Computational and Rational Stabilization of Toll-Like Receptors for the Development of Novel Tools

Cameron Criswell

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2025

Reading Committee:

Neil King, Chair

Jessica A. Hamerman

Thomas R. Hawn

Program Authorized to Offer Degree:

Biochemistry

University of Washington

Abstract

Computational and Rational Stabilization of Toll-Like Receptors for the Development of Novel Tools

Cameron Ann Criswell

Chair of the Supervisory Committee:

Neil King

Department of Biochemistry

Toll-like receptors (TLRs) are membrane-bound pattern recognition receptors essential for innate immune sensing, but their large, glycosylated extracellular domains and intrinsic instability make them notoriously difficult to express and purify recombinantly. These challenges have historically limited structural and functional studies, as well as the development of therapeutic reagents targeting TLRs. This dissertation presents a computational design framework for stabilizing and functionally interrogating TLRs, enabling the development of synthetic immunomodulatory tools. Using AI-guided design with ProteinMPNN, AlphaFold2 and RosettaFold Diffusion in addition to physics and rational approaches, we generated stabilized and expressible variants of TLR2 and TLR5, facilitating downstream applications including de novo minibinder generation and antibody development. Overall, this work highlights a generalizable approach to stabilizing immune receptors and advancing rational immunotherapy design.

Table of Contents

Acknowledgements

As I look back on the six years it took for me to get here I am overwhelmed with gratitude for everyone that helped me get to this point. First and foremost, I want to thank Neil for letting me work in one of the coolest labs on the planet. From the moment you presented at the Biochemistry Retreat back in 2019 I knew your lab was where I wanted to be. You allowed me to grow as a scientist on my own terms, giving me the independence to try (and fail, and fail and fail) without losing faith that one day it was going to all come together.

I also want to thank all the professors at Texas A&M University as well as the amazing friends (Karissa, Bavnah, and Morgan) I made while there. Without your dedication and fantastic teaching I would not have been able to make it to this point. I distinctly remember my junior year where I had back to back classes in the same room, followed by lunch, then another class in the same room. Spending almost 8 hours in the same building with the same people learning about science is something I look back on with fondness. To the roommates I had at A&M, Katie Naiser and Darla Deckard, thanks for always being there as we struggled through our respective STEM fields together.

I would also like to thank the mentors in the King Lab that nurtured the young scientist just out of undergrad. Jung Ho Chun taught me everything a rotation student needs to know about protein design, from computational design to yeast display and BLI. Looking back, I learned more skills in those first 10 weeks than any other portion of my life.

The King lab culture is one of the best a young scientist could hope for. I am constantly humbled by the willingness people have to stop what they are doing to help. To Annie, and Cara for always helping with my EM questions, Mark, Chelsea and Naveen for constantly managing the AKTAs (I am sorry for letting the buffers run out). An unbelievable shout out to Sebastian, who introduced the antibody portion of the project to me and helped in every step to make sure I was successful. To everyone else who is always there for either a laugh (or cry) Grace, Cyrus, Meg, Marti, Sam, Helen, JoJo, Maggie, Hyojin and Svea. I want to also thank Ho Min Kim, a visiting scientist from Korea who helped me believe in myself during the darkest part of my PhD. I also want to thank RUNCLUB: Maggie, Hoyjin, Marti, Cara, Chloe, and Jojo for helping me blow off steam while getting fit.

I want to give a special mention to Chloe Adams, my biggest mentor and closest friend in the lab. There are so many things you taught me and there is no way I could ever repay you. As the only other member working on TLRs, there are certain struggles only

you would understand. These last couple months since you have gone to greener pastures have made me realize how much you have done for me and I greatly appreciate all of it.

To the IPD Management and their Admin staff (Kandise, Hernan, Luki, Patrick, Zari, and Sylvia), I thank the tireless (and sometimes thankless) work you do. Ordering packages, setting up meetings, fixing bugs, you are the lifeblood of the IPD and I thank you for making my job easier every day. In addition, I want to thank Erin Kirschner, Tonya A Hirtzel and the entire Biochemistry staff. You got me into this program and you'll make sure I get out, and I appreciate everything you do.

I am so lucky to have the support from so many people outside the lab as well. Thank you to all my friends Reed for always asking how it's going. Even when I got mad when it wasn't going well, I appreciated the gesture. To Gwen for driving me to the lab that one year during COVID. To Emily for the positive vibes and cute dog Piper. And to Taylor, for never charging me rent for the 1.5 I basically lived in your apartment during lockdown. Finally to Jay and Joelynn for throwing some amazing parties keeping my mind away from the stress of lab.

Since I was young I have always known I wanted to be a scientist (although I didn't know what that meant). To my parents Ann and Rex Criswell, I am so grateful for you fostering my passion and guiding me in the right direction for all these years. You allowed me every opportunity to be the best version of myself that I could be and always knew I was going to make the right decision, even if it meant being over 3,000 miles away from home. And of course, to my brother Ryan, even though we are total opposites, we are always on the wavelength.

Last but never least, I want to thank my fiance Nicholas Tiner. To say I couldn't have asked for a better person to scrap through the last six years of my PhD would be an understatement. No one believes in me more and I could not be more grateful to have you in my life. You were there for some of my biggest disappointments and breakthroughs, never knowing what I was talking about, but always knowing how much it meant to me. Even on my worst days, you reminded me that I was doing well and to believe in myself. You pulled me out of slumps and helped me take my mind off the lab, whether it was going well or not. Your love and support always motivated me even when times were tough.

Chapter 1. Introduction

Toll-like receptors (TLRs) are a conserved family of pattern recognition receptors (PRRs) that play a pivotal role in the innate immune system by recognizing pathogen-associated molecular patterns (PAMPs) derived from microbes[1–3]. In humans, ten TLRs (TLR1–TLR10) have been identified, each exhibiting specificity for distinct microbial components, including lipopolysaccharides, nucleic acids, and lipoproteins. While some TLRs are located on the surface of cells; TLR1,2,4,5,6 and 10, responding to microbial cell wall or motile components[1,4].While unique in their location both cell surface and endosomal TLRs share a highly conserved structural architecture and therefore mechanism of activation. Structurally, TLRs are type I transmembrane glycoproteins characterized by an extracellular domain composed of leucine-rich repeats (LRRs) arranged in a “horseshoe” shaped structure[4–6]. This domain is responsible for ligand recognition and binding.

TLRs are predominantly expressed on innate immune cells, including macrophages, dendritic cells, and epithelial cells, and are localized to either the plasma membrane or endosomal compartments, aligning with their ligand specificities[1,3,7,8]. Given their central role in immune activation and their ability to bridge innate and adaptive immunity, TLRs are critical targets for both immunostimulatory and immunomodulatory therapeutic strategies[9–12].

Due to TLRs key role in bridging the innate and adaptive immune system, researchers have targeted this family for various therapeutic uses[10,13–15]. There are a myriad of different approaches from both activation and repression that are currently being investigated. Vaccine adjuvants using TLR4 and TLR9 have been widely used to great success to activate the innate immune system in a controlled manner. Namely TLR4 based adjuvant AS01 and AS04 have shown great efficacy of multiple vaccines from shingles to malaria and HPV[16,17]. Additionally, TLR agonists have been incorporated into cancer immunotherapy regimens to stimulate tumor-infiltrating immune cells and reverse the immunosuppressive tumor microenvironment[11,18]. Agents such as TLR7/8 agonists (e.g., imiquimod) have demonstrated efficacy in treating skin cancers[10], while TLR9 agonists are under investigation for enhancing tumor antigen presentation and synergizing with immune checkpoint inhibitors[18,19] .

While potent immune modulating agents, TLRs when left unchecked can cause autoimmune responses. TLR signaling has been implicated in the pathogenesis of autoimmune and chronic inflammatory diseases through aberrant or sustained activation. Both TLR7 and 9 have been implicated in Systemic Lupus Erythematosus (SLE)[12,20]. Additionally, TLR4 overactivation in the glia has been correlated with increased Alzhiemers risk[9]. TLRs have also been shown to be misregulated on the surface of cells within tumors[21,22 23].

Given the critical role of TLRs in disease and the necessity for precise modulation, there is a growing demand for reliable, high-quality ectodomain

preparations to enable the development of targeted therapeutic tools. Unfortunately, TLRs are extremely difficult to produce recombinantly, especially in mammalian cells[6,24,25]. This has created a large gap in the field as the ability to obtain reliable, high quality ectodomain is highly difficult. The combination of these factors shows the need for highly specific tuning of TLRs for the use needed. Due to this, the development of TLR focused tools, including novel antibodies, agonists and antagonists, has been extremely difficult.

Recent advances in artificial intelligence (AI), particularly in large language models (LLMs) and deep learning frameworks, have revolutionized the field of protein design, accelerating the generation of novel proteins with desired structural and functional properties. Models such as ProteinMPNN[29,30] can predict optimized amino acid sequences compatible with a given backbone, allowing the rapid generation of diverse, stable, and functionally competent proteins[29,31,32]. In addition AlphaFold2 (AF2) has provided near-experimental accuracy for a wide array of protein folds[33], including difficult targets such as membrane proteins or complexes[29,31,34]. This capability has significantly streamlined the design-validation loop, enabling in silico prediction of the structural consequences of designed sequences, and assessment of their folding accuracy.

Computational design has enabled de novo protein design, wherein entirely new protein backbones and folds are created without relying on naturally existing templates[26,29,30,35]. Tools like RFdiffusion[36], a diffusion-based generative model, have been utilized to generate novel protein topologies and scaffolds with tunable properties. These approaches have also been applied to modeling de novo protein protein interactions enabling the rational engineering of binding sites with high affinity and specificity. Deep learning models can predict favorable interaction interfaces[26,29,34], guide the redesign of existing scaffolds[31], and optimize sequences for enhanced binding kinetics . This is exemplified by the successful application of AI models in the design of minibinders against viral targets[26].

To harness the recent advances in AI as well as the power of TLRs for immune modulation, we first stabilized human TLR2 (2S5) then TLR5 (H4) for the development of various therapeutic tools. Once stable, 2S5 was used to create both monoclonal antibodies (mAbs) and de novo minibinders against TLR2. The final three chapters describe the computational approaches that were also used to develop both minibinders and de novo mAbs against TLR1 and TLR4. While the relative success of these projects varies, we describe a successful model going forward for stabilizing TLRs and the development of their tools for future uses.

Chapter 2. Computational design of stabilized TLR2 ectodomains and their use in generation of novel antibodies

Abstract

Toll-like Receptor 2 (TLR2) plays a critical role in innate immunity by recognizing microbial components[5,18,37–39], yet its poor recombinant expression and structural instability have impeded therapeutic and antibody development. Here, we report the design and characterization of a computationally stabilized variant of human TLR2, termed 2S5, engineered using deep learning-based tools (ProteinMPNN[30], AlphaFold2[33]) in conjunction with rational design principles[40–42]. Key stabilizing mutations were introduced to both termini and structurally unstable regions, enhancing solubility and thermostability without disrupting critical functional interfaces[6,25,39,43,44]. The 2S5 variant demonstrated robust expression in mammalian systems as both monomeric and multimeric forms and exhibited a > 20°C increase in melting temperature relative to wild-type TLR2. Structural analyses confirmed the preservation of canonical leucine-rich repeat architecture while functional assays further validated binding to known TLR2 antibodies and the Pam3CSK4 ligand, confirming retention of native epitopes. When displayed on a self assembling nanoparticle I53-50 platform[45,46], 2S5 elicited strong humoral responses, underscoring its utility as a high-quality immunogen for monoclonal antibody discovery. This work establishes a generalizable strategy for stabilizing innate immune receptors and provides a valuable tool for TLR2-targeted therapeutic development.

Introduction

Toll-like receptors (TLRs) constitute a central family of pattern recognition receptors within the innate immune system[1–3]. In humans, ten TLRs have been identified, eight of which recognize distinct pathogen-associated molecular patterns (PAMPs), enabling discrimination between self and non-self entities[1]. Structurally, TLRs are type I transmembrane proteins composed of 20–30 leucine-rich repeats (LRRs) that form a characteristic horseshoe-shaped solenoid, a single transmembrane helix, and a cytoplasmic Toll/IL-1 receptor (TIR) domain[5,25]. Upon engagement with their respective PAMPs, TLRs dimerize, leading to multimerization of the TIR domains and subsequent recruitment of adaptor proteins via the MyD88 or TRIF pathways[6,25,47]. This triggers downstream activation of NF-κB and IRF3/7, culminating in the induction of pro-inflammatory cytokines, chemokines, and type I interferons[47–52]. Given the well-characterized signaling pathways and immunological functions of TLRs, they represent attractive targets for therapeutic strategies aimed at modulating immune responses[9,10,22,53]. However, the intrinsic challenges associated with the recombinant

expression of TLRs have hindered the development of biologics, including high-quality monoclonal antibodies (mAbs), against these receptors.

Recent advances in both computational and rational protein engineering approaches have significantly improved the stabilization and expression of difficult proteins[30,32,54,55]. In particular, deep learning-based tools such as ProteinMPNN and AlphaFold2 have been successfully employed to guide the redesign of structurally complex or unstable targets . Integration of these artificial intelligence (AI) methodologies with rational design has facilitated the identification of stabilizing mutations that enhance expression, stability, and functional activity of clinically relevant, yet difficult-to-manufacture, proteins[32,41,54–56]. These developments have rapidly accelerated progress in the field of protein stabilization and therapeutic design.

One of the most impactful applications of these strategies is in the optimization of antigenic proteins for mAb discovery[40,41,57]. A major bottleneck in mAb development lies in the production of high-quality antigens capable of eliciting robust B cell responses, a challenge particularly pronounced for TLRs, for which few reliable mAbs are available[58]. In light of the recent success of computational stabilization techniques and the persistent difficulties in TLR expression, we sought to apply a combination of deep learning-based and rational approaches to redesign human TLR2 (hTLR2) for improved expression and immunogenicity.

Here, we demonstrate that the application of these methodologies yielded a stabilized hTLR2 variant, designated 2S5, which exhibits robust expression in mammalian systems and supports the generation of mAbs against hTLR2. The 2S5 variant can be expressed both as a monomer and in multimeric forms when fused to Fc or the trimeric component of self assembling nanoparticle scaffold 153-50[45,59]. Notably, 2S5 displays enhanced thermal stability while retaining the structural integrity and functional properties characteristic of native TLR2. When presented on the self-assembling nanoparticle platform I53-50, 2S5 elicited high-affinity antibodies targeting hTLR2, highlighting its utility as an optimized antigen for mAb discovery.

Results

Stabilizing TLR2 with AI and holistic approaches

Our initial goal was to redesign human TLR2 (hTLR2) ectodomain to improve both its expression and overall stability while retaining its surface and functional characteristics. To do this we developed an approach using both ProteinMPNN[30] to optimize the side chains of TLR2 without altering the protein backbone. Key characteristics of the TLRs were identified to mutate or mask from redesign (Fig. 1a). To start, we masked five disulfide bonds as well as four likely N-linked glycosylation sites (PNGS). Additionally,

the interface where TLR2 heterodimerizes with TLRs 1, 6, and 10[6,39,60] as well as the acyl chain binding pocket within the core of TLR2 were held constant[5,6,61]. In addition, we forced the mutation of unpaired cysteines as this has been cited to hinder protein expression[41,62,63]. Finally, excepting the restrictions noted above, we allowed ProteinMPNN to fully redesign the 75 N-terminal and 50 C-terminal residues of the protein, as several previous studies have demonstrated that specifically the C-terminal region of TLR ectodomains is suboptimal and that replacing it with heterologous leucine-rich repeats derived from variable lymphocyte receptors (VLRs) can improve expression and stability[6,25,39,43]. While the C-terminal has been frequently noted as being suboptimal, our studies have shown the N-terminal is also highly unstable and was thus included in the full redesign.

An increase in surface conservation allows for larger areas of similarity for future development of tools such as antibodies[40,64,65]**.** Thus a balance needed to be struck for optimal surface conservation and predicted expression success based on non templated AF2 plddt scores. Due to the lack of an experimentally determined hTLR2 structure, we used ColabFold[66,67], a version of AF2[33], that utilized a MMseqs2[68] clustering data for modeling, to generate a model of residues 26-588 of full length TLR2 as a substrate for redesign. These models produced significantly improved plddts as well as RMSDs compared to AF2 models (56.7 vs. 93.9 and 8.54Å vs. 1.19Å). The model generated recapitulated the characteristic LRR tertiary structure and an RMSD of 1.19Å compared to partial length hTLRs with VLR domains. While AF2 failed to properly model TLR2, we continued to use its plddt scores as a key metric as improvements here indicated we were moving toward more predictable (and likely more stable) designs.

As a test of the ability of side chain optimization alone to improve the confidence of AF2 prediction we allowed ProteinMPNN to redesign all residues except the TLR1/6/10 interface and acyl chain binding pocket in the model. Predicting the structures of 1,500 redesigned sequences resulted in large deviations from the original prediction (2.05Å-26.6Å RMSD) and considerable improvements in pLDDT (60.8-81.5) based on AF2 predictions. The large range in both RMSD and pLDDT is unsurprising given the freedom we allowed during redesign in this stage. While scoring well and expressing well (Extended Fig1. a), we inferred these unconstrained models would likely not function well as a proxy for development of future TLR tools due to lack of sequence similarity to TLR2, which hovered below 60% on average. To solve this, rational redesign[42,57,69] was employed to identify extraneous mutations from unconstrained computational redesign **(**Fig. 1b**)**. To create a conservatively mutated TLR that is simultaneously functional and developable in mammalian cells, we developed an

iterative ProteinMPNN-based approach that combined only the most promising mutations.

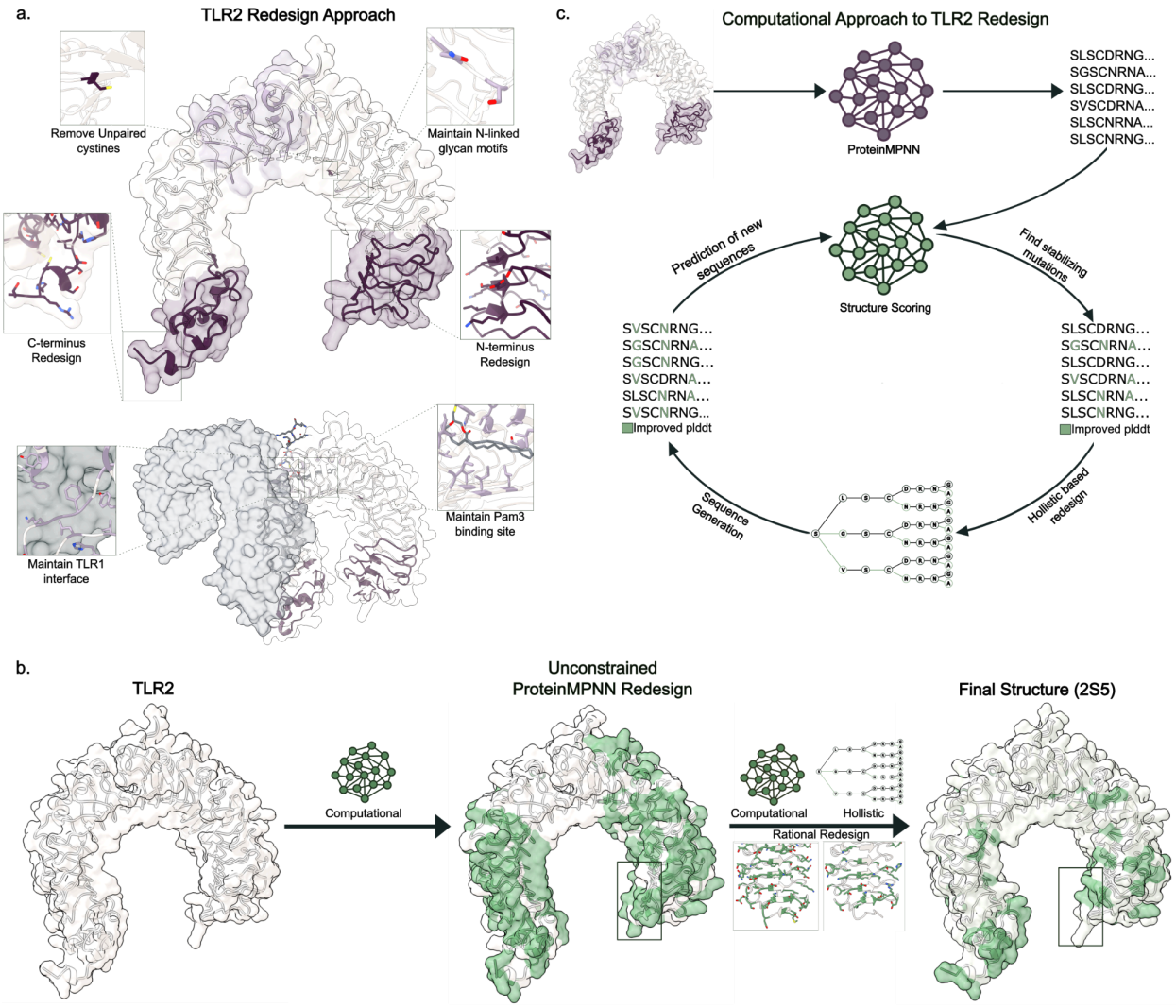


Figure 1. Computational and rational approach to hTLR2 redesign

a, Schematic showing redesign approach for hTLR2. Unpaired cysteines and terminal domains were removed to increase stability while PNGS domains and TLR1/Pam3 interface residues were held to maintain functionality. **b,** Pipeline for improving hTLR2 from Unconstrained design to Final Structure (2S5). hTLR2 is put through AI based computational methods to generate Unconstrained Models. Once well scoring designs were identified, rational redesign was used to revert mutations along computational and holistic approaches to stabilize reversions to generate the final structures. **c,** Holistic design method generating diverse sequences. Protein models are fed into ProteinMPNN, sequences are scored and the best residues are

combined to generate novel sequences for prediction. This process is iterable and will converge on a set of sequences.

Once unconstrained models were created first we looked deeper into the highest scoring of these models and selected a subset for increasing conservation to hTLR2. Firstly, mutations that were unnecessary or were inferred to have small benefits were reverted back to improve overall sequence conservation to TLR2. As a general rule, surface mutations were heavily scrutinized, while all core mutations except for the acyl chain binding site were allowed. The sequence conservation increased from approximately 59% to 83% identical to hTLR2. Once we created these reversions we re-scored these designs. As expected, reversion of these designs back to partially native states led to a large reduction in pLDDT scores, indicating a need for further redesign with ProteinMPNN. To do this, while maintaining most of the conservation, obtained from reversions we then allowed the subset of still mutated residues and some of their neighboring residues to be re-submitted to MPNN, generating 1500 sequences for each of the four parent designs (Fig. 1b). While these results were good, we felt the scores could be improved upon. This led us to look for additional methods to increase pLDDT scores for these designs. A holistic approach was then used to generate a large number of novel sequences based on these MPNN outputs. This method took the top scoring residues based on pLDDT of each residue and shuffled these residues from MPNN outputs (Fig. 1c). Because this method is capable of exponentially large sequence generation, a random subset of 1,500-500 was taken and rescored with AF2 and repeated over multiple rounds. The holistic approach consistently produced higher scoring outputs that improved over multiple rounds of iteration and led to higher plddt's on average than MPNN alone with greater sequence diversity. After the final round of iteration the top designs were screened for sequence conservation to each other as well as the types of mutations made. While all versions of stabilized TLR2 expressed well (Extended Data Fig1. b), 2S5 was chosen for further characterization due to its high pLDDT score (69.580) and low RMSD (1.740Å) to hTLR2.

Key mutations for stabilization of hTLR2

In order to increase expression of 2S5 various mutations at key locations were made (Fig. 2a). To stabilize the N-terminus, an increase in polar residues on the surface contribute to the stability of 2S5. More precisely, a high amount of mutations were made within the inner beta sheets making up the first four LRR domains (Fig. 2b). While most of these interactions were mutations from nonpolar to polar residues. The most impactful mutation to pLDDT was H78V. This mutation led to a favorable hydrophobic interaction with F103 in the LRR above, while still being small enough to not disrupt the polar surface interactions (Fig. 2b). An additional key mutation we identified was G74T on the surface of 2S5. This mutation interacts with V50T below to create a more stable

interaction on the surface between the LRRs. Other notable mutations that improved pLDDT were changing a large hydrophobic to a small polar residue (F149S) and a proline mutation that stabilizes a loop (Q172P). These represent small optimizations that could be made via rational design and were confirmed by computational techniques (Fig. 2a).

The two bends within the TLR2 that contribute to its canonical “horseshoe” structure were an area of major instability for TLR2. There were multiple mutations within the core of this N-terminal bend (F191I, L203I, T206M, F228V, I250F). Of the mutations, T206M is likely the most key to increasing the stability of this region by replacing a polar to a hydrophobic in the protein core. In addition alterations where large aromatics and smaller hydrophobics are placed indicate that TLR2 may have been poorly packed (Fig. 2b). The remaining mutation to L203I assists in the repacking of the core to fill in the voids created from removing the phenylalanines from the core (Fig. 2b). In the concave region, the change from cysteine to histidine resulted in a large improvement in pLDDT. Likely due to the improved polar interactions, including the addition of a hydrogen bond spanning from H200 to S229 (Fig. 2b).

Toward the C-terminal side of the Pam3 binding domain is another bend with multiple key mutations in the core (Fig. 2a). While not as dramatic as the repacking from residues 190-250, there are subtle changes that increase the stability of this bend. Most notably, the replacement of I392L, and M412L make room for the large W408 while L373I and I394V open more room for F399 (Fig. 2b). Lastly, while not a notable mutation on the graph, the rational mutation of T420N in the core reverts the threonine back to its conserved residue at that position. This improves the hydrogen bonding network within the core and introduces a more evolutionarily conserved residue back into this position[1,2,6,70].

To stabilize the C-term, ProteinMPNN was allowed to fully redesign the C-terminus from residue 515 to the end, with the exception of the disulfide bonds at 538 and 559 (Fig. 2b). A large portion of the surface was mutated to larger, more polar residues. These changes contribute to an improved hydrogen bonding network at the C-terminus which will help prevent flexibility in the end of the protein. Additionally, F509Y creates multiple interactions within the core with polar residues that have not been mutated (N507, Y536). Mutations repack the C-terminus by replacing many large amino acids for more compact ones (F515L, W532Y). At the very end of the C-terminus, larger and more polar residues are used (I510D, H432K, V544Y) (Fig. 2b) to cap the protein. These residues fill the pores in the C-terminal cap and creates a more polar surface. This likely protects the very bottom of the terminus from water entering the hydrophobic core resulting in unfolding and increased instability.

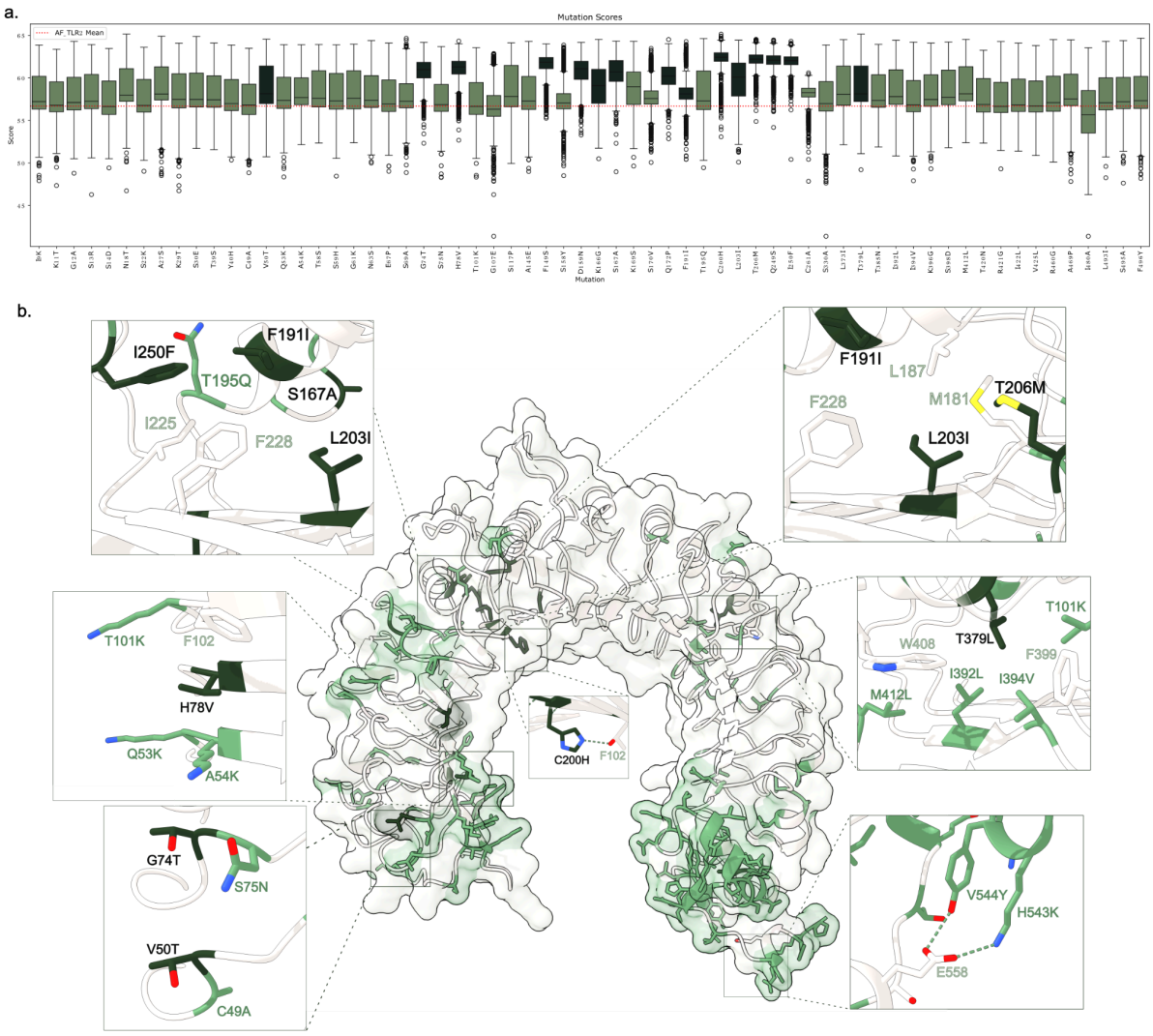


Figure 2. Structural comparison of hTLR2 and 2S5.

a, Table displaying the key mutations from residues 1-496. Dark green bars indicate high impact mutations that resulted in large changes in pLDDT. **b,** Full display of 2S5 with mutations. Non-mutated portions are in gray, while high impact mutations are in dark green. Other mutations that had less impact are displayed in light green. Cutouts display regions where multiple high impact mutations are localized.

Characterization of structure and functionality of 2S5

Due to the difficulty to express stable native TLR2, most expression protocols call for insect cells. We sought to create a stabilized TLR2 capable of expression in mammalian suspension cells. Initially, all stabilized TLR2 proteins were ordered with an Fc domain at the C-terminus as it is known to increase expression of TLRs[6,25,39]. The addition of the Fc domain forces 2S5 into a rigid dimer (Fig. 3a). While this likely helps expression due

to the increased stability at the C-terminus, we feared that the dimer prevents adequate binding of Pam3 and TLR dimerization partners and blocks ideal binding locations for future therapeutic design[61,71]. This led us to develop a stabilized TLR that did not require an Fc domain to be successfully expressed in mammalian cells. Once we saw expression in 2S5-Fc, we examined if 2S5 could be successfully expressed without the Fc domain. To do this, we designed two additional forms of 2S5: Avi-Tag (monomeric) and fused to nanoparticle I53-50A (trimeric)[46,72] (Fig. 3a).

To fully characterize whether 2S5 was capable of expression and was structurally stable, 2S5 was expressed in Expi293 cells and purified via IMAC and SEC (Fig. 3b, Extended data Fig. 2a). Each of the 2S5 versions expressed successfully, albeit at different rates with monomeric, 2S5-Fc and 2S5-50A expressing at 1.125mg/L, 3.52mg/L, and 1.95mg/L respectively. This is unsurprising given that Fc is well known to increase expression and 50A has also been shown to assist in secretion of proteins[6,25,39]. The size of 2S5 was then confirmed to correspond to previous literature[5,6,61] by Dynamic Light Scattering (DLS) of 2S5 and 2S5-Fc which showed a monodisperse protein with minimal aggregation at room temperature and a size of 7Å and 13Å respectively (Fig. 3c) This differs greatly from the native TLR2, which shows an extremely polydisperse particle with high levels of aggregation, even at room temperature.

