INVESTIGATING THE MECHANISM OF ACTIVATION BY INFLAMMATORY SIGNAL MOLECULES S100A9 AND LIPOPOLYSACCHARIDES

by

HANNAH MURAWSKY

A THESIS

Presented to the Department of Biology
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

May 2021

An Abstract of the Thesis of

Hannah Murawsky for the degree of Bachelor of Science
in the Department of Biology to be taken June 2021

Title: Investigating the Mechanism of Activation by Inflammatory Signal Molecules S100A9 and Lipopolysaccharides

Approved: *Michael J. Harms, Ph.D.*
Primary Thesis Advisor

Inflammation is the most common cause of death worldwide. Mediated by the innate immune system, inflammation serves as one of the first lines of defense against infection and injury. However, when this natural and healthy response runs amok, it can cause and exacerbate many diseases. The innate immune receptor Toll-like Receptor 4 (TLR4), along with its cofactors, MD-2 and CD14, activates inflammation in response to external and internal danger signals. TLR4 induced inflammation is implicated in sepsis and many other inflammatory disease states. Despite the significance of this receptor complex for human health, much is yet to be understood regarding the biochemical mechanism by which it recognizes danger signals. In particular, the role of the cofactor CD14 is well understood for external danger signals, but its role in internal danger signal recognition is poorly understood. By introducing mutations to CD14 and measuring their effect on TLR4 induced inflammation, I have determined the importance of specific amino acids and regions of CD14 involved in activation by internal danger signals.

Acknowledgements

I would like to thank Professor Mike Harms and Ph.D. student Lauren Lehmann for their continued support and guidance during my time at the Harms Lab. I want to thank them for helping me grow not only as a student, but also as a budding scientist, and I want to extend my gratitude for their immense help in fully examining my thesis topic and the related subject matter.

I also wish to thank them for serving on my thesis committee and thank Professor Casey Shoop for serving on the committee and for his support and guidance as well. I also grateful for the support I received from the Clark Honors College, and for the privilege of having so many excellent professors who were willing to guide me through this strenuous but rewarding process.

I also wish to thank Professor Steve Vacchi. His constant support and guidance throughout my years in college have helped me overcome the difficulties of obtaining degrees in both music and biology, and I have taken much of his advice with me and will cherish it wherever I go.

Lastly, I wish to extend my sincerest gratitude to my family. To my Mom, for her unwavering support and love, and consistent openness to phone calls; to my Dad, for always pushing me to be the best person I can be and helping me through my hardest times, and to my brother, who has always been there for me with love, support, and a good laugh. I love you all, and words cannot express how much I am grateful for your faith in me and constant support.

Table of Contents

List of Figures

List of Tables

Introduction

Inflammation is a double-edged sword. When faced with injury or infection, inflammation serves to protect us. When we get sick with a flu or cold, common symptoms such as fever and chills are not due to the infection itself. Instead, this response is our body activating inflammation, which helps clear the virus from our bodies. When a tissue is inflamed, it recruits white blood cells. These white blood cells secrete highly reactive molecules that non-specifically damage both germs and our own cells. Our bodies sacrifice a few replaceable cells in an effort to kill the invading viruses. This healthy and natural response produced by our immune system allows us to handle various challenges. Without inflammation as an initial response, our bodies would have a difficult time staving off even the smallest of infections. However, inflammation becomes less beneficial the longer it occurs—while acute inflammation can prevent serious infection, chronic inflammation can have grave consequences for human health. Many diseases either cause or are marked by chronic inflammation (Pahwa et. al, 2020). Diabetes, cardiovascular disease, autoimmune diseases, Alzheimer's, and cancer are all chronic inflammatory diseases, and contribute to 3 in 5 deaths worldwide (Pahwa et. al, 2020). In each of these disease states, chronic inflammation leads to chronic tissue damage and thus, poor health outcomes. How our bodies maintain this delicate balance of clearing infections without inducing chronic inflammation has been a topic of intensive study for many years.

Inflammation is controlled by the innate immune system. The innate immune system is thought to have evolved as early as 500-600 million years ago (Buchmann, 2014). It uses the strategy of "molecular pattern recognition" to identify when

pathogens are present (Buchmann, 2014). Composed of a variety of different Pattern Recognition Receptors (PRRs), the innate immune system can recognize and respond to molecules that are unique to microbes (Buchmann, 2014). These Microbe-Associated Molecular Patterns (MAMPs) are thus chemical hallmarks of pathogens. Upon recognition, innate immune receptors trigger a non-specific immune response to create an inhospitable environment for the pathogen, such as inflammation and fever (Buchmann, 2014). It is also sometimes advantageous for human cells to activate a local inflammatory response in the absence of pathogens. For example, in wound healing inflammation occurs even in a sterile environment. To activate inflammation, immune cells produce molecules known as Damage Associated Molecular Patterns (DAMPs) that activate PRRs in a fashion analogous to MAMPs (Alberts, 2008).

Toll-like receptors

While there are many different classes of innate immune receptors, one group that is heavily studied and can recognize a wide variety of MAMPs is known as the Toll-like receptors (TLRs). TLRs first arose in multicellular invertebrates, such as cnidarians and sponges, but have expanded in number and specificity throughout evolutionary time (Buchmann, 2014). Humans have 10 TLR receptors, each specialized to recognize a specific MAMP (Christmas, 2010). For example, the heterodimer of TLR2 and TLR1 recognizes triacylated lipopeptides (found in bacteria), whereas the heterodimer of TLR7 and TLR8 recognizes ssRNA (found in viruses) (Uematsu & Akira, 2008). The homodimer of TLR4, the receptor complex which I am studying, is the canonical receptor for lipopolysaccharide (LPS), which is a cell-wall component of gram-negative bacteria (Molteni et. al, 2016).

TLR4 has important implications for human health—its response to LPS contributes to sepsis, which caused 1 in 5 deaths worldwide from 1990 – 2017. TLR4 induced inflammation can also be hijacked by cancerous cells, which produce DAMPs that activate TLR4 to promote blood vessel growth in growing tumors. The discovery of TLR4 in 1998 by the Beutler lab earned the Nobel Prize in Physiology or Medicine in 2011 (Rudd et. al, 2020). TLR4 is present on the cell membranes of a variety of cell types, including blood-cell progenitors and immune cells (Molteni. et. al, 2016). TLR4 acts as a complex with two accessory proteins, MD-2 and CD14; together, they recognize and respond to both MAMPs and DAMPs.

LPS is a molecule made exclusively by gram-negative bacteria. It contains three main parts: the O antigen, core oligosaccharides, and Lipid A tails (Figure 1), with the lipid A portion being primarily responsible for pathogenicity. The size and structure of LPS varies greatly between bacterial species, with modifications occurring to the O antigen, core oligosaccharides, and lipid A tails. While various forms of LPS exist, the two main forms used to test activation of TLR4 are LPS-S and LPS-R derived from *S. enterica* and *E. coli* bacteria, respectively. LPS-R lacks the O antigen, whereas LPS-S contains the O antigen; LPS-R is the form of LPS used in my experiments.

