flip a binary switch to initiate rapid inflammasome signaling.

4.5 Methods

4.5.1 Preparation of inflammasome complexes

Mouse NAIP5, NLRC4, and *Legionella pneumophila* flagellin (FlaA) were expressed in HEK293T from the MSCV2.2 retroviral vector. FLAG-NAIP5, 6myc-FlaA, NLRC4, and GFP-NLRC4 have been described[3,123]. Residues F79 and D83 of GFP-NLRC4 were mutated to alanine using Quickchange PCR (F: GAGTCTTGAAAACTGGGACTATgcTG TGTATCAGGcCTTAACTGGACAAAATCTTTCTTATC, R: GATAAGAAAGATTTTGTCC AGTTAAGgCCTGATACACAgcATAGTCCCAGTTTTCAAGACTC). HEK293T cells were grown in DMEM (supplemented with 10% FBS, 2 mM L-glutamine, 100 U/mL penicillin, and 100 µg/mL streptomycin), and confluent cells were diluted 1:4 onto fresh 15 cm TC-treated plates 1 day prior to transfection. Each plate (typically 8-12 plates) was transfected with 16 µg each of GFP-NLRC4-F79A/D83A and FLAG-NAIP5 and 8 µg of 6myc-FlaA using Lipofectamine 2000 (Invitrogen). After 36-48 hr, cells were harvested in cold PBS and lysed in cold, buffered detergent (50 mM HEPES, pH 7.6, 150 mM NaCl, 10 mM KCl, 5 mM $MgCl_2$, 5% glycerol, 1% Triton-X-100, 1x Roche protease inhibitor cocktail). Lysates were clarified (14,000 x g, 30 min, 4 ºC) and incubated 2 hr at 4 ºC with 200 µL of equilibrated FLAG M2 resin (Sigma). Bound resin was washed with 100 column volumes of cold SEC buffer (50 mM HEPES, pH 7.6, 150 mM NaCl, 10 mM KCl, 5 mM $MgCl_2$, 5% glycerol, 0.02% NP-40) and eluted with 0.5 column volume fractions of SEC buffer supplemented with 0.15 mg/mL FLAG peptide (Sigma). Peak fractions (#3 – 7) were pooled, supplemented with 1 mM buffered TCEP (Thermo Fisher), and centrifuged (14,000 x g, 15 min, 4 ºC) to remove aggregates. Purified proteins were separated on a Superose6 size exclusion chromatography column (GE Healthcare) equilibrated in SEC buffer containing 1 mM TCEP, and eluates were collected in 0.5 mL fractions and tracked by absorbance at 280 nm. Peak eluates were used for EM analysis.

4.5.2 Electron microscopy

Cryo-EM grids were prepared on continuous carbon-coated C-flat holey carbon grids (Protochips) that had been plasma cleaned for 8s in air using a Solarus Plasma Cleaner (Gatan) operated at 10W. Each grid was placed carbon-side down on top of a 20 uL droplet of prepared inflammasome complexes and incubated at room temperature for 10 min. After incubation, the grid was washed carbon-side down on a droplet of EM washing buffer (50 mM HEPES, pH 7.6, 150 mM NaCl, 10 mM KCl, 5 mM $MgCl_2$, 1% trehalose, 0.02% NP-40) and loaded into a Vitrobot Mark II (FEI) at 22°C and 100% relative humidity, then immediately blotted and plunge-frozen in liquid ethane. Grids were inserted into a Titan Krios transmission electron microscope (FEI) that was operated at 300 keV and equipped with a K2 direct electron detector (Gatan). Data was collected using SerialEM at a magnification corresponding to 1.31 Å/pixel, with a defocus range of –1.8 µm to –4.0 µm. Each 6-second exposure consisted of 20 frames and the total dose was 10 electrons/pixel/second.

4.5.3 Image processing

MotionCor2[221] was used for frame alignment and dose weighting and the CTF was estimated using Gctf[222]. All other processing, including particle picking, 2D and 3D classification, and 3D refinement were conducted using RELION 1.4[223]. 865,358 particles were picked and subjected to initial 2D classification. 626,460 particles representing the best 2D classes were used for 3D classification into five classes. The highest-resolution class, containing 252,214 particles, was selected for 3D refinement. A mask around the first three subunits (NAIP5 and two NLRC4 subunits) and flagellin was then applied when the refinement began local searches. The overall resolution of the map was 7.2 Å according to the gold-standard FSC = 0.143 criterion[224,225]. The final map was filtered based on local resolution using BSoft[226].

4.5.4 Structural modeling

For the atomic interpretation of the high resolution cryo-EM map we employed the protein structure prediction server, I-TASSER[214], to generate reliable independent homology models for the BIR1 (54-121), BIR2 (159-232), BIR3 (272-361), NDB (397-601), HD1 (603-655), WHD (663-763), HD2 (766-921), and LRR (984-1389) domains of NAIP5. The main structural template identified by I-TASSER for NAIP5 was the crystal structure of NLRC4 in the inactive conformation[159] (PDB ID: 4KXF) that covered all the domains except the N-terminal (Nt) BIR region, where homology models from several BIR domains were recognized (PDB IDs: 1SE0, 2VM5, and 1OXQ for BIR1, BIR2, and BIR3, respectively). NAIP5 homology modeling is challenging, as the overall sequence identity with respect to NLRC4 is low (15%). However, the high-resolution cryo-EM NAIP5 reconstruction evidences a well-conserved fold, in particular, in domains NDB, HD1, and WHD, and to a lesser extent in the LRR and HD2 domains. The identities of BIR domains with their templates were 25, 75, and 45% for BIR1, BIR2, and BIR3, respectively. Also using I-TASSER, we modeled *L. pneumophila* flagellin from the flagellar filament structure of *Salmonella enterica*[140] (PDB ID: 1UCU).

The map densities were initially assigned to specific components by rigid-body fitting of the homology models using UCSF Chimera[227], ADP_EM[228], or Situs[229]. These fitted models were used as starting point for flexible refinement using iMODFIT[215] when necessary. The small loops connecting domains not accounted for in the homology modeling were *ab-initio* generated by RCD+ server[230]. The loops that better fitted the map were included in the complete model, which was finally refined with PHENIX[216].

The fitting of NDB, HD1, and WHD models was straightforward. The HD2 domain (sequence identity 20%) was modeled in three parts: Nt (766-817), horseshoe-like (818-850), and Ct (851-921) to better account for the conformational differences with respect to the HD2 NLRC4 template. In the N-terminal region of NAIP5 we were only able to model the three BIR domains plus one N-terminal helix (31-48), as the lack of structural homologs for the connecting regions and its slightly lower resolution prevented further atomic interpretation. The most challenging region was the LRR (sequence identity 20%), and in particular modeling the two insertions with respect to NLRC4 sequence.

The first insertion, the inserted domain (ID), comprised 62 residues near the start of the LRR (922-983) that partially overlap with 23 missing residues (622-644) of the NLRC4 crystal structure. In this case, the low coverage and poor sequence identity with known protein structures precluded the complete modeling. We maximally extended the fitting procedure following the continuous density, but the majority of this insertion remained unsolved leaving unassigned densities in the cryo-EM map. However, given spatial proximity, we hypothesized that one of these densities with apparent cylindrical shape corresponds to part of the ID with a strong secondary structure prediction of α-helix (964-976). The second insertion, approximately localized between residues 1102 and 1138, was consistently identified as an extra leucine-rich repeat in several homology models considering variable length fragments of the LRR domain.

After assigning all NAIP5 domains and NLRC4, the location of the flagellin D0 domain (1-33, 441-475) was unambiguously determined since the density that would account for the remaining flagellin domains rapidly vanishes, a likely consequence of their intrinsic flexibility. Finally, we refined NLRC4 with iMODFIT and Phenix because small but significant differences from previous structures (PDB IDs: 4kxf and 3jbl) were found (see Figure 4.10).

To obtain the inactive/closed NAIP5 model, the NBD, HD1, WHD, HD2, and LRR domains of NAIP5 model were independently superimposed in the corresponding domains of crystallographic inactive/closed NLRC4 structure (PDB ID: 4KXF, chain K). The NAIP5 domains can adopt similar dispositions except in the HD2 inserted helices (corresponding to the NLRC4 phosphorylation loop), where important steric clashes with the NBD were apparent. This region was manually moved away to obtain a clash-free NLRC4 inactive-like model. Finally, to obtain a complete inactive/closed model, our EM NAIP5 active conformation was morphed into the superimposed domains using iMorph1 tool from the iMODS2 server[231,232].

4.5.5 Mutagenesis and immunoprecipitation

NAIP5 and NLRC4 mutants were generated by Quickchange PCR (Extended Data Table 1). Mutants were assessed for function by transfection into HEK293T with either 6myc-FlaA or GFP-FlaN65 and FlaC35-GFP, followed by anti-FLAG co-immunoprecipitation, as described in Chapter 3.

Table 4.1. Primers used to generate constructs in this chapter

Primer	Use	Sequence
Quickchange mutagenesis of NAIP5 - NLRC4 interface		
oJT484	NAIP5 R590A (F)	CAAACAGGGTCAGAGACATCgcCCTATACCTAGGTACAAGTCTAG
oJT485	NAIP5 R590A (R)	CTAGACTTGTACCTAGGTATAGGgcGATGTCTCTGACCCTGTTTG
oJT492	NAIP5 Q735A (F)	CTTGATGAGCAAATTCACCGCCgcGAGACTGAGACCAGTCTAtCG
oJT493	NAIP5 Q735A (R)	CGaTAGACTGGTCTCAGTCTCgcGGCGGTGAATTTGCTCATCAAG
oJT594	NAIP5 L737D (F)	GCAAATTCACCGCCCAGAGAgacAGACCAGTCTAtCGaTTTTTAGG

oJT595	NAIP5 L737D (R)	CCTAAAAAtCGaTAGACTGGTCTgtcTCTCTGGGCGGTGAATTTGC
oJT514	NLRC4 N116A (F)	GTACAACAGCCCTGCTTTTCTGgcCTTCTACCCCCTGGGTGAAG
oJT515	NLRC4 N116A (R)	CTTCACCCAGGGGGTAGAAGgcCAGAAAAGCAGGGCTGTTGTAC
oJT600	NLRC4 Y118 (F)	AGCCCTGCTTTTCTGAACTTCgcCCCCCTGGGTGAAGATATC
oJT601	NLRC4 Y118 (R)	GATATCTTCACCCAGGGGGgcGAAGTTCAGAAAAGCAGGGCT
oJT516	NLRC4 L120A (F)	GCTTTTCTGAACTTCTACCCCgcGGGTGAAGATATCGACATAATTTTTAA
oJT517	NLRC4 L120A (R)	TTAAAAATTATGTCGATATCTTCACCCgcGGGGTAGAAGTTCAGAAAAGC
oJT520	NLRC4 D123A (F)	TCTACCCCCTGGGTGAAGcTATCGACATAATTTTTAATCTGGAGAAAAC
oJT521	NLRC4 D123A (R)	GTTTTCTCCAGATTAAAAATTATGTCGATAgCTTCACCCAGGGGGTAGA
oJT522	NLRC4 I124A (F)	CTACCCCCTGGGTGAAGATgcCGACATAATTTTTAATCTGGAGAAAACC
oJT523	NLRC4 I124A (R)	GGTTTTCTCCAGATTAAAAATTATGTCGgcATCTTCACCCAGGGGGTAG
oJT524	NLRC4 D125A (F)	CCCCTGGGTGAAGATATCGcCATAATTTTTAATCTGGAGAAAACCTTC
oJT525	NLRC4 D125A (R)	GAAGGTTTTCTCCAGATTAAAAATTATGgCGATATCTTCACCCAGGGG
oJT602	NLRC4 I127A (F)	CCTGGGTGAAGATATCGACATAgcTTTTAATCTGGAGAAAACCTTCACAG
oJT603	NLRC4 I127A (R)	CTGTGAAGGTTTTCTCCAGATTAAAAgcTATGTCGATATCTTCACCCAGG
oJT530	NLRC4 M349A (F)	GGTGATCACCTGTGCAATTCAGgcGGGCAGACAGGAATTCCAAG
oJT531	NLRC4 M349A (R)	CTTGGAATTCCTGTCTGCCCgcCTGAATTGCACAGGTGATCACC
Quickchange mutagenesis of NAIP5 flagellin-binding residues		
oJT550	F32A/Q33A/V34A (F)	TCTTCTCGGGGTGGATGCAgcTgcGGcGGCAAAGAGCCAAGAAGAAGAA
oJT551	F32A/Q33A/V34A (R)	TTCTTCTTCTTGGCTCTTTGCCgCCgcAgcTGCATCCACCCCGAGAAGA
oJT554	N107A/S108A (F)	CTGTAGCTTGATCCTCTTTGGTgcCgcCCTCAGGAAGCTTCCCATAG
oJT555	N107A/S108A (R)	CTATGGGAAGCTTCCTGAGGgcGgcACCAAAGAGGATCAAGCTACAG
oJT650	I626A/I627A (F)	GACATAATCTGTGTGGAAAAGCTTAgcAgcTACTTTATTGATAATAAAGATTTACAGGG
oJT651	I626A/I627A (R)	CCCTGTAAATCTTTATTATCAATAAAGTAgcTgcTAAGCTTTTCCACACAGATTATGTC
oJT570	Q837A/T838A (F)	GATTACATGAAGCTCCATCCAgcAgCTTTTCTATGGTTTCAGTTTGTTAG
oJT571	Q837A/T838A (R)	CTAACAAACTGAAACCATAGAAAAGcTgcTGGATGGAGCTTCATGTAATC
oJT572	F839A/F842A (F)	CATGAAGCTCCATCCACAAACTgcTgcATGGTTTCAGTTTGTTAGAGGG
oJT573	F839A/F842A (R)	CCCTCTAACAAACTGAAACCATgcAgcAGTTTGTGGATGGAGCTTCATG
oJT574	W841A/F842A (F)	GCTCCATCCACAAACTTTTCTAgcGgcTCAGTTTGTTAGAGGGTTGTGG
oJT575	W841A/F842A (R)	CCACAACCCTCTAACAAACTGAgcCgcTAGAAAAGTTTGTGGATGGAGC
oJT576	Q843A/F844A (F)	CCATCCACAAACTTTTCTATGGTTTgcGgcTGTTAGAGGGTTGTGGCTG
oJT577	Q843A/F844A (R)	CAGCCACAACCCTCTAACAgcCgcAAACCATAGAAAAGTTTGTGGATGG
oJT580	G847K (F)	CTTTTCTATGGTTTCAGTTTGTTAGAaaGTTGTGGCTGGTGTCTCC
oJT581	G847K (R)	GGAGACACCAGCCACAACttTCTAACAAACTGAAACCATAGAAAAG
oJT582	L848A/W849A (F)	CTATGGTTTCAGTTTGTTAGAGGGgcGgcGCTGGTGTCTCCTGAATCTT
oJT583	L848A/W849A (R)	AAGATTCAGGAGACACCAGCgcCgcCCCTCTAACAAACTGAAACCATAG
oJT622	N922A/N924A (F)	GAGGAGCTTAAAGGTTTCCATAgcTGGAgcTAAAATGTCATCTTATGTAGATTATTCATT
oJT623	N922A/N924A (R)	AATGAATAATCTACATAAGATGACATTTTAgcTCCAgcTATGGAAACCTTTAAGCTCCTC
oJT624	K925A/M926A/S927A (F)	GCTTAAAGGTTTCCATAAATGGAAATgcAgcGgCATCTTATGTAGATTATTCATTCAAGA

oJT625	K925A/M926A/S927A (R)	TCTTGAATGAATAATCTACATAAGATGcCgcTgcATTTCCATTTATGGAAACCTTTAAGC
oJT626	S928A/Y929A/V930A (F)	TTTCCATAAATGGAAATAAAATGTCAgCTgcTGcAGATTATTCATTCAAGACATATTTTG
oJT627	S928A/Y929A/V930A (R)	CAAAATATGTCTTGAATGAATAATCTgCAgcAGcTGACATTTTATTTCCATTTATGGAAA
oJT628	D931A/Y932A/S933A (F)	ATGGAAATAAAATGTCATCTTATGTAGcTgcTgCATTCAAGACATATTTTGAAAACTTAC
oJT629	D931A/Y932A/S933A (R)	GTAAGTTTTCAAAATATGTCTTGAATGcAgcAgCTACATAAGATGACATTTTATTTCCAT
oJT630	F934A/K935A/T936A (F)	ATAAAATGTCATCTTATGTAGATTATTCAgcCgcGgCATATTTTGAAAACTTACAGCCAC
oJT631	F934A/K935A/T936A (R)	GTGGCTGTAAGTTTTCAAAATATGcCgcGgcTGAATAATCTACATAAGATGACATTTTAT
oJT632	Y937A/F938A/E939A (F)	TCTTATGTAGATTATTCATTCAAGACAgcTgcTGcAAACTTACAGCCACCAGCTAT
oJT633	Y937A/F938A/E939A (R)	ATAGCTGGTGGCTGTAAGTTTgCAgcAgcTGTCTTGAATGAATAATCTACATAAGA
oJT634	N940A/L941A/Q942A (F)	GTAGATTATTCATTCAAGACATATTTTGAAgcCgcAgcGCCACCAGCTATAGATGAG
oJT635	N940A/L941A/Q942A (R)	CTCATCTATAGCTGGTGGCgcTgcGgcTTCAAAATATGTCTTGAATGAATAATCTAC
oJT636	P943A/P944A (F)	ATTCAAGACATATTTTGAAAACTTACAGgCAgCAGCTATAGATGAGGAGTATAC
oJT637	P943A/P944A (R)	GTATACTCCTCATCTATAGCTGcTGcCTGTAAGTTTTCAAAATATGTCTTGAAT
oJT638	I946A/D947A/E948A (F)	GAAAACTTACAGCCACCAGCTgcAGcTGcGGAGTATACATCTGCCTTTGA
oJT639	I946A/D947A/E948A (R)	TCAAAGGCAGATGTATACTCCgCAgCTgcAGCTGGTGGCTGTAAGTTTTC
oJT640	E949A/Y950A/T951A (F)	CCACCAGCTATAGATGAGGcGgcTgCATCTGCCTTTGAGCATATAAG
oJT641	E949A/Y950A/T951A (R)	CTTATATGCTCAAAGGCAGATGcAgcCgCCTCATCTATAGCTGGTGG
oJT642	S952A/F954A (F)	CAGCTATAGATGAGGAGTATACAgCTGCCgcTGAGCATATAAGTGAATGGAGG
oJT643	S952A/F954A (R)	CCTCCATTCACTTATATGCTCAgcGGCAGcTGTATACTCCTCATCTATAGCTG
oJT644	E95A5/H956A/I957A (F)	GAGGAGTATACATCTGCCTTTGcGgcTgcAAGTGAATGGAGGAGAAATTTTGC
oJT645	E95A5/H956A/I957A (R)	GCAAAATTTCTCCTCCATTCACTTgcAgcCgCAAAGGCAGATGTATACTCCTC
oJT646	S958A/E959A/W960A (F)	GTATACATCTGCCTTTGAGCATATAgcTGcAgcGAGGAGAAATTTTGCTCAAGATG
oJT647	S958A/E959A/W960A (R)	CATCTTGAGCAAAATTTCTCCTCgcTgCAgcTATATGCTCAAAGGCAGATGTATAC
oJT648	R961A/R962A/N963A (F)	CCTTTGAGCATATAAGTGAATGGgcGgcAgcTTTTGCTCAAGATGAGGAGATC
oJT649	R961A/R962A/N963A (R)	GATCTCCTCATCTTGAGCAAAAgcTgcCgcCCATTCACTTATATGCTCAAAGG
oJT616	D967A/E968A (F)	TGGAGGAGAAATTTTGCTCAAGcTGcGGAGATCATAAAAAACTATGAAAATATC
oJT617	D967A/E968A (R)	GATATTTTCATAGTTTTTTATGATCTCCgCAgCTTGAGCAAAATTTCTCCTCCA
oJT618	I970A/I971A (F)	GAAATTTTGCTCAAGATGAGGAGgcCgcAAAAAACTATGAAAATATCCGACCCA
oJT619	I970A/I971A (R)	TGGGTCGGATATTTTCATAGTTTTTTgcGgcCTCCTCATCTTGAGCAAAATTTC
oJT620	Y974A/E975A (F)	TCAAGATGAGGAGATCATAAAAAACgcTGcAAATATCCGACCCAGAGCC
oJT621	Y974A/E975A (R)	GGCTCTGGGTCGGATATTTgCAgcGTTTTTTATGATCTCCTCATCTTGA
oJT586	C1329A/R1330A (F)	CAAACCTACAAGAGCTGAACATCgcCgcGAATATCCCAGGACGCATTCA
oJT587	C1329A/R1330A (R)	TGAATGCGTCCTGGGATATTCgcGgcGATGTTCAGCTCTTGTAGGTTTG
oJT588	H1360A/L1362A/S1363A (F)	CCCAGCCTCATCAGACTGgcCATGgcCgcTTGGCTCCTGGATGAAGAGG
oJT589	H1360A/L1362A/S1363A (R)	CCTCTTCATCCAGGAGCCAAgcGgcCATGgcCAGTCTGATGAGGCTGGG

4.6 Acknowledgments

Patricia Grob and Nicole Haloupek performed all EM analysis and generated reconstructions, with support from Eva Nogales. Jose Ramon Lopez Blanco modeled proteins into EM density, with support from Pablo Chacon. Nicole Haloupek wrote the

initial draft for the manuscript that constitutes most of this chapter. We thank Tom Houweling and Abhiram Chintangal for computer support; Zhiheng Yu, and Chuan Hong for electron microscopy support; Keiichi Namba for the coordinates of the R-type flagellar filament[140]; Ashley Truxal for illustration (Figure 4.12c); and Basil Greber, Robert Louder, Avinash Patel, and Andrew Sandstrom for discussion.

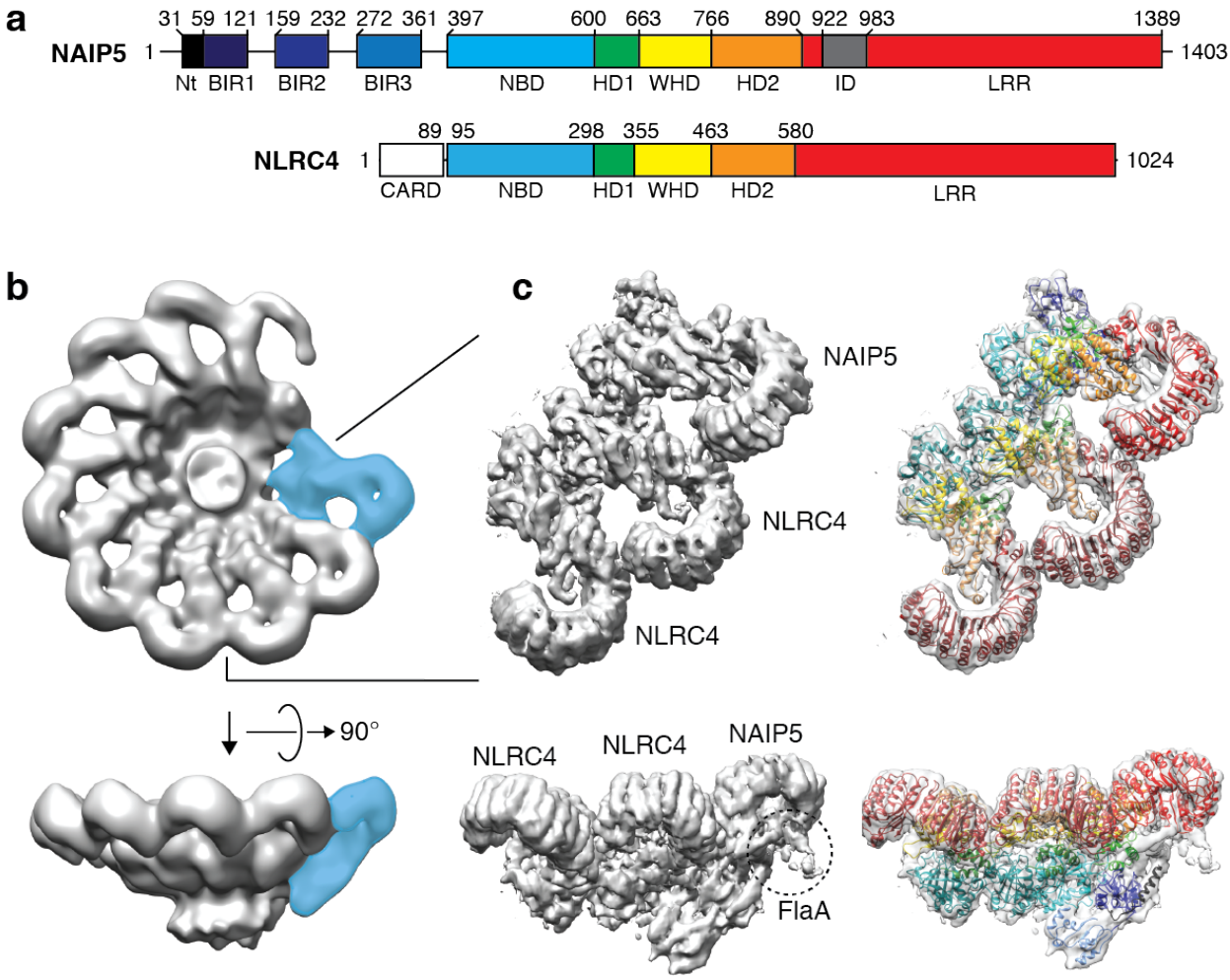


Figure 4.1. Structure of the NAIP5–NLRC4 inflammasome
a, Linear map of NAIP5 and NLRC4 domains. NAIP5 domains were defined here (see Methods); NLRC4 domains were previously defined[159]. Numbers indicate positions of domain boundaries. **b**, 3D reconstruction of inflammasomes containing a single NAIP5–flagellin (blue) and nine NLRC4 protomers (gray). **c**, A 3D refinement on NAIP5 and the first two NLRC4 protomers yielded higher resolution. Right, transparent surface and ribbon diagrams of NAIP5 and NLRC4, obtained by modeling, are superimposed.