Once the expression yield and relative size was confirmed, we sought to ensure 2S5 secondary structure corresponded to the cited in the literature[5,6,61], namely the presence of beta sheets which make up the horseshoe shape. To further characterize the secondary structure Circular Dichroism (CD) was done to confirm the presence of beta sheets[73], a defining aspect of the LRR tertiary structure and the TLR as a whole. CD data confirmed the presence of beta sheets in 2S5 (Fig. 3d), confirming the secondary structure of 2S5 matches that of hTLR2. We then wanted to identify if 2S5 is in fact more thermostable than TLR2. To do this Differential Scanning Fluorimetry (DSF) analysis showed a dramatic increase in melting temperatures between hTLR2 and 2S5. 2S5 exhibits an almost 25C increase in Tm over TLR2 based on DSF (58.1℃ vs. 82.6℃) (Fig. 3e, Extended data Fig. 2b). Thermal melts of 2S5-Fc displayed two melting points, one at approximately 68℃-the melting temperature of the mouse IgG[74,75]-and another at 81℃ (Fig. 3e), corresponding to the monomeric form of 2S5-Avi.

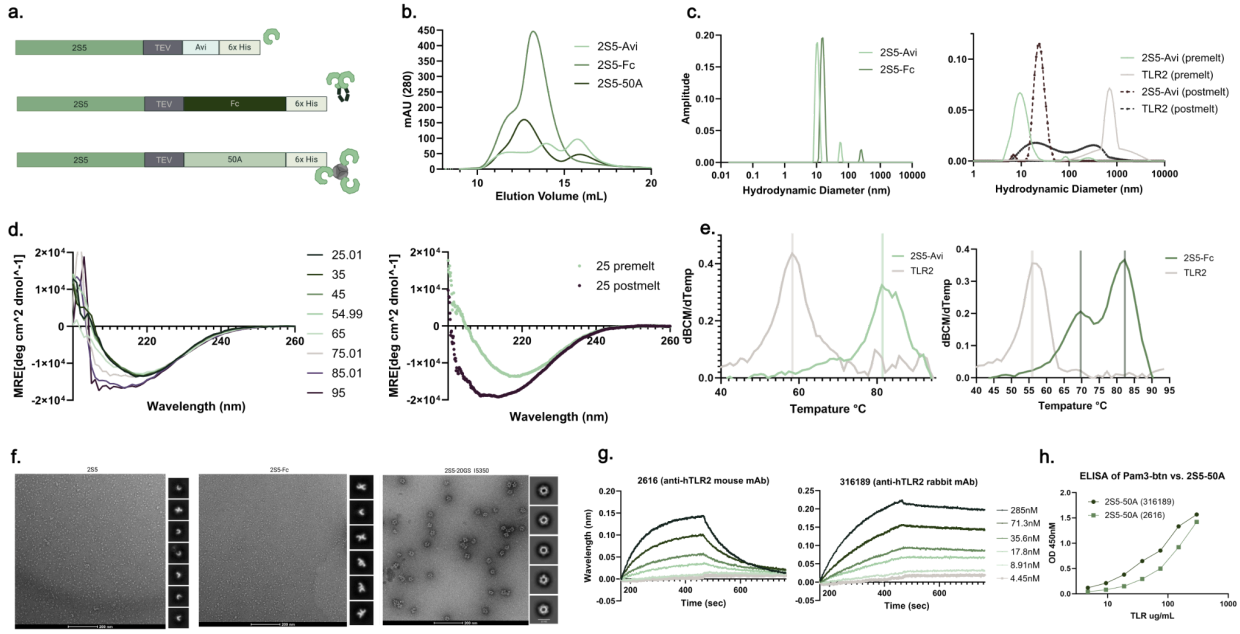


Figure 3. Biochemical characterization of 2S5.

a, Map of various tags added to 2S5 versions for characterization. **b,** SEC of 2S5 versions displaying differences in elution volumes and expression levels. **c,** DLS of 2S5 versions and hTLR2 pre and post melt. Pre-melts are indicated in solid lines, and post-melts are shown with dotted lines **d,** CD of pre and post melt of 2S5 and temperature ramp of melting of 2S5. **e,** DFS data of 2S5 and 2S5-Fc compared to monomeric hTLR2. DFS shown is smoothed first derivatives used to calculate melting temperatures. Vertical lines indicate the melting temperatures of each protein. **f,** nsEM and 2D averages of 2S5, 2S5-Fc and 2S5 fused to I53-50. **g,** BLI of monomeric 2S5 against two mAb; 2616 and 316189. Kd values are given. **h,** ELISA of 2S5-50A against biotinylated Pam3 using both 2616 and 316189 mAbs.

To ensure 2S5 maintains proper folding we looked at structures under negative stain Electron Microscopy (nsEM). The 2D class averages of each show the monomeric 2S5 maintains the canonical “horseshoe” shape in and the dimeric form maintains the “M” form in averages (Fig. 3f). It should be noted that the monomeric form of 2S5 can form dimers stochastically. This is seen in both SEC traces and 2D averages (Fig. 3b, 3f). This dimerization may be some form of aggregation, but it has been shown that native TLR2 is capable of dimerizing[61,76], albeit this was mainly seen in the presence of a small molecule.

To further ensure 2S5 maintains key functional properties of TLR2, we tested its binding to various antibodies and its canonical agonist Pam3[6,38]. Two antibodies were tested, an

antibody with an unknown binding modality 316189 and an antagonising antibody 2616[77–79]. This led us to infer that this antibody binds near the Pam3 - and likely TLR1-interface[6]. BioLayer Interferometry (BLI) shows both 316189 and 2616 bind tightly to 2S5 (Fig. 3g) indicating that 2S5 maintains surface functionality similar to that of hTLR2. To go further, we wanted to ensure the core of 2S5 was functional as well. To do this we ran an ELISA of Pam3-btn against 2S5. ELISA data shows that 2S5 binds to Pam3 (Fig. 3h). This indicates that 2S5 is functional on both its surface and within the key binding domain inside 2S5.

2S5 fused I53-50 antigens elicit potent antibodies against both 2S5 and hTLR2

Once it was confirmed that 2S5 displayed high levels of expression with increased stability to that of TLR2 we looked to show the applicability of 2S5 by generating antibodies against that target hTLR2 using 2S5 as a proxy. It has been shown previously that multimeric display of antigens to the immune system is capable of inducing more robust B-cell responses[42,59,72,80,81]. With this in mind we leveraged the self assembling nanoparticle platform to co-display 60 copies of 2S5 on the surface of I53-50. To do this 2S5 was flexibly fused to the 50A component of I53-50 with a 20GS linker (Fig. 3a). As described previously, 2S5-50A trimer showed the expected size by SEC and DLS with no loss in functionality (Fig. 3f, 3h). Full cage assembly was confirmed by SEC and DLS with nsEM 2D class averages verifying cage formation (Fig. 3f, Extended Data Fig. 3a,b), the antigenicity of the 2S5 when assembled into a cage was confirmed with BLI to 2616 and 316189.

Mice were immunized with 5ug of 2S5-cage formulated with Addavax at 0, 4, and 8 weeks with bleeds at 2 and 5 weeks (Fig. 4a). At the two week mark, serum was taken to measure the antibodies generated pre-boost (Extended Data Fig. 3. c,d). After 10 weeks the spleens and sera were harvested. In order to produce antibodies against 2S5 and hTLR2, B-cell receptors (BCRs) from B-cells that bound 2S5 were identified, cloned and sequenced[72,82]. To do this, we sorted B-cells that were generated against 2S5 using Fluorescence-Activated Cell Sorting (FACS) with streptavidin beads bound to biotinylated 2S5 was used as a proxy to mimic a cell expressing 2S5 (Fig. 4b). The 2S5 binding B-cells had their heavy and light chains cloned and variable regions as well as CD3 domain to identify unique sequences against 2S5[82].

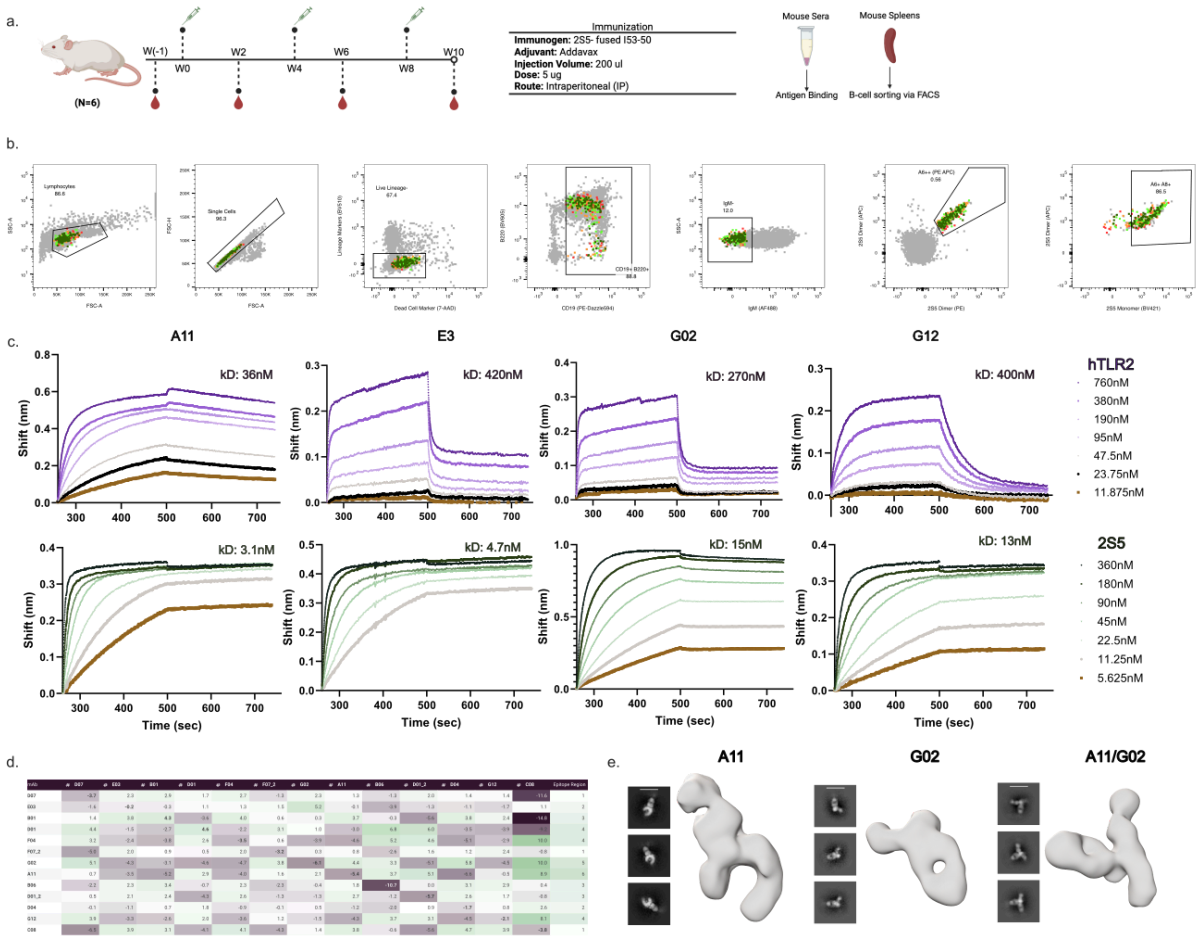


Figure 4. Development of antibodies against hTLR2 using 2S5 fused 153-50.

a, Mice were immunized with 5ug of 2S5 fused to I53-50 adjuvanted with Addavax and boosted 3 times. Bleeds were done every 2 weeks until the termination of the study at week 10 when spleens were harvested. **b,** FACS of B-cells harvested to identify binders to both monomeric and dimeric 2S5. **c,** BLI of cross reactive minibinders against hTLR2 and 2S5. Kd values for each given. **d,** All-by-all epitope mapping of binders that successfully bound to 2S5. Diagonal and bolded cells indicate binders that are competing against each other. Binning reveals six total bins for epitopes. Furthest left column indicates “bin” each epitope has been sorted in. **e,** 2D averages and 3D maps of mAbs bound to 2S5. Combining both mAbs shows the two binding modalities of the mAbs.

After cloning was completed, 38 unique BCR sequences were identified by cloning as 2S5 binders. We chose to down select the number of antibodies to characterize based on the highest occurring clonotypes and their diversity, in addition two antibodies were

identified to bind to the dimeric version of 2S5 (Fig. 4b). This led to 22 sequences being ordered for further testing. Unfortunately two of the sequences were unsuccessful during development while another three showed no loading to either ProA or AntiMouse Fc capture tips for BLI. When tested on BLI, 15 of the remaining 17 mAbs bound to monomeric 2S5**.** Once the known binders were identified, we looked to see if any bound with notably higher affinity dimeric form of 2S5. Four of the binders showed improved binding to the dimer over the monomeric 2S5, including the two noted during the sorting phase as binding specifically to dimeric 2S5.

Once we identified the mAbs capable of binding to 2S5, we sought to identify cross reactive antibodies to hTLR2. Of those that bound to 2S5, only four showed binding to hTLR2. Of the four, one mAb (A11) displayed a higher affinity of 36nM to hTLR2 than the other three which ranged from 270nM-420nM. The affinity of A11 is comparable to the 2616 and 316189 which have binding affinities of 63nM and 9.9nM respectively. Together this shows that the redesign of hTLR2 to a more stable form allowed for the development of high quality antibodies against TLR2 as well as the high immunogenicity induced by multivalent display of antigen on a cage[42,83–86].

We inferred that the drastic reduction in antibodies that are cross reactive was due to an immunodominant mutation made to 2S5[83,87]. In order to determine if the mAbs were binding to a common mutated epitope, we used an all-by-all competition matrix by BLI. When comparing mAbs that successfully bound to 2S5 in an all-by-all epitope screen we identified 6 unique epitopes. Although when tested against hTLR2 four were found to be cross reactive: A11, and E03, G02, and G11. Of the four cross reactive epitopes, G02 and A11 appear to bind to unique epitopes by binning (Extended Data Fig. 3e), while the remaining two might also have unique epitopes, poor loading made it difficult to parse these groups.

To better understand how each of these epitopes bound to 2S5, A11 and G02 Fab domains were bound to 2S5 and looked at under nsEM (Fig. 4e). The 2D averages revealed A11 clearly binds to the upper portion of 2S5 closer to the N-terminus. Conversely, while difficult to confirm without better resolution, G02 appears to bind to a terminal edge of 2S5, although additional angles will need to be obtained to confirm this. When A11 and G02 are combined as observed under nsEM, both mAbs bind to similar termini of 2S5, but different epitopes (Fig. 4e) although different angles will be needed to confirm epitope locations[88,89].

Discussion

This study demonstrates that computational protein design can be effectively leveraged to stabilize hTLR2, enabling the production of high-quality recombinant protein suitable

for downstream applications such as antibody generation. Historically, efforts to develop antibodies against TLRs have been hampered by poor protein expression and solubility[6,25,44]. By engineering a stabilized variant of TLR2, 2S5, we provide a model for overcoming these limitations, with implications for other structurally related but difficult-to-express receptors.

A key aspect of our strategy was the integration of artificial intelligence[30,32,33], rational design[40,41,57], and holistic assessment. AI tools such as ProteinMPNN and AF2 facilitated large-scale generation and scoring of sequence variants, rapidly narrowing down candidates with improved structural compatibility[55,65]. However, these models can produce opaque or non-intuitive solutions[90]. Rational design played a critical complementary role by identifying and excluding non-beneficial or superficial mutations from AI-generated outputs. A holistic approach—incorporating combinations of the highest scoring residues non specifically—further allowed us to rescue favorable variants that might otherwise have been overlooked. This combined methodology yielded a design process more robust than any individual approach alone.

While the C-terminal region of TLRs has previously been identified as a modifiable site for improving expression[6,25,39,43,44], our results indicate that stabilization of additional structural features—particularly the N-terminus and concave regions of the ectodomain—is key to achieve expression levels suitable for high-throughput screening. Given the structural conservation across the TLR family[1,70,91], these principles are likely to be transferable to other members, offering a generalizable framework for future stabilization efforts.

The functional significance of 2S5 was underscored by its utility in antigen presentation. Fusion of 2S5 to the self-assembling nanoparticle scaffold I53-50 markedly enhanced its immunogenicity compared to monomeric forms[72,80,81,88]. This technique would not be possible with native hTLR2, as fusion to the 50A component resulted in a marked decrease in expression compared to fusing the Fc. Using this nanoparticle display platform, we identified over 20 monoclonal antibodies via FACS, of which 15 bound specifically to 2S5 and four showed cross-reactivity across species. Epitope mapping revealed five distinct epitopes on 2S5, compared to only two on wild-type hTLR2. This discrepancy may reflect an immunodominant region on 2S5, potentially concentrating the humoral response. While promising, further studies including high-resolution structural characterization and functional assays are needed to confirm these mechanisms of action.

In summary, we present a design and screening pipeline that enables the stabilization of structurally challenging membrane-associated receptors like TLR2. The successful generation of both a functional, stable TLR2 variant and a diverse panel of antibodies

highlights the potential of this approach. Beyond TLR2, this methodology may be broadly applicable to other immunologically relevant receptors, facilitating the development of therapeutic tools and advancing our ability to modulate innate immune responses.

Acknowledgments

The authors are Cameron Criswell, Sebastian Ols, Craig Dobbins, Adri Tran-Pearson, Elizabeth Sorberg, Chloe S. Adams and Neil King. N.P.K. is a corresponding author. C.C designed stabilized TLR2. C.C. performed computational analysis of designs. C.C. and C.S.A screened designs and carried out biochemical characterization. C.C designed and carried out electron microscopy experiments. S.O and L.S designed and mouse experiments and L.S carried out mouse injections, bleeds and harvests. S.O designed B-cell screening design and carried out FACS sorting experiments and sequencing. A.T.P and C.D. performed cell-based assays. N.P.K. supervised research. In addition, I would like to thank Jung Ho Chun, Chloe Adams, Cara Chao, and Anne Dosey for helpful discussions. This work was supported by grants from the Bill & Melinda Gates Foundation (INV-010680 and INV-043758 to N.P.K.).

Materials and Methods

Computational Sequence Generation of hTLR2 with ProteinMPNN

A structural model of residues 26–588 was generated using ColabFold[67], an implementation of AF2[33] that incorporates MMseqs2-based sequence clustering[68]. Regions corresponding to the TLR1/2 dimerization interface, the Pam3 binding site, conserved disulfide bonds, and putative N-linked glycosylation sites were excluded from redesign for reasons listed in the main text. Sequence design was performed using ProteinMPNN with a backbone perturbation noise of 0.2 Å. Three temperature values (0.01, 0.1, and 0.25) were sampled, and 500 sequences were generated per temperature condition, yielding a total of 1,500 candidate sequences. Each designed sequence was structurally predicted using AlphaFold2, and the resulting models were assessed based on predicted local distance difference test (pLDDT) scores and root-mean-square deviation (RMSD) relative to the original hTLR2 model. The top 50 designs, ranked by pLDDT and structural diversity, were selected for further evaluation. The five highest-scoring sequences from this initial round were advanced to subsequent rounds of targeted redesign.

Holistic Sequence Generation

A holistic approach was employed to generate high-confidence sequence variants by integrating per-residue performance metrics. For each position in the aligned sequences, the top amino acids were identified based on per-residue pLDDT scores derived from AF2 predictions. In round one, 1,500 sequences were generated by

randomly sampling from the top 150 best-scoring combinations across positions. These sequences were structurally predicted using AF2, and the highest-scoring models were carried forward.

In round two, sampling was refined by selecting from the top 75 amino acids per position, generating an additional 1,000 sequences. In subsequent rounds (round three and beyond), the sampling pool was further narrowed to the top 50 amino acids per position, with 500 sequences generated per round. This iterative design process continued until some convergence in sequences was met indicating a relative maximum of scores and sequence diversity was achieved.

Mammalian Expression of 2S5

All constructs contained a TEV cleavage site as well as a C-terminal histidine affinity tag and were codon optimized by GenScript for mammalian cell expression. Expi293F cells were transiently transfected using PEI MAX and cultured for three days 37°C. Cells were harvested, centrifuged, and filtered to obtain clarified supernatants. Clarified cell supernatants were buffered with 5 mL of 5M NaCl and 7 mL of Tris pH 8.0 per 100 mL of supernatant and batch bound with Ni Sepharose Excel resin while shaking for 1 to 2 hours at room temperature. Resin were collected in a gravity column, washed with five column volumes of wash buffer containing 25 mM Tris pH 8.0, 250 mM NaCl, 200 mM L-Arginine and 20 mM imidazole, and eluted with three column volumes of elution buffer comprised of 25 mM Tris pH 8.0, 250 mM NaCl, 200 mM L-Arginine and 200 mM imidazole. Eluted component constructs were further purified using SEC on either a Superose 200 Increase 10/300 gel filtration column.

Western Blot

Protein samples were resolved on SDS-PAGE gels and subsequently transferred to nitrocellulose membranes using the Trans-Blot Turbo Transfer System (Bio-Rad). Gels were equilibrated in 1X Trans-Blot Turbo Transfer Buffer (diluted from 5X stock, stored at 4°C) prior to transfer. The transfer sandwich was assembled in the blotting cassette and all components were fully saturated with Transfer Buffer and carefully assembled to ensure a flat, bubble-free interface. Following transfer, membranes were blocked for 15–30 minutes at room temperature in TBS-T blocking buffer (10 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% Tween-20) supplemented with 3% non-fat dry milk (Cytiva). The same buffer was used for primary and secondary antibody dilutions.

For detection of His-tagged proteins, membranes were incubated with HRP-conjugated anti-His antibody (Thermo Fisher Scientific) diluted 1:20,000 in blocking buffer for 1 hour at room temperature. If non-HRP-conjugated primary antibodies were used, membranes were incubated with the primary antibody diluted in blocking buffer for 1 hour at room

temperature or overnight at 4°C on a tilting platform, followed by three washes in TBS-T. Secondary HRP-conjugated antibodies (anti-mouse or anti-human; Jackson ImmunoResearch) were used at 1:20,000 dilution and incubated under the same conditions as the primary. For ladder visualization, 2 µL of prestained molecular weight marker was loaded, and membranes were probed with HRP-conjugated Strep-Tactin antibody (IBA Lifesciences) at 1:50,000 dilution in blocking buffer.

After antibody incubation, membranes were washed three times with TBS-T (5–10 minutes per wash), then incubated in Clarity Western ECL substrate (Bio-Rad). The substrate was prepared by mixing equal volumes of the peroxide solution and luminol/enhancer solution.. Blots were imaged using a Gel Imager (Bio-Rad)

Thermal Denaturation (nanoDSF)
RBD-based samples were prepared in a buffer containing 50 mM Tris pH 8, 250 mM NaCl, 200 mM L-arginine, at 1.0 mg/mL for nanoDSF analysis. Non-equilibrium melting temperatures were determined using an UNcle (UNchained Labs) based on the barycentric mean of intrinsic tryptophan fluorescence emission spectra collected from 20–95°C using a thermal ramp of 1°C per minute. Melting temperatures were defined as the maximum point of the first derivative of the melting curve, with first derivatives calculated using GraphPad Prism software after smoothing with four neighboring points using 2nd order polynomial settings.

Circular dichroism
CD measurements were carried out on a JASCO J1500 spectrometer at 25–95°C, using a 1 mm path-length cuvette, at wavelengths from 200 to 260 nm. Proteins were measured at 0.4 mg/mL in TBS buffer.

Bio-Layer Interferometry
BLI was performed on an Octet RED96 or Octet R8. All biosensors were hydrated in kinetics buffer a made of the 2S5 purification buffer (25 mM Tris pH 8.0, 250 mM NaCl, 200 mM L-Arginine) with 0.05% v/v TWEEN-20 and 0.5% w/v non-fat dry milk (Cytiva). Antibodies were diluted to a final concentration of 0.01 mg/mL in kinetics buffer and 316189 or 2616 loaded onto either ProA tips or Anti-Mouse Capture biosensors (Sartorius) respectively. 2S5 was diluted in kinetics buffer and its association was measured for 300s, followed by a dissociation for 300s in kinetics buffer.

Enzyme-linked immunosorbent assay (ELISA) for Pam3
Peirce Streptavidin 96-well high protein binding plates were coated for one hour at room temperature at with biotinylated Pam3 (Invivogen) at 10 µg/mL and blocked in 100 µL of blocking buffer composed of (TBST: 1x Tris-buffered saline with 25 mM Tris pH 8.0, 150

mM NaCl, 0.2% Tween 20, and 5% nonfat milk). Plates were then washed 3x with TBST in an automated plate washer (Biotek); all washing steps follow the same protocol. 2-fold serial dilutions of 2S5 were made starting at 200ug/ml and 100 µL were plated and incubated for 1 h at room temperature shaking at 500 rpm. Plates were washed then 100 µL of a mouse anti-TLR2 antibody 2616 (RnD Biosystems) for 40 minutes shaking at 500 rpm. Secondary anti-mouse (Cell Signaling Technology) (1:2,000 dilution) IgG-HRP were added to each well and incubated shaking at room temperature. Plates were washed before 100 µL per well of TMB were added and developed for 3 min, then quenched with 100 µL of 1N HCl. Reading at absorbance at 450 nM was carried out with an Epoch plate reader (Biotek).

Enzyme-linked immunosorbent assay (ELISA) for Antibodies

Peirce Streptavidin 96-well high protein binding plates were coated for one hour at room temperature with biotinylated 2S5 at 2 µg/mL and blocked in 100 µL of blocking buffer composed of (TBST: 1x Tris-buffered saline with 25 mM Tris pH 8.0, 150 mM NaCl, 0.2% Tween 20, and 5% nonfat milk). Plates were then washed 3x with TBST in an automated plate washer (Biotek); all washing steps follow the same protocol. 3-fold serial dilutions of 2616 (RnD Biosystems) or 316189 (Abcam) were made starting at 100ug/ml and 100 µL were plated and incubated for 1 h at room temperature shaking at 500 rpm. Secondary anti-mouse (Cell Signaling Technology) (1:2,000 dilution) IgG-HRP or anti-rabbit (Cell Signaling Technology) (1:2,000 dilution) IgG-HRP were added to each well and incubated shaking at room temperature. Plates were washed before 100 µL per well of TMB were added and developed for 3 min, then quenched with 100 µL of 1N HCl. Reading at absorbance at 450 nM was carried out with an Epoch plate reader (Biotek).

Assembly of 2S5 I53-50 for mouse studies

In vitro assembly of nanoparticles was conducted using a 1:1 M ratio of each component and incubated at 4°C overnight with rocking. Following nanoparticle assembly, all particles were purified one last time through a Superose 6 Increase 10/300 column. 2S5 assembled nanoparticles were purified into 50 mM Tris pH 8.0, 500 mM NaCl, and 200mM L-Arginine.

Antigen-specific GC B cell probing by flow cytometry

To assess 2S5-specific GC B cell responses, fresh LN cell suspensions were stained with live/dead fixable blue viability dye (Invitrogen) followed by 400 ng of 2S5 probes in BV421 for 30 min at 4°C. Cells were subsequently stained with anti-human CD20 BV570 (2H7, Biolegend), and CD3 APC-Cy7 (SP34-2, BD Biosciences) for an additional 20 min at 4°C. Cells were permeabilized using the transcription factor buffer set (BD Biosciences) and stained intracellularly for anti-human IgG BV786 (G18-145, BD

Biosciences), BCL6 PE-Cy7 (K112-91, BD Biosciences), and Ki67 PE (B56, BD Biosciences). 2S5 probes were prepared by incubation of 4-fold molar excess of Avi-tag biotinylated 2S5 protein with streptavidin-conjugated BV421 (Biolegend).

Antigen-specific memory B cell sorting

Cryopreserved PBMCs collected two weeks after the third immunization (i.e., boost 2) were thawed at 37°C and washed twice with warm R10. Cells were stained with 400 ng of fluorescent 2S5-BV421 probes in 100 µl R10 for 20 min at 4°C. Following a wash with R10, cells were stained with a panel of antibodies to identify memory B cells in a total volume of 100 µl R10 for 30 minutes at room temperature. The panel contained anti-human CD3 BV510 (SP34-2, BD Biosciences), CD14 BV510 (M5E2, Biolegend), CD16 BV510 (3G8, BD Biosciences), CD56 BV510 (B159, BD Biosciences), CD19 ECD (J3-119, Beckman Coulter), CD20 BV605 (2H7, Biolegend), IgD FITC (polyclonal, Southern Biotech), and IgG BV786 (G18-145, BD Biosciences). After another wash with R10, cells were resuspended in approximately 1 ml R10 with 7 aminoactinomycin D (7-AAD; Invitrogen) at a dilution of 1:4,000 and kept on ice before acquisition on a BD Aria III Fusion cell sorter. RSV-F-specific memory B cells (Lymphocytes/Singlets/Live/CD3- CD14- CD16- CD56-/CD20+ CD19+/IgD- IgG+/2S5+ PostF- and 2S5+; see also Fig. 4b) were index-sorted into 96-well plates, immediately placed on dry ice and frozen at -80°C. 2S5 was prepared by incubation of 4-fold molar excess of Avi-tag biotinylated protein with streptavidin-conjugated BV421 (Biolegend) or streptavidin-conjugated APC (Biolegend).

Single BCR amplification and sequencing

cDNA was retrieved in one step by single cell mRNA reverse transcription. Reverse transcription mix containing 5 µl 5× First Strand buffer (Invitrogen), 4.81 mM DTT (Invitrogen), 0.77 mM dNTPs (Invitrogen), 0.24 % IGEPAL CA-630 (Sigma-Aldrich), 150 ng random hexamers (Invitrogen), 20 U RNaseOUT (Invitrogen), 100 U SuperScript III Reverse Transcriptase (Invitrogen) and nuclease-free water in a total volume of 26 µl was added to single sorted cells. Reverse transcription was performed by initial 10 min incubation at 42°C, 10 min at 25°C, 60 min at 50°C, 5 min at 94°C and final cooling to 4°C. The plates with cDNA were stored at -80°C.

Variable regions from antibody heavy, lambda, and kappa chains were amplified by nested PCR protocols adapted from Sundling et al[92]. where forward and reverse primers anneal to the leader sequence and constant region, respectively, enabling amplification of the whole VDJ segment. Briefly, 1st PCR master mix contained 2.6 µl 10× Hot Star PCR buffer (Qiagen) with 15 mM MgCl2, 0.25 mM dNTPs, 1.50 mM MgCl2 (Qiagen), 0.10 µM 5'-primer mix (primers were supplied at 200 µM each and mixed at equal volumes), 0.10 µM 3'-primer and 1.3 U HotStarTaq Plus DNA polymerase (Qiagen). The

master mix was filled to 23.7 µl with nuclease-free water and added 2.3 µl of the template cDNA. The 2nd PCR master mix consisted of 2.6 µl 10× Hot Star PCR buffer, 0.25 mM dNTPs, 5.2 µl 5× Q-solution (Qiagen), 0.63 µM 5'-primer mix (primers ware supplied at 200 µM each and mixed at equal volumes), 0.63 µM 3'-primer and 1.3 U HotStarTaq Plus DNA polymerase. Nuclease-free water was filled to 23.7 µl and 2.3 µl of the 1st PCR product was used as a template. The PCR amplification program for both rounds of nested PCR was performed as follows: 5 min at 95°C, 50 cycles of 30 s at 95°C, 30 s at 50°C and 55 s at 72°C, then 10 min at 72°C and cooling to 4°C. After the 2nd PCR, selected wells were tested on a 2 % agarose gel, with expected product size ~400 bp. PCR products were sequenced using Sanger sequencing by Genewiz, Leipzig, Germany.

Antibody binding affinity studies

BLI was performed on an Octet RED96 or Octet R8. All biosensors were hydrated in kinetics buffer HBS-EP+ (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% v/v surfactant P20) with 0.5% w/v non-fat dry milk (Cytiva). Antibodies were diluted to a final concentration of 0.01 mg/mL in kinetics buffer and loaded onto ProA biosensors (Sartorius) for at least 10 minutes prior to use. 2S5 or hTLR2 were diluted in kinetics buffer and its association was measured for 300 s, followed by a dissociation for 300 s in kinetics buffer.

Epitope binning BLI competition studies

Epitope binning was performed using biolayer interferometry (BLI) on the Octet R8 instrument (Sartorius) to assess potential competitive binding between monoclonal antibodies and the target protein, 2S5. All steps were conducted at 25°C in kinetics buffer HBS-EP+ (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% v/v surfactant P20) with 0.5% w/v non-fat dry milk (Cytiva) to minimize non-specific interactions. Antibodies were diluted to a final concentration of 0.01 mg/mL in kinetics buffer and loaded onto ProA biosensors (Sartorius) for at least 10 minutes prior to use.

Monoclonal antibody 1 (mAb1) was loaded onto the biosensor tips at a concentration of 0.01 mg/mL for 120 seconds. Following a brief baseline step of 30 seconds, sensors were exposed to an excess of 2S5 at 0.02ug/mL for association, allowing complex formation between mAb1 and 2S5. After association equilibrium was reached, another baseline step of 30 seconds was performed to remove unbound analyte.