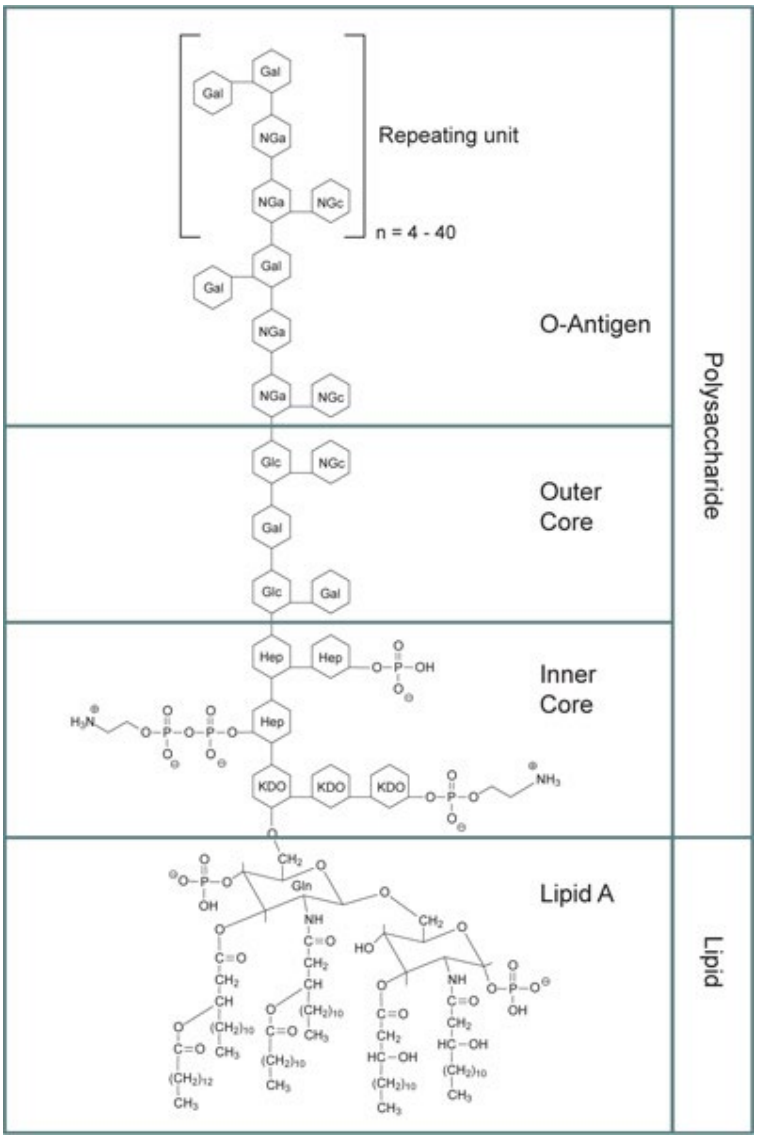


Figure 1: Subsections of LPS, labeled. Obtained from Sigma Aldrich

TLR4 is also activated by a variety of DAMPs. One important TLR4 DAMP is S100A9. This protein is 14 kDa in size, acts as a calcium binding protein, and frequently occurs as a homodimer (Markowitz & Carson, 2013). It is much larger than LPS and does not contain any analogous structure to the long hydrophobic tails of the lipid-A portion of LPS. Very little is understood regarding how S100A9 activates TLR4. By studying S100A9 and its interaction with TLR4 and its other cofactors, we

can gain insight into not just how this specific system works, but also learn more about how DAMPs and TLRs may interact.

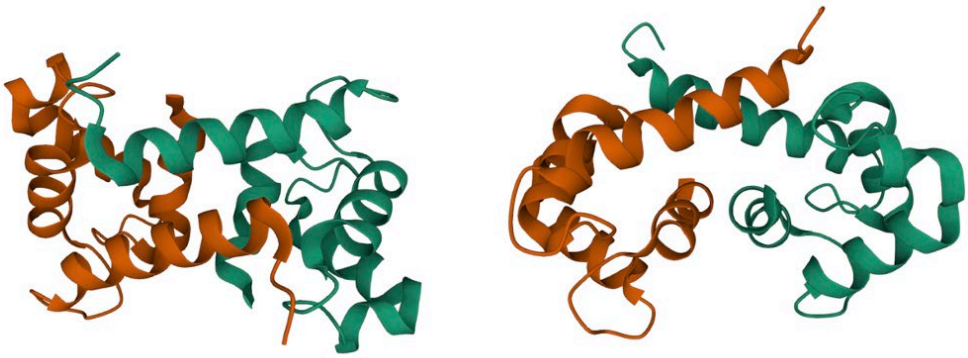


Figure 2: Two views of the crystal structure of human S100A9 Homodimer.

Each individual S100A9 molecule labeled in green or orange. Obtained from RSCB PDB. PDB ID: 1IRJ

Mechanism of TLR4 activation

TLR4 is composed of an intra- and extracellular domain, connected by a transmembrane helix (Molteni et. al, 2016). The extracellular domain of TLR4, along with MD-2, recognizes both DAMPs and MAMPs (Molteni et. al, 2016). In the case of LPS, the hydrophobic lipid tails of LPS bind inside a pocket of MD-2, while its polar charged head and O antigen interact with the outside of MD-2 and parts of TLR4 (Park et. al, 2009). A crystal structure of LPS bound to the TLR4/MD-2 complex has been solved, which shows the direct interaction of TLR4/MD-2 with LPS in detail (Park et. al, 2009). This binding event promotes dimerization with a second molecule of TLR4/MD-2 bound to LPS (Molteni et. al, 2016). Dimerization induces a conformational change in the intracellular domains of TLR4 and facilitates the recruitment of adaptor proteins, initiating a molecular cascade that ultimately causes

inflammation through the production of NF-kB, a promoter of many genes involved in inflammation, and various pro-inflammatory cytokines (Molteni et. al, 2016).

The cofactor CD14 is also essential for TLR4's ability to recognize and respond to LPS. CD14 is anchored to the cell membrane on one side and has an LPS binding pocket on the other side. CD14 shuttles LPS to the TLR4/MD-2 complex, where it is passes from CD14 to its binding location on the TLR4/MD-2 complex. CD14 is necessary for the recognition of LPS by the TLR4 complex for two main reasons—it allows TLR4 to recognize LPS at lower concentrations promoting a swift immune response to microbes and protects the hydrophobic tails of LPS from interacting with other hydrophobic molecules in the cell (Granucci & Zanoni, 2013). Previous research has shown that LPS binds to CD14 in the N-terminal pocket (Juan et. al, 1995, He et. al, 2016). While both binding and activation assays have shown this portion of CD14 to contain the LPS binding site, and specific residues have been found to play important roles, there is no crystal structure of LPS bound to CD14.

While a great deal of research has been conducted regarding the interaction of LPS with the TLR4 complex, the exact mechanism of activation of TLR4 via S100A9 is largely unknown. *In vitro* activity studies have shown that both CD14 and MD-2 are necessary for activation via S100A9. Intriguingly S100A9 and CD14 co-localize on cell membranes *in vivo* and have even been shown to directly bind *in vitro* (He et. al, 2009). Despite this clear evidence for the existence of an interaction between S100A9 and CD14, little is known about the nature of the interaction. Thus, the main questions driving my research have been: how does S100A9 interact with CD14; what are the

specific sites and regions involved; and, is this interaction different than the LPS/CD14 interaction (Figure 3)?

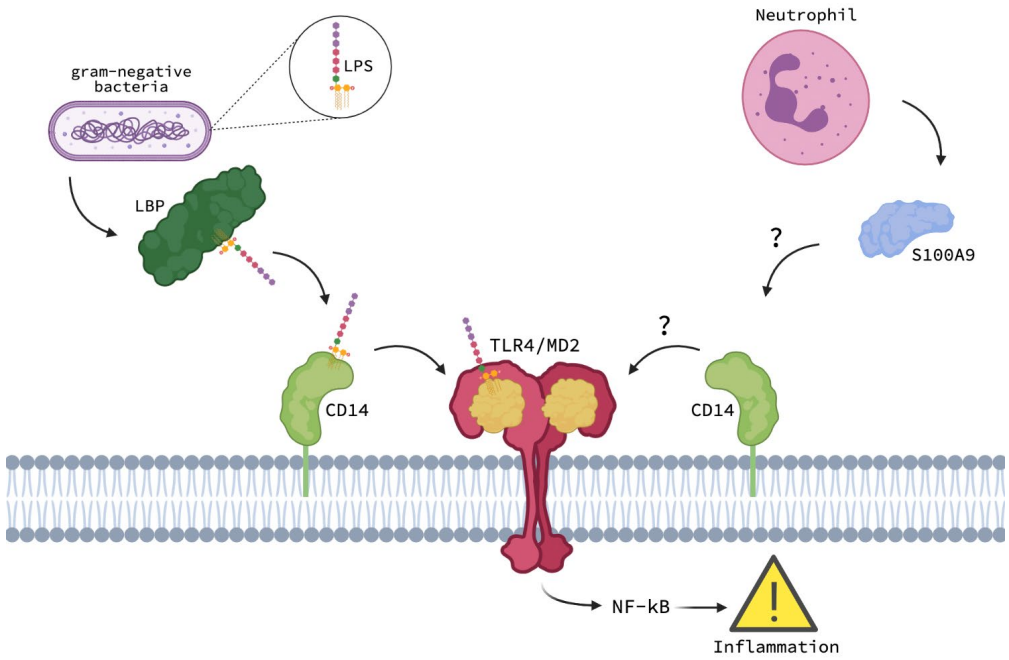


Figure 3: Illustration of research question and system.