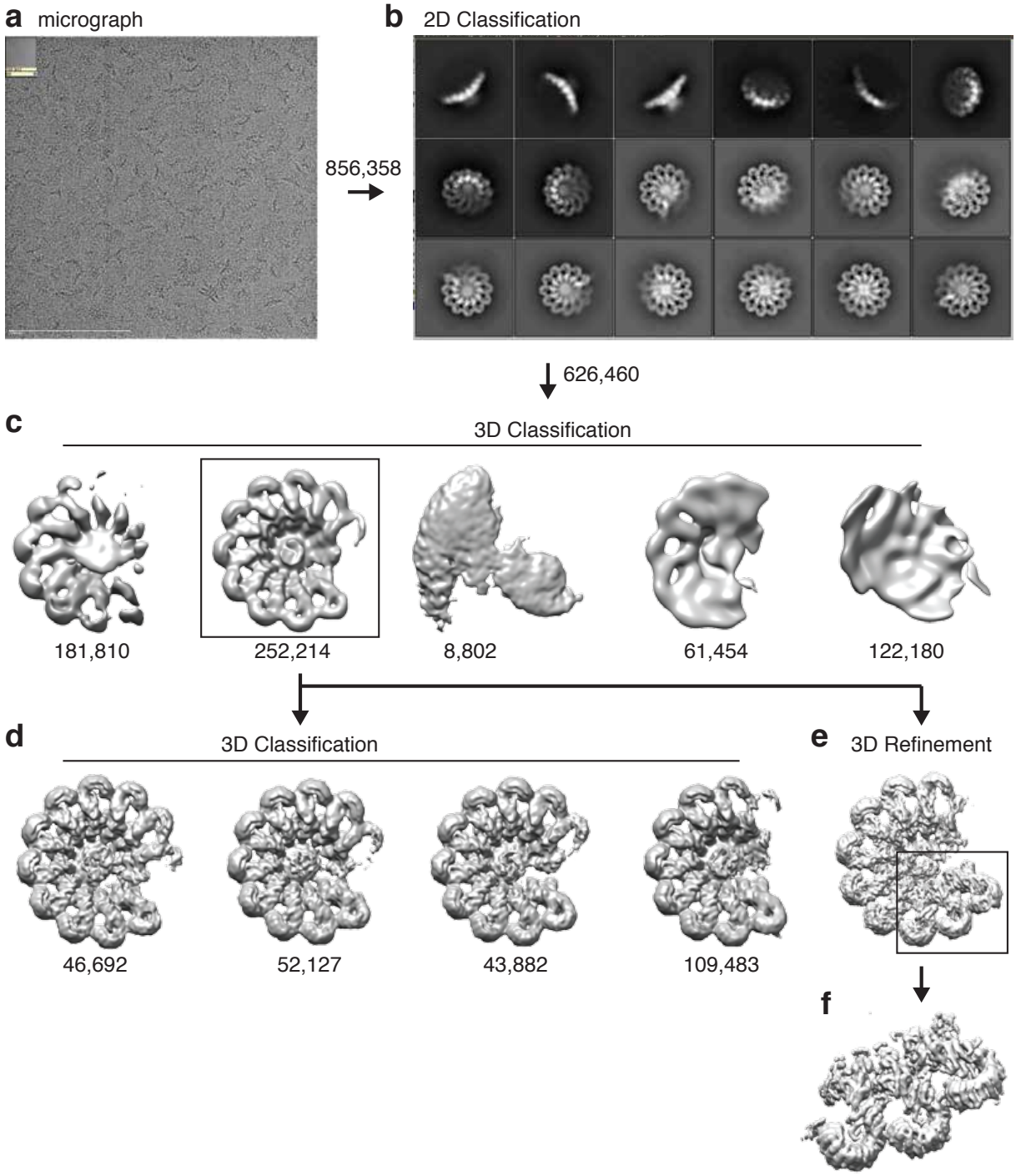


Figure 4.2. Data collection and processing pipeline
a, Representative micrograph. **b**, Representative 2D class averages. **c**, Results from 3D classification into five classes. **d**, Further classification of the most populated class into four new classes did not produce results with higher resolution. **e**, The most populated class (boxed) was used for refinement. **f**, A mask was applied around the first three subunits at the start of local searches, resulting in improved resolution. The number of particles used in each processing step are indicated.

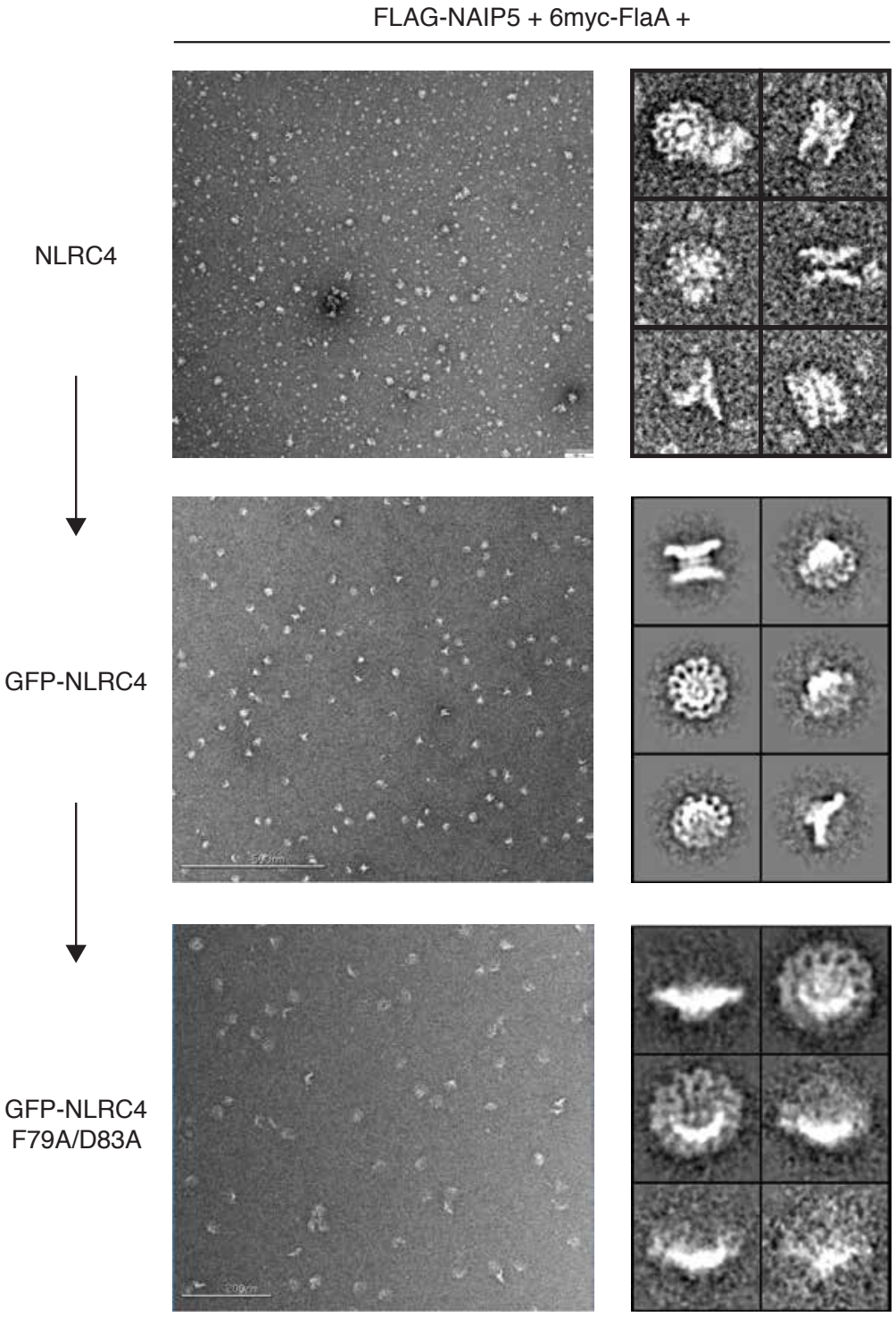


Figure 4.3. Strategy to remove CARD-mediated ring stacking
Micrographs of NLRC4 inflammasomes display stacked rings. An N-terminal GFP, adjacent to the CARD of NLRC4, partially disrupts stacking. Mutation of F79A and D83A further disrupts stacking to yield stable, unstacked rings suitable for cryo-EM analysis.

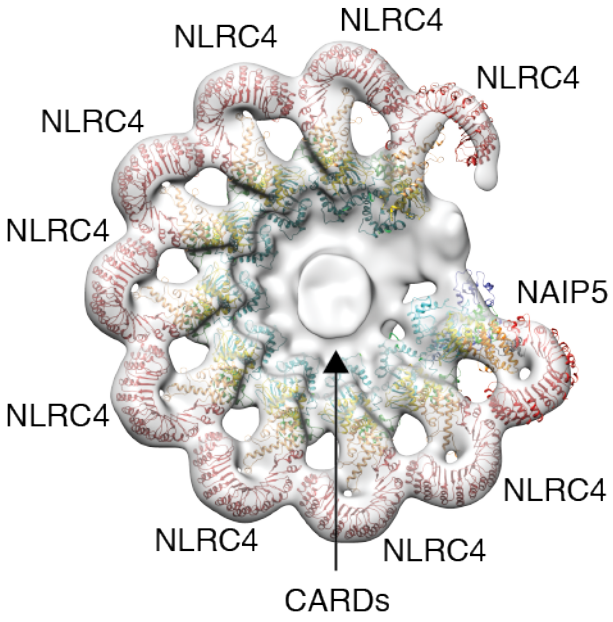


Figure 4.4. CARD density is intact
Superimposing the EM density on the models of NAIP5 and NLRC4 demonstrates the location of the NLRC4 CARDs at the center of the ring.

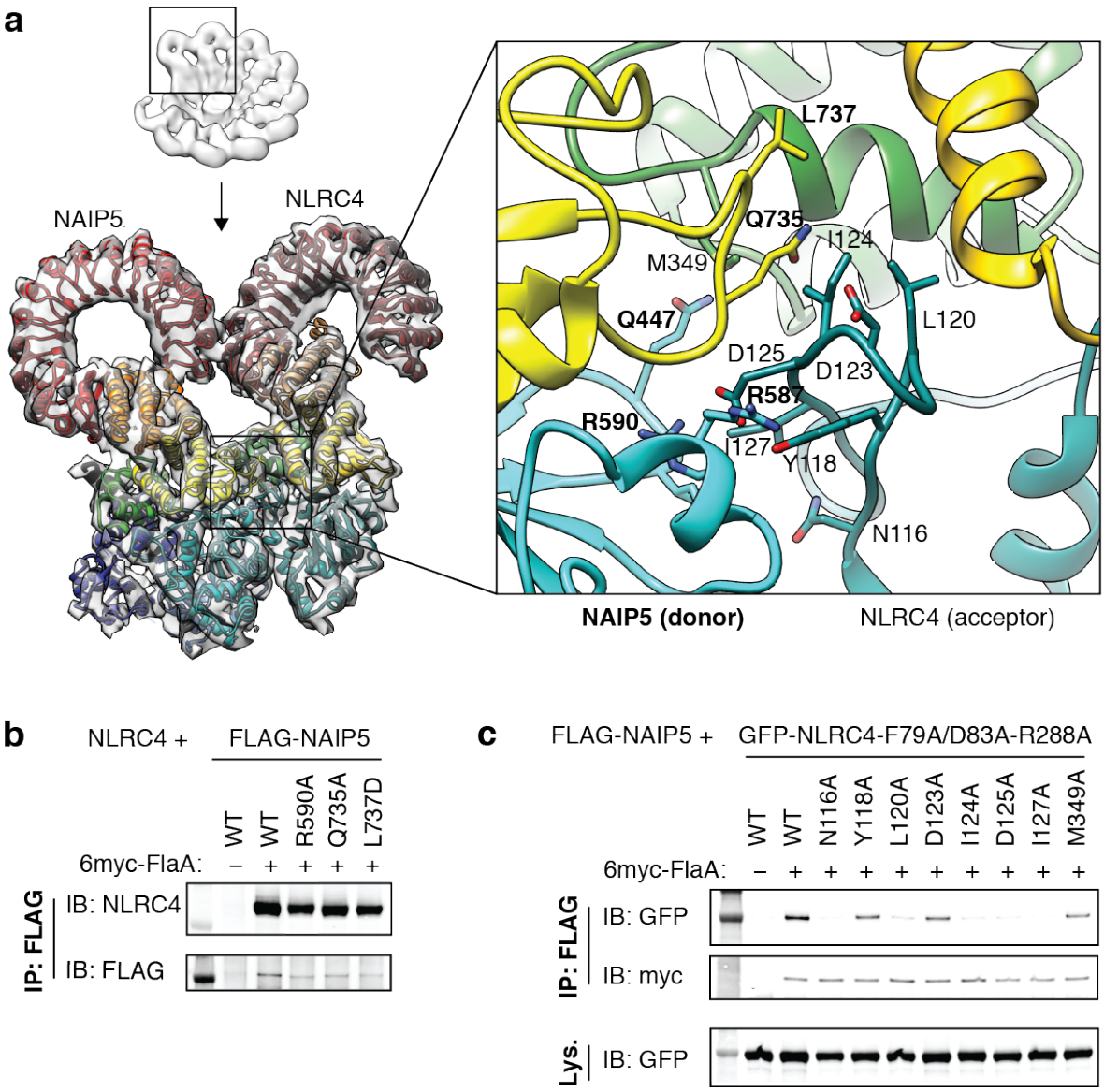


Figure 4.5. Interface between NAIP5 and NLRC4
a, Transparent surface overlaid on ribbon diagrams for NAIP5 and NLRC4. Residues confirmed by mutagenesis (**b-c**) to be important for binding are highlighted, with NAIP5 residues bolded. Side chains shown are from modeling-predicted positions. **b-c**, The indicated constructs were transfected into HEK293T and subjected to anti-FLAG IP. Results are representative of at least three independent experiments. **b**, Mutations in the NAIP5 donor oligomerization surface partially disrupt NLRC4 binding. **c**, Mutations in the NLRC4 acceptor oligomerization surface disrupt binding to NAIP5. An R288A mutation in the NLRC4 donor oligomerization surface prevents recruitment of additional NLRC4 protomers[5] and isolates the binding of a single NLRC4 to NAIP5.

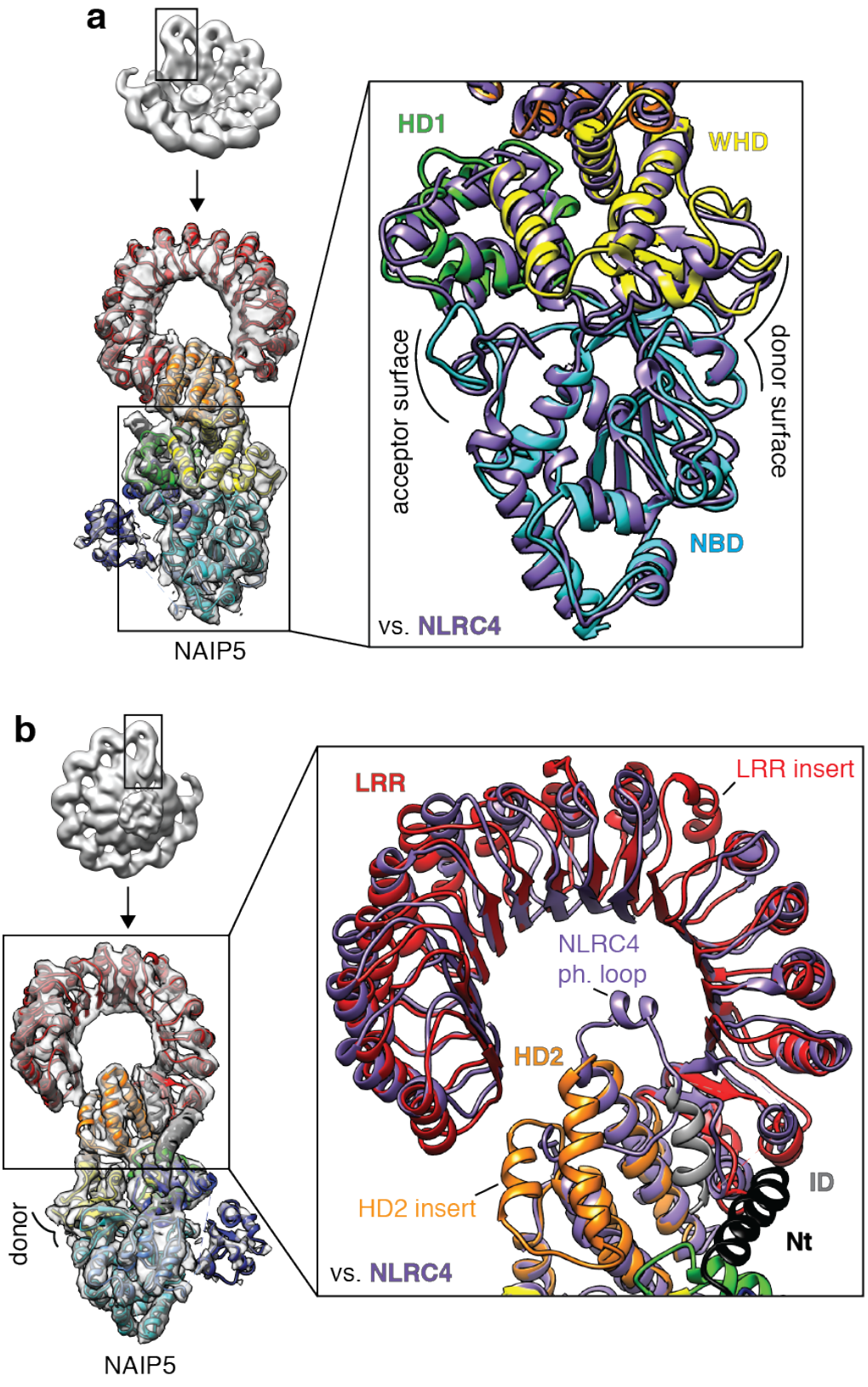


Figure 4.6. Comparison of NAIP5 and NLRC4 in the inflammasome

The structure of NAIP5 (rainbow) was aligned with that of an NLRC4 protomer (purple). **a,** The NBD, HD1 and WHD oligomerization domains of NAIP5 and NLRC4 are highly similar. Oligomerization donor and acceptor surfaces (see [5,112] and Figure 4.5) are indicated. For clarity, NAIP5 BIR domains were omitted from the inset view (right). **b**, The HD2 and LRR of NAIP5 diverge from NLRC4. NAIP5-specific insertions, including an extra leucine-rich repeat and the modeled helix of the ID, are indicated. The NLRC4 S533 phosphorylation loop has been replaced by two alpha helices in NAIP5 (HD2 insert).

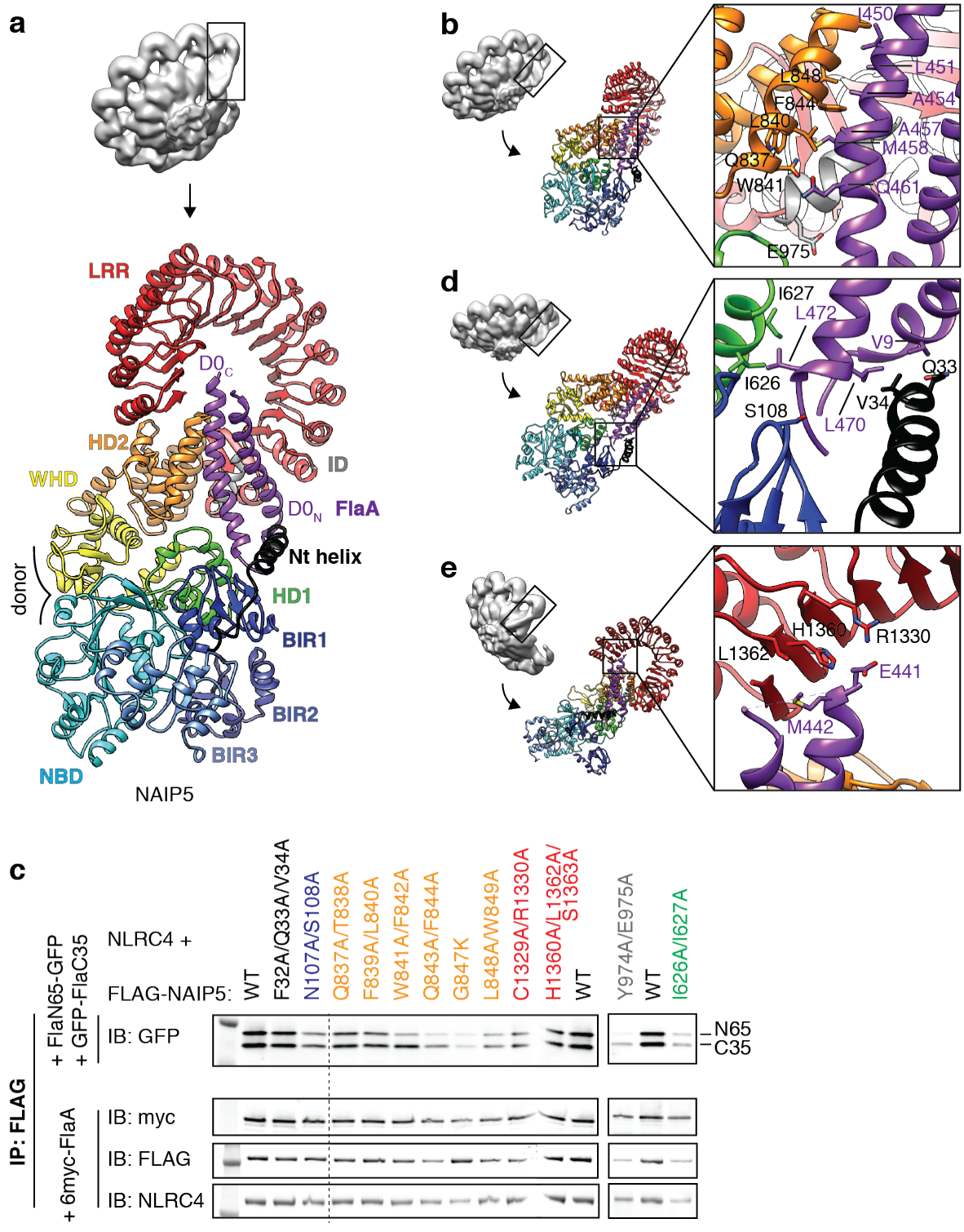


Figure 4.7. Multiple NAIP5 domains contact extended surfaces on both helices of the flagellin D0 domain

a, The flagellin D0 domain (purple) is locked into place by multiple NAIP5 domains. **b, d, e**, Detailed interactions between flagellin D0 helices and NAIP5 domains HD2 and ID (**b**), Nt, BIR1, and HD1 (**d**), and LRR (**e**). Side chains shown are from modeling-predicted positions. **c**, Mutagenesis confirms the importance of NAIP5 residues in binding flagellin. Full-length flagellin (FlaA), or GFP-fused D0 helices (FlaN65 and FlaC35) expressed as separate polypeptides, were tested for co-immunoprecipitation (IP) with NAIP5 from transfected HEK293T cells. FlaN65 binding to NAIP5 requires FlaC35, whereas FlaC35 co-IP is independent of FlaN65 (Chapter 3). IB, immunoblot. Results are representative of at least three independent experiments.

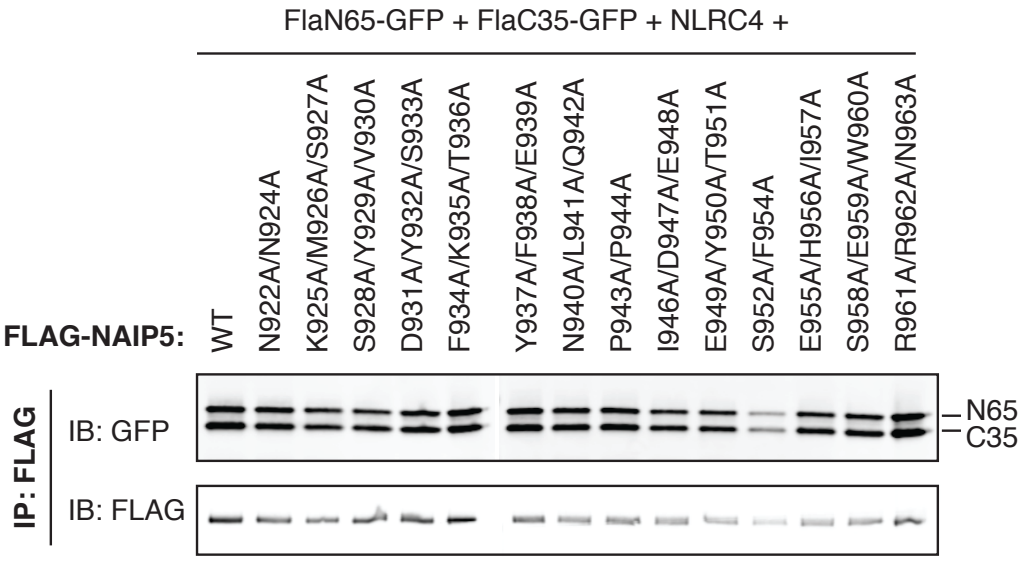


Figure 4.8. Alanine scanning of the unmodeled portions of the ID
The indicated NAIP5 mutants were tested for their ability to co-IP either the $D0_N$ or $D0_C$ helix of flagellin, as in Figure 4.7. The S952A and F954A mutations appear to specifically affect $D0_C$ binding, as $D0_N$ binding was disrupted equivalently. Similar levels of FLAG-NAIP5 IP indicate that all mutants were expressed.