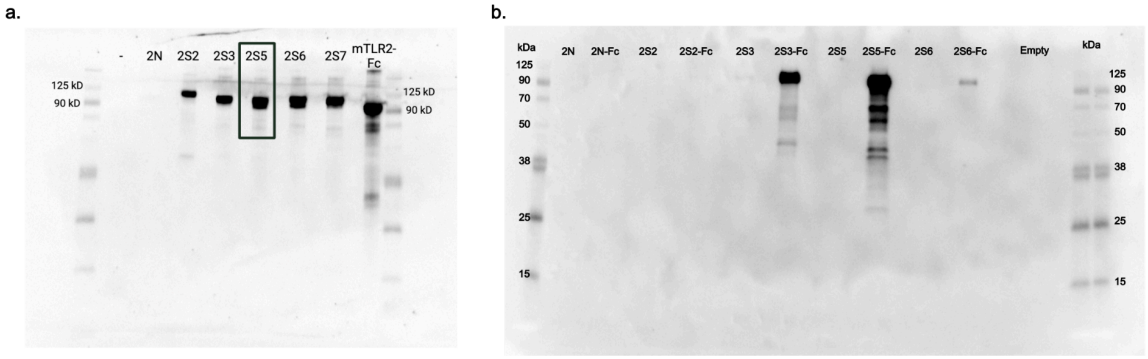
To assess epitope competition, monoclonal antibody 2 (mAb2) was then introduced at 0.01 mg/mL to determine whether it could bind the mAb1–2S5 complex. A positive binding signal indicated that mAb2 recognized a distinct, non-overlapping epitope on

2S5, while the absence or significant reduction in signal suggested competitive or steric interference between mAb1 and mAb2.

Data were reference-subtracted and aligned using Octet Data Analysis software (Sartorius), and binding profiles were interpreted to determine epitope overlap or uniqueness.

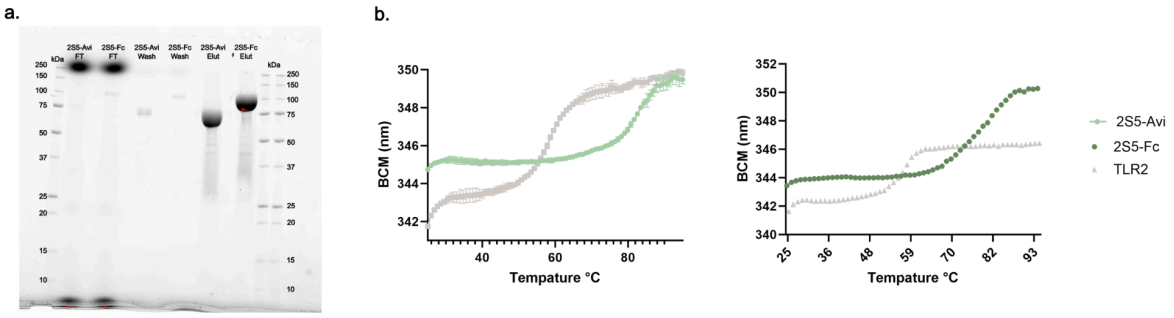
Negative Stain Electron Microscopy

400-mesh carbon coated grids (Electron Microscopy Sciences) were glow-discharged and 3 uL of TLR-FC at 20 ug/mL was applied and then stained with 2% (w/v) uranyl formate. All nsEM data were collected with a BM-Ceta camera at 57,000x magnification using EPU 2.0 on a 120 kV Talos L120C transmission electron microscope (Thermo Scientific). CTF processing, particle picking, particle extraction, 2D classification, and 3D refinement steps were all performed with CryoSPARC[93].

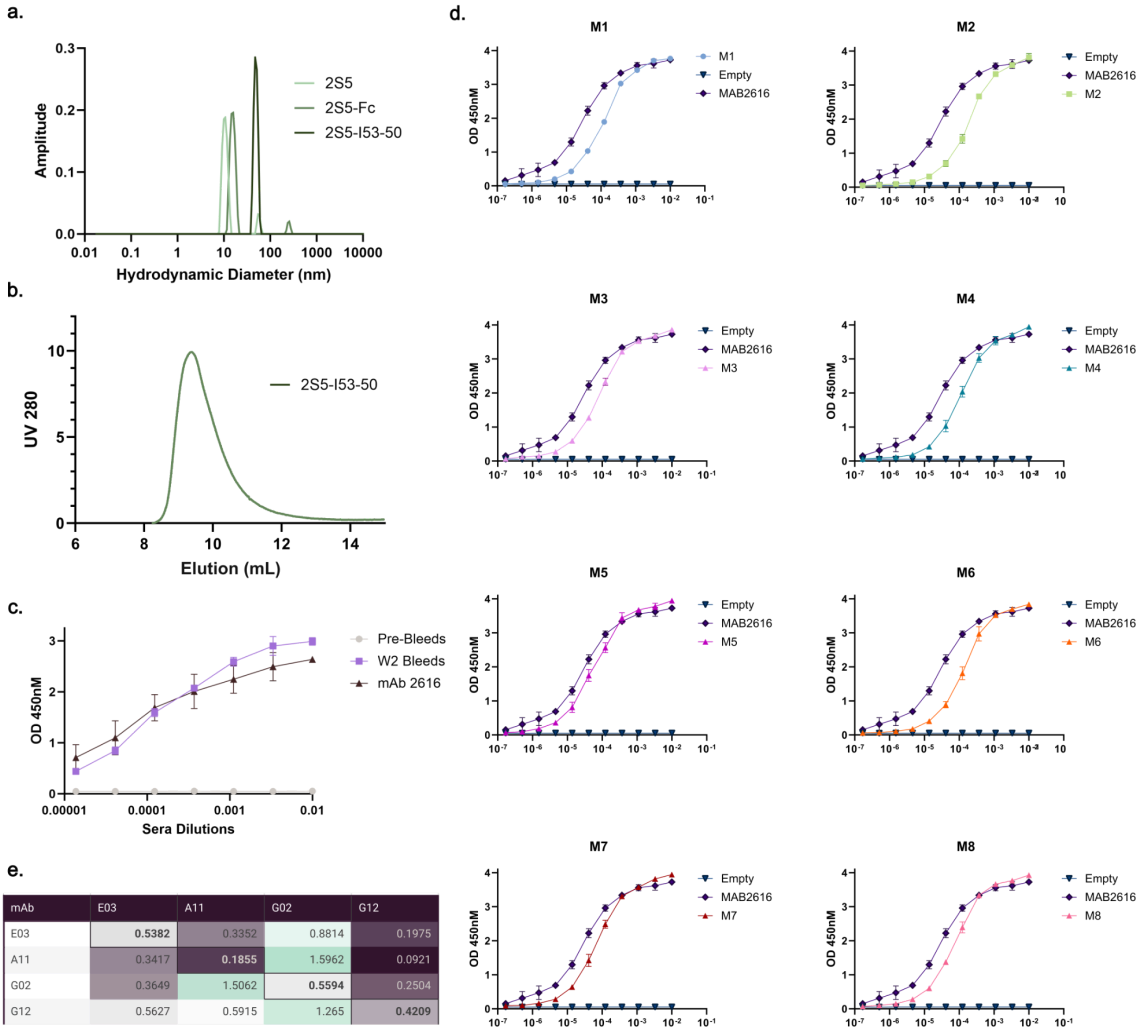


Extended Data Fig. 1: Western blot of unconstrained and constrained stabilized TLR2.

a, Western blot depicting 3mL screen of stabilized TLR2 designs. The "-" well is an empty vector, 2N is hTLR2-Fc, mTLR2-Fc was purchased from a supplier and serves as a control for size. **b,** Western blot of 3mL screen of constrained designs with and without Fc fusions.



Extended Data Fig. 2: Expression and stability screen of 2S5-Avi and 2S5-Fc.
a, SDS-PAGE gels of 250mL Flow Through, Wash and Elution of 2S5–Avi and 2S5-Fc prior to SEC purification. **b,** Raw DFS data prior to first derivative and smoothing.



mAb	E03	A11	G02	G12
E03	**0.5382**	0.3352	0.8814	0.1975
A11	0.3417	**0.1855**	1.5962	0.0921
G02	0.3649	1.5062	**0.5594**	0.2504
G12	0.5627	0.5915	1.265	**0.4209**

Extended Data Fig. 3: 2S5-I53-50 mouse ELISAs and epitope binning of cross reactive antibodies.

a, DLS data of 2S5-I53-50 compared to 2S5-Fc and 2S5-Avi (seen in Fig 3b). **b,** SEC trace of 2S5-I53-50 showing single peak at correct elution volume. **c-d,** Initial two week bleeds mice indicate development of high affinity antibodies against 2S5. Panel c is the pool of each mouse represented in panel d. **e,** Table of epitope binning for the cross reactive species of mAbs.

Supplementary Table 1. Amino acid sequences of stabilized TLR2

Construct	Amino Acid Sequence
hTLR2 AlphaFold2 Model	SLSCDRNGICKGSSGSLNSIPSGLTEAVKSLDLSNNRITYISNSDLQR CVNLQALVLTSNGINTIEEDSFSSLGSLEHLDLSYNYLSNLSSSWFK PLSSLTFLNLLGNPYKTLGETSLFSHLTKLQILRVGNMDTFTKIQRKD FAGLTFLEELEIDASDLQSYEPKSLKSIQNVSHLILHMKQHILLLEIFV DVTSSVECLELRDTDLDTFHFSELSTGETNSLIKKFTFRNVKITDESL FQVMKLLNQISGLLELEFDDCTLNGVGNFRASDNDRVIDPGKVETL TIRRLHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLSQHLKS LEYLDLSENLMVEEYLKNSACEDAWPSLQTLILRQNHLASLEKTGE TLLTLKNLTNIDISKNSFHSMPETCQWPEKMKYLNLSSTRIHSVTGCI PKTLEILDVSNNNLNLFSLNLPQLKELYISRNKLMTLPDASLLPMLLV LKISRNAITTFSKEQLDSFHTLKTLEAGGNNFICSCEFLSFTQEQQAL AKVLIDWPANYLCDSPSHVRGQQVQDVRLSVSECHRTSGGHHHH HH
2S2-Fc Unconstrained	SLVCDSNGVCTATDGSRTAIPKGLTDSVKELNLSNNKISYISKSDLAN ATNLEKLNLSNNNISYIEDDAFANLTKLKVLDLSSNKLTKLSKAWFEP LTSLEELNLLDNPYKSFGDEPLFKNLTKLEKLSVGNKTNFEEIRKKDF EGLTKLKELTLDAYNLKKYEPGALSVIPNVEKVIIRMKNYDHFLTMLE DIKGNVKELEIEDMDLTDFVFKPLKSSSKNSKIKKVVFKRVKITDYSF REILKVLLTLSGLLELEFDDATLNGVGNFRASDNDRVLDPGKVETLTI RRLHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLAQHCKSL EYLDLSENLMVEEYLKNSACEDAFPSLQTLILRQNHIKDLKKLGETL LPLKSLEKLDVSGNPFKSIPSSAKWPSKLQHLNLSSNGLTELTNAIP SSLKELDVSNNKLTSVTLPLPNLQKLYISGNQLTSLPNASNYPNLQE LVASNNKIKSFTKAEIDAYKNLKKLRVGGNQWDCSCELLDLVQTNPE LEKIIEGYPEELLCASPEKYKGQRIRDVVLSEEECNAESGGHHHHH H

2S3-Fc Unconstrained	SLECDASGVCTATDGSLTAIPAGLTSSVTHLDLSNNQISYISASDLAN ATNLETLILSHNQISHIEDNAFSNLTNLKHLDLSSNQLTTLSPAWFAPL TSLEHLDLTDNPYTSFGDAPLFKNLKNLKVLHIGNKTNFKEIRKKDF EGLTKLKELHIDAYSLEKYEKGALSVIPDVEKVILRMKNSDHLLTIFED IKDNVKELEIEDMDLTNFKFKKLKSSSKNSKIEKVRFKRVKMTDYSF RELLKLLSTLSGLLELEFDDATLNGVGNFRASDNDRVIDPGKVETLTI RRLHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLAQHCKSL EYLDLSENLMVEEYLKNSACEDAFPSLQTLILRQNHIKSLKKLGETLL PLKSLEHLDVSGNPFKELPETVKWPKSLKVLNLSSNGLTKLTNAIPK SLEVLDVSNNQLTSITLSLPNLKELYASGNQLTSLPDPSLFPNLEVLV VSNNQIKSFTKEEIEAYKNLKELEIGGNKYDCSCEFLDLYQSNPELE KIIKGYPEKLLCASPEKYKGQPVRDVVLSEEECNKESGGHHHHHH
2S5-Fc Unconstrained	SLECDDEGKCTATDGTLTAIPKGLTDSVTELDLSNNKISHISKTDLSN ATNLEKLILSNNEISYIEDDAFSNLTNLKVLDLSSNKLTSLKPSWFEPL TSLEYLDLLDNPYTSFGSSPLFKNLTNLKVLKVGNKTNFKTLTKKDF EGLTKLKELELDAYNLQSYEPGALSVIPDVEKFILRMKNSDHLLTILE DIKSNVKHLEIEDMDLTNFKFEKLKSKDKNSKIKKVVFKRVKITDYSF RQILKLLNTLSGLLELEFDDATLNGVGNYRASDNDRVIDPGKVETLTI RRLHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLAQHCKSL EYLDLSENLMVEEYLKNSACEDAFPSLQTLILRQNHITDLKKLGETLL PLKSLTHLDVSHNPFKELPESCEWPSSLQHLNLSHCGLTKLTNAIPS SLKTLDVSHNQLTSFTLSLPNLQELYISHNQLTSLPDPSLLPNLQTLV ASHNKIKSFTKEEIDAYKNLKKLEVTGNQYDCSCELLDLIQSNPELK EIIVGYPEGLLCATPEKYKGQRITDVVLSEEECNKESGGHHHHHH
2S6-Fc Unconstrained	SLSCDSSGVCTATDGTLTAIPKGLTSSVTSLDLSNNKISYISSEDLKN ATNLKTLLLSNNQISHIEDDAFANLTNLQVLDLSSNKLTKLSPSWFAP LTSLKYLNLENNPYKTTGDAPLFKNLTNLQHLSIGNDTNFKKIGKKDF EGLTNLKELTINAYALEEYEKGALSVIPDVEKFILRMKNADHLLTILEDI KSNVKELEIEDMDLTNFKFKKLSSKDKDSKIEKVRFKRVKITDYSFRE ILKVLSTLSGLLELEFDDATLNGVGNFRASDNDRVIDPGKVETLTIRR LHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLAQHCKSLEY LDLSENLMVEEYLKNSACEDAFPSLQTLILRQNHITDLKKLGETLLPL KSLEHLDVSHNPIKEIPEECKWPPSLKYLNLSSNGITKLTNAIPKSLE VLDVSHNKLKEVTIELPNLKELYIDHNELTELPDAKLYPNLEVLIASHN KIKKFTKEEIEAYKNLKELEIGGNEFDCSCEFLDLYQTNPELKKIIKGY PEKLLCATPEKYKGQRIQDVVLSDEECFKESGGHHHHHH

2S2-Fc SLSCDRNGNCKKTDGSLTSIPKGLTEAVKSLDLSNNRITYISNSDLQ
RATNLKKLVLSHNGISTIEPDSFSSLGSLEHLDLSYNYLSNLSSSWF
KPLSSLKFLNLLENPYKTLGETPLFSHLTKLQILRVGNMDTFTKIQRK
DFEGLTSLEELEIDAYNLQSYEPGALKSIQNVSHLILHMKQHILFLEM
LVDVKGSVEELEIRDVDLDTFHFSELSTGETNSLIKKVTFRNVKITDE
SLFQIFKLLNTFSGLLELEFDDATLNGVGNFRASDNDRVIDPGKVET
LTIRRLHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLAQHAK
SLEYLDLSENLMVEEYLKNSACEDAFPSLQTLILRQNHIASLEKLGE
TLLNLKNLTNLDVSGNPFHSMPETCQWPEKLKYLNLSSNNLHSLTG
AIPKTLEILDVSGNNLNLFSLNLPNLKELYISSNKLTTLPDASLFPMLL
KLVASRNAITTFSKEQVDAYHTLKELEAGGNNFICSCEFLSFTQEQQ
ALAKVLIDWPANYLCDSPSHVRGQQVQDVRLSVSECHRTSGGHHH
HHH

2S3-Fc SLSCDRNGHCVATDGSLTSIPKGLTESVTHLDLSNNRISYISNSDLQ
RATNLKALVLSHNRISTIEPNAFSSLTNLEHLDLSYNYLSNLSSSWFK
PLSSLKFLNLLENPYKTLGETPLFSHLTKLQILRVGNMDTFTKIQRKD
FEGLTSLEELEIDAYNLQSYEPGALSVIPNVSHLILHMKQHILLLEIFV
DVKDSAEELEIRDMDLDTFHFSELSTGETNSLIKKVTFRNVKTTDES
FRNILKLLNTLSGLLELEFDDATLNGVGNFRASDNDRVIDPGKVETL
TIRRLHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLAQHLKS
LEYLDLSENLMVEEYLKNSACEDAFPSLQTLILRQNHIASLEKLGETL
LNLKNLTNLDVSGNPFHSMPETCQWPEKLKYLNLSSNNLHSLTGAI
PKTLEILDVSNNNLNLFSLNLPQLKELYASGNKLTTLPDPSLFPMLLV
LVVSRNAITTFSKEQIDAYHTLKKLEAGGNNYDCSCELIKTIKNNPEL
KEILVNYPDKYLCASPEEYKGKRITEVELSEEECNKESGGHHHHHH

2S5-Fc SLSCDRNGKCTARDGSLTSIPKGLTESVTELDLSNNRISHISNSDLQ
RATNLKKLVLSHNKISTIEPDAFSSLTNLEVLDLSYNYLSNLSSSWFK
PLSSLKFLNLLENPYKTLGETPLFSHLTKLQILRVGNMDTFTKIQRKD
FEGLTSLEELEIDAYNLQSYEPGALSVIPNVSHLILHMKQHILLLEIIVD
VQSSVEHLEIRDMDLDTFHFSELSTGETNSLIKKVTFRNVKITDESL
RSIFKLLNSFSGLLELEFDDATLNGVGNFRASDNDRVIDPGKVETLTI
RRLHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLAQHLKSL
EYLDLSENLMVEEYLKNSACEDAFPSLQTLILRQNHIASLEKLGETLL
NLKNLTNLDVSGNDFHSMPETCQWPEKLKYLNLSSNGLHSLTGCIP
KTLEILDVSNNNLNLFSLNLPQLKELYISGNKLMTLPDPSLLPMLLVL
KASRNAITTFSKEQIDAYHTLKELEAGGNNYDCSCELIDTVQSNPEL
KEILVDYPDKYLCASPPKYKGKLLHEVTLSEEECNKPSGGHHHHHH

2S6-Fc SLSCDRNGRCVATDGKLTSIPKGLTESVVSLDLSNNRISYISNSDLQ
RATNLKVLVLSHNRISTIEPDAFSSLTNLEVLDLSYNYLSNLSSSWFK
PLSSLTFLNLLGNPYKTLGETPLFSHLTKLQILRVGNMDTFTKIQRKD
FAGLTFLEELEIDASDLQSYEPGALSVIPNVSHLILHMKQHILLLEIIVD
VTSSVEELEIRDVDLDTFHFSELSTGETNSLIKKVTFRNVKITDESLR
SIFKLLNTLSGLLELEFDDATLNGVGNFRASDNDRVIDPGKVETLTIR
RLHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLAQHAKSLE
YLDLSENLMVEEYLKNSACEDAFPSLQTLILRQNHIASLEKLGETLL
NLKNLTNLDISGNPIHSMPETCQWPEKLKYLNLSSNNIHSLTGCIPKT
LEILDVSNNNLNLFSLNLPQLKELYISGNKLMTLPDASLFPMLLVLKA
SRNAITTFSKEQVDAYHTLKRLEAGGNNFICSCEFLSFTQEQQALAK
VLIDWPANYLCDSPSHVRGQQVQDVRLSVSECHRTSGGHHHHHH

2S5-Avi SLSCDRNGKCTARDGSLTSIPKGLTESVTELDLSNNRISHISNSDLQ
RATNLKKLVLSHNKISTIEPDAFSSLTNLEVLDLSYNYLSNLSSSWFK
PLSSLKFLNLLENPYKTLGETPLFSHLTKLQILRVGNMDTFTKIQRKD
FEGLTSLEELEIDAYNLQSYEPGALSVIPNVSHLILHMKQHILLLEIIVD
VQSSVEHLEIRDMDLDTFHFSELSTGETNSLIKKFTFRNVKITDESLF
QVMKLLNSFSGLLELEFDDATLNGVGNFRASDNDRVIDPGKVETLTI
RRLHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLAQHLKSL
EYLDLSENLMVEEYLKNSACEDAWPSLQTLILRQNHIASLEKLGETL
LNLKNLTNLDVSGNDFHSMPETCQWPEKLKYLNLSSNGLHSLTGCI
PKTLEILDVSNNNLNLFSLNLPQLKELYISGNKLMTLPDPSLLPMLLV
LKASRNAITTFSKEQIDAYHTLKELEAGGNNYDCSCELIDTVQSNPE
LKEILVDYPDKYLCASPPKYKGKLLHEVTLSEEECNKPLVPRGSGSG
GLNDIFEAQKIEWHESGGHHHHHH**

2S5-Fc SLSCDRNGKCTARDGSLTSIPKGLTESVTELDLSNNRISHISNSDLQ
RATNLKKLVLSHNKISTIEPDAFSSLTNLEVLDLSYNYLSNLSSSWFK
PLSSLKFLNLLENPYKTLGETPLFSHLTKLQILRVGNMDTFTKIQRKD
FEGLTSLEELEIDAYNLQSYEPGALSVIPNVSHLILHMKQHILLLEIIVD
VQSSVEHLEIRDMDLDTFHFSELSTGETNSLIKKFTFRNVKITDESLF
QVMKLLNSFSGLLELEFDDATLNGVGNFRASDNDRVIDPGKVETLTI
RRLHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLAQHLKSL
EYLDLSENLMVEEYLKNSACEDAWPSLQTLILRQNHIASLEKLGETL
LNLKNLTNLDVSGNDFHSMPETCQWPEKLKYLNLSSNGLHSLTGCI
PKTLEILDVSNNNLNLFSLNLPQLKELYISGNKLMTLPDPSLLPMLLV
LKASRNAITTFSKEQIDAYHTLKELEAGGNNYDCSCELIDTVQSNPE
LKEILVDYPDKYLCASPPKYKGKLLHEVTLSEEECNKPLVPRGSGSG
SGGSGGSEPKSSDKTHTCPPCPAPELLGGPSVFLFPPKPKDTLMIS

	RTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNS TYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPR EPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENN YKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNH YTQKSLSLSPGKGSGGLNDIFEAQKIEWHESGGHHHHHH
2S5-50A	SLSCDRNGKCTARDGSLTSIPKGLTESVTELDLSNNRISHISNSDLQ RATNLKKLVLSHNKISTIEPDAFSSLTNLEVLDLSYNYLSNLSSSWFK PLSSLKFLNLLENPYKTLGETPLFSHLTKLQILRVGNMDTFTKIQRKD FEGLTSLEELEIDAYNLQSYEPGALSVIPNVSHLILHMKQHILLLEIIVD VQSSVEHLEIRDMDLDTFHFSELSTGETNSLIKKFTFRNVKITDESLF QVMKLLNSFSGLLELEFDDATLNGVGNFRASDNDRVIDPGKVETLTI RRLHIPRFYLFYDLSTLYSLTERVKRITVENSKVFLVPCLLAQHLKSL EYLDLSENLMVEEYLKNSACEDAWPSLQTLILRQNHIASLEKLGETL LNLKNLTNLDVSGNDFHSMPETCQWPEKLKYLNLSSNGLHSLTGCI PKTLEILDVSNNNLNLFSLNLPQLKELYISGNKLMTLPDPSLLPMLLV LKASRNAITTFSKEQIDAYHTLKELEAGGNNYDCSCELIDTVQSNPE LKEILVDYPDKYLCASPPKYKGKLLHEVTLSEEECNKPLVPRGSGSG GSGGSGSGSGSGSGSEKAAKAEEAARKMEELFKKHKIVAVLRANS VEEAIEKAVAVFAGGVHLIEITFTVPDADTVIKALSVLKEKGAIIGAGTV TSVEQCRKAVESGAEFIVSPHLDEEISQFCKEKGVFYMPGVMTPTE LVKAMKLGHTILKLFPGEVVGPQFVKAMKGPFPNVKFVPTGGVNLD NVCEWFKAGVLAVGVGSALVKGTPDEVREKAKAFVEKIRGCTEHH HHHH

Supplementary Table 2. Amino acid sequences of mAbs

Construct	Heavy Chain	Light Chain
D07	QVQLQESGAELAKPGASVKMSCKASG YTFTDYWIHWVRQRPGQGLEWFGYINP RTDYSEYNQEFKDKATFTADKSSNTAY MQLSSLTSEDSAVYFCTRYQLERYWYF DVWGAGTTVTVSS	NIVMTQSPKSMSMSVGERVTLTC KASQNVVSYVAWYQQKPGQSPK LLIYGASNRYTGVPDRFTGSGSA TDFTLTISSVQAEDLADYHCGQT YRYPYTFGGGTKLEIK
F05	EVQLQESGPELVRPGASVKIPCKASGY TFTAYNMDWVKQSHGKSLDWIGYINPN NGGTIYNQKFKGKATLTVDKSSSTAFME LRSLTSEDTAVYYCARQLGLREDYWGQ GTTLTVSS	DIQMTQFPASLSVFVGETVTITCR ATENIYSNLAWYQQKQGLSPQLL VYAATNLADGVPSRFSGSGSGT QFSLRINSLQSEDFGNYYCQHFW GTPRTFGGGTKLEIK
F07	QVQLQESGPQLDRPGASVKISCKASGY	QIVLTQSPAIMSAFPGERVTMTCS

	SFTSYWMHWVIQRPGQGLQWIGMIDP SDSETRLNQKFTDKAILTVDRSSSTAYM QLRSPTFQDSAVYYCARGNFRYDDGD MDYWGQGTSVTVSS	ASSNFSYMHWFQQRSDTSPKR WIYDTSRLPSGVPARFSGSGSGT SYSLTISSMEAEDVATYYCHQRS CYPYTFGGGTKLEIK
B01	EVQLQESGPELVKPGTSVKISCKASGYS FSGYFMNWVMQTRGKSLEWIGHINPYN GNIFYNPNFKGKATLTVDKSSDTAHMEL RSLASEDSAVYYCTRSSQSLLRFFDYW GHGTTLTVSS	DIQMTQTTSSLSASLGDRVTISCR ASQDISNYLNWYQQKPDGTVKLL ISYTSTLHSGVPSRFSGSGSGTD FSLTISNLEQEDVATYFCQQGNTL PFTFGSGTKLEIK
D01	QVQLQESGPQLVRPGSSVKISCKASGD SFTSYWIHWVKQRPGQGLEWIGMIDPS DSETRLNQKFKDKATLTVDKSSSTAYMQ LSSPTSEDSAVYYCARDYYYSNLDSWG HGTSVTVSS	DIQMTQTPSSLSASLGERVSLTC RASQDIGSSLNWLQQEPDGTIKR LIYATSNLDSGVPKRFSGSRSGS DYSLTISSLESEDFVDYYCLQYSS SPYTFGGGTKLEIK
E03	EVQLQESGPDLVKPGASVKISCKASSD SFTAYYMHWMKQSHGKSLEWIGRVHP NNGGTFYNQKFKDKAILTVDKSSSTAYV ELRSLTSEDSAVYYCAPTTTAWFAFWG QGTLVTVSA	DVQITQTPSYLAASPGETITINCR ASKRISKYLAWYQEKPGKTNKLLI YSGSTVQSGIPSRFSGSGSGTDF TLTISSLEPEDFAMYYCQQYNEY PYTFGGGTKLEIK
E04	DVQLQESGPDLMKPSQSLSLTCTVTGY SITSGNSWLWIRQFPGNKLEWMGYIHY SGITNYNPSLKSRISITRDTSKNQFFLQL NSVTTEDTATYYCARGTVSYGDYWGQ GTTLTVSS	DIQMTQTTSSLSASLGDRVTISCR ASQDIGNYLNWYQQKPDGTVKLL IYSTSRLHSGVPSRFSGSGSGTD YSLTISNLEQEDIATYFCQQGDTL PPTFGAGTKLELK
F11	DVQLQESGPGLVKPSQSLSLTCSVTGY SITSGYHWNWIRQFPGNKLQWMGYIRY DGSNYYNPSLRNRISITRDTSKNQVFLR LNSVTADDTSTYYCATSRRDYFDFWGQ GTTLTVSS	ENVLTQSPAIMSASPGEKVTMTC SAISSVSYMHWYQQKSGTSPKL WIYDTSKLTSGVPGRFSGSGSGH SYSLTISSVEAEDVATYYCFQGSG YPQTFGGGTKLEIK
F04	EVQLQESGPDLVKPGASVKISCKASSD SFTGYYMHWMKQSHGKSLEWIGRVNP NNGGTFYNQKFKDKAILTVDKSSSTAYV ELRSLTSEDSAVYYCAPTTTAWFVFWG QGTLVTVSA	DVQITQSPSYLAASPGETITINCR ASKRISKYLAWYQEKPGKTNKLLI YSGSTVQSGIPSRFSGSGSGTDF TLTISSLEPEDFAIYYCQQYNEYP YTFGGGTKLEIK
F07_2	QVQLQQSGPDLVKPGTSVRISCKASDY TFTRYYIHWVKQRPGQGLEWIGWINPG NFDTKFNEKFKGKATLTADKSSSTAHMQ LSTLTSEDSAVYFCTREGASGIWYFDV	ENVLTQSPAIMSASPGEKVTMTC SASSSVSYMHWYQKKSTTSPKL WIYDTSKLASGVPGRFSGSGSG YSFSLTISSMEAEDVATYYCFQGS

	WGAGTTVTVSP	GYPLTFGAGTQLELK
G02	QVQLQESGPHLVRPGTSLKISCKASGY SFTSYWIHWVKERPGQGLQWIGMIDPS DSETRLNQKFWDTATLTVDISSSIAYMQ LRSPTSDDSAVYYCARGGYRYDDFVM DYWGQGTSVTVSS	DIVMTQAVPSVPVTPGESVSISCR SSKSLLRSNGNTYLSWFLQRPG QSPQLLIYRMSNLASGVPDRFSG SGSGSVFTLRISRVAAEDVGIYYC MQRLDYPYTFGGGTKLEIK
A11	QVQLQESGPELVRPGESVKISCKGSGY RFTDYAMHWVKQSHAKSLEWIGVISFY YDNTNYNQKFKGKATMTVDKPSSTAYM ELARLTSEESAIYYCARSTNYGNSLWFF DVWGAGTTVTVSS	DIKMTQSPSSMYASLGERVTITCK ASQDINYYLSWFQQKPGKSPKTL IYRVKRLVDGVPSRFSGSGSGQD YSLTISSLEYEDMGTYYCLQYDEL PYTFGGGTKLEIK
B06	QVQLQESGAELAKPGASVKMSCKASG YIFTNYWMHWVKQRPGQGLEWIGNINP SSGNIEFNQRFKDKATLTADKSSSTASM QLSSLTSEDSAVYYCAYDYAYAMEYWG QGTSVTVSS	QIVLTQTPAIMSASPGEKVTISCS ASSSVSYMYWFQQKPGSSPKP WIYRTSNLAYGVPARFSGSGSGT SYSLTISSMEAEDAATYYCQQYH SYPYTFGGGTKLEIK
C10	QVQLQESGAELAKPGASVKMSCKASG YTFTSYWMHWVKQRPGQGLEWIGYIN PSTGYTEYNQKFKDKATLTADKSSSTAY MQLSSLTSEDSAVYYCARGALWSFAYW GQGTLVTVSA	DIVMSQSPSSLAVSAGEKVTMSC KSSQSMLNSRTRKNYLAWYQQK PGQSPKLLIYWASTRESGVPDRF TGSGSGTDFTLTISSVQAEDLAVY YCKQSYNLFTFGSGTKLEIK
C02	EVQLQESGPDLVKPGTSMKISCEASGY SFIDYNMEWVKQSHGKNLEWIGLINPY NAITLYSQKFKGKAIFTVDKSSSTAYMEL LSLTSEDSAVYYCARLGKLGLQETLDY WGQGISVTVSS	QFVLTQSPTIMSASPGEKVTMSC SASSSLSYMHWYQQKPGSSPKP WIYRTSNLASGVPSRFSGSGSGT SYSLTISSMEAEDAATYYCQQYH SYPPTFGAGSKLELK
D01_2	QVQLQESGPQLVRPGASVKISCKASAD SFSTNWMHWVKQRPGQGLEWIGMIDP SDSEIRLNQKFKDKATLTVDKSSSTAYM QLNSPTSEDSAVYYCARDNHLAWLTFW GQGTLVAVSA	DIQMTQSPSSLSASLGERVSLTC RASQDIGTSLIWLQREPDGTIKRL IYATSTLDSGVPKRFSGSRSGSD FSLTISSLESEDFVDYFCLQYSGS PYTFGGGTKLEIK
D04	EVQLQESGPELVKTGASVKISCKASGYS FTSYYMHWVKQSHGKSLEWIGFISCYN GATIYNQKFKDRATFTVDTSSSTAYMQF NSLTSEDSAVYFCARKEGGNHAMDYW GQGTSVTVSS	DIQMTQTTSSLSASLGDRVTISCR ASQDISNYLNWYQQKPDGTVKLL IYYTSRLHSGVPSRFSGSGSGTD YSLTISNLEQEDIATYFCQQGNTL PFTFGSGTKLEIK
G12	EVQLQESGPEVVKPGASVKISCKSSGY	DVQITQSPSYLAASPGESITINCR

	SFTGYYIHWVKQSHVKSLEWIGRVNPY NGVTWYNQIFKDKASLTVDKSSTTAYME LHSLTSEDSAVYYCARDYDSSLYAMDY WGQGTSVTVSS	TSKRISKFLAWYQEKVGKTNKLLI YSGSTLQSGIPSRFSGSGSGTDF TLTISGLEPEDFAMYYCQQHNEF PWTFGGGTQLEIK
H10	EVQLQQSGPELVKPGASVNISCKASGY TFTAFNMQWVKQSHGKSLEWIGYIYPY SGITGQNQKFKTKATLTVDTSSSTAYME LRSLTFEESAVYYCARSSYGSDFDYWG QGTTLTVSS	DIKMTQSPSSMYASLGERVTITCK ASQDINNYLGWFQQKPGKSPKTL IYRANRLVDGVPSRFSGSGSGQD YSLTISSLDYEDMGFYYCLQYDE FPPTFGSGTKLEIK
B09	QVQLQESGAELAKPGASVKMSCKASG YTFTSYWMHWVKQRPGQGLEWIGYIN PSSGYTEYNQKFKDKATLTADKSSSTAY MQLSSLTSEDSAVYYCARLRLYTMDYW GQGTSVTVSS	DIVMTQSPSSLAVSAGEKVTMICK SSQSLLNSRTRKNYLAWYQQKP GQSPELLIYWASTRESGVPDRFT GSGSGTDFTLTISSVQAEDLAVYY CKQSYNLYTFGGGPSLEIK
C08	QVQLQESGPEVVKPGASVKISCKTSGS AFSTSWMNWVKQRPGQGLEWLGRIYP GDGDTLYNGKFKGKATLTADKSSSTAFI QFNSLTSVDSAVYFCARSELRQPYYFD SWGQGTSLTVAS	DIVMTQSHRFMSTSVGDRVSITC KAGQDVDTAVAWYQQKPGQSPK LLIYWASSRHTGVPNHFTGSGSG TDFTLTISNVQSEDLADYFCQQY NNYPYTFGGGTKLEIK
E05	QVQLQESGPGLVAPSQSLSITCTVSGFS LTSYGVHWVRQPPGKGLEWLGVIWAG GSTNYNSALMSRLSISKDNSKSQVFLK MNSLQTDDTAMYYCARDNYGRAMDYW GQGTSVTVSS	QIVLTQSPAIMSASLGERVTMTCT ASSSVSSSYLHWYQQKPGSSPK LWIYSTSNLASGVPARFSGSGSG TSYSLTISSMEAEDAATYYCHQY HRSPYTFGGGTKLEIK

Chapter 3. De novo design of minibinders against both hTLR2 and stabilized 2S5

Abstract

To address the need for targeted antagonists of TLR2, we employed a structure-guided, computational approach to design *de novo* minibinders that target a conserved epitope at the TLR1/2 interface[26,27,36]. Using 2S5, a stabilized surrogate of TLR2, we applied RosettaFold Diffusion[36], ProteinMPNN[30,36], and AlphaFold2[33] to generate and refine high-confidence binder candidates. High-throughput screening in *E. coli* yielded purified minibinders, of which eight demonstrated strong binding to 2S5 by ELISA. Five of these also bound human TLR2, grouping into two structural scaffold families distinguished by helix usage at the interaction interface[6]. Knockout mutagenesis confirmed binding at or near the Pam3-binding site. Functional testing in HEK-Blue™ reporter cells showed no inhibition of Pam3-induced NF-κB activation at saturating ligand concentrations, though modest antagonism was observed at reduced Pam3 levels. Collectively, these results demonstrate the feasibility of de novo minibinder design for targeting TLRs and lay the groundwork for developing novel immunomodulatory therapeutics[11,13,15].