Figure created using BioRender.

Investigating this question will provide valuable insight into exactly how S100A9 interacts with the TLR4 complex. Since elevated levels of S100A9 are associated with a variety of conditions, including but not limited to arthritis, diabetes, septic shock, Alzheimer's, cancer, and even serious COVID-19 infection outcomes, learning more about where S100A9 interacts with CD14 can help inform the creation of possible drug targets (Pahwa et. al, 2020; Wang et. al, 2013; Källberg et. al, 2012; Gong et. al, 2019; Averill et. al, 2012, Barnes & Karin, 1997). Previous research has attempted to make anti-S100 family drugs, but such drugs failed during clinical trials due to lack of specificity (Pelletier, 2018).

Understanding how S100A9 directly interacts with CD14, could provide insights to a possible drug target—one that allows for activation of TLR4 via LPS to maintain immune function, but prevents activation of TLR4 via S100A9 to help lower inflammation.

Methods/Experimental Approach

To attempt to answer my research question, a truncation mutant of CD14 informed by existing literature was generated to determine the overall region required for CD14 function. I then used site-directed mutagenesis to introduce single amino acid substitutions within that region (Table 1). Like all proteins, CD14 is a long chain of specific amino acids connected end-to-end. Different amino acids have different physiochemical properties, thus allowing proteins to interact specifically with other molecules that have matched physiochemical properties. In site-directed mutagenesis, I swapped out one amino acid at a specific point in the chain with another amino acid with different properties. This allowed me to test whether the amino acid in the protein is important for a given function. The creation of these mutants was informed by three main strategies: previous literature, conservation analysis, and structural analysis.

Mutation[1]	*Cofactor*	*Reasoning*	*Literature*
D29K	CD14	Mutation found to reduce LPS binding by >50%.	Cunningham et al, 2000
R33E	CD14	Mutation found to reduce LPS binding by >50%.	Cunningham et al, 2000
E56K	CD14	Mutation found to reduce LPS binding by >50%.	Cunningham et al, 2000
W45A	CD14	Encircles the rim with its side chain overlaying the entrance to the LPS binding pocket.	Kelly et al., 2013
F49A	CD14	Encircles the rim with its side chain overlaying the entrance to the LPS binding pocket.	Kelly et al., 2013
V52A	CD14	Encircles the rim with its side chain overlaying the entrance to the LPS binding pocket.	Kelly et al., 2013
F69A	CD14	Encircles the rim with its side chain overlaying the entrance to the LPS binding pocket.	Kelly et al., 2013

[1] A list of the twenty amino acids, along with their abbreviated names and chemical properties can be found in the glossary, Table 1a.

Y82A	CD14	Encircles the rim with its side chain overlaying the entrance to the LPS binding pocket.	Kelly et al., 2013
L89A	CD14	Encircles the rim with its side chain overlaying the entrance to the LPS binding pocket.	Kelly et al., 2013
N-terminal 152 AA	CD14	Truncation of CD14 only containing N-terminal 152 amino acids has been shown to retain activation ability	Juan et. al, 1995

Table 1: List of mutations with reasoning and referenced literature.

Previous Literature

Previous research has tested the interaction of CD14 with LPS; some of these studies include direct binding assays or activation assays and utilize the introduction of mutations or blocking antibodies with known binding sites. After looking at previous literature that introduced mutations to CD14, I chose to select a few mutants from these studies. Research conducted by Juan et. al in 1995 indicated that a truncation of CD14 than only contained the N-terminal 152 amino acids was still able to activate TLR4 via LPS, however, this mutant has not been tested with S100A9. Additionally, research conducted by Cunningham et. al in 2000 demonstrated that three mutants, D29K, R33E, and E56K all decreased binding of LPS with CD14 by greater than 50%. Like the truncation of CD14, these mutants were also not tested with S100A9. By treating these mutants with both S100A9 and LPS, I hoped to gain some insight into whether the two ligands may bind CD14 in a similar location.

Structural Analysis

Some of the mutants were created based on the known crystal structure of CD14. It is known that LPS binds to CD14 in the N-terminal hydrophobic pocket. By looking at the structure of individual amino acid residues, and analyzing their position

in the pocket, I chose to mutate residues in the rim of the pocket whose side chains stuck into the entrance (Figure 4).

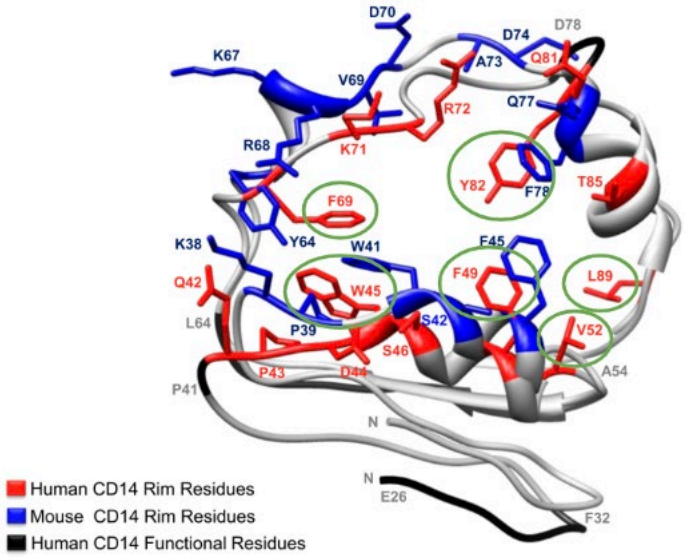


Figure 4: Crystal Structure of human and mouse CD14 with mutations labeled.

Human rim residues are colored in red, and mouse rim residues are colored in blue. Mutations are circled in green and are occurring on human CD14. Original image obtained from Kelly et. al, 2003, with annotations of selected residues added by author.

These mutations were: W45A, F49A, V52A, F69A, Y82A, and L89A. By seeing how these mutants affected the activation of TLR4 via S100A9, I hoped to gain additional insight into whether S100A9 and LPS may share the N-terminal binding pocket.

Conservation Analysis

To further inform mutant generation, I also performed a conservation analysis on the residues of CD14 (Figure 5). Residues that stayed as the same amino acid across evolutionary time were considered highly conserved, whereas residues that did not stay as the same amino acid across evolutionary time were considered not conserved. Then, I

looked at those results to see if any of the residues I chose to mutate on CD14 were highly conserved.

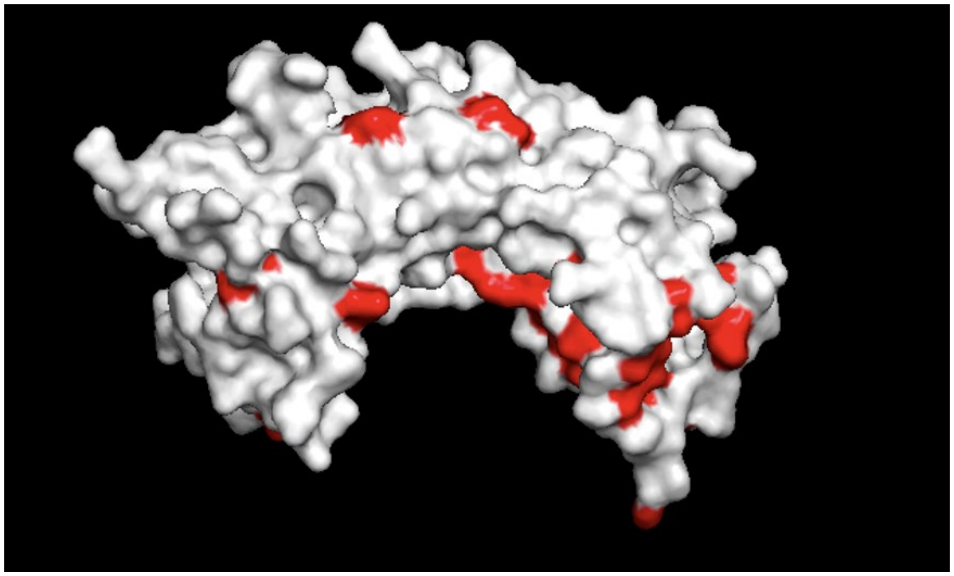


Figure 5: Crystal structure of human CD14 as surface representation.