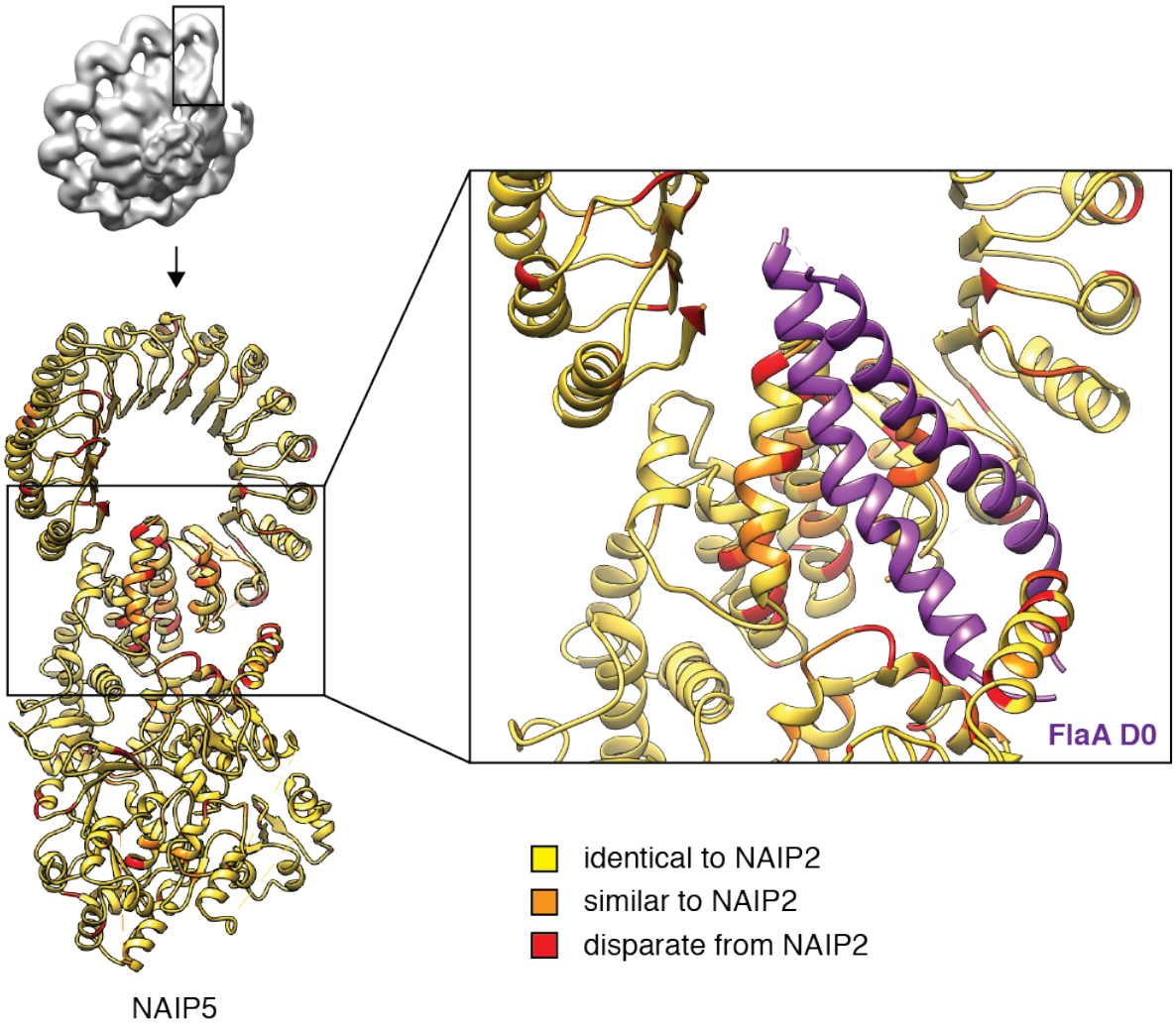


Figure 4.9. NAIP5 divergence from NAIP2 is highest in flagellin-binding regions
Alignment of NAIP5 (Q9R016) and NAIP2 (Q9QUK4) sequences were carried out using the Needle server. The identity is 1171/1448 (80.9%) and similarity 1263/1448 (87.2%). Identical residues are colored yellow, similar residues are colored orange, and disparate residues are colored red. A cluster of the most disparate residues map to sites that form the binding surface for flagellin, suggesting that rod protein ligands may bind NAIP2 in a similar orientation.

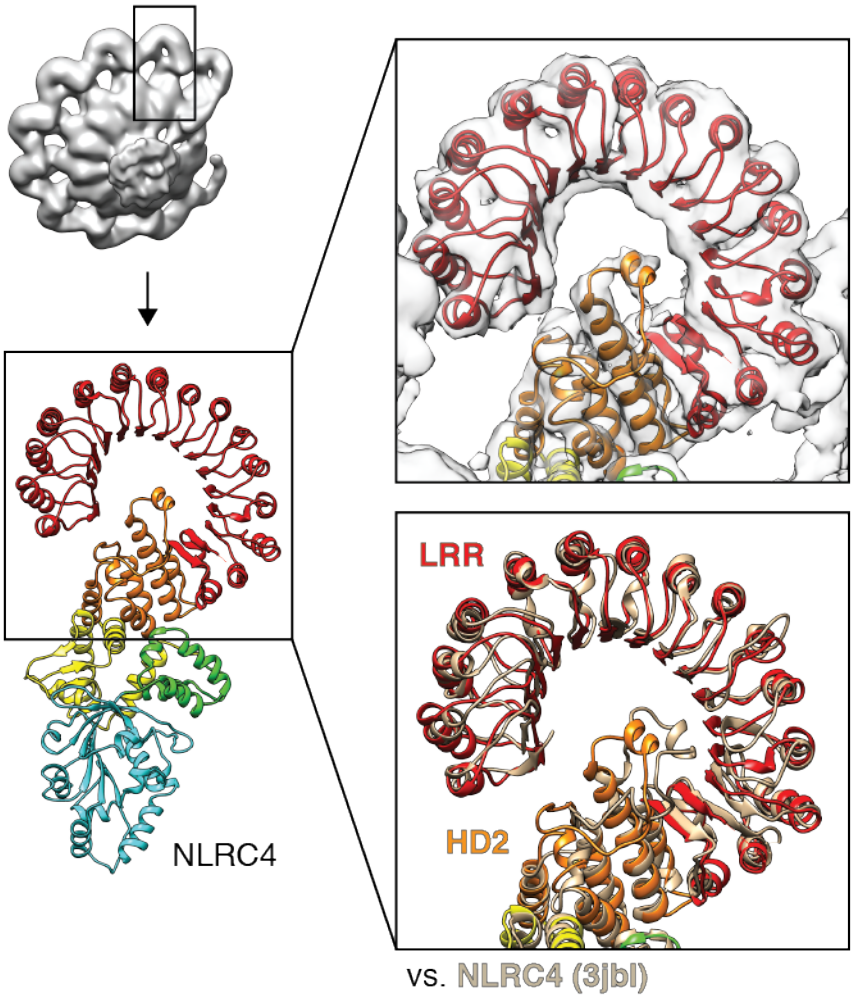


Figure 4.10. Our NLRC4 structure differs from previous NLRC4 structures
In particular, our NLRC4, which has not been averaged with NAIP (as in PDB 3JBL), shows differences in the position of the phosphorylation loop.

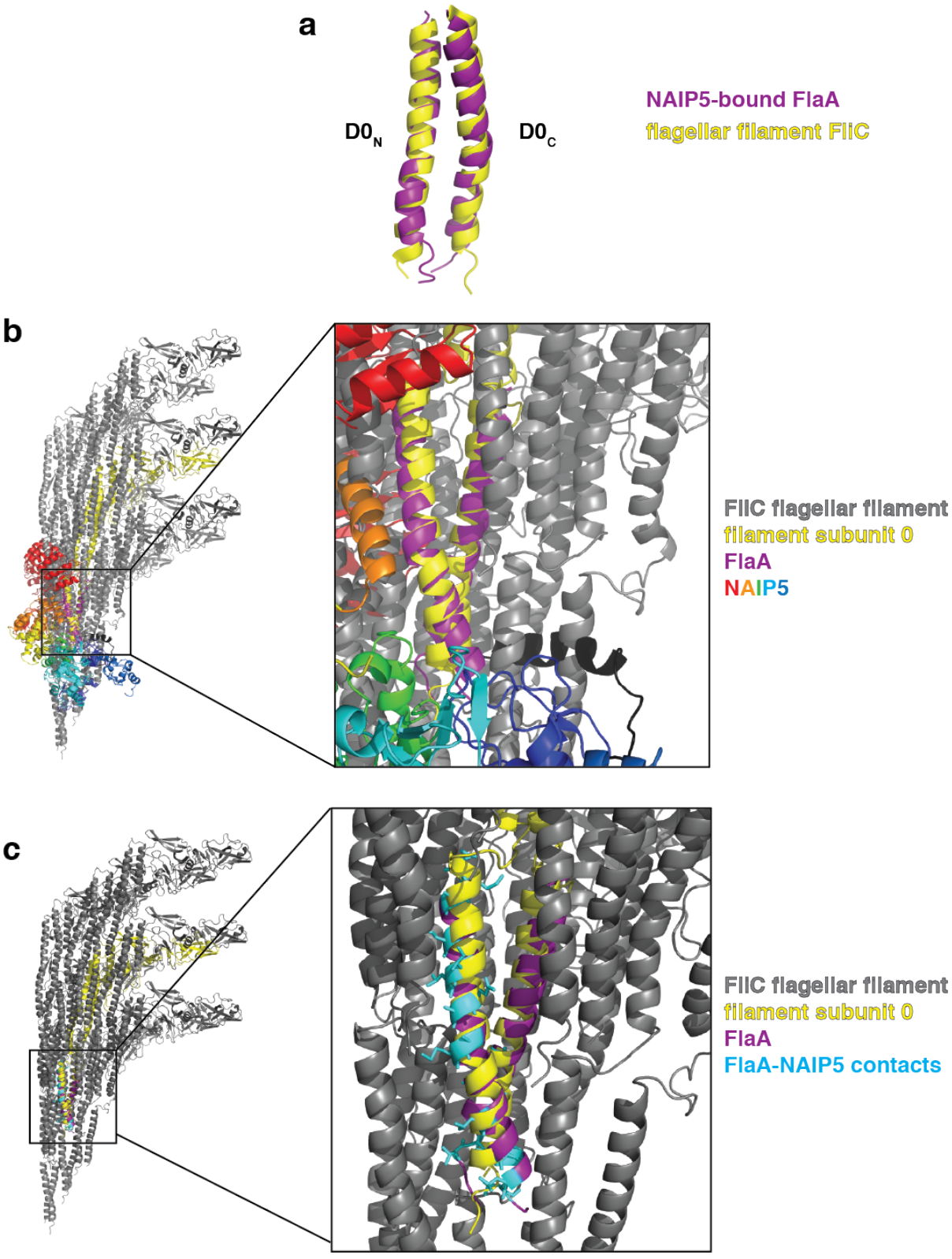


Figure 4.11. Flagellin bound to NAIP5 adopts a similar conformation to flagellin in the flagellar filament

a, Overlay of the D0 domain of *L. pneumophila* flagellin (FlaA) bound to NAIP5 (purple) and *Salmonella typhimurium* flagellin (FliC) in a flagellar filament (PDB 1UCU, yellow). **b**, NAIP5 contacts residues buried in the flagellar filament. NAIP5 and flagellin were aligned to the FliC subunit 0 (yellow) in the flagellar filament (gray)[140]. **c**, For clarity, NAIP5 was removed from the view in (**b**), and NAIP5-contacted residues of flagellin were colored in cyan.

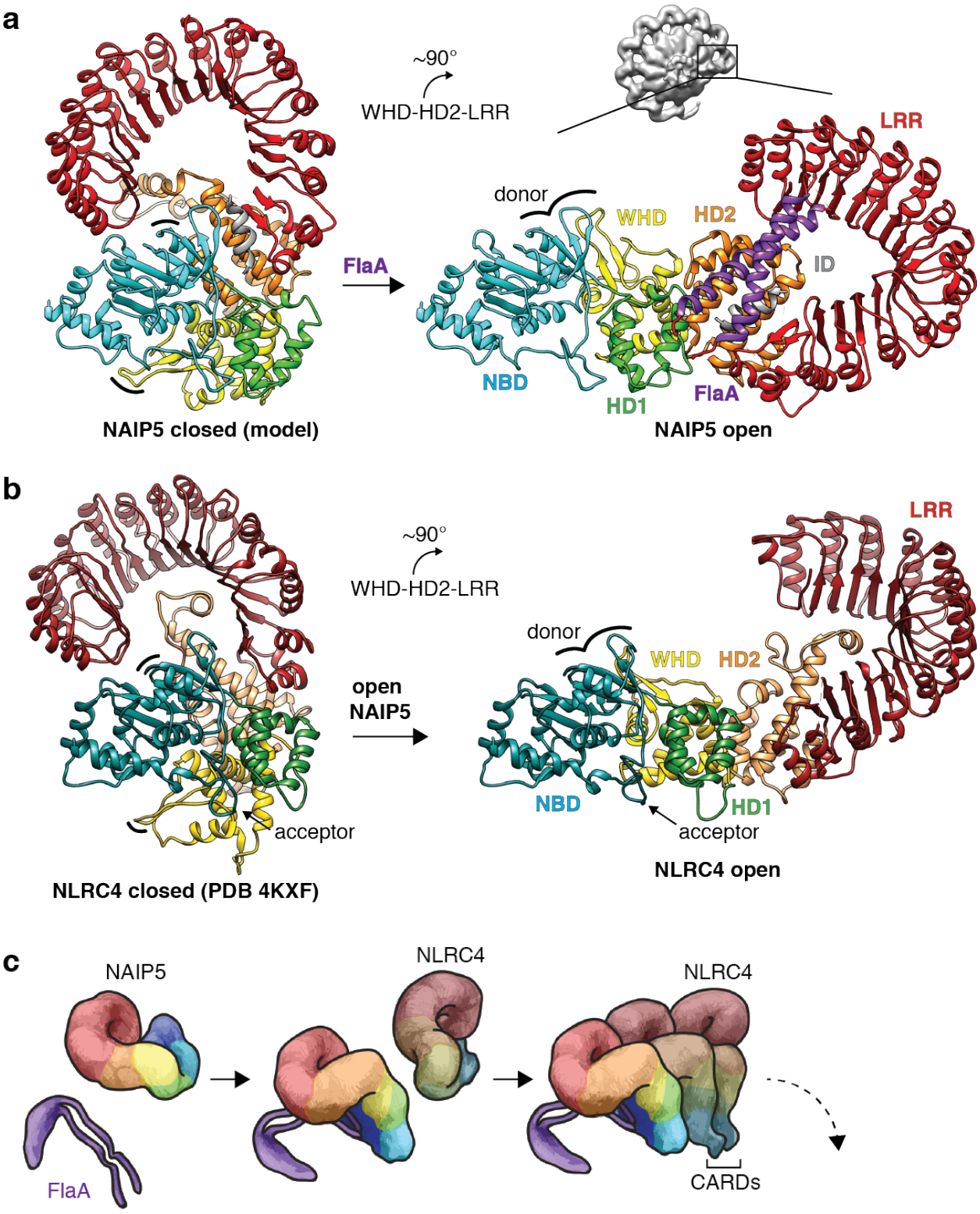


Figure 4.12. Model of NAIP5–NLRC4 inflammasome assembly
a, The NAIP5 HD2, important for flagellin binding, is sterically occluded by the NBD and LRR in the predicted NAIP5 closed confirmation (see Methods). Binding induces a ~90º

rigid-body rotation of the WHD-HD2-LRR module to complete and expose the donor oligomerization surface. **b**, Activated NAIP5 induces an analogous rotation of the NLRC4 WHD-HD2-LRR module[5,112]. **c**, Events of inflammasome assembly. The flagellin D0 domain (purple) binds to NAIP5 and unfurls the protein for subsequent NLRC4 recruitment and activation. Active NLRC4 recruits further NLRC4 protomers for self-propagating oligomerization and completion of a Caspase-1 recruitment platform.

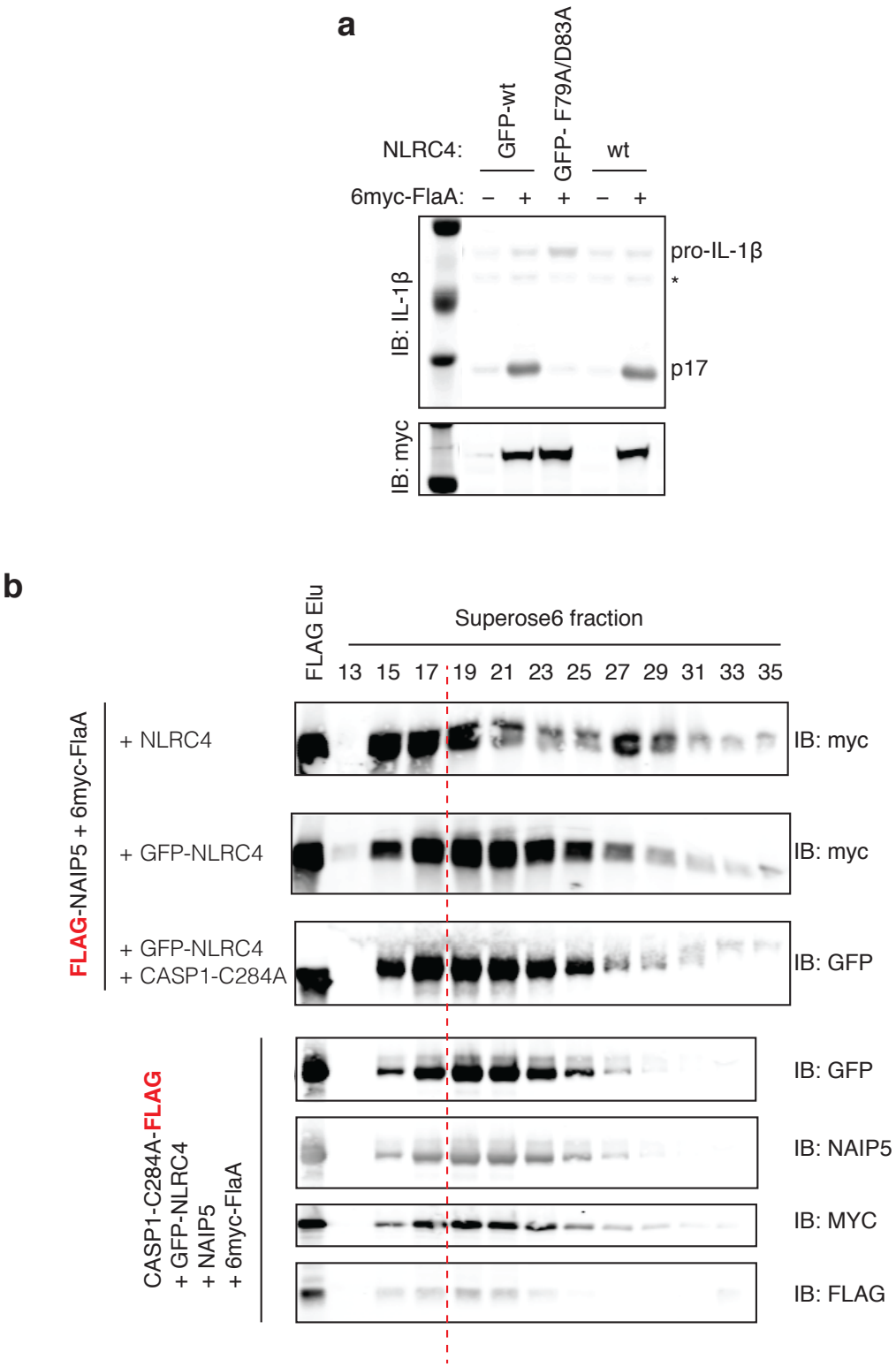


Figure 4.13. Modified strategy to purify CASP1-associated inflammasomes
a, F79A/D83A-NLRC4 cannot activate CASP1, but GFP-NLRC4 can. The indicated constructs were transfected into HEK293T, and cells lysates were probed for IL-1β processing 24 hr after transfection. *, background. Results representative of 3

independent experiments. **b**, Purification of inflammasomes by FLAG-tagged CASP1-C284A is sufficient to disrupt residual aggregation of GFP-NLRC4 inflammasomes, whereas co-expression of CASP1-C284A is insufficient. The indicated constructs were transfected into HEK293T, and complexes were purified via FLAG immunoaffinity purification at 48 hr after transfection. FLAG purification eluates (FLAG Elu) were separated by Superose6 size exclusion chromatography. Particles eluting in fractions 19-21, right of the red line, are typically single discs.

CASP1-C284A-FLAG + GFP-NLRC4 + HA-NAIP5 + 6myc-FlaA

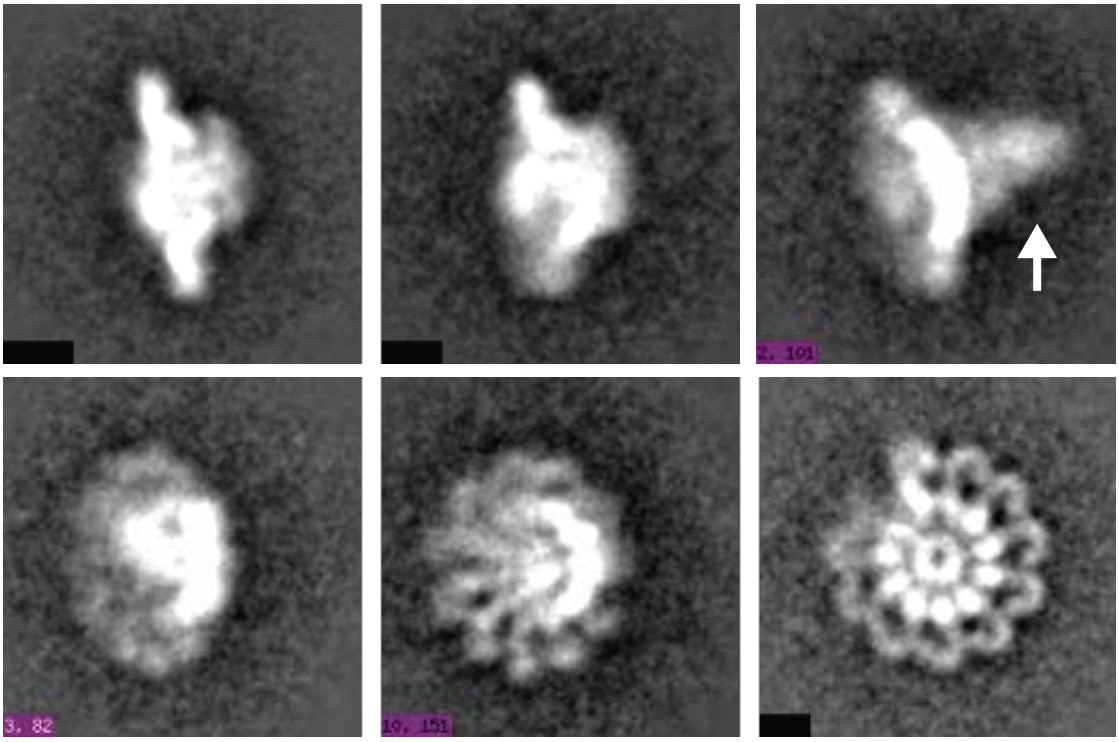


Figure 4.14. CASP1-associated inflammasomes reveal CASP1 filaments
The indicated components were transfected into HEK293T, and inflammasomes were isolated by FLAG immunoaffinity purification followed by Superose6 size exclusion chromatography. Representative class averages from negative stain EM are shown. Arrow indicates a CASP1 filament below the NLRC4 CARD ring.

Chapter Five: Tools to Identify Novel Cofactors and Regulators of NAIP–NLRC4 Inflammasomes

5.1 Introduction

The above studies have investigated the interactions of the minimal set of proteins required to initiate NAIP–NLRC4 inflammasome signaling in response to the cytosolic presence of bacterial pathogens. Several lines of evidence suggest that additional proteins may be involved in either augmenting or regulating that response, in order to ensure the appropriate level of signaling, as discussed below. This chapter therefore aims to generate tools to identify such associated or regulatory factors in an unbiased manner.

5.1.1 Recruitment of non-cognate NLRs

Mammalian genomes encode numerous NLRs that share a core domain architecture. The most conserved of these shared domains is the NBD, the domain that mediates NLR oligomerization to initiate signaling (Figure 5.1a). Thus, it is possible that other NLRs are capable of associating with an assembling NAIP–NLRC4 inflammasome, in much the same way that NLRC4 interacts heterotypically with NAIP. There is some evidence that NLRC4 can associate with heterotypic NBDs from other NLRs, at least in an overexpression setting[233]. However, the 'acceptor' and 'donor' residues that are critical for NLRC4 homotypic interactions are poorly conserved among NLRs (Figure 5.1b). It is therefore unclear whether heterologous NLR interactions occur at endogenous levels in immune sentinel cells.

It is further unclear whether recruitment of non-cognate NLRs would facilitate or inhibit NLRC4 signaling. More promiscuous assembly of a CASP1-activating platform might increase the rate of assembly or the total number of inflammasomes formed following infection. However, it is equally plausible that heterotypic NLRs might be incapable of dimerizing CASP1 with NLRC4. Many NLRs contain a PYD signaling domain, and this domain, unlike the NLRC4 CARD, requires the adaptor ASC to mediate interactions with CASP1[81,113]. Furthermore, even those NLRs that contain a CARD may not position this domain appropriately relative to the NLRC4 CARD, as linker length or even domain ordering can vary. Several NLRs are not basally expressed and require transcriptional priming[80]. Thus, the prior exposure of immune cells to priming signals might also affect the output of NAIP–NLRC4 inflammasome signaling.

In support of the possibility that other NLRs can be recruited to the NAIP–NLRC4 inflammasome, *S. typhimurium* infection activates both NAIP–NLRC4 and NLRP3, and both NLRs have been reported to colocalize in a ~1 μm ASC-mediated inflammasome aggregate[234]. Activated NLRC4 can also co-immunoprecipitate NLRP3 from macrophages, though it is unclear whether this association is direct or mediated by ASC[235]. Interestingly, some pathogens appear to activate NAIP–NLRC4 but not NLRP3, even after NLRP3 has been transcriptionally primed[236]. This may be indicative of a

requirement for each NLRP3 protomer to detect its cognate stimulus, analogous to the NAIP sensor (Chapter 2) rather than the NLRC4 signal-propagation adaptor.

5.1.2 Recruitment of additional caspases

Oligomerization of NLRC4 promotes the activation of CASP1, but several groups have established that it can also recruit and activate CASPASE-8 through the adaptor ASC. CASPASE-8 was found to colocalize with CASP1 and ASC following NAIP–NLRC4 activation[237]. This recruitment induced CASPASE-8-dependent apoptosis that was uncovered in the absence of CASP1-dependent pyroptosis[153]. CASPASE-8 also allowed for the release of pro-inflammatory mediators like eicosanoids following NAIP5 stimulation[153]. Thus, CASPASE-8 activation may be an important backup response against pathogens that inhibit CASP1 signaling. In contrast, CASPASE-12 has been reported to bind CASP1 and suppress its enzymatic activity, possibly by preventing CASP1 homodimerization[238]. However, *Caspase-12*$^{-/-}$ mice do not exhibit enhanced CASP1 signaling after stimulation of NAIP–NLRC4 or other inflammasomes, calling this finding into question[239].

5.1.3 Negative regulators of inflammasomes

Several CARD- or PYD-only proteins (COPs and POPs, respectively) have been reported to inhibit inflammasome activity by sequestering either the CASP1 CARD or the ASC PYD[240]. These include ICEBERG, COP1, and INCA, all of which are highly similar to the CASP1 CARD domain, as well as POP1 and POP2, which are related to the ASC PYD. Intriguingly, COP and POP negative regulators are found in humans but not in mice[240], suggesting that mice require less regulation of inflammasomes or possess alternate negative regulators.

