Targeting the initiating proteases of the complement system using structure-based small molecule drug discovery

By

Denise Rohlik

July 2023

Director of Thesis: Brandon Garcia, PhD

Major Department: Microbiology and Immunology

The complement system is an essential element of host development and immune homeostasis. This immunoregulatory enzyme cascade is driven by serine protease activation and plays critical roles in the recognition of danger signals associated with pathogens and damaged or apoptotic self-molecules. Due to its potent effector functions and ability to drive overwhelming innate and adaptive immune responses against a particular target, tight regulation is vital to prevent damage to host tissues. In the event of complement dysregulation, humans are known to become susceptible to an ever-growing number of autoimmune, inflammatory, and neurodegenerative disorders. A detailed understanding of the molecular mechanisms that drive protein function, and specifically complement activation, is fundamental to the design and development of specific complement-directed therapeutics.

The complement system is comprised of three distinct pathways that differ in the molecular targets that drive pathway activation. The initiating proteases of the classical and lectin pathways are set apart from other serine proteases in that they undergo autocatalytic activation in the

presence of a molecular trigger without necessitating extrinsic factors for full enzymatic activation. Using structural data derived from the enzyme-product complex of self-associated lectin pathway MASP-2 proteases, we elucidate the conformational changes that occur during autoactivation in the initiating proteases of the classical and lectin pathway. Our structural model provides an empirically derived alternative mechanism for autocatalysis that has not been previously described in the complement proteases.

Using structural information derived from X-ray crystallography and autocatalysis as well as known interactions between bacterial inhibitors utilized for host complement evasion, we pursued multiple avenues for rational small-molecule drug design aimed at C1r and C1s, which are initiating proteases of the classical pathway. Molecular dynamics studies and chemical modification of compounds were used to guide the optimization and development of competitive small-molecule inhibitors against the active site of C1r. Structure-activity relationship data was generated through the bioisosteric replacement of chemical moieties of another lead compound to drive noncompetitive C1r inhibition. Moreover, advancements in technology building off protein-ligand interaction fingerprints derived from known structures were utilized in collaborative machine learning and artificial intelligence methods for the rational design of novel C1s inhibitors. Additional optimization of small-molecule compounds may ultimately lead to novel therapeutic options for diseases mediated by aberrant classical pathway activation.

To aid in future endeavors in small-molecule drug screening and design, we developed a novel platform and methodology for evaluating targets outside of the complement system that requires the presence of lipid bilayers for activity. Using surface plasmon resonance, we established a bilayer model to mimic biological membranes to aid in protein characterization and screen for inhibitors capable of high affinity binding to the target surface as well as those that block

protein association to the membrane, thereby inhibiting protein function. As such, biofunctional assays, such as surface plasmon resonance, can serve as powerful tools for aiding structure-based drug design strategies. Together, the structural insights gleaned from the enzyme-product complex of MASP-2, as well as surface plasmon resonance binding assays, allowed us to identify three small-molecule compound hits against both C1r and C1s using different methodologies and means of inhibition, validating differing strategies of inhibitor design. The identification of these hits, in addition to the development of a novel SPR screening platform, has led to the generation of in-depth structural data and method systems to guide future continued optimization efforts that may one day lead to the design of a small-molecule therapeutic option for diseases marked by autoimmune and inflammatory pathologies.

Targeting the initiating proteases of the complement system using structure-based small molecule drug discovery

A Dissertation

Presented to the Faculty of the Microbiology and Immunology Program

Brody School of Medicine at East Carolina University

In Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy in Microbiology and Immunology

By

Denise L. Rohlik

July 2023

Director of Dissertation: Brandon L. Garcia, PhD

Thesis Committee Members:

Everett C. Pesci, PhD

Mark D. Mannie, PhD

Michelle L. Ratliff, PhD

Qun Lu, PhD

Acknowledgments

I would like to express my deepest gratitude to my advisor, Dr. Garcia, for his invaluable guidance and limitless patience throughout my development as a researcher in his lab. Additionally, I would like to thank my department chair, Dr. Pesci, without whom I would have never reached the twin peaks of both my master's and doctorate degrees. I am also grateful to Dr. Polakowski who never failed to remind me that when opportunity strikes, grab it by the ankles and don't let go. Thank you also to Dr. Martin who helped me keep my target in sight throughout this journey.

I must also thank my wonderful lab mates for their support both in and out of the lab, especially Dr. CJ Booth who entered this marathon with me, and reminded me to stop and smell the roses, then make detailed observations about local pollinators, and note their microbiome for the win. Additional thanks to my lab mate Ryan Garrigues for also reminding me to savor all the moments, Dr. Sourav Roy for imparting upon me the importance of organization and labelling in the lab, and Kimson Hoang for making the entire experience an adventure.

I am extremely grateful for the support and guidance of my defense committee who have helped broaden my scientific knowledge and develop my project.

Lastly, thank you to my friends and family, especially my mom, my dad, and my bother Marc, who have provided support, encouragement, and motivation throughout this journey.

TABLE OF CONTENTS

LIST OF TABLES ..... **vi**

LIST OF FIGURES ..... **vii**

CHAPTER 1: INTRODUCTION **1**

1.1 Complement system background *1*

1.2 Complement-mediated disorders *13*

1.3 Autoactivation of complement proteases *17*

1.4 Complement therapeutics *20*

1.5 Small-molecular inhibitors: Competitive vs noncompetitive inhibition *24*

1.6 Fragment-based drug discovery *34*

1.7 Artificial intelligence-based drug design and AlphaFold *38*

CHAPTER 2: AUTOACTIVATION OF COMPLEMENT PROTEASES: MASP-2 ..... **40**

2.1 Autoactivating enzymes *41*

2.2 Enzyme-product complex *47*

2.3 S1-P1 interactions *49*

2.4 Catalytic mechanisms *51*

2.5 Alternate catalytic serine conformations *52*

2.6 Buried surface area *53*

CHAPTER 3: SMALL-MOLECULE INHIBITOR DEVELOPMENT: CMP-1696 ..... **57**

3.1 C1r-CCP2-SP structure-guided approach *59*

3.2 CMP-1696 interactions with active site *72*

3.3 Analog synthesis from molecular docking studies *82*

3.4 Further optimization *93*

CHAPTER 4: SMALL-MOLECULE INHIBITOR DEVELOPMENT: CMP-1611 ..... **100**

4.1 Compound-structure guided approach *104*

4.2 SAR elucidation *107*

4.3 Analog synthesis for noncompetitive inhibition *110*

4.4 Future directions *112*

CHAPTER 5: SMALL-MOLECULE INHIBITOR DEVELOPMENT: A1 115

5.1 Protein-structure guided approach using AI *116*

5.2 Atomwise for in silico compound screening and synthesis *120*

5.3 A1 characterization *125*

5.4 Analog synthesis and characterization *134*

CHAPTER 6: HUMAN LIPOXYGENASE STUDY WITH 2D LP SENSOR CHIP 141

6.1 Liposome capture *147*

6.2 Characterization of 8R-LOX *150*

6.3 Identification of binding characteristics in other model LOX species *155*

CHAPTER 7: DISCUSSION AND SUMMARY OF THESIS 162

REFERENCES 165

APPENDIX 197

X-ray crystallography *197*

LIST OF TABLES

1. MASP-2-CCP2-SP data collection and refinement statistics……...…… 46
2. Contact residues present at the enzyme-product interface of MASP-2……55

LIST OF FIGURES

1. The complement pathways 1
2. Structure of C1 and MBL in complex with their initiating serine proteases, C1r/C1s and MASP-2, respectively. 4
3. The complement system is involved in a wide range of functions and roles involved in immune homeostasis. 11
4. Involvement of classical complement activation in various disease settings. 14
5. Autocatalytic activation of zymogens to enzymes 19
6. *B. burgdorferi* lipoproteins, BBK32 and ElpQ, provide a basis for complement protease inhibition 33
7. Diagram illustrating the process involved in fragment-based drug discovery 35
8. MASP-2-CCP2-SP crystals in 0.2 M $AmmSO_4$, PEG 4,000, and 20% glycerol 45
9. Structure of crystallized MASP-2-CCP2-SP enzyme 47
10. Enzyme-product complex formed between MASP-2 and symmetry mate. 49
11. Two-dimensional interaction map of activation peptide with -nitrogen and -carboxylate residues of substrate binding pocket 50
12. Residues that form the catalytic triad found in serine proteases 51
13. Electron-density representation of the MASP-2-CCP2-SP catalytic triad 53
14. MASP-2-CCP2-SP enzyme-product structure aligned and superimposed on classical pathway initiating serine protease, C1r-CCP2-SP (PDB:1MD8), enzyme-product complex. 54
15. Schematic diagram showing hit-to-lead generation to identify CMP-1696 58
16. C1r-CCP2-SP crystals grown in condition from the Hampton crystal screen 68
17. Selectivity and mechanistic analysis of CMP-1611 and CMP-1696 76

18. CMP-1696 structure-activity relationship..... 81
19. Activity of analogs synthesized from parent compound, CMP-1696..... 85
20. Additional R-group substitutions drive activity changes in analog behavior..... 88
21. R-group substitutions alter proposed H-bond network and further elucidate the structure-activity relationship..... 90
22. Ring replacement of the oxadiazole group in CMP-1696 results in improvements in biological activity..... 93
23. Schematic diagram showing hit-to-lead generation to identify CMP-1611 and generate optimized analogs 102
24. Compound structure of CMP-1611..... 104
25. Dose-dependent binding and inhibition of C1r..... 105
26. Analogs derived from parent CMP-1611 with retained core pyrimidine scaffold retain inhibitory characteristics against classical pathway activation..... 106
27. Design strategy for CMP-1611 analogs was performed using the Craig plot and ring replacement..... 108
28. Structurally similar compounds to CMP-1611 analogs designed by ring replacement 110
29. Evaluation of second-generation analogs for inhibition against the classical pathway of complement..... 111
30. The structure of C1 complex with C1s 118
31. Dose-dependent inhibition of complement in classical pathway-specific assays by A1 128
32. Compound A1 displays specific dose-dependent inhibition of C1s, and to a lesser extent, MASP-2 130

33. Compound A1 shows differential activity against other pathways of the complement system and activity against murine-derived complement 132
34. Initial co-crystallization efforts towards soaking inhibitor with C1s-CCP2-SP 133
35. Evaluation of A1-like newly synthesized compounds in classical pathway-specific assay systems 136
36. Characterization of two A1-like analogs 137
37. Liposome capture on a 2D LP SPR sensor chip 148
38. 8*R*-LOX membrane binding 151
39. 8*R*-LOX WT dose-response over PC immobilized biosensors 153
40. Binding activity of 8R-LOX to different lipid membranes 155
41. Human 15-LOX-2 and soybean SLO binding activity over PC:PS surface 158
42. Schematic diagram of X-ray crystallography 199

CHAPTER 1: INTRODUCTION

1.1 Complement system background

The complement system is a complex series of plasma-circulating precursor proteins involved in innate immune surveillance that, when active, play a critical role in early development, defense against pathogens, and host homeostasis. To perform these functions, complement activation relies on the recognition of danger-associated signals or triggers, which include but are not limited to pathogens, misfolded proteins, damaged cells, and immune complexes. Activation can occur through three distinct pathways: the classical pathway, the lectin pathway, and the alternative pathway, which are distinguished by pathway-specific molecular initiation events [1]. Despite differences in initiation mechanism, each pathway is comprised of a series of amplifying, sequential enzymatic reactions that possess common effector functions that promote inflammation and converge at the level of C3, ultimately resulting in the formation of the lytic membrane attack complex (MAC) (**Fig. 1**) [2].

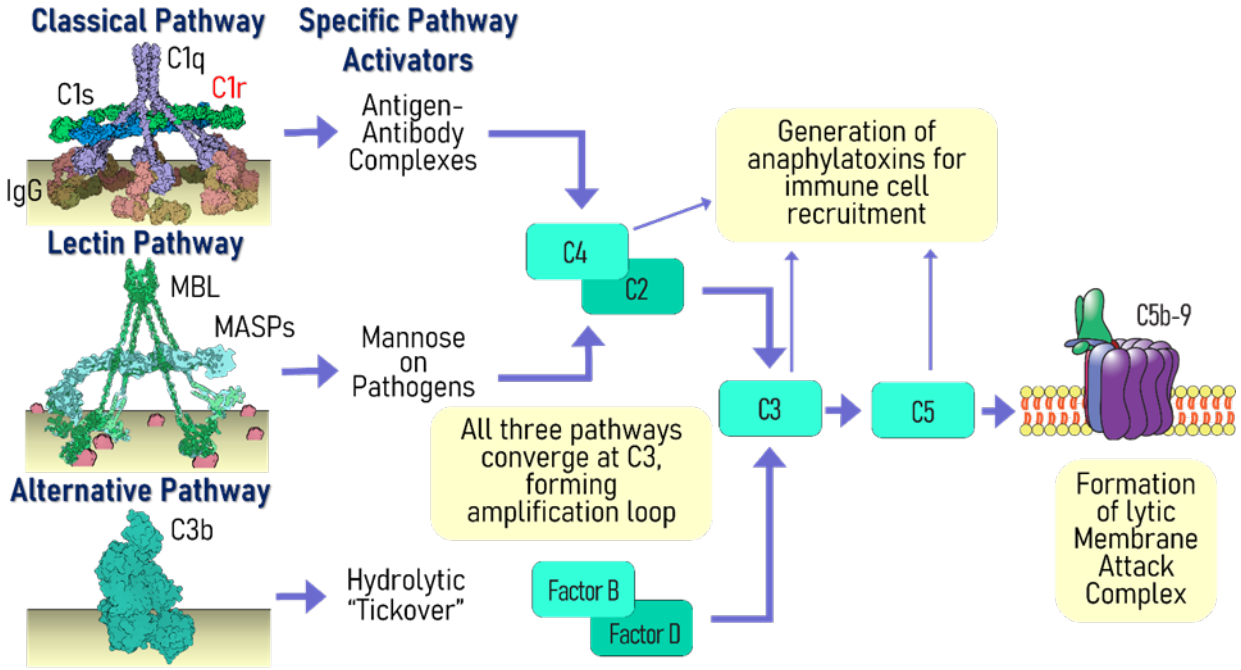


FIGURE 1. The complement pathways. Complement is activated by three canonical pathways: the classical, lectin, or alternative pathways. Activation of the classical complement pathway begins when the pattern recognition protein C1q binds to target surfaces resulting in the autoactivation of the zymogen C1r proteases to proteolytically cleave and activate C1s within the C1 complex. Similarly, the lectin pathway is activated by lectin pathway-specific pattern recognition proteins in complex with mannose-binding associated serine proteases (MASPs). In contrast, the alternative pathway is constitutively activated at low levels by a spontaneous hydrolytic event known as 'tick-over.' Both the classical and lectin pathways converge at the cleavage of C2 and C4 to generate the classical/lectin pathway C3 convertases, C4b2b. Alternative pathway activation results in the formation of C3 convertases in the form of C3bBb. C3 convertases cleave the central molecule of the cascade, C3, into C3a and C3b, resulting in an amplification loop that produces increasing quantities of surface-bound C3b. At high surface concentrations of C3b, C3 convertases bind an additional C3b molecule, resulting in a switch of substrate specificity to C5. Cleavage of C5 by these C5 convertases (i.e., C4b2bC3b and C3bBbC3b) results in the release of the anaphylatoxin C5a and the formation of the pore-like lytic structure called the membrane attack complex (MAC) (i.e., C5b-C9).

For many years following its discovery, the complement system had been solely attributed to the antimicrobial effector roles it played to "complement" antibody-antigen complexes. Recognized as a critical arm of innate immunity, complement acts as both a sentinel and first line of defense against pathogens. While possessing common effector functions across all three pathways, the classical and lectin pathways share more common features starting with the

recognition of danger-associated signals from pathways initiating pattern recognition molecules (PRMs).

The PRM of the classical pathway, C1q, is a highly complex multimeric protein composed of six globular heads, each formed by interacting collagen-like stems with the globular regions. Hundreds of molecular determinants ranging from pathogen-derived macromolecules, eat-me signals stemming from cellular damage and debris, altered self-structures such as misfolded or aggregated proteins, and immune complex formation, among others have been identified as ligands by the C1q globular heads or antigen recognition domains [3]. In canonical classical pathway activation, C1q functions in complex with heterotetrameric serine proteases, C1r and C1s, which rely on C1q for activity (**Fig. 2A**) [4].

Initiation of the classical pathway begins with the recognition and binding event of C1q to a target surface resulting in a binding-induced conformational change that allows for C1r autoactivation. Activated C1r then cleaves its substrate, proenzyme C1s, which results in C1s activation. This activation of the C1 complex allows for the cleavage of complement components C4 and C2, wherein C4a and C2a fragments are released, and C4b and C2b associate to form C3 convertase (C4bC2b) for the cleavage of C3 on the pathogenic surface, resulting in the generation and release of potent anaphylatoxin, C3a, and covalent binding of C3b to hydroxyl groups present on the target surface through the exposure of an internal thioester bond [5].

The lectin pathway follows similar methodologies of C3 convertase generation as that of the classical pathway. Lectin pathway activation is initiated by the binding event of associated pattern recognition molecules, resulting in subsequent serine protease activation. Mannan-binding lectin (MBL), ficolins (M-ficolin, L-ficolin, and H-ficolin), and collectins (collectin-K1 and collectin-L1) recognize and bind carbohydrate ligands containing terminal mannose or N-

acetylglucosamine residues through their carbohydrate recognition domains [6,7]. Lectin pathway PRMs differ from C1q in that they can form a complex with both homodimers and heterodimers of their pathway-specific initiating proteases, MBL-associated serine protease-1 (MASP-1) and MASP-2 (**Fig. 2C**) [8]. Upon binding to a pathogenic surface, a conformational change is induced in the PRM, allowing for autoactivation of the MASPs and then MASP-2 cleavage of downstream complement components C4 and C2 for the generation of the C3 convertase complex (C4bC2b) [9].

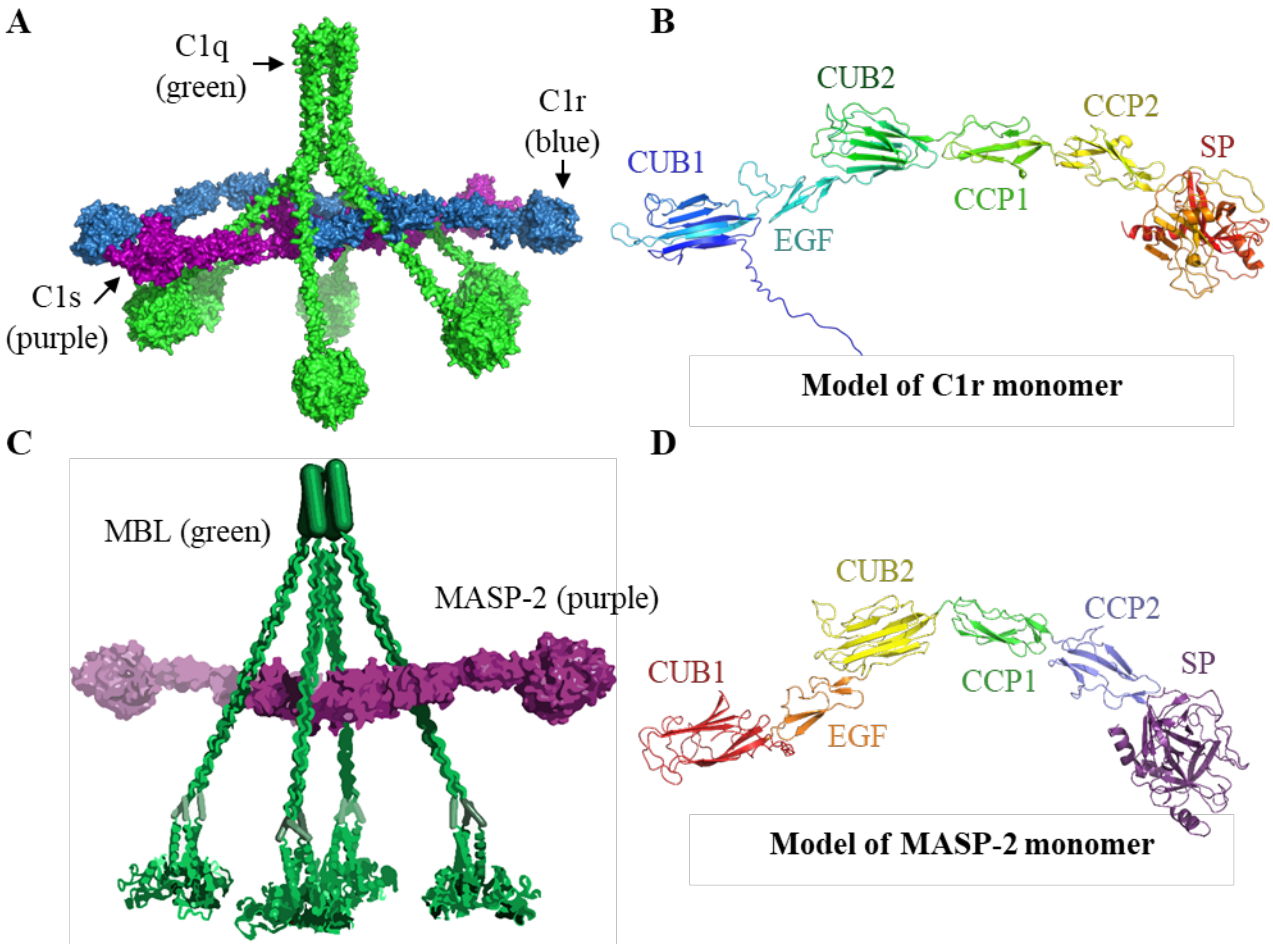


FIGURE 2. Structure of C1 and MBL in complex with their initiating serine proteases, C1r/C1s and MASP-2, respectively. (A) The C1 complex is made up of the recognition molecule, C1q, with six globular heads for multivalent attachment and the proteolytic homodimers,

$C1r_2C1s_2$. Upon C1 fixation to a target site, C1r autoactivates and cleaves pro-C1s, yielding enzymatically active C1s with its serine protease domains extending beyond the C1q stalks to provide a binding site for its natural substrates C4 and C2. **(B)** The C1r monomer comprises six domains, starting with the CUB1, EGF, and CUB2 at the N- terminus and the proteolytically-active CCP2 and the serine protease domains terminating at the C-terminus. The CCP2-SP domains can be recombinantly expressed to form stable and functionally active truncations.

The initiating pathway proteases C1r, C1s, and the MASP-1, and MASP-2 proteins of the classical and lectin pathways, respectively, are structurally and functionally homologous serine proteases sharing many structural features. Each protein is comprised of six domains, including the N-terminal noncatalytic domains responsible for PRM complex interactions facilitating pathway activation: CUB1 (complement C1r/C1s, Uegf, and bone morphogenetic protein-1) domain, epidermal growth factor-like (EGF) domain, CUB2 domain, and the C-terminal domain consisting of two complement control protein (CCP1 and CCP2) domains, and the catalytic serine protease (SP) domain (**Fig. 2 B, D**). The essential catalytic function of the SP domain originates from the presence of the catalytic triad composed of serine, histidine, and aspartate residues.

The alternative pathway is constitutively active and performs spontaneous hydrolysis of C3 into anaphylatoxin C3a and covalently-linked opsonin C3b(H_20) on surfaces lacking complement regulatory molecules promoting phagocytosis by immune cells [10]. Bound C3b fragments associate with serum-phase Factor B, which undergoes cleavage by Factor D into fragments Ba and Bb [11]. This event results in the formation of the alternative pathway C3 convertase, C3bBb. All three pathways converge at the level of C3, wherein the alternative

pathway serves as an amplification loop, enhancing classical and lectin pathway activation and subsequent recruitment of immune cells through chemotactic means and opsonization.

At sufficiently high densities within the classical and lectin pathways, C3b associates with the existing C4bC2b complex to form the C5 convertase (C4bC2bC3b) [12]. Similarly, C5 convertase is formed through the alternative pathway when C3b molecules covalently bind the C3bBb complex, C3bBbC3b. Regardless of pathway origin, the C5 convertase cleaves complement component C5 into potent chemoattractant C5a and into C5b, initiating the assembly of the multimeric membrane attack complex (C5b-C9 or MAC), leading to osmotic lysis of the target membrane through perforations of the cell membrane by polymerized C8 and C9 molecules [1]. Moreover, the generation of anaphylatoxins C3a and C5a molecules recruit and activate immune cells, such as mast cells, basophils, eosinophils, and neutrophils, to the site of infection or tissue injury, resulting in local inflammation [5]. This recruitment event requires recognition of highly conserved pentapeptide sequences, LGLAR for C3a and MQLGR for C5a, present on the C-terminus of C3a and C5a by their cognate receptors [13]. Anaphylatoxins are also responsible for increased vasodilation, vascular permeability, induce contraction of smooth muscles, and expression of proinflammatory cytokines [13–15].

Beyond its established effector roles in innate immunity, the complement system has recently been recognized as an essential player early development, especially during embryogenesis, organogenesis, and tissue remodeling. During the early phase of pregnancy, extensive tissue remodeling of the endometrium of the uterus into decidua occurs for embryo attachment and fetal growth to take place [16]. In early development, complement proteins contribute to placental immune defense against pathogens and promote trophoblast migration to the decidua, supplying the embryo with nourishment [16]. Additionally, studies have shown that

complement is heavily involved in proper development by maintaining tissue integrity and homeostasis through the clearance of cellular debris, which occurs during extensive cell proliferation, differentiation and remodeling, hallmarks of the early stages of development [12].

During neural development, an excessive number of neuronal connections are made, which must undergo synaptic pruning to ensure efficient signaling and prevent enhanced excitatory synaptic connectivity, which could result in disorders relating to epilepsy [17]. Classical pathway components C1q and C3 have been shown to play critical roles in tagging unwanted synapses for elimination by microglia expressing complement receptor 3 (CR3) [18–20]. Complement expression in early CNS tissue, especially that of C1q, C3, C5, their respective complement receptors (C1qR, C3aR, and C5aR), and MASPs have been shown to be critical for neural crest cell migration, and knockdowns or mutations in these genes results in impairment of neural crest migration and craniofacial defects [21–23].

Additional studies have shown that anaphylatoxins C3a and C5a promote cell migration and adhesion, especially during angiogenesis, through induced expression of growth factors such as vascular endothelial growth factor [24,25]. Moreover, it has been suggested that complement plays a direct role in cell adhesions through its effects on cell-matrix interactions [26]. There, too, is mounting evidence that C3 and C5 may have influential roles in the differentiation of skeletal muscle at sites of tissue development and regeneration, demonstrated in studies where there was increased expression of these proteins at sites involved in tissue regeneration [27,28]. This is further supported by a number of studies showing that mice deficient in C3 and C5 were lacking in their ability to regenerate damaged livers and administration of exogenous C5 proteins allowed for liver repair [29–32]. Overwhelmingly, complement activation during early development is

critical for embryonic implantation, defense against infection, promotion of cell migration and adhesion, synaptic plasticity, tissue homeostasis and cell differentiation, and angiogenesis.

As previously implicated, the complement system also interacts and crosstalks with several other physiologically relevant pathways, resulting in a pronounced influence over other biological systems. Best described is the intricate network of interactions and cooperativity between the complement and coagulation systems to coordinate rapid immune responses, inflammation, and hemostasis to sites of injury and prevent infection via vascular injury through those sites. Thrombin, involved in blood coagulation, most notably through the conversion of fibrinogen to fibrin, is a serine protease structurally similar to those found to activate the classical and lectin pathways [33]. In addition to its interactions with blood coagulation factors, thrombin can directly activate complement components through the cleavage of C3 and C5 to induce proinflammatory effects, and it is not alone in its ability to do so. Human coagulation factors XI, Xa, IXa, and plasmin have also been identified as serine proteases capable of C3 and C5 cleavage by acting as natural convertases [34]. Following cleavage, C3a and C5a, as well as MAC, can induce procoagulant responses through the induction of increased expression of tissue factors and platelet activation, resulting in aggregation and the exposure of prothrombinase assembly sites amplifying the inflammatory response [33,35]. This connection is best highlighted in the treatment of paroxysmal nocturnal hemoglobinuria (PNH) and atypical hemolytic uremic syndrome (aHUS). Both diseases are marked by complement dysregulation and thrombosis, where inhibition of C3 by compstatin or C5 with α-C5 monoclonal eculizumab, results in a decrease in hemolysis, thrombosis, and tissue factor expression in patients with either disease [33,36]. Interestingly, *in vitro* studies have demonstrated free-C1q inhibition of coagulation factor XII, suggesting a noncanonical role for C1q in the inhibition of coagulation [37].

As previously discussed, complement activation and the generation of anaphylatoxins can have profound effects on inflammasome activation through the recruitment of immune cells, such as neutrophils, monocytes, macrophages, and dendritic cells to the site of injury or infection (**Fig. 3**). The inflammasome is made up of innate immune system receptors and sensors, such as complement, whose role is to mount immune responses against danger-associated stimuli [38]. Anaphylatoxin-mediated activation of leukocytes can also cause increased inflammation through the triggering of cellular responses, such as degranulation, cytokine production, upregulation of adhesion molecules, and enhancement of phagocytic activity via complement receptor activation [13,39,40]. Target surfaces opsonized by C1q, C3b, and iC3b promote increased phagocytosis through receptor interactions present on neutrophils and macrophages, which also results in the increased production of proinflammatory cytokines and recruitment of additional immune cell support, whereas C5a receptor activation results in immune cell activation, polarization, and chemotaxis [40]. One feature common to immune cell receptor expression is the co-expression of toll-like receptors (TLRs) and complement receptors, which can also drive magnified inflammatory effects through enhanced pathway signaling and immune cell activation [41]. For example, costimulation of complement C3a receptor (C3aR) and TLR-4 present on human monocytes and monocyte-derived macrophages induces the production of proinflammatory molecules like IL-1β, tumor necrosis factor-alpha (TNF-α), IL-6, and prostaglandin E2 (PGE2) [14]. *In vivo* studies further support complement and TLR co-activation wherein cobra venom factor (CVF) synergizes with LPS to cause both an increase in complement activation and IL-6 production [14].

Outside of its roles in innate immunity, these interactions between complement and its cognate receptors present on immune cells, such as antigen-presenting cells (APCs), T cells, and

B cells also serves to prime and modulate the adaptive immune response through increased antigen presentation, synergistic co-stimulatory signaling, and modulation of B cell activation and antibody production [5,11,14,42,43]. Complement opsonization of target surfaces plays a critical role in complement receptor interactions on APCs to promote the uptake and processing of antigenic molecules, which can then be presented to T cells via major histocompatibility complex (MHC) molecules, resulting in their activation [42]. C3a and C5a fragments induce the upregulation of co-stimulatory molecules, such as CD80 and CD86, on APCs providing additional signals required for T cell activation through C3aR and C5aR receptor interactions on the T cell surface [42,44]. In addition to recruiting T cells to the site of complement activation, they have pronounced influence on their differentiation into effector subsets, such as Th1, Th2, or regulatory T cells, depending on the microenvironment, with studies having demonstrated that APC deficiencies in C3, C3aR, and C5aR have abrogated T cell proliferation and differentiation [43,45]. Moreover, complement components such as C3 and C5 are constitutively expressed by T cells, and C3a and C5a signaling through C3aR and C5aR enhance T cell proliferation and diminish apoptosis, actuating the expansion of the effector T cell repertoire following antigenic stimulation [42,44]. Studies have also shown that C3aR and C5aR signaling is essential for T cell homeostasis, and deficiencies in both result in accelerated cell death *in vitro* and *in vivo* [43].

Interactions between complement opsonins C3 and C4 present on antigenic surfaces presented by follicular dendritic cells with receptors found on B cells (CR1 and CR2) also serve to activate B cells and promote antibody production. These interactions have been shown to lower the threshold of B cell activation, enhance B cell receptor signaling, and drive B cell maturation and differentiation [11,42,46]. One study demonstrated that C3 cleavage products (C3d or C3dg) bind to complement receptor 2 (CR2 or CD21, which associates with CD19) on B cells and

amplifies B cell stimulation and activation. When mice were depleted of C3 or CR2 activation was blocked, the threshold dose of antigen required to elicit antibody was raised 10-fold [11,46]. This co-stimulation can promote isotype switching, leading to the production of antibodies with different immunoglobulin subclasses, allowing for induction of long-lasting antibody responses and improved immunological memory, while deficiencies in CR2 have been shown to impair humoral immune response [47].

Complement interactions with APCs, T cells and B cells aid in priming and improving protective adaptive immune memory and response for mounting effective immune responses against danger signals [11,14,42,46,47]. What is more, complement crosstalk with the coagulation system, inflammatory pathways, and adaptive immune cells form a complex network of interactions that orchestrate powerful immune responses and host defense against pathogens and altered self-cells [5,34,37,40].

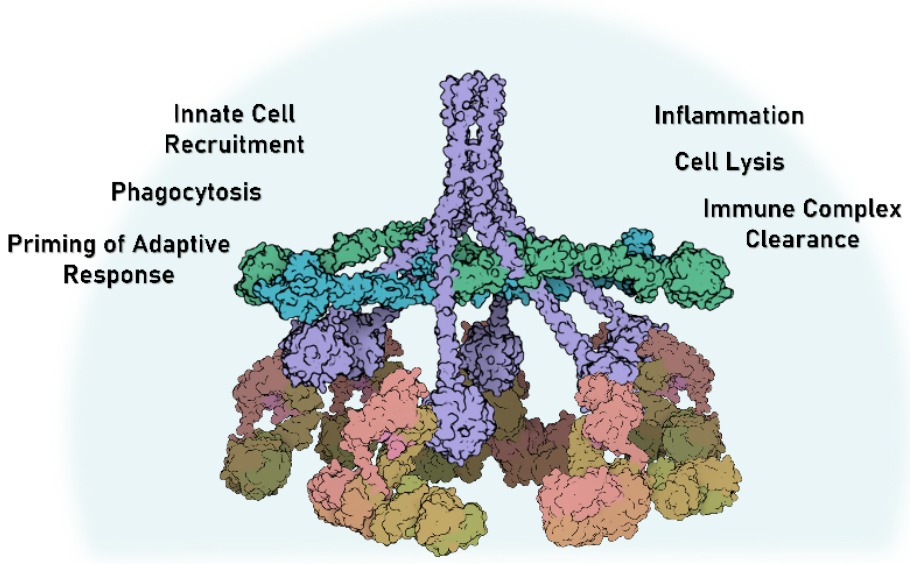


FIGURE 3. The complement system is involved in a wide range of functions and roles involved in immune homeostasis.

An interesting and relatively new area of research involves the presence and role of intracellular complement. Traditionally, complement has been studied in the context of extracellular expression. However, emerging evidence suggests that several complement components and regulatory proteins are also present within cells, contributing to intracellular processes [48]. While the functions of the complosome (intracellular complement) area still an area of active research, and the full extent of its involvement is speculative and not yet fully understood, studies have implicated intracellular complement components C1q, C3, and C5 in maintaining cellular activation and homeostasis and regulating cellular metabolism through cellular growth, autophagy of damaged organelles and proteins, lipid metabolism, and mitochondrial function [48–50]. In response to cellular stress stemming from oxidative stress, inflammation, and cellular damage, intracellular complement has been shown to be upregulated or have increased activity, contributing to both the initiation and resolution of cellular injury. However, inappropriate or overactivation of the complosome may contribute to host cell damage, as is implied with age-related macular degeneration (AMD) [51]. Moreover, during autoinflammatory disease settings, a recent study suggested that intracellular C3 and C3aR signaling in synovial fibroblasts may be an essential element for inflammatory tissue priming [52]. In the cases of intracellular infection, complement has, perhaps unsurprisingly, been suggested to undergo activation for the clearance and lysis of pathogens with some suggesting a role for cytosolic C3. Interestingly, one study demonstrated that C3-opsonized cytoinvasive pathogens can trigger cell autonomous immunity via NFkB and IRF signaling [53]. Additionally, intracellular complement has been shown to be present in immune cells, such as macrophages, dendritic cells, and T cells, where they are suggested to influence immune cell activation, maturation, and cytokine production [45,49], thus having a more pivotal role in immunity than what was previously believed

and continues to be an area of significant interest and research. A better understanding of the complosome and its functions may provide deeper insights into new therapeutic targets and strategies for disorders associated with inflammation and impaired immune responses, especially as it relates to C3.

1.2 Complement-mediated disorders

The complement system is a powerful tool for maintaining health and homeostasis; however, to ensure proper function, tight regulation must occur to prevent damage to host cells. Owing to the fact that most cells are capable of expressing complement components and those proteins present powerful effector roles with the ability to communicate with other biological pathways, dysregulation or overactivation of complement can lead to overwhelming or uncontrolled complement activation that can spell catastrophic consequences for the host [54].

It is important to remember that the complement cascade generates a series of inflammatory mediators and cytotoxic molecules that can promote the attack of host cells and tissues by immune cells distinct from the complement system, contributing to further injury and/or the pathogenesis of various diseases, including autoimmune disorders, inflammatory conditions, and organ damage [40]. Complement involvement in disease is complex and multifaceted and can range from acute to chronic and local to systemic conditions (**Fig. 4**).

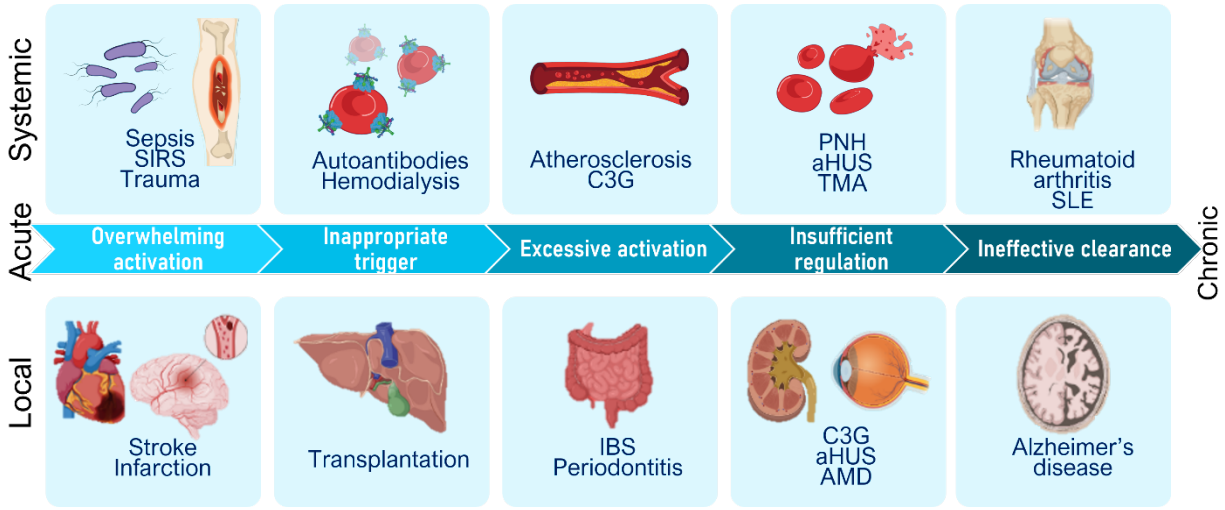


FIGURE 4. Involvement of classical complement activation in various disease settings. Dysregulation of the classical complement pathway can produce catastrophic disease in the individual at the local or systemic level. Autoimmune, inflammatory, and neurodegenerative diseases caused by classical pathway overactivation can produce both chronic and acute diseases in any bodily tissue [18]. Anti-complement drug treatment options are available for the treatment of complement 3 glomerulopathy (C3G), paroxysmal nocturnal hemoglobinuria (PNH), atypical hemolytic uremic syndrome (aHUS), and thrombotic microangiopathy (TMA) (Figure adapted from Ricklin, 2018 [55]).

Dysregulation of complement can occur through several means. Inherited genetic mutations in genes encoding complement proteins or regulatory proteins can disrupt the regulation and control of the complement system leading to either excessive complement activation or impaired regulation, causing damage to host tissues. One such example of this is hereditary angioedema (HAE), which is caused by a deficiency or dysfunction in complement regulatory protein C1-INH, leading to angio-edematous swelling that can occur anywhere throughout the body through the crosstalk and overactivation of the complement, coagulation and fibrinolytic

systems [56]. Tissue damage can also occur through classical pathway-mediated disorders where the presence of autoantibodies against self-tissues or organs can drive overactivation of the classical pathway through recognition and binding of C1q to the Fc region of autoantibody immune complexes [57]. Autoimmune disorders involving complement are exacerbated by increased proliferation of memory B cells propagated through complement co-stimulatory signaling, contributing to chronic or recurrent autoimmunity. Autoimmunity can also occur when the complement system inappropriately recognizes self-antigens as foreign and initiates an immune response against them and driving inflammation. In both cases, the effects of complement overactivation or impaired regulation serve as fuel to stoke increased immune imbalance [58]. Autoimmune diseases, such as systemic lupus erythematosus (SLE), rheumatoid arthritis (RA), atypical hemolytic uremic syndrome (aHUS), and paroxysmal nocturnal hemoglobinuria (PNH), among many others are often associated with complement dysregulation as the result of genetic mutations [57].

Chronic inflammatory conditions, such as the autoimmune diseases previously mentioned as well as inflammatory bowel disease, can lead to the persistent activation of the complement system through a positive feedback amplification loop, as inflammation can trigger complement activation pathways. These sustained inflammatory signals can result in chronic complement activation, perpetuating the inflammatory response and contributing to tissue damage, leading to the generation of additional danger patterns, and by these means, feeding into a vicious cycle [58]. Alternatively, in systems where complement is essential for clearance of cellular debris, deficiencies in complement can disrupt these processes, leading to impaired clearance of apoptotic cells and cellular waste and the accumulation of cellular debris, such is the case with Alzheimer's disease [59].

Similarly, persistent microbial infections that can activate the complement system can lead to dysregulation of complement activation. Notably, some pathogens, such as *Borrelia burgdorferi*, the causative agent of Lyme disease, can very effectively evade complement regulatory mechanisms, leading to further tissue injury and inflammation [60–63]. Prolonged or uncontrolled inflammation can disrupt normal organ functions and contribute to the development or progression of various diseases and/or comorbidities, such as arthritis, vasculitis, and neurodegenerative disorders [57,60]. On the other side of the coin, excessive or dysregulated complement activation can also impair the body's ability to fight infections, leading to an imbalance in the immune response, making individuals more susceptible to certain infections and impairing the clearance of pathogens. Individuals with deficiencies in complement components, and especially those deficient in C3 or C5, are predisposed to severe recurrent infections, notably of infections associated with Gram-negative bacteria, which are more sensitive to MAC lysis [64].

Acute traumatic injury that can lead to cellular stress or disruption of blood flow, such as decompression sickness/arterial gas embolism, traumatic brain injury, or ischemia-reperfusion injury, among others can also drive complement activation as blood flow is temporarily interrupted to a tissue or organ and then restored. The subsequent restoration of blood flow can then trigger complement activation and exacerbate inflammation, worsening the inflammatory response and initial injury [65–68]. Moreover, the potential for gas bubbles in the bloodstream is heightened and has been shown to activate both complement and platelets triggering inflammation, thrombocytopenia, and coagulopathy in *in vitro* and *in vivo* studies [69,70]. In such cases, complement components, and especially those specific to the classical pathway, show significant long-lasting, temporospatial-dependent increases in mRNA levels of C1q, C1r, and C1s proteins following injury [66,71].

Efficient regulation of the complement system is crucial to maintain a delicate balance between protective immune responses and avoiding collateral damage to healthy tissues. Dysregulation or overactivation of complement can have widespread detrimental effects on the body's tissues, immune responses, and overall health, contributing to the development, exacerbation, and progression of various diseases, not limited to autoimmune diseases, infections, inflammatory disorders, and complement-associated tissue damage.

1.3 Autoactivation of complement proteases

Most enzymes are synthesized as inactive zymogens, also known as proenzymes or precursor enzymes, which require specific activation steps and limited proteolytic cleavage events to become fully physiologically functional [72]. This feature of enzymes is essential for regulating enzyme activity and preventing unwanted or uncontrolled reactions as zymogen activation is an irreversible reaction to stimulus, e.g., complications arising from overactivation of the complement system. The synthesis of enzymes as zymogens ensures that the enzymes are produced in an inactive form, preventing them from catalyzing reactions prematurely or inappropriately within the cell or during protein synthesis, maintaining cellular homeostasis and preventing host cell damage [73]. This allows for precise regulation and control of enzyme activity through limited proteolysis, which is characterized by the ability of the enzyme to complex with a specific macromolecular substrate, cleaving only one or two predetermined peptide bonds (as distinct from the process of digestion) and releasing the product for subsequent activity within the system [74]. This activation step is irreversible as there are no biological means by which to repair the cleaved peptide bond [73]. Proteolysis is an exergonic reaction, resulting in the release of free energy (G); however, specific physiological or environmental cues, such as changes in pH or the presence of

specific enzymes or cofactors are required [75] to finely tune and regulate the activity of enzymes in response to the specific temporal and spatial needs of the individual or specific cellular events, and, importantly, prevent deleterious effects, such as tissue damage, inflammation, thrombosis, or metabolic disorders [76]. Proteases that are capable of autoactivation do not necessitate the involvement of other enzymes for activation and instead rely on self-association and proteolysis via active site generation through conformational changes [77] and very few are capable of autoactivation in the absence of extrinsic factors.

In the complement system, pathway-initiating serine proteases, C1r, MASP-1, and MASP-2 can undergo autocatalytic activation to rapidly respond to danger-associated signals and propagate activation of their respective pathways [77–80]. This autocatalysis process involves the self-association and cleavage of specific peptide bonds found within the proenzyme forms of these proteases (**Fig. 5**). Under normal conditions, C1r circulates in serum bound within the C1 complex and conformational changes incurred upon C1q-antigen binding allow for activation. Autocatalysis of C1r involves the cleavage of the peptide bond between Arg^{463} and Ile^{464} within the flexible scissile loop, also known as the activation peptide. This self-cleavage event generates an active form of C1r that can further activate downstream components of the complement system. Similarly, MASP-1 and MASP-2, of the lectin pathway, also undergo autocatalytic activation. In their inactive proenzyme forms, MASP-1 and MASP-2 exist as either a homodimer or heterodimer in complex with MBL, ficolin, or collectin. Upon recognition and binding to carbohydrate antigen by the PRM, the MASP complex is activated through the Arg-Ile peptide cleavage of the activation peptide, initiating downstream complement activation events.

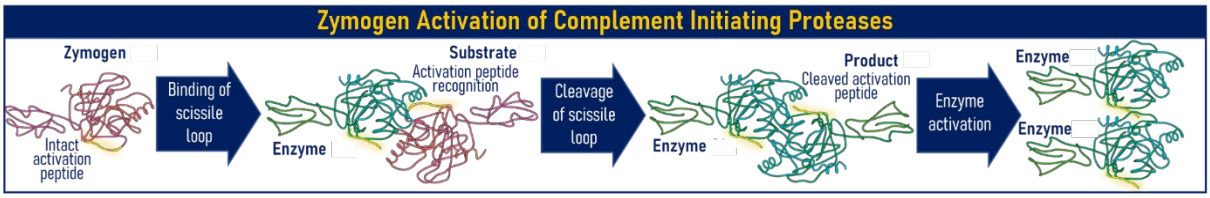


FIGURE 5. Autocatalytic activation of zymogens to enzymes. Enzymes engage in the process of autoactivation when the inactive precursor enzyme, also known as a zymogen, undergoes limited proteolysis, wherein a nearby "self" enzyme recognizes and self-associates with the zymogen. The flexible scissile loop of the zymogen binds within the substrate binding groove of the enzyme, forming an enzyme-substrate complex. The bond between the P1 and P1' residues is hydrolyzed and at this state the reaction is completed, and the zymogen is transformed from substrate to product [81]. This cleavage event triggers conformational changes within the zymogen, resulting in the formation of the oxyanion pocket and fully formed active site, thereby converting the zymogen to its active enzyme form[82]. The newly cleaved activation peptide is released from the active site of the original enzyme and the enzyme-product complex dissociates with the generation of activated enzyme [83].

Proteolytic cleavage is characterized by the cleavage of the peptide bond occurring between residues deemed P1-P1'. The P1 residue refers to the amino acid residue immediately preceding the scissile bond, while the P1' residue refers to the amino acid immediately following. On the enzyme side of the reaction, the S1 residue refers to the amino acid residue within the active site of the protease that interacts with the P1 residue. S residue nomenclature follows that of the P residues where the P2 residue preceding the P1 residue interacts with the S2 residues, and so on

[82]. The specific amino acid sequences surrounding the cleavage site determine the substrate specificity of the protease while guiding the autocatalytic cleavage process.

Following autoactivation and cleavage of the scissile loop, the activation peptide forms the leaving group and exits the enzyme active site. Conformational changes are incurred by the newly activated enzyme allowing for the formation of the oxyanion hole and substrate specificity pocket that was previously unformed and occluded in the zymogen state [84]. In a healthy individual, the autoactivation mechanism of the classical and lectin pathways ensure that the pathway remains inert until the presence of specific antigenic stimuli triggers the proteolytic cascade.