Conserved residues labeled in red. The left side of the molecule in the above image is the N-terminus. Image generated with PyMol.

Individual amino acids that are more highly conserved across evolutionary time are likely facing selective pressure; the amino acid that is highly conserved may participate in an important interaction (such as binding with a ligand) or have an important structural role. If the mutants I chose were important for binding with either LPS or S100A9, they also may be highly conserved. Ultimately, I found that of the residues I had chosen to mutate, only residues D29 and L89 were highly conserved. However, the conservation analysis I performed only indicated that the specific amino acid was conserved, not whether the chemical properties of the amino acid (which can also be satisfied by other chemically similar residues) was conserved.

Activity Assay

To test TLR4 activation I performed a widely used transfection-based activity assay. The mutants of CD14 I created were transfected into Human Embryonic Kidney (HEK) 293T cells, along with the wild-type TLR4 and MD-2, pcDNA, renilla, and ELAM-luciferase. The pcDNA plasmid acted as "junk" DNA to improve transfection efficiency, whereas the renilla and ELAM-luciferase plasmid acted as luminescent reporters. These reporters were used to measure transfection efficiency and TLR4 activation, respectively.

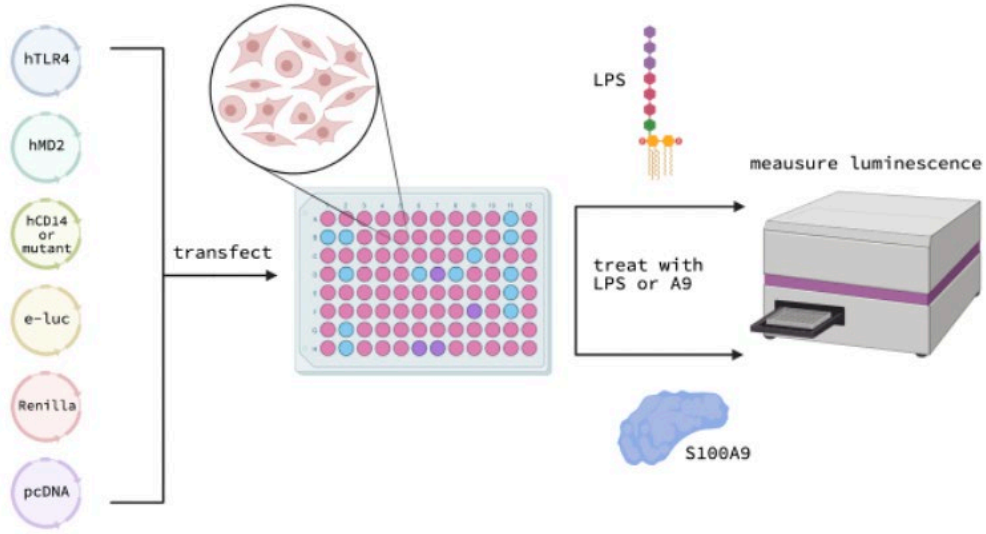


Figure 6: Illustration of transfection, treatment, and activation assay.

Image created with BioRender.

The ELAM-luciferase plasmid contained an NF-kB promoter for the firefly luciferase enzyme gene. When NF-kB was not present, the firefly luciferase gene was not transcribed, and firefly luciferase was not created. However, when TLR4 was activated, NF-kB was generated, which promoted expression of the luciferase firefly enzyme. After 3-4 hours of expression, I added the luciferase substrate. The firefly

luciferase enzyme cleaved this protein, causing it to luminesce (Figure 7). I then measured the intensity of this emitted light to determine the degree of activation; cells that had higher levels of activation created more NF-kB, resulting in the production of more firefly luciferase enzyme and a greater luminescence intensity.

Unlike the firefly luciferase, all cells that were successfully transfected expressed renilla luciferase from the renilla plasmid. Then, when I analyzed the cells, I also added the renilla substrate. This substrate was cleaved by the renilla luciferase enzyme and luminesced (Figure 7). I was then able to measure the intensity of the emitted light to determine how many cells were present in the assay. This allowed me to normalize the output of my firefly luciferase reporter.

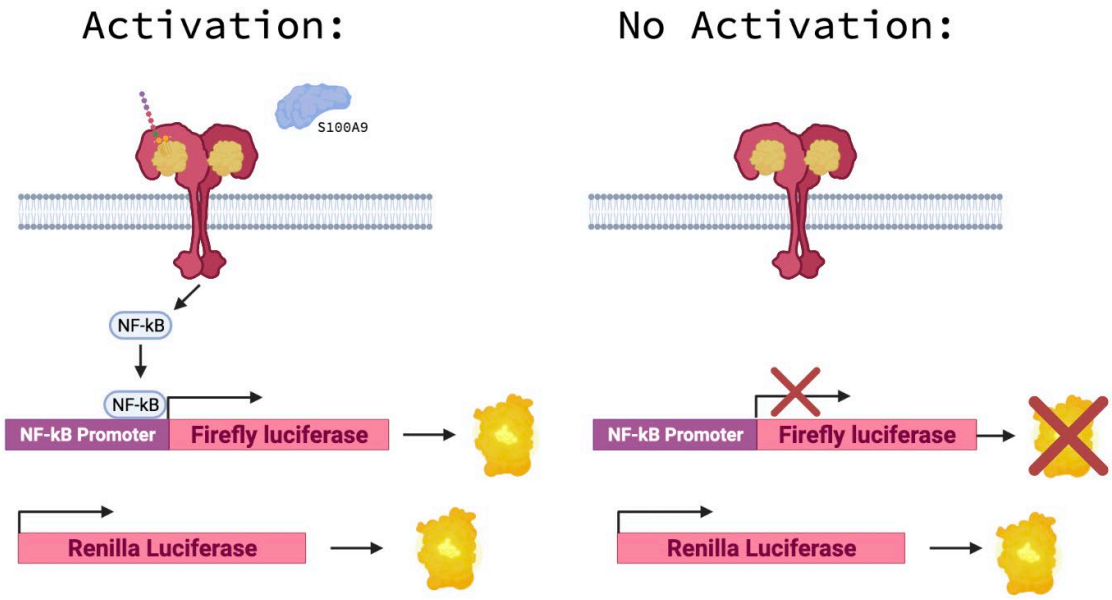


Figure 7: Firefly luciferase and Renilla luciferase reporter assay

Illustration of Firefly luciferase and Renilla luciferase reporter assay. Image created with BioRender.

To determine how each mutation to CD14 affects the ability of LPS and S100A9 to activate TLR4, I added a variety of treatments to the transfected cells. The 200

ng/mL LPS and 2uM S100A9 treatments were used to measure how the mutants affect activation of TLR4 at saturating concentrations of either ligand. Treatment of 200 ng/mL LPS and Polymyxin B (PB) acted as a negative control; PB binds to and sequesters LPS, preventing activation of TLR4. The treatment of 2uM S100A9 and PB acted as a more accurate measure of how the mutants affect the activation of TLR4 via S100A9. Because S100A9 is expressed and purified from *E. coli*, which contains LPS, contaminating LPS is likely present. The addition of PB to the S100A9 treatment removes any LPS that may be present in the sample. This allowed me to see the effect of the mutant on the activation of TLR4 via S100A9 only. The last treatment type, PBS (phosphate buffered saline) Buffer, acted as a negative control. (Table 2).