In contrast to these species-specific regulators, inflammasome signaling in both mice and humans appears to be temperature-sensitive. This temperature dependence was uncovered in human patients with constitutively active alleles of NLRP3 or NLRC4, whose symptoms are often initiated or aggravated by cold exposure[155,241]. Strikingly, mice with an autoactive *Nlrc4* transgene also exhibited cold-induced autoinflammation[155]. It is currently unclear why cold temperatures exacerbate inflammasome signaling. One possibility is that NLRs are intrinsically more prone to spontaneous aggregation when cold. In support of this possibility, constitutively active alleles of plant NLRs can also exhibit increased signaling at cold temperatures[242]. Alternatively, the interaction of NLRs with sequestering chaperones or negative regulators may be disrupted by cold exposure. Not all autoactive inflammasome alleles induce cold-responsive inflammation[154,156], suggesting that there is some specificity to the cold-sensitive phenotype. A plausible mechanism for this specificity could be the alteration of post-translational modifications at distinct temperatures. For example, mutations adjacent to a reported phosphorylation site in NLRP3 are associated with cold-induced autoinflammation in humans[243].

5.1.4 NLRC4 Phosphorylation as a Potential Licensing Strategy

Phosphorylation is by far the most frequently reported post-translational modification of inflammasomes[244]. NLR phosphorylation can have opposite effects on inflammasome assembly or signaling. Phosphorylation of pyrin by PKN1 or PKN2 increases its association with negative regulatory proteins, and the manipulation of Rho GTPases inhibits PKN1/2 activity to trigger pyrin activation[245]. NLRP3 phosphorylation by PKA is also inhibitory[243], and dephosphorylation of a different NLRP3 residue by PTPN22 is required for efficient signaling[246]. Conversely, multiple kinases have been implicated in positively regulating NLRP3 responses, although it is unclear whether these kinases directly phosphorylate NLRP3[246]. Recently, CRISPR and ENU mutagenesis screens, as well as NLRP3 pull-down with mass spectrometry, identified the NEK7 kinase as required for NLRP3 inflammasome signaling[247-249]. Surprisingly, the catalytic activity of NEK7 is dispensable for NLRP3 activity, indicating that NEK7 physical association was critical for its role in inflammasome assembly.

Activated NLRC4 was found to contain a single post-translational modification, the phosphorylation of Ser533[218]. Phosphorylation of this residue, possibly mediated by PKCδ, was reported to occur only upon infection and to promote NLRC4 signaling[218,250]. Given that NAIP5 is dispensable for NLRC4 phosphorylation in response to flagellin[250], phosphorylation may be downstream of TLR priming signals. Indeed, mutation of NLRC4 Ser533 to Ala reduced the inflammasome response to flagellin only when macrophages were primed with TLR ligands. Thus, NLRC4 phosphorylation appears to license signaling in primed macrophages but is completely dispensable in unprimed macrophages. One possible way to explain this apparent discrepancy is the requirement for phosphorylation to disrupt NLRC4 interaction with a negative regulator that is only present in primed cells.

Confusing matters further, monomeric NLRC4 purified from insect cells is phosphorylated in its inactive and closed conformation[159]. These data indicate that phosphorylation can occur in the absence of infection or priming, at least in heterologous expression settings, and that it is insufficient to induce NLRC4 oligomerization and signaling. Due to resolution constraints, it is unclear whether NLRC4 remains phosphorylated in assembled inflammasomes ([5,112] and Chapter 4). Thus, it remains possible that the observed phosphorylation of NLRC4 is not a critical regulatory component of the inflammasome but simply an artifact of stress- or lysis-induced kinase activation.

5.2 Results

5.2.1 NLRC4 phosphorylation does not affect signaling in response to NAIP ligands

Phosphorylation of NLRC4 at Ser533 was reported to be both necessary and sufficient for inflammasome activation[218]. To determine at what stage phosphorylation affects inflammasome signaling, S533A (non-phosphorylatable) and S533D (phosphomimetic)

mutants of NLRC4 were tested for inflammasome assembly and signaling in HEK293T cells. Based on previous reports, the S533A mutation is predicted to be incapable of assembling an inflammasome, while the S533D mutant should oligomerize regardless of the presence or absence of an activating ligand. However, both the S533A and S533D mutants of NLRC4 assembled an inflammasome normally in response to the NAIP5 ligand, flagellin (Figure 5.2a). Unlike truncation of the LRR domain that results in constitutive NLRC4 oligomerization[3], S533D did not induce spontaneous NLRC4 inflammasome assembly. It is possible that phosphorylation affects NLRC4 signaling downstream of inflammasome assembly, potentially via recruitment of CASP1. However, neither the S533A nor S533D mutation affected CASP1 activation, as assessed by the cleavage of pro-IL-1β into its mature p17 fragment (Figure 5.2b). Furthermore, neither mutation affected the constitutive signaling of NLRC4ΔLRR. Thus, phosphorylation of NLRC4 appears to be dispensable for inflammasome signaling in a heterologous expression system.

We reasoned that overexpression of inflammasome components might mask a partial requirement for NLRC4 phosphorylation. Using CRISPR/Cas9 targeting with homology-directed repair, we therefore generated mice in which both alleles of the endogenous *Nlrc4* gene were edited to either S533A or S533D. These mice enabled us to analyze the effect of non-phosphorylatable or phosphomimetic mutants of NLRC4 at endogenous expression levels. We first noted that both the S533A and S533D mouse lines were viable as homozygotes and displayed no obvious autoimmune phenotypes like joint swelling or insufficient weight gain. The lack of overt phenotype for S533D heterozygous or homozygous mice was somewhat unexpected given the reported inability to transduce bone marrow macrophages (BMM) with an S533D *Nlrc4* allele, purportedly due to the constitutive induction of cell death[218].

We next assessed the ability of these NLRC4 mutants to signal in BMM. Both NLRC4 alleles were expressed at roughly wild-type levels (Figure 5.3a). Flagellin was delivered to the cytosol of BMM using *Bacillus anthracis* protective antigen (PA) pores and a targeting sequence from lethal factor (LF), a PA translocation target[107]. At a low dose of flagellin (LFn-FlaA) that induced ~70% of B6 BMM to undergo pyroptosis, S533A/A BMM exhibited a partial defect in NLRC4-mediated cell death (Figure 5.3b). However, it should be noted that this is a preliminary result, and a similar defect was also observed for S533D/D BMM in this experiment only (compare Figure 5b, c). The partial defect in S533A/A signaling is consistent with a recent report suggesting that NLRC4 phosphorylation is only partially required for the response to flagellin in BMM[235]. However, in contrast to this report, we observed this slight defect even in the absence of priming with TLR ligands. Careful dissection of NLRC4 S533A signaling with flagellin dose response curves, with or without prior TLR priming, will be required to confirm these findings.

Consistent with the HEK293T inflammasome reconstitution assay and the viability of S533D/D mice, S533D/D BMM did not induce constitutive pyroptosis in the absence of

flagellin (Figure 5.3b, c). Indeed, the S533D mutation appeared to have little to no effect on NLRC4 responses to cytosolic flagellin. There was a very slight defect in flagellin responsiveness in primed S533D BMM, although this result was not consistent between experiments (compare Figure 5.3b, c). We therefore conclude that the NLRC4 S533D mutant is not constitutively active, as previously suggested. Formally, we cannot conclude that phosphorylation does not result in constitutive NLRC4 signaling, as an Asp phosphomimetic does not always function equivalently to a phosphorylated Ser. However, the observation of phosphorylated S533 in the crystal structure of monomeric NLRC4 strongly suggests that phosphorylation is insufficient to induce signaling[159].

5.2.2 Nlrc4 H443P is a hypomorphic allele

Recently, several constitutively active alleles of *Nlrc4* were identified in human patients, but only one allele was associated with cold-induced signaling[154-158]. This pattern is strikingly reminiscent of autoactive *Nlrp3* alleles, only some of which cause FCAS (familial cold autoinflammatory syndrome)[243]. The cold-associated H443P allele of *Nlrc4* was also shown to induce inflammation within minutes of cold exposure in mice, at least when expressed as a transgene in addition to endogenous *Nlrc4*[155]. To investigate the hypothesis that the temperature dependence of NLRC4 H443P signaling is due to the loss of interaction with a negative regulator at cold temperatures, we generated *Nlrc4* H443P knock-in mice at the endogenous *Nlrc4* locus using CRISRP/Cas9 and homology-directed repair.

H443P mice were viable and did not exhibit any overt signs of inflammation, either as heterozygous or homozygous knock-in mice of either sex. Surprisingly, BMM from homozygous H443P/P were largely unable to signal in response to cytosolic flagellin (Figure 5.3b, c). This defect may be due to significantly decreased steady state levels of NLRC4 H443P in BMM (Figure 5.3a), suggesting that this non-conservative mutation destabilizes the protein. It is also possible, though unlikely, that CRISRP targeting induced an additional mutation in *Nlrc4* distal to the repaired H443P mutation. It will therefore be necessary to confirm the sequence of the entire *Nlrc4* coding sequence, outside of the ~600 base pairs sequenced surrounding codon 443, and to confirm that mRNA levels are normal.

Despite the apparent inability of the H443P knock-in allele to signal, overexpression of this allele induced at least some signaling in BMM[155], suggesting that our knock-in allele might be hypomorphic rather than a null allele. We therefore investigated the possibility that NLRC4 H443P was still capable of mediating ligand-independent signaling at cold temperatures. Incubation of B6 macrophages at temperatures ranging from 22 – 37 ºC did not induce ligand-independent cell death, although flagellin-dependent pyroptosis was temperature-sensitive (Figure 5.4). Incubation of H443P heterozygous or homozygous BMM at < 37 ºC also did not induce constitutive cell death, although it is possible that colder temperatures are required for significant response. Notably, H443P/P BMM exhibited increased cell death in response to flagellin at 27 – 32 ºC relative to 37 ºC. This preliminary finding suggests that H443P may retain slight

functionality and hints that this signaling may be cold-responsive. However, further testing is needed to verify these results.

5.2.3 New Nlrc4$^{-/-}$ mouse lines as a byproduct of knock-in CRISPR targeting

As a byproduct of Cas9 targeting of *Nlrc4* at S533 and H443, we recovered one allele from each targeting that failed to repair the double strand break by homology directed repair and instead accrued either a +1 or -1 frameshift mutation in *Nlrc4* (Figure 5.5). These mutations resulted in premature stop codons at position 542 or 459, yielding truncations in the middle of or just before the NLRC4 HD2, respectively. Both *Nlrc4* 542stop and 459stop appear to be null alleles, as indicated by the inability of these truncated proteins to induce cell death in response to cytosolic flagellin (Figure 5.3, 5.4). Indeed, both alleles behaved indistinguishably in this assay from a previously reported *Nlrc4* null allele[207]. Unlike the established *Nlrc4*$^{-/-}$ line, however, the new *Nlrc4*$^{-/-}$ mice were generated on a C57BL/6J rather than C57BL/6N background.

The inability of NLRC4 459stop BMM to induce cell death is somewhat at odds with a report that NLRC4 residues 1-463 are capable of constitutive signaling when overexpressed[159]. It is formally possible that the 459stop allele is spontaneously active and results in the death of any BMM that expresses NLRC4 prior to flagellin stimulation, although normal numbers of BMM were obtained during differentiation of 459stop bone marrow. Furthermore, the lack of autoimmune phenotype for heterozygous and homozygous *Nlrc4* 459stop mice, in contrast to the severe phenotypes observed upon systemic activation of wild-type NLRC4[107], suggests that this scenario is highly unlikely. By contrast, truncation of NLRC4 in the middle of a domain is expected to destabilize the protein, so the inability of the 542stop allele to respond to flagellin is unsurprising.

We used these new NLRC4-deficient mice to address an unexpected report that NLRC4 suppresses the growth of B16F10 melanomas[162]. In this study, the authors observed that *Nlrc4*$^{-/-}$ mice had enhanced tumor growth, whereas *Casp1*$^{-/-}$ mice did not. Because NLRC4 is involved in sensing of bacterial pathogens, we hypothesized that the observed difference may have been an indirect consequence of a change in the microbiota in this study's *Nlrc4*$^{-/-}$ colony but not in the separately maintained B6 or *Casp1*$^{-/-}$ colonies. We therefore crossed heterozygous *Nlrc4* 459stop and 542stop mice and maintained littermates in cohoused cages to normalize the microbiota between NLRC4-sufficient and -deficient mice. Consistent with the hypothesis, in a pilot experiment *Nlrc4*$^{-/-}$ mice were equally susceptible to melanoma as their cohoused littermate controls (Figure 5.6).

5.2.4 Purification of endogenous inflammasomes to identify associated cofactors

NAIP inflammasome signaling can be reconstituted in HEK293T cells by the expression of a NAIP, its cognate ligand, NLRC4, and CASP1 ([3] and Figure 5.2). These data suggest that no additional cofactors are strictly required for signaling, but it is difficult to rule out that HEK293T express the required cofactors. Furthermore, overexpression of inflammasome constituents may override the need for cofactors. Indeed, IL-1β

processing appears to be independent of ASC in HEK293T but not in BMM[113]. The recent identification of NEK7 as a required binding partner for the NLRP3 inflammasome[247-249] provides precedent for the idea that inflammasome assembly or signaling in endogenous settings requires non-NLR proteins. To identify any such factors that are required for NAIP–NLRC4 inflammasome signaling, we took a biochemical approach and attempted to purify assembled NAIP–NLRC4 inflammasomes, along with any associated cofactors, from BMM.

In order to purify inflammasomes in sufficient quantity, we used primary or immortalized *Casp1*$^{-/-}$ macrophages to block pyroptosis downstream of inflammasome assembly. We additionally removed ASC to prevent the aggregation of inflammasomes into a biochemically intractable ~1 μm 'ASC speck'. ASC-deficiency also prevents CASPASE-8-mediated apoptosis downstream of NAIP activation[153]. These BMM are hereafter referred to as *ASC/C1*$^{-/-}$ or *iASC/C1*$^{-/-}$. In Chapter 4 we showed that it was possible to purify inflammasomes using an N-terminal FLAG epitope tag on NAIP5, followed by size exclusion chromatography to separate inflammasomes from monomeric NAIP5. Rather than modifying the genomic *Naip5* locus in *iASC/C1*$^{-/-}$ BMM to include a FLAG tag, we took a simpler approach and introduced an epitope tag on exogenously delivered flagellin. An N-terminal epitope tag on flagellin does not interfere with inflammasome assembly and can be used to purify the NAIP5 inflammasome[3].

We first introduced flagellin into *iASC/C1*$^{-/-}$ BMM by lentiviral or retroviral transduction. A dox-inducible FLAG-tagged flagellin was capable of inducing cell death in transduced (GFP^{+}) iB6 but not *iASC/C1*$^{-/-}$ BMM (Figure 5.7a). The loss of GFP^{+} cells was specific to NAIP5 inflammasome activation, as death required the C-terminal 35 amino acids of flagellin. However, induction of flagellin expression in *iASC/C1*$^{-/-}$ BMM was undetectable by anti-FLAG immunoblot of lysates or by silver stain following FLAG purification (Figure 5.7b). To increase flagellin levels we next used a constitutively expressing, puromycin-selectable retrovirus. Pull down of constitutively expressed FLAG-FlaA, but not FLAG-FlaAΔC35, captured endogenous NAIP5 to an amount detectable by immunoblot but not colloidal Coomassie (Figure 5.8). Thus flagellin transduction seems to yield insufficient expression for mass spectrometry analysis of purified inflammasome components.

As an alternate strategy, we delivered flagellin via PA and fusion to LFn[107]. His_6-LFn-FlaA was translocated into *ASC/C1*$^{-/-}$ BMM and then purified out of these cells by nickel resin chromatography. Although some His_6-LFn-FlaA was recovered, very little NAIP5 or NLRC4 co-purified with flagellin (Figure 5.9a). There was also little evidence of other proteins that co-purified with assembled inflammasomes but not with a FlaA-3A negative control (Figure 5.9b). PA-mediated delivery of LFn-FlaA to the cytosol appeared to be very inefficient (Figure 5.9a, IB: His). We reasoned that the relatively large size of LFn-FlaA might prevent its efficient translocation through the PA pore, and that smaller NAIP ligands like T3SS components might be more efficiently delivered. Consistent with this hypothesis, LFn-mediated delivery of a T3SS rod (PrgJ) was

dramatically more potent than FlaA (Table 5.1). Furthermore, a truncated form of FlaA that lacks the $D0_N$ (Fla166) induced cell death at the same dose as full-length FlaA, despite its incomplete complement of NAIP5 recognition motifs. The isolated D0 domain of flagellin (LFn-FlaN52-SGSGSG-FlaC45) might therefore be expected to be of similar potency to LFn-PrgJ.

We attempted to increase yield by purifying NAIP2 inflammasomes with LFn-PrgJ. Nickel resin purification of LFn-PrgJ out of *iASC/C1*$^{-/-}$ BMM was still inefficient at isolating NAIP–NLRC4 inflammasomes (Figure 5.9c). While His_6 tagged proteins can be purified from *E. coli* with reasonable specificity, nickel resin purification is typically less efficient in eukaryotic cells. We therefore added a FLAG epitope tag to His_6-LFn-ligands for more specific FLAG immunoaffinity purification. The addition of a FLAG epitope tag did not prevent LFn-FlaA or -PrgJ from inducing cell death in B6 BMM (Table 5.1). Therefore, FLAG-LFn-PrgJ, FLAG-LFn-FlaA, or FLAG-LFn-FlaA-D0 may be useful tools for isolation of NAIP inflammasomes from *ASC/C1*$^{-/-}$ BMM.

5.3 Future directions

The above preliminary results suggest that phosphorylation of NLRC4 at Ser533 is largely dispensable for CASP1-dependent signaling in response to NAIP ligands. However, several independent groups have observed a significant fraction of NLRC4 to be phosphorylated at this residue, albeit under different conditions[159,218,250]. If phosphorylation is not required for inflammasome signaling, why then is NLRC4 phosphorylated? One possibility is that phosphorylation may regulate which caspases are efficiently recruited to the inflammasome. A recent report suggests that S533 phosphorylation is necessary and sufficient to induce CASPASE-8-dependent apoptosis when NLRC4 is overexpressed or constitutively activated by removal of the LRR domain[251]. It remains to be seen whether CASPASE-8 activation following bacterial infection also requires this phosphorylation event. It is unclear how phosphorylation at S533, a residue distal to the CARD, would affect caspase recruitment but not the oligomerization status of the adjacent NBD. Nevertheless, this hypothesis should be readily testable by examining unstimulated S533D/D BMM for apoptotic markers and by inhibiting CASP1 prior to flagellin stimulation to isolate CASPASE-8 activity.

Alternatively, phosphorylation of S533 may be involved in enhancing the duration of inflammasome signaling by shifting the oligomerization equilibrium to disfavor dissociation of NLRC4 protomers. In the NLRC4 crystal structure, the phosphate group coordinates several positively charged residues in the LRR domain of the same NLRC4 protomer[159]. This coordination may help to keep the HD2 and LRR locked as a rigid body that is rotated away from the NBD in the assembled inflammasome, whereas dephosphorylation might promote flexibility that would increase the likelihood of dissociation. In support of this possibility, non-phosphorylatable S533A NLRC4 exhibited delayed responses to *S. typhimurium* infection at low MOI[235]. If this hypothesis is correct, flagellin dose response curves or time course experiments should reveal a difference between wild-type and S533A NLRC4 signaling in BMM. A last possibility is

that phosphorylation of NLRC4 on S533 is an 'accident', explained by the promiscuity of kinases, and is of no functional importance.

In contrast to the constitutive signaling observed upon overexpression of NLRC4 H443P[155,251], we found that H443P is either a null or hypomorphic allele. Nevertheless, consistent with the cold-induced symptoms of patients, NLRC4 H443P signaling appeared to be enhanced at temperatures < 37 ºC. If this result holds, it may be possible to identify temperature-sensitive negative regulators of inflammasome signaling in H443P BMM. Given the difficulty of purifying inflammasomes out of BMM when NLRC4 is expressed at wild-type levels (Figures 5.7-5.9), purification of H443P inflammasomes activated at different temperatures followed by mass spectrometry of associated proteins is unlikely to prove fruitful. However, the advent of high-throughput CRISPR screens in eukaryotic cells provides an alternative strategy for identification of negative regulators. A forward genetic screen might identify CRISPR-targeted genes that are selectively lost in *Nlrc4*$^{+/+}$ or sensitized H443P/P BMM but not in *Nlrc4*$^{-/-}$ BMM, either in the presence or absence flagellin administration.

CRISPR screens have also proven useful for the identification of cofactors required for NAIP–NLRC4 to induce pyroptosis, such as GSDMD[89]. Such screens are especially powerful in selecting for cofactor-deficient cells that survive inflammasome activation. However, the induction of CASPASE-8-mediated apoptosis in the absence of the more rapid CASP1-induced pyroptosis[153] may limit the ability to recover non-pyroptotic cells. Indeed, *Gsdmd*$^{-/-}$ cells are only delayed in their induction of cell death following inflammasome activation[88,89,153]. A CRISPR screen for *Caspase-8/RIPK3*$^{-/-}$ BMM that survive cytosolic flagellin challenge should provide a larger window of selection to isolate additional factors that are partially or fully required for the induction of pyroptosis.

5.4 Methods

5.4.1 Inflammasome reconstitution in HEK293T

HEK293T were grown in DMEM supplemented with 10% FBS, 2 mM L-glutamine, 100 U/mL penicillin, and 100 μg/mL streptomycin. Cells were transfected with inflammasome components using Lipofectamine-2000. Inflammasome assembly was assessed by native PAGE at 48 hr after transfection, and IL-1β processing was assessed in cell lysates at 24 hr, as described in Chapter 2. NLRC4 S533A and S533D mutations were introduced by Quickchange PCR using the following primers: S533A (F: GGCCTCTCTGGAGGCAGGAAgCAATCCAGAGTCTGAGAAATACCACTG, R: CAGTGGTATTTCTCAGACTCTGGATTGcTTCCTGCCTCCAGAGAGGCC), S533D (F: GGCCTCTCTGGAGGCAGGAAgacATCCAGAGTCTGAGAAATACCACTGAG, R: CTCAGTGGTATTTCTCAGACTCTGGATgtcTTCCTGCCTCCAGAGAGGCC). All other constructs have been described[3,123].

5.4.2 CRISPR targeting in mice

Fertilized embryos from C56BL/6J (B6) mice were injected with Cas9 mRNA and sgRNA, as described[127], along with DNA oligonucleotides (ssODN) for homology-directed repair. Guide RNAs were designed using MIT and Benchling CRISPR design tools and chosen to optimize targeting efficiency relative to efficiency of off-target sites. The following guides were selected: S533 (ATTGATTCCTGCCTCCAGAG, non-coding PAM, MIT targeting score = 48, highest off-target score = 5.2), H443 (TTTATGAAAGAATTTATACG, non-coding PAM, MIT targeting score = 55, highest off-target score = 6.5). S533 sgRNA was co-injected with S533A (TGCAATGGTTTATCAG CACGGCAGCCTACAAGGACTTTCAGTCACCAAGAGGCCTCTCTGGAGaCAaGAAgC AATtCAGAGTCTGAGAAATACCACTGAGCAAGATGTTCTGAAAGCCATCAATGTAAA TTCCTTC) and S533D (tgcaatggtttatcagcacggcagcctacaaggactttcagtcaccaagaggcctctct ggagacaagaggatatccagagtctgagaaataccactgagcaagatgttctgaaagccatcaatgtaaattcc) ssODNs. The H443 sgRNA was coinjected with H443P ssODNs (CGTCCTGGTGACAA TAGGGCTCCTCTGTAAGTACACAGCTCAGAGGCTGAAGCCCACGTAcAAgTTtTTTCc TAAgTCATTTCAGGAGTACACGGCAGGTCGGAGACTCAGCAGTTTGCTGACGTCCA AAGAGCCA). ssODNs contained the indicated coding change and several silent point mutations to prevent continued targeting of the repaired allele. ssODNs were synthesized with a terminal 5' and 3' phosphorothioate bond for stability and PAGE purified. Founder mice were genotyped by PCR amplification of *Nlrc4* exon 4 using Ipaf-GenoF (ATGGGTCCAGCATGAACGAG) and Ipaf-GenoR primers (TCTGAGAACAAAT TGATGCCACAC). PCR products were digested with fast alkaline phosphatase (Thermo Fisher) and exonuclease I (NEB) and sequenced with Ipaf-GenoF or Ipaf-GenoR primers. Founders were backcrossed to B6 mice and then crossed to homozygosity. Animal experiments were approved by the UC Berkeley Animal Care and Use Committee.

5.4.3 FlaTox and cell death measurement

B. anthracis PA and the N-terminus of LF fused to *L. pneumophila* flagellin (LFn-FlaA, LFn-FlaA-3A) were purified from bacteria using Nickel NTA resin, essentially as described[107]. His_6-LFn-PrgJ and -YscF were purified from insect cells as described[127]. An N-terminal FLAG tag was introduced 5' of the His_6-LFn using Quickchange PCR with the following primers for His_6-LFn-FlaA and His_6-LFn-FlaA-3A (F: ctttaagaaggagatatacc ATGggcgattacaaggacgacgatgacaagGGCAGCAGCCATCATCAT, R: ATGATGATGGCT GCTGCCcttgtcatcgtcgtccttgtaatcgccCATggtatatctccttcttaaag). A FLAG tag was added to His_6-LFn-PrgJ by conventional PCR using a forward primer (caaccATGggcgattacaaggac gacgatgacaagGGCAGCAGCCATCATCATCATC) with reverse primers for PrgJ (ggtcctcgagTCATGAGCGTAATAGCGTTTCAACAGCC) or PrgJΔC4 (ggtcctcgagTCAC GTTTCAACAGCCCCGACTC). PCR amplicons were cloned into the NcoI and XhoI sites of pET28a. FLAG-His_6-LFn-ligands were also purified from bacteria using Nickel NTA resin. BMM were differentiated in 5% CSF for 7 days and then seeded in 96-well TC-treated plates at 10^5 cells/well. Where indicated, BMM were primed with 0.5 μg/mL Pam_3CSK_4 for 4 hr. Media was replaced with fresh media containing 4 μg/mL PA and the indicated concentrations of LFn-FlaA or other LFn-ligand. Treated cells were

incubated 4 hr at 37ºC or the indicated temperature, and supernatants were analyzed for LDH release as in Chapter 3. Cell death was normalized to Triton-lysed wells of the same genotype.