1.4 Complement therapeutics

Complement-directed therapeutics have emerged as a rapidly growing and promising area of research and development for the treatment of various complement-mediated diseases. For its part in maintaining immune homeostasis, dysregulation of complement can significantly contribute to the pathogenesis of several diseases, including autoimmune, inflammatory, and neurodegenerative diseases. Within the complement-targeted therapeutic space, there are several complement inhibitors either on the market or in advanced clinical trials. Most widely recognized in the field, recombinant anti-C5 monoclonal antibody, eculizumab, inhibits the generation of anaphylatoxin C5a and the formation of terminal pathway MAC, which is responsible for the lysis of red blood cells. Eculizumab, developed by Alexion Pharmaceuticals, has been approved for the treatment of paroxysmal nocturnal hemoglobinuria (PNH), atypical hemolytic uremic syndrome (aHUS), myasthenia gravis, and neuromyelitis optica (NMO) through the prevention of prothrombotic effects and inflammation [85–88]. Similarly, ravulizumab is a humanized

monoclonal antibody targeting C5 and is approved for both the treatment of PNH and aHUS, providing longer dosing intervals as compared to eculizumab [89].

Anti-C1s monoclonal antibody, Sutimlimab, has undergone FDA approval for treatment of CAD, an autoimmune hemolytic anemia, due to its ability to prevent opsonization by complement of red blood cells following recognition of IgM autoantibodies against RBC antigens (cold agglutinins) active at cold temperatures. It is has been shown that inhibition at the level of C1s prevents the generation and opsonization of RBCs by C3b and the subsequent recruitment of receptor-specfic macrophages resulting in temperature-dependent hemolysis [90]. Additional support for therapeutics targeting specifically classical pathway-mediated diseases is demonstrated by anti-C1q monoclonal antibody, ANX005, which is currently in phase 2 clinical trials evaluated by Annexon Biosciences for Huntington's disease (HD), warm autoimmune hemolytic anemia (wAIHA), amyotrophic lateral sclerosis (ALS), and phase 3 clinical trials for Guillain-Barré syndrome (GBS) [91]. Applied broadly across a number of conditions, inhibition of C1q can have therapeutic use by blocking C1q interactions with host targets both dependent and independent on the presence of autoantibodies. While classical pathway targeted therapeutics are showing promise across a wide swath of pathologies, the development of additional classes of inhibitors, including small-molecule inhibitors, is essential for improved treatment outcomes for patients.

Small-molecule inhibitors targeting C1s provide an additional tool to move the needle in the treatment of complex systems. Advantages to small-molecule drug development is increased tissue permeability and patient compliance in an orally bioavailable drug. Moreover, detailed SAR analyses of currently available small-molecule inhibitors in complex with protein targets with structures available in the PDB show that most replicate some component of human biology and targeted interactions already present in nature [92]. Protein modification by enzymes, such as

phosphorylation or glycosylation, results in changes to structure and therefore protein function, which is what small-molecule drugs frequently replicate through molecular mimicry. These protein-ligand interactions compete for natural ligands or cofactors for enzymes, and the utilization of these SAR data and structural information derived individually from protein target and ligand are heavily leaned upon for drug development [93].

Complement therapeutics directed towards targeting downstream complement components, C3 and C5, or MAC, common across all three pathways, may mitigate complement-mediated tissue damage but are challenged by higher concentrations of those downstream components and higher turnover rates [87]. Complement proteins have a relatively short half-life, and the body has a robust system for replenishing them, especially those of C3 and C5, so even if a therapeutic agent effectively inhibits or modulates the activity of C3 or C5, their levels can quickly rebound once the treatment is discontinued, providing a challenge for the long-term efficacy of complement-targeted therapeutics [87].

Targeted therapeutic intervention of complement relies on the specific involvement of complement in driving pathophysiological mechanisms underlying disease as well as the desired therapeutic effect. Targeting upstream complement proteins at the level of the initiating proteases allows for the selective inhibition of that pathway, where earlier intervention has a broader impact on downstream complement activity. Specific targeting of C1r, C1s, or the MASPs may potentially minimize off-target effects and reduce the risk of interfering with other crucial immune functions. This pathway selectivity may offer a more favorable safety profile when compared to targeting downstream complement components, C3 or C5, due to the fact that the other pathways still remain active against infections. Furthermore, targeting the upstream initiating proteases also affords selectivity for diseases modulated by specific pathways, such is the case with classical pathway-

mediated diseases triggered by the recognition and binding of immune complexes by C1, like that observed in heparin-induced thrombocytopenia (HIT) [94]. Studies have confirmed that bacterial lipoprotein, BBK32, derived from *B. burgdorferi* specifically blocks C1r activity and downstream activation of the classical pathway thereby attenuating prothrombotic effects of the condition [95].

Additionally, specificity at the level of the pathway-initiating serine proteases rather than C1q may further enhance the safety profile to mitigate potential off-target effects and minimize interference with other essential immune functions mediated by C1q. C1q has diverse roles beyond its involvement in complement activation, including its interactions with various immune cells and its contribution to tissue repair [96]. C1q opsonization of microbial pathogens and damaged host cells facilitates the interaction between the target and CR1 receptors on phagocytes, leading to phagocytosis and elimination of the target. In transgenic mice lacking C1q (APPQ-/-) and overexpressing the human mutant amyloid precursor protein (APP), C1q was shown to have an anti-inflammatory effect through enhanced phagocytosis of apoptotic cells and blebs, modulating cytokine expression and inflammation. However, in the presence of C1r and C1s, in which it formed complexes, C1q displayed a proinflammatory phenotype resulting in enhanced neurodegeneration [97]. The beneficial roles of C1q outside of the C1 complex and classical pathway activation are not well understood but highlight its multifaceted functions in anti-inflammatory responses, consequently preserving the beneficial functions of C1q while specifically inhibiting the proteases involved in excessive complement activation may offer fewer off-target effects in therapeutic interventions.

Recent advancements in medicine and data processing have helped to increase our understanding of the intracacies of disease manfiestation and mechanisms. Drug development as it currently exists, functions as an outcome of the clinical presentation of patient symptoms for

target identification, but there are many complex biological processes underlying clinician reports that cannot fully take into account individual factors, such as temporality involving disease evolution and trajectory, patient comorbidities, genetic profiles, epigenetic factors, and clinical history, among others. The overemphasis on symptomology that ends up boxing or categorizing patient cohorts into highly defined categories minimizes individual patient complexity and leads to inadequate accounting for critical patient heterogeneity, which can then lead to the misidentification of drug targets. It is becoming increasingly clear that for various diseases, additional drug classes are necessary for effective treatment of disease, and in the complement realm, treatment options outside of C3 and C5 and intravenous administration of biologics is essential for handling a hetergenous patient population.

1.5 Small-molecule inhibitors: Competitive vs. noncompetitive inhibition

Small molecules play a major role in drug synthesis and development. In fact, according to industry reports, nearly 90% of the approved drugs on the market are small molecules [98]. Small-molecule drugs widely utilized due to their ease of synthesis, oral bioavailability, and ability to target intracellular pathways [99]. Small-molecule inhibitors represent a class of drugs characterized by organic compounds with relatively low molecular weight, typically less than 1 kDa, and are synthesized or obtained through various methods, including chemical synthesis, natural product extraction, and combinatorial chemistry [100]. Such small-molecule drugs have been successful in treating a wide range of diseases, including cardiovascular disorders, cancer, infectious diseases, and metabolic disorders, among others through the binding of active site or orthosteric pockets found on the surface of target proteins to modulate activity, conformation, and/or function of the target [101].

The process of small-molecule drug development consists of several steps, including target identification, hit identification, and hit to lead and lead optimization. Hit identification involves the screening of a large number of small molecules to identify compounds that exhibit the desired pharmacological and bioactivity against a specific target, traditionally typified by high throughput screening [100]. After identification of a hit compound, lead optimization using structure-activity relationship (SAR) studies aims to modify the chemical structure of the compound to improve chemical qualities, including binding affinity to the target, inhibition, selectivity, and pharmacokinetics [100].

During compound synthesis and optimization, many guidelines exist to both lead and assess the drug-likeness of small molecules. Lipinski's Rule of Five is one such set of criteria, in which all values are multiples of five, that predict the likelihood of a small-molecule compound having favorable pharmacokinetic properties, such as solubility and permeability, and oral bioavailability [102]. The rules state that a compound is more likely to be orally active and have favorable molecular properties, such as absorption, distribution, metabolism, and excretion (ADME), if it has no more than one violation of the following criteria: the compound retains a molecular weight of no more than 500 Da, possesses a logP (partition coefficient) value of less than or equal to 5, has no more than five hydrogen bond donors and 10 hydrogen bond acceptors [102]. Lipinski's Rule of Five has been widely used as a starting point for evaluating the drug-likeness of small molecules; however, other factors, such as target specificity, toxicity, and pharmacodynamic properties, also play crucial roles in the overall evaluation and optimization of small-molecule drugs.

Their low molecular weight affords them several advantages over other drug classes. One such feature is that they are typically designed for oral bioavailability, which is a more comfortable

and convenient method of administration for patients, allowing for self-administration and greater patient compliance . In contrast, other drug classes like biologics often require intravenous or subcutaneous administration, which can be more complex and require healthcare professional assistance [103]. Furthermore, small molecules have greater ability to efficiently penetrate various tissues and cross cellular membranes, including the blood-brain barrier, and can be optimized to do so, enabling them to reach their targets in different parts of the body, which is essential for treating diseases that may involve multiple organs or tissues, such as cancer or infectious diseases [104].

In terms of synthesis, small-molecule inhibitors are generally easier and less expensive to develop and manufacture compared to other drug classes like biologics. The process of synthesizing small molecules can be standardized and scaled up, facilitating large-scale production, while the stability and storage requirements for small molecules are often more straightforward than for biologics, reducing logistical challenges [105]. Additionally, this inhibitor class can be utilized combinatorically with other small molecules or drug classes to generate oral combination therapies allowing for synergistic effects, improved efficacy, and the potential to target multiple disease pathways and comorbidities simultaneously [106,107].

An important consideration for the development of small-molecule drugs to ensure long-term efficacy and patient safety for the treatment of chronic conditions is antigenicity over time. While small molecules are less likely to elicit an immune response compared to larger biological molecules, such as modified proteins or antibodies, it is possible that they can trigger an immune response and potentially become antigenic over time, due to metabolic processing, wherein the generation of reactive metabolites can form neoantigens with endogenous proteins [108]. Similarly, these smaller metabolites may possess chemical properties that lend themselves to

binding larger carrier proteins within the body. This hapten-protein complex in circulation can trigger antibody-mediated immune responses against the complex [109,110]. Concerns with small molecule antigenicity can be addressed through compound structural modifications, wherein antigenicity is reduced or eliminated. Chemical modifications may involve altering specific functional groups or substituents that are known to trigger immune responses; however, drug-likeness screening guidelines for compound development and optimization should factor in immunogenicity and toxicity prior to synthesis [111]. Computational tools evaluating compound pharmacokinetics, such as SwissADME, are readily available; however, *in vitro* and *in vivo* assays screening for immunogenicity should still be employed to evaluate potential immunogenic effects of the compound [111]. Detection of immunogenicity or especially immunogenic chemical entities should guide the optimization and substituent replacement of those chemical groups to reduce the likelihood of an immune response.

Rational design and considerations for the chemical properties of small molecules is an essential part of the drug development process; however, the first great challenge to be overcome in the process is identification of a target protein within disease and identification of "druggable" binding pockets on the solvent exposed surface of the protein [112]. To reach that point, sufficient analysis of protein structure and function must be undertaken to further knowledge regarding the pathophysiological roles of the target in disease mediation and progression. Surface plasmon resonance (SPR) is one such tool for elucidating the mechanisms underpinning protein function through targeted binding assays characterizing the biomolecular interactions between protein-protein complexes, protein-small-molecule interactions, protein-cofactor dependence, and/or protein-bilayer interactions in the case of membrane proteins [113]. Determination of factors that necessitate protein activity is essential for the development of efficacious drug candidates.

Following target discovery, small molecule or fragment identification of molecules that interact with the target and inhibit or alter activity must occur. This “hit” identification generally takes the form of high-throughput screening, such as that performed by SPR or *in silico* virtual screening of compound libraries [114]. Once a hit(s) has been identified, lead compounds undergo a series of preclinical biofunctional evaluations to determine bioactivity against the target, efficacy, safety, pharmacokinetic properties, and toxicology (ADME) [115]. Based on these results in combination with SAR data, compounds are optimized through structural modifications to improve favorable or druglike characteristics, such as the incorporation of polar groups to improve solubility [116]. Once drug-likeness and bioactivity of the compound has reached optimal levels, an investigational new drug (IND) application including preclinical data, proposed clinical trial plans, and manufacturing details is submitted to regulatory authorities (such as the FDA in the United States) to request permission to begin clinical trials, which are conducted in multiple phases [117]. Phase one trials involve a small number of healthy volunteers to assess safety, dosage, and pharmacokinetics. Once phase two trials are reached, clinical evaluation is ramped up and involves a larger number of patients and assess effectiveness and further safety. The final phases, phase three and phase four clinical trials, are large-scale trials conducted on a larger patient population to confirm efficacy, monitor side effects, and compare the drug to existing treatments over long periods of time [118].

Provided the results from clinical trials are positive, the drug developer submits a new drug application containing comprehensive data on the drug's safety, efficacy, manufacturing, labeling, and proposed use to the FDA or regulatory authorities [119]. Regulatory authorities review the data to determine if the benefits of the drug outweigh the risks and if they are deemed to be safe and effective, approval is granted for marketing and commercialization allowing for the drug to be

made available to healthcare providers and patients. Regarding immunogenicity, post-marketing surveillance is conducted to monitor the drug's safety and effectiveness in a larger patient population, with adverse events and new information about the drug's risks or benefits evaluated and reported [108]. The traditional timeline and process for a biopharmaceutical company developing a small-molecule drug from discovery to market will span over a decade.

Unfortunately, the process of drug development is rife with challenges and the majority of drug candidates never pass clinical trials. In fact, in the last year alone, 20,109 drug candidates were undergoing development, and yet only about 30-40 drugs are approved each year [120,121]. The greatest hurdle for drug developers is maintaining safety and efficacy in drug candidates. During the phases of preclinical and clinical development, candidates undergo extensive testing to evaluate the drug safety profile and effectiveness in treating the disease target, and those that exhibit unexpected toxicities or off-target effects are terminated from further development and trials. This is largely due to insufficient target specificity, in which the drug candidate interacts not only with the target but also with unintended targets, leading to adverse and/or toxic reactions [122].

A large portion of drugs in development are for those targeting serine proteases for inhibition. Serine proteases are attractive drug targets as they play critical roles in numerous physiological processes and disease conditions, such as blood clotting, digestion, inflammation, and as in the case of complement proteins - immune response, making them potential therapeutic targets for a wide range of pathologies [123]. As previously described, proteases must undergo tight regulation to avoid damage to host cells. Dysregulation or overactivation of serine proteases has been implicated in various diseases, including cancer, cardiovascular disorders, autoimmunity, inflammatory diseases, and infectious diseases, consequently by selectively inhibiting the activity

of specific serine proteases, it is possible to modulate these disease processes and potentially develop effective treatments.

Additionally, the structural homology and catalytic mechanism of serine proteases has been extensively studied and well defined offering the advantage of immense, available SAR data from which to develop an inhibitor. The active site of serine proteases contains a conserved catalytic triad of a serine, histidine, and aspartate residue, which are responsible for catalyzing the hydrolysis of conserved peptide bonds [82,83,124]. This catalytic mechanism provides a well-characterized target for designing small-molecule inhibitors that can specifically bind to and block the active site of serine proteases. Despite this wealth of information regarding serine proteases, they stand as difficult targets for drug development. The most notable challenge is achieving selectivity and specificity in light of structural homology and conserved active site residues between serine proteases [123]. There are many different serine proteases present in the human body representing biologically diverse pathways. Developing inhibitors that selectively target one specific serine protease while sparing others to avoid off-target effects and potential toxicity remains a challenge to the field. Furthermore, serine proteases often exist in zymogen or inactive forms and require proteolytic activation to become functional. This activation step adds another layer of complexity to developing effective inhibitors as many inhibitors are designed to selectively target the active form of the protease and knowledge of which specific protease is involved in disease populations is essential [125]. While serine proteases are implicated in a growing number of pathologies, overcoming obstacles associated with selectivity and specificity, as well as addressing enzyme activation state, make them a complex and challenging class of drug targets.

At present, most serine protease inhibitors function to alter or inhibit target activity through means of competitive inhibition or binding of the active site to outcompete native substrates. In addition to the challenges that come with active site specificity for a particular serine protease target, achieving effective inhibition often requires high concentrations of the competitive inhibitor to outcompete the substrate for binding to the active site, especially in the case of proteins with high turnover rate [123]. While there are obstacles that must be overcome in the development of active site inhibitors, recent advances in technology have allowed for exceptionally high-resolution structural data to visualize unique molecular features present in specific protein targets aiding in structure-guided drug development.

Another method for overcoming structural homology between the family of serine proteases and the drug target is noncompetitive or allosteric inhibition. This occurs through the identification and binding of solvent-exposed pockets present on the target surface that exist distally from the active site, yet through cooperative motions, can modulate protein activity [126,127]. Pockets that exist outside the proximity of the active site have far greater potential to confer architectural and structural heterogeneity from other related proteins [74]. Binding of these sites can modulate protein function by blocking conformational changes that are requisite for activity or by distally altering the active site, making it less accessible to the substrate or reducing its catalytic activity [101]. Unfortunately, identification of allosteric sites of inhibition and designing small-molecule inhibitors that bind specifically to those sites can be challenging. Additionally, understanding allosteric interactions through correlated and anti-correlated protein motions can be complex and context-dependent [128], making it more difficult to predict the effect of an allosteric inhibitor on the enzyme's activity.

One method of identifying noncompetitive sites of inhibition is to examine serine protease inhibitors found in nature, which can present valuable starting data for rational drug design. Natural inhibitors are compounds produced by organisms that have evolved to interact with specific target proteins or enzymes, and accordingly, exhibit high affinity and specificity for their target as they have undergone extensive selective pressures and evolution in their function and design [129]. Identifying established natural inhibitors and their modes of action can provide valuable insights into the design of synthetic inhibitors through molecular mimicry, especially of those targeting noncompetitive inhibition (**Fig. 6**). By determining the three-dimensional structure of the inhibitor-target complex, researchers can understand the key molecular interactions that contribute to inhibitor potency and selectivity. This knowledge can guide the design of synthetic inhibitors with improved binding affinity and target specificity. Additionally, they can aid in the identification of structural or functional elements necessary for binding to the target protein and blocking activity, allowing medicinal chemists to leverage the structural insights to employ rational drug design strategies through functional group modification or replacement to improve compound potency, selectivity, and drug-like properties.

Borrelia is a genus of bacteria that includes the species *Borrelia burgdorferi*, the causative agent of Lyme disease, and has emerged as a model system for studying complement evasion strategies due to its ability to evade and block the host complement system [61]. The *B. burgdorferi* life cycle involves transmission between ticks and mammalian hosts, including humans. Once it reaches the human host, the bacteria must inherently interact with elements of the host innate immune system, most notably the complement system. The ability of *B. burgdorferi* to evade complement-mediated killing is critical for its survival, infectivity, and persistence in the host [62].

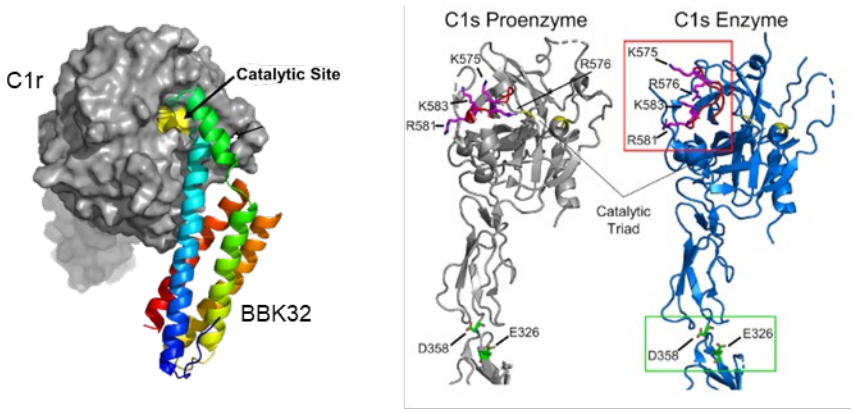


Figure 6. *B. burgdorferi* lipoproteins, BBK32 and ElpQ, provide a basis for complement protease inhibition. Lipoproteins expressed on the surface of borrelial pathogens serve as guides for the development of competitive and noncompetitive inhibitors by driving target validation and elucidating SAR information through the binding mechanisms of protein-inhibitor complexes formed during *B. burgdorferi* complement evasion.

B. burgdorferi has evolved multiple evasion mechanisms to counteract complement-mediated immune responses, including inhibition of classical pathway initiating proteases C1r and C1s through targeted active site and noncompetitive site inhibition [62]. By studying these mechanisms, we can gain insights into the molecular mechanisms and evasion strategies used by pathogens to block classical pathway activation for the development of therapeutic inhibition of the complement proteases, C1r and C1s, by small-molecule drug development (**Fig. 6**).

These pathogen-host interactions also provide valuable insights for structure-based drug design (SBDD), in which utilization of the three-dimensional structure of a target protein is used to guide the development of small-molecule inhibitors. Various computational techniques, such as X-ray crystallography, molecular docking, and quantitative structure-activity relationship (QSAR)

analysis, can be employed to identify lead compounds predicted to possess the desired bioactivity against the target. When crystal structures of the target protein are available, leveraging information from the crystallographic waters present is also an effective strategy in structure-based drug design. This information can be utilized for the identification of noncompetitive binding sights, as waters will occupy cavities or clefts within the protein structure, indicating potential interaction sites for small-molecule ligands [130]. By analyzing the positions and orientations of crystallographic waters, insights into the location and geometry of the binding site can be developed, guiding the design of ligands that can interact with the protein [131]. In the case of competitive inhibition, crystallographic waters can mediate interactions between the protein and ligand molecules and inform where hydrogen bonds or other types of interactions with both the protein and the ligand may occur, contributing to the increased stability and specificity of the protein-ligand complex [131]. Crystallographic waters can provide valuable information about the displacements or rearrangements of water molecules upon ligand binding, which is crucial for accurately predicting the energetics and thermodynamics of ligand binding and for designing small molecules that can efficiently displace or compete with the crystallographic waters in the binding site [130]. Valuable structural information that can guide structure-based drug design can be gained through analyses of interactions between naturally derived inhibitors and waters to support the design and optimization of potent and selective small-molecule inhibitors.

1.6 Fragment-based drug discovery

Fragment-based drug discovery (FBDD) is an effective strategy in drug development that focuses on identifying small, low molecular weight fragments that bind to a target protein. These fragments serve as starting points for the development of more potent and selective drug

candidates. FBDD differs from traditional high-throughput screening (HTS) approaches, which typically screen larger, more complex compounds, whereas a chemically and structurally diverse library of fragments is screened by *in silico* or biophysical techniques, such as SPR or computation molecular docking. Similar to small-molecule hit identification, fragments are identified and scored based on predicted free energy binding affinities. Those hits that are identified undergo optimization to generate the next generation of analogs with improved chemical and physical properties, especially as it relates to improved binding affinity and selectivity (**Fig. 7**). Due to their inherent low molecular weight and complexity, fragments can be optimized through fragment growing by which functional groups are added to generate a larger, more potent molecule, fragment linking in which two fragments that bind the target in close proximity to one another are linked via synthetic linker, or by fragment merging wherein fragment binding conformations that overlapped are merged to create one molecule with improved binding affinity [132].

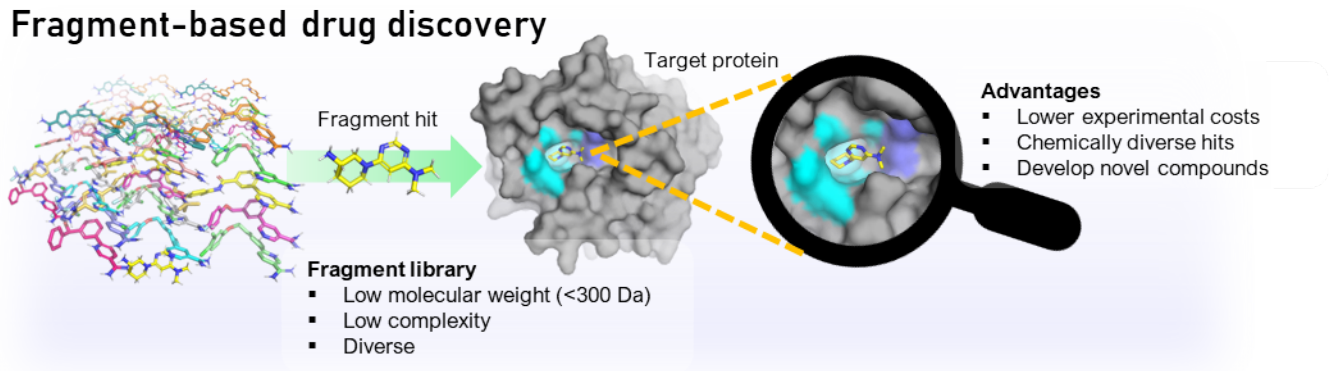


FIGURE 7. Diagram illustrating the process involved in fragment-based drug discovery.

One strategy for improved bioactivity of the fragment is the utilization of the Craig plot or the Lipophilicity Efficiency (LipE) plot, which is a graphical representation that aids in the

optimization of drug-like compounds through the visualization of lipophilicity and potency of chemical entities and substituent groups within the compound. When examining a Craig plot, the lower left quadrant represents groups with favorable low lipophilicity but low potency, whereas the upper left quadrant represents compounds with low lipophilicity and desirable high potency. Meanwhile, the lower right quadrant represents chemical entities with unfavorable high lipophilicity and low potency, and the upper right quadrant represents compounds with unfavorable high lipophilicity and high potency. Consequently, the compounds that possess substituent groups falling in the upper left quadrant of the Craig plot, known as the "Golden Triangle," are considered the most promising candidates for further development [133]. These compounds exhibit optimal balance, with sufficient potency and favorable lipophilicity and are more likely to possess desired drug-like properties such as good bioavailability, appropriate pharmacokinetics, and reduced risk of off-target effects [133]. During compound optimization, medicinal chemists can guide the optimization of lead compounds through bioisoteric or ring replacement towards regions of the plot associated with improved drug-like properties, increasing the chances of identifying successful drug candidates.

Similar to the design of small-molecule compounds, fragment-based drug design also follows loose guidelines that aid in the development of inhibitors with favorable physicochemical properties. Designated the rule of three, these guidelines define a fragment as that which is no larger than 300 Da, has a cLogP value of less than three, and possesses no more than three hydrogen donors and acceptors [134]. The employment of these measures for fragment designation ensure that the chemical space occupied by these molecules remains less complex so as to increase the binding interactions available to these compounds and allow for the subsequent development of

more potent leads [135]. The objective in compound optimization is to identify compounds that possess both favorable potency and lipophilicity properties.

Utilization of FBDD allows for the employ of fragment libraries that cover a wide chemical space allowing for increased chemical diversity and chances of identifying hits with novel and unique binding interactions. The low molecular weight and complexity also offers the advantage of identifying and characterizing fragments that have high ligand efficiency with strong binding interactions per atom, potentially yielding lead compounds that can be developed into potent and selective drugs with optimal pharmacokinetic properties [136].

At present, there are several drugs on the market developed using FBDD or had FBDD as part of their discovery process, validating this approach to guiding the development of novel fragment-based small-molecule drug candidates. Notable examples include vemurafenib, which became the first FDA-approved fragment-derived drug in 2011 and is a kinase inhibitor developed by Plexxikon for the treatment of metastatic melanoma with specific BRAF mutations [137]; venetoclax,, which is a BCL-2-selective inhibitor that inhibitors the growth of BCL-2-dependent tumors in patients with refractory chronic lymphocytic leukemia [138]; erdafitinib, approved in 2019 with highly selectively, potent tyrosine kinase inhibitory activity against fibroblast growth factor receptor (FGFR) family members involved in proliferation of oncogenic tumors [139]; and pexidartinib, which inhibits colony-stimulating factor 1 receptor (CSF1R) by trapping the kinase in the autoinhibited conformation in patients with tenosynovial giant-cell tumors [140]. Technological advancements and improved computational methods for hit-to-lead generation have garnered significant interest for the adoption of FBDD and its impacts on the development of new therapeutics is expected to see continued growth in the future.

1.7 Artificial intelligence-based drug design and AlphaFold

Machine learning and artificial intelligence (AI) have become increasingly important in drug design and discovery. At present, the race is on for pharmaceutical companies to produce the most effective algorithms designed the identify disease targets and small molecule/fragment hits and guide optimize for synthesis of potent drug candidates thereby reducing costs incurred through prolonged research and development and drug candidates terminated during preclinical and clinical trials. Machine learning algorithms can be trained on large datasets of known molecules and their corresponding biological activities to build predictive models that can be used to predict the activity, properties, or safety profiles of new lead compounds [120] . These predictions help to score and prioritize compounds for further evaluation, reducing the need for extensive experimental testing. In this process, AI algorithms provide improved efficiency for virtual screening efforts, where libraries of millions of compounds and chemical scaffolds are screened computationally against a target protein to identify potential hits [141]. Machine learning models can learn from known active and inactive compounds to predict the likelihood of a compound being active against a specific target [142]. An alternative approach is utilizing machine learning algorithms to generate novel molecules with desired properties. By mining medical and scientific literature of known drugs and learning from existing compounds and their structure/function, AI can generate new chemical structures that are likely to be active against a given target or have specific pharmacological properties with improved drug-likeness [92].

On a structural level, AlphaFold, an AI-based technology, has made significant contributions to rational drug design. AlphaFold uses deep learning techniques to predict protein structures based on their amino acid sequences and by accurately predicting protein structures, AlphaFold has aided in understanding protein functions and interactions, including the

identification of binding sites for small molecules [143]. This information is invaluable in the rational design of drugs, as it provides insights into how potential drug candidates may interact with their target proteins.

The use of AI and machine learning has improved drug design efficiency by accelerating the screening and optimization of large compound libraries, reducing the number of compounds that need to be synthesized and tested experimentally, and prioritizing the most promising candidates for further development. AI can also help optimize lead compounds by predicting their pharmacokinetic properties and toxicity profiles, reducing the likelihood of failures in later stages of drug development [120]. As mentioned previously, the traditional state drug development means that fragment hit identification to potent lead compound generation is a process that can take many years. Advancements in drug development through AI has seen that process reduced to mere months in some cases. With the continuous advancements in AI algorithms, increased availability of large-scale datasets, and improved computing power, AI is expected to further enhance drug development efficiency, reduce costs, and lead to the discovery of novel therapeutics for various diseases.

CHAPTER 2: Structural insights towards the molecular mechanism of autoactivation of MASP-2

Abstract

Enzyme activation and regulation, particularly of proteases involved in cascade pathways, has been of intense interest to the fields of structural biology and therapeutic drug development, in part due to propensity and the devastating consequences of dysregulation or overactivation in these proteases in human pathologies [127,144]. Serine proteases represent nearly one third of all known proteases, and appropriately, extensive research has delved into the molecular mechanisms underlying their function and activation processes [144,145]. Here, the X-ray crystal structure of human complement serine protease MASP-2 informs a more detailed view of the structural dynamics that occur during the end stages of autocatalytic activation wherein an enzyme-product complex is formed between self-associated MASP-2 proteins. The 1.6 Å resolved structure captures a static glimpse of the binding interactions between the cleaved scissile loop of the leaving MASP-2 protein product and the substrate binding groove of the enzyme MASP-2, which impart a significant conformational rearrangement of the catalytic Ser^{633} that is essential for proteolysis. The protein-protein interface that is formed provides an additional model of proposed MASP-2 autoactivation, which aligns with the known enzyme-product complex formed by structurally and functionally homologous classical complement pathway protease, C1r. The derived structural data highlight novel insights into the requirements of complement activation and provide additional structural support for the development of complement-targeted therapeutics.

Introduction

Enzymatic cascades participate in a number of reaction pathways and amplification systems involved in regulation and homeostasis by driving essential physiological processes relating to digestion, coagulation, fibrinolysis, metabolism, and immunity [72,146]. Though disparate in biological consequence, all cascades require an initial stimulus to initiate a series of concurrent steps of enzyme activation primed to meet the cellular demands of that pathway [72]. Due to the irreversible and often cooperative nature of enzymatic cascades, tight regulation of these systems is vital for maintaining bodily health and homeostasis, especially at the level of the pathway-initiating enzymes [33,73,123]. One measure of intrinsic regulation is the condition that most enzymes are synthesized as inactive precursor molecules that require site-selective proteolytic cleavage for activation and most activation processes also necessitate extrinsic factors [73].

The complement system is a proteolytic cascade of approximately 50 proteins present in the blood that retain the primary role of serving as the first line of defense against the deleterious effects of pathogenic microbes, misfolded proteins, apoptotic cells, and cellular debris, among others [147]. The complement pathways synergize with many other biological processes and in a dysregulated system, this potent effector function can rather serve to drive pathologies associated with inflammation, autoimmunity, and neurodegeneration [40,54,57]. The complement system is divided into three different pathways: the classical, the alternative, and the lectin, defined by the select initiating proteases that catalyze their respective pathway cascade [1]. Both the classical and lectin pathways possess distinctive pattern recognition molecules (PRM) that upon binding their specified target, enable autoactivation of the serine proteases housed in complex with the PRM [148]. These initiating proteases of the classical and lectin pathway, C1r and mannan-binding

lectin-associated serine protease (MASP)-1/2, respectively, possess the idiosyncratic ability to undergo autoactivation without the presence of extrinsic activating factors, which distinguish them and their molecular processes from other autoactivating enzymes [80,149].

MASP-2 is a critical complement component for the initiation of the lectin pathway and has been shown to undergo activation through proteolytic cleavage of its scissile loop at residues Arg^{444}-Ile^{445} by MASP-1 but also via autocatalytic activation demonstrated in *in vitro* studies by activated MASP-2 on its zymogen [150,151]. Overexpression and overactivation of MASP-2 have been implicated as driving and/or compounding factors in a number of inflammatory pathologies, such as venous thromboembolisms, IgA nephropathy (IgAN), and C3 glomerulopathy, among others [152,153]. At present, human monoclonal antibody, Narsoplimab, is pending-FDA approval for the treatment of hematopoietic stem cell transplant-associated thrombotic microangiopathy (HSCT-TMA) [154].

The full-length MASP-2 protein is comprised of six domains; however, it is the C-terminal serine protease domain that possesses the protein's catalytic activity [8]. Previous studies have employed molecular dynamics for modeling of the mechanism of MASP-2 autoactivation by utilizing individual MASP-2 enzyme (PDB entry 1Q3X [155]) and zymogen (PDB entry 1ZJK) crystal structures as structural input [80]. More recently, a 1.6 Å X-ray crystal structure containing MASP-2 forming an enzyme-product (EP) complex with its MASP-2 symmetry mate has been determined, in which the molecular interactions between the cleaved scissile loop of the product with key active site residues are herein described. Additionally, conformational flexibility of the catalytic Ser^{633} interacting with both catalytic His^{483} and the P1 Arg^{444} of the product activation peptide is shown, confirming previous hypotheses that in biologically derived complexes, the catalytic serine undergoes critical conformational changes to accommodate the substrate and

perform the catalytic functions within the protease [156]. Further analyses revealed that similar contacts are potentially conserved between EP interfaces in the autoactivating initiating protease of the classical pathway, C1r [157].

Materials & Methods

Expression and purification of MASP-2-CCP2-SP domain truncation

Recombinant MASP-2-CCP2-SP (rMASP-2) expression plasmid was transformed into BL21 (DE3) *Escherichia coli.* The cells were grown to an optical density of 0.6-0.8 OD_{600} at 37°C in Terrific Broth supplemented with kanamycin. After overnight induction with isopropyl β-D-thiogalactoside, the cells were collected by centrifugation (4,000 x *g*, 10 min). Supernatants were discarded and the inclusion bodies solubilized in 100 mL of lysis buffer (6 M guanidine HCl, 100 mM Tris pH 8.0, 10 mM imidazole) for 30 min and clarified by centrifugation (11,000 x *g*, 30 min). Nickel NTA beads (GoldBio) (5 mL column volume) were washed with 50 mL of denaturing binding buffer (8 M urea, 500 mM NaCl, 20 mM sodium phosphate pH 6.0, 10 mM imidazole). Samples were then passed over columns at a flow of 1-2 drops per second and afterwards washed with another 25 mL of denaturing binding buffer. Samples were eluted with 7.5 mL of denaturing elution buffer (8 M urea, 500 mM sodium chloride, 20 mM sodium phosphate pH 6.0, 200 mM imidazole) and then diluted into refold buffer containing 50 mM Tris pH 8.3, 3 mM reduced glutathione, 1 mM oxidized glutathione, 5 mM ethylenediaminetetraacetic acid (EDTA), 0.5 M arginine. The renatured protein solutions were dialyzed against 4 L of 50 mM Tris-HCl (pH 9.0), 140 mM sodium chloride twice for four hours at 25ºC and concentrated to less than 12 mL using 10 kDa molecular weight cutoff ultracentrifugation filters (Millipore). Refolded protein was then purified using an ÄKTA pure Fast Pressure Liquid Chromatograph (FPLC, GE Healthcare) connected to a HiLoad 26/600 Superdex 75 pg column previously equilibrated in 10 mM HEPES

(pH 7.3), 140 mM sodium chloride. To remove affinity tags, tobacco etch virus (TEV) enzyme (previously activated with 1 mM DTT) was incubated with pooled refolded protein overnight at 25 ºC. FPLC nickel affinity chromatography was carried out the following day and the unbound fraction was collected, concentrated by ultracentrifugation, buffer exchanged into 20 mM HEPES, 140 mM NaCl, aliquoted, and stored at -80 °C until use.

Protein crystallization

MASP-2-CCP2-SP was crystallized by hanging drop vapor diffusion technique. Crystals were obtained by mixing a 2 mL:2 mL protein to precipitant ratio of 3.2 mg ml^{-1} MASP-2 in 0.2M ammonium sulfate, 30% w/v polyethylene glycol 4000 sourced from the Hampton crystal screen condition. The mixture was equilibrated against a reservoir containing 500mL of precipitant at 25°C. Crystals were harvested and exchanged into crystallographic conditions supplemented with 20% glycerol for cryoprotection prior to vitrification in liquid nitrogen (**Fig. 8**).

FIGURE 8. MASP-2-CCP2-SP crystals in 0.2 M Ammonium sulfate, PEG 4,000, and 20% glycerol.

Structure determination and refinement

MASP-2-CCP2-SP X-ray diffraction data were collected on SER-CAT beamline 22-ID at the Advanced Photon Source, Argonne National Laboratory. Crystals occupied the C121 space group and contained one copy of the protein forming a complex its symmetry mate within the asymmetric unit. The resulting enzyme-product complex was solved by molecular replacement using Phaser in the PHENIX suite [158] using MASP-2-CCP2-SP PDB entry 1QX3 [155]. Model building was performed using AutoBuild and the structure refined using iterative cycles of refinement in Phenix to a resolution of 1.6 Å [158]. Manual model building was performed in

Coot [159]. Structure solution, properties, and refinement statistics are shown in Table 1. Structural representations and electron density maps were generated using PyMol.

Data collection and refinement	
Data collection	**MASP-2-2SP** (366-686)
Space group	C 1 2 1
Cell dimensions	
a, b, c, Å	87.85, 56.08, 86.83
α, β, γ, °	90.00, 118.95, 90.00
Resolution, Å	37.61-1.62 (1.68-1.62)
R_{meas}	0.097 (0.537)
R_{pim}	0.042 (0.244)
$CC_{1/2}$	0.996 (0.843)
I/σ|	12.52 (1.90)
Completeness, %	96.92
Redundancy	4.7 (4.1)
Refinement	
Resolution, Å	1.621(1.679-1.621)
No. reflections	45,569
R_{work}/R_{free}	0.1924, 0.2262
No. non-hydrogen atoms	2753
Protein	2483
Water	250
B-factors	32.12
Protein	31.39
Water	38.57
Rmsd	
Bond lengths, Å	0.007
Bond angles, °	0.94

Table 1. Structure properties, solution, and refinement. Values in parentheses are for the outer shell.

Interaction maps and buried surface area

The two-dimensional interaction map to visualize the binding interactions between the product activation peptide and residues of the active site and substrate-binding pocket of the enzyme MASP-2 was generated using LigPlot+ [160]. The residue interaction network was elaborated from the input of the PDB file containing MASP-2-CCP2-SP and residues Ser^{437}-Arg^{444} in the activation peptide present in product MASP-2 symmetry mate.

The enzyme-product interface details were generated from the buried surface area (BSA) calculations based on finite element analysis performed by PDBePISA v1.52 [161]. The structure analysis was performed from the input of the coordinate PDB file of the solved MASP-2-CCP2-SP crystal structure.

Results and discussion

The MASP-2 domain truncation mutant (MASP-2-CCP2-SP) containing the serine protease (SP) and complement control protein 2 (CCP2) domains was recombinantly expressed and purified to stably maintain the catalytic structure and function of the protein for enzymatic activity/inhibitory studies and X-ray crystallographic experiments for small-molecule inhibitor development. Size-exclusion chromatography, SDS-PAGE gels, and enzyme activity assays validated that the MASP-2-CCP2-SP mutant forms a monomer during purification and both refolded properly and retained proteolytic activity against native substates, complement components C4 and C2, and synthetic substate Z-Gly-Arg-S-Bzl.

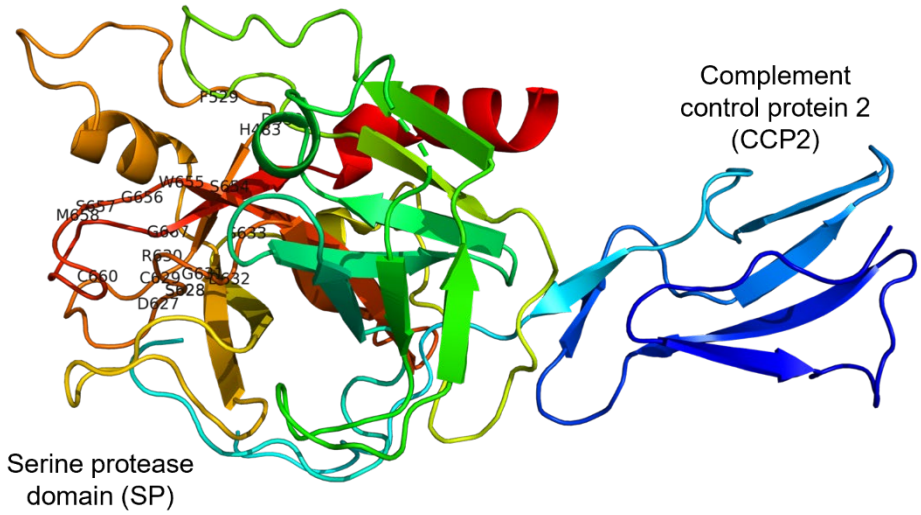


FIGURE 9. **Structure of crystallized MASP-2-CCP2-SP enzyme**. The C-terminal serine protease domain holds the protease's catalytic activity. Active site residues containing the catalytic residues His^{483}, Asp^{532}, and Ser^{633} are labeled.

The enzyme MASP-2-CCP2-SP crystal was solved by molecular replacement and refined to 1.6 Å resolution with a final R_{work} and R_{free} of 18.2% and 19.5%, respectively (**Fig. 9)**. These values represent measures for assessing model quality of the structure based on model-to-crystallographic data agreement, where a random set of values will yield an R-value of approximately 0.63, the perfect fit will have a value of 0, and most high-quality structures available will have values of approximately 0.20 [162]. MASP-2 is one of few enzymes capable of *in vitro* autoactivation without necessitation of extrinsic cleavage factors [80,151]. As such, during the autoactivation process, a temporary dimer is formed between the two catalytic serine protease domains of the active enzyme and the zymogen structure that results in the release of the newly activated zymogen-to-enzyme leaving product [80]. The autoactivation reaction of MASP-2 undergoes the same events that define the catalytic cycle of serine proteases, in which tetrahedral and acyl-enzyme intermediates are formed during proteolysis [84,163]. The crystal structure produced an enzyme-product complex with a MASP-2-CCP2-SP symmetry mate within the crystallographic asymmetric unit, allowing for a more detailed study of the binding interface between molecules. Within this complex, the cleaved scissile loop of the product MASP-2-CCP2-SP was bound within the active site of the enzymatic MASP-2-CCP2-SP. This captured binding event resembles the final steps in proteolysis characterized by serine proteases preceding the release of the product acyl-enzyme intermediate and subsequent active site regeneration [163], which is also representative of the MASP-2 autoactivation process (**Fig. 10**) [80].

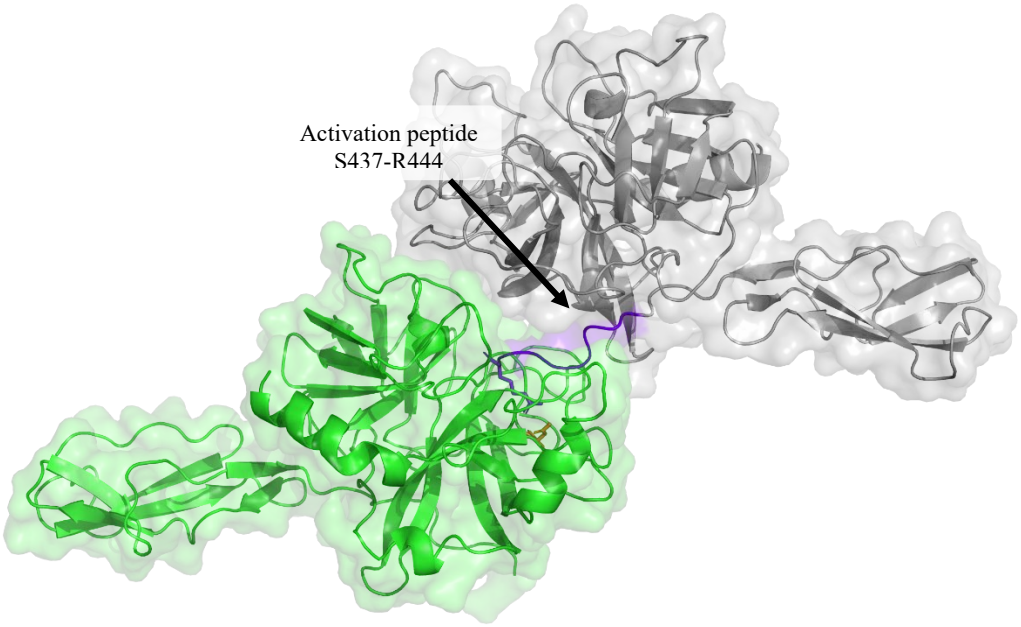


FIGURE 10. Enzyme-product complex formed between MASP-2 and symmetry mate.

Proteolytic cleavage at the P1 arginine, Arg^{444}, of the zymogen activation loop by the active enzyme molecule allows for subsequent conformational changes within the zymogen structure that results in formation of the S1 site and oxyanion hole allowing for activation of the protein and generation of the product molecule [163–165], as observed in the EP structure. Closer structural analysis of the interactions that form between the cleaved scissile loop and substrate-binding pocket generated using LigPlot+ [160] reveal that the P6-P1 amino acid residues, Arg^{439}-Arg^{444}, of the product molecule form several contacts via hydrogen bond formation within the substrate binding groove of the enzyme (**Fig. 11**). The charged, basic side chain of the P1 Arg^{444} forms a salt bridge and hydrogen bonds with specificity-determinant, S1 Asp^{627}, as well as hydrogen and hydrophobic bonds with loop 2 residues making up the S1 pocket (Gly^{656}, Ser^{657}, enzyme). The carboxy group of the P1 Arg^{444} also forms hydrogen and hydrophobic interactions with amino acid

residues comprising loop 1 (Ser^{628}, Cys^{629}, Arg^{630}, Gly^{631}, Ser^{633}, enzyme) and a hydrogen bond with His^{483} of the catalytic triad (enzyme). P2, P3, and P5 residues form hydrogen bonds with residues found within loop 2 and P4 Thr^{441} is hydrogen-bonded to Tyr^{607} of loop 1 (enzyme).

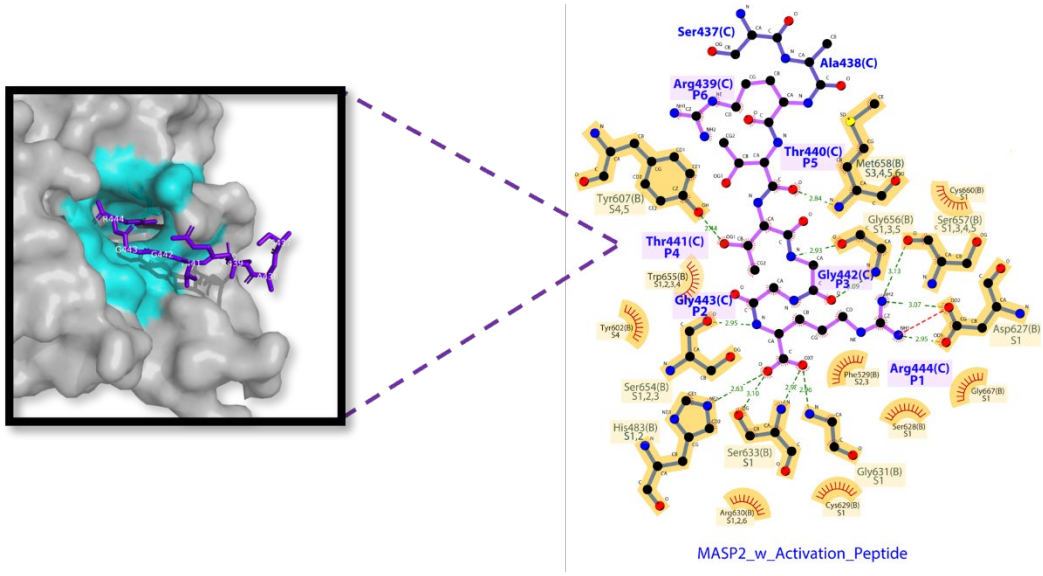


Figure 11. Two-dimensional interaction map of activation peptide with -nitrogen and -carboxylate residues of substrate binding pocket. P1-6 residues are represented as purple sticks, whereas S1-6 residues are shown as yellow sticks. Carbon (black), oxygen (red), and nitrogen (blue) atoms are represented as spheres. The green dashed lines labeled with distance (Å) represent hydrogen bonds. The red semicircles with bristles represent hydrophobic interactions. Figure generated using LigPlot. Inset shows surface representation of substrate-binding pocket in cyan with insertion of P1 Arg^{444} of activation peptide as purple stick representation.

