University of Nevada, Reno

Elucidating the Role of Quorum Sensing (QS) and Developing QS Modulators for *Streptococcus pneumoniae* and Its Close Commensal Relative, *Streptococcus mitis*

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemistry

By

Tahmina Ahmed Milly

Dr. Yftah Tal-Gan/Dissertation Advisor

December 2022

THE GRADUATE SCHOOL

We recommend that the dissertation

prepared under our supervision by

Tahmina Ahmed Milly

entitled

Elucidating the Role of Quorum Sensing (QS) and Developing QS Modulators for *Streptococcus pneumoniae* and Its Close Commensal Relative, *Streptococcus mitis*

be accepted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

Yftah Tal-Gan, Ph.D.

Advisor

Thomas W. Bell, Ph.D.

Committee Member

Laina M. Geary, Ph.D.

Committee Member

Subhash Verma, Ph.D.

Committee Member

David P. Aucoin, Ph.D.

Graduate School Representative

Markus Kemmelmeier, Ph.D., Dean, Graduate School

December 2022

Abstract

Quorum sensing (QS) is a ubiquitous communication mechanism in bacteria that controls bacterial group behavior phenotypes, such as competence, biofilm formation, and virulence factor production. *Streptococcus pneumoniae*, an opportunistic pathogen, and *Streptococcus mitis* are prototypes of commensal bacteria in the Mitis group and share >80% of their genes. Both *S. mitis* and *S. pneumoniae* utilize a peptide pheromone (competence stimulating peptide, CSP), which binds to its membrane bound ComD receptor to induce QS responses and different pathogenic phenotypes. The aim of this study was to target this non- essential bacterial communication pathway through impediment of the peptide-receptor interaction by using synthetic CSP analogs, thereby circumventing a key issue with traditional antibiotics, the introduction of selective pressure for resistance development. To this end, I have conducted comprehensive structure-activity relationship (SAR) analyses of both *S. pneumoniae* and *S. mitis* CSPs to gain a deeper understanding of the molecular mechanisms that drive these QS circuitries. The SAR results revealed several interesting activity trends and uncovered several CSP-based QS modulators with distinct activity profiles. Additionally, I evaluated several phenotypes that can be utilized to assess the effect of lead peptide analogs on the expression of group behavior genes. In addition to yielding a series of new QS activators and inhibitors, key SAR knowledge of the CSP pheromones obtained through this study, can be utilized for the rational design of highly potent, pharmacologically stable CSP-based QS modulators with therapeutic potential.

Acknowledgments

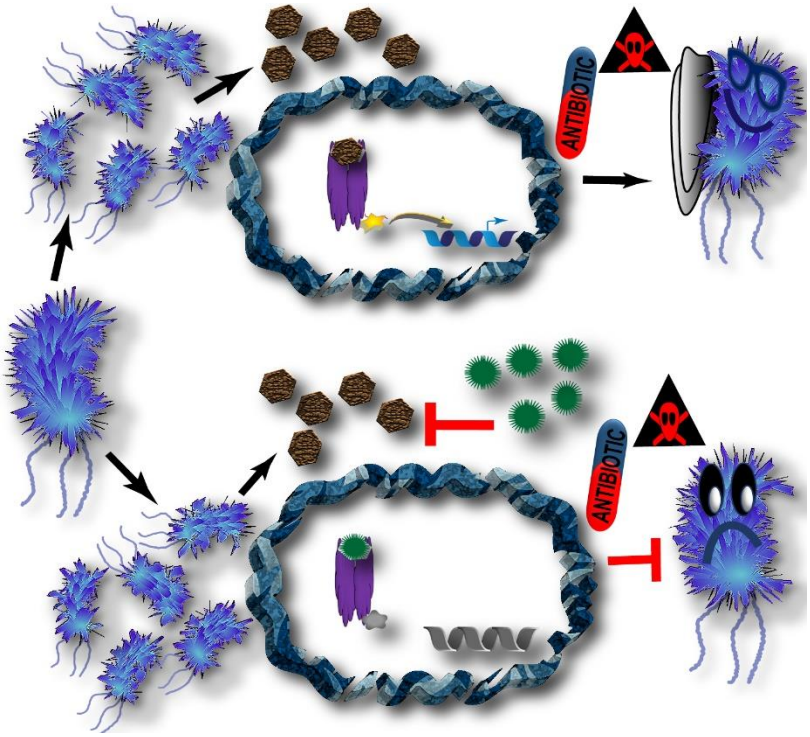
First and foremost, I would like to thank my supervisor, Dr. Yftah Tal-Gan, for his invaluable advice, continuous support, and guidance during my Ph.D. studies. I am grateful to him for opening my eyes to new stages of opportunity and strength.

I would like to express the gratitude to my husband for his uncountable sacrifice, help and support throughout my graduate life. It wouldn't have been possible to complete my studies without his immense support. I am also pleased to say thank you to my one-year-old son for making me stronger than I could ever imagine. I consider myself nothing without them. I would like to thank my father, who always gave me encouragement and motivation to accomplish my goals.

I am also thankful to my collaborator, Dr. Bertucci, for his help and insightful suggestions. I would like to thank all my committee members: Dr. Thomas W. Bell, Dr. Laina M. Geary, Dr. David P. Aucoin, and Dr. Subhash Verma for their time and support during my Ph.D.

In addition, this work was supported by grants from the National Institutes of Health (R35GM128651 and R01HL142626) and the National Science Foundation (CHE-1808370).

Table of Contents



Chapter 1: Introduction (Targeting peptide-based quorum sensing systems for the treatment of gram-positive bacterial infections)[a]

[a]Reprinted with permission from Milly, T. A.; Tal-Gan, Y. Targeting peptide-based quorum sensing systems for the treatment of gram-positive bacterial infections. *J. Pept. Sci.* **2022**, e24298. Copyright 2022 John Wiley and Sons Ltd.

Targeting Peptide-Based Quorum Sensing Systems for the Treatment of Gram-Positive Bacterial Infections

Tahmina A. Milly and Yftah Tal-Gan

Department of Chemistry, University of Nevada, Reno, 1664 North Virginia Street, Reno, Nevada, 89557, United States

Abstract

Bacteria utilize a cell density-dependent communication system called quorum sensing (QS) to coordinate group behaviors. In Gram-positive bacteria, QS involves the production of and response to auto-inducing peptide (AIP) signaling molecules to modulate group phenotypes, including pathogenicity. As such, this bacterial communication system has been identified as a potential therapeutic target against bacterial infections. More specifically, developing synthetic modulators derived from the native peptide signal paves a new way to selectively block the pathogenic behaviors associated with this signaling system. Moreover, rational design and development of potent synthetic peptide modulators allows in depth understanding of the molecular mechanisms that drive QS circuits in diverse bacterial species. Overall, studies aimed at understanding the role of QS in microbial social behavior could result in the accumulation of significant knowledge of microbial interactions, and consequently lead to the development of alternative therapeutic agents to treat bacterial infectivity. In this review, we discuss recent advances in the development of peptide-based modulators to target QS systems in Gram-positive pathogens, with a focus on evaluating the therapeutic potential associated with these bacterial signaling pathways.

Introduction

In nature bacteria regulate a vast range of social behaviors in a synchronized manner with members of their community using a universal cell-cell signaling mechanism commonly referred to as quorum sensing (QS).[1,2] Bacteria utilize this communication system to monitor their population density in each environment and to trigger population-wide changes in gene expression when the population reaches a critical cell density.[3] This population size-dependent pathway has appeared as a popular model for understanding bacterial sociality as this signaling system allows the microbial community to behave in a coordinated manner to direct processes contributing to virulence factor production, biofilm formation, competence development, and other pathogenic or symbiotic interactions.[1-5]

Initiation of bacterial communication depends on the production, secretion and detection of small chemical signals called autoinducers. There are several types of autoinducers: Gram-negative bacteria generally utilize small molecules known as *N*-acyl-homoserine lactones (AHLs), whereas Gram-positive bacteria generally utilize peptide-based molecules named auto-inducing peptides (AIPs). These two classes of autoinducers are thought to be highly specific and are generally used for intraspecies cell-cell communication. However, it is becoming increasingly evident that these two signal classes are not only limited to intraspecies chemical language, rather, bacteria also use these signaling molecules as interspecies communication devices.[3,4,6] Moreover, auto-inducer 2 (AI-2) has been found in over 70 species of Gram-negative and Gram-positive bacteria, and has been recognized as a universal interspecies signaling molecule. QS circuits rely on the concentration of the secreted signal in the extracellular environment, and once it

reaches a threshold level, the signal binds and activates a receptor protein, initiating a coordinated change in gene expression in the population.[4,5,7]

There is a great deal of research showing the link between QS and bacterial pathogenicity.[5,7,8] Increased rate of antibiotic resistance development among bacteria due to the overuse and misuse of antibiotics has also been well-documented. These two factors: involvement of QS circuits in bacterial infectivity and increased rate of resistance development, offer a novel approach to traditional antimicrobials that would allow the treatment of bacterial infections through targeting of QS systems while avoiding selective pressure for resistance development. Both traditional methods and bioinformatics techniques have revealed the presence of homologues, analogues or similar QS systems that use small peptide signals across Gram-positive genera.[9] The potential for commonality in aspects of these communication systems paves the way to identifying therapeutics that could target multiple pathogens. In addition, understanding how bacteria interact with one another using QS within polymicrobial communities can improve our current understanding of microbial community interactions and guide the design of new antibacterial therapeutics to treat bacterial co-infections. In this review, we will focus on the advances that have been made in understanding the structure and function of several peptide-based Gram-positive bacterial QS systems with the intent to highlight the therapeutic potential of these QS systems to treat microbial infections.

Peptide-Based QS mechanisms in Gram-positive bacteria

Based on features of the peptide signaling molecules and their receptors, Gram-positive bacterial QS signaling pathways have been classified into two groups. The peptide

signaling molecules, which are central to QS activation, are ribosomally synthesized as precursor peptides and oftentimes undergo post-translational modifications. Following transcription and translation, the precursor peptide is generally processed to the mature peptide signal and secreted into the extracellular environment via a dedicated ATP-binding cassette (ABC) transporter. These mature peptide signals can be linear or macrocyclic.[10] As the cell density increases, the concentration of processed peptide signal increases, and once it reaches a certain threshold concentration, the peptide binds to its cognate receptor, either in a self-signaling pathway or a two-component pathway. The self-signaling pathway is known as the RRNPP protein family, which stands for **R**ap (*Bacillus subtilis*), **R**gg (*Streptococcus*), **N**prR (*Bacillus cereus*), **P**lcR (*B. cereus*), **P**rgX (*Enterococcus faecalis*),[11,12] as these are the response regulators that were identified from various founding species of this circuit class. In this system class, either during or following export, the peptide signal is processed to its mature form. Then, the mature peptide signal is being imported back into the cell by a transmembrane oligopeptide permease (Opp) transporter where it directly binds to its cytoplasmic response regulator, resulting in upregulation of the corresponding QS genes (**Figure 1b**).

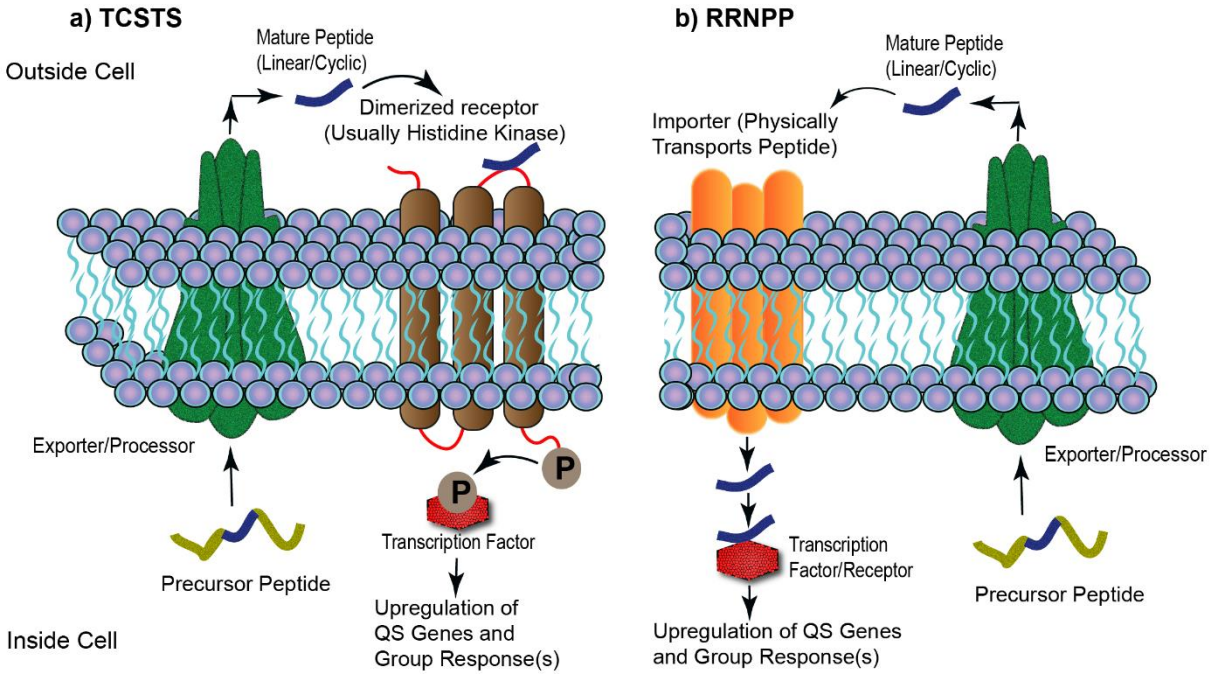


Figure 1: Generalized Common QS circuits in Gram-positive bacteria. a) TCSTS-like QS circuits involve the transduction of mature peptide signal across the membrane without physical transport of the signal; b) RRNPP-like QS circuits involve the physical import of the mature peptide signal before it binds to and activates the cytosolic transcription factor/receptor.

The other most widely studied pathway, which was first described in *Streptococcus pneumoniae* (Com system) and later in many other Gram-positive bacteria, including *Staphylococcus aureus* (Agr system), and *E. faecalis* (Fsr system), is commonly referred to as two-component signal transduction system (TCSTS).[5] This pathway is comprised of a membrane-bound histidine kinase receptor and a separate response regulator that is associated with the induction of QS-dependent phenotypes. In this pathway, the peptide signaling molecule is primarily encoded as precursor peptide and then modified post-translationally. The mature peptide is secreted into the outside of the cell and upon reaching a threshold concentration, it is recognized by the specific histidine kinase receptor. The activated receptor then phosphorylates the downstream response regulator. The activated

regulator upregulates the transcription of specific bacterial genes associated with group behaviors and phenotypes as well as transcription of the respective genes related to the AIP secretion pathway itself (**Figure 1a**). TCSTS-based QS systems have been found to regulate a myriad of group behaviors in many Gram-positive species.

PlcR-PapR QS system in *B. cereus*

B. cereus is a food-poisoning pathogen and is genetically close to two other human pathogens: *B. anthracis* and *B. thuringiensis*.[13] Some strains of *B. cereus* can colonize the intestines and cause more severe diseases, such as endophthalmitis or meningitis, through the production and secretion of several virulence factors, including hemolysins, phospholipases, proteases, and enterotoxins.[13,14] In *B. cereus*, the expression of virulence factors is controlled by a QS system that consists of an intracellular PapR-derived AIP (**Table 1**) and the transcription factor PlcR (Phospholipase C Regulator).[15] The initiation of the QS signaling pathway begins with the production of the PapR precursor peptide, which needs to be transported outside the cell by a membrane carrier protein and processed by the secreted neutral protease B (NprB) into the mature AIP signal. The processed AIP is reimported back into the cell through the Opp, where it binds to PlcR, stimulating the expression of several genes involved in the secretion of virulence factors (**Figure 2**).[13-15] Previous studies showed that inactivation of the *plcR* gene reduces the secretion of virulence factors but cannot fully eliminate virulence factor secretion. This is due to the involvement of several additional QS systems in the regulation of virulence factors.[8,14]

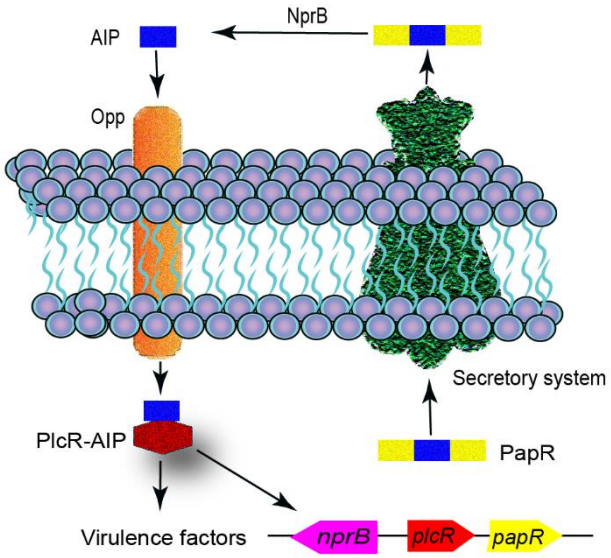


Figure 2: *B. cereus* PlcR-PapR QS system. The precursor peptide PapR is secreted and is then processed to the mature AIP by the extracellular protease, NprB. Opp transports the mature AIP back into the cell and then the intracellular AIP binds to the transcription factor, PlcR. The resultant PlcR-AIP complex upregulates virulence factor production and activates the expression of *papR*.

Table 1. Signaling peptides of Gram-positive bacteria. In the peptide sequences, cysteine (red) is marked to highlight the residue required for thiolactone ring formation. Macrocyclic part marked in parenthesis.

Species	Peptide regulator family	QS System	Peptide Name	Peptide Sequence
Bacillus cereus	RRNPP family	*papR-plcR*	PapR I(PapR$_7$) PapR II PapR III PapR IV PapRa	ADLPFEF SDMPFEF NEVPFEF SDLPFEH CSIPYEY
		nprR-nprX	NprRB	SKPDIVG
Bacillus anthracis *Bacillus thuringiensis*	RRNPP family	*nprR-nprX*	NprRB	SKPDI SDIYG
Staphylococcus aureus	TCSTS (Agr)	*agrBDCA*	Sa AIP 1 Sa AIP 2 Sa AIP 3 Sa AIP 4	YST(CDFIM) GVNA(CSSLF) IN(CDFLL) YST(CYFIM)
Streptococcus pneumoniae	TCSTS(Com)	*comCDE*	CSP1 CSP2	EMRLSKFFRDFILQRKK EMRISRIILDFLFLRKK
	Rgg family	*rgg-shp*	SHP	DIIIIVGG
Streptococcus mitis	TCSTS(Com)	*comCDE*	*S. mitis*-CSP-2	EIRQTHNIFFNFFKRR
	Rgg family	*rgg-shp*	SHP	DIIIVGG
Streptococcus pyogenes	Rgg family	*rgg2-shp2* *rgg3-shp3*	SHP2 SHP3	DILIIVGG DIIIIVGG
Streptococcus agalactiae	Rgg family	*rovS- shp2*	SHP1520	DILIIVGG
Streptococcus dysgalactiae subsp. equisimilis	Rgg family	*rgg-shp*	SHP SHP	DILIIVGG DILIIVGG
Streptococcus porcinus	Rgg family	*rgg-shp*	SHP	DIIIIAGG
Streptococcus thermophilus	Rgg family	*rgg-shp*	SHP	DIIIIVGG

In sporulating *Bacillus*, a similar signaling pathway, the NprR-NprX cell–cell communication system, has been shown to regulate sporulation and the expression of virulence genes, including enterotoxins, phospholipases, proteases and chitinases.[11] Like the PlcR-PapR QS system, this system is also comprised of the NprR transcriptional regulator, whose activity depends on the signaling peptide, NprX. The imported mature signal activates the NprR regulator, resulting in the expression of an extracellular protease gene, *nprA*, during the sporulation process.[11,16,17] This QS system was found to be strain specific, with possible cross talk between some phylogenetic groups.[17]

Rgg/SHP QS systems in Streptococci

The Rgg (regulator gene of glucosyltransferase) and SHP (short hydrophobic peptide) cell-to-cell communication system is widespread in the *streptococcus* genus and has been gaining increased attention due to its involvement in controlling a variety of functions in several species of streptococci.[5,18] The Rgg transcriptional regulator was first described in *S. gordonii* and later found in nearly all streptococci species.[19-21] Rgg regulators are regulated by re-internalized SHPs acting as pheromones. This QS system regulates different streptococcal phenotypes, such as bacteriocin production in *S. mutans*,[21] as well as the production of the exotoxin SpeB, virulence regulation and biofilm development in *S. pyogenes*.[22-24] Moreover, this pheromone system is widely utilized by members of the mitis group of streptococci, and an Rgg/SHP system (**Figure 3**) regulating surface polysaccharide expression in *S. pneumoniae* has been reported.[25] In this QS system, first the SHP precursor peptide is produced, processed and exported with the help of the PptAB transporter system and the membrane protease Eep. The mature peptide, SHP, is then transported from the extracellular environment to the cytosol through an Opp. Once the

mature peptide is bound by the Rgg regulator, the Rgg:SHP complex activates the pheromone-feedback loop and an operon consisting of 12 genes. Activation of this regulon upregulates polysaccharide synthesis and downregulates biofilm formation (**Figure 3**).

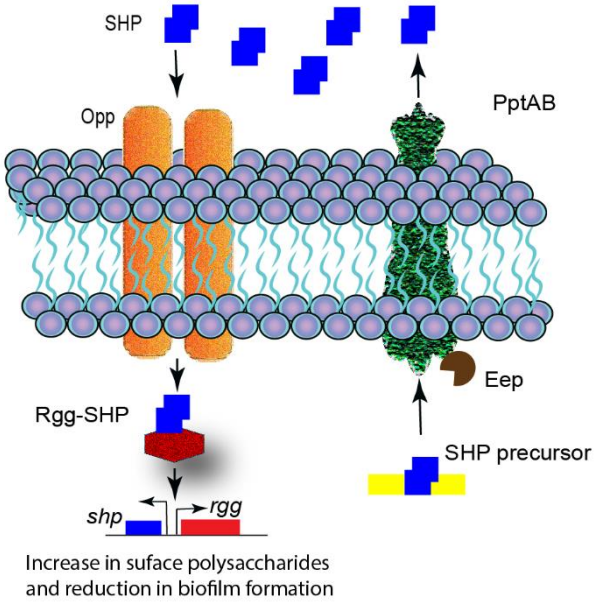


Figure 3: *S. pneumoniae* Rgg/SHP QS system. The SHP precursor peptide is processed and exported by PptAB and Eep. When the SHP concentration reaches a certain threshold, the mature peptide is imported by an Opp and binds to the Rgg regulator. The resultant Rgg:SHP complex upregulates polysaccharide synthesis and downregulates biofilm formation.

Interspecies cross-communication has been found in some of these systems, making such systems an important tool to study bacterial communities, which, in turn assists in the development of new anti-virulence strategies to treat bacterial infections.[26-28]

QS in *S. aureus* (Agr system)

S. aureus is a human pathogen responsible for bacteremia, sepsis, endocarditis, toxic shock syndrome and a wide spectrum of infections of skin and soft tissues. The outbreaks of multidrug-resistant *S. aureus* strains, known as methicillin-resistant *S. aureus* (MRSA) makes treatment of *S. aureus* extremely difficult.[9,29,30] This pathogen secretes a wide range

of virulence determinants to attack the host, including various tissue degrading enzymes, pore-forming toxins, and immune evasion factors. The accessory gene regulator (*agr*) operon-encoded QS system regulates the expression of different virulence factors and is directly related to the pathogenicity of this Gram-positive bacterium.[9,30,31]

The Agr system is known to contain two adjacent transcripts named RNAII and RNAIII. The RNAII transcript comprises of four genes in the transcriptional order of *agrBDCA*. The signaling cascade begins with the production of a 46-amino acid precursor peptide of the AIP signal, AgrD. AgrB is an integral membrane protein that processes the AgrD precursor peptide into its mature cyclic form and exports the cyclic peptide as the functional signal to the extracellular environment. AgrC and AgrA form a classical bacterial TCSTS where AgrC acts as the histidine kinase sensor and AgrA is a response regulator. Upon binding of AIP to AgrC, AgrA gets activated and binds to two promoter regions (P2 for RNAII and P3 for RNAIII) to autoactivate the Agr system and upregulate RNAIII transcription (**Figure 4**).

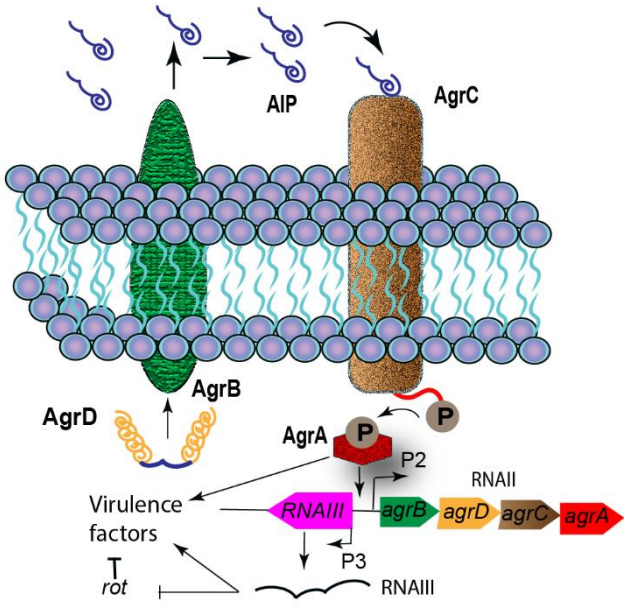


Figure 4: *S. aureus* Agr QS system. The precursor peptide AgrD is processed and secreted across the cell membrane as the mature AIP signal by the AgrB transporter system. Mature AIP signal then binds to the extracellular domain of the AgrC receptor. Upon binding, the histidine kinase domain of AgrC phosphorylates the response regulator, AgrA. Phosphorylated AgrA binds the P2 and P3 promoters to autoactivate the *agr* operon (called RNAII) and the RNAIII regulatory RNA, respectively. RNAIII promotes virulence factor production and represses the expression of *rot*, leading to derepression of virulence factor production.

Upregulation of toxins and virulence determinants depends on either the activation of RNAIII or AgrA.[30,32] RNAIII is the *agr*-induced major regulatory RNA (*rRNA*) molecule and is responsible for the expression of most of the virulence genes through directly affecting mRNA stability and stimulating or inhibiting mRNA translation, or indirectly through RNAIII-mediated inhibition of the Rot transcriptional regulator translation.[32] The toxin repressor protein, Rot, is a member of the Sar and MarR families of transcriptional regulators that can promote or repress the expression of hundreds of toxins, proteases, adhesion factors and metabolic pathways.[29-31] This combination of direct and indirect

RNAIII-mediated gene regulation transforms *S. aureus* cells into a hostile form capable of invasive infection.

AgrA acts as a key transcription factor and positively regulates the expression of the *agr* operon mainly via activating the *agr* P2 and P3 promoters.[9,33] AgrA can additionally activate the expression of the *α*- and *β*-phenol soluble modulins (PSMs).[29] PSMs are short amphipathic cytolytic peptides that have been shown to be critical for staphylococcal pathogenesis by enhancing the survival and dissemination of *S. aureus* in invasive infection.[9,30] PSMβs are important for staphylococcal biofilm maturation, and within the mature biofilm structures, *agr*-mediated PSM promoter activation can be observed.[34] AgrA may be the most important element in the initiation of transcription at P2 and a previous study has shown that the deletion of *agrA* completely abolishes both RNAIII and *agr* mRNA transcription.[33]

The *S. aureus* Agr QS system has been classified into four different specificity groups. Each Agr system (referred to as Agr-I, Agr-II, Agr-III, and Agr-IV) recognizes a different AIP signal (referred to as AIP-I, AIP-II, AIP-III, and AIP-IV, **Table 1**). These AIP signals vary in overall peptide length (7-9 amino acids) and sequences.[9] All the *S. aureus* AIP signals share several structural features: (1) The last five residues in each AIP are constrained as a thiolactone macrocycle between the side chain of a conserved cysteine residues and the *C*-terminal carboxylate. (2) All AIPs contain an *N*-terminal exocyclic "tail" that varies in length (2-4 residues). (3) All AIPs contain 2-3 hydrophobic residues at the *C*-terminal region (within the macrocycle) that are critical for receptor binding. Due to the differences in peptide sequences, these signaling molecules function as cross-type antagonists for AgrC activation in a process named "Agr interference".[9,30] This

phenomenon is typical in this species and leads to competition between different specificity groups of *S. aureus*. The involvement of Agr or analogues QS circuitries in the regulation of virulence in other Gram-positive human pathogens makes the Agr QS system an attractive anti-virulence target for the treatment of multiple pathogen infections.

QS in *S. pneumoniae* (Com system)

S. pneumoniae is a notorious opportunistic human pathogen that causes pneumonia, bacteremia, otitis media, and meningitis in immune compromised human hosts.[35,36] *S. pneumoniae* was the first species of bacteria in which natural transformation was reported in 1928 by Frederick Griffith.[37] Natural transformation ensures genomic plasticity of pneumococcus, which is associated with better adaptation to different environmental stressors and spread crucial features, such as antibiotic resistance. This process takes place only in pneumococcal competent cells, which can acquire DNA from the surrounding environment.[38-40] The competence regulon ComCDE QS circuitry plays a decisive role in regulating competence, biofilm formation and virulence factor production in *S. pneumoniae*.[39,41,42]

In the *S. pneumoniae* ComCDE QS system, the signaling molecule is a 17-residue oligopeptide CSP (competence stimulating peptide, **Table 1**), which is formed from the precursor peptide, called ComC. The CSP is exported out of the bacterial cell by an ATP-binding cassette (ABC) transporter, termed ComAB. Once accumulation of extracellular mature CSP reaches a threshold concentration, it binds to its cognate ComD histidine kinase receptor, which triggers the phosphorylation and activation of a response regulator protein, ComE. Activated ComE then acts as a transcription factor for the *comAB* and *comCDE* genes. ComE also activates the transcription of the *comX* gene, an alternative

sigma factor, also known as *sigX*, which is involved in the expression of a large regulon of effector genes involved in the acquisition of competence within the cells (**Figure 5**).[43-45] Previous studies have shown that deletion of the *comA*, *comB*, *comC*, *comD*, *comE*, as well as *comX* genes reduces competence induction and the severity of pneumococcal infections.[39,42,46]

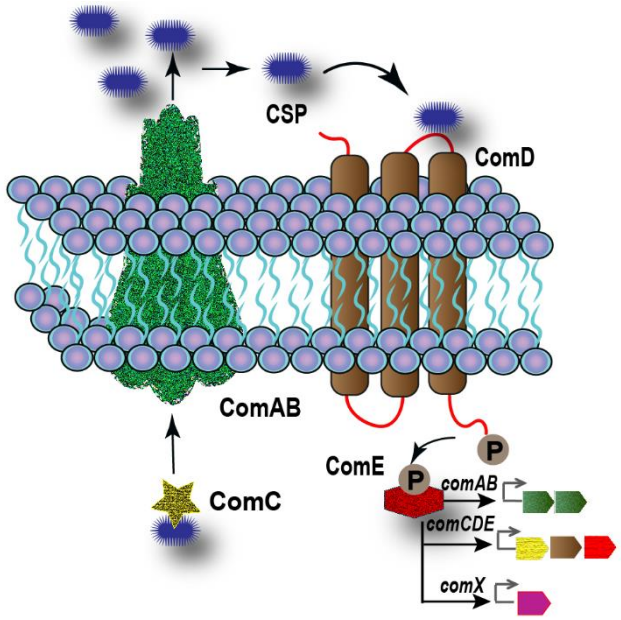


Figure 5: *S. pneumoniae* ComCDE QS system. The precursor peptide ComC is processed to the mature peptide signal, CSP, and transported outside the cell by the ComAB trasporter system. When the CSP concentration reaches a certain threshold, it binds and activates the transmembrane histidine kinase receptor, ComD. Upon activation, ComD undergoes autophosphorylation, leading to phosphorylation of the response regulator, ComE. Phosphorylated ComE then acts a transcription factor to upregulate the expression of the QS genes, *comABCDE*, as well as genes involved in group phenotypes, such as *comX*.

Different strains of pneumococcus can produce different CSP signals and the two major forms of CSP, namely CSP1 and CSP2 (**Table 1**), are used by the majority of pneumococcus strains. Although these two CSP signals share 50% homology in their sequence, they are highly specific toward their respective cognate receptors, ComD1 and ComD2, respectively, and only at higher signal concentration can activate their non-

cognate receptor.[45,47] Streptococci belonging to the mitis and anginosus groups use this ComCDE system to regulate different phenotypes associated with infectivity, making this QS circuit a prime target for the attenuation of pathogenicity in multiple pathogens.

Peptide-based QS systems as therapeutic targets

QS interference is an attractive strategy to limit both the spread of infectious organisms and selective growth pressure that results in the proliferation of resistant organisms. Targeting virulence factors, rather than bacterial growth, has been shown to be effective in controlling bacterial infections without engendering resistance development.[9,48,49] The close connection between QS and pathogenicity has led to the development of novel antimicrobial therapeutic approaches, such as quorum quenching and other QS-blocking methodologies, that focus on controlling virulence factor production.[49,50] Targeting the ability of the microbe to recognize and respond to the AIP signaling molecules has proven to be an effective approach for the treatment of several human pathogens.[9,48] Current technologies, such as solid-phase peptide synthesis (SPPS) and high-performance liquid chromatography (HPLC), facilitate the design and synthesis of numerous peptides or modified peptide analogs that can modulate QS-mediated phenotypic responses *in vitro*.[51,52] By having superior properties like excellent selectivity, remarkable potency, improved stability and low toxicity, peptides, the biological mediators, have gained significant attention to consider them as promising therapeutic candidates.[51] Several reports demonstrated the importance of developing synthetic peptide analogs capable of inhibiting QS circuits in many of the human pathogens, including *S. aureus*, *S. pneumoniae*, and *E. faecalis*.[8,9,14,48-50]

Although the development of potent synthetic peptide analogs capable of inhibiting QS communication systems has been reported, many of these compounds suffer from poor affinity and low pharmacokinetic profiles. To overcome these barriers, the stability of lead peptide sequences can be improved through the utilization of several chemical alterations to the peptide, such as, incorporation of non-proteinogenic amino acids, *N*-methylated residues, or D-amino acids.[53] These modifications could lead to improved protease resistance and enhanced biological activity of the parent compound.[51,52] Peptide cyclization has also been shown to be an effective approach to increase peptide stability, binding affinity, and specificity to the receptor.[54-57]

In vivo studies have been reported determining the efficacy of the peptide modulators in preventing or reducing diseases. For example, our group, together with the Lau lab, reported the development of potent synthetic pneumococcal QS modulators that can attenuate pneumococcal infections *in vivo*.[55,56,58] Blocking the Agr QS system to control virulence in *S. aureus* could be a fruitful target as the pathology of an Agr group IV strain can be attenuated through vaccination with hapten-linked AIP-4.[59] Recently, Blackwell and co-workers reported a stable and potent synthetic peptide-based inhibitor of the *S. aureus* agr system attenuating *S. aureus* (MRSA) infection in an *in vivo* mouse model of skin infection.[60] This work present a new, useful, and modular anti-virulence approach to controlling bacterial skin infections *in vivo*. The following section will detail the use of peptide-based probes in modulating QS circuitries and their related phenotypes in important Gram-positive human pathogenic bacteria.

Agr QS system as therapeutic target

As the expression of many virulence genes is regulated by the Agr QS system, targeting Agr could result in the reduction in *S. aureus* pathogenicity.[61,62] Various methods have been reported to block Agr function and attracted significant attention as potential anti-infective therapies to prevent *S. aureus* infections.[9,63-72] Among them, interference of AIP:receptor interactions represents a direct strategy to block QS, and may have the potential to prevent virulence in *S. aureus*.[66,73] Development of non-native synthetic ligands, such as, small peptides and macromolecules capable of inhibition of the AgrC receptor have been promoted by many researchers.[59,71,73-77] Initial studies have shown that each of the four native staphylococcal AIPs (**Table 1**) can antagonize the other three, noncognate AgrC receptors to prevent activation of the Agr regulon.[30,71,74,78] Previously, an *in vivo* study conducted by Wright et al. showed that *S. aureus* pathologies can be reduced by blocking the agr function.[79] The authors showed that by injecting inhibitory concentrations of the AIP-II peptide in a mouse dermonecrosis model, ulcers and abscesses caused by an Agr-I strain could be impeded.

There are numerous examples of developing AIP:AgrC modulators based on the AIP-I and AIP-II signals.[69,71,74,78,80-83] Detailed Structure-Activity Relationship (SAR) studies of AIPs -I and -II were conducted by Muir, Novick, Williams, and co-workers and several potent inhibitors capable of inhibiting both cognate and noncognate AgrC receptors were developed.[30,84] The SAR studies revealed that the AIP macrocycle is important for initial receptor recognition/binding, and the involvement of the AIP exocyclic tail in interactions is responsible for receptor activation. Hydrophobic residues at the *C*-terminal region (residues 6−8 for AIP-I and residues 8 and 9 for AIP-II) were found to be important for

both AgrC cognate and noncognate receptor binding, and based on this analysis, the most potent inhibitor, tAIP-I D2A, was developed, which is a truncated version of AIP-I lacking an exocyclic tail and having an aspartic acid to alanine mutation in the macrocyclic core.

Studies by Blackwell and co-workers closely examined the SAR of AIP-III (**Figure 6a**) and through the design, synthesis, and biological testing of a series of first- and second-generation AIP-III mimetics, the authors identified a set of non-native AIP-III analogs that can inhibit all four staphylococcal AgrC receptors with picomolar IC_{50} values.[85] The systematic SAR analysis of the AIP-III analogs revealed several important structural features that are responsible for the ability of AIP-III to modulate the four AgrC receptors. The AIP-III SAR analysis was consistent with the previous AIPs -I and -II SAR analyses, emphasizing the role of the endocyclic hydrophobic residues (residues 5-7 in AIP-III) in receptor binding and the exocyclic tail in receptor activation. Moreover, the most potent global AgrC inhibitor, AIP-III D4A (**Figure 6b**), was developed through the inclusion of a previously reported identical mutation (replacement of Asp4 with alanine), emphasizing the previous findings regarding the importance of the Asp4 side chain in AgrC receptor activation. This global inhibitor was able to block QS-regulated hemolysis in all four *S. aureus* specificity groups (Agr-I−IV) at sub-nanomolar concentrations. The production of toxic shock syndrome toxin-1 (TSST-1) is a hallmark QS phenotype in group-III *S. aureus*, and this analog was also able to attenuate TSST-1 production in a *S. aureus* group-III strain by over 80% at nanomolar concentrations. However, the presence of a hydrolytically unstable thioester linkage in AIP-III D4A makes this most potent inhibitor less stable and soluble in aqueous media. To address these issues, a new AIP-III analog containing a non-

native amide bridge, AIP-III D4A Amide (**Figure 6c**), was developed, and exhibited improved stability and solubility in aqueous media.[75]

As studies suggested that AIP-III might provide a superior scaffold for the development of potent AgrC inhibitors, Blackwell and co-workers further conducted in depth structural analysis of all four native *S. aureus* AIPs, along with several AIP-III analogs, using 2D NMR to obtain valuable insight as to the structural features that drive AIP-mediated AgrC activation and inhibition.[73,76,86] The NMR data revealed two important structural motifs within AIP-type ligands that are required for AgrC receptors modulation: (i) the presence of a triangular, hydrophobic knob motif on the macrocycle that is needed for receptor binding, and (ii) a fourth hydrophobic contact or anchor on the *N*-terminal tail that is essential for receptor activation. These structural findings can be utilized for the design of new staphylococcal modulators with better potency and selectivity.

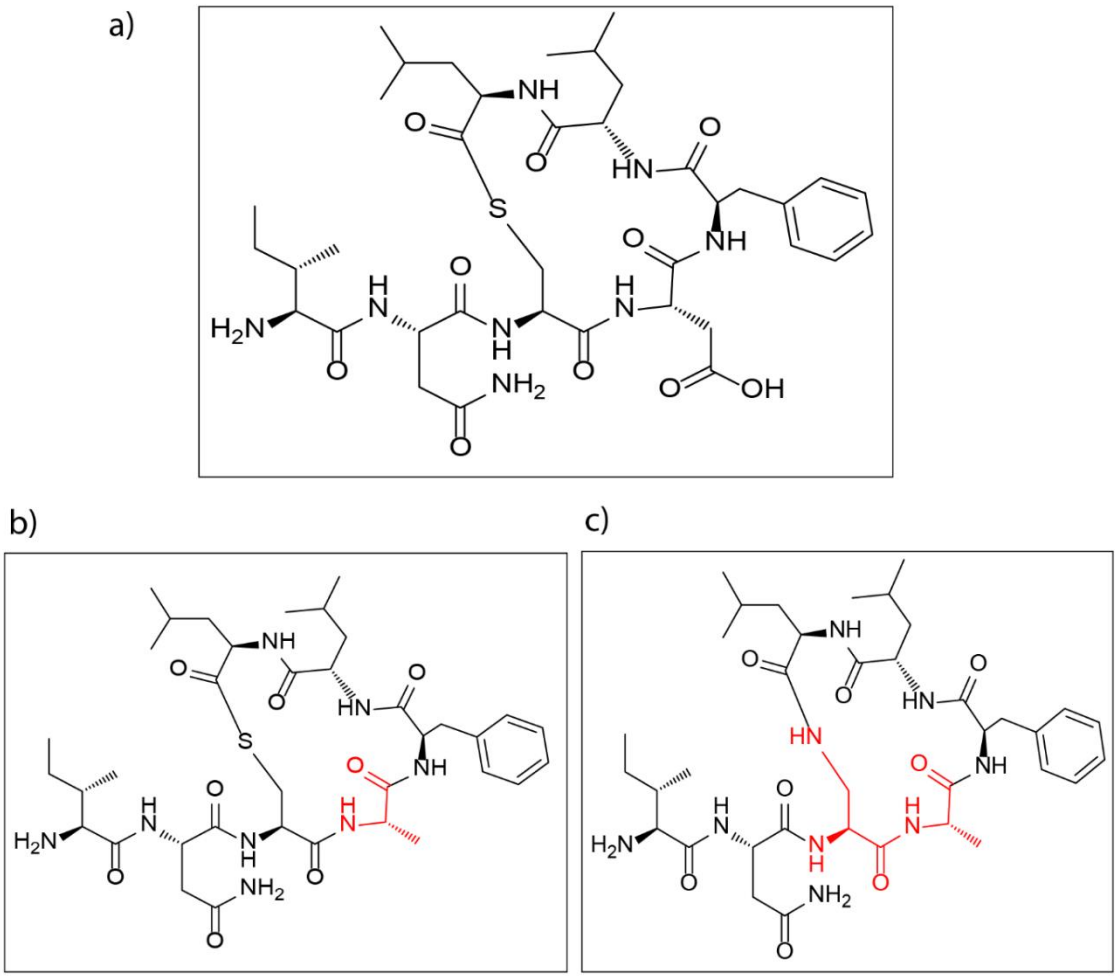


Figure 6. Lead Inhibitors of the *S. aureus* Agr QS circuitry. a) Structure of the native AIP-III; b) Structure of the lead inhibitor, AIP-III D4A; and c) Structure of the inhibitor AIP-III D4A Amide, which was found to inhibit all four *S. aureus* specificity groups (Agr-I-IV) while exhibiting improved pharmacological properties. Red color indicates the modifications at specific amino acids in AIP-III. Here, D4A represents that Asp at position 4 was replaced by Ala.

Com QS system as therapeutic target

The streptococcal Com QS system controls bacterial competence and biofilm formation in numerous streptococci. While competence is not directly related to bacterial pathogenesis in most of the streptococcal species, *S. pneumoniae* ability to utilize exogenous genetic material has been recognized as an important factor for genetic variability and subsequent

evolution. For instance, previous work has shown that non-opsonizing antibodies against pneumococcal polysaccharide capsular antigens alter the expression of QS genes and thus improve bacterial viability and pathogenesis.[87] Another significant QS-regulated phenotype, biofilm formation, enhances pneumococcal survival by facilitating evasion of host immune responses during pneumococcal colonization of the nasopharynx.[41,88] A novel biofilm regulating peptide, BriC, which is upregulated by the ComE transcriptional regulator, has been shown to act as a molecular linker of pneumococcal competence, biofilm formation, and colonization.[88] Lastly, *S. pneumoniae* produces a wide range of toxins and virulence factors, including Ply, autolysin, choline-binding proteins, lipoproteins, LPXTG cell wall bound proteins, capsule polysaccharide, cell wall polysaccharide, and Immunoglobulin A1 (IgA1) protease, and the production of many of these virulence factors is influenced by the competence regulon (ComCDE system).[40,89] Thus, to control pneumococcus infections, targeting this nonessential communication system could be considered as a potential anti-virulence approach.

Extensive work has been done by our lab focusing on the development of CSP-based QS-modulators in *S. pneumoniae.* To do so, systematic alanine and epimer scans were conducted on both pneumococcal CSPs (CSP1 and CSP2, **Table 1**) to evaluate the structural and functional properties of each residue, eventually leading to the development of several potent QS-modulators of both *S. pneumoniae* serotype groups.[47,90] It has been found that *N*-terminal amino acids play critical roles in both CSPs activity, whereas the last three *C*-terminal amino acids are unnecessary for activity. Evaluation of the CSP1 central region revealed that the hydrophobic residues (L4, F7, F8, F11, I12) are critical for receptor binding (**Figure 7**). Follow-up studies have focused on optimizing the degree of occupancy

of the CSP1 binding pockets within the ComD1 receptor by utilizing conservative substitutions at these key hydrophobic residues of CSP1.[91,92,93] Through the incorporation of bulkier, hydrophobic nonproteogenic amino acids, Milly et al. were able to develop pneumococcal QS modulators with high potency and superior metabolic stability, while remaining nontoxic against mammalian cells.[93]

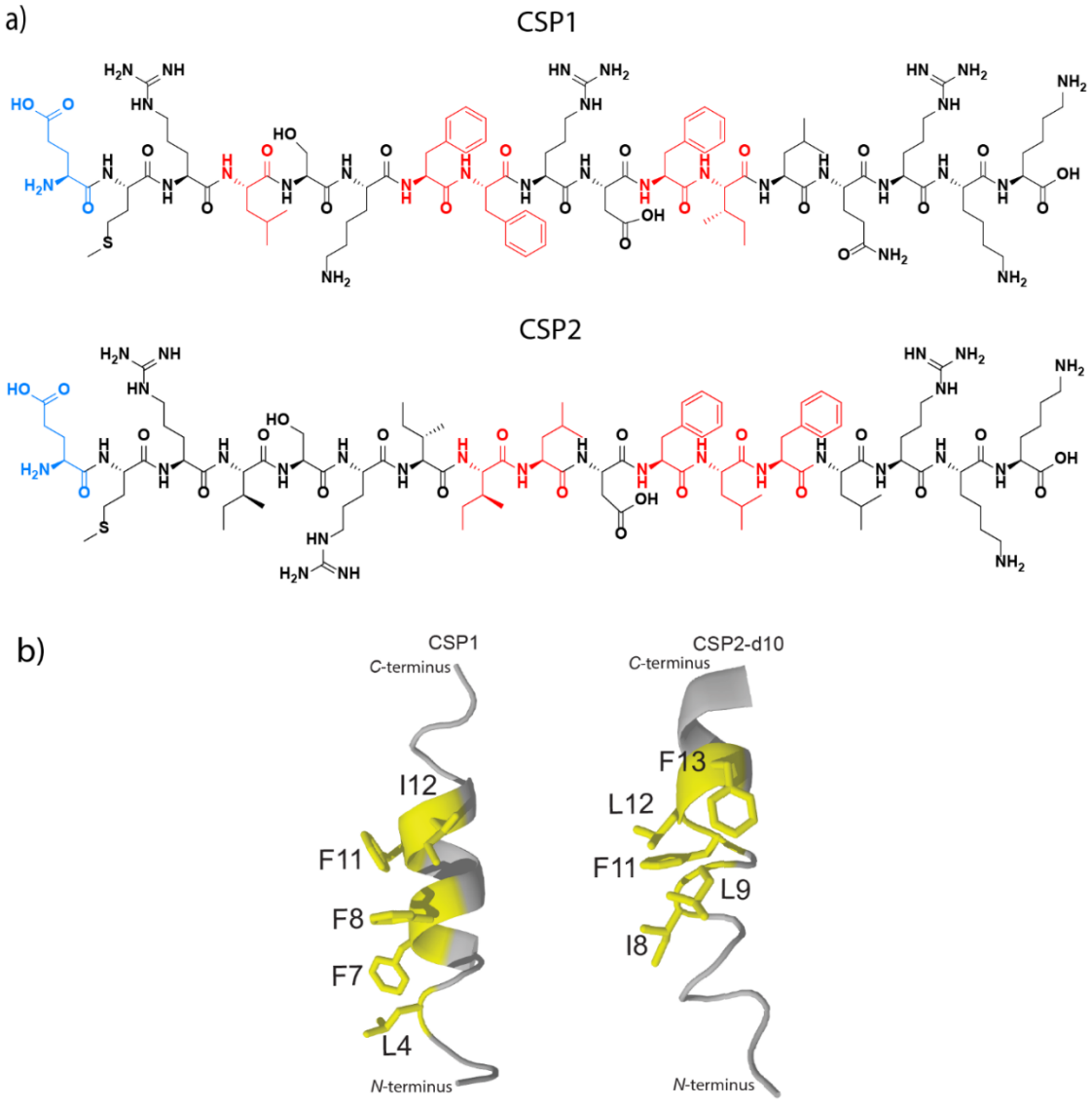


Figure 7. Key structural features required for CSP:ComD binding. a) The structures of CSP1 and CSP2 highlighting the hydrophobic residues that constitute the hydrophobic patches (in red) required for effective ComD1 and ComD2 binding, respectively. The Glu1 residue that is critical for receptor activation is marked in blue; b) 3D structures of CSP1

(PDB 6COW) and CSP2-d10 (PDB 6COT) exhibiting the two distinct hydrophobic patches. The two structures were obtained through structural NMR studies.[94]

The first position on the pherotype 1 AIP (CSP1, **Table 1**; with its unique primary structure, each CSP represents a separate pheromone type – pherotype) was found to be necessary for competence development, as Glu1 to Ala mutation (to afford CSP1-E1A) was sufficient to convert the signaling peptide into a competitive inhibitor capable of inhibiting competence induction and regulation of down-stream genes in *S. pneumoniae*.[42] The *N*-terminal first residue (Glu1) was further investigated through the implication of amino acids (AAs) with different side-chain length, polarity, and chirality.[95] Structural analysis of both native CSPs, along with several synthetic abiotic analogs, revealed that an α-helical character is important for receptor binding.[94] Results from these studies directed the design of cyclic CSP1 AIPs with enhanced binding characteristics. To this end, Yang et al. first applied ring position scan of the CSP1 scaffold, followed by ring size conformational scan to afford two potent pan-group activators with low nanomolar potency, (CSP1-cyc(Dab6E10) and CSP1-cyc(Dap6E10)).[57] Then, by incorporating the E1A modification, the lead pan-group inhibitor, CSP1-E1A-cyc(Dap6E10) (**Figure 8**), was developed, which can inhibit the ComD1 receptor with an IC_{50} value of 75.8 nM and ComD2 receptor with an IC_{50} value of 182 nM. Structural analysis revealed that this peptide also contains the previously identified important hydrophobic “patches” that are crucial for both ComD1 and ComD2 receptor binding.

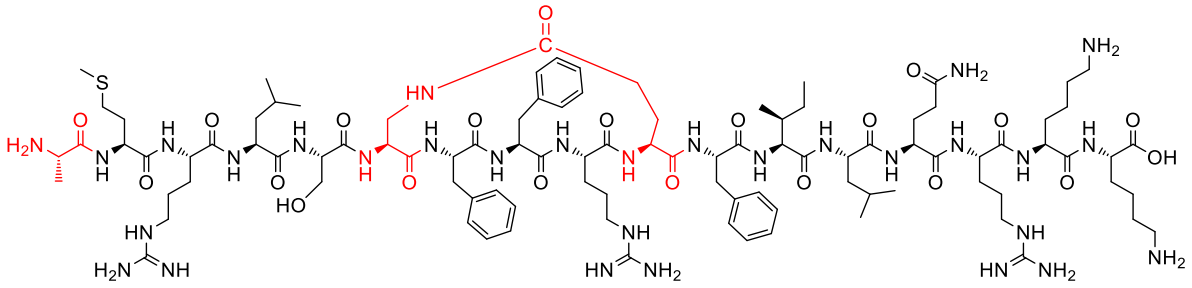


Figure 8. Pan-group inhibitor of the *S. pneumoniae* competence regulon QS circuitry.

The most potent inhibitor, CSP1-E1A-cyc(Dap6E10), was found to effectively inhibit both group 1 and group 2 pneumococcal competence regulon. The modifications at specific amino acids in CSP1 are highlighted as red color. Here, E1A represents that Glu at position 1 was replaced by Ala, Dap represents 2,3-diaminopropionic acid, and cyclization was introduced by connecting the 6th and 10th amino acid side chains.

The use of dominant-negative competence-stimulating peptide (dnCSP) analogs is considered to be an attractive therapeutic strategy, as apart from their high receptor interference capability, these agents also exhibited reduced antibiotic resistance and capsule biosynthesis genes acquisition, decreased allolytic factors LytA and CbpD expression and PLY release, and mouse mortality attenuation in *in vivo* studies.[42,57,58] Furthermore, the pharmacological properties and safety profiles of CSP1-E1A-cyc(Dap6E10) were recently evaluated, demonstrating the superior safety and pharmacokinetics profiles of this lead analog.[56] Biostability of this dnCSP, along with the native CSP1 and CSP2 signals, was examined using an IVIS Spectrum *in vivo* imaging system. To this end, the Cy7.5 fluorophore was attached to the *N*-terminus of the peptide to construct Cyanine7.5 (Cy7.5)-labeled dnCSP and both native CSP analogs. Comparison between the biostability of the dnCSP and native signals revealed the superior biostability property of the cyclized dnCSP, demonstrating the importance of macrocyclization of the

peptide. *In vitro* cytotoxicity and *in vivo* toxicity assays further supported the strong safety and pharmacological profiles of this promising peptide-based drug candidate.

As previous studies evidenced that nonproteogenic amino acid substitutions and cyclization resulted in improved modulators of the pneumococcal competence regulon,[57] Lella et al. intended to design and develop pharmacologically enhanced cyclic peptidomimetic scaffolds by utilizing the side chain-to-side chain urea-bridge cyclization chemistry.[55] This strategy led to the development of the first pneumococcus dual-action CSP modulator that can block group 1 while activating group 2 pneumococcus competence regulon. This is an interesting finding showing that analogs bearing the E1A modification, a reported key modification in the conversion of CSPs into competitive ComD inhibitors, can activate noncognate pneumococcal ComD receptors. The lead dual-action urea-bridged cyclic peptide, CSP1-E1A-cyc(Dab6Dab10) (**Figure 9**), can attenuate group 1 pneumococcal infections without affecting the bacterial burden, as was shown by *in vivo* studies using a mouse model of infection. These findings further highlighted the therapeutic potential of utilizing peptide-based competitive inhibitors to block the competence regulon without exposing the bacteria to selective pressure for resistance development.

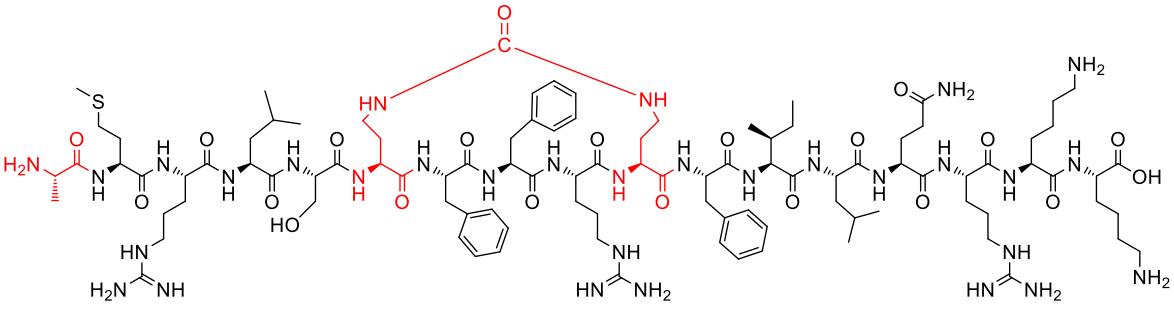


Figure 9. The dual-action urea-bridged cyclic CSP. The dual-action analog, CSP1-E1A-cyc(Dab6Dab10), was found to inhibit group 1 while activating group 2 pneumococcus competence regulon. The modifications at specific amino acids in CSP1 are highlighted as red color. Here, E1A represents that Glu at position 1 was replaced by Ala, Dab represents

2,4-diaminobutyric acid, and the urea bridge cyclization was introduced by connecting the 6^{th} and 10^{th} amino acid side chains.