5.4.4 Melanoma growth

Age-matched 8-12 week female mice were injected subcutaneously with 10^5 B16-F10 cells. Littermates from heterozygous *Nlrc4*-459stop and *Nlrc4*-542 stop crosses were used cohoused with 2 *Nlrc4*$^{+/+}$, 2 *Nlrc4*$^{+/-}$, 2 *Nlrc4*$^{-/-}$ mice per cage for at least 3 weeks prior to injection. Tumor volume was monitored by caliper measurement.

5.4.5 Macrophage immortalization

ASC/Casp1/Casp11$^{-/-}$ bone marrow was harvested and differentiated in 50% BMM media (5% CSF) and 50% filtered (0.45 µm) spent media from *Casp1/Casp11*$^{-/-}$ immortalized cells. Each day all cells (adherent and non-adherent) were harvested, pelleted, and given a fresh mixture of 1:1 BMM media / spent media. After 6 days immortalized BMM were split 1:5 in BMM media alone. After an additional 2 weeks of passaging, immortalized BMM were weaned into CSF-free BMM (IMM media) and passaged on non-TC-treated plates.

5.4.6 Retroviral and lentiviral transductions

Retroviral transductions were performed as described in Chapter 3. Lentiviral transductions using pFG42 (dox-inducible promoter, IRES-GFP reporter) were similar, except that HEK293T were used as packaging cells and transfected with 2 µg of pFG42 derivative, 0.5 µg of vsv-g, and 1.5 µg of pSPAX2. *L. pneumophila* FlaA was cloned into the BamHI and XbaI sites of pFG42, and an N-terminal FLAG tag was added to FlaA by PCR using the forward primer FLAG-FlaF (caaggatccgccaccATGgattacaaggacgacgatga caagGCTCAAGTAATCAACACTAATGTG) with reverse primers ggtgttctagaCTATCGAC CTAACAAAGATAATACAG (for FlaA) or gtcgtTCTAGActaAGCAGCGTAATCAGCATCT TGAATAC (for FlaAΔC35). Transduced cells were passaged in BMM or IMM media, as appropriate, and flagellin expression was induced by overnight incubation in 5 µg/mL doxycycline. GFP expression was analyzed by flow cytometry, and gates were set on untransduced cells. A subset of transductants were lysed in RIPA buffer and analyzed by SDS-PAGE. For puromycin-selectable transduction, an MSCV2.2 derivative containing an IRES-puro-T2A-mCherry cassette was modified to remove a SalI site in the puromycin-resistance cassette using Quickchange (F: CCACGCGCCACACCGTtGA CCCGGACCGCCAC, R: GTGGCGGTCCGGGTCaACGGTGTGGCGCGTGG). FLAG-FlaA or FLAG-FlaAΔC35 were amplified using the FLAG-FlaF primer (BamHI site) with reverse primers ggttgcggccgcCTATCGACCTAACAAAGATAATACAGATTGC (for FlaA) or gtatGCGGCCGCctaAGCAGCGTAATCAGCATCTTGAATAC (for FlaAΔC35) and cloned into the BglII and NotI sites of MSCV2.2. Transductants were selected with 5 µg/mL puromycin, based on the sensitivity of untransduced cells to 2-4 µg/mL puromycin, starting at 4 days post-transduction.

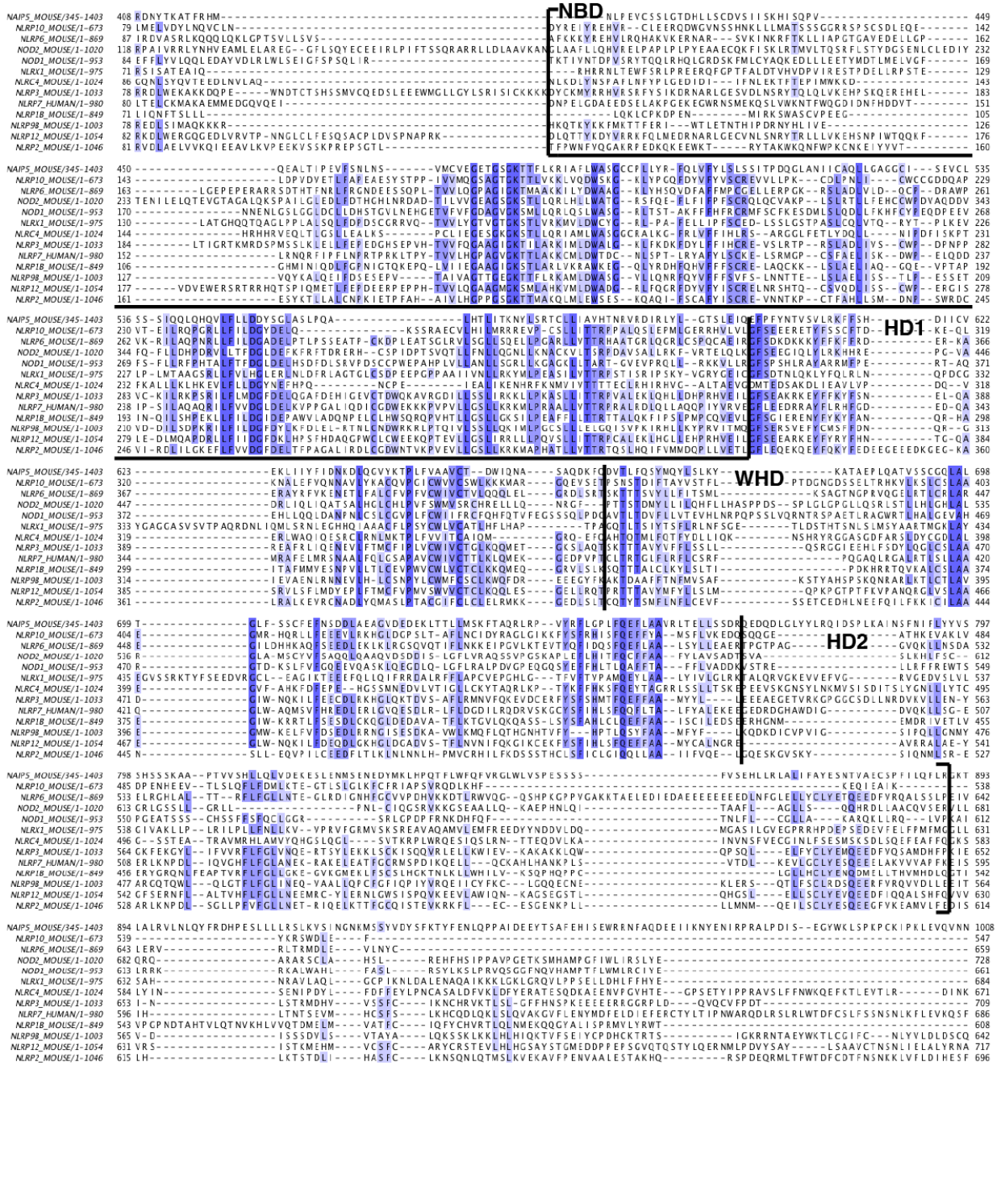
5.4.7 Purification of inflammasomes out of BMM

Inflammasomes assembled with FLAG-tagged ligands were purified using FLAG M2 resin as described in Chapter 4. Inflammasomes assembled with His_6-tagged ligands were purified using Nickel-NTA resin with the following buffers: cell lysis buffer (25 mM Tris, pH 8, 150 mM NaCl, 20 mM imidazole, 1 mM $MgCl_2$, 1x EDTA-free protease inhibitor cocktail), resin wash buffer (25 mM Tris, pH 8, 500 mM NaCl, 20 mM imidazole, 1 mM $MgCl_2$), and resin elution buffer (25 mM Tris, pH 8, 500 mM NaCl, 500 mM imidazole, 1 mM $MgCl_2$).

5.5 Acknowledgments

Angus Lee injected mouse zygotes to generate CRISPR-targeted mice. Steve Wilson provided advice for sgRNA design. Isabella Rauch provided advice on genotyping and colony maintenance. Thornton Thompson performed the B16 melanoma experiment. Lentiviral vectors were a gift of Benjamin Gowen and David Raulet. Livy Majer modified MSCV2.2 to replace the IRES-GFP with IRES-puromycin-T2A-mCherry.

a



a (cont'd)

b

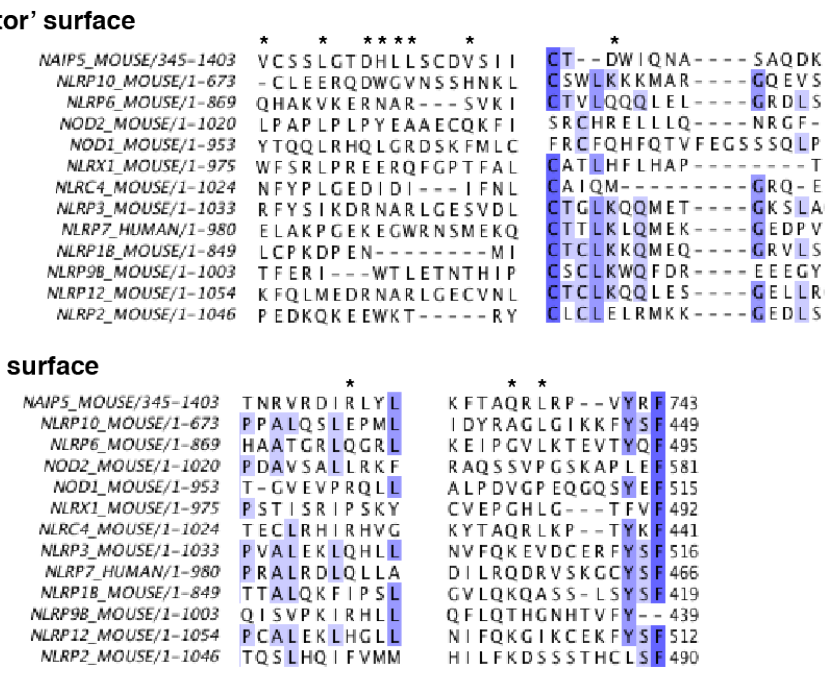


Figure 5.1. Conservation of NBD motifs among mouse NLRs

a, The NBD and HD1 are relatively conserved among mouse NLRs. NAIP5, NLRC4, NLRP1b, NLRP2, NLRP3, NLRP6, NLRP7, NLRP9B, NLRP10, NLRP12, NLRX1, NOD1, and NOD2 protein sequences were aligned using Muscle with default settings. The

NAIP5 BIRs (1-344) and NLRP1b FIIND and CARD (850-1003) were removed prior to alignment. Alignments are colored by percent identity using Jalview. The NBD, HD1, WHD, and HD2 (as defined by PDB 4KXF) are bracketed and indicated; the NBD is underlined. **b**, NLRC4 ‘donor’ and ‘acceptor’ motifs are not conserved among mouse NLRs. Residues important for NAIP5-NLRC4 and NLRC4-NLRC4 interactions, as defined in [5,112] and Chapter 4, are indicated by asterisk.

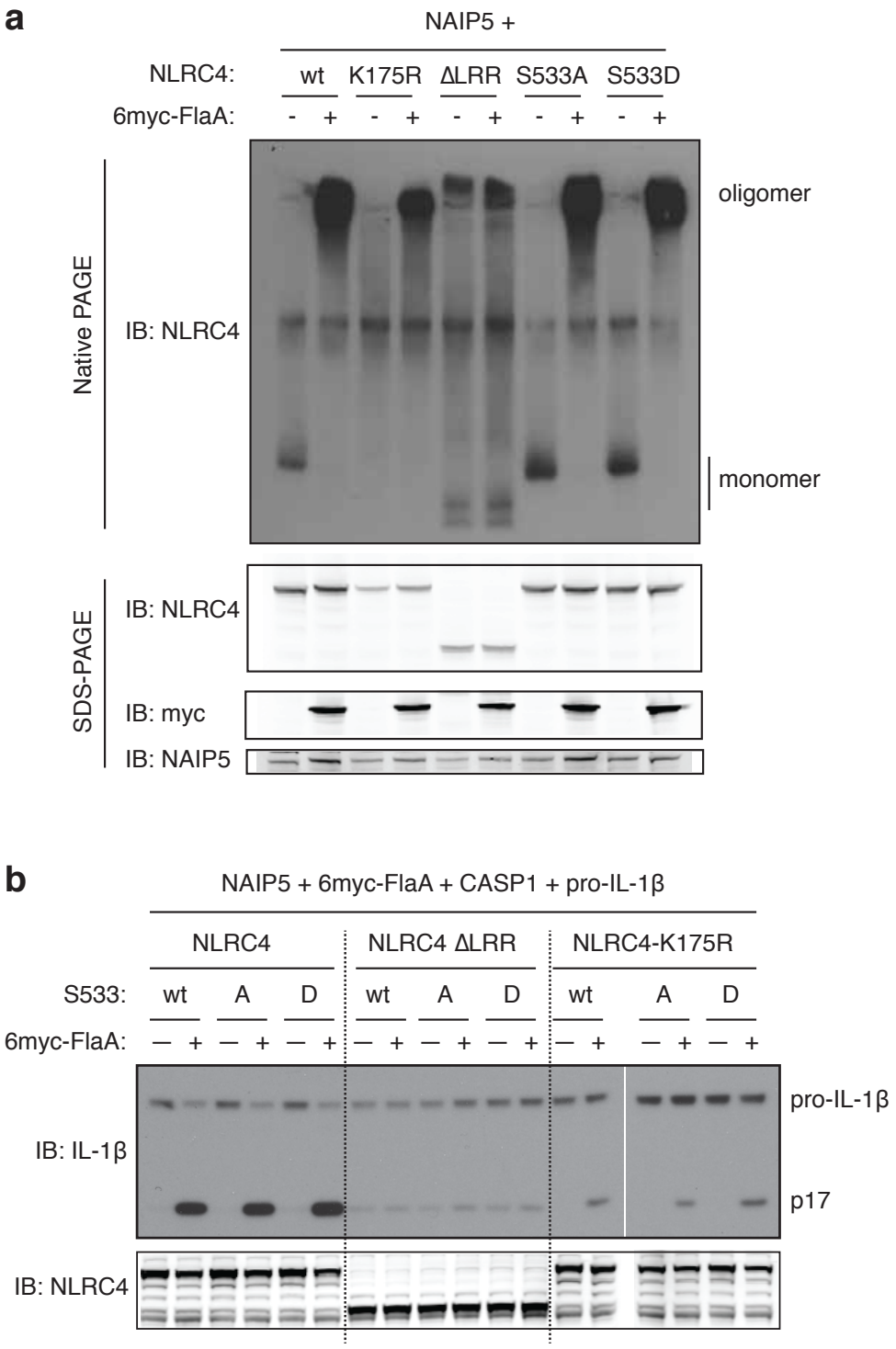


Figure 5.2. NLRC4 S533 phosphorylation is not necessary for signaling
a, The indicated constructs were transfected into HEK293T and analyzed for oligomerization by native PAGE or expression by denaturing PAGE. **b**, Transfected HEK293T cell lysates were analyzed by SDS-PAGE for cleavage of pro-IL-1β into the p17 fragment. Results representative of 2 (**a**) or 1 (**b**) independent experiments.

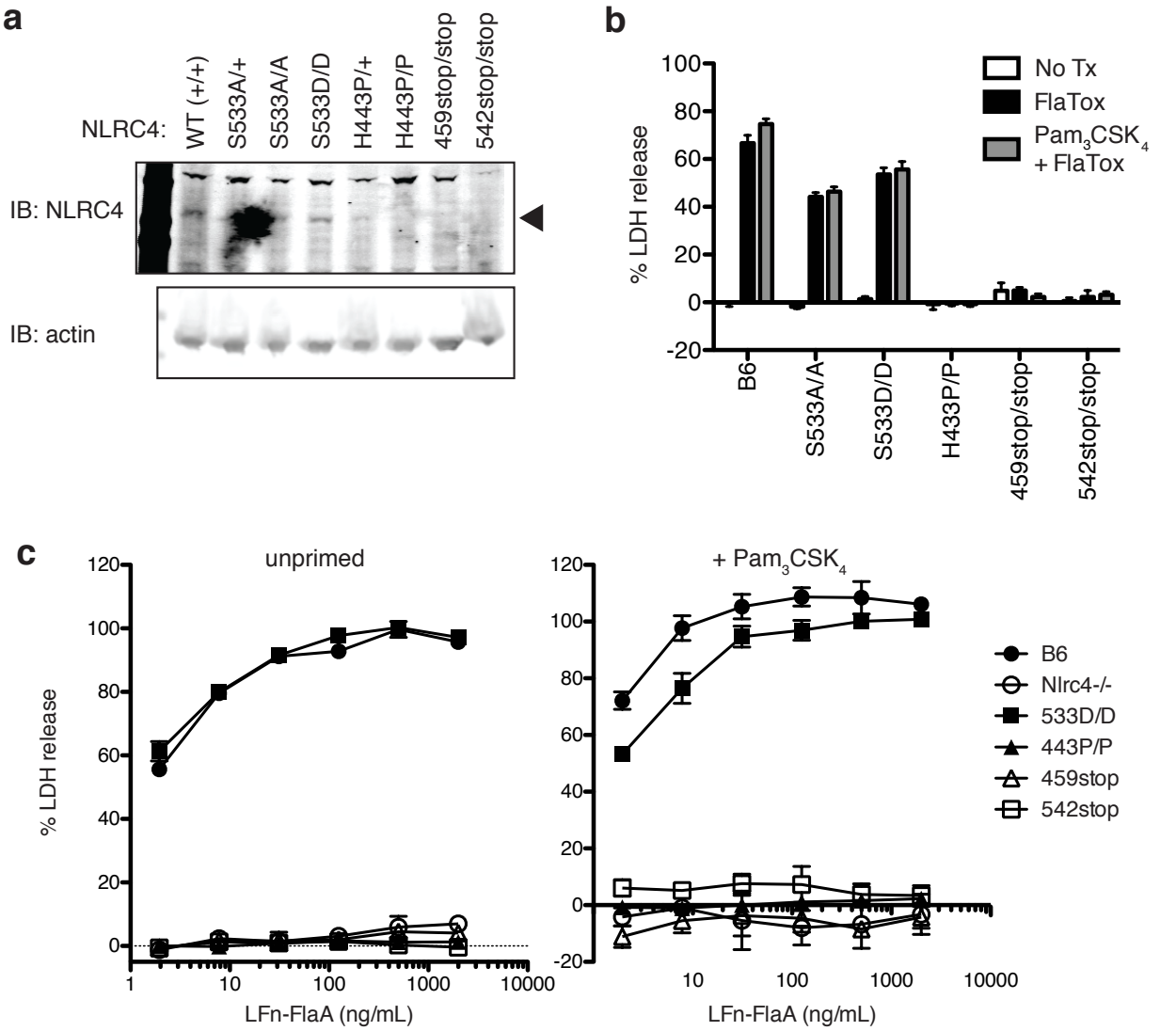


Figure 5.3. *Nlrc4* alleles do not behave as expected in BMM
a, BMM from the indicated strains were lysed and probed for NLRC4 expression by immunoblot (IB). Arrow indicates the size of full-length NLRC4. **b**, BMM were left untreated or primed with 0.5 µg/mL Pam_3CSK_4 for 4 hr, then treated for 4 hr with 4 µg/mL PA and 0 or 20 ng/mL LFn-FlaA. Supernatants were analyzed for LDH release and normalized to untreated wells of the same genotype lysed with 1% Triton. **c**, BMM were treated as in **b** but with 4 µg/mL PA and the indicated concentration of LFn-FlaA. Results are from a single experiment and have not yet been repeated.

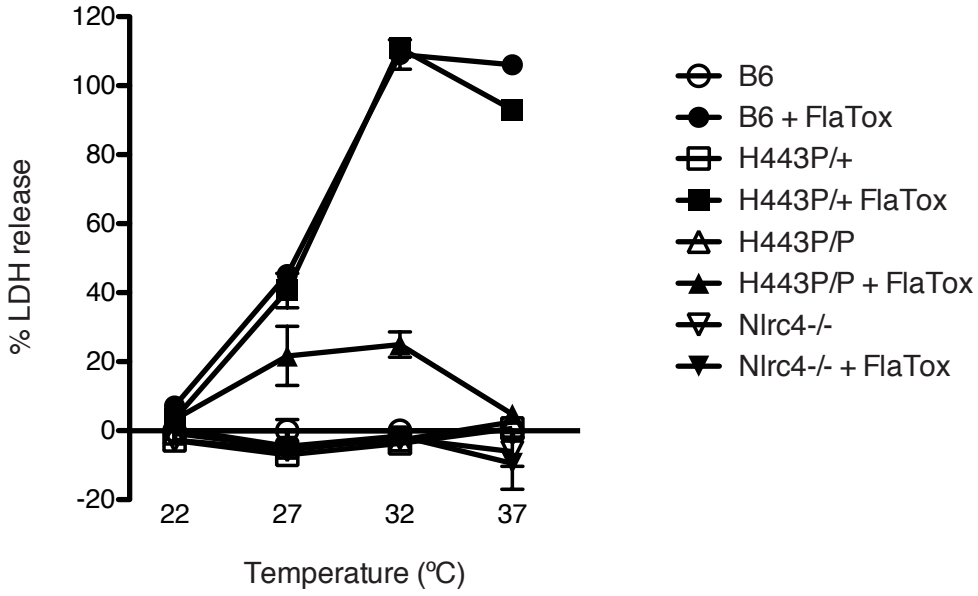


Figure 5.4. NLRC4 H443P may be cold-responsive
BMM of the indicated genotype were incubated 4 hr at the indicated temperature in media containing 4 μg/mL PA and 0 or 2 μg/mL LFn-FlaA. Cell death was assessed by LDH release, normalized to untreated B6 BMM at the indicated temperature. Results are from a single experiment and have not yet been repeated.

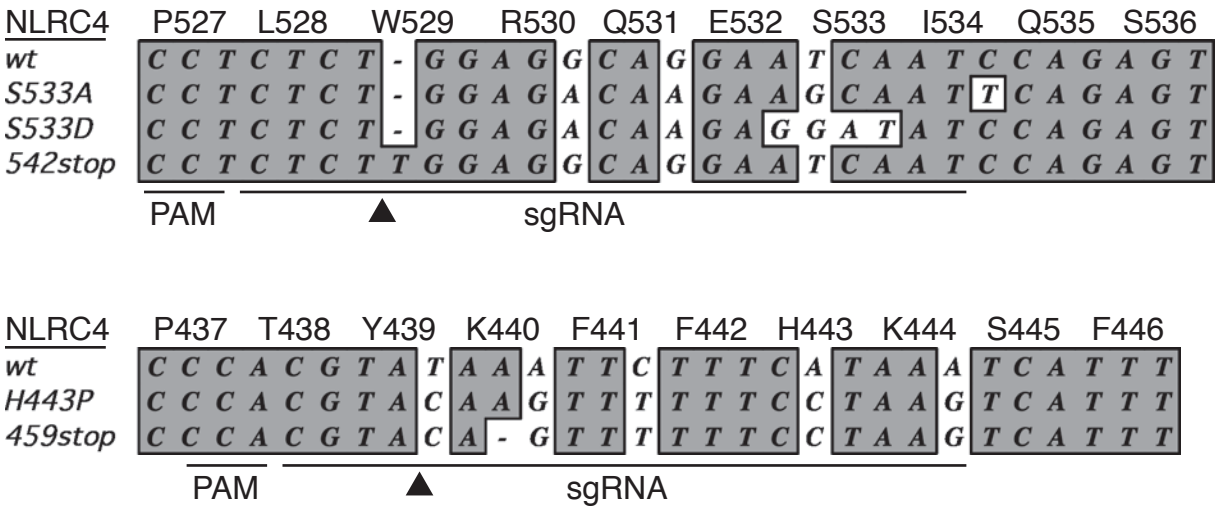


Figure 5.5. *Nlrc4* frameshift alleles generated by CRISPR targeting
Sequence of *Nlrc4* alleles recovered from CRISPR founder mice. Guide RNAs and the associated PAM site are indicated for targeting of S533 and H443. Arrowheads indicate predicted Cas9 cut sites.

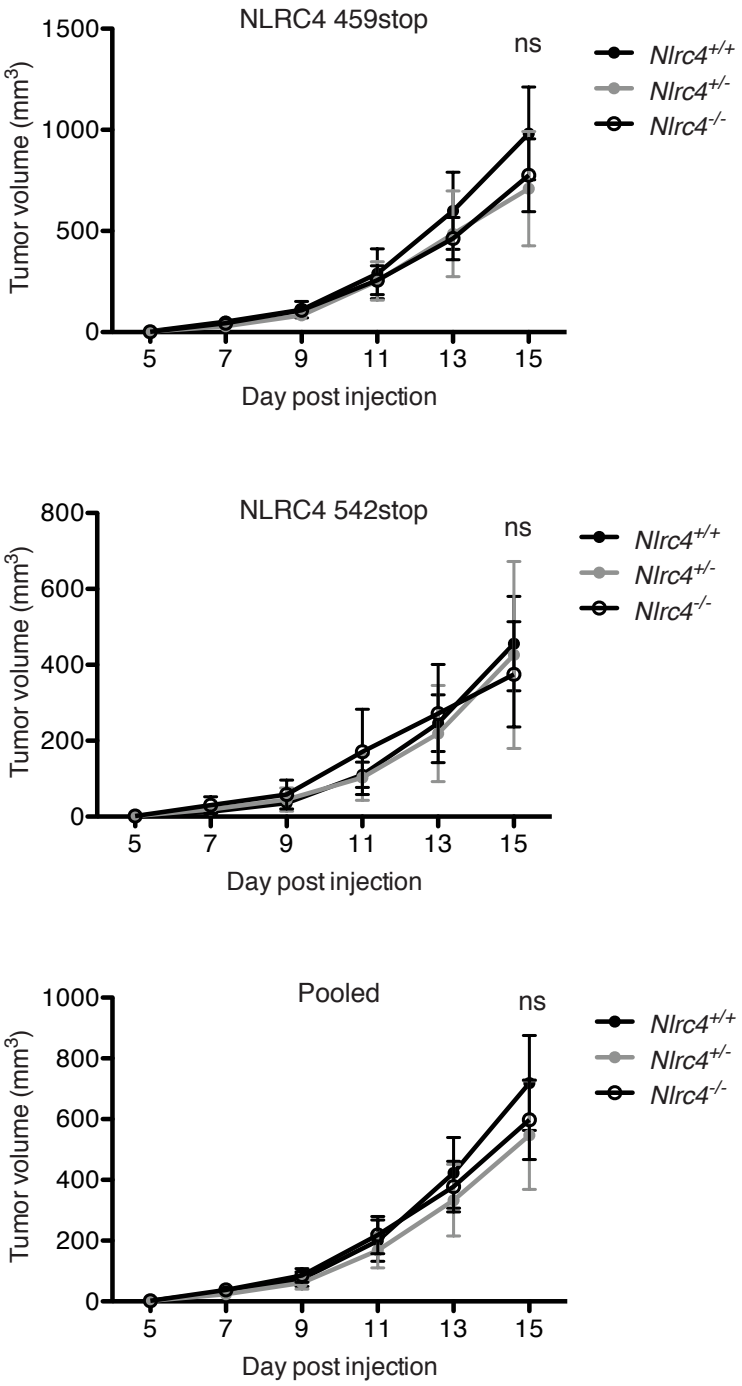


Figure 5.6. NLRC4 does not affect melanoma growth
Littermates of the indicated genotype were injected subcutaneously with B16-F10 melanoma cells. The same data from *Nlrc4* 459stop and *Nlrc4* 542stop mice are presented separately and pooled. *Nlrc4*$^{+/-}$ or *Nlrc4*$^{-/-}$ mice were compared to *Nlrc4*$^{+/+}$ mice using repeated measures 2-way ANOVA. Results are from a single experiment.