Comprehensive studies of the catalytic triad of serine proteases, typified as containing an Asp-His-Ser within the active site, have shown significant and conserved structural rigidity of those residues as required for substrate proteolysis [166]. The catalytic mechanism entails the

nucleophilic attack by the hydroxyl group of the catalytic serine (S1) on the carbonyl carbon (P1) atom of the substrate peptide bond. This step results in the formation of a tetrahedral intermediate, which proceeds into an acyl-enzyme intermediate upon the termination of the carbon-nitrogen bond. The acyl-enzyme intermediate undergoes hydrolysis so that the free hydroxyl group of the catalytic serine is regenerated during deacylation, and the product released (**Fig. 12**) [83,165,166]. The crystallized EP complex captures the final stages of this process in which the product/activated MASP-2 is poised for release from the enzyme active site [165]. During this stepwise process of proteolysis or autocatalytic activation, the catalytic serine plays pivotal dual roles in not only enacting the nucleophilic attack but also in stabilizing the charge distribution of the catalytic triad required for the formation of the oxyanion hole, particularly the catalytic histidine that serves as both general acid and base in this reaction [166].

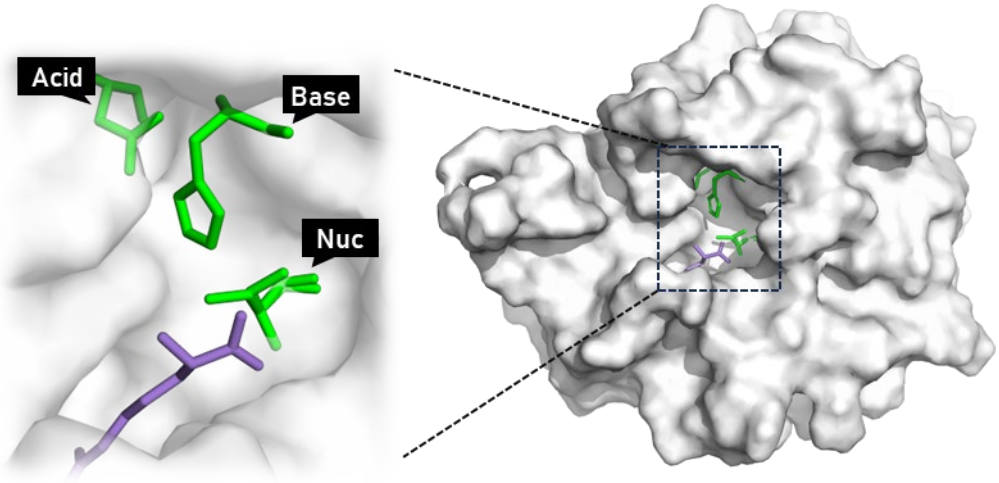


FIGURE 12. Residues that form the catalytic triad found in serine proteases.

To perform these dual functions, it has been proposed through previously described crystal structures of a number of serine proteases in both unliganded (bacterial α-lytic protease, PDB entry 1SSX; porcine elastase, PDB entry 1QNJ; *Fusarium oxysporum* trypsin, PDB entry 1XVO; *Streptomyces griseus* trypsin (SGT), PDB entries 1OS8 and 4M7G [167–170]) and liganded states (bovine trypsin, PDB entries 3UNR, 4I8H, 1BTZ, 2AGE [165,171,172]) and through analyses of the geometry of the catalytic residues [170], that while the catalytic histidine, aspartate, and waters near the catalytic histidine remain in conserved or homologous positions, the γ-oxygen of the catalytic serine can be observed in two conformations [170]. These serine conformers have been shown to be dependent upon the unliganded or deacylated state in which the serine is positioned to minimize the hydrogen bond distance of 2.8 Å between the serine γ-oxygen and the ε1-carbon of catalytic histidine in order to stabilize the charge distribution. This acts to further open the substrate binding fold to accommodate substrate binding and is named conformation A. The alternate conformation is shown to be reliant on the presence of substrate with which it will interact to generate an acylated state, named conformation B, wherein the catalytic serine is turned facing into the active site [170].

The high resolution of the MASP-2 EP structure allowed for visualization of the two conformers of the catalytic serine, wherein the major conformer of $Ser^{633}{}_{B}$, has an occupancy of 60%, is hydrogen bonded within 3.1 Å to the substrate Arg^{444} and is observed facing into the active site. The minor conformer of $Ser^{633}{}_{A}$ has an occupancy of 40% and is hydrogen-bonded within 3.5 Å to the catalytic His^{433} while facing away from the active site (**Fig. 13C**). Dual conformations of the catalytic serine in model serine proteases using synthetic ligands have been previously shown [170]; however, a physiologically relevant structure of enzyme-substrate or enzyme-product presenting two serine conformers as a static capture of autoactivation has not hitherto been shown.

The MASP-2-CCP2-SP structure, PDB entry 1Q3X, contains two copies of the enzyme in the asymmetric unit, however, electron density maps indicate but a single conformation of the Ser^{633} in the unliganded- or conformation A state (**Fig. 13A, B**).

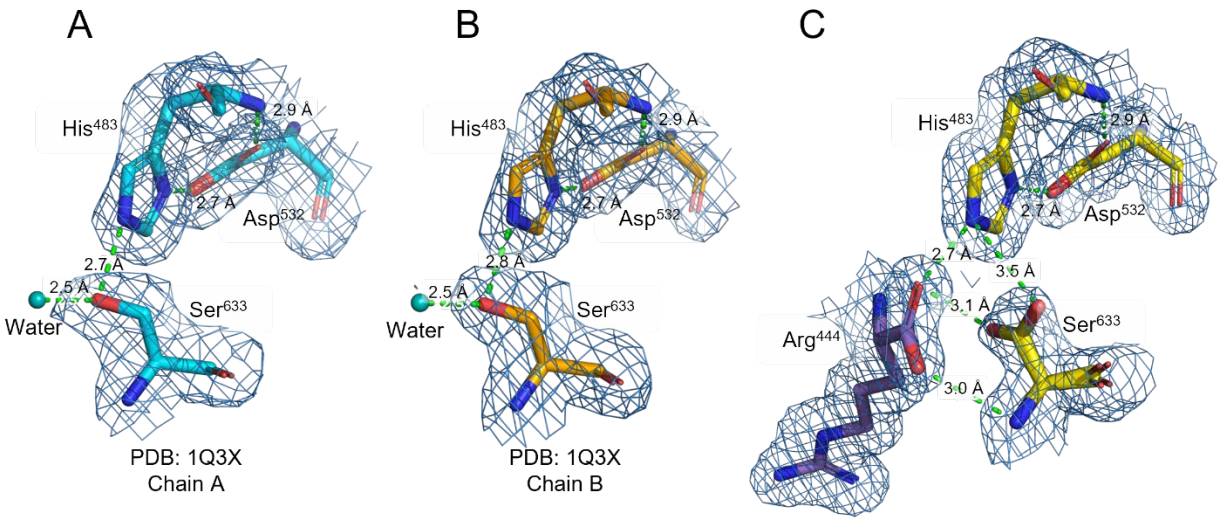


FIGURE 13. Electron-density representation of the MASP-2-CCP2-SP catalytic triad. A) 2*Fo* − *Fc* electron-density maps of residues Ser^{633}, Asp^{582}, His^{483} of PDB entry 1Q3X show a single catalytic serine conformer, Ser$^{633}{}_{A}$. The asymmetric unit possesses two copies of the enzyme, represented as chain A and chain B. The stick representation of the catalytic residues in chain A are shown in cyan. B) The catalytic triad of chain B is represented by orange sticks. C) The 1.6 Å MASP-2-CCP2-SP EP complex structure reveals high resolution electron density maps, demonstrating conformational flexibility of the catalytic serine, wherein Ser^{633} is present in both conformation A and B. The catalytic residues of the enzyme are shown as yellow sticks, and the P1 Arg^{444} shown as purple sticks. Hydrogen bonds are represented by green dashed lines.

Furthermore, earlier proposed mechanisms of autoactivation have used molecular modeling approaches to guide analyses of the binding interface that forms between the crystal

structures of MASP-2 enzyme and zymogen structure, forming the enzyme-substrate (ES) complex [80]. Analysis of the EP complex interface results in a buried surface area of 728.9 $Å^2$ resulting from 25 product and 24 enzyme residues involved in the formation of the interface, as determined by structural analysis of the BSA by PDBePISA [161]. The orientation of the two protein structures proposed for the ES complex differs from what is observed in the 1.6 Å crystal structure of MASP-2-CCP2-SP EP complex, revealing previously unknown solvent-exposed amino acids that form the basis of the contact residues of the enzyme-product interface (**Table 2**). Structural alignment of the EP complex formed by MASP-2-CCP2-SP with the C1r-CCP2-SP EP complex, PDB entry 1MD8, shows a conserved interface between the two serine proteases that provides further support for this mechanism of autoactivation (**Fig. 14**).

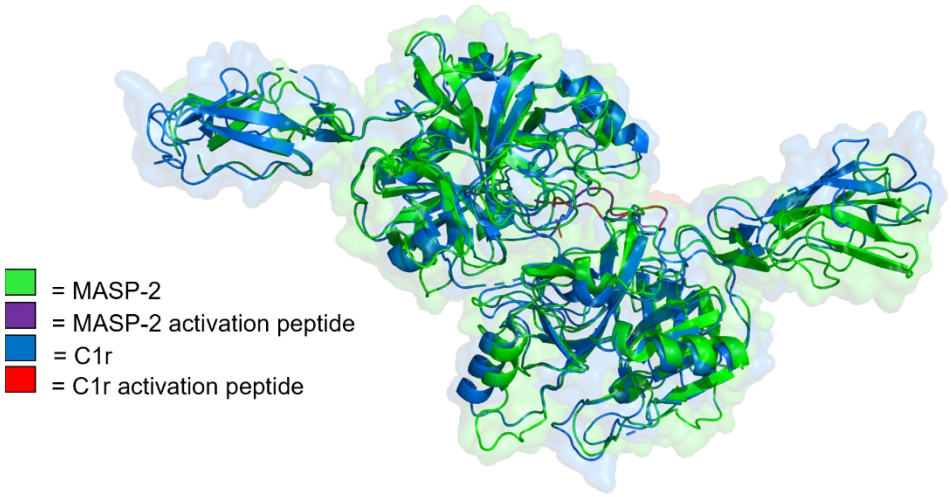


FIGURE 14. MASP-2-CCP2-SP enzyme-product structure aligned and superimposed on classical pathway initiating serine protease, C1r-CCP2-SP (PDB:1MD8), enzyme-product complex. The activation peptides of both protein products are positioned within the substrate-binding pocket of their respective enzymes. The structural features and enzyme-product interface of both MASP-2 and C1r are well conserved. MASP-2 is shown as a green ribbon structure and

its activation peptide in purple. C1r is represented by a blue-ribbon structure and its activation peptide shown in red.

CHAIN A	CHAIN B
464(GLY)	623(SER)
465(GLY)	622(GLU)
466(THR)	621(LEU)
467(THR)	591(PRO)
487(GLU)	567(ILE)
488(GLN)	589(ASP)
491(ASP)	626(LYS)
504(ARG)	565(ASP)
525(HIS)	624(GLY)
526(ASP)	625(GLY)
575(LEU)	449(GLN)
579(GLY)	447(GLY)
580(PHE)	566(ASP)
581(LEU)	642(SER)
630(ARG)	561(PHE)
658(MET)	560(SER)
659(ASN)	563(ARG)
662(GLU)	564(THR)
	647(TRP)
	640(LEU)
	569(THR)
	645(GLU)

TABLE 2. Contact residues present at the interface between the enzyme and product.

The MASP-2-CCP2-SP EP crystal structure revealed novel structural insights into the mechanisms of MASP-2 autoactivation through the *in vitro* formation and static capture of an EP complex. Detailed analyses into the binding interactions that occur between the cleaved scissile

loop of the product and the active site of the enzyme allow mapping of the key interactions that provide molecular clues for self-association in autoactivating enzymes. Additionally, high resolution electron density maps of the active site revealed dual conformers of the catalytic serine not previously captured by X-ray crystallography in biologically derived protein-protein complexes. Autoactivation modeled using molecular dynamics simulations have been formerly used to describe the MASP-2 ES complex interface as one mode of autoactivation [80]. The structure of the MASP-2 EP complex provides an additional model for which to describe complement autoactivation. Comparisons with structurally and functionally homologous complement serine protease C1r captured in an EP complex demonstrate a conserved interface present during autoactivation that aligns with our observations. The structural details herein derived from the MASP-2-CCP2-SP EP complex provide additional insights into the dynamic nature of complement proteases further providing a more in-depth understanding of the molecular underpinnings of serine protease autoactivation.

CHAPTER 3: Identification of Novel C1r-Binding Small-Molecule Inhibitors of the Classical Complement Pathway Using Fragment-Based Drug Discovery

Abstract: The initiating protease of the complement classical pathway, C1r, represents an upstream and pathway-specific intervention point for complement-related autoimmune and inflammatory diseases. Yet, C1r-targeted therapeutic development is currently underrepresented relative to other complement targets. In this study, we developed a fragment-based drug discovery approach using surface plasmon resonance (SPR) and molecular modeling to identify and characterize novel C1r-binding small-molecule fragments. SPR was used to screen a 2,000-compound fragment library for binding to human C1r. This led to the identification of 24 compounds that bound C1r with equilibrium dissociation constants ranging between 180 – 1,700 μM. Two fragments, termed CMP-1611 and CMP-1696, directly inhibited classical pathway-specific complement activation in a dose-dependent manner (**Fig. 15**). CMP-1611 was selective for classical pathway inhibition, while CMP-1696 also blocked the lectin pathway but not the alternative pathway. Direct binding experiments mapped the CMP-1696 binding site to the serine protease domain of C1r and molecular docking and molecular dynamics studies, combined with C1r autoactivation assays, suggest that CMP-1696 binds within the C1r active site. Additional structural modification to generate structurally similar analogs validated computationally derived structure-activity relationship data. The group of structurally distinct fragments identified here, along with the structure-activity relationship profiling of two lead fragments, form the basis for future development of novel high-affinity C1r-binding, classical pathway-specific, small-molecule complement inhibitors.

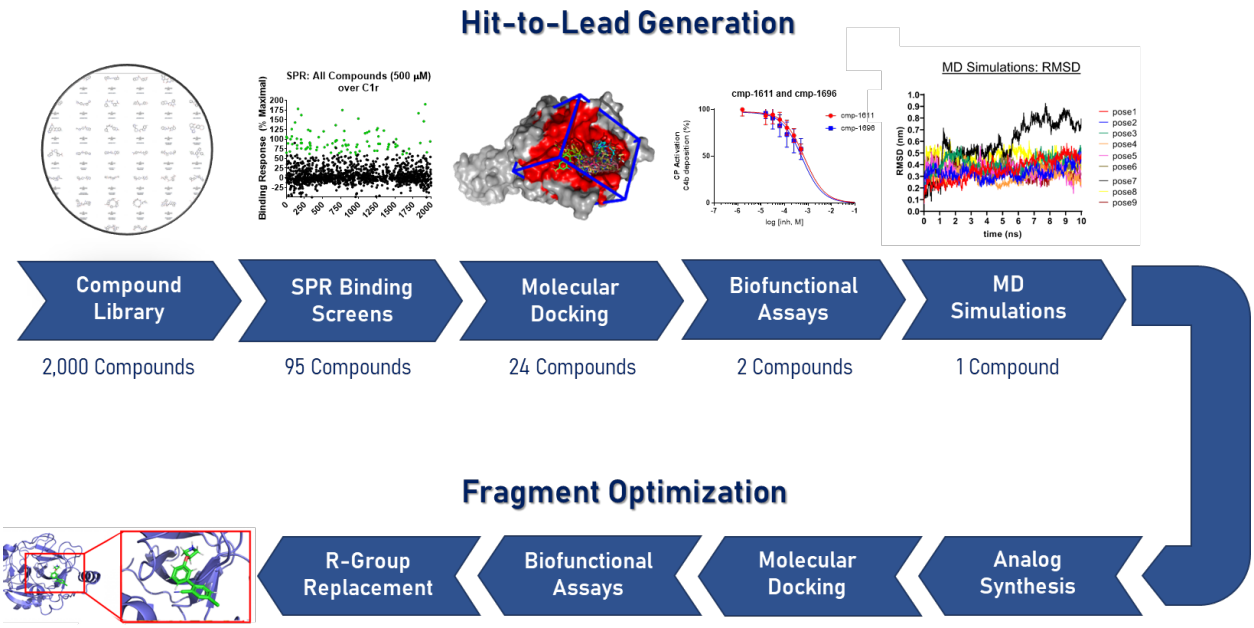


FIGURE 15. Schematic diagram showing hit-to-lead generation to identify CMP-1696. Beginning with an initial library of 2,000 compounds, SPR binding screens served as a platform for eliminating all but 95 compounds for additional evaluation. Using molecular docking techniques, 24 compounds were selected for characterization in a number of biofunctional assays, in which two were selected for having the most favorable bioactivity against target C1r. Molecular dynamics simulations elucidated the SAR data for single compound, CMP-1696, allowing for the synthesis and characterization of analogs.

Introduction

The developmental landscape of anti-complement therapies is now robust and includes many drugs in preclinical and clinical stages [55,87,173]. Two complement drugs, C1 esterase inhibitor (C1-INH) and the anti-C5 antibody drug eculizumab/Soliris, have reached the market, although C1-INH is FDA-approved for an indication not primarily mediated by complement (hereditary

angioedema (HAE)). Eculizumab/Soliris has now gained FDA-approval for use in in paroxysmal nocturnal hemoglobinuria (PNH), atypical hemolytic uremic syndrome (aHUS), and refractory generalized myasthenia gravis (gMG)) [87,173,174]. Importantly, both drugs are also being evaluated in clinical trials for other complement-related diseases [87,173,174]. Eculizumab and C1-INH are joined by dozens of new complement-directed drugs that include small molecules, antibodies, biologics, peptides, and nucleotide-based therapies for the treatment of numerous complement-related diseases, such as aHUS, complement 3 glomerulopathy (C3G), PNH, antibody-mediated rejection, IgA nephropathy, age-related macular degeneration (AMD) and many others [87,173,174].

Among the proteins that comprise the complement system are a small set of serine proteases with highly restricted substrate specificity. These include the initiating proteases of the alternative pathway (i.e., MASP-3/factor D), lectin pathway (i.e., MASP-1), and the classical pathway (i.e., C1r). Due to their far upstream position, initiating proteases represent a promising complement intervention point. Targeting complement at the level of MASP-1 or C1r provides a second advantage due to the potential of providing a pathway-specific blockade. While factor D-mediated formation of alternative pathway C3 convertases (i.e., C3bBb) amplifies all three pathways (Figure 1A), MASP-1 and C1r activities are specific for the lectin or classical pathways, respectively. Thus, selective inhibition of MASP-1 or C1r may be ideal in a disease setting where pathophysiological activation of complement is specific to the lectin pathway (LP) or classical pathway (CP), as it would theoretically leave two activation pathways available for immune surveillance. Furthermore, proteases are regarded as highly druggable targets with some estimates placing them as targets in 5-10% of all drug development [123]. Not surprisingly, proteases of the complement system have garnered attention as therapeutic targets [87,173–175]. However, to date

there has been no significant pharmaceutical development of inhibitors specifically directed towards MASP-1 or C1r. In this regard, C1r is particularly attractive as inappropriate classical pathway activation is implicated in an increasing number of diseases, including HIT [176], neuromyelitis optica [177], bullous pemphigoid [178,179], and autoimmune hemolytic anemias [180], among others. The classical pathway has also been shown to play causal roles in mouse models of Alzheimer's disease [181] and has been genetically linked to other neurological disorders, such as schizophrenia [182].

A common challenge for the development of drugs that target serine proteases like C1r, especially with small molecules, is addressing specificity. Interestingly, the role of C1r as the initiator protease of the classical pathway requires the molecular context of the C1 complex. Furthermore, its physiological autoactivation and subsequent catalytic activity on C1s proenzyme involves a precise set of coordinated intramolecular events within the C1 complex (i.e. $C1qC1r_{2}s_{2}$) [183–185]. For example, C1r is autoactivated only when C1q binds to immune complexes or activating non-antibody ligands, which is then followed by C1r cleavage of C1s into its active form. C1r is positioned in the C1 complex via intermolecular contacts with both C1q and C1s [186,187] and disruption or displacement of C1r from C1 has been described as a complement evasion strategy employed by human microbial pathogens [62]. Therefore, we hypothesized that along with orthosteric C1r-binding inhibitors, small molecules that target C1r at sites required for C1 complex stability could lead to the development of highly selective classical pathway inhibitors.

With these considerations in mind and with the long-term goal of producing a highly specific small-molecule C1 inhibitor, we report here the results of a C1r-targeted fragment-based drug discovery (FBDD) campaign (**Fig. 15**). FBDD employs compound screening libraries of very low

molecular weight small-molecule fragments ≤ 300 Da [115,188,189]. Fragment screening offers vastly increased hit rates and allows more efficient sampling of chemical space relative to traditional high throughput screening of large elaborated compound libraries [115,188,189]. While fragment-based hits are most often characterized by low-affinity binders with equilibrium dissociation constants (K_D) and half-maximal inhibitory concentrations (IC_{50}) ≥ 100 μM, their small starting size gives greater flexibility during compound optimization stages [115,188,189]. Our FBDD campaign ultimately yielded 24 small-molecule fragments that reversibly bind full-length human C1r with affinities that ranged from 180 μM to 1,700 μM. Four of these C1r-binding fragments had direct inhibitory activity in serum-based classical pathway activation assays. We selected two lead fragments, termed CMP-1611 and CMP-1696, and carried out detailed structure-function analysis which showed that these two compounds differ in their selectivity profile and possibly in their mechanisms of action. The discovery here of structurally distinct classes of C1r-binding fragments represents a significant step forward in the development of novel small-molecule inhibitors of the classical pathway.

Materials and Methods

Recombinant expression, purification, and refolding of C1r-domain truncations

Purified C1r and C1r proenzyme were purchased from Complement Technology. Recombinant C1r-domain truncations were produced by sub-cloning an *Escherichia coli* codon-optimized synthetic oligonucleotide (IDT Technologies) flanked with a 5' BamHI site, a 3' NotI site and a stop codon into the pT7HMT vector [190]. The C1r-CUB1 construct corresponds to amino acid residues 18-141, whereas the C1r-CCP2-SP construct corresponds to C1r residues 307-705 (UNIPROT numbering: #P00736). C1r-domain truncations were purified under denaturing

conditions using previously published protocols with some modifications [63,191]. Plasmids were transformed into *E. coli* BL21 (DE3) and cells were grown to an optical density of 0.6-0.8 OD600 at 37°C in Terrific Broth supplemented with kanamycin. Cultures were then induced overnight with isopropyl β-D-thiogalactoside and cells were collected by spinning for 10 min at 4000 x g. Supernatants were discarded and resuspended in 100 mL of lysis buffer (6 M guanidine HCl, 100 mM Tris pH 8.0, 10 mM imidazole) for 30 min and clarified by centrifugation for 30 min at 16000 x g. Nickel NTA beads (GoldBio) (5 mL column volume) were washed with 25 mL of denaturing binding buffer (8 M urea, 500 mM NaCl, 20 mM sodium phosphate pH 6.0, 10 mM imidazole). Samples were then passed over columns at a flow rate of 1-2 drops per second and afterwards washed with another 25 mL of denaturing binding buffer. Samples were eluted with 5 mL of denaturing elution buffer (8 M urea, 500 mM sodium chloride, 20 mM sodium phosphate pH 6.0, 200 mM imidazole) and then rapidly diluted 1:10 into refold buffer containing 50 mM Tris pH 8.3, 3 mM reduced glutathione, 1 mM oxidized glutathione, 5 mM ethylenediaminetetraacetic acid (EDTA), 500 mM arginine. The next day, samples were dialyzed twice for four hours at 25°C against 2 L of 50 mM Tris-HCl pH 7.4, 145 mM NaCl and concentrated to less than 12 mL using 10 kDa molecular weight cutoff ultracentrifugation filters (Millipore). Refolded protein was then purified using an ÄKTA pure Fast Pressure Liquid Chromatograph (FPLC, GE Healthcare) connected to a HiLoad 26/600 Superdex 75 pg column previously equilibrated in 200 mM sodium chloride and 20 mM Tris pH 8.0. For purification of C1r-CUB1, 5 mM $CaCl_2$ was added to all native buffers. Fractions were evaluated by SDS-PAGE and those containing the C1r truncation mutant were pooled. To remove affinity tags, tobacco etch virus (TEV) enzyme (previously activated with 1 mM DTT) was incubated with pooled refolded protein overnight at 25°C. FPLC nickel affinity chromatography was carried out the following day and the unbound fraction was

collected, concentrated by ultracentrifugation, buffer exchanged into 20 mM HEPES, 140 mM NaCl, aliquoted, and stored at -80°C until use. Sequence alignments were performed with EMBL-EBI Clustal Omega [192].

Compound library

Small-molecule compounds were obtained from ChemDiv Inc. All compounds were received in pre-weighed vials as powder, dissolved in dimethyl sulfoxide (DMSO) (GoldBio) to a final concentration of 10 mM and stored at -20 °C. Representations of compounds were prepared using ChemDraw Prime 19.1 (Perkin Elmer). Structural comparisons of compounds were carried out with the ChemMine software suite [193]. Hierarchical clustering using a single linkage was used to generate a distance matrix using Cluster, and atom pair Tanimoto coefficients [194,195] were calculated using the Similarity Workbench.

Surface plasmon resonance

All SPR experiments were carried out using a Biacore T200 (GE Healthcare) at 25°C. For all experiments, a running buffer (HBS-T) of 20 mM HEPES, 140 mM NaCl, 0.005% (v/v) Tween-20, 5% DMSO (v/v) was used. For each experiment, DMSO calibration curves ranging from 4 to 5.5% DMSO (v/v) were collected at the beginning, end, and every 50 cycles throughout the duration of the experiment. All experiments were performed using HC1500M sensor chips (Xantec). In total, 10 sensor chips and 16 ligand immobilized surfaces were created over the course of this study (detailed in Supplementary Table 1). In all cases, C1r proteins were immobilized onto chip surfaces using amine-coupling chemistry with 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide) (EDC), N-hydroxysuccinimide (NHS), followed by ethanolamine. Briefly, 100 mM of EDC and NHS were mixed and injected to activate the chip surface followed by the C1r protein

ligand captured at a defined resonance unit (RU) level, followed by reaction quenching with 100 mM ethanolamine pH 8.5. Flow cell 1 was always used as a reference cell (no ligand), unless otherwise noted. All sensorgrams were reference and blank subtracted and analyzed using T200 Evaluation Software (GE Healthcare).

Initial SPR screening

Using 96-well plates, 5 µL of each of the 2,000 fragment compounds were diluted with 95 µL of 1.05x HBS-T. This provided individual 500 µM solutions of each compound with a final DMSO concentration of 5% (v/v) to match the running buffer. All compounds were first visually inspected for solubility in running buffer, and insoluble compounds were not included in the subsequent "clean screen." To determine if compounds nonspecifically interacted with the chip surface, a "clean screen" was performed using a single flowcell on an uncoupled HC1500M. Compounds which exhibited > 5.0 RU of residual binding at 10 s post injection were removed from the subsequent screen. Approximately 1,600 compounds passed this selection criteria. These compounds were then screened at 500 µM for binding to full-length C1r enzyme immobilized at high density. Baseline noise was accounted for using T200 evaluation software. A theoretical maximal binding signal (R_{max}) was calculated using the equation R_{max} = (C1r immobilization level (RU) x (mol. wt. compound / mol. wt. C1r) * n), where the C1r immobilization for each surface used, mol. wt. C1r = 92,000 Da, and n is the binding stoichiometry, assumed here to be 1. Compounds were considered as hits if all the following criteria were met: i) injections fell within the DMSO calibration curve, ii) did not exhibit > 5.0 RU of residual binding to the reference surface, iii) did not exhibit abnormal sensorgram shape, and iv) did not exhibit superstoichiometric binding [196].

Evaluation of dose-dependent binding by SPR

To further evaluate hit compounds from the initial SPR screen, dose-dependent SPR binding assays were carried out. Three replicate flow cells were used containing both low and high immobilization densities of C1r. Hit compounds were tested for dose-dependent binding to full-length C1r using a compound concentration range of 7.8 to 500 μM. T200 evaluation software was used to calculate steady-state affinities for each compound by fitting sensorgrams from each variable concentration injection dataset using a 1:1 Langmuir model of interaction constrained by a theoretical experimental binding response based on the mol. wt. of each compound and derived as stated above. K_D values are reported as the mean ± the standard deviation from at least three independent measurements.

Complement Inhibition Assay

To carry out classical pathway ELISA-based inhibition assays, IgM was dissolved in coating buffer (100 mM Na_2CO_3/$NaHCO_3$, pH 9.6) to a final concentration of 3 μg/mL and was dispensed into 96-well plates at 100 μL per well. Mannan was dissolved to 2 μg/mL and lipopolysaccharides (LPS) dissolved to 10 μg/mL in coating buffer for use in lectin and alternative pathway inhibition assays, respectively. Immobilization was allowed to proceed overnight at 25°C. The immobilization reagent was then removed, followed by three washes in washing buffer (50 mM Tris pH 8.0, 200 mM NaCl, 0.05% (v/v) Tween-20), and blocked using 100 μL of 1% (w/v) bovine serum albumin in PBS-T for 1 hour at 37°C. Blocking buffer was removed and plates were washed three times in washing buffer. Next, 500 μM of each compound, 5% (v/v) DMSO, and 1% (v/v) normal human serum (Innovative Research) (final concentrations) were dissolved in classical/lectin pathway ELISA buffer (20 mM HEPES (pH 7.3), 0.1% (w/v) gelatin, 140 mM

NaCl, 2 mM $CaCl_2$, 0.5 mM $MgCl_2$) and added to the IgM-coated plate for 1 hour at 37°C. For the lectin pathway-specific ELISAs shown in Figure 17B, we used C1q-depleted serum from CompTech at 2% (v/v) and thus the classical pathway assays for this analysis were also performed using serum from CompTech at 2% (v/v). For the alternative pathway, 20% (v/v) normal human serum (CompTech) (final concentration) and alternative pathway ELISA buffer (20mM HEPES, (pH 7.5), 0.1% (w/v) gelatin, 140 mM NaCl, 2 mM $CaCl_2$, 0.5 mM $MgCl_2$) were used in place of classical/lectin pathway ELISA buffer. Plates were then washed three times in wash buffer and each well was filled with 100 µL of α-C4 (HYB 162-02, Santa Cruz Biotechnology) diluted 1:300 for classical and lectin pathway assays and 100 µL of C3b (WM-1, Sigma) diluted 1:300 for alternative pathway assays. Plates were again incubated at 37°C for 1 hour, washed three times, filled with 100 µL of HRP-conjugated goat-α-mouse secondary antibody (diluted 1:3000) (Thermo Scientific), and rocked gently for 1 hour at 25°C. After three more washes, 50 µl of substrate (1-step Ultra TMB, Thermo Scientific) was added to each well and rocked for 15 minutes in the dark at 25°C. The reaction was stopped with 50 µL of 0.16 M sulfuric acid, and the plate was read at 450 nm using an EnSight multimode plate reader (Perkin-Elmer). Each column contained a positive control for full complement activity (5% (v/v) blank DMSO with 1% (v/v) serum) and a negative control for no complement activity (5% (v/v) blank DMSO and no serum). Positive controls (1% (v/v) serum, no inhibitor) were defined as 100% C4b or C3b signal whereas negative controls (no serum) were defined as 0% C4b or C3b signal for each column. Compounds that inhibited complement deposition relative to the control were then evaluated using the same assay setup across a two-fold variable concentration series for each compound (7.8 to 500 µM). These data were used to obtain a half maximal inhibitory concentration (IC_{50}) by fitting dose-response

curves using inhibitor vs. response models in GraphPad Prism 8. All experiments were performed no less than three times.

X-ray crystallography

Using a 3K MWCO spin filter, C1r-CCP2-SP was spun at 13,000 x g for 5 minutes to concentrate the sample volume down to 100 – 200 μL and a final concentration of at least 3 mg/mL, after correcting for the extinction coefficient. The protein solution was pipetted in a crystal sample tube and with the lead compound and dispensed onto the tray at a 1:1 mix of protein to mother liquor in the top well and a mix of 2:1 in the bottom well. Crystals were grown in 0.2 M magnesium chloride hexahydrate, 0.1 M Tris hydrochloride, pH 8.5, 30% w/v polyethylene glycol 4,000. Crystals were harvested and cryoprotected with supplementation of 5% glycerol to the precipitant solution. Crystals in this condition grew in the space group P $2\ 2_1\ 2_1$ with one copy of the C1r-CCP2-SP molecule in the asymmetric unit; however, no inhibitor was present (**Fig. 16**).

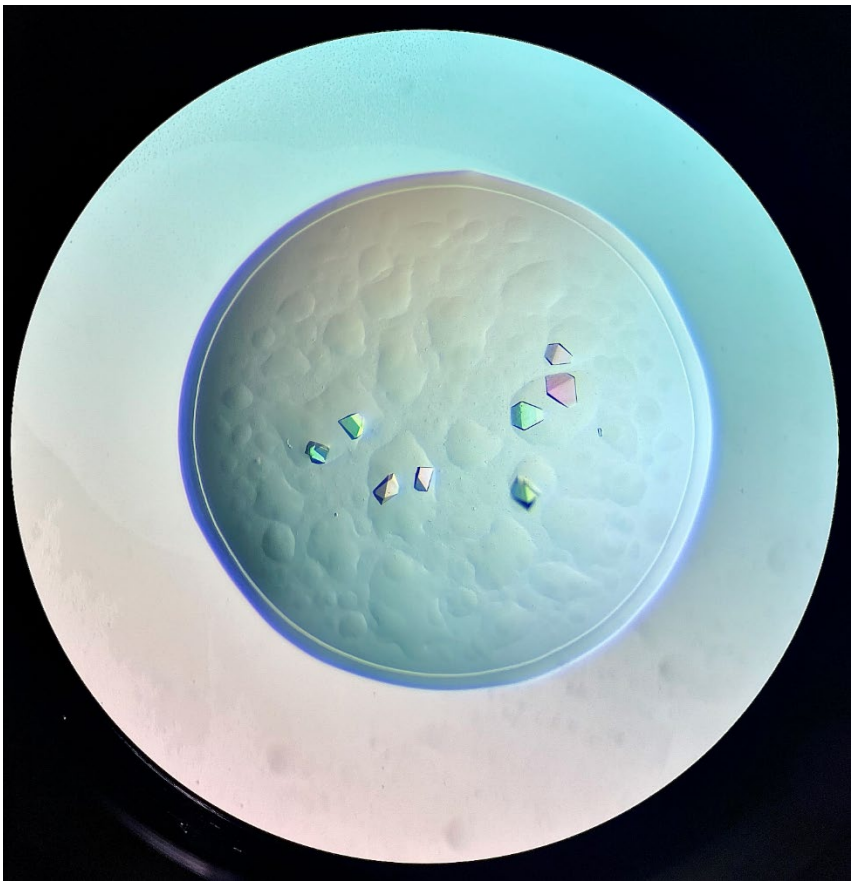


FIGURE 16. C1r-CCP2-SP crystals grown in conditions from the Hampton crystal screen.

Molecular Docking

Compound structures in 2D SDF formatted images were converted to 3D MOL2 formatted files using OpenBabel v 2.4.1 [197]. MOL2 formatted files were then converted to PDBQT file format using the prepare_ligand.py script from AutoDockTools1.5.6 [198]. Protein structure files for the N-terminal portion of C1r (CUB1-EGF-CUB2; PDB: 6F39) and the C-terminal region of C1r (CCP1-CCP2-SP; PDB: 1GPZ) were prepared for docking by removing water molecules and adding hydrogens using AutoDockTools1.5.6 before generating protein structure PDBQT files. Binding boxes for molecular docking were designed to encompass the entirety of the protein structures. Coordinates for the center (x,y,z) and size (x,y,z in Å) of target boxes are; N-terminal

region of C1r: center (135, 100, 40), size (82, 62, 92), and C-terminal region of C1r: (55, 10, 33), size (98, 70, 66). Molecular docking was then performed using AutoDock Vina 1.1.2 [46] and as detailed previously [199]. Nine poses for each compound were collected and included in the full dataset. Molecular docking of CMP-1696 onto the C1r-CCP2-SP crystal structure was performed using PDB:1MD8. A grid box encompassing only the SP domains of the receptor were chosen with center (26, -2, 12) and a box length of (40, 44, 38). Molecular docking was then performed with Vina executable embedded with PyRx v0.9.8 [200] resulting in 9 binding poses for each compound. All representations of protein structures were prepared using PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC (www.pymol.org/).

C1r proenzyme activation assay

Individual stock solutions of 1 mM of Z-Gly-Arg-sBzl (MP Biomedicals) and 5,5′-Dithiobis(2-nitrobenzoic acid) (DTNB) (Sigma-Aldrich) were made by dissolving powder in DMSO to give 10 mM stocks of each compound, which were then diluted to 1 mM in assay buffer (20 mM HEPES pH 7.3, 140 mM NaCl, 0.005% (v/v) Tween-20, 2 mM $CaCl_2$). In 96-well plates, 5 µL of 100 mM compound stocks (in DMSO) were added into wells in triplicate along with 25 µL of 8.1 nM proenzyme C1r (stored over ice until use). To these wells, 5 µL of 1 mM DTNB and 5 µL of 1 mM Z-Gly-Arg-sBzl were added simultaneously. Another 10 µL of assay buffer were added to each well to give a total reaction volume of 50 µL. Reactions were carried out at 37°C. Plates were read at 450 nm using an EnSight multimode plate reader (Perkin-Elmer) and the reaction progress was tracked for 1 hour.

Molecular Dynamics

Molecular dynamics simulation of C1r-CCP2-SP/CMP-1696 binding were performed with Gromacs 2019.3 [201]. For C1r simulation, the high-resolution crystal structure with PDB code:1MD8 was used and missing residues were built using Coot v0.8.9.2 [202]. Simulations of the nine top poses from the CMP-1696 docking onto C1r-CCP2-SP were performed at 300K for a duration of 10 ns, while pose 1 was eventually carried out to 50 ns. The topology files for the receptors were prepared with Gromos 43a1 united atom force field, while that of compound CMP-1696 was built using the PRODRG server [203]. The protein-ligand complex system was then solvated in a dodecahedron box with dimensions of 49.9 Å, 46.6 Å, 76.4 Å with TIP3P model waters (26,734 water molecules) using periodic boundary conditions of 10 Å from the edge of the solvation box. The system was charge neutralized by replacing eight Na+ ions in place of TIP3P water molecules. Bond lengths were constrained by the LINCS algorithm [204] and all long-range electrostatics were determined using the smooth Particle Mesh Ewald (PME) method [205]. Energy minimization was performed with the steepest descent algorithm until convergence (~1000 steps). Temperature equilibration was conducted by the isochoric-isothermal NVT ensemble (constant Number of particles, Volume, Temperature) with a Berendsen thermostat [206] for 100 ps. The system was then subjected to pressure equilibration in the NPT ensemble (constant Number of particles, Pressure, and Temperature) for 100 ps using the Parrinello-Rahman protocol [207], maintaining a pressure of 1 bar. The top-pose of CMP-1696, which also turned out to yield the highest binding energy through MM/PBSA [208] energy calculations, was chosen for a production MD run of 50 ns with snapshots being saved at 2 ps intervals. Backbone RMSD, intra-protein hydrogen bonds and trajectory analyses were performed with GROMACS programs ‘gmx rms,’ ‘gmx hbond,’ and ‘gmx trjconv.’

Binding energy calculations for the receptor-compound complex were carried out with g_mmpbsa program [208]. Initially, for all the docked poses, the entire simulation trajectories were subjected to these calculations with an interval of 100 ps, amounting to 100 snapshots. For the top binding pose within the 50 ns MD simulation, trajectories were sampled at 100 ps, amounting to 500 snapshots. All components of binding energy categorized into non-polar, polar and solvent-accessible surface area were calculated using standard protocols employed by the g_mmpbsa program. Error estimates for MM/PBSA calculations were performed via bootstrapping methods for 2000 steps according to guidelines outlined in g_mmpbsa program. All subsequent data analyses were performed using in-house codes written in Fortran 90, Python and C-shell scripts.

Statistics

Statistical analyses were performed using GraphPad Prism version 8. For complement inhibition assays, measures of statistical significance for single-dose experiments were assessed using unpaired Student's t tests compared to no inhibitor controls (pathway ELISAs) or non-binding control compounds (classical pathway ELISA). Statistical significance was defined as $p < 0.05$. Calculation of IC_{50} values from dose-dependent complement assays were obtained by non-linear regression analysis using a using an inhibitor vs. response model and with 95% confidence intervals reported. For non-linear regression analyses, the top and bottom of each curve were constrained to 100 and 0, respectively.

Ring replacement strategy

Ring Replacement Recommender v2002.02 was utilized to identify ring analogs derived from the compound structure of CMP-1696. The oxadiazole group, computationally shown to interact transiently with C1r residues, was selectively swapped with aromatic ring structures with

improved biological activity based on information mined from scientific literature. The average improved bioactivity improved by replacement in log10 scale is reported along with the number of literature references from which this data sourced. Optimized compounds were visualized in ChemDraw 20.1.1 and SMILES generated from the new chemical structures. Those SMILES files generated by Ring Replacement were used to search and select for structurally similar compounds using hierarchical clustering from atom pair Tanimoto coefficients that were commercially available through ChemDiv.

Results

Small-molecule fragment library design

This study began with the design and acquisition of a custom-selected, commercially synthesized 2,000-compound small-molecule library (ChemDiv Inc.). The library was composed of five subsets: i) 250 "2-dimensional fragments" (2D-FL), ii) 250 "3-dimensional fragments (3D-FL), iii) 250 "serine-protease inhibitor" compounds (SPI), iv) 250 "natural product scaffolds" (NPB), and v) 1,000 "protein-protein interaction inhibitor" compounds (PPI). Compounds originating from the 2D-FL subset represent traditional small-molecule fragments ranging in size between 99 – 330 Da, while the 3D-FL library were fragment-sized compounds (109–372 Da) with increased three-dimensional character, as judged by increased sp^3 hybridized carbons [209]. The SPI library consisted of small molecules and fragments (164–575 Da) with scaffold similarity to known protease inhibitors, the NPB library (221–568 Da) included compounds with scaffolds that share similarity to natural product-derived compounds, and the PPI library (138–580 Da) included compounds inspired by the chemical features of known protein-protein interaction

inhibitors. Collectively, these sub-libraries encompassed diverse chemical scaffolds, functional groups, stereochemistry, conformers, and substituent moieties.

Initial library screening of C1r-binding by surface plasmon resonance.

To identify compounds with non-specific binding behavior in our SPR screening platform and/or low solubility in aqueous SPR buffers, each compound was initially subjected to a "clean screen." Compounds were first diluted to a final concentration of 500 μM in SPR running buffer. Compounds that were visibly insoluble at this concentration were eliminated from further testing. The remaining compounds were then tested for non-specific binding to a blank flowcell on an SPR sensor chip. Compounds were injected at 500 μM for 30 seconds and those that exhibited > 5.0 RU residual binding signal at 10 seconds post injection were considered as non-specific binders and were eliminated from further testing. The remaining 1,619 compounds that passed the clean screen were carried forward to direct C1r-binding assays. Each compound was injected at a concentration of 500 μM over immobilized full-length human C1r. Injections that fell outside of the DMSO solvent correction curve and those that exhibited superstoichiometric binding [114] were eliminated from further consideration. In total, 95 compounds of the original 2,000 (4.8%) exhibited ≥ 60% of the theoretical maximal binding response.

C1r-binding properties of hit compounds

We obtained 65 hit compounds in larger quantities for further characterization. Of these, 24 compounds bound to C1r dose-dependently and fit well to a 1:1 steady-state binding model. Steady-state affinities for these compounds ranged from (K_D = 180 - 1,700 μM) . Clustering analysis revealed that nearly all 24 compounds were structurally distinct from one another, which was expected due to the design criterion of chemical diversity for the fragment library (see above).

However, we noted very high similarity between CMP-24 and CMP-202 as judged by an atom pair Tanimoto coefficient of 0.83 [194,195]. Likewise, compounds CMP-761 vs. CMP-981 (Tanimoto coefficient = 0.64) and CMP-618 vs. CMP-638 (Tanimoto coefficient = 0.50) share a high level of similarity.

To better understand where each fragment binds on C1r, we used molecular docking. C1r is a 92 kDa modular serine protease composed of six sequentially arranged domains named Complement C1r/C1s, Uegf, Bmp1 (CUB), Epidermal Growth Factor-like (EGF), Complement Control Protein (CCP), and Serine Protease (SP). While a crystal structure of full-length C1r (i.e. CUB1-EGF-CUB2-CCP1-CCP2-SP) has not been solved, atomic resolution structures of several C1r-domain truncations have been reported, including the N-terminal domains (CUB1-EGF-CUB2; PDB: 6F39) and the remaining C-terminal domains (CCP1-CCP2-SP; PDB: 1GPZ) [78,164,191,210,211].

To limit the conformational search space and to restrict the docking to experimentally derived structures rather than a model of C1r, we carried out two independent *in silico* experiments. To this end, each of the 24 lead compounds were docked onto the available structures of the N-terminal half of C1r (PDB:6F39) and separately to the C-terminal half of C1r (PDB:1GPZ). In the N-terminal docking experiment, each compound bound to one of two pockets either on the CUB1 domain or at the interface between CUB1-EGF. In the C-terminal docking experiment, four potential binding pockets were identified across all lead compounds, all of which were found on the C1r-SP domain. Most of the compounds bound to the S1 subsite near the C1r catalytic site. In general, the respective N- and C-terminal binding pockets for the top scored poses of each compound were similarly favorable, as judged by the calculated docking binding energies. Notable exceptions included CMP-4, CMP-59, and CMP-981, which have more favorable predicted

binding energies of ≤ -2.8 kcal/mol for the N-terminal site compared to the corresponding C-terminal site.

The identification of two structurally distinct C1r-binding and complement inhibitory lead fragments

To determine if our lead fragments had direct inhibitory activity against the classical pathway, each compound was tested in a serum-based in vitro assay of complement function using conditions specific for classical pathway activation. Compared to a non-C1r-binding control compound, CMP-4, CMP-778, CMP-1611, and CMP-1696 exhibited statistically significant inhibitory activity when used at a single 500 μM final concentration. Two of these compounds, CMP-1611 and CMP-1696, behaved dose-dependently and exhibited half-maximal inhibitory concentrations of 480 and 670 μM, respectively (**Fig. 17A**). Compounds CMP-1611 and CMP-1696 are structurally distinct from one another (Tanimoto coefficient = 0.09). CMP-1611 is composed of a central pyrimidine group linked to aminopiperdine and pyrrolidine rings, whereas CMP-1696 is composed of a central phenol group linked to an oxadiazole group on one side and an azetidine on the other side. Each compound is best classified as a small-molecule fragment and each exhibited favorable "rule-of-three" compliant physicochemical properties (i.e., ≤ 300 Da, ≤ 3 hydrogen bond donors, and ≤ 3 clogP) [189].