RRNPP QS system as therapeutic target

The RRNPP protein family has been found in several members of the *Bacillus* genus and is a critical factor in regulating the expression of several genes responsible for biofilm formation, sporulation, and pathogenic responses.[14] The transcription factor PlcR and its associated AIP regulate the secretion of several virulence factors, including hemolysin, in *B. cereus*. Similar to the four Agr groups in *S. aureus*, five PapR signal peptides (**Table 1**) have been reported so far, and cross-reactivity has been shown in four of these pherotypes.[96] Binding of the signaling heptapeptide $PapR_7$ (ADLPFEF) triggers PlcR activity.[97,98] Alanine and epimer scans of $PapR_7$ provided valuable SAR insights, including the importance of stereochemistry of the side chains in positions 3-7 for PlcR activation. This analysis eventually led to the development of one activator and five inhibitors of the PlcR-PapR QS circuitry in *B. cereus*.

B. thuringiensis produces an insecticidal toxin and deletion of PapR resulted in a reduction of toxin production and the associated killing of susceptible insects.[98] The NprR transcriptional regulator activity was analyzed in *B. thuringiensis* and findings revealed that it is important for biofilm formation and sporulation in insect cadavers.[99] Analysis of NprR and its associated AIPs, termed NprRB (six known pherotypes), in *B. thuringiensis* revealed that the native peptide SKPDI (**Table 1**), along with another pherotype, SKPDT, and its K2A substituted synthetic analog, SAPDT, can induce the expression of the insecticidal toxin, *cry1Aa* in this species.[100] Among these three peptides, only SKPDT was able to induce a response above background levels at 100 nM. Further analysis with the

heptapeptide, SKPDIVG, and octapeptide, SSKPDIVG, revealed that both could upregulate bacterial sporulation (efficiency was increased by 2.1- and 1.6-fold, respectively). Three heptapeptide analogs were developed based on SKPDIVG: YSSKPDI, SSKPDIV, and SKPDIVG, and all exhibited mild inhibitory effects. The octapeptide, SSKPDIVG (**Figure 10**) was found to be the best inhibitor (35% drop in signal relative to the untreated control) whereas the nonapeptide YSSKPDIVG had no effect.[100] These virulence-regulating QS systems in *Bacillus* species could be targeted as an anti-virulence strategy.

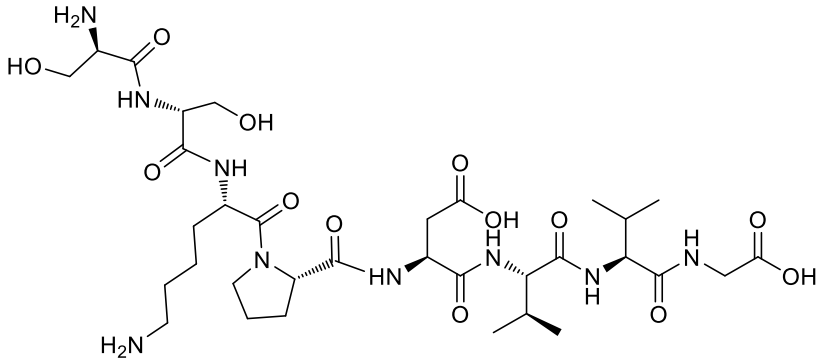


Figure 10. Inhibitor of the *B. thuringiensis* NprR QS circuitry. The octapeptide SSKPDIVG was found to be the most potent inhibitor of *B. thuringiensis cry1Aa* expression.

Highly homologous Rgg/SHP pairs, which act as an activator and a repressor, respectively, have been reported in several species of streptococci and it has been established that cross-communication could be taking place between these different streptococcal Rgg/SHP systems.[5,18,26-28] A bidirectional signaling associated with SHP pheromones was documented between human pathogens, including Group A *streptococcus* (GAS, *S. pyogenes*), group B *streptococcus* (GBS, *S. agalactiae*), and *S. dysgalactiae* subsp. *equisimilis*.[27] GAS can cause a verity of illnesses and life-threatening infections in humans.[101] In GAS, two competing transcriptional regulator pairs, Rgg2/SHP2 and

Rgg3/SHP3, are required to regulate the expression of a wide range of genes, including genes involved in biofilm formation.[23] GAS is often found to colonize with GBS at the same sites within the human host, and in GBS, virulence genes are regulated by an ortholog of the Rgg2 regulator, termed RovS, which is associated with a peptide pheromone signal nearly identical to SHP2.[27] Spent culture supernatant and co-culture experiments demonstrated that Rgg2/3-regulated gene expression, specifically genes involved in biofilm formation in GAS could be modulated by the production and secretion of GBS-produced SHPs, and on the other hand, GAS-produced SHP2/3 signals (**Table 1**) can stimulate RovS-mediated gene regulation in GBS. The results also revealed that both GAS and GBS Rgg/SHP QS circuits can be induced by using an orthologous Rgg2/SHP2 pair of *S. dysgalactiae* subsp. *equisimilis* and *S. porcinus*. Rgg/SHP systems of *S. mutans*, and *S. thermophilus* were also shown to be triggered by using slightly different and non-cognate synthetic SHPs from similar Rgg classes of streptococci.[28] These Rgg/SHP mediated cross-talks among streptococci provide a means to influence the regulation of a variety of pathogenic behaviors of Rgg/SHP-carrying pathogens by deceiving, diverting, or dissuading the Rgg/SHP QS systems of competing bacteria.

Crosstalk between pathogenic and non-pathogenic streptococci has been evidenced by a recent study conducted by Junges et.al.[26] *In silico* analyses revealed the presence of a Rgg/SHP cell-to-cell communication system in a commensal species from the mitis group, *S. mitis*. This *S. mitis*-Rgg regulator exhibited greater similarity to a *S. pyogenes*-Rgg repressor, however, the *S. mitis*-Rgg regulator functions as an activator, as confirmed by genetic mutation analysis and autoinducing assays. The *S. mitis*-Rgg regulator exhibited 74% identity and 87% similarity to the Rgg transcriptional regulator of the closely related

human pathogen, *S. pneumoniae*. The predicted mature SHP sequence in *S. mitis* is DIIIVGG (**Table 1**), having a unique feature of containing one less isoleucine residue than the other streptococcal SHPs (DIIIIVGG or DILIIVGG). The *S. mitis*-Rgg regulator can respond to other streptococcal SHPs and the similar noncognate peptides (DIIIIVGG and DILIIVGG) exhibited the highest activity potential of *S. mitis shp* induction. Not only the *S. pneumoniae* SHP (DIIIIVGG) can activate the QS system in *S. mitis*, but *S. mitis* SHP can also trigger the activation of the Rgg/SHP system in *S. pneumoniae*, which regulates pneumococcal surface polysaccharide production. Cross-communication of Rgg/SHP systems can provide important insights regarding the host-microbiome relationships, and interference with these signaling pathways may have therapeutic potential by controlling bacterial pathological behavior.

Conclusions and Outlook

With the increasingly recognized importance of tackling antimicrobial resistance, the interference and alteration of QS systems have the potential to promote the development of novel anti-virulence therapeutic strategies for treating harmful diseases. The attenuation of these non-vital communication systems with the use of peptide-based modulators in Gram-positive bacteria may aid in the fight to treat bacterial infections by preventing the spread of antimicrobial resistance while inhibiting bacterial pathogenesis. Therapeutically relevant peptide-based QS modulators that can be utilized as promising alternatives to current antibiotic therapy have been developed for the treatment of several Gram-positive bacterial pathogens, including *S. pneumoniae* and *S. aureus*. Such QS modulators can both act as drug leads and as probes to delineate the molecular mechanisms that drive QS activation as well as interspecies interference.

Although several promising preliminary *in vivo* studies have been reported in recent years showcasing the potential of QS modulation as a therapeutic approach; before these lead compounds could be brought to clinical trials extensive systematic *in vivo* studies as well as comprehensive ADME pharmacokinetic studies must be completed. To the best of our knowledge, such studies have not been completed yet. Furthermore, it is still not clear how QS modulators could best be utilized, as a standalone therapeutic or as an adjuvant as part of a combination therapy with traditional antibiotics. Both strategies have their advantages, yet failure in the pursue of either one may discourage future attempts to pursue the other approach. Therefore, the specific system and therapeutic strategy selected for advancement may determine the fate of QS modulation as a therapeutic approach in the near future.

Acknowledgments

The National Institutes of Health (R35GM128651 and R01HL142626) and the National Science Foundation (CHE-1808370) are acknowledged for the generous support of research in our laboratory.

Conflict of interest

The authors declare no competing interests.

Data Availability Statement

Data sharing not applicable – no new data generated.

References

1 Fuqua, W. C., Winans, S. C. & Greenberg, E. P. Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *J Bacteriol* **176**, 269-275 (1994). https://doi.org:10.1128/jb.176.2.269-275.1994

2 Waters, C. M. & Bassler, B. L. Quorum sensing: cell-to-cell communication in bacteria. *Annu Rev Cell Dev Biol* **21**, 319-346 (2005). https://doi.org:10.1146/annurev.cellbio.21.012704.131001

3 Abisado, R. G., Benomar, S., Klaus, J. R., Dandekar, A. A. & Chandler, J. R. Bacterial Quorum Sensing and Microbial Community Interactions. *mBio* **9** (2018). https://doi.org:10.1128/mBio.02331-17

4 Ryan, R. P. & Dow, J. M. Diffusible signals and interspecies communication in bacteria. *Microbiology (Reading)* **154**, 1845-1858 (2008). https://doi.org:10.1099/mic.0.2008/017871-0

5 Cook, L. C. & Federle, M. J. Peptide pheromone signaling in *Streptococcus* and *Enterococcus*. *FEMS Microbiol Rev* **38**, 473-492 (2014).

6 Lowery, C. A., Dickerson, T. J. & Janda, K. D. Interspecies and interkingdom communication mediated by bacterial quorum sensing. *Chem Soc Rev* **37**, 1337-1346 (2008). https://doi.org:10.1039/b702781h

7 Hawver, L. A., Jung, S. A. & Ng, W. L. Specificity and complexity in bacterial quorum-sensing systems. *FEMS Microbiol Rev* **40**, 738-752 (2016). https://doi.org:10.1093/femsre/fuw014

8 Rutherford, S. T. & Bassler, B. L. Bacterial quorum sensing: its role in virulence and possibilities for its control. *Cold Spring Harb Perspect Med* **2** (2012). https://doi.org:10.1101/cshperspect.a012427

9 Gray, B., Hall, P. & Gresham, H. Targeting agr-and agr-like quorum sensing systems for development of common therapeutics to treat multiple gram-positive bacterial infections. *Sensors* **13**, 5130-5166 (2013).

10 Mull, R. W., Harrington, A., Sanchez, L. A. & Tal-Gan, Y. Cyclic Peptides that Govern Signal Transduction Pathways: From Prokaryotes to Multi-Cellular Organisms. *Curr Top Med Chem* **18**, 625-644 (2018). https://doi.org:10.2174/1568026618666180518090705

11 Rocha-Estrada, J., Aceves-Diez, A. E., Guarneros, G. & de la Torre, M. The RNPP family of quorum-sensing proteins in Gram-positive bacteria. *Appl Microbiol Biotechnol* **87**, 913-923 (2010).

12 Neiditch, M. B., Capodagli, G. C., Prehna, G. & Federle, M. J. Genetic and Structural Analyses of RRNPP Intercellular Peptide Signaling of Gram-Positive Bacteria. *Annu Rev Genet* **51**, 311-333 (2017). https://doi.org:10.1146/annurev-genet-120116-023507

13 Haque, S. *et al.* Developments in strategies for Quorum Sensing virulence factor inhibition to combat bacterial drug resistance. *Microb Pathog* **121**, 293-302 (2018).

14 Prazdnova, E. V. *et al.* Quorum-Sensing Inhibition by Gram-Positive Bacteria. *Microorganisms* **10**, 350 (2022).

15 Gohar, M. *et al.* The PlcR virulence regulon of *Bacillus* cereus. *PLoS One* **3**, e2793 (2008).

16 Zouhir, S. *et al.* Peptide-binding dependent conformational changes regulate the transcriptional activity of the quorum-sensor *NprR*. *Nucleic Acids Res* **41**, 7920-7933 (2013).

17 Perchat, S. *et al.* A cell–cell communication system regulates protease production during sporulation in bacteria of the *Bacillus cereus* group. *Mol Microbiol* **82**, 619-633 (2011).

18 Perez-Pascual, D., Monnet, V. & Gardan, R. Bacterial cell–cell communication in the host via RRNPP peptide-binding regulators. *Front Microbiol* **7**, 706 (2016).

19 Sulavik, M., Tardif, G. & Clewell, D. Identification of a gene, *rgg*, which regulates expression of glucosyltransferase and influences the Spp phenotype of *Streptococcus gordonii Challis*. *J Bacteriol* **174**, 3577-3586 (1992).

20 Kreikemeyer, B., McIver, K. S. & Podbielski, A. Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen–host interactions. *Trends Microbiol* **11**, 224-232 (2003).

21 Qi, F., Chen, P. & Caufield, P. W. Functional analyses of the promoters in the lantibiotic mutacin II biosynthetic locus in *Streptococcus mutans*. *Appl Environ Microbiol* **65**, 652-658 (1999).

22 Chaussee, M. S., Ajdic, D. & Ferretti, J. J. The *rgg* gene of *Streptococcus pyogenes NZ131* positively influences extracellular SPE B production. *Infect Immun* **67**, 1715-1722 (1999).

23 Chang, J. C., LaSarre, B., Jimenez, J. C., Aggarwal, C. & Federle, M. J. Two group A streptococcal peptide pheromones act through opposing Rgg regulators to control biofilm development. *PLoS Pathog* **7**, e1002190 (2011).

24 Lyon, W. R., Gibson, C. M. & Caparon, M. G. A role for trigger factor and an rgg-like regulator in the transcription, secretion and processing of the cysteine proteinase of *Streptococcus pyogenes*. *EMBO J* **17**, 6263-6275 (1998).

25 Junges, R. *et al.* A quorum-sensing system that regulates *Streptococcus pneumoniae* biofilm formation and surface polysaccharide production. *Msphere* **2**, e00324-00317 (2017).

26 Junges, R. *et al.* Characterization of a signaling system in *Streptococcus mitis* that mediates interspecies communication with *Streptococcus pneumoniae*. *Appl Environ Microbiol* **85**, e02297-02218 (2019).

27 Cook, L. C., LaSarre, B. & Federle, M. J. Interspecies communication among commensal and pathogenic streptococci. *MBio* **4**, e00382-00313 (2013).

28 Fleuchot, B. *et al.* Rgg-associated SHP signaling peptides mediate cross-talk in *Streptococci*. *PLoS One* **8**, e66042 (2013).

29 Queck, S. Y. *et al.* RNAIII-independent target gene control by the agr quorum-sensing system: insight into the evolution of virulence regulation in *Staphylococcus aureus*. *Mol Cell* **32**, 150-158 (2008).

30 Thoendel, M., Kavanaugh, J. S., Flack, C. E. & Horswill, A. R. Peptide signaling in the *staphylococci*. *Chem Rev* **111**, 117-151 (2011).

31 Cheung, A. L., Bayer, A. S., Zhang, G., Gresham, H. & Xiong, Y.-Q. Regulation of virulence determinants *in vitro* and *in vivo* in *Staphylococcus aureus*. *FEMS Microbiol Immunol* **40**, 1-9 (2004).

32 Felden, B., Vandenesch, F., Bouloc, P. & Romby, P. The *Staphylococcus aureus* RNome and its commitment to virulence. *PLoS Pathog* **7**, e1002006 (2011).

33 Reyes, D. *et al.* Coordinated regulation by AgrA, SarA, and SarR to control agr expression in *Staphylococcus aureus*. *J Bacteriol* **193**, 6020-6031 (2011).

34 Periasamy, S. *et al.* How *Staphylococcus aureus* biofilms develop their characteristic structure. *Proc Natl Acad Sci U S A* **109**, 1281-1286 (2012).

35 Mehr, S. & Wood, N. *Streptococcus pneumoniae*--a review of carriage, infection, serotype replacement and vaccination. *Paediatr Respir Rev* **13**, 258-264 (2012). https://doi.org:10.1016/j.prrv.2011.12.001

36 Kadioglu, A., Weiser, J. N., Paton, J. C. & Andrew, P. W. The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol* **6**, 288-301 (2008). https://doi.org:10.1038/nrmicro1871

37 Griffith, F. The Significance of *Pneumococcal* Types. *J Hyg (Lond)* **27**, 113-159 (1928). https://doi.org:10.1017/s0022172400031879

38 Croucher, N. J. *et al.* Rapid *pneumococcal* evolution in response to clinical interventions. *Science* **331**, 430-434 (2011). https://doi.org:10.1126/science.1198545

39 Guiral, S., Mitchell, T. J., Martin, B. & Claverys, J. P. Competence-programmed predation of noncompetent cells in the human pathogen *Streptococcus pneumoniae*: genetic requirements. *Proc Natl Acad Sci U S A* **102**, 8710-8715 (2005). https://doi.org:10.1073/pnas.0500879102

40 Lella, M. & Tal-Gan, Y. Strategies to attenuate the competence regulon in *Streptococcus pneumoniae*. *J Pept Sci* **113**, e24222 (2021).

41 Chao, Y., Marks, L. R., Pettigrew, M. M. & Hakansson, A. P. *Streptococcus pneumoniae* biofilm formation and dispersion during colonization and disease. *Front Cell Infect Microbiol* **4**, 194 (2014). https://doi.org:10.3389/fcimb.2014.00194

42 Zhu, L. & Lau, G. W. Inhibition of competence development, horizontal gene transfer and virulence in *Streptococcus pneumoniae* by a modified competence stimulating peptide. *PLoS Pathog* **7**, e1002241 (2011).

43 Hui, F. M., Zhou, L. & Morrison, D. A. Competence for genetic transformation in *Streptococcus pneumoniae*: organization of a regulatory locus with homology to two lactococcin A secretion genes. *Gene* **153**, 25-31 (1995). https://doi.org:10.1016/0378-1119(94)00841-f

44 Peterson, S. N. *et al.* Identification of competence pheromone responsive genes in *Streptococcus pneumoniae* by use of DNA microarrays. *Mol Microbiol* **51**, 1051-1070 (2004). https://doi.org:10.1046/j.1365-2958.2003.03907.x

45 Milly, T. A. & Tal-Gan, Y. Biological evaluation of native *streptococcal* competence stimulating peptides reveals potential crosstalk between *Streptococcus mitis* and *Streptococcus pneumoniae* and a new scaffold for the development of *S. pneumoniae* quorum sensing modulators. *RSC Chem Biol* **1**, 60-67 (2020).

46 Kowalko, J. E. & Sebert, M. E. The *Streptococcus pneumoniae* competence regulatory system influences respiratory tract colonization. *Infect Immun* **76**, 3131-3140 (2008). https://doi.org:10.1128/IAI.01696-07

47 Yang, Y. *et al.* Structure–activity relationships of the competence stimulating peptides (CSPs) in *Streptococcus pneumoniae* reveal motifs critical for intra-group and cross-group ComD receptor activation. *ACS Chem Biol* **12**, 1141-1151 (2017).

48 McBrayer, D. N. & Tal-Gan, Y. Deciphering bacterial signalling. *Nat Chem* **11**, 398-399 (2019).

49 Wu, S., Liu, J., Liu, C., Yang, A. & Qiao, J. Quorum sensing for population-level control of bacteria and potential therapeutic applications. *Cell Mol Life Sci* **77**, 1319-1343 (2020). https://doi.org:10.1007/s00018-019-03326-8

50 Kalia, V. C. Quorum sensing inhibitors: an overview. *Biotechnol Adv* **31**, 224-245 (2013). https://doi.org:10.1016/j.biotechadv.2012.10.004

51 Henninot, A., Collins, J. C. & Nuss, J. M. The Current State of Peptide Drug Discovery: Back to the Future? *J Med Chem* **61**, 1382-1414 (2018). https://doi.org:10.1021/acs.jmedchem.7b00318

52 Angell, Y., Holford, M. & Moos, W. H. Building on Success: A Bright Future for Peptide Therapeutics. *Protein Pept Lett* **25**, 1044-1050 (2018). https://doi.org:10.2174/0929866525666181114155542

53 Ding, Y. *et al.* Impact of non-proteinogenic amino acids in the discovery and development of peptide therapeutics. *Amino Acids* **52**, 1207-1226 (2020). https://doi.org:10.1007/s00726-020-02890-9

54 Vinogradov, A. A., Yin, Y. & Suga, H. Macrocyclic Peptides as Drug Candidates: Recent Progress and Remaining Challenges. *J Am Chem Soc* **141**, 4167-4181 (2019). https://doi.org:10.1021/jacs.8b13178

55 Lella, M., Oh, M. W., Kuo, S. H., Lau, G. W. & Tal-Gan, Y. Attenuating the *Streptococcus pneumoniae* Competence Regulon Using Urea-Bridged Cyclic Dominant-Negative Competence-Stimulating Peptide Analogs. *J Med Chem* (2022).

56 Oh, M. W., Lella, M., Kuo, S. H., Tal-Gan, Y. & Lau, G. W. Pharmacological Evaluation of Synthetic Dominant-Negative Peptides Derived from the Competence-Stimulating Peptide of *Streptococcus pneumoniae*. *ACS Pharmacol Transl Sci* (2022).

57 Yang, Y. *et al.* Designing cyclic competence-stimulating peptide (CSP) analogs with pan-group quorum-sensing inhibition activity in *Streptococcus pneumoniae*. *Proc Natl Acad Sci U S A* **117**, 1689-1699 (2020).

58 Koirala, B., Lin, J., Lau, G. W. & Tal-Gan, Y. Development of a Dominant Negative Competence-Stimulating Peptide (dnCSP) that Attenuates *Streptococcus pneumoniae* Infectivity in a Mouse Model of Acute Pneumonia. *ChemBioChem* **19**, 2380-2386 (2018).

59 Park, J. *et al.* Infection control by antibody disruption of bacterial quorum sensing signaling. *Chem Biol* **14**, 1119-1127 (2007). https://doi.org:10.1016/j.chembiol.2007.08.013

60 West, K. H. *et al.* Sustained Release of a Synthetic Autoinducing Peptide Mimetic Blocks Bacterial Communication and Virulence *In Vivo*. *Angew Chem Int Ed*, e202201798 (2022).

61 Haseeb, A., Ajit Singh, V., Teh, C. S. J. & Loke, M. F. Addition of ceftaroline fosamil or vancomycin to PMMA: an in vitro comparison of biomechanical properties and anti-MRSA efficacy. *Orthop Surg* **27**, 2309499019850324 (2019).

62 Guo, Y., Song, G., Sun, M., Wang, J. & Wang, Y. Prevalence and therapies of antibiotic-resistance in *Staphylococcus aureus*. *Front Cell Infect* **10**, 107 (2020).

63 Khan, B. A., Yeh, A. J., Cheung, G. Y. & Otto, M. Investigational therapies targeting quorum-sensing for the treatment of *Staphylococcus aureus* infections. *Expert Opin Investigat Drugs* **24**, 689-704 (2015).

64 Parlet, C. P. *et al.* Apicidin attenuates MRSA virulence through quorum-sensing inhibition and enhanced host defense. *Cell Rep* **27**, 187-198. e186 (2019).

65 Rasko, D. A. & Sperandio, V. Anti-virulence strategies to combat bacteria-mediated disease. *Nat Rev Drug Discov* **9**, 117-128 (2010).

66 Gordon, C. P., Williams, P. & Chan, W. C. Attenuating *Staphylococcus aureus* virulence gene regulation: a medicinal chemistry perspective. *J Med Chem* **56**, 1389-1404 (2013).

67 Leonard, P. G., Bezar, I. F., Sidote, D. J. & Stock, A. M. Identification of a hydrophobic cleft in the LytTR domain of AgrA as a locus for small molecule interactions that inhibit DNA binding. *Biochem* **51**, 10035-10043 (2012).

68 Rowe, S. E., Mahon, V., Smith, S. G. & O'Gara, J. P. A novel role for SarX in *Staphylococcus epidermidis* biofilm regulation. *Microbiol* **157**, 1042-1049 (2011).

69 Scott, R. J. *et al.* Side-chain-to-tail thiolactone peptide inhibitors of the *staphylococcal* quorum-sensing system. *Bioorg Med Chem Lett* **13**, 2449-2453 (2003).

70 Benjamin, D., Colombi, M., Moroni, C. & Hall, M. N. Rapamycin passes the torch: a new generation of mTOR inhibitors. *Nat Rev Drug Discov* **10**, 868-880 (2011).

71 Lyon, G. J., Mayville, P., Muir, T. W. & Novick, R. P. Rational design of a global inhibitor of the virulence response in *Staphylococcus aureus*, based in part on localization of the site of inhibition to the receptor-histidine kinase, AgrC. *Proc Natl Acad Sci U S A* **97**, 13330-13335 (2000).

72 Beenken, K. E. *et al.* Epistatic relationships between sarA and agr in *Staphylococcus aureus* biofilm formation. *PLoS One* **5**, e10790 (2010).

73 Tal-Gan, Y., Ivancic, M., Cornilescu, G., Cornilescu, C. C. & Blackwell, H. E. Structural characterization of native autoinducing peptides and abiotic analogues reveals key features essential for activation and inhibition of an AgrC quorum sensing receptor in *Staphylococcus aureus*. *J Am Chem Soc* **135**, 18436-18444 (2013).

74 Mayville, P. *et al.* Structure-activity analysis of synthetic autoinducing thiolactone peptides from *Staphylococcus aureus* responsible for virulence. *Proc Natl Acad Sci U S A* **96**, 1218-1223 (1999).

75 Tal-Gan, Y., Ivancic, M., Cornilescu, G., Yang, T. & Blackwell, H. E. Highly stable, amide-bridged autoinducing peptide analogues that strongly inhibit the agrc quorum sensing receptor in *Staphylococcus aureus*. *Angew Chem Int Ed* **128**, 9059-9063 (2016).

76 Tal-Gan, Y., Ivancic, M., Cornilescu, G. & Blackwell, H. E. Characterization of structural elements in native autoinducing peptides and non-native analogues that permit the differential modulation of AgrC-type quorum sensing receptors in *Staphylococcus aureus*. *Org Biomol Chem* **14**, 113-121 (2016).

77 Vasquez, J. K., Tal-Gan, Y., Cornilescu, G., Tyler, K. A. & Blackwell, H. E. Simplified AIP-II Peptidomimetics Are Potent Inhibitors of *Staphylococcus aureus* AgrC Quorum Sensing Receptors. *ChemBioChem* **18**, 413-423 (2017).

78 Ji, G., Beavis, R. & Novick, R. P. Bacterial interference caused by autoinducing peptide variants. *Science* **276**, 2027-2030 (1997).

79 Wright, J. S., Jin, R. & Novick, R. P. Transient interference with *staphylococcal* quorum sensing blocks abscess formation. *Proc Natl Acad Sci U S A* **102**, 1691-1696 (2005).

80 Lyon, G. J., Wright, J. S., Muir, T. W. & Novick, R. P. Key determinants of receptor activation in the agr autoinducing peptides of *Staphylococcus aureus*. *Biochemistry* **41**, 10095-10104 (2002).

81 MDowell, P. *et al.* Structure, activity and evolution of the group I thiolactone peptide quorum-sensing system of *Staphylococcus aureus*. *Mol Microbiol* **41**, 503-512 (2001).

82 George, E. A., Novick, R. P. & Muir, T. W. Cyclic peptide inhibitors of *staphylococcal* virulence prepared by Fmoc-based thiolactone peptide synthesis. *J Am Chem Soc* **130**, 4914-4924 (2008).

83 Fowler, S. A., Stacy, D. M. & Blackwell, H. E. Design and synthesis of macrocyclic peptomers as mimics of a quorum sensing signal from *Staphylococcus aureus*. *Org Lett* **10**, 2329-2332 (2008).

84 Chan, W. C., Coyle, B. J. & Williams, P. Virulence regulation and quorum sensing in *staphylococcal* infections: competitive AgrC antagonists as quorum sensing inhibitors. *J Am Chem Soc* **47**, 4633-4641 (2004).

85 Tal-Gan, Y., Stacy, D. M., Foegen, M. K., Koenig, D. W. & Blackwell, H. E. Highly potent inhibitors of quorum sensing in *Staphylococcus aureus* revealed through a systematic synthetic study of the group-III autoinducing peptide. *J Am Chem Soc* **135**, 7869-7882 (2013).

86 Tal-Gan, Y., Stacy, D. M. & Blackwell, H. E. N-Methyl and peptoid scans of an autoinducing peptide reveal new structural features required for inhibition and activation of AgrC quorum sensing receptors in *Staphylococcus aureus*. *Chem Comm* **50**, 3000-3003 (2014).

87 Yano, M., Gohil, S., Coleman, J. R., Manix, C. & Pirofski, L.-a. Antibodies to *Streptococcus pneumoniae* capsular polysaccharide enhance pneumococcal quorum sensing. *MBio* **2**, e00176-00111 (2011).

88 Aggarwal, S. D. *et al.* Function of BriC peptide in the *pneumococcal* competence and virulence portfolio. *PLoS Pathog* **14**, e1007328 (2018).

89 Kohler, S., Voß, F., Gómez Mejia, A., Brown, J. S. & Hammerschmidt, S. *Pneumococcal* lipoproteins involved in bacterial fitness, virulence, and immune evasion. *FEBS Lett* **590**, 3820-3839 (2016).

90 McBrayer, D. N., Cameron, C. D. & Tal-Gan, Y. Development and utilization of peptide-based quorum sensing modulators in Gram-positive bacteria. *Org Biomol Chem* **18**, 7273-7290 (2020). https://doi.org:10.1039/d0ob01421d

91 Koirala, B., Hillman, R. A., Tiwold, E. K., Bertucci, M. A. & Tal-Gan, Y. Defining the hydrophobic interactions that drive competence stimulating peptide (CSP)-ComD binding in *Streptococcus pneumoniae*. *Beilstein J Org Chem* **14**, 1769-1777 (2018).

92 Milly, T. A. *et al.* Harnessing Multiple, Nonproteogenic Substitutions to Optimize CSP: ComD Hydrophobic Interactions in Group 1 *Streptococcus pneumoniae*. *ChemBioChem* **22**, 1940-1947 (2021).

93 Milly, T. A. *et al.* Optimizing CSP1 analogs for modulating quorum sensing in *Streptococcus pneumoniae* with bulky, hydrophobic nonproteogenic amino acid substitutions. *RSC Chem Biol* **3**, 301-311 (2022).

94 Yang, Y., Cornilescu, G. & Tal-Gan, Y. Structural characterization of competence-stimulating peptide analogues reveals key features for ComD1 and ComD2 receptor binding in *Streptococcus pneumoniae*. *Biochemistry* **57**, 5359-5369 (2018).

95 Yang, Y. & Tal-Gan, Y. Exploring the competence stimulating peptide (CSP) *N*-terminal requirements for effective ComD receptor activation in group1 *Streptococcus pneumoniae*. *Bioorg Chem* **89**, 102987 (2019).

96 Bouillaut, L. *et al.* Molecular basis for group-specific activation of the virulence regulator PlcR by PapR heptapeptides. *Nucleic Acids Res* **36**, 3791-3801 (2008).

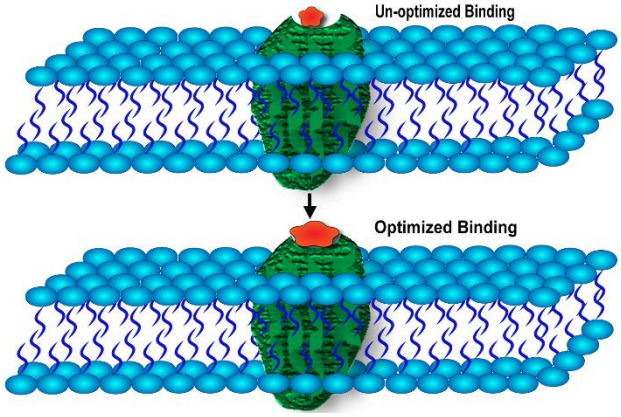
97 Yehuda, A. *et al.* Turning off Bacillus cereus quorum sensing system with peptidic analogs. *Chem Comm* **54**, 9777-9780 (2018).

98 Slamti, L. & Lereclus, D. A cell–cell signaling peptide activates the PlcR virulence regulon in bacteria of the *Bacillus cereus* group. *EMBO J* **21**, 4550-4559 (2002).

99 Dubois, T. *et al.* Necrotrophism is a quorum-sensing-regulated lifestyle in *Bacillus thuringiensis*. *PLoS Pathog* **8**, e1002629 (2012).

100 Rocha, J. *et al.* Evolution and some functions of the NprR–NprRB quorum-sensing system in the *Bacillus cereus* group. *Appl Microbiol Biotechnol* **94**, 1069-1078 (2012).

101 Cunningham, M. W. Pathogenesis of group A *streptococcal* infections and their sequelae. *Hot Topics in Infection and Immunity in Children IV*, 29-42 (2008).



Chapter 2: Harnessing Multiple, Nonproteogenic Substitutions to Optimize CSP:ComD Hydrophobic Interactions in Group 1 *Streptococcus pneumoniae*

[a]Reprinted with permission from Milly, T. A.; Engler, E. R.; Chichura, K. S.; Buttner, A. R.; Koirala, B.; Tal-Gan, Y.; Bertucci, M. A. Harnessing Multiple, Nonproteogenic Substitutions to Optimize CSP: ComD Hydrophobic Interactions in Group 1 *Streptococcus pneumoniae*. *ChemBioChem* **2021**, 22 (11), 1940-1947.

Harnessing Multiple, Non-proteogenic Substitutions to Optimize CSP:ComD Hydrophobic Interactions in Group1 *Streptococcus pneumoniae*

Tahmina A. Milly,[†] Emilee R. Engler,[‡] Kylie S. Chichura,[‡] Alec R. Buttner,[‡] Bimal Koirala,[†] Yftah Tal-Gan,[†] and Michael A. Bertucci[‡]

[†] Department of Chemistry, University of Nevada, Reno, 1664 North Virginia Street, Reno, Nevada, 89557, United States

[‡] Department of Chemistry, Moravian College, 1200 Main Street, Bethlehem, Pennsylvania, 18018, United States

Abstract

Streptococcus pneumoniae (pneumococcus) is a human pathobiont that causes drastic antibiotic-resistant infections and is responsible for millions of deaths universally. Pneumococcus pathogenicity relies on the competence stimulating peptide (CSP) - mediated quorum sensing (QS) pathway that controls competence development for genetic transformation and, consequently, the spread of antibiotic resistance and virulence genes. Modulation of QS in *S. pneumoniae* can therefore be utilized to enervate pneumococcal infectivity as well as minimize the susceptibility for resistance development. In this work, we sought to optimize the interaction of CSP1 with its cognate transmembrane histidine kinase receptor (ComD1) through substitution of proteogenic and non-proteogenic amino acids on the hydrophobic binding face of CSP1. The findings from this study not only provided additional structure-activity data that are significant in optimizing CSP1 potency, but also led to the development of potent QS modulators. These CSP-based QS modulators could be used as privileged scaffolds for the development of antimicrobial agents against pneumococcal infections.

Introduction

Streptococcus pneumoniae, or pneumococcus, is a Gram-positive, commensal bacterium that populates the nasopharyngeal cavity of many healthy humans but can behave as an opportunistic pathogen if it extends to other parts of the body (ears, sinuses, and lungs). This results in a variety of life-threatening diseases, including deadly pneumonia, meningitis and bacteremia.[102,103] Pathogenic pneumococcus was recently estimated to be responsible for over 22,000 deaths and 445,000 hospitalizations, with a $3.5 billion direct medical cost annually in the United States.[104] Moreover, recombinogenic pneumococcus is proficient at expeditiously developing resistance against diverse antimicrobial agents such as linezolid, vancomycin, and β-lactams, as well as avoiding several pneumococcal conjugate vaccines by switching its capsular type.[105-108] Intra- and inter-species genetic material exchange through natural genetic transformation actively contributes to the rapid emergence of resistant strains by promoting genetic flexibility in the pneumococcus community.[105,109-111] Competent pneumococci are able to directly acquire exogenous DNA from noncompetent pneumococcal cells as well as closely related species inhabiting the same microbial niche, such as *S. mitis*, providing a reservoir of both antimicrobial resistance and virulence genes.[112-114] This recombination-mediated genetic plasticity grants a significant advantage to *S. pneumoniae* by enabling rapid evolution of the genome and capsular diversity.

In *S. pneumoniae*, the rapid acquisition of genetic material from the surroundings is mediated by the competence regulon, a quorum-sensing (QS) circuit that is regulated by a peptide pheromone named the competence-stimulating peptide (CSP, **Figure 11**).[115] After being processed and exported outside the cell, the mature signaling molecule, CSP,

initiates the QS signaling cascade by binding to its cognate transmembrane histidine kinase receptor, ComD.[43, 116] This binding of the receptor leads to activation of a cytoplasmic cognate response regulator, ComE. Once activated, phosphorylated ComE can act as a transcriptional activator to initiate the transcription of the QS circuitry genes (*comABCDE*) together with the alternative sigma factor (*comX*) gene (**Figure 11**). ComX is a key regulator of the QS circuitry that controls the different QS-regulated phenotypes, such as biofilm formation, virulence factor production and competence for genetic transformation.[117-121] The involvement of the CSP-mediated QS system in pneumococcal pathogenesis has resulted in considerable efforts to interfere with this QS circuitry as a potential anti-infective therapeutic approach.[42,47,58,122-125] Furthermore, QS interference provides the dual advantage of attenuating the infectivity of the bacteria while minimizing the selection pressure for the emergence of drug resistant strains.[85,126-129]

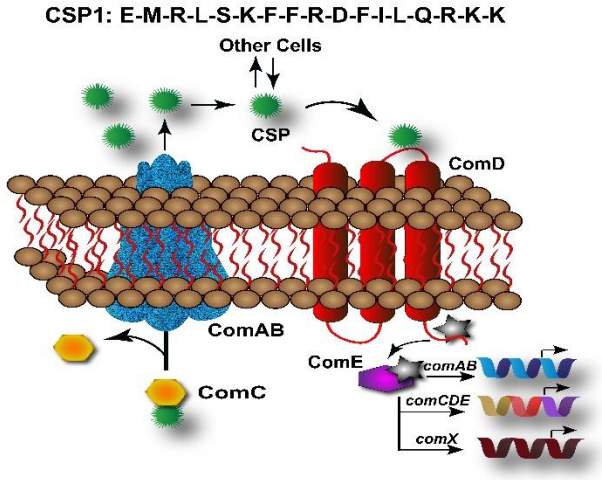


Figure 11: Pneumococcus CSP-mediated QS circuit. CSP is synthesized as a propeptide, ComC, and is processed and exported out of the cell as the mature peptide with the help of an ABC transporter, ComAB. At threshold concentration, mature CSP can effectively bind and activate a membrane-bound histidine-kinase receptor, ComD. Upon activation, ComD phosphorylates the transcription factor, ComE. Phosphorylated ComE then autoactivates the QS circuit and triggers the transcription of several genes involved in achieving genetic competence, virulence factor production and biofilm formation. The sequence of CSP1 is shown at the top.

The vast majority of pneumococcus strains use one of two major forms of the CSP pheromone, namely CSP1 and CSP2, with their respective receptors, ComD1 and ComD2, leading to the classification of two pherotypes or specificity groups, group1 and group2, among *S. pneumoniae*.[130,131] Although these two CSP signals share 50% homology, they are highly specific toward their compatible receptor.[130] Hydrophobic interactions have been identified as a critical driving force for the association of the CSP signals with their cognate receptors.[47,94,132] Previously, Yang et al. conducted an extensive structure-activity relationship (SAR) study of the native CSP1 signal and revealed that several key hydrophobic residues in the central region of the CSP1 sequence, specifically positions 4, 7, 8, 11, 12 and 13, are important for ComD1 receptor binding and activation.[47] Moreover, in our previous work, we have incorporated highly conservative point mutations in these key hydrophobic positions of CSP1, utilizing both proteogenic and non-proteogenic amino acids to define the hydrophobic binding pockets within the ComD1 receptor.[91] In the present study, we continued our efforts to optimize the interactions of CSP1 with ComD1 by rationally designing a multiple mutation library of CSP1 intended to achieve more potent CSP-based QS modulators. Special attention was given to positions 4, 12, and 13, as these positions, naturally containing the aliphatic, hydrophobic side-chains of leucine and isoleucine, were found to tolerate mutations better than positions 7, 8 and 11. Improved ComD1 binding was observed when these residues were substituted with amino acids differing in carbon chain count – namely leucine (L), isoleucine (I), norleucine (NL), and norvaline (NV).[91] Herein, to determine whether these substitutions would act cooperatively at the receptor surface, we combined these modifications to create a library of multiply mutated CSP1 analogs. Through this analysis, we were able to develop a potent ComD1

activator (CSP1-L4NV/I12NL/L13NV), containing a triplet non-proteogenic amino acids, which may provide a pharmacokinetic advantage. We then sought to combine the lead CSP1 analogs identified in this study with the E1A mutation known to act as an activator-to-inhibitor switch to afford potent CSP1-based QS inhibitors of *S. pneumoniae*. Indeed, several group1 pneumococcal ComD inhibitors with activities in the nanomolar range were identified. The results of this study provide an avenue for the future development of CSP-based drug leads with improved pharmacological properties.

Results & Discussion

As noted above, the single mutations at positions 4, 12, and 13 previously determined to have the lowest EC_{50} values were integrated into our new multiple-mutant CSP1 library (**Table 2**).[38] All of the CSP1 analogs (**Table 2** and **3**) were built using solid-phase peptide synthesis (SPPS) protocols,[133] followed by purification by semi-preparative RP-HPLC (for full details see the Experimental Section). Biological assays were conducted using a *S. pneumoniae* strain harboring a pComX-*lacZ* reporter plasmid that was incorporated into the bacterial genome, previously constructed by Lau and co-workers,[42] to assess how well CSP1 mutants bind and activate the ComD1 receptor. In this reporter strain, the *comX* promoter is fused to the *lacZ* gene. Thus, at high cell densities, which correspond to high CSP concentrations, the CSP:ComD complex will phosphorylate ComE, which will then bind pcomX and transcribe *lacZ* (in addition to upregulation of ComX). Therefore, by measuring β-galactosidase activity, ComD activation can be quantified. Initial activation assays were conducted at high analog concentration, followed by more rigorous dose-dependent trials to calculate EC_{50} values for each derivative (see the Appendix 1 for dose response curves).

Screening of Multiply-Substituted CSP1 Derivatives for ComD1 Activation

According to the previous SAR results of singly-substituted CSP1 analogs, norvaline (NV) may be the linear limit of how far the binding pocket will allow the 4th position's side chain to protrude.[91] The norleucine (NL) substitution, which elongates the carbon chain in this position, resulted in a slight loss in potency. Considering the data of single substitutions at L4, the length of the carbon chain appears to be the most significant factor in optimizing CSP activity. Beta-branching at this position is unnecessary as seen by the comparable EC_{50} values between the native CSP1 and the L4I substitution (10.3 nM and 10.2 nM, respectively).[91] Since introducing NV in position 4 yielded the highest potency with a tight margin of error as a single substitution, it was integrated into our library of multiply-substituted CSP1 mutants. Our data (**Table 2**) suggests that NV is the optimal amino acid in this position, especially when combined with the L13NV substitution (L4NV/L13NV, EC_{50} = 4.3 nM). As a straight-chain, non-proteogenic amino acid, NV appears to have improved binding. Aside from providing an optimized carbon chain length, we suspect this substitution increases side-chain flexibility at the binding interface without disrupting the secondary structure of the peptide. However, the efficacy of modifications at L4 is highly dependent on changes at I12 and L13. For example, the EC_{50} of L4I/I12L was 78 nM, whereas the EC_{50} of L4I/I12NL was 7.3 nM. Thus, the impact of a single amino acid substitution at position 4 on potency cannot be determined without considering other modifications made to the CSP1 primary structure.

Table 2: EC_{50} values of multiply-substituted CSP1 derivatives against the ComD1 receptor[a].

Peptide name	Peptide Sequence	EC_{50} [nM][b]	95% CI[c]
CSP1[d]	EMRLSKFFRDFILQRKK	10	6.3-17
CSP1-L4I/I12L	EMR**I**SKFFRDF**L**LQRKK	78	67-90
CSP1-L4I/I12NL	EMR**I**SKFFRDF(**NL**)LQRKK	7.3	5.4-9.8
CSP1-L4I/L13I	EMR**I**SKFFRDFI**I**QRKK	41	19-88
CSP1-L4I/L13NL	EMR**I**SKFFRDFI(**NL**)QRKK	26	13-52
CSP1-L4I/L13NV	EMR**I**SKFFRDFI(**NV**)QRKK	22	12-40
CSP1-L4NV/I12L	EMR(**NV**)SKFFRDF**L**LQRKK	25	12-53
CSP1-L4NV/I12NL	EMR(**NV**)SKFFRDF(**NL**)LQRKK	7.0	3.3-15
CSP1-L4NV/L13I	EMR(**NV**)SKFFRDFI**I**QRKK	80	55-120
CSP1-L4NV/L13NL	EMR(**NV**)SKFFRDFI(**NL**)QRKK	42	37-49
CSP1-L4NV/L13NV	EMR(**NV**)SKFFRDFI(**NV**)QRKK	4.3	3.3-5.5
CSP1-I12L/L13I	EMRLSKFFRDF**LI**QRKK	7.2	5.6-9.2
CSP1-I12L/L13NL	EMRLSKFFRDF**L**(**NL**)QRKK	11	8.5-14
CSP1-I12L/L13NV	EMRLSKFFRDF**L**(**NV**)QRKK	16	11-24
CSP1-I12NL/L13I	EMRLSKFFRDF(**NL**)**I**QRKK	21	11-38
CSP1-I12NL/L13NL	EMRLSKFFRDF(**NL**)(**NL**)QRKK	5.1	3.6-7.2
CSP1-I12NL/L13NV	EMRLSKFFRDF(**NL**)(**NV**)QRKK	46	23-95
CSP1-L4NV/I12NL/L13NL	EMR(**NV**)SKFFRDF(**NL**)(**NL**)QRKK	8.1	4.8-14
CSP1-L4NV/I12NL/L13NV	EMR(**NV**)SKFFRDF(**NL**)(**NV**)QRKK	3.3	3.1-3.5

[a] See experimental section for details of reporter strain and methods. See Appendix 1 for plots of agonism dose response curves. All bioassays were performed in triplicate. [b] EC_{50} values determined by testing peptides over a range of concentrations. [c] 95% confidence interval. [d] data from ref. [25].

Based on activation data of single mutations at position I12, the non-proteogenic I12NL substitution was shown to have the highest potency upon interacting with the receptor (I12NL, EC_{50} = 6.7 nM).[91] The persistence of this trend (observed in the 4th position as well) is evident in the SAR data of the CSP1 double mutants (**Table 2**). With the exception of L13NV, norleucine outperformed leucine in each analogous double mutant. The addition of an L4NV or L13NL substitution to the I12NL mutation resulted in two new ComD1 activators, L4NV/I12NL and I12NL/L13NL, exhibiting EC_{50} values of 7.0 nM and 5.1 nM, respectively (**Table 2**).

L13 does not reside directly within the hydrophobic binding pocket based on helical wheel analysis. Yet, it is still a vital residue, as replacement of leucine with alanine prevents

CSP1:ComD1 interaction.[47] The position does accommodate a variety of hydrophobic natural and non-natural amino acids; isoleucine, norleucine, and norvaline were all substituted and exhibited less than a two-fold change in potency.[91] Thus, these substitutions were carried forward into our library of CSP1 double mutants. The resultant bioassay data demonstrates that the L13 substitutions significantly alter potency when combined with mutations at other positions (**Table 2**). Even after accounting for 95% CI overlap, we observe a 10-fold change in potency changing from L13I to L13NL to L13NV. As observed with L4 and I12, L13 favors non-proteogenic substitutions. Interestingly, NV is the preferred substitution at position 13 when combined with a 4th position modification, while NL is the preferred substitution when combined with a 12th position modification. These results suggest that small alterations in branching and carbon number at the 13th position substantially impact CSP binding affinity.

To test for further synergy amongst the amino acids screened, the single substitutions present in the best ComD1 activators were combined to create triply-mutated CSP1 analogs. The most potent activators found were L4NV/I12NL (EC_{50} = 7.0 nM), L4NV/L13NV (EC_{50} = 4.3), and I12NL/L13NL (EC_{50} = 5.1 nM). These ComD1 activators were combined to produce two triply-substituted CSP1 analogs – L4NV/I12NL/L13NL and L4NV/I12NL/L13NV. These two analogs resulted in increased activation relative to the native CSP1 (**Table 2**). However, L4NV/I12NL/L13NL exhibited slightly reduced potency (EC_{50} = 8.1 nM) relative to the aforementioned double mutants, whereas L4NV/I12NL/L13NV exhibited increased potency (EC_{50} = 3.3 nM), providing us with the most potent ComD1 activator in this study.

Assessing ComD1 Inhibition by Singly and Multiply-Substituted CSP1 Derivatives

After identifying the most potent ComD1 activators in our multiple mutant library, the *N*-terminus of each peptide was modified in order to convert it into an inhibitor. A single substitution to the *N*-terminus of the native CSP1 peptide, the replacement of glutamic acid with alanine (E1A), results in the conversion to a potent ComD1 inhibitor, with an IC_{50} value of 85.7 nM.[47] We were interested to see if the enhancement in activation observed in remodeling the hydrophobic face of the peptide could be leveraged to enhance the inhibition when combined with the E1A substitution. We assessed the quality of our inhibitory peptide variants through initial competitive inhibition assays in the presence of the native signal. Multiply-substituted CSP1 derivatives containing the E1A modification that exhibited greater than 50% inhibition were carried forward in dose-response assays to obtain the corresponding IC_{50} values (for full details see the Experimental Section).The initial substitution of L4I produced an analog with an EC_{50} value of 10.2 nM, while the initial substitution of L4NV produced an analog with an EC_{50} value of 5.74 nM.[91] As evidenced by the data in **Table 3**, L4NV also yielded a more potent inhibitor when combined with the E1A modification compared to the inhibitor formed by the combination of the L4I and E1A modifications (IC_{50} of 140 nM and 520 nM, respectively). By comparing the IC_{50} data for E1A/L4NV and E1A/L4I, a correlation between the potency of the activating peptides and their corresponding inhibitors and the preference of ComD1 for non-proteogenic unbranched amino acids is evident.

Table 3: IC_{50} values of singly and multiply-substituted CSP1-E1A derivatives against the ComD1 receptor[a].

Peptide name	Peptide Sequence	IC_{50} [nM][b]	95% CI[c]
CSP1-E1A[e]	**A**MRLSKFFRDFILQRKK	86	51-150
CSP1-E1A/L4I	**A**MR**I**SKFFRDFILQRKK	520	340-810
CSP1-E1A/L4NV	**A**MR(**NV**)SKFFRDFILQRKK	140	81-230
CSP1-E1A/I12L	**A**MRLSKFFRDF**L**LQRKK	370	250-540
CSP1-E1A/I12NL	**A**MRLSKFFRDF(**NL**)LQRKK	130	60-280
CSP1-E1A/L13NL	**A**MRLSKFFRDFI(**NL**)QRKK	68	34-140
CSP1-E1A/L13I	**A**MRLSKFFRDFI**I**QRKK	320	200-510
CSP1-E1A/L13NV	**A**MRLSKFFRDFI(**NV**)QRKK	---[d]	---
CSP1-E1A/L4NV/I12NL	**A**MR(**NV**)SKFFRDF(**NL**)LQRKK	55	37-80
CSP1-E1A/I12NL/L13NL	**A**MRLSKFFRDF(**NL**)(**NL**)QRKK	430	190-970
CSP1-E1A/L4NV/L13NV	**A**MR(**NV**)SKFFRDFI(**NV**)QRKK	---[d]	---
CSP1-E1A/L4NV/I12NL/L13NL	**A**MR(**NV**)SKFFRDF(**NL**)(**NL**)QRKK	---[d]	---
CSP1-E1A/L4NV/I12NL/L13NV	**A**MR(**NV**)SKFFRDF(**NL**)(**NV**)QRKK	420	310-570

[a] See experimental section for details of reporter strain and methods. See Appendix 1 for plots of antagonism dose response curves. All bioassays were performed in triplicate. [b] IC_{50} values determined by testing peptides over a range of concentrations. [c] 95% confidence interval. [d] IC_{50} not determined due to the derivative low induction in primary antagonism screening assay. [e] data from ref. [25].

Evaluating the results of the 12th position analogs reinforce the notion that, generally, peptides with potent ComD1 activation can be effectively converted to potent inhibitors, especially with the addition of non-proteogenic amino acids. The initial I12L substitution afforded a moderate activator, while its inhibitor counterpart, E1A/I12L, displayed moderate inhibitory activity (E1A/I12L, IC_{50} = 370 nM). However, the E1A variants of the peptides containing non-natural substitutions exhibited greater potency. The combination of E1A and I12NL improved QS inhibition (E1A/I12NL, IC_{50} = 130 nM). Moreover, it was the combination of NV at the 4th position and NL at the 12th position that resulted in the most potent ComD1 inhibitor in our library (E1A/L4NV/I12NL, IC_{50} = 55 nM) (**Table 3**).

In the 13th position, E1A/L13NL yielded an IC_{50} value of 68 nM (**Table 3**). Yet, in combination with substitutions in other positions, we observe the inhibitory effects

decrease. For example, E1A/I12NL/L13NL resulted in a sharp increase in IC_{50} value to 430 nm. The results underscore the importance of the 13th position to ComD1 binding as well as the sensitivity of the position to additional substitutions. This idea is further underscored by data garnered after introducing the E1A substitutions into our triply-mutated activators, L4NV/I12NL/L13NL and L4NV/I12NL/L13NV. E1A/L4NV/I12NL/L13NL did not display any inhibitory activity, while E1A/L4NV/I12NL/L13NV produced an IC_{50} value of 420 nM (**Table 3**).

Ability of lead analogs to modulate competence induction

Lastly, we set out to evaluate the ability of our lead compounds (best activator and best inhibitor) to modulate a QS-regulated phenotype, competence induction. To this end, we utilized an antibiotic resistance transformation assay. A chloramphenicol resistance plasmid (pEVP3) was introduced to wildtype *S. pneumoniae* (D39) in the presence of CSP1, our lead activator CSP1-L4NV/I12NL/L13NV, or CSP1 together with our lead inhibitor CSP1-E1A/L4NV/I12NL, and incubated for 2 h. The bacteria were then spread-plated on THY agar plates containing 5% horse serum and chloramphenicol, and competence induction was assessed through colony formation. As expected, CSP1 was able to induce competence, resulting in many transformants (**Figure 12**, left section). Our lead activator, CSP1-L4NV/I12NL/L13NV, was also able to effectively induce competence, leading to many transformants (**Figure 12**, right section). Contrary, the addition of CSP1 together with our lead inhibitor, CSP1-E1A/L4NV/I12NL, did not yield any observed transformants, suggesting that our lead inhibitor was able to block CSP1-mediated competence induction (**Figure 12**, bottom section).

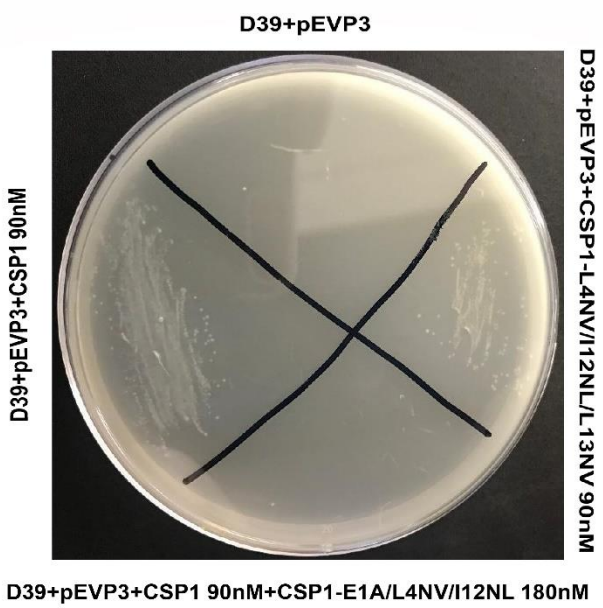


Figure 12. Transformation assay of *S. pneumoniae* D39 in the presence of CSP1, CSP1-L4NV/I12NL/L13NV, and CSP1 together with CSP1-E1A/L4NV/I12NL. The ability of D39 to internalize a chloramphenicol resistance plasmid (pEVP3) was evaluated following treatment with CSP1, CSP1-L4NV/I12NL/L13NV, or CSP1 + CSP1-E1A/L4NV/I12NL. Following CSP1 or CSP1-L4NV/I12NL/L13NV treatment, D39 was able to internalize chloramphenicol resistance, as can be seen by the number of transformants (left and right sections, respectively). Contrary, D39 was unable to internalize chloramphenicol resistance following treatment with CSP1 + CSP1-E1A/L4NV/I12NL, as determined by the lack of apparent transformants (bottom section).