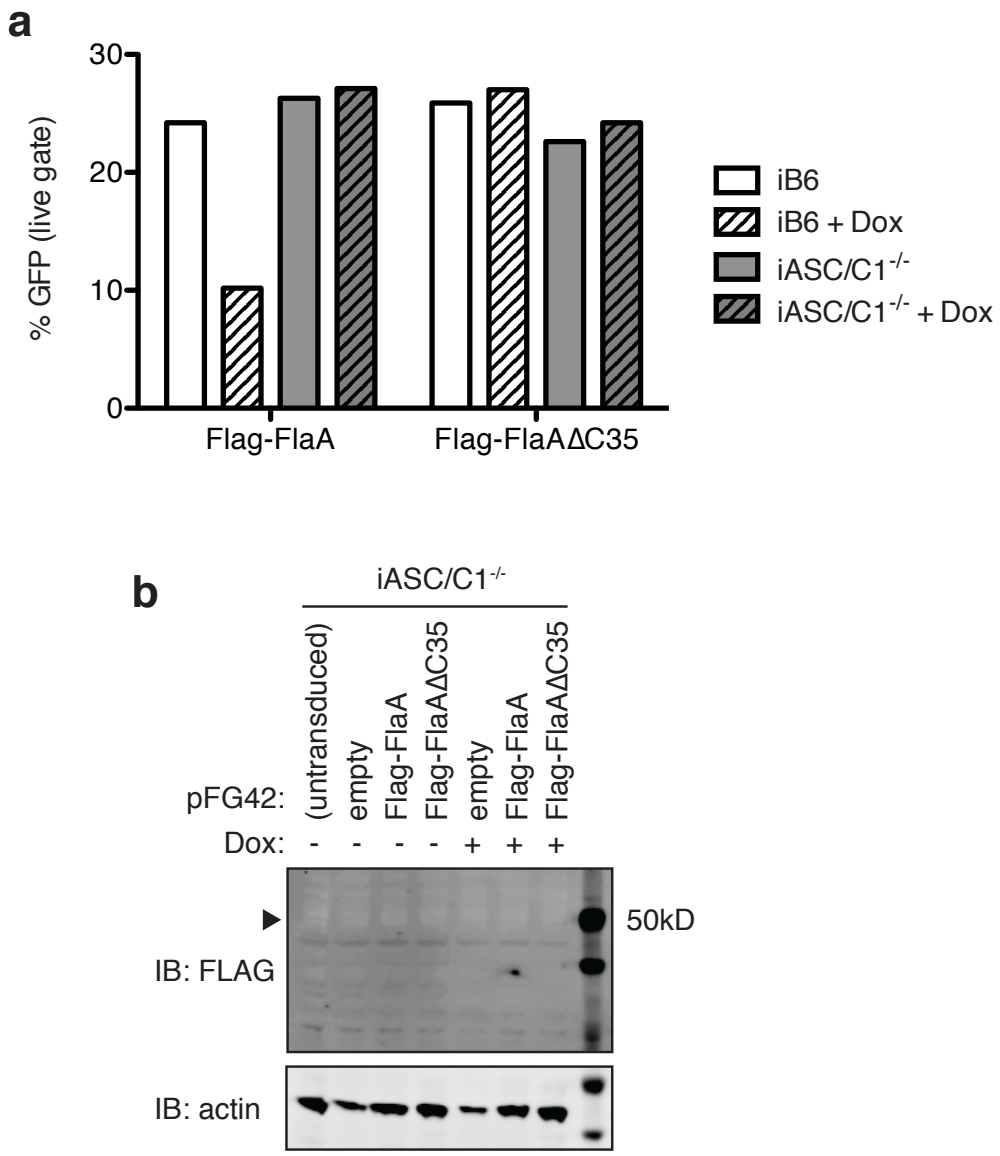


Figure 5.7. Dox-induction of flagellin of transduced BMM is weak
Immortalized BMM of the indicated genotype were transduced with FLAG-tagged flagellin under a doxycycline-inducible promoter (pFG42 vector). At 4 days post-transduction, flagellin expression was induced with 5 μg/mL doxycycline overnight. **a**, Cell death was monitored by loss of GFP^+ cells, analyzed by flow cytometry. **b**, In parallel, iBMM were lysed and analyzed for flagellin expression by anti-FLAG immunoblot.

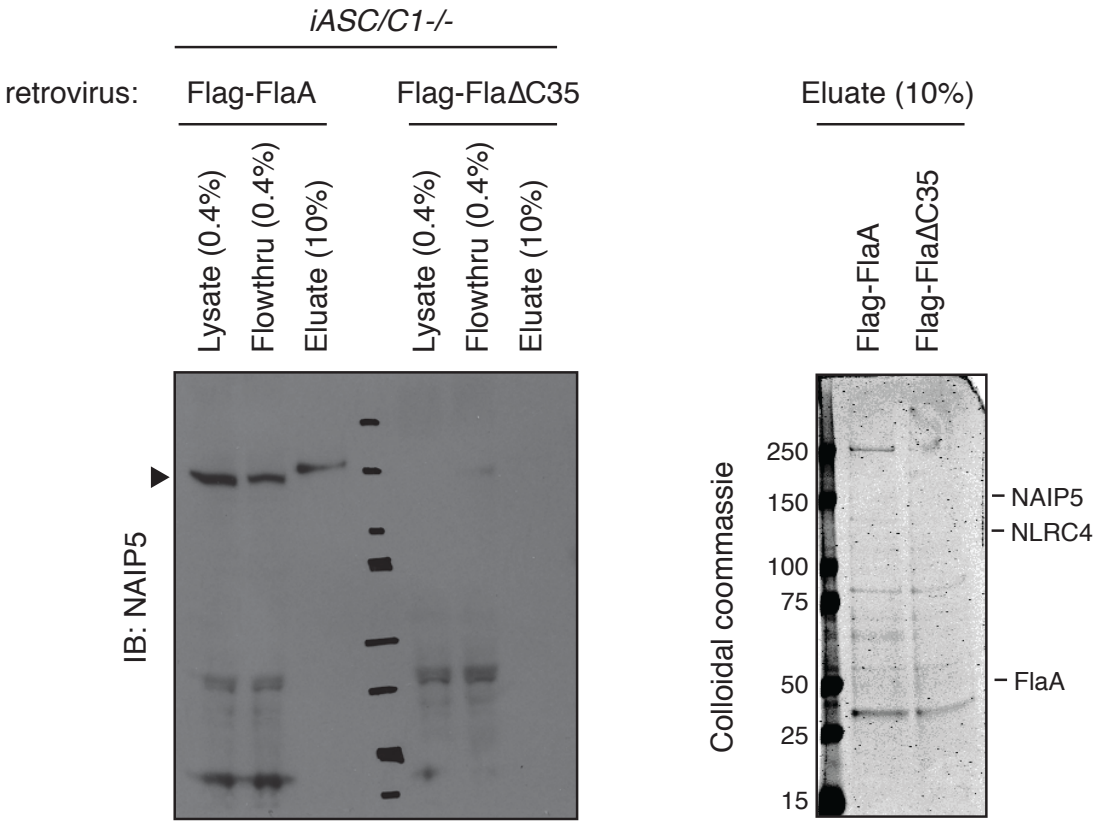


Figure 5.8. Constitutively expressed flagellin pulls down low amounts of NAIP5
Immortalized *ASC/C1*$^{-/-}$ BMM were transduced with N-terminally FLAG-tagged flagellin in a puromycin-selectable vector. Transductants were selected with 5 µg/mL puromycin starting at day 4 post-transduction and expanded to 6 x 15cm plates. Cells were lysed and subjected to FLAG immunoaffinity purification. FLAG eluates were analyzed by anti-NAIP5 immunoblot and colloidal Coomassie stain. The expected positions of NAIP5, NLRC4, and flagellin are indicated.

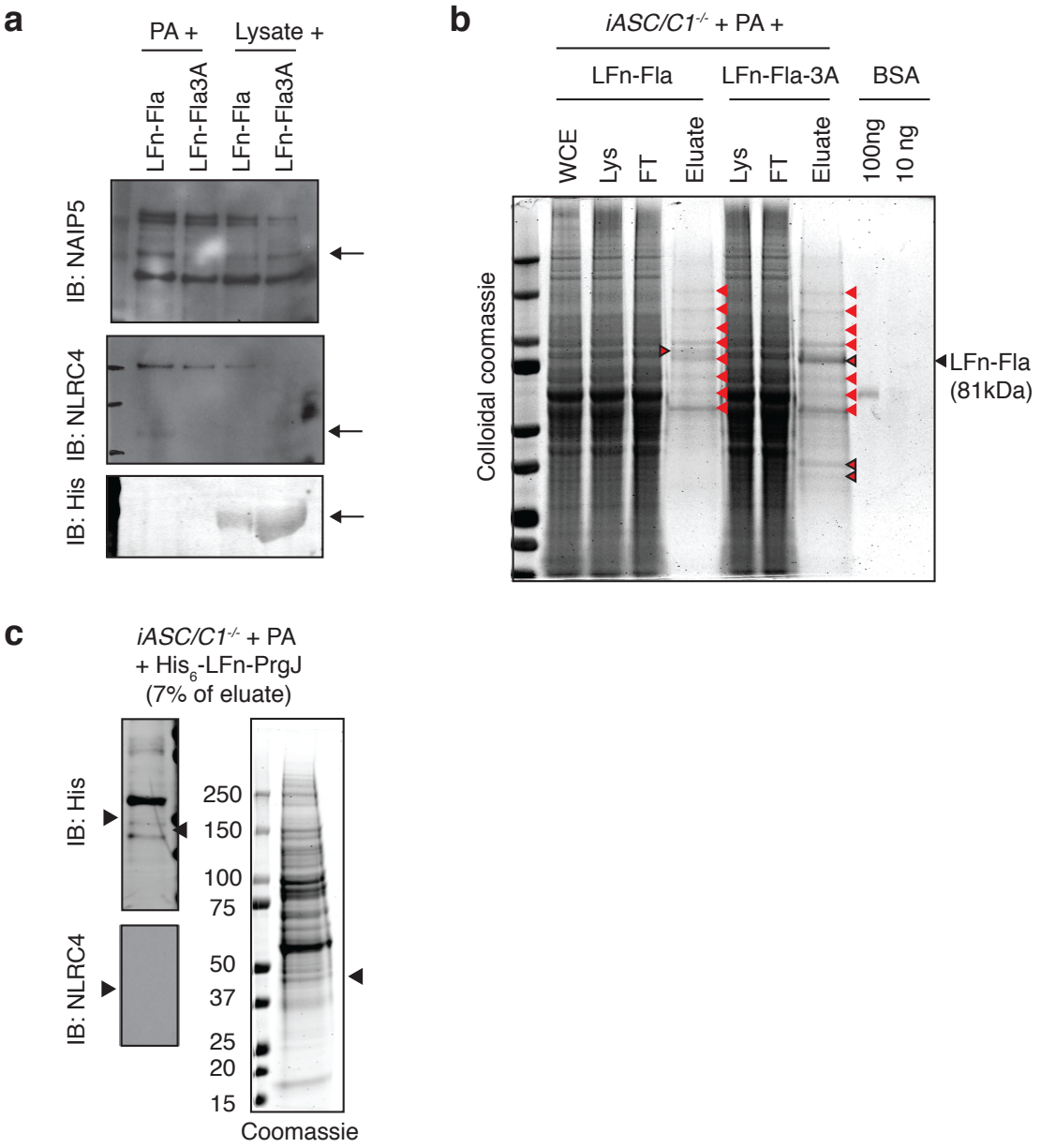


Figure 5.9. Purification of inflammasomes by LFn-delivered NAIP ligands
a, Primary *ASC/C1*$^{-/-}$ BMM (2 x 15 cm plates) were treated with 10 µg/mL PA and 10 µg/mL His_6-LFn-FlaA or His_6-LFn-FlaA-3A (L470A/L472A/L473A) for 4 hr at 37 ºC. In parallel, BMM were lysed and incubated with 10 µg/mL LFn-FlaA without PA for 4 hr at 4 ºC. FlaA was purified from cell lysates by nickel resin. **b,** Immortaliezd *ASC/C1*$^{-/-}$ BMM (2 x 10 cm plates of confluent cells) were treated as in **a**, and eluates were analyzed by colloidal Coomassie stain. Red arrows indicate co-purified bands that are identical for FlaA and FlaA-3A; outlined red arrows indicate bands that differ. **c**, Immortaliezd *ASC/C1*$^{-/-}$ BMM (2 x 10 cm plates of confluent cells) were treated with 10 µg/mL PA and 3.3 µg/mL His_6-LFn-Fla166 or His_6-LFn-PrgJ for 4 hr at 37 ºC. Lysates were subjected to nickel resin purification, and 7% of total eluate was analyzed by immunoblot or Coomassie stain. Arrows indicated expected position of NAIP5, NLRC4, Fla166, or PrgJ.

Table 5.1. Smaller ligands are more potent in LFn delivery

LFn-Ligand	Size (kDa)	EC_{50} (ng/mL)*	EC_{50} (pM)
His_6-LFn-FlaA	80.9	2.3	28
His_6-LFn-FlaA-3A	80.9	n.d.	
FLAG-His_6-LFn-FlaA	82.0	0.93	11
FLAG-His_6-LFn-FlaA-3A	82.0	n.d.	
FLAG-His_6-LFn-FlaA-D0	44.8	not made	
His_6-LFn-FlaA166	50.1	2.2	44
His_6-LFn-PrgJ	44.0	0.027	0.61
FLAG-His_6-LFn-PrgJ	45.1	0.31	6.9
FLAG-His_6-LFn-PrgJΔC4	45.1	not converged (no death)	
His_6-LFn-YscF	42.9	2.1	49

*Calculated using serial dilutions of LFn-ligand with 4 μg/mL PA. EC_{50} for LDH release from B6 BMM was calculated using the log(agonist) vs. response (3 parameters) non-linear regression fit in Prism.

Chapter Six: Conclusions — How to Safely Play with Fire

Inflammasomes are sentinel PRRs that induce potent pro-inflammatory responses upon detection of cell-invasive pathogens. While these responses are effective at clearing would-be invaders[14], they are also harmful to the health and well-being of the host[107,154-158]. The danger is especially pronounced for NAIP–NLRC4 inflammasomes, as they are constitutively expressed and poised for immediate activation, without a requirement for priming, in both immune cells and the epithelial cells that form a physical barrier against invading pathogens[153]. The challenge for NAIP inflammasomes is to maintain robust recognition of bacterial pathogens, which are constantly evolving to evade detection, without sacrificing the specificity of responding only during *bona fide* infections. These two contradictory goals are both critical for the survival and, ultimately, reproductive success of the host.

The work presented here provides several insights into how NAIP inflammasomes walk the line to achieve highly specific recognition of bacterial pathogens that is also resilient to mutation of cognate ligands. First, NAIPs use distinct surfaces to bind bacterial ligands and to mediate autoinhibition, and this separation allows the ligand-binding domain to rapidly evolve ligand specificity without risk of compromising autoinhibition (Chapter 2). The LRR domain[3] and inter-domain surfaces coordinating ADP[107,154-157,159] keep NAIP domains locked together in a closed conformation to prevent constitutive signaling. Ligand binding occurs on surfaces distal to the ATP binding pocket and does not disrupt the orientation of the LRR relative to its adjacent domains. Instead, ligand binding activates NAIPs by sterically inducing a rigid body rotation of the locked WHD-HD2-LRR segment (Chapter 4). Altering the ligand-binding domain is therefore unlikely to affect autoinhibition, allowing NAIPs to safely sample mutations with better ligand recognition. This freedom is reflected in the evolutionary history of NAIPs, which have undergone repeated rounds of selection for amino acid changes in the ligand-binding domain (Chapter 2).

Interestingly, the ligand specificity of NAIPs in different lineages suggests that their history of diversifying selection may have been driven by different selective pressures. Duplicated mouse NAIP paralogs have diverged to recognize unique ligands, and coding changes may have played a role in refining and optimizing the binding of specific ligands. The sole human NAIP appears to recognize all known mouse NAIP ligands, suggesting that this broader specificity may be the ancestral NAIP phenotype. Diversifying selection is also operating on single-copy primate NAIPs and thus may serve to optimize generalized, rather than specific, ligand binding. Alternately, in both lineages coding changes may increase recognition of species-relevant bacterial pathogens that are evolving to evade detection. To understand which of these selective pressures have driven the evolutionary history of NAIPs, it will be necessary to determine the effect of specific amino acid changes on recognition of multiple ligands from different bacterial pathogens.

A second strategy employed by NAIPs to maintain recognition of rapidly evolving bacterial ligands is the detection of multiple ligand surfaces. When deployed to recognize several of the most conserved sites on bacterial ligands, this strategy constrains bacterial immune evasion by increasing the required mutational threshold above the level tolerated for ligand function (Chapter 3). The requirement for multiple ligand mutations to disrupt NAIP recognition is likely due to the properties of the NAIP–ligand binding interface (Chapter 4). NAIP5 contacts multiple amino acids along the length of the flagellin $D0_C$ alpha helix and wedges both ends of the $D0_N$ alpha helix to lock the D0 domain into place. This distributed binding surface is largely composed of hydrophobic contacts with no grooves or pockets that might constitute 'hot spot' binding sites. The lack of particular residues that contribute most of the binding energy allows NAIPs to tolerate single mutations in any of the contacted residues. Thus, the structure of flagellin-bound NAIP5 answers the critical question of how ligand binding can be simultaneously specific and resilient to ligand mutations, through the use of generic shape complementarity for multiple alpha helices of the right length and hydrophobicity. This mode of binding likely also explains why closely related NAIP paralogs or a single human NAIP can recognize several unrelated ligands, so long as the ligand comprises two alpha helices with the requisite hydrophobic contacts.

The generic nature of ligand binding raises the troubling possibility that endogenous host proteins resembling bacterial ligands might aberrantly activate NAIPs to induce constitutive inflammation and disease. Multi-surface recognition may serve a second critical function as a stringency filter to prevent NAIPs from recognizing such proteins. Even in *bona fide* bacterial ligands with the dominant C-terminal motif intact, a second motif on the distal end of the ligand is required for NAIP activation (Chapter 3). Thus a cytosolic host protein with a C-terminal helix containing three terminal leucines should fail to activate NAIPs due to its lack of this secondary motif. To test this hypothesis, it should be possible to identify at least one such host protein and show that (a) its C-terminal helix is capable of stimulating NAIP5 when it replaces the $D0_C$ of flagellin and (b) the addition of $D0_N$ to this host protein induces constitutive NAIP5 activation.

In support of the hypothesis that multi-surface recognition can serve to prevent self-recognition, in addition to constraining bacterial immune evasion, multi-surface PAMP detection has been observed even in PRRs that detect non-variable ligands like DNA. For example, human TLR9 is selective for two adjacent CpG motifs, multiplicatively decreasing the chances of responding to mammalian genomic DNA that contains lower frequencies of CpG motifs than microbial DNA[197]. Determining the contribution of multi-surface recognition to selectivity and/or robust ligand recognition, for NAIPs and other PRRs, will help to elucidate the rules by which eukaryotes build their innate immune systems.

Control of dangerous immune responses is of course not limited to how PRRs discriminate non-self from self ligands. Strict control of where and under what circumstances PRRs elicit specific responses plays an important role in preventing

autoimmune disease. There is some evidence that the outcome of NAIP inflammasome signaling may differ depending on the cell type in which it is induced[87]. Setting-specific differences in the licensing cofactors or regulatory proteins that affect NLRC4 signaling, which are currently almost entirely unknown (Chapter 5), are likely to dictate whether NAIP activation induces pyroptosis or apoptosis, cytokine or eicosanoid release, or no signal at all.

Additionally, the inherent biochemical properties of NLRs may dictate the duration of signaling. Assembly of the related apoptosome from APAF1 monomers requires exchange of ADP for ATP, though it is unclear whether ATP hydrolysis within an assembled apoptosome favors disassembly of the complex[217]. There is some evidence to suggest that the nucleotide bound to NLRC4 is important for inflammasome assembly. Mutation of the NLRC4 Walker A motif to disrupt nucleotide binding reduces oligomerization as well as protein stability (Figure 5.2), although the NAIP5 Walker A and B motifs are not required (data not shown). Furthermore, mutation of H443, which contacts the nucleotide in a Sensor 2-like manner[252], induces constitutive oligomerization of NLRC4[159]. The Walker B motif of NLRC4 is missing the catalytic second Asp or Glu residue critical for ATP hydrolysis, although it does contain a Glu three residues downstream that may contribute to the low basal ATPase rate of monomeric NLRC4[159]. It is unclear whether inflammasome assembly would increase the ATPase activity of NLRC4, as the ATP binding pocket is sequestered from arginine fingers in the adjacent NAIP5 or NLRC4 protomer in the inflammasome. The role of nucleotide binding and hydrolysis in inflammasome assembly is not a purely academic question, as disassembly of inflammasome complexes may be necessary to prevent a feed-forward signaling cascade. Indeed, inflammasomes released from pyroptotic cells can be taken up by neighboring phagocytes, where they further propagate inflammasome signaling[253,254].

The affinity of NAIPs for their cognate ligands has been difficult to experimentally determine. Nevertheless, indirect lines of evidence suggest that NAIP surveillance for cytosolic bacterial ligands is extremely sensitive. First, pathogen secretion of ligand monomers into the cytosol is thought to be accidental, suggesting that cytosolic concentrations of NAIP ligands during bacterial infection are likely to be extremely low[118]. Second, the robust co-immunoprecipitation of NAIP5 and flagellin suggests that the affinity is at least sub-micromolar. Finally, a single flagellin is sufficient to nucleate an inflammasome capable of recruiting CASP1 (Chapter 4). It is extremely challenging to test whether assembly of one inflammasome ring is sufficient to induce cell death or cytokine processing, but the prion-like recruitment of ASC and CASP1[220,254] certainly suggest that it is possible.

All told, the sensitivity of NAIP detection may help to at least partially counteract pathogen attempts to evade recognition through downregulation of NAIP ligands. For example, *S. typhimurium* evasion of NAIP–NLRC4 detection appears to be incomplete, particularly in the earliest stages of invasion from the intestinal lumen[109,153]. Thus it is

important to remember that even incomplete victories against pathogens in the battle for innate immune surveillance can have profound consequences for the host, through both direct bacterial restriction and promotion of downstream adaptive immune responses.

Collectively, the work described here elucidates some of the strategies in play during the conflict between an innate immune receptor and bacterial pathogens, both within an individual mammalian host and over evolutionary time. Furthermore, these studies illustrate the utility of biochemical dissection of immunological questions. Here, these biochemical tools have provided a platform to understand how the innate immune system maximizes the health of the host, retaining stringent non-self specificity while flexibly adapting to competing bacterial pathogens. Continued biochemical study of the remaining open questions, only some of which are discussed above, is certain to yield further surprising insights.

References

1 Endrizzi, M. G., Hadinoto, V., Growney, J. D., Miller, W. & Dietrich, W. F. Genomic sequence analysis of the mouse Naip gene array. *Genome research* **10**, 1095-1102 (2000).

2 Daugherty, M. D. & Malik, H. S. Rules of engagement: molecular insights from host-virus arms races. *Annual review of genetics* **46**, 677-700, doi:10.1146/annurev-genet-110711-155522 (2012).

3 Kofoed, E. M. & Vance, R. E. Innate immune recognition of bacterial ligands by NAIPs determines inflammasome specificity. *Nature* **477**, 592-595, doi:10.1038/nature10394 (2011).

4 Zhao, Y. *et al.* The NLRC4 inflammasome receptors for bacterial flagellin and type III secretion apparatus. *Nature* **477**, 596-600, doi:10.1038/nature10510 (2011).

5 Hu, Z. *et al.* Structural and biochemical basis for induced self-propagation of NLRC4. *Science* **350**, 399-404, doi:10.1126/science.aac5489 (2015).

6 Darwin, C. *On the Origin of Species by Means of Natural Selection, Or, the Preservation of Favoured Races in the Struggle for Life*. (J. Murray, 1859).

7 Goldberg, G. W. & Marraffini, L. A. Resistance and tolerance to foreign elements by prokaryotic immune systems - curating the genome. *Nature reviews. Immunology* **15**, 717-724, doi:10.1038/nri3910 (2015).

8 Boehm, T. & Swann, J. B. Origin and evolution of adaptive immunity. *Annual review of animal biosciences* **2**, 259-283, doi:10.1146/annurev-animal-022513-114201 (2014).

9 Mojica, F. J., Diez-Villasenor, C., Garcia-Martinez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *Journal of molecular evolution* **60**, 174-182, doi:10.1007/s00239-004-0046-3 (2005).

10 Dunkelberger, J. R. & Song, W. C. Complement and its role in innate and adaptive immune responses. *Cell research* **20**, 34-50, doi:10.1038/cr.2009.139 (2010).

11 Aachoui, Y. *et al.* Canonical Inflammasomes Drive IFN-gamma to Prime Caspase-11 in Defense against a Cytosol-Invasive Bacterium. *Cell host & microbe* **18**, 320-332, doi:10.1016/j.chom.2015.07.016 (2015).