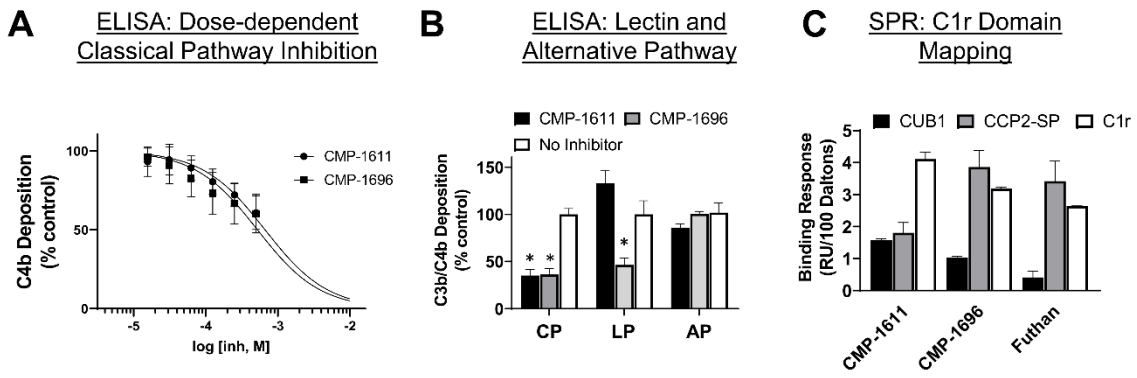


FIGURE 17. Selectivity and mechanistic analysis of CMP-1611 and CMP-1696. (A) Dose-dependent inhibition by CMP-1611 and CMP-1696 in a classical pathway-specific ELISA. Data were fit with GraphPad Prism using non-linear regression with a log(inhibitor) vs. response model. For CMP-1611, an IC_{50} value of 660 µM with an associated 95% confidence interval of (560 – 790 µM, n=9). For CMP-1696, an IC_{50} value of 520 µM with an associated 95% confidence interval of (410 – 680 µM, n=7) was calculated. **(B)** Complement pathway selectivity of CMP-1611 and CMP-1696. Compounds were assessed for their ability to inhibit activation of complement via the lectin and alternative pathways using single doses of 500 µM compound in triplicate. To ensure only lectin pathway activation, 2% (v/v) C1q-depleted serum (CompTech) was used and mannan was used as the activator. To match serum sources and amounts for this assay, the clasical pathway assays were repeated here using serum from CompTech at 2% (v/v) final concentration. Alternative pathway activation assays were performed using 20% (v/v) serum (CompTech), alternative pathway buffers, and C3b detection (see Methods and Materials for details). CMP-1611 had no effect on the lectin or alternative pathway, whereas CMP-1696 blocked lectin but not alternative pathway activation. **(C)** C1r, C1r-CUB1 and C1r-CCP2-SP, were immobilized on an SPR sensor chip and binding responses for 500 µM CMP-1611 and CMP-1696

or 10 μM futhan were each injected in duplicate over all surfaces. Binding responses were corrected for the molecular weight of each analyte and the immobilization level and molecular weight of each surface ligand. Measures of statistical significance in (B) were obtained by comparison of vehicle control using an unpaired t-test (*$p < 0.05$).

To assess selectivity of CMP-1611 and CMP-1696, we performed pathway-specific ELISA-based inhibition assays. Neither CMP-1611 nor CMP-1696 inhibited complement under conditions that selectively activate the alternative pathway, suggesting a level of specificity for the classical pathway that would be expected based on the C1r-binding properties of each compound (**Fig. 17B**). However, CMP-1696, but not CMP-1611, blocked lectin pathway activation at a similar level as it did the classical pathway (**Fig. 17B**). To further investigate the divergent selectivity profiles of CMP-1611 and CMP-1696, we first sought to clarify the binding site on C1r for each compound. Both compounds were predicted to bind the same pocket on the CUB1 domain in the N-terminal C1r molecular docking experiment. Likewise, the C-terminal docking experiment predicts that CMP-1611 and CMP-1696 bind within the S1 subsite of the C1r-SP domain. To test these predictions, we produced recombinant C1r-domain truncations that included the predicted N-terminal binding site (C1r-CUB1 domain) or the C-terminal site (C1r-CCP2-SP). On the same biosensor chip we immobilized C1r-CUB1, C1r-CCP2-SP, and full-length C1r on three different flowcells. We then injected each compound simultaneously over the three surfaces at 500 μM and measured binding responses. As a control, we also injected 10 μM Nafamostat/FUT-175/futhan, which is a promiscuous serine protease inhibitor previously reported to inhibit C1r and which is proposed to bind the active site of its serine protease targets [212]. Indeed, comparison of the relative binding responses showed that futhan bound C1r and C1r-CCP2-SP similarly, whereas very little binding response was seen for the C1r-CUB1 domain (**Fig.**

17C). The domain mapping profile for CMP-1696 was similar to that of futhan, with CMP-1696 exhibiting high binding to C1r and C1r-CCP2-SP but low binding to CUB1 (**Fig. 17C**). In contrast, CMP-1611 bound weakly to both C1r-CUB1 and C1r-CCP2-SP relative to full-length C1r (**Fig. 17C**).

C1r-binding mode of CMP-1696

Collectively, the data above showed that CMP-1611 is selective for classical pathway inhibition but that its interaction site on C1r and mechanism of action remain ambiguous. By contrast, an analogous set of experiments for CMP-1696 is consistent with this compound binding to the S1 subsite on C1r and potentially to serine proteases of the lectin pathway. In an effort to support future development of CMP-1696-based compounds into larger drug-like molecules that selectively inhibit C1r, we carried out a series of molecular docking and molecular dynamics (MD) simulations. Given the experimental data obtained showing that CMP-1696 binds directly to C1r-CCP2-SP (**Fig. 17C**), we redocked CMP-1696 onto the available crystal structure of this region of C1r (PDB:1MD8) (**Fig. 18A**). The top nine poses for CMP-1696 bound within the S1 pocket and have similar calculated docking energies. Consistent with an S1 binding site on C1r, 10 mM CMP-1696 directly inhibited the autocatalysis of C1r proenzyme (**Fig. 18B**).

Identification of the correct binding pose resulting from a molecular docking experiment may be improved by rescoring the ensemble of poses using a combination of MD simulations and molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) free binding energy calculations [213,214]. To further investigate the true CMP-1696 binding mode, we carried out short 10 ns MD simulations for each of the top nine scored docking poses (**Fig. 18C, D**) and used MM/PBSA to calculate free binding energies. Aside from pose 7, all CMP-1696 docking poses

are predicted to be energetically favorable with pose 1, ultimately agreeing with the docking scoring as the most favorable (**Fig. 18D**). This binding pose, shown in **Fig. 18E**, forms hydrogen bonds between the CMP-1696 azetidine group and sidechain atoms of Asp^{648}. The Asp^{648} residue in C1r is functionally important as it forms the bottom of the S1 pocket in C1r [78]. The CMP-1696 oxadiazole group forms several hydrogen bonds including with the catalytic serine (Ser^{654}) and catalytic histidine (His^{502}) as well as with sidechain atoms of Gln^{651} and Thr^{503} (**Fig. 18E**). The interactions with Gln^{651} and Thr^{503} are notable as these residues are not conserved in either C1s or MASP-2 . To explore these differences further, we docked CMP-1696 onto the crystal structures of C1s-CCP2-SP and MASP-2-CCP2-SP. Analysis of the top scored binding pose predicts that CMP-1696 binds MASP-2 in a nearly identical conformation as to that observed in C1r-CCP2-SP involving six homologous contact residues (four identical). In contrast, the docking of CMP-1696 onto C1s indicates an alternative binding mode near, but not in, the S1 subsite and with only a single contact residue (Lys^{629}) in a homologous position to C1r.

To further analyze the binding mode of CMP-1696, we carried out a longer 50 ns simulation using docking pose 1. MM/PBSA analysis of this simulation further highlighted the importance of the polar interactions mentioned above in the CMP-1696/C1r interaction (**Fig. 18F**). A more detailed analysis of the simulation reveals a number of interesting features that may help guide the development of CMP-1696 for improved C1r affinity and specificity. From the ligand side of the interaction, we noted that all atoms of CMP-1696 interact with one or more protein atoms in ≥ 50% of the simulation with the notable exceptions of the oxygen (atom label: OAM) and terminal carbon (atom label: CAN) atoms on the 5-methoxymethyl subsituent of the oxadizaole group. These atoms form transient contacts across 14 and 21 C1r residues respectively. Moreover, several transient interactions were identified near this moeity that were not present in

the original docking pose and include Ile^{483}, His^{484}, Gly^{485}, Gly^{487}, Ala^{500}, Pro^{506}, Glu^{508}, and His^{509}. In contrast, the interaction of the nitrogen atom (atom label: NAR) forms an extremely stable interaction with the Asp^{648} (97% of the simulation) and stable interactions with Ala^{649} (84%) and Gly^{679} (53%), while forming only one transient interaction (Tyr^{684}, 4%). Collectively, the results of the MD simulation suggest that modification of the oxadiazole ring may be a promising fragment growth strategy.

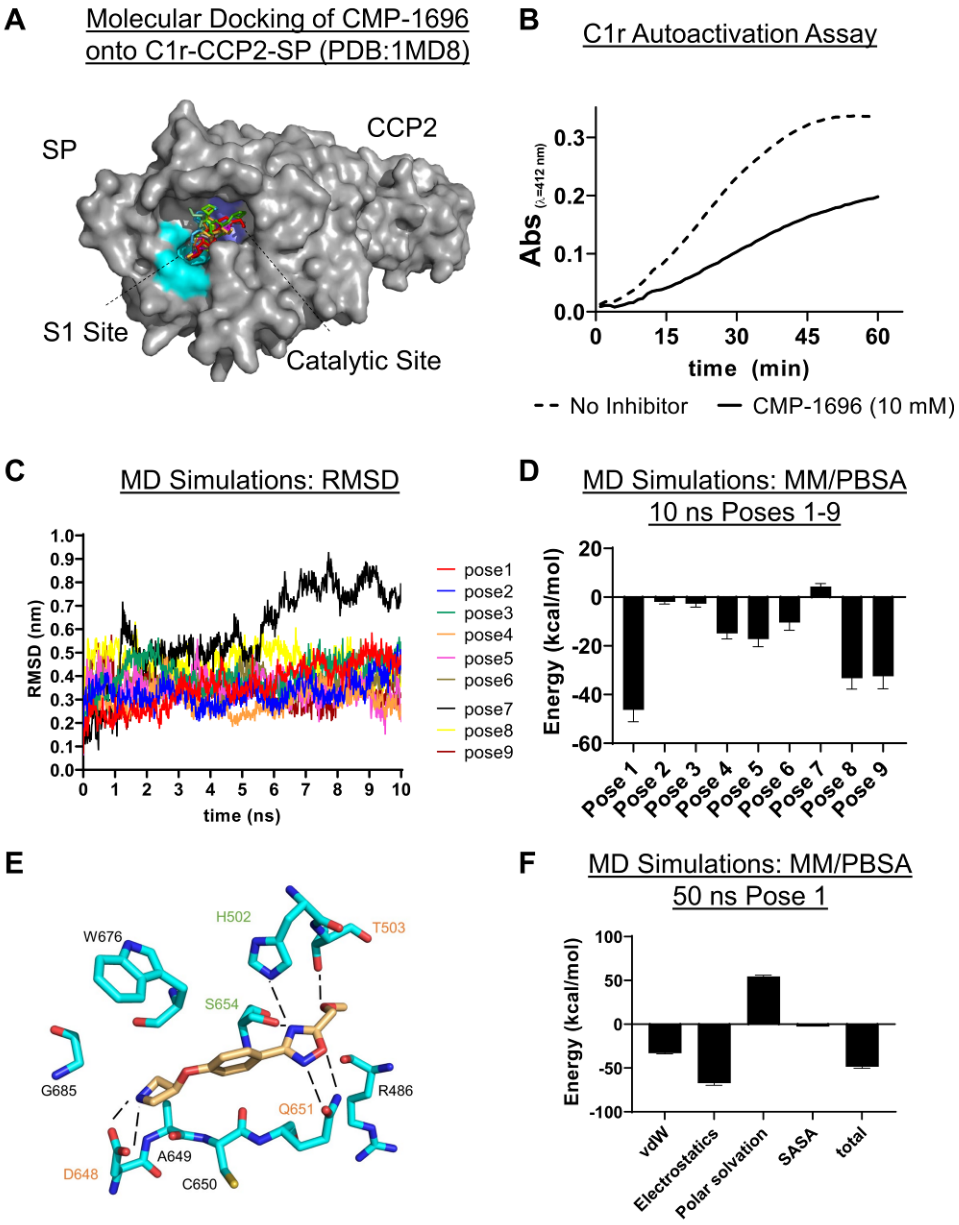


FIGURE 18. CMP-1696 structure activity relationship. **(A)** CMP-1696 was redocked onto C1r-CCP2-SP (PDB: 1MD8, grey surface representation) and the top nine scored poses are shown. All CMP-1696 poses dock into the S1 subsite (cyan) near the catalytic triad (blue). **(B)** C1r autoactivation assay. C1r proenzyme undergoes time-dependent autoactivation at 37 °C. Autoactivation was measured using a synthetic substrate for C1r enzyme. The reaction progress of

vehicle control (dashed line) or in the presence of 10 mM CMP-1696 (solid line) was monitored for 1 hour. **(C)** Molecular dynamics (MD) simulations of CMP-1696/C1r-CCP2-SP. Root mean square deviation (RMSD) in nm for each of the CMP-1696 poses measured over the 10 ns molecular dynamics simulation. **(D)** MM/PBSA energy calculations for each pose in the 10 ns MD simulations indicate that pose 1 is the most energetically favorable. **(E)** All residues within 4.0 Å of CMP-1696 in pose 1 are shown using UNIPROT numbering (ID# P00736). Hydrogen bonding interactions are shown as dashed lines. **(F)** A 50 ns MD simulation was carried out for pose 1 and MM/PBSA was used to calculate total energy. Subcategorized energy contributions are also shown where vDW is van der Waals forces and SASA is solvent-accessible surface area.

The SAR data that was generated provided helpful theoretical information for which to move forward with compound optimization. Due to the lack of empircal binding evidence, such as that which could be obtained through a co-crystal structure of the compound bound within the substrate binding groove of C1r, we sought to validate the obtained SAR data using bioisoteric replacement of the oxadiazole group to generate analogs whose bioactivity could be characterized. In collaboration with medicinal chemists of the John Walker lab at Saint Louis University, School of Medicine, analogs were synethesized where the stable azetidine ring was retained and the less interactive oxadiazole group was replaced with chemically and structurally diverse substituent groups (**Fig. 19A**). Diversity within these newly replaced groups is essential for continuing to drive SAR for compound optimization. Compounds were computationally-generated using substituent groups readily available to expedite the process of SAR elucidation, and docked against the active site of C1r. We then obtained the top-scoring compounds that were synthesized for biofunctional assay evaluation.

This first generation of analogs derived from parent compound CMP-1696 were comprised of seven novel small molecules that were prioritized and selected for based on their predicted binding affnities for C1r via molecular docking studies. The compounds were first evaluated for inhibition of the classical pathway utilizing a classical pathway-specific ELISA and deposition of C4c on the surface of the ELISA plate was measured as a representation of classical pathway activity. At 500 μM, four compounds displayed statistically significant inhibition of complement activation relative to CMP-1696. However, it is most critical to note that all compounds retained the ability to block the classical pathway, serving as a evidence that maintaining the azetidine ring is critical for competitive inhibition of the compound.

The two compounds with the greatest inhibition at 500 μM (32 and 29) were selected for evaluation of dose-dependent inhibition of the classical pathway, as well as analog 35, due to its structural similarity to analog 32, differing only in the replacement of a meta bond for a para bond at the junction of the azetidine ring and replacement group, yet significantly reduced activity against complement as displayed in the single-dose classical pathway ELISA. In two-fold serial dilutions of concentrations ranging from 500 μM to 15.6 μM, the analogs were assessed for dose-dependent inhibition of classical pathway complement. Consistent with the previous single-dose ELISA results, analogs 29 and 32 performed similarly and displayed greater dose-dependent inhibition of complement (2229 IC_{50} = 601 μM, 2232 IC_{50} = 738 μM) than that of analog 35 (IC_{50} = 3.2 mM).

As previously described, one of the greatest challenges in designing small-molecule inhibitors targeting serine proteases is compound specificity against a single target. Classical-pathway specific ELISAs answer to pathway selectivity, however, information about which

complement proteins within the classical pathway are undergoing inhibition cannot be determined. To assess specificity for C1r over structurally similar serine ptorease C1s, we performed C1s-mediated C2 gel cleavage assays, wherein analogs were incubated with 6.25 nM C1s at concentrations of 500 μM. Following incubation, C1s native substrate, C2, was introduced and the reaction allowed to proceed at 37 °C for one hour to ensure protease activity. The reactions were then halted and ran on SDS-PAGE to assess C2 cleavage by C1s. Intact C2 is indicative of C1s inhibition by the compound, whereas the presence of a C2b and C2a band represents C1s activity. Unsurprisingly, all three analogs had activity against structurally homologous C1s, though curiously, analog 2235, which thus far had the worst bioactivity of the three, had the greatest activity against C1s (**Fig. 19D, E**). These substantial differences in activity as seen in analogs 32 and 35 demonstrate how small alterations in bond conformation (meta vs. para) within the structure of the small molecule can have profound effects on the binding interactions between inhibitor and target.

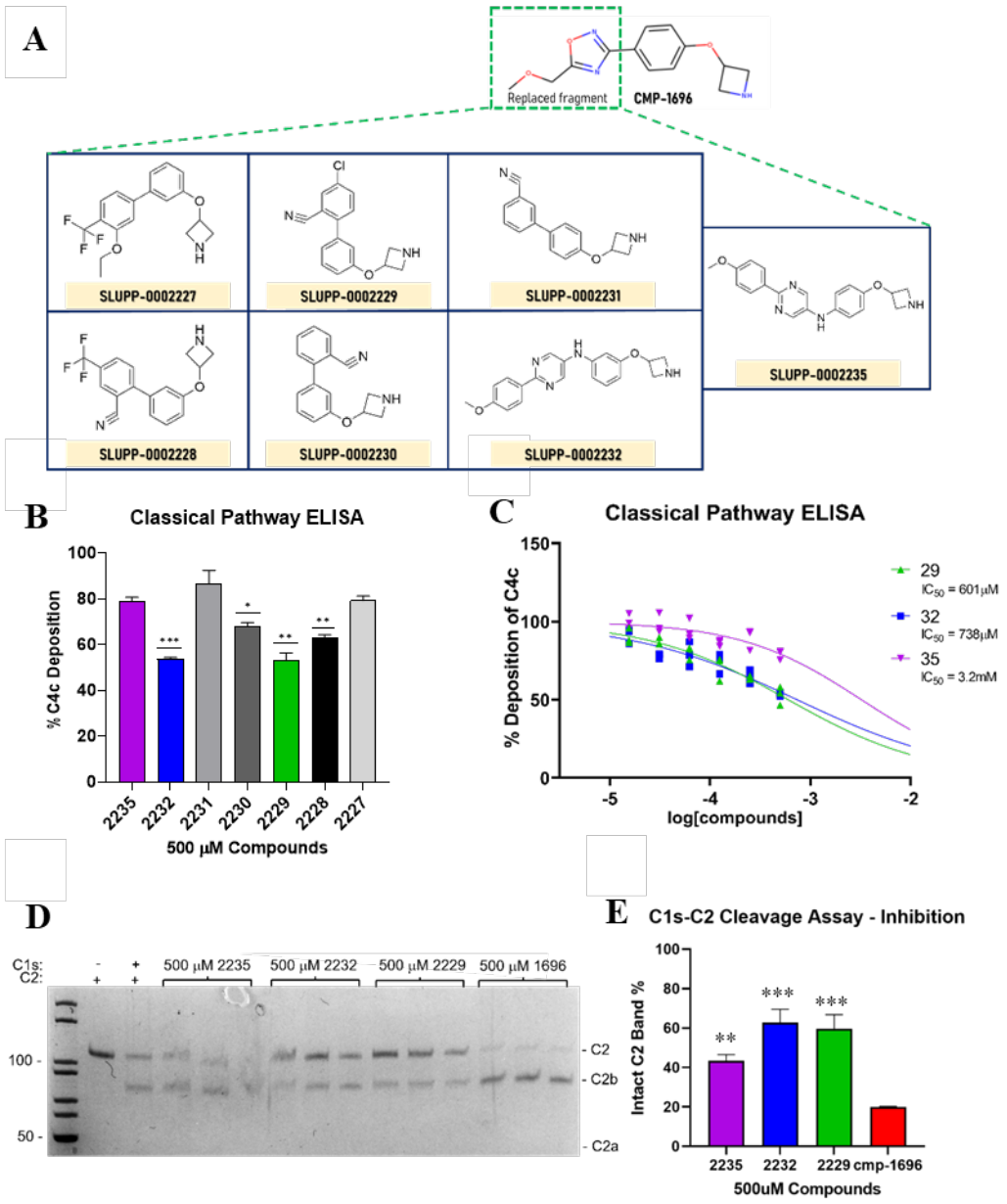


Figure 19. Activity of analogs synthesized from parent compound, CMP-1696. A) Seven newly synthesized compounds formed via the replacement of the unstable oxadiazole methoxyethane group. Compounds were selected for and synthesized post-in silico C1r-CCP2-SP docking screens. **B)** Analogs underwent evaluation via classical pathway-based ELISA at triplicate single 500 µM concentrations. Statistical analysis was performed with GraphPad 8.0 using unpaired Student's t test, where $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All data are presented as mean

± SEM. **C)** Compounds SLUPP-0002229 (29), SLUPP-0002232 (32), and SLUPP-0002235 (35) were tested for dose-dependent inhibition of classical pathway activation in a classical pathway-specific ELISA at concentrations ranging from 500 μM – 15.6 μM. **D)** Additional testing was performed to examine inhibitor specificity by examining the ability of the compounds to block downstream C1s. In C1s-mediated C2 cleavage gel assays, analogs 29, 32, and 35 were tested alongside parent compound, CMP-1696, at 500 μM concentrations and ran on SDS-PAGE. Experiments were performed in triplicate. **E)** Bar chart represents percentage of intact C2 present, demonstrating C1s inhibition against C2 cleavage into fragments C2a and C2b. Statistical analysis was performed with GraphPad 8.0 using unpaired Student's t test, where $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All data are presented as mean ± SEM.

To gain improved insight into the binding mechanisms of these three analogs and increase our pool of SAR data, the John Walker lab synthesized a second generation of analogs using analogs 29, 32, and 35 as the new parent compounds. Analog 29 underwent chemical modification of the cyano group to a carbonyl group to generate second-generation analog 29-A, whereas analogs 32 and 35 had replacement of a methoxyl group to a hydoxyl group, designated 32-A and 35-A, respectively (**Fig. 20A**). When injected in two-fold serial dilution series ranging from 125 or 250 μM to 15.6 μM concrentations over full-length C1r, the chemical modifications resulted in notable changes in binding affinity, as quantified by kinetic fit analysis. Differences in top concentration and lack of binding data for analog 32 and 32-A were due to poor dose-response curves as a result of nonspecific binding to the reference surface. The bioisosteric replacement of the nitrile group to the cyano group in analogs 29 and 29-A resulted in a decrease in binding affinity

from a K_D of 250.7 μM to 475 μM (**Fig. 20B, C**). However, when the methoxyl group was replaced with a hydroxyl group in analog 35 to 35-A, an increase in binding affinity was observed that outperformed that of analog 29, shifting from a K_D of 500 μM to 172 μM (**Fig. 20D, E**, once more demonstrating that subtle chemical modifications can have acute effects on binding interactions taking place between target and inhibitor.

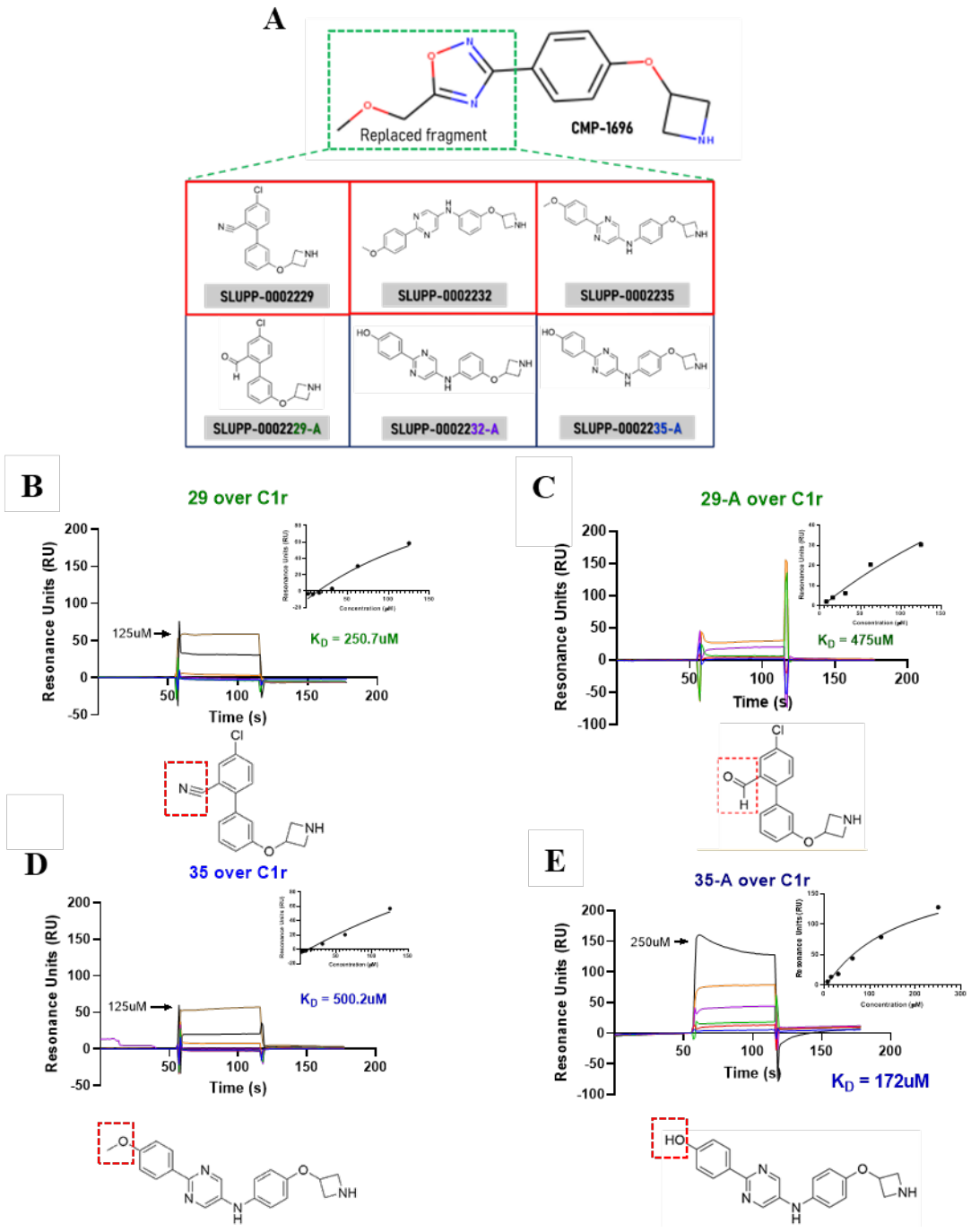


FIGURE 20. Additional R-group substitutions drive activity changes in analog behavior. A) Results from the first generation of analog synthesis, specifically from compounds 29, 32, and 35, were used to synthesize a second generation of analogs with additional R-group substitutions. Molecular docking experiments were performed to provide informational SAR data for the new

chemical modifications. **B-E)** Surface plasmon resonance was performed in which first generation analogs and their successive second generation analogs were injected in two-fold serially diluted concentrations ranging from either 125 μM or 250 μM to 15.6 mM concentrations over a sensor chip surface immobilized with full-length C1r to assess dose-dependent binding. Top immobilization densities differed based on nonspecific binding to the reference surface cell. Chemical modifications between structures are highlighted by red boxes with dashed line. K_D values were calculated using kinetic fit analysis with Biacore Evaluation software and reported as the mean ± S.D.

To delve deeper into the mechanisms underpinning these changes in bioactivity as observed by SPR binding analysis, we performed molecular docking studies to examine the molecular interactions taking place and how they differ between analogs 35 and 35-A. To account for variations in binding results that may occur between algorithms used in docking software, we utilized both SAMSON's ligand docking platform and LigPlot to dock against C1r-CCP2-SP structure, 1MD8 (**Fig. 21**). While analog 35 displays three hydrogen bond interactions that take place between heteroatoms (O1, N2, and N4) present on the compound and residues Gly^{468}, $Glyn^{634}$, and Asp^{631}, found within the C1r active site (**Fig. 21A, B**), the second-generation analog 35-A forms five hydrogen bond interactions with active site residues (**Fig. 21C, D**). Notably, three of those hydrogen bond formations occur on O1 of the newly replaced hydroxyl group and one on N4 of the retained azetidine moiety, generating theoretically stable interactions throughout the whole of the compound with active site residues. The single replacement of the terminal methoxyl group, which forms a single hydrogen bond interaction between O1 of analog 35 and Gly^{468}, to the hydroxyl seen in second-generation analog 35-A, increases the hydrogen bond formation by two

at the newly substituted group, leading to observable increases in binding affinity (nearly three times greater) to active site residues found in C1r.

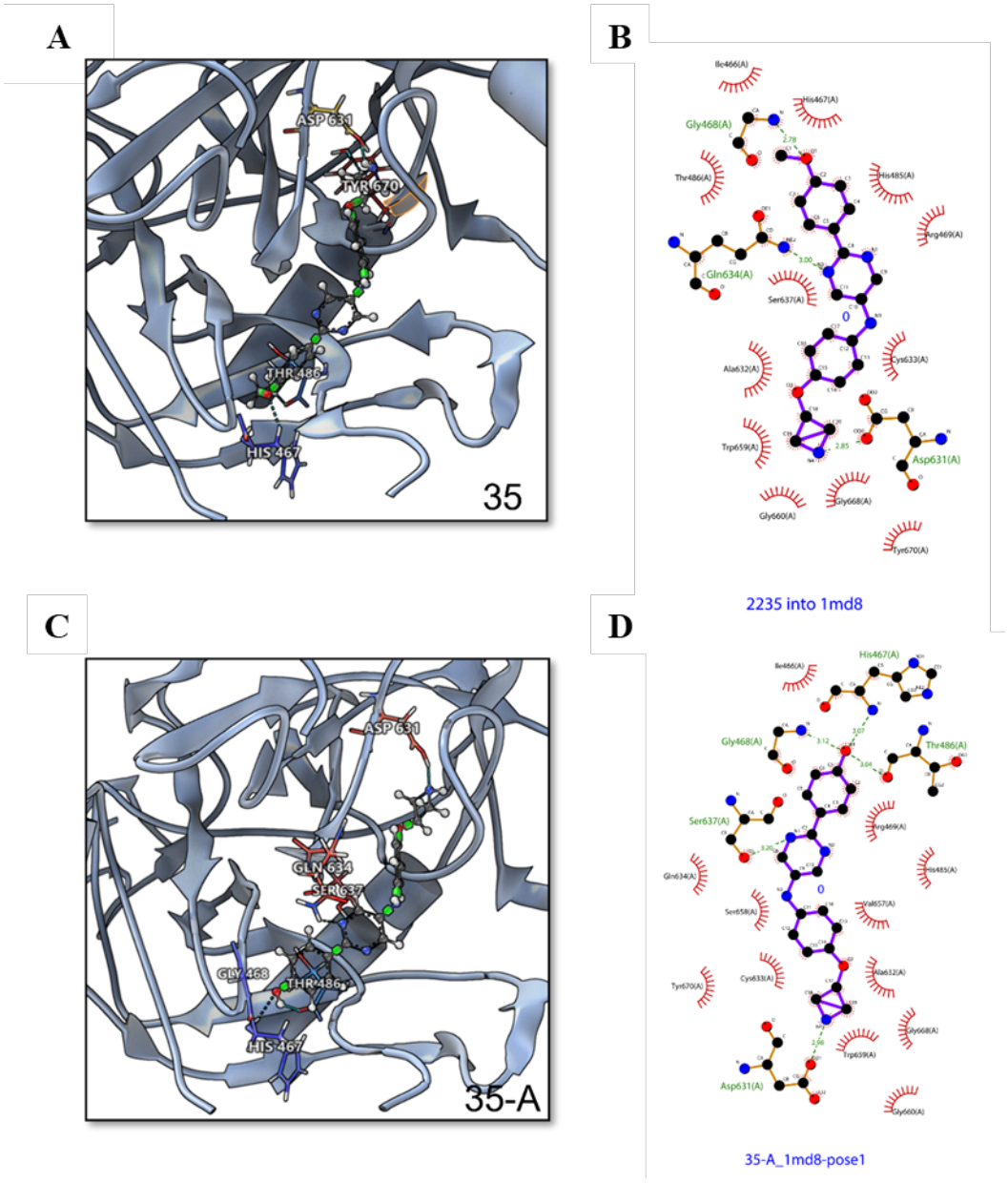


FIGURE 21. **R-group substitutions alter proposed H-bond network and further elucidate structure-activity relationship. A)** Three-dimensional representation of molecular docking of compound 35 (black sticks) show the top binding conformation within the substrate binding groove of C1r (gray ribbon structure). Active site residue contacts are labeled and shown as sticks with

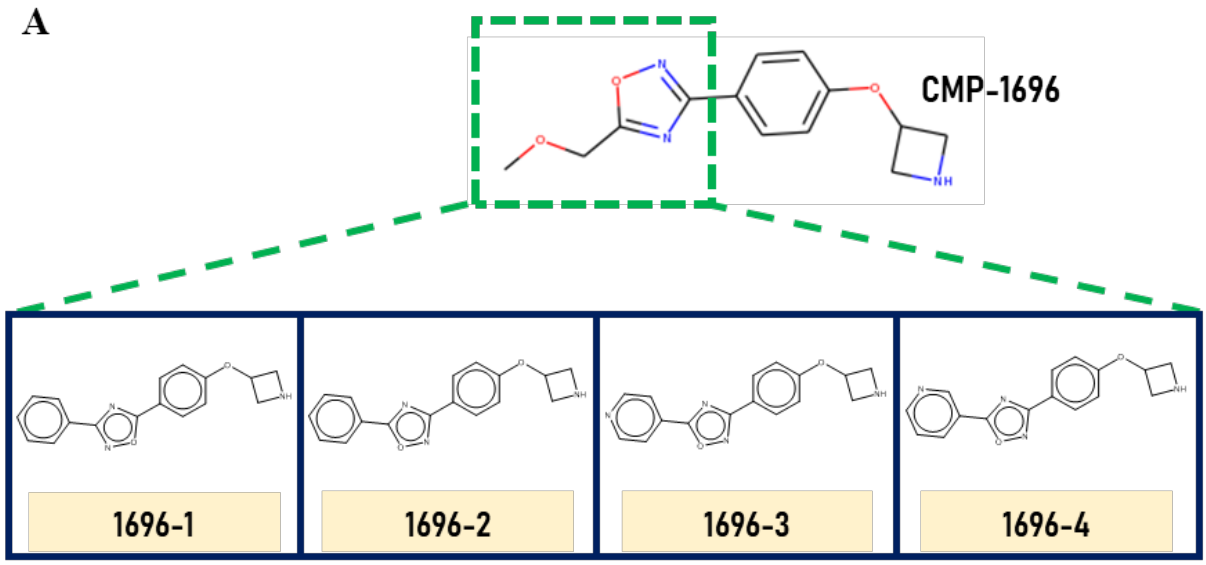
proposed hydrogen bond formation shown as blue dashed lines. **B)** A two dimensional representation of compound 35 interactions within the substrate binding groove are shown. Interacting protein residues are shown as yellow sticks and compound 35 is shown as a purple stick structure. **C)** The three-dimensional representation of second generation analog, compound 35-A (black sticks), bound within the active site of C1r (gray ribbon structure) is shown. Active site residue contacts are labeled and shown as sticks with proposed hydrogen bond formation shown as blue dashed lines. **D)** Compound 35-A (purple sticks) binding interactions with C1r (yellow sticks and red crescents) shown in a two-dimensional rendering. Black spheres represent carbon atoms, white spheres represent hydrogen atoms, blue spheres represent nitrogen atoms, and red spheres represent oxygen atoms. Hydrophobic interactions are shown as red crescents with bristles and hydrogen bonds shown as green dashed lines.

While subtle modifications to substituent groups led to increases in hydrogen bond interactions and increases in binding affinity for second-generation analogs, we next wanted to explore the chemical space involving aromatic ring systems. Much research has gone into structural and chemical analysis of existing drugs available on the market today as a tenable track for drug optimization and design through the improvement of hydrophobic, electronic and steric properties to improve drug-likeness [215]. X-ray crystal structures of protein-small molecule complexes detail the noncovalent interactions that occur between aromatic amino acid side chains of the protein and/or the aromatic and heteroaromatic rings of the inhibitor [216]. Indeed, ring systems form the building blocks of most small-molecule drugs available on the market today [217], which has allowed for data mining and the identification of universal ring replacements for 245 of the most common heterocyclic rings suggested for increased compound bioactivity [215].

This Ring Replacement Recommender developed by Peter Ertl and colleagues at Novartis is a web tool available for researchers and medicinal chemists to explore ring replacement strategies during the early stages of lead optimization. Using this tool, we replaced the oxadiazole group of CMP-1611 with aromatic ring systems with high bioactivity scores and screened the ChemDiv library for available compounds that were structurally similar to those that were designed by ring replacement using hierarchical clustering from atom pair Tanimoto coefficients. Ultimately, we identified four compounds that were structurally related to CMP-1696 that possessed a ring system in place of the oxadiazole group (**Fig. 22A**) to validate whether this platform for CMP-1696 lead optimization was a viable strategy for active site inhibition in C1r.

The compounds obtained from ChemDiv were designated 1696-1, 1696-2, 1696-3, and 1696-4, with the only structural differences between them being the presence/location of heteroatoms in the replaced ring systems. To evaluate efficacy in inhibition of the classical pathway as quantified by C4c deposition, these compounds were employed in a classical pathway-specific ELISA at single doses of 500 μM concentrations and performed in triplicate. Interestingly, not only did each compound retain activity, three of the four exhibited statistically significant improvements in pathway inhibition with 1696-1 displaying approximately twice the improvement in activity as compared to CMP-1696 (1696-1 exhibited approximately 60% inhibition of classical pathway activation as compared to roughly 30% inhibition by CMP-1696) (**Fig. 22B**). Replacement of the oxadiazole group with aromatic ring structures resulted in an increase in

biological activity against classical pathway activation and demonstrates an additional pathway for lead optimization.



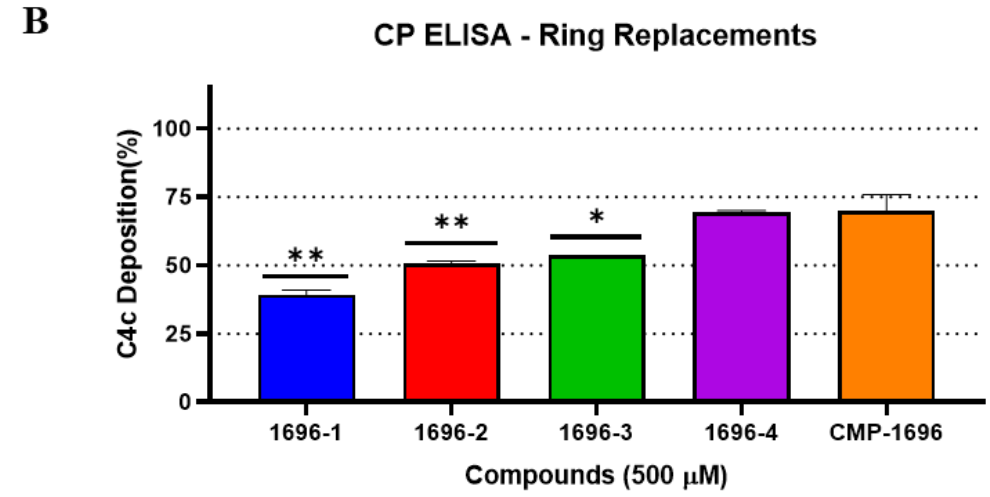


FIGURE 22. Ring replacement of the oxadiazole group in CMP-1696 results in improvements in biological activity. A) Compounds commercially available in the ChemDiv database and structurally similar to compounds generated through the Ring Replacement Recommender web tool were obtained wherein the only modifications were the replacement of the oxadiazole group for a ring system with variances in the heteroatom placement. **B)** Compounds were evaluated for complement inhibition in single-dose classical pathway-specific ELISAs at 500 μM concentrations. Statistical analysis comparing ring replacement analog activity to CMP-1696 was performed with GraphPad 8.0 using unpaired Student's t test, where $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. All data are presented as mean ± SEM.

Discussion

The field of complement-directed therapeutics has grown considerably over the past decade, which is exemplified by latest estimates of over three dozen complement-targeted drugs in various stages of clinical development [55,87,88,173,174]. This growth has been sparked by an increased understanding of the relationship between complement and human diseases, along with the remarkable clinical successes of eculizumab/Soliris. Despite the burgeoning pipeline, drugs targeting the activation pathways are underrepresented, with a vast majority of current drug discovery efforts being devoted to targeting complement at the level of C3 or C5 [87,88]. Complement inhibitors that halt the cascade at the most upstream initiation steps may be ideal therapies for pathway specific-mediated complement conditions and impart the added advantage of leaving other activation pathways available for complement's critical role as a sentinel against pathogens.

The classical pathway of complement has been implicated as a driver of several human diseases [177–182,218]. Development of drugs that specifically target classical pathway components is currently limited to recombinant and native preparations of C1-INH (Cinryze/Berinert/Ruconest), an anti-C1q antibody ANX005 (Annexon Biosciences) and an anti-C1s antibody TNT009/BIVV009 (True North Therapeutics). C1-INH replacement therapy is FDA approved for the treatment of hereditary angioedema, a disorder not primarily driven by complement; however, evaluation of off-label uses for C1-INH in complement-mediated diseases are ongoing [88,173]. ANX005/ANX007 is being investigated in complement-mediated neurodegenerative disorders, while TNT009/BIVV009 has reached phase 3 clinical trials for cold agglutin disease and is being evaluated in other antibody-driven complement-related conditions, such as bullous pemphigoid [88,173]. In sum, the current landscape of classical pathway-specific drug development is characterized by a limited but promising class of large proteins/antibodies in early clinical development for treatment of an increasing number of complement-related conditions.

The mechanisms underlying the role of complement in human diseases are often complex, disease-specific, and in many cases, poorly understood. In some classical pathway-related diseases, autoantibodies activate complement via C1q/immune complex binding and cause the attack of healthy host tissues. In other cases, the involvement of the classical pathway may arise from C1q binding to disease-associated non-antibody ligands capable of activating C1 and initiating downstream complement activation. While the long-term goal of developing therapeutics is paramount, successful identification of high-affinity small-molecule inhibitors specific for C1r would allow complement researchers to better understand the role of the classical pathway initiator protease in numerous models of disease.

In this study, we set out to expand the classical pathway-specific therapeutic toolkit by targeting C1 activation at the level of C1r with small-molecule inhibitors. As the initiating protease of the classical pathway, C1r catalyzes the first proteolytic cleavage event and thus represents the most upstream enzymatic target of the pathway. While complement-directed therapeutics continue to be dominated by antibody-based drugs, small-molecules have recently seen two major breakthroughs in the successful development of factor B and factor D-specific inhibitors [219–221]. In general, small-molecule drugs are afforded greater tissue penetrance relative to antibodies/biologics and can more easily cross the blood-brain barrier [222,223]. In certain complement-related disease settings, such as neurological disorders, these are likely critical considerations. Additionally, small molecules are more easily formulated into oral medications, which could increase patient compliance and lower costs associated with these therapies. However, there are currently no small-molecule inhibitors that specifically target the classical pathway proteases.

To begin to address this gap, we carried out an unbiased screen to identify novel small-molecule fragments that bound directly to C1r. We chose a FBDD approach, as fragment hits are often better positioned than larger, more chemically complex hits, to be optimized with both affinity and specificity considerations in mind. Additionally, we sought to potentially exploit the fact that C1r's function as the initiating protease of the classical pathway is constrained by its unique orientation within the C1 complex, which could allow for the discovery of noncompetitive small-molecule inhibitors. Our initial SPR-based FBDD screen resulted in the identification of 95 compounds with C1r-binding capacity. Ultimately, 24 compounds dose-dependently bound full-length C1r with affinities ranging between 180 – 1,700 μM and ligand efficiencies between 0.16 and 0.41 [136]. Two fragments emerged as leads and exhibited similar C1r affinities (K_D, CMP-

1611 = 480 and K_D, CMP-1696 = 670 μM) and potencies (IC_{50}, CMP-1611 = 660 μM and IC_{50}, CMP-1696 = 520 μM) (**Fig. 17**). CMP-1611 was selective for the classical pathway; however, both its binding and inhibitory modes are currently unclear. In contrast, the data for CMP-1696 suggests that this fragment binds within the S1 pocket of C1r and directly blocks C1r autoactivation (**Fig. 18**). Although CMP-1696 also blocks the lectin pathway, our molecular docking and molecular dynamics studies have provided insight into fragment optimization strategies for C1r-specific compounds.

While the compounds identified here represent high quality fragment hits, we note that several limitations remain for the continued development of these leads. The first challenge is common to nearly all FBDD projects in the need to optimize affinity and potency. In this regard, all 24 C1r-binding fragments identified here benefit from having favorable physiochemical properties, making them generally well-suited for further hit-to-lead development. Moreover, while directly growing an individual fragment by adding substituent groups is one strategy for fragment optimization, it is also possible to link or merge separate hits [132]. Thus, the identification here of structurally distinct compounds, many of which are predicted to bind C1r within close proximity to one another, strongly supports this possibility. Another key challenge to overcome will be that of specificity. Due to the conserved nature of the catalytic domains of serine proteases, compounds like CMP-1696 that target the specificity pockets or catalytic sites, increase the possibility of off-target effects. Indeed, CMP-1696 inhibits the lectin pathway and is predicted to adopt a nearly identical binding mode on MASP-2 (**Fig. 17B**). However, given their low molecular complexity, it is not necessarily expected that fragment hits themselves will exhibit high selectivity [114]. Nonetheless, it is encouraging that active-site targeted small-molecule inhibitors of other complement serine proteases, factor B and factor D, have ultimately overcome this same challenge

of specificity [219–221]. Although our C1r-domain mapping and molecular dynamics studies have provided information about CMP-1696, defining the binding mode of each fragment hit identified here – including CMP-1611- using empirical experimental methods, such as X-ray crystallography coupled with rigorous MD simulations, is a key next step being actively pursued in our laboratory.

In lieu of empirical binding evidence, we used bioisosteric replacement of substituent groups to validate the molecular dynamics simulation results, in which it was predicted that the azetidine group is required to drive activity of the compound against C1r. Replacement of the unstable oxadiazole group with structurally diverse chemical moieties did not result in a loss of activity in several biofunctional assays performed against C1r, supporting computational evidence of the pivotal role the azetidine ring plays in protein-inhibitor interactions. Moreover, specific structural and chemical modifications to the newly synthesized analogs to produce second-generation analogs resulted in a modest increase in bioactivity that could be modulated with subtle changes to terminal chemical entities (*i.e.,* methoxyl group substituted for a hydroxyl group), resulting in an increase in predicted hydrogen bond formation between heteroatoms of the inhibitor with C1r active site residues. Furthermore, employing ring replacement strategies of the oxadiazole group also resulted in statistically significant improvements in biological activity against the classical pathway. The generation of these analogs not only supports MD-derived SAR data but also points to avenues of lead compound optimization.

In summary, we report the identification of 24 novel small-molecule fragments that bind directly to the initiator protease of the classical pathway of complement, C1r. Two of these fragments directly inhibit the classical pathway but exhibit different selectivity and may act by different mechanisms. Each hit compound is small (161 – 406 Da) and has favorable physicochemical properties, and therefore has strong potential to be optimized independently or in

combination. Chemical modification through bioisosteric replacement may serve in the early stages of compound optimization and the developmental evolution of one or more of these compounds may ultimately yield valuable research tools and potentially novel treatment options for diseases associated with aberrant classical pathway activation.

CHAPTER 4: SMALL-MOLECULE INHIBITOR DEVELOPMENT: CMP-1611

Introduction

Serine proteases play a significant role in a number of physiological processes. Many of these processes are involved in signaling pathways, such as blood coagulation, cell signaling, inflammation, digestion, fibrinolysis, development, fertilization, apoptosis, and immune function [81]. Dysregulation of these systems can leave individuals susceptible to disorders involved in coagulopathies, inflammation, infectious diseases, cancer, and degenerative diseases, making serine proteases an attractive target for therapeutic intervention [76]. Efforts have long focused on the development of selective inhibitors to modulate the activity of these enzymes in order to influence disease progression with some of the earliest work dating back to the 1960s with aprotinin, also known as bovine pancreatic trypsin inhibitor (BPTI), which is used to prevent fibrinolysis during cardiopulmonary bypass (CPB) by competitively inhibiting of a broad swath of serine proteases, including plasmin and kallikrein [224–226].

Indeed, structural data reveals that serine proteases share conserved structural and physicochemical similarities, especially among functionally related proteins. These homologous features are especially prevalent in the catalytic domains and active sites, where conserved residues are essential for proteolytic function [124]. As such, development of selective and specific serine protease inhibitors is a challenge for drug design, as it necessitates the identification of subtle molecular nuances or unique features present on the target protein. To add to complications surrounding drug discovery efforts targeting serine proteases, much of drug discovery efforts have focused on competitive inhibition of the highly conserved active site [76]. Overcoming undesirable off-target effects is one of the greatest hurdles in the development of small-molecule inhibitors against serine proteases. This is particularly relevant when dealing with proteases that share

identical catalytic mechanisms and substrate specificities, such as those found within the classical and lectin pathways of complement, e.g., C1r, C1s, MASP-1, and MASP-2 [148,227,228]. Functional dichotomy, as serine proteases perform both essential beneficial roles in regulation and signaling as well as in driving human pathologies, and homologous biomolecular architecture has resulted in the failure of many drug candidates in clinical trials [125].

Despite structural and functional homology, serine proteases, like those found in the complement system, maintain a high degree of macromolecule specificity in limited proteolysis [124], which underscores the importance of identifying molecular nuances, such as those which characterize intermediary states during zymogen activation [74]. High resolution structure-based drug design and molecular modeling techniques are important elements for identifying the molecular and mechanistic frameworks governing protease specificity and, consequently, intelligent drug design. A large body of work has gone into characterizing and developing structural models for understanding serine protease function at an atomic level, including that which was discussed in chapter 2. Discovery of topographical and conformational features outside of the active site that are unique to target serine proteases may provide a means of driving inhibitor specificity.