Conclusion

In this work, we sought to further characterize the interactions between CSP1 and the cognate ComD1 receptor by assessing how multiple, coordinated substitutions in the CSP1 amino acid sequence can affect CSP1 binding. We developed a library of CSP1 peptide mutants that further defines the hydrophobic binding surface of the cognate ComD1 receptor in *S. pneumoniae*. Through point mutations that elongated, truncated, and rearranged the side chains of residues on the hydrophobic face of the QS autoinducer, we were able to gain more information about the steric limitations of each position. For example, if a substituted amino acid has a side chain that is too large, it causes a steric clash and CSP1 is no longer able to fit into the ComD1 receptor as tightly. This ultimately leads to unoptimized binding and a decrease in dose-response potency. When an amino acid side

chain is too small, that position will bind too loosely to the receptor and reduction in potency is observed. Notably, several of the non-proteogenic substitutions incorporated into our library increased potency, suggesting that the binding interaction between CSP1 and ComD1 is unoptimized at certain locations. Thus, the use of non-proteogenic amino acids has more clearly defined the spatial limitations (or lack thereof) on the ComD1 binding surface, revealing new avenues for enhanced QS control. Additionally, non-proteogenic amino acids often provide pharmacokinetic advantage, such as reducing rates of proteolytic degradation by enzymes. The inability of some enzymes to recognize these non-native amino acids increases the longevity of the peptides that contain them, an important consideration for peptide-drug candidates.[134]

The introduction of non-proteogenic amino acids at key positions (4, 12, and 13) in the primary structure enhanced potency relative to the native peptide. Position 4 was optimized through the substitution of leucine with norvaline (NV), which changed the side chain from a hydrophobic, branched amino acid to a hydrophobic, linear amino acid. Similarly, position 12 was optimized through the substitution of isoleucine with norleucine (NL), which again replaced a hydrophobic, branched amino acid with a hydrophobic, linear amino acid. We believe both substitutions provided an additional degree of flexibility to maximize the noncovalent interactions between the peptide and ComD1 without disrupting other critical interactions with other residues along the peptide backbone. Finally, we were able to observe the effect of mutations of the 13th position on the binding of CSP1. Although the 13th position does not reside within the CSP1 binding face, our data suggest that mutating this position has significant effects on the binding of CSP1 when combined with other mutations. In sum, three new doubly-mutated potent ComD1 activators were

developed, L4NV/I12NL, L4NV/L13NV, and I12NL/L13NL. We combined our most successful mutations to develop the first triple-substituted CSP1 analogs, which resulted in our most potent ComD1 activator in this library, L4NV/I12NL/L13NV, three times more potent on average than the native CSP1 itself (**Figure 13**).

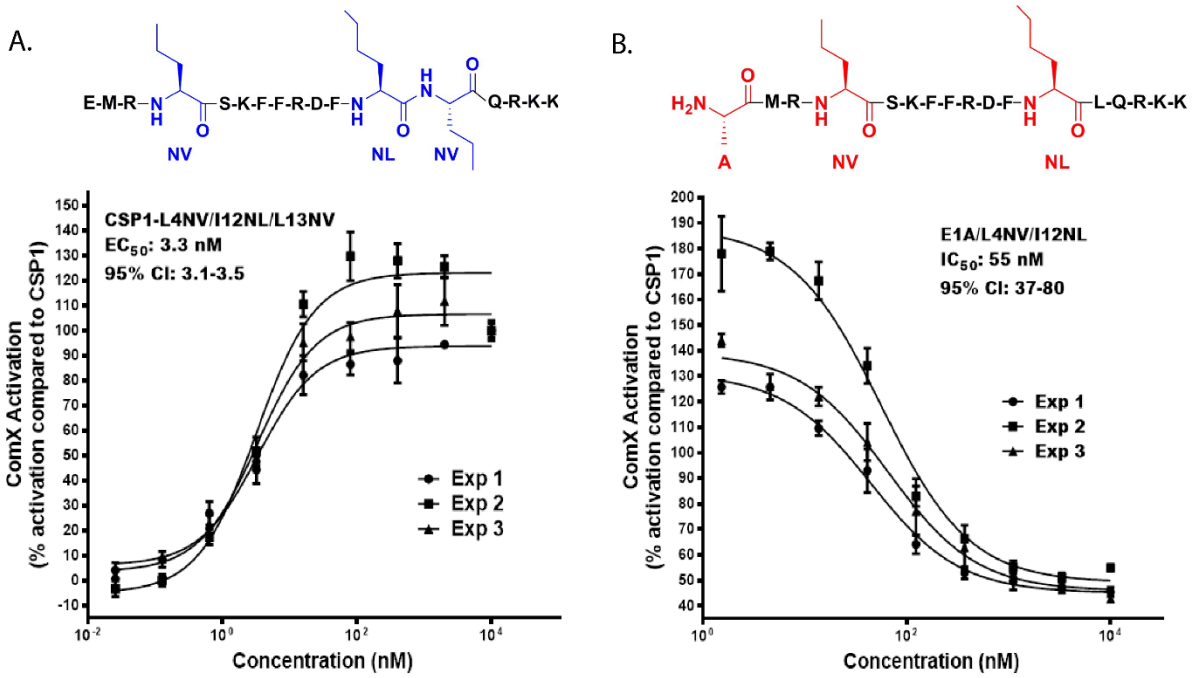


Figure 13. Dose response curves of the lead ComD1 activator and inhibitor identified in this study and their structures, highlighting the changes in structure compared to CSP1.

Starting from the most potent activating peptides, we attempted to construct potent inhibitory peptides using the E1A substitution. Through this analysis, several low nanomolar inhibitors were discovered as a result, including E1A/L4NV/I12NL with an IC_{50} = 55 nM (**Figure 13**). Finally, we evaluated the ability of our lead activator and inhibitor to modulate the hallmark QS phenotype regulated by the competence regulon in *S. pneumoniae*, genetic transformation. Indeed, our results indicate that our lead activator can effectively induce the acquisition of genetic material from the environment whereas our lead inhibitor can effectively block transformation induced by CSP1.

Beyond providing further structure-activity data regarding the CSP1:ComD1 interaction, the most potent activator and inhibitor developed through this study contain at least two non-proteogenic amino acids. Based on the precedent for such amino acids to thwart degradation by proteolytic enzymes, these derivatives may present a pharmacokinetic advantage over peptides composed of only natural residues. Thus, this study is an important contribution to the consideration of QS regulation in general and ComX regulation in particular as a potential drug target.

Experimental Details

Chemical reagents and instrumentation: All chemical reagents and solvents were purchased from Sigma-Aldrich or Chem-Impex and used without further purification. Water (18 MΩ) was purified using a Millipore Analyzer Feed System. Solid-phase resins were purchased from Advanced Chem Tech or P3 Biosystems.

Reversed-phase high-performance liquid chromatography (RP-HPLC) was performed using two Shimadzu systems each equipped with a CBM-20A communications bus module, two LC-20AT pumps, an SIL-20A auto sampler, an SPD-20A UV/VIS detector, a CTO-20A column oven, one with an FRC-10A fraction collector and one without. All RP-HPLC solvents (18 MΩ water and HPLC-grade acetonitrile (ACN)) contained 0.1% trifluoroacetic acid (TFA). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) data were obtained on either a Bruker Autoflex or Bruker Microflex spectrometer equipped with a 60 Hz nitrogen laser and a reflectron. In positive ion mode, the acceleration voltage on Ion Source 1 was 19.01 kV. Exact mass (EM) data were obtained on an Agilent Technologies 6230 TOF LC/MS spectrometer. The samples

were sprayed with a capillary voltage of 3500 V and the electrospray ionization (ESI) source parameters were as follows: gas temperature of 325 °C at a drying gas flow rate of 8 L/min at a pressure of 35 psi.

Peptide Synthesis: All the CSP1 analogs were synthesized using standard Fmoc-based solid-phase peptide synthesis (SPPS) procedures on preloaded Fmoc-L-Lys (Boc) 4-Benzyloxybenzyl alcohol (Wang) resin (0.59 mmol/g) by using the Discover Microwave Synthesizer or the Liberty1 automated peptide synthesizer (CEM Corporation). For peptides synthesized on the automated synthesizer, the resin (0.1 g) was first swelled by suspension in *N*,*N*-dimethylformamide (DMF) for 15 minutes at room temperature and then drained. Fmoc-protecting group removal was achieved with treatment of the resin by 5 mL of 20% piperidine in DMF (90 sec, 90 °C) followed by another 5 mL of 20% piperidine in DMF (90 sec, 90 °C). The resin was washed with DMF (3 x 5 mL) after each deprotection cycle. To couple each amino acid, Fmoc-protected amino acids (5 equiv. relative to the overall loading of the resin) were dissolved in DMF (5 mL) and mixed with 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU; 5 equiv.) and diisopropylethylamine (DIPEA; 5 equiv.). All amino acids were coupled for 5 minutes (30 W, 75 °C). After each coupling step, the resin was drained and washed with DMF (2 × 5 mL). This process was repeated until the desired peptide sequence was obtained. The same swelling and deprotection protocols were followed for peptides synthesized on the manual microwave synthesizer. However, 0.2 g of Wang resin was used and couplings were completed using diisopropyl carbodiimide (DIC) and Oxyma Pure with a 3.6:3:3 ratio of DIC:Oxyma:AA in DMF for a final DIC concentration of 0.2 M. Each coupling was run at 75 °C for 8 minutes (50 W).

Cleavage: The resin containing the final peptide product was washed with diethyl ether (2 mL) and dried under nitrogen stream for 3 minutes before transferring it into a 15 mL falcon tube. Three mL cleavage cocktail of 2.5% 18 MΩ water and 2.5% triisopropylsilane (TIPS) in 95% trifluoroacetic acid (TFA) for every 0.1 g of resin was then added and the tube was shaken for 3 hours. The resulting cleavage product solution was filtered through a cotton ball and the filtrate was transferred into a new 50 mL falcon tube. A cooled solution of diethyl ether:hexane (1:1, 45 mL, -20 °C) was added to the tube, and the tube was kept in a freezer at -20 °C for 10 minutes in order to precipitate the crude peptide. The pellet of the crude peptide was obtained by centrifugation of the 50 mL tube in a Beckman Coulter Allegra 6 centrifuge equipped with a GH3.8 rotor at 3000 RPM for 5 minutes. The supernatant was poured off and the remaining crude peptide was dissolved in 10 mL acetonitrile (ACN):water (1:1) and lyophilized before HPLC purification.

Peptide purification: Crude peptides were purified with RP-HPLC. The crude peptide was dissolved in ACN:H_2O (1:4) (volume of ACN in water depends on the solubility of the peptide) and purified in 4 mL portions on either a Phenomenex Luna 5 μm C18 semi-preparative column (10 mm × 150 mm, 100 Å) or a Phenomenex Kinetex 5 μm C18 semi-preparative column (10 mm × 250 mm, 110 Å) with a flow rate of 5 mL/min; mobile phase A = 18 MΩ water + 0.1 % TFA and mobile phase B = ACN + 0.1 % TFA. The collected fraction was lyophilized overnight and dissolved again in 5 mL ACN:H_2O (1:4) for a secondary purification run. Preparative HPLC methods were used for the crude purification using a linear gradient (first prep 5% B → 45% B over 40 min and second prep 25% B → 38% B over 45 min). Fraction purity was determined through analysis on either a Phenomenex Luna 5 μm analytical C18 column (4.6 mm × 150 mm, 100 Å) or a

Phenomenex Kinetex 5 μm analytical C18 column (4.6 mm × 250 mm, 110 Å) using a linear gradient (5% B → 95% B over 22 min or 27 min, respectively). Purities were determined by integration of peaks with UV detection at 220 nm. Only peptide fractions that were purified to homogeneity (>95%) were used for the biological assays. For the final verification of the peptides, a high-resolution MS was used to verify the exact masses of the peptides. The observed mass-to-charge (*m*/*z*) ratio of the peptide was compared to the expected *m*/*z* ratio for each peptide.

Biological Reagents and Strain Information: All standard biological reagents were purchased from Sigma-Aldrich. Donor horse serum (defibrinated) was purchased from Sigma-Aldrich and stored at -20 °C until use in pneumococcus bacterial culture. To examine the ability of the synthesized CSP1 analogs to modulate the ComD1 receptor, and thus, the QS circuit in *S. pneumoniae,* beta-galactosidase assays were performed using the D39pcomX::lacZ reporter strain whereas transformation assays were performed using *S. pneumoniae* D39.[42,47] For the transformation assays, pEVP3 (*Escherichia coli-Streptococcus* shuttle vector containing *cat* cassette; Cm^r)[135] was used to evaluated competence induction.

Bacterial Growth Conditions: Freezer stocks were created from 1.5 mL aliquots of overnight cultures (0.2 OD 600_{nm}) in Todd-Hewitt broth supplemented with 0.5% yeast extract (THY) and 0.5 mL glycerol, and stored at -80 °C. For the Beta-galactosidase experiments, Frozen stocks of pneumococcal strain were streaked into a THY agar plate containing 5% horse serum and chloramphenicol at a final concentration of 4 μg/mL. The plate was incubated for 8-9 hours in 37 °C with 5% CO_2. An isolated fresh single colony was picked into sterilized cultural tube containing 5 mL of THY broth supplemented with

a final concentration of 4 μg/mL chloramphenicol, and the culture was incubated in a CO_2 incubator overnight (15 h). Overnight culture was then diluted (1:50) with THY and the resulting solution was incubated in a CO_2 incubator for 3-4 hours, until the bacteria reached early exponential stage (0.30-0.35) as determined by using a plate reader. For the transformation assays, bacteria from frozen stocks were plated on THY agar plate containing 5% horse serum without antibiotics and incubated for 12-24 h in a CO_2 incubator. Isolated fresh colonies were transferred using a sterile transfer loop to 5 mL of THY broth and incubated overnight (15 h) in a CO_2 incubator. Overnight culture was then diluted 1:100 with THY and the diluted culture was incubated in a CO_2 incubator for 1-2 hours, until the bacteria reached an OD_{600} value of 0.10, as determined by using a plate reader.

Beta-Galactosidase Activation Assays: The ability of synthetic CSP1 analogs to activate the expression of *S. pneumoniae comX* was determined using a reporter strain grown in THY (pH 7.3). An initial activation screening was performed at high concentration (10 μM) for all CSP1 analogs. Two μL of 1 mM solution of CSP1 analogs in dimethyl sulfoxide (DMSO) were added in triplicate to a clear 96-well microtiter plate. Two μL of 20 μM solution of CSP1 (200 nM final concentration) were added in triplicate and served as the positive control. This concentration was chosen to afford full activation of the quorum sensing (QS) circuit, as determined from the dose-dependent curve created for the native CSP1. Two μL of DMSO were added in triplicate and served as the negative control. Then, 198 μL of pneumococcal cultures were added to each well containing CSP1 or analogs, the plate was incubated at 37 ºC for 30 minutes, and the absorbance at 600_{nm} was measured. In order to measure the beta-galactosidase activity in the pneumococcal culture, the wells

were then treated with 20 μL of 0.1% Triton X-100 for 30 minutes at 37 °C to lyse the bacterial cells. Then, 100 μL of Z-buffer solution (60.2 mM Na_2HPO_4, 45.8 mM NaH_2PO_4, 10 mM KCl, and 1.0 mM $MgSO_4$ in 18 MΩ H_2O; pH was adjusted to 7.0 and sterilized before use) containing 2-Nitrophenyl-Beta-D-galactopyranoside (ONPG) at a final concentration of 0.4 mg/mL was added in a new plate, followed by 100 μL of lysate, and the plate was incubated for 3 hours at 37 °C. The reaction was quenched by adding 50 μL of 1 M sodium carbonate solution, and the OD 420_{nm} and OD 550_{nm} were measured using a plate reader. The results were reported as percent activation, which is the ratio between the Miller units of the analog and that of the positive control. For calculation of Miller units, please see data analysis below. Analogs that exhibited high activity (>75% activation compared to the native signal) in the initial screening were further evaluated using a dose-dependent assay in which peptide stock solutions were diluted with DMSO in serial dilutions (either 1:2, 1:3, or 1:5) and assayed as described above. The EC_{50} (the concentration of a drug that gives half maximal response) value was then determined through fitting using nonlinear regression with GraphPad Prism 5.

Beta-Galactosidase Inhibition Assays: Analogs that exhibited low competence activation in the initial screening were evaluated for competitive inhibition. The ability of synthesized CSP1 analogs to inhibit the expression of *comX* by outcompeting CSP1 for the receptor binding site was evaluated using the same assay conditions as described above, except that in this case native CSP1 (2 μL, 50 nM final concentration) was added to every well as a competitor standard. This concentration was chosen to afford full activation of the QS circuit, as determined from the dose-dependent curve created for CSP1. Two μL of native CSP1 (5 μM solution) and 2 μL of 1 mM solution of CSP1 analogs were added to the same

well in triplicate in a clear 96-well microtiter plate. Two μL of CSP1 (5 μM solution) and 2 μL of DMSO were added to the same well in triplicate and served as the positive control. Four μL of DMSO were added in triplicate and served as the negative control. Then, 196 μL of pneumococcal cultures were added to each well and the plate was incubated at 37 ºC for 30 minutes. The procedure for lysis, incubation with ONPG and all the measurements were as described in the beta-galactosidase activation assay. Analogs that exhibited significant competitive inhibition in the initial screening were further evaluated using a dose-dependent assay where a series of dilutions of the CSP1 peptide analogs was prepared similarly to those made for the EC_{50} assay, and the IC_{50} (the concentration of an inhibitor where the response or binding is reduced by half) was then determined by using GraphPad Prism 5.

Analysis of Activation/Inhibition Data: Miller units were calculated using the following formula:

$$Miller\ Unit = 1000 \times \frac{[Abs_{420} - (1.75 \times Abs_{550})]}{(t \times v \times Abs_{600})}$$

Abs_{420} is the absorbance of o-nitrophenol (ONP). Abs_{550} is the scatter from cell debris, which, when multiplied by 1.75 approximates the scatter observed at 420 nm. *t* is the duration of incubation with ONPG in minutes, *v* is volume of lysate in milliliters, and Abs_{600} reflects cell density.

Transformation Assays: *S. pneumonaie* D39 cells were centrifuged at 5000 rpm for 10 min and the THY medium was discarded. Cells were then washed two times and centrifuged (5 min at 5000 rpm) with 5 mL sterile PBS buffer. PBS buffer was discarded,

and the cells were resuspended in THY containing 5% horse serum media (pH 8.3) and vortexed for 10 s followed by 30-60 min of incubation in a CO_2 incubator (37 °C with 5% CO_2). CSP1 (at a final concentration of 90 nM), the lead activator, CSP1-L4NV/I12NL/L13NV (at a final concentration of 90 nM), or CSP1 (at a final concentration of 90 nM) together with the lead inhibitor, CSP1-E1A/L4NV/I12NL (at a final concentration of 180 nM) were added to the culture. After the addition of the pEVP3 plasmid (3-4 μg) to the final 1 mL culture, the resulting transformation mix was incubated for 2 h in a CO_2 incubator before being spread-plated on THY plates supplemented with 5% horse serum and 4 μg/mL of chloramphenicol. The plates were then incubated at 37 °C in a CO_2 incubator for 20 to 48 h, and colony formation was evaluated. This experiment was performed in triplicate and on three separate days.

Corresponding Authors

bertuccim@moravian.edu; ytalgan@unr.edu

Supporting Information (Appendix 1)

HPLC traces and MS data, initial screening results, and dose response curves for CSP1 analogs. This information is available online.

Notes

The authors declare no competing financial interest.

Acknowledgements

This work was supported by a grant from the National Science Foundation (CHE-1808370). The *S. pneumoniae* D39 wildtype strain and D39pcomX::lacZ reporter strain were a generous gift from G. W. Lau (University of Illinois at Urbana−Champaign).

References

[1] S. Mehr, N. Wood, *Paediatric respiratory reviews* **2012**, *13*, 258-264.

[2] K. L. O'Brien, L. J. Wolfson, J. P. Watt, E. Henkle, M. Deloria-Knoll, N. McCall, E. Lee, K. Mulholland, O. S. Levine, T. Cherian, *The Lancet* **2009**, *374*, 893-902.

[3] S. S. Huang, K. M. Johnson, G. T. Ray, P. Wroe, T. A. Lieu, M. R. Moore, E. R. Zell, J. A. Linder, C. G. Grijalva, J. P. Metlay, *Vaccine* **2011**, *29*, 3398-3412.

[4] N. J. Croucher, S. R. Harris, C. Fraser, M. A. Quail, J. Burton, M. van der Linden, L. McGee, A. von Gottberg, J. H. Song, K. S. Ko, *science* **2011**, *331*, 430-434.

[5] J. Cornick, S. Bentley, *Microbes and infection* **2012**, *14*, 573-583.

[6] D. R. Feikin, E. W. Kagucia, J. D. Loo, R. Link-Gelles, M. A. Puhan, T. Cherian, O. S. Levine, C. G. Whitney, K. L. O'Brien, M. R. Moore, *PLoS Med* **2013**, *10*, e1001517.

[7] P. K. Mitchell, M. Lipsitch, W. P. Hanage, *Philosophical Transactions of the Royal Society B: Biological Sciences* **2015**, *370*, 20140342.

[8] W. P. Hanage, C. Fraser, J. Tang, T. R. Connor, J. Corander, *Science* **2009**, *324*, 1454-1457.

[9] C. Chewapreecha, S. R. Harris, N. J. Croucher, C. Turner, P. Marttinen, L. Cheng, A. Pessia, D. M. Aanensen, A. E. Mather, A. J. Page, *Nature genetics* **2014**, *46*, 305-309.

[10] D. J. Engelmoer, I. Donaldson, D. E. Rozen, *PLoS Pathog* **2013**, *9*, e1003758.

[11] L. S. Håvarstein, R. Hakenbeck, P. Gaustad, *Journal of Bacteriology* **1997**, *179*, 6589-6594.

[12] G. Salvadori, R. Junges, D. A. Morrison, F. C. Petersen, *Frontiers in cellular and infection microbiology* **2019**, *9*, 94.

[13] T. A. Milly, Y. Tal-Gan, *RSC Chemical Biology* **2020**.

[14] E. Pestova, L. Håvarstein, D. Morrison, *Molecular microbiology* **1996**, *21*, 853-862.

[15] F. M. Hui, L. Zhou, D. A. Morrison, *Gene* **1995**, *153*, 25-31.

[16] J.-P. Claverys, B. Martin, P. Polard, *FEMS microbiology reviews* **2009**, *33*, 643-656.

[17] D. G. Cvitkovitch, Y.-H. Li, R. P. Ellen, *The Journal of clinical investigation* **2003**, *112*, 1626-1632.

[18] L. S. Håvarstein, G. Coomaraswamy, D. A. Morrison, *Proceedings of the National Academy of Sciences* **1995**, *92*, 11140-11144.

[19] D. L. Hava, A. Camilli, *Molecular microbiology* **2002**, *45*, 1389-1406.

[20] M. S. Lee, D. A. Morrison, *Journal of Bacteriology* **1999**, *181*, 5004-5016.

[21] O. Ween, P. Gaustad, L. S. Håvarstein, *Molecular microbiology* **1999**, *33*, 817-827.

[22] L. Zhu, G. W. Lau, *PLoS pathogens* **2011**, *7*, e1002241.

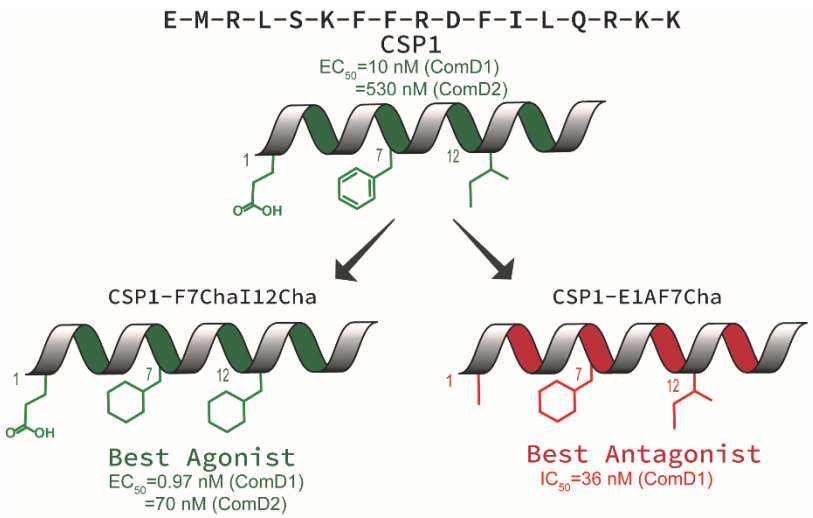
[23] C. Duan, L. Zhu, Y. Xu, G. W. Lau, *PloS one* **2012**, *7*, e44710.

[24] B. Koirala, J. Lin, G. W. Lau, Y. Tal-Gan, *ChemBioChem* **2018**, *19*, 2380-2386.

[25] Y. Yang, B. Koirala, L. A. Sanchez, N. R. Phillips, S. R. Hamry, Y. Tal-Gan, *ACS chemical biology* **2017**, *12*, 1141-1151.

[26] D. N. McBrayer, C. D. Cameron, Y. Tal-Gan, *Organic & Biomolecular Chemistry* **2020**, *18*, 7273-7290.

[27] Y. Yang, J. Lin, A. Harrington, G. Cornilescu, G. W. Lau, Y. Tal-Gan, *Proceedings of the National Academy of Sciences* **2020**.
[28] B. Koirala, Y. Tal-Gan, *ChemBioChem* **2020**, *21*, 340-345.
[29] J. R. Brannon, M. Hadjifrangiskou, *Drug design, development and therapy* **2016**, *10*, 1795.
[30] Y. Tal-Gan, D. M. Stacy, M. K. Foegen, D. W. Koenig, H. E. Blackwell, *Journal of the American Chemical Society* **2013**, *135*, 7869-7882.
[31] Y. Tal-Gan, M. Ivancic, G. Cornilescu, T. Yang, H. E. Blackwell, *Angewandte Chemie International Edition* **2016**, *55*, 8913-8917.
[32] D. N. McBrayer, B. K. Gantman, C. D. Cameron, Y. Tal-Gan, *Organic letters* **2017**, *19*, 3295-3298.
[33] L. Cegelski, G. R. Marshall, G. R. Eldridge, S. J. Hultgren, *Nature Reviews Microbiology* **2008**, *6*, 17-27.
[34] G. Pozzi, L. Masala, F. Iannelli, R. Manganelli, L. Havarstein, L. Piccoli, D. Simon, D. Morrison, *Journal of Bacteriology* **1996**, *178*, 6087-6090.
[35] A. M. Whatmore, V. A. Barcus, C. G. Dowson, *Journal of bacteriology* **1999**, *181*, 3144-3154.
[36] Y. Yang, G. Cornilescu, Y. Tal-Gan, *Biochemistry* **2018**, *57*, 5359-5369.
[37] O. Johnsborg, P. E. Kristiansen, T. Blomqvist, L. S. Håvarstein, *Journal of bacteriology* **2006**, *188*, 1744-1749.
[38] B. Koirala, R. A. Hillman, E. K. Tiwold, M. A. Bertucci, Y. Tal-Gan, *Beilstein journal of organic chemistry* **2018**, *14*, 1769-1777.
[39] W. Chan, P. White, *Fmoc solid phase peptide synthesis: a practical approach, Vol. 222*, OUP Oxford, **1999**.
[40] J. S. Khara, M. Priestman, I. Uhía, M. S. Hamilton, N. Krishnan, Y. Wang, Y. Y. Yang, P. R. Langford, S. M. Newton, B. D. Robertson, *Journal of Antimicrobial Chemotherapy* **2016**, *71*, 2181-2191.
[41] J. P. Claverys, A. Dintilhac, E. V. Pestova, B. Martin, D. A. Morrison, *Gene* **1995**, *164*, 123-128.



Chapter 3: Optimizing CSP1 Analogs for Modulating Quorum Sensing in *Streptococcus pneumoniae* with Bulky, Hydrophobic Nonproteogenic Amino Acid Substitutions

[a]Reprinted with permission from Milly, T. A.; Buttner, A. R.; Rieth, N.; Hutnick, E.; Engler, E. R.; Campanella, A. R.; Lella, M.; Bertucci, M. A.; Tal-Gan, Y. Optimizing CSP1 analogs for modulating quorum sensing in Group 1 *Streptococcus pneumoniae* with bulky, hydrophobic nonproteogenic amino acid substitutions.. *RSC Chem Biol* **2022**, 3, 301-311.

Optimizing CSP1 Analogs for Modulating Quorum Sensing in *Streptococcus pneumoniae* with Bulky, Hydrophobic Nonproteogenic Amino Acid Substitutions

Tahmina A. Milly,[†] Alec R. Buttner,[§] Naomi Rieth,[§] Elizabeth Hutnick,[§] Emilee R. Engler,[§] Alexandra R. Campanella,[‡] Muralikrishna Lella,[†] Michael A. Bertucci*[,‡] and Yftah Tal-Gan,*[,†]

[†] Department of Chemistry, University of Nevada, Reno, 1664 North Virginia Street, Reno, Nevada, 89557, United States

[‡] Department of Chemistry, Lafayette College, 701 Sullivan Rd., Easton, PA 18042, United States

[§] Department of Chemistry, Moravian University, 1200 Main St., Bethlehem, PA 18018, United States

Abstract

The prompt appearance of multiantibiotic-resistant bacteria necessitates finding alternative treatments that can attenuate bacterial infections while minimizing the rate of antibiotic resistance development. *Streptococcus pneumoniae*, a notorious human pathogen, is responsible for severe antibiotic-resistant infections. Its pathogenicity is influenced by a cell-density communication system, termed quorum sensing (QS). As a result, controlling QS through the development of peptide-based QS modulators may serve to attenuate pneumococcal infections. Herein, we set out to evaluate the impact of the introduction of bulkier, nonproteogenic side-chain residues on the hydrophobic binding face of CSP1 to optimize receptor-binding interactions in both of the *S. pneumoniae* specificity groups. Our results indicate that these substitutions optimize the peptide-protein binding interactions, yielding several pneumococcal QS modulators with high potency. Moreover, pharmacological evaluation of lead analogs revealed that the incorporation of nonproteogenic amino acids increased the peptides' half-life towards enzymatic degradation while remaining nontoxic. Overall, our data conveys key considerations for SAR using nonproteogenic amino acids, which provide analogs with better pharmacological properties.

Introduction

Bacteria coordinate their various physiological behaviors and control gene expression in response to changes in cell density by utilizing intercellular chemical signaling pathways in a process known as quorum sensing (QS).[136] QS involves the detection of a small signaling molecule known as an autoinducer that is synthesized and then actively or passively secreted. In Gram-positive bacteria, the signals that mediate this self-propagating mechanism are generally peptides and, thus, are referred to as autoinducing peptides (AIPs).[90] The AIPs are detected by a membrane-bound receptor, typically a histidine-kinase receptor, prompting a response when the AIP reaches a certain threshold concentration.[136,137] This sensing enables bacterial cells to communicate with the other neighboring cells and establish different group-beneficial traits, including the production of virulence factors, sporulation, bioluminescence, root nodulation, swarming, biofilm formation, and competence development.[118,119,138,139] Specifically, many human pathogenic species utilize QS to effectively attack their host and establish an infection. As such, QS has gained significant attention as an anti-pathogenic drug target for the development of novel therapeutics, especially considering the ongoing emergence of multidrug resistant bacterial pathogens.[42,45,57,58,140,141] Interception of QS circuits' signal-receptor interactions using signal molecule-based drugs would lead to an attenuation of bacterial pathogenicity rather than induction of cell death, thus limiting the potential for resistance development while still preventing many bacterial pathogenic traits.[126,142]

Streptococcus pneumoniae, or pneumococcus, is a Gram-positive bacterium that populates the nasopharyngeal cavity and upper respiratory tract of humans. As an opportunistic

pathogen, it is responsible for more than 1 million pneumococcal infections including bacteremia, sepsis, meningitis, and pneumonia in the United States alone.[102,103] Moreover, recombinogenic pneumococcus is intrinsically resistant to several antibiotics such as vancomycin, linezolid, quinolones, and beta lactams.[105,106,109] A major contributor to the development of antibiotic resistant strains is the ability of pneumococci to exchange intra- and inter-species genetic material with other neighboring species.[45,109,112,143] Specifically, competent pneumococci have been shown to acquire antimicrobial resistance and virulence genes through transformation from its closely related species, *Streptococcus mitis*, which provides a significant advantage to pneumococci by enabling rapid evolution of the genome and capsular diversity.[45,112,143] Therefore, the costs associated with pneumococcal infections as well as high-speed accumulation of multi-drug resistance by *S. pneumoniae* necessitates treating this pathogen using alternative approaches. In *S. pneumoniae,* the acquisition of antibiotic-resistance genes and pathogenicity is directly associated with the activation of the pheromone-responsive competence regulon, a conserved QS circuit (**Figure. 14**).[115] Thus, this QS circuitry can be utilized as an excellent target for the design of anti-virulence drug leads to control pneumococcal infections.

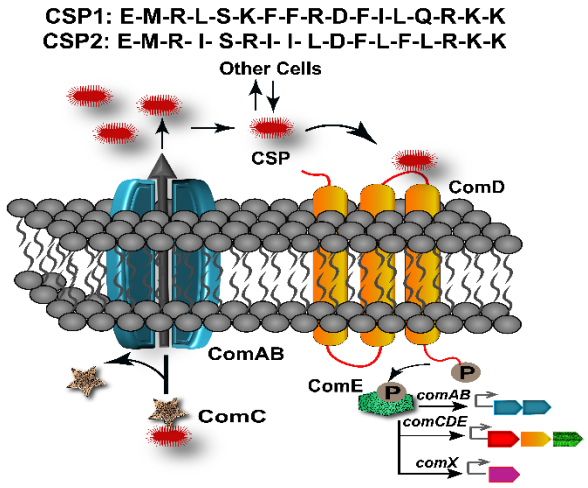


Figure 14: *S. pneumoniae* CSP-mediated QS circuit. The pre-CSP peptide, ComC, is being processed and secreted by the ComAB transporter as the mature CSP signal. At high concentration, CSP can effectively bind and activate a transmembrane histidine kinase receptor, ComD, which, after being activated, transfers a phosphate group to its cognate response regulator, ComE. Phosphorylated ComE then triggers the transcription of numerous genes, including the effector molecule of the circuitry, ComX, which regulates QS-mediated phenotypes.

In *S. pneumoniae*, the competence regulon is triggered by a 17-amino acid AIP termed the competence stimulating peptide (CSP, **Figure 14**).[118] With the help of a proteolytic ATP binding cassette (ABC) transporter, (ComAB, **Figure 14**), the CSP pro-peptide, ComC, is processed and the mature signaling molecule, CSP, is exported out of the cell.[118] Upon reaching a threshold concentration, CSP can effectively bind and activate a membrane-bound histidine kinase receptor, ComD, resulting in phosphorylation of the response regulator, ComE.[43,115,121] Phosphorylated ComE then acts as a transcription factor and initiates the transcription of the *comAB* and *comCDE* genes, resulting in autoinduction of the QS circuitry. ComE also initiates the transcription of the gene for the effector molecule of the QS circuit, the alternative sigma factor, ComX, which controls different QS-regulated phenotypes.[121,144,145] The majority of *S. pneumoniae* strains can be divided into two main pherotypes or specificity groups based on the AIP they produce (CSP1 or CSP2, **Figure 14**), along with their own specific ComD receptor (ComD1 or ComD2, respectively).[131] These two peptide pheromones share approximately 50% sequence

similarity, with most of the variation occurring amongst hydrophobic residues in the central region of the pheromone, allowing them to confer high selectivity toward their respective cognate receptors.[47,94,131,146]

The pneumococcal competence regulon communication pathway can be modulated through impediment of the peptide-receptor interaction by using synthetic AIP analogs. To improve the potency and pharmacological properties of CSP-based QS modulators, Yang et al. previously performed a systematic structure-activity relationship (SAR) analysis of the native CSP1 signal.[47,94] The results of their studies suggested that an α-helix is the bioactive conformation of CSP1, and that the hydrophobic side of the helix plays a crucial role in the binding of CSP1 to ComD1. Specifically, hydrophobic residues in positions 4, 7, 8, 11 and 12 of the CSP1 sequence form a hydrophobic patch that spans two full helical turns and stabilizes CSP1-ComD1 binding. These results suggest that any structural change affecting this hydrophobic patch will alter the CSP1-receptor binding interaction.[94] In a previous study, we incorporated highly conservative point mutations to the hydrophobic side-chain residues in these positions of the CSP1 sequence using both proteogenic and nonproteogenic amino acids.[91] Information gained from this work suggested that the side-chain residues do not fully occupy the hydrophobic binding pockets and, thus, the CSP1-ComD1 binding interactions could be further optimized utilizing elongated side-chain residues. We advanced these results by incorporating multiple mutations containing several of the nonproteogenic amino acids in the CSP1 sequence.[92] The findings from this study provided several important structural insights, specifically the preference of ComD1 for linear, hydrophobic, nonproteogenic amino acids. Combined, the results obtained from these two studies revealed strong potential for even further optimization of the binding

interaction between CSP1 and ComD1. In this work, we focused on assessing the hydrophobic pockets within the ComD1 receptor through the introduction of bulkier and more hydrophobic nonproteogenic amino acids, namely a non-natural Phe-derivative, cyclohexylalanine (Cha), and an extended aliphatic hydrophobic residue, homoleucine (HLeu), in key hydrophobic positions (4, 7, 8, 11, and 12). To this end, we rationally designed and chemically synthesized a library of singly and multiply mutated CSP1 peptides intended to develop novel CSP-based QS modulators with enhanced activities against both pneumococcal ComD receptors. Our analysis revealed several nanomolar- and picomolar-range ComD1 and ComD2 activators containing these two nonproteogenic amino acids, suggesting that size and hydrophobicity, rather than the aromaticity of the amino acid side chains, dictate the stabilization of the binding interaction. Our next goal was to construct potent inhibitory peptides of the *S. pneumoniae* competence regulon by combining the lead CSP1 analogs with a previously characterized E1A substitution.[42] Indeed, through this analysis, we were able to develop several low nanomolar-range ComD1 inhibitors. Moreover, we evaluated the pharmacological properties of lead analogs and observed a significant increase in stability towards enzymatic degradation while maintaining low toxicity. In addition to yielding a series of new QS activators and inhibitors, our results provide valuable information regarding the ComD1 hydrophobic binding pockets and key SAR knowledge of the CSP1 pheromone. This information can be utilized for the rational design of highly potent, pharmacologically stable CSP-based QS modulators with therapeutic potential.

Results and Discussion

Design and synthesis of nonproteogenic CSP1 analogs

Designing optimized peptide ligands that interact with receptor proteins involves the enhancement of individual peptide-protein binding interactions.[147] For instance, an unoccupied binding site within the protein due to the steric limits of a proteogenic amino acid side chain could be optimized through substitution of a bulkier amino acid side chain (**Figure 15**). Contrary, when there is an unfavorable steric clash within the binding pocket due to the presence of larger amino acid side chain, the binding interactions could be optimized through utilization of a smaller amino acid side chain.

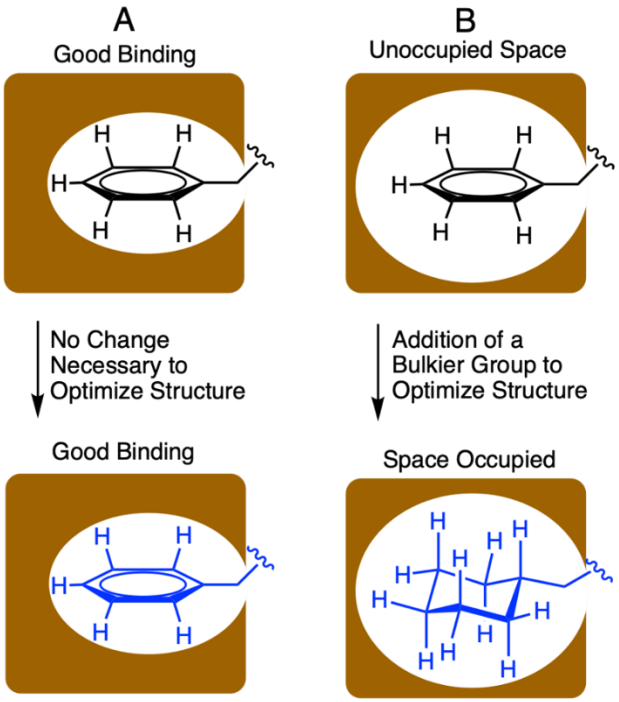


Figure 15: Portrayal of optimized and unoptimized protein-peptide binding interactions. (A) represents an optimized binding interaction where the side chain residue fills the binding pocket entirely. (B) represents an unoptimized binding interactions due to limited contacts of small side chain with the binding pocket. This interaction was improved through incorporation of a bulkier, nonproteogenic side chain.

In the context of *S. pneumoniae* CSP-1, our previous investigations of the CSP1-ComD1 binding interaction revealed that positions 4, 7, 8, 11, and 12 in CSP-1 have unoccupied space within the receptor binding site.[91,92] We aimed to assess the steric limit of the CSP1-ComD1 interaction by utilizing bulkier, hydrophobic nonproteogenic substituents. To this end, aliphatic and aromatic hydrophobic residues (Leu, Ile or Phe) in positions 4, 7, 8, 11, and 12 were substituted with the nonproteogenic amino acids, cyclohexylalanine (Cha) or homoleucine (HLeu) (**Figure 16**).

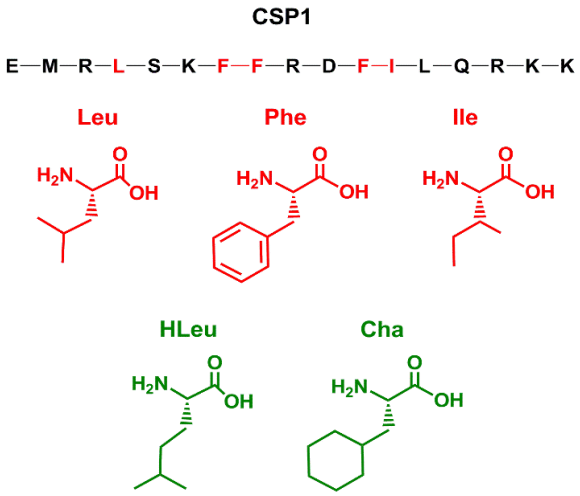


Figure 16: Conservative point mutations of CSP1 performed in this study. The CSP1 sequence is presented using the one letter amino acid code. Residues in red were replaced by the green residues. HLeu, homoleucine; Cha, cyclohexylalanine.

Utilization of Cha substitutions in positions 7, 8 and 11 allowed us to increase the side chain hydrophobic surface area, while minimizing alterations to chain length and polarity. This mutation also removes the aromaticity of the Phe residue, abolishing forces governed by π electrons, such as π-π stacking. In parallel, we employed a series of HLeu substitutions intended to explore the spatial extremes of each position, reaching the upper limit of carbon chain length with a nonproteogenic amino acid substitution. At positions 7, 8, and 11, HLeu substitution similarly provided an opportunity to explore the effects of converting aromatic

side chains to aliphatic residues. Overall, by systematically substituting HLeu and Cha at these five positions (4, 7, 8, 11, and 12), we created a library of analogs that explores the effects of interconverting cyclic and aliphatic residue sidechains, while also assessing the effects of size, hydrophobicity, and aromaticity. The analogs were built using standard solid-phase peptide synthesis (SPPS) protocols on Wang resin,[133] followed by purification using semi-preparative RP-HPLC to >95% purity and their identity confirmed by mass spectrometry (for full details see the Appendix 2).

Nonproteogenic singly substituted CSP1 analogs

To evaluate the ability of the CSP1 analogs to modulate the activity of the pneumococcal ComD receptors (both ComD1 and ComD2), we utilized a β-galactosidase cell-based bacterial reporter assay with the two previously constructed reporter strains, D39pcomx::lacZ and TIGR4pcomx::lacZ.[42] These are two wild type strains capable of producing their native CSPs and carry the *lacZ* gene under the control of the *comX* promoter. Activation of the ComD receptors can therefore be assessed by measuring β-gal activity. All the CSP1 analogs were initially screened for their ability to modulate both ComD receptors at high analog concentration, followed by the determination of their EC_{50}/IC_{50} values through dose-response curves (see the Appendix 2 for initial screening and dose-response curves).

Biological evaluation of the singly substituted CSP1 analogs revealed that substitution of either HLeu or Cha at all but the 11th position resulted in comparable or higher potency against the ComD1 receptor relative to the native CSP1. Each substitution also maintained activity against the ComD2 receptor (see **Table 4**). The data from single, nonproteogenic substitutions at L4 revealed that increased size and hydrophobicity does not lead to a

significant change in activity compared to the native side chain at this position. The EC_{50} values are comparable between the native CSP1, L4Cha, and L4Hleu against both the ComD1 and ComD2 receptors (about 2-fold reduction and 2-fold increase in potency against ComD1 and ComD2, respectively). Regarding the 12^{th} position, substitution of Cha or HLeu for Ile resulted in two potent ComD1 activators, I12Cha and I12HLeu, exhibiting EC_{50} values of 3.1 and 5.1 nM, respectively (**Table 4**). These data support the previous observation that the binding pocket of the 12^{th} residue in the ComD1 receptor is not fully occupied by Ile and thus the CSP1-ComD1 binding interaction can be optimized utilizing bulkier hydrophobic side chain residues at this position.[91] Contrary to the ComD1 receptor, these two mutations were not as well tolerated against the ComD2 receptor, resulting in a reduction in potency compared to CSP1 (EC_{50} values of I12Cha and I12HLeu are >1000 nM and 870 nM, respectively, **Table 4**). This suggests that the introduction of bulky side chains in the ComD2 binding pocket for this position is not as permissible.

Table 4. EC_{50} values of singly substituted CSP1 analogs against the ComD1 and ComD2 receptors[a]

Peptide name	Peptide Sequence	ComD1		ComD2	
		EC_{50} (nM) [b]	95% CI [c]	EC_{50} (nM) [b]	95% CI [c]
CSP1 [d]	EMRLSKFFRDFILQRKK	10	6.3-17	530	500-560
CSP1-L4Cha	EMR(**Cha**)SKFFRDFILQRKK	17	12-23	590	300-1200
CSP1-L4HLeu	EMR(**HLeu**)SKFFRDFILQRKK	11	7.0-17	220	130-360
CSP1-F7Cha	EMRLSK(**Cha**)FRDFILQRKK	1.5	0.95-2.3	780	550-1100
CSP1-F7HLeu	EMRLSK(**HLeu**)FRDFILQRKK	0.82	0.77-0.87	75	35-160
CSP1-F8Cha	EMRLSKF(**Cha**)RDFILQRKK	4.8	2.7-8.3	490	230-1100
CSP1-F8HLeu	EMRLSKF(**HLeu**)RDFILQRKK	3.3	2.0-5.2	520	390-710
CSP1-F11Cha	EMRLSKFFRD(**Cha**)ILQRKK	110	67-200	210	110-420
CSP1-F11HLeu	EMRLSKFFRD(**HLeu**)ILQRKK	67	32-140	540	300-980
CSP1-I12Cha	EMRLSKFFRDF(**Cha**)LQRKK	3.1	2.0-5.0	>1000	--
CSP1-I12HLeu	EMRLSKFFRDF(**HLeu**)LQRKK	5.1	5.0-5.3	870	800-950

[a] See the Experimental details for methods and the Appendix 2 for plots of agonism dose response curves. [b] EC_{50} values were determined by testing peptides over a range of concentrations. [c] 95% confidence interval. [d] Data from ref.[47]

According to data from previous structural studies of CSP1, the positions bearing Phe (7, 8, and 11) are located on one side of the helix and mostly occupy the hydrophobic binding site within the ComD1 receptor.[94] The Phe side chains were assumed to effectively interact with the ComD1 binding pocket through both hydrophobic and aromatic interactions. The results of our singly substituted CSP1 analogs revealed that substitution of Cha or HLeu for Phe at positions 7 and 8 resulted in more potent analogs against ComD1. With the exception of F7Cha, all the resulting analogs at these positions exhibited enhanced or comparable activity against ComD2 compared to CSP1 (**Table 4**). These results suggest that these sites in the ComD1 binding pocket may be larger than originally thought. Specifically, the activation data of single mutations revealed that the F7HLeu derivative is the most potent ComD1 agonist reported to date and a highly potent ComD2 agonist (>12-

fold increase in potency against ComD1, EC_{50} = 0.82 nM; and >7-fold increase in potency against ComD2, EC_{50} = 75 nM; **Table 4**). As observed with F7, the F8 position favors both the Cha and the HLeu substitutions, affording modestly improved ComD1 activators (more than 2-fold increase in potency) and exhibiting similar activity to CSP1 against the ComD2 receptor. This result is consistent with our previous observation, suggesting that the binding pocket for the eighth residue can accommodate elongated hydrophobic side chains.[31] On the contrary, an opposite trend was observed for the 11th residue. An 11-fold (F11Cha) and ~7-fold (F11HLeu) reduction in potency against ComD1 was observed, depicting that this position is either more spatially restricted than the other positions or that the aromaticity of the Phe side-chain residue plays a critical role at this site on the receptor.

Nonproteogenic poly substituted CSP1 analogs

The single mutant library revealed several potent CSP-based pneumococcal QS activators. Therefore, we sought to utilize the recently acquired SAR insight to rationally design more potent QS modulators. To this end, a library of poly-substituted peptides was synthesized to examine potential synergistic effects on ComD activation. These mutations focused on positions 7, 8, and 12, based on the results of the singly substituted analogs (**Table 5**). Biological evaluation revealed that all the doubly substituted peptides possess EC_{50} values against ComD1 at or below the EC_{50} value of the native CSP1 signal. Our results reveal that a combination of two Cha substitutions at positions 7 and 12 yielded the most potent pan-group activator in our library (F7Cha/I12Cha, EC_{50} = 0.97 nM against ComD1 and 70 nM against ComD2; **Table 5**). This dual-modified analog displayed a cumulative effect, as the recorded activation was greater than that of the single replacement analogs (compare the EC_{50} values of F7Cha and I12Cha against the ComD1 and ComD2 receptors in **Table**

4 with the EC_{50} values of F7Cha/I12Cha against the ComD1 and ComD2 receptors in **Table 5**). This result suggests that the increased hydrophobicity and steric bulk in both the 7th and 12th positions, brought upon by the introduction of the Cha residues, allows for optimized interactions in both binding pockets. Likewise, HLeu was incorporated into several doubly substituted peptides. This resulted in another potent pan-group activator formed by combining Cha and HLeu substitutions at the 7th and 12th positions, respectively (F7Cha/I12HLeu), which exhibited a 3- to 4-fold increase in potency against both ComD receptors. While some doubly substituted analogs exhibited increased potency relative to CSP1 against both pneumococcal receptors, most resulted in similar activities to CSP1 against ComD2, supporting our previous observation that there are different binding requirements for the ComD1 and ComD2 receptors (**Table 5**).[91]

Table 5. EC_{50} values of poly substituted CSP1 analogs against the ComD1 and ComD2 receptors[a]

Peptide name	Peptide Sequence	ComD1		ComD2	
		EC_{50} (nM) [b]	95% CI [c]	EC_{50} (nM) [b]	95% CI [c]
CSP1 [d]	EMRLSKFFRDFILQRKK	10	6.3-17	530	500-560
CSP1-F7Cha/F8Cha	EMRLSK**(Cha)** **(Cha)**RDFILQRKK	2.2	1.4-3.5	590	330-1000
CSP1-F7Cha/F8HLeu	EMRLSK**(Cha)** **(HLeu)**RDFILQRKK	2.5	1.5-4.4	550	280-1100
CSP1-F7HLeu/F8Cha	EMRLSK**(HLeu)** **(Cha)**RDFILQRKK	4.8	2.9-8.1	340	190-640
CSP1-F7HLeu/F8HLeu	EMRLSK**(HLeu)** **(HLeu)**RDFILQRKK	7.2	3.4-15	710	530-940
CSP1-F7Cha/I12Cha	EMRLSK**(Cha)**FRDF**(Cha)**LQRKK	0.97	0.44-2.2	70	41-120
CSP1-F7Cha/I12HLeu	EMRLSK**(Cha)**FRDF**(HLeu)**LQRKK	3.0	1.4-6.4	140	68-300
CSP1-F7HLeu/I12Cha	EMRLSK**(HLeu)**FRDF(**Cha**)LQRKK	3.4	2.1-5.4	320	190-540
CSP1-F7HLeu/I12HLeu	EMRLSK**(HLeu)**FRDF**(HLeu)**LQRKK	1.2	0.80-1.8	350	180-680
CSP1-F8Cha/I12Cha	EMRLSKF**(Cha)**RDF**(Cha)**LQRKK	8.4	7.9-9.0	840	730-980
CSP1-F8Cha/I12HLeu	EMRLSKF**(Cha)**RDF**(HLeu)**LQRKK	5.4	3.7-7.8	380	250-600
CSP1-F8HLeu/I12Cha	EMRLSKF**(HLeu)**RDF**(Cha)**LQRKK	10	5.2-21	680	560-840
CSP1-F8HLeu/I12HLeu	EMRLSKF**(HLeu)**RDF**(HLeu)**LQRKK	3.3	2.7-4.0	400	230-720

[a] See the Experimental details for methods and the Appendix 2 for plots of agonism dose response curves. [b] EC_{50} values were determined by testing peptides over a range of concentrations. [c] 95% confidence interval. [d] Data from ref.[47]

E1A substituted CSP1 analogs

Previously, Zhu et al. reported that the replacement of glutamic acid at position 1 with alanine in CSP1 resulted in an analog that exhibits competitive inhibition against ComD1.[42] Having identified several pan-group QS activators in both our single and double mutant libraries, we set out to evaluate whether these pan-group activators could be converted into pan-group QS inhibitors by applying the same key modification. For this analysis, we chose the most potent pan-group activators in both the single and double mutant libraries and incorporated the E1A modification to afford a library of eight E1A-containing analogs. This library included six singly substituted CSP1 analogs containing the Cha or HLeu substitutions at positions 7, 8 or 12, as well as the two most potent doubly substituted pan-group activators (F7Cha/I12Cha and F7Cha/I12HLeu). Biological evaluation revealed that all the resultant analogs can only effectively inhibit the ComD1 receptor (**Table 6** and **Figures S-4 and S-8 in Appendix 2**). The lack of inhibitory activity against the ComD2 receptor highlights the strict and different requirements for receptor inhibition compared to receptor activation. These results are consistent with a previous study demonstrating that direct conversion of pan-group activators into pan-group inhibitors requires further modification rather than just a single substitution.[125]

Table 6. IC_{50} values of E1A substituted CSP1 analogs against the ComD1 receptor[a]

Peptide name	Peptide Sequence	ComD1	
		IC_{50} (nM) [b]	95% CI [c]
CSP1-E1A[d]	**A**MRLSKFFRDFILQRKK	86	51-150
CSP1-E1A/F7Cha	**A**MRLSK**(Cha)**FRDFILQRKK	36	16-79
CSP1-E1A/F7HLeu	**A**MRLSK**(HLeu)**FRDFILQRKK	72	42-120
CSP1-E1A/F8Cha	**A**MRLSKF**(Cha)**RDFILQRKK	210	95-470
CSP1-E1A/F8HLeu	**A**MRLSKF**(HLeu)**RDFILQRKK	340	200-580
CSP1-E1A/I12Cha	**A**MRLSKFFRDF**(Cha)**LQRKK	590	380-910
CSP1-E1A/I12HLeu	**A**MRLSKFFRDF**(HLeu)**LQRKK	41	29-56
CSP1-E1A/F7Cha/I12Cha	**A**MRLSK**(Cha)**FRDF(**Cha**)LQRKK	72	39-140
CSP1-E1A/F7Cha/I12HLeu	**A**MRLSK**(Cha)**FRDF(**HLeu**)LQRKK	57	46-72

[a] See the Experimental details for methods and the Appendix 2 for plots of antagonism dose response curves. [b] IC_{50} values were determined by testing peptides over a range of concentrations. [c] 95% confidence interval. [d] Data from ref.[47]

Comparing the bioactivities of the resulting single-substitution inhibitors reveals that there is no correlation between EC_{50} and IC_{50} values. However, several novel ComD1 inhibitors with low nanomolar potency were discovered. E1A/F7Cha (IC_{50} = 36 nM), E1A/I12HLeu (IC_{50} = 41 nM), and E1A/F7HLeu (IC_{50} = 72 nM) displayed higher inhibitory potency against the ComD1 receptor compared to the E1A substitution alone (IC_{50} = 86 nM). The remaining single-substituted analogs displayed inhibitory activity but were less potent (IC_{50} > 86 nM) (**Table 6**). Although single substitution of HLeu or Cha at the 12th position yielded two potent ComD1 activators (**Table 4)**, only the I12HLeu combined with the E1A modification generated a more potent ComD1 inhibitor. Conversely, the I12Cha substitution combined with the E1A substitution yielded a ComD1 inhibitor with ~7-fold reduced activity, exhibiting the sensitivity of the 12th position for ComD1 inhibition. This

sensitivity is further underscored by data garnered after introducing the E1A substitution into our doubly mutated activators, F7Cha/I12Cha and F7Cha/I12HLeu. E1A/F7Cha/I12Cha displayed an 8-fold increase in potency relative to E1A/I12Cha, but was only half as potent as E1A/F7Cha. Similarly, E1A/F7Cha/I12HLeu was a less potent inhibitor than its doubly substituted precursors, E1A/F7Cha and E1A/I12HLeu. Thus, select substitutions of a single bulky, nonproteogenic amino acid in combination with the E1A modification proved to be more effective in harnessing inhibitory activity than combining multiple substitutions.