Introduction

TLR2 is a pattern recognition receptor involved in the innate immune response. It recognizes a variety of microbial components, particularly lipoproteins and lipopeptides[1,2,91]. In the presence of a triacylated lipopeptide that mimics gram positive bacterial cell wall components, Pam3CSK4 (Pam3) TLR2/TLR1 heterodimers[5,6,39,60]. The signaling cascade due to dimerization activates NF-κB and MAPKs, leading to the transcription of pro-inflammatory cytokines (e.g., TNF-α, IL-6)[1,49,51,94]. This results in inflammation, antimicrobial responses, and activation of the adaptive immune system.

While crucial to careful regulation of the immune system, overactivation of TLR2 can have many severe consequences. Persistent TLR2 activation can drive prolonged production of pro-inflammatory cytokines like TNF-α, IL-1β, and IL-6[21,95]. This can contribute to tissue damage, autoimmune-like conditions, or worsen existing inflammation-related diseases (e.g., rheumatoid arthritis or inflammatory bowel disease)[96,97]. Continuous stimulation may disrupt immune tolerance, potentially promoting autoimmune responses such as lupus[18,98], multiple sclerosis[99–102], or type 1 diabetes[103]. In addition, TLR2 is expressed on glial cells in the central nervous system[99,101,102]. Overactivation by Pam3 (or endogenous TLR2 ligands released during injury) can promote neuroinflammation[102], contributing to diseases like Alzheimer's or Parkinson's disease[104–106]. Lastly, In extreme cases, especially with systemic exposure to TLR2 agonists, the immune system can enter a cytokine storm—an over exuberant

inflammatory response resembling sepsis[107,108], leading to multi-organ failure. This shows the clear and obvious need for therapeutics that can target and antagonize TLR2 activation.

Minibinders are small, engineered proteins (typically ~60 amino acids) that are designed to bind with high affinity and specificity to a particular molecular interface[26,27,109,110]. Unlike antibodies, minibinders are often created de novo using computational protein design rather than being derived from natural immune responses [109,111]. This makes them highly specific to key epitopes researchers want to target[26,109] including cryptic or conserved sites inaccessible to antibodies while simultaneously being optimized for affinity, specificity, or reduced off-target effects[26,29,110]. In certain contexts, minibinders can have marked advantages over antibodies, namely regarding specificity, stability and expression. Computational design enables quick iteration and response to emerging targets and can be expressed in prokaryotic systems (like E. coli) with extremely high yields, making them cheaper and faster to produce than antibodies, which often require mammalian systems[109].

Here we show the use of AI models such as RossettaFold Diffusion (RFDiffusion)[36], ProteinMPNN[29,30] and AF2[33] to develop de novo binders against hTLR2. Using 2S5 as a proxy for design, we characterize and test these binders to identify multiple binders. Lastly, these binders are further tested as antagonists to TLR2 activation in the presence of Pam3. These studies show the usefulness of de novo minibinder design to quickly identify therapeutic targets for TLRs.

Results

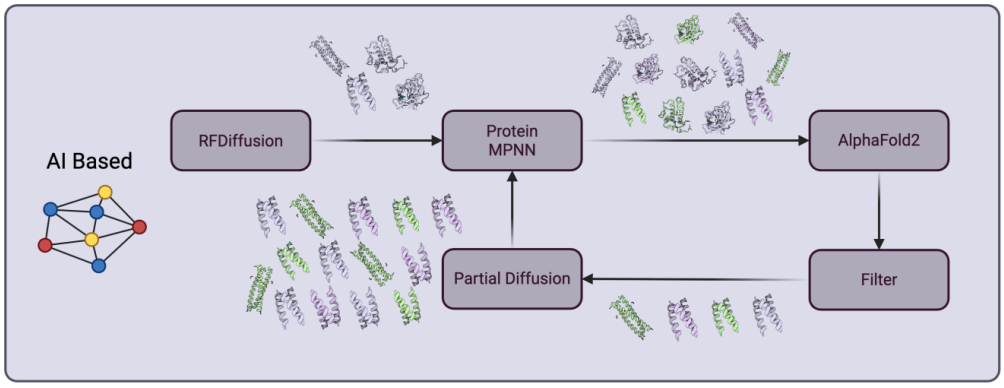
Computational design of TLR2 minibinders

Following confirmation that 2S5 retained surface functionality against human hTLR2 (see Chapter 2), de novo minibinders were designed to target conserved regions between the conserved and stabilized versions. Backbone scaffolds were generated using RosettaFold Diffusion (RFDiffusion)[36], yielding a library of candidate de novo protein frameworks (Fig. 1a). Sequences were assigned to these scaffolds using ProteinMPNN[30], a step that also increased scaffold diversity by applying multiple sequences to each structural framework. Structural models were evaluated using AF2[33] and filtered based on predicted alignment error (PAE_interaction) and average predicted local distance difference test (pLDDT) scores. High-scoring scaffold sets were further refined using Partial Diffusion, a method involving localized perturbations of the backbone to enhance shape complementarity at the interface[36]. The iterative application of RFDiffusion, ProteinMPNN, and Partial Diffusion resulted in a robust pipeline capable of generating a diverse set of structurally viable minibinders (Fig. 1a). The RFDiffusion pathway was done for both 2S5 and the AF2 model of hTLR2, with a high correlation of

pae_interaction for predicted complexes between the two (Extended Data Fig. 1a), implying that 2S5 is a good computational proxy for hTLR2 at this interface. Through successive iterations, a final set of 159 high-confidence minibinders were produced.

To ensure dual specificity toward both 2S5 and hTLR2, a conserved binding site was selected at the TLR1 interaction interface, which overlaps with the Pam3 binding region (Fig. 1b)[6]. As this loop resides within the TLR1/2 interface, it was preserved throughout the design process. This site was chosen due to its high surface hydrophobicity, a feature that enhances design tractability. In particular, residues 302–307, forming the loop sequence FYLFYD, represent a hydrophobic hotspot critical for interaction with TLR1 (Fig. 1b). Targeting this epitope not only promotes cross-reactivity with both hTLR2 and 2S5, but also offers the potential to antagonize TLR2 activation[61], thereby presenting a safer and potentially inhibitory design target (Fig. 1b).

a.



b.

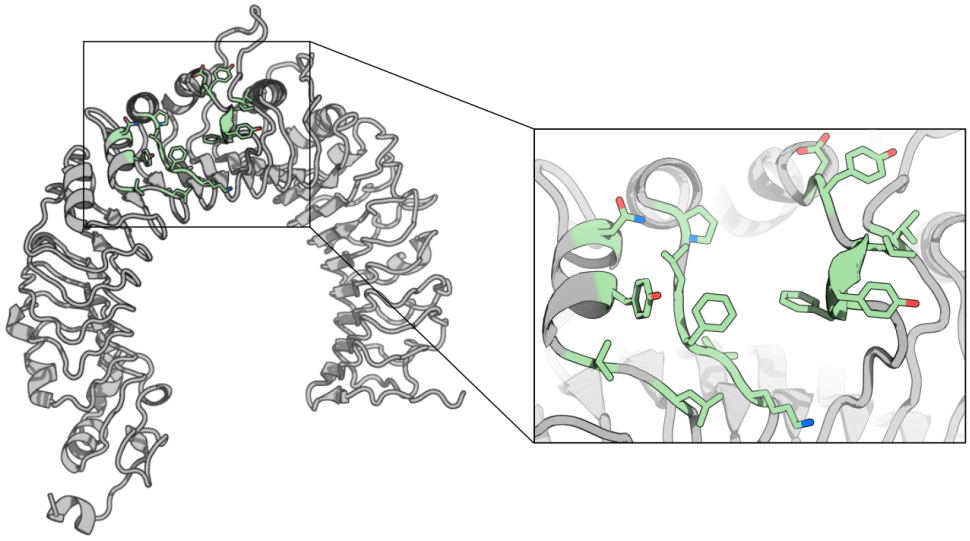


Figure 1. Computational Design of TLR2 minibinders

a, Schematic depicting the pipeline of AI based approach for design of TLR2 minibinders. **b,** TLR2 (PDB: 2Z7X) depicting the interface chosen to develop de novo minibinder against TLR2.

High Throughput screening and Characterization of hTLR2/2S5 minibinders

To facilitate the cloning of 159 distinct designs minibinder constructs were cloned into *E.coli* and expressed using a high-throughput purification workflow (Fig. 2a). This enabled efficient parallel insertion of DNA fragments into expression vectors in a 96-well plate format for downstream studies. Of the 159 designs, 142 were successfully expressed and purified, corresponding to an overall success rate of 89% (Fig. 2a). The molecular size and oligomeric state of each purified minibinder were verified by SEC (Fig. 2b), and eluates were subsequently subjected to binding analysis via ELISA.

Initial ELISA screening against 1 μM 2S5 identified eight minibinders with absorbance values exceeding the threshold of OD_{450} > 0.75 (Fig. 2c, Extended Data Fig. 1b). These candidates were further characterized for cross-reactivity with hTLR2 using BLI. Of the eight screened, five exhibited binding to hTLR2 and were grouped into two distinct structural scaffold families. The first included designs A1 and F6, while the second consisted of G8, G9, and G11 (Fig. 2d). Although both scaffolds share a common three-helix topology common in minibinders[26,27,109–111], the A1/F6 variants engage the target interface using all three helices, whereas the G-series scaffolds utilize only two. This broader interaction surface in the A1 family may enhance binding affinity but could potentially compromise protein stability and expression yield.

To validate binding site specificity, knock-out (KO) mutations were introduced at the putative interface[110]. Comparative ELISAs between wild-type and KO variants confirmed that binding of the G-family minibinders to hTLR2 was abolished following mutation (Fig. 2e), indicating that binding occurs at or near the Pam3/TLR1 interaction interface. KO analysis of the A1 and F6 designs is ongoing to determine whether a similar binding mechanism is employed.

One unique facet of the G scaffolds is their similarity to one another, in particular, G9 vs. G11. Notably, high sequence and structural similarity was observed among the G-family scaffolds, particularly between G9 and G11. While G9 demonstrated strong binding to both hTLR2 and 2S5, G11 exhibited reduced binding activity, consistent with a knockdown phenotype. Structural inspection revealed several mutations at non-interface locations, but one key difference at residue 37 within the binding interface: G9 possesses a tyrosine, whereas G11 contains a glycine at this position, located adjacent to the critical TLR2 loop (residues 302–307). This substitution underscores the

functional importance not only of the loop region itself but also of surrounding residues that contribute to the overall binding interface.

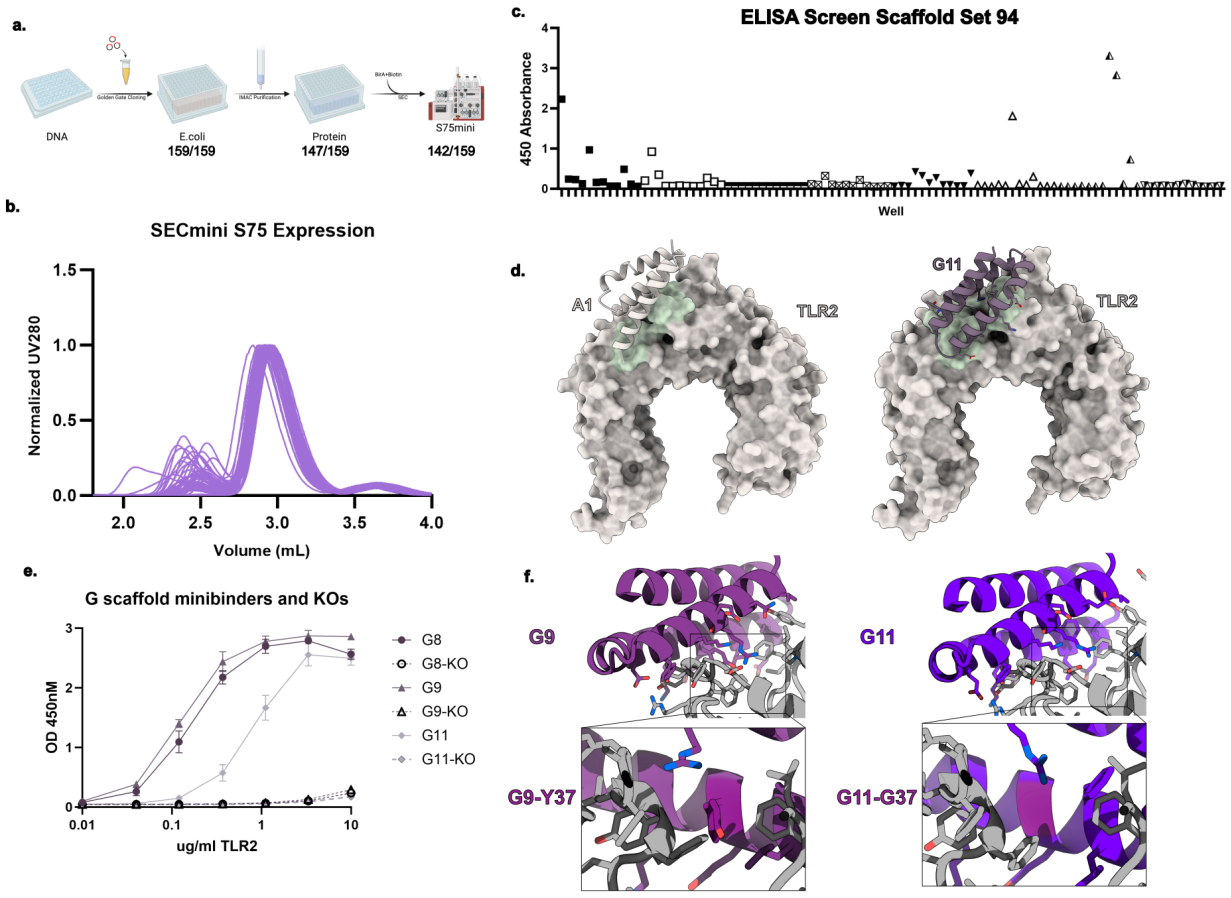


Figure 2. High Throughput Expression and Characterization of TLR2 minibinders.

a, Schematic representing how minibinder DNA was transformed, expressed and purified in 96 well format. **b,** SEC traces depicting successful expression of designs. **c,** ELISA data of scaffold set 94 and their binding to 2S5. **d,** AF2 models of minibinders A1 and G9 bound to TLR2. **e,** ELISA of G9, G8, and G11, which share very similar backbones and sequences, and their KO sequences against hTLR2. **f,** Computational model of subtle difference of Y37 vs. A37 at the interface between G9 (high affinity) and G11 (low affinity) binders.

In Vitro antagonism of TLR1/2 dimerization with G9 minibinders

Following successful characterization of minibinders and confirmation of their binding sites, we next evaluated their potential to antagonize Pam3-induced activation of the

TLR1/2 signaling complex. To this end, we employed the HEK-Blue™ reporter cell line , which stably expresses TLR1 and TLR2 and contains a secreted embryonic alkaline phosphatase (SEAP) reporter under the control of the NF-κB promoter[71,104]. Upon stimulation with a TLR1/2 agonist such as Pam3, activation of the NF-κB pathway results in SEAP secretion, which can be quantitatively measured as an indicator of TLR2 activation (Fig. 3a).

In initial experiments, 250 µg/mL of minibinder was co-incubated with 100 ng/mL of Pam3, and SEAP activity was measured by absorbance at OD_{655}. Under these conditions, no significant inhibition of Pam3-mediated activation was observed. Subsequently, while maintaining a constant minibinder concentration (250 µg/mL), a titration of Pam3 was performed, starting at 80 ng/mL and proceeding through a 4-fold serial dilution (Fig. 3c). Under reduced Pam3 concentrations, modest attenuation of TLR2 activation was observed relative to the G9-KO control, suggesting potential weak antagonistic activity. However, additional experiments are required to validate these preliminary findings and to elucidate the mechanistic basis of the observed inhibition.

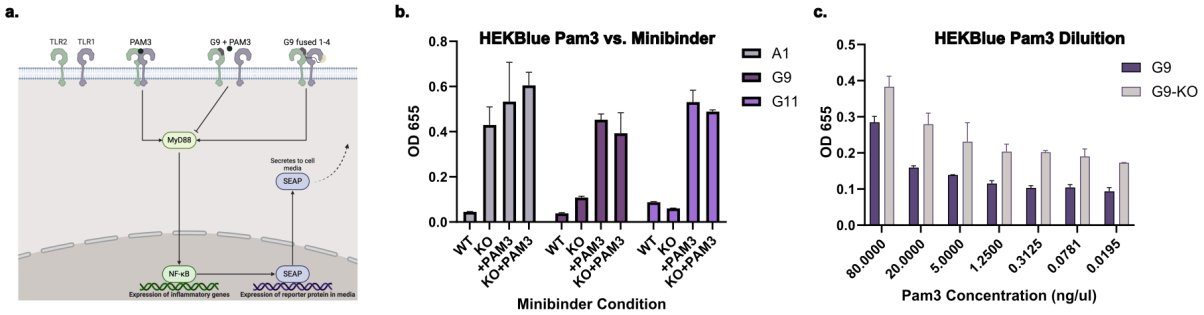


Figure 3. HEKBlue Cell assay for TLR2 antagonism.

a, Schematic representing the high level functionality of the TLR1/2 HEKBlue assay. When activated by Pam3, the NF-κB pathway induces the secretion of the SEAP reporter into the media. In the presence of monomeric minibinder TLR1/2 activation is inhibited, leading to lower NF-κB and thus SEAP expression. **b,** Initial assay indicating G9 is capable of reducing TLR1/2 signaling in the presence of Pam3. **c,** Dilution series of G9 with lowering levels of Pam3 showing a reduction in signal as Pam3 levels decrease compared to that of G9-KO.

Discussion

In this study, we demonstrate that the use of the stabilized TLR2 variant, 2S5, enabled the successful development of de novo minibinders that are cross-reactive with both

hTLR2 and 2S5. This represents a significant advancement, particularly in light of prior efforts that failed to yield effective TLR2-targeting minibinders. Several key factors contributed to this success. First, careful selection of the target interface was critical. We focused on a region predicted to be conserved and structurally similar between hTLR2 and 2S5, which likely overlaps with the binding sites of known antibodies such as MAB2616 and mAb A11 (as discussed in Chapter 2).

Second, the use of RFDiffusion facilitated the generation of scaffold sets with enhanced shape complementarity to structurally important features of the TLR1/2 interface, particularly the curved surface topology of TLR2 and the hydrophobic loop spanning residues 302–307. Third, the superior expression and stability of 2S5 relative to wild-type TLR2 significantly improved the throughput and fidelity of downstream screening assays. Notably, ELISA-based validation at 1 μM required substantial amounts of target protein—quantities that would not have been attainable using native TLR2.

While these minibinders show promise for future therapeutic and synthetic biology applications, further in vitro characterization is necessary to elucidate their functional properties. Preliminary data suggest that one binder, G9, may bind with sufficient affinity to compete with Pam3 at the TLR1/2 interface. Ongoing work will involve titrating minibinder concentrations to define the threshold required for effective antagonism and to assess whether higher-affinity variants can achieve complete inhibition of receptor activation. Additionally, future studies (see Chapter 5) will explore the fusion of these minibinders to TLR1-targeting domains as a strategy to modulate or induce receptor activation.

Acknowledgments

Thanks to the entire RFDiffusion team for their development of this great tool. Thanks to Justas Dupras for ProteinMPNN development. I would like to thank Craig Dobbins and Adri Tran-Pearson for their work with cell based assays.

Materials and Methods

Computational Design

The ectodomain of TLR2 (PDB ID: 2Z7X)[6] and 2S5 were used for the design. The PDBs were truncated to save computational time. Minibinders scaffolds were designed with RFDiffusion outlined in ref[36]. To generate hundreds of de novo backbones, where the final ten time steps were taken from each scaffold in the diffusion process. Once scaffolds were generated ProteinMPNN was used to generate ten sequences per backbone, resulting in hundreds of thousands of designs. Complexes were then predicted using AF2 and filtered using various metrics (pae_interaction, binder_plddt).

Small Scale protein expression and purification

DNA encoding the designed proteins was ordered from IDT as eblocks and cloned into vector 515 LM1452 (Addgene) which contains a C-terminal 6xHis tag for IMAC purification and an Avi tag for biotinylation. Plasmids were transformed into BL21(DE3) (NEB) and 1 mL of single-colony culture was grown in Terrific Broth II (MP Biomedical) at 37°C overnight. Cultures were spun down for 30 min at 4000 g and pellets were resuspended in autoinduction medium and grown at 37°C for 2 hr and 18°C overnight[112]. Cells were collected by spinning for 30 min at 4000 g. Pellets were resuspended in lysis buffer (50 mM Tris pH 8.0, 250 mM NaCl, 20 mM imidazole, 0.04 mg/mL RNase, 0.1 mg/mL lysozyme, 0.1 mg/mL DNase, 1 mM PMSF). The cells were lysed via a 24 pronged sonicator translated 4 times to sonicate all wells. The lysed cells were clarified via centrifugation at 14,000g for 30 min. The supernatant was run over Nickel-NTA resin and washed with 3 CV of wash buffer (50 mM Tris pH 8.0, 250 mM NaCl, 20 mM imidazole). Protein was eluted with an elution buffer (50 mM Tris pH 8.0, 250 mM NaCl, 500 mM imidazole) and further purified by size exclusion chromatography using a Superdex S75mini 10/300 GL column (Cytiva).

Protein expression and purification

Genes encoding minibinders were synthesized and cloned by Integrated DNA Technologies into a pET-29b(+) vector. Plasmids were transformed into BL21(DE3) (NEB) and 30 mL of single-colony culture was grown in Terrific Broth II (MP Biomedical) at 37°C overnight. 25 mL was transferred to 0.5L of autoinduction medium and grown at 37°C for 2 hr and 18°C overnight[112]. Cells were collected by spinning for 30 min at 4000 g. Pellets were resuspended in lysis buffer (50 mM Tris pH 8.0, 250 mM NaCl, 20 mM imidazole, 0.04 mg/mL RNase, 0.1 mg/mL lysozyme, 0.1 mg/mL DNase, 1 mM PMSF). The cells were lysed via a sonicator. Purification of cells was performed by methods described above.

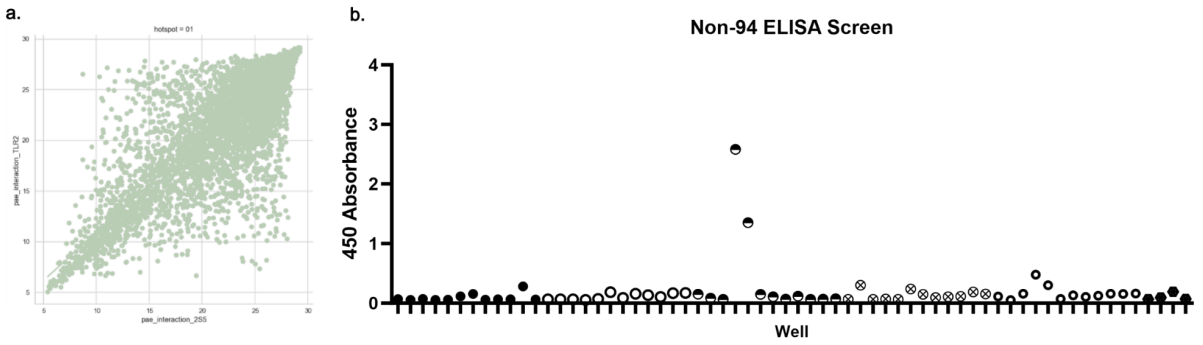
Enzyme-linked immunosorbent assay (ELISA) for minibinders

Peirce Streptavidin 96-well high protein binding plates were coated for one hour at room temperature at with biotinylated minibinders at 10 µg/mL and blocked in 100 µL of blocking buffer composed of (PBST: 1x Phosphate-buffered saline with 25 mM Tris pH 8.0, 150 mM NaCl, 0.2% Tween 20, and 5% nonfat milk). Plates were then washed 3x with PBST in an automated plate washer (Biotek); all washing steps follow the same protocol. Minibinders were incubated with 1uM 2S5 for 2 hours at 600 rpm. Plates were washed then 100 µL of a mouse anti-TLR2 antibody 316189 (Abcam) for 40 minutes shaking at 500 rpm. Secondary anti-rabbit (Cell Signaling Technology) (1:1,000 dilution) IgG-HRP were added to each well and incubated shaking at room temperature. Plates

were washed before 100 µL per well of TMB were added and developed for 3 min, then quenched with 100 µL of 1N HCl. Reading at absorbance at 450 nM was carried out with an Epoch plate reader (Biotek).

HEK-Blue™ antagonism assays

Cells were cultured as described in the manufacturer's technical data sheet (InvivoGen). For the antagonism assays with purified minibinders, the proteins were diluted in the HEK-TLR2 medium at 250ug/ml into 200-µl medium containing around 1 × 10^5 HEK-TLR2 cells prior to application for the test. Pam3 was purchased from Invivogen dissolved in water at either 1 mg/ml and aliquoted into 200uL quantities for easier dilutions in the above volume antagonism test. Stimulation was carried out for 18 h. After 18h incubation, activation levels were quantified using a Neo2 plate reader (Agilent) at a wavelength of 650nm.



Extended Data Fig. 1: 2S5 minibinder design and additional scaffold set.
a, Graph depicting scores of pae_interaction of hTLR2 vs. 2S5. Data indicates scores for 2S5 match well to hTLR2 for this interface (R^2 = 0.8). **b,** ELISA data of second scaffold set used for minibinder design. While two did appear to bind to 2S5, they did not bind to hTLR2 and were therefore disregarded for further screening.

Supplementary Table 1. Amino acid sequences

Construct*	Amino Acid Sequence
A1	MEEAKKKAEEVKPKAETDLYARLEYLRYMYAVAVAEGDKKKAEELE KEAKKVYEELIKKLE
B2	AEAEEYKAKAESLKPKLDDPLYQNKYYAYLYHYYKLLGDKEKAAEY EEKAKAAYEKIVEEIE

F6	EAEEHLAKAEEYKAKSSTLLEKWLAAAHLYAAYKELGDEEKAKEAE KKAKELYEELVKSIE
G8	NPVEAYEAYKEALKKLYEKGDEKALKELEPKLKELFEKALKAIEAGD LEGAAKYFEEAAKLVEEAL
G9	SPVEAYNRYREALKKLYEKGDKKALEELDPKLKELYEKALEAIEKG DLEGAAKYFQEAADLVEKAI
G11	SLVAALEAYREALKALYDAGKEEELKAIEPELKALGEKFLAAIEAGD KEGAIAYLQAATDLVKKAA
A1-KO	MEEAKKKAEEVKPKAETDLAARLAYLAAMYAVAVAEGDKKKAEELA KEAAKVAEELIKKLE
B2-KO	AAAEEAKAKAESLKPKLDDPAYQNAYYAALYHYYKLLGDKEKAAEY AEKAAAAAEKIVEEAE
F6-KO	AAEEALAKAEEYKAKSSTLAEKWLAAAALYAAYKELGDEEKAKEAE KKAKELAEELVKSAE
G8-KO	NPVEAAEAYAEALAKLAEKGDEKALKELEPKLKELAEKALKAIEAG DLEGAAKYFEEAAKLVEEAL
G9-KO	SPVEAANRYAEALAKLAEKGDKKALEELDPKLKELAEKALEAIEKG DLEGAAKYFQEAADLVEKAI
G11-KO	SLVAAAEAYAEALAALADAGKEEELKAIEPELKALGEKFLAAIEAGD KEGAIAYLQAATDLVKKAA

Chapter 4. Computational and Rational Redesign of TLR5

Abstract

As previously discussed, TLRs are difficult to produce recombinantly in mammalian cells. This is especially true of TLR5, which relies on flexible loops as the binding modality for the agonist flagellin[43,44]. Here we use the methods from Chapter 2 to redesign TLRs while maintaining these flexible binding loops and key interfaces. While many surface mutations were required to increase expression, including full redesign of N and C termini as previously used in TLR2, TLR5 maintains binding to both native flagellin as well as a modified flagellin that is more stable in vitro. This shows the applicability of this approach to other TLRs to increase expression of various forms of TLR.

Introduction

Toll-like receptor 5 (TLR5) is a pattern recognition receptor that plays a crucial role in the innate immune system by detecting bacterial flagellin[1,113], a structural protein of the flagellum involved in motility. TLR5 is primarily expressed on the surface of various antigen-presenting cells (APCs), including dendritic cells and macrophages[50,114]. Unlike TLR2 activation which only functions as a heterodimer, TLR5 functions as a homodimer[44]. During TLR5 activation, flagellin interacts with a highly conserved, flexible loop within LRR9, as well as additional contact points in the C-terminal LRRs (including LRR18–LRR21) on the partnering TLR5 monomer[1,50,113]. This dual interaction bridges the two TLR5 molecules, promoting receptor dimerization. The ligand-induced dimerization brings together the intracellular TIR domains, which initiates downstream signaling through the adaptor protein MyD88, leading to the assembly of the Myddosome complex and activates key transcription factors such as NF-κB and AP-1, resulting in the production and secretion of proinflammatory cytokines[1,3,51].