Noncompetitive inhibition of sites distinct from the active site can modify protease activity through correlated backbone motions of amino acid residues mediated by hydrogen-bond networks at both the local and long-distance range, locking in or blocking structural conformational changes requisite for activity [126,229]. This process is evident in the mechanism of autocatalytic zymogen activation, which occur with MASP-2, where dual conformations and occupancies of residues and disulfide bonds represent different catalytic transition states of the enzyme. These networks of

correlated or anti-correlated motions mediate both noncompetitive transitions and substrate binding [128].

Successful strategies for applied discovery of noncompetitive protease inhibitors targeting these features requires high resolution protein structure determination and binding site analysis, which can be revealed by computational tools such as molecular docking and molecular dynamics simulations approaches [93]. Previously, we had identified two compounds with similar bioactivity against the complement serine protease, C1r. Molecular dynamics simulations revealed a mechanistic binding motif within the active site for one compound, CMP-1696; however, the binding details of the second compound, CMP-1611, remained undetermined. In the absence of robust binding data, mechanistic studies were undertaken to further develop and validate SAR mode of action (**Fig. 23**).

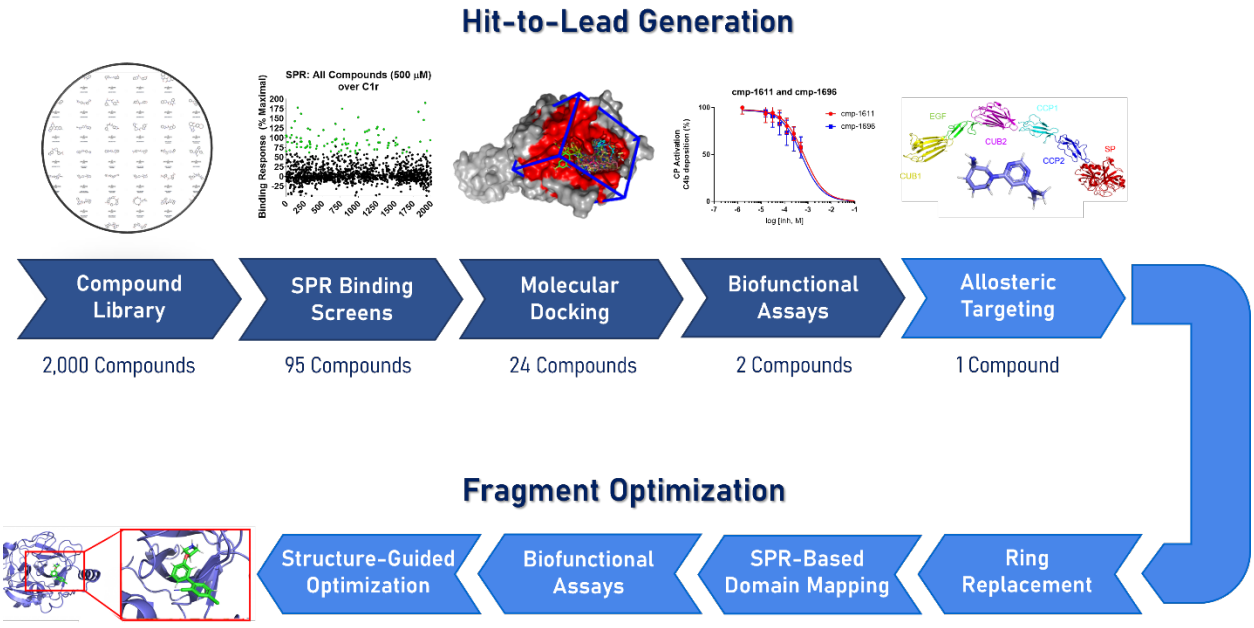


FIGURE 23. Schematic diagram showing hit-to-lead generation to identify CMP-1611 and generate optimized analogs. Initial screening began with a 2,000 small molecule, SPR screening binding screens, molecular docking, biofunctional assays for compound characterization were

utilized to identify two small-molecule hits with favorable bioactivity against target C1r. Molecular dynamics simulations elucidated the SAR data for single compound, CMP-1696, and led to a divergent path for compound development of CMP-1611. Ring replacement strategies for the generation of analogs were employed to further improve SAR and the mode of binding for CMP-1611.

Methods

Molecular docking

Molecular docking was performed as described in chapter 3 methods. Additional docking studies were performed utilizing the AlphaFold full-length structure of C1r, including domains CUB1, EGF, CUB2, CCP1, CCP2, and SP.

Surface plasmon resonance binding experiments

All SPR experiments were carried out using a Biacore T200 (GE Healthcare) at 25°C and performed as previously described in chapter 3 methods.

Complement Inhibition Assay

Complement pathway-specific ELISAs were performed as previously described in chapter 3 methods.

DataWarrior Scaffold Analysis

Chemistry data analysis and visualization platform, *DataWarrio*r, was used for the identification of the core chemical scaffold primarily responsible for activity of the compound. Chemical and structural descriptors were generated through input of the CMP-1611 SMILES code.

Scaffold Analysis identified core structures and categorized them based on scaffold type. From the identified scaffolds, SAR tables were generated wherein the chemical features of scaffolds were assigned a column and the scaffold with attached R-groups and substituents was labeled core-fragment.

Ring replacement strategy

Ring Replacement Recommender v2002.02 was utilized to identify ring analogs derived from the compound structure of CMP-1611. Ring features present on CMP-1611 could be theoretically and selectively swapped with an analogous ring structure with improved biological activity based on information derived from medicinal chemistry literature sources. The average improved bioactivity improved by replacement in log10 scale is reported along with the number of literature references from which this data sourced. Optimized compounds were visualized in ChemDraw 20.1.1 and SMILES generated from the new chemical structures. Those SMILES files generated by Ring Replacement were used to search and select for structurally similar compounds using hierarchical clustering from atom pair Tanimoto coefficients that were commercially available through ChemDiv.

Results

Compound screening and triaging

An initial screen of 2,000 small-molecule compounds sourced from ChemDiv compound libraries encompassing five subsets: 2-dimensional fragments, 3-dimensional fragments, serine-protease inhibitor compounds, natural product scaffolds, and protein-protein interaction inhibitor compounds (PPI). Those compounds were screened and triaged via SPR binding studies in which non-specific binders and compounds exhibiting superstoichiometric binding over full-length C1r were eliminated, reducing the compound candidates from 2,000 to 95. Molecular docking studies

evaluating and ranking the predicted binding energies of compounds modeled to bind near or within the active site of C1r further narrowed compound candidates to 24. Extensive biofunctional assays were performed to identify and select the two most promising compound candidates for further optimization, CMP-1696 and CMP-1611 (**Fig. 24**).

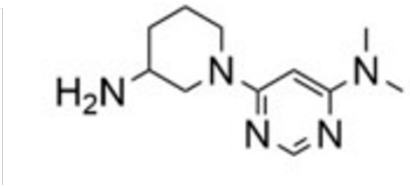


FIGURE 24. Compound structure of CMP-1611.

Dose-dependent binding of full-length C1r

A critical element for hit compound identification was the ability of the compound to dose-dependently bind full-length C1r. In our SPR binding experiments, CMP-1611 displayed dose-dependent binding in triplicate injection series of two-fold serial dilution concentrations ranging from 500 μM – 7.8 μM (**Fig. 25A**). The steady-state affinity of 480 μM was calculated using Biacore Evaluation software to fit sensorgrams from each variable concentration injection dataset using a 1:1 Langmuir model of interaction constrained by a theoretical experimental binding response. SPR binding experiments assessing CMP-1611 binding over C1r-CCP1-CCP2-SP yielded no binding, interestingly suggesting that the site of binding is located on one of the N-terminal domains, CUB1, EGF, or CUB2.

Dose-dependent inhibition of the classical pathway

To assess dose-dependent inhibition of the classical pathway, inhibitor concentrations of CMP-1611 ranging from 500 mM – 15.6 mM were incubated with 2% (v/v) NHS and C4b deposition was measured as a metric for pathway activation (**Fig. 25B**). The half-maximal

inhibitory concentration of 660 µM with an associated 95% confidence interval of (560 – 790 µM, n=9) was calculated fitting dose-response curves using inhibitor vs. response models.

Delineating pathway-specificity

Developing selective serine proteases remains a challenge to the field due to the structural similarity found between serine proteases. To test for selectivity, CMP-1611 was tested in triplicate at 500 µM concentrations in ELISAs selective for either the lectin or alternative pathway based on measurable C4b or C3c deposition, respectively, resulting in colorimetric differences detectable by optical sensor. Surprisingly, despite C1r structural homology to MASP-1 and MASP-2, in the lectin pathway and alternative pathway systems, little to no complement inhibition was observed (**Fig. 25C**). These results suggest that noncompetitive inhibition by CMP-1611 occurs at a structural facet unique to C1r that is present on the one of the N-terminal domains.

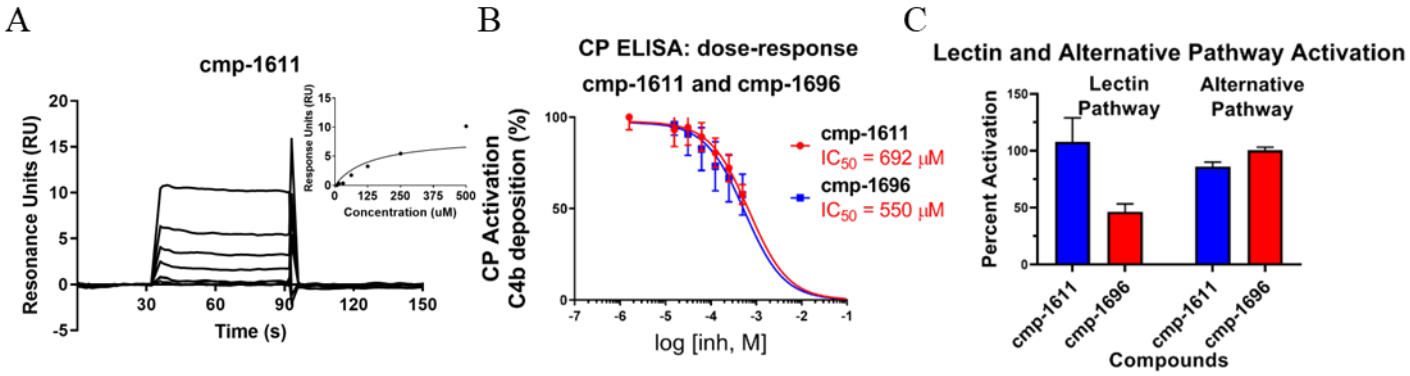


FIGURE 25. Dose-dependent binding and inhibition of C1r.

Molecular docking studies

Previous molecular docking experiments utilizing each domain of C1r as the macromolecular structure for which to dock and triage compounds from the initial screen had identified CMP-1611 as having preferential binding to the SP domain. Indeed, SPR experiments comparing binding interactions between full-length C1r and C1r-1, 2SP reveal that while CMP-

1611 does bind C1r, it does not interact with the catalytic C-terminal domains. Using the full-length structure of C1r derived from AlphaFold 2.0 as the macromolecular target for molecular docking program PyRx, we see inconsistent binding predictions across the protein surface. These low-binding interactions do not account from the data obtained pointing to N-terminal binding that occurs between CMP-1611 and C1r. This may be explained by conformational changes that are present in C1r when in C1 complex or due to the absence of surface pockets that may not be evident or present in the state that AlphaFold presents C1r, as proteins are known to inherently sample differing low-energy conformation states.

Ring replacement and what it informs about SAR and previous use of Data Warrior to identify core scaffold

There are several established methods for improving the activity of a fragment against a target using structure-guided approaches employing fragment growing, merging, linking, or a combination of those methods. Optimization approaches can originate from available ligand or protein structural information in lieu of a complexed structure. Previously, we utilized DataWarrior's Scaffold Analysis to identify the core pyrimidine fragment of CMP-1611 believed to engage in the most stable interactions between compound and protein residues. From there, we obtained two structurally similar, commercially available compounds that contained the pyrimidine scaffold from ChemDiv's compound library. Those compounds retained similar dose-dependent inhibitory capabilities against the activation of the classical pathway, supporting the SAR analysis and hypothesis that the pyrimidine scaffold drives activity against C1r in the inhibitor (**Fig. 26A, B)**.

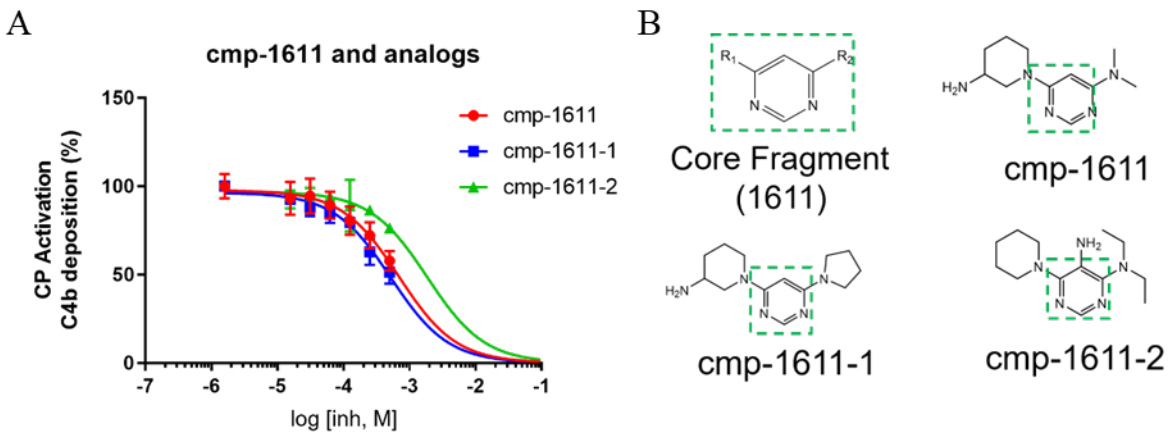


Figure 26. Analogs derived from parent CMP-1611 with retained core pyrimidine scaffold retain inhibitory characteristics against classical pathway activation. A) Compound analogs of CMP-1611 display dose-dependent inhibition against the classical pathway in a pathway-specific ELISA. **B)** Two-dimensional representation of core scaffold and compound structures.

Building off the results that suggest that the core pyrimidine is an essential element for C1r inhibition, we decided upon a novel ring replacement optimization approach for the next generation of CMP-1611 analogs. Because the initial SAR information is derived from theoretical binding modes and *in vitro* biofunctional assays, the further development of a structurally diverse iteration of analogs is vital for improved elucidation of SAR data.

The electronic, hydrophobic, and steric properties of substituent groups can dramatically alter the bioactivity of a compound, and exploiting this facet can lend itself to effective hit optimization. One such available method of visualizing these chemical properties is via a two-dimensional Craig plot, named for its developer, Paul N. Craig, in which the polarizability, parachor, or molar attraction constraints of such substituent groups is plotted as a function of their relative independence in the bioisoteric space [230,231]. This tool allows the substituent property space to be further explored based on groups that are expected to possess similar physicochemical

properties, such as size, shape, and electronic properties, and therefore similar bioactivity. To a non-medicinal chemist, bioisoteric replacement of substituent groups may appear as minor structural alterations; however, they may be utilized to favorably alter ligand-target interactions. Several approaches specifically target ring systems for replacement due to the critical and easily navigable role they play in compound activity and, specifically, potency, selectivity, and pharmacokinetics.

Increased efforts have been made to derive compound bioactivity data from scientific literature analyzing known drugs, as well as X-ray structures of protein-ligand complexes, allowing for the estimation of average contributions of functional groups, particularly aromatic ring substituents, and identification of the most common binding interactions as to how they influence the binding energy of the compound [215]. One particularly effective tool for early lead optimization through heterocyclic ring replacements is the Ring Replacement Recommender web tool, in which parent rings are selected and ring substitutions suggested based on predicted improvements in biological activity, which may result from face-to-face π-π stacking, edge-to-face π-π stacking and π-amide stacking. From this web tool, the associated biological activity of over two million molecules was utilized from the ChEMBL database, which extracts data from available scientific literature to divine universal building blocks that may increase binding affinity. Using the core pyrimidine scaffold identified from DataWarrior's Scaffold Analysis as the parent ring, we employed the Ring Replacement Recommender and Craig plot to design structurally distinct fragment compounds to explore and increase SAR data (**Fig. 27**).

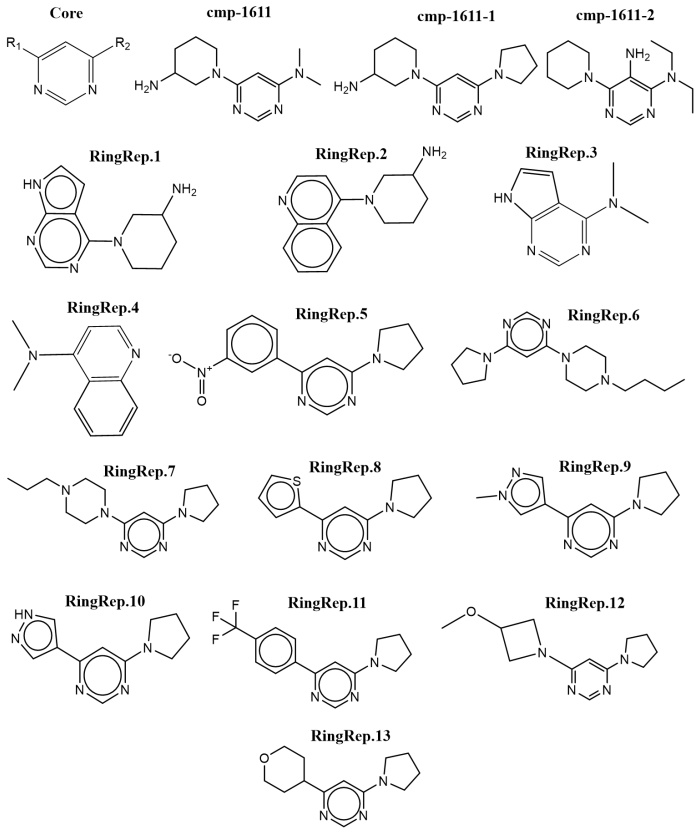


FIGURE 27. Design strategy for CMP-1611 analogs was performed using the Craig plot and ring replacement. Replacement of substituent aromatic ring groups was guided by chemical property and bioactivity analysis of compound chemical entities.

For more rapid turnaround, we then screened the ChemDiv library for available compounds that were structurally similar to those that were designed by ring replacement using hierarchical clustering from atom pair Tanimoto coefficients. While ChemDiv did not have any compounds with exact structural matches to those the designed compounds, we were able to select for compounds that provided ring substituents with diverse electronic and hydrophobic properties to broaden SAR information (**Fig. 28**).

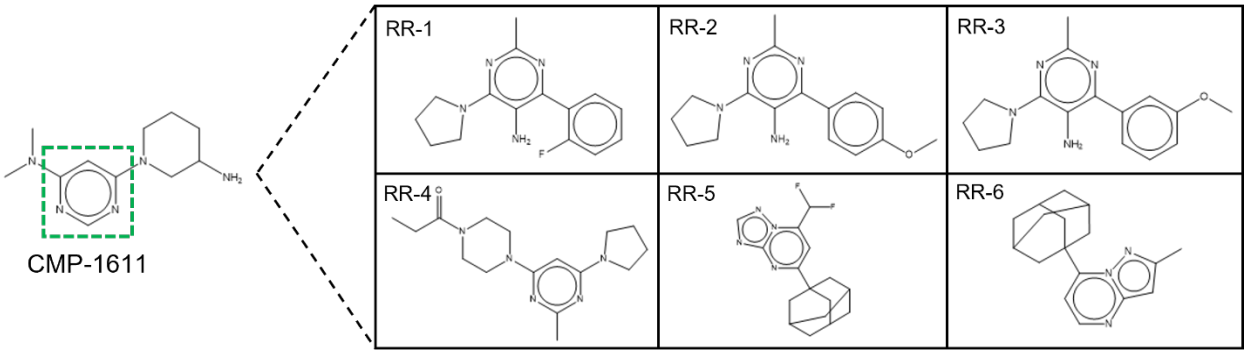


FIGURE 28. Structurally similar compounds to CMP-1611 analogs designed by ring replacement. All compounds retain the core pyrimidine ring and were obtained from ChemDiv.

Compounds were then tested at 500 μM in a classical pathway-based ELISA for their ability to block pathway activation as quantified through C4c deposition on the plate surface in comparison to parent compound CMP-1611. Despite retention of the core pyrimidine ring, second-generation analogs did not exhibit increased bioactivity against target C1r (**Fig. 29**). This is likely owing to the fact that the actual structures of the designed ring replacement compounds were not commercially available through ChemDiv and would instead require synthesis.

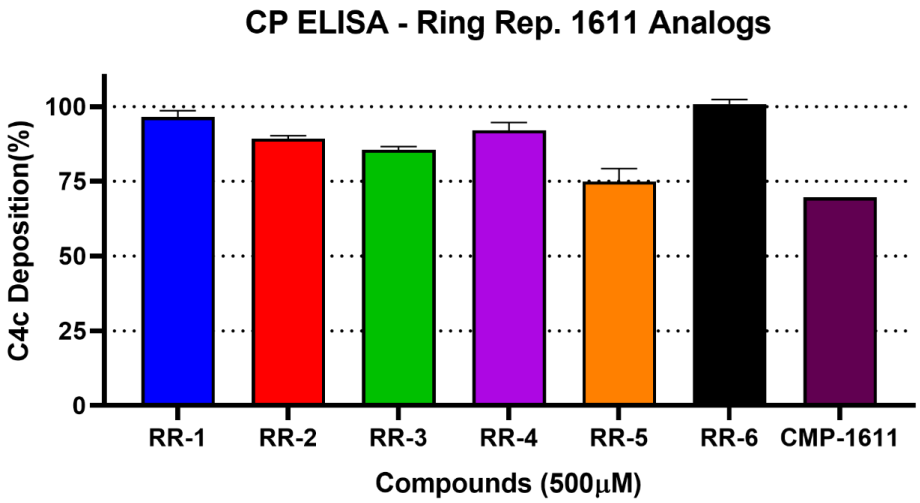


FIGURE 29. Evaluation of second-generation analogs for inhibition against the classical pathway of complement. Compounds were tested in triplicate at single concentrations of 500 μM for their ability to block C4c deposition.

While improved activity was not achieved, most analogs did retain marginal activity against classical pathway activation (**Fig. 29**). Furthermore, strict emphasis on the pyrimidine core scaffold as the most important element responsible for CMP-1611 activity may need revision. Further work to elucidate SAR details is necessary for continued optimization of CMP-1611.

Discussion

The vast majority of protease-targeted therapeutics available on the market are designed with the intent of competitive binding to the active site, thereby blocking or modifying enzymatic activity [101]. This becomes a challenge in designing effective serine protease inhibitors owing to the fact that there is overlapping functionality and substate specificity, as well as structural homology between different serine proteases. Accordingly, interest in noncompetitive sites of

inhibition has gained momentum, especially through technological advancements in molecular dynamic simulations, allowing for the elucidation of cooperative/correlated protein motions at the atomic level.

Beyond the realm of computational molecular approaches for identifying and designing specific small-molecule inhibitors, high-throughput and fragment-based screens also allow for the identification of novel compounds that have the ability to bind distal pockets present on the biological target and interfere with protease activity. Inherent to small molecules and fragment-based drug design is low molecular complexity, allowing for biochemical optimization that is targeted towards increased binding affinity and specificity for the target [132]. Low molecular weight and complexity also provides for the ability to engage in bioisosterism and scaffold hopping to build SAR data and explore new chemical spaces for identifying novel drug candidates [217].

At present, much headway has been made in characterizing the atomistic elements and molecular interactions that occur between protein-small-molecule inhibitor complex structures available in the PDB. With this wealth of structural information, tools have been developed to identify high-frequency patterns and common molecular interaction fingerprints present in these datasets, aiding in the rational design and optimization of small-molecule drugs. Furthermore, these tools improve the information available for increasing the efficiency and design of virtual screening platforms for both target and hit identification, with the most frequent non-covalent interactions occurring between aromatic ring structures present in either receptor or ligand [232]. Building from available strategies for SAR elucidation and compound optimization, we can improve efficiency and potency in drug development.

In the development of small-molecule inhibitors targeting classical pathway protease C1r, we identified fragment hit, CMP-1611, with favorable binding and inhibitory characteristics

against C1r. Structural elucidation studies failed to derive the mode of action of this compound, and using scaffold analysis and ring replacement strategies for the generation of analogs we expanded upon SAR data and contributed to future efforts to develop and rationally design an noncompetitive small-molecule inhibitor with improved pharmacokinetics against C1r.

CHAPTER 5: SMALL-MOLECULE INHIBITOR DEVELOPMENT: A1

Abstract

Each year, the number of drug candidates in the research and development pipeline increases, yet despite advancements driving this ongoing growth, most candidates never progress to clinical trials. Of those that do, about 90% of those drugs in clinical development fail primarily due to toxicity and lack of clinical efficacy [120], highlighting the challenges that come with developing a highly selective drug with favorable physicochemical properties. Adding to the obstacles that must be overcome, many families of drug targets, such as serine proteases, share significant structural and functional homology/redundancy between related proteins, necessitating the development of multiple drug options and classes for disease treatment. This requisite for therapeutic variety is further heightened when accounting for the intricacies involved with treating disease from the standpoint of patient complexity as well as disease evolution and trajectory. Machine learning and artificial intelligence aid in the rational drug design process and therapeutic efficacy by rapidly identifying scaffold and compound "hits" with favorable drug-like properties from libraries of millions of compounds and chemical entities, delivering compounds with promising bioactivity against its target, precipitously cutting down on time and cost expenditures to drug developers [141]. In this study, in collaboration with the Geisbrecht lab at Kansas State University, and AI-driven pharmaceutical company, Atomwise, we identified a fragment with low micromolar affinity and half-maximal inhibitory values against complement protease C1s. Using the information derived from our biofunctional assays and molecular docking studies, the John Walker lab at St. Louis University optimized the fragment lead to generate a next-generation analog with similar bioactivity, establishing the structure-activity relationship essential for

optimizing a small-molecule compound with even greater drug-likeness for future therapeutic design in diseases-mediate by the classical pathway of complement.

Introduction

The ever growing number diseases with known involvement from the classical pathway of complement has drawn increased interest from biomedical and drug development groups for the specific therapeutic targeting of the classical pathway. While C1r-based complement inhibitors are notably absent from the therapeutic space, efforts have been made to selectively target the ensuing protease in line in the classical pathway, C1s, for inhibition. To date, there is one FDA-approved anti-C1s monoclonal antibody, Sutimlimab – formerly known as BIVV009 or TNT009, for the treatment of adults with cold agglutinin disease (CAD) with several other anti-C1s targeted candidates at various stages of clinical trials [90]. These drug candidates rely on successful inhibition of C1s to prevent further activation of the classical pathway and ultimately obviscate downstream effector activities involving damage-associated inflammation and/or extravascular hemolysis. It is important to note that complement involvement in clinical conditions is complex, and every condition must be assessed individually to the extent that the recent successes of Sutimlimab may not encompass all individual patients. Targeted inhibition of C1s using small-molecule inhibitors may add to the therapeutic arsenal against dysregulation of complement in classical pathway-mediated diseases. Herein, we have identified a potent fragment hit that selectively targets the active site and displays dose-dependent binding and inhibition of C1s.

The classical pathway of complement is initiated through the binding event of C1q to a ligand following recognition of damage- or pathogen-associated molecular patterns, which can have both host or pathogenic origins, by C1q in complex with heterotetrameric serine proteases,

C1r and C1s, or through the recognition and binding of immune complexes [1](**Fig. 30A**). The structural features of C1r and MASP-2 have been previously discussed, and as a point of fact, C1s retains a homologous quarternary structure in that it is comprised of six domains that from N-terminus to C-terminus are: CUB1, EGF, CUB2, CCP1, CCP2, and the serine protease (SP) domain, respectively [233] (**Fig. 30B**). The 79.8 kDa protein similarly relies on its proteolytic function through the catalytic region CCP1-CCP2-SP, in which the SP domain possesses trypsin and MASP-2-like substrate specificity for positively charged lysine and arginine side chains [234]. When C1 complex is bound to a surface, C1r undergoes the process of autoactivation, as previously discussed, and subsequently cleaves the Arg^{437}-Ile^{438} bond found within C1s for its activation from proenzyme to active enzyme. This cleavage event is an important distinction from C1r and MASP-2, as C1s is not capable of autoactivation and therefore must be acted upon by C1r for enzymatic activity [148].

This cleavage event is also critical for the forward progression of the classical pathway, wherein like MASP-2, C1s cleaves its native substrates, C4 and C2, to form C3 convertase on the target surface, C4b2b [235]. This complex then serves to cleave C3, from which C3a is released as a potent anaphylatoxin, and C3b remains associated with the complex through a newly formed thioester bond to form C5 convertase, C4b2b3b. The C3b fragment serves as a binding site for C5, wherein C5 is cleaved into the released chemotactant agent, C5a, and the first component of the terminal pathway and lytic membrane attack complex (MAC), C5b. C5b serves to bind one unit of C6, which then allows for C7 binding and dissociation of MAC from C5 convertase to the plasma membrane through the insertion of a hydrophobic site present on C7 into the lipid bilayer of the cell surface [1]. From initiation to termination, the classical pathway serves as a powerful weapon in the immune arsenal to opsonize danger-associated molecules, generate potent

chemotactant agents for the recruitment of professional phagocytes to the target site, and lyse cellular membranes via MAC.

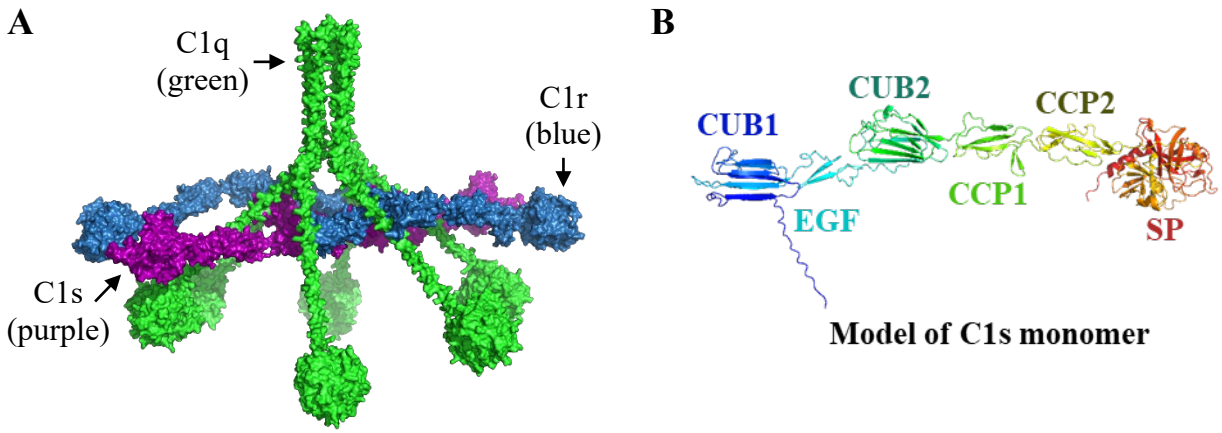


FIGURE 30. The structure of C1 complex with C1s. A) The model of C1q (green) in complex with serine proteases, C1r (blue) and C1s (purple), to form heteromeric C1 complex. **B)** The AlphaFold model of full-length C1s. The N-terminal domains are represented by CUB1, EGF, and CUB2, whereas, the catalytic C-terminal region is made up of CCP1, CCP2, and the SP domain.

During normal homeostatic events, these effector functions of the classical pathway and crosstalk with the coagulation and contact pathways provide protection against pathogens and cellular damage. However, as a consequence of overactivation or dysregulation, these functions can play key roles in a long list of detrimental processes and pathologies involved with autoimmunity, inflammation, and neurodegeneration. Current therapeutics targeting classical pathway-mediated disorders, impart particular focus on those associated with the generation of autoantibodies and hemolytic syndromes, such as cold agglutinin disease-associated hemolytic

anemia (CAD) [90], Huntington's disease (HD) , warm autoimmune hemolytic anemia (wAIHA), amyotrophic lateral sclerosis (ALS), and Guillain-Barré syndrome (GBS) [236], among others.

Anti-C1s monoclonal antibody, Sutimlimab, has undergone FDA approval for treatment of CAD, an autoimmune hemolytic anemia, due to its ability to prevent opsonization by complement of red blood cells following recognition of IgM autoantibodies against RBC antigens (cold agglutinins) active at cold temperatures. It is has been shown that inhibition at the level of C1s prevents the generation and opsonization of RBCs by C3b and the subsequent recruitment of receptor-specific macrophages resulting in temperature-dependent hemolysis [90]. Additional support for therapeutics targeting specifically classical pathway-mediated diseases is demonstrated by anti-C1q monoclonal antibody, ANX005, which is currently being evaluated in phase 2 clinical trials by Annexon Biosciences for Huntington's disease (HD), warm autoimmune hemolytic anemia (wAIHA), amyotrophic lateral sclerosis (ALS), and phase 3 clinical trials for Guillain-Barré syndrome (GBS) [91]. Applied broadly across a number of conditions, inhibition of C1q can have therapeutic use by blocking C1q interactions with host targets both dependent and independent on the presence of autoantibodies. While classical pathway targeted therapeutics are showing promise across a wide swath of pathologies, the development of additional classes of inhibitors, including small-molecule inhibitors, is essential for improved treatment outcomes for patients.

With advancements in medicine, data processing, and our understanding of disease presentation and mechanisms, it has become increasingly apparent that classic phenotyping, by which patients are categorized, is inadequate for therapeutic treatment. The current state of drug development relies heavily on patient cohorts generated from disease presentation that may belie underlying biological process by not taking into account individual factors, such as temporality

involving disease evolution and trajectory, patient comorbidities, genetic profiles, epigenetic factors, and clinical history among others. The high degree of complexity represented in individual patients and attempts to simplify disease as a linear progression with clinical granularity provides inadequate accounting for critical patient heterogeneity that goes into much of drug development. Additional treatment options for complement-mediated disorders that are represented outside of intravenous administration of monoclonal antibodies, is a necessary element for delivering options to a hetergenous patient population.

Small-molecule inhibitors targeting C1s provide an additional tool to move the needle in the treatment of complex systems. Advantages to small-molecule drug development are increased tissue permeability and patient compliance in an orally bioavailable drug [99]. Moreover, detailed SAR analyses of currently available small-molecule inhibitors in complex with protein targets with structures available in the PDB show that most replicate some component of human biology and targeted interactions already present in nature. Protein modification by enzymes, such as phosphorylation or glycosylation, results in changes to structure and therefore protein function, which is what small-molecule drugs frequently replicate through molecular mimicry. These protein-ligand interactions compete for natural ligands or cofactors for enzymes, and the utilization of these SAR data and structural information derived individually from protein target and ligand are heavily leaned upon for drug development.

Machine learning and artificial intelligence have exploited these data sets to gain novel insights into favorable biochemical features present in both protein and small-molecule inhibitors at a level not previously attainable in the history of medicine. Biopharmaceutical companies employing differing algorithms and strategies now have the ability to rationally screen billions of small molecules and compound scaffolds for hit-to-lead generation [142]. Atomwise, based out of

San Francisco, utilizes its deep learning technology to advance structure-based drug design for the rapid identification and synthesis of small-molecule drug candidates against a wide-variety of proteins [237]. Additionally, it offers its research platform to academic, government, and non-profit research institutes to further structure-based drug development campaigns. In collaboration with the Brian Geisbrecht lab at Kansas State University-partnered with Atomwise, we identified a highly potent fragment hit against the active site of C1s that exhibits dose-dependent binding and inhibition of the classical pathway of complement and were able to synthesize two analogs via bioisosterism with similar bioactivity profiles and modes of action.

Methods

Expression and purification of recombinant C1s domain truncation mutants

Purified C1s was purchased from Complement Technology. Recombinant C1s-domain truncations were produced by sub-cloning an *Escherichia coli* codon-optimized synthetic oligonucleotide (IDT Technologies) flanked with a 5' BamHI site, a 3' NotI site and a stop codon into the pT7HMT vector [190]. The C1s-CCP2-SP construct corresponds to C1s residues 356-688, whereas the C1s-CCP1-CCP2-SP construct corresponds to residues 292-688 (UNIPROT numbering: #P00736). C1s-domain truncations were purified under denaturing conditions using previously published protocols with some modifications [63,191]. Plasmids were transformed into *E. coli* BL21 (DE3) and cells were grown to an optical density of 0.6-0.8 OD600 at 37°C in Terrific Broth supplemented with kanamycin. Cultures were then induced overnight with isopropyl β-D-thiogalactoside and cells were collected by spinning for 10 min at 4000 x g. Supernatants were discarded and resuspended in 100 mL of lysis buffer (6 M guanidine HCl, 100 mM Tris pH 8.0, 10 mM imidazole) for 30 min and clarified by centrifugation for 30 min at 16000 x g. Nickel NTA

beads (GoldBio) (5 mL column volume) were washed with 25 mL of denaturing binding buffer (8 M urea, 500 mM NaCl, 20 mM sodium phosphate pH 6.0, 10 mM imidazole). Samples were then passed over columns at a flow rate of 1-2 drops per second and afterwards washed with another 25 mL of denaturing binding buffer. Samples were eluted with 5 mL of denaturing elution buffer (8 M urea, 500 mM sodium chloride, 20 mM sodium phosphate pH 6.0, 200 mM imidazole) and then rapidly diluted 1:10 into refold buffer containing 50 mM Tris pH 8.3, 3 mM reduced glutathione, 1 mM oxidized glutathione, 5 mM ethylenediaminetetraacetic acid (EDTA), 500 mM arginine. The next day, samples were dialyzed twice for four hours at 25°C against 2 L of 50 mM Tris-HCl pH 7.4, 145 mM NaCl and concentrated to less than 12 mL using 10 kDa molecular weight cutoff ultracentrifugation filters (Millipore). Refolded protein was then purified using an ÄKTA pure Fast Pressure Liquid Chromatograph (FPLC, GE Healthcare) connected to a HiLoad 26/600 Superdex 75 pg column previously equilibrated in 200 mM sodium chloride and 20 mM Tris pH 8.0. Fractions were evaluated by SDS-PAGE and those containing the C1r truncation mutant were pooled. To remove affinity tags, tobacco etch virus (TEV) enzyme (previously activated with 1 mM DTT) was incubated with pooled refolded protein overnight at 25°C. FPLC nickel affinity chromatography was carried out the following day and the unbound fraction was collected and concentrated by ultracentrifugation. The sample was then incubated with 5 mM $CaCl_2$ and 50 nM purified C1r for enzyme activation overnight at 37°C. The activated C1s mutants were further purified to remove C1r by size-exclusion chromatography. The purified fractions were collected, concentrated by centrifugation, and buffer exchanged into 20 mM HEPES, 140 mM NaCl, aliquoted, and stored at -80°C until use. Sequence alignments were performed with EMBL-EBI Clustal Omega [192].

Crystallography

Attempts to co-crystalize the inhibitor bound to C1s were made to be able to empirically visualize protein-ligand interactions. Crystals formed in the presence of 100 μM inhibitor were crystallized by hanging drop vapor diffusion technique. Crystals were obtained by mixing a 2 mL:2 mL protein to precipitant ratio of 3.28 mg ml^{-1} C1s-CCP1-CCP2-SP in 49.5% (v/v) Tacsimate, pH 7.0 modified from the Hampton index screen condition and from a precipitant mixture of 0.1 M HEPES pH 7.3, 1.2 M Sodium citrate, 20% glycerol modified from the Hampton crystal screen condition mixed with 3.08 mg ml^{-1} C1s-CCP2-SP. The mixtures were equilibrated against a reservoir containing 500 mL of precipitant at 4°C. Crystals were harvested and exchanged into crystallographic conditions supplemented with 100 μM A1 and 30% and 20% glycerol, respectively, for cryoprotection prior to vitrification in liquid nitrogen.

Pathway-specific ELISAs

Complement pathway-specific ELISAs were performed as previously described in chapter 3 methods.

Gel-Based Inhibition of C1s-/MASP-2-Mediated C2 Cleavage Assay

To demonstrate inhibition of C1s- or MASP-2- mediated cleavage of C2, a 10 μL reaction in HBS-Ca^{2+} (10 mM HEPES (pH 7.3), 140 mM NaCl, 5 mM $CaCl_2$) was made by adding 6.25 nM C1s enzyme or, alternatively, 25 nM MASP-2-CCP2-SP with two-fold dilutions of compound A1 from 100 μM to 1.56 μM with subsequent addition of 1.25 μL of C2 (1 mg/mL) (Complement Technologies). In experiments using single concentrations in triplicate, 500 μM A1 was used. The reaction proceeded at 37º C for 1 hour and was stopped by the addition of 5 μL Laemmli buffer followed by boiling for 5 minutes. Subsequently, 10% SDS-PAGE gels were utilized with

Coomassie staining. Gel imaging was completed on a ChemiDocTM XRS+ (Bio-Rad). Gels are representative of three independent experiments.

Gel-based C2 inhibition assays were subjected to further quantitative analysis with Image Lab^{TM} (Bio-Rad). Lanes and bands were manually selected and analyzed as follows. C2b bands were in lane corrected for total C2 (C2 + C2b + C2a). 100% cleavage was constrained to C1s + C2 control. A normalized four-parameter nonparametric response was analyzed in GraphPad v8.4.

Classical pathway hemolysis assays

To assess the lytic inhibitory capabilities of A1 in both mouse and human serum, 50 mLs of gelatin veronal buffer (GVB^{++}) (20 mM HEPES (pH 7.3), 140mM NaCl, 0.1% gelatin (w/v), 0.15mM $CaCl_2$, and 0.5 mM $MgCl_2$) (CompTech) was prepared. In a 96-well nonbinding plate, triplicate series of 20 μL of A1 at 500 μM and 100 μM concentrations were incubated for 10 minutes with 10 μL of 20% normal human serum diluted into GVB^{++} or, alternatively, 100% normal mouse serum co-incubated with 10 μL mouse complement assay reagent (MCAR) (CompTech) to prevent loss of activity due to the inherent instability of mouse serum. Following incubation, 45 μL GVB^{++} was added to all wells. Control wells lacked inhibitor and was compensated for by the addition of 10 μL GVB^{++}, and the negative control wells, in the absence of serum, saw the addition of 10 μL GVB^{++}.

Sheep red blood cells coated with rabbit anti-sheep erythrocyte antiserum (EA (Ab-sensitized sheep E)) were washed thrice with GVB^{++} in by centrifugation at 500 x g and resuspended in GVB^{++} at least 25μL per well and wash with GHB^{++} 3X in centrifuge at 500 x g. Following wash, 25 μL of the EA cells were added to each well and gently mixed. The plates were then sealed with a top seal and incubated at 37°C for 1 hour with agitation every 15 minutes.

Cells that were not lysed were spun down in a plate rotor at 500 x g for 3 minutes. Supernatant was then transferred in 50 µL volume into non-binding polystyrene plates. The optical density of each plate was read at 412 nm in EnSight multimode plate reader (Perkin-Elmer). These data were then normalized with 100% lysis represented by the positive controls in GraphPad Prism 8. All experiments were performed no less than three times.

Molecular docking

Molecular docking was performed as previously described in chapter 3 methods using protein structure files for the C-terminal region of C1s (CCP2-SP; PDB: 1ELV) and were prepared for docking by removing water molecules and adding hydrogens using AutoDockTools1.5.6 before generating protein structure PDBQT files. Molecular docking was then performed using AutoDock Vina 1.1.2 [46] and as detailed previously [199]. Nine poses for each compound were collected and included in the full dataset. Molecular docking of A1 onto the C1s-CCP2-SP crystal structure was performed using PDB:1ELV. A grid box encompassing only the SP domain of the receptor was chosen. Molecular docking was then performed with Vina executable embedded with PyRx v0.9.8 [200] resulting in 9 binding poses for each compound. All representations of protein structures were prepared using PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC (www.pymol.org/).

Results

Atomwise technology and compound screening

Collaborative efforts to identify small-molecule hits that bind and inhibit C1s catalytic function, were undertaken with the Geisbrecht lab at Kansas State University and pharmaceutical company, Atomwise, to generate a starting point from which to develop compounds with

therapeutic potential through the AIMS Awards Program. Atomwise utilizes a deep convolutional neural network model, AtomNet, to elucidate quantitative structure-activity relationships and ligand-based bioactivity predictions to screen small-molecule libraries and triage compounds based on their predicted bioactivity. Utilizing information derived from both the ligand and target structure, molecular features are detected in order to describe favorable and unfavorable interactions between ligands and the target [237].

Structural information from which to identify C1s-directed small molecule inhibitors was drawn from the C1s-CCP2-SP crystal structure, PDB entry 1ELV. The high resolution (1.7 Å) structure allowed for the clear electron density mapping of the protease active site, which was utilized to form the parameters of the ligand binding site. Crystallographic waters present within the crystal structure provided a molecular fingerprint and coordinates of features from essential active site residues for which to bind based on interactions with the solvent network, which could then be exploited for structure-based ligand-binding affinity predictions [130].

A molecular library of several million small-molecule compounds was screened against the assigned target binding site and ranked hierarchically based on favorable biochemical features and predicted binding energies. In total, 96 compounds were selected and synthesized by Atomwise and subsequently blinded for *in vitro* testing. Compounds were assigned names based off their well position within the 96-well plate. Dose-dependent SPR binding experiments against C1s-CCP2-SP and C1s-CCP1-CCP2-SP and single concentration enzyme activity assays were performed for all 96 compounds at a top concentration of 100 μM by the Geisbrecht lab at Kansas State University. From the original plate of 96 compounds, 13 were highlighted as having either relative high binding affinity to C1s or were inhibitory against C1s cleavage of a synthetic peptide substrate, with the compound A1 having favorable activities in all three assays.

Classical pathway-specific ELISA screening

To assess compound inhibition against the classical pathway, 100 μM concentrations of compounds were tested in triplicate in the presence of 2% normal human serum in classical pathway specific ELISAs, in which C4b deposition on the plate surface was quantified as a measure of pathway activation. In this assay system, only four compounds displayed greater than 25% inhibition of classical pathway activation, with A1 again presenting itself as one of the four compounds displaying favorable inhibition at 100 μM.

Evaluation of dose-dependent inhibition by A1 in the classical pathway ELISA

Based off the promising binding and inhibitory qualities exhibited by A1 across four different assay systems, additional biofunctional assays were undertaken with selective focus on A1 for further characterization of the IC_{50} value, pathway specificity, and potential for use in murine model systems of disease. To this end, we examined the ability of compound A1 to dose-dependently block classical pathway activation using two-fold serial dilutions of A1 ranging from 100 μM to 1.56 μM in a classical pathway-specific ELISA using 2% normal human serum. C4b deposition was quantified as a measure of pathway activity. In this assay, A1 exhibited dose-dependent inhibition of the classical pathway, and by fitting dose-response curves using inhibitor vs. response models in GraphPad Prism 8, the data were used to obtain an IC_{50} value of 20 μM (**Fig. 31A**).

Evaluation of dose-dependent inhibition by A1 in the classical pathway-specific hemolysis assay

Another metric to evaluate the ability of A1 to block complement activity is through the use of pathway-specific hemolysis assays, which provide a sensitive, quantitative means of calculating classical pathway-mediated hemolysis of IgG sensitized sheep red blood cells.

Concentrations of A1 were serially diluted in GVB^{++} in concentrations ranging from 100 μM to 1.56 μM prior to incubation with 2% (v/v) NHS to allow for complement inhibition. Reactions were then incubated with EA cells for one hour, in which unlysed cells were spun down and the supernatants containing cell lysis had optical densities measured and normalized against 100% hemolysis. From these values, an IC_{50} of 16 μM was calculated using GraphPad Prism, closely matching that which was observed in classical pathway ELISAs (**Fig. 31B**).

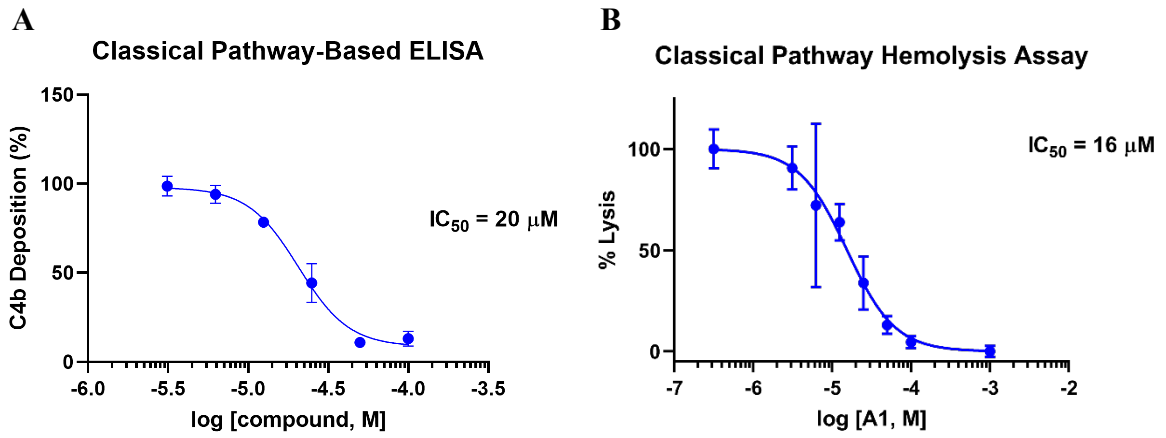


FIGURE 31. Dose-dependent inhibition of complement in classical pathway-specific assays by A1. A) Compound A1 displayed inhibition of the classical pathway in a dose-dependent manner in a classical pathway-specific ELISA, as quantified by C4b deposition. **B)** In classical pathway-specific hemolysis assays, A1 blocked complement-mediated hemolysis in a dose-dependent manner. Both assays were performed in triplicate and concentrations of A1 were serially-diluted in buffer at concentrations ranging from 100 μM – 1.56 μM prior to incubation with 2% (v/v) NHS. Data were fit with GraphPad Prism using non-linear regression with a log(inhibitor) vs. response model with an associated 95% confidence interval.