Treatment	*Description and Reasoning*
200 ng/mL LPS	Saturating concentration of LPS; test effect of CD14 mutants on TLR4 activation via LPS.
2 uM S100A9	Saturating concentration of S100A9; test effect of CD14 mutants on TLR4 activation via S100A9.
200 ng/mL LPS + PB	Negative control; PB sequesters LPS, preventing TLR4 activation.
2 uM hS100A9 + PB	Purification of S100A9 sample; removes any LPS contaminant reducing background noise generated by possible activation of TLR4 via LPS.
PBS	Negative control; buffer.

Table 2: Description and reasoning of various treatments.

Transfection of cells with no CD14 also acted as an additional negative control, as both LPS and S100A9 are unable to cause activation of TLR4 in its absence.

Following 3-4 hours of treatment, I removed the treatment mix, and analyzed the cells for activity using the renilla luciferase and firefly luciferase reporters as previously described.

Data Analysis

I measured the normalized activation intensity by using the equations presented in Equations 1. I calculated the normalized luminescence by dividing the raw firefly luminescence intensity by the raw renilla luminescence intensity. I then subtracted a blank (the data for No CD14 + LPS, which was a negative control) from the normalized luminescence to get the corrected luminescence value for each well. Next, I normalized this value by dividing it by the average corrected luminescence of the WT + LPS or WT + A9 + PB treatment. Finally, to get the normalized activity for each mutant and treatment, I averaged three technical replicates for each mutant and treatment (Equations 1).

$$normalized\ luminescence = \frac{raw\ firefly}{raw\ renilla}$$

$$blank\ luminescence = avg.(normalized\ luminescence_{No\ CD14+LPS})$$

$$luminescence = normalized\ luminescence - blank\ luminescence$$

$$normalized\ activity_{mutant+LPS} = avg.\left(\frac{luminescence_{mutant+LPS}}{avg.\ luminescence_{WT+LPS}}\right)$$

$$normalized\ activity_{mutant+A9+PB} = avg.\left(\frac{luminescence_{mutant+A9+PB}}{avg.\ luminescence_{WT+A9+PB}}\right)$$

Equations 1: Equations used to determine activity for various mutants and treatments.

I also averaged three biological replicates of each mutant with each treatment and calculated the standard error across biological replicates. I calculated the P-values using a one sample, two-tailed t-test comparing each value to wild-type, which was normalized to 1. Any p-values < 0.05 were considered statistically significant.

Results

Amino acids 1-152 of CD14 are sufficient for activation of TLR4 via LPS and S100A9

Previous research has shown that only the N-terminal 152 amino acids of CD14 are sufficient for activation of TLR4 via LPS, but this truncation mutation has not been tested with S100A9 (Juan et. al, 1995). To test whether S100A9 binding with CD14 also occurs in the N-terminal 152 amino acids, HEK293T cells were transfected with the N-terminal 152 amino acid truncation of CD14 along with WT TLR4 and MD, and luminescent reporter plasmids. The transfected cells were treated with various treatments including S100A9 (Table 2). As expected, we observed that both LPS and S100A9 were able to activate TLR4 with CD14, but not without (Figure 8, dark grey and light grey bars). Addition of truncated protein was able to restore activity for both LPS and S100A9 (Figure 8, medium grey bars). These preliminary results indicate that S100A9 requires only the N-terminal 152 amino acids of CD14 to activate TLR4 (Figure 8). This also suggests that S100A9 and LPS share a binding region on CD14.

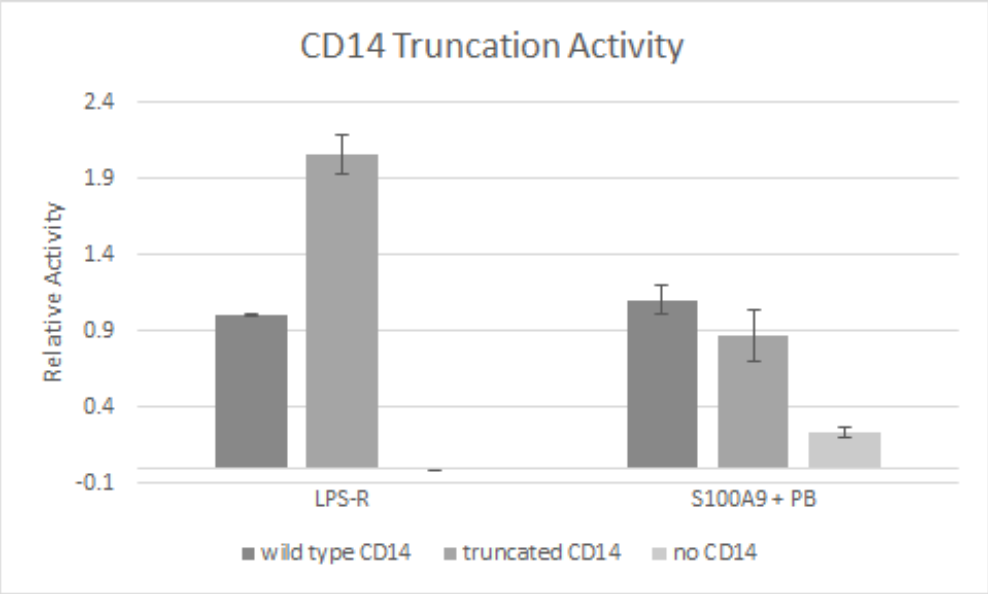


Figure 8: Relative Activity of TLR4 with truncated CD14 when treated with S100A9 or LPS.

Observed relative activity of TLRR4 with N-terminal 152 amino acid truncation of CD14 when treated with 2uM S100A9 + PB and 200ng/mL LPS-R. Bars show mean fold activation in the assay relative to the 200 ng/mL LPS control. Error bars are the standard error of three technical replicates. Treatments are shown below each set of bars. Dark, medium, and light grey bars indicate the WT, truncated, and no CD14, respectively. Data collected and figured created by Lauren Lehmann.

Mutations D29K and W45A to CD14 decrease activation ability of LPS and S100A9

HEK293T cells were also transfected with a series of single amino acid substitution CD14 mutants, and measured for activity following treatment (Table 1, Table 2). The two mutants with the most striking changes to activation were D29K and W45A. Cells transfected with these CD14 mutations showed a statistically significant decrease in TLR4 activity upon treatment with either S100A9 or LPS (Figures 9, 10). These preliminary results show that S100A9 and LPS require the same sites on CD14 to

activate TLR4 and suggest that S100A9 and LPS may share these amino acids for binding.

Mutation F49A to CD14 decreases activation of TLR4 via S100A9 by 50% but has no effect on LPS activation ability

Unlike mutations D29K and W45A, we observed one mutation that showed varying effects of TLR4 activation dependent on treatment. Cells transfected with the F49A mutation to CD14 showed a statistically significant decrease of around 50% in TLR4 activation when treated with S100A9 and PB, but no effect when treated with LPS (Figures 9, 10). These results suggest that while LPS and S100A9 may share some residues for binding, there still may be some slight variations in all the residues important for binding each ligand.

Select mutations to CD14 do not affect the ability of LPS or S100A9 to activate TLR4

HEK293T cells were also transfected with the following mutations: R33E, E56K, V52A, F69A, Y82A, and L89A. These mutations resulted in no significant change in activation when treated with either ligand when compared to WT (Figures 9, 10). Notably, mutations R33E and E56K, which were shown in previous research to decrease LPS binding by 50%, showed no effect on activation of TLR4 via either ligand (Figure 9). These results further suggest that LPS and S100A9 may share a similar binding location, or that while these residues may participate in binding both ligands, they are not necessary for the interaction.