Although TLR5 plays an essential role in initiating innate immune responses, its recombinant expression has proven particularly challenging. This difficulty is largely attributed to its extensive flexible loop regions, which are critical for flagellin recognition but hinder structural stability and expression yields[43,44]. Consequently, the development of molecular tools targeting TLR5—including monoclonal antibodies and de novo binding proteins—has been significantly impeded[113,115]. Motivated by our success in stabilizing TLR2, we sought to determine whether similar strategies could be applied to TLR5. Specifically, we explored the integration of structure-based design tools, including ProteinMPNN and AlphaFold2, in conjunction with both rational and holistic design principles, to enhance the stability of this recalcitrant receptor.

In this study, we demonstrate that TLR5 can be successfully stabilized using the computational and experimental strategies outlined in Chapter 2. Although its expression levels remain substantially lower compared to the stabilized TLR2 construct (2S5), the engineered TLR5 variant was expressed in sufficient quantities for structural and functional characterization. These results underscore the potential of this stabilized TLR5 construct as a platform for the development of antibodies and de novo minibinders targeting this innate immune receptor.

Results

Computational methods to redesign TLR5

TLR5 presents a greater challenge for recombinant expression relative to other TLR family members, primarily due to the presence of flexible loop regions essential for flagellin recognition and subsequent signal transduction. Building on the methodologies described in Chapter 1, we applied our prior design principles to this more structurally complex target. Key structural features—including N-linked glycosylation sites[74], disulfide bonds[41], and flagellin-interacting regions[2,44,115,116]—were preserved to maintain functional integrity. In contrast, regions exhibiting low pLDDT scores, indicative of conformational instability, were subjected to extensive redesign. Notably, the C-terminal domain underwent a comprehensive structural revision[6,25,43], complemented by targeted rational mutations at select sites to enhance overall protein stability and expression.

To initiate the redesign process, an AF2[33] structural model of human TLR5 (hTLR5) was used as input for ProteinMPNN[29] to generate a set of 200 computationally designed sequences. These designs were largely unconstrained, aside from previously defined restrictions on key functional residues and structurally critical motifs. Based on prior success in stabilizing TLR2 using a similar approach, we anticipated that unconstrained designs would exhibit favorable expression profiles despite the relatively small sequence pool. The resulting sequences displayed an average identity of 67.57% compared to wild-type TLR5, with extensive regions undergoing complete sequence replacement. From this initial pool, the top four candidates were selected for rational refinement. During this step, numerous mutations were reverted to the wild-type sequence when they were predicted to confer no clear structural or functional advantage. Guided by insights from TLR2 optimization, we focused redesign efforts on the structurally flexible concave surface of TLR5 and adjacent regions near the flagellin-binding loops, while preserving residues critical for ligand recognition. This strategy was intended to enhance overall protein stability without compromising native binding functionality.

Due to the need to stabilize both the concave portion and loops, more surface mutations were allowed than that of TLR2. Once reversions were made, the four redesigned TLR5

were fed back into ProteinMPNN to generate another 5000 sequences for each candidate which we then applied the same key holistic approach that was used to generate the best scores for each of the final candidate structures.

In addition to the methods previously mentioned, consensus design was also attempted to see if this could be used to improve expression[32]. This entailed reverting non canonical residues of the TLR back to their most common ancestral residue observed in both TLR and LRR families with the exception of the binding loops (which was maintained) and C-terminus (which was redesigned). To do this an TLR5 was put through the UniRef100 database[117]. From here a weighted score was provided for each residue at a given position. These scores were then used to identify residues that could be reverted to canonical residues. While most of these residues were located on the surface (Fig. 1a,b) of the TLR and therefore were avoided for redesign (Fig. 1c), a collection of core residues were identified (Fig. 1b), and combinations of reversions were made for redesign. The highest scoring combinations were ordered along with the original method described for expression studies in mammalian cells.

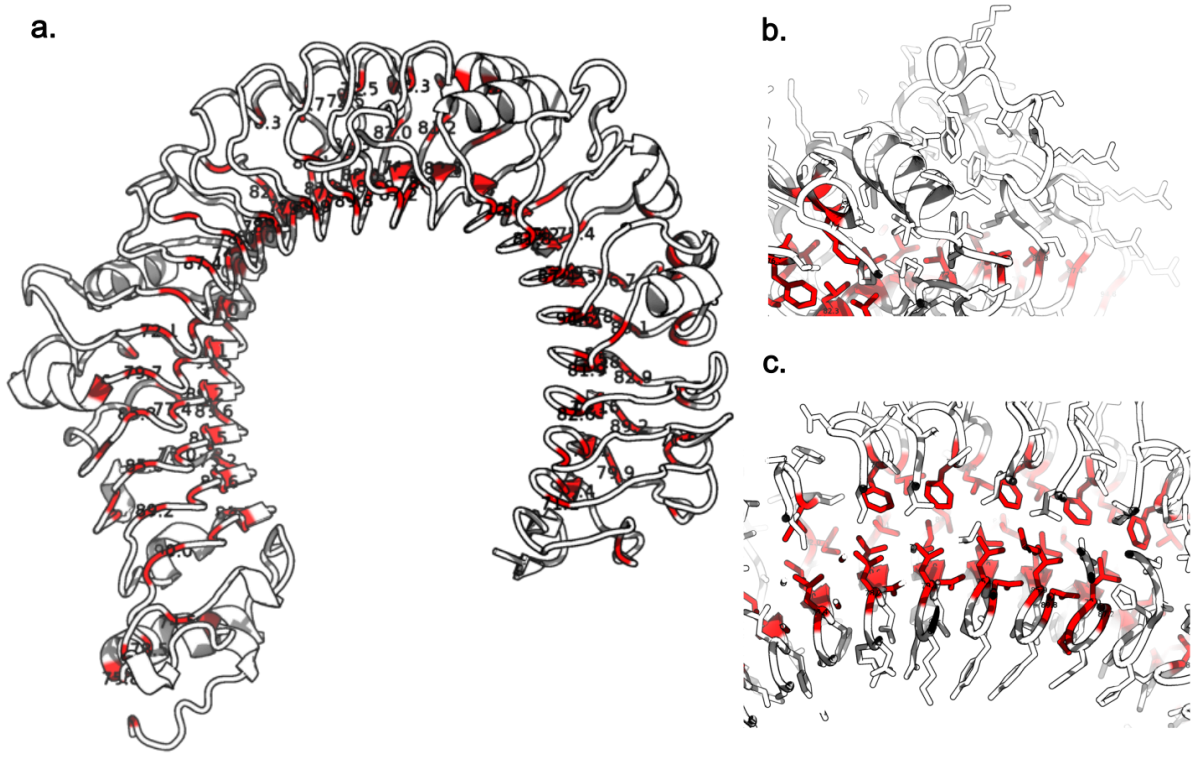


Figure 1. Consensus regions of hTLR5.

a, Full model of hTLR5 (from AF2 model) consensus regions with weighted conservation > 0.75 (red), while the surface and N/C-termini remain largely unconserved (white). **b,** Zoomed in look at flagellin binding loop of hTLR5 and the

lack of conservation this region has compared to the core. **c,** A cross section focusing on the core of hTLR5, noting the highly conserved and repeated LRR motif LXXNXLXXXF.

Large scale mammalian screen of successfully redesigned TLR5

Due to the difficulty to generate stabilized TLR5, we chose to screen a significantly larger number of designs than with TLR2. We screened 95 stabilized designs (48 consensus designs, and 47 computationally redesigned sequences, and native TLR5 fused to Fc). Designs were ordered as two sets of DNA fragments inserted into a CMVR plasmids with a C-terminal Fc domain as part of the backbone (Fig. 2a). The 96 designs were transfected with an Fc domain in 3mL of Expi293 cells. As expected, most of the designs did not successfully express in mammalian cells with 9/95 indicating some expression after 3 days (Extended Data Fig. 1). Of the nine designs, two (A7, A8) came from the consensus designs; while the rest were from the full redesign (E12, F2, F11, H1-H4). The nine successful designs were scaled up to 35 mL for full purification and yield testing. The SEC data indicated that few of these express at high levels (Fig. 2b) with A7 producing significantly higher expression than the others. The ones that eluted at the correct volume were scaled again to 125mL to identify these could be produced at practical yields. After SEC, both A7 and H4 indicated viable, albeit low, levels of expression.

Sequence and structural comparison between A7 and H4 to hTLR5

After the two final candidate sequences were identified, their sequences were looked at in more detail. A7 has an overall sequence identity of 72.11% with an 82% similarity, compared to hTLR5, while H4 showed an identical sequence of 79.70% and 87% similar. The results of this correlate with the expression data where A7 expresses at a higher rate in medium scale SEC traces than that of H4.

When looking at the models for A7 vs. H4 the difference in approaches between consensus and AI based methods becomes very clear. While consensus design created a largely stochastic set of mutations all along the TLR, AI and reversion methods are far more targeted (Fig. 3)**.** Similar areas to TLR2 are targeted in the computational stabilization of TLR5, namely the core residues contributing to the curvature of the TLR (185-225, 425-475). Additionally, for both structures the C-terminus underwent drastic mutations to stabilize TLR5 (Fig. 3b,d). Namely the contribution of polar residues on the final five LRRs appears in both models suggesting it confers a high level of stability to the end of the TLR.

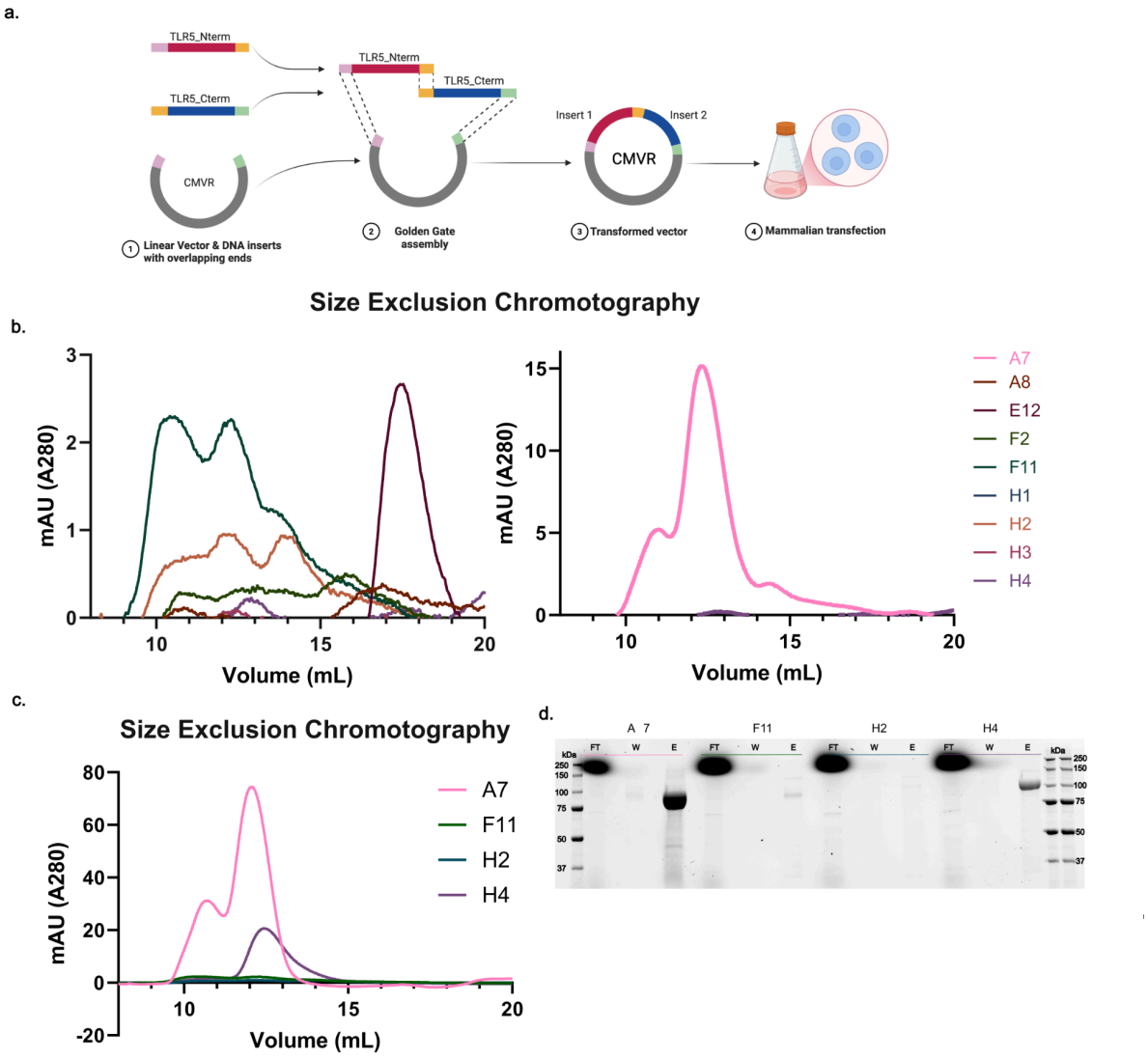


Figure 2. Characterization of stabilized hTLR5 models.

a, In order to screen 96 designs, high throughput plasmid preparation was utilized. Two component TLR5 designs were cloned into CMVR backbones with 6xHis tags on the C-terminus for low volume expression tests. **b,** SEC graphs depicting the nine candidates that appeared to express at 35mL volume. The SEC trace to the right is A7 which showed significantly higher traces at the same transfection volume compared to the other 8. For comparison, H4 is shown as well for comparison. **c,** SEC Traces of the remaining four candidates at 125mL volume. **d,** SDS-PAGE gels depicting purification of each of the remaining candidates with Flow Through (FT), Wash (W) and Elutions (E), indicating designs are the correct size.

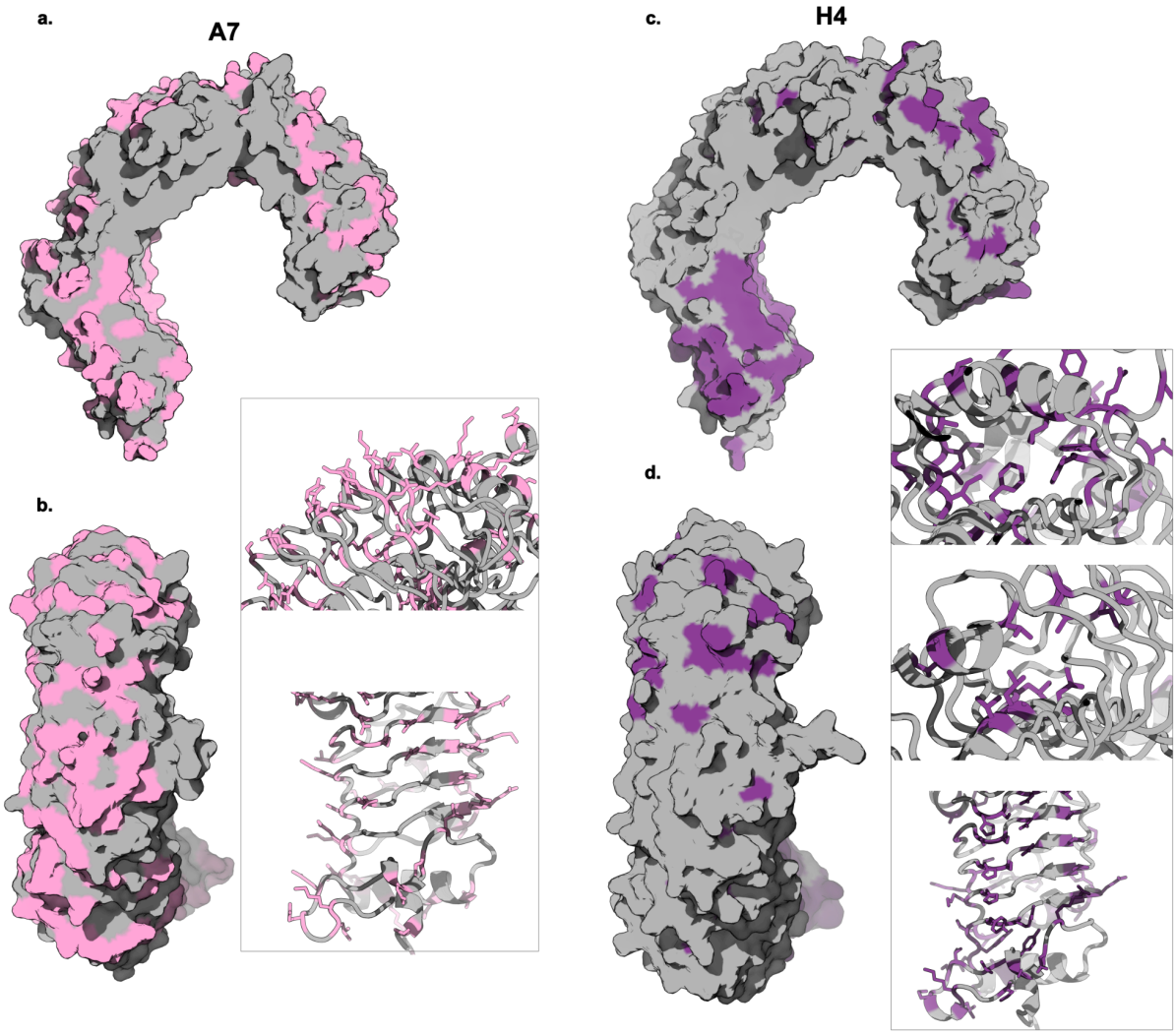


Figure 3. Structural Comparison of A7 and H4.

a, Whole TLR view of A7 with mutations denoted in pink and conserved residues in gray. **b,** Top down and zoomed view better depicting a wide variety of mutations on the surface and C-terminus. **c,** Whole TLR view of H4 with mutation denoted in purple. **d,** Top down view of H4 and zoomed in portion showing stabilizing mutations in the protein core and C-terminus.

Determination of overall stability and shape of stabilized TLR5

Once expression of TLRs was established, we sought to identify if these proteins maintained both stability and their canonical shape. In addition to eluting at the correct

volume, both A7 and H4 appear at the proper size by DLS (Fig. 4a). Using SYPRO melts, we observed an approximate melting point of 55℃ and 68℃ for A7 and H4 respectively (Fig. 4b). This indicates that H4 may be significantly more stable than A7, although further testing is needed to confirm this inference. Finally, the overall shape of the proteins were determined by nsEM. The 2D class averages show that A7 is unclear if proper folding is maintained, while H4 maintains a tight dimer due to expression with an Fc (Fig. 4c). This indicates that while A7 may express and maintain initial proper folding, it is likely not as stable as H4 which displays a higher melting point as well as far better folding when observed under nsEM.

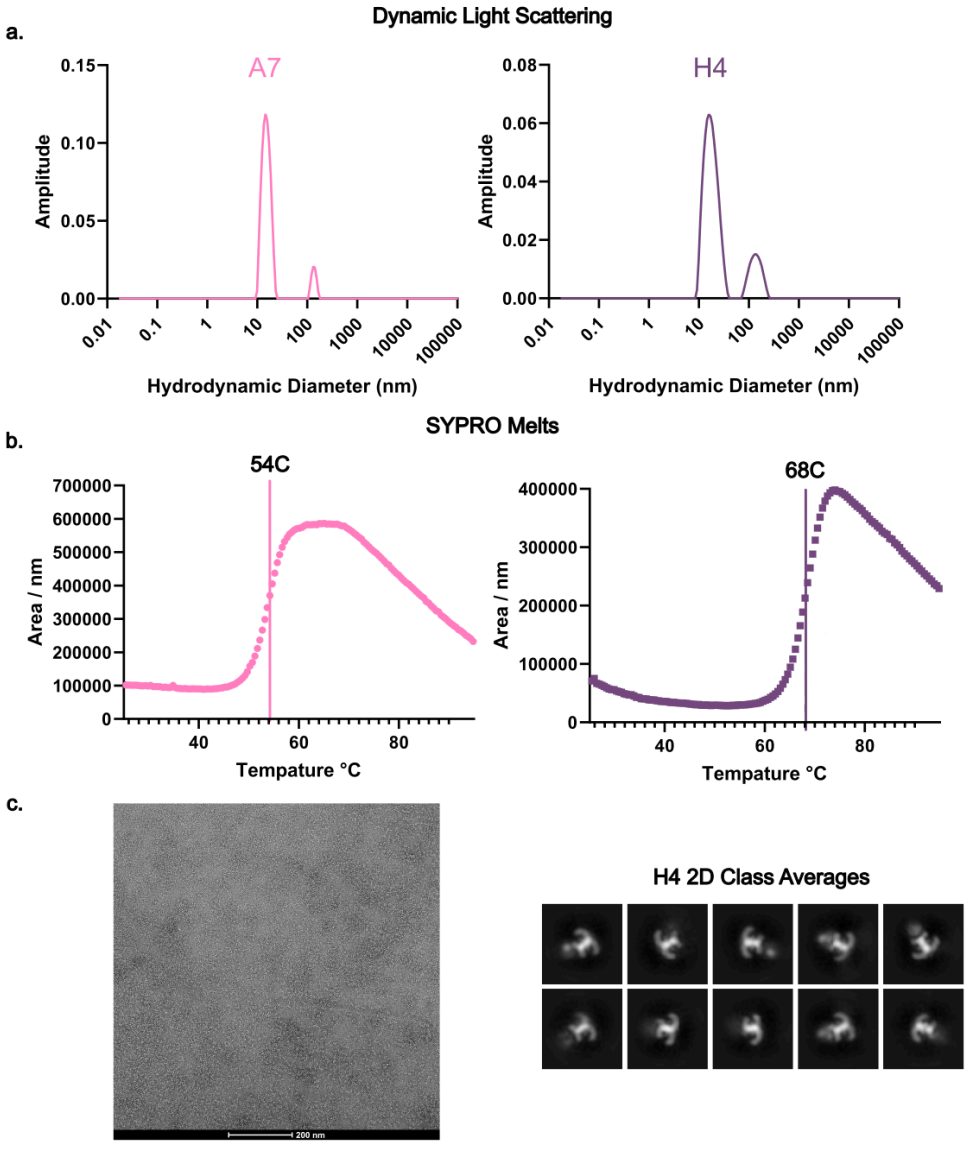


Figure 4. Stability screens of A7 and H4.

a, Dynamic Light Scattering (DLS) data depicting A7 and H4 as approximately 14Å in length with minimal aggregation (i.e Hydrodynamic Radius > 100). **b,** SYPRO melts showing melting temperatures of A7 and H4 at 54℃ and 68℃ respectively. **c,** nsEM data indicating proper folding of H4.

Stabilized TLR5 bind to flagellin and flagellin mimetics

In order to determine if H4 maintained the functional characteristics of TLR5, namely binding to TLR5 agonist flagellin. To do this we took multiple approaches due to the instability of flagellin as well as the variety of species it can be derived from[113,115,116]. Firstly we measured the binding of TLR5 to computationally stabilized flagellin designed in-house. These flagellin maintain much of the surface identity to native flagellin, but have either large truncations in insatiable regions, or subtle mutations to increase expression and stability (Fig. 5a,b). In vitro studies indicate that many of these flagellin express well and activate TLR5 at native levels (Fig. 5c). To ascertain if H4 maintained flagellin binding functionality, ELSAs against native flagellin as well as the high performing mimetics was done. ELISA data indicated that while H4 does maintain some binding affinity to the stabilized flagellins, it is unable to bind to native flagellin. In addition, the binding that is observed in the redesigned flagellin, is weaker than that of mTLR5-Fc. This indicates that while some functionality is maintained, more work is required to get full functionality in H4.

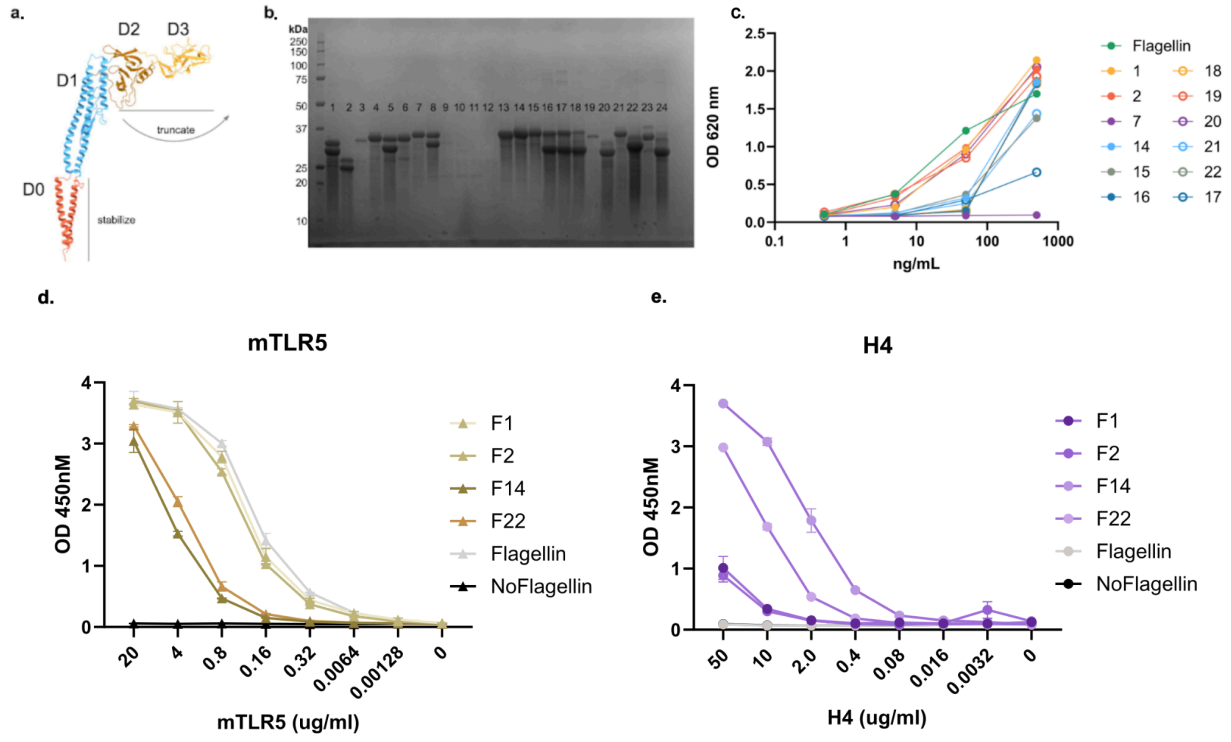


Figure 5. ELISA Binding of mTLR5-Fc and H4 to Flagellin and Flagellin mimetics.

a, (Data for figures **a-c** kindly provided by Naveen Jasti and Chloe Adams) Schematic depicting how flagellin was stabilized. The D2 and D3 regions that cause aggregation were truncated and the D0 region was stabilized using ProteinMPNN. **b,** SDS-PAGE of all flagellin proteins. **c,** Stabilized flagellin was tested for activation in a 4x dilution series using a SEAP reporter Ramos cell line. **d,** ELISA data showing successful binding of both native flagellin and flagellin mimics to mTLR5-Fc. **e,** ELISA data of stochastic binding of flagellin and its mimics to H4 with no binding of native flagellin.

Discussion

Here, we show that the previously described methods allow us to extrapolate our findings—specifically the location of key mutations and C-terminal reversions—to another member of the TLR family. Previous attempts to stabilize full length hTLR5 have been largely unsuccessful or produce low quality products[44]. While consensus design did result in some expression in A7, it did not produce strong enough indications of stable expression to move forward with this design. There are many reasons for this, the most prominent being that most of the core of TLR5 is still highly conserved. This means that the majority of mutations are surface mutations that likey ablate the functionality of the stabilized TLR, thus defeating the purpose of stabilization efforts. Lastly, it appears that while the C-terminus inner surface was mutated heavily for stability[6,25,43], the core residues did not undergo much in the way of modifications. This indicates that the core of C-terminus is well packed, and the rigidity of the C-terminus on the surface was key to stabilizing this region of the protein.

While A7 did produce slightly higher yield than H4, the lower melting temperature indicating that the quality of mutations in H4 is significantly higher than that of A7. With nearly 50 less mutations, H4 was able to show slightly lower expression, but far better indications of stability in nsEM data. The consensus approach has been used to stabilize difficult proteins in the past. It is likely a combination of AI and consensus techniques could provide an optimal balance of developability and functionality.

As for the stability and functionality of H4, further work is needed to verify the stability studies shown. While structurally H4 is quite stable, the inability for it to bind to native flagellin[115,116] brings the surface-and more pertinently binding loop-conservation into question[44]. We did not have access to a structure of hTLR5 bound to flagellin so the likely culprit for the unsuccessful binding to flagellin is the use of a computational model[67] to represent the key binding domains.

Lastly, we note the lack of antibody binding data to verify surface antigenicity of H4 and A7. This is because while there are currency antibodies being sold that bind to hTLR5, we did not find their quality to be high enough to convincingly confirm the stabilized antibodies maintained surface antigenicity.

Acknowledgments

Thank you to Chloe Adams, Abby Burtner, and Naveen Jasti for the development of the flagellin mimetics used in this Chapter. Thank you to Helen Eisenach for her help developing CMVR plasmid backbones.

Materials and Methods

Computational Design of stabilized hTLR5

Designs were stabilized following the methods outlined in Chapter 2.

Consensus Stabilization of hTLR5

Consensus design was performed based on methods described here in ref[32]. In brief, TLR5 was input into BLAST servers to generate MSAs that would be used to identify areas of consensus based on their weights of conservation. Conservation is defined as frequency of most-frequent AA at each position of an MSA obtained by hhblits against uniclust30 with max E-value cutoff of 1e-4, hhfiltered at 90% identity, 50% coverage, and min query identity 30%. From here regions were scored on a range from 0.0-1.0 (1.0 being the fully conserved) based on how often they appear at a given position. Scores were mapped to their subsequent residues in the AF2 modeled hTLR5 (shown in Fig. 3a). Once conserved residues were identified and an MSA generated residues that were present in 30, 50, 70 or 80% of MSAs based were held, while the remaining residues were allowed to undergo computational redesign mentioned above.

Library Preparation of Stabilized hTLR5

A custom CMVR plasmid was developed that contained a signal peptide in the N-termial domain and a 6x His tag in the C-terminal domain for IMAC purification. DNA encoding the designed hTLR5 proteins was ordered from IDT as eblocks separated into two regions (N and C Terminal) due to the size constraints for eblocks. These three elements (plasmid, hTLR5-Nterm and hTLR5-Cterm) were cloned then ligated and cloned together and sequenced to ensure plasmid quality.

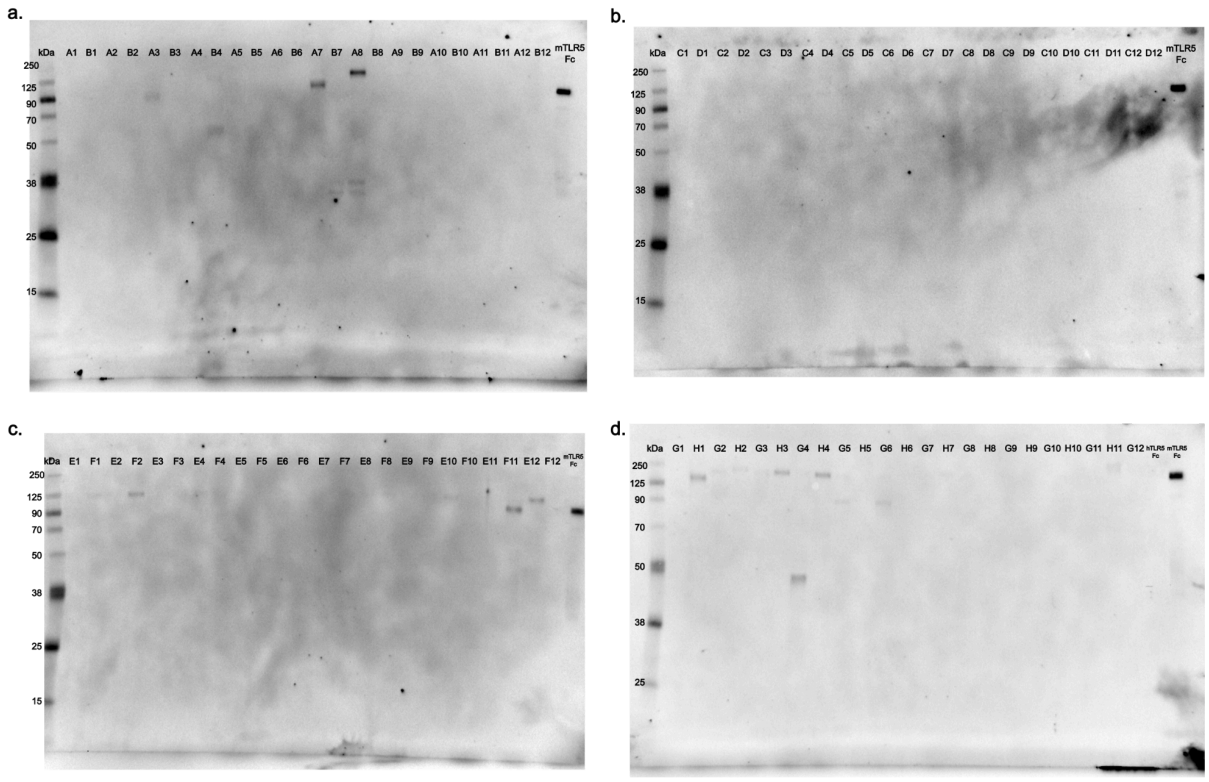
Small Scale Mammalian Expression of Stabilized hTLR5

All constructs contained a C-terminal histidine affinity tag and were codon optimized by GenScript for mammalian cell expression. Four, 24 well plates were transfected with 2mL Expi293F cells were transiently transfected using PEI MAX and cultured for three

days 37°C. Cells were harvested, centrifuged, and filtered to obtain clarified supernatants for initial western blot expression studies.