Compound A1 selectivity for C1s of the classical pathway

To test for A1's ability to dose-dependently block the proteolytic activity of C1s on its native substrate C2, C1s-mediated C2 cleavage gel assays were performed using two-fold serial dilutions of A1 ranging from 100 μM to 1.56 μM were incubated with 6.25 nM full-length C1s, after which, 1 mg ml^{-1} C2 was introduced and the reaction allowed to proceed for 1 hour at 37º C to ensure enzyme activity. The reactions were then halted through the introduction of 5 μL Laemmli buffer followed by boiling for 5 minutes and run on 10% SDS-PAGE gels resolved with Coomassie staining (**Fig. 32A**). The gels were then imaged and quantitative analysis performed in which lanes and bands were manually selected and analyzed with C2b bands lane corrected for total C2 (C2 + C2b + C2a). A normalized four-parameter nonparametric response was analyzed in GraphPad v8.4 from which IC_{50} values of 84 μM could be derived from both intact C2 and C2b deposition (**Fig. 32B**).

Next, we examined selectivity of A1 for C1s of the classical pathway over structurally and functionally related MASP-2 of the lectin pathway, given that both proteases share substrate specificity and have structurally similar substrate binding grooves. C2 gel cleavage assays were performed with triplicate concentrations of 500 μM and 100 μM A1 incubated with either 6.25 nM C1s or 25 nM MASP-2 prior to the addition of 1 mg ml^{-1} C2. Again, reactions were allowed to proceed for 1 hour at 37º C at which point they were halted and run on SDS-PAGE (**Fig. 32C**). Following analysis, it was revealed that A1 exhibited nearly complete inhibition of C2 cleavage by C1s at both concentrations of 500 μM and 100 μM. At 500 μM of A1, MASP-2 cleavage of C2 to C2a and C2b inhibited by 13% ± 0.6 and 47.5% ± 2.0 when incubated with 100 μM A1 (**Fig. 32D**). The results demonstrate that A1 has some selectivity for MASP-2, though to a lesser extent than that of C1s.

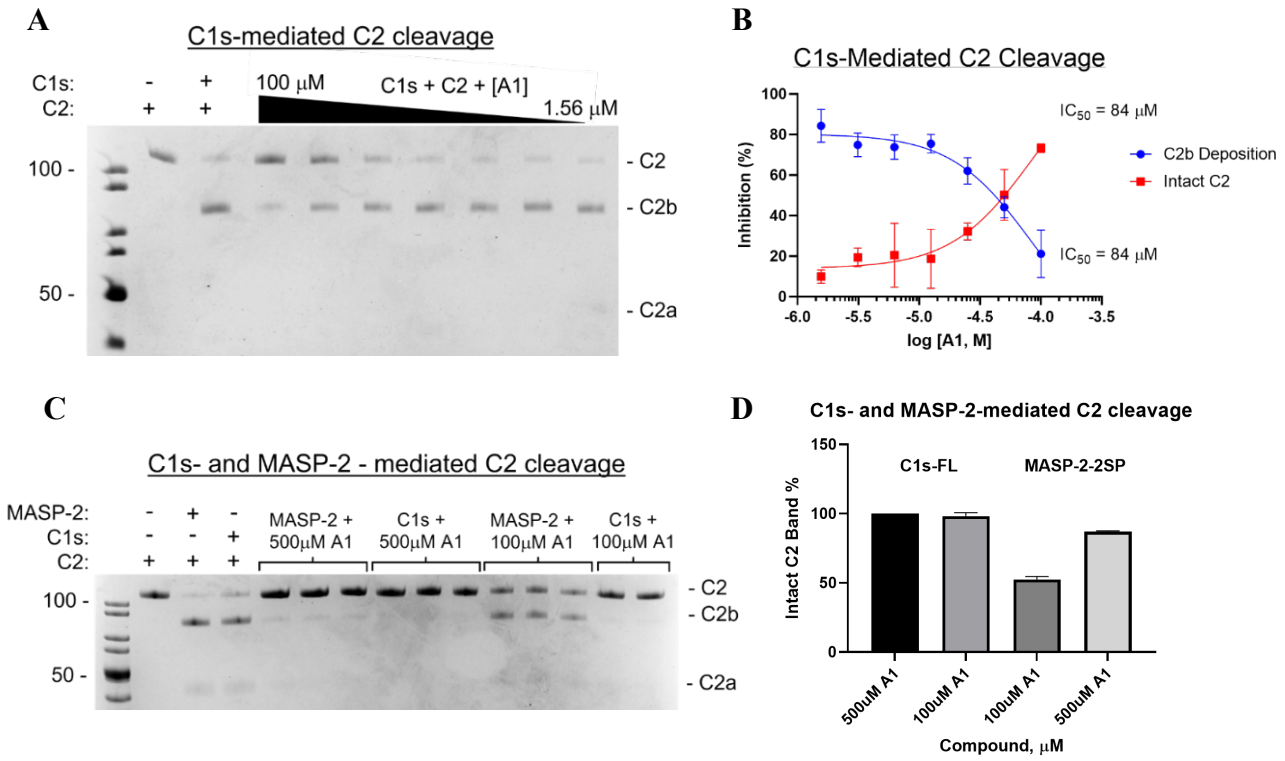


FIGURE 32. Compound A1 displays specific dose-dependent inhibition of C1s, and to a lesser extent, MASP-2. A) A representative SDS-PAGE gel showing dose-dependent C1s inhibition by A1. Intact C2 bands represent C1s inhibition, whereas the presence of C2b and C2a bands represent C1s proteolytic activity against substrate C2. Concentrations of A1 ranged from 100 µM to 1.56 µM and were incubated with 6.25 nM full-length C1s. **B)** A representative graph displaying both C1s inhibition (red line) and activity (blue line) in the presence of A1. IC_{50} values were calculated by fitting dose-response curves with a non-linear regression with a log(inhibitor) vs. response model. **C)** Inhibition of both 6.25 nM C1s and 25 nM MASP-2 against triplicate concentrations of 500 µM and 100 µM A1 in the presence of C2. **D)** Bar graph showing C1s- and MASP-2-mediated cleavage of C2 by A1 at 500 µM and 100 µM concentrations.

Having established that A1 blocks cleavage of native substrates and displays some preference for MASP-2, we next evaluated the activity of A1 in lectin pathway and alternative pathway-specific ELISAs. Using 100 μM concentrations of inhibitor, A1 was incubated with 2% (v/v) C1q-depleted serum in the lectin pathway-specific ELISA and 20% NHS in the alternative pathway-specific ELISA. C4d and C3b deposition was quantified to represent pathway activation in each respective pathway and the experiments performed in triplicate. At 100 μM concentrations of A1, lectin pathway activity was reduced to 72.7% ± 3.0, which corresponds with the MASP-2 inhibition seen in the C2 cleavage gel assays. Interestingly, the alternative pathway showed increased activity at 349.3% ± 53.6. Further analysis is necessary to understand the mechanism of action in the alternative pathway (**Fig. 33A**).

Looking ahead to the potential of A1 in future *in vivo* experiments, we examined whether A1 can not only block complement in human serum but also in normal mouse serum for murine studies. To establish the ability of A1 to block the classical pathway to the extent of MAC formation, sheep red blood cells sensitized with rabbit anti-sheep erythrocyte antiserum (EA cells) were used in classical pathway hemolysis assays. At concentrations of 500 mM and 100 mM, A1 was incubated on a 96-well microtiter plate with either 10% NHS or 100% NMS co-incubated with MCAR to maintain serum stability. Following incubation, washed EA cells were introduced and allowed to continue to incubate at 37°C for 1 hour with agitation every 15 minutes. The plates were centrifuged in a plate rotor and the optical density of the supernatant containing lysed cells was read at 412 nm in EnSight multimode plate reader (Perkin-Elmer). These data were then normalized with 100% lysis represented by the positive controls in GraphPad Prism 8. While NMS inherently shows lower activity than NHS, A1 displayed inhibition of hemolysis by MAC in both human and mouse serum. At concentrations of 500 μM A1, the percent lysis of EA cells was

completely abolished in NHS and 18.42% ± 7.78 in NMS. By reducing the concentration of A1 to 100 μM, the percent lysis became 17.78% ± 10.26 in NHS and 44.91% ± 13.42 in NMS (**Fig. 33B**). The ability of A1 to inhibit hemolysis in both NHS and NMS will prove valuable in future *in vivo* studies of disease models, especially those marked by autoimmunity or hemolytic dysfunction.

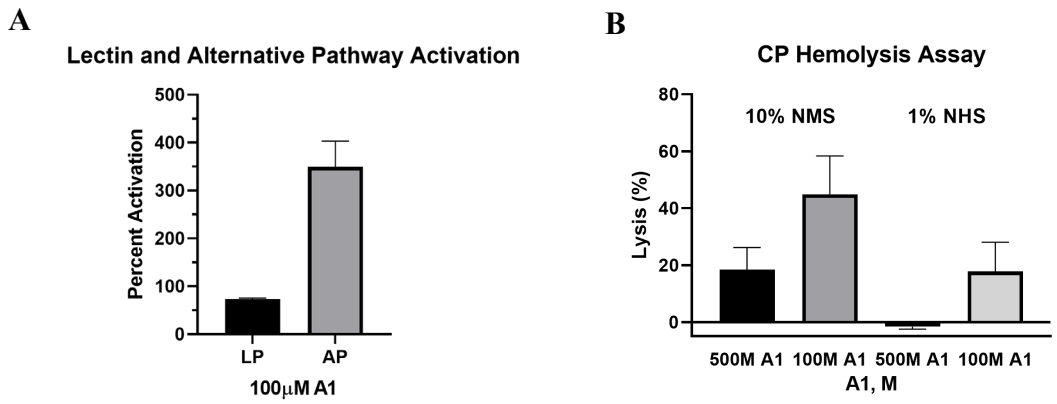


FIGURE 33. Compound A1 showing differential activity against other pathways of the complement system and activity against murine-derived complement. A) In lectin and alternative-pathway specific ELISAs, A1 displayed differential inhibition of complement at single concentrations of 100 μM. A1 exhibited some inhibition of the lectin pathway, whereas appeared to increase complement activity in the alternative pathway. **B)** Measured responses of 500 mM and 100 mM concentrations of A1 against complement of the classical pathway derived from murine serum (10% (v/v) NMS) and 2% (v/v) NHS are represented by bar graph. In all assay systems, experiments were performed in triplicate and represented as mean ± S.E.

Co-crystallography efforts

Gaining insights into the molecular details of the structure-activity relationship of A1 with C1s is a necessary element for optimization to increase compound specificity and inhibitory capabilities. Attempts were made to co-crystallize A1 in complex with C1s domain truncation mutants. Hampton crystal and index screens were employed to crystallize both C1s-1, 2SP and C1s-CCP2-SP in the presence of 100 μM A1, as well as crystallize those two forms of C1s for subsequent compound soaks with A1. Under two precipitant conditions, 49.5% (v/v) Tacsimate, pH 7.0 and 0.1 M HEPES pH 7.3, 1.2 M Sodium citrate, 20% glycerol, C1s-CCP1-CCP2-SP and C1s-CCP2-SP crystals formed in the presence of A1, and additional crystals were produced by use of hanging drop vapor diffusion technique (**Fig. 34**). Crystals were harvested and exchanged into crystallographic conditions supplemented with additional 100 μM A1 to try to ensure ligand-protein complex formation. Despite having well-formed crystals, attempts to collect X-ray diffraction data were unsuccessful.

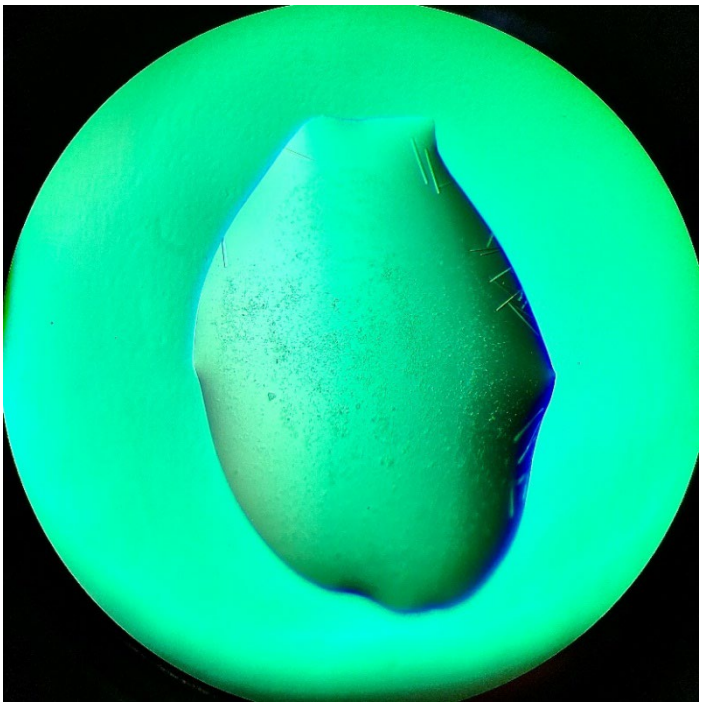


FIGURE 34. Initial co-crystallization efforts towards soaking inhibitor with C1s-CCP2-SP.

Molecular docking

While additional work to generate high resolution diffracting crystals of C1s complexed with A1 must necessarily be continued, advancements in molecular modeling technology lend itself to reasonably reliable SAR data generation. Following comprehensive biofunctional analysis of A1, the compound structure was unblinded and the PDBQT structure file was docked against full-length C1s and C1s-CCP2-SP (PDB entry, 1ELV). The binding parameters encompassed the entirety of the protein to allow for unbiased generation of the most energetically favorable binding sites and compound conformations within those sites. Molecular docking was then performed using AutoDock Vina 1.1.2 and the top scoring nine poses were collected. Within each model, A1 consistently bound within the active site of the serine protease domain.

Analog synthesis

Utilizing the predicted molecular binding interactions between A1 and the C1s active site, we were able to identify functional groups present on the compound that we believed to be responsible for driving activity. For synthesis of analogs, ring replacement, bioisosteric substitution, and pharmacokinetic characterization was performed computationally using commercially available substiuent groups that could be readily synthesized in lab by Dr. John Walker and Hinman Zhou at Saint Louis University. Newly designed compounds underwent molecular docking against C1s-CCP2-SP and analogs were subsequently ranked according to free binding energies. Compounds with favorable biochemical/functional characteristics were

synthesized, wherein they were sent to the Geisbrecht lab for *in vitro* evaluation by SPR binding assays and enzyme inhibition assays.

To assess inhibition of the classical pathway, we subjected the 11 compounds to classical pathway-specific ELISAs at triplicate concentrations of 100 μM and tested for C4c deposition at the plate surface. Of the 11 compounds, three (compounds -92, -95, and -98) had greater than 50% inhibition of classical pathway activation using 2% (v/v) NHS (**Fig. 35A**). The most promising lead in terms of binding affinity and inhibition of C1s against synthetic substrate Z-S-Lys from the Geisbrecht lab was compound -93. To replicate more physiologically relevant conditions to assess compound activity, we assessed dose-dependent inhibition of the classical pathway via a classical pathway-specfic hemolysis assays, in which compound concentrations were serially diluted in GVB^{++} and ranged from 100 μM – 15.6 μM. Compounds were incubated with 20% (v/v) NHS and sensitized sheep red blood cells (EA cells), and cell lysis was measured by optical density in a microplate reader. Of the 11 compounds, compound -98, and to a lesser extent compound -93, showed dose-dependent inhibition of complement-mediated lysis of RBCs (**Fig. 35B**). Dose-response of compounds -93 and -98 were further evaluated by classical pathway-specific ELISA in which concentrations of the two compounds were serially diluted over a range of 100 μM – 15.6 μM concentrations and C4c deposition quantified. A half maximal inhibitory concentration (IC_{50}) was determined to be 16 μM for compound -98 as determined by fitting dose-response curves using inhibitor vs. response models in GraphPad Prism 8; however, an IC_{50} for compound -93 could not be determined (**Fig. 35C**).

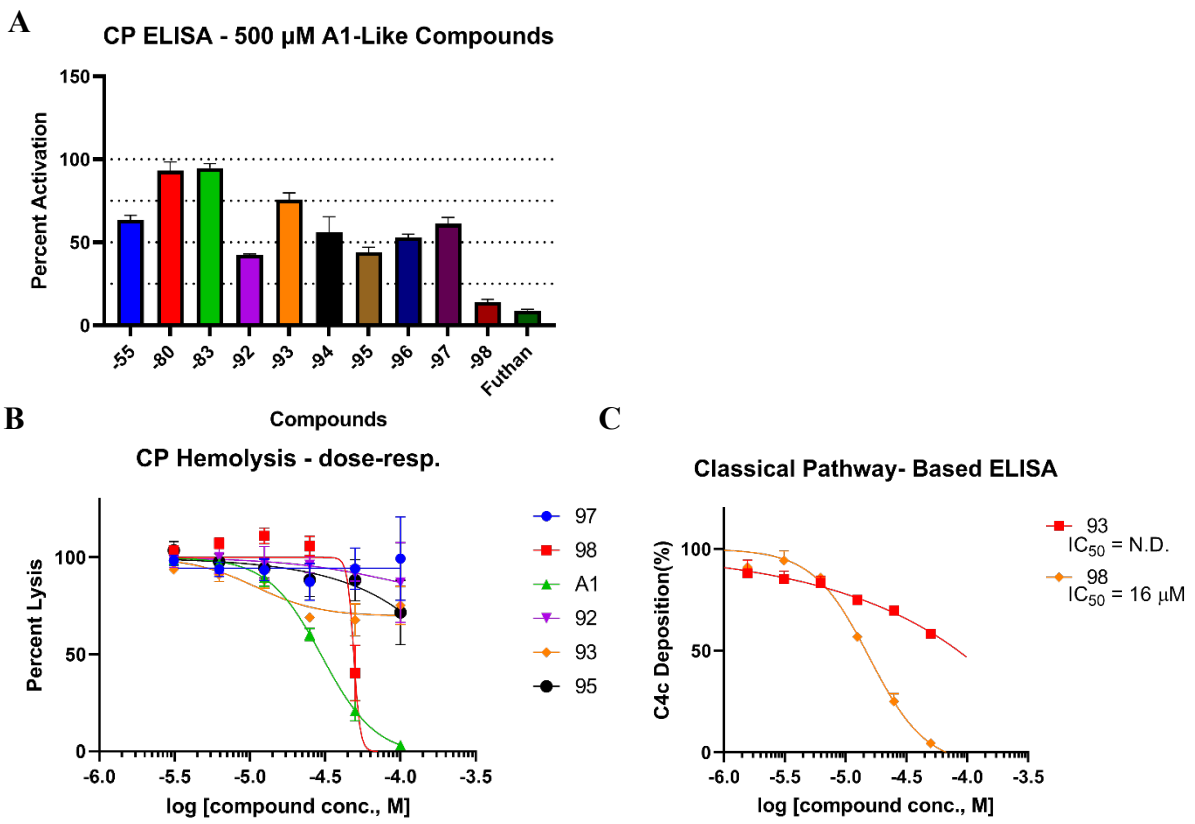


Figure 35. Evaluation of A1-like newly synthesized compounds in classical pathway-specific assay systems. A) Eleven newly synthesized analogs were evaluated for classical pathway inhibition at 100 µM concentrations. The percentage of pathway inhibition is quantified by C4c deposition. **B)** Dose-dependent activity of analogs was tested for inhibition of complement-mediated lysis in a classical pathway hemolysis assay. **C)** Two compounds, -93 and -98, were selected for further evaluation of dose-dependent inhibition of the classical pathway through a pathway-specific ELISA utilizing 2% (v/v) NHS. All experiments were performed in triplicate and the IC_{50} value of compound -98 calculated by fitting dose-response curves using an inhibitor vs. response model.

To isolate activity specifically against C1s, compounds -93 and -98 were tested for dose-dependent inhibition in C1s-mediated C2 cleavage assays. The assays were performed with two-fold serially diluted concentrations of inhibitor ranging from 100 µM to 1.56 µM, which were incubated with 6.25 nM C1s prior to incubation with 1 mg ml^{-1} C2. After reactions had incubated for 1 hour at 37º C, they were halted and run on representative SDS-PAGE gels (**Fig. 36 A, C**). The presence of C2b indicated cleavage by C1s, whereas the intact C2 band represented C1s inhibition. From the dose-response curves, IC_{50} values were calculated as 30 µM and 48 µM for compound -93 and -98, respectively (**Fig. 36 B, D**). These results show similar analog bioactivity and IC_{50} values to that of parent compound A1 (IC_{50} = 84 µM).

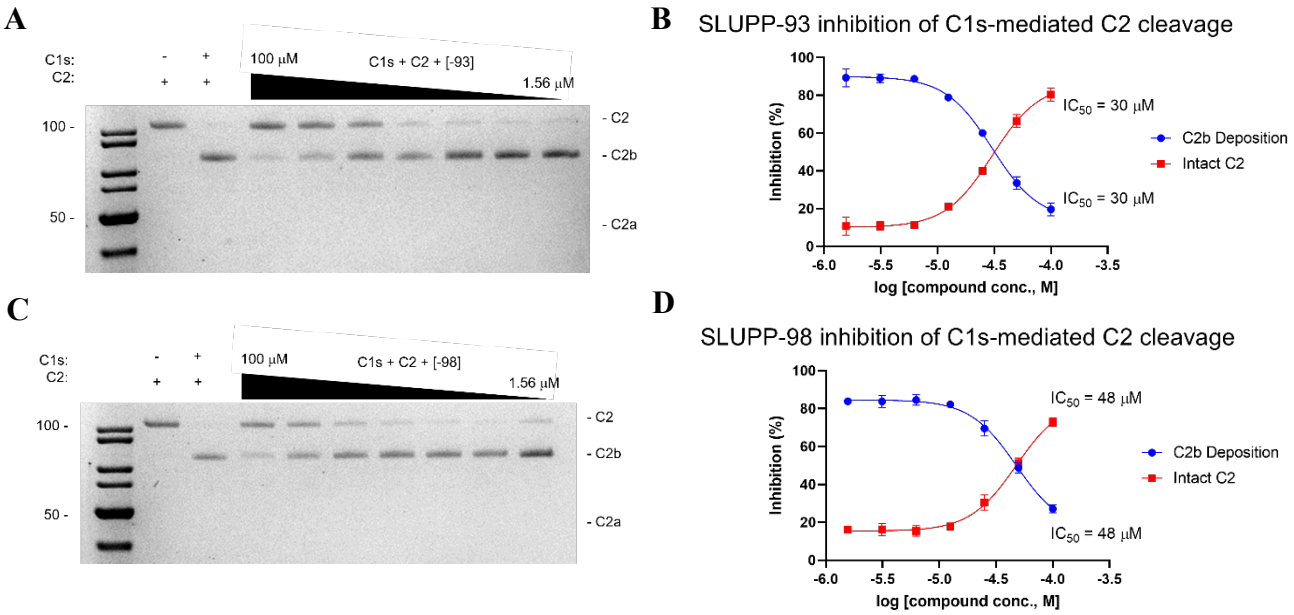


FIGURE 36. Characterization of two A1-like analogs. A) A representative SDS-PAGE gel showing dose-dependent C1s inhibition by analog 93. Intact C2 bands represent C1s inhibition, whereas the presence of C2b and C2a bands represent C1s proteolytic activity against substrate C2. Concentrations of analog 93 ranged from 100 µM to 1.56 µM and were incubated with 6.25 nM full-length C1s. **B)** A representative graph displaying both C1s inhibition (red line) and activity

(blue line) in the presence of analog 93. **C)** Dose-dependent inhibition of 6.25 nM C1s by analog 98 is shown by representative SDS-PAGE. **D)** Graph displaying both C1s inhibition (red line) and activity (blue line) in the presence of analog 98. IC_{50} values were calculated by fitting dose-response curves with a non-linear regression with a log(inhibitor) vs. response model.

Discussion

Presently, the landscape of classical pathway-based therapeutics relies heavily n the development of monoclonal antibodies for the treatment of complement-mediated disorders with small molecules being underrepresented as a therapeutic treatment option. The clinical successes of anti-C1s Sutimlimab and anti-C1q ANX005 demonstrate the wide swath of classical pathway-mediated diseases that can be targeted therapeutically; however, no one therapeutic can deal with the complexities that come with the biological mechanisms underpinning disease or those that come with the heterogeneity represented by patient populations. Complications that come with diagnosis and treatment of a number of diseases associated with complement, such as Lyme disease-associated rheumatoid arthritis, neuromyelitis optica, and myesthenia gravis, resulting in delayed, missed, or misdiagnoses underscore problems faced in disease diagnosis that funnel downward into the drug development process. These challenges can be overcome to a greater degree through drug class variety that exists outside the bounds of a single drug as a treat-all cure. Therapeutic small-molecule inhibitors against C1s provide an alternative therapeutic strategy in addition to established small advantages, including increased tissue permeability and patient compliance.

Technological advancements in machine learning and AI have lead to a marked increase in small-molecule drug leads in development. Biochemical properties and pharmacometrics factored into rational drug design increase the chances of a small-molecule drug candidate advancing to clinical testing. AI structure-based technology and deep learning affords the benefit of accelerated discovery and development of novel small-molecule therapeutics.

In this study, we collaborated with the Geisbrecht lab who was partnered with pharmaceutical company, Atomwise, to identify a potent fragment hit from an in silico screen of millions of compound scaffolds for the targeted inhibition of C1s. Using structure-based virtual screening methods, crystallographic waters present in the C1s-CCP2-SP structure (PDB entry, 1ELV) the active site solvent network and essential active site residues and interactions could be mapped out to select for small-molecule compounds that could occupy a similar molecular fingerprint. From the initial screen, 96 compounds were synthesized and after detailed biofunctional characterization of the compounds, one compound named A1 emerged singularly as the most promising candidate for its ability to dose-dependently bind and inhibit the classical pathway at the level of C1s.

Future structural studies, such as X-ray crystallography, will guide efforts to elucidate the mechanisms of action and key interactions of A1. However, in lieu of an empirical protein-ligand complex structure, molecular docking studies suggest that A1 is bound within the substrate binding groove of C1s and has key hydrogen-bond interactions with catalytic Ser628. To achieve synthesis of analogs with similar or improved bioactivity, molecular docking studies were performed on newly designed analogs with the objective of retaining a related mode of action. Of 11 analogs synthesized and evaluated, two compounds (compounds -93 and -98) emerged retaining similar inhibitory capabilities to parent compound A1 and a similar mode of binding, as demonstrated in

molecular docking studies. Using this newly derived structural data, optimization to increase specificity and inhibitory characteristics can be undertaken to strengthen the biochemical profile of A1 and its analogs and may provide a future therapeutic option for the broad spectrum of classical pathway-mediated disorders.

CHAPTER 6: Investigating membrane-binding properties of lipoxygenases using surface plasmon resonance

Abstract

Lipoxygenases (LOXs) catalyze the oxidation of polyunsaturated fatty acids and synthesize oxylipin products that drive important cellular signaling processes in plants and animals. While there has been indirect evidence presented for the interaction of mammalian LOXs with membranes, a quantitative study of the molecular details of LOX-membrane interactions is lacking. Here, we mimicked biological membranes using surface plasmon resonance (SPR) sensor chips derivatized with 2-D planar lipophilic anchors (2D LP) to capture liposomes of varying phospholipid compositions that self-assemble into lipid bilayers on the SPR chip. The sensor chip surfaces were then used to investigate the membrane-binding properties of model LOX enzymes. SPR binding assays displayed reproducible and stable liposome capture to the sensor chip surface that allowed for the detailed characterization of LOX-membrane interactions. Our studies demonstrate a calcium-dependence for the membrane binding activities of coral 8*R*-LOX and human 15-LOX-2. Furthermore, our data confirm the importance of key membrane insertion loop residues in each of these LOX enzymes for membrane binding activity. Experiments utilizing model plant and human LOXs reveal differences in membrane-binding specificities. Our study establishes and validates a robust SPR-based platform using 2D LP sensor chips that allows for the detailed study of LOX-membrane interactions under different experimental conditions, including altered membrane compositions. Collectively, this investigation improves our overall understanding of LOX-membrane interaction properties, and our SPR-based approach holds potential for future use in the development of LOX-based therapeutics.

Introduction

Lipoxygenases (LOXs) are a family of non-heme, iron-containing enzymes that catalyze the regio- and stereospecific hydroperoxidation of polyunsaturated fatty acids (PUFAs), such as arachidonic acid and linoleic acid, to oxylipins[238]. LOXs are broadly found in plants, animals, fungi, and select bacteria, and while there are significant differences in size, sequence, and substrate-specificity between organism-specific enzymes, the overall fold and geometry of the catalytic, non-heme, iron-binding site is conserved [239]. Within plants, LOXs play important roles in seed germination, senescence, growth, development, and defense against pathogens [240]. In humans, LOX function presents a double-edged sword as they have known involvement in maintaining homeostasis and the pro-resolution response, as well as being implicated in a wide range of chronic inflammatory pathologies and cancers in a similar manner as to what is observed with the dualities associated with complement in health and disease [241–244]. Unfortunately, development of therapeutics targeting human LOXs has lagged behind cyclooxygenases (COX) with only one LOX-targeting drug, Zileuton, gaining FDA approval [245]. Successful drug design efforts have been further complicated by factors such as protein allostery – including that from lipid interactions and membrane association – that influence the structure and catalytic activity of mammalian LOXs [242,246]. Accordingly, a deeper understanding of protein function and substrate specificity is required for the initial stages of hit-to-lead generation and the downstream development of a strong drug candidate.

Plant and mammalian LOXs are often monomeric, two-domain single polypeptide enzymes. The domains include a structurally conserved, primarily alpha helical C-terminal catalytic domain and an N-terminal, C2-like Polycystin-1, Lipoxygenase, Alpha-Toxin (PLAT) domain. The latter has been shown to have Ca^{2+}-binding sites that confer Ca^{2+}-dependent membrane binding

[247,248]. There is increasing evidence for interactions of LOXs at the membrane interface, mediated by the binding of Ca^{2+} or small molecules, that influence their reactivity and/or product distribution [249,250]. In select cases, such as human 5-LOX involved in leukotriene biosynthesis, LOXs are expected to be fully active only at the membrane interface [242,251]. Other non-human, mammalian lipoxygenases also show membrane-dependent activation properties [247,251–253]. Thus, further information about the molecular interactions of the respective LOXs with the membrane is needed to complement solution structure-function studies.

Synthetic membrane environments using liposomes offer a tractable experimental approach to studying protein-membrane interactions as their lipid composition can be explicitly defined. For example, liposomes can be generated consisting of phosphatidylcholine lipid, 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (PC), which is a primary component of the extracellular leaflet [254], or modified to include other lipids that are enriched at the intracellular leaflet such as phosphatidylethanolamine (PE) lipids, phosphatidylinositol lipids (PI), and negatively charged phosphatidylserine (PS) lipids [254]. Qualitative measures of phospholipid interactions with LOXs have included size exclusion detected co-elution with biomimetic membranes, enzyme activity measurements in the presence of liposomes or nanodiscs, or immunofluorescence studies showing localization of LOXs, via the PLAT domain, to the membrane [252,255,256]. SPR has become a gold standard for characterizing biomolecular interactions as it allows for the real time measurement of binding affinity and kinetics [257,258].

SPR sensor chips offer varying matrices grafted to a metallic surface, which provide the capacity to bind a variety of ligands. Previous studies utilizing SPR sensor chips for liposome immobilization have most frequently used Biacore hydrophobic association (HPA) chips, consisting of an octadecane-thiol self-assembled monolayer surface; streptavidin-coated sensor

chips (SA sensor chip) that rely on tagging ligands with biotin moieties; and L1 sensor chips, consisting of a carboxymethylated dextran surface modified with lipophilic groups that are capable of inserting into and non-covalently capturing three-dimensional lipid vesicles [254,257]. In this study, we utilized sensor chips derivatized with planar, two-dimensional lipophilic anchors linked to a carboxymethyl dextran layer (2D LP, XanTec bioanalytics GmbH) that allow for the fusion of purified, extruded liposomes with the sensor chip surface and the formation of planar phospholipid bilayers, thus mimicking a cell membrane [258]. We then used this system to study the interaction of model LOX enzymes with membranes composed of various lipid mixtures and under different experimental conditions. These results improve our understanding of the membrane-binding properties of LOX enzymes and highlight the utility of this SPR-based assay platform for the study of LOX function and in the future development of LOX-targeted therapeutics.

Materials and Methods

Structural modeling of membrane-bound 8R-LOX

The Positioning of Proteins in Membranes (PPM) v3.0 Web Server of the Orientations of Proteins in Membranes Database (OPM), to visualize coral 8*R*-LOX putative interactions with a membrane [259]. The primary crystal structure of wild-type 8*R*-LOX (PDB: 2FNQ, chain A) was used as input using the PPM v3.0 server to model plasma membrane binding. Figures were made using The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

LOX expression and purification

Coral 8*R*-LOX was expressed in *E. coli* BL21(DE3) cells with autoinduction ZYM-5052 media. Cells were lysed and the N-terminal His_6 tagged protein was purified by Ni-NTA affinity

chromatography as described previously [253]. In a final purification step, 8*R*-LOX was purified to homogeneity by a HiLoad 16/600 Superdex 200 size exclusion column using 50 mM HEPES (pH 7.4), 150 mM NaCl buffer. The protein was flash frozen and stored at -80°C until use. An 8*R*-LOX mutant with a loop deletion (LM; Δ41-45:GS amino acids) was also prepared in a similar manner to 8*R*-LOX WT. Human 15-LOX-2 was expressed in *E. coli* BL21(DE3) Rosetta 2 cells with 2xYT [248]. These the N-terminal His_6-tagged proteins were purified by a combination of Co-HisTrap and SEC in a similar manner to 8*R*-LOX. Soybean lipoxygenase (SLO) was expressed in and purified from BL21 (DE3) Codon Plus bacterial cultures using a series of ion exchange columns, as described previously [260]. All proteins were characterized by SDS-PAGE for purity.

Liposome preparation

Liposomes were prepared from solid forms of the appropriate phospholipids, purchased from Avanti Polar Lipids (Birmingham, AL). For this study, all phospholipids contained unreactive 16:0 (palmitic acid) and 18:1 (oleic acid) fatty acids. The dried phospholipids were dissolved in 50 mM Tris pH 8 buffer at 50-100 mM concentrations. The solutions were agitated by a vortex mixer and maintained above the phospholipid transition temperatures. For phospholipid mixtures, PC:X (where X = PE, PI, or PS) ratios were constructed at 2:1 (m/m). To maintain a similar size, liposome solutions were extruded prior to SPR immobilization.

Surface plasmon resonance instrument and reagents

All experiments utilized T200 style XanTec planar lipophilic anchor 2D LP chips (XanTec bioanalytics GmbH) which are carboxymethyl dextran surfaces coupled to lipophilic alkyl chains. SPR experiments were conducted on a Biacore T200 at 25ºC at a flowrate of 20 µL/min unless otherwise noted. The instrument was first cleaned with 0.5% SDS followed by 50 mM glycine pH

9.5. The instrument was then equilibrated in double distilled water overnight in standby mode. For all experiments utilizing Ca^{2+}, a running buffer of 10 mM HEPES pH 7.3, 140 mM sodium chloride, 0.5 mM EDTA, and 2 mM calcium chloride was used. For experiments performed in the absence of calcium, the instrument was cleaned once more with 0.5% SDS and 50 mM glycine, equilibrated overnight with double distilled water, and performed using a running buffer of 10 mM HEPES pH 7.3, 140 mM sodium chloride, and 0.5 mM EDTA.

Liposome capture and regeneration protocol

Prior to each experiment, the surface of the sensor chip was regenerated with two 20-second injections of 50 mM sodium hydroxide, one 20-second injection of 40 mM CHAPS, and a single 20-second injection of 50 mM sodium hydroxide. Freshly prepared 1 mM liposome solution was injected over the surface for 8 minutes (unless otherwise noted), at a flowrate of 20 µL/min, followed by a 20-second injection of 50 mM sodium hydroxide to remove loosely bound lipids and stabilize the baseline. All liposome capture experiments were performed in triplicate using PC, 2:1 (m/m) phosphatidylcholine:phosphatidylserine (PC:PS), 2:1 (m/m) phosphatidylcholine:phosphatidylethanolamine (PC:PE), and 2:1 (m/m) phosphatidylcholine:phosphatidylinositol (PC:PI). Running buffer was flowed over the surface for two minutes to ascertain baseline stability before injecting LOX analytes. Flow cell 1 was used as a reference cell (*i.e.*, blank/no ligand). Immobilized lipids were removed, and the surface regenerated following a three-minute injection of 40 mM CHAPS and a one-minute injection of 50 mM sodium hydroxide. All sensorgrams were reference and blank subtracted and analyzed using T200 Evaluation Software. Graphpad Prism v9 was used to prepare figures.

SPR binding experiments

To test for differential binding of the LOX variants at a single concentration, LOX enzyme solutions were diluted to a concentration of 500 nM in running buffer and triplicate injections were performed over each liposomal surface. 8*R*-LOX and 15-LOX-2 proteins were tested in the presence and absence of calcium in the running buffer. Dose-dependent binding experiments were subsequently carried out for all 8*R*-LOX proteins across all four liposome surfaces, or for SLO and 15-LOX-2 for select surfaces, in triplicate using a two-fold dilution series ranging from 500 nM – 15.6 nM and a flow rate of 20 μL/min. The impact of surface immobilization density was also measured for 8*R*-LOX by capturing PC liposomes for 30 seconds. Equilibrium dissociation constants were calculated using Biacore T200 evaluation software and steady-state affinities for each LOX variant were calculated by fitting sensorgrams from each variable concentration injection dataset using both 1:1 binding kinetic fit and steady state affinity models, respectively. The sensor chip surface was regenerated in between each binding experiment as described above and a fresh liposome capture was performed.

Results

Liposomal membranes reproducibly and stably captured on 2D LP SPR sensor chips

To assess the reproducibility of liposomal capture on 2D LP sensor chips, each sensor chip was first subjected to two 20-second injections of 50 mM NaOH, one 20-second injection of 40 mM CHAPS, and a final 20-second injection of 50 mM NaOH. The following liposomes: PC, 2:1 (m/m) PC:PE, 2:1 (m/m) PC:PI, and 2:1 (m/m) PC:PS, were prepared at a final concentration of 1 mM and injected over individual flow cells at a flow rate of 20 μL/min for 480 seconds. The sensor chip surface was then washed with 50 mM NaOH for 20 seconds to remove any loosely bound lipids (**Fig. 37A**). Running buffer was allowed to flow over the sensor chip surface until a baseline

signal was achieved, as this was an important consideration for performing subsequent LOX binding experiments. All liposome surfaces were stably captured and could be reproducibly regenerated for each of the liposome compositions used (**Fig. 37B**). Biacore conventionally uses an arbitrary unit termed resonance units (RU) to measure the binding response and relative RUs were calculated by comparing final RU values after immobilization to the baseline just prior to liposome injection. Total capture for each liposome was as follows: 4190 ± 730 RU for PC, 5870 ± 460 RU for PC:PE, 5410 ± 720 RU for PC:PI, and 4410 ± 670 RU for PC:PS (**Fig. 37B**).

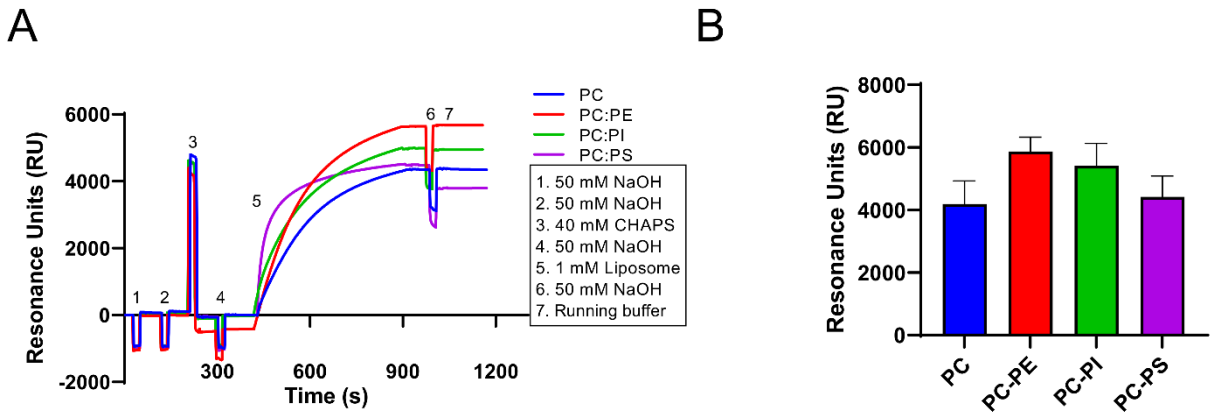


FIGURE 37. Liposome capture on a 2D LP SPR sensor chip. A) Capture of each liposome mixture began by preconditioning a Xantec 2D LP sensor chip with two 20-second washes of 50 mM NaOH, one 20-second wash of 40 mM CHAPS, and a final 20-second injection of 50 mM NaOH. Liposome mixtures (1 mM) were then injected over each flow cell for eight minutes at a flowrate of 20 µL/min. Liposome capture experiments were performed in triplicate 1 mM final concentrations of PC, 2:1 (m/m) PC:PS, 2:1 (m/m) PC:PE, and 2:1 (m/m) PC:PI. A running buffer of 10 mM HEPES pH 7.3, 140 mM sodium chloride, 0.5 mM EDTA, and 2 mM calcium chloride then flowed over the surface for two minutes to ensure baseline stability. **B)** Relative response measured in RU of final capture density for each liposome. Data are from three independent

experiments are shown and reported as the mean ± S.D. Sensorgrams were analyzed with Biacore Evaluation software and GraphPad Prism.

Modeling LOX-membrane interactions

We began the investigation of LOX-membrane interactions using a model animal LOX, the coral 8*R*-LOX from *Plexaura homomalla*, which is a homolog of human 5-LOX [253]. Previous kinetic analyses have demonstrated that 8*R*-LOX activity is stimulated in the presence of Ca^{2+} and liposomes [247]. The N-terminal PLAT domain of 8*R*-LOX possesses three crystallographically-resolved calcium-binding sites believed to promote membrane interactions between the enzyme and lipid bilayer [247]. We modeled the interaction of 8*R*-LOX with a eukaryotic plasma membrane using PPM v3.0 web server [261], which provides a computational approach to predict the association of proteins at a membrane surface (**Fig. 38A, B**). Using the published crystal structure of wild-type 8*R*-LOX (PDB:2FNQ, chain A) [247] as input, a favorable transfer energy ($\Delta G_{transfer}$ = -7.5 kcal/mol) was calculated and lipid contacts were formed by hydrophobic loop residues 43-44 and 78-79. This computational model is consistent with the importance of PLAT domain loop residues and Ca^{2+} in the membrane-binding activity of 8*R*-LOX.

Lipid membrane binding by 8R-LOX is calcium-dependent

We next sought to test aspects of this model using 2D LP SPR sensor chips. In consideration of phospholipid interactions, PC is among the most highly abundant lipids found to compose lipid bilayers in plasma membranes, especially that of the extracellular leaflet. Due to its high stability, PC is also widely used as a model lipid system for liposome studies [254,262]. Given these

considerations, we chose 8*R*-LOX/PC as an initial model system for testing the binding properties of LOXs to membranes by SPR. Following capture of PC liposomes (**Fig. 37**), 500 nM 8*R*-LOX was injected in triplicate over the PC surface in a running buffer containing 2 mM Ca^{2+} for 2 minutes and dissociation was monitored for 3 minutes. A strong binding response was observed as shown in the representative sensorgram in **Fig. 38C** (blue). In contrast, when an identical injection was performed in a buffer that was supplemented with excess EDTA to remove free Ca^{2+}, no detectable binding response with the PC surface was observed (**Fig. 38C,** green).

Loop mutations ameliorate membrane binding by 8R-LOX

The relationship of Ca^{2+} to membrane interaction by 8*R*-LOX is proposed to arise through the binding of Ca^{2+} ions that stabilize key loops in the PLAT domain of 8*R*-LOX, including the membrane insertion loops (**Fig. 38A, B**). To directly test the importance of these residues, an 8*R*-LOX variant was produced in which this loop was deleted by mutagenesis (*i.e.*, Δ41-45:GS). Triplicate injections of 500 nM 8*R*-LOX loop-deletion mutant (8*R*-LOX-LM) were performed on a PC-surface in the presence and absence of Ca^{2+}. No binding response was observed for 8*R*-LOX-LM in buffers supplemented with Ca^{2+} or EDTA, which strongly suggests the structural and functional importance of these loop residues in 8*R*-LOX membrane-binding activity (**Fig. 38C, D**).

8R-LOX displays dose-dependent membrane binding in the presence of calcium

Next, we assessed the dose-dependence of 8*R*-LOX binding to PC biosensors in the presence of Ca^{2+} by injecting a two-fold serial dilution of 8*R*-LOX ranging from 500 nM to 15.6 nM (**Fig. 39A**). 8*R*-LOX exhibited clear dose-dependent interactions over PC surfaces (**Fig. 39A**). Equilibrium dissociation constants using a steady-state binding affinity fit model ($K_{D,\ s.s.}$) were

calculated using BiaEvaluation Software for three independent injection series with an average value of $K_{D,\ s.s.}$ = 260 ± 58 nM (**Fig. 39A**). However, attempts to fit these data to a kinetic 1:1 Langmuir model of binding produced poor fits as judged by residual plots and elevated chi-squared values.

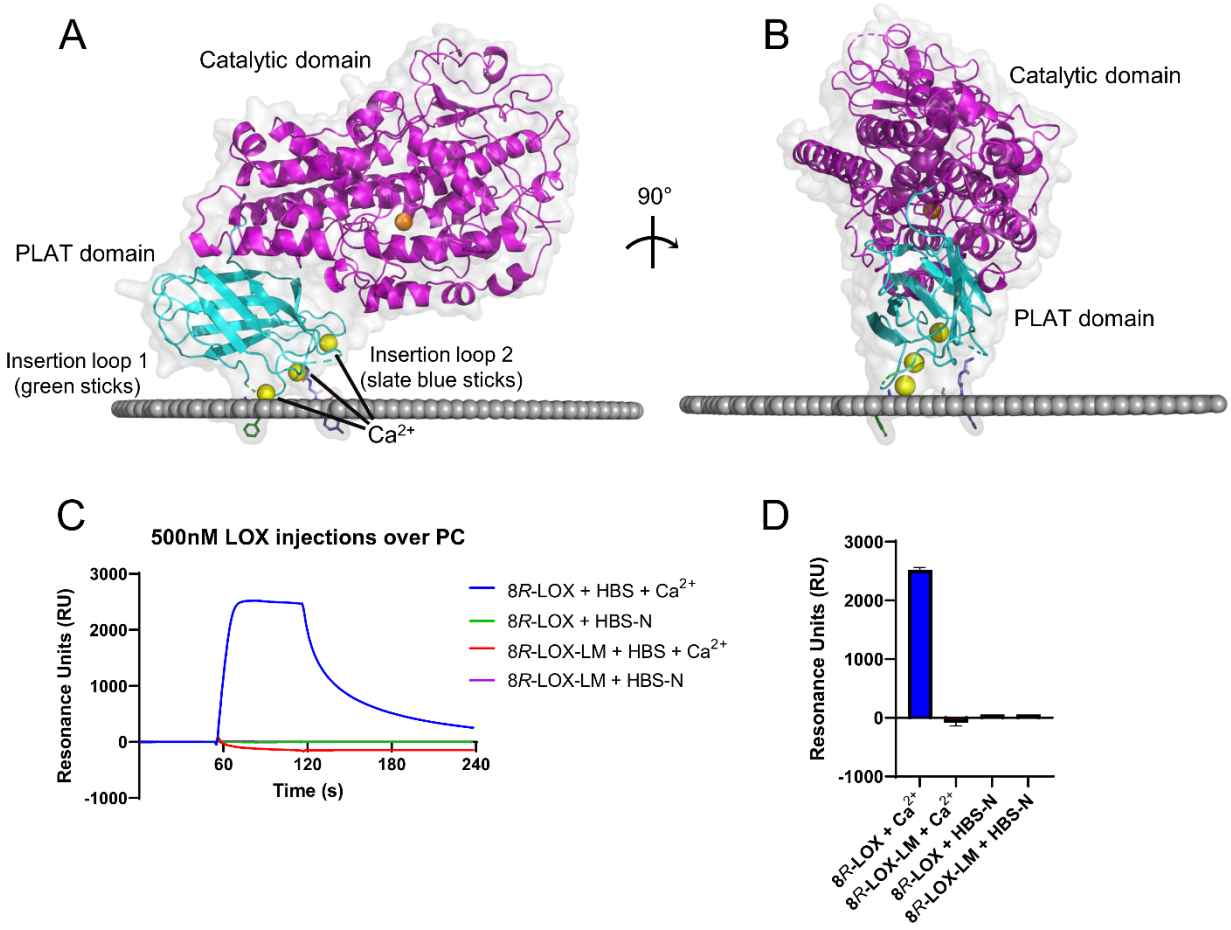


FIGURE 38. 8*R*-LOX membrane binding. **A-B**) The 8*R*-LOX crystal structure (PDB: 2FNQ, chain A) was used as input for the PPM v3.0 web server to model its interaction with a membrane. 8*R*-LOX is drawn with secondary structure shown in cartoon representation and the hydrophobic leaflet shown as grey rods. The catalytic domain is colored magenta, and the catalytic iron is displayed as an orange sphere. The PLAT domain is colored cyan and two membrane insertion

loops as green and slate blue. Calcium ions are displayed as yellow spheres. **C)** Calcium-dependent membrane binding, as well as loop-dependent membrane binding, was measured by injecting 500 nM concentrations of control 8*R*-LOX and 8*R*-LOX with a loop-deletion mutation (8*R*-LOX-LM) over the PC surface using running buffer in the presence of free calcium (HBS+Ca^{2+}) or absence (HBS-N) in three independent injection series. A representative SPR sensorgram is shown. 8*R*-LOX binds the PC surface only in the presence of calcium and with the insertion loop intact. **D)** The relative binding response of 8*R*-LOX and 8*R*-LOX-LM triplicate 500 nM injections in the presence or absence of Ca^{2+} are represented in GraphPad by bar graph and shown as mean ± S.D.