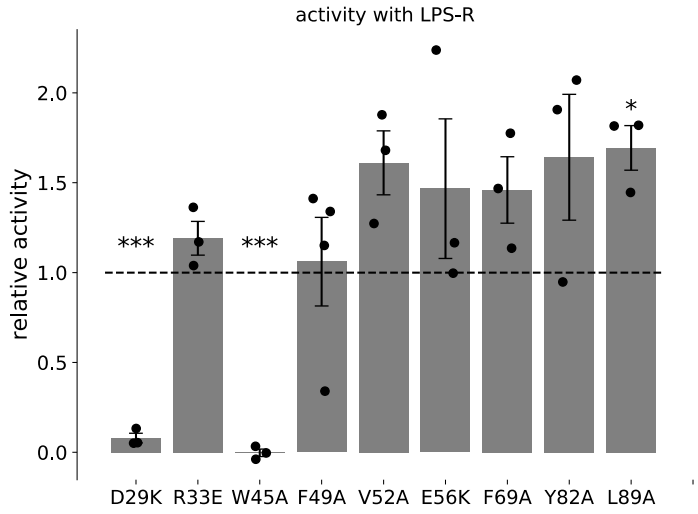


Figure 9: Relative activity of TLR4 with various CD14 mutants when treated with 200 ng/mL LPS-R

Observed relative activity of TLR4 with CD14 mutants D29K, R33E, W45A, F49A, V52A, E56K, F69A, Y82A, and L89A when treated with 200 ng/mL LPS-R. *** indicates a p value < 0.001. Bars show mean fold activation in the assay relative to the 200 ng/mL LPS control. Error bars are the standard error of three biological replicates.

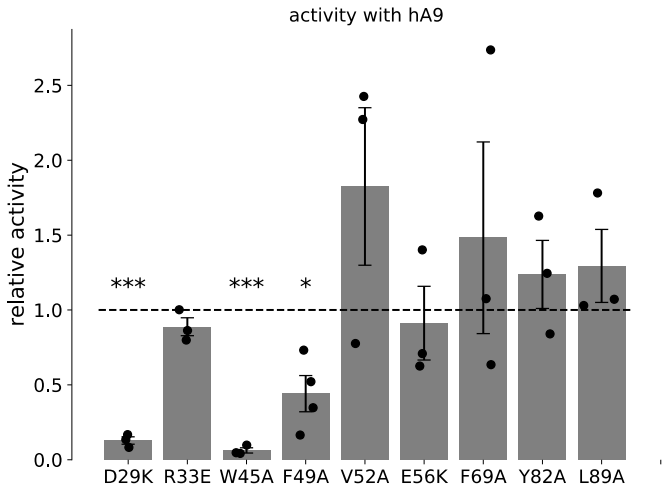


Figure 10: Relative activity of TLR4 with various CD14 mutants when treated with 2uM S100A9 + PB

Observed relative activity of TLR4 with CD14 mutants D29K, R33E, W45A, F49A, V52A, E56K, F69A, Y82A, and L89A when treated with 2 uM S100A9 + PB. *** indicates a p value < 0.001, * indicates a p value < 0.05. Bars show mean fold activation in the assay relative to the 2uM S100A9 + PB control. Error bars are the standard error of three biological replicates.

Discussion

Treatment of HEK293T cells containing the CD14 truncation with S100A9 or LPS showed that activity of TLR4 via both ligands was retained. This data suggests that both ligands only require the N-terminal 152 amino acids for activation, and that within this region lies their binding site(s). These results are consistent with previous data regarding the truncation with LPS treatment and align with our understanding of the role of the C-terminus. In full-length versions of CD14, the C-terminus acts as an anchoring point between CD14 and the cell membrane; due to this, only the N-terminal amino acids would be available for ligand binding and transfer to TLR4/MD-2.

Of the nine CD14 point mutants tested, two mutants had striking effects on activation. Both D29K and W45A showed a significant decrease in activation in the presence of both mutants. Based on the crystal structure of CD14, blocking antibody tests, and MD (molecular dynamics) simulations regarding its interaction with LPS and the TLR4 complex, it is likely that the residue D29 is involved in the interaction between CD14 and MD-2, rather than as a binding partner for S100A9 or LPS (Juan et. Al, 1995, Kim et. Al, 2005). Because this assay measures activation and not direct ligand binding, the results collected do not inform us of whether the mutation affected binding directly or affected activation through some other means. Due to this, the results collected for the D29K mutation could indicate that it interfered with the ability of CD14 to bind with TLR4/MD-2, rather than disrupting binding with both ligands.

The mutation W45A also seemed to affect the activation of TLR4/MD-2 via LPS and S100A9. Based on our understanding of the interaction between CD14 and the TLR4 complex, it is unlikely that the mutation disrupts this interaction; however, there

are other possible explanations that could account for these results. W45A could disrupt the binding of S100A9 and LPS to CD14; this would mean that W45 was a residue necessary for both interactions. Another possible explanation is that the mutation could have affected expression of CD14. When genes are transcribed from DNA to RNA and then translated from RNA to proteins, the process can occasionally fail. Certain sequences and chemical properties can alter this efficiency, resulting in very little protein production, despite ample gene expression. It is possible that the mutation W45A may have altered the chemical properties of the protein that makes it more difficult to translate. This would cause lower levels of CD14 expression, and therefore lower levels of activation in the presence of either ligand.

Unlike D29K and W45A, F49A seemed to affect LPS and A9 differently. This could indicate a few possibilities. If D29K and W45A are not within the binding site of both S100A9 and LPS, it is possible that F49A only lies within the binding site of LPS, and that S100A9 occupies a different binding site. If D29K and/or W45A are found to be within the binding sites of both molecules, it could instead indicate that while S100A9 and LPS may share the same overall binding site, there may be differences between specific residues involved in the interaction between the two ligands.

Future Directions

One main limitation of this experiment is that it only utilizes a TLR4 activity assay. Because it does not directly test binding between CD14 and ligand, interpreting the results to inform binding interactions can be somewhat limited. As seen with D29K, it is possible that some mutations may affect the ability of CD14 to interact with MD-2/TLR4, and therefore lower activation, but not disrupt binding. To directly investigate

how these mutants affect binding, a coimmunoprecipitation experiment could be performed. By conducting a pull-down assay with various CD14 mutants and S100A9 or LPS, we could gain some insight into whether these mutations directly disrupt binding. Furthermore, we could also perform the experiment with wild type CD14 in the presence of LPS, rather than a mutant CD14, and titrate in S100A9. This would help elucidate whether S100A9 and LPS compete for a binding site on CD14. The use of antibodies with known binding locations on CD14 could also be used to determine whether S100A9 and LPS share a binding site, and *in vitro* measurements of direct binding could be collected with SPR(surface plasmon resonance) or BLI (biolayer interferometry). Another important test that could be conducted would be to test for CD14 expression levels using a western blot. Should the explanation regarding W45A and its effect on protein translation be true, measuring the expression levels of various CD14 mutants could answer whether the mutation affects protein expression, rather than binding of either ligand.