12 Maltez, V. I. *et al.* Inflammasomes Coordinate Pyroptosis and Natural Killer Cell Cytotoxicity to Clear Infection by a Ubiquitous Environmental Bacterium. *Immunity* **43**, 987-997, doi:10.1016/j.immuni.2015.10.010 (2015).

13 Sturdevant, G. L. & Caldwell, H. D. Innate immunity is sufficient for the clearance of Chlamydia trachomatis from the female mouse genital tract. *Pathogens and disease* **72**, 70-73, doi:10.1111/2049-632x.12164 (2014).

14 Maltez, V. I. & Miao, E. A. Reassessing the Evolutionary Importance of Inflammasomes. *J Immunol* **196**, 956-962, doi:10.4049/jimmunol.1502060 (2016).

15 Iwasaki, A. & Medzhitov, R. Control of adaptive immunity by the innate immune system. *Nature immunology* **16**, 343-353, doi:10.1038/ni.3123 (2015).

16 Schenten, D. & Medzhitov, R. The control of adaptive immune responses by the innate immune system. *Advances in immunology* **109**, 87-124, doi:10.1016/b978-0-12-387664-5.00003-0 (2011).
17 Janeway, C. A., Jr. Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harbor symposia on quantitative biology* **54 Pt 1**, 1-13 (1989).
18 Poltorak, A. *et al.* Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science* **282**, 2085-2088 (1998).
19 Schwandner, R., Dziarski, R., Wesche, H., Rothe, M. & Kirschning, C. J. Peptidoglycan- and lipoteichoic acid-induced cell activation is mediated by toll-like receptor 2. *The Journal of biological chemistry* **274**, 17406-17409 (1999).
20 Yoshimura, A. *et al.* Cutting edge: recognition of Gram-positive bacterial cell wall components by the innate immune system occurs via Toll-like receptor 2. *J Immunol* **163**, 1-5 (1999).
21 Kato, H. *et al.* Differential roles of MDA5 and RIG-I helicases in the recognition of RNA viruses. *Nature* **441**, 101-105, doi:10.1038/nature04734 (2006).
22 Brown, G. D. & Gordon, S. Immune recognition. A new receptor for beta-glucans. *Nature* **413**, 36-37, doi:10.1038/35092620 (2001).
23 Perrigoue, J. G., Marshall, F. & Artis, D. On the hunt for helminths: Innate immune cells in the recognition and response to helminth parasites. *Cellular microbiology* **10**, 1757-1764 (2008).
24 Vance, R. E., Isberg, R. R. & Portnoy, D. A. Patterns of pathogenesis: discrimination of pathogenic and nonpathogenic microbes by the innate immune system. *Cell host & microbe* **6**, 10-21, doi:10.1016/j.chom.2009.06.007 (2009).
25 Mallick, E. M. *et al.* The ability of an attaching and effacing pathogen to trigger localized actin assembly contributes to virulence by promoting mucosal attachment. *Cellular microbiology* **16**, 1405-1424 (2014).
26 Tilney, L. G. & Portnoy, D. A. Actin filaments and the growth, movement, and spread of the intracellular bacterial parasite, Listeria monocytogenes. *The Journal of cell biology* **109**, 1597-1608 (1989).
27 Goldberg, M. B. Actin-Based Motility of Intracellular Microbial Pathogens. *Microbiology and Molecular Biology Reviews* **65**, 595-626 (2001).
28 Xu, H. *et al.* Innate immune sensing of bacterial modifications of Rho GTPases by the Pyrin inflammasome. *Nature* **513**, 237-241, doi:10.1038/nature13449 (2014).
29 Waite, A. L. *et al.* Pyrin and ASC co-localize to cellular sites that are rich in polymerizing actin. *Experimental biology and medicine (Maywood, N.J.)* **234**, 40-52, doi:10.3181/0806-rm-184 (2009).
30 Chavarria-Smith, J., Mitchell, P. S., Ho, A. M., Daugherty, M. D. & Vance, R. E. Functional and Evolutionary Analyses Identify Proteolysis as a General Mechanism for NLRP1 Inflammasome Activation. *PLoS pathogens* **12**, e1006052, doi:10.1371/journal.ppat.1006052 (2016).

31 Ewald, S. E., Chavarria-Smith, J. & Boothroyd, J. C. NLRP1 is an inflammasome sensor for Toxoplasma gondii. *Infection and immunity* **82**, 460-468, doi:10.1128/iai.01170-13 (2014).

32 Levinsohn, J. L. *et al.* Anthrax lethal factor cleavage of Nlrp1 is required for activation of the inflammasome. *PLoS pathogens* **8**, e1002638, doi:10.1371/journal.ppat.1002638 (2012).

33 Girardin, S. E. *et al.* Nod2 is a general sensor of peptidoglycan through muramyl dipeptide (MDP) detection. *The Journal of biological chemistry* **278**, 8869-8872, doi:10.1074/jbc.C200651200 (2003).

34 Magalhaes, J. G. *et al.* Murine Nod1 but not its human orthologue mediates innate immune detection of tracheal cytotoxin. *EMBO reports* **6**, 1201-1207, doi:10.1038/sj.embor.7400552 (2005).

35 Keestra-Gounder, A. M. *et al.* NOD1 and NOD2 signalling links ER stress with inflammation. *Nature* **532**, 394-397, doi:10.1038/nature17631 (2016).

36 Celli, J. & Tsolis, R. M. Bacteria, the endoplasmic reticulum and the unfolded protein response: friends or foes? *Nature reviews. Microbiology* **13**, 71-82, doi:10.1038/nrmicro3393 (2015).

37 He, B. Viruses, endoplasmic reticulum stress, and interferon responses. *Cell death and differentiation* **13**, 393-403, doi:10.1038/sj.cdd.4401833 (2006).

38 Inacio, P. *et al.* Parasite-induced ER stress response in hepatocytes facilitates Plasmodium liver stage infection. *EMBO reports* **16**, 955-964, doi:10.15252/embr.201439979 (2015).

39 Bronner, D. N. *et al.* Endoplasmic Reticulum Stress Activates the Inflammasome via NLRP3- and Caspase-2-Driven Mitochondrial Damage. *Immunity* **43**, 451-462, doi:10.1016/j.immuni.2015.08.008 (2015).

40 Menu, P. *et al.* ER stress activates the NLRP3 inflammasome via an UPR-independent pathway. *Cell death & disease* **3**, e261, doi:10.1038/cddis.2011.132 (2012).

41 Munoz-Planillo, R. *et al.* K(+) efflux is the common trigger of NLRP3 inflammasome activation by bacterial toxins and particulate matter. *Immunity* **38**, 1142-1153, doi:10.1016/j.immuni.2013.05.016 (2013).

42 Ip, W. K. & Medzhitov, R. Macrophages monitor tissue osmolarity and induce inflammatory response through NLRP3 and NLRC4 inflammasome activation. *Nature communications* **6**, 6931, doi:10.1038/ncomms7931 (2015).

43 Kono, H. & Rock, K. L. How dying cells alert the immune system to danger. *Nature reviews. Immunology* **8**, 279-289, doi:10.1038/nri2215 (2008).

44 Matzinger, P. Tolerance, danger, and the extended family. *Annual review of immunology* **12**, 991-1045, doi:10.1146/annurev.iy.12.040194.005015 (1994).

45 Jozefowski, S. The danger model: questioning an unconvincing theory. *Immunology and cell biology* **94**, 525, doi:10.1038/icb.2016.29 (2016).

46 Chisholm, S. T., Coaker, G., Day, B. & Staskawicz, B. J. Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* **124**, 803-814, doi:10.1016/j.cell.2006.02.008 (2006).

47 Jones, J. D. & Dangl, J. L. The plant immune system. *Nature* **444**, 323-329, doi:10.1038/nature05286 (2006).
48 van der Hoorn, R. A. L. & Kamoun, S. From Guard to Decoy: A New Model for Perception of Plant Pathogen Effectors. *The Plant cell* **20**, 2009-2017 (2008).
49 Axtell, M. J. & Staskawicz, B. J. Initiation of RPS2-specified disease resistance in Arabidopsis is coupled to the AvrRpt2-directed elimination of RIN4. *Cell* **112**, 369-377 (2003).
50 Chung, E. H. *et al.* Specific threonine phosphorylation of a host target by two unrelated type III effectors activates a host innate immune receptor in plants. *Cell host & microbe* **9**, 125-136, doi:10.1016/j.chom.2011.01.009 (2011).
51 Chung, E. H., El-Kasmi, F., He, Y., Loehr, A. & Dangl, J. L. A plant phosphoswitch platform repeatedly targeted by type III effector proteins regulates the output of both tiers of plant immune receptors. *Cell host & microbe* **16**, 484-494, doi:10.1016/j.chom.2014.09.004 (2014).
52 Mackey, D., Holt, B. F., 3rd, Wiig, A. & Dangl, J. L. RIN4 interacts with Pseudomonas syringae type III effector molecules and is required for RPM1-mediated resistance in Arabidopsis. *Cell* **108**, 743-754 (2002).
53 Le Roux, C. *et al.* A receptor pair with an integrated decoy converts pathogen disabling of transcription factors to immunity. *Cell* **161**, 1074-1088, doi:10.1016/j.cell.2015.04.025 (2015).
54 Sarris, P. F. *et al.* A Plant Immune Receptor Detects Pathogen Effectors that Target WRKY Transcription Factors. *Cell* **161**, 1089-1100, doi:10.1016/j.cell.2015.04.024 (2015).
55 Jones, J. D., Vance, R. E. & Dangl, J. L. Intracellular innate immune surveillance devices in plants and animals. *Science* **354**, doi:10.1126/science.aaf6395 (2016).
56 Gomez-Gomez, L. & Boller, T. FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in Arabidopsis. *Molecular cell* **5**, 1003-1011 (2000).
57 Sun, Y. *et al.* Structural basis for flg22-induced activation of the Arabidopsis FLS2-BAK1 immune complex. *Science* **342**, 624-628, doi:10.1126/science.1243825 (2013).
58 Spoel, S. H. & Dong, X. How do plants achieve immunity? Defence without specialized immune cells. *Nature reviews. Immunology* **12**, 89-100, doi:10.1038/nri3141 (2012).
59 Ausubel, F. M. Are innate immune signaling pathways in plants and animals conserved? *Nature immunology* **6**, 973-979, doi:10.1038/ni1253 (2005).
60 Rescigno, M. *et al.* Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria. *Nature immunology* **2**, 361-367, doi:10.1038/86373 (2001).
61 Luzio, J. P., Pryor, P. R. & Bright, N. A. Lysosomes: fusion and function. *Nature reviews. Molecular cell biology* **8**, 622-632, doi:10.1038/nrm2217 (2007).
62 Hemmi, H. *et al.* A Toll-like receptor recognizes bacterial DNA. *Nature* **408**, 740-745, doi:10.1038/35047123 (2000).

63 Kato, H. *et al.* Cell type-specific involvement of RIG-I in antiviral response. *Immunity* **23**, 19-28, doi:10.1016/j.immuni.2005.04.010 (2005).
64 Fernandes-Alnemri, T., Yu, J. W., Datta, P., Wu, J. & Alnemri, E. S. AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. *Nature* **458**, 509-513, doi:10.1038/nature07710 (2009).
65 Hornung, V. *et al.* AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* **458**, 514-518, doi:10.1038/nature07725 (2009).
66 Cai, X., Chiu, Y. H. & Chen, Z. J. The cGAS-cGAMP-STING pathway of cytosolic DNA sensing and signaling. *Molecular cell* **54**, 289-296, doi:10.1016/j.molcel.2014.03.040 (2014).
67 Burdette, D. L. & Vance, R. E. STING and the innate immune response to nucleic acids in the cytosol. *Nature immunology* **14**, 19-26, doi:10.1038/ni.2491 (2013).
68 Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* **339**, 786-791, doi:10.1126/science.1232458 (2013).
69 Alexopoulou, L., Holt, A. C., Medzhitov, R. & Flavell, R. A. Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature* **413**, 732-738, doi:10.1038/35099560 (2001).
70 Yoneyama, M. *et al.* The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nature immunology* **5**, 730-737, doi:10.1038/ni1087 (2004).
71 Takeuchi, O. & Akira, S. Pattern recognition receptors and inflammation. *Cell* **140**, 805-820, doi:10.1016/j.cell.2010.01.022 (2010).
72 Shi, J. *et al.* Inflammatory caspases are innate immune receptors for intracellular LPS. *Nature* **514**, 187-192, doi:10.1038/nature13683 (2014).
73 Kayagaki, N. *et al.* Non-canonical inflammasome activation targets caspase-11. *Nature* **479**, 117-121, doi:10.1038/nature10558 (2011).
74 Hayashi, F. *et al.* The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature* **410**, 1099-1103, doi:10.1038/35074106 (2001).
75 Ren, T., Zamboni, D. S., Roy, C. R., Dietrich, W. F. & Vance, R. E. Flagellin-deficient Legionella mutants evade caspase-1- and Naip5-mediated macrophage immunity. *PLoS pathogens* **2**, e18, doi:10.1371/journal.ppat.0020018 (2006).
76 Zamboni, D. S. *et al.* The Birc1e cytosolic pattern-recognition receptor contributes to the detection and control of Legionella pneumophila infection. *Nature immunology* **7**, 318-325, doi:10.1038/ni1305 (2006).
77 Fontana, M. F. & Vance, R. E. Two signal models in innate immunity. *Immunological reviews* **243**, 26-39, doi:10.1111/j.1600-065X.2011.01037.x (2011).
78 Broz, P. & Dixit, V. M. Inflammasomes: mechanism of assembly, regulation and signalling. *Nature reviews. Immunology* **16**, 407-420, doi:10.1038/nri.2016.58 (2016).

79 Dinarello, C. A. Interleukin-1, interleukin-1 receptors and interleukin-1 receptor antagonist. *International reviews of immunology* **16**, 457-499, doi:10.3109/08830189809043005 (1998).
80 von Moltke, J., Ayres, J. S., Kofoed, E. M., Chavarria-Smith, J. & Vance, R. E. Recognition of bacteria by inflammasomes. *Annual review of immunology* **31**, 73-106, doi:10.1146/annurev-immunol-032712-095944 (2013).
81 Martinon, F., Burns, K. & Tschopp, J. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL-beta. *Molecular cell* **10**, 417-426 (2002).
82 Shi, Y. Caspase activation: revisiting the induced proximity model. *Cell* **117**, 855-858, doi:10.1016/j.cell.2004.06.007 (2004).
83 Miller, L. S. *et al.* Inflammasome-mediated production of IL-1beta is required for neutrophil recruitment against Staphylococcus aureus in vivo. *J Immunol* **179**, 6933-6942 (2007).
84 Schroder, K., Hertzog, P. J., Ravasi, T. & Hume, D. A. Interferon-gamma: an overview of signals, mechanisms and functions. *Journal of leukocyte biology* **75**, 163-189, doi:10.1189/jlb.0603252 (2004).
85 Bergsbaken, T., Fink, S. L. & Cookson, B. T. Pyroptosis: host cell death and inflammation. *Nature reviews. Microbiology* **7**, 99-109, doi:10.1038/nrmicro2070 (2009).
86 Case, C. L. & Roy, C. R. Asc modulates the function of NLRC4 in response to infection of macrophages by Legionella pneumophila. *mBio* **2**, doi:10.1128/mBio.00117-11 (2011).
87 Chen, K. W. *et al.* The neutrophil NLRC4 inflammasome selectively promotes IL-1beta maturation without pyroptosis during acute Salmonella challenge. *Cell reports* **8**, 570-582, doi:10.1016/j.celrep.2014.06.028 (2014).
88 Kayagaki, N. *et al.* Caspase-11 cleaves gasdermin D for non-canonical inflammasome signalling. *Nature* **526**, 666-671, doi:10.1038/nature15541 (2015).
89 Shi, J. *et al.* Cleavage of GSDMD by inflammatory caspases determines pyroptotic cell death. *Nature* **526**, 660-665, doi:10.1038/nature15514 (2015).
90 Ding, J. *et al.* Pore-forming activity and structural autoinhibition of the gasdermin family. *Nature* **535**, 111-116, doi:10.1038/nature18590 (2016).
91 Liu, X. *et al.* Inflammasome-activated gasdermin D causes pyroptosis by forming membrane pores. *Nature* **535**, 153-158, doi:10.1038/nature18629 (2016).
92 Miao, E. A. *et al.* Caspase-1-induced pyroptosis is an innate immune effector mechanism against intracellular bacteria. *Nature immunology* **11**, 1136-1142, doi:10.1038/ni.1960 (2010).
93 Sauer, J. D. *et al.* Listeria monocytogenes engineered to activate the Nlrc4 inflammasome are severely attenuated and are poor inducers of protective immunity. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 12419-12424, doi:10.1073/pnas.1019041108 (2011).
94 Jorgensen, I., Zhang, Y., Krantz, B. A. & Miao, E. A. Pyroptosis triggers pore-induced intracellular traps (PITs) that capture bacteria and lead to their clearance

by efferocytosis. *The Journal of experimental medicine* **213**, 2113-2128, doi:10.1084/jem.20151613 (2016).

95 Antonopoulos, C. *et al.* Caspase-8 as an Effector and Regulator of NLRP3 Inflammasome Signaling. *The Journal of biological chemistry* **290**, 20167-20184, doi:10.1074/jbc.M115.652321 (2015).

96 Rauch, I. *et al.* NAIP-NLRC4 Inflammasomes Coordinate Intestinal Epithelial Cell Expulsion with Eicosanoid and IL-18 Release via Activation of Caspase-1 and -8. *Immunity* (2017).

97 Mostowy, S. & Shenoy, A. R. The cytoskeleton in cell-autonomous immunity: structural determinants of host defence. *Nature reviews. Immunology* **15**, 559-573 (2015).

98 Chavarria-Smith, J. & Vance, R. E. Direct proteolytic cleavage of NLRP1B is necessary and sufficient for inflammasome activation by anthrax lethal factor. *PLoS pathogens* **9**, e1003452, doi:10.1371/journal.ppat.1003452 (2013).

99 Hellmich, K. A. *et al.* Anthrax lethal factor cleaves mouse nlrp1b in both toxin-sensitive and toxin-resistant macrophages. *PloS one* **7**, e49741, doi:10.1371/journal.pone.0049741 (2012).

100 Hornung, V. *et al.* Silica crystals and aluminum salts activate the NALP3 inflammasome through phagosomal destabilization. *Nature immunology* **9**, 847-856, doi:10.1038/ni.1631 (2008).

101 Iyer, S. S. *et al.* Mitochondrial cardiolipin is required for Nlrp3 inflammasome activation. *Immunity* **39**, 311-323, doi:10.1016/j.immuni.2013.08.001 (2013).

102 Jin, T. *et al.* Structures of the HIN domain:DNA complexes reveal ligand binding and activation mechanisms of the AIM2 inflammasome and IFI16 receptor. *Immunity* **36**, 561-571, doi:10.1016/j.immuni.2012.02.014 (2012).

103 Rathinam, V. A. *et al.* The AIM2 inflammasome is essential for host defense against cytosolic bacteria and DNA viruses. *Nature immunology* **11**, 395-402, doi:10.1038/ni.1864 (2010).

104 Jones, J. W. *et al.* Absent in melanoma 2 is required for innate immune recognition of Francisella tularensis. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 9771-9776, doi:10.1073/pnas.1003738107 (2010).

105 Sauer, J. D. *et al.* Listeria monocytogenes triggers AIM2-mediated pyroptosis upon infrequent bacteriolysis in the macrophage cytosol. *Cell host & microbe* **7**, 412-419, doi:10.1016/j.chom.2010.04.004 (2010).

106 Khare, S. *et al.* An NLRP7-containing inflammasome mediates recognition of microbial lipopeptides in human macrophages. *Immunity* **36**, 464-476, doi:10.1016/j.immuni.2012.02.001 (2012).

107 von Moltke, J. *et al.* Rapid induction of inflammatory lipid mediators by the inflammasome in vivo. *Nature* **490**, 107-111, doi:10.1038/nature11351 (2012).

108 Jorgensen, I., Rayamajhi, M. & Miao, E. A. Programmed cell death as a defence against infection. *Nature reviews. Immunology* **17**, 151-164, doi:10.1038/nri.2016.147 (2017).

109 Sellin, M. E. *et al.* Epithelium-intrinsic NAIP/NLRC4 inflammasome drives infected enterocyte expulsion to restrict Salmonella replication in the intestinal mucosa. *Cell host & microbe* **16**, 237-248, doi:10.1016/j.chom.2014.07.001 (2014).
110 Zhong, F. L. *et al.* Germline NLRP1 Mutations Cause Skin Inflammatory and Cancer Susceptibility Syndromes via Inflammasome Activation. *Cell* **167**, 187-202 e117, doi:10.1016/j.cell.2016.09.001 (2016).
111 Erzberger, J. P. & Berger, J. M. Evolutionary relationships and structural mechanisms of AAA+ proteins. *Annual review of biophysics and biomolecular structure* **35**, 93-114, doi:10.1146/annurev.biophys.35.040405.101933 (2006).
112 Zhang, L. *et al.* Cryo-EM structure of the activated NAIP2-NLRC4 inflammasome reveals nucleated polymerization. *Science* **350**, 404-409, doi:10.1126/science.aac5789 (2015).
113 Broz, P., von Moltke, J., Jones, J. W., Vance, R. E. & Monack, D. M. Differential requirement for Caspase-1 autoproteolysis in pathogen-induced cell death and cytokine processing. *Cell host & microbe* **8**, 471-483, doi:10.1016/j.chom.2010.11.007 (2010).
114 Urbach, J. M. & Ausubel, F. M. The NBS-LRR architectures of plant R-proteins and metazoan NLRs evolved in independent events. *Proceedings of the National Academy of Sciences of the United States of America* **114**, 1063-1068, doi:10.1073/pnas.1619730114 (2017).
115 Glass, N. L. & Kaneko, I. Fatal attraction: nonself recognition and heterokaryon incompatibility in filamentous fungi. *Eukaryotic cell* **2**, 1-8 (2003).
116 Dyrka, W. *et al.* Diversity and variability of NOD-like receptors in fungi. *Genome biology and evolution* **6**, 3137-3158, doi:10.1093/gbe/evu251 (2014).
117 Daskalov, A. *et al.* Signal transduction by a fungal NOD-like receptor based on propagation of a prion amyloid fold. *PLoS biology* **13**, e1002059, doi:10.1371/journal.pbio.1002059 (2015).
118 Lightfield, K. L. *et al.* Critical function for Naip5 in inflammasome activation by a conserved carboxy-terminal domain of flagellin. *Nature immunology* **9**, 1171-1178, doi:10.1038/ni.1646 (2008).
119 Molofsky, A. B. *et al.* Cytosolic recognition of flagellin by mouse macrophages restricts Legionella pneumophila infection. *The Journal of experimental medicine* **203**, 1093-1104, doi:10.1084/jem.20051659 (2006).
120 Franchi, L. *et al.* Cytosolic flagellin requires Ipaf for activation of caspase-1 and interleukin 1beta in salmonella-infected macrophages. *Nature immunology* **7**, 576-582, doi:10.1038/ni1346 (2006).
121 Miao, E. A. *et al.* Cytoplasmic flagellin activates caspase-1 and secretion of interleukin 1beta via Ipaf. *Nature immunology* **7**, 569-575, doi:10.1038/ni1344 (2006).
122 Miao, E. A. *et al.* Innate immune detection of the type III secretion apparatus through the NLRC4 inflammasome. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 3076-3080, doi:10.1073/pnas.0913087107 (2010).

123 Tenthorey, J. L., Kofoed, E. M., Daugherty, M. D., Malik, H. S. & Vance, R. E. Molecular basis for specific recognition of bacterial ligands by NAIP/NLRC4 inflammasomes. *Molecular cell* **54**, 17-29, doi:10.1016/j.molcel.2014.02.018 (2014).

124 Lightfield, K. L. *et al.* Differential requirements for NAIP5 in activation of the NLRC4 inflammasome. *Infection and immunity* **79**, 1606-1614, doi:10.1128/IAI.01187-10 (2011).

125 Rayamajhi, M., Zak, D. E., Chavarria-Smith, J., Vance, R. E. & Miao, E. A. Cutting Edge: Mouse NAIP1 Detects the Type III Secretion System Needle Protein. *J Immunol*, doi:10.4049/jimmunol.1301549 (2013).

126 Yang, J., Zhao, Y., Shi, J. & Shao, F. Human NAIP and mouse NAIP1 recognize bacterial type III secretion needle protein for inflammasome activation. *Proceedings of the National Academy of Sciences of the United States of America*, doi:10.1073/pnas.1306376110 (2013).

127 Rauch, I. *et al.* NAIP proteins are required for cytosolic detection of specific bacterial ligands in vivo. *The Journal of experimental medicine* **213**, 657-665, doi:10.1084/jem.20151809 (2016).

128 Zhao, Y. *et al.* Genetic functions of the NAIP family of inflammasome receptors for bacterial ligands in mice. *The Journal of experimental medicine* **213**, 647-656 (2016).

129 Romanish, M. T., Nakamura, H., Lai, C. B., Wang, Y. & Mager, D. L. A novel protein isoform of the multicopy human NAIP gene derives from intragenic Alu SINE promoters. *PloS one* **4**, e5761, doi:10.1371/journal.pone.0005761 (2009).

130 Boniotto, M. *et al.* Population variation in NAIP functional copy number confers increased cell death upon Legionella pneumophila infection. *Human immunology* **73**, 196-200, doi:10.1016/j.humimm.2011.10.014 (2012).

131 Kortmann, J., Brubaker, S. W. & Monack, D. M. Cutting Edge: Inflammasome Activation in Primary Human Macrophages Is Dependent on Flagellin. *J Immunol* **195**, 815-819, doi:10.4049/jimmunol.1403100 (2015).

132 Diepold, A. & Armitage, J. P. Type III secretion systems: the bacterial flagellum and the injectisome. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **370**, doi:10.1098/rstb.2015.0020 (2015).

133 Sun, Y. H., Rolan, H. G. & Tsolis, R. M. Injection of flagellin into the host cell cytosol by Salmonella enterica serotype Typhimurium. *The Journal of biological chemistry* **282**, 33897-33901, doi:10.1074/jbc.C700181200 (2007).

134 Aizawa, S. I., Vonderviszt, F., Ishima, R. & Akasaka, K. Termini of Salmonella flagellin are disordered and become organized upon polymerization into flagellar filament. *Journal of molecular biology* **211**, 673-677, doi:10.1016/0022-2836(90)90064-s (1990).

135 Vonderviszt, F., Kanto, S., Aizawa, S. & Namba, K. Terminal regions of flagellin are disordered in solution. *Journal of molecular biology* **209**, 127-133 (1989).