Determining the effect of PL capture density on 8R-LOX binding

Mass transport limitations may impact the ability to accurately measure rate constants in SPR experiments [263]. In these cases, the contribution of mass transport kinetics can be reduced by utilizing lower surface densities and/or higher flow rates; however, incomplete membrane coverage of the sensor chip may conceivably impact specific binding responses to the sensor chip surface, especially as it relates to lipid density within the lipid bilayer [264]. To test for these possibilities directly, the 8*R*-LOX/PC dose-response binding assays were repeated on a lower density PC surface and at an increased flow rate of 50 μL/min. A lower density surface was prepared by injecting a short 30-second pulse of freshly prepared 1mM PC liposome solution (**Fig. 39B**). This yielded a capture density of 854 RU, or ~12% of the average capture density for our standard PC surfaces (**Figs. 37B**, **39A**). Steady-state K_D values were calculated for a triplicate injection series, $K_{D,\,s.s.}$ = 180 ± 32 nM (**Fig. 39C**), which was comparable to the value calculated for the higher-density PC surfaces (*i.e.*, $K_{D,\,s.s.}$ = 260 ± 58 nM) (**Fig. 39A**). When all sensorgrams

were simultaneously fit to a kinetic 1:1 model, the calculated rate constants were (k_a = 7.96 x 10^5 $M^{-1}s^{-1}$ ± 1.70 x 10^4 (S.E.)) and (k_d = 0.086 s^{-1} ± 9.40 x 10^{-4} (S.E.)), which were well fit and within the dynamic range of the instrument. These microscopic rate constants result in a calculated dissociation equilibrium constant of $K_{D, kin} = k_d / k_a$ = 110 ± 2.6 (S.E.) nM, which is similar to that calculated using the steady-state fitting approach on this surface.

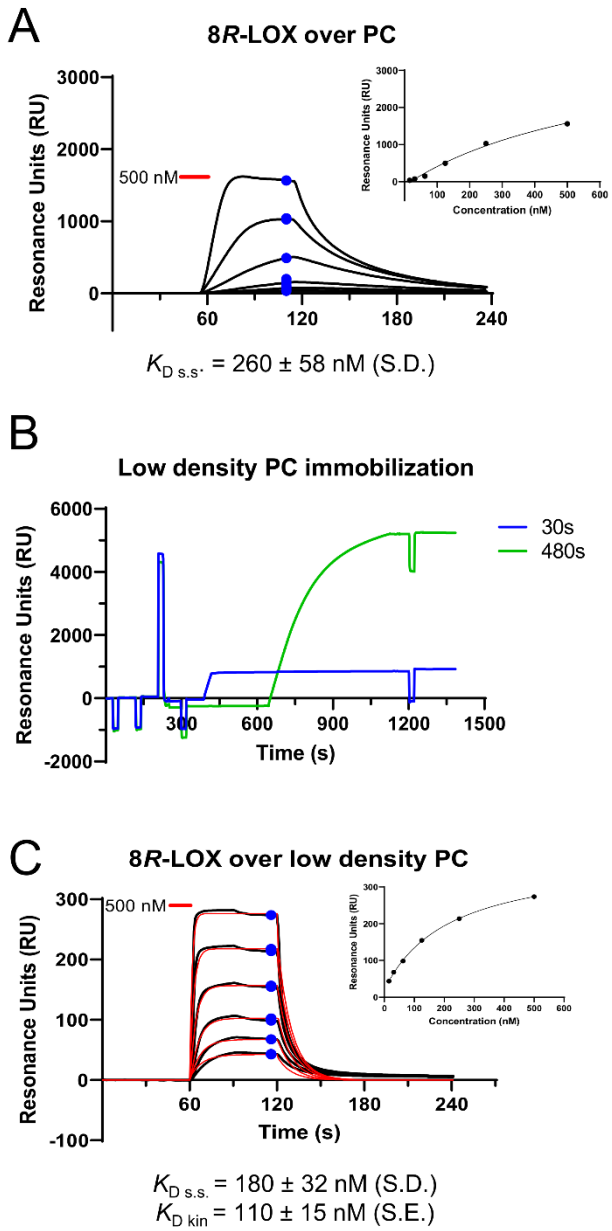


FIGURE 39. 8*R*-LOX WT dose-response over PC immobilized biosensors. A) Dose-dependent binding of 8*R*-LOX to membranes of PC lipid composition was measured by SPR using a two-fold variable concentration series ranging from 500 nM to 15.6 nM over. K_D values were calculated using steady-state analysis ($K_{D,\ s.s.}$) with Biacore Evaluation software and reported as the mean ± S.D. Blue circles indicate the measured value for steady-state response and are plotted on the inset concentration vs. response diagram. **B)** PC liposomes were captured at "low" density by using a 30-second injection of 1 mM PC (green) and standard density using a 480-second injection. **C)** 8*R*-LOX was injected over the low-density surface using a two-fold variable concentration series ranging from 500 nM to 15.6 nM in triplicate injection series. Data were fit to a steady-state model of binding ($K_{D,s.s.}$**)** using Biacore Evaluation software. Blue circles indicate the measured value for steady-state response and are plotted on the inset concentration vs. response diagram. Data were also fit to a kinetic model of binding ($K_{D,\ kin}$**)** with fits shown as red traces.

8R-LOX displays similar binding preferences across differing lipid membrane compositions

Previous work has shown that certain LOXs display a preference for specific lipid membrane compositions due to the electrostatic and hydrophobic interactions that occur at the interface of the membrane contact plane of the enzyme [255,265,266]. To investigate the membrane interactions between 8*R*-LOX and alternative liposome compositions, we performed identical dose-dependent SPR binding assays of variable 8*R*-LOX concentrations to flow cells immobilized with PC:PE, PC:PI, and PC:PS. We note that original capture densities and steady-state evaluations were utilized to facilitate these comparative studies. Interestingly, 8*R*-LOX showed similar binding affinities across these three surfaces, as assessed by calculated steady-state

K_D values ($K_{D,\ s.s.,\ PC:PE.}$ = 280 ± 42 nM , $K_{D,\ s.s.,\ PC:PI}$ = 360 ± 71 nM, and $K_{D,\ s.s.,\ PC:PS}$ = 210 ± 22 nM) (**Fig. 4 A-C**).

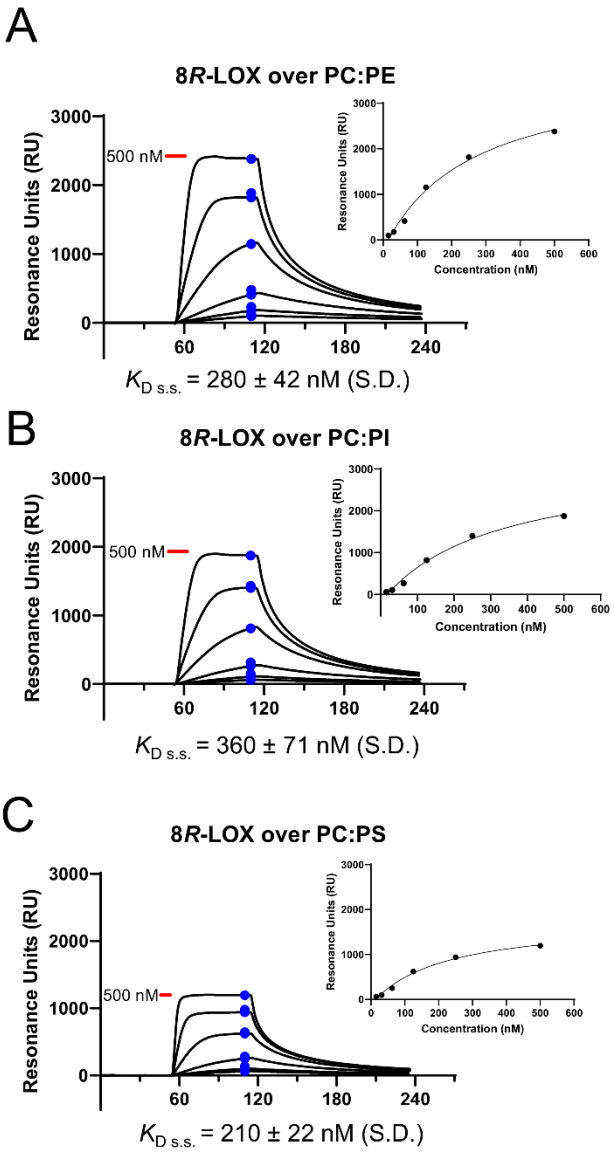


FIGURE 40. Binding activity of 8*R*-LOX to different lipid membranes. Dose-dependent binding of 8*R*-LOX to membranes of different lipid compositions was measured by SPR using a two-fold variable concentration series ranging from 500 nM to 15.6 nM over **A)** PC:PE, **B)** PC:PI, and **C)** PC:PS membrane surfaces in a running buffer containing free calcium. K_D values were

calculated using steady-state analysis ($K_{D,\ s.s.}$) with Biacore Evaluation software and reported as the mean ± S.D. Blue circles indicate the measured value for steady-state response and are plotted on the inset concentration vs. response diagram.

Evaluation of human 15-LOX-2 and soybean SLO membrane-binding by SPR

Having established that an SPR-based liposome capture method could be utilized to study a model animal LOX (*i.e.*, 8*R*-LOX), we next evaluated other important model LOX enzymes. We selected the human isozyme, 15-LOX-2, for its implications in a number of inflammatory disease pathologies, including atherosclerosis [248,255], and soybean lipoxygenase-1 (SLO) as a model plant LOX [267], and ran parallel binding assays to those of 8*R*-LOX to serve as a comparison between LOXs across species.

We first assessed dose-dependence of 15-LOX-2 with PC and PC:PS by injecting two-fold serial dilutions ranging from 500 to 15.6 nM of 15-LOX-2 at a flow rate of 20 µL/min over PC:PS immobilized to the sensor chip surface (**Figs. 41A**). 15-LOX-2 exhibited dose-dependent binding interactions and the steady-state binding affinity fit was calculated using BiaEvaluation Software for three independent injection series over the PC surface ($K_{D\ s.s.}$ = 1,400 ± 200 nM) and the PC:PS surface ($K_{D\ s.s.}$ = 630 ± 190 nM). These results are consistent with previous reports that 15-LOX-2 binds to PC and PS lipids [265]. Similar to 8*R*-LOX, 15-LOX-2 is also expected to exhibit Ca^{2+}-dependence for membrane interactions that involve residues from a key membrane insertion loop [268]. Using a PC:PS surface, we tested this using single 500 nM concentration injections in the presence or absence of Ca^{2+}. Injections without Ca^{2+} completely abrogated the binding response

(**Fig. 5B**). Furthermore, a variant of 15-LOX-2, in which a key membrane insertion loop is mutagenized (*i.e.*, 15-LOX-2-LM), also exhibited no detectable binding response (**Fig. 41B**).

Next we examined the membrane association of a plant LOX, SLO, which is also a 15-LOX [238], by assessing dose-dependent binding to a PC:PS surface. Unlike 15-LOX-2 and 8*R*-LOX, SLO did not exhibit dose-dependent binding, and an equilibrium dissociation constant could not be determined under the concentration range employed (**Fig. 41C**). These results are in line with the expectation that plant LOXs do not associate strongly with membranes [269]. Collectively, these experiments demonstrate that like 8*R*-LOX, 15-LOX-2 binds directly to PC:PS membranes in a Ca^{2+}-dependent manner and requires residues located on a loop in the PLAT domain. Furthermore, while dose-dependent binding is observed for 8*R*-LOX and 15-LOX-2, the plant LOX SLO does not exhibit dose-dependent interactions with membranes under the same conditions.

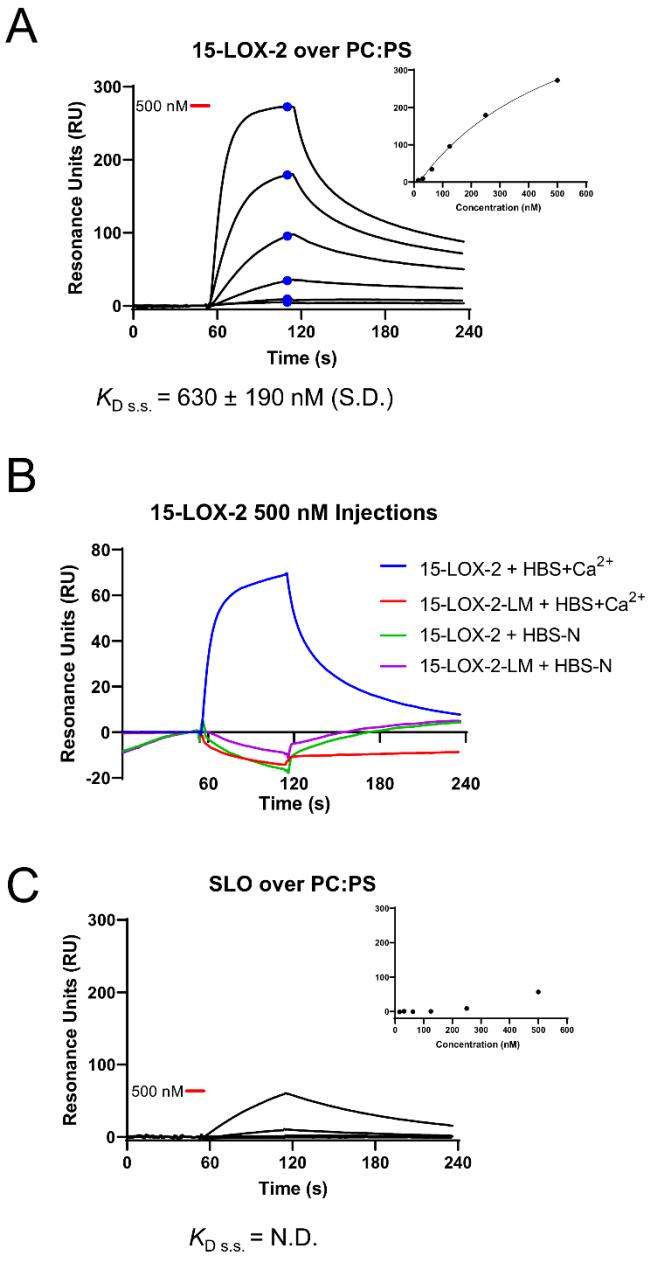


FIGURE 41. Human 15-LOX-2 and soybean SLO binding activity over PC:PS surface. A) Dose-dependent binding of 15-LOX-2 to a standard-capture PC:PS surface was measured by SPR using a two-fold variable concentration series ranging from 500 nM to 15.6 nM. K_D values were calculated using steady-state analysis ($K_{D,\ s.s.}$) with Biacore Evaluation software and reported as the mean ± S.D of three independent injection series. A representative sensorgram is shown. Blue

circles indicate the measured value for steady-state response and are plotted on the inset concentration vs. response diagram. **B)** Binding of 15-LOX-2 to the PC:PS surface is Ca^{2+}- and loop-dependent, as indicated by 500 nM injections of wild-type 15-LOX-2 and 15-LOX-2-LM that contains a loop-deletion mutation. Triplicate independent injection series were carried out. A representative SPR sensorgram is shown. **C)** SPR response for SLO association over PC:PS membrane using the same concentration range for 15-LOX-2. $K_{D,\ s.s.}$ values could not be determined (N.D.) for SLO (inset).

Discussion

Liposomes composed of phospholipids naturally found within cellular membranes provide a simplified model of biological membranes. In this study, we immobilized liposomes on the surface of 2D LP SPR sensor chips to allow for the quantitative characterization of LOX enzymes with membranes. Several LOXs have been reported to depend on the presence of Ca^{2+} for membrane interaction [247,266,270]. Consistent with this, our SPR experiments showed that 8*R*-LOX required Ca^{2+} for interaction with liposome captured membranes on the 2D LP sensor chip (**Fig. 38C, D**). The PPM-predicted interaction of 8*R*-LOX with the membrane (**Fig. 38A, B**) exhibits a $\Delta G_{transfer}$ of -7.5 kcal/mol, which is a similar binding energy observed for the homologous human 5-LOX [271]. Our data confirm the importance of loop residues in 8*R*-LOX-membrane binding as 8*R*-LOX-LM showed no quantifiable membrane interactions with the liposome-captured biosensors (**Fig. 38C, D)**.

We found that 8*R*-LOX interacted with all four membranes tested to a comparable degree as judged by steady-state affinities (**Fig. 40**). A previous Forster resonance energy transfer (FRET)

assay of WT 8*R*-LOX on PC:PS liposomes estimated a K_D value of 11 ± 3 μM [247]. This indirect fluorescence-based method relies on the quenching of the dansyl fluorophore (embedded in the liposomes) by the interaction of tryptophan sidechains in the putative membrane binding loop. The difference in affinities reported here could reflect the altered shape of the biomimetic membranes (convex vs flat) and the challenge in controlling the liposome size and fluorophore loading/distribution. While the overall magnitude of the K_D varied, we note that the FRET technique also showed similar mutational trends in abolishing membrane association that were observed here. Overall, our study of LOX-membrane interactions demonstrates that this SPR method can be applied to the general study of LOX-membrane interactions, even in cases where Trp residues are not located in or adjacent to membrane loops.

We found that capture of lower membrane densities and use of higher flow rates may better facilitate kinetic measurements for SPR-based LOX-membrane-binding assays (**Fig. 39B, C**). Further optimization could likely also be achieved in certain LOX systems by identifying an inert liposomal capture surface to serve as the reference surface, rather than the blank reference cell used here. Despite these limitations, we demonstrated the utility of the SPR-based system employed here in: i) quantitatively characterizing LOX-membrane interactions; ii) evaluating the requirement for Ca^{2+} in these interactions; and iii) probing the importance of specific LOX residues in promoting membrane binding. To this point, experiments carried out for human 15-LOX-2 confirmed both Ca^{2+}-dependence and the importance of loop residues for membrane binding (**Fig. 41B**).

In addition to being useful for elucidating the molecular requirements of LOX enzymes for direct membrane association, the SPR method utilized here could be readily adapted to high-throughput screening formats that may be useful in drug discovery efforts of LOX-based

therapeutics. To date, a number of clinically available therapeutics have utilized SPR for rapid quantification of binding interactions between target protein and inhibitor, as well as protein-substrate binding kinetics in the presence of inhibitor [272]. In the LOX field, SPR studies have previously supported the development of several small-molecule inhibitors targeting human LOXs to target their role in the synthesis of lipid oxidation products involved in exacerbation of chronic inflammatory disorders and cancers, such as rheumatoid arthritis, asthma, pancreatic cancer, among many others [273–275].

In conclusion, we have investigated the utility of 2D LP sensor chips for the purpose of studying the membrane-binding activities of lipoxygenases. These sensor chips allowed for a reproducible method by which to capture liposomes of different phospholipid composition and provided a system to quantitatively study the direct interaction of lipoxygenases with lipid membranes. Our results confirm a clear role of Ca^{2+}-dependence in the membrane-associated activity of 8*R*-LOX and 15-LOX-2 and demonstrate a requirement for key loop residues in the PLAT domain for direct membrane binding. The use of 2D LP sensor chips provides a robust platform that may aid future efforts to better understand the membrane-associated enzymatic activities of LOXs and the impact of small molecule effectors or inhibitors to support the development of novel LOX-targeted therapeutics.

CHAPTER 7: DISCUSSION AND SUMMARY OF THESIS

Enzyme cascades play critical physiological roles in a number of biological processes involved in digestion, coagulation, fibrinolysis, metabolism, and immunity [72,146], and whose ability to crosstalk with other host pathways make them instrumental in maintaining homeostasis. Dysregulation of the proteins involved in these enzyme cascades can lead to catastrophic consequences for the host that cover a wide array of pathologies. The complement system is one such biological system with heavy involvement in both innate and adaptive immune responses that can drive host tissue damage associated with a growing number of diseases involved in autoimmunity, inflammatory disorders, and neurodegeneration. Due to the diverse assortment of pathway activators, understanding the molecular mechanisms guiding complement activation is a necessary element to developing therapeutic options to combat aberrant complement activation.

The initiating serine proteases of the classical and lectin pathways have been extensively studied; however, the intricate molecular details involved in their autocatalytic behavior are not well understood. Using the structural information gleaned from the X-ray crystallographic enzyme-product complex of self-associated MASP-2, we were able to glean insights into the catalytic mechanisms governing autoactivation. These structural details help provide a framework from which to identify target elements present on and within the proteases that could be leveraged for structure-based drug discovery.

High level knowledge of the underpinnings of serine protease biology is instrumental in small-molecule drug design as the greatest hurdle in the preclinical stages of drug development is accounting for target specificity. The molecular binding events and conformational changes that occur within the active site can be targeted for ligand interactions and, consequently, inhibition of the protease when designing competitive small-molecule inhibitors. In this study, we identified a

compound hit predicted to bind within the active site of C1r and inhibit activation of the classical pathway. Using molecular dynamics simulations, we designed analogs of the parent compound to further elucidate SAR to guide future optimization efforts designed to increase binding affinity and specificity.

The intricacies of catalysis are also vital for understanding how concerted correlated and anti-correlated protein motions at distal sites may be involved in catalytic activity. In expanding our knowledge of protein dynamics, we can look to avenues of small-molecule inhibitor development through allostery. Our initial library screen of compounds designed to identify C1r-binding hits identified a fragment capable of noncompetitive C1r inhibition. Rational drug design through ring replacement strategies added to the SAR repertoire available for forthcoming endeavors in noncompetitive small-molecule inhibitor design.

Limitations imposed by off-target effects and low bioactivity inherent in small-molecule/fragment-based drug design can be ameliorated by leveraging machine learning and artificial intelligence to design novel inhibitors based on algorithms utilizing existing protein-inhibitor structures to identify unique biochemical fingerprints to synthesize high affinity hits curated from fragment scaffolds. In a parallel study to identify inhibitors against serine protease target C1s, we collaborated with the Geisbrecht lab at Kansas State University, which was partnered with biopharmaceutical company, Atomwise, to identify a potent small-molecule hit predicted to bind within the substrate binding groove of C1s. This compound was characterized, and analogs synthesized with the medicinal chemistry of the Walker lab of the Saint Louis University School of Medicine to develop novel leads with similar biochemical and pharmacokinetic properties of the parent compound, A1. These efforts lend themselves to the

additional optimization of A1 for the potential development of an orally bioavailable drug designed to mitigate pathologies associated to classical pathway overactivation and dysregulation.

Additional efforts were undertaken to develop a novel platform for elucidating protein function in molecules involved in other pathways with known implications in proinflammatory disorders. Collaboration with the Offenbacher lab from the department of chemistry at East Carolina University lent itself to the development of a bilayer model for which to simulate protein-membrane interactions between captured liposomes and integral membrane lipoxygenase proteins. This led to the characterization of several model lipoxygenases with the membrane bilayer, as well as provided a tool for the future screening and evaluation of small-molecule inhibitors blocking critical membrane association requisite for mammalian lipoxygenase activity.

Our work has produced high resolution mechanistic details involved in autocatalysis in the initiating proteases of the complement system and provided additional information to the field of serine protease biology. Moreover, using structure-based drug discovery approaches, we have identified three small-molecule compound hits that have activity against both C1r and C1s using different methodologies and means of inhibition, which provided for differing strategies of inhibitor design. The identification of these hits, in addition to the development of a novel SPR screening platform, has led to the generation of in-depth structural data and method systems to guide future continued optimization efforts that may one day lead to the design of a small-molecule therapeutic option for diseases marked by autoimmune and inflammatory pathologies.

References

[1] Merle NS, Church SE, Fremeaux-Bacchi V, Roumenina LT. Complement system part I - molecular mechanisms of activation and regulation. Front Immunol 2015;6. https://doi.org/10.3389/fimmu.2015.00262.

[2] Parra-Medina R, Quintero-Ronderos P, Rodríguez ÉG. The complement system 2013.

[3] Gaboriaud C, Frachet P, Thielens NM, Arlaud GJ, Quax P. The human C1q globular domain: structure and recognition of non-immune self ligands 2012. https://doi.org/10.3389/fimmu.2011.00092.

[4] Thiel S, Petersen S V., Vorup-Jensen T, Matsushita M, Fujita T, Stover CM, et al. Interaction of C1q and Mannan-Binding Lectin (MBL) with C1r, C1s, MBL-Associated Serine Proteases 1 and 2, and the MBL-Associated Protein MAp19. J Immunol 2000;165:878–87. https://doi.org/10.4049/JIMMUNOL.165.2.878.

[5] Dunkelberger JR, Song WC. Complement and its role in innate and adaptive immune responses. Cell Res 2010 201 2009;20:34–50. https://doi.org/10.1038/cr.2009.139.

[6] Kjaer TR, Le LTM, Pedersen JS, Sander B, Golas MM, Jensenius JC, et al. Structural Insights into the Initiating Complex of the Lectin Pathway of Complement Activation. Structure 2015;23:342–51. https://doi.org/10.1016/J.STR.2014.10.024.

[7] Holmskov U, Thiel S, Jensenius JC. Collectins and Ficolins: Humoral Lectins of the Innate Immune Defense. Https://DoiOrg/101146/AnnurevImmunol21120601140954 2003;21:547–78. https://doi.org/10.1146/ANNUREV.IMMUNOL.21.120601.140954.

[8] TR K, S T, GR A. Toward a structure-based comprehension of the lectin pathway of complement. Mol Immunol 2013;56:413–22. https://doi.org/10.1016/J.MOLIMM.2013.05.007.

[9] Matsushita M, Endo Y, Fujita T. Structural and functional overview of the lectin complement pathway: Its molecular basis and physiological implication. Arch Immunol Ther Exp (Warsz) 2013;61:273–83. https://doi.org/10.1007/S00005-013-0229-Y.

[10] Ricklin D, Lambris JD. Complement in immune and inflammatory disorders: therapeutic interventions. J Immunol (Baltimore, Md 1950) 2013;190:3839–47. https://doi.org/10.4049/jimmunol.1203200.

[11] Carroll MC, Isenman DE. Regulation of Humoral Immunity by Complement. Immunity 2012;37:199–207. https://doi.org/10.1016/j.immuni.2012.08.002.

[12] Ricklin D, Hajishengallis G, Yang K, Lambris JD. Complement: a key system for immune surveillance and homeostasis. Nat Immunol 2010;11:785–97. https://doi.org/10.1038/ni.1923.

[13] Klos A, Tenner AJ, Johswich K-O, Ager RR, Reis ES, Köhl J. The role of the anaphylatoxins in health and disease. Mol Immunol 2009;46:2753–66. https://doi.org/10.1016/j.molimm.2009.04.027.

[14] Merle NS, Noe R, Halbwachs-Mecarelli L, Fremeaux-Bacchi V, Roumenina LT. Complement system part II: Role in immunity. Front Immunol 2015;6. https://doi.org/10.3389/fimmu.2015.00257.

[15] Ricklin D, Reis ES, Mastellos DC, Gros P, Lambris JD. Complement component C3 - The "Swiss Army Knife" of innate immunity and host defense. Immunol Rev 2016;274:33–58. https://doi.org/doi:10.1111/imr.12500.

[16] Bulla R, Bossi F, Tedesco F. The Complement System at the Embryo Implantation Site: Friend or Foe? Front Immunol 2012;3. https://doi.org/10.3389/FIMMU.2012.00055.

[17] Boulanger LM. Immune Proteins in Brain Development and Synaptic Plasticity. Neuron 2009;64:93–109. https://doi.org/10.1016/J.NEURON.2009.09.001.

[18] Boulanger LM, Shatz CJ. Immune signalling in neural development, synaptic plasticity and disease. Nat Rev Neurosci 2004;5:521–31. https://doi.org/10.1038/nrn1428.

[19] Schafer DP, Lehrman EK, Kautzman AG, Koyama R, Mardinly AR, Yamasaki R, et al. Microglia Sculpt Postnatal Neural Circuits in an Activity and Complement-Dependent Manner. Neuron 2012;74:691–705. https://doi.org/10.1016/j.neuron.2012.03.026.

[20] Stevens B, Allen NJ, Vazquez LE, Howell GR, Christopherson KS, Nouri N, et al. The Classical Complement Cascade Mediates CNS Synapse Elimination. Cell 2007;131:1164–78. https://doi.org/10.1016/J.CELL.2007.10.036.

[21] Coulthard LG, Hawksworth OA, Woodruff TM. Complement: The Emerging Architect of the Developing Brain. Trends Neurosci 2018;41:373–84. https://doi.org/10.1016/J.TINS.2018.03.009.

[22] Rooryck C, Diaz-Font A, Osborn DPS, Chabchoub E, Hernandez-Hernandez V, Shamseldin H, et al. Mutations in lectin complement pathway genes COLEC11 and MASP1 cause 3MC syndrome. Nat Genet 2011;43:197–203. https://doi.org/10.1038/ng.757.

[23] Gorelik A, Sapir T, Haffner-Krausz R, Olender T, Woodruff TM, Reiner O. Developmental activities of the complement pathway in migrating neurons. Nat Commun 2017;8:1–12. https://doi.org/10.1038/ncomms15096.

[24] Nozaki M, Raisler BJ, Sakurai E, Sarma JV, Barnum SR, Lambris JD, et al. Drusen complement components C3a and C5a promote choroidal neovascularization. Proc Natl Acad Sci U S A 2006;103:2328. https://doi.org/10.1073/PNAS.0408835103.

[25] Kurihara R, Yamaoka K, Sawamukai N, Shimajiri S, Oshita K, Yukawa S, et al. C5a promotes migration, proliferation, and vessel formation in endothelial cells. Inflamm Res 2010;59:659–66. https://doi.org/10.1007/S00011-010-0178-4/FIGURES/7.

[26] HAUTANEN A, KESKI-OJA J. Interaction of Fibronectin with Complement Component C3. Scand J Immunol 1983;17:225–30. https://doi.org/10.1111/J.1365-3083.1983.TB00785.X.

[27] Kimura Y, Madhavan M, Call MK, Santiago W, Tsonis PA, Lambris JD, et al. Expression of Complement 3 and Complement 5 in Newt Limb and Lens Regeneration. J Immunol 2003;170:2331–9. https://doi.org/10.4049/JIMMUNOL.170.5.2331.

[28] Del Rio-Tsonis K, Tsonis PA, Zarkadis IK, Tsagas AG, Lambris JD. Expression of the Third Component of Complement, C3, in Regenerating Limb Blastema Cells of Urodeles .

J Immunol 1998;161:6819–24. https://doi.org/10.4049/jimmunol.161.12.6819.

[29] Mastellos D, Papadimitriou JC, Franchini S, Tsonis PA, Lambris JD. A Novel Role of Complement: Mice Deficient in the Fifth Component of Complement (C5) Exhibit Impaired Liver Regeneration. J Immunol 2001;166:2479–86. https://doi.org/10.4049/jimmunol.166.4.2479.

[30] He S, Atkinson C, Qiao F, Cianflone K, Chen X, Tomlinson S. A complement-dependent balance between hepatic ischemia/reperfusion injury and liver regeneration in mice. J Clin Invest 2009;119:2304–16. https://doi.org/10.1172/JCI38289.

[31] Markiewski MM, DeAngelis RA, Strey CW, Foukas PG, Gerard C, Gerard N, et al. The Regulation of Liver Cell Survival by Complement. J Immunol 2009;182:5412–8. https://doi.org/10.4049/jimmunol.0804179.

[32] Clark A, Weymann A, Hartman E, Turmelle Y, Carroll M, Thurman JM, et al. Evidence for non-traditional activation of complement factor C3 during murine liver regeneration. Mol Immunol 2008;45:3125–32. https://doi.org/10.1016/j.molimm.2008.03.008.

[33] Foley JH. Examining coagulation-complement crosstalk: complement activation and thrombosis. Thromb Res 2016;141:S50–4. https://doi.org/10.1016/S0049-3848(16)30365-6.

[34] Amara U, Flierl MA, Rittirsch D, Klos A, Chen H, Acker B, et al. Molecular Intercommunication between the Complement and Coagulation Systems. J Immunol 2010;185:5628–36. https://doi.org/10.4049/jimmunol.0903678.

[35] Bumiller-Bini V, Oliveira-Toré C de F, Carvalho TM, Kretzschmar GC, Gonçalves LB, Alencar N de M, et al. Masps at the crossroad between the complement and the coagulation cascades-the case for covid-19. Genet Mol Biol 2021;44:1–21. https://doi.org/10.1590/1678-4685-GMB-2020-0199.

[36] Øvstebø R, Hellum M, Aass HCD, Trøseid AM, Brandtzaeg P, Mollnes TE, et al. Microparticle-associated tissue factor activity is reduced by inhibition of the complement protein 5 in Neisseria meningitidis-exposed whole blood. Innate Immun 2014;20:552–60.

https://doi.org/10.1177/1753425913502099/ASSET/IMAGES/LARGE/10.1177_1753425913502099-FIG5.JPEG.

[37] Rehmus EH, Greene BM, Everson BA, Ratnoff OD. Inhibition of the activation of Hageman factor (factor XII) by complement subcomponent C1q. J Clin Invest 1987;80:516–21. https://doi.org/10.1172/JCI113100.

[38] Guo H, Callaway JB, Ting JPY. Inflammasomes: Mechanism of action, role in disease, and therapeutics. Nat Med 2015;21:677–87. https://doi.org/10.1038/nm.3893.

[39] Jagels MA, Daffern PJ, Hugli TE. C3a and C5a enhance granulocyte adhesion to endothelial and epithelial cell monolayers: Epithelial and endothelial priming is required for C3a-induced eosinophil adhesion. Immunopharmacology 2000;46:209–22. https://doi.org/10.1016/S0162-3109(99)00178-2.

[40] Holers VM. Complement and Its Receptors: New Insights into Human Disease. Annu Rev Immunol 2014;32:433–59. https://doi.org/10.1146/annurev-immunol-032713-120154.

[41] Arbore G, Kemper C. A novel "complement–metabolism–inflammasome axis" as a key regulator of immune cell effector function. Eur J Immunol 2016;46:1563–73. https://doi.org/10.1002/eji.201546131.

[42] Kwan W hong, van der Touw W, Heeger PS. Complement regulation of T cell immunity. Immunol Res 2012;54:247–53. https://doi.org/doi:10.1007/s12026-012-8327-1.

[43] Strainic MG, Liu J, Huang D, An F, Lalli PN, Muqim N, et al. Locally Produced Complement Fragments C5a and C3a Provide Both Costimulatory and Survival Signals to Naive CD4+ T Cells. Immunity 2008;28:425–35. https://doi.org/doi:10.1016/j.immuni.2008.02.001.

[44] Kwan W hong, van der Touw W, Paz-Artal E, Li MO, Heeger PS. Signaling through C5a receptor and C3a receptor diminishes function of murine natural regulatory T cells. J Exp Med 2013;210:257–68. https://doi.org/10.1084/jem.20121525.

[45] West EE, Afzali B, Kemper C. Unexpected Roles for Intracellular Complement in the

Regulation of Th1 Responses. Adv. Immunol., vol. 138, Academic Press Inc.; 2018, p. 35–70. https://doi.org/10.1016/bs.ai.2018.02.001.

[46] Dempsey PW, Allison MED, Akkaraju S, Goodnow CC, Fearon DT. C3d of complement as a molecular adjuvant: Bridging innate and acquired immunity. Science (80-) 1996;271:348–50. https://doi.org/10.1126/science.271.5247.348.

[47] Fang Y, Xu C, Fu Y-X, Holers VM, Molina H. Expression of Complement Receptors 1 and 2 on Follicular Dendritic Cells Is Necessary for the Generation of a Strong Antigen-Specific IgG Response. J Immunol 1998;160:5273–9. https://doi.org/10.4049/jimmunol.160.11.5273.

[48] King BC, Blom AM. Intracellular complement: Evidence, definitions, controversies, and solutions. Immunol Rev 2023;313:104–19. https://doi.org/10.1111/imr.13135.

[49] Kunz N, Kemper C. Complement Has Brains—Do Intracellular Complement and Immunometabolism Cooperate in Tissue Homeostasis and Behavior? Front Immunol 2021;12:154. https://doi.org/10.3389/FIMMU.2021.629986/BIBTEX.

[50] Arbore G, Kemper C, Kolev M. Intracellular complement − the complosome − in immune cell regulation. Mol Immunol 2017;89:2–9. https://doi.org/10.1016/j.molimm.2017.05.012.

[51] Ishii M, Beeson G, Beeson C, Rohrer B. Mitochondrial C3a Receptor Activation in Oxidatively Stressed Epithelial Cells Reduces Mitochondrial Respiration and Metabolism. Front Immunol 2021;12:1–16. https://doi.org/10.3389/fimmu.2021.628062.

[52] Friščić J, Böttcher M, Reinwald C, Bruns H, Wirth B, Popp SJ, et al. The complement system drives local inflammatory tissue priming by metabolic reprogramming of synovial fibroblasts. Immunity 2021;54:1002-1021.e10. https://doi.org/10.1016/j.immuni.2021.03.003.

[53] Tam JCH, Bidgood SR, McEwan WA, James LC. Intracellular sensing of complement C3 activates cell autonomous immunity. Science (80-) 2014;345:1134. https://doi.org/10.1126/science.1256070.

[54] Ricklin D, Reis ES, Lambris JD. Complement in disease: a defence system turning offensive. Nat Rev Nephrol 2016;12:383–401. https://doi.org/10.1038/nrneph.2016.70.

[55] Ricklin D, Mastellos DC, Reis ES, Lambris JD. The renaissance of complement therapeutics. Nat Rev Nephrol 2017;14:26–47. https://doi.org/10.1038/nrneph.2017.156.

[56] Csuka D, Füst G, Farkas H, Varga L. Parameters of the classical complement pathway predict disease severity in hereditary angioedema. Clin Immunol 2011;139:85–93. https://doi.org/10.1016/j.clim.2011.01.003.

[57] Garred P, Tenner AJ, Mollnes TE. Therapeutic Targeting of the Complement System: From Rare Diseases to Pandemics. Pharmacol Rev 2021;73:792–827. https://doi.org/10.1124/pharmrev.120.000072.

[58] Ricklin D, Lambris JD. Complement in Immune and Inflammatory Disorders: Pathophysiological Mechanisms. J Immunol Ref 2019;190:3831–8. https://doi.org/10.4049/jimmunol.1203487.

[59] Morgan BP. Complement in the pathogenesis of Alzheimer's disease. Semin Immunopathol 2018;40:113–24. https://doi.org/10.1007/s00281-017-0662-9.

[60] Lawrenz MB, Wooten RM, Zachary JF, Drouin SM, Weis JJ, Wetsel RA, et al. Effect of complement component C3 deficiency on experimental lyme borreliosis in mice. Infect Immun 2003;71:4432–40. https://doi.org/10.1128/IAI.71.8.4432-4440.2003.

[61] Skare JT, Garcia BL. Complement Evasion by Lyme Disease Spirochetes. Trends Microbiol 2020;28:889–99. https://doi.org/10.1016/J.TIM.2020.05.004.

[62] Garcia BL, Zwarthoff SA, Rooijakkers SHM, Geisbrecht B V. Novel Evasion Mechanisms of the Classical Complement Pathway. J Immunol 2016;197:2051–60. https://doi.org/10.4049/jimmunol.1600863.

[63] Xie J, Zhi Id H, Garrigues Id RJ, Keightley A, Garcia BL, 2☯ ID, et al. Structural determination of the complement inhibitory domain of Borrelia burgdorferi BBK32 provides insight into classical pathway complement evasion by Lyme disease spirochetes.

PLOS Pathog 2019:1–29. https://doi.org/10.1371/journal.ppat.1007659.

[64] Morgan BP, Harris CL. Complement, a target for therapy in inflammatory and degenerative diseases. Nat Rev Drug Discov 2015;14:857–77. https://doi.org/10.1038/nrd4657.

[65] Castellano G, Curci C, Racaniello D, Melchiorre R, Loverre A, Montinaro V, et al. Therapeutic targeting of classical and lectin pathways of Complement protects from ischemia-reperfusion induced renal damage inhibiting the activation of dendritic cells. Abstr / Mol Immunol 2009;46:2818–71. https://doi.org/10.1016/j.molimm.2009.05.262.

[66] Silverman SM, Kim BJ, Howell GR, Miller J, John SWM, Wordinger RJ, et al. C1q propagates microglial activation and neurodegeneration in the visual axis following retinal ischemia/reperfusion injury. Mol Neurodegener 2016;11. https://doi.org/10.1186/s13024-016-0089-0.

[67] Phuan P-W, Zhang H, Asavapanumas N, Leviten M, Rosenthal A, Lukmanee ·, et al. C1q-targeted monoclonal antibody prevents complement-dependent cytotoxicity and neuropathology in in vitro and mouse models of neuromyelitis optica. Acta Neuropathol 2013;3:829–40. https://doi.org/10.1007/s00401-013-1128-3.

[68] Zhang J, Fife CE, Currie MS, Moon RE, Piantadosi CA. Venous gas emboli and complement activation after deep repetitive air diving. Undersea Biomed Res 1991:293–302. https://pubmed.ncbi.nlm.nih.gov/1887517/ (accessed June 12, 2023).

[69] Storm BS, Christiansen D, Fure H, Ludviksen JK, Lau C, Lambris JD, et al. Air Bubbles Activate Complement and Trigger Hemostasis and C3-Dependent Cytokine Release Ex Vivo in Human Whole Blood. J Immunol 2021;207:2828–40. https://doi.org/10.4049/jimmunol.2100308.

[70] Storm BS, Ludviksen JK, Christiansen D, Fure H, Pettersen K, Landsem A, et al. Venous Air Embolism Activates Complement C3 Without Corresponding C5 Activation and Trigger Thromboinflammation in Pigs. Front Immunol 2022;13:1–13. https://doi.org/10.3389/fimmu.2022.839632.

[71] Ciechanowska A, Ciapała K, Pawlik K, Oggioni M, Mercurio D, De Simoni MG, et al. Initiators of classical and lectin complement pathways are differently engaged after traumatic brain injury—time-dependent changes in the cortex, striatum, thalamus and hippocampus in a mouse model. Int J Mol Sci 2021;22:1–30. https://doi.org/10.3390/ijms22010045.

[72] Roach PJ. Functional significance of enzyme cascade systems. Trends Biochem Sci 1977;2:87–90. https://doi.org/10.1016/0968-0004(77)90045-7.

[73] Neurath H, Walsh KA. Role of proteolytic enzymes in biological regulation (a review). Proc Natl Acad Sci U S A 1976;73:3825. https://doi.org/10.1073/PNAS.73.11.3825.

[74] Liebman MN. Analysis of the biomacromolecular architecture of eukaryotic and prokaryotic serine proteases. J Ind Microbiol 1988;3:127–37. https://doi.org/10.1007/BF01569519.

[75] Hamori E, Muldrey JE. Use of the word eager instead of spontaneous for the description of exergonic reactions. J Chem Educ 1984;61:710. https://doi.org/10.1021/ed061p710.

[76] Drag M, Salvesen GS. Emerging principles in protease-based drug discovery. Nat Rev Drug Discov 2010;9:690–701. https://doi.org/10.1038/nrd3053.

[77] Kasahara Y, Odai H, Takahashi K, Nagasawa S, Koyama J. Autocatalytic Activation of Clr Subcomponent of the First Component of Human Complement 1 1985;97:365–72.

[78] Kardos J, Harmat V, Palló A, Barabás O, Szilágyi K, Gráf L, et al. Revisiting the mechanism of the autoactivation of the complement protease C1r in the C1 complex: Structure of the active catalytic region of C1r. Mol Immunol 2008. https://doi.org/10.1016/j.molimm.2007.09.031.

[79] Vorup-Jensen T, Petersen S V., Hansen AG, Poulsen K, Schwaeble W, Sim RB, et al. Distinct Pathways of Mannan-Binding Lectin (MBL)- and C1-Complex Autoactivation Revealed by Reconstitution of MBL with Recombinant MBL-Associated Serine Protease-2. J Immunol 2000;165:2093–100. https://doi.org/10.4049/JIMMUNOL.165.4.2093.

[80] Gál P, Harmat V, Kocsis A, Bián T, Barna L, Ambrus G, et al. A true autoactivating enzyme: Structural insight into mannose-binding lectin-associated serine protease-2 activations. J Biol Chem 2005;280:33435–44. https://doi.org/10.1074/jbc.M506051200.

[81] Cera E Di. Serine Proteases. IUBMB Life 2009;61:510. https://doi.org/10.1002/IUB.186.

[82] Hedstrom L. Serine protease mechanism and specificity. Chem Rev 2002;102:4501–23. https://doi.org/10.1021/cr000033x.

[83] Nakagawa S, Yu H -A, Karplus M, Urneyama H. Active site dynamics of acyl-chymotrypsin. Proteins Struct Funct Bioinforma 1993;16:172–94. https://doi.org/10.1002/prot.340160205.

[84] Radisky ES, Lee JM, Lu CJK, Koshland DE. Insights into the serine protease mechanism from atomic resolution structures of trypsin reaction intermediates. Proc Natl Acad Sci U S A 2006;103:6835–40. https://doi.org/10.1073/PNAS.0601910103/SUPPL_FILE/01910FIG3.PDF.

[85] Al-Ani F, Chin-Yee I, Lazo-Langner A. Eculizumab in the management of paroxysmal nocturnal hemoglobinuria: Patient selection and special considerations. Ther Clin Risk Manag 2016;12:1161–70. https://doi.org/10.2147/TCRM.S96720.

[86] Waters AM, Licht C. AHUS caused by complement dysregulation: New therapies on the horizon. Pediatr Nephrol 2011;26:41–57. https://doi.org/10.1007/s00467-010-1556-4.

[87] Harris CL. Expanding horizons in complement drug discovery: challenges and emerging strategies. Semin Immunopathol 2018;40:125–40. https://doi.org/10.1007/s00281-017-0655-8.

[88] Mastellos DC, Ricklin D, Lambris JD. Clinical promise of next-generation complement therapeutics. Nat Rev Drug Discov 2019;18:707–29. https://doi.org/10.1038/s41573-019-0031-6.

[89] Risitano AM, Marotta S, Ricci P, Marano L, Frieri C, Cacace F, et al. Anti-complement Treatment for Paroxysmal Nocturnal Hemoglobinuria: Time for Proximal Complement

Inhibition? A Position Paper From the SAAWP of the EBMT. Front Immunol 2019;10:1157. https://doi.org/10.3389/fimmu.2019.01157.

[90] Röth A, Barcellini W, D'Sa S, Miyakawa Y, Broome CM, Michel M, et al. Sutimlimab in Cold Agglutinin Disease. N Engl J Med 2021;384:1323–34. https://doi.org/10.1056/nejmoa2027760.

[91] ANX005 - List Results - ClinicalTrials.gov n.d. https://clinicaltrials.gov/ct2/results?cond=&term=anx005&cntry=&state=&city=&dist= (accessed June 12, 2023).

[92] Ferreira De Freitas R, Schapira M. A systematic analysis of atomic protein-ligand interactions in the PDB. Medchemcomm 2017;8:1970–81. https://doi.org/10.1039/c7md00381a.

[93] Renaud JP, Chung CW, Danielson UH, Egner U, Hennig M, Hubbard RE, et al. Biophysics in drug discovery: Impact, challenges and opportunities. Nat Rev Drug Discov 2016;15:679–98. https://doi.org/10.1038/NRD.2016.123.

[94] Barnes AP, Khandelwal S, Sartoretto S, Myoung S, Francis SJ, Lee GM, et al. Minimal role for the alternative pathway in complement activation by HIT immune complexes. J Thromb Haemost 2022;20:2656–65. https://doi.org/10.1111/jth.15856.

[95] Khandelwal S, Barnes A, Rauova L, Sarkar A, Rux AH, Yarovoi S V., et al. Complement mediates binding and procoagulant effects of ultralarge HIT immune complexes. Blood 2021;138:2106–16. https://doi.org/10.1182/blood.2020009487.

[96] Zabel MK, Kirsch WM. From development to dysfunction: Microglia and the complement cascade in CNS homeostasis. Ageing Res Rev 2013;12:749–56. https://doi.org/10.1016/j.arr.2013.02.001.