Overview

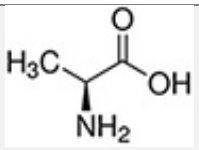
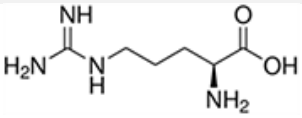
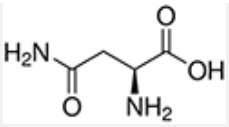
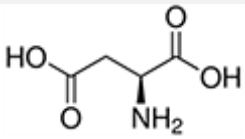
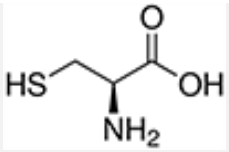
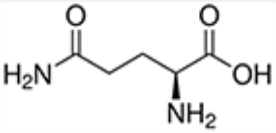
My work suggests that LPS and S100A9 may use essentially the same binding pocket. This is surprising given their very different structures. However, given these differences, and because one mutation to CD14 (F49A) lowered activation of TLR4 via S100A9 and had no effect on LPS, I speculate that they may still have slightly different binding mechanisms. Since these experiments only demonstrated an effect on activation and not on direct binding, it is possible that there may be differences in the binding of CD14 to LPS and S100A9. It is also possible that binding does not influence the activity

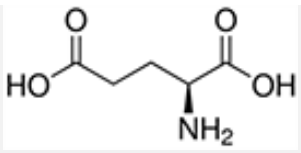
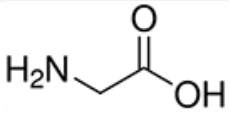
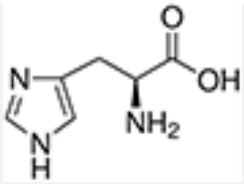
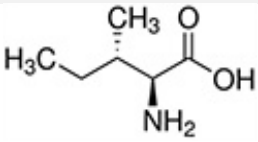
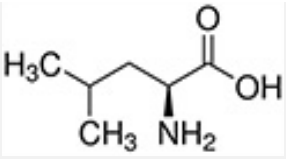
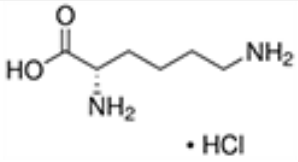
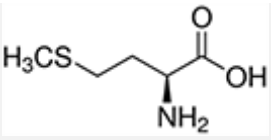
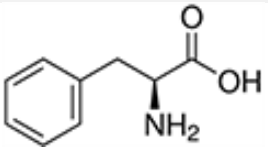
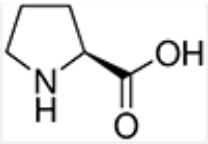
of CD14, and these mutants only affect activity. Further experimentation could be performed to help narrow down these possibilities.

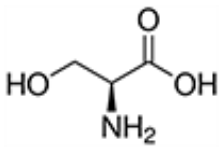
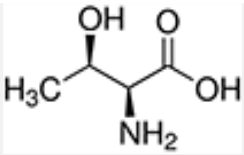
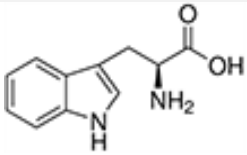
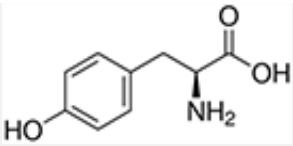
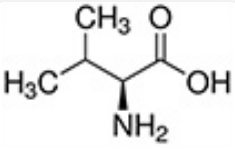
These findings have implications for future drug development. Without a differing binding site, creating a drug that inhibits the binding of one ligand without affecting the binding of the other will be difficult. Because of this, drug design may need to focus up- or downstream of CD14 to find a suitable target that involves one ligand and not the other. In addition to informing drug creation, these findings also have implications for our understanding of other TLRs; since other TLRs also recognize various MAMPs and DAMPs in an analogous fashion to LPS and S100A9, it is possible that they too have their ligands sharing a similar binding site. Experiments like the ones performed in this research could be applied to these systems and help inform us of their similarities and differences. Additionally, because of these findings, drug designers attempting to target these receptors and their ligands may also need to change their strategies. However, more research still needs to be conducted to improve our understanding of how these complex systems function.

Glossary

Amino acid/residue: Organic compounds that are the base unit of proteins. They contain amine and carboxyl functional groups, along with a side chain that is specific to each amino acid. There are twenty different naturally occurring amino acids, which have the following names, properties, and chemical structures (Sigma).

Full Name	*1-letter abbreviation*	*Chemical properties*	*Chemical structure*
Alanine	A	Hydrophobic, aliphatic	
Arginine	R	Charged, basic	
Asparagine	N	Polar, neutral	
Aspartic acid	D	Charged, acidic	
Cysteine	C	Polar, neutral	
Glutamine	Q	Polar, neutral	

Glutamic acid	E	Charged, acidic	
Glycine	G	Unique, no side chain	
Histidine	H	Charged, basic	
Isoleucine	I	Hydrophobic, aliphatic	
Leucine	L	Hydrophobic, aliphatic	
Lysine	K	Charged, basic	
Methionine	M	Hydrophobic, aliphatic	
Phenylalanine	F	Hydrophobic, aromatic	
Proline	P	Unique	

Serine	S	Polar, neutral	
Threonine	T	Polar, neutral	
Tryptophan	W	Hydrophobic, aromatic	
Tyrosine	Y	Hydrophobic, aromatic	
Valine	V	Hydrophobic, aliphatic	

Assay: A procedure for an experiment.

BLI: Biolayer interferometry. A way to measure molecular interactions.

Buffer: An aqueous solution, in this case used as a negative control.

Canonical receptor: Common receptor used in the biological system.

CD14: A human protein that is produced by macrophages, and functions as part of the innate immune system as a cofactor of TLR4. It can exist in a membrane bound or soluble form and is known to bind with LPS and deliver it to TLR4. It is also known to be necessary for the activation of TLR4 by S100A9, but the mechanism of this interaction is unclear.

Cnidarians: A phylum (evolutionary group) of animals that are mostly marine. Includes jellyfish, corals, and sea anemones, among others.

Co-immunoprecipitation: A way to identify protein-protein interactions.

Complex: A group of molecules that work together. For example, the TLR4 complex involves TLR4, MD-2, and CD14—all of which work together to recognize danger signals.

Conserved: Maintained across evolutionary time.

Crystal structure: A representation of the location of atoms. Shows the structure of a molecule in 3-dimensional space.

Cytokines: Small proteins released by the immune system that trigger inflammation.

Damage-associated molecular pattern (DAMP): Biomolecules of host origin that can initiate a noninfectious/sterile inflammatory response. In this case, S100A9 is a DAMP.

Dimerization: The process in which two molecular subunits join to form one complex.

ELAM-luciferase: The plasmid that contains the firefly luciferase gene that is promoted by NF-KB.

Endogenous: A substance that originates from within the organism. In this case, S100A9 is endogenous.

Exogenous: A substance that originates from something outside the organism. In this case, LPS is exogenous.

Gram-negative bacteria: A group of bacteria that do not have peptidoglycan (a molecule) in their cell membrane.

Heterodimer: Two different molecules that come together to form a complex. For example, the combination of TLR1 and TLR2 is a heterodimer that recognizes triacylated lipopeptides.

Homodimer: Two of the same molecules that come together. For example, S100A9 often exists as two of S100A9 molecules bound together, and two TLR4 molecules come together to form the TLR4 complex.

Hydrophobic: The quality of repelling or failing to mix with water. Molecules that are hydrophobic are often nonpolar.

In vitro: Latin for in a test tube. Is used to refer to experiments that do not occur in living things. For example, the transfection assay I performed is an *in vitro* assay.

In vivo: Latin for in a living organism. Is used to refer to experiments that occur in living things. For example, experiments performed on mice are considered to be *in vivo.*

kDa: Kilodaltons, a unit of size for molecules.

Ligand: An endogenous or exogenous molecule that binds with another biomolecule. In this case, LPS and S100A9 are ligands for TLR4, MD-2, and CD14.

Lipid A tails: A chemical component of LPS. Is what binds to the hydrophobic pockets of CD14 and MD-2.

Lipopolysaccharide (LPS): A major component of the cell wall of Gram-negative bacteria. It contains an O antigen, core oligosaccharide, and lipid tails. It is known to activate TLR4, and its mechanism of binding with TLR4, MD-2, and CD14 is understood.

Macrophages: A type of immune cell, specifically a type of white blood cell.

MD-2: Also known as lymphocyte antigen 96. It is a cofactor of TLR4, and binds to LPS. It is known to be necessary for the activation of TLR4 by S100A9, but the precise mechanism is unclear.

Microbial-associated molecular pattern (MAMP): A biomolecule of microbial origin that is recognized by TLRs and can initiate an inflammatory response. In this case, LPS is a MAMP.

Mutagenesis: The process by which genetic information is altered.

NF-KB: A protein that controls the transcription of specific genes. It becomes activated following TLR4 activation. When this protein is activated, it causes the transcription of genes that contribute to inflammation.

O antigen: A chemical portion of some variants of the LPS molecule.

Oligosaccharides: A large sugar molecule made up of many repeating small sugar molecules. Is a chemical portion of LPS.