Mammalian Expression of Stabilized hTLR5

Plasmids were transfected into Expi293 cells with expression and purification methods described in Chapter 2.



Extended Data Fig. 1: Western blot of small scale TLR5 stabilized designs.
a-d, Four SDS-PAGE gels depicting 3mL screen of stabilized TLR5 designs. Wells A1-D12 represent consensus design, E1-G12 represent design methods used in Chapter 2. The final well is TLR5-Fc. All gels use purchased mTLR5-Fc as a control.

Supplementary Table 1. Amino acid sequences

Construct*	Amino Acid Sequence
hTLR5 AlphaFold2 Model	SCSFDGRIAFYRFCNLTQVPQVLNTTERLLLSFNYIRTVTASSFPFL EQLQLLELGSQYIPLTIDKEAFRNLPNLRILDLGSSKIYFLHPDAFQG LFHLFELRLYFCGLSDAVLKDGYFRNLKALTRLDLSKNQIRSLYLHP SFGKLNSLKSIDFSSNQIFLVCEHELEPLQGKTLSFFSLAANSLYSR VSVDWGKCMNPFRNMVLEILDVSGNGWTVDITGNFSNAISKSQAF

	SLILAHHIMGAGFGFHNIKDPDQNTFAGLARSSVRHLDLSHGFVFS LNSRVFETLKDLKVLNLAYNKINKIADEAFYGLDNLQVLNLSYNLLG ELYSSNFYGLPKVAYIDLQKNHIAIIQDQTFKFLEKLQTLDLRDNALT TIHFIPSIPDIFLSGNKLVTLPKINLTANLIHLSENRLENLDILYFLLRVP HLQILILNQNRFSSCSGDQTPSENPSLEQLFLGENMLQLAWETELC WDVFEGLSHLQVLYLNHNYLNSLPPGVFSHLTALRGLSLNSNRLTV LSHNDLPANLEILDISRNQLLAPNPDVFVSLSVLDITHNKFICECELS TFINWLNHTNVTIAGPPADIYCVYPDSFSGVSLFSLSTEGCDGGSH HHHHH
A7	SCVYTGTIAFYRFVNLTEVPEVLSTTEELYLSFNYISVVTSSSFPLLE KLRVLVLGSQYVPLTIEDHAFINLPNLEVLDLGSSPIYYLHPDAFLGL EKLKELYLYFTGLDDAVLKLGYFRTLSALTKLDLSKNKITHLNLHPHF AKLEALKEIDFSSNQIFLVCESELKPLQGKSLDYFSLAANSLFSRVS VDWDKCMNPFRNMTLDVLDVSGNGFTVDITRRFSKAIAKSNATSL KLAHHIMGAGFGFHNIKDPDVNTFAGLASSSVKSLDLSHGFVFTLK SNVFKSLTDLEVLNLSYNKINIIEKNAFSGLENLKYLNLSYNLLGELY SFNFEGLPNVEFIDLSKNHIAIIQDNTFSGLSKLKTLNLRDNALETIN NIPKIPEIYLSGNKLTTLPDIELTATKIHLSENRLENLDILYQLLSVPVL QELILNSNRFSRCDGTSSPVLNPTLKKLFLGDNRLQEAWETELCW DMFDGLTTLEELDLSGNGLTSLPPGVFSHLTALKKLNLNSNKLTKLN HSDLPKNLEVLDISNNSLKAPNPDVFTHLEELDITHNEFVCECDLK AFIDWLNTTNVTILGSPDDIYCVEPESLKGVSLFTLSTDE
A8	SGSCSYSGRIANYRFQNLTTVPPVLSSTEELLLSFNYISVVTSSSFP LLEKLRYLDLGSQYVPLTIESDAFKNLPNLEVLDLGSSPIYYLHPDA FKGLEKLEELNLYFTGLDDAVLRLGYFKDLSALTKLDLSKNKITRLD LHPHFAKLKSLKEINFSSNQIFLVCEEELKPLQGKTLSYFSLAANSL ASRVSVDWSKCMNPFRNMTLEVLDVSGNGWTVDITRKFSRAIAK SNATSLKLAHHIMGAGFGFHNIKDPDEDTFAGLSSSSVKSLDLSHG FVFTLKSNVFRDLTNLEVLNLSYNKINIIEKDAFSGLTNLRYLNLSYN LLGELYSFNFDGLPNVEYIDLSKNHIAIIQDETFSGLSKLKTLNLRDN ALESISNIPKIPYIYLSGNKLTTLPDIELTAVEIHLSENRLENLDILYQLL SVPVLQKLYLGSNRFSKCSGTEKPVENPTLKELFLSDNLLQLAWET EKCWDMFDGLTSLKELDLSGNGLKSLPPGVFSHLTALEKLNLNSN ELTELRHEDLPKNLKKLDISSNKLKAPNPEVFTHLDELDITNNDFVC ECDLKSFIDWLNTTNVTIVGSPEDIYCVEPESLKGTSLFTLSTDGCD GSGSHHWGSTHHHHHH

E12 SGSCSYTGKNAFFRFCNLTQVPKVPNTTERLLLSFNYISTITSSSFS
PLPNLKLLEIGSQYVPLTIDKEAFENLPNLEILDLGSTKIYFLHPDAF
QGLPKLFELRLYFCGLSDSILTDGYFSNLKALTRLDLSKNQITSLKL
HPSFSKLNSLKSIDFSSNQIFEVCEHELEPLQGKTLSHFSLAANSLF
SRSSVDFSKCGNPFRNITIGHFDVSGNGFTVDILGNFSNLISGAQV
FSLNLAHGIFGAGFGFHNIKDPDQNTFSGLARSNVRHLDLSHGFIF
TLNSRVFETLKDLKVLNLSYNKINKIADEAFYGLDNLQVLNLSYNLL
GEIYSSDFYGLPKLAYLDLQKNHIAIIQDQAFKFLEKLQTLDLRDNAL
TTLHFLPSIPDLFLSGNKLVTLPKINLTANLVHLSENRLENLDILFFLL
RIPHLQILILNQNRLSSCSGDQTPSENPSLELLDLSHNMLALAFETG
LCWDIFEGLSHLQSLDLSSNGLNSLPPGLFSHLTALRVLDLSHNRIK
KLEPNDLPASLEVLDISHNKLTTPDPEVFTHLEELDIRHNEIEATPG
GAKFFEFVKNSSAKIIGDLSEIKISGSKEYEGTPLSDLDISKIKGSGS
HHWGSTHHHHHH

F2 MSGSCSYTGKNAFFRFCNLTQVPKVPNTTERLLLSFNYISTITSSSF
SPLPNLKLLEIGSQYVPLTIDKEAFENLPNLEILDLGSTKIYFLHPDA
FQGLPKLFELRLYFCGLSDSILTDGYFSNLKALTRLDLSKNQITSLKL
HPSFSKLNSLKSIDFSSNQIFEVCEHELEPLQGKTLSHFSLAANSLF
SRSSVDFSKCGNPFRNITIGHFDVSGNGFTVDILGNFSNLISGAQV
FSLNLAHGIFGAGFGFHNIKDPDQNTFSGLARSNVRHLDLSHGFIF
TLNSRVFETLKDLKVLNLSYNKINKIADEAFYGLDNLQVLNLSYNLL
GEIYSSDFYGLPKLAYLDLQKNHIAIIQDQAFKFLEKLQTLDLRDNAL
TTLHFLPSIPDLFLSGNKLVTLPKINLTANLVHLSENRLENLDILFFLL
RIPHLQILILNQNRLSSCSGDQTPSENPSLELLDLSHNMLALAFETG
LCWDIFEGLSHLQSLDLSSNGLNSLPPGLFSHLTALRVLDLSHNRIK
KLEPNDLPASLEVLDISHNKLTTPDPEVFTHLEELDIRYNELAATPG
GAKLFEFLKNSKAKIIGDLDDIKIAGSKEYEGTPLSELDIEKIKGSGS
HHWGSTHHHHHH

F11 SGSCSYTGKNAFFRFCNLTQVPKVPNTTERLLLSFNYISTITSSSFS
PLPNLKLLEIGSQYVPLTIDKEAFENLPNLEILDLGSTKIYFLHPDAF
QGLPKLFELRLYFCGLSDSILTDGYFSNLKALTRLDLSKNQITSLKL
HPSFSKLNSLKSIDFSSNQIFEVCEHELEPLQGKTLSHFSLAANSLF
SRSSVDFSKCGNPFRNITIGHFDVSGNGFTVDILGNFSNLISGAQV
FSLNLAHGIFGAGFGFHNIKDPDQNTFSGLARSNVRHLDLSHGFIF
TLNSRVFETLKDLKVLNLSYNKINKIADEAFYGLDNLQVLNLSYNLL
GEIYSSDFYGLPKLAYLDLQKNHIAIIQDQAFKFLEKLQTLDLRDNAL
TTLHFLPSIPDLFLSGNKLVTLPKINLTANLVHLSENRLENLDILFFLL
RIPHLQILILNQNRLSSCSGDQTPSENPSLELLDLSHNMLALAFETG

LCWDIFEGLSHLQSLDLSSNGLNSLPPGLFSHLTALRVLDLSHNRIK
KLEPNDLPASLEVLDISHNKLTTPDPEVFTHLEELDIRDNELELSEG
TARLATLLREGGTRLLSPPESYVLKKSPEYKGTSLLDLDLEKVKGS
GSHHWGSTHHHHHH

H1 MSGSCSYTGKNAFFRFCNLTQVPKVPNTTERLLLSFNYISTITSSSF
SPLPNLKLLEIGSQYVPLTIDKEAFENLPNLEILDLGSTKIYFLHPDA
FQGLPKLFELRLYFCGLSDSILTDGYFSNLKALTRLDLSKNQITSLKL
HPSFSKLNSLKSIDFSSNQIFEVCEHELEPLQGKTLSHFSLAANSLF
SRSSVDFSKCGNPFRNITIGHFDVSGNGFTVDILGNFSNLISGAQV
FSLNLAHGIFGAGFGFHNIKDPDQNTFSGLARSNVRHLDLSHGFIF
TLNSRVFETLKDLKVLNLSYNKINKIADEAFYGLDNLQVLNLSYNLL
GEIYSSDFYGLPKLAYLDLQKNHIAIIQDQAFKFLEKLQTLDLRDNAL
TTLHFLPSIPDLFLSGNKLVTLPKINLTANLVHLSENRLENLDILFFLL
RIPHLQILILNQNRLSSCSGDQTPSENPSLEHLHLSGNMLALAFET
GLCWDIFEGLSHLQSLDLSGNGLNSLPPGLFSHLTALRHLNLSGNR
IKKLEPNDLPASLESLDLSGNKLTTPDPEVFTHLESVNLTGNEFDCS
CAAAKFIDWYKTTSTNLLGDPADRTCTTPEALKGTPIESLSTAGCE
GSGSHHWGSTHHHHHH

H2 MSGSCSYTGKNAFFRFCNLTQVPKVPNTTERLLLSFNYISTITSSSF
SPLPNLKLLEIGSQYVPLTIDKEAFENLPNLEILDLGSTKIYFLHPDA
FQGLPKLFELRLYFCGLSDEILKGPYFAGLKNLEKLDLSKNQATSVE
FGESIAKLENLSYLDLSSNQIFEVCEHEFAALHGKKLDYLSLAANSL
FSRSSVDFSKCGNPFKNITINHLDVSGNGFTVDILGNFSNLISGAQ
VFSLNLAHGIFGAGFGFHNIKDPDQNTFSGLARSNVRHLDLSHGFI
FTLNSRVFETLKDLKVLNLSYNKINKIADEAFYGLDNLQVLNLSYNL
LGEIYSSDFYGLPKLAYLDLQKNHIAIIQDQAFKFLEKLQTLDLRDNA
LTTLHFLPSIPDLFLSGNKLVTLPKINLTANLVHLSENRLENLDILFFLL
RIPHLQILILNQNRLSSCSGDQTPSENPSLELLDLSHNMLALAFETG
LCWDIFEGLSHLQSLDLSSNGLNSLPPGLFSHLTALRVLDLSHNRIK
KLEPNDLPASLEVLDISHNKLTTPDPEVFTHLEELDIRHNEFDCSCA
SARFINWLKTTSTKLLGDPSEYTCSTPASLKGTPLVSLSTAGCEGS
GSHHWGSTHHHHHH

H3 MSGSCSYTGKNAFFRFCNLTQVPKVPNTTERLLLSFNYISTITSSSF
SPLPNLKLLEIGSQYVPLTIDKEAFENLPNLEILDLGSTKIYFLHPDA
FQGLPKLFELRLYFCGLSDEITLSGYFAGLPSLESLDLSKNQISRLK
FHPAFASLKNLRHLDLSSNQIFEVCEHDLAALRGKSLEYLSLAANS
LFSRSSVDFSKCGNPFRDISIEHLDVSGNGFTVDILGNFSNLISGAQ

	VFSLNLAHGIFGAGFGFHNIKDPDQNTFSGLARSNVRHLDLSHGFI FTLNSRVFETLKDLKVLNLSYNKINKIADEAFYGLDNLQVLNLSYNL LGEIYSSDFYGLPKLAYLDLQKNHIAIIQDQAFKFLEKLQTLDLRDNA LTTLHFLPSIPDLFLSGNKLVTLPKINLTANLVHLSENRLENLDILFFLL RIPHLQILILNQNRLSSCSGDQTPSENPSLELLDLSHNMLALAFETG LCWDIFEGLSHLQSLDLSSNGLNSLPPGLFSHLTALRVLDLSHNRIK KLEPNDLPASLEVLDISHNKLTTPDPEVFTHLEELDIRHNDFDCSCK SAKFISWLKTTSTKLLGDPSDYTCVSPSSLKGTPLISLSTSGCEGS GSHHWGSTHHHHHH
H4	SCSYTGKNAFFRFCNLTQVPKVPNTTERLLLSFNYISTITSSSFSPL PNLKLLEIGSQYVPLTIDKEAFENLPNLEILDLGSTKIYFLHPDAFQG LPKLFELRLYFSGLSDSILTDGYFSNLKALTRLDLSKNQITSLKLHPS FSKLNSLKSIDFSSNQIFEVCEHELEPLQGKTLSHFSLAANSLFSRS SVDFSKCGNPFRGITIGHFDVSGNGFTVDILGNFSNLISGAQVFSL NLAHGIFGAGFGFHNIKDPDQNTFSGLARSNVRHLDLSHGFIFTLN SRVFETLKDLKVLNLAYNKINKIADEAFYGLDNLQVLNLSYNLLGEIY SSDFYGLPKLAYLDLQKNHIAIIQDQAFKFLEKLQTLDLRDNALTTLH FLPSIPDLFLSGNKLVTLPKINLTANLVHLSENRLENLDILFFLLRIPHL QILILNQNRLSSCSGDQTPSENPSLELLDLSHNMLALAFETGLCWD IFEGLSHLQSLDLSSNGLNSLPPGLFSHLTALRVLDLSHNRIKKLEP NDLPASLEVLDISHNKLTTPDPEVFTHLEELDIRYNSYDCGCESATF INWLNHTNVTIKGDPSDYVCTSPPSLKGVPLFSLSTEE
F1 Flagellin mimetic	KFDINYNKELEEEKENLNKSQSSLSSAIERLSSGCRINSAKDDAAG QAIANRFTSNIKGLTQASRNANDGISIAQTTEGALNEINNNLQRVRE LSVQATNGTNSDSDLKSIQDEIQQRLEEIDRVSNQTQFNGVKVLSQ DNQMKIQVGANDGETITIDLQKIDVKSLGLDGFNVNGSGSTANPLA SIDSALSKVDAVRSSLGAIQNRFDSAITNLGNTVTNLNSARSRIECA DYATEVSNMSKAQILQQAGTEKLKEAEEKIEKYKELLKGGSGLNDI FEAQKIEWHE
F2 Flagellin mimetic	NLNKSQSSLSSAIERLSSGCRINSAKDDAAGQAIANRFTSNIKGLT QASRNANDGISIAQTTEGALNEINNNLQRVRELSVQATNGTNSDS DLKSIQDEIQQRLEEIDRVSNQTQFNGVKVLSQDNQMKIQVGAND GETITIDLQKIDVKSLGLDGFNVNGSGSTANPLASIDSALSKVDAVR SSLGAIQNRFDSAITNLGNTVTNLNSARSRIECADYATEVSNMSKA QILQQAGTGGSGLNDIFEAQKIEWHE
F14 Flagellin	RINHNIAALNTKNRLGSNNGAAQKNMEKLSSGLRINRAGDDAAGL ALSEKMRGLIRGLEQASKNSQDGISLIQTAEGALTRTHAILQRMREL

mimetic	TVQAGNTGTQQEEELGAIKDEMDALIEEIDGISNRTEFNGKKLLDG TNSTDGFTFQIGALAGQQLNVKIDSMSSTALGVNALDVTDFAATAF DDQLKSIDTAINTVSTQIAKLGAVINRLEHTISNLGASGENLTAAESSI RDVDMEKMREEFDKNNKLTEQAEKDLKKANEIPINELKKLEGGSG LNDIFEAQKIEWHE
F22 Flagellin mimetic	RLNHNIKALNLQNRLGSANGAAQKNEEKLSSGCRINRAGDDAAGL ALSEKMRGLIRGLEEASKNSQDGISLIQTAEGALTETHAILQRMKEL TVQAGNTGTQQPEQLGAIKDEMDALIERIDGISNRTTFNGKKLLDG TNSTDGFTFQIGAEAGQQLNVKIDSMSSTALGVNALDVTDFAATAF DDQLKSIDTAINTVSTQQAKLGAVINRLEHTIRNLGASGENLTAAES KIQCVDMEKVQEELDKQNEEIKQAQKELEEAQKEPLNLLELLKGG SGLNDIFEAQKIEWHE

*Flagellin mimetic sequences provided by Naveen Jasti and Chloe Adams

Chapter 5. De novo design of TLR1 minibinders

Abstract

TLR1 is one of three receptors confirmed to dimerize with TLR2. It is required for the downstream activation of the TLR2. From this we inferred minibinders that were capable of binding to TLR1 could be linked to TLR2 to force the activation of TLR2. Computational design of minibinders against TLR1 yielded 8 successful binders ranging in binding affinity from 130nM to 1000nM. Of the eight identified, 1-4 (Kd 460nM) was used in subsequent studies to induce dimerization of TLR1/2.

Introduction

TLR1 is a member of the human TLR family, which consists of 10 receptors (TLR1–TLR10) that detect PAMPs and initiate innate immune responses[1–3,91]. Within the broader TLR family, TLR1 contributes to host defense against Gram-positive bacteria, mycobacteria, and certain parasites, emphasizing its role in sensing extracellular microbial components and shaping the early immune response[6,38,118,119]. Unlike some other TLRs that can signal as homodimers (e.g., TLR4 or TLR5), TLR1 requires TLR2 for ligand recognition and signal transduction. Together, the TLR1/2 complex activates the MyD88-dependent pathway, leading to proinflammatory cytokine production[5,6,49,51,53]. There is no indication that TLR1 or TLR2 can activate as homodimers[61], thus heterodimerization is required for activation of innate immune pathways. This proximity enables the recruitment of TIR-domain-containing adaptor proteins, particularly MyD88, through homotypic TIR-TIR interactions[1,5,6].

To harness the immunostimulatory potential of TLR1/2 signaling for use in vaccine adjuvant development, it is necessary to artificially induce receptor dimerization at the cell surface, mimicking pathogen-induced activation[13,17,69,110]. This strategy enables controlled activation of innate immune pathways, particularly NF-κB-mediated transcription of pro-inflammatory cytokines, which are essential for shaping robust adaptive immune responses[53,120,121].

The development of de novo minibinders has aided in careful modulation of the immune system as is required with adjuvant development[26,27,110,111]. Minibinders can be rationally designed to target specific protein surfaces or conformations, enabling precise modulation of protein-protein interactions, enzyme activity, or receptor signaling pathways[26,27,32]. For example, minibinders can act as agonists or antagonists by stabilizing particular conformational states of receptors, as shown in synthetic activation of immune receptors like TLRs[110].

Here, we designed minibinders targeting both TLR1 and TLR2, based on prior studies demonstrating that flexibly linked binding domains can effectively induce receptor dimerization and downstream signaling[110]. Following the successful generation of TLR2-specific minibinders (see Chapter 2), we engineered complementary TLR1-binding minibinders with the aim of forming a synthetic TLR1/2 dimer upon fusion.

Results

Computational design of TLR1 minibinders

Unlike the approach used for TLR2 minibinders, we employed the RifDock[122] pipeline to perform the initial docking of binder scaffolds against this hydrophobic region. Following docking, ProteinMPNN[29,30] was used to generate thousands of sequence variants for each scaffold set (Fig. 1a). Since TLR1 is not known to function independently[2,5], it was essential to identify binding sites that would not disrupt its interaction with TLR2[6,61,108]. Accordingly, regions involved in the TLR1/2 dimerization interface and the Pam3CSK4 ligand-binding site were excluded from consideration[6,61]. Instead, we selected a hydrophobic patch located at the distal N-terminus of TLR1, characterized by several large aromatic residues favorable for de novo binder design (Fig. 1b).

These designs were subsequently evaluated using AF2 to predict structure and binding compatibility. Although experimentally testing the entire set is theoretically feasible, computational filtering significantly enhances the likelihood of identifying effective binders while conserving experimental resources. A total of 13,500 designs were input into AF2, and filtering was performed based on two key criteria: a minibinder pLDDT > 80, indicating confident and reliable folding, and a PAE_interaction < 8, suggesting high-confidence interface predictions between TLR1 and the minibinder. These filters reduced the initial pool to 1,266 high-confidence candidates (9.4%), which were synthesized on a DNA chip for downstream screening and validation.

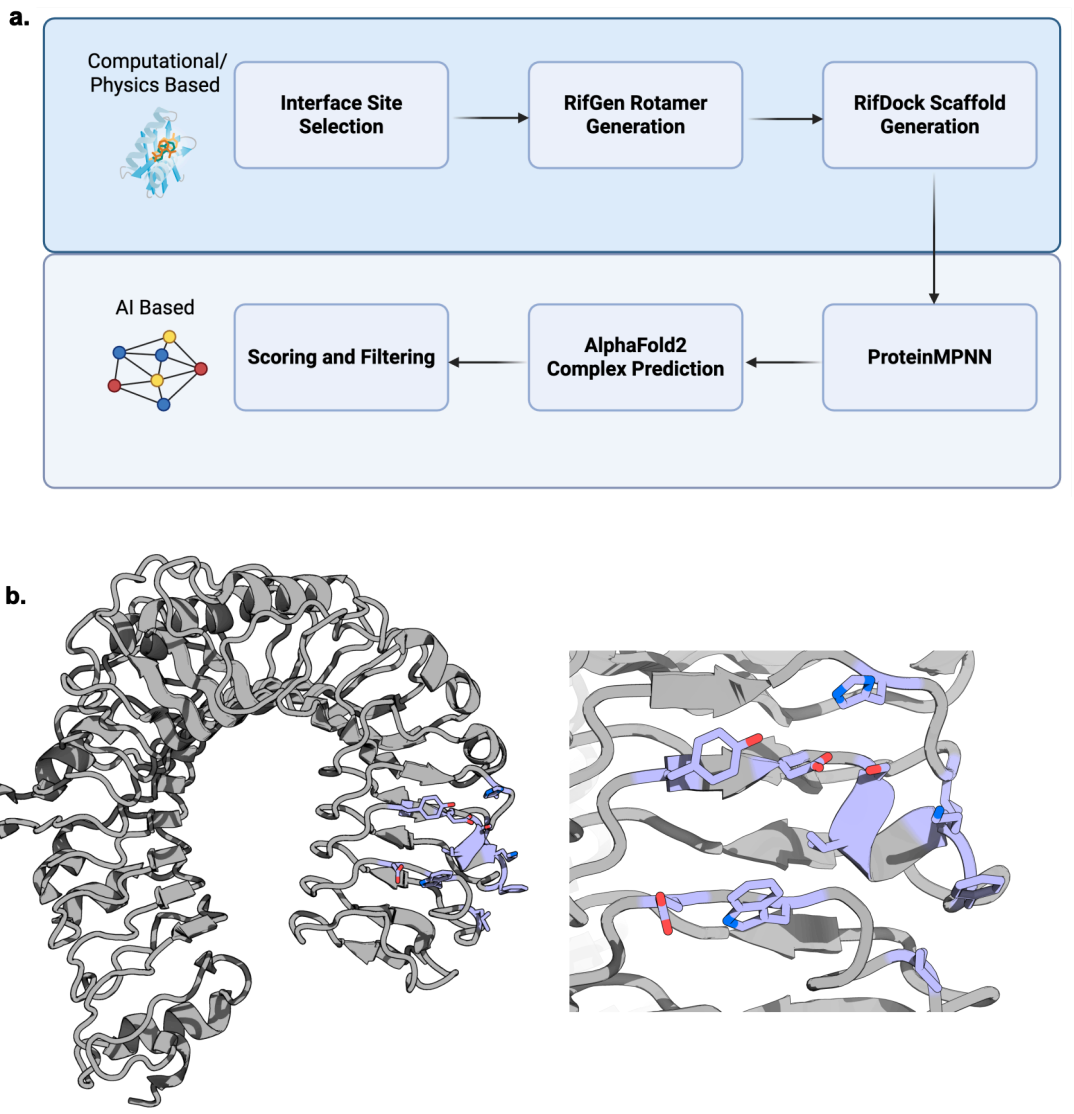


Figure 1. Computational Design of TLR1 minibinders

a, Schematic depicting the pipeline of the physics based (RifGen, RifDock) approach coupled with AI approaches for design of TLR1 minibinders. **b,** TLR1 (PDB: 6NIH) depicting the N-terminal interface chosen to develop de novo minibinders against TLR1.

Sorting of minibinders that bind TLR1

Designed sequences were cloned into a yeast surface display[109] expression vector for screening via FACS. Binding was assessed using 1 μM of commercially available human TLR1-Fc protein, complexed with a Protein A–streptavidin fluorescent probe. Following two rounds of sorting, a distinct population of double-positive cells emerged,

indicating successful enrichment of TLR1-binding clones (Fig. 2a). Sequencing of this population identified seven unique minibinders from the original pool of 1,266 designs (Fig. 2b).

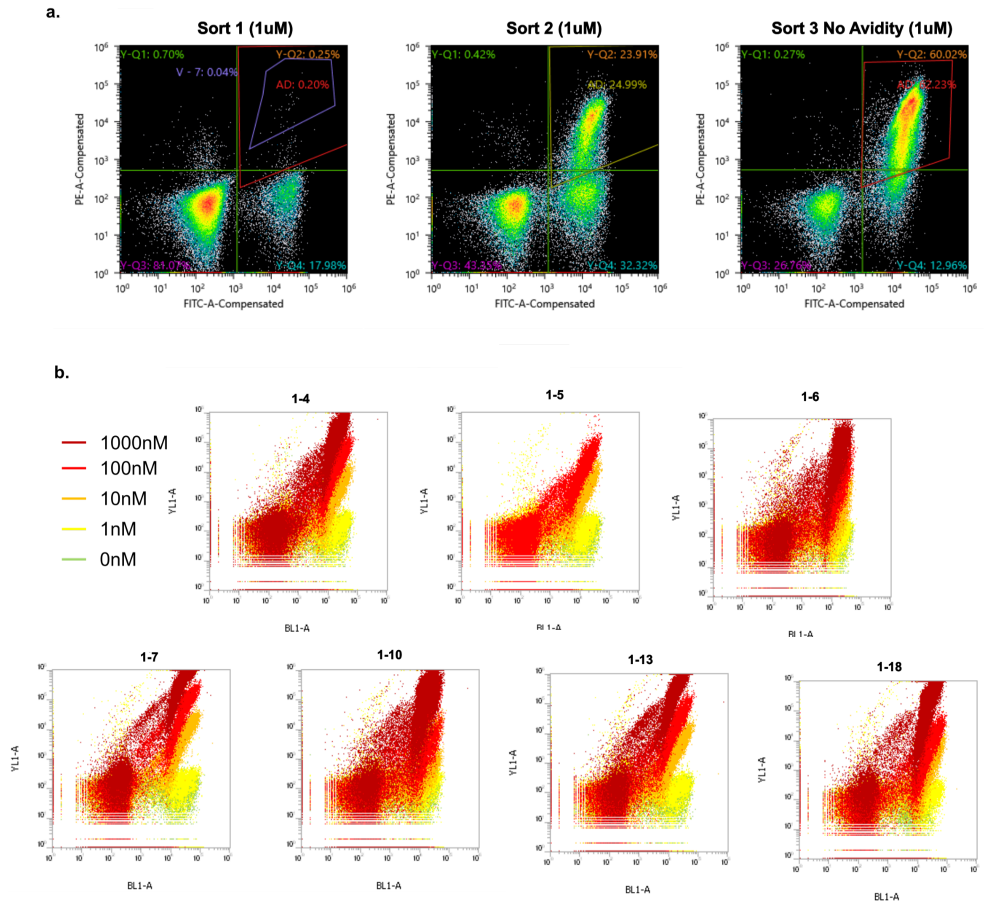


Figure 2. Yeast Display of TLR1 minibinders.

a, 1,266 designs were tested on Yeast Display with TLR1-Fc at 1uM. After 2 sorts, a large population of TLR1-Fc binders emerged. After a third sort with no avidity this population was confirmed. **b,** After sequencing seven binders were identified and screened with a 10x dilution of TLR1-Fc, each showing reducing affinity after 100nM.

Biochemical characterization of TLR1 minibinders

Each of the binders was identified to bind in a similar way to the N-terminus of TLR1 (Fig. 3a). The identified binders were subsequently expressed in *E.coli* with elution volumes matching SDS PAGE sizes (Fig. 3b) and biotinylated for binding affinity measurements using BLI. Five of the seven candidates exhibited moderate binding to TLR1, with dissociation constants (Kd) ranging from 130 nM to 1 µM (Fig. 3c). The remaining two binders showed weaker binding signals under the same conditions. While

these results confirm the ability of the designed binders to interact with TLR1, further optimization will be necessary to improve binding affinity and functional performance.