[97] Tenner AJ, Fonseca MI. The double-edged flower: Roles of complement protein C1q in neurodegenerative diseases. Adv. Exp. Med. Biol., vol. 586, Springer New York; 2006, p. 153–76. https://doi.org/10.1007/0-387-34134-X_11.

[98] Bigal M. Are Small Molecule Drugs Still Relevant? Med Mak 2021. https://themedicinemaker.com/discovery-development/a-reawakening-of-small-molecule-drug-development (accessed June 12, 2023).

[99] Gurevich E V, Gurevich V V. Therapeutic potential of small molecules and engineered proteins. Handb Exp Pharmacol 2014;219:1–11. https://doi.org/10.1007/978-3-642-41199-1.

[100] Li Q, Kang C. Mechanisms of action for small molecules revealed by structural biology in drug discovery. Int J Mol Sci 2020;21:1–18. https://doi.org/10.3390/ijms21155262.

[101] Abdel-Magid AF. Allosteric Modulators: An Emerging Concept in Drug Discovery. ACS Med Chem Lett 2015;6:104. https://doi.org/10.1021/ML5005365.

[102] Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 2012;64:4–17. https://doi.org/10.1016/j.addr.2012.09.019.

[103] Baryakova TH, Pogostin BH, Langer R, McHugh KJ. Overcoming barriers to patient adherence: the case for developing innovative drug delivery systems. Nat Rev Drug Discov 2023;22. https://doi.org/10.1038/s41573-023-00670-0.

[104] Zhong L, Li Y, Xiong L, Wang W, Wu M, Yuan T, et al. Small molecules in targeted cancer therapy: advances, challenges, and future perspectives. Signal Transduct Target Ther 2021;6. https://doi.org/10.1038/s41392-021-00572-w.

[105] Ngo HX, Garneau-Tsodikova S. What are the drugs of the future? Medchemcomm 2018;9:757–8. https://doi.org/10.1039/c8md90019a.

[106] Mu X, Tang Y, Wu F, Ma H, Huang S, Liang M, et al. A Simple Small Molecule with Synergistic Passive and Active Dual-Targeting Effects for Imaging-Guided Photothermal Cancer Therapy. ACS Appl Mater Interfaces 2021;13:36958–66. https://doi.org/10.1021/acsami.1c11798.

[107] Chacon AC, Melucci AD, Qin SS, Prieto PA. Thinking small: Small molecules as

potential synergistic adjuncts to checkpoint inhibition in melanoma. Int J Mol Sci 2021;22:1–26. https://doi.org/10.3390/ijms22063228.

[108] Sernoskie SC, Jee A, Uetrecht JP. The emerging role of the innate immune response in idiosyncratic drug reactionss. Pharmacol Rev 2021;73:861–96. https://doi.org/10.1124/pharmrev.120.000090.

[109] Günther S, Hempel D, Dunkel M, Rother K, Preissner R. SuperHapten: A comprehensive database for small immunogenic compounds. Nucleic Acids Res 2007;35:906–10. https://doi.org/10.1093/nar/gkl849.

[110] Baker MP, Reynolds HM, Lumicisi B, Bryson CJ. Immunogenicity of protein therapeutics: The key causes, consequences and challenges. Self/Nonself - Immune Recognit Signal 2010;1:314–22. https://doi.org/10.4161/self.1.4.13904.

[111] Daina A, Michielin O, Zoete V. SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep 2017;7:1–13. https://doi.org/10.1038/srep42717.

[112] Watkins AM, Arora PS. Structure-based inhibition of protein - protein interactions. Eur J Med Chem 2015;94:480–8. https://doi.org/10.1016/j.ejmech.2014.09.047.

[113] Giannetti AM. From eperimental design to validated hits: A comprehensive walk-through of fragment lead identification using surface plasmon resonance. Methods Enzymol., 2011. https://doi.org/10.1016/B978-0-12-381274-2.00008-X.

[114] Davis BJ, Erlanson DA. Learning from our mistakes: The “unknown knowns” in fragment screening. Bioorganic Med Chem Lett 2013. https://doi.org/10.1016/j.bmcl.2013.03.028.

[115] Baker M. Fragment-based lead discovery grows up. Nat Rev Drug Discov 2013;12:5–7. https://doi.org/10.1038/nrd3926.

[116] Schultes S, de Graaf C, Haaksma EEJ, de Esch IJP, Leurs R, Kramer O. Ligand efficiency as a guide in fragment hit selection and optimization | Elsevier Enhanced Reader. Drug Discov Today 2010:157–62. https://doi.org/10.1016/j.ddtec.2010.11.003.

[117] Investigational New Drug (IND) Application | FDA n.d. https://www.fda.gov/drugs/types-applications/investigational-new-drug-ind-application (accessed June 12, 2023).

[118] Types and Phases of Clinical Trials | What Are Clinical Trial Phases? n.d. https://www.cancer.org/cancer/managing-cancer/making-treatment-decisions/clinical-trials/what-you-need-to-know/phases-of-clinical-trials.html (accessed June 12, 2023).

[119] New Drug Application (NDA) | FDA n.d. https://www.fda.gov/drugs/types-applications/new-drug-application-nda (accessed June 12, 2023).

[120] Sun D, Gao W, Hu H, Zhou S. Why 90% of clinical drug development fails and how to improve it? Acta Pharm Sin B 2022;12:3049–62. https://doi.org/10.1016/J.APSB.2022.02.002.

[121] Mullard A. 2021 FDA approvals. Nat Rev Drug Discov 2022;21:83–8. https://doi.org/10.1038/D41573-022-00001-9.

[122] Yamaguchi S, Kaneko M, Narukawa M. Approval success rates of drug candidates based on target, action, modality, application, and their combinations. Clin Transl Sci 2021;14:1113–22. https://doi.org/10.1111/CTS.12980.

[123] Drag M, Salvesen GS. Emerging principles in protease-based drug discovery. Nat Rev Drug Discov 2010;9:690–701. https://doi.org/10.1038/nrd3053.

[124] Liebman MN. Structural Organization in the Serine Proteases 1986.

[125] Deu E, Verdoes M, Bogyo M. New approaches for dissecting protease functions to improve probe development and drug discovery. Nat Struct Mol Biol 2012;19:9–16. https://doi.org/10.1038/nsmb.2203.

[126] Fenwick RB, Esteban-Martiín S, Richter B, Lee D, Walter KFA, Milovanovic D, et al. Weak long-range correlated motions in a surface patch of ubiquitin involved in molecular recognition. J Am Chem Soc 2011;133:10336–9. https://doi.org/10.1021/ja200461n.

[127] Gohara DW, Di Cera E. Allostery in trypsin-like proteases suggests new therapeutic

strategies. Trends Biotechnol 2011;29:577–85. https://doi.org/10.1016/J.TIBTECH.2011.06.001.

[128] Yu H, Dalby PA. A beginner's guide to molecular dynamics simulations and the identification of cross-correlation networks for enzyme engineering. vol. 643. 1st ed. Elsevier Inc.; 2020. https://doi.org/10.1016/bs.mie.2020.04.020.

[129] Atanasov AG, Zotchev SB, Dirsch VM, Orhan IE, Banach M, Rollinger JM, et al. Natural products in drug discovery: advances and opportunities. Nat Rev Drug Discov 2021;20:200–16. https://doi.org/10.1038/s41573-020-00114-z.

[130] Tošović J, Fijan D, Jukič M, Bren U. Conserved Water Networks Identification for Drug Design Using Density Clustering Approaches on Positional and Orientational Data. J Chem Inf Model 2022;62:6105–17. https://doi.org/10.1021/acs.jcim.2c00801.

[131] Lukac I, Wyatt PG, Gilbert IH, Zuccotto F. Ligand binding: evaluating the contribution of the water molecules network using the Fragment Molecular Orbital method. J Comput Aided Mol Des 2021;35:1025–36. https://doi.org/10.1007/s10822-021-00416-3.

[132] Lamoree B, Hubbard RE. Current perspectives in fragment-based lead discovery (FBLD). Essays Biochem 2017;61:453–64. https://doi.org/10.1042/EBC20170028.

[133] Ritchie TJ, Ertl P, Lewis R. The graphical representation of ADME-related molecule properties for medicinal chemists. Drug Discov Today 2011;16:65–72. https://doi.org/10.1016/j.drudis.2010.11.002.

[134] Osborne J, Panova S, Rapti M, Urushima T, Jhoti H. Fragments: where are we now? Biochem Soc Trans 2020;48:271–80. https://doi.org/10.1042/BST20190694.

[135] Jhoti H, Williams G, Rees DC, Murray CW. The "rule of three" for fragment-based drug discovery: Where are we now? Nat Rev Drug Discov 2013;12:644. https://doi.org/10.1038/nrd3926-c1.

[136] Hopkins AL, Groom CR, Alex A. Ligand efficiency: a useful metric for lead selection. Drug Discov Today 2004;9:430–1. https://doi.org/10.1016/S1359-6446(04)03069-7.

[137] Bollag G, Tsai J, Zhang J, Zhang C, Ibrahim P, Nolop K, et al. Vemurafenib: The first drug approved for BRAF-mutant cancer. Nat Rev Drug Discov 2012;11:873–86. https://doi.org/10.1038/nrd3847.

[138] Leverson JD, Boghaert ER, Ackler SL, Catron ND, Chen J, Dayton BD, et al. ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. Nat Med 2013;19. https://doi.org/10.1038/nm.3048.

[139] Perera TPS, Jovcheva E, Mevellec L, Vialard J, De Lange D, Verhulst T, et al. Discovery & pharmacological characterization of JNJ-42756493 (Erdafitinib), a functionally selective small-molecule FGFR family inhibitor. Mol Cancer Ther 2017;16:1010–20. https://doi.org/10.1158/1535-7163.MCT-16-0589.

[140] Tap WD, Wainberg ZA, Anthony SP, Ibrahim PN, Zhang C, Healey JH, et al. Structure-Guided Blockade of CSF1R Kinase in Tenosynovial Giant-Cell Tumor. N Engl J Med 2015;373:428–37. https://doi.org/10.1056/nejmoa1411366.

[141] Patel V, Shah M. Artificial intelligence and machine learning in drug discovery and development. Intell Med 2022;2:134–40. https://doi.org/10.1016/j.imed.2021.10.001.

[142] Çelik IN, Arslan FK, Tun R, Yildiz I. Artificial Intelligence on Drug Discovery and Development. Ankara Univ Eczac Fak Derg 2022;46:400–27. https://doi.org/10.33483/jfpau.878041q.

[143] Borkakoti N, Thornton JM. AlphaFold2 protein structure prediction: Implications for drug discovery. Curr Opin Struct Biol 2023;78:102526. https://doi.org/10.1016/J.SBI.2022.102526.

[144] Page MJ, Di Cera E. Serine peptidases: Classification, structure and function. Cell Mol Life Sci 2008;65:1220–36. https://doi.org/10.1007/s00018-008-7565-9.

[145] Ekici ÖD, Paetzel M, Dalbey RE. Unconventional serine proteases: Variations on the catalytic Ser/His/Asp triad configuration. Protein Sci 2008;17:2023. https://doi.org/10.1110/PS.035436.108.

[146] Krem MM, Cera E Di. Evolution of enzyme cascades from embryonic development to blood coagulation. Trends Biochem Sci 2002;27:67–74. https://doi.org/10.1016/S0968-0004(01)02007-2.

[147] Nesargikar PN, Spiller B, Chavez R. The Complement System: History, Pathways, Cascade and Inhibitors. Eur J Microbiol Immunol 2012;2:103–11. https://doi.org/10.1556/EuJMI.2.2012.2.2.

[148] Gál P, Barna L, Kocsis A, Závodszky P. Serine proteases of the classical and lectin pathways: Similarities and differences. Immunobiology 2007;212:267–77. https://doi.org/10.1016/j.imbio.2006.11.002.

[149] Vorup-Jensen T, Petersen S V., Hansen AG, Poulsen K, Schwaeble W, Sim RB, et al. Distinct Pathways of Mannan-Binding Lectin (MBL)- and C1-Complex Autoactivation Revealed by Reconstitution of MBL with Recombinant MBL-Associated Serine Protease-2. J Immunol 2000;165:2093–100. https://doi.org/10.4049/JIMMUNOL.165.4.2093.

[150] Megyeri M, Harmat V, Major B, Végh Á, Balczer J, Héja D, et al. Quantitative Characterization of the Activation Steps of Mannan-binding Lectin (MBL)-associated Serine Proteases (MASPs) Points to the Central Role of MASP-1 in the Initiation of the Complement Lectin Pathway. J Biol Chem 2013;288:8922–34. https://doi.org/10.1074/JBC.M112.446500.

[151] Héja D, Harmats V, Fodor K, Wilmanns M, Dobó J, Kékesi KA, et al. Monospecific inhibitors show that both mannan-binding lectin-associated serine protease-1 (MASP-1) and -2 are essential for lectin pathway activation and reveal structural plasticity of MASP-2. J Biol Chem 2012;287:20290–300. https://doi.org/10.1074/JBC.M112.354332.

[152] Zipfel PF, Wiech T, Rudnick R, Afonso S, Person F, Skerka C. Complement inhibitors in clinical trials for glomerular diseases. Front Immunol 2019;10:2166. https://doi.org/10.3389/fimmu.2019.02166.

[153] Damoah CE, Snir O, Hindberg K, Garred P, Ludviksen JK, Brækkan SK, et al. High Levels of Complement Activating Enzyme MASP-2 Are Associated With the Risk of

Future Incident Venous Thromboembolism. Arterioscler Thromb Vasc Biol 2022;42:1186–97. https://doi.org/10.1161/ATVBAHA.122.317746.

[154] Khaled SK, Claes K, Goh YT, Kwong YL, Leung N, Mendrek W, et al. Narsoplimab, a Mannan-Binding Lectin-Associated Serine Protease-2 Inhibitor, for the Treatment of Adult Hematopoietic Stem-Cell Transplantation-Associated Thrombotic Microangiopathy. J Clin Oncol 2022;3. https://doi.org/10.1200/JCO.21.02389.

[155] Harmat V, Gál P, Kardos J, Szilágyi K, Ambrus G, Végh B, et al. The Structure of MBL-associated Serine Protease-2 Reveals that Identical Substrate Specificities of C1s and MASP-2 are Realized Through Different Sets of Enzyme–Substrate Interactions. J Mol Biol 2004;342:1533–46. https://doi.org/10.1016/J.JMB.2004.07.014.

[156] Blankenship E, Vukoti K, Miyagi M, Lodowski DT. Conformational flexibility in the catalytic triad revealed by the high-resolution crystal structure of Streptomyces erythraeus trypsin in an unliganded state. Acta Crystallogr Sect D Biol Crystallogr 2014;70:833–40. https://doi.org/10.1107/S1399004713033658.

[157] Budayova-Spano M, Grabarse W, Thielens NM, Hillen H, Lacroix M, Schmidt M, et al. Monomeric structures of the zymogen and active catalytic domain of complement protease C1r: Further insights into the C1 activation mechanism. Structure 2002. https://doi.org/10.1016/S0969-2126(02)00881-X.

[158] Liebschner D, Afonine P V., Baker ML, Bunkoczi G, Chen VB, Croll TI, et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Urn:Issn:2059-7983 2019;75:861–77. https://doi.org/10.1107/S2059798319011471.

[159] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and Development of Coot. Acta Crystallogr Sect D - Biol Crystallogr 2010;66:486–501.

[160] Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J Chem Inf Model 2011;51:2778–86. https://doi.org/10.1021/CI200227U.

[161] Krissinel E, Henrick K. Inference of macromolecular assemblies from crystalline state. J

Mol Biol 2007:774–97.

[162] Wang J. Estimation of the quality of refined protein crystal structures. Protein Sci 2015;24:661–9. https://doi.org/10.1002/pro.2639.

[163] Bizzozero SA, Dutler H. Stereochemical aspects of peptide hydrolysis catalyzed by serine proteases of the chymotrypsin type. Bioorg Chem 1981;10:46–62. https://doi.org/10.1016/0045-2068(81)90042-0.

[164] Budayova-Spano M, Grabarse W, Thielens NM, Hillen H, Lacroix M, Schmidt M, et al. Monomeric Structures of the Zymogen and Active Catalytic Domain of Complement Protease C1r: Further Insights into the C1 Activation Mechanism. Structure 2002;10:1509–19. https://doi.org/10.1016/S0969-2126(02)00881-X.

[165] Zakharova E, Horvath MP, Goldenberg DP. Structure of a serine protease poised to resynthesize a peptide bond. Proc Natl Acad Sci U S A 2009;106:11034–9. https://doi.org/10.1073/pnas.0902463106.

[166] Derewenda ZS, Derewenda U, Kobos PM. (His)Cε-H···O=C< Hydrogen Bond in the Active Sites of Serine Hydrolases. J Mol Biol 1994;241:83–93. https://doi.org/10.1006/JMBI.1994.1475.

[167] Fuhrmann CN, Kelch BA, Ota N, Agard DA. The 0.83 Å resolution crystal structure of α-lytic protease reveals the detailed structure of the active site and identifies a source of conformational strain. J Mol Biol 2004;338:999–1013. https://doi.org/10.1016/j.jmb.2004.03.018.

[168] Würtele M, Hahn M, Hilpert K, Höhne W. Atomic resolution structure of native porcine pancreatic elastase at 1.1 A. Acta Crystallogr D Biol Crystallogr 2000;56:520–3. https://doi.org/10.1107/S0907444900000299.

[169] Schmidt A, Lamzin VS. Extraction of functional motion in trypsin crystal structures. Acta Crystallogr D Biol Crystallogr 2005;61:1132–9. https://doi.org/10.1107/S0907444905016732.

[170] Blankenship E, Vukoti K, Miyagi M, Lodowski DT. Conformational flexibility in the catalytic triad revealed by the high-resolution crystal structure of Streptomyces erythraeus trypsin in an unliganded state. Acta Crystallogr Sect D Biol Crystallogr 2014;70:833–40. https://doi.org/10.1107/S1399004713033658.

[171] Fisher SJ, Blakeley MP, Cianci M, McSweeney S, Helliwell JR. Protonation-state determination in proteins using high-resolution X-ray crystallography: effects of resolution and completeness. Acta Crystallogr D Biol Crystallogr 2012;68:800–9. https://doi.org/10.1107/S0907444912012589.

[172] Liebschner D, Dauter M, Brzuszkiewicz A, Dauter Z. On the reproducibility of protein crystal structures: Five atomic resolution structures of trypsin. Acta Crystallogr Sect D Biol Crystallogr 2013;69:1447–62. https://doi.org/10.1107/S0907444913009050.

[173] Harris CL, Pouw RB, Kavanagh D, Sun R, Ricklin D. Developments in anti-complement therapy; from disease to clinical trial. Mol Immunol 2018. https://doi.org/10.1016/j.molimm.2018.06.008.

[174] Morgan BP, Harris CL. Complement, a target for therapy in inflammatory and degenerative diseases. Nat Rev Drug Discov 2015;14:857–77. https://doi.org/10.1038/nrd4657.

[175] Dobó J, Kocsis A, Gál P. Be on Target: Strategies of Targeting Alternative and Lectin Pathway Components in Complement-Mediated Diseases. Front Immunol 2018;9:1851. https://doi.org/10.3389/fimmu.2018.01851.

[176] Arepally GM. Heparin-induced thrombocytopenia. Blood 2017;129:2864–72. https://doi.org/10.1182/blood-2016-11-709873.

[177] Phuan P-W, Zhang H, Asavapanumas N, Leviten M, Rosenthal A, Tradtrantip L, et al. C1q-targeted monoclonal antibody prevents complement-dependent cytotoxicity and neuropathology in in vitro and mouse models of neuromyelitis optica. Acta Neuropathol 2013;125:829–40. https://doi.org/10.1007/s00401-013-1128-3.

[178] Li Q, Ujiie H, Shibaki A, Wang G, Moriuchi R, Qiao H, et al. Human IgG1 monoclonal

antibody against human collagen 17 noncollagenous 16A domain induces blisters via complement activation in experimental bullous pemphigoid model. J Immunol (Baltimore, Md 1950) 2010;185:7746–55. https://doi.org/10.4049/jimmunol.1000667.

[179] Edwards G, Diercks GFH, Seelen MAJ, Horvath B, van Doorn MBA, Damman J. Complement Activation in Autoimmune Bullous Dermatoses: A Comprehensive Review. Front Immunol 2019;10:1477. https://doi.org/10.3389/fimmu.2019.01477.

[180] Berentsen S. Complement Activation and Inhibition in Autoimmune Hemolytic Anemia: Focus on Cold Agglutinin Disease. Semin Hematol 2018;55:141–9. https://doi.org/10.1053/j.seminhematol.2018.04.002.

[181] Hong S, Beja-Glasser VF, Nfonoyim BM, Frouin A, Li S, Ramakrishnan S, et al. Complement and microglia mediate early synapse loss in Alzheimer mouse models. Science (80-) 2016;352:712–6. https://doi.org/10.1126/science.aad8373.

[182] Sekar A, Bialas AR, de Rivera H, Davis A, Hammond TR, Kamitaki N, et al. Schizophrenia risk from complex variation of complement component 4. Nature 2016;530:177-+. https://doi.org/10.1038/nature16549.

[183] Deniz U, Stuart CH, de Kreuk Bart-Jan, Roman IK, de Jong Rob N, Frank JB, et al. Structures of C1-IgG1 provide insights into how danger pattern recognition activates complement. Science (80-) 2018;359:794–7. https://doi.org/10.1126/science.aao4988.

[184] Diebolder CA, Beurskens FJ, de Jong RN, Koning RI, Strumane K, Lindorfer MA, et al. Complement is activated by IgG hexamers assembled at the cell surface. Science 2014;343:1260–3. https://doi.org/10.1126/science.1248943.

[185] Sharp TH, Boyle AL, Diebolder CA, Kros A, Koster AJ, Gros P. Insights into IgM-mediated complement activation based on in situ structures of IgM-C1-C4b. Proc Natl Acad Sci U S A 2019;116:11900–5. https://doi.org/10.1073/pnas.1901841116.

[186] Girija UV, Gingras AR, Marshall JE, Panchal R, Sheikh MA, Harper JAJ, et al. Structural basis of the C1q/C1s interaction and its central role in assembly of the C1 complex of complement activation. Proc Natl Acad Sci U S A 2013;110:13916–20.

https://doi.org/10.1073/pnas.1311113110.

[187] Phillips AE, Toth J, Dodds AW, Girija UV, Furze CM, Pala E, et al. Analogous Interactions in Initiating Complexes of the Classical and Lectin Pathways of Complement. J Immunol 2009. https://doi.org/10.4049/jimmunol.0900666.

[188] Erlanson DA, Fesik SW, Hubbard RE, Jahnke W, Jhoti H. Twenty years on: The impact of fragments on drug discovery. Nat Rev Drug Discov 2016;15:605–19. https://doi.org/10.1038/nrd.2016.109.

[189] Congreve M, Carr R, Murray C, Jhoti H. A "Rule of Three" for fragment-based lead discovery? Drug Discov Today 2003. https://doi.org/10.1016/S1359-6446(03)02831-9.

[190] Geisbrecht B V., Bouyain S, Pop M. An optimized system for expression and purification of secreted bacterial proteins. Protein Expr Purif 2006. https://doi.org/10.1016/j.pep.2005.09.003.

[191] Kardos J, Gal P, Szilagyi L, Thielens NM, Szilagyi K, Lorincz Z, et al. The Role of the Individual Domains in the Structure and Function of the Catalytic Region of a Modular Serine Protease, C1r. J Immunol 2001;167:5202–8.

[192] Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 2011. https://doi.org/10.1038/msb.2011.75.

[193] Backman TWH, Cao Y, Girke T. ChemMine tools: An online service for analyzing and clustering small molecules. Nucleic Acids Res 2011. https://doi.org/10.1093/nar/gkr320.

[194] Chen X, Reynolds CH. Performance of similarity measures in 2D fragment-based similarity searching: Comparison of structural descriptors and similarity coefficients. J Chem Inf Comput Sci 2002. https://doi.org/10.1021/ci025531g.

[195] Cao Y, Jiang T, Girke T. A maximum common substructure-based algorithm for searching and predicting drug-like compounds. Bioinformatics 2008. https://doi.org/10.1093/bioinformatics/btn186.

[196] Giannetti AM. From Experimental Design to Validated Hits. In: Kuo LC, editor. vol. 493, Elsevier Science & Technology; 2011, p. 169–218. https://doi.org/10.1016/B978-0-12-381274-2.00008-X.

[197] O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. Open Babel: An open chemical toolbox. J Cheminform 2011;3:33. https://doi.org/10.1186/1758-2946-3-33.

[198] Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, et al. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem 2009;30:2785–91. https://doi.org/10.1002/jcc.21256.

[199] DA GJSW. Big Data Analysis for Bioinformatics and Biomedical Discoveries. Chapman and Hall/CRC; 2016. https://doi.org/10.1201/b19129.

[200] Dallakyan S, Olson AJ. Small-molecule library screening by docking with PyRx. Methods Mol Biol 2015. https://doi.org/10.1007/978-1-4939-2269-7_19.

[201] Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: A message-passing parallel molecular dynamics implementation. Comput Phys Commun 1995. https://doi.org/10.1016/0010-4655(95)00042-E.

[202] Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 2010;66:486–501. https://doi.org/10.1107/S0907444910007493.

[203] Van Aalten DMF. PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. J Comput Aided Mol Des 1996. https://doi.org/10.1007/BF00355047.

[204] Hess B, Bekker H, Berendsen HJC, Fraaije JGEM. LINCS: A Linear Constraint Solver for molecular simulations. J Comput Chem 1997. https://doi.org/10.1002/(SICI)1096-987X(199709)18:12<1463::AID-JCC4>3.0.CO;2-H.

[205] Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle

mesh Ewald method. J Chem Phys 1995. https://doi.org/10.1063/1.470117.

[206] Berendsen HJC, Postma JPM, Van Gunsteren WF, Dinola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984. https://doi.org/10.1063/1.448118.

[207] Parrinello M, Rahman A. Polymorphic transitions in single crystals: A new molecular dynamics method. J Appl Phys 1981. https://doi.org/10.1063/1.328693.

[208] Kumari R, Kumar R, Lynn A. G-mmpbsa -A GROMACS tool for high-throughput MM-PBSA calculations. J Chem Inf Model 2014. https://doi.org/10.1021/ci500020m.

[209] Lovering F, Bikker J, Humblet C. Escape from flatland: Increasing saturation as an approach to improving clinical success. J Med Chem 2009;52:6752–6. https://doi.org/10.1021/JM901241E.

[210] Budayova-Spano M, Lacroix M, Thielens NM, Arlaud GJ, Fontecilla-Camps JC, Gaboriaud C. The crystal structure of the zymogen catalytic domain of complement protease C1r reveals that a disruptive mechanical stress is required to trigger activation of the C1 complex. EMBO J 2002;21:231–9. https://doi.org/10.1093/emboj/21.3.231.

[211] Almitairi JOM, Girija UV, Furze CM, Simpson-Gray X, Badakshi F, Marshall JE, et al. Structure of the C1r-C1s interaction of the C1 complex of complement activation. Proc Natl Acad Sci U S A 2018;115:768–73. https://doi.org/10.1073/pnas.1718709115.

[212] Plummer JS, Cai C, Hays SJ, Gilmore JL, Emmerling MR, Michael W, et al. Benzenesulfonamide derivatives of 2-substituted 4H-3,1-benzoxazin-4-ones and benzthiazin-4-ones as inhibitors of complement C1r protease. Bioorg Med Chem Lett 1999;9:815–20. https://doi.org/10.1016/s0960-894x(99)00095-5.

[213] Thompson DC, Humblet C, Joseph-McCarthy D. Investigation of MM-PBSA rescoring of docking poses. J Chem Inf Model 2008. https://doi.org/10.1021/ci700470c.

[214] Greenidge PA, Kramer C, Mozziconacci JC, Sherman W. Improving docking results via reranking of ensembles of ligand poses in multiple X-ray protein conformations with MM-

GBSA. J Chem Inf Model 2014. https://doi.org/10.1021/ci5003735.

[215] Ertl P, Altmann E, Racine S, Lewis R. Ring replacement recommender: Ring modifications for improving biological activity. Eur J Med Chem 2022;238:114483. https://doi.org/10.1016/j.ejmech.2022.114483.

[216] Meyer EA, Castellano RK, Diederich F. Interactions with aromatic rings in chemical and biological recognition. Angew Chemie - Int Ed 2003;42:1210–50. https://doi.org/10.1002/ANIE.200390319.

[217] Taylor RD, Maccoss M, Lawson ADG. Rings in drugs. J Med Chem 2014;57:5845–59. https://doi.org/10.1021/JM4017625.

[218] Khandelwal S, Ravi J, Rauova L, Johnson A, Lee GM, Gilner JB, et al. Polyreactive IgM initiates complement activation by PF4/heparin complexes through the classical pathway. Blood 2018;132:2431–40. https://doi.org/10.1182/blood-2018-03-834598.

[219] Schubart A, Anderson K, Mainolfi N, Sellner H, Ehara T, Adams CM, et al. Small-molecule factor B inhibitor for the treatment of complement-mediated diseases. Proc Natl Acad Sci U S A 2019;116:7926–31. https://doi.org/10.1073/pnas.1820892116.

[220] Karki RG, Powers J, Mainolfi N, Anderson K, Belanger DB, Liu D, et al. Design, Synthesis, and Preclinical Characterization of Selective Factor D Inhibitors Targeting the Alternative Complement Pathway. J Med Chem 2019;62:4656–68. https://doi.org/10.1021/acs.jmedchem.9b00271.

[221] Maibaum J, Liao SM, Vulpetti A, Ostermann N, Randl S, Rüdisser S, et al. Small-molecule factor D inhibitors targeting the alternative complement pathway. Nat Chem Biol 2016;12:1105–10. https://doi.org/10.1038/nchembio.2208.

[222] Pardridge WM. The Blood-Brain Barrier: Bottleneck in Brain Drug Development. J Am Soc Exp Neurother 2005;2:3–14.

[223] Pardridge WM. Drug transport across the blood-brain barrier. J Cereb Blood Flow Metab Off J Int Soc Cereb Blood Flow Metab 2012;32:1959–72.

https://doi.org/10.1038/jcbfm.2012.126.

[224] Koller F. Intravascular Clotting and Spontaneous Fibrinolysis. Acta Haematol 1964;31:239–46. https://doi.org/10.1159/000209633.

[225] Haddadin AS, Faraday N. Hemostasis, Coagulopathy, and Tamponade. Johns Hopkins Man Card Surg Care 2008:237–56. https://doi.org/10.1016/B978-0-323-01810-4.10010-0.

[226] Aprotinin - Pentapharm n.d. https://www.pentapharm.com/aprotinin/ (accessed May 30, 2023).

[227] Harmat V, Gál P, Kardos J, Szilágyi K, Ambrus G, Végh B, et al. The structure of MBL-associated serine protease-2 reveals that identical substrate specificities of C1s and MASP-2 are realized through different sets of enzyme-substrate interactions. J Mol Biol 2004;342:1533–46. https://doi.org/10.1016/j.jmb.2004.07.014.

[228] Dobó J, Harmat V, Beinrohr L, Sebestyén E, Závodszky P, Gál P. MASP-1, a Promiscuous Complement Protease: Structure of Its Catalytic Region Reveals the Basis of Its Broad Specificity. J Immunol 2009;183:1207–14. https://doi.org/10.4049/JIMMUNOL.0901141.

[229] Fenwick RB, Orellana L, Esteban-Martín S, Orozco M, Salvatella X. Correlated motions are a fundamental property of β-sheets. Nat Commun 2014;5:1–9. https://doi.org/10.1038/ncomms5070.

[230] Ertl P. Craig plot 2.0: An interactive navigation in the substituent bioisosteric space. J Cheminform 2020;12. https://doi.org/10.1186/S13321-020-0412-1.

[231] Craig PN. Interdependence between Physical Parametess and Selection of Substituent Groups for Correlation Studies. J Med Chem 1971;14:680–4. https://doi.org/10.1021/JM00290A004/ASSET/JM00290A004.FP.PNG_V03.

[232] Ferreira De Freitas R, Schapira M. A systematic analysis of atomic protein–ligand interactions in the PDB. Medchemcomm 2017;8:1970–81. https://doi.org/10.1039/C7MD00381A.

[233] Kocsis A. Serine proteases of the classical and lectin pathways : Similarities and differences. Immunobiology 2007;212:267–77. https://doi.org/10.1016/j.imbio.2006.11.002.

[234] Kardos J, Gál P, Szilágyi L, Thielens NM, Szilágyi K, Lõrincz Z, et al. The Role of the Individual Domains in the Structure and Function of the Catalytic Region of a Modular Serine Protease, C1r. J Immunol 2001;167:5202–8. https://doi.org/10.4049/jimmunol.167.9.5202.

[235] Perry AJ, Wijeyewickrema LC, Wilmann PG, Gunzburg MJ, D'Andrea L, Irving JA, et al. A Molecular Switch Governs the Interaction between the Human Complement Protease C1s and Its Substrate, Complement C4. J Biol Chem 2013;288:15821–9. https://doi.org/10.1074/JBC.M113.464545.

[236] Lansita JA, Mease KM, Qiu H, Yednock T, Sankaranarayanan S, Kramer S. Nonclinical Development of ANX005: A Humanized Anti-C1q Antibody for Treatment of Autoimmune and Neurodegenerative Diseases. Int J Toxicol 2017;36:449–62. https://doi.org/10.1177/1091581817740873.

[237] Wallach I, Dzamba M, Heifets A. AtomNet : A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery n.d.:1–11.

[238] Brash AR. Lipoxygenases: Occurrence, functions, catalysis, and acquisition of substrate. J Biol Chem 1999;274:23679–82. https://doi.org/10.1074/JBC.274.34.23679.

[239] Newcomer ME, Brash AR. The structural basis for specificity in lipoxygenase catalysis. Protein Sci 2015;24:298–309. https://doi.org/10.1002/PRO.2626.

[240] Viswanath KK, Varakumar P, Pamuru RR, Basha SJ, Mehta S, Rao AD. Plant Lipoxygenases and Their Role in Plant Physiology. J Plant Biol 2020;63:83–95. https://doi.org/10.1007/s12374-020-09241-x.

[241] Haeggström JZ, Funk CD. Lipoxygenase and leukotriene pathways: Biochemistry, biology, and roles in disease. Chem Rev 2011;111:5866–96. https://doi.org/10.1021/cr200246d.

[242] Gilbert NC, Newcomer ME, Werz O. Untangling the web of 5-lipoxygenase-derived products from a molecular and structural perspective: The battle between pro- and anti-inflammatory lipid mediators. Biochem Pharmacol 2021;193:114759. https://doi.org/10.1016/J.BCP.2021.114759.

[243] Dobrian AD, Morris MA, Taylor-Fishwick DA, Holman TR, Imai Y, Mirmira RG, et al. Role of the 12-lipoxygenase pathway in diabetes pathogenesis and complications. Pharmacol Ther 2019;195:100–10. https://doi.org/10.1016/J.PHARMTHERA.2018.10.010.

[244] Mashima R, Okuyama T. The role of lipoxygenases in pathophysiology; new insights and future perspectives. Redox Biol 2015;6:297–310. https://doi.org/10.1016/j.redox.2015.08.006.

[245] Berger W, De Chandt MTM, Cairns CB. Zileuton: Clinical implications of 5-Lipoxygenase inhibition in severe airway disease. Int J Clin Pract 2007;61:663–76. https://doi.org/10.1111/j.1742-1241.2007.01320.x.

[246] Offenbacher AR, Holman TR. Fatty Acid Allosteric Regulation of C-H Activation in Plant and Animal Lipoxygenases. Molecules 2020:244–54. https://doi.org/10.3390/molecules25153374.

[247] Oldham ML, Brash AR, Newcomer ME. Insights from the x-ray crystal structure of coral 8R-lipoxygenase: Calcium activation via a C2-like domain and a structural basis of product chirality. J Biol Chem 2005;280:39545–52. https://doi.org/10.1074/jbc.M506675200.

[248] Kobe MJ, Neau DB, Mitchell CE, Bartlett SG, Newcomer ME. The structure of human 15-lipoxygenase-2 with a substrate mimic. J Biol Chem 2014;289:8562–9. https://doi.org/10.1074/jbc.M113.543777.

[249] Anthonymuthu TS, Kenny EM, Shrivastava I, Tyurina YY, Hier ZE, Ting HC, et al. Empowerment of 15-Lipoxygenase Catalytic Competence in Selective Oxidation of Membrane ETE-PE to Ferroptotic Death Signals, HpETE-PE. J Am Chem Soc

2018;140:17835–9. https://doi.org/10.1021/jacs.8b09913.

[250] Börner F, Pace S, Jordan PM, Gerstmeier J, Gomez M, Rossi A, et al. Allosteric Activation of 15-Lipoxygenase-1 by Boswellic Acid Induces the Lipid Mediator Class Switch to Promote Resolution of Inflammation. Adv Sci 2022;2205604:1–14. https://doi.org/10.1002/advs.202205604.

[251] Dixon RAF, Diehl RE, Opas E, Rands E, Vickers PJ, Evans JF, et al. Requirement of a 5-lipoxygenase-activating protein for leukotriene synthesis. Nat 1990 3436255 1990;343:282–4. https://doi.org/10.1038/343282a0.

[252] Järving R, Lõokene A, Kurg R, Siimon L, Järving I, Samel N. Activation of 11R-lipoxygenase is fully Ca 2+-dependent and controlled by the phospholipid composition of the target membrane. Biochemistry 2012;51:3310–20. https://doi.org/10.1021/bi201690z.

[253] Gilbert NC, Neau DB, Newcomer ME. Expression of an 8R-lipoxygenase from the Coral Plexaura homomalla. Methods Enzymol 2018;605:33–49. https://doi.org/10.1016/bs.mie.2018.02.010n.

[254] Shahane G, Ding W, Palaiokostas M, Orsi M. Physical properties of model biological lipid bilayers: insights from all-atom molecular dynamics simulations. J Mol Model 2019;25:76. https://doi.org/10.1007/s00894-019-3964-0.

[255] Bender G, Schexnaydre EE, Murphy RC, Uhlson C, Newcomer ME. Membrane-dependent activities of human 15-lox-2 and its murine counterpart implications for murine models of atherosclerosis. J Biol Chem 2016;291:19413–24. https://doi.org/10.1074/jbc.M116.741454.

[256] Chen XS, Funk CD. The N-terminal "β-barrel" domain of 5-lipoxygenase is essential for nuclear membrane translocation. J Biol Chem 2001;276:811–8. https://doi.org/10.1074/jbc.M008203200.

[257] Hodnik V, Anderluh G. Capture of Intact Liposomes on Biacore Sensor Chips for Protein-Membrane Interaction Studies. Life Sci 2010:255. https://doi.org/10.1007/978-1-60761-670-2_13.

[258] Erb EM, Chen X, Allen S, Roberts CJ, Tendler SJB, Davies MC, et al. Characterization of the surfaces generated by liposome binding to the modified dextran matrix of a surface plasmon resonance sensor chip. Anal Biochem 2000;280:29–35. https://doi.org/10.1006/abio.1999.4469.

[259] Lomize AL, Todd SC, Pogozheva ID. Spatial arrangement of proteins in planar and curved membranes by PPM 3.0. Protein Sci 2022;31:209. https://doi.org/10.1002/PRO.4219.

[260] Hu S, Sharma SC, Scouras AD, Soudackov A V., Carr CAM, Hammes-Schiffer S, et al. Extremely elevated room-temperature kinetic isotope effects quantify the critical role of barrier width in enzymatic C-H activation. J Am Chem Soc 2014;136:8157–60. https://doi.org/https://doi.org/10.1021/ja502726s.

[261] Lomize MA, Pogozheva ID, Joo H, Mosberg HI, Lomize AL. OPM database and PPM web server: resources for positioning of proteins in membranes. Nucleic Acids Res 2012;40:D370–6. https://doi.org/10.1093/nar/gkr703.

[262] Saeedimasine M, Montanino A, Kleiven S, Villa A. Role of lipid composition on the structural and mechanical features of axonal membranes: a molecular simulation study. Sci Rep 2019;9:8000. https://doi.org/10.1038/s41598-019-44318-9.

[263] Schuck P, Zhao H. The Role of Mass Transport Limitation and Surface Heterogeneity in the Biophysical Characterization of Macromolecular Binding Processes by SPR Biosensing. Humana Press 2010;627. https://doi.org/10.1007/978-1-60761-670-2_2.

[264] Myres GJ, Kitt JP, Harris JM. Inter-Leaflet Phospholipid Exchange Impacts the Ligand Density Available for Protein Binding at Supported Lipid Bilayers 2023;21:57. https://doi.org/10.1021/acs.langmuir.2c00526.

[265] Droege KD, Keithly ME, Sanders CR, Armstrong RN, Thompson MK. Structural Dynamics of 15-Lipoxygenase-2 via Hydrogen−Deuterium Exchange. Biochemistry 2023;21:1. https://doi.org/10.1021/acs.biochem.7b00559.

[266] Walther M, Wiesner R, Kuhn H. Investigations into calcium-dependent membrane

association of 15-lipoxygenase-1: Mechanistic roles of surface-exposed hydrophobic amino acids and calcium. J Biol Chem 2004;279:3717–25. https://doi.org/10.1074/JBC.M309564200.

[267] Klinman JP, Offenbacher AR. Understanding Biological Hydrogen Transfer Through the Lens of Temperature Dependent Kinetic Isotope Effects. Acc Chem Res 2018;51:1966–74. https://doi.org/10.1021/acs.accounts.8b00226.

[268] Choi J, Chon JK, Kim S, Shin W. Conformational flexibility in mammalian 15S-lipoxygenase: Reinterpretation of the crystallographic data. Proteins Struct Funct Bioinforma 2007;70:1023–32. https://doi.org/10.1002/prot.21590.

[269] Tatulian SA, Steczko J, Minor W. Uncovering a calcium-regulated membrane-binding mechanism for soybean lipoxygenase-1. Biochemistry 1998;37:15481–90. https://doi.org/10.1021/BI981062T/ASSET/IMAGES/LARGE/BI981062TF00008.JPEG.

[270] Neau DB, Gilbert NC, Bartlett SG, Boeglin W, Brash AR, Newcomer ME. The 1.85 Å structure of an 8R-lipoxygenase suggests a general model for lipoxygenase product specificity. Biochemistry 2009;48:7906–15. https://doi.org/10.1021/bi900084m.

[271] Pande AH, Moe D, Nemec KN, Qin S, Tan S, Tatulian SA. Modulation of human 5-lipoxygenase activity by membrane lipids. Biochemistry 2004;43:14653–66. https://doi.org/10.1021/bi048775y.

[272] Schneider CS, Bhargav AG, Perez JG, Wadajkar AS, Winkles JA, Woodworth GF, et al. Surface plasmon resonance as a high throughput method to evaluate specific and non-specific binding of nanotherapeutics. J Control Release 2015;219:331–44. https://doi.org/10.1016/j.jconrel.2015.09.048.

[273] Du L, Zhang Z, Luo X, Chen K, Shen X, Jiang H. Binding investigation of human 5-lipoxygenase with its inhibitors by SPR technology correlating with molecular docking simulation. J Biochem 2006;139:715–23. https://doi.org/10.1093/jb/mvj084.

[274] Somvanshi RK, Singh AK, Saxena M, Mishra B, Dey S. Development of novel peptide inhibitor of Lipoxygenase based on biochemical and BIAcore evidences. Biochim

Biophys Acta - Proteins Proteomics 2008;1784:1812–7. https://doi.org/10.1016/J.BBAPAP.2008.07.004.

[275] Ak S, Singh R, Naz F, Chauhan SS, Dinda A. Structure Based Design and Synthesis of Peptide Inhibitor of Human LOX-12: In Vitro and In Vivo Analysis of a Novel Therapeutic Agent for Breast Cancer. PLoS One 2012;7:32521. https://doi.org/10.1371/journal.pone.0032521.

APPENDIX

X-ray crystallography

X-ray crystallography utilizes the diffraction patterns of a crystal sample to yield a three-dimensional molecular structure for making structural determinations of proteins or biological macromolecules. Protein crystallography relies on the high quality and high concentration of homogenous, soluble protein. Crystals are grown from protein solutions containing a minimum protein concentration of 2 mg/mL that are induced to slowly precipitate out of solution in the form of a crystal under given crystallization conditions [83]. These conditions can be obtained through commercially available "crystal screen" packages containing 50 solutions of differing precipitant, buffer, pH, and salt [83].

Crystallization of the protein occurs through vapor diffusion of the 2 μL protein drop solution exchanged with the 50 μL crystallization solution. There are two techniques for establishing vapor diffusion. One method is manually introducing a "hanging drop" of the sample suspended from a glass coverslip over a reservoir of the crystallization condition while the other method is mechanically establishing vapor diffusion through a "sitting drop," which utilizes a Crystal Gryphon instrument that dispenses the proteins into 96-channel crystal plates containing the crystallization condition [83]. To co-crystallize a compound in complex with the protein, small volumes of compound solution are added to the protein sample prior to vapor diffusion.

Crystal growth can take days or weeks at room temperature, but once crystals are obtained, they are introduced into a cryoprotectant buffer so that they may be collected and stored in liquid nitrogen to prevent degradation. In our study, our proteins are then sent to the Southeast Regional Collaborative Access Team (SER-CAT) located at the Advanced Photon Source at the Argonne

National Laboratory for diffraction analysis.

When the crystals are ready for analysis of their X-ray diffraction behavior, they are remotely and mechanically mounted within a rotating spindle and adjusted for proper alignment with the X-ray beam using what is called a goniometer head (**Fig. 42**). The X-rays are generated by accelerating electrons in a synchrotron storage ring and the X-rays focused into a beam. The diameter of the incident beam, as well as alignment of the crystal must be carefully adjusted to ensure proper exposure and prevent radiation damage [83].

The crystal rotates within the spindle perpendicularly to the incident beam to provide the maximum amount of diffraction data. After exposure, diffraction spots are collected on an imaging plate revealing a digitized image (**Fig. 42**). The diffraction spots are resolved based on the distance from the crystal to the detector with the highest resolution accounting for the edge of the detection. The diffraction image becomes strongest at lower resolutions, with a resolution of 3 Å sufficient for detecting amino acid side chains in an electron density map. The resolution of the image is dictated by how well-ordered the molecular subunit is. Within the crystal, the unit cell is defined as the smallest repeating unit that makes up the crystal with dimensions given by three lengths and angles. The dimensions of the unit cells are proportional to the number of diffraction spots for a given unit area. The space group of the crystal is given by the symmetry of the diffraction pattern, which is defined by the configuration and packing of the molecules within the crystal lattice. The more symmetrical a crystal structure, the more ordered the diffraction spots appear in image [83].

Computer programs with well-established algorithms for making structural determinations are employed to resolve the unit cell dimensions and orientation, measure the intensity of

diffraction spots, and normalize spot intensity using a scale factor to relate all the information within the data set. Given the availability of published models of related protein structures, molecular replacement is utilized to reconstruct amino acid sequences within the protein. Combining the "borrowed" model and diffraction data produces electron density maps that account for the location and bond angles of amino acid sequences of the protein. The crystal structure must undergo several rounds of model building and refinement to establish the structural molecular nuances of a given protein. The quality of the protein model is calculated from the reflection data and given by an R-factor [83]. The lower the R-value, the closer the fit is to "perfect." In general, a random set of atoms produces an R-value of approximately 0.63, whereas high-quality published structures aim for values of about 0.20.

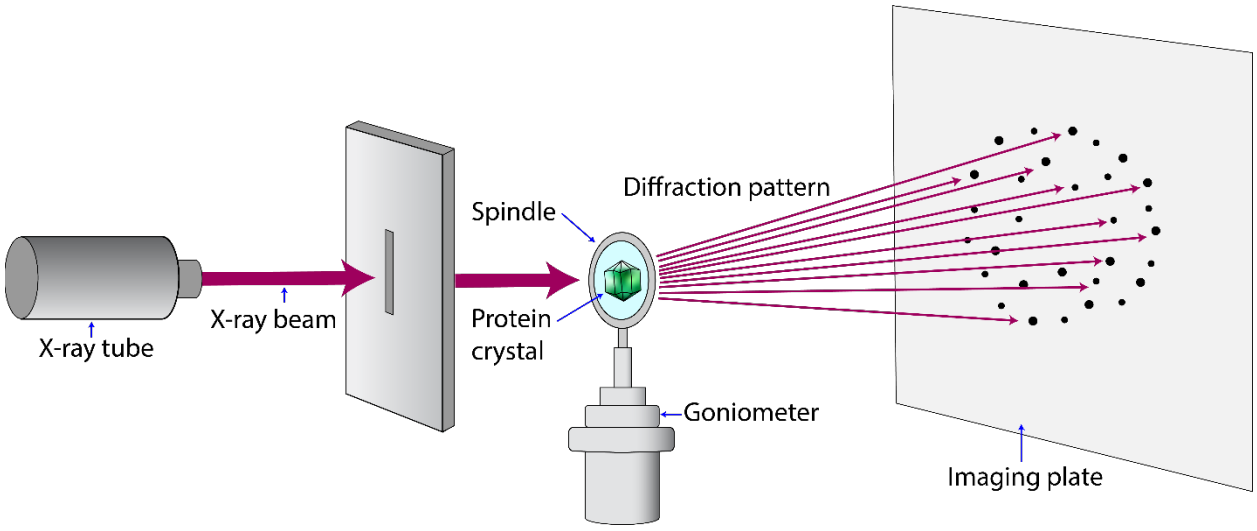


FIGURE 42. Schematic diagram of X-ray crystallography.