136 Zhong, D. *et al.* The Salmonella type III secretion system inner rod protein PrgJ is partially folded. *The Journal of biological chemistry* **287**, 25303-25311, doi:10.1074/jbc.M112.381574 (2012).

137 Zhang, L., Wang, Y., Picking, W. L., Picking, W. D. & De Guzman, R. N. Solution structure of monomeric BsaL, the type III secretion needle protein of Burkholderia pseudomallei. *Journal of molecular biology* **359**, 322-330, doi:10.1016/j.jmb.2006.03.028 (2006).

138 Poyraz, O. *et al.* Protein refolding is required for assembly of the type three secretion needle. *Nature structural & molecular biology* **17**, 788-792, doi:10.1038/nsmb.1822 (2010).

139 Loquet, A. *et al.* Atomic model of the type III secretion system needle. *Nature* **486**, 276-279, doi:10.1038/nature11079 (2012).

140 Yonekura, K., Maki-Yonekura, S. & Namba, K. Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424**, 643-650, doi:10.1038/nature01830 (2003).

141 Kathariou, S., Kanenaka, R., Allen, R. D., Fok, A. K. & Mizumoto, C. Repression of motility and flagellin production at 37 degrees C is stronger in Listeria monocytogenes than in the nonpathogenic species Listeria innocua. *Canadian journal of microbiology* **41**, 572-577 (1995).

142 Warren, S. E. *et al.* Generation of a Listeria vaccine strain by enhanced caspase-1 activation. *European journal of immunology* **41**, 1934-1940, doi:10.1002/eji.201041214 (2011).

143 Brundage, R. A., Smith, G. A., Camilli, A., Theriot, J. A. & Portnoy, D. A. Expression and phosphorylation of the Listeria monocytogenes ActA protein in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 11890-11894 (1993).

144 Choe, J. E. & Welch, M. D. Actin-based motility of bacterial pathogens: mechanistic diversity and its impact on virulence. *Pathogens and disease*, doi:10.1093/femspd/ftw099 (2016).

145 Monack, D. M., Detweiler, C. S. & Falkow, S. Salmonella pathogenicity island 2-dependent macrophage death is mediated in part by the host cysteine protease caspase-1. *Cellular microbiology* **3**, 825-837 (2001).

146 Deiwick, J., Nikolaus, T., Erdogan, S. & Hensel, M. Environmental regulation of Salmonella pathogenicity island 2 gene expression. *Molecular microbiology* **31**, 1759-1773 (1999).

147 Hautefort, I. *et al.* During infection of epithelial cells Salmonella enterica serovar Typhimurium undergoes a time-dependent transcriptional adaptation that results in simultaneous expression of three type 3 secretion systems. *Cellular microbiology* **10**, 958-984 (2008).

148 Winter, S. E. *et al.* Gut inflammation provides a respiratory electron acceptor for Salmonella. *Nature* **467**, 426-429, doi:10.1038/nature09415 (2010).

149 Brodsky, I. E. *et al.* A Yersinia effector protein promotes virulence by preventing inflammasome recognition of the type III secretion system. *Cell host & microbe* **7**, 376-387, doi:10.1016/j.chom.2010.04.009 (2010).

150 LaRock, C. N. & Cookson, B. T. The Yersinia virulence effector YopM binds caspase-1 to arrest inflammasome assembly and processing. *Cell host & microbe* **12**, 799-805, doi:10.1016/j.chom.2012.10.020 (2012).

151 Chung, L. K. *et al.* IQGAP1 is important for activation of caspase-1 in macrophages and is targeted by Yersinia pestis type III effector YopM. *mBio* **5**, e01402-01414, doi:10.1128/mBio.01402-14 (2014).
152 Ratner, D. *et al.* The Yersinia pestis Effector YopM Inhibits Pyrin Inflammasome Activation. *PLoS pathogens* **12**, e1006035, doi:10.1371/journal.ppat.1006035 (2016).
153 Rauch, I. *et al.* NAIP-NLRC4 Inflammasomes Coordinate Intestinal Epithelial Cell Expulsion with Eicosanoid and IL-18 Release via Activation of Caspase-1 and -8. *Immunity*, doi:10.1016/j.immuni.2017.03.016 (2017).
154 Canna, S. W. *et al.* An activating NLRC4 inflammasome mutation causes autoinflammation with recurrent macrophage activation syndrome. *Nature genetics* **46**, 1140-1146, doi:10.1038/ng.3089 (2014).
155 Kitamura, A., Sasaki, Y., Abe, T., Kano, H. & Yasutomo, K. An inherited mutation in NLRC4 causes autoinflammation in human and mice. *The Journal of experimental medicine* **211**, 2385-2396, doi:10.1084/jem.20141091 (2014).
156 Romberg, N. *et al.* Mutation of NLRC4 causes a syndrome of enterocolitis and autoinflammation. *Nature genetics* **46**, 1135-1139, doi:10.1038/ng.3066 (2014).
157 Liang, J. *et al.* Novel NLRC4 Mutation Causes a Syndrome of Perinatal Autoinflammation With Hemophagocytic Lymphohistiocytosis, Hepatosplenomegaly, Fetal Thrombotic Vasculopathy, and Congenital Anemia and Ascites. *Pediatric and developmental pathology : the official journal of the Society for Pediatric Pathology and the Paediatric Pathology Society*, 1093526616686890, doi:10.1177/1093526616686890 (2017).
158 Volker-Touw, C. M. *et al.* Erythematous nodes, urticarial rash and arthralgias in a large pedigree with NLRC4-related autoinflammatory disease, expansion of the phenotype. *The British journal of dermatology* **176**, 244-248, doi:10.1111/bjd.14757 (2017).
159 Hu, Z. *et al.* Crystal structure of NLRC4 reveals its autoinhibition mechanism. *Science* **341**, 172-175, doi:10.1126/science.1236381 (2013).
160 Allam, R. *et al.* Epithelial NAIPs protect against colonic tumorigenesis. *The Journal of experimental medicine* **212**, 369-383, doi:10.1084/jem.20140474 (2015).
161 Hu, B. *et al.* Inflammation-induced tumorigenesis in the colon is regulated by caspase-1 and NLRC4. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 21635-21640, doi:10.1073/pnas.1016814108 (2010).
162 Janowski, A. M. *et al.* NLRC4 suppresses melanoma tumor progression independently of inflammasome activation. *The Journal of clinical investigation* **126**, 3917-3928, doi:10.1172/jci86953 (2016).
163 Kolb, R. *et al.* Obesity-associated NLRC4 inflammasome activation drives breast cancer progression. *Nature communications* **7** (2016).
164 Lin, C. & Zhang, J. Inflammasomes in Inflammation-Induced Cancer. *Frontiers in immunology* **8**, 271, doi:10.3389/fimmu.2017.00271 (2017).

165 Halff, E. F. *et al.* Formation and structure of a NAIP5-NLRC4 inflammasome induced by direct interactions with conserved N- and C-terminal regions of flagellin. *The Journal of biological chemistry* **287**, 38460-38472, doi:10.1074/jbc.M112.393512 (2012).
166 Schroder, K. & Tschopp, J. The inflammasomes. *Cell* **140**, 821-832, doi:10.1016/j.cell.2010.01.040 (2010).
167 Yang, H., Antoine, D. J., Andersson, U. & Tracey, K. J. The many faces of HMGB1: molecular structure-functional activity in inflammation, apoptosis, and chemotaxis. *Journal of leukocyte biology* **93**, 865-873, doi:10.1189/jlb.1212662 (2013).
168 Bonardi, V., Cherkis, K., Nishimura, M. T. & Dangl, J. L. A new eye on NLR proteins: focused on clarity or diffused by complexity? *Curr Opin Immunol* **24**, 41-50, doi:10.1016/j.coi.2011.12.006 (2012).
169 Leipe, D. D., Koonin, E. V. & Aravind, L. STAND, a class of P-loop NTPases including animal and plant regulators of programmed cell death: multiple, complex domain architectures, unusual phyletic patterns, and evolution by horizontal gene transfer. *Journal of molecular biology* **343**, 1-28, doi:10.1016/j.jmb.2004.08.023 (2004).
170 Danot, O., Marquenet, E., Vidal-Ingigliardi, D. & Richet, E. Wheel of Life, Wheel of Death: A Mechanistic Insight into Signaling by STAND Proteins. *Structure* **17**, 172-182, doi:10.1016/j.str.2009.01.001 (2009).
171 Qi, S. *et al.* Crystal structure of the Caenorhabditis elegans apoptosome reveals an octameric assembly of CED-4. *Cell* **141**, 446-457, doi:10.1016/j.cell.2010.03.017 (2010).
172 Poyet, J. L. *et al.* Identification of Ipaf, a human caspase-1-activating protein related to Apaf-1. *The Journal of biological chemistry* **276**, 28309-28313, doi:10.1074/jbc.C100250200 (2001).
173 Tanabe, T. *et al.* Regulatory regions and critical residues of NOD2 involved in muramyl dipeptide recognition. *The EMBO journal* **23**, 1587-1597, doi:10.1038/sj.emboj.7600175 (2004).
174 Song, D. H. & Lee, J. O. Sensing of microbial molecular patterns by Toll-like receptors. *Immunological reviews* **250**, 216-229, doi:10.1111/j.1600-065X.2012.01167.x (2012).
175 Faustin, B. *et al.* Reconstituted NALP1 inflammasome reveals two-step mechanism of caspase-1 activation. *Molecular cell* **25**, 713-724, doi:10.1016/j.molcel.2007.01.032 (2007).
176 Mitchell, P. S. *et al.* Evolution-guided identification of antiviral specificity determinants in the broadly acting interferon-induced innate immunity factor MxA. *Cell host & microbe* **12**, 598-604, doi:10.1016/j.chom.2012.09.005 (2012).
177 Ellis, J. G., Lawrence, G. J., Luck, J. E. & Dodds, P. N. Identification of regions in alleles of the flax rust resistance gene L that determine differences in gene-for-gene specificity. *The Plant cell* **11**, 495-506 (1999).
178 Dodds, P. N., Lawrence, G. J. & Ellis, J. G. Six amino acid changes confined to the leucine-rich repeat beta-strand/beta-turn motif determine the difference

between the P and P2 rust resistance specificities in flax. *The Plant cell* **13**, 163-178 (2001).

179 Jia, Y., McAdams, S. A., Bryan, G. T., Hershey, H. P. & Valent, B. Direct interaction of resistance gene and avirulence gene products confers rice blast resistance. *The EMBO journal* **19**, 4004-4014, doi:10.1093/emboj/19.15.4004 (2000).

180 Krasileva, K. V., Dahlbeck, D. & Staskawicz, B. J. Activation of an Arabidopsis resistance protein is specified by the in planta association of its leucine-rich repeat domain with the cognate oomycete effector. *The Plant cell* **22**, 2444-2458, doi:10.1105/tpc.110.075358 (2010).

181 Girardin, S. E. *et al.* Identification of the critical residues involved in peptidoglycan detection by Nod1. *The Journal of biological chemistry* **280**, 38648-38656, doi:10.1074/jbc.M509537200 (2005).

182 Kiefer, F., Arnold, K., Kunzli, M., Bordoli, L. & Schwede, T. The SWISS-MODEL Repository and associated resources. *Nucleic acids research* **37**, D387-392, doi:10.1093/nar/gkn750 (2009).

183 McDowell, J. M. *et al.* Intragenic recombination and diversifying selection contribute to the evolution of downy mildew resistance at the RPP8 locus of Arabidopsis. *The Plant cell* **10**, 1861-1874 (1998).

184 Reubold, T. F., Wohlgemuth, S. & Eschenburg, S. Crystal structure of full-length Apaf-1: how the death signal is relayed in the mitochondrial pathway of apoptosis. *Structure* **19**, 1074-1083, doi:10.1016/j.str.2011.05.013 (2011).

185 Kofoed, E. M. & Vance, R. E. Blue native polyacrylamide gel electrophoresis to monitor inflammasome assembly and composition. *Methods Mol Biol* **1040**, 169-183, doi:10.1007/978-1-62703-523-1_13 (2013).

186 Wright, E. K. *et al.* Naip5 affects host susceptibility to the intracellular pathogen Legionella pneumophila. *Current biology : CB* **13**, 27-36 (2003).

187 Kelley, L. A. & Sternberg, M. J. Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols* **4**, 363-371, doi:10.1038/nprot.2009.2 (2009).

188 Kosakovsky Pond, S. L., Posada, D., Gravenor, M. B., Woelk, C. H. & Frost, S. D. GARD: a genetic algorithm for recombination detection. *Bioinformatics* **22**, 3096-3098, doi:10.1093/bioinformatics/btl474 (2006).

189 Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Molecular biology and evolution* **24**, 1586-1591, doi:10.1093/molbev/msm088 (2007).

190 Comeron, J. M. K-Estimator: calculation of the number of nucleotide substitutions per site and the confidence intervals. *Bioinformatics* **15**, 763-764 (1999).

191 Quintana-Murci, L. & Clark, A. G. Population genetic tools for dissecting innate immunity in humans. *Nature reviews. Immunology* **13**, 280-293, doi:10.1038/nri3421 (2013).

192 Tiffin, P. & Moeller, D. A. Molecular evolution of plant immune system genes. *Trends in genetics : TIG* **22**, 662-670, doi:10.1016/j.tig.2006.09.011 (2006).

193 Viljakainen, L. Evolutionary genetics of insect innate immunity. *Briefings in functional genomics* **14**, 407-412, doi:10.1093/bfgp/elv002 (2015).

194 Barton, G. M., Kagan, J. C. & Medzhitov, R. Intracellular localization of Toll-like receptor 9 prevents recognition of self DNA but facilitates access to viral DNA. *Nature immunology* **7**, 49-56, doi:10.1038/ni1280 (2006).
195 Ewald, S. E. *et al.* The ectodomain of Toll-like receptor 9 is cleaved to generate a functional receptor. *Nature* **456**, 658-662, doi:10.1038/nature07405 (2008).
196 Stetson, D. B. & Medzhitov, R. Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. *Immunity* **24**, 93-103, doi:10.1016/j.immuni.2005.12.003 (2006).
197 Pohar, J. *et al.* Selectivity of Human TLR9 for Double CpG Motifs and Implications for the Recognition of Genomic DNA. *J Immunol* **198**, 2093-2104, doi:10.4049/jimmunol.1600757 (2017).
198 Gray, E. E., Treuting, P. M., Woodward, J. J. & Stetson, D. B. Cutting Edge: cGAS Is Required for Lethal Autoimmune Disease in the Trex1-Deficient Mouse Model of Aicardi-Goutieres Syndrome. *J Immunol* **195**, 1939-1943, doi:10.4049/jimmunol.1500969 (2015).
199 Hughes, D. & Andersson, D. I. Evolutionary consequences of drug resistance: shared principles across diverse targets and organisms. *Nature reviews. Genetics* **16**, 459-471, doi:10.1038/nrg3922 (2015).
200 Smith, K. D. *et al.* Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility. *Nature immunology* **4**, 1247-1253, doi:10.1038/ni1011 (2003).
201 Gugolya, Z., Muskotal, A., Sebestyen, A., Dioszeghy, Z. & Vonderviszt, F. Interaction of the disordered terminal regions of flagellin upon flagellar filament formation. *FEBS letters* **535**, 66-70 (2003).
202 Chou, S. *et al.* Hyaloperonospora arabidopsidis ATR1 effector is a repeat protein with distributed recognition surfaces. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 13323-13328, doi:10.1073/pnas.1109791108 (2011).
203 Yoon, S. I. *et al.* Structural basis of TLR5-flagellin recognition and signaling. *Science* **335**, 859-864, doi:10.1126/science.1215584 (2012).
204 Song, W. S., Jeon, Y. J., Namgung, B., Hong, M. & Yoon, S. I. A conserved TLR5 binding and activation hot spot on flagellin. *Scientific reports* **7**, 40878, doi:10.1038/srep40878 (2017).
205 Andersen-Nissen, E. *et al.* Evasion of Toll-like receptor 5 by flagellated bacteria. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 9247-9252, doi:10.1073/pnas.0502040102 (2005).
206 Smith, S. A. *et al.* Adaptive evolution of Toll-like receptor 5 in domesticated mammals. *BMC evolutionary biology* **12**, 122, doi:10.1186/1471-2148-12-122 (2012).
207 Mariathasan, S. *et al.* Differential activation of the inflammasome by caspase-1 adaptors ASC and Ipaf. *Nature* **430**, 213-218, doi:10.1038/nature02664 (2004).
208 Kim, J., Webb, A. M., Kershner, J. P., Blaskowski, S. & Copley, S. D. A versatile and highly efficient method for scarless genome editing in Escherichia coli and

Salmonella enterica. *BMC biotechnology* **14**, 84, doi:10.1186/1472-6750-14-84 (2014).

209 Decker, T. & Lohmann-Matthes, M. L. A quick and simple method for the quantitation of lactate dehydrogenase release in measurements of cellular cytotoxicity and tumor necrosis factor (TNF) activity. *Journal of immunological methods* **115**, 61-69 (1988).

210 Kersse, K., Verspurten, J., Vanden Berghe, T. & Vandenabeele, P. The death-fold superfamily of homotypic interaction motifs. *Trends in biochemical sciences* **36**, 541-552, doi:10.1016/j.tibs.2011.06.006 (2011).

211 Lamkanfi, M. & Dixit, V. M. Mechanisms and functions of inflammasomes. *Cell* **157**, 1013-1022, doi:10.1016/j.cell.2014.04.007 (2014).

212 Rathinam, V. A. & Fitzgerald, K. A. Inflammasome Complexes: Emerging Mechanisms and Effector Functions. *Cell* **165**, 792-800, doi:10.1016/j.cell.2016.03.046 (2016).

213 Kersse, K., Lamkanfi, M., Bertrand, M. J., Vanden Berghe, T. & Vandenabeele, P. Interaction patches of procaspase-1 caspase recruitment domains (CARDs) are differently involved in procaspase-1 activation and receptor-interacting protein 2 (RIP2)-dependent nuclear factor kappaB signaling. *The Journal of biological chemistry* **286**, 35874-35882, doi:10.1074/jbc.M111.242321 (2011).

214 Yang, J. *et al.* The I-TASSER Suite: protein structure and function prediction. *Nat Methods* **12**, 7-8, doi:10.1038/nmeth.3213 (2015).

215 Ramon Lopez-Blanco, J. & Chacon, P. iMODFIT: Efficient and robust flexible fitting based on vibrational analysis in internal coordinates. *Journal of Structural Biology* **184**, 261-270, doi:10.1016/j.jsb.2013.08.010 (2013).

216 Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221, doi:10.1107/S0907444909052925 (2010).

217 Yuan, S. & Akey, C. W. Apoptosome structure, assembly, and procaspase activation. *Structure* **21**, 501-515, doi:10.1016/j.str.2013.02.024 (2013).

218 Qu, Y. *et al.* Phosphorylation of NLRC4 is critical for inflammasome activation. *Nature* **490**, 539-542, doi:10.1038/nature11429 (2012).

219 Lu, A. *et al.* Molecular basis of caspase-1 polymerization and its inhibition by a new capping mechanism. *Nature structural & molecular biology* **23**, 416-425, doi:10.1038/nsmb.3199 (2016).

220 Lu, A. *et al.* Unified polymerization mechanism for the assembly of ASC-dependent inflammasomes. *Cell* **156**, 1193-1206, doi:10.1016/j.cell.2014.02.008 (2014).

221 Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Meth* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).

222 Zhang, K. Gctf: Real-time CTF determination and correction. *Journal of structural biology* **193**, 1-12, doi:10.1016/j.jsb.2015.11.003 (2016).

223 Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *Journal of structural biology* **180**, 519-530, doi:10.1016/j.jsb.2012.09.006 (2012).
224 Scheres, S. H. W. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Meth.* **9**, 853-854 (2012).
225 Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721-745 (2003).
226 Heymann, J. B. Bsoft: image and molecular processing in electron microscopy. *Journal of structural biology* **133**, 156-169, doi:10.1006/jsbi.2001.4339 (2001).
227 Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612, doi:10.1002/jcc.20084 (2004).
228 Garzon, J. I., Kovacs, J., Abagyan, R. & Chacon, P. ADP_EM: fast exhaustive multi-resolution docking for high-throughput coverage. *Bioinformatics* **23**, 427-433, doi:10.1093/bioinformatics/btl625 (2007).
229 Chacon, P. & Wriggers, W. Multi-resolution contour-based fitting of macromolecular structures. *Journal of Molecular Biology* **317**, 375-384, doi:10.1006/jmbi.2002.5438 (2002).
230 Ramon Lopez-Blanco, J., Jesus Canosa-Valls, A., Li, Y. & Chacon, P. RCD plus : Fast loop modeling server. *Nucleic Acids Research* **44**, W395-W400, doi:10.1093/nar/gkw395 (2016).
231 Lopez-Blanco, J. R., Garzon, J. I. & Chacon, P. iMod: multipurpose normal mode analysis in internal coordinates. *Bioinformatics (Oxford, England)* **27**, 2843-2850, doi:10.1093/bioinformatics/btr497 (2011).
232 Lopez-Blanco, J. R., Aliaga, J. I., Quintana-Orti, E. S. & Chacon, P. iMODS: internal coordinates normal mode analysis server. *Nucleic acids research* **42**, W271-276, doi:10.1093/nar/gku339 (2014).
233 Damiano, J. S., Oliveira, V., Welsh, K. & Reed, J. C. Heterotypic interactions among NACHT domains: implications for regulation of innate immune responses. *The Biochemical journal* **381**, 213-219, doi:10.1042/bj20031506 (2004).
234 Man, S. M. *et al.* Inflammasome activation causes dual recruitment of NLRC4 and NLRP3 to the same macromolecular complex. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 7403-7408, doi:10.1073/pnas.1402911111 (2014).
235 Qu, Y. *et al.* NLRP3 recruitment by NLRC4 during Salmonella infection. *The Journal of experimental medicine* **213**, 877-885, doi:10.1084/jem.20132234 (2016).
236 Miao, E. A., Ernst, R. K., Dors, M., Mao, D. P. & Aderem, A. Pseudomonas aeruginosa activates caspase 1 through Ipaf. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 2562-2567, doi:10.1073/pnas.0712183105 (2008).

237 Man, S. M. *et al.* Salmonella infection induces recruitment of Caspase-8 to the inflammasome to modulate IL-1beta production. *J Immunol* **191**, 5239-5246, doi:10.4049/jimmunol.1301581 (2013).
238 Saleh, M. *et al.* Enhanced bacterial clearance and sepsis resistance in caspase-12-deficient mice. *Nature* **440**, 1064-1068, doi:10.1038/nature04656 (2006).
239 Vande Walle, L. *et al.* Does caspase-12 suppress inflammasome activation? *Nature* **534**, E1-4, doi:10.1038/nature17649 (2016).
240 Stehlik, C. & Dorfleutner, A. COPs and POPs: modulators of inflammasome activity. *J Immunol* **179**, 7993-7998 (2007).
241 Hoffman, H. M., Mueller, J. L., Broide, D. H., Wanderer, A. A. & Kolodner, R. D. Mutation of a new gene encoding a putative pyrin-like protein causes familial cold autoinflammatory syndrome and Muckle-Wells syndrome. *Nature genetics* **29**, 301-305, doi:10.1038/ng756 (2001).
242 Yang, H. *et al.* A mutant CHS3 protein with TIR-NB-LRR-LIM domains modulates growth, cell death and freezing tolerance in a temperature-dependent manner in Arabidopsis. *The Plant journal : for cell and molecular biology* **63**, 283-296, doi:10.1111/j.1365-313X.2010.04241.x (2010).
243 Mortimer, L., Moreau, F., MacDonald, J. A. & Chadee, K. NLRP3 inflammasome inhibition is disrupted in a group of auto-inflammatory disease CAPS mutations. *Nature immunology* **17**, 1176-1186, doi:10.1038/ni.3538 (2016).
244 Yang, J., Liu, Z. & Xiao, T. S. Post-translational regulation of inflammasomes. *Cellular and Molecular Immunology* **14**, 65-79 (2017).
245 Park, Y. H., Wood, G., Kastner, D. L. & Chae, J. J. Pyrin inflammasome activation and RhoA signaling in the autoinflammatory diseases FMF and HIDS. *Nature immunology* **17**, 914-921, doi:10.1038/ni.3457 (2016).
246 Spalinger, M. R. *et al.* NLRP3 tyrosine phosphorylation is controlled by protein tyrosine phosphatase PTPN22. *The Journal of clinical investigation* **126**, 1783-1800, doi:10.1172/jci83669 (2016).
247 He, Y., Zeng, M. Y., Yang, D., Motro, B. & Nunez, G. NEK7 is an essential mediator of NLRP3 activation downstream of potassium efflux. *Nature* **530**, 354-357, doi:10.1038/nature16959 (2016).
248 Schmid-Burgk, J. L. *et al.* A Genome-wide CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) Screen Identifies NEK7 as an Essential Component of NLRP3 Inflammasome Activation. *The Journal of biological chemistry* **291**, 103-109, doi:10.1074/jbc.C115.700492 (2016).
249 Shi, H. *et al.* NLRP3 activation and mitosis are mutually exclusive events coordinated by NEK7, a new inflammasome component. *Nature immunology* **17**, 250-258, doi:10.1038/ni.3333 (2016).
250 Matusiak, M. *et al.* Flagellin-induced NLRC4 phosphorylation primes the inflammasome for activation by NAIP5. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 1541-1546, doi:10.1073/pnas.1417945112 (2015).
251 Raghawan, A. K. *et al.* A Disease-associated Mutant of NLRC4 Shows Enhanced Interaction with SUG1 Leading to Constitutive FADD-dependent Caspase-8

Activation and Cell Death. *The Journal of biological chemistry* **292**, 1218-1230, doi:10.1074/jbc.M116.763979 (2017).

252 Proell, M., Riedl, S. J., Fritz, J. H., Rojas, A. M. & Schwarzenbacher, R. The Nod-like receptor (NLR) family: a tale of similarities and differences. *PloS one* **3**, e2119, doi:10.1371/journal.pone.0002119 (2008).

253 Baroja-Mazo, A. *et al.* The NLRP3 inflammasome is released as a particulate danger signal that amplifies the inflammatory response. *Nature immunology* **15**, 738-748, doi:10.1038/ni.2919 (2014).

254 Franklin, B. S. *et al.* The adaptor ASC has extracellular and 'prionoid' activities that propagate inflammation. *Nature immunology* **15**, 727-737, doi:10.1038/ni.2913 (2014).