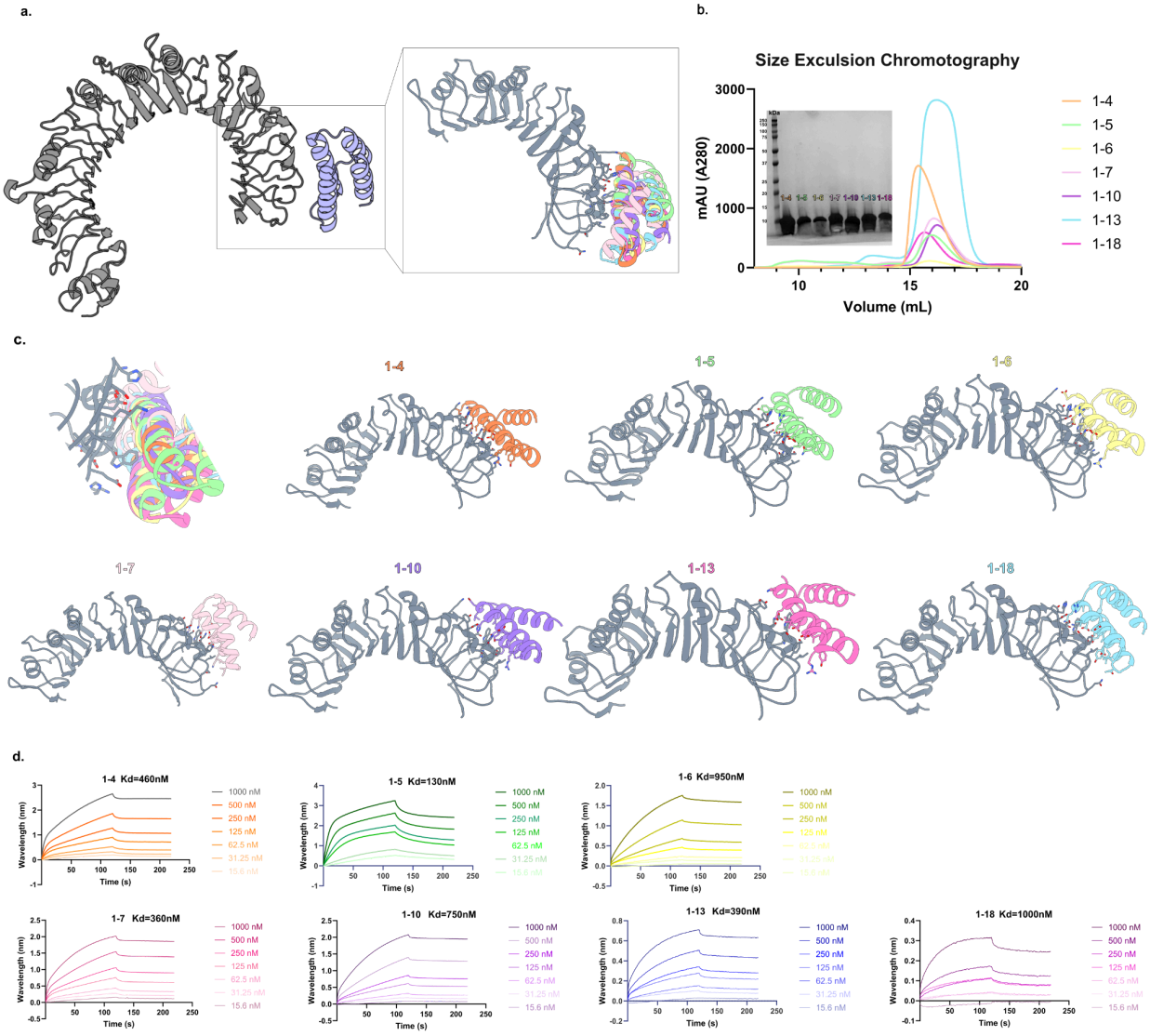


Figure 3. Biochemical Characterization of TLR1 minibinders

a, Computational model showing location of all 7 binders at the N-terminal domain of TLR1. **b,** SEC traces and SDS-PAGE of minibinders at correct size and their expression levels. **c,** Minibinder specific view of each complex with TLR1. **d,** BLI data for each minibinder with Kd values. Binders were immobilized and a 2x serial dilution of TLR1-Fc was performed to obtain binding affinities.

Flexibly linking TLR1 and TLR2 minibinders for activation of TLR1/2

Following determination of binding affinities, we next evaluated whether genetic fusion of the TLR1 minibinders to previously characterized TLR2 minibinders could induce TLR1/2 activation. Among the candidates, minibinder 1-5 (shown in green) exhibited the highest affinity; however, concerns regarding expression efficiency, particularly when expressed as fusion constructs (Fig. 3b), influenced our selection. Minibinder 1-4 was chosen for further studies due to its relatively high expression level—second only to the top-expressing variant 1-13, despite having slightly lower affinity (Fig. 3b, d).

To physically bridge the distance between the N-terminal binding site on TLR1 and the target site on TLR2, a flexible 24-residue Gly-Ser (GS) linker was employed to connect the two minibinders (Fig. 4a)[110]. To assess functional activity, we used the HEK-Blue™ cells stably expressing TLR1 and TLR2, and the SEAP reporter system[71,104]. Initial experiments demonstrated that fusion constructs comprising either A1–1-4 or G9–1-4 induced measurable NF-κB activation. Notably, the A1–1-4 fusion elicited a stronger response than 100 ng/μL Pam3CSK4, a known TLR1/2 agonist, whereas G9–1-4 induced a more modest, but still above-background, level of activation (Fig. 4b). Interestingly, G9 has previously been characterized as an antagonist in the context of Pam3 stimulation (see Chapter 3), suggesting it may exhibit context-dependent agonist or antagonist activity when genetically fused to a TLR1 binder.

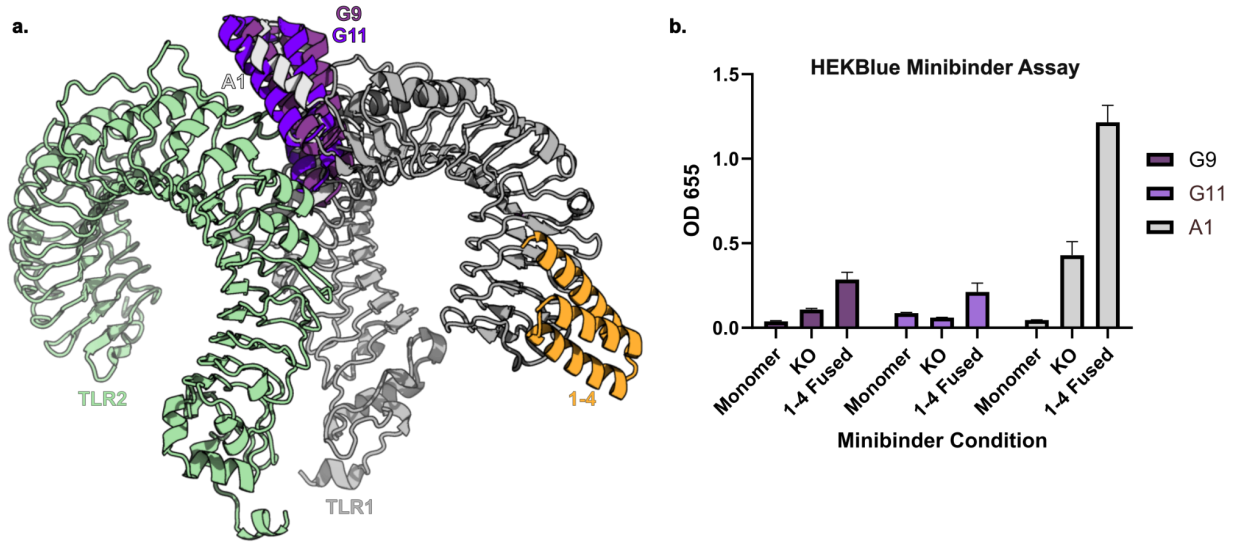


Figure 4. Biochemical Characterization of TLR1 minibinders

a, Computational model showing location of G9 and G11 (see Chapter 3) bound to TLR2 aligned with 1-4 at the N-terminal domain of TLR1 (GS linker not shown). **b,**

HEK-Blue™ assay showing high activation of A1 linked minibinders and modest activation of G9 and G11 from GS linking to TLR1 minibinder.

Discussion

While the current findings are encouraging, several limitations remain that warrant further investigation. Most notably, the absence of TLR1 or TLR2 knockout (KO) controls limits our ability to definitively attribute observed activation to specific receptor engagement rather than confounding factors such as cellular stress[26,27,110]. Although monomeric minibinder controls suggest that off-target activation is unlikely, validation using KO cell lines will be essential to confirm the specificity of the observed responses.

Beyond establishing specificity, further work is required to characterize the parameters that modulate TLR1/2 activation. In particular, delineating the dose–response relationship will be important to evaluate how these synthetic constructs can tune immune signaling in a physiologically relevant manner. Parameters such as linker length, minibinder pairing, and the stoichiometry of fusion constructs may all influence receptor clustering and activation dynamics, offering a platform to systematically dissect the mechanistic underpinnings of TLR1/2 signaling[110].

Finally, the current minibinders have not yet been validated for cross-reactivity with murine TLR1 or TLR2, posing a significant challenge for in vivo studies in mouse models[27,111]. The absence of a published structure for murine TLR1 (mTLR1), along with the limited availability of high-quality protein, hinders structure-guided design and experimental validation[6,61]. Future efforts will focus on engineering cross-reactive variants, leveraging computational tools such as AlphaFold3 to model mTLR1 and guide binder redesign. Stabilization efforts discussed in Chapter 2 and 4 could also be applied to stabilize TLR1 for development of TLR1 in both mouse and human species. These challenges underscore the broader need for stabilized Toll-like receptor reagents to enable the development and translation of synthetic immunomodulatory tools.

Acknowledgments

Thank you to Brain Coventry, Nathaniel Bennet and Justas Dupras for leading development of RifDock and MPNN minibinder pipelines. Thank you to Inna Goreshnik, Aza Allen, Cami Cordray, Samer Halabiya, and Dionne Vafeados for help with yeast library preparation. Thanks to Craig Dobbins and Adri Tran-Pearson for their contribution to cell based assays.

Materials and Methods

Computational design

The ectodomain of TLR1 (PDB ID: 6NIH)[61] was used for the design. The PDB was truncated to save computational time. Minibinders were designed as outlined in ref[122]**.** In brief, RifGen[35], PatchDock[123], RifDock[35], were used. Hundreds of scaffolds were created using this pipeline then fed into ProteinMPNN[30] for sequence design, then predicted and filtered on various AF2 metrics (pae_interaction, binder_plddt)[29].

Library preparation

Designs were padded to 65 amino acids with serine. Then, the designs were codon-optimized for expression in *Saccharomyces cerevisiae*. Oligonucleotides encoding the designs and SSM mutations were purchased from Agilent technologies. Libraries were amplified as previously described[122].

Yeast surface display

Minibinders were displayed on the surface of yeast with a C-terminal myc tag to enable detection of cell surface expression using a FITC-labeled anti-myc antibody. hTLR1-Fc ectodomain was recognized by ProA fused to streptavidin-PE. An initial sort was done to collect cells that express the designs (FITC+). These cells were subsequently sorted against decreasing concentrations of hTLR1-Fc with and without avidity.

The *Saccharomyces cerevisiae* EBY100 strain was cultured in C-Trp-Ura medium containing 2% (w/v) glucose (CTUG). For induction, cells were harvested by centrifugation at 5,000 g for 5 minutes and subsequently resuspended in SGCAA medium supplemented with 0.2% (w/v) glucose at a cell density of 1×10^7 cells per mL. The cells grew at a temperature of 30°C for 16-24 hours. After induction, the cells were washed using PBSF (PBS with 1% (w/v) BSA). The cells were labeled with anti-c-Myc fluorescein isothiocyanate (FITC, Miltenyi Biotech) and streptavidin-phycoerythrin (SAPE, ThermoFisher).

During the initial sorting of the naive library, an avidity-based approach was employed. This entailed simultaneous incubation of the biotinylated target, SAPE, and FITC. In contrast, sorts performed without avidity involved a sequential procedure. Specifically, cells were incubated with the biotinylated target, washed with PBSF, and then incubated with SAPE and FITC.

Biolayer interferometry

BLI was performed on an Octet RED96 or Octet R8. All biosensors were hydrated in kinetics buffer HBS-EP+ (10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% v/v surfactant P20) with 0.5% w/v non-fat dry milk (Cytiva). Mibnibinders were biotinylated

and excess biotin was purified out by SEC. Biotinylated minibinders were diluted to a final concentration of 0.01 mg/mL in kinetics buffer and loaded onto streptavidin biosensors (Sartorius). hTLR2 were diluted in kinetics buffer to 250 nM and its association was measured for 300 s, followed by a dissociation for 300 s in kinetics buffer.

HEK-Blue™ agonism assays

HEK-Blue™ agonism assays are described in Chapter 3 of this work. Cells were incubated as either monomers, or GS linked dimers to TLR2 minibinders characterized in Chapter 3.

Supplementary Table 1. Amino acid sequences

Construct*	Amino Acid Sequence
1-4	AEEIEKAIKEIEELGFPDLAEKARELYKLYKETGDERYLEAVKIDLEIA KKKKEL
1-5	SIIDRGYERLKELLKEGNFEAIEIDLRQLEYHAKELNDPEALKKVEEI KEALAKLK
1-6	MEEVAEKAEELVKKLTDKYRRDKLEELLARLKATGDPRLAEALAIDL RIAEE
1-7	GELEEKVESKIKELLKAGKIEEAKKLADLWDEYKRSGKEEIKKKLEE KLKEV
1-10	AEARERARRLIEETRAAIAAGDIETAHRLESEALELALAVGDQELID ELRRAIEEAR
1-13	MELEERVRELSAKAREAGKPWIADSLEDLIEEYKKTGDPRYKEALE IDAAQAEKVLE
1-18	LERFDRLLEVARAEDKPKLVEVIEYLKSLVEEGTITEEEALEAALIDA RQVGDPRTEALIREELA
A1/1-4 dimer	MEEAKKKAEEVKPKAETDLYARLEYLRYMYAVAVAEGDKKKAEELE KEAKKVYEELIKKLEGSGGGSGSGSGGSGSSGSGSGSGGAEEIE KAIKEIEELGFPDLAEKARELYKLYKETGDERYLEAVKIDLEIAKKKK EL
B2/1-4 dimer	AEAEEYKAKAESLKPKLDDPLYQNKYYAYLYHYYKLLGDKEKAAEY EEKAKAAYEKIVEEIEGSGGGSGSGSGGSGSSGSGSGSGGAEEIE

	KAIKEIEELGFPDLAEKARELYKLYKETGDERYLEAVKIDLEIAKKKK EL
F6/1-4 dimer	EAEEHLAKAEEYKAKSSTLLEKWLAAAHLYAAYKELGDEEKAKEAE KKAKELYEELVKSIEGSGGGSGSGSGGSGSSGSGSGSGGAEEIE KAIKEIEELGFPDLAEKARELYKLYKETGDERYLEAVKIDLEIAKKKK EL
G8/1-4 dimer	NPVEAYEAYKEALKKLYEKGDEKALKELEPKLKELFEKALKAIEAGD LEGAAKYFEEAAKLVEEALGSGGGSGSGSGGSGSSGSGSGSGG AEEIEKAIKEIEELGFPDLAEKARELYKLYKETGDERYLEAVKIDLEIA KKKKEL
G9/1-4 dimer	SPVEAYNRYREALKKLYEKGDKKALEELDPKLKELYEKALEAIEKG DLEGAAKYFQEAADLVEKAIGSGGGSGSGSGGSGSSGSGSGSG GAEEIEKAIKEIEELGFPDLAEKARELYKLYKETGDERYLEAVKIDLEI AKKKKEL
G11/1-4 dimer	SLVAALEAYREALKALYDAGKEEELKAIEPELKALGEKFLAAIEAGD KEGAIAYLQAATDLVKKAAGSGGGSGSGSGGSGSSGSGSGSGGA EEIEKAIKEIEELGFPDLAEKARELYKLYKETGDERYLEAVKIDLEIAK KKKEL

Chapter 6. De novo design of TLR4 minibinders and antibodies

Abstract

TLR4 homodimerizes with its coreceptor MD2 in the presence of Lipopolysaccharide (LPS) on both the surface and endosome of cells to induce the MyD88 and TRIF pathways respectively[5,25,50,124,125]. Much like the previously described desire to create minibinders against TLR1 and 2 to induce activation of NF-κB, we developed de novo minibinders at the N-terminal region similar to that of TLR1 (see Chapter 5). Using the same approach as mentioned above for TLR1 minibinder development, we identified two binders capable of binding at relatively low affinity (kD of 300nM) to TLR4. Once successful in minibinder development, we sought to use the same location to create de novo antibodies against TLR4. This is a novel approach to development of antibodies against TLRs which differs from the methods discussed above that forgoes the need for high expressing protein for mouse antibodies[42,126]. While antibodies scored well and looked promising in silico, first screens did not yield successful binders to TLR4.

Introduction

TLR4 is a member of the TLR family and shares structural similarities with other TLRs[1–3,127]. It contains approximately 20 LRRs flanked by N-terminal and C-terminal capping domains[5,25,50]. In the presence of its canonical ligand, LPS, TLR4 forms a complex with MD-2, which facilitates receptor dimerization and subsequent signaling[25,50,124]. Unlike some TLRs, TLR4 signaling strictly requires this heteromeric complex formation with MD-2 and does not function effectively as a homodimer without ligand-induced structural reorganization as it contains a hydrophobic pocket that specifically accommodates the lipid A moiety of LPS, which is the bioactive portion of LPS[5,25].

TLR4 activation has played a pivotal role in modern adjuvant development because it powerfully stimulates both innate and adaptive immunity—making it ideal for boosting vaccine efficacy[13,16,17,128–130]. Leveraging the TLR4 pathway allows adjuvants to mimic PAMPs, triggering robust immune responses while avoiding actual infection[16]. Several notable TLR4-based adjuvants have been developed to enhance vaccine efficacy by safely harnessing innate immune activation. Monophosphoryl lipid A (MPL or MPLA), a detoxified derivative of lipopolysaccharide, is the most widely used and is featured in licensed adjuvant systems such as AS04 (MPLA + aluminum salt), and AS01 (MPLA + QS-21 in liposomes)[16,17,128,129]. Additionally, synthetic TLR4 agonists like GLA-SE are under clinical development for vaccines targeting tuberculosis, HIV, and cancer[131,132]. These adjuvants stimulate strong humoral and cellular immunity while maintaining a

favorable safety profile through TRIF-biased signaling[1,50,94]. Prior work[110] has demonstrated that flexibly linking minibinders can effectively induce dimerization and activation of immune receptors which could serve as a protein based method for activation of TLR4 for use in adjuvants.

In this study, we developed two candidate minibinders targeting TLR4 at a site analogous to the previously characterized TLR1 minibinder interface (see Chapter 5). These TLR4 binders exhibited moderate affinities, ranging from 330 to 400nM. Based on this binding site, we proceeded to design de novo single-chain variable fragments (scFvs) against TLR4 using computational modeling. Although these antibodies received high in silico affinity scores, they did not demonstrate detectable binding to TLR4 when evaluated by yeast surface display.

Results

Computational design of TLR4 minibinders

Minibinders targeting TLR4 were generated employing the same methodology previously described for TLR1 minibinders[30,122]. Initial scaffold generation was performed using the Rifdock pipeline[122], which identified four putative binding sites on TLR4 (Fig. 1a). These included one at the N-terminal domain (Fig. 1b); corresponding to the binding region successfully exploited for TLR1 minibinders; one situated on an upper looped segment (Fig. 1c), and two regions within the inner C-terminal domain (Fig. 1d). Following scaffold generation, sequence design was conducted using ProteinMPNN[30], yielding approximately 165,000 candidate designs across the identified sites. These designs were subsequently evaluated using AF2[33], with filtering criteria adjusted to accommodate the lower predictive scores observed for this target. Specifically, the pae_interaction threshold was relaxed from < 8 to 12, while maintaining a pLDDT score of > 80 for the minibinder regions[29]. Application of these criteria resulted in a refined pool of approximately 11,767 designs, which were advanced for further characterization using the methods detailed in Chapter 5.

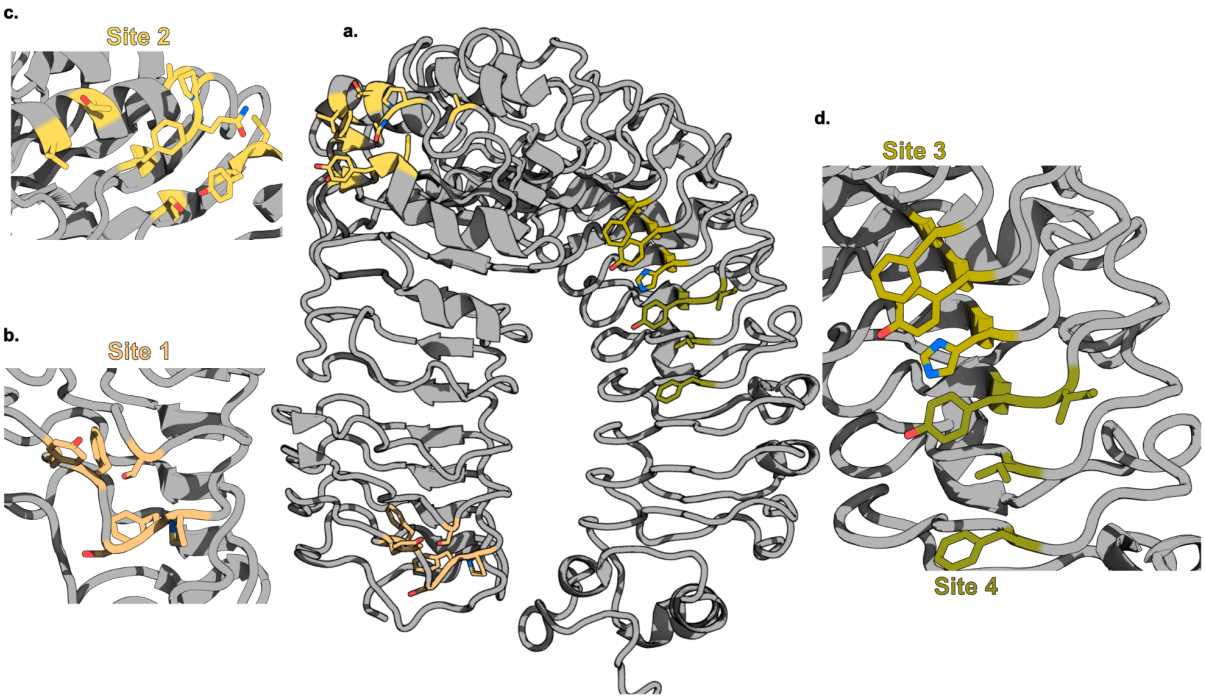


Figure 1. Site selection of TLR4 minibinders

a, TLR4 (PDB: 3FXI)[25] PDB model showing all four sites for minibinder design. **b,** Site 1 is a the first 4 LRR domains and is similar to the interface of TLR1 minibinders. **c,** Site 2 is located at the top of TLR4 where many aromatic residues reside on the surface. **d,** Sites 3 and 4 occupy a similar region of the concave portion of the TLR, due to the shape of the TLR many hydrophobic residues are packed into this region.

Sorting of minibinders that bind TLR4

As previously described, designed sequences were cloned into a yeast surface display[133] expression vector for screening via FACS. Binding was assessed using 1μM of commercially available human TLR4 protein with a 10xHis tag on the C-terminus. This is complexed with an antiHis antibody fused to a streptavidin fluorescent probe. Following four rounds of sorting, a distinct population of double-positive cells emerged, indicating successful enrichment of TLR4-binding clones (Fig. 1a). Sequencing of this population identified two unique minibinders from the original pool of 11,767 designs capable of binding via yeast surface display at 31nM (Fig. 1b).

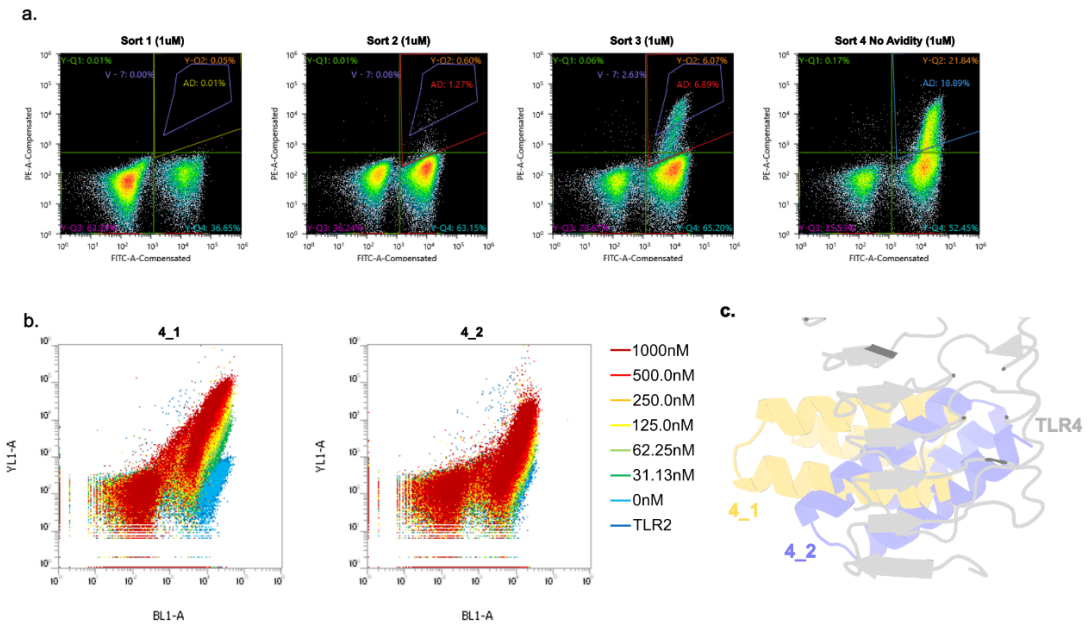


Figure 2. Yeast display of TLR4 minibinders

a, 11,000 designs were tested on yeast display with TLR4 at 1uM. After 3 sorts, a large population of TLR4 binders emerged. After a third sort with no avidity, this population was confirmed **b,** Once sequenced, two binders were identified. Binding was re-confirmed with a screen of yeast display using a 2x serial dilution of TLR4 against each binder. **c,** Model overlapping both binders bound to TLR4 showing the difference in binding modality to the interface.

Biochemical characterization of TLR4 minibinders

Similar to the observations made with TLR1, the identified minibinders targeting TLR4 were found to engage the N-terminal domain (Fig. 3c). Once identified, both binders were expressed in *E.coli* and purified, with elution profiles consistent with their predicted molecular weights (Fig. 3d). Although the binding sites are analogous, the binding modes differ between the two designs. The first binder, designated 4_1 (shown in yellow), exhibited an extended interaction along the length of TLR4 and demonstrated a strong dependence on a single residue, Phe75. In contrast, the second binder, 4_2 (shown in blue), adopted a more vertically oriented binding mode, relying on interactions with both Phe75 and Phe54 (Fig. 4b). The proteins were subsequently biotinylated and assessed for binding affinity using BLI. Both binders demonstrated moderate affinity for TLR4, with Kd's of 330nM and 570μM, respectively (Fig. 4c). Consistent with the initial TLR1 minibinders, further optimization through site-saturation mutagenesis (SSM) will

be required to enhance binding affinity and facilitate their application in downstream functional assays[122].

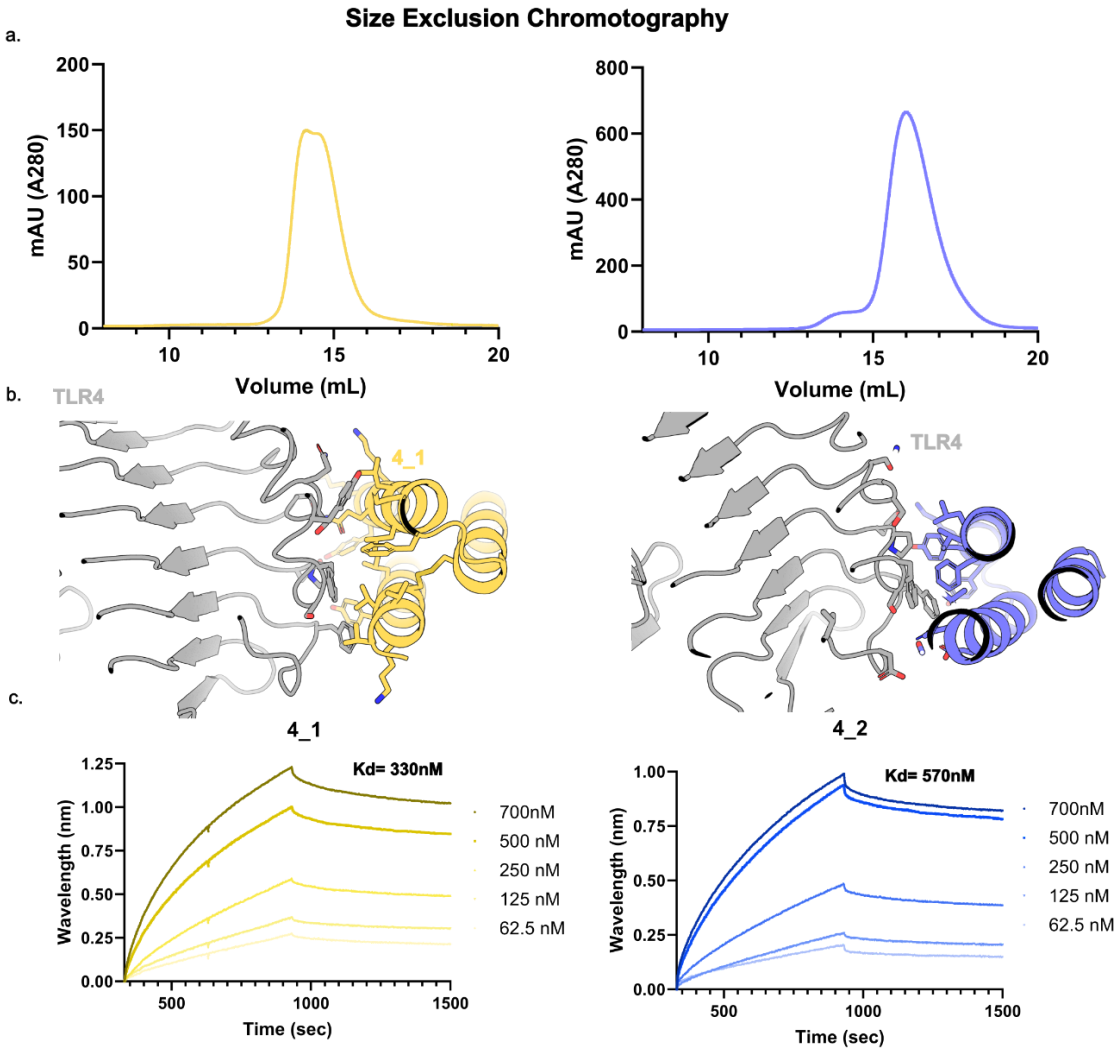


Figure 3. Biochemical characterization of 4_1 and 4_2

a, SEC traces of 4_1 and 4_2 eluting on at the correct volume with high levels of expression. **b,** Up close model displaying side view of the binding interface between each binder. **c,** BLI of both binders with immobilized binder and a 2x serial dilution of 500uM of TLR4 after an initial concentration of 700uM.

Computational design of TLR4 antibodies with RFDiffusion

De novo antibody development represents a promising avenue for advancing therapeutic design by enabling precise targeting of specific epitopes to elicit desired biological responses[36,134,135]. Although conceptually related to de novo minibinder generation, the development of antibodies using fully AI-driven approaches involves

several distinct considerations. This process entails the design of both the antibody backbone, which must be structurally compatible with the target epitope, and the hypervariable loop regions, which mediate specific epitope recognition (Fig. 5a)[135]. Following the structural definition of the backbone and loop regions in the context of the target epitope, ProteinMPNN is employed to generate a diverse set of candidate sequences for the variable regions. These candidate antibody sequences are subsequently evaluated and refined using AF2, which facilitates structural prediction and enables filtering of the final designs based on predicted conformational and binding properties[36,135].

To initiate the development of antibodies targeting TLR4, we selected the previously characterized interface that had yielded effective binders. This epitope comprises two surface-exposed phenylalanine residues and one tyrosine, all located within three consecutive LRR domains (Fig. 5b). The designed single-chain variable fragments (scFvs) engage the first four LRR domains at the N-terminus of TLR4 (Fig. 5c), exhibiting a binding mode analogous to that observed in the 4_1 construct. A notable advantage of scFvs over minibinders is their enhanced shape complementarity to the TLR4 surface. From a top-down perspective (Fig. 5d), the scFv loop regions conform closely to the curvature of TLR4, while a side view reveals that the bent geometry of the Fab-like regions allows for more extensive interaction compared to the limited contact area of the dual α-helical minibinder scaffold (Fig. 5d). Candidate designs were filtered based on predicted interaction quality (PAE_interaction < 10) and the spatial proximity of loop regions to the target epitope, as assessed by root mean square deviation (RMSD_loop < 1 Å) (Fig. 5a). The RMSD criterion ensures that the predicted loop conformations align with the intended epitope in a structurally consistent manner[135].

Given the relatively large size of scFvs and the associated challenges in high-throughput antibody production, we selected the top 60 candidate designs for expression using an in-house yeast surface display platform. Despite successful cloning, none of these designs exhibited detectable binding to TLR4 under the conditions tested (data not shown). Efforts are currently underway to optimize the screening platform to enable more robust and scalable de novo antibody characterization in future studies.