Pharmacological evaluation of lead CSP1 analogs

Our next goal was to evaluate the impact the non-proteogenic substitutions have on key pharmacological properties, i.e., metabolic stability and toxicity. To this end, we first evaluated the stability of CSP1, CSP1-F7Cha/I12Cha (lead activator), and CSP1-E1A/F7Cha (lead inhibitor) towards trypsin and chymotrypsin degradation. Our results indicate that CSP1 is highly susceptible to enzymatic degradation, exhibiting half-lives of 30 min and 1 h against trypsin and chymotrypsin, respectively (**Figure 17**). The introduction of Cha residues at either position 7 or positions 7 and 12 resulted in improved metabolic stability against both proteases. CSP1-E1A/F7Cha exhibited a half-life of 4 h against both enzymes, whereas CSP1-F7Cha/I12Cha exhibited half-lives of 3 h and 6 h against trypsin and chymotrypsin, respectively (**Figure 17**). Further analysis of the degradation products revealed that the introduction of the Cha residues did not alter the trypsin cleavage sites (R3, K6, R9, and R15), but it did slow the cleavage kinetics (**Figures S-9–S-11 in Appendix 2**). As for chymotrypsin, the analysis of the degradation products revealed that, for CSP1, chymotrypsin can cleave the peptide in all three aromatic residues

(F7, F8, and F11; **Figure S-12 in Appendix 2**). Furthermore, the Cha residues in both the 7 and 12 positions could be recognized and cleaved by chymotrypsin (Figures **S-13–S-14 in Appendix 2**). Interestingly, introduction of Cha at the 7th position resulted in elimination of cleavage at F8, whereas introduction of Cha at the 12th position resulted in elimination of cleavage at F11. Overall, the introduction of Cha residues resulted in altered cleavage sites and improved peptide stability.

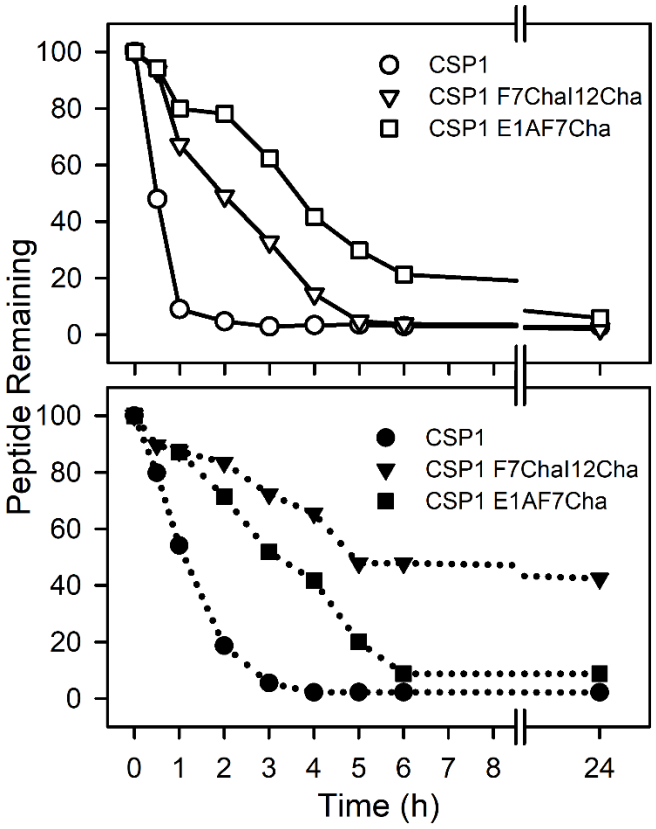


Figure 17: Metabolic Stability of CSP1 analogs. Peptides were incubated with trypsin (top panel) or chymotrypsin (bottom panel) and peptide degradation was monitored using analytical RP-HPLC. In the trypsin assay, CSP1 degraded first (half-life of 30 min), CSP1-F7ChaI12Cha degraded second (half-life of 3 h), and CSP1-E1AF7Cha was most stable (half-life of 4 h). In the chymotrypsin assay, CSP1 again degraded first (half-life of 1 h), however in this case CSP1-E1AF7Cha degraded second (half-life of 4 h), whereas CSP1-F7ChaI12Cha was most stable (half-life of 6 h).

We next set out to evaluate the toxicity of CSP1 and the lead analogs towards mammalian cells. To this end, we performed a hemolysis assay of red blood cells (RBCs). Our results indicate that all three peptides are nontoxic, resulting in only minimal hemolysis, similar

to the negative control, DMSO (**Figure 18**). Overall, our results highlight the potential of introducing non-proteogenic amino acids to the CSP1 scaffold as a means to improve the pharmacological properties of the peptide without eliciting toxicity.

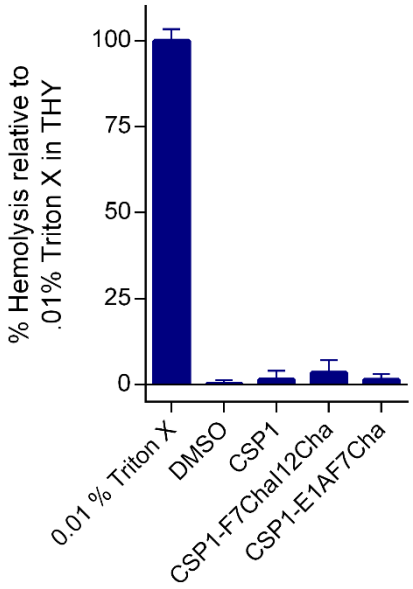


Figure 18: Hemolytic activity of CSP-derived QS modulators on defibrinated RBCs. The CSP analogs exhibit no toxicity against RBCs.

Conclusion

In this work, we set out to stabilize the CSP-ComD binding interactions through the development of single and multiple mutant CSP1 libraries consisting of bulkier, hydrophobic, and nonproteogenic substituents. Our analysis indicates that Cha and HLeu substitutions are less tolerated at positions 4 and 11 while being beneficial at positions 7, 8, and 12. These results suggest that the Cha and HLeu substitutions at 7^{th} ,8^{th}, and 12^{th} positions provided a more optimal residue size and structure than the original proteogenic side chains. We hypothesize that the cyclic amino acids in the hydrophobic binding face are more open to Cha substitutions as Cha has a larger hydrophobicity index and steric

value than Phe.[148] These properties are not available with other natural amino acids. Additionally, Cha may be adopting conformations that Phe cannot as a planar, aromatic side chain in order to optimize hydrophobic interactions with the binding pocket of ComD1. For these reasons, our work highlights the value of using Cha in routine screens for improving peptide-protein interactions and assessing SAR in this QS system and beyond.

Our rationally designed CSP1-based point and multiple mutant analogs yielded some of the most potent agonists of pneumococcal QS to date. The improved binding resulting from Cha and HLeu substitutions at positions 7, 8, and 12 indicates that there is unoccupied space in the ComD1 hydrophobic binding pocket that can be optimized. The success of some of our poly-substituted mutants demonstrate that optimized residues can be combined to further occupy this space. For example, F7Cha by itself had an EC_{50} value of 1.5 nM but, when combined with the I12Cha substitution, the EC_{50} value was reduced to 0.97 nM.

To test for pan-group activity, our CSP1 derivatives were screened against ComD2. In general, the derivatives were much less potent against ComD2. Importantly, the two most potent pan-group activators found in our study are F7HLeu and F7Cha/I12Cha, both with EC_{50} values below 1 nM against ComD1 and EC_{50} values around 70 nM against ComD2. Even after accounting for 95% CI overlap, these two substituted analogs exhibited more than 7 to 12-fold increase in potency against both ComD receptors. However, there does not seem to be a strong correlation between the ComD1 and ComD2 activation data. For example, F8HLeu is a stronger ComD1 activator than F11Cha, yet F11Cha is a stronger ComD2 activator than F8HLeu. It is important to note that F7Cha and I12Cha by

themselves were ineffective as ComD2 activators but in combination make the best ComD2 activator identified in this study. Previously, it was determined that CSP1 and CSP2 form two distinctive hydrophobic patches that are optimal for ComD1 and ComD2 binding, respectively, and that a hybrid hydrophobic patch can be achieved by a single peptide.[57] Therefore, it could be that the combined substitution of Cha at F7 and I12 result in a peptide that exhibits such a hybrid hydrophobic patch and binds both receptors effectively.

All of the single and double mutants that were resynthesized with the E1A substitution displayed some degree of inhibitory activity against ComD1. E1A/F7Cha and E1A/I12HLeu proved to be the most successful combinations, displaying IC_{50} values 2 - 3 times more potent than E1A alone. Our analysis revealed that some of the peptides that showcased agonistic activity were less effective as inhibitors. For instance, E1A/I12HLeu made a better inhibitor than E1A/I12Cha and E1A/F7HLeu, even though it was less potent than both as an activator. The lack of ComD2 inhibition activity for all the E1A-based analogs further supports the previous observation completed by our lab highlighting the strict requirements for pan-group activation and inhibition, as opposed to the simple E1A modification that was found to be sufficient in converting ComD1 or ComD2 activators into competitive inhibitors.

Finally, pharmacological evaluation of the lead analogs, CSP1-F7Cha/I12Cha (lead activator) and CSP1-E1A/F7Cha (lead inhibitor), revealed that the incorporation of the Cha residues resulted in analogs that exhibit superior metabolic stability while remaining nontoxic against mammalian cells.

In conclusion, our systematic study of the hydrophobic binding surface of CSP1 revealed that larger, nonproteogenic amino acids produced improved binding relative to the proteogenic amino acids present in the native sequence. We highlighted the importance of pushing the steric limit in peptide-protein SAR and discovered the most potent ComD1 agonist and several potent *S. pneumoniae* QS inhibitors as a result.

Experimental Section

Chemical reagents and instrumentation: All chemical reagents and solvents were purchased from Sigma-Aldrich or Chem-Impex and used without further purification. Water (18 MΩ) was purified using a Thermo Scientific Smart2Pure Pro UV/UF 16 LPH water purification system. Solid-phase resins were purchased from Chem-Impex or P3 Biosystems.

Reversed-phase high-performance liquid chromatography (RP-HPLC) was performed using two Shimadzu systems each equipped with a CBM-20A communications bus module, two LC-20AT pumps, an SIL-20A auto sampler, an SPD-20A UV/VIS detector, a CTO-20A column oven, one with an FRC-10A fraction collector and one without. All RP-HPLC solvents (18 MΩ water and HPLC-grade acetonitrile (ACN)) contained 0.1% trifluoroacetic acid (TFA). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) data were obtained on either a Bruker Autoflex or Bruker Microflex spectrometer equipped with a 60 Hz nitrogen laser and a reflectron. In positive ion mode, the acceleration voltage on Ion Source 1 was 19.01 kV. Exact mass (EM) data

were obtained on an Agilent Technologies 6230 TOF LC/MS spectrometer. The samples were sprayed with a capillary voltage of 3500 V and the electrospray ionization (ESI) source parameters were as follows: gas temperature of 325 °C at a drying gas flow rate of 8 L/min at a pressure of 35 psi.

Solid Phase Peptide Synthesis: All the CSP1 analogs were synthesized using standard Fluorenyl methoxycarbonyl (Fmoc)-based solid-phase peptide synthesis (SPPS) procedures on preloaded Fmoc-L-Lys(Boc)-OH 4-benzyloxybenzyl alcohol (Wang) resin (0.59 mmol/g) by using a Discover microwave synthesizer or Liberty1 automated peptide synthesizer (CEM Corporation). For peptides synthesized on the automated synthesizer, the resin (0.1 g) was first swelled by suspension in *N*,*N*-dimethylformamide (DMF) for 15 min at room temperature and then drained. Fmoc-protecting group removal was accomplished with treatment of the resin by 5 mL of 20% piperidine in DMF (90 sec, 90 °C) followed by another 5 mL of 20% piperidine in DMF (90 sec, 90 °C). The resin was then washed with DMF (3 x 5 mL) after each deprotection cycle. To couple each amino acid, Fmoc-protected amino acids (5 equiv. relative to the overall loading of the resin) were dissolved in DMF (5 mL) and mixed with 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU; 5 equiv.) and diisopropylethylamine (DIPEA; 5 equiv.). All amino acids were coupled for 5 min (30 W, 75 °C). After each coupling step, the resin was drained and washed with DMF (2 × 5 mL). This process was repeated until the desired peptide sequence was obtained. The same swelling and deprotection protocols were followed for peptides synthesized on the manual microwave synthesizer. However, 0.2 g of Wang resin was used and couplings were completed using

diisopropyl carbodiimide (DIC) and Oxyma Pure with a 3.6:3:3 ratio of DIC:Oxyma:AA in DMF for a final DIC concentration of 0.2 M. Each coupling was run at 75 °C for 8 min (50 W).

Peptide Cleavage from Solid Support: Following coupling of the final residue, the resin was washed three times with DCM (2 mL) with manual shaking for 1 min. The resin was then washed with diethyl ether (2 mL) and dried under nitrogen stream for 3 min before transferring it into a 15 mL falcon tube. A 3 mL solution of 2.5% 18 MΩ water and 2.5% triisopropylsilane (TIPS) in 95% trifluoroacetic acid (TFA) for every 0.1 g of resin was added and the tube was shaken for 3 h at 200 rpm. Following completion of the cleavage reaction, the resin was filtered through a cotton ball, or a fritted syringe and the filtrate was transferred into a new 50 mL falcon tube. A cooled solution of diethyl ether:hexane (1:1, 45 mL, -20 °C) was added to the tube, and the tube was kept in a freezer at -20 °C for 10 min in order to precipitate the crude peptide. The pellet of the crude peptide was obtained by centrifugation of the 50 mL tube in a Beckman Coulter Allegra 6 centrifuge equipped with a GH3.8 rotor at 3000 RPM for 5 min. The supernatant was poured off and the solid peptide product was re-dissolved in 10 mL acetonitrile (ACN):water (1:1) and lyophilized for a minimum of 24 h before HPLC purification.

Peptide purification by HPLC: Crude peptides were purified using RP-HPLC. The crude peptide was dissolved in ACN:H_2O (1:4; volume of ACN in water depends on the solubility of the peptide) and purified in 4 mL portions on either a Phenomenex Luna 5 μm C18 semi-preparative column (10 mm × 150 mm, 100 Å) or a Phenomenex Kinetex 5 μm C18 semi-preparative column (10 mm × 250 mm, 110 Å) with a flow rate of 5 mL/min; mobile phase

A = 18 MΩ water + 0.1 % TFA and mobile phase B = ACN + 0.1 % TFA. The collected fraction was lyophilized overnight and dissolved again in 5 mL ACN:H_2O (1:4) for a secondary purification run. Preparative HPLC methods were used for the crude purification using a linear gradient (first prep 5% B → 45% B over 40 min and second prep 25% B → 38% B over 45 min). Fraction purity was determined through analysis on either a Phenomenex Luna 5 μm analytical C18 column (4.6 mm × 150 mm, 100 Å) or a Phenomenex Kinetex 5 μm analytical C18 column (4.6 mm × 250 mm, 110 Å) using a linear gradient (5% B → 95% B over 22 min or 27 min, respectively). Purities were determined by integration of peaks with UV detection at 220 nm. Only peptide fractions that were purified to homogeneity (>95%) were used for the biological assays. Following purification, peptides were frozen using a dry ice-acetone bath, and then lyophilized for a minimum of 24 h. Before the final masses and yields of purified peptides were determined, peptides were dissolved in 25% acetic acid in up to 1:1 ACN:water to allow removal of any residual TFA. The solution was then frozen and lyophilized for at least 24 h before the yield of the peptide was determined.

Peptide Verification with Mass Spectrometry: During primary and secondary purification of crude peptide, peaks were verified to contain the desired peptide mass by MALDI-TOF MS. Samples were prepared using 1 μL α-Cyano-4-hydroxycinammic acid (10 μg) in 1:1 H_2O:ACN as a matrix and 1 μL of the desired peptide fraction. For the final verification of the peptides, a high resolution ESI-TOF MS **(Tables S1 and S2 in Appendix 2)** was used to verify the exact masses of the peptides. The observed mass-to-charge (*m*/*z*) ratio of the peptide was compared to the expected *m*/*z* ratio for each peptide.

Biological Reagents and Strain Information: All standard biological reagents were purchased from Sigma-Aldrich. Donor horse serum (defibrinated) was purchased from Sigma-Aldrich and stored at -20 °C until use in pneumococcus bacterial culture. To examine the ability of the synthesized CSP1 analogs to modulate the ComD receptors, and thus, the QS circuit in *S. pneumoniae*, β-galactosidase assays were performed using D39pcomX::lacZ (group I) and TIGR4pcomX::lacZ (group II) reporter strains.[42]

Bacterial Growth Conditions: Freezer stocks of individual pneumococcal strains, D39pcomX::lacZ and TIGR4pcomX::lacZ, were created from 1.5 mL aliquots of overnight cultures (0.2 OD_{600nm}) in Todd-Hewitt broth supplemented with 0.5% yeast extract (THY) and 0.5 mL glycerol, and stored at -80 °C. For the β-galactosidase experiments, bacteria from the freezer stocks were streaked into a THY agar plate containing 5% horse serum and chloramphenicol at a final concentration of 4 μg/mL. The plate was incubated for 8-9 h in 37 °C with 5% CO_2. An isolated fresh single colony was picked into sterilized cultural tube containing 5 mL of THY broth supplemented with a final concentration of 4 μg/mL chloramphenicol, and the culture was incubated in a CO_2 incubator overnight (15 h). Overnight cultures were then diluted (1:50 for D39pcomX::lacZ; 1:10 for TIGR4pcomX::lacZ) with THY and the resulting solution was incubated in a CO_2 incubator for 3-4 h, until the bacteria reached early exponential stage (OD_{600} values of 0.30–0.35 for D39pcomX::lacZ and 0.20–0.25 for TIGR4pcomX::lacZ) as determined by using a plate reader.

β-Galactosidase Activation Assays: The ability of synthetic CSP1 analogs to activate the expression of *S. pneumoniae comX* was determined using indicated reporter strains grown

in THY (pH 7.3). An initial activation screening was performed at high concentration (10 μM) for all CSP1 analogs. Two μL of 1 mM solution of CSP1 analogs in dimethyl sulfoxide (DMSO) were added in triplicate to a clear 96-well microtiter plate. Two μL of 20 μM solution of native CSP1 (200 nM final concentration) were added in triplicate and served as the positive control for the *S. pneumoniae* group I strain (D39pcomX::lacZ), while two μL of 100 μM solution of CSP2 (1000 nM final concentration) were added as the positive control for the *S. pneumoniae* group II strain (TIGR4pcomX::lacZ). These concentrations were chosen to afford full activation of the QS circuit, as determined from the dose-dependent curves created for the native *S. pneumoniae* CSPs.[47] Two μL of DMSO were added in triplicate and served as the negative control. Then, 198 μL of pneumococcal cultures were added to each well, the plate was incubated at 37 ºC for 30 min, and the absorbance at 600 nm was measured. In order to measure the β-galactosidase activity in the pneumococcal culture, the cells were then lysed by incubating the culture for 30 min at 37 °C with 20 μL 0.1% Triton X-100 in water. Then, 100 μL of Z-buffer solution (60.2 mM Na_2HPO_4, 45.8 mM NaH_2PO_4, 10 mM KCl, and 1.0 mM $MgSO_4$ in 18 MΩ H_2O; pH was adjusted to 7.0 and sterilized before use) containing 2-nitrophenyl-β-$_D$-galactopyranoside (ONPG) at a final concentration of 0.4 mg/mL was added in a new plate, followed by 100 μL of lysate, and the plate was incubated for 3 h at 37 °C. The reaction was quenched by adding 50 μL of 1 M sodium carbonate solution, and the OD_{420nm} and OD_{550nm} were measured using a plate reader. The results were reported as percent activation, which is the ratio between the Miller units of the analog and that of the positive control. For calculation of Miller units, please see data analysis below. Analogs that exhibited high activity (>75% activation compared to the native CSP) in the initial screening were further evaluated using

a dose-dependent assay in which peptide stock solutions were diluted with DMSO in serial dilutions (either 1:2, 1:3, or 1:5) and assayed as described above. The EC_{50} (the concentration of a drug that gives half maximal response) value was then determined through fitting using nonlinear regression with GraphPad Prism 5.

β-Galactosidase Inhibition Assays: Analogs that exhibited low competence activation in the initial screening were evaluated for competitive inhibition. The ability of synthesized CSP1 analogs to inhibit the expression of *comX* by outcompeting native pneumococcal CSPs (CSP1 or CSP2) for the receptor binding site was evaluated using the same assay conditions as described above, except that in this case native CSP was added to every well at a set concentration (2 μL, 50 nM final concentration of CSP1 for group I; 250 nM final concentration of CSP2 for group II) that was chosen to afford full activation of the QS circuit, as determined from the dose-dependent curves created for the native pneumococcal CSPs. Two μL of native CSP (5 μM solution of CSP1 for group I; 25 μM solution of CSP2 for group II) and 2 μL of DMSO were added to the same well in triplicate and served as the positive control. Four μL of DMSO were added in triplicate and served as the negative control. Then, 196 μL of bacterial cultures were added to each well and the plate was incubated at 37 ºC for 30 min. The procedure for lysis, incubation with ONPG and all the measurements were as described in the β-galactosidase activation assay. Analogs that exhibited significant competitive inhibition in the initial screening were further evaluated using a dose-dependent assay where a series of dilutions of the CSP1 analogs was prepared similarly to those made for the EC_{50} assay, and the IC_{50} (the concentration of an inhibitor

where the response or binding is reduced by half) was then determined by using GraphPad Prism 5.

Analysis of Activation/Inhibition Data: Miller units were calculated using the following formula:

$$Miller\ Unit = 1000 \times \frac{[Abs_{420} - (1.75 \times Abs_{550})]}{(t \times v \times Abs_{600})}$$

Abs_{420} is the absorbance of o-nitrophenol (ONP). Abs_{550} is the scatter from cell debris, which, when multiplied by 1.75 approximates the scatter observed at 420 nm. *t* is the duration of incubation with ONPG in minutes, *v* is volume of lysate in milliliters, and Abs_{600} reflects cell density.

Metabolic Stability: Enzymatic stability studies of CSP1 analogs was carried out in aqueous PBS solution (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, and 1.8 mM KH_2PO_4; pH 7.4). Peptide stocks (1 mM) were made in 18 mΩ H_2O and final working concentration was 0.33 mM in PBS. Protease (Trypsin or Chymotrypsin) stock solution (25 μg x mL^{-1}; diluted from a 2.5 mg x mL^{-1} solution) was made in PBS solution. Protease solution was added to the peptide solution to afford a final concentration of 0.05 μg x mL^{-1} and then the solution was incubated at 37 °C with shaking (200 rpm) for 24 h. Aliquots (100 μL) were taken at 0, 0.5, 1, 2, 3, 4, 5, 6, and 24 h time points and mixed with 20 μL acetonitrile to stop the enzymatic activity. Then, 100 μL were injected via an autosampler and analyzed immediately for peptide degradation by analytical RP-HPLC. During the

analytical RP-HPLC runs, the degradation products were manually collected and analyzed by MALDI-TOF MS. The digested peptide fractions at the end of the experiment (24 h time point) were also analyzed on a high-resolution ESI-TOF LC-MS to elucidate the peptide degradation pattern and confirm the identity of the peptide fragments.

Hemolysis Assay: The toxicity of synthetic CSP analogs was evaluated through hemolysis of red blood cells (RBCs). The hemolysis assay was performed as previously described with minor modifications.[149] Briefly, 2 μL of a 1 mM CSP analog stock solution in DMSO were plated in triplicate in a clear bottom 96-well microtiter plate. Then, 198 μL of fresh THY media was added to each well. A positive control was prepared by adding 2 μL of a 1% Triton X solution to 198 μL THY media, and a negative control was prepared by adding 2 μL DMSO to 198 μL fresh THY media. One mL of defibrinated rabbit RBCs were centrifuged down, then the supernatants containing plasma and pre-lysed RBCs were pipetted out and the pelleted RBCs were resuspended in approximately 1 mL PBS. The process was repeated until the pre-lysed RBCs were completely removed (two to three washes, until the supernatant was clear). The washed RBCs were resuspended in PBS to a total volume of 1 mL, then 15 μL aliquots were added to each sample in the 96-well plate, and the plate was incubated for 30 min at 37 °C. Following incubation, the 96-well plate was centrifuged at 2000 rpm for 5 min at 4 °C and 25 μL of supernatants were transferred to a new plate containing 175 μL of 18 mΩ H_2O, and the absorbance at 541 nm was recorded. Three independent trials were performed, and data is presented as the percent hemolysis relative to the .01% Triton X positive control.

Corresponding Authors

ytalgan@unr.edu, bertuccm@layette.edu

Supporting Information (Appendix 2)

HPLC traces and MS data, initial screening results, dose response curves for CSP1 analogs, and additional metabolic stability figures. This information is available online.

Notes

The authors declare no competing financial interest.

Acknowledgements

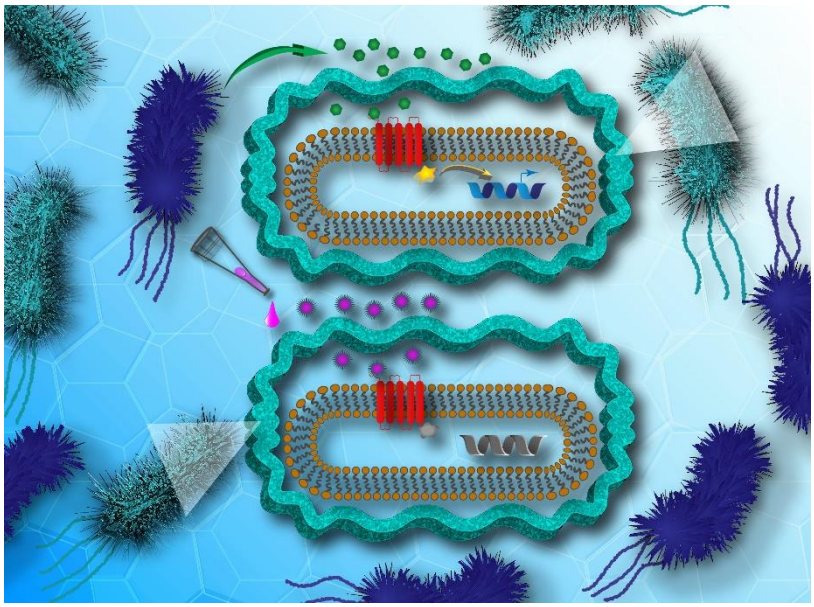
This work was supported by a grant from the National Science Foundation (CHE-1808370). The *S. pneumoniae* D39pcomX::lacZ and TIGR4pcomX::lacZ reporter strains were a generous gift from G. W. Lau (University of Illinois at Urbana−Champaign).

References

1. C. M. Waters and B. L. Bassler, *Annu. Rev. Cell Dev. Biol.*, 2005, **21**, 319-346.
2. D. N. McBrayer, C. D. Cameron and Y. Tal-Gan, *Org Biomol Chem*, 2020, **18**, 7273-7290.
3. B. L. Bassler and R. Losick, *Cell*, 2006, **125**, 237-246.
4. M. R. Parsek and E. Greenberg, *Trends in microbiology*, 2005, **13**, 27-33.
5. L. S. Håvarstein, G. Coomaraswamy and D. A. Morrison, *Proceedings of the National Academy of Sciences*, 1995, **92**, 11140-11144.
6. D. L. Hava and A. Camilli, *Molecular microbiology*, 2002, **45**, 1389-1406.
7. E. A. Meighen, *Microbiological reviews*, 1991, **55**, 123-142.
8. L. Zhu and G. W. Lau, *PLoS pathogens*, 2011, **7**, e1002241.
9. B. Koirala, J. Lin, G. W. Lau and Y. Tal-Gan, *ChemBioChem*, 2018, **19**, 2380-2386.

10. Y. Yang, J. Lin, A. Harrington, G. Cornilescu, G. W. Lau and Y. Tal-Gan, *Proceedings of the National Academy of Sciences*, 2020, **117**, 1689-1699.
11. C. Duan, L. Zhu, Y. Xu and G. W. Lau, 2012.
12. M. Lella and Y. Tal-Gan, *Peptide Science*, 2021, e24222.
13. T. A. Milly and Y. Tal-Gan, *RSC chemical biology*, 2020, **1**, 60-67.
14. J. R. Brannon and M. Hadjifrangiskou, *Drug design, development and therapy*, 2016, **10**, 1795.
15. B. LaSarre and M. J. Federle, *Microbiology and molecular biology reviews*, 2013, **77**, 73-111.
16. S. Mehr and N. Wood, *Paediatric respiratory reviews*, 2012, **13**, 258-264.
17. K. L. O'brien, L. J. Wolfson, J. P. Watt, E. Henkle, M. Deloria-Knoll, N. McCall, E. Lee, K. Mulholland, O. S. Levine and T. Cherian, *The Lancet*, 2009, **374**, 893-902.
18. N. J. Croucher, S. R. Harris, C. Fraser, M. A. Quail, J. Burton, M. Van Der Linden, L. McGee, A. Von Gottberg, J. H. Song and K. S. Ko, *science*, 2011, **331**, 430-434.
19. J. Cornick and S. Bentley, *Microbes and infection*, 2012, **14**, 573-583.
20. W. P. Hanage, C. Fraser, J. Tang, T. R. Connor and J. Corander, *Science*, 2009, **324**, 1454-1457.
21. L. S. Håvarstein, R. Hakenbeck and P. Gaustad, *Journal of bacteriology*, 1997, **179**, 6589-6594.
22. S. Shekhar, R. Khan, D. M. Ferreira, E. Mitsi, E. German, G. H. Rørvik, D. Berild, K. Schenck, K. Kwon and F. Petersen, *Frontiers in immunology*, 2018, **9**, 747.
23. E. Pestova, L. Håvarstein and D. Morrison, *Molecular microbiology*, 1996, **21**, 853-862.
24. F. M. Hui, L. Zhou and D. A. Morrison, *Gene*, 1995, **153**, 25-31.
25. O. Ween, P. Gaustad and L. S. Håvarstein, *Molecular microbiology*, 1999, **33**, 817-827.
26. M. S. Lee and D. A. Morrison, *Journal of Bacteriology*, 1999, **181**, 5004-5016.
27. A. Dagkessamanskaia, M. Moscoso, V. Hénard, S. Guiral, K. Overweg, M. Reuter, B. Martin, J. Wells and J. P. Claverys, *Molecular microbiology*, 2004, **51**, 1071-1086.
28. A. M. Whatmore, V. A. Barcus and C. G. Dowson, *Journal of bacteriology*, 1999, **181**, 3144-3154.
29. Y. Yang, B. Koirala, L. A. Sanchez, N. R. Phillips, S. R. Hamry and Y. Tal-Gan, *ACS chemical biology*, 2017, **12**, 1141-1151.
30. Y. Yang, G. Cornilescu and Y. Tal-Gan, *Biochemistry*, 2018, **57**, 5359-5369.
31. O. Johnsborg, P. E. Kristiansen, T. Blomqvist and L. S. Håvarstein, *Journal of bacteriology*, 2006, **188**, 1744-1749.
32. B. Koirala, R. A. Hillman, E. K. Tiwold, M. A. Bertucci and Y. Tal-Gan, *Beilstein journal of organic chemistry*, 2018, **14**, 1769-1777.
33. T. A. Milly, E. R. Engler, K. S. Chichura, A. R. Buttner, B. Koirala, Y. Tal-Gan and M. A. Bertucci, *ChemBioChem*, 2021, **22**, 1940-1947.
34. D. Rooklin, A. E. Modell, H. Li, V. Berdan, P. S. Arora and Y. Zhang, *Journal of the American Chemical Society*, 2017, **139**, 15560-15563.

35. W. Chan and P. White, *Fmoc solid phase peptide synthesis: a practical approach*, OUP Oxford, 1999.
36. B. Koirala and Y. Tal-Gan, *ChemBioChem*, 2020, **21**, 340-345.
37. J. L. Fauchere, M. Charton, L. B. Kier, A. Verloop and V. Pliska, *Int J Pept Protein Res*, 1988, **32**, 269-278.
38. R. W. Mull and Y. Tal-Gan, *ACS Chem Biol*, 2021, **16**, 2834-2844.



Chapter 4: Biological Evaluation of Native Streptococcal Competence Stimulating Peptides Reveals Potential Crosstalk between *Streptococcus mitis* and *Streptococcus pneumoniae* and a New Scaffold for the Development of *S. pneumoniae* Quorum Sensing Modulators

[a]Reprinted with permission from Milly, T. A.; Tal-Gan, Y. Biological evaluation of native streptococcal competence stimulating peptides reveals potential crosstalk between *Streptococcus mitis* and *Streptococcus pneumoniae* and a new scaffold for the development of *S. pneumoniae* quorum sensing modulators. *RSC Chem. Biol.* **2020**, 1, 60-67.

Biological Evaluation of Native Streptococcal Competence Stimulating Peptides Reveal Potential Crosstalk Between *Streptococcus mitis* and *Streptococcus pneumoniae* and a New Scaffold for the Development of *S. pneumoniae* Quorum Sensing Modulators

Tahmina Ahmed Milly and Yftah Tal-Gan

Department of Chemistry, University of Nevada, Reno, 1664 N. Virginia Street, Reno, Nevada 89557, United States

ABSTRACT

Streptococcus pneumoniae, an opportunistic human pathogen, acquires genes from its neighboring species of the mitis group of streptococci that confer antibiotic resistances and allow it to produce diverse virulence factors. Most species of the mitis group are naturally competent, and they utilize the competence stimulating peptide (CSP) and the CSP-dependent competence regulon, a conserved quorum sensing (QS) circuit, to regulate their competence behavior. The dependence of the mitis group on this communication pathway makes QS a potential target to control their behavior. In this work, we sought to evaluate the impact of native pheromones of the adjacent species of *S. pneumoniae* to modulate the activity of the *S. pneumoniae* competence regulon. Our results revealed the potential role of *S. mitis* as a modulator of QS in *S. pneumoniae*. Most importantly, our analysis also revealed that by using the native pheromone of *S. mitis* as a template, highly potent pan-group agonists and antagonists of the pneumococcal competence regulon could be developed. The newly developed QS modulators may have therapeutic utility in treating pneumococcus infections.

Streptococcus pneumoniae (pneumococcus), a member of the mitis group of streptococci, is an opportunistic human pathogen that is a major cause of a variety of diseases, including deadly pneumonia, bacteremia, meningitis, sepsis and otitis media.[150] It has been estimated that more than 22,000 deaths and 445,000 hospitalizations, with $3.5 billion direct medical costs, occur annually in the United States as a result of pneumococcus infections.[104] Furthermore, the increasing rate of resistance development against several antibiotics, including vancomycin, quinolones, macrolides and beta lactams by recombinogenic pneumococcal strains necessitates the development of alternative strategies to attenuate pneumococcus infections.[105,151] Pneumococcus utilizes the competence regulon, a conserved quorum sensing (QS) circuit (**Figure 19**), to acquire genetic material from the neighboring environment, including genes that confer antibiotic resistance, as well as regulate a wide range of functions related to pathogenicity such as virulence factor production and biofilm formation.[152-156] Thus, intercepting this QS circuitry could be utilized as a potential non-antibiotic therapeutic strategy to control pneumococcus infections without affecting the survival of the bacteria, thus minimizing the selective pressure for resistance development.[10,157-162]

The competence regulon in pneumococcus relies on the production, secretion and response to a 17-amino-acid competence stimulating peptide (CSP, **Figure 19**) signal.[152] Different strains of pneumococcus can produce different CSP signals and the two major forms of CSP (CSP1 and CSP2), share 50% homology with each other.[163] Each CSP signal is associated with a cognate transmembrane histidine kinase receptor, termed ComD (ComD1 and ComD2, respectively).[164] The competence regulon QS circuitry (**Figure 19**) involves the secretion of mature CSP signal to the extracellular environment with the help of a

dedicated ABC transporter, ComAB.[152] As the bacteria grow, the concentration of CSP increases, and once a threshold concentration is attained, the CSP pheromone can bind to and activate the ComD receptor.[155] ComD activation prompts the phosphorylation of the cytoplasmic response regulator, ComE, which then acts as a transcription factor to upregulate the transcription of the QS circuitry genes (*comABCDE*) as well as the *comX* gene, the master regulator of the QS circuitry that controls the different QS-regulated phenotypes.[153-155]

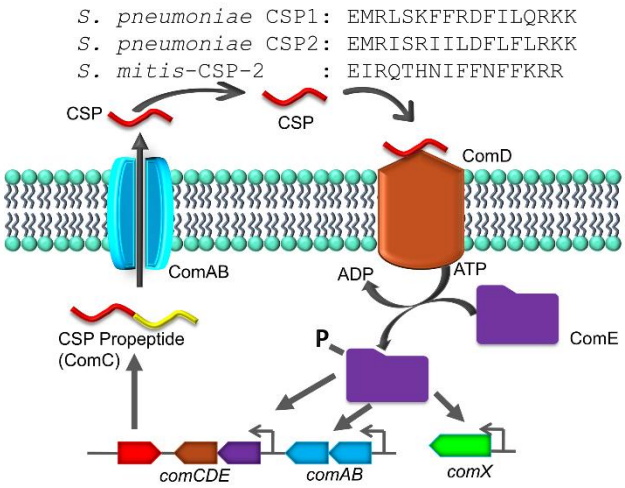


Figure 19. CSP-mediated QS circuit of the mitis group of streptococci. The CSP propeptide, ComC, is processed and secreted by the ComAB transporter as the mature peptide, CSP, that interacts with the cognate receptor, ComD. Activated ComD phosphorylates ComE, which then autoactivates the competence QS circuit. The sequences for *S. pneumoniae* CSP1, *S. pneumoniae* CSP2 and *S. mitis*-CSP-2 are shown at the top.

Similar to *S. pneumoniae*, other species of the mitis and the anginosus streptococcal groups are naturally competent. Regardless of the specific CSP pheromone involved, these species utilize a similar pheromone-dependent competence regulon to upregulate the expression of competence genes required for DNA uptake and recombination.[165] In the anginosus group, different species utilize the same CSP signaling molecule, while a large variation of CSP pheromones have been reported within the same species in the mitis group, specifically

among *Streptococcus mitis* strains.[166-168] The mitis group of streptococci is comprised of 13 species, most of them are commensal bacteria, with *S. pneumoniae* also being a notorious opportunistic pathogen.[166,167] The recombination events among the mitis group of streptococci have been shown to be instrumental to the development of antibiotic resistance. Specifically, interspecies gene transfer of both antimicrobial resistance and virulence genes were observed between *S. pneumoniae* and its neighboring species, *S. mitis*, who shares more than 80% of its genes with *S. pneumoniae*.[167,169] In this study we aimed to evaluate the impact native CSP signals of streptococci of the mitis and anginosus groups have on the activity of the *S. pneumoniae* competence regulon. To this end, we chemically synthesized different reported native streptococcal CSP pheromones of the mitis and anginosus groups and tested their ability to activate/inhibit the *S. pneumoniae* ComD1 and ComD2 receptors. Our initial analysis revealed that one of the CSP pheromones of *S. mitis*, *S. mitis*-CSP-2 (**Figure 19**), can activate both *S. pneumoniae* ComD receptors. We therefore chose *S. mitis*-CSP-2 as the lead scaffold for the development of pan-group *S. pneumoniae* QS modulators. Indeed, using *S. mitis*-CSP-2 as a template, we were able to develop potent pan group QS inhibitors of *S. pneumoniae* QS. Moreover, our results revealed potential crosstalk between *S. mitis* and *S. pneumoniae*.

RESULTS AND DISCUSSION

Synthesis and initial screening of native streptococcal competence pheromones (CSPs). In this work, we first set out to evaluate potential crosstalk between species of the mitis group by testing the ability of native streptococcal competence pheromones (CSPs) to modulate the activity of the *S. pneumoniae* competence regulon QS circuitry. To this end, we chemically synthesized different native CSPs of the most common commensal

bacterial strains of the mitis group (see **Table 7** for strain information and CSP sequences). To explore potential crosstalk between different groups of streptococci, we decided to include CSP pheromones from the anginosus group. Since most of the species of the anginosus group, namely *S. anginosus*, *S. constellatus* and *S. intermedius* produce the same CSP signal,[167] we synthesized only one CSP signal (*S. intermedius*-CSP) representing the anginosus group of streptococci (**Table 7**). Lastly, in the initial screening assay, we included the two CSP forms of *S. mutans*, *S. mutans*-21-CSP (pre-mature extracellular signal) and *S. mutans*-18-CSP (mature extracellular signal),[170] to test for potential crosstalk between opportunistic pathogens. All of the CSP pheromones (**Table 7**) were constructed using established solid-phase peptide synthesis (SPPS) protocols on 4-benzyloxybenzyl alcohol (Wang) resin.[171] The peptides were then cleaved and purified to homogeneity using semipreparative RP-HPLC (>95% purity), and their identity was confirmed using mass spectrometry (see the Experimental section and Appendix 3 for full experimental details). Next, we tested the ability of the native streptococcal CSPs to modulate the activity of the *S. pneumoniae* ComD receptors (both ComD1 and ComD2) using a β-galactosidase cell-based bacterial reporter assay. Two β-gal reporter strains, D39pcomX::lacZ and TIGR4pcomX::lacZ, previously constructed by Lau and co-workers, were utilized to evaluate QS modulation in both groups of *S. pneumoniae*.[172] These reporter systems carry the *lacZ* gene under the control of the *comX* promoter; thus, upon QS activation, *lacZ* will be expressed, allowing for monitoring of the QS response by measuring β-galactosidase activity (see the Experimental section for details). We then conducted an initial screening of all the native CSP pheromones at high peptide concentration (10 μM; see **Figures S-1, S-6, S-9, and S-14 in Appendix 3**). The initial screening results revealed that of the twelve

CSP signals tested, only one native CSP pheromone, *S. mitis*-CSP-2, was able to fully activate both *S. pneumoniae* ComD receptors (>75% activation compared to the native *S. pneumoniae* CSPs; **Figures S-1 and S-9 in Appendix 3**). The EC_{50} values of this peptide were determined through dose response curves and found to be 663 nM and 635 nM when activating the *S. pneumoniae* ComD1 and ComD2 receptors, respectively (**Table 8**). Although these EC_{50} values are 12- to 64-fold higher than the native pneumococcus CSPs, the values are still biologically relevant as native CSP concentrations in streptococci supernatants were found to be in the micromolar range.[162] Since none of the other native peptides were capable of activating/inhibiting the *S. pneumoniae* QS circuitry, we selected the *S. mitis*-CSP-2 pheromone as a template for the development of pan-group *S. pneumoniae* QS modulators that could be utilized to attenuate pneumococcal infections.

Table 7. Sequence of streptococcal competence pheromones (CSPs) from select strains of streptococci for which the pheromone has been synthesized in this study

Group	Name of CSP	Strain	CSP sequence
Mitis	*S. pneumoniae* CSP1	*S. pneumoniae* D39	EMRLSKFFRDFILQRKK
	S. pneumoniae CSP2	*S. pneumoniae* TIGR4	EMRISRIILDFLFLRKK
	S. mitis-CSP-1	*S. mitis* NCTC 8033 (SK321)	ESRLPKIRFDFIFPRKK
	S. mitis-CSP-2	*S. mitis* CCUG 31611^{TG} (NCTC 12261)	EIRQTHNIFFNFFKRR
	S. cristatus-CSP	*S. cristatus* ACTC 51100^{G}	DLRNIFLKIKFKKK
	S. oligofermentans-CSP	*S. oligofermentans* LMG 21535^{T}	DSRNIFLKIKFKKK
	S. oralis-CSP	*S. oralis* SK255	DWRISETIRNLIFPRRK
	S. gordonii-CSP-1	*S. gordonii* NCTC 3165	SQKGVYASQRSFVPSWFRKIFRN
	S. gordonii-CSP-2	*S. gordonii* NCTC 7865^{T}	DIRHRINNSIWRDIFLKRK
	S. gordonii challis-CSP	*S. gordonii challis* CHI^{G}	DVRSNKIRLWWENIFFNKK
	S. sanguinis-CSP	*S. sanguinis* SK36^{G} (*S. sanguinis* NCTC 7863)	DLRGVPNPWGWIFGR
Anginosus	*S. intermedius*-CSP	*S. intermedius* NCTC 11324^{T}	DSRIRMGFDFSKLFGK
Mutans	*S. mutans*-18-CSP	*S. mutans* UA159^{G}	SGSLSTFFRLFNRSFTQA
	S. mutans-21-CSP		SGSLSTFFRLFNRSFTQALGK

^{T} type strain, ^{G} Genome sequence available

Point Modification of *S. mitis*-CSP-2 analogues. Previously, the Tal-Gan laboratory has conducted a detailed structure-activity relationship (SAR) analysis of the two *S. pneumoniae* signal variants, CSP1 and CSP2, against the two *S. pneumoniae* ComD

receptors, ComD1 and ComD2, and identified the critical residues involved in ComD1 and ComD2 receptor binding and activation.[173-175] Based on the SAR results of both the CSP1 and CSP2 scaffolds in *S. pneumoniae*, we designed ten *S. mitis*-CSP-2 analogues bearing a point modification in key positions to match the corresponding residue in *S. pneumoniae* CSP1 or CSP2 pheromones and examined their ability to modulate the competence regulon of *S. pneumoniae* (**Table 8**). Since the *S. mitis*-CSP-2 and the two CSP signals of *S. pneumoniae* differ in many positions (**Figure 19 and Table 7**), we chose the key positions contributing to the *S. pneumoniae* receptor binding and activation, all of which are in the *N*-terminus and the central region of *S. pneumoniae* CSPs, and systematically replaced each residue in *S. mitis*-CSP-2 with the residues found in *S. pneumoniae* CSP1 or CSP2. We excluded the modifications related to the *C*-terminus of the pheromones as the *C*-terminal residues of *S. pneumoniae* CSP1 and CSP2 were found to be dispensable.[173]

Table 8. EC_{50} values of *S. mitis*-CSP-2 point-modification analogues against *S. pneumoniae* ComD1 and ComD2 receptors[a]

Peptide name	Peptide Sequence	*S. pneumoniae* ComD1		*S. pneumoniae* ComD2	
		EC_{50} (nM)[b]	95% CI[c]	EC_{50} (nM)[b]	95% CI[c]
S. pneumoniae CSP1[d]	EMRLSKFFRDFILQRKK	10.3	6.27-16.8	526	498-556
S. pneumoniae CSP2[d]	EMRISRIILDFLFLRKK	1650	1190-2300	50.7	40.6-63.2
S. mitis-CSP-2	EIRQTHNIFFNFFKRR	663	608-722	635	426-947
S. mitis-CSP-2-I2M	E**M**RQTHNIFFNFFKRR	87.7	79.1-97.3	136	89.0-208
S. mitis-CSP-2-F10D	EIRQTHNIF**D**NFFKRR	--[e]	--	>1000	--
S. mitis-CSP-2-Q4L	EIR**L**THNIFFNFFKRR	88.2	64.6-121	252	231-275
S. mitis-CSP-2-N7F	EIRQTH**F**IFFNFFKRR	140	113-172	471	459-482
S. mitis-CSP-2-I8F	EIRQTHN**F**FFNFFKRR	128	93.2-176	209	139-313
S. mitis-CSP-2-N11F	EIRQTHNIFF**F**FFKRR	4.63	2.43-8.84	220	194-248
S. mitis-CSP-2-F12I	EIRQTHNIFFN**I**FKRR	--[e]	--	>1000	--
S. mitis-CSP-2-Q4I	EIR**I**THNIFFNFFKRR	--[e]	--	--	--
S. mitis-CSP-2-N7I	EIRQTH**I**IFFNFFKRR	101	79.5-129	211	147-303
S. mitis-CSP-2-F12L	EIRQTHNIFFN**L**FKRR	54.3	34.7-84.7	813	539-1230

[a] See the Experimental section for full experimental details and the Appendix 3 for plots of agonism dose response curves. [b] EC_{50} values were determined by testing peptides over a range of concentrations. [c] 95% confidence interval. [d] EC_{50} values of *S. pneumoniae* CSP1 and CSP2 from ref [27]. [e] EC_{50} not determined due to the analogue's low induction in primary agonism screening assay.

The initial screening of the *S. mitis*-CSP-2 point-modification analogues revealed that most of the mutated analogues exhibited higher potency against both *S. pneumoniae* ComD1 and ComD2 receptors compared to the native *S. mitis*-CSP-2 (**Figures S-2 and S-10 in Appendix 3, and Table 8**). Specifically, this analysis revealed that *S. mitis*-CSP-2-I2M, a point-mutated analogue that introduced the conserved residue, Met2, that is present in both *S. pneumoniae* pheromones, has increased potency against both *S. pneumoniae* ComD receptors compared to the native *S. mitis*-CSP-2. These results support previous observations that Met2 plays a critical role in binding and activation in both *S. pneumoniae* receptors.[94,173] Contrary, the second *S. mitis*-CSP-2-based analogue bearing a conserved residue present in both *S. pneumoniae* CSPs, *S. mitis*-CSP-2-F10D, was found to activate only the *S. pneumoniae* ComD2 receptor, although only at high concentration (>1000 nM, **Table 8**). Looking at the *S. pneumoniae* CSP1-based modifications, alterations of the hydrophobic residues in the central region of *S. mitis* at positions 4, 7, 8, 11 and 12 in *S. mitis*-CSP-2 with the corresponding residues in *S. pneumoniae* CSP1 revealed that, with the exception of *S. mitis*-CSP-2-F12I, all the resulting analogues exhibit enhanced activity against ComD1 compared to the native *S. mitis*-CSP-2. Importantly, the replacement of the hydrophilic residue Asn at position 11 in *S. mitis*-CSP-2 with the hydrophobic residue Phe resulted in an analogue (*S. mitis*-CSP-2-N11F) with >140-fold increased potency compared to the native *S. mitis*-CSP-2 signal against *S. pneumoniae* ComD1 receptor, exhibiting comparable activity to the pneumococcal native peptide, *S. pneumoniae* CSP1. With the exception of *S. mitis*-CSP-2-F12I, all the resulting analogues also exhibited increased potency compared to the native *S. mitis*-CSP-2 signal against the *S. pneumoniae* ComD2 receptor (**Table 8**).

Moving to the *S. pneumoniae* CSP2-based modifications (Q4I, N7I, F12L), the initial screening revealed that only *S. mitis*-CSP-2-N7I exhibits enhanced potency compared to native *S. mitis*-CSP-2 against both pneumococcal ComD receptors, highlighting the important role the hydrophobic residues play in both pneumococcal receptor binding and activation. Contrary, *S. mitis*-CSP-2-Q4I was found to be relatively inactive against both *S. pneumoniae* ComD receptors (**Table 8**).

Multiple Modifications of *S. mitis*-CSP-2 analogues. The point modification study revealed several important *S. mitis*-CSP-2-based *S. pneumoniae* QS modulators. We therefore set out to utilize the recently acquired SAR insight to rationally design *S. pneumoniae* QS modulators with enhanced potencies. To this end, we synthesized a library of double and triple modification analogues where we incorporated multiple beneficial point modifications into a single analogue. Expectedly, the initial evaluation revealed that most of the doubly modified analogues are more active than the native *S. mitis*-CSP-2 signal against both *S. pneumoniae* ComD receptors. However, while some dual-modified analogues exhibited slightly increased or similar potency relative to the single replacement analogues, other double modification analogues resulted in reduction in potency compared to the single point modification analogues (**Table 9**). Specifically, the combination of the Q4L modification with other central region hydrophobic modification (N7F, N7I, I8F, or N11F) resulted in doubly-modified analogues with reduced potency compared to the monosubstituted analogues against both pneumococcal ComD receptors. Most importantly, we identified one double modification analogue, *S. mitis*-CSP-2-N7II8F, that can bind and activate both pneumococcal ComD receptors with enhanced potencies

compared to the single replacement analogues (compare the EC_{50} values of *S. mitis*-CSP-2-N7I and *S. mitis*-CSP-2-I8F against the pneumococcal ComD1 and ComD2 receptors in **Table 8** with the EC_{50} values of *S. mitis*-CSP-2-N7II8F against the pneumococcal ComD1 and ComD2 receptors in **Table 9**), making it the most potent *S. mitis*-CSP-2-based pan-group pneumococcal QS activator identified in the double modification library. Interestingly, a similar double-modified analogue where Asn7 was replaced with phenylalanine, the native residue in *S. pneumoniae* CSP1, resulted in a double-modified analogue, *S. mitis*-CSP-2-N7FI8F, that lost its activity against the pneumococcal ComD1 receptor, while maintaining its activity against the pneumococcal ComD2 receptor (**Table 9**).

Table 9. EC_{50} values of *S. mitis*-CSP-2 multiple-modification analogues against *S. pneumoniae* ComD1 and ComD2 receptors[a]

Peptide name	Peptide Sequence	*S. pneumoniae* ComD1		*S. pneumoniae* ComD2	
		EC_{50} (nM)[b]	95% CI[c]	EC_{50} (nM)[b]	95% CI[c]
S. pneumoniae CSP1[d]	EMRLSKFFRDFILQRKK	10.3	6.27-16.8	526	498-556
S. pneumoniae CSP2[d]	EMRISRIILDFLFLRKK	1650	1190-2300	50.7	40.6-63.2
S. mitis-CSP-2	EIRQTHNIFFNFFKRR	663	608-722	635	426-947
S. mitis-CSP-2-I2MQ4L	E**M**R**L**THNIFFNFFKRR	49.9	42.4-58.9	141	76.8-257
S. mitis-CSP-2-I2MN7F	E**M**RQTH**F**IFFNFFKRR	131	76.5-225	123	91.7-166
S. mitis-CSP-2-I2MN7I	E**M**RQTH**I**IFFNFFKRR	448	345-582	54.8	51.7-58.0
S. mitis-CSP-2-I2MI8F	E**M**RQTHN**F**FFNFFKRR	151	87.1-262	83.2	65.9-105
S. mitis-CSP-2-I2MN11F	E**M**RQTHNIFF**F**FFKRR	17.7	11.5-27.4	753	573-990
S. mitis-CSP-2-I2MF12L	E**M**RQTHNIFFN**L**FKRR	25.8	13.1-50.8	146	81.3-261
S. mitis-CSP-2-Q4LN7F	EIR**L**TH**F**IFFNFFKRR	371	293-469	322	257-402
S. mitis-CSP-2-Q4LN7I	EIR**L**TH**I**IFFNFFKRR	348	278-436	315	229-433
S. mitis-CSP-2-Q4LI8F	EIR**L**THN**F**FFNFFKRR	280	153-513	370	275-499
S. mitis-CSP-2-Q4LN11F	EIR**L**THNIFF**F**FFKRR	9.38	8.44-10.4	434	308-609
S. mitis-CSP-2-Q4LF12L	EIR**L**THNIFFN**L**FKRR	68.7	50.6-93.2	202	193-211
S. mitis-CSP-2-N7FI8F	EIRQTH**FF**FFNFFKRR	--[e]	--	24.6	21.9-27.6
S. mitis-CSP-2-N7FN11F	EIRQTH**F**IFF**F**FFKRR	--[e]	--	--[e]	--
S. mitis-CSP-2-N7FF12L	EIRQTH**F**IFFN**L**FKRR	156	84.9-287	533	262-1083
S. mitis-CSP-2-N7II8F	EIRQTH**IF**FFNFFKRR	87.2	56.8-134	22.8	13.1-40.0
S. mitis-CSP-2-N7IN11F	EIRQTH**I**IFF**F**FFKRR	--[e]	--	--[e]	--
S. mitis-CSP-2-N7IF12L	EIRQTH**I**IFFN**L**FKRR	613	506-743	321	216-475
S. mitis-CSP-2-I8FN11F	EIRQTHN**F**FF**F**FFKRR	12.2	5.73-26.0	284	188-429
S. mitis-CSP-2-I8FF12L	EIRQTHN**F**FFN**L**FKRR	160	104-248	202	145-281
S. mitis-CSP-2-N11FF12L	EIRQTHNIFF**FL**FKRR	4.97	4.12-5.99	127	117-137
S. mitis-CSP-2-I2MQ4LN7F	E**M**R**L**TH**F**IFFNFFKRR	137	105-178	75.6	69.2-82.7
S. mitis-CSP-2-I2MI8FN11F	E**M**RQTHN**F**FF**F**FFKRR	6.95	4.69-10.3	26.2	14.1-49.0
S. mitis-CSP-2-I2MN7FF12L	E**M**RQTH**F**IFFN**L**FKRR	72.8	46.7-114	112	76.4-163
S. mitis-CSP-2-I2MN7II8F	E**M**RQTH**IF**FFNFFKRR	61.6	46.1-82.3	2.67	1.91-3.73
S. mitis-CSP-2-I2MQ4LF12L	E**M**R**L**THNIFFN**L**FKRR	17.1	15.8-18.5	139	92.6-210
S. mitis-CSP-2-I2MQ4LI8F	E**M**R**L**THN**F**FFNFFKRR	42.0	25.7-68.5	30.0	15.1-59.3
S. mitis-CSP-2-I2MQ4LN7I	E**M**R**L**TH**I**IFFNFFKRR	155	141-170	187	150-233
S. mitis-CSP-2-I2MQ4LN11F	E**M**R**L**THNIFF**F**FFKRR	14.8	11.3-19.4	188	126-281
S. mitis-CSP-2-Q4LN7FI8F	EIR**L**TH**FF**FFNFFKRR	427	402-454	87.2	69.1-110
S. mitis-CSP-2-Q4LN7II8F	EIR**L**TH**IF**FFNFFKRR	242	120-489	--[e]	--
S. mitis-CSP-2-Q4LN7FN11F	EIR**L**TH**F**IFF**F**FFKRR	--[e]	--	--[e]	--
S. mitis-CSP-2-I2MN7FI8F	E**M**RQTH**FF**FFNFFKRR	63.6	34.1-119	13.5	11.7-15.5
S. mitis-CSP-2-Q4LN7FF12L	EIR**L**TH**F**IFFN**L**FKRR	426	292-622	--[e]	--
S. mitis-CSP-2-N7FI8FF12L	EIRQTH**FF**FFN**L**FKRR	101	63.5-160	67.6	36.2-126
S. mitis-CSP-2-N7II8FN11F	EIRQTH**IF**FF**F**FFKRR	--[e]	--	--[e]	--
S. mitis-CSP-2-N7II8FF12L	EIRQTH**IF**FFN**L**FKRR	891	793-1001	--[e]	--
S. mitis-CSP-2-I2MN7IN11F	E**M**RQTH**I**IFF**F**FFKRR	568	429-750	--[e]	--
S. mitis-CSP-2-N7FI8FN11F	EIRQTH**FF**FF**F**FFKRR	--[e]	--	--[e]	--

[a] See the Experimental section for full experimental details and the Appendix 3 for plots of agonism dose response curves. [b] EC_{50} values were determined by testing peptides over a range of concentrations. [c] 95% confidence interval. [d] EC_{50} values of *S. pneumoniae* CSP1

and CSP2 from ref [27]. [e] EC_{50} not determined due to the analogue's low induction in primary agonism screening assay. * Key analogues discussed in the text are highlighted with a grey shadow.

Moving to the triple-modified analogues, it appears that the majority of combinations were not tolerated well, as most analogues exhibited similar or reduced activities compared to the single- or doubly-modified analogues against the pneumococcal ComD receptors (**Table 9**). However, the triple-modified library also revealed four *S. mitis*-CSP-2-based analogues, *S. mitis*-CSP-2-I2MI8FN11F, *S. mitis*-CSP-2-I2MN7II8F, *S. mitis*-CSP-2-I2MQ4LI8F, and *S. mitis*-CSP-2-I2MN7FI8F, with activities at the low nanomolar range (EC_{50} values <100 nM) against both pneumococcal ComD receptors (**Table 9**). All four analogues share the same two modifications, I2M and I8F, suggesting that these two modifications are critical for pan-group reactivity of the *S. mitis*-CSP-2 scaffold against the pneumococcus ComD receptors.

Development of *S. mitis*-CSP-2-based pneumococcal pan-group QS inhibitors. It has been shown that interception of the pneumococcal competence regulon by either targeting CSP export or disrupting native CSP:ComD interactions can be utilized to attenuate QS-dependent pneumococcus pathogenicity.[58,172,173,176,177] All previously developed CSP-based inhibitors share the same key modification, the replacement of glutamic acid at position 1 with alanine.[58,172-175,177-179] We therefore reasoned that the *S. mitis*-CSP-2 template can also be converted into a pneumococcus ComD inhibitor by incorporating the same E1A modification. To this end, we first replaced Glu1 with alanine in the native *S. mitis*-CSP-2 to afford *S. mitis*-CSP-2-E1A and found that the resulting analogue can effectively inhibit the pneumococcus ComD1 receptor, albeit at 6-fold higher

concentration than *S. pneumoniae* CSP1-E1A (IC_{50} value of 497 nM), while losing all activity against the *S. pneumoniae* ComD2 receptor (**Table 10**).

Next, we selected the top *S. mitis*-CSP-2 modified analogues identified in this study and incorporated the E1A modification. We hypothesized that combining these modifications together with the E1A mutation would result in *S. mitis*-CSP-2-based analogues that maintain the binding affinity to pneumococcal ComD receptors but may not be able to activate the receptors, leading to competitive pan-group antagonists of the *S. pneumoniae* competence regulon. To test our hypothesis, we synthesized eleven *S. mitis*-CSP-2 analogues bearing the E1A modification. These analogues included two *S. mitis*-CSP-2 single-modified analogues, five *S. mitis*-CSP-2 double-modified analogues and four *S. mitis*-CSP-2 triple-modified analogues, all bearing also the E1A modification, and tested their ability to modulate the QS circuitry in both pneumococcal specificity groups (**Table 10**). Our analysis revealed that all the *S. mitis*-CSP-2-E1A-based analogues that had the N11F modification were inactive against both pneumococcal ComD receptors (**Table 10**). Since the N11F modification resulted in highly potent QS activators, the lack of inhibitory activity for all the E1A-based analogues bearing this modification highlights the stringent and different requirements for ComD receptor inhibition compared to receptor activation. Most importantly, from the *S. mitis*-CSP-2-E1A-based library we identified three nanomolar range antagonists of both pneumococcus pherotypes, *S. mitis*-CSP-2-E1AN7II8F, *S. mitis*-CSP-2-E1AI2MN7II8F and *S. mitis*-CSP-2-E1AI2MN7FI8F (**Table 10**). These potent *S. mitis*-CSP-2-based pan-group inhibitors of the pneumococcus competence regulon are an important addition to the arsenal of chemical tools available to study pneumococcus behavior and attenuate pneumococcal infections.

Table 10. IC_{50} values of *S. mitis*-CSP-2-E1A modification analogues against *S. pneumoniae* ComD1 and ComD2 receptors[a]

Peptide name	Peptide Sequence	*S. pneumoniae* ComD1		*S. pneumoniae* ComD2	
		IC_{50} (nM)[b]	95% CI[c]	IC_{50} (nM)[b]	95% CI[c]
S. pneumoniae CSP1-E1A[e]	**A**MRLSKFFRDFILQRKK	85.7	50.8-145	--[d]	--[d]
S. mitis-CSP-2-E1A	**A**IRQTHNIFFNFFKRR	497	422-585	--[d]	--
S. mitis-CSP-2-E1AI2M	**AM**RQTHNIFFNFFKRR	85.4	57.4-127	--[d]	--
S. mitis-CSP-2-E1AN11F	**A**IRQTHNIFF**F**FFKRR	--[d]	--	--[d]	--
S. mitis-CSP-2-E1AI2MQ4L	**AM**R**L**THNIFFNFFKRR	54.2	40.6-72.2	--[d]	--
S. mitis-CSP-2-E1AN11FF12L	**A**IRQTHNIFF**FL**FKRR	--[d]	--	--[d]	--
S. mitis-CSP-2-E1AN7II8F	**A**IRQTH**IF**FFNFFKRR	204	133-311	135	82.5-222
S. mitis-CSP-2-E1AI2MF12L	**AM**RQTHNIFFN**L**FKRR	--[d]	--	--[d]	--
S. mitis-CSP-2-E1AI2MI8F	**AM**RQTHN**F**FFNFFKRR	143	88.9-229	--[d]	--
S. mitis-CSP-2-E1AI2MN7FI8F	**AM**RQTH**FF**FFNFFKRR	294	263-328	418	235-744
S. mitis-CSP-2-E1AI2MN7II8F	**AM**RQTH**IF**FFNFFKRR	141	81.8-243	32.9	16.4-66.0
S. mitis-CSP-2-E1AI2MQ4LI8F	**AM**R**L**THN**F**FFNFFKRR	317	238-423	--[d]	--
S. mitis-CSP-2-E1AI2MI8FN11F	**AM**RQTHN**F**FF**F**FFKRR	--[d]	--	--[d]	--

[a] See the Experimental section for full experimental details and the Appendix 3 for plots of antagonism dose response curves. [b] IC_{50} values were determined by testing peptides over a range of concentrations. [c] 95% confidence interval. [d] IC_{50} not determined due to the analogue's low induction in primary antagonism screening assay. [e] IC_{50} value of *S. pneumoniae* CSP1-E1A from ref [27].

CONCLUSIONS

In conclusion, we set out to investigate potential crosstalk between streptococci species and utilize the identified signaling molecules to develop novel pneumococcal QS modulators. First, our analysis revealed that the majority of CSP signals from closely related streptococci of the mitis group, as well as CSPs from other groups of streptococci were unable to modulate QS in *S. pneumoniae*, highlighting the relatively high specificity of the pneumococcus ComD receptors to their cognate native CSP signals. Second, our analysis also revealed a potential role of *S. mitis* as a modulator of the competence regulon in *S. pneumoniae*. Our rationally-designed *S. mitis*-CSP-2-based analogues, generated using the SAR results of both the CSP1 and CSP2 scaffolds in *S. pneumoniae*,[173] yielded several potent pan-group agonists of pneumococcal QS. Moreover, our analysis revealed that by using the *S. mitis*-CSP-2 pheromone as a template, it is possible to develop potent

antagonists of the pneumococcal competence regulon. Most importantly, our work revealed the first pan-group *S. mitis*-CSP-2-based ComD inhibitors of *S. pneumoniae* with activities in the nanomolar range.

Since *S. mitis* has the potential to influence the regulation of pneumococcus phenotypes associated with virulence and infectivity, developing highly potent QS modulators using the native CSP pheromone of *S. mitis*, *S. mitis*-CSP-2, with enhanced potencies could lead to a complementary strategy to attenuate pneumococcal infections. Furthermore, these privileged scaffolds could influence the competence regulon in *S. mitis*, thus providing novel chemical tools capable of modulating QS in multiple species. Indeed, experiments aimed at identifying the effect of the *S. mitis*-CSP-2 analogues on the competence regulon in *S. mitis* are ongoing in our laboratory and will be reported in due course.

EXPERIMENTAL SECTION

Chemical reagents and instrumentation. All chemical reagents and solvents were purchased from Sigma-Aldrich or Chem-Impex and used without further purification. Water (18 MΩ) was purified using a Millipore Analyzer Feed System. Solid-phase resins were purchased from Advanced Chem Tech.

Reversed-phase high-performance liquid chromatography (RP-HPLC) was performed using a Shimadzu system equipped with a CBM-20A communications bus module, two LC-20AT pumps, an SIL-20A auto sampler, an SPD-20A UV/VIS detector, a CTO-20A column oven, and an FRC-10A fraction collector. Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) data were obtained on a

Bruker Microflex spectrometer equipped with a 60 Hz nitrogen laser and a reflectron. In positive ion mode, the acceleration voltage on Ion Source 1 was 19.01 kV. Exact mass (EM) data were obtained on an Agilent Technologies 6230 TOF LC/MS spectrometer. The samples were sprayed with a capillary voltage of 3500 V and the electrospray ionization (ESI) source parameters were as follows: gas temperature of 325 °C at a drying gas flow rate of 8 L/min at a pressure of 35 psi.

Peptide Synthesis. All the CSP analogues were synthesized using standard Fmoc-based solid-phase peptide synthesis (SPPS) procedures on 4-Benzyloxybenzyl alcohol (Wang) resin. Pre-loaded Fmoc-L-Arg (Pbf) Wang resin (0.305 mmol/g), Fmoc-L-Lys (Boc) Wang resin (0.59 mmol/g), or Fmoc-Asn (Trt) Wang resin (0.332 mmol/g) was used for peptides that required an arginine, lysine, or asparagine at the *C*-terminus, respectively. CSP and analogues were synthesized by using the Liberty1 automated peptide synthesizer (CEM Corporation) (for the full procedures, see the **Appendix 3**).

Peptide purification by HPLC. Crude peptides were purified with RP-HPLC. The crude peptide was dissolved by using ACN:H_2O (1:3) (volume of ACN in water depends on the solubility of the peptide) and purified in 5 mL portions on a semi-preparative Phenomenex Kinetex C18 column (5 μm, 10 mm × 250 mm, 110 Å) with a 5 mL/min flow rate; mobile phase A= 18 MΩ water + 0.1 % TFA and mobile phase B = ACN + 0.1 % TFA. The collected fraction was frozen using an acetone/dry ice bath, lyophilized overnight, and dissolved again in 5 mL ACN:H_2O (1:3) for a secondary purification run. Preparative HPLC methods were used to separate the crude peptide mixture to different chemical components using the following linear gradient: (first prep 5% B → 45% B over 40 min and second prep 25% B → 38% B over 45 min). Fraction purity was determined through

analysis on a Phenomenex Kinetex analytical C18 column (5 μm, 4.6 mm × 250 mm, 110 Å) using a linear gradient (5% B → 95% B over 27 min). Purities were determined by integration of peaks with UV detection at 220 nm. Only peptide fractions that were purified to homogeneity (>95%) were used for the biological assays.

Biological Assays

Biological Reagents and Strain Information. All standard biological reagents were purchased from Sigma-Aldrich and used according to enclosed instructions. Donor horse serum (defibrinated) was stored at 4 °C until use in bacterial growth conditions. To examine the ability of the synthesized CSP analogues to modulate the ComD receptors, and thus, the QS circuit in *S. pneumoniae*, beta-galactosidase assays were performed using D39pcomX::lacZ (group I) and TIGR4pcomX::lacZ (group II) reporter strains.[172]

Bacterial Growth Conditions. Freezer stocks of individual pneumococcal strains, D39pcomX::lacZ and TIGR4pcomX::lacZ were created from 1.5 mL aliquots of bacteria (0.2 OD 600_{nm}) in Todd-Hewitt broth supplemented with 0.5% yeast extract (THY) and 0.5 mL glycerol, and stored at -80 °C. For experiments, bacteria from the freezer stocks were streaked onto a THY agar plate supplemented with 5% horse serum with chloramphenicol at a final concentration of 4 μg/mL. The plate was incubated for 8 h in a CO_2 incubator (37 °C with 5% CO_2). Isolated fresh colonies were then transferred to 5 mL of THY broth containing 4 μg/mL of chloramphenicol and incubated in a CO_2 incubator overnight (15 h). These overnight cultures were then diluted (1:50 for D39pcomX::lacZ; 1:10 for TIGR4pcomX::lacZ) with THY and the resulting solution was incubated in a CO_2 incubator for 3-4 hours, until the bacteria reached early exponential stage (OD_{600} of 0.30-

0.35 for D39pcomX::lacZ and 0.20-0.25 for TIGR4pcomX::lacZ) as determined by using a plate reader.

Beta-Galactosidase Assays. The ability of synthetic CSP analogues to activate the QS circuit was evaluated as previously described (see the **Appendix 3** for full experimental details).[173,174]

EC_{50} Experiments. Analogues that exhibited high activity in the initial screening were further evaluated using a dose-dependent assay in which peptide stock solutions were diluted with DMSO in serial dilutions (either 1:2, 1:3 or 1:5) and assayed as described in the **Appendix 3**. GraphPad Prism 5 was used to calculate the EC_{50} values, which are the concentration of a drug that gives half-maximal response.

IC_{50} Experiments. Analogues that exhibited low *S. pneumoniae comX* activation in the initial screening were evaluated for competitive inhibition. The ability of synthesized CSP analogues to inhibit the expression of *S. pneumoniae comX* by outcompeting *S. pneumoniae* CSP for the receptor binding site was evaluated using the same assay conditions as described in the **Appendix 3**, except that in the initial inhibition screening, the native *S. pneumoniae* CSP was added to every well in a set concentration (50 nM *S. pneumoniae* CSP1 for group I; 250 nM *S. pneumoniae* CSP2 for group II) that was chosen to afford full activation of the QS circuit, as determined from the dose-dependent curves created for the native *S. pneumoniae* CSPs. Two μL of native CSP (5 μM solution of CSP1 for group I; 25 μM solution of CSP2 for group II) and 2 μL of 1 mM solution of synthetic CSP analogues were added to the same well in triplicate in a clear 96-well microtiter plate. Two μL native CSP (5 μM solution of CSP1 for group I; 25 μM solution of CSP2 for group II) and 2 μL DMSO were added to the same well in triplicate and served as the positive

control. Four µL DMSO were added in triplicate and served as the negative control. Then, 196 µL bacterial culture were added to the wells and the plate was incubated at 37 °C for 30 min. The procedure for lysis, incubation with ONPG and all the measurements were as described in the beta-galactosidase assay section in the **Appendix 3**. Analogs that exhibited significant competitive inhibition in the initial screening were further evaluated using a dose-dependent assay where peptide stock solutions were diluted with DMSO in serial dilutions (either 1:2, 1:3, or 1:5) and assayed as described above. GraphPad Prism 5 was used to calculate the IC_{50} values, which are the concentration of an inhibitor where the response (or binding) is reduced by half.

ASSOCIATED CONTENT

Supporting Information (Appendix 3)

The Supporting Information is available free of charge:

Full details of experimental procedures, peptide characterization, initial screening results, and dose-response curves for all CSP analogues (PDF)

AUTHOR INFORMATION

Corresponding Author

*E-mail: ytalgan@unr.edu

ORCID

Yftah Tal-Gan: 0000-0003-2052-6782

Notes

The authors declare no competing financial interest.

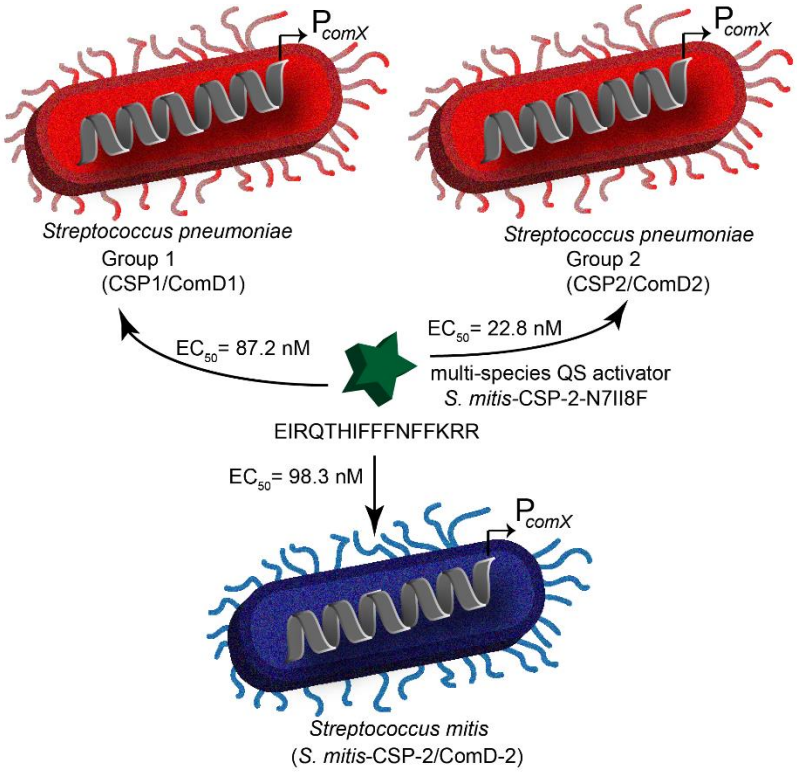
ACKNOWLEDGEMENTS

This work was supported by a grant from the National Institutes of Health (R35GM128651). *S. pneumoniae* D39pcomX::lacZ and TIGR4pcomX::lacZ reporter strains were generous gifts from G. W. Lau (University of Illinois at Urbana-Champaign).

REFERENCES

1. S. Mehr and N. Wood, *Paediatr Respir Rev*, 2012, **13**, 258-264.
2. S. S. Huang, K. M. Johnson, G. T. Ray, P. Wroe, T. A. Lieu, M. R. Moore, E. R. Zell, J. A. Linder, C. G. Grijalva and J. P. Metlay, *Vaccine*, 2011, **29**, 3398-3412.
3. J. Cornick and S. Bentley, *Microbes Infect*, 2012, **14**, 573-583.
4. N. J. Croucher, S. R. Harris, C. Fraser, M. A. Quail, J. Burton, M. van der Linden, L. McGee, A. von Gottberg, J. H. Song and K. S. Ko, *Science*, 2011, **331**, 430-434.
5. L. Håvarstein, G. Coomaraswamy and D. A. Morrison, *Proc Natl Acad Sci U S A*, 1995, **92**, 11140-11144.
6. G. W. Lau, S. Haataja, M. Lonetto, S. E. Kensit, A. Marra, A. P. Bryant, D. McDevitt, D. A. Morrison and D. W. Holden, *Mol Microbiol*, 2001, **40**, 555-571.
7. L. Zhu, J. Lin, Z. Kuang, J. E. Vidal and G. W. Lau, *Mol Microbiol*, 2015, **97**, 151-165.
8. M. R. Oggioni, C. Trappetti, A. Kadioglu, M. Cassone, F. Iannelli, S. Ricci, P. W. Andrew and G. Pozzi, *Mol Microbiol*, 2006, **61**, 1196-1210.
9. D. L. Hava and A. Camilli, *Mol Microbiol*, 2002, **45**, 1389-1406.
10. S. T. Rutherford and B. L. Bassler, *Cold Spring Harb Perspect Med*, 2012, **2**, a012427.
11. D. N. McBrayer, B. K. Gantman, C. D. Cameron and Y. Tal-Gan, *Org Lett*, 2017, **19**, 3295-3298.
12. B. LaSarre and M. J. Federle, *Microbiol Mol Biol Rev*, 2013, **77**, 73-111.
13. Y. Tal-Gan, D. M. Stacy, M. K. Foegen, D. W. Koenig and H. E. Blackwell, *J Am Chem Soc*, 2013, **135**, 7869-7882.
14. Y. Tal-Gan, M. Ivancic, G. Cornilescu, T. Yang and H. E. Blackwell, *Angew Chem Int Ed Engl*, 2016, **55**, 8913-8917.
15. A. Harrington and Y. Tal-Gan, *J Bacteriol*, 2018, **200**, e00709-00717.
16. R. W. Mull, A. Harrington, L. A. Sanchez and Y. Tal-Gan, *Curr Top Med Chem*, 2018, **18**, 625-644.
17. O. Johnsborg, P. E. Kristiansen, T. Blomqvist and L. S. Håvarstein, *J Bacteriol*, 2006, **188**, 1744-1749.
18. G. Pozzi, L. Masala, F. Iannelli, R. Manganelli, L. Havarstein, L. Piccoli, D. Simon and D. Morrison, *J Bacteriol*, 1996, **178**, 6087-6090.
19. G. Salvadori, R. Junges, R. Khan, H. A. Åmdal, D. A. Morrison and F. C. Petersen, *Methods Mol Biol*, 2017, **1537**, 219-232.

20. M. Kilian, K. Poulsen, T. Blomqvist, L. S. Håvarstein, M. Bek-Thomsen, H. Tettelin and U. B. Sørensen, *PLoS One*, 2008, **3**, e2683.
21. L. S. Håvarstein, R. Hakenbeck and P. Gaustad, *J Bacteriol*, 1997, **179**, 6589-6594.
22. G. Salvadori, R. Junges, D. A. Morrison and F. C. Petersen, *Front Microbiol*, 2016, **7**, 1009.
23. S. Shekhar, R. Khan, D. M. Ferreira, E. Mitsi, E. German, G. H. Rørvik, D. Berild, K. Schenck, K. Kwon and F. Petersen, *Front Immunol*, 2018, **9**, 747.
24. C. R. Bikash, S. R. Hamry and Y. Tal-Gan, *ACS Infect Dis*, 2018, **4**, 1385-1394.
25. W. C. Chan and P. D. White, *Fmoc solid phase peptide synthesis: a practical approach*, Oxford University Press Oxford, 2000.
26. L. Zhu and G. W. Lau, *PLoS Pathog*, 2011, **7**, e1002241.
27. Y. Yang, B. Koirala, L. A. Sanchez, N. R. Phillips, S. R. Hamry and Y. Tal-Gan, *ACS Chem Biol*, 2017, **12**, 1141-1151.
28. B. Koirala, N. R. Phillips and Y. Tal-Gan, *ACS Med Chem Lett*, 2019, **10**, 880-886.
29. Y. Yang and Y. Tal-Gan, *Bioorg Chem*, 2019, **89**, 102987.
30. Y. Yang, G. Cornilescu and Y. Tal-Gan, *Biochemistry*, 2018, **57**, 5359-5369.
31. A. Domenech, A. R. Brochado, V. Sender, K. Hentrich, B. Henriques-Normark, A. Typas and J. W. Veening, *Cell Host Microbe*, 2020, **27**, 544-555 e543.
32. B. Koirala, J. Lin, G. W. Lau and Y. Tal-Gan, *ChemBioChem*, 2018, **19**, 2380-2386.
33. Y. Yang, J. Lin, A. Harrington, G. Cornilescu, G. W. Lau and Y. Tal-Gan, *Proc Natl Acad Sci U S A*, 2020, **117**, 1689-1699.
34. B. Koirala, R. A. Hillman, E. K. Tiwold, M. A. Bertucci and Y. Tal-Gan, *Beilstein J Org Chem*, 2018, **14**, 1769-1777.
35. B. Koirala and Y. Tal-Gan, *Chembiochem*, 2020, **21**, 340-345.



Chapter 5: Developing Multi-Species Quorum Sensing Modulators Based on the *Streptococcus mitis* Competence-Stimulating Peptide[a]

[a]Submitted

Developing Multi-Species Quorum Sensing Modulators Based on the *Streptococcus mitis* Competence-Stimulating Peptide

Tahmina A. Milly, Clay P. Renshaw, and Yftah Tal-Gan

Department of Chemistry, University of Nevada, Reno, 1664 North Virginia Street, Reno, Nevada, 89557, United States

ABSTRACT.

Quorum sensing (QS) is a cell-cell communication mechanism utilized by bacteria to coordinate group behaviors, including pathogenic phenotypes. As such, QS has attracted significant attention as a potential mean to attenuate bacterial infectivity without introducing selective pressure for resistance development. *Streptococcus mitis*, a human commensal, acts as a genetic diversity reservoir for *Streptococcus pneumoniae*, one of the most important human pathogens. *S. mitis* possesses a typical *comABCDE* competence regulon QS circuitry found in most members of the mitis group of streptococci. However, the competence-stimulating peptide (CSP) responsible for QS activation and the regulatory role of the competence regulon QS circuitry in *S. mitis* are yet to be explored. Furthermore, as competence for genetic transformation is a hallmark phenotype controlled by the competence regulon and is responsible for the transfer of genetic material between *S. mitis* and *S. pneumoniae*, this QS circuitry likely has a major role in the crosstalk between these two bacterial species. Therefore, herein, we set out to delineate the competence regulon QS circuitry in *S. mitis*, including confirming the identity of the native CSP signal, evaluating the molecular mechanism that governs CSP:ComD interactions and leads to ComD activation, and defining the regulatory roles of the competence regulon QS circuitry in initiating various *S. mitis* phenotypes. Our analysis revealed important SAR insights of

the CSP signal and facilitated the development of novel CSP-based QS modulators, both enhanced QS activators and potent competitive inhibitors. Our analysis also revealed the involvement of the competence regulon in modulating competence development and biofilm formation. Furthermore, our analysis revealed that the native *S. mitis* CSP signal, *S. mitis*-CSP-2, can modulate the QS response in *S. pneumoniae*, but that the native CSP signals in *S. pneumoniae* are not effective modulators of QS response in *S. mitis*. Capitalizing on this potential crosstalk, we developed a *S. mitis*-CSP-2-based QS modulator capable of activating both the pneumococcus ComD receptors and the *S. mitis* ComD-2 receptor with high potencies (*S. mitis*-CSP-2-N7II8F; EC_{50} values in the low nanomolar range against all three receptors). To the best of our knowledge, this is the first example of a CSP-based multi-species QS activator. Overall, the novel CSP-based scaffolds identified in this study can be utilized to evaluate the effects temporal QS modulation has on the ability of *S. mitis* to inhabit its natural niche.

INTRODUCTION

Streptococcus mitis is a pioneer oral commensal species that inhabits the human mouth and is an integral member of the oral microbiome. *S. mitis* colonizes virtually all oral sites, such as dental hard tissues as well as mucous membranes, and is found in the throat and nasopharynx, where it may exist, side-by-side, in biofilms with *Streptococcus pneumoniae* (pneumococcus), a closely related major human pathogen.[1-3] *S. pneumoniae* is also a nasopharyngeal commensal, however, unlike *S. mitis*, has high pathogenic potential and is responsible for an estimated 445 000 hospitalizations and 22 000 deaths annually in the United States alone.[4] Most of the deaths caused by *S. pneumoniae* are a result of bacteremia, meningitis, or pneumonia.[5] Conversely, *S. mitis* presents mostly a nonvirulent behavior, and is only occasionally associated with diseases,[5-8] however, concerns about its pathogenic potential have recently been raised.[9] The differences in pathogenic potential in these two species are striking, as they share more than 80% gene similarity.[1,3] Evolutionary analyses suggest that although both species share a common ancestor, loss of virulence genes may have contributed to the less pathogenic potential in *S. mitis*.[1] The ability to persistently colonize oral surfaces and the nasopharynx, as well as the ability to induce mucosal antibody responses, are unique biological features of *S. mitis* that make this commensal bacterium a potential mucosal vaccine or therapeutic delivery vehicle.[10] Despite the universal prevalence of *S. mitis* as a human colonizer, very few studies have been conducted to address the molecular mechanisms involved in *S. mitis* colonization and host interactions, making it of particular interest to study.[6,11]

S. mitis acts as a genetic pool to competent pneumococci, including genes responsible for antibiotic resistance and virulence factor production.[12] It has been postulated that the

acquisition of *S. mitis* genes by *S. pneumoniae* has contributed to the survival of *S. pneumoniae* during stress conditions, suggesting that *S. mitis* plays an important role in pneumococcal evolution.[13,14] This process was found to be bidirectional, as genome analysis of six *S. mitis* strains revealed the presence of a pneumococcal-like capsule locus in four of them including in the *S. mitis* type strain CCUG 31611 (also known as NCTC 12261 or ATCC 49456).[15,16] The exchange of serotype 4 capsule locus has been reported between the *S. mitis* type strain and *S. pneumoniae* TIGR4, following induction of competence for natural transformation.[11] Natural transformation has been considered as a dominant force for the evolution of bacteria. This well-known process is used by competent streptococci to acquire genetic material from the surroundings, including genes that confer antibiotic resistance, as well as genes that regulate a wide range of functions related to pathogenicity, such as virulence factor production and biofilm formation. In both *S. mitis* and *S. pneumoniae*, competence is regulated by a density dependent quorum-sensing (QS) system, which is centered on competence pheromones, also known as competence stimulating peptides (CSPs).[17-20] This QS circuitry, termed the competence (or ComABCDE) regulon, launches the competence development cascade: first, the CSP precursor-peptide (ComC) is processed and exported to the extracellular environment by an ABC transporter (ComAB) as the mature CSP signal.[21] Once CSP reaches a threshold concentration, the CSP binds and activates a membrane-bound histidine kinase receptor (ComD).[21] This binding event results in the phosphorylation of the cytoplasmic response regulator (ComE). Phosphorylated ComE acts as a transcription factor to upregulate the transcription of the QS genes (*comABCDE*) as well as the alternative sigma factor gene

(*comX*), the master regulator of the QS circuitry that controls the different QS-regulated phenotypes, including competence (**Figure 20**).

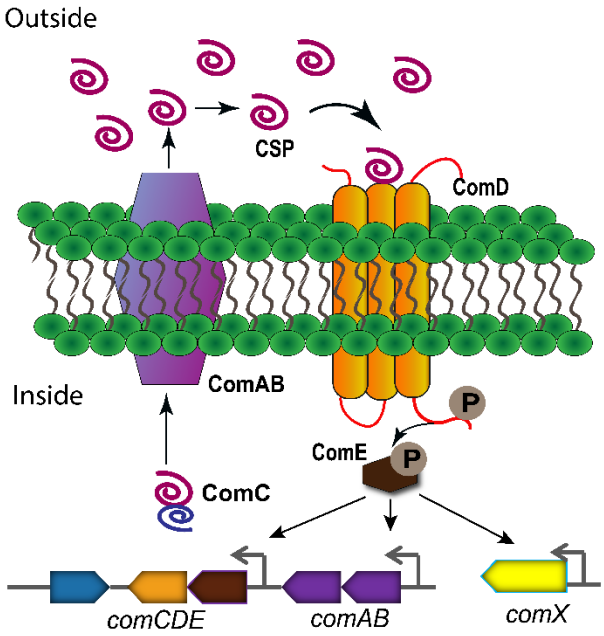


Figure 20. General streptococcal CSP-mediated QS pathway. ComC is processed and exported to the extracellular environment as the mature CSP signal by ComAB. At high concentration, the CSP binds to its cognate histidine-kinase receptor ComD. Activation of ComD leads to phosphorylation of ComE, a response regulator, resulting in further activation of the *comABCDE* operon and expression of *comX*, the master regulator of group-behavior genes.

S. mitis has the potential to influence pneumococcal QS-regulated phenotypes. Indeed, in a previous study we have shown that a native *S. mitis* CSP pheromone, *S. mitis*-CSP-2, can effectively activate the competence regulon in both group 1 and group 2 pneumococcus.[2] Capitalizing on this crosstalk and utilizing rational design, we were able to develop *S. mitis*-based CSP analogs capable of modulating the pneumococcal competence regulon at low nanomolar concentrations. This study demonstrated a complementary strategy to attenuate pneumococcal infections by using the native CSP signals of other related species. Furthermore, the privileged scaffolds identified in this study could potentially influence the *S. mitis* competence regulon, thereby providing novel chemical tools capable of

modulating QS in multiple species. An additional evidence of potential crosstalk between *S. mitis* and *S. pneumoniae* has been recently reported by Junges et.al.[22] In this study, the authors found that the Rgg (regulator gene of glucosyltransferase) / SHP (short hydrophobic peptide) cell-to-cell communication system in *S. mitis* activates an Rgg/SHP system in *S. pneumoniae* associated with the regulation of pneumococcal surface polysaccharide synthesis. The cross-communication between QS systems may provide important insights regarding interspecies interactions within the microbiome, and between commensal and pathogenic species. However, the activity, regulation, and possible role of the *S. mitis* competence regulon in the flow of genetic information across species remain unknown, and the ability of this QS circuitry to regulate other bacterial phenotypes has yet to be explored.

The native CSP sequence for the *S. mitis* type strain was previously deduced using genetic prediction, however its identity was not confirmed through isolation from bacterial supernatants.[23] Another study exhibited that with the use of this proposed synthetic 16-amino acid peptide, termed *S. mitis*-CSP-2, the challenges associated with genetically transforming *S. mitis* in the laboratory setting can be resolved.[3] Furthermore, several studies reported the existence of a large variety of CSP pheromones in the mitis group of streptococci, specifically among *S. mitis* strains.[3,23] We started our study by confirming the identity of the native CSP signal for our working *S. mitis* type strain (*S. mitis* ATCC 49456) by isolating it from *S. mitis* supernatants. We then set out to characterize the regulatory role of the *S. mitis* competence regulon, and to gain a deeper understanding of the molecular mechanisms that drive signal/receptor binding and subsequent activation of the QS circuitry. To this end, we first performed a full alanine and D-amino acid scans of the *S.*

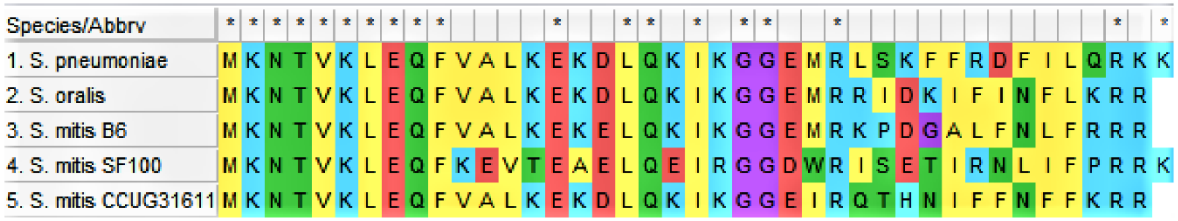
mitis-CSP-2 sequence to determine the structural motifs required for ComD binding and activation, and developed a luciferase-based reporter gene system to measure the activity profiles of all the *S. mitis*-CSP-2 analogs. We then utilized circular dichroism (CD) spectroscopy to analyze the overall structural features of all the alanine and D-amino acid *S. mitis*-CSP-2 substituted analogs. Through the structure-function insights we gained, we designed and synthesized a library of second-generation analogs and identified a potent CSP-based inhibitor of the *S. mitis* competence regulon. We also evaluated the activities of our previously designed *S. mitis*-CSP-2-based pneumococcal QS modulators[2] against the *S. mitis* QS system and identified a highly potent multi-species *S. mitis*-CSP-2-based QS activator. Finally, we utilized several phenotypic assays to evaluate the role the *S. mitis* competence regulon QS circuitry plays in modulating different pathogenic phenotypes and found that this circuitry has a regulatory role in both bacterial competence and biofilm formation.

RESULTS AND DISCUSSION

Prediction and Isolation of the *S. mitis*-CSP-2 Signal from Cell-free Supernatants. As previous studies reported the likelihood of multiple CSP variants in the mitis group of streptococci, particularly among *S. mitis* strains, we wanted to validate the predicted CSP sequence of our tested *S. mitis* strain (*S. mitis* ATCC 49456).[3,23] Based on genomic data, Salvadori et al. have previously predicted the native CSP sequence for the *S. mitis* ATCC 49456 strain, synthesized the predicted 16-amino acid CSP with the sequence EIRQTHNIFFNFFKRR, and exhibited that this synthetic peptide can induce competence development in *S. mitis* ATCC 49456.[3] We first identified the CSP-encoding gene (*comC*) of our tested *S. mitis* strain by using primers annealing to conserved Arg-tRNA and Glu-

tRNA flanking the *comCDE* operon (see **Figure 21** for primer sequences). Our sequencing results obtained through both forward and reverse primers confirmed the previously reported predicted 16-amino acid peptide sequence (**Figure 21**). Most often, after translation of the PCR product, the deduced mature CSP peptide sequence is preceded by a double glycine cleavage site. Previously, several studies have reported that to afford a mature CSP signal, ComC sometimes undergoes additional processing, either through further processing by the membrane bound ComAB transporter or through cleavage of the exported CSP signal with the help of an extracellular SepM protease.[23-25] We therefore sought to validate the identity of the mature *S. mitis* CSP by isolating the processed CSP from bacterial supernatants. Following ammonium sulfate precipitation of the excreted crude peptide mixture from cell-free *S. mitis* supernatants, we used semi-preparative RP-HPLC to fractionate the total crude mixture and observed exact masses similar to those of the predicted CSP sequence (**Figure 22**).

(A)



(B)

Primer	Sequence
SmATCCComCfwd	5′-GGCGGTGTCTTAACCCCTTGACCAACGGACC-3′
SmATCCComCrev	5′-CATAGCTCAGCTGGATAGAGCATTCGCCTTC-3′

(C)

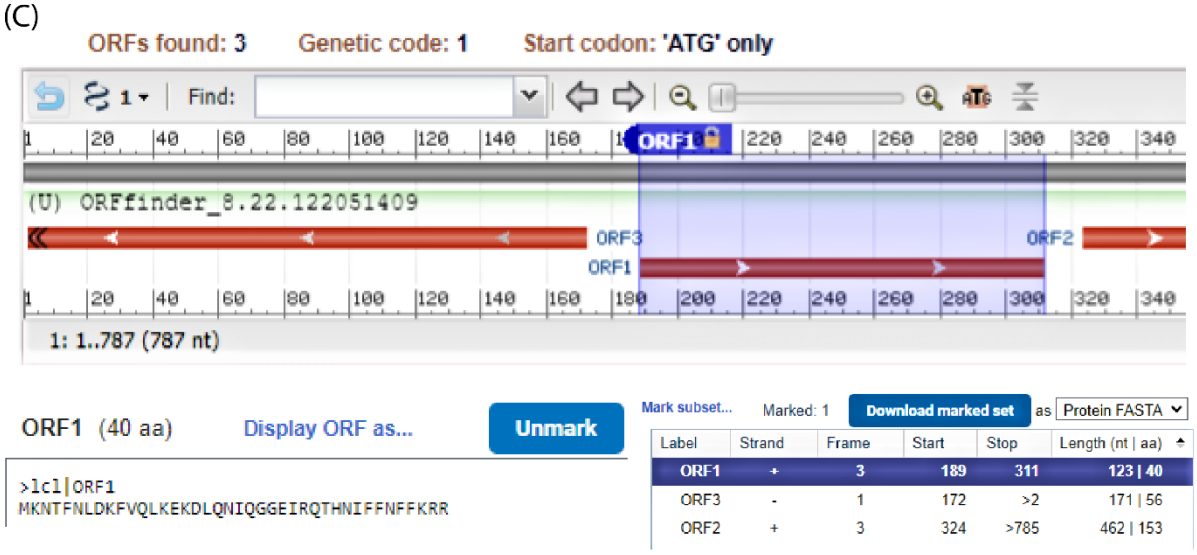


Figure 21. (**A**) ClustalW alignment of the *comC* gene product of several mitis group streptococci. ComC is typically cleaved at the double glycine site to afford a mature CSP signal with a negatively charged *N*-terminal residue. (**B**) Primers used for *comC* amplification. (**C**) Identification of *S. mitis*-CSP-2 following *comC* sequencing.

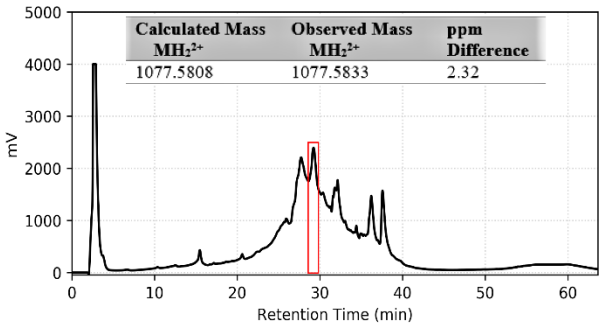
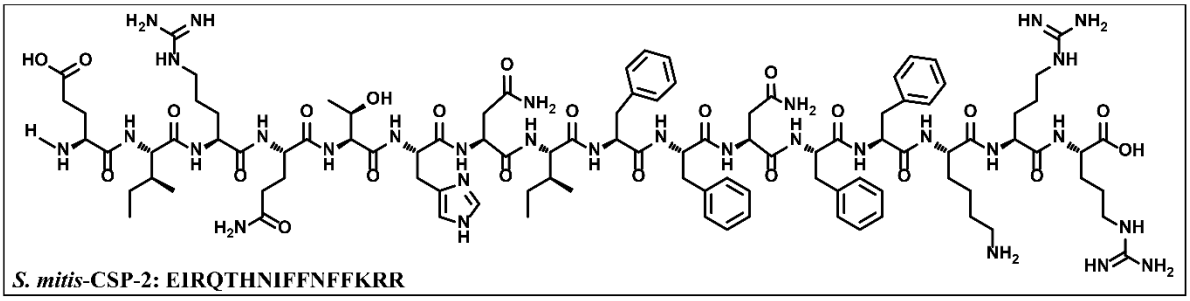


Figure 22. Isolation and detection of the *S. mitis*-CSP-2 from cell-free supernatants. The RP-HPLC chromatogram of total proteins isolated from the supernatant sample and High-Resolution ESI-TOF MS of the fraction collected from 28 to 30 min (red). See the **Supporting Information** for full experimental details.

Comparison of Isolated and Synthetic *S. mitis*-CSP-2. To further validate the identity of the *S. mitis* native CSP signal, the predicted CSP was synthesized using standard solid-phase peptide synthesis (SPPS) protocols (for full details see the **Supporting Information**). We then compared the synthetic purified peptide to the isolated purified peptide obtained from *S. mitis* cell-free supernatants. Analytical HPLC analysis of both individual peptides revealed a dominant peak with the same retention time, and a combined fraction of the synthetic and isolated peptides in a 1:1 ratio resulted in a single peak that possessed the same retention time as the individual fractions (**Figure 23**). The exact masses of the isolated and synthetic peptides were obtained and matched the expected exact mass of the 16 amino acid CSP sequence (**Figure 23**). Furthermore, MS/MS analysis of the isolated and synthetic peptides validated the connectivity of the CSP sequence (**Figures S-23-S-24 in Appendix 4**). Overall, our results confirmed the native *S. mitis* CSP sequence, which is the 16-amino acid peptide, EIRQTHNIFFNFFKRR.

(A)



(B)

S. mitis-CSP-2	Calculated Mass MH_2^{2+}	Observed Mass MH_2^{2+}	ppm Difference
Extracted	1077.5808	1077.5833	2.32
Synthetic	1077.5808	1077.5829	1.94

(C)

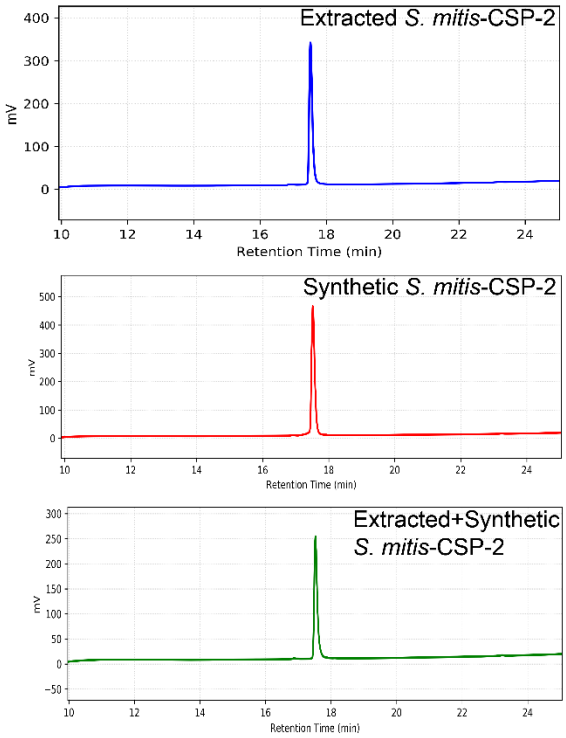


Figure 23. Comparison of purified synthetic and isolated *S. mitis*-CSP-2. (**A**) Proposed structure of the 16-amino acid *S. mitis*-CSP-2. (**B**) Comparison of observed masses of isolated and synthetic peptides by ESI-TOF MS (**C**) Comparison of analytical RP-HPLC chromatograms of the purified isolated, synthetic, and isolated and synthetic *S. mitis*-CSP-2.

Development of *S. mitis* Luciferase QS Reporter Strain. A luciferase-based *S. mitis* ATCC 49456 QS reporter strain was constructed to test the ability of the native *S. mitis*-CSP-2 and *S. mitis*-CSP-2 analogs to modulate the *S. mitis* ComD-2 receptor activity. We utilized a similar approach performed by Salvadori and co-workers to construct a

luciferase-based *S. mitis* reporter strain to directly quantify the degree of *comX* expression upon CSP/ComD binding.[3] First, the primer pair SmATCCComXfwd-SmATCCComXrev (**Table S-1 in Appendix 4**) was used to amplify the 911 base-pair (bp) *comX* promoter region from the *S. mitis* ATCC 49456 strain. Both the amplified *comX* promoter region and the plasmid pFW5-luc ($Spec^R$) were then restriction-digested using BamHI and NheI. The digested amplicon was ligated to pFW5-luc ($Spec^R$) using the T4 DNA ligase enzyme and cloned into *Escherichia coli*, as previously described by Green et al.[26] The purified pFW5-luc plasmid containing the *S. mitis comX* promoter region was then transformed into wildtype *S. mitis* ATCC 49456. Incorporation of the reporter plasmid was validated by performing colony PCR followed by sequencing. Additionally, reporter plasmid incorporation was further verified by observing an increase in luminescence following treatment of *S. mitis* reporter culture ($ATCC_{49456}$ P*comX* luc::Spc) with the native *S. mitis*-CSP-2 signal and 15 μg/mL D-luciferin, compared to a control untreated with the *S. mitis*-CSP-2 (see **Appendix 4** for full experimental details).

Design and Synthesis of First-Generation *S. mitis*-CSP-2 Analogs. To quantitatively evaluate the contribution of each side chain and chiral center of the 16-amino acid *S. mitis*-CSP-2, we conducted full alanine and D-amino acid scans of the *S. mitis*-CSP-2 sequence. All the *S. mitis*-CSP-2 analogs were constructed using standard SPPS protocols on 4-benzyloxybenzyl alcohol (Wang) resin (see **Supporting Information in Appendix 4** for full experimental details).[27] Following on-resin synthesis of each *S. mitis*-CSP-2 analog, peptides were cleaved from the solid support, purified to homogeneity (>95% purity) by semipreparative RP-HPLC and isolated in acceptable yields (see the **Supporting Information in Appendix 4** for peptide characterization details). To evaluate the activity

of each synthesized *S. mitis*-CSP-2 analog, we utilized a luciferase reporter gene assay using the *S. mitis* luciferase QS reporter strain we constructed (see **Materials and Methods** for protocol details).

Structure-Activity Relationship (SAR) Analysis of the *S. mitis*-CSP-2 Alanine-Screen Analogs. We first quantified the ability of the native *S. mitis*-CSP-2 to activate its cognate ComD receptor through determination of its EC_{50} value (half maximal effective concentration). Initial bioassays revealed that *S. mitis*-CSP-2 activates *S. mitis* ComD-2 with an EC_{50} value of 148 nM (**Table 11**). All the *S. mitis*-CSP-2 analogs were then initially screened, at high analog concentration (10,000 nM), for their ability to activate the *S. mitis* ComD-2 receptor to a level comparable to the native *S. mitis*-CSP-2 signal (see **Figure S-1 in Appendix 4**). *S. mitis*-CSP-2 analogs that revealed high receptor activation, as determined in the initial screening (>75% activation compared to *S. mitis*-CSP-2), were further assessed to determine their EC_{50} values, while *S. mitis*-CSP-2 analogs that failed to activate the receptor (<50% activation compared to *S. mitis*-CSP-2) were assessed for their ability to competitively inhibit the receptor (see **Figure S-8 in Appendix 4**).

Table 11: EC_{50} or IC_{50} Values of *S. mitis*-CSP-2 Alanine-Screen Analogs against the *S. mitis* ComD-2 Receptor[a]

Peptide Name	Peptide Sequence	EC_{50} or IC_{50}*(nM) [b]	95% CI [c]
S. mitis-CSP-2	EIRQTHNIFFNFFKRR	148	86.0-255
S. mitis-CSP-2-E1A	**A**IRQTHNIFFNFFKRR	456*	338-617
S. mitis-CSP-2-I2A	E**A**RQTHNIFFNFFKRR	---[d]	---
S. mitis-CSP-2-R3A	EI**A**QTHNIFFNFFKRR	---[d]	---
S. mitis-CSP-2-Q4A	EIR**A**THNIFFNFFKRR	227	210-246
S. mitis-CSP-2-T5A	EIRQ**A**HNIFFNFFKRR	178	108-292
S. mitis-CSP-2-H6A	EIRQT**A**NIFFNFFKRR	>1000	---
S. mitis-CSP-2-N7A	EIRQTH**A**IFFNFFKRR	114	74.3-175
S. mitis-CSP-2-I8A	EIRQTHN**A**FFNFFKRR	605	558-655
S. mitis-CSP-2-F9A	EIRQTHNI**A**FNFFKRR	>1000	---
S. mitis-CSP-2-F10A	EIRQTHNIF**A**NFFKRR	579	497-675
S. mitis-CSP-2-N11A	EIRQTHNIFF**A**FFKRR	46.3	41.0-53.0
S. mitis-CSP-2-F12A	EIRQTHNIFFN**A**FKRR	121	92.4-158
S. mitis-CSP-2-F13A	EIRQTHNIFFNF**A**KRR	142	112-180
S. mitis-CSP-2-K14A	EIRQTHNIFFNFF**A**RR	163	143-187
S. mitis-CSP-2-R15A	EIRQTHNIFFNFFK**A**R	303	248-370
S. mitis-CSP-2-R16A	EIRQTHNIFFNFFKR**A**	251	184-344

[a] See the Materials and Methods section for experimental details and the Appendix 4 for details of the reporter strain and plots of agonism or antagonism dose-response curves. All assays were performed in triplicate. [b] EC_{50} or IC_{50} values were determined by testing peptides over a wide range of concentrations. [c] 95% confidence interval. [d] EC_{50} not determined due to the analog's low induction in primary agonism screening assay. See the Supporting Information for details.

The reporter gene data revealed several interesting SAR trends regarding *S. mitis*-CSP-2:ComD-2 binding and consequently ComD-2 activation. The *S. mitis*-CSP-2 can be divided into three distinct regions, the *N*-terminus (first three amino acid residues), the central region (residues 4-13), and the *C*-terminus (last three amino acid residues, residues 14-16). Alanine screening of the *S. mitis*-CSP-2 *N*-terminus revealed that the first three residues (Glu1, Ile2, and Arg3) have a critical role in *S. mitis* ComD-2 receptor binding and activation. Alanine replacements of Ile2 and Arg3 resulted in a complete loss of activity, whereas alanine replacement of Glu1 resulted in the conversion of the *S. mitis*-CSP-2 into a ComD-2 inhibitor (**Table 11**). These results suggest that the *N*-terminus of *S. mitis*-CSP-2 participates in very specific binding interactions that cannot withstand significant side-chain modifications. Our results of the *N*-terminus region of *S. mitis*-CSP-2 are consistent with what was previously observed in other members of the mitis group of

streptococci, such as *S. pneumoniae* and *S. oligofermentans*, where the first negatively charged amino acid (Glu or Asp) was found to play a critical role in receptor activation while the third residue (Arg) was found to play a critical role in receptor binding, suggesting the presence of highly conserved motifs at these two positions.[28,25]

In the *S. mitis*-CSP-2 core region, the alanine scan results revealed that changes in residues H6, I8, F9 and F10 resulted in a significant reduction in potency, while modification in residues Q4, T5, N7, N11, F12 and F13 resulted in analogs with similar activities to the native CSP (**Table 11**). Specifically, the alanine substitution results suggest that the H6 and F9 side chains are either critical for receptor binding, or are key to stabilizing the peptide bioactive conformation. Thus, modifications to either one of these residues resulted in the most significant loss in potency (**Table 11**). In contrast, alanine replacement of N11 resulted in an analog with increased potency (EC_{50} = 46.3 nM) compared to the native signal (EC_{50} = 148 nM), suggesting that the removal of this side chain either directly improved the binding interactions with the ComD-2 receptor (likely through elimination of electrostatic or steric clashes), or allowed the peptide to assume a more favorable conformation for receptor binding.

As for the *S. mitis*-CSP-2 *C*-terminal region (residues K14, R15, and R16), alanine mutation resulted in analogs with similar activities to the parent *S. mitis*-CSP-2 signal (**Table 11**), suggesting that these residues are not critical for receptor binding or activation. These findings are in agreement with the results obtained in other members of the mitis group of streptococci. Yet, since these three residues are conserved in most mitis group of streptococci CSPs, it is likely that these residues have an important evolutionary role in the competence regulon QS circuitry. Indeed, it was hypothesized that these three positively

charged residues increase the solubility of these overall relatively hydrophobic peptide sequences, thereby facilitating their departure from the cell membrane to the extracellular environment following export.

SAR Analysis of the *S. mitis*-CSP-2 D-amino acid Scan Analogs. Similarly to the *S. mitis*-CSP-2 *N*-terminus alanine substitutions, examination of the D-amino acid scan of the *S. mitis*-CSP-2 *N*-terminus further confirmed the importance of the first three residues in ComD-2 receptor binding and activation. Alanine and D-amino acid replacement of Ile2 and Arg3 resulted in complete loss of activity. However, in the case of Glu1, while alanine substitution resulted in the conversion of *S. Mitis*-CSP-2 into a potent ComD-2 inhibitor, D-amino acid replacement completely abolished the peptide activity (see both **Table 11** and **Table 12**). These results suggest that the *N*-terminus of the *S. mitis*-CSP-2 cannot accommodate chirality alterations, likely due to the inability of the differently oriented side chains to fit into their respective receptor binding pockets.

Table 12: EC_{50} Values of *S. mitis*-CSP-2 D-Amino Acid Scan Analogs against the *S. mitis* ComD-2 Receptor[a]

Peptide Name	Peptide Sequence	EC_{50} (nM) [b]	95% CI [c]
S. mitis-CSP-2	EIRQTHNIFFNFFKRR	148	86.0-255
S. mitis-CSP-2-e1	**e**IRQTHNIFFNFFKRR	---[d]	---
S. mitis-CSP-2-i2	E**i**RQTHNIFFNFFKRR	---[d]	---
S. mitis-CSP-2-r3	EI**r**QTHNIFFNFFKRR	---[d]	---
S. mitis-CSP-2-q4	EIR**q**THNIFFNFFKRR	>1000	---
S. mitis-CSP-2-t5	EIRQ**t**HNIFFNFFKRR	>1000	---
S. mitis-CSP-2-h6	EIRQT**h**NIFFNFFKRR	>1000	---
S. mitis-CSP-2-n7	EIRQTH**n**IFFNFFKRR	>1000	---
S. mitis-CSP-2-i8	EIRQTHN**i**FFNFFKRR	90.0	49.0-165
S. mitis-CSP-2-f9	EIRQTHNI**f**FNFFKRR	438	328-583
S. mitis-CSP-2-f10	EIRQTHNIF**f**NFFKRR	64.0	42.2-96.4
S. mitis-CSP-2-n11	EIRQTHNIFF**n**FFKRR	38.3	22.4-66.0
S. mitis-CSP-2-f12	EIRQTHNIFFN**f**FKRR	245	152-394
S. mitis-CSP-2-f13	EIRQTHNIFFNF**f**KRR	338	155-735
S. mitis-CSP-2-k14	EIRQTHNIFFNFF**k**RR	95.0	47.3-189
S. mitis-CSP-2-r15	EIRQTHNIFFNFFK**r**R	143	73.1-278
S. mitis-CSP-2-r16	EIRQTHNIFFNFFKR**r**	64.0	59.0-69.0

[a] See the Materials and Methods section for experimental details and Appendix 4 for details of the reporter strain and plots of agonism or antagonism dose-response curves. All assays were performed in triplicate. [b] EC_{50} values were determined by testing peptides over a wide range of concentrations. [c] 95% confidence interval. [d] EC_{50} not determined due to the analog's low induction in primary agonism screening assay. See the Appendix 4 for details.

Looking at the hydrophilic residues at the central region of *S. mitis*-CSP-2 (Gln4, Thr5, His6, Asn7, and Asn11, **Table 12**), it appears that side chain orientation changes are less tolerated than alanine mutations, as, with the exception of Asn11, D-amino acid substitution analogs exhibited significant loss of activity. Thus, we concluded that in these cases the side chain orientation and concomitantly induced local peptide conformation is more important than the actual side chain residue. Interestingly, like the alanine substitution, the D-amino acid replacement of Asn11 resulted in an analog with improved potency (EC_{50} = 38.3 nM, **Table 12**), producing the most potent analog in these two libraries. Since position 11 of *S. mitis*-CSP-2 tolerates well both chirality and side chain changes, it is likely not directly participating in important receptor binding interactions.

Our D-amino acid scan analysis of the hydrophobic residues at the core region of *S. mitis*-CSP-2 (I8, F9, F10, F12, and F13) revealed an opposite trend than the one for the hydrophilic residues. That is, in the case of the hydrophobic residues, the side chain identity is more important than the actual side chain orientation, as most of the D-amino acid substitutions in these positions yielded analogs with similar or only slightly reduced activities compared to the native *S. mitis*-CSP-2 signal, whereas most alanine modifications in these positions resulted in significant reduction in potency (see both **Table 11** and **Table 12**). The results of both the alanine and D-amino acid replacements suggest that the Phe12 and Phe13 residues do not play an important role in *S. mitis* ComD-2 receptor binding.

Lastly, chirality changes at the *C*-terminal residues of *S. mitis*-CSP-2 confirmed that these three positions (Lys14, Arg15, and Arg16) are highly modifiable, as D-amino acid substitutions resulted in analogs with similar (Arg15) or even better activities (Lys14 and Arg16) compared to the native *S. mitis*-CSP-2 signal.

Second-Generation CSP Analogs. Our initial alanine and D-amino acid substitutions analysis revealed several interesting activity trends. For instance, our alanine scan analysis revealed that alanine replacements of Asn7, Asn11, Phe12, and Phe13 resulted in analogs with improved activities compared to the native peptide (**Table 11**). Furthermore, replacement of the negatively charged glutamic acid at position 1 with alanine resulted in an analog that exhibits competitive inhibition against the *S. mitis* ComD-2 receptor (**Table 11**), a trend consistent with other mitis group streptococci. Similarly, the D-amino acid scan revealed that replacements of Ile8, Phe10, Asn11, Lys14, and Arg16 with their enantiomers resulted in analogs with enhanced activities compared to native *S. mitis*-CSP-2 (**Table 12**). Therefore, to further explore these trends and identify enhanced CSP-based QS modulators of the *S. mitis* competence regulon, we designed and synthesized a second-generation library of *S. mitis*-CSP-2 analogs that included both sequential truncations of either the three *N*-terminal residues or the three *C*-terminal residues of the native CSP signal (**Table 13**), as well as multiple-modification analogs combining the E1A modification with the modifications that led to enhanced activators (**Table 14**).

Table 13: EC_{50} or IC_{50} Values of *S. mitis*-CSP-2 Truncated Analogs against the *S. mitis* ComD-2 Receptor[a]

Peptide Name	Peptide Sequence	EC_{50} or IC_{50}*(nM) [b]	95% CI [c]
S. mitis-CSP-2	EIRQTHNIFFNFFKRR	148	86.0-255
S. mitis-CSP-2-des-E1	IRQTHNIFFNFFKRR	>1000*	---
S. mitis-CSP-2-des-E1I2	RQTHNIFFNFFKRR	---[d]	---
S. mitis-CSP-2-des-E1I2R3	QTHNIFFNFFKRR	---[d]	---
S. mitis-CSP-2-des-R16	EIRQTHNIFFNFFKR	39.2	21.4-72.0
S. mitis-CSP-2-des-R15R16	EIRQTHNIFFNFFK	242	185-315
S. mitis-CSP-2-des-K14R15R16	EIRQTHNIFFNFF	---[d]	---

[a] See the Materials and Methods section for experimental details and Appendix 4 for details of the reporter strain and plots of agonism or antagonism dose-response curves. All assays were performed in triplicate. [b] EC_{50} or IC_{50} values were determined by testing peptides over a wide range of concentrations. [c] 95% confidence interval. [d] EC_{50} not determined due to the analog's low induction in primary agonism screening assay. See Appendix 4 for details.

Truncation of the first amino acid at the peptide *N*-terminus resulted in a weak QS inhibitor (*S. mitis*-CSP-2-des-E1, IC_{50} > 1000 nM, **Table 13**). Removal of any additional residue

from the *N*-terminus resulted in complete loss of activity, highlighting the importance of the *S. mitis*-CSP-2 *N*-terminus in both receptor binding and activation. As for the *C*-terminus, removal of the last *C*-terminal residue produced a peptide with a four-fold enhanced activity compared to the native peptide (compare the EC_{50} value of *S. mitis*-CSP-2-des-R16, 39.2 nM, with that of *S. mitis*-CSP-2, 148 nM, **Table 13**). Furthermore, truncation of the last two residues at the *C*-terminus resulted in an analog with comparable activity to the native signal. Together these results suggest that the last two *C*-terminal residues within the *S. mitis*-CSP-2 scaffold do not play a major role in the peptide activity. Contrary, truncation of three residues at the *C*-terminus, to afford *S. mitis*-CSP-2-des-K14R15R16, has led to a complete loss of activity (**Table 13**), revealing the minimal sequence required for effective *S. mitis* ComD-2 binding. Of note, truncation of either two or three *C*-terminal residues produced analogs that were difficult to dissolve in aqueous solutions, supporting the hypothesis that the purpose of the positively charged *S. mitis*-CSP-2 *C*-terminus is to enhance peptide solubility.

Biological evaluation of the four double-alanine mutation analogs, *S. mitis*-CSP-2-E1AN7A, *S. mitis*-CSP-2-E1AN11A, *S. mitis*-CSP-2-E1AF12A, and *S. mitis*-CSP-2-E1AF13A, revealed four *S. mitis* ComD-2 inhibitors (**Table 14**). However, when comparing the bioactivities of the resulting double mutation inhibitors with the corresponding initial single mutation analogs, it appears that there is no correlation between EC_{50} and IC_{50} values. Although alanine replacement of all these residues, with the exception of the E1A modification, yielded potent *S. mitis* ComD-2 activators (**Table 11**), only the double mutant bearing the N7A modification combined with the E1A modification generated a more potent *S. mitis* ComD-2 inhibitor (IC_{50} = 242 nM) compared to the E1A

substitution alone (IC_{50} = 456 nM). The remaining three analogs were only weak inhibitors (**Table 14**).

Table 14: IC_{50} Values of *S. mitis*-CSP-2-E1A Modification Analogs against the *S. mitis* ComD-2 Receptor[a]

Peptide Name	Peptide Sequence	IC_{50} (nM) [b]	95% CI [c]
S. mitis-CSP-2-E1A	**A**IRQTHNIFFNFFKRR	456	338-617
S. mitis-CSP-2-E1AN7A	**A**IRQTH**A**IFFNFFKRR	242	139-422
S. mitis-CSP-2-E1AN11A	**A**IRQTHNIFF**A**FFKRR	513	322-819
S. mitis-CSP-2-E1AF12A	**A**IRQTHNIFFN**A**FKRR	>1000	---
S. mitis-CSP-2-E1AF13A	**A**IRQTHNIFFNF**A**KRR	>1000	---
S. mitis-CSP-2-E1Ai8	**A**IRQTHN**i**FFNFFKRR	---[d]	---
S. mitis-CSP-2-E1Af10	**A**IRQTHNIF**f**NFFKRR	372	215-643
S. mitis-CSP-2-E1An11	**A**IRQTHNIFF**n**FFKRR	258	121-548
S. mitis-CSP-2-E1Ak14	**A**IRQTHNIFFNFF**k**RR	159	106-238
S. mitis-CSP-2-E1Ar16	**A**IRQTHNIFFNFFKR**r**	236	147-380
S. mitis-CSP-2-E1A-des-R16	**A**IRQTHNIFFNFFKR	347	208-579
S. mitis-CSP-2-E1AN7Af10	**A**IRQTH**A**IF**f**NFFKRR	172	123-240
S. mitis-CSP-2-E1AN7An11	**A**IRQTH**A**IFF**n**FFKRR	197	127-306
S. mitis-CSP-2-E1AN7Ak14	**A**IRQTH**A**IFFNFF**k**RR	440	301-642
S. mitis-CSP-2-E1AN7Ar16	**A**IRQTH**A**IFFNFFKR**r**	290	151-555
S. mitis-CSP-2-E1Af10n11	**A**IRQTHNIF**fn**FFKRR	>1000	---
S. mitis-CSP-2-E1Af10k14	**A**IRQTHNIF**f**NFF**k**RR	521	380-715
S. mitis-CSP-2-E1Af10r16	**A**IRQTHNIF**f**NFFKR**r**	87.3	50.0-154
S. mitis-CSP-2-E1An11k14	**A**IRQTHNIFF**n**FF**k**RR	248	148-416
S. mitis-CSP-2-E1An11r16	**A**IRQTHNIFF**n**FFKR**r**	238	135-420
S. mitis-CSP-2-E1Ak14r16	**A**IRQTHNIFFNFF**k**R**r**	661	537-814

[a] See the Materials and Methods section for experimental details and Appendix 4 for details of the reporter strain and plots of antagonism dose-response curves. All assays were performed in triplicate. **[b]** IC_{50} values were determined by testing peptides over a wide range of concentrations. **[c]** 95% confidence interval. **[d]** IC_{50} not determined due to the analog's low activity in primary antagonism screening assay. See Appendix 4 for details.

A different trend was observed for the double mutation analogs generated from the combination of the potent activators revealed through the D-amino acid replacement and the E1A mutation. In this case all the resultant analogs, with the exception of *S. mitis*-CSP-2-E1Ai8, exhibited *S. mitis* ComD-2 inhibition with higher potency compared to the E1A substitution alone (**Table 14)**. Interestingly, D-amino acid replacement of residue Ile8 displayed two-fold higher agonistic activity compared to the native signal (**Table 12)**, whereas the combination with the E1A substitution resulted in an analog with complete loss of activity. Overall, our results suggest that there are different requirements for

receptor inhibition compared to receptor activation. These results are consistent with results obtained in other members of the mitis group of streptococci demonstrating that direct conversion of potent activators into potent inhibitors is not necessarily straightforward and require additional fine-tuning.[18]

Our sequential truncation analysis revealed that removal of R16 produced a peptide that exhibited a 4-fold enhanced activity compared to that of the native peptide (**Table 13**), suggesting that this side chain residue may in fact cause some steric clashes that reduces the potency of the native CSP. Thus, we incorporated this truncation together with the E1A modification to afford *S. mitis*-CSP-2-E1A-des-R16 (**Table 14).** This analog exhibited similar *S. mitis* ComD-2 inhibitory activity to that of the analog containing only the E1A modification, suggesting that for *S. mitis* ComD-2 inhibition, the presence of this side chain residue is not detrimental and can thus be incorporated for the design of potent inhibitors of the *S. mitis* competence regulon.

To produce enhanced inhibitors of the competence regulon, a library of tri-substituted analogs was designed and constructed. To this end, the E1A modification was combined with two of the following modifications: alanine replacement of position 7, or D-amino acid replacement at positions 10, 11, 14, or 16, which were found to be the most beneficial in the doubly-substitution library (**Table 14**). As expected, all the tri-substituted analogs exhibited *S. mitis* ComD-2 inhibitory activity. Analogs containing the E1A modification together with the N7A substitution and D-amino acid replacement of either position 10, 11, or 16 resulted in three analogs, *S. mitis*-CSP-2-E1AN7Af10, *S. mitis*-CSP-2-E1AN7An11, and *S. mitis*-CSP-2-E1AN7Ar16, that exhibited a 2- to 3-fold improvement in inhibitory potency compared to the E1A substitution alone, whereas the combination of the E1A,

N7A and D-amino acid substitution at position 14 produced an analog, *S. mitis*-CSP-2-E1AN7Ak14, that exhibited comparable activity to that of the E1A substitution alone (**Table 14**). The combination of D-amino acid substitution at position Phe10 and D-amino acid in either Asn11 or Lys14, together with the E1A modification, resulted in significant decrease in inhibitory potency (*S. mitis*-CSP-2-E1Af10n11, IC_{50} > 1000 nM, and *S. mitis*-CSP-2-E1Af10k14, IC_{50} = 521 nM; **Table 14**). Contrary, the combination of D-amino acid at positions Phe10 and Arg16, together with the E1A modification resulted in an analog, *S. mitis*-CSP-2-E1Af10r16, that exhibited a 5-fold increase in potency compared to the E1A modification alone, yielding the most potent inhibitor identified in this study (IC_{50} = 87.3 nM, **Table 14**). Overall, several potent *S. mitis* ComD-2 inhibitors were discovered through the multiple mutation library.

Structural Analysis Using Circular Dichroism (CD) Spectroscopy. Though our SAR analysis of the *S. mitis*-CSP-2 revealed the importance of several residues in receptor binding or activation, this analysis was not enough to explain how these different modifications affected the overall peptide structure. To gain insight as to how the different modifications affected the overall peptide structure and correlate it to bioactivity, we set out to assess the structures of the native *S. mitis*-CSP-2 pheromone and its alanine and D-amino acid scan analogs using CD spectroscopy. We assessed all the synthetic *S. mitis*-CSP-2 analogs under both aqueous (phosphate-buffered saline (PBS) buffer, pH 7.4) (**Figure S-14 to S-17 in Appendix 4**) and membrane-mimicking conditions (20% trifluoroethanol (TFE) in PBS, pH 7.4) (**Figure S-18 to S-21 in Appendix 4**). In both aqueous and membrane mimicking conditions, the native *S. mitis*-CSP-2 was found to be unstructured, exhibiting a random coil characteristic. Like the native pheromone, most of

the alanine and D-amino acid scan analogs were found to be unstructured, exhibiting a random coil pattern in both aqueous and membrane mimicking conditions, with the exception of a few analogs that exhibited some β-sheet pattern. These results were surprising, as for CSPs produced by other members of the mitis group of streptococci, including *S. pneumoniae* and *S. oligofermentans*, an α-helix conformation was proposed to be important for biological activity.[28,25]

Identifying *S. mitis*-CSP-2-Based Multi-Species QS Modulators. In a previous study, we sought to uncover potential crosstalk between streptococci species by evaluating the ability of native streptococci CSPs to modulate the *S. pneumoniae* competence regulon.[2] Our analysis revealed a potential role of *S. mitis* in modulating QS in *S. pneumoniae*, as the native *S. mitis*-CSP-2 pheromone was found to activate both pneumococcal ComD receptors (ComD1 and ComD2). We therefore sought to evaluate whether native signals from closely related streptococci species, namely CSP signals produced by members of the mitis and anginosus groups of streptococci, are able to modulate the *S. mitis* competence regulon. Our cell-based luciferase assay revealed that only one native signal produced by *S. pneumoniae*, termed CSP1, is capable of activating the *S. mitis* ComD-2 receptor, albeit at high peptide concentration (EC_{50} = 3020 nM, **Table 15**; see the **Supporting Information** in **Appendix 4** for initial screening).

In our previous work, we utilized the *S. mitis*-CSP-2 scaffold to develop analogs capable of modulating the two pneumococcal ComD receptors at low nanomolar concentrations.[2] We wanted to evaluate whether the optimization of the *S. mitis*-CSP-2 scaffold to the pneumococcus ComD receptors affected the ability of the modified analogs to interact with the *S. mitis* ComD-2 receptor. We therefore set out to test the 20 most potent pneumococcus

activators identified in our previous study against the *S. mitis* ComD-2 receptor (**Table 15**). Our analysis revealed that several potent *S. mitis*-CSP-2-based activators of the pneumococcal QS circuitry are also able to activate the *S. mitis* competence regulon with high potency. In addition to identifying QS modulators that are capable of activating both the pneumococcal and the mitis competence regulon, our analysis also highlighted several important SAR insights regarding ComD receptor binding in both *S. pneumoniae* and *S. mitis*. For example, the singly substituted *S. mitis*-CSP-2-based analog, *S. mitis*-CSP-2-I2M, exhibited high activity against both pneumococcal ComD receptors (EC_{50} values of 87.7 nM and 136 nM against ComD1 and ComD2, respectively, **Table 15**).[2] However, this mutation resulted in reduced activity of this analog against the *S. mitis* ComD-2 receptor (EC_{50} = 591 nM, **Table 15)**. The same trend was observed for all the other *S. mitis*-CSP-2-based analogs bearing the I2M mutation that were evaluated in this study. These results are in line with previous SAR studies and highlight the importance of the side chain residue in position 2 in both pneumococcal and mitis CSPs for receptor binding and activation. These results also suggest that position 2 has an important role in conferring receptor specificity between the *S. pneumoniae* and *S. mitis* native CSP signals. Importantly, our analysis revealed a highly potent QS activator, *S. mitis*-CSP-2-N7II8F, capable of activating both the pneumococcus ComD receptors and the *S. mitis* ComD-2 receptor at low nanomolar concentrations (EC_{50} values of 87.2 nM and 22.8 nM against the pneumococcal ComD1 and ComD2 receptors, respectively,[2] and EC_{50} value of 98.3 against the *S. mitis* ComD-2 receptor, **Table 15**). To the best of our knowledge, this is the first example of a CSP-based multi-species QS activator, highlighting the potential of these privileged scaffolds as tools

to study the role of QS in the competition between bacterial species in complex bacterial milieu.

Contrary to the *S. mitis*-CSP-2-based pneumococcal QS activators, none of the previously designed *S. mitis*-CSP-2-based pneumococcal QS inhibitors were found to effectively inhibit the *S. mitis* ComD-2 receptor. From 11 *S. mitis*-CSP-2-based potential pneumococcal inhibitors, only *S. mitis*-CSP-2-E1AN11FF12L was found to weakly inhibit the *S. mitis* ComD-2 receptor (**Table 15)**. However, this analog exhibited no inhibitory activity against the *S. pneumoniae* ComD1 and ComD2 receptors (**Table 15**),[2] highlighting the different structural requirements for receptor inhibition compared to receptor activation in different streptococcal species.

Table 15: EC_{50} or IC_{50} Values of Select *S. mitis*-CSP-2-based pneumococcal QS Modulators against the pneumococcus ComD1 and ComD2, and the *S. mitis* ComD-2 Receptors[a]

Peptide Name	Peptide Sequence	*S. mitis*		*S. pneumoniae*[e]	
		ComD-2		ComD1	ComD2
		EC_{50} or IC_{50}* (nM) [b]	95% CI[c]	EC_{50} or IC_{50}* (nM)	EC_{50} or IC_{50}* (nM)
S. mitis-CSP-2	EIRQTHNIFFNFFKRR	148	86.0-255	663	635
S. pneumoniae-CSP1	EMRLSKFFRDFILQRKK	3020	1550-5870	10.3	526
S. pneumoniae-CSP2	EMRISRIILDFLFLRKK	---	---	1650	50.7
S. mitis-CSP-2-I2M	E**M**RQTHNIFFNFFKRR	591	385-908	87.7	136
S. mitis-CSP-2-Q4L	EIR**L**THNIFFNFFKRR	54.2	25.0-118	88.2	252
S. mitis-CSP-2-I8F	EIRQTHN**F**FFNFFKRR	470	234-945	128	209
S. mitis-CSP-2-N11F	EIRQTHNIFF**F**FKRR	47.1	22.0-102	4.63	220
S. mitis-CSP-2-N7I	EIRQTH**I**IFFNFFKRR	105	57.0-194	101	211
S. mitis-CSP-2-I2MQ4L	E**M**R**L**THNIFFNFFKRR	332	212-522	49.9	141
S. mitis-CSP-2-I2MI8F	E**M**RQTHN**F**FFNFFKRR	---[d]	---	151	83.2
S. mitis-CSP-2-I2MF12L	E**M**RQTHNIFFN**L**FKRR	738	475-1150	25.8	146
S. mitis-CSP-2-Q4LF12L	EIR**L**THNIFFN**L**FKRR	201	129-315	68.7	202
S. mitis-CSP-2-N7II8F	EIRQTH**IF**FFNFFKRR	98.3	59.2-163	87.2	22.8
S. mitis-CSP-2-N11FF12L	EIRQTHNIFF**FL**FKRR	289	202-415	4.97	127
S. mitis-CSP-2-I2MQ4LN7F	E**M**R**L**TH**F**IFFNFFKRR	>1000	---	137	75.6
S. mitis-CSP-2-I2MI8FN11F	E**M**RQTHN**F**FF**F**FFKRR	---[d]	---	6.95	26.2
S. mitis-CSP-2-I2MN7FF12L	E**M**RQTH**F**IFFN**L**FKRR	>1000	---	72.8	112
S. mitis-CSP-2-I2MN7II8F	E**M**RQTH**IF**FFNFFKRR	>1000	---	61.6	2.67
S. mitis-CSP-2-I2MQ4LF12L	E**M**R**L**THNIFFN**L**FKRR	878	688-1120	17.1	139
S. mitis-CSP-2-I2MQ4LI8F	E**M**R**L**THN**F**FFNFFKRR	---[d]	---	42.0	30.0
S. mitis-CSP-2-I2MQ4LN11F	E**M**R**L**THNIFF**F**FFKRR	>1000	---	14.8	188
S. mitis-CSP-2-I2MN7FI8F	E**M**RQTH**FF**FFNFFKRR	>1000	---	63.6	13.5
S. mitis-CSP-2-N7FI8FF12L	EIRQTH**FF**FFN**L**FKRR	548	434-693	101	67.6
S. mitis-CSP-2-E1AN11FF12L	**A**IRQTHNIFF**FL**FKRR	>1000*	---	---[d]	---[d]

[a] See the Materials and Methods section for experimental details and Appendix 4 for details of the reporter strain and plots of agonism or antagonism dose-response curves. All assays were performed in triplicate. **[b]** EC_{50} or IC_{50} values were determined by testing peptides over a wide range of concentrations. **[c]** 95% confidence interval. [d] EC_{50} not determined due to the analog's low induction in primary agonism screening assay. See Appendix 4 for details. **[e]** Data from ref. 2.

Phenotypic Evaluation. Following the confirmation of the native CSP sequence for the *S. mitis* ATCC 49456 strain, we wanted to evaluate the regulatory role of this native CSP signal and the competence regulon QS circuitry. We focused our analysis on three phenotypes: competence, biofilm formation, and virulence factor production, as these three

phenotypes have been found to be regulated by the competence regulon QS circuitry in several other streptococcal species.

The first phenotype evaluated was competence induction. To this end, an antibiotic resistance transformation assay was conducted (see **Supporting Information** in **Appendix 4** for details), where a spectinomycin resistance plasmid (pFW5-luc) was introduced to wildtype *S. mitis* ATCC 49456 and the effectiveness of *S. mitis*-CSP-2, a lead *S. mitis*-CSP-2-based QS activator (*S. mitis*-CSP-2-n11) and a lead *S. mitis*-CSP-2-based inhibitor (*S. mitis*-CSP-2-E1Af10r16) in modulating this phenotype was evaluated. Following a 4-hour incubation period with the plasmid and native or synthetic CSPs, the bacteria were spread-plated on THY agar plates containing 5% horse serum and spectinomycin at a final concentration of 200 μg/mL, and competence induction was assessed through colony formation. *S. mitis* ATCC 49456 incubated with the pFW5-luc plasmid alone (no exogenous CSP addition) resulted in no apparent transformants (**Figure 24, left**). As expected, incubation with the native *S. mitis*-CSP-2 resulted in many transformants, suggestive of competence induction (**Figure 24, top**). Our lead synthetic activator, *S. mitis*-CSP-2-n11, was also able to effectively induce competence, leading to many transformants (**Figure 24, right**). Conversely, the addition of native *S. mitis*-CSP-2 together with our lead inhibitor, *S. mitis*-CSP-2-E1Af10r16, did not yield any observed transformants, suggesting that our lead inhibitor was able to block *S. mitis*-CSP-2-mediated competence induction (**Figure 24, bottom**).

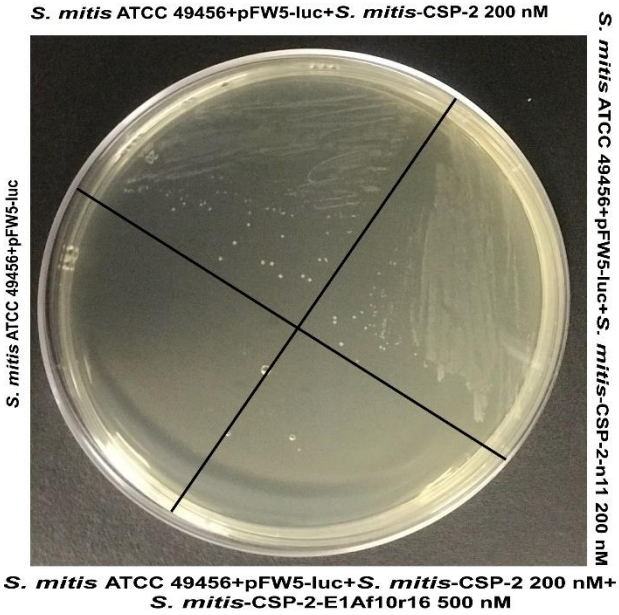


Figure 24. Transformation assay of *S. mitis* ATCC 49456 in the presence of *S. mitis*-CSP-2, *S. mitis*-CSP-2-n11, and *S. mitis*-CSP-2 together with *S. mitis*-CSP-2-E1Af10r16. The ability of ATCC 49456 to internalize a spectinomycin resistance plasmid (pFW5-luc) was evaluated following treatment with *S. mitis*-CSP-2, *S. mitis*-CSP-2-n11, and *S. mitis*-CSP-2 + *S. mitis*-CSP-2-E1Af10r16. Following treatment with *S. mitis*-CSP-2 or *S. mitis*-CSP-2-n11, ATCC 49456 was able to internalize spectinomycin resistance, as can be seen by the number of transformants (top and right, respectively). In contrast, ATCC 49456 was unable to internalize spectinomycin resistance following treatment with *S. mitis*-CSP-2 + *S. mitis*-CSP-2-E1Af10r16, as determined by the lack of apparent transformants (bottom). The incubation of *S. mitis* ATCC 49456 with the pFW5-luc plasmid without the addition of synthetic CSP was used as a negative control (left). The experiment was repeated three times in triplicate for a total of nine experiments. See the Supporting Information for full experimental details.

Next, we sought to determine whether biofilm formation, a trait associated with enhanced bacterial pathogenesis, was also under the control of the *S. mitis* competence regulon. To this end, we conducted a crystal violet biofilm quantification assay (see **Supporting Information in Appendix 4** for experimental details) in the presence of the native *S. mitis*-CSP-2, our lead activator, *S. mitis*-CSP-2-n11, our lead inhibitor, *S. mitis*-CSP-2-E1Af10r16, or *S. mitis*-CSP-2 together with our lead inhibitor, *S. mitis*-CSP-2-E1Af10r16. Our results revealed that exogenous addition of either the native *S. mitis*-CSP-2 or the lead activator did not significantly alter the amount of biofilm formed compared to untreated *S. mitis* ATCC 49456 (**Figure 25**). Contrary, regardless of whether exogenous *S. mitis*-CSP-

2 was included or not, inclusion of the lead inhibitor, *S. mitis*-CSP-2-E1Af10r16, at a concentration 5-fold greater than the IC_{50} value (IC_{50} = 87.3 nM, **Table 15**) resulted in a statistically significant decrease in the amount of biofilm formed compared to the untreated control (**Figure 25**). These results demonstrate the ability of our lead inhibitor to dramatically reduce *S. mitis* biofilms, even when a sufficient concentration of native *S. mitis*-CSP-2 signal is present to induce a QS response.

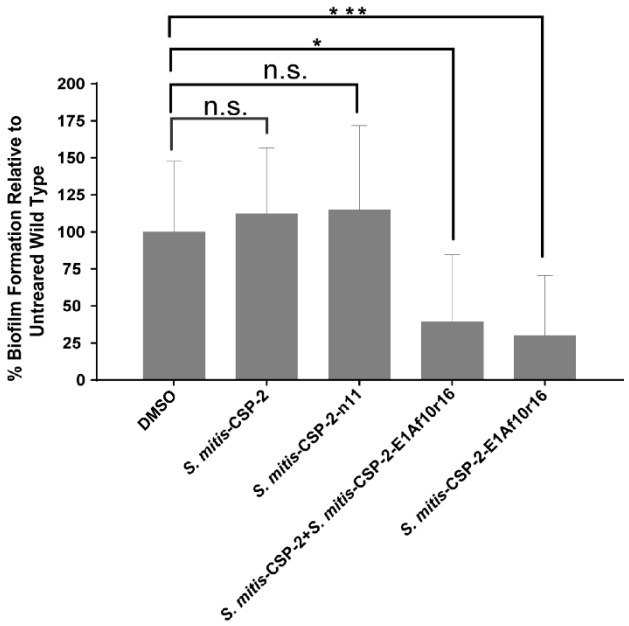


Figure 25. Biofilm formation of *S. mitis* ATCC 49456 in the presence of *S. mitis*-CSP-2, *S. mitis*-CSP-2-n11, *S. mitis*-CSP-2 together with *S. mitis*-CSP-2-E1Af10r16, and *S. mitis*-CSP-2-E1Af10r16 alone. The lead inhibitor, *S. mitis*-CSP-2-E1Af10r16, with or without exogenous addition of the native *S. mitis*-CSP-2 signal, was found to decrease the amount of biofilm formed compared to the untreated control. Statistical significance was determined using a one-way ANOVA with Bonferroni's correction; n.s., not significant; *P $\leq$ 0.1; ***P $\leq$ 0.001. The experiment was repeated three times in triplicate for a total of nine experiments. See Appendix 4 for full experimental details.

Lastly, we set out to assess the cytotoxicity of both *S. mitis* and the lead *S. mitis*-CSP-2 analogs. *S. mitis* predominantly colonizes and persists in the human mouth and other mucosal surfaces, and was found to have the ability to induce the production of oral mucosal antibodies. These unique biological features categorize *S. mitis* as a potential

mucosal vaccine or therapeutic delivery vector.[10] Since previously it was reported that many streptococci species produce virulence factors associated with cytotoxicity,[29] we evaluated the potential toxicity of *S. mitis* and the lead *S. mitis*-CSP-2 analogs (native signal, *S. mitis*-CSP-2, lead activator, *S. mitis*-CSP-2-n11, and lead inhibitor, *S. mitis*-CSP-2-E1Af10r16) toward mammalian cells. To this end, we conducted hemolysis assays against defibrinated rabbit red blood cells (RBCs) in both THY media alone and THY media containing the *S. mitis* ATCC 49456 strain. Our results indicate that *S. mitis* as well as all the tested peptides are nontoxic, exhibiting the same degree of hemolysis as the media only negative control (**Figure 26**). These results highlight the potential utility of *S. mitis* as a therapeutic delivery vehicle that does not elicit any apparent cytotoxicity.

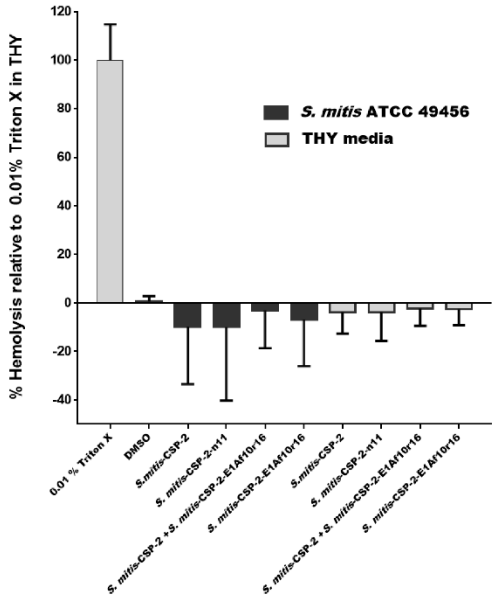


Figure 26. Hemolytic activity of *S. mitis*-CSP-2-derived QS modulators against defibrinated rabbit red blood cells in THY media (white) or in the presence of *S. mitis* ATCC 49456 (black). Both *S. mitis* and the lead *S. mitis*-CSP-2-based QS modulators exhibited no apparent cytotoxicity. Experiments were performed in triplicate on three separate days. See Appendix 4 for full experimental details.

SUMMARY AND CONCLUSIONS

It has become evident that bacteria utilize population size-dependent QS pathways to coordinate different phenotypes, including processes involved in pathogenicity. As such, delineating QS circuits, including their regulatory roles, could be used to study bacterial sociality and to develop novel anti-virulence therapeutic strategies. In this work, we set out to study the competence regulon QS circuitry in *S. mitis*. Our analysis involved confirming the identity of the native CSP signal utilized by *S. mitis* to assess its population density and initiate the QS response, evaluating the SARs that govern CSP:ComD interactions and lead to ComD, and consequently, QS activation, and determining the regulatory roles of the competence regulon QS circuitry in initiating potentially pathogenic phenotypes. Our analysis revealed important SAR insights of the CSP signal and facilitated the development of novel CSP-based QS modulators, both enhanced QS activators and potent competitive inhibitors. Our analysis further demonstrated the involvement of the competence regulon QS circuitry in regulating competence development and biofilm formation, as well as the ability of the CSP-based tools developed in this study to either enhance or attenuate these phenotypes. The CSP-based QS modulators identified in this study can thus be utilized to evaluate the effects temporal QS modulation has on the ability of *S. mitis* to inhabit its natural niche.

It is also now clear that bacteria utilize QS not only for intra-species communication, but also to mediate inter-species communication and interactions. Through the evaluation of the ability of the *S. pneumoniae* and *S. mitis* native CSP signals to modulate the ComD receptor, and thus QS response, of the other species, we found that the native *S. mitis* CSP signal, *S. mitis*-CSP-2, can modulate the QS response in *S. pneumoniae*, but that the native

CSP signals in *S. pneumoniae* are not effective modulators of QS response in *S. mitis*. It would be interesting to speculate which species would benefit from this potential crosstalk: whether it would be the receiver species, *S. pneumoniae*, who eardrops into *S. mitis* population size, and at high *S. mitis* population density releases virulence determinants to eliminate the competition and acquire potentially useful genetic material; or whether it would be the producer species, *S. mitis*, who intentionally triggers the competence regulon QS circuitry in *S. pneumoniae* to take advantage of the highly pathogenic nature of pneumococcus and the associated production and release of a myriad of virulence determinant by *S. pneumoniae* to effectively acquire genetic material from the environment. To test these hypotheses in complex bacterial communities or in *in vivo* settings, temporal regulation of QS in either species, or both species, would be needed. As such, the *S. mitis*-CSP-2-based QS modulator capable of activating both the pneumococcus ComD receptors and the *S. mitis* ComD-2 receptor with high potencies (*S. mitis*-CSP-2-N7II8F; EC_{50} values in the low nanomolar range against all three receptors) that was identified in this study, to the best of our knowledge the first example of a CSP-based multi-species QS activator, would be of great value to study this potential crosstalk between a commensal (*S. mitis*) and pathogenic (*S. pneumoniae*) species. Such studies are ongoing in our laboratory and will be reported in due course.

MATERIALS AND METHODS.

General. The peptide analogs were synthesized using standard Fmoc solid-phase peptide synthesis protocols,[27] followed by purification using RP-HPLC to >95% purity. Masses of purified peptides were confirmed by mass spectrometry. See Tables **S-2** to **S-5** in

Appendix 4 for peptide masses and purities. See the Supporting Information for full experimental details.

Isolation of Crude Peptides from Bacterial Supernatants. The native *S. mitis*-CSP-2 was isolated from cell-free bacterial supernatants using previously described methods with minor modifications.[25] For full experimental details, see the Supporting Information.

Development of *S. mitis* Luciferase-based Reporter System. The *S. mitis* luciferase-based reporter strain was constructed using previously described methods with some modifications.[3] See the Supporting Information in Appendix 4 for full experimental details.

Luminescence Reporter Assay. Activation Assays. The ability of synthetic *S. mitis*-CSP-2 analogs to activate the expression of *S. mitis comX* was determined using the constructed *S. mitis* luciferase reporter strain. First, bacteria from the *S. mitis* luciferase reporter strain freezer stock were streaked onto a Todd-Hewitt broth supplemented with 0.5% yeast extract (THY) agar plate containing 5% horse serum and spectinomycin at a final concentration of 200 μg/mL. The plate was incubated overnight in a CO_2 incubator (37 °C with 5% CO_2). Fresh colonies were transferred to Tryptic Soy Broth (TSB**)** supplemented with spectinomycin at a final concentration of 200 μg/mL, and the culture was incubated at 37 °C in a 5% CO_2-supplemented atmosphere and grown overnight (16 hrs) until the culture reached an absorbance at 600 nm (optical density at 600 nm [OD_{600}]) of 0.45. Overnight cultures of the *S. mitis* reporter strain at OD_{600} 0.45 were diluted 1:10 in TSB and the liquid cultures were incubated in a CO_2 incubator for 30-45 min, until the bacteria reached early exponential stage ([OD_{600} ~ 0.2) as determined by using a plate reader. During incubation, clear-bottom white 96-well microtiter plates were prepared for the activation assays. An initial activation screening was performed at a high peptide

concentration (10,000 nM) for all the *S. mitis*-CSP-2 analogs. For each experimental sample, a total of 2 μL of a 1 mM CSP stock solution in dimethyl sulfoxide (DMSO) was added in triplicate to the clear-bottom white 96-well microtiter plate. A total of 2 μL of DMSO was added in triplicate and served as the negative control, and 2 μL of a 1 mM stock of native *S. mitis*-CSP-2 was added in triplicate as a positive control. In the dark, 2 μL of a 15 mg/mL D-luciferin stock in distilled water was added to each well. Following the addition of synthetic CSPs and D-luciferin, 196 μL of the diluted bacterial culture was added to each well, and the plate was incubated for 30 min at 37 °C. Following incubation, the OD_{600} and luminescence of each experimental well were measured. Results were reported as percent activation, which is the ratio between the luminescence (presented as relative luminescence units, RLU/OD_{600}) of the analog and that of the positive control. Analogs that exhibited activity > 75% compared to the positive control were further evaluated using a dose dependent assay in which peptide stock solutions were diluted with DMSO in serial dilutions (1:2 ,1:3, or 1:5) and assayed as described above. EC_{50} values, the concentration of the peptide to achieve a half maximal response, for each activator were calculated using GraphPad Prism. Experiments were performed in triplicate on three separate days.

Inhibition Assays. Analogs that exhibited less than 50% activation in the initial screening were further evaluated for potential competitive inhibition of the competence regulon. The ability of synthetic *S. mitis*-CSP-2 analogs to competitively inhibit *comX* expression by outcompeting the native *S. mitis*-CSP-2 for the receptor binding site was evaluated using the same assay conditions as those for the activation assays, except that for the inhibition screening the native *S. mitis*-CSP-2 was introduced to each well at a set concentration (1

μM) that was chosen to afford complete activation of the competence regulon, as determined from dose−dependent curves created for the native *S. mitis*-CSP-2. As a positive control, 2 μL of *S. mitis*-CSP-2 and 2 μL of DMSO were added to the same well in triplicate. As a negative control, 4 μL of DMSO was added in triplicate. Then, 2 μL of 15 mg/mL D-luciferin in distilled water was added to each experimental or control well, after which 194 μL of bacteria was added, and the plate was incubated for 30 min at 37 °C. Following incubation, the OD_{600} and luminescence of each experimental well were measured. Results were reported as percent activation, which is the ratio between the luminescence (presented as relative luminescence units, RLU/OD_{600}) of the analog and that of the positive control. Analogs that exhibited significant competitive inhibition in the initial screening were further evaluated using a dose-dependent assay where peptide stock solutions were diluted with DMSO in serial dilutions (either 1:2, 1:3, or 1:5) and assayed as described above. GraphPad Prism was used to calculate the IC_{50} values, which are the concentration of inhibitor where the response (or binding) is reduced by half. Experiments were performed in triplicate on three separate days.

Transformation Assay. A single colony of *S. mitis* ATCC 49456 was grown overnight in 5 mL of THY media (pH 7.3) at 37 °C with 5% CO_2. Following incubation, *S. mitis* ATCC 49456 was diluted 1:100 into fresh THY media containing 5% horse serum and the culture was incubated at 37 °C in a 5% CO_2-supplemented atmosphere for 4 to 5 h (until the culture reached an OD_{600} of 0.30). Then, 100 μL of this diluted culture was inoculated to 900 μL of fresh THY media containing 5% horse serum. Following inoculation, synthetic *S. mitis*-CSP-2 (at a final concentration of 200 nM), the lead activator, *S. mitis*-CSP-2-n11 (at a final concentration of 200 nM), or *S. mitis*-CSP-2 (at a final concentration of 200 nM)

together with the lead inhibitor, *S. mitis*-CSP-2-E1Af10r16 (at a final concentration of 500 nM) were added to the culture. After that, pFW5-luc ($Spec^R$) plasmid at a final concentration of 1 μg/mL was added to the culture. Parallel assays without the *S. mitis*-CSP-2 or without both the *S. mitis*-CSP-2 and plasmid were used as controls to assess the indigenous competence or antibiotic resistance of *S. mitis* under the tested conditions. After 3 to 4 h of incubation at 37 °C, 200 μL of the culture was plated on THY agar containing 200 μg/mL spectinomycin and incubated at 37 °C with 5% CO_2 for 24 to 48 h to identify positive transformants. Experiments were performed in triplicate on three separate days.

Biofilm-Formation Assay. Biofilm formation was assessed using previously described protocols with minor modifications.[25] See the Supporting Information in Appendix 4 for full experimental details.

Hemolysis Assay. Hemolysis was assessed using previously described protocols with minor modifications.[25] See the Supporting Information in Appendix 4 for full experimental details.

CD Spectroscopy. CD spectra were recorded using an Aviv Biomedical CD spectrometer (model 202-01). For experimental details, see the Supporting Information in Appendix 4.

Acknowledgments

This work was supported by a grant from the National Institutes of Health (R35GM128651). The *S. mitis* ATCC 49456 strain was a generous gift from L. McGee (CDC Streptococcus Lab), and the pFW5-luc plasmid was a generous gift from D. Bernd Kreikemeyer (University of Rostock). We would like to thank M. J. Tucker (University of Nevada, Reno) for the use of the CD spectrometer.

Conflict of interest

The authors declare no competing interests.

REFERENCES

1 Kilian, M. *et al.* Evolution of *Streptococcus pneumoniae* and Its Close Commensal Relatives. *PLos One* **3**, e2683 (2008) .

2 Milly, T. A. & Tal-Gan, Y. Biological evaluation of native streptococcal competence stimulating peptides reveals potential crosstalk between *Streptococcus mitis* and *Streptococcus pneumoniae* and a new scaffold for the development of *S. pneumoniae* quorum sensing modulators. *RSC chem. biol.* **1**, 60-67 (2020).

3 Salvadori, G., Junges, R., Morrison, D. A. & Petersen, F. C. Overcoming the barrier of low efficiency during genetic transformation of *Streptococcus mitis*. *Front. Microbiol.* **7**, 1009 (2016).

4 Huang, S. S. *et al.* Healthcare utilization and cost of pneumococcal disease in the United States. *Vaccine* **29**, 3398-3412 (2011).

5 Mehr, S. & Wood, N. *Streptococcus pneumoniae*--a review of carriage, infection, serotype replacement and vaccination. *Paediatr. Respir. Rev.* **13**, 258-264 (2012).

6 Mitchell, J. *Streptococcus mitis*: walking the line between commensalism and pathogenesis. *Mol. Oral Microbiol.* **26**, 89-98 (2011).

7 Tunkel, A. R. & Sepkowitz, K. A. Infections caused by viridans streptococci in patients with neutropenia. *Clin. Infect. Dis.* **34**, 1524-1529 (2002).

8 Beighton, D., Carr, A. & Oppenheim, B. A. Identification of viridans streptococci associated with bacteraemia in neutropenic cancer patients. *J. Med. Microbiol.* **40**, 202-204 (1994).

9 Wescombe, P. A., Hale, J. D., Heng, N. C. & Tagg, J. R. Developing oral probiotics from *Streptococcus salivarius*. *Future Microbiol.* **7**, 1355-1371 (2012).

10 Xie, E. *et al.* Correction: Oral Delivery of a Novel Recombinant *Streptococcus mitis* Vector Elicits Robust Vaccine Antigen-Specific Oral Mucosal and Systemic Antibody Responses and T Cell Tolerance. *PLos One* **11**, e0147781 (2016).

11 Rukke, H. V. *et al.* Protective role of the capsule and impact of serotype 4 switching on *Streptococcus mitis*. *Infect. Immun.* **82**, 3790-3801 (2014).

12 Shekhar, S. *et al.* Antibodies reactive to commensal *Streptococcus mitis* show cross-reactivity with virulent *Streptococcus pneumoniae* serotypes. *Front. immunol.* **9**, 747 (2018).

13 Donati, C. *et al.* Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. *Genome Biol.* **11**, 1-19 (2010).

14 Hakenbeck, R. *et al.* Acquisition of Five High-M rPenicillin-Binding Protein Variants during Transfer of High-Level β-Lactam Resistance from *Streptococcus mitis* to *Streptococcus pneumoniae*. *J. Bacteriol.* **180**, 1831-1840 (1998).

15 Rukke, H., Hegna, I. & Petersen, F. Identification of a functional capsule locus in *Streptococcus mitis*. *Mol. Oral Microbiol.* **27**, 95-108 (2012).

16 da Gloria Carvalho, M. *et al.* Non-pneumococcal mitis-group streptococci confound detection of pneumococcal capsular serotype-specific loci in upper respiratory tract. *PeerJ* **1**, e97 (2013).

17 Milly, T. A. *et al.* Harnessing Multiple, Nonproteogenic Substitutions to Optimize CSP: ComD Hydrophobic Interactions in Group 1 *Streptococcus pneumoniae*. *ChemBioChem* **22**, 1940-1947 (2021).

18 Milly, T. A. *et al.* Optimizing CSP1 analogs for modulating quorum sensing in *Streptococcus pneumoniae* with Bulky, Hydrophobic Nonproteogenic Amino Acid Substitutions. *RSC chem. biol.* **3**, 301-311 (2022).

19 Brennan, A. A., Mehrani, M. & Tal-Gan, Y. Modulating streptococcal phenotypes using signal peptide analogues. *Open Biol.* **12**, 220143 (2022).

20 Milly, T. A. & Tal-Gan, Y. Targeting peptide-based quorum sensing systems for the treatment of gram-positive bacterial infections. *J. Pept. Sci.* e24298 (2022).

21 Cvitkovitch, D. G., Li, Y.-H. & Ellen, R. P. Quorum sensing and biofilm formation in *Streptococcal* infections. *J. Clin. Investig.* **112**, 1626-1632 (2003).

22 Junges, R. *et al.* Characterization of a signaling system in *Streptococcus mitis* that mediates interspecies communication with *Streptococcus pneumoniae*. *Appl. Environ. Microbiol.* **85**, e02297-02218 (2019).

23 Salvadori, G. *et al.* Natural transformation of oral Streptococci by use of synthetic pheromones. *Oral. Biol.* 219-232 (2017).

24 Harrington, A. *et al.* Secretion, maturation, and activity of a quorum sensing peptide (GSP) inducing bacteriocin transcription in *Streptococcus gallolyticus*. *Mbio* **12**, e03189-03120 (2021).

25 Mull, R. W. & Tal-Gan, Y. Elucidating the Role and Structure–Activity Relationships of the *Streptococcus oligofermentans* Competence-Stimulating Peptide. *ACS Chem. Biol.* **16**, 2834-2844 (2021).

26 Green, M. R. & Sambrook, J. The Inoue method for preparation and transformation of competent *Escherichia coli*:"Ultracompetent" cells. *Cold Spring Harb. Protoc.* **2020**, pdb. prot101196 (2020).

27 Chan, W. & White, P. *Fmoc solid phase peptide synthesis: a practical approach*. Vol. 222 (OUP Oxford, 1999).

28 Yang, Y. *et al.* Structure–activity relationships of the competence stimulating peptides (CSPs) in *Streptococcus pneumoniae* reveal motifs critical for intra-group and cross-group ComD receptor activation. *ACS Chem. Biol.* **12**, 1141-1151 (2017).

29 Kadioglu, A., Weiser, J. N., Paton, J. C. & Andrew, P. W. The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol* **6**, 288-301 (2008).

Chapter 6: Conclusion

The growing prevalence of multi-drug resistant bacteria necessitates the development of nonantibiotic therapeutics that reduce the rate of antibiotic resistance development while still attenuating bacterial infections by reducing the organism's pathogenicity. One potential approach to attenuate the development of antibiotic resistant bacteria is modulating nonessential bacterial pathways that are not required for cell survival. Interception of the competence regulon QS circuitry could therefore be used to attenuate bacterial infectivity and to develop novel anti-virulence therapeutic strategies.

In our initial collaborative studies, we highlighted the development of highly potent pneumococcal QS modulators with better pharmacological properties. We have optimized the binding interactions between CSP1 and ComD1 in Group 1 *S. pneumoniae* through the incorporation of bulkier, hydrophobic, non-proteogenic amino acids in key hydrophobic positions of CSP1. We were able to develop highly potent pneumococcal QS modulators with superior pharmacological properties, including improved resistance to enzymatic degradation, all while remaining nontoxic against mammalian cells. Overall, this analysis provided valuable information regarding the ComD1 hydrophobic binding pockets and key SAR knowledge of the pneumococcal CSP1 pheromone, which can be utilized for the rational design of highly potent, pharmacologically stable CSP-based pneumococcal QS modulators. Potential future studies can focus on *in vivo* assessment of the lead analogs identified in these studies to determine their efficacy in attenuating pneumococcus virulence during host infection. Additional assessment of pharmacological properties of the lead analogs would increase the possibility of utilization of these lead QS modulators as promising drug lead scaffolds against *S. pneumoniae* infections, either as the sole agent or in combination with current antimicrobial agents.

In a different study, we set out to study the competence regulon QS circuitry in the oral commensal, *S. mitis*. Through this analysis, we were able to determine the regulatory roles of this complex QS circuitry in initiating potentially pathogenic phenotypes. Our phenotypic analysis revealed that both competence development and biofilm formation are regulated by the *S. mitis* competence regulon. Additionally, in-depth SAR analysis of the native *S. mitis*-CSP-2 pheromone revealed several structural motifs that are critical to receptor binding and activation. The critical SAR insights obtained through this study eventually led to the development of potent CSP-based QS modulators of the *S. mitis* competence regulon. Most importantly, biological evaluation of the ability of the native *streptococci* CSPs to modulate the competence regulon in *S. pneumoniae* and *S. mitis* revealed that the native *S. mitis*-CSP-2 signal has the potential to influence pneumococcal QS-regulated phenotypes, but that the native pneumococcal CSP signals are not effective modulators of QS response in *S. mitis*. Capitalizing on this crosstalk and utilizing rational design, we were able to identify the first example of a CSP-based multi-species QS activator capable of activating both the pneumococcus ComD receptors and the *S. mitis* ComD-2 receptor with high potencies. This multi-species QS modulator could be used as a potential tool in future studies to further investigate the potential crosstalk between a commensal (*S. mitis*) and pathogenic (*S. pneumoniae*) species. As *S. mitis* and *S. pneumoniae* are often found in the same habitat, it would be interesting to determine the possible effects this QS-mediated crosstalk may have on colonization or virulence adaptation for these commensal and pathogenic species. The human pathogen *S. pneumoniae* may take benefit from this crosstalk to infiltrate the *S. mitis* population, and upon sensing a high *S. mitis* population density release virulence factors to eliminate the

competition and obtain resistance genes from the environment. Alternatively, this agonistic crosstalk could be an effective strategy for the commensal *S. mitis* to gain advantage in a complex microbial environment. *S. mitis* may utilize several virulence factors produced and released by *S. pneumoniae* to effectively obtain resistance genes from the environment. The understanding of such QS-mediated crosstalk could provide important insights into the dynamics of bacterial communities. Further investigations focused on temporal regulation of QS in either species, or both species in either complex bacterial communities or in *in vivo* settings would provide key insights to answer these important fundamental questions regarding the relationship between *S. mitis* and *S. pneumoniae*. In future studies, the potent *S. mitis*-CSP-2-based QS modulators developed in this study could be used as potential tools to evaluate QS regulation in other *S. mitis* strains, such as *S. mitis* NCTC 8033. A cell-based luminescence assay using a *S. mitis* NCTC 8033 luciferase QS reporter strain could be utilized to understand QS modulation in this *S. mitis* strain. This biological evaluation would provide valuable insights about intraspecies interactions between strains belonging to different *S. mitis* specificity groups (pherotypes). Besides, the native CSP signals of *S. pneumoniae* could be used as templates and using the critical SAR insights obtained through this study, potential pneumococcal-based modulator could be developed that can modulate the QS system in *S. mitis* Group II. This would provide additional information about the QS circuits and interactions between *S. mitis* and *S. pneumoniae*, and aid in developing pneumococcal-based QS modulators capable of modulating the *S. mitis* ComD-2 receptor that could be applied to assess potential interspecies interactions. Further future work could be used to validate the SAR insights obtained in this study. For example, *S. mitis*-CSP-2-I2M, the *S. mitis*-CSP-2-based pneumococcal QS activator developed in

this study, capable of activating both ComD pneumococcal receptors with high potency and weakly activating the *S. mitis* ComD-2 receptor, revealed the critical role of position 2 in receptor specificity for both pneumococcal and *S. mitis* CSPs. To further validate this result, *S. pneumoniae*-CSP1-M2I could be synthesized, and biological assays could be carried out to determine how this analog affects QS regulation in both *S. pneumoniae* and *S. mitis*.

Overall, beyond providing important SAR insights, our studies on the competence regulons in *S. pneumoniae* and *S. mitis* revealed potential CSP-based QS modulators that can both act as drug leads and as probes to delineate the molecular mechanisms that drive QS activation as well as interspecies interactions.

Appendix 1: Harnessing Multiple, Nonproteogenic Substitutions to Optimize CSP:ComD Hydrophobic Interactions in Group 1 *Streptococcus pneumoniae*

[a]Reprinted with permission from Milly, T. A.; Engler, E. R.; Chichura, K. S.; Buttner, A. R.; Koirala, B.; Tal-Gan, Y.; Bertucci, M. A. Harnessing Multiple, Nonproteogenic Substitutions to Optimize CSP: ComD Hydrophobic Interactions in Group 1 *Streptococcus pneumoniae*. *ChemBioChem 2021*, 22 (11), 1940-1947.

Harnessing Multiple, Non-proteogenic Substitutions to Optimize CSP:ComD Hydrophobic Interactions in Group1 *Streptococcus pneumoniae*

Supporting Information

Tahmina A. Milly,[†] Emilee R. Engler,[‡] Kylie S. Chichura,[‡] Alec R. Buttner,[‡] Bimal Koirala,[†] Yftah Tal-Gan,*[,†] and Michael A. Bertucci*[,‡]

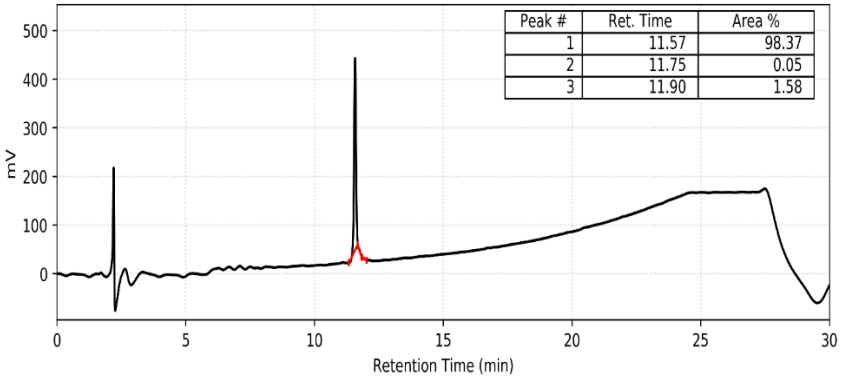
[†] Department of Chemistry, University of Nevada, Reno, 1664 North Virginia Street, Reno, Nevada, 89557, United States

[‡] Department of Chemistry, Moravian College, 1200 Main Street, Bethlehem, Pennsylvania, 18018, United States

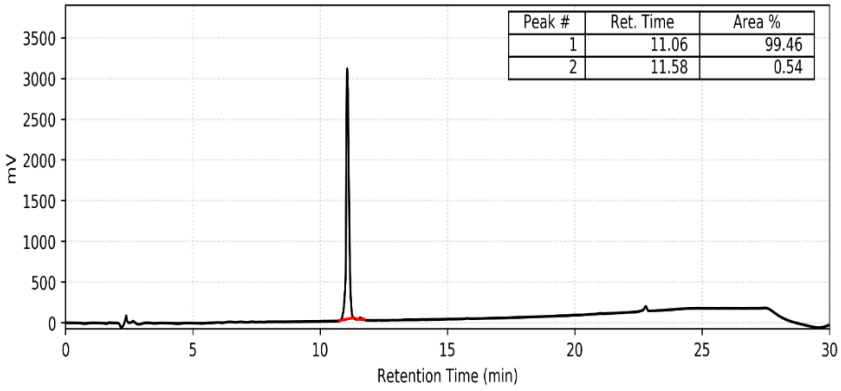
[*] To whom correspondence should be, addressed. ytalgan@unr.edu, bertuccim@moravian.edu

HPLC Traces for CSP1 analogs

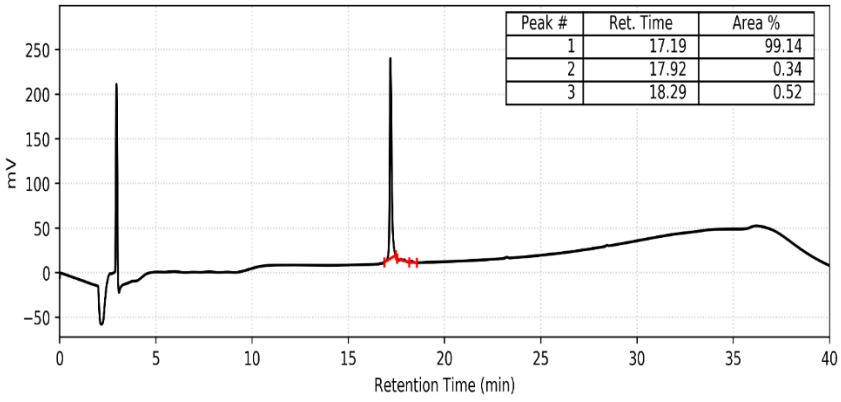
CSP1-L4I/I12L



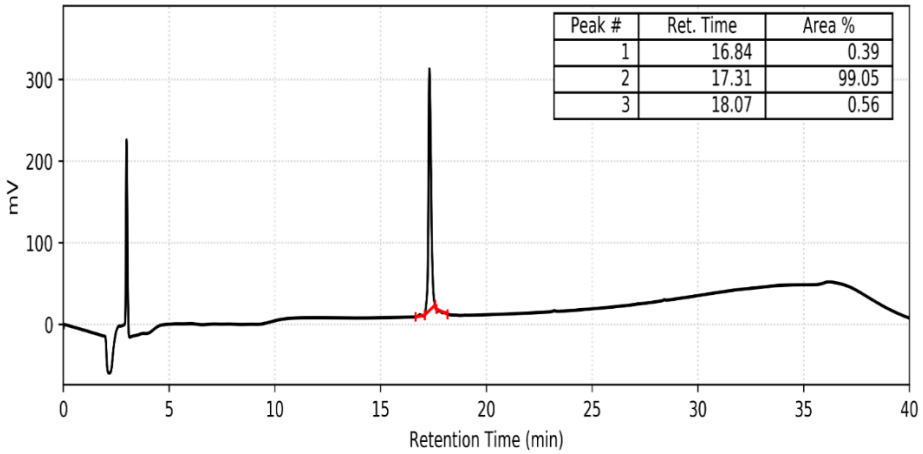
CSP1-L4I/I12NL



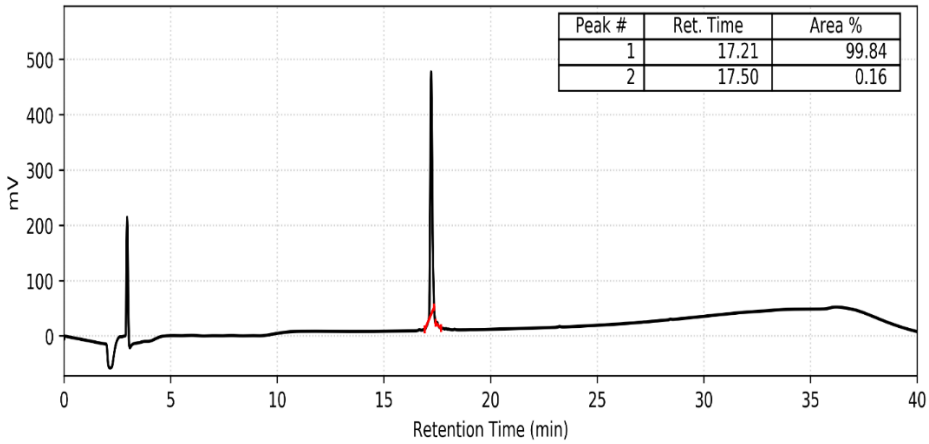
CSP1-L4I/L13I



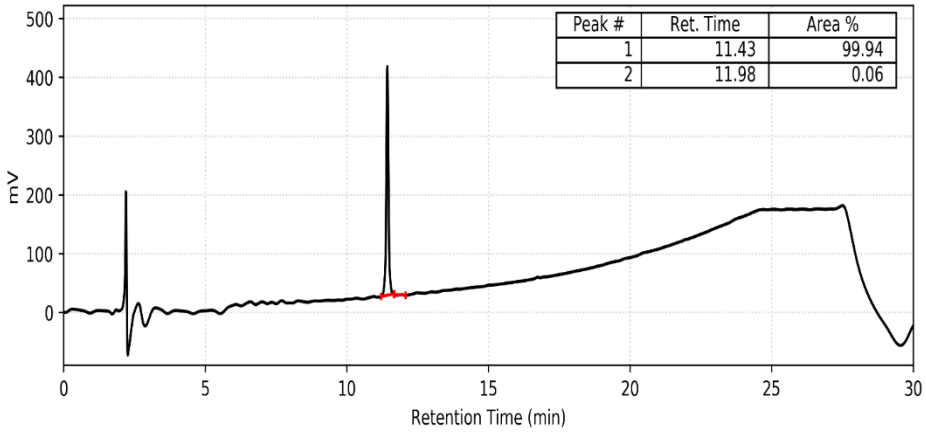
CSP1-L4I/L13NL



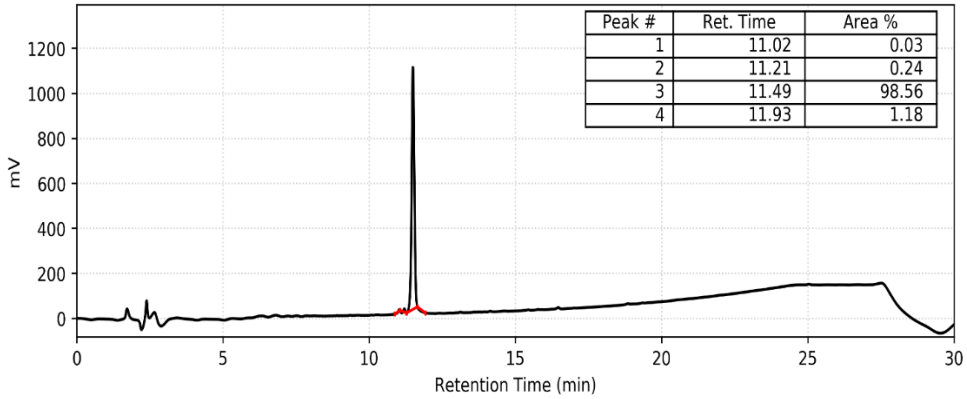
CSP1-L4I/L13NV



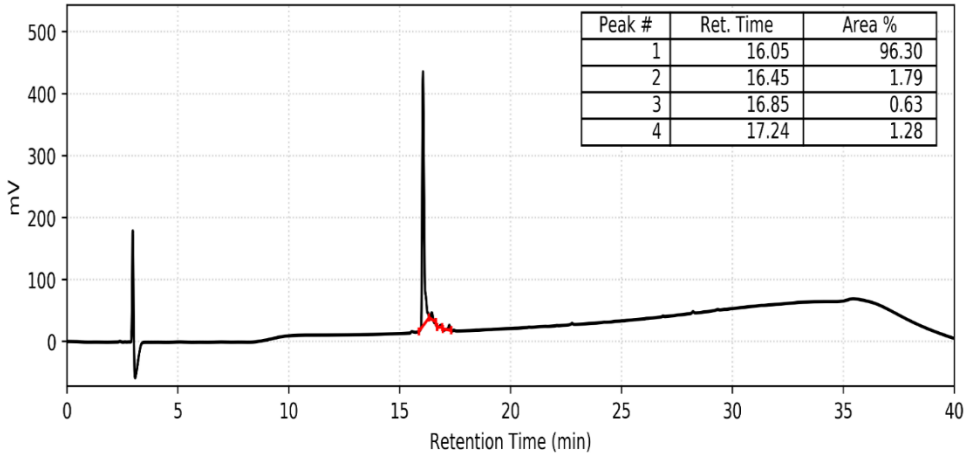
CSP1-L4NV/I12L



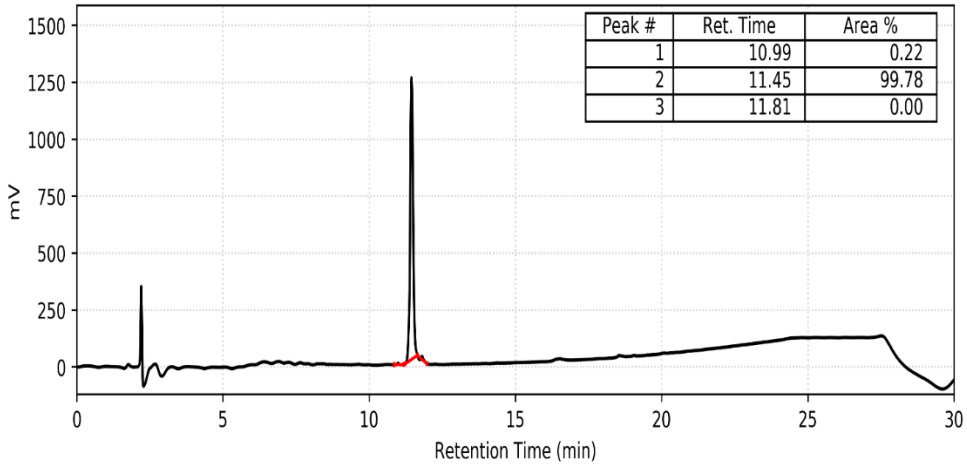
CSP1-L4NV/I12NL



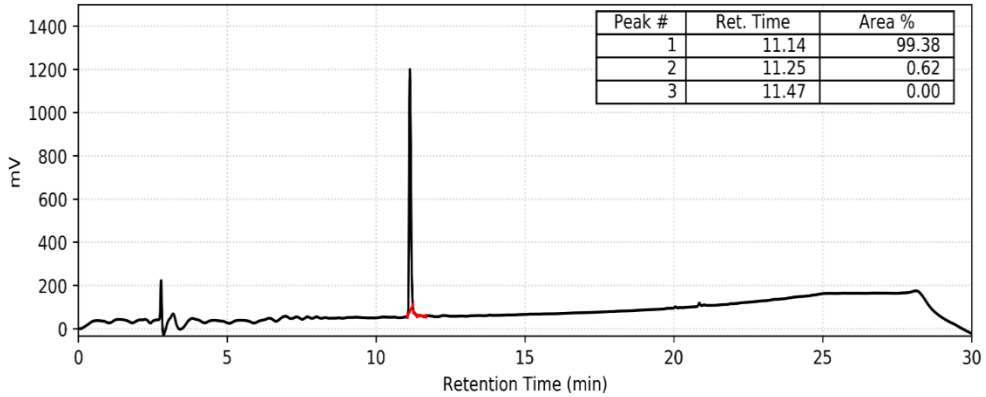
CSP1-L4NV/L13I



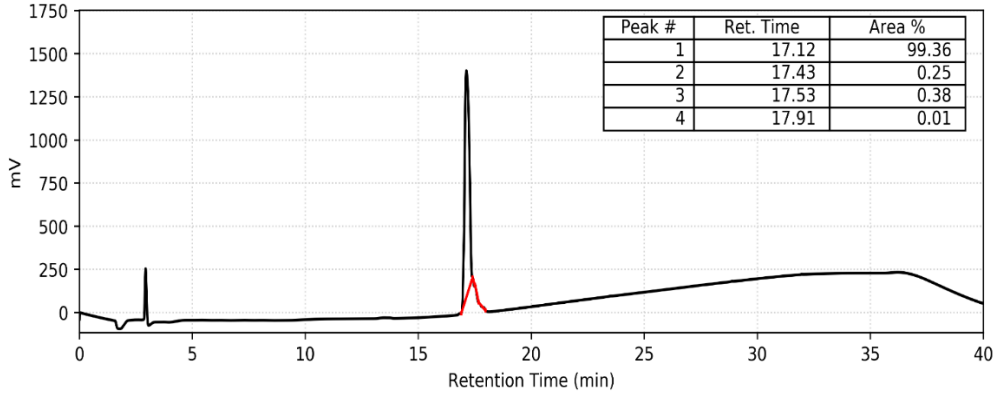
CSP1-L4NV/L13NL



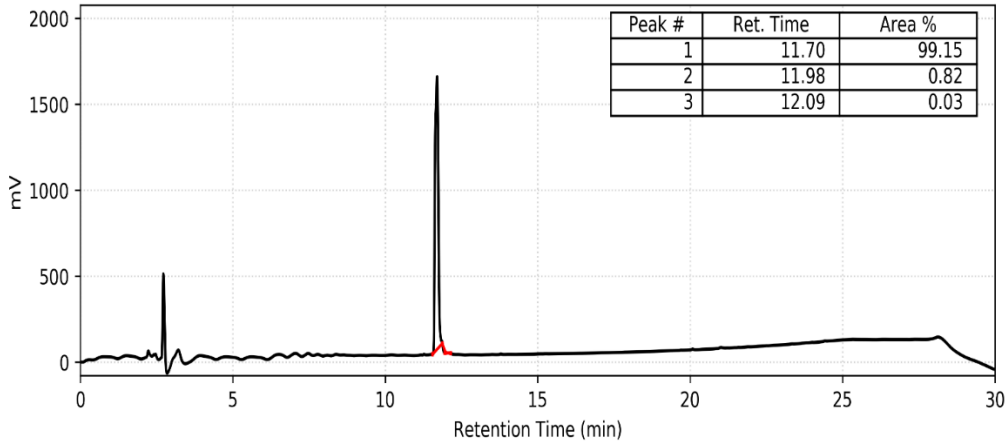
CSP1-L4NV/L13NV



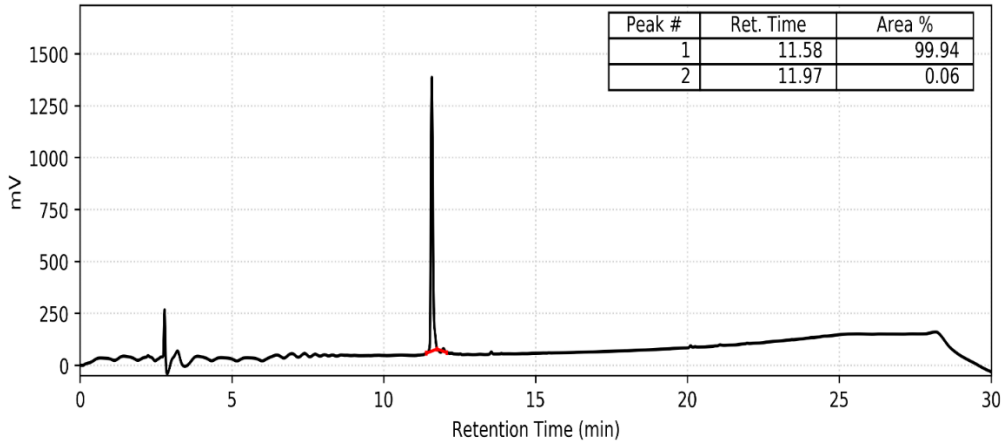
CSP1-I12L/L13I



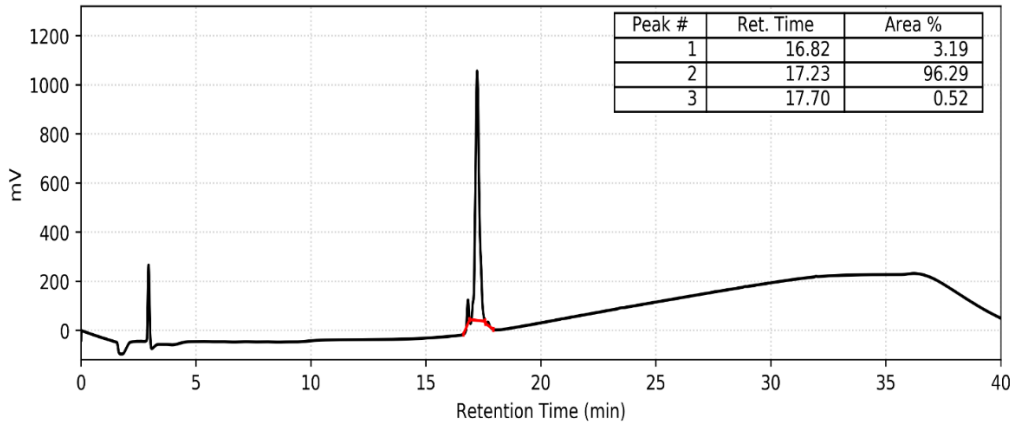
CSP1-I12L/L13NL



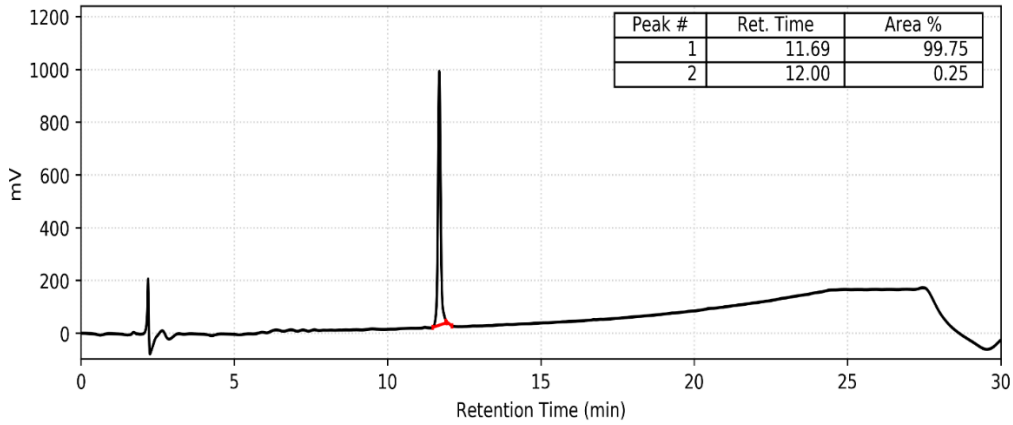
CSP1-I12L/L13NV



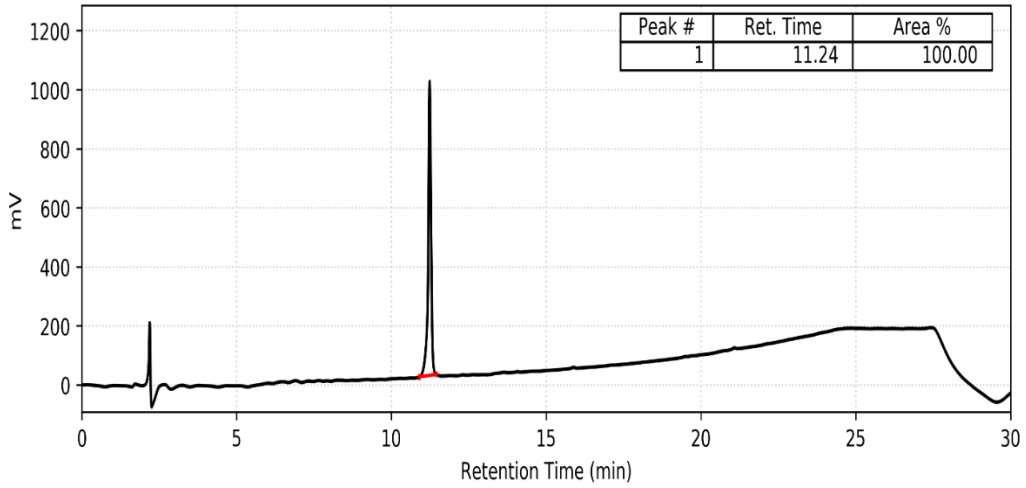
CSP1-I12NL/L13I



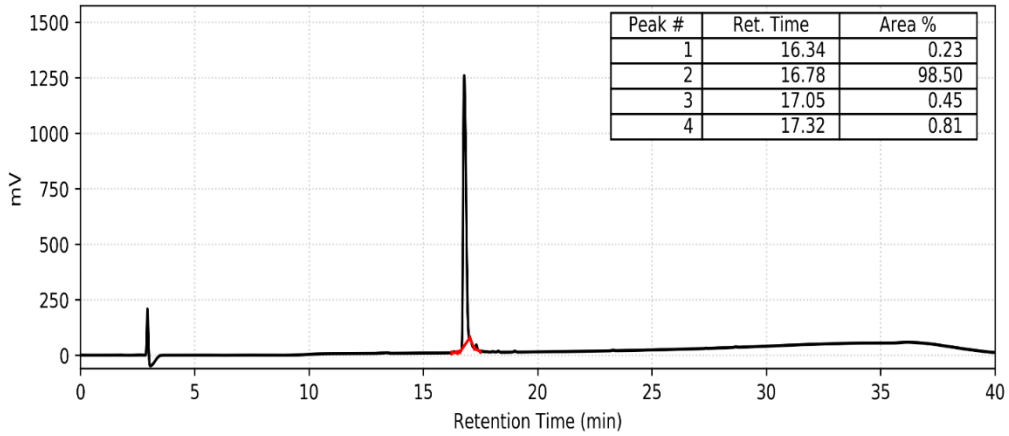
CSP1-I12NL/L13NL



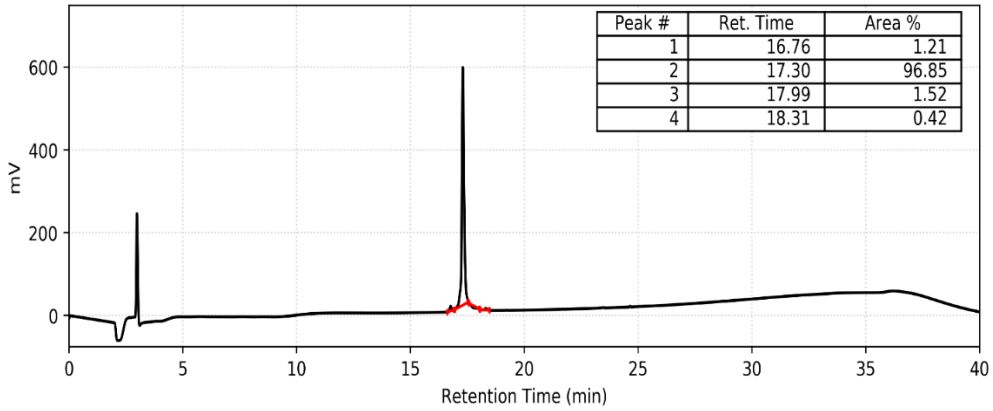
CSP1-I12NL/L13NV



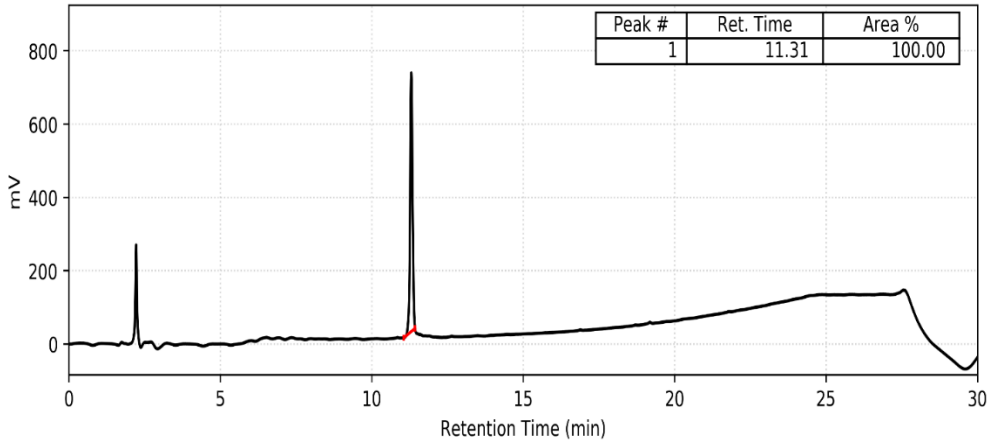
CSP1-L4NV/I12NL/L13NL



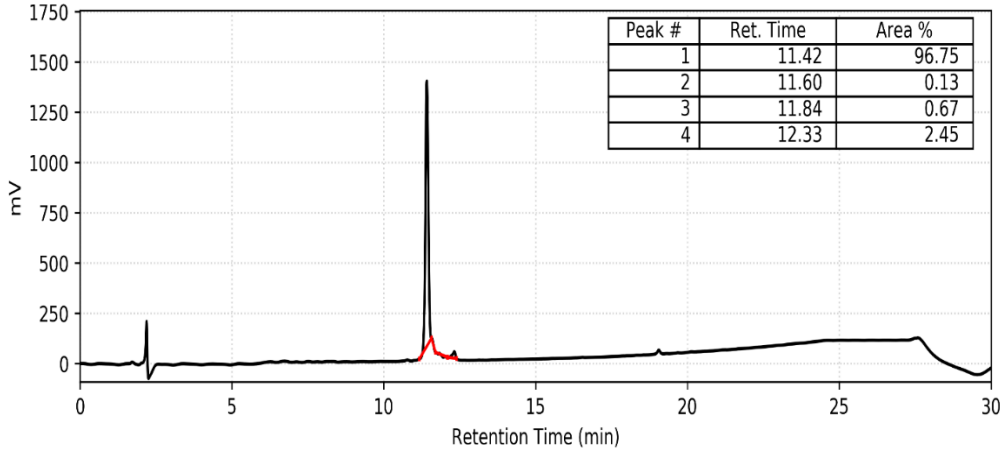
CSP1-E1A/L4I



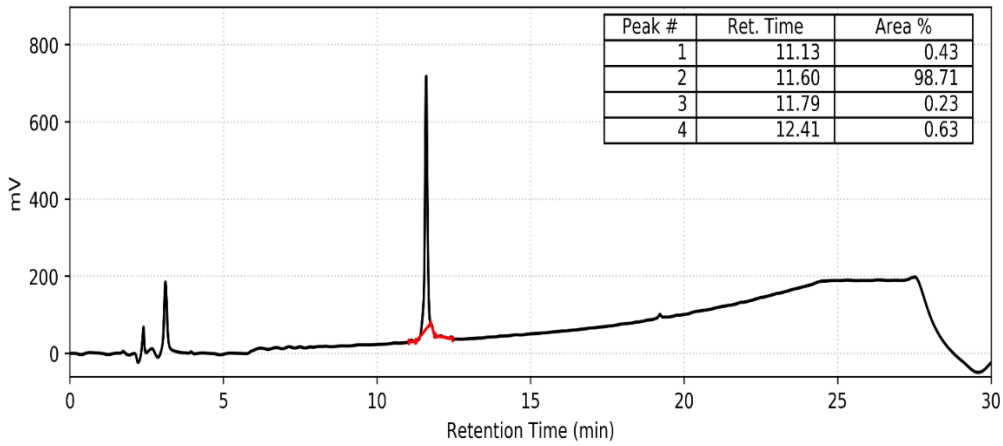
CSP1-E1A/L4NV



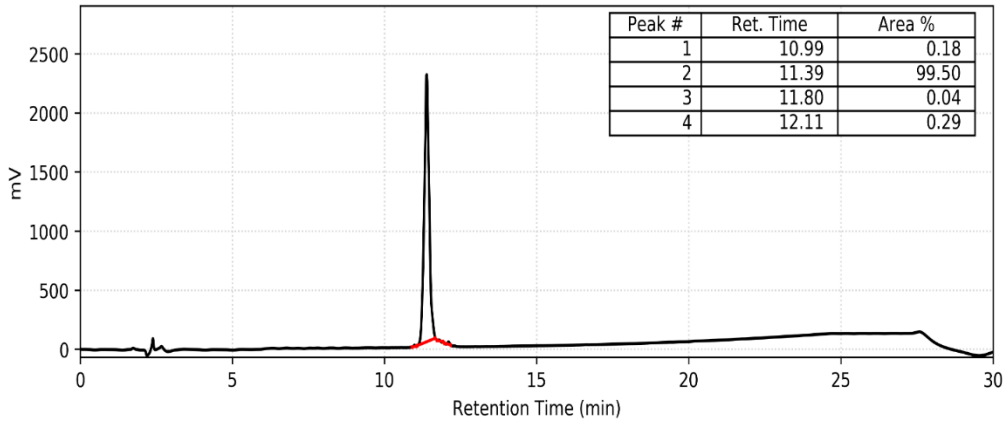
CSP1-E1A/I12L



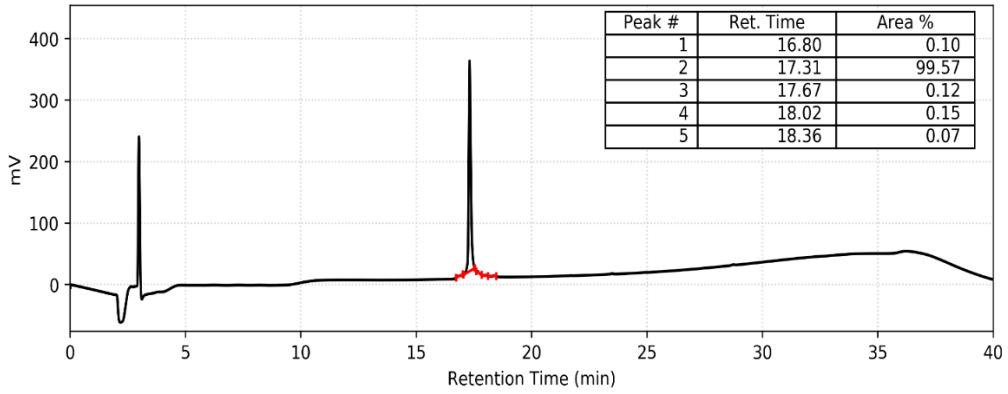
CSP1-E1A/I12NL



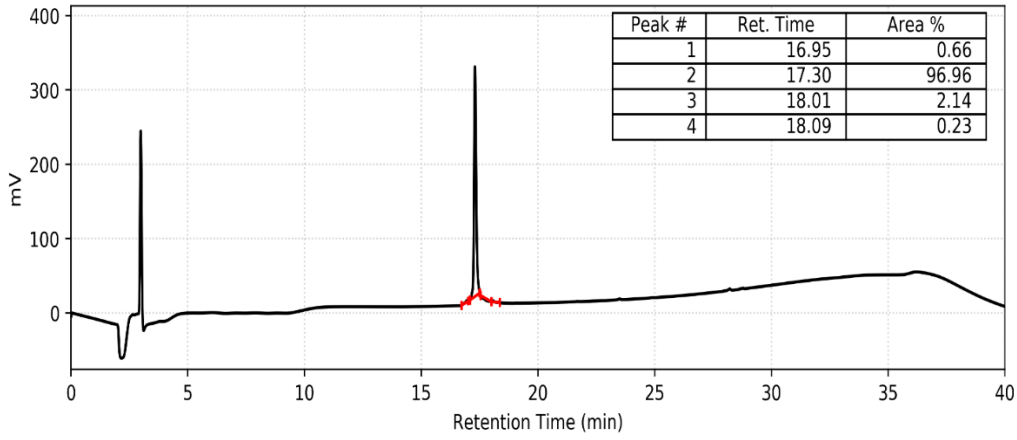
CSP1-E1A/L13NL



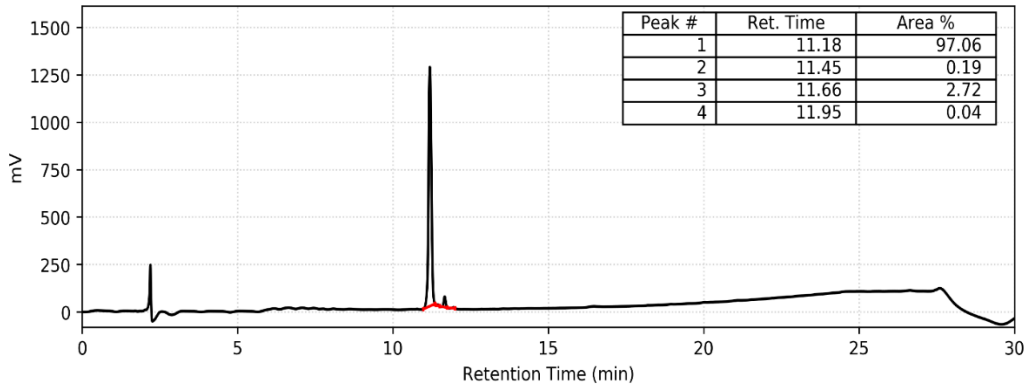
CSP1-E1A/L13I



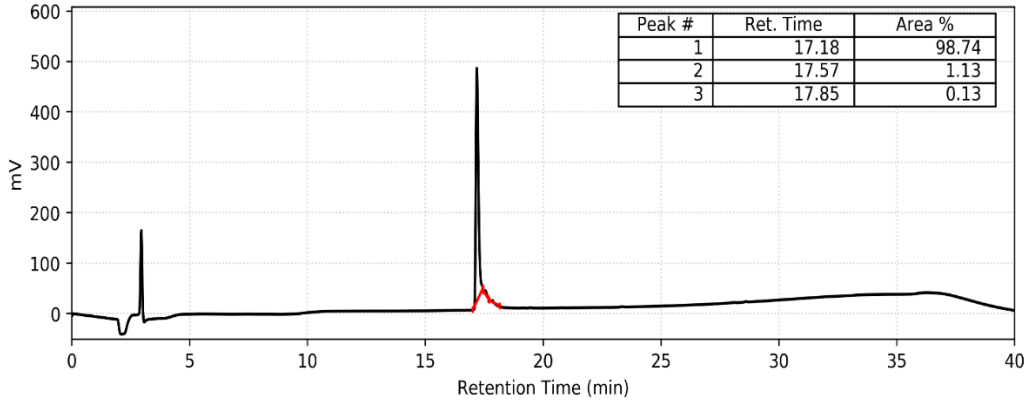
CSP1-E1A/L13NV



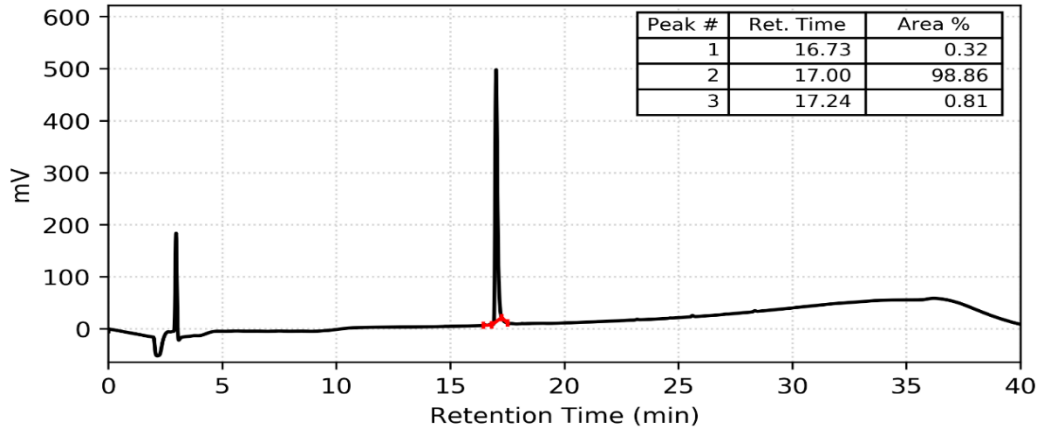
CSP1-E1A/L4NV/I12NL



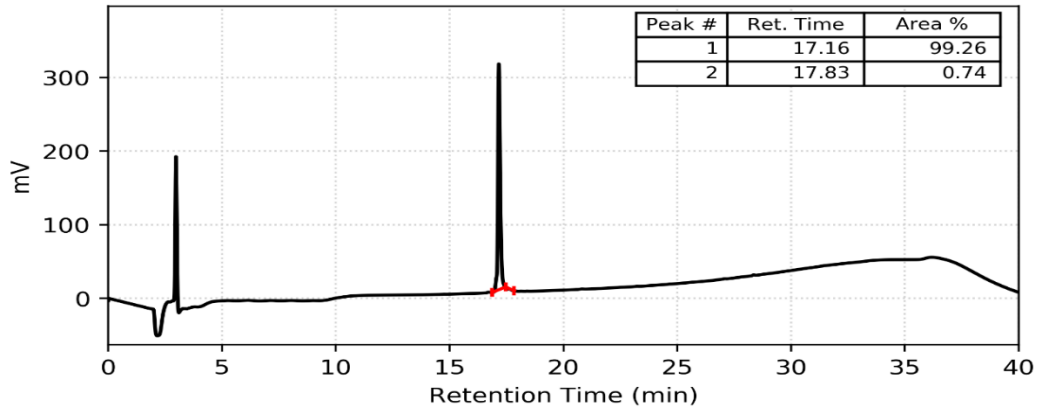
CSP1-E1A/I12NL/L13NL



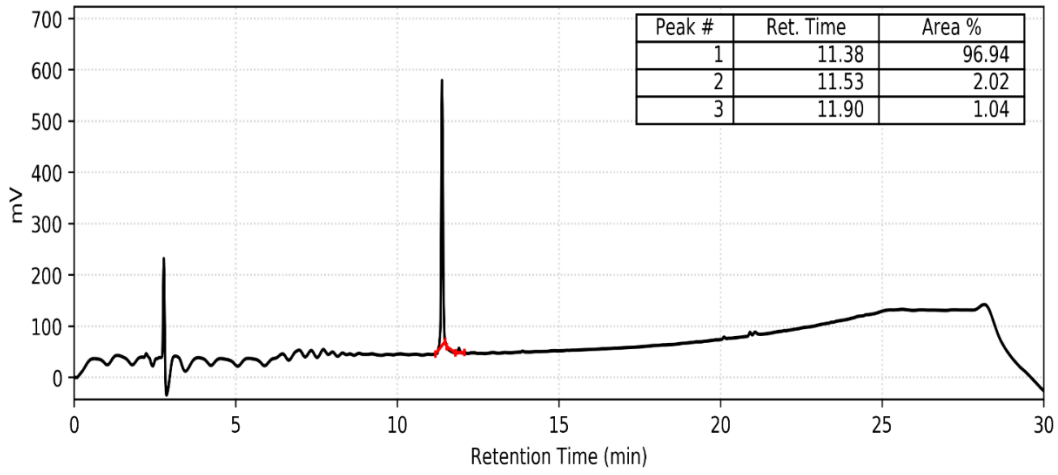
CSP1-E1A/L4NV/L13NV



CSP1-E1A/L4NV/I12NL/L13NV



CSP1-E1A/L4NV/I12NL/L13NL



MS and HPLC data for CSP1 analogs

Table S1. MS and HPLC data for multiply-substituted CSP1 derivatives.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
L4I/I12L	1122.1371	1122.1368	≥98
L4I/I12NL	1122.1371	1122.1362	≥99
L4I/L13I	1122.1371	1122.1397	≥99
L4I/L13NL	1122.1371	1122.1409	≥99
L4I/L13NV	1115.1293	1115.1326	≥99
L4NV/I12L	1115.1293	1115.1277	≥99
L4NV/I12NL	1115.1293	1115.1288	≥98
L4NV/L13I	1115.1293	1115.1293	≥96
L4NV/L13NL	1115.1293	1115.1300	≥99
L4NV/L13NV	1108.1215	1108.1237	≥99
I12L/L13I	1122.1371	1122.1368	≥99
I12L/L13NL	1122.1371	1122.1387	≥99
I12L/L13NV	1115.1293	1115.1307	≥99
I12NL/L13I	1122.1371	1122.1364	≥96
I12NL/L13NL	1122.1371	1122.1360	≥99
I12NL/L13NV	1115.1293	1115.1278	≥99
L4NV/I12NL/L13NL	1115.1293	1115.1295	≥98
L4NV/I12NL/L13NV	1108.1215	1108.1227	≥99

EM = Exact Mass. See methods above.

Table S2. MS and HPLC data for singly and multiply-substituted CSP1-E1A derivatives.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
E1A/L4I	1093.1344	1093.1383	≥96
E1A/L4NV	1086.1265	1086.1256	≥99
E1A/I12L	1093.1344	1093.1314	≥96
E1A/I12NL	1093.1344	1093.1292	≥98
E1A/L13NL	1093.1344	1093.1322	≥99
E1A/L13I	1093.1344	1093.1344	≥99
E1A/L13NV	1086.1265	1086.1261	≥96
E1A/L4NV/I12NL	1086.1265	1086.1238	≥97
E1A/L4NV/L13NV	1079.1187	1079.1219	≥98
E1A/I12NL/L13NL	1093.1344	1093.1347	≥98
E1A/L4NV/I12NL/L13NL	1086.1265	1086.1289	≥96
E1A/L4NV/I12NL/L13NV	1079.1187	1079.1200	≥99

EM = Exact Mass. See methods above.

Primary reporter gene assay data

S. pneumoniae D39pcomX::lacZ (ComD1)

Agonism assays were performed at 10 μM concentration of synthetic CSP1 analogs. CSP1 was used as the positive control (100%) while DMSO as the negative control (0%). Percent (%) ComD1 activation was measured by normalizing the Miller units obtained for each peptide to that of the native CSP1. All peptides were screened in triplicate over three separate trials. Error bars indicate standard error of the mean of nine values.

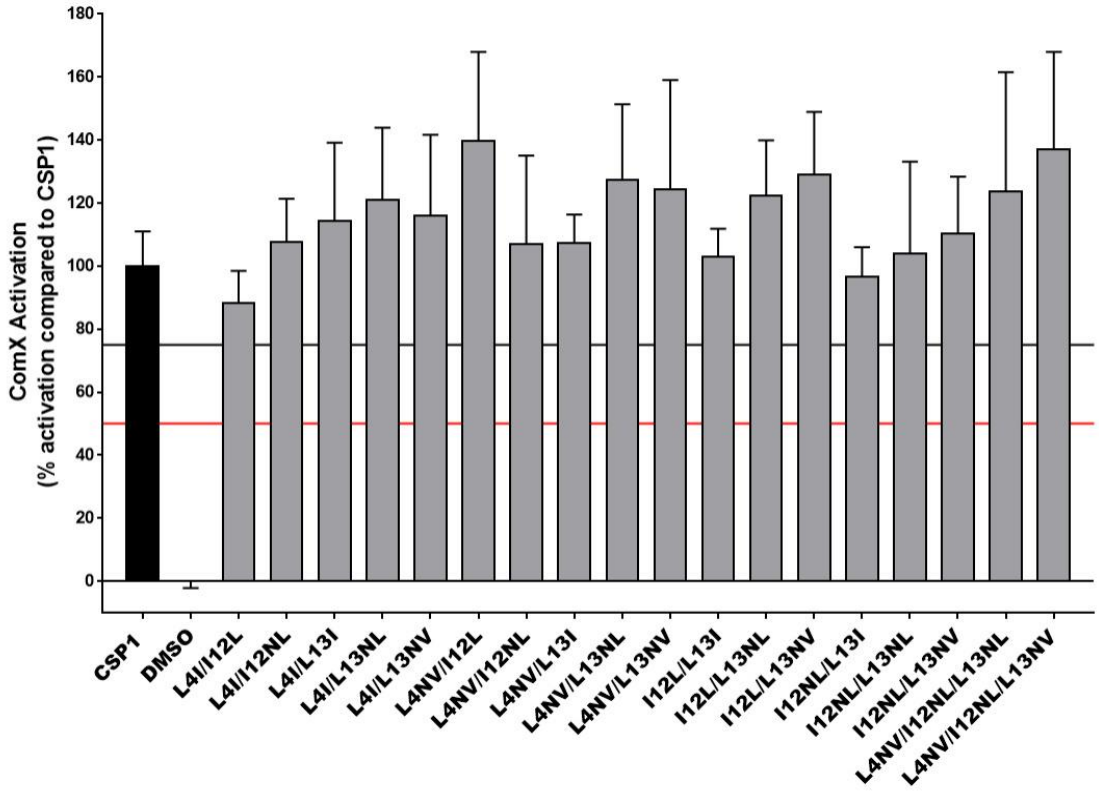


Figure S-1. Primary agonism screening assay data for multiply-substituted CSP1 derivatives. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} values.

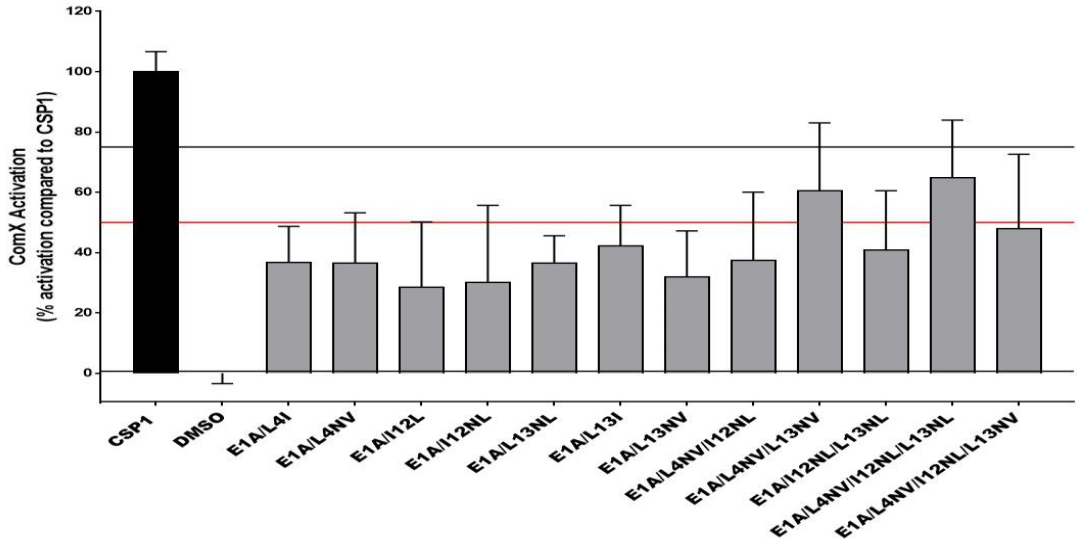


Figure S-2. Primary agonism screening assay data for singly and multiply-substituted CSP1-E1A derivatives. None of the peptides exhibited activation of the ComD1 receptor and peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

Antagonism assays were performed at 10 μM concentration of peptides against 50 nM concentration of CSP1. CSP1 (50 nM) was used as the positive control (100%) while DMSO as the negative control (0%). Percent (%) *comX* activation was measured by normalizing the Miller units obtained for each peptide to that of CSP1. All peptides were screened in triplicate over three separate trials. Error bars indicate standard error of the mean of nine values.

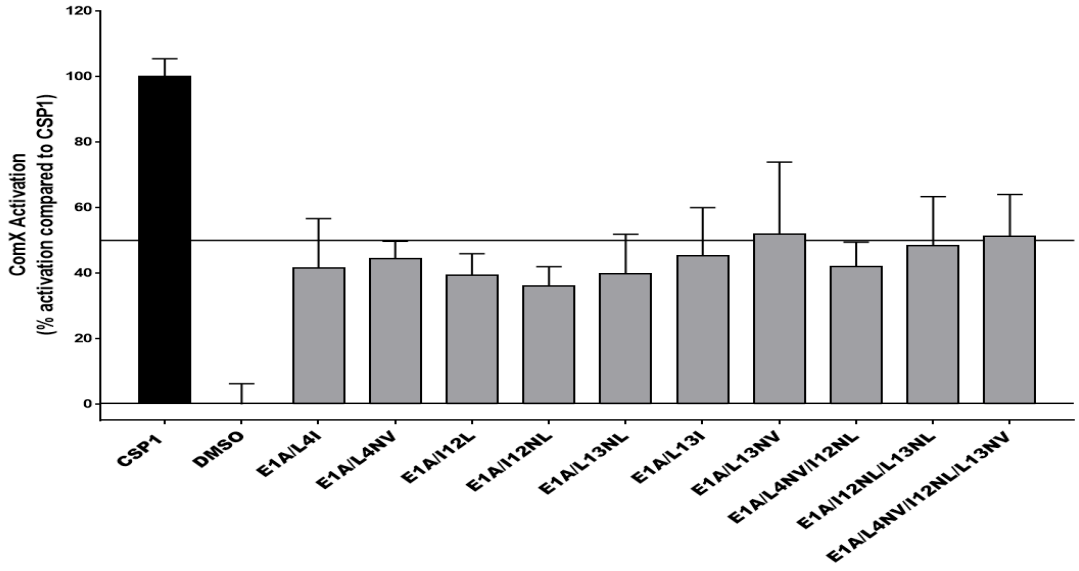


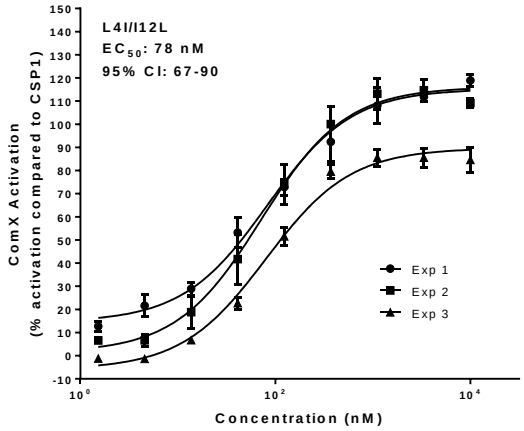
Figure S-3. Primary antagonism screening assay data for singly and multiply-substituted CSP1-E1A derivatives. Peptides that exhibited less than 50% activation were further evaluated to determine their IC_{50} values.

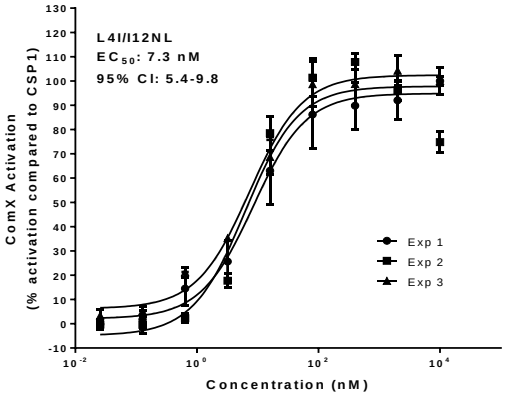
Agonism and antagonism dose response curves

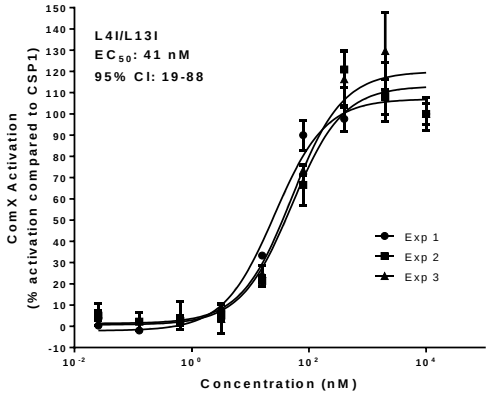
CSP1 analogs were tested to determine their EC_{50} or IC_{50} values over varying concentrations in the D39 pcomX::lacZ reporter strain. Each dose response experiment was performed in triplicate on three separate occasions (i.e., experiments (Exp.) #1-3; shown for each peptide below). Error bars indicate standard error of the mean of triplicate values. In each plot, the peptide, as well as its EC_{50} or IC_{50} value (in nM) and 95% confidence interval (95% CI) values (in nM), are indicated at top left or top right.

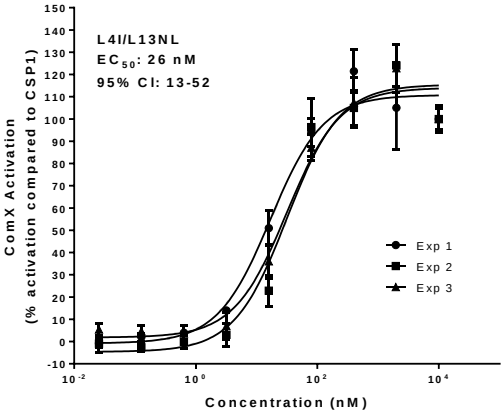
S. pneumoniae D39 pcomX::lacZ (ComD1)

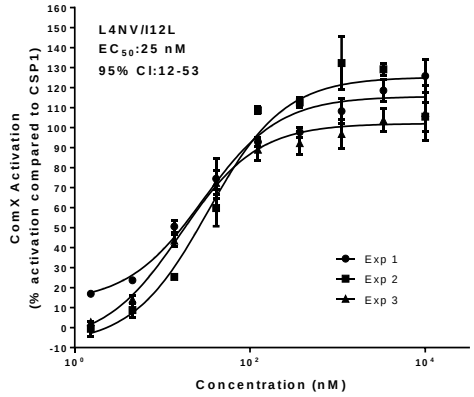
Activation dose response curves

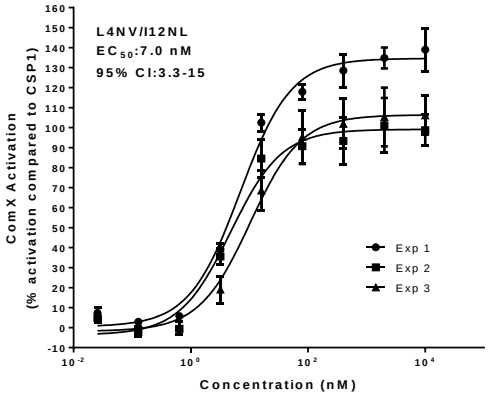


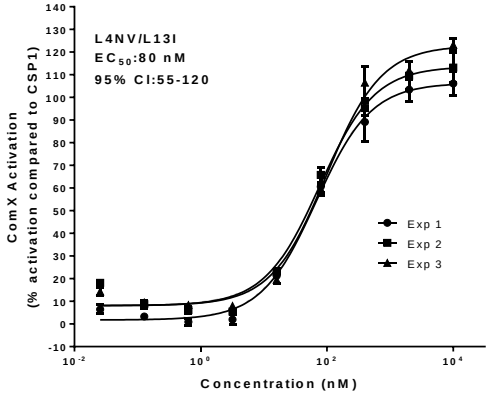


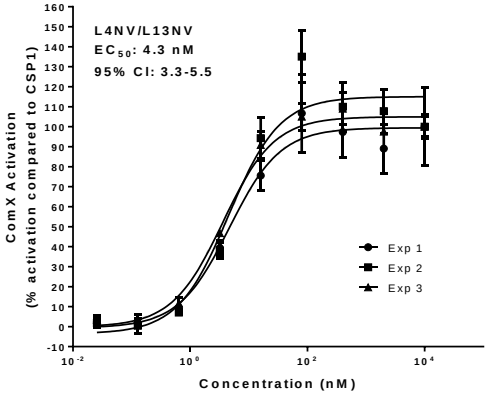


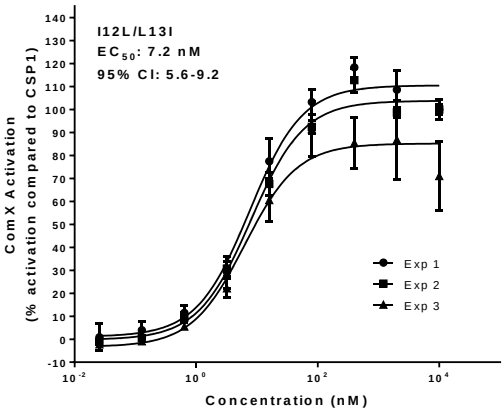


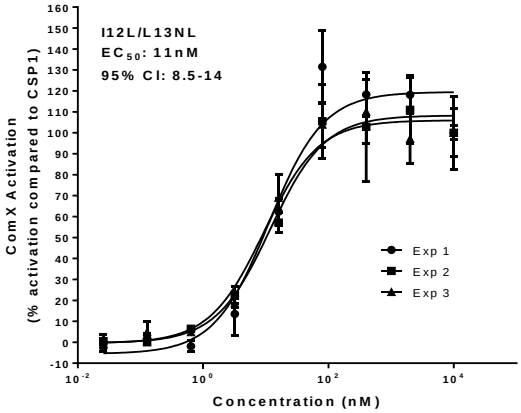

L4NV/I12L
EC50:25 nM
95% CI:12-53
ComX Activation
(% activation compared to CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

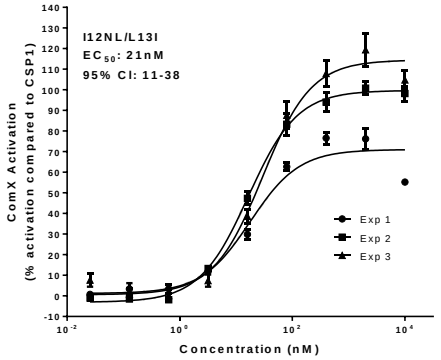

L4NV/I12NL
EC50:7.0 nM
95% CI:3.3-15
ComX Activation
(% activation compared to CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

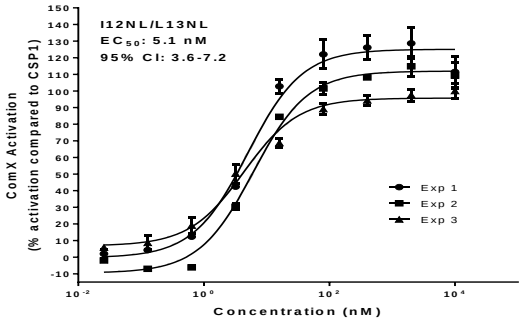

L4NV/L13I
EC50:80 nM
95% CI:55-120
ComX Activation
(% activation compared to CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

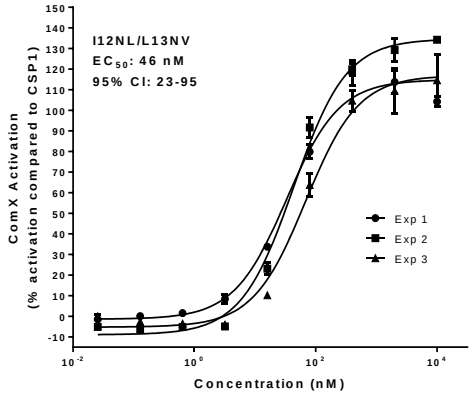

L4NV/L13NV
EC50: 4.3 nM
95% CI: 3.3-5.5
ComX Activation
(% activation compared to CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

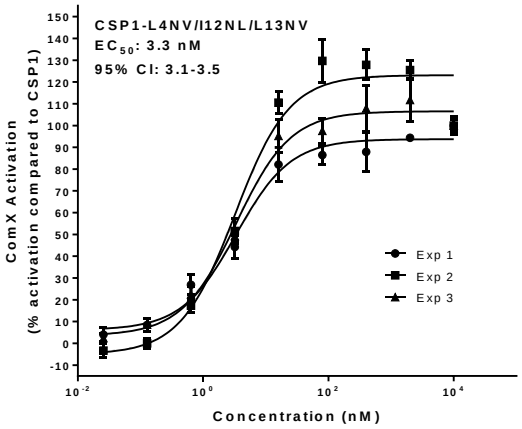

I12L/L13I
EC50: 7.2 nM
95% CI: 5.6-9.2
ComX Activation
(% activation compared to CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3


I12L/L13NL
EC50: 11nM
95% CI: 8.5-14
ComX Activation
(% activation compared to CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

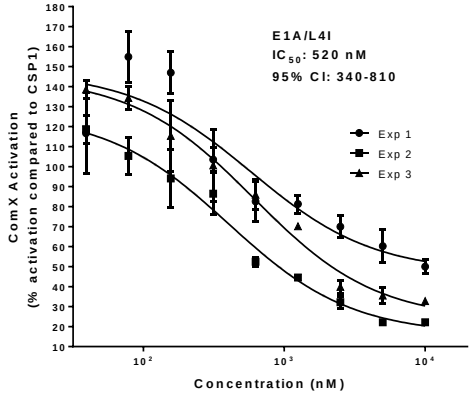


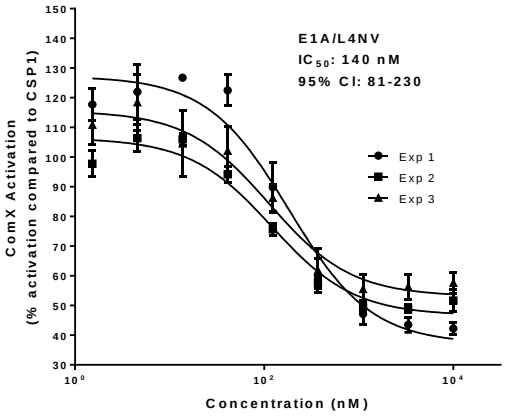


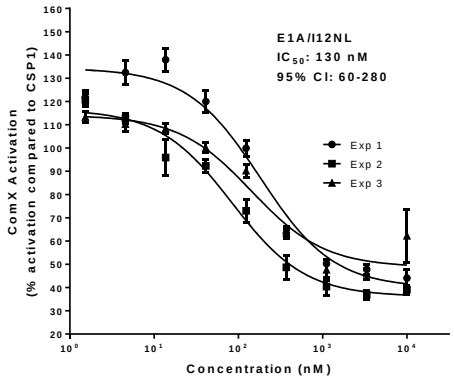


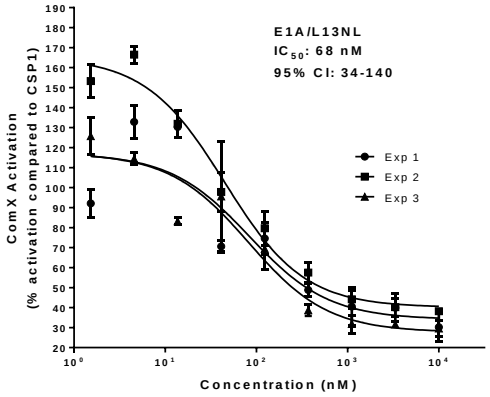


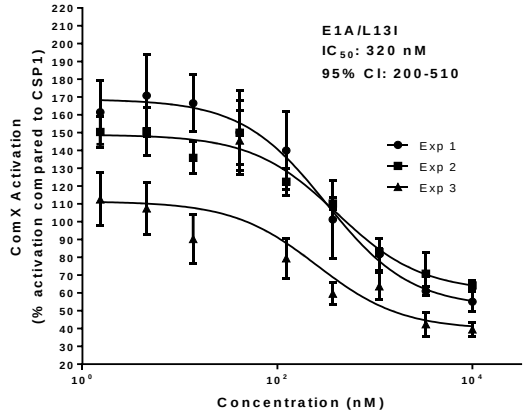
Inhibition dose response curves

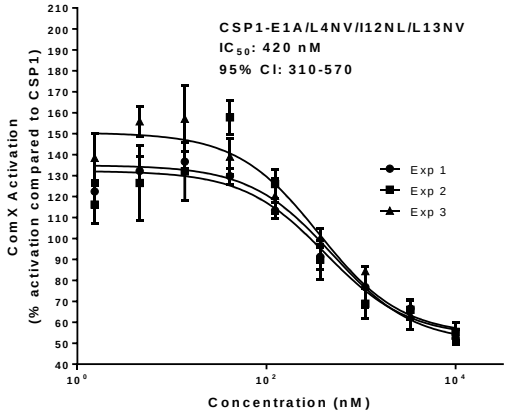


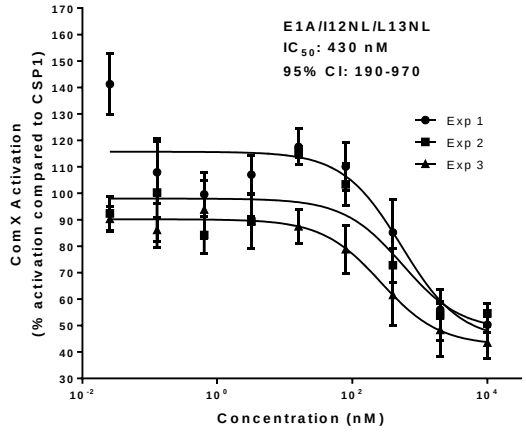



E1A/I12NL
IC50: 130 nM
95% CI: 60-280
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to CSP1)
Concentration (nM)


E1A/L13NL
IC50: 68 nM
95% CI: 34-140
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to CSP1)
Concentration (nM)


E1A/L13I
IC50: 320 nM
95% CI: 200-510
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to CSP1)
Concentration (nM)


CSP1-E1A/L4NV/I12NL/L13NV
IC50: 420 nM
95% CI: 310-570
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to CSP1)
Concentration (nM)


E1A/I12NL/L13NL
IC50: 430 nM
95% CI: 190-970
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to CSP1)
Concentration (nM)

Appendix 2: Optimizing CSP1 Analogs for Modulating Quorum Sensing in *Streptococcus pneumoniae* with Bulky, Hydrophobic Nonproteogenic Amino Acid Substitutions

[a]Reprinted with permission from Milly, T. A.; Buttner, A. R.; Rieth, N.; Hutnick, E.; Engler, E. R.; Campanella, A. R.; Lella, M.; Bertucci, M. A.; Tal-Gan, Y. Optimizing CSP1 analogs for modulating quorum sensing in Group 1 *Streptococcus pneumoniae* with bulky, hydrophobic nonproteogenic amino acid substitutions.. *RSC Chem Biol* **2022**, 3, 301-311.

Optimizing CSP1 Analogs for Modulating Quorum Sensing in *Streptococcus pneumoniae* with Bulky, Hydrophobic Nonproteogenic Amino Acid Substitutions

Supporting Information

Tahmina A. Milly,[†] Alec R. Buttner,[§] Naomi Rieth,[§] Elizabeth Hutnick,[§] Emilee R. Engler,[§] Alexandra R. Campanella,[‡] Muralikrishna Lella,[†] Michael A. Bertucci*[,‡] and Yftah Tal-Gan,*[,†]

[†] Department of Chemistry, University of Nevada, Reno, 1664 North Virginia Street, Reno, Nevada, 89557, United States

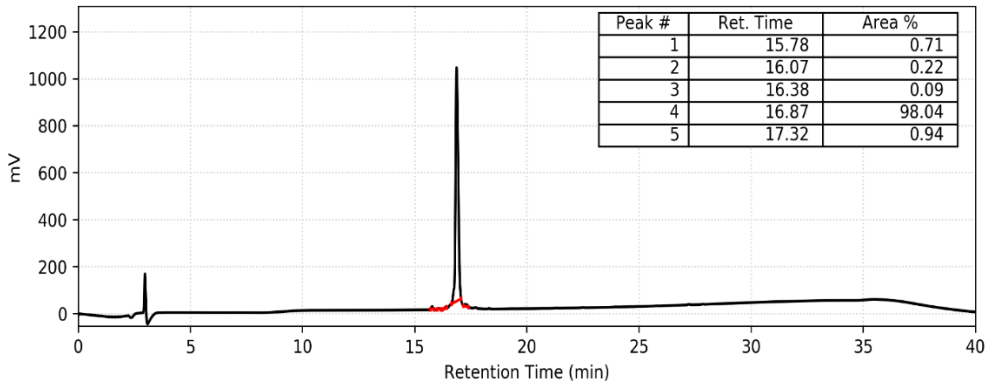
[‡] Department of Chemistry, Lafayette College, 701 Sullivan Rd., Easton, PA 18042, United States

[§] Department of Chemistry, Moravian University, 1200 Main St., Bethlehem, PA 18018, United States

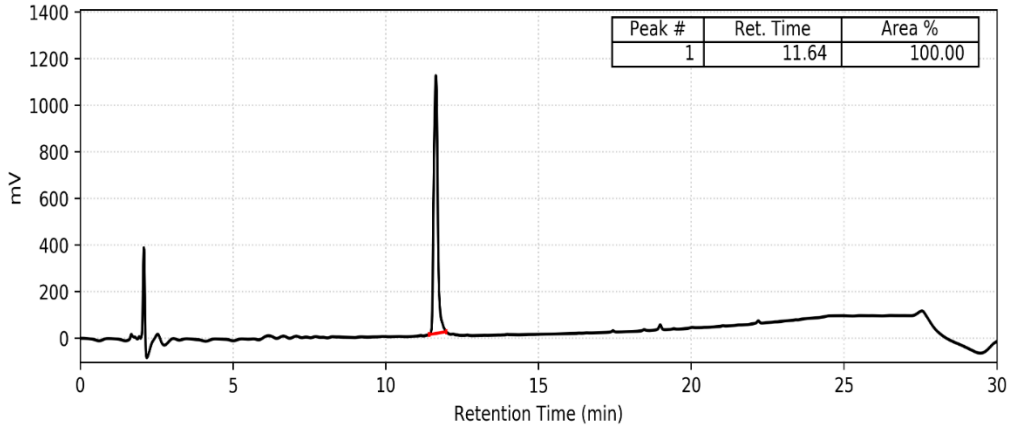
[*] To whom correspondence should be addressed. ytalgan@unr.edu, bertuccm@lafayette.edu

HPLC Traces for CSP1 Analogs

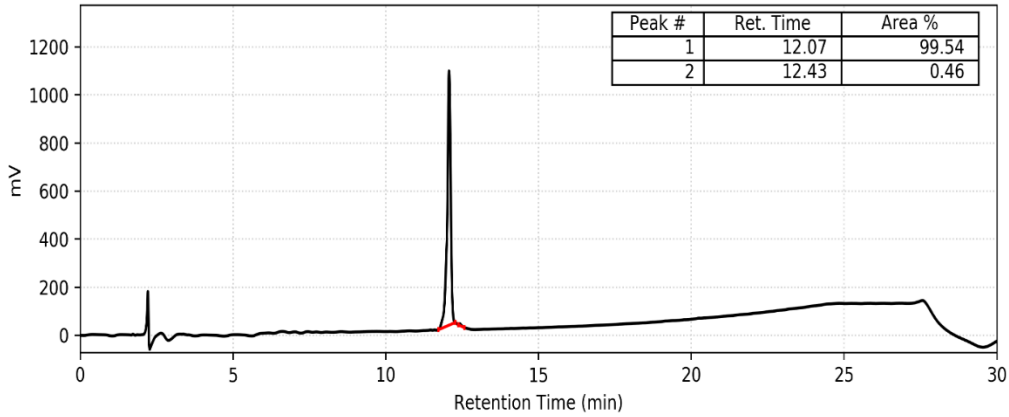
CSP1-L4Cha



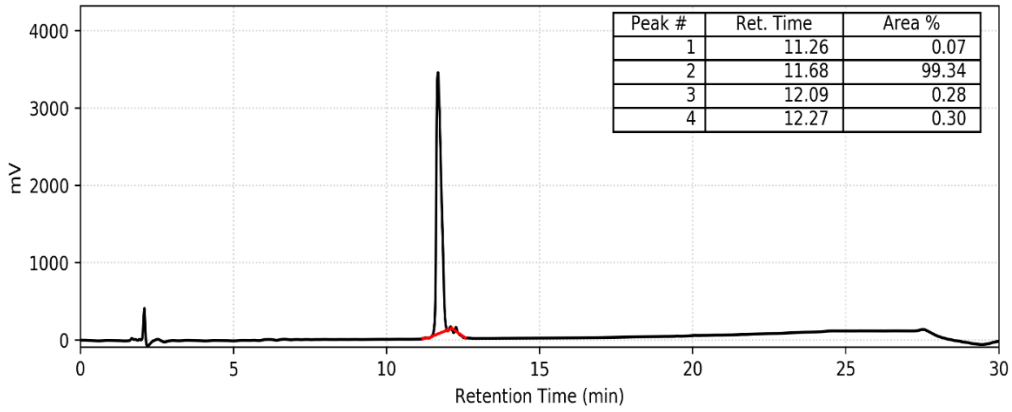
CSP1-L4HLeu



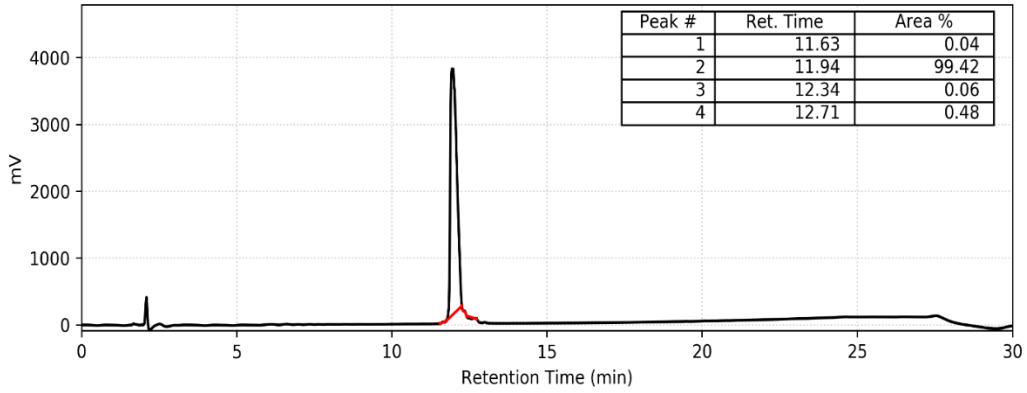
CSP1-F7Cha



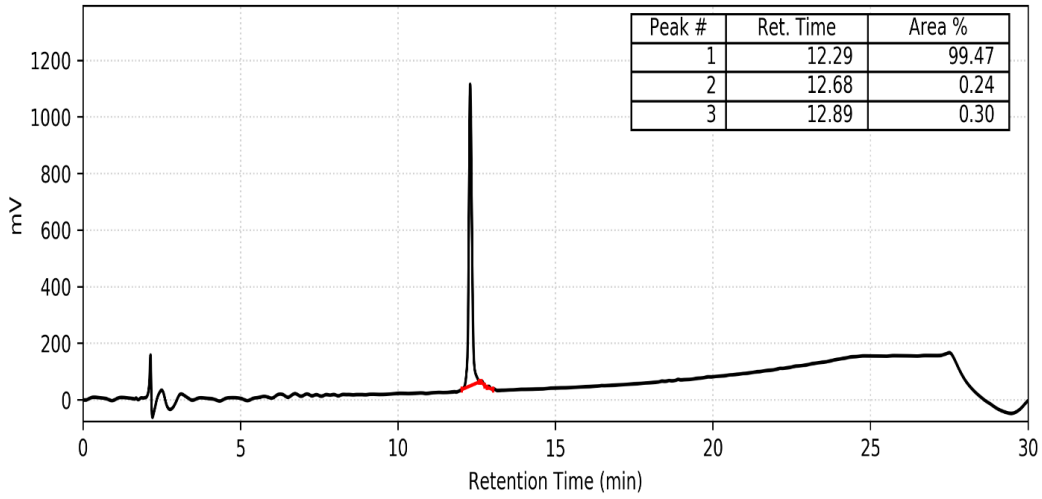
CSP1-F7HLeu



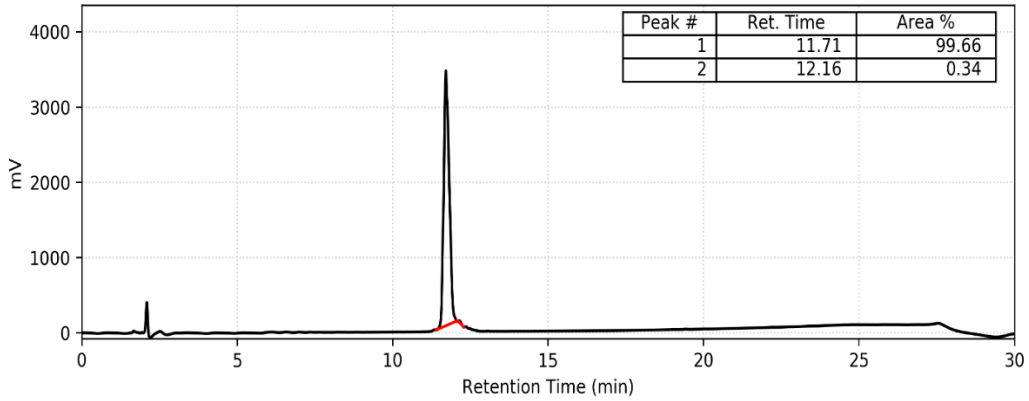
CSP1-F8Cha



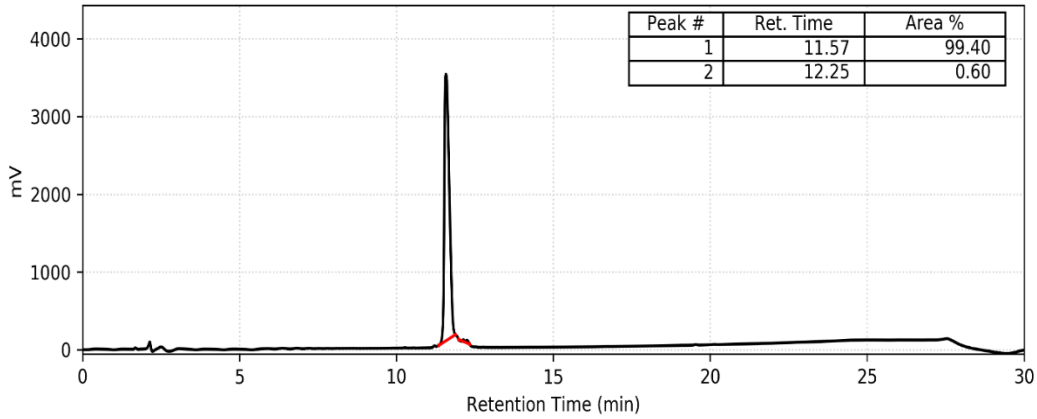
CSP1-F8HLeu



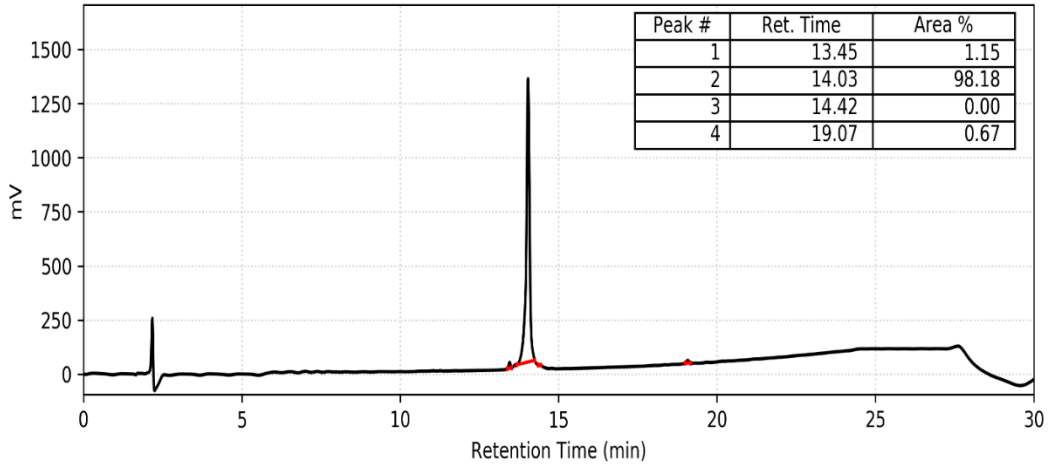
CSP1-F11Cha



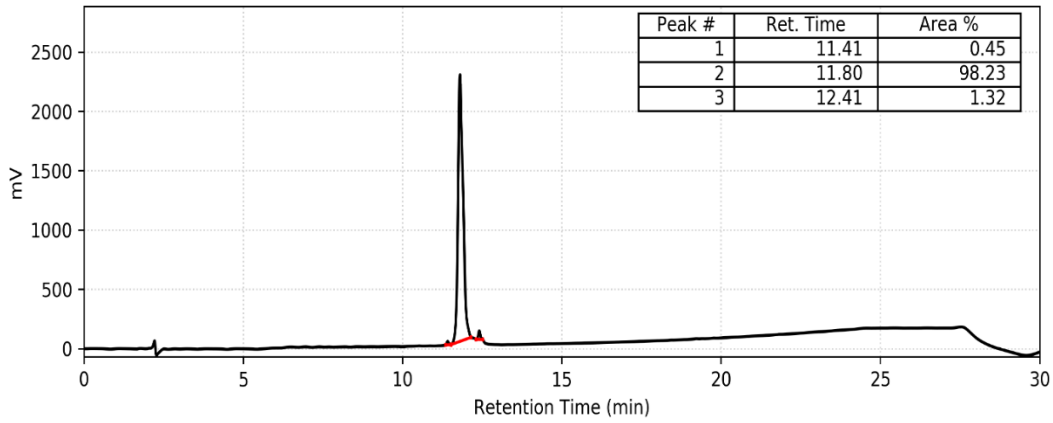
CSP1-F11HLeu



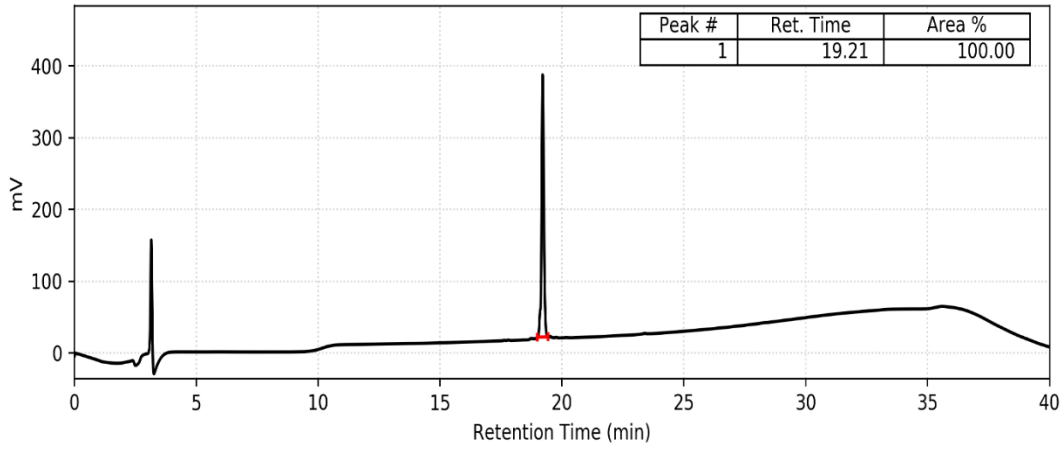
CSP1-I12Cha



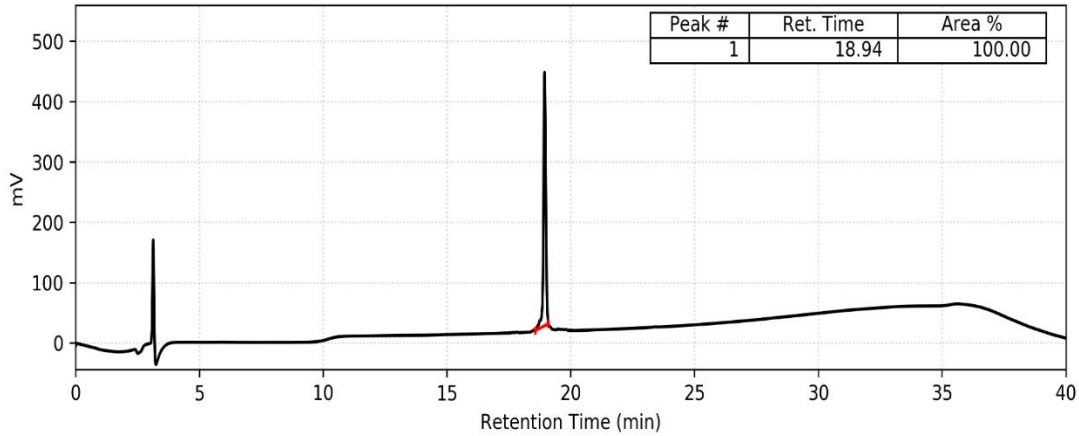
CSP1-I12HLeu



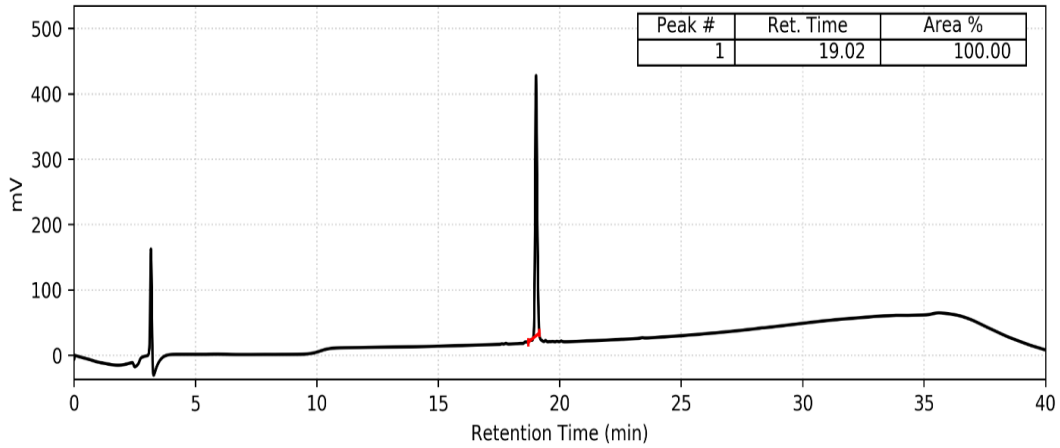
CSP1-F7Cha/F8Cha



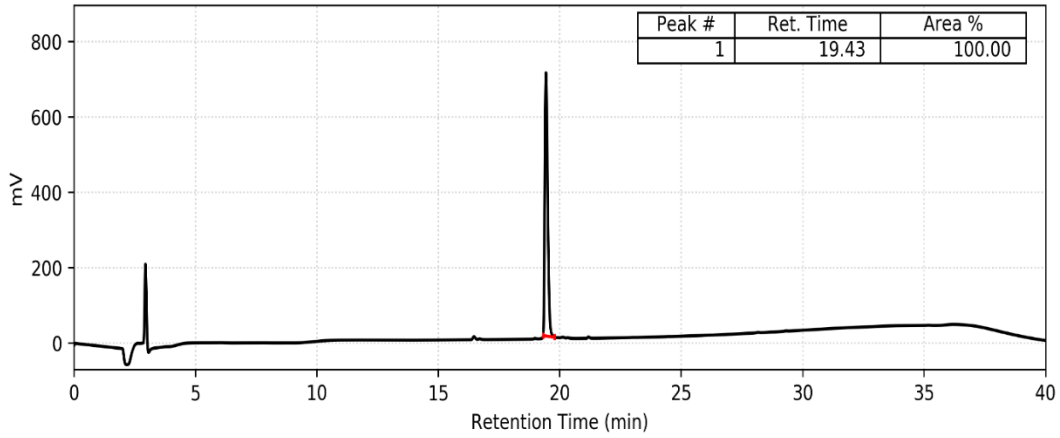
CSP1-F7Cha/F8HLeu



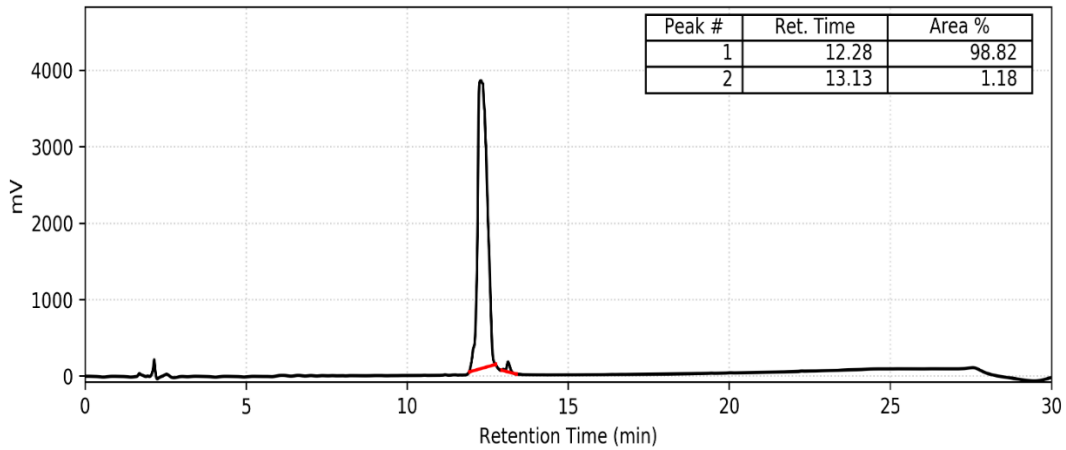
CSP1-F7HLeu/F8Cha



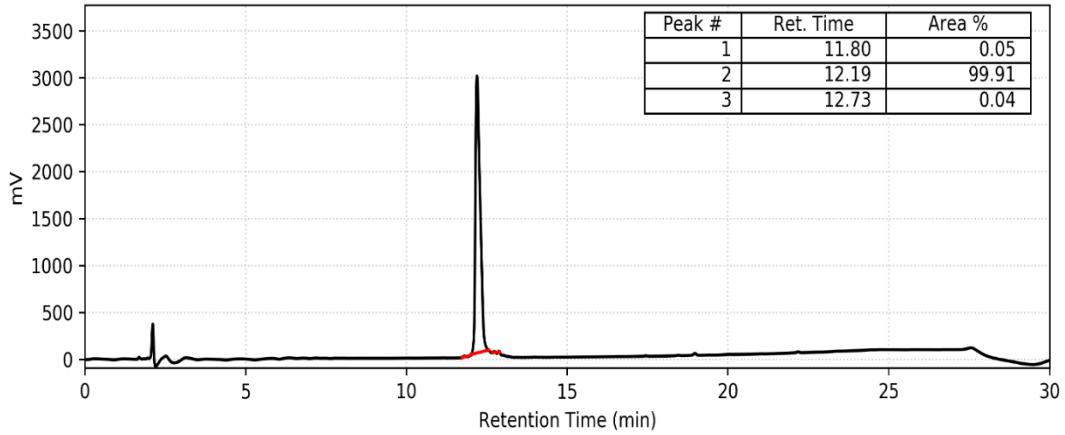
CSP1-F7HLeu/F8HLeu



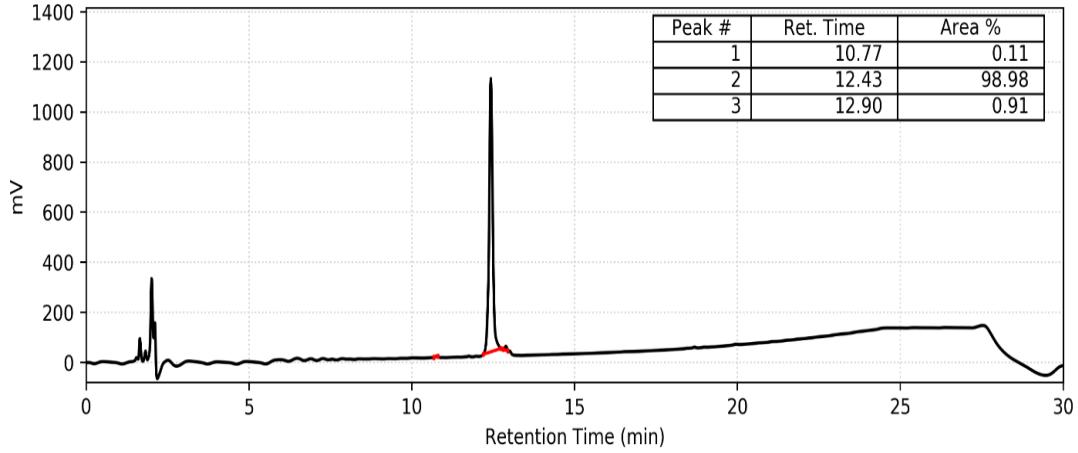
CSP1-F7Cha/I12Cha



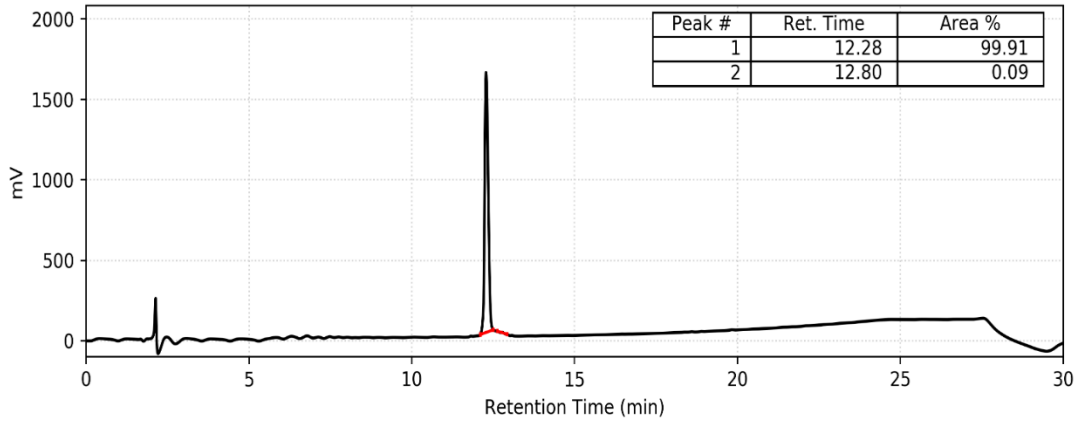
CSP1-F7Cha/I12HLeu



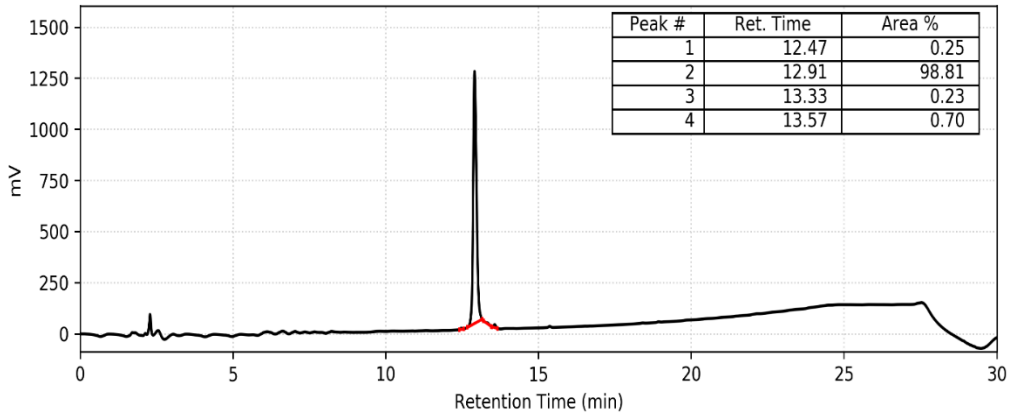
CSP1-F7HLeu/I12Cha



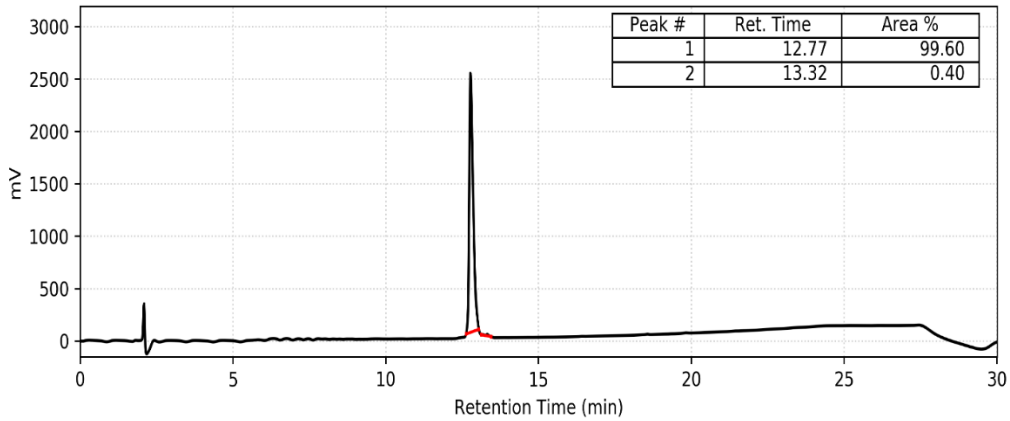
CSP1-F7HLeu/I12HLeu



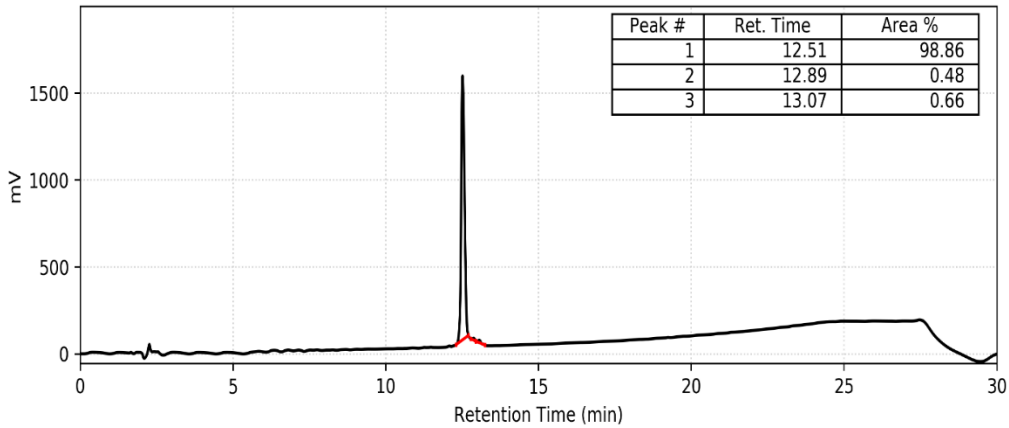
CSP1-F8Cha/I12Cha



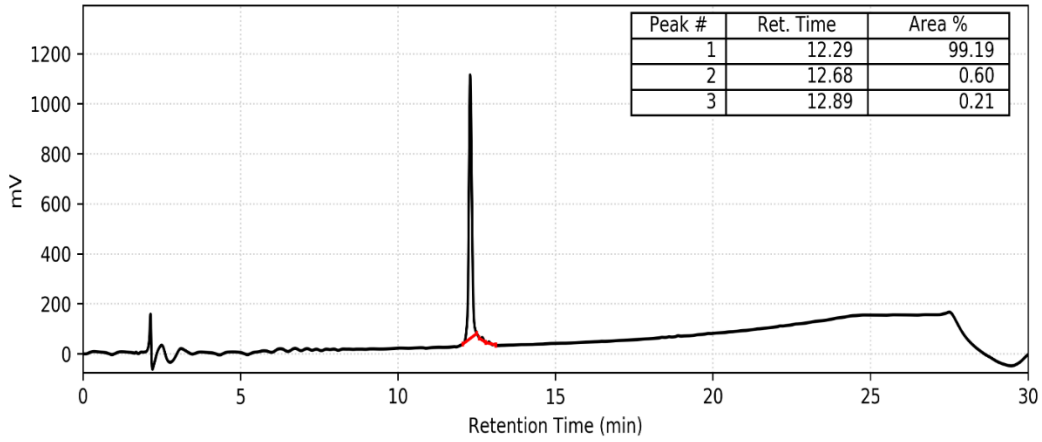
CSP1-F8Cha/I12HLeu



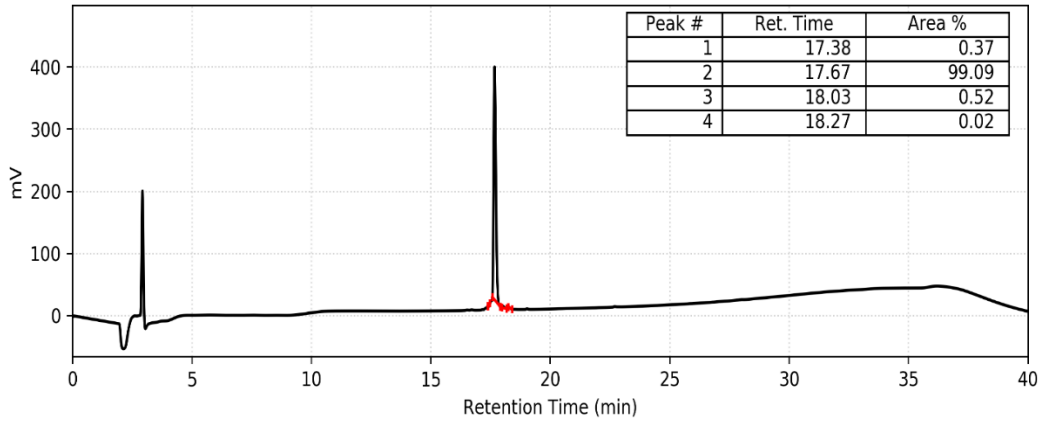
CSP1-F8HLeu/I12Cha



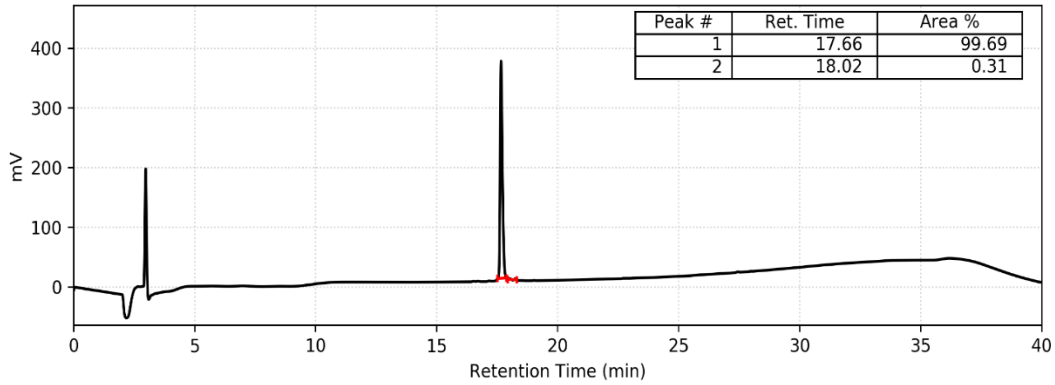
CSP1-F8HLeu/I12HLeu



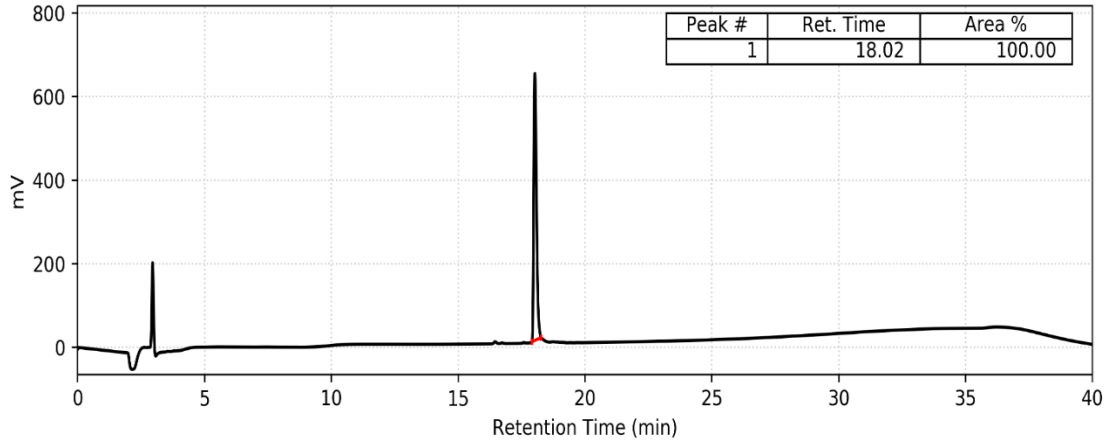
CSP1-E1A/F7Cha



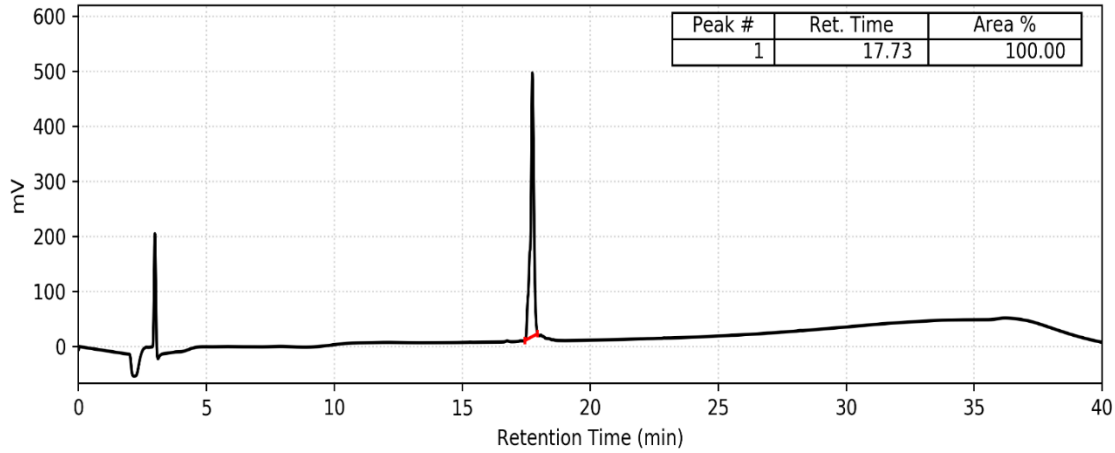
CSP1-E1A/F7HLeu



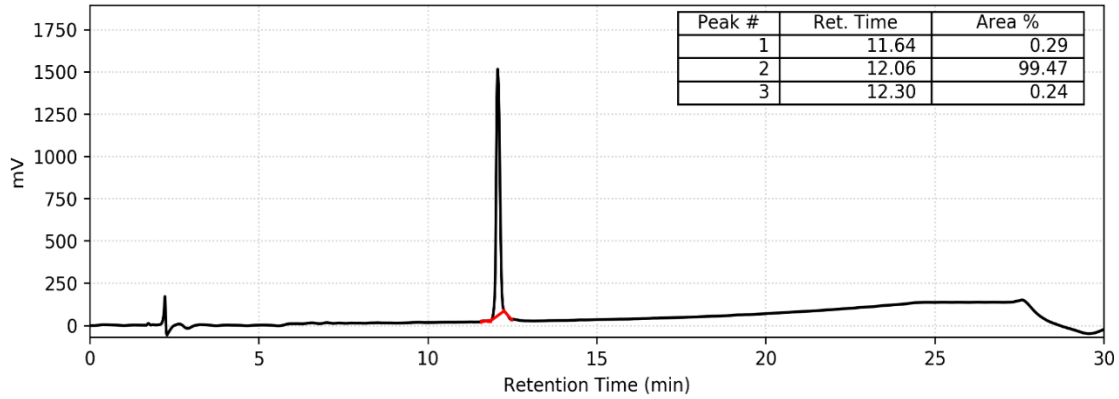
CSP1-E1A/F8Cha



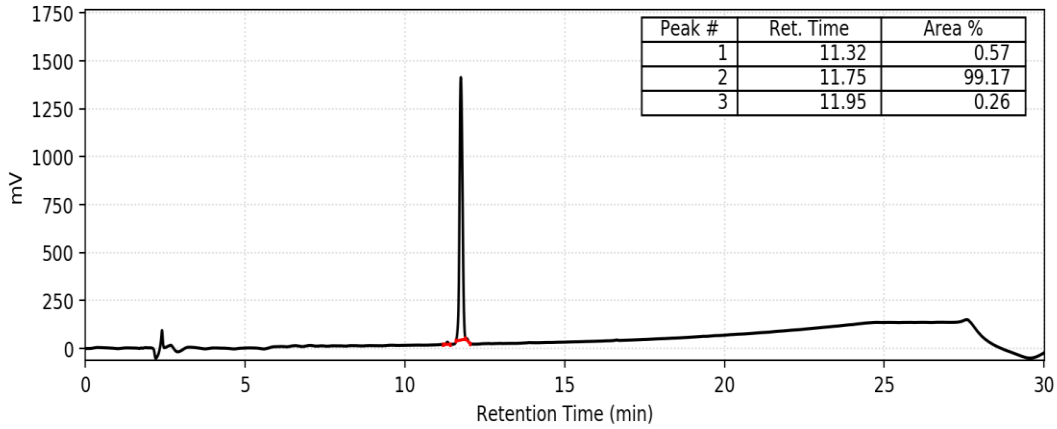
CSP1-E1A/F8HLeu



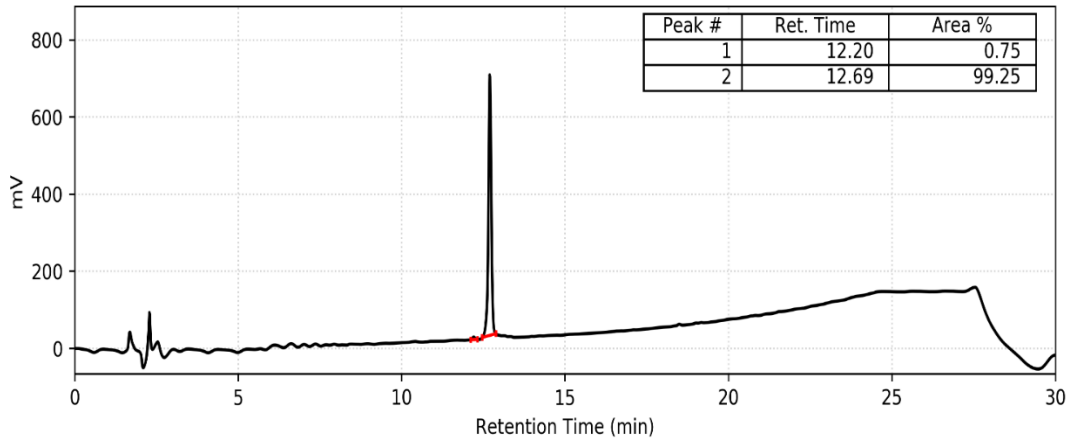
CSP1-E1A/I12Cha



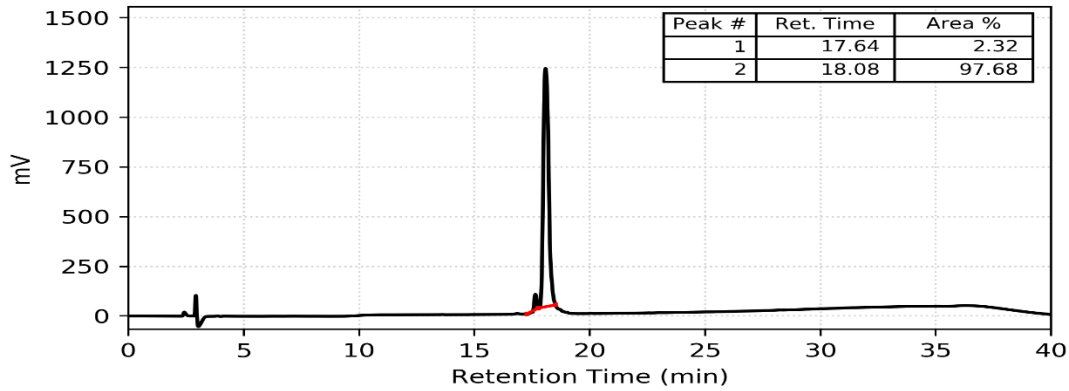
CSP1-E1A/I12HLeu



CSP1-E1A/F7Cha/I12Cha



CSP1-E1A/F7Cha/I12HLeu



MS and HPLC Data for CSP1 Analogs

Table S1. MS and HPLC data for CSP1 single and double mutant analogs.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
CSP1-L4Cha	1142.1528	1142.1529	≥98
CSP1-L4HLeu	1129.1449	1129.1451	≥99
CSP1-F7Cha	1125.1606	1125.1561	≥99
CSP1-F7HLeu	1112.1528	1112.1537	≥99
CSP1-F8Cha	1125.1606	1125.1616	≥99
CSP1-F8HLeu	1112.1528	1112.1535	≥99
CSP1-F11Cha	1125.1606	1125.1614	≥99
CSP1-F11HLeu	1112.1528	1112.1532	≥99
CSP1-I12Cha	761.7709*	761.7678*	≥98
CSP1-I12HLeu	1129.1449	1129.1437	≥98
CSP1-F7Cha/F8Cha	1128.1841	1128.1838	≥99
CSP1-F7Cha/F8HLeu	1115.1762	1115.1743	≥99
CSP1-F7HLeu/F8Cha	1115.1762	1115.1754	≥99
CSP1-F7HLeu/F8HLeu	1102.1684	1102.1718	≥99
CSP1-F7Cha/I12Cha	1145.1762	1145.1771	≥98
CSP1-F7Cha/I12HLeu	1132.1684	1132.1691	≥99
CSP1-F7HLeu/I12Cha	1132.1684	1132.1725	≥98
CSP1-F7HLeu/I12HLeu	1119.1606	1119.1620	≥99
CSP1-F8Cha/I12Cha	1145.1762	1145.1775	≥98
CSP1-F8Cha/I12HLeu	1132.1684	1132.1720	≥99
CSP1-F8HLeu/I12Cha	1132.1684	1132.1711	≥98
CSP1-F8HLeu/I12HLeu	1119.1606	1119.1622	≥99

EM = Exact Mass. See methods above, *$\mathbf{MH_3^{3+}}$.

Table S2. MS and HPLC data for CSP1-E1A mutant analogs.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
CSP1-E1A/F7Cha	1096.1578	1096.1562	≥99
CSP1-E1A/F7HLeu	1083.1500	1083.1493	≥99
CSP1-E1A/F8Cha	1096.1578	1096.1598	≥99
CSP1-E1A/F8HLeu	1083.1500	1083.1488	≥99
CSP1-E1A/I12Cha	1113.1500	1113.1506	≥99
CSP1-E1A/I12HLeu	1100.1422	1100.1461	≥99
CSP1-E1A/F7Cha/I12Cha	1116.1735	1116.1759	≥99
CSP1-E1A/F7Cha/I12HLeu	1103.1657	1103.1672	≥97

EM = Exact Mass. See methods above.

Primary Reporter Gene Assay Data

S. pneumoniae D39pcomX::lacZ (ComD1)

Agonism assays were performed at 10 µM concentration of synthetic CSP1 analogs. CSP1 was used as the positive control (100%) while DMSO as the negative control (0%). Percent (%) ComD1 activation was measured by normalizing the Miller units obtained for each peptide to that of the native CSP1. All peptides were screened in triplicate over three separate trials. Error bars indicate standard error of the mean of nine values.