Pathogenicity: The state of being pathogenic; having the ability to produce disease.

Pattern recognition receptor (PRR): Host sensors that recognize molecules from specific pathogens. In this case, TLR4 is a PRR.

pcDNA: A "junk" plasmid used to improve transfection efficiency.

Plasmid: Small, extrachromosomal DNA that can be replicated independently from cellular DNA.

Progenitors: Early descendants of stem cells that have some classifications of cell type but have not fully specialized yet.

P-value: The probability of the observed event occurring, given that the null hypothesis is true. P-values < 0.05 indicate that the results are significant, meaning that they likely occurred because the mutation had an effect and not because of chance.

Renilla: The plasmid containing the renilla luciferase gene.

S100A9: A human protein, elevated in many inflammatory diseases, and known to activate TLR4. The precise mechanism of activation is unknown.

Saturating concentration: The concentration of a ligand when the receptor will always be bound to it.

SPR: Surface plasmon resonance. A way to measure molecular interactions.

ssRNA: Single stranded RNA; a type of molecule that stores the information to make proteins. Is found in many viruses.

Surface representation: A model of a molecule that represents the space taken up by the molecule/its chemical components. For example, while a protein is composed of amino acids, and those amino acids are often drawn with a 2-D representation (see amino acids table), these structures take up 3-D space, and a surface representation aims to show that.

Titrate: To slowly add more of something while measuring the response due to its addition.

Toll-like Receptor 4 (TLR4): A transmembrane receptor, which when activated triggers the intracellular NF-KB pathway, ultimately causing inflammation. It is expressed in select immune cells, and functions as a part of the innate immune system. Its cofactors are MD-2 and CD14, and it is known to be activated by LPS and S100A9. It is thought that it exists *in vivo* as a homodimer.

Transcription/Transcribed: The process by which DNA is converted to RNA.

Transfection: The process of introducing exogenous DNA into cells. The contents of this DNA are then read by the cells, and the proteins that it encodes are produced.

Triacylated lipopetides: A type of molecule found in the cell wall of certain bacteria.

Truncation: The removal of a string of amino acids. For example, if a protein sequence normally has 100 amino acids, and a truncation mutation of the first 50 amino acids is created, the resulting mutant only has the first 50 amino acids.

Western blot: A method used to determine what proteins are present in a sample.

Wild-type (WT): The naturally occurring form; in this case, non-mutagenized CD14.

Bibliography

Alberts, B. *Molecular biology of the cell*. New York: Garland Science Taylor & Francis. (2008).

Amino Acids Reference Charts. Sigma. (n.d.). https://www.sigmaaldrich.com/life-science/metabolomics/learning-center/amino-acid-reference-chart.html.

Averill, M. M., Kerkhoff, C., & Bornfeldt, K. E. S100A8 and S100A9 in Cardiovascular Biology and Disease. *Arteriosclerosis, Thrombosis, and Vascular Biology* **32**, 223–229 (2012) doi: 10.1161/atvbaha.111.236927.

Barnes, P. J., & Karin, M. Nuclear Factor-κB — A Pivotal Transcription Factor in Chronic Inflammatory Diseases. *New England Journal of Medicine* **336**, 1066–1071 (1997) doi: 10.1056/nejm199704103361506.

Buchmann, K. (2014). Evolution of innate immunity: clues from invertebrates via fish to mammals. *Frontiers in immunology*, *5*, 459.

Christmas, P. (2010). Toll-like receptors: sensors that detect infection. *Nature Education*, *3*(9), 85.

Cunningham, M. D., Shapiro, R. A., Seachord, C., Ratcliffe, K., Cassiano, L., & Darveau, R. P. CD14 employs hydrophilic regions to "capture" lipopolysaccharides. *The Journal of Immunology* **164**, 3255-3263 (2000).

Gong, T., Liu, L., Jiang, W., & Zhou, R. DAMP-sensing receptors in sterile inflammation and inflammatory diseases. *Nature Reviews Immunology*, 1-18 (2019).

Granucci, F., & Zanoni, I. (2013). Role of CD14 in host protection against infections and in metabolism regulation. *Frontiers in cellular and infection microbiology*, *3*, 32.

He, Z., Riva, M., Björk, P., Swärd, K., Mörgelin, M., Leanderson, T., & Ivars, F. CD14 Is a Co-Receptor for TLR4 in the S100A9-Induced Pro-Inflammatory Response in Monocytes. *Plos One* **11**, (2016) doi: 10.1371/journal.pone.0156377.

Juan, T. S. C., Kelley, M. J., Johnson, D. A., Busse, L. A., Hailman, E., Wright, S. D., & Lichenstein, H. S. (1995). Soluble CD14 truncated at amino acid 152 binds lipopolysaccharide (LPS) and enables cellular response to LPS. *Journal of Biological Chemistry*, *270*(3), 1382-1387.

Källberg, E., Vogl, T., Liberg, D., Olsson, A., Björk, P., Wikström, P., … Leanderson, T. S100A9 Interaction with TLR4 Promotes Tumor Growth. *PLoS ONE* **7**, (2012). doi: 10.1371/journal.pone.0034207

Kelley, S., Lukk, T., Nair, S., & Tapping, R. The crystal structure of soluble human CD14 reveals a bent solenoid with a hydrophobic amino-terminal pocket. *The Journal of Immunology* **190**, 1304-1311 (2012).

Kim, J. I., Lee, C. J., Jin, M. S., Lee, C. H., Paik, S. G., Lee, H., & Lee, J. O. (2005). Crystal structure of CD14 and its implications for lipopolysaccharide signaling. *Journal of Biological Chemistry*, *280*(12), 11347-11351.

Markowitz, J., & Carson III, W. E. (2013). Review of S100A9 biology and its role in cancer. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer*, *1835*(1), 100-109.

Molteni, M., Gemma, S., & Rossetti, C. (2016). The role of toll-like receptor 4 in infectious and noninfectious inflammation. *Mediators of inflammation*, *2016*.

Park, B. S., Song, D. H., Kim, H. M., Choi, B. S., Lee, H., & Lee, J. O. The structural basis of lipopolysaccharide recognition by the TLR4–MD-2 complex. *Nature* **458**, 1191-1195 (2009).

Pahwa, R., Goyal, A., Bansal, P., & Jialal, I. (2020). Chronic inflammation. *StatPearls [Internet]*.

Pelletier, M., Simard, J. C., Girard, D., & Tessier, P. A. Quinoline-3-carboxamides such as tasquinimod are not specific inhibitors of S100A9. *Blood advances* **2**, 1170-1171 (2018).

Poltorak, A., He, X., Smirnova, I., Liu, M. Y., Van Huffel, C., Du, X., ... & Beutler, B. (1998). Defective LPS signaling in C3H/HeJ and C57BL/10ScCr mice: mutations in Tlr4 gene. *Science*, *282*(5396), 2085-2088.

Rudd, K. E., Johnson, S. C., Agesa, K. M., Shackelford, K. A., Tsoi, D., Kievlan, D. R., ... & Naghavi, M. (2020). Global, regional, and national sepsis incidence and mortality, 1990–2017: analysis for the Global Burden of Disease Study. *The Lancet*, *395*(10219), 200-211.

Uematsu, S., & Akira, S. (2008). Toll-Like receptors (TLRs) and their ligands. *Toll-like receptors (TLRs) and innate immunity*, 1-20.

Vogl, T., Stratis, A., Wixler, V., Völler, T., Thurainayagam, S., Jorch, S. K., ... & Paruzel, P. Autoinhibitory regulation of S100A8/S100A9 alarmin activity locally restricts sterile inflammation. *The Journal of clinical investigation* **128**, 1852-1866 (2018).

Wang, C., Klechikov, A. G., Gharibyan, A. L., Wärmländer, S. K. T. S., Jarvet, J., Zhao, L., … Morozova-Roche, L. A. The role of pro-inflammatory S100A9 in Alzheimer's disease amyloid-neuroinflammatory cascade. *Acta Neuropathologica* **127**, 507–522 (2013) doi: 10.1007/s00401-013-1208-4.