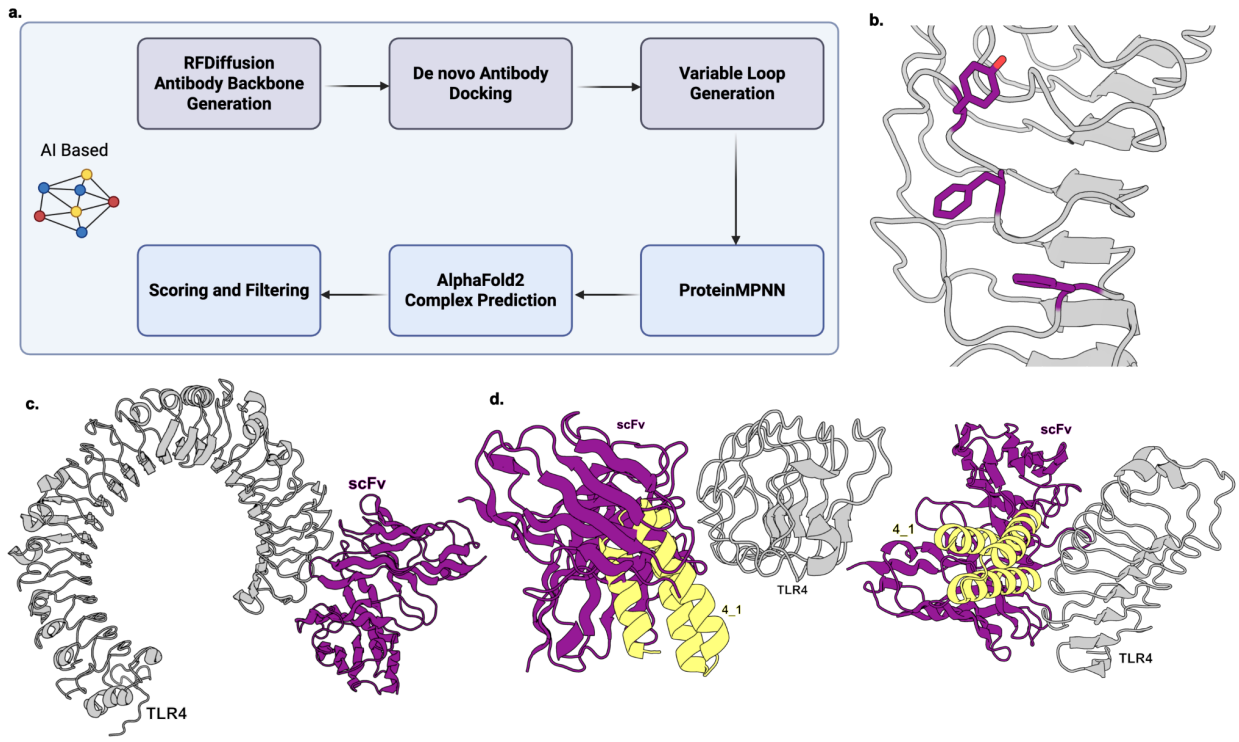


Figure 4. Computational Design of TLR4 de novo scFv

a, Schematic depicting the pipeline of the AI approach for design of scFvs utilizing RFDiffusion for backbone generation and CDR loop docking. **b,** TLR4 (PDB: 3FXI) depicting the N-terminal interface chosen to develop de novo antibodies. This region corresponds to TLR1 and TLR4 minibinder interfaces. **c,** Whole view of TLR4 bound to modeled scFV. **d,** Comparison of modeled scFv binding vs. 4_1 bound to TLR4. While larger, the scFv has far better shape complementarity to the curvature of the TLR.

Discussion

In this study, the Rifdock[36,122] pipeline was successfully employed to generate de novo minibinders targeting TLR4. The resulting binders exhibited moderate affinity toward TLR4 and represent promising candidates for further optimization to enhance binding potency. Consistent with previous observations, the N-terminal domain of TLR4 was confirmed as a conserved and accessible site for the development of both minibinders and antibody-based therapeutics.

Conversely, the antibody designs generated during this project did not exhibit detectable binding in yeast surface display assays. Several factors may have contributed to this outcome. Notably, the antibodies were designed and cloned in bacteria or yeast using

emerging methodologies that had not yet undergone complete experimental validation at the time of testing, potentially compromising the stability or folding of the constructs[136,137]. Additionally, the computational scoring of these designs appeared to be disproportionately influenced by contributions from the constant regions incorporated into the redesigned variable domains, potentially allowing lower-quality designs to progress through the filtering stages.

While both Rifdock[122] and RFdiffusion[36] have been applied successfully in numerous contexts to generate high-affinity minibinders against diverse targets[26,27,110,135], the reliable computational design of de novo antibodies remains a significant and unresolved challenge. This area continues to be the focus of intensive investigation. Importantly, insights gained from the present work—particularly regarding target interface selection and the rigorous screening of computational designs—are expected to inform and refine future approaches in this field.

Acknowledgments

Thank you to Brain Coventry, Nathaniel Bennet and Justas Dupras for leading development of RifDock and MPNN minibinder pipelines. Thank you to Inna Goreshnik, Aza Allen, Cami Cordray, Samer Halabiya, and Dionne Vafeados for help with yeast library preparation. Thank you to Joe Watson for developing the scFv and de novo antibody pipeline. Thank you to Homin Kim for his helpful discussions and partnership in designing antibodies against TLR4.

Materials and Methods

Computational design

The ectodomain of TLR4[25] (PDB ID: 3FXI) was used for the design. The PDB was truncated to save computational time. Minibinders were designed as outlined in ref **(REF).** In brief, RifGen[35], PatchDock[123], RifDock[35], were used. Hundreds of scaffolds were created using this pipeline then fed into ProteinMPNN for sequence design, then predicted and filtered on various AF2 metrics (pae_interaction, binder_plddt).

Library preparation

Designs were padded to 65 amino acids with serine. Then, the designs were codon-optimized for expression in *Saccharomyces cerevisiae*. Oligonucleotides encoding the designs and SSM mutations were purchased from Agilent technologies. Libraries were amplified as previously described in Chapter 2.

Yeast surface display

Yeast surface display was performed as described in Chapter 5[133]. Sort 4 was done without avidity.

Biolayer interferometry

BLI was performed as described in Chapter 5. Initial binding was tested at 700nM followed by 2x serial dilutions of hTLR4-Fc (RnD Biosystems) were preformed starting at 500nM.

Computational design of de novo scFvs

The design of scFVs was performed as outlined in ref[135]**.** In brief scFv backbones were generated with RFDiffusion to map to specific epitope regions. From here CDR loops are generated around epitope domains to increase specificity to target regions. Designs are initially filtered here to remove scFv's where loops are not in close proximity to epitope regions. Then those that pass through are given sequences with ProteinMNN[30]. Complexes are then predicted and scored with AF2 on various metrics (pae_interaction, loop_rmsd, complexity) before ordering[36].

Supplementary Table 1. Amino acid sequences

Construct*	Amino Acid Sequence
4_1	GVVEKVCEYLYKKIKELIKEGKIEEAKKYLKQLELLAEKFNSPYCKK LVEEIKKEL
4_2	DEVWKAVEEAEKLAKEGDLEGCVKLFLKVYKIYKETGDPEAKEAL EKIRKIVTELRE

Chapter 7. Conclusions

In this work, we present the development and application of advanced computational strategies to enhance the functional utility of TLRs, encompassing both the generation of stabilized recombinant ectodomains for improved screening and the design of novel binders in the form of antibodies and de novo minibinders.

A central achievement of this dissertation was the successful stabilization of TLR2 through the design of the 2S5 variant, which allowed for significantly improved protein expression, solubility, and functional activity. The integration of AI-based tools such as ProteinMPNN[30,36] and AF2[33,36] facilitated large-scale sequence generation and structural scoring, dramatically accelerating the design process. However, it became evident that these AI models exhibit certain biases, such as the over-selection of specific residue types or motifs. This limitation necessitated a complementary rational design strategy, incorporating both targeted mutations and holistic sequence diversification to enhance structural and functional diversity[42,57,59,63]. The resulting stabilized protein, 2S5, served as a robust platform for the subsequent development of functional binders, highlighting the translational potential of such engineered receptors.

Building upon the methodologies established with TLR2, similar approaches were applied to the more structurally complex TLR5, which contains a flexible loop critical for flagellin recognition[44]. Despite the challenges associated with this region, we successfully engineered a TLR5 variant with improved expression and biophysical stability. This construct demonstrated partial binding to select flagellin variants, though the observed affinities were low. These findings underscore the need for further structural studies to elucidate the determinants of differential flagellin recognition and to refine binder design accordingly.

Beyond receptor stabilization, this work also focused on the development of de novo binders using both physics-based (e.g., RifDock[122]) and AI-driven (e.g., RFDiffusion, ProteinMPNN, AF2) computational methods[29,36,135]. These approaches enabled the successful generation of minibinders and antibodies targeting conserved interface regions on TLR1 and TLR4. In particular, minibinder designs demonstrated promising binding activity at predicted hotspot regions. While de novo antibody design also yielded encouraging results, further improvements in screening throughput and in silico filtering pipelines are required to fully realize their potential.

Taken together, this dissertation illustrates the power of integrating AI-guided modeling, physics-based design, and rational intervention across multiple TLR targets. As computational protein design continues to scale, the importance of high-throughput, high-fidelity experimental validation will become increasingly critical. Furthermore,

strategic understanding of the target protein's biology, structural landscape, and intended application will be essential to reduce inefficiencies and maximize the functional success of future designs. In this context, the iterative synergy between computational prediction and experimental evaluation emerges as a key paradigm for the next generation of synthetic immunobiology.

References

1. Takeda, K. & Akira, S. Toll-like receptors in innate immunity. *Int. Immunol.* **17**, 1–14 (2005).
2. Akira, S. & Takeda, K. Toll-like receptor signalling. *Nat. Rev. Immunol.* **4**, 499–511 (2004).
3. Kawai, T. & Akira, S. The role of pattern-recognition receptors in innate immunity: update on Toll-like receptors. *Nat. Immunol.* **11**, 373–384 (2010).
4. Bell, J. K. *et al.* The molecular structure of the Toll-like receptor 3 ligand-binding domain. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 10976–10980 (2005).
5. Jin, M. S. & Lee, J.-O. Structures of the toll-like receptor family and its ligand complexes. *Immunity* **29**, 182–191 (2008).
6. Jin, M. S. *et al.* Crystal Structure of the TLR1-TLR2 Heterodimer Induced by Binding of a Tri-Acylated Lipopeptide. *Cell* **130**, 1071–1082 (2007).
7. Underhill, D. M. *et al.* The Toll-like receptor 2 is recruited to macrophage phagosomes and discriminates between pathogens. *Nature* **401**, 811–815 (1999).
8. Kadowaki, N. *et al.* Subsets of human dendritic cell precursors express different toll-like receptors and respond to different microbial antigens. *J. Exp. Med.* **194**, 863–869 (2001).
9. Zhu, M. *et al.* TLR4/Rac1/NLRP3 pathway mediates amyloid-β-induced neuroinflammation in Alzheimer's disease. *J. Alzheimers. Dis.* **99**, 911–925 (2024).
10. Sun, H. *et al.* Targeting toll-like receptor 7/8 for immunotherapy: recent advances and prospectives. *Biomark. Res.* **10**, 89 (2022).
11. Wang, J. *et al.* Small-molecule modulators targeting toll-like receptors for potential anticancer therapeutics. *J. Med. Chem.* **66**, 6437–6462 (2023).
12. David, C. *et al.* Interface gain-of-function mutations in TLR7 cause systemic and neuro-inflammatory disease. *J. Clin. Immunol.* **44**, 60 (2024).

13. Kirtland, M. E., Tsitoura, D. C., Durham, S. R. & Shamji, M. H. Toll-like receptor agonists as adjuvants for Allergen Immunotherapy. *Front. Immunol.* **11**, 599083 (2020).
14. Katelaris, A. L. *et al.* Effectiveness of BCG Vaccination Against Mycobacterium tuberculosis Infection in Adults: A Cross-sectional Analysis of a UK-Based Cohort. *J. Infect. Dis.* **221**, 146–155 (2020).
15. Zheng, H. *et al.* The TLR2 is activated by sporozoites and suppresses intrahepatic rodent malaria parasite development. *Sci. Rep.* **5**, 18239 (2015).
16. Laupèze, B., Hervé, C., Di Pasquale, A. & Tavares Da Silva, F. Adjuvant Systems for vaccines: 13 years of post-licensure experience in diverse populations have progressed the way adjuvanted vaccine safety is investigated and understood. *Vaccine* **37**, 5670–5680 (2019).
17. Didierlaurent, A. M. *et al.* AS04, an Aluminum Salt- and TLR4 Agonist-Based Adjuvant System, Induces a Transient Localized Innate Immune Response Leading to Enhanced Adaptive Immunity. *The Journal of Immunology* vol. 183 6186–6197 Preprint at https://doi.org/10.4049/jimmunol.0901474 (2009).
18. O'Neill, L. A. J., Bryant, C. E. & Doyle, S. L. Therapeutic targeting of Toll-like receptors for infectious and inflammatory diseases and cancer. *Pharmacol. Rev.* **61**, 177–197 (2009).
19. Sagiv-Barfi, I. *et al.* Eradication of spontaneous malignancy by local immunotherapy. *Sci. Transl. Med.* **10**, eaan4488 (2018).
20. Nickerson, K. M. *et al.* TLR9 regulates TLR7- and MyD88-dependent autoantibody production and disease in a murine model of lupus. *J. Immunol.* **184**, 1840–1848 (2010).
21. Lowe, E. L. *et al.* Toll-like receptor 2 signaling protects mice from tumor development in a mouse model of colitis-induced cancer. *PLoS One* **5**, e13027 (2010).
22. Li, T.-T., Ogino, S. & Qian, Z. R. Toll-like receptor signaling in colorectal cancer: carcinogenesis to cancer therapy. *World J. Gastroenterol.* **20**, 17699–17708 (2014).
23. Chung, K. Y. *et al.* Toll-like receptor 3 signaling attenuated colitis-associated cancer

development in mice. *Sci. Rep.* **14**, 30308 (2024).

24. Choe, J., Kelker, M. S. & Wilson, I. A. Crystal structure of human toll-like receptor 3 (TLR3) ectodomain. *Science* **309**, 581–585 (2005).
25. Kim, H. M. *et al.* Crystal structure of the TLR4-MD-2 complex with bound endotoxin antagonist Eritoran. *Cell* **130**, 906–917 (2007).
26. Cao, L. *et al.* De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *bioRxiv* (2020) doi:10.1101/2020.08.03.234914.
27. Silva, D.-A. *et al.* De novo design of potent and selective mimics of IL-2 and IL-15. *Nature* **565**, 186–191 (2019).
28. Wang, Y. *et al.* Human monoclonal anti-TLR4 antibody negatively regulates lipopolysaccharide-induced inflammatory responses in mouse macrophages. *Mol. Med. Rep.* **22**, 4125–4134 (2020).
29. Bennett, N. R. *et al.* Improving de novo protein binder design with deep learning. *Nat. Commun.* **14**, 2625 (2023).
30. Dauparas, J. *et al.* Robust deep learning based protein sequence design using ProteinMPNN. *bioRxiv* (2022) doi:10.1101/2022.06.03.494563.
31. de Haas, R. J. *et al.* Rapid and automated design of two-component protein nanomaterials using ProteinMPNN. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2314646121 (2024).
32. Sumida, K. H. *et al.* Improving protein expression, stability, and function with ProteinMPNN. *J. Am. Chem. Soc.* **146**, 2054–2061 (2024).
33. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
34. Anishchenko, I. *et al.* De novo protein design by deep network hallucination. *Nature* **600**, 547–552 (2021).
35. Dou, J. *et al.* De novo design of a fluorescence-activating β-barrel. *Nature* **561**, 485–491 (2018).

36. Watson, J. L. *et al.* De novo design of protein structure and function with RFdiffusion. *Nature* **620**, 1089–1100 (2023).
37. Roeder, A., Kirschning, C. J. & Schaller, M. Induction of Nuclear Factor-κB and c-Jun/Activator Protein-1 via Toll-Like Receptor 2 in Macrophages by Antimycotic-Treated Candida albicans. *The Journal of* (2004).
38. Cheng, K. *et al.* Specific activation of the TLR1-TLR2 heterodimer by small-molecule agonists. *Sci Adv* **1**, (2015).
39. Kang, J. Y. *et al.* Recognition of Lipopeptide Patterns by Toll-like Receptor 2-Toll-like Receptor 6 Heterodimer. *Immunity* **31**, 873–884 (2009).
40. McFadden, E. *et al.* Engineering and structures of Crimean-Congo hemorrhagic fever virus glycoprotein complexes. *Cell* **188**, 303–315.e13 (2025).
41. Sponholtz, M. R. *et al.* Structure-based design of a soluble human cytomegalovirus glycoprotein B antigen stabilized in a prefusion-like conformation. *Proc. Natl. Acad. Sci. U. S. A.* **121**, e2404250121 (2024).
42. Sanders, R. W. *et al.* A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. *PLoS Pathog.* **9**, e1003618 (2013).
43. Gunn, R. J., Herrin, B. R., Acharya, S., Cooper, M. D. & Wilson, I. A. VLR recognition of TLR5 expands the molecular characterization of protein antigen binding by non-ig-based antibodies. *J. Mol. Biol.* **430**, 1350–1367 (2018).
44. Yoon, S.-I. *et al.* Structural basis of TLR5-flagellin recognition and signaling. *Science* **335**, 859–864 (2012).
45. King, N. P. *et al.* Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510**, 103–108 (2014).
46. Bale, J. B. *et al.* Accurate design of megadalton-scale two-component icosahedral protein complexes. *Science* **353**, 389–394 (2016).

47. Liu, L. *et al.* Structural basis of Toll-like receptor 3 signaling with double-stranded RNA. *Science* **320**, 379–381 (2008).
48. West, A. P., Koblansky, A. A. & Ghosh, S. Recognition and signaling by toll-like receptors. *Annu. Rev. Cell Dev. Biol.* **22**, 409–437 (2006).
49. Stack, J. *et al.* TRAM is required for TLR2 endosomal signaling to type I IFN induction. *J. Immunol.* **193**, 6090–6102 (2014).
50. Yamamoto, M. *et al.* Role of adaptor TRIF in the MyD88-independent toll-like receptor signaling pathway. *Science* **301**, 640–643 (2003).
51. Balka, K. R. & De Nardo, D. Understanding early TLR signaling through the Myddosome. *J. Leukoc. Biol.* **105**, 339–351 (2019).
52. Gay, N. J. Role of self-organising myddosome oligomers in inflammatory signalling by Toll-like receptors. *BMC Biol.* **17**, 15 (2019).
53. Chaudhari, A., Axelsson, C., Mattsson Hultén, L. & Rotter Sopasakis, V. Toll-like receptors 1, 3 and 7 activate distinct genetic features of NF-κB signaling and γ-protocadherin expression in human cardiac fibroblasts. *Inflammation* (2025) doi:10.1007/s10753-025-02238-z.
54. Hsieh, C.-L. *et al.* Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **369**, 1501–1505 (2020).
55. Yang, Z., Zeng, X., Zhao, Y. & Chen, R. AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduct. Target. Ther.* **8**, 115 (2023).
56. Ellis, D., Lederhofer, J., Acton, O. J., Tsybovsky, Y. & Kephart, S. Structure-based design of stabilized recombinant influenza neuraminidase tetramers. *bioRxiv* (2021).
57. Greaney, A. J. *et al.* Complete Mapping of Mutations to the SARS-CoV-2 Spike Receptor-Binding Domain that Escape Antibody Recognition. *Cell Host Microbe* **29**, 44–57.e9 (2021).
58. Ranbhor, R. Advancing monoclonal antibody manufacturing: Process optimization, cost

reduction strategies, and emerging technologies. *Biologics* **19**, 177–187 (2025).

59. Ellis, D. *et al.* Stabilization of the SARS-CoV-2 Spike Receptor-Binding Domain Using Deep Mutational Scanning and Structure-Based Design. *Front. Immunol.* **12**, 710263 (2021).
60. Jiang, S., Li, X., Hess, N. J., Guan, Y. & Tapping, R. I. TLR10 Is a Negative Regulator of Both MyD88-Dependent and -Independent TLR Signaling. *The Journal of Immunology* **196**, 3834–3841 (2016).
61. Su, L. *et al.* Structural Basis of TLR2/TLR1 Activation by the Synthetic Agonist Diprovocim. *J. Med. Chem.* **62**, 2938–2949 (2019).
62. Freedman, R. B. Eukaryotic protein disulfide-isomerases and their potential in the production of disulfide-bonded protein products: What we need to know but do not! in *Oxidative Folding of Peptides and Proteins* 121–157 (The Royal Society of Chemistry, Cambridge, 2008).
63. Guo, W., Cleveland, B., Davenport, T. M., Lee, K. K. & Hu, S.-L. Purification of recombinant vaccinia virus-expressed monomeric HIV-1 gp120 to apparent homogeneity. *Protein Expr. Purif.* **90**, 34–39 (2013).
64. Xu, Z., Davila, A., Wilamowski, J., Teraguchi, S. & Standley, D. M. Improved antibody-specific Epitope prediction using AlphaFold and AbAdapt. *Chembiochem* **23**, e202200303 (2022).
65. Bui, H.-H., Sidney, J., Li, W., Fusseder, N. & Sette, A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. *BMC Bioinformatics* **8**, 361 (2007).
66. Bryant, P. *et al.* Predicting the structure of large protein complexes using AlphaFold and Monte Carlo tree search. *Nat. Commun.* **13**, 6028 (2022).
67. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
68. Steinegger, M. & Söding, J. MMseqs2 enables sensitive protein sequence searching for the

analysis of massive data sets. *Nat. Biotechnol.* **35**, 1026–1028 (2017).

69. Christensen, D. Vaccine adjuvants: Why and how. *Hum. Vaccin. Immunother.* **12**, 2709–2711 (2016).
70. Kobe, B. & Kajava, A. V. The leucine-rich repeat as a protein recognition motif. *Curr. Opin. Struct. Biol.* **11**, 725–732 (2001).
71. Mistry, P. *et al.* Inhibition of TLR2 signaling by small molecule inhibitors targeting a pocket within the TLR2 TIR domain. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 5455–5460 (2015).
72. Marcandalli, J. *et al.* Induction of Potent Neutralizing Antibody Responses by a Designed Protein Nanoparticle Vaccine for Respiratory Syncytial Virus. *Cell* **176**, 1420–1431.e17 (2019).
73. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nat. Protoc.* **1**, 2876–2890 (2006).
74. Jefferis, R. Recombinant antibody therapeutics: the impact of glycosylation on mechanisms of action. *Trends Pharmacol. Sci.* **30**, 356–362 (2009).
75. Borrok, M. J., Jung, S. T., Kang, T. H., Monzingo, A. F. & Georgiou, G. Revisiting the role of glycosylation in the structure of human IgG Fc. *ACS Chem. Biol.* **7**, 1596–1602 (2012).
76. Botos, I., Segal, D. M. & Davies, D. R. The structural biology of Toll-like receptors. *Structure* **19**, 447–459 (2011).
77. Decout, A. *et al.* Lactobacillus crispatus S-layer proteins modulate innate immune response and inflammation in the lower female reproductive tract. *Nat. Commun.* **15**, 10879 (2024).
78. Hari, P. *et al.* The innate immune sensor Toll-like receptor 2 controls the senescence-associated secretory phenotype. *Sci. Adv.* **5**, eaaw0254 (2019).
79. Sharma, R., Ghasparian, A., Robinson, J. A. & McCullough, K. C. Dendritic cell sensing of hydrophobic Di- and triacylated lipopeptides self-assembled within synthetic virus-like particles. *J. Immunol.* **199**, 734–749 (2017).
80. Brouwer, P. J. M. *et al.* Two-component spike nanoparticle vaccine protects macaques from

SARS-CoV-2 infection. *Cell* **184**, 1188–1200.e19 (2021).

81. Boyoglu-Barnum, S. *et al.* Quadrivalent influenza nanoparticle vaccines induce broad protection. *Nature* **592**, 623–628 (2021).
82. Tiller, T., Busse, C. E. & Wardemann, H. Cloning and expression of murine Ig genes from single B cells. *J. Immunol. Methods* **350**, 183–193 (2009).
83. Cirelli, K. M. *et al.* Slow delivery immunization enhances HIV neutralizing antibody and germinal center responses via modulation of immunodominance. *Cell* **177**, 1153–1171.e28 (2019).
84. Lashgari, D. *et al.* From affinity selection to kinetic selection in Germinal Centre modelling. *PLoS Comput. Biol.* **18**, e1010168 (2022).
85. McLellan, J. S. *et al.* Structure-based design of a fusion glycoprotein vaccine for respiratory syncytial virus. *Science* **342**, 592–598 (2013).
86. Lee, J. *et al.* A broadly generalizable stabilization strategy for sarbecovirus fusion machinery vaccines. *Nat. Commun.* **15**, 5496 (2024).
87. Haynes, B. F. *et al.* Immune-correlates analysis of an HIV-1 vaccine efficacy trial. *N. Engl. J. Med.* **366**, 1275–1286 (2012).
88. Chao, C. W. *et al.* Protein nanoparticle vaccines induce potent neutralizing antibody responses against MERS-CoV. *Cell Rep.* **43**, 115036 (2024).
89. Johnson, N. V. *et al.* Discovery and characterization of a pan-betacoronavirus S2-binding antibody. *Structure* **32**, 1893–1909.e11 (2024).
90. Villalobos-Alva, J. *et al.* Protein science meets artificial intelligence: A systematic review and a biochemical meta-analysis of an inter-field. *Front. Bioeng. Biotechnol.* **10**, 788300 (2022).
91. Akira, S., Uematsu, S. & Takeuchi, O. Pathogen recognition and innate immunity. *Cell* **124**, 783–801 (2006).
92. Sundling, C., Phad, G., Douagi, I., Navis, M. & Karlsson Hedestam, G. B. Isolation of

antibody V(D)J sequences from single cell sorted rhesus macaque B cells. *J. Immunol. Methods* **386**, 85–93 (2012).

93. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, (2017).
94. Häcker, H. *et al.* Specificity in Toll-like receptor signalling through distinct effector functions of TRAF3 and TRAF6. *Nature* **439**, 204–207 (2006).
95. Asea, A. *et al.* Novel signal transduction pathway utilized by extracellular HSP70: role of toll-like receptor (TLR) 2 and TLR4. *J. Biol. Chem.* **277**, 15028–15034 (2002).
96. Seibl, R. *et al.* Expression and regulation of Toll-like receptor 2 in rheumatoid arthritis synovium. *Am. J. Pathol.* **162**, 1221–1227 (2003).
97. Cario, E., Gerken, G. & Podolsky, D. K. Toll-like receptor 2 controls mucosal inflammation by regulating epithelial barrier function. *Gastroenterology* **132**, 1359–1374 (2007).
98. Elloumi, N. *et al.* Role of innate immune receptors TLR4 and TLR2 polymorphisms in systemic lupus erythematosus susceptibility. *Ann. Hum. Genet.* **86**, 137–144 (2022).
99. Oliveira, I. B. N. *et al.* Interferon-beta treatment differentially alters TLR2 and TLR4-dependent cytokine production in multiple sclerosis patients. *Neuroimmunomodulation* **26**, 77–83 (2019).
100. Wasko, N. J., Nichols, F. & Clark, R. B. Multiple sclerosis, the microbiome, TLR2, and the hygiene hypothesis. *Autoimmun. Rev.* **19**, 102430 (2020).
101. Miranda-Hernandez, S. & Baxter, A. G. Role of toll-like receptors in multiple sclerosis. *Am. J. Clin. Exp. Immunol.* **2**, 75–93 (2013).
102. Squillace, S. & Salvemini, D. Toll-like receptor-mediated neuroinflammation: relevance for cognitive dysfunctions. *Trends Pharmacol. Sci.* **43**, 726–739 (2022).
103. Burrows, M. P., Volchkov, P., Kobayashi, K. S. & Chervonsky, A. V. Microbiota regulates type 1 diabetes through Toll-like receptors. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 9973–9977 (2015).

104. Liu, S. *et al.* TLR2 is a primary receptor for Alzheimer's amyloid β peptide to trigger neuroinflammatory activation. *J. Immunol.* **188**, 1098–1107 (2012).

105. Li, Y. *et al.* Targeting microglial α-synuclein/TLRs/NF-kappaB/NLRP3 inflammasome axis in Parkinson's disease. *Front. Immunol.* **12**, 719807 (2021).

106. Kim, C. *et al.* Neuron-released oligomeric α-synuclein is an endogenous agonist of TLR2 for paracrine activation of microglia. *Nat. Commun.* **4**, 1562 (2013).

107. van der Poll, T., van de Veerdonk, F. L., Scicluna, B. P. & Netea, M. G. The immunopathology of sepsis and potential therapeutic targets. *Nat. Rev. Immunol.* **17**, 407–420 (2017).

108. Meng, G. *et al.* Antagonistic antibody prevents toll-like receptor 2-driven lethal shock-like syndromes. *J. Clin. Invest.* **113**, 1473–1481 (2004).

109. Chevalier, A. *et al.* Massively parallel de novo protein design for targeted therapeutics. *Nature* **550**, 74–79 (2017).

110. Adams, C. S. *et al.* De novo design of protein minibinder agonists of TLR3. *Nat. Commun.* **16**, 1234 (2025).

111. Fleishman, S. J. *et al.* Computational design of proteins targeting the conserved stem region of influenza hemagglutinin. *Science* **332**, 816–821 (2011).

112. Studier, F. W. Stable expression clones and auto-induction for protein production in E. coli. *Methods Mol. Biol.* **1091**, 17–32 (2014).

113. Afzal, H., Murtaza, A. & Cheng, L.-T. Structural engineering of flagellin as vaccine adjuvant: quest for the minimal domain of flagellin for TLR5 activation. *Mol. Biol. Rep.* **52**, 104 (2025).

114. O'Mahony, D. S., Pham, U., Iyer, R., Hawn, T. R. & Liles, W. C. Differential constitutive and cytokine-modulated expression of human Toll-like receptors in primary neutrophils, monocytes, and macrophages. *Int. J. Med. Sci.* **5**, 1–8 (2008).

115. Song, W. S., Jeon, Y. J., Namgung, B., Hong, M. & Yoon, S.-I. A conserved TLR5 binding and activation hot spot on flagellin. *Sci. Rep.* **7**, 40878 (2017).

116. Hajam, I. A., Dar, P. A., Shahnawaz, I., Jaume, J. C. & Lee, J. H. Bacterial flagellin-a potent immunomodulatory agent. *Exp. Mol. Med.* **49**, e373 (2017).
117. Suzek, B. E., Huang, H., McGarvey, P., Mazumder, R. & Wu, C. H. UniRef: comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* **23**, 1282–1288 (2007).
118. Takeuchi, O. *et al.* Cutting edge: role of Toll-like receptor 1 in mediating immune response to microbial lipoproteins. *J. Immunol.* **169**, 10–14 (2002).
119. Carpenter, S. & O'Neill, L. A. J. How important are Toll-like receptors for antimicrobial responses? *Cell. Microbiol.* **9**, 1891–1901 (2007).
120. Yan, B. *et al.* A review: The significance of toll-like receptors 2 and 4, and NF-κB signaling in endothelial cells during atherosclerosis. *Front. Biosci. (Landmark Ed.)* **29**, 161 (2024).
121. Sharp, F. A. *et al.* Uptake of particulate vaccine adjuvants by dendritic cells activates the NALP3 inflammasome. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 870–875 (2009).
122. Cao, L. *et al.* Design of protein-binding proteins from the target structure alone. *Nature* **605**, 551–560 (2022).
123. Schneidman-Duhovny, D., Inbar, Y., Nussinov, R. & Wolfson, H. J. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res.* **33**, W363 (2005).
124. Hoshino, K. *et al.* Cutting edge: Toll-like receptor 4 (TLR4)-deficient mice are hyporesponsive to lipopolysaccharide: evidence for TLR4 as the Lps gene product. *J. Immunol.* **162**, 3749–3752 (1999).
125. Zanoni, I. *et al.* CD14 controls the LPS-induced endocytosis of Toll-like receptor 4. *Cell* **147**, 868–880 (2011).
126. Mitra, S. & Tomar, P. C. Hybridoma technology; advancements, clinical significance, and future aspects. *J. Genet. Eng. Biotechnol.* **19**, 159 (2021).
127. O'Neill, L. A. J. & Bowie, A. G. The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nat. Rev. Immunol.* **7**, 353–364 (2007).
128. De Mot, L. *et al.* Transcriptional profiles of adjuvanted hepatitis B vaccines display variable

interindividual homogeneity but a shared core signature. *Sci. Transl. Med.* **12**, (2020).

129. Tait, D. R. *et al.* Final Analysis of a Trial of M72/AS01E Vaccine to Prevent Tuberculosis. *N. Engl. J. Med.* **381**, 2429–2439 (2019).

130. Akira, S., Takeda, K. & Kaisho, T. Toll-like receptors: critical proteins linking innate and acquired immunity. *Nat. Immunol.* **2**, 675–680 (2001).

131. *Safety, Tolerability, and Immunogenicity of the Vaccine Candidates ID93 + AP10-602 and ID93 + GLA-SE Administered Intramuscularly in Healthy Adult Subjects*.

132. *Temperature-Stable TB Vaccine Safe, Prompts Immune Response in NIH-Supported Study*.

133. Liu, B. *Yeast Surface Display: Methods, Protocols, and Applications*. (Springer New York, 2016).

134. He, X.-H. *et al.* AI-driven antibody design with generative diffusion models: current insights and future directions. *Acta Pharmacol. Sin.* **46**, 565–574 (2025).

135. Bennett, N. R. *et al.* Atomically accurate de novo design of antibodies with RFdiffusion. *bioRxivorg* 2024.03.14.585103 (2025) doi:10.1101/2024.03.14.585103.

136. Lee, Y. J. & Jeong, K. J. Challenges to production of antibodies in bacteria and yeast. *J. Biosci. Bioeng.* **120**, 483–490 (2015).

137. Sandomenico, A., Sivaccumar, J. P. & Ruvo, M. Evolution of Escherichia coli expression system in producing antibody recombinant fragments. *Int. J. Mol. Sci.* **21**, 6324 (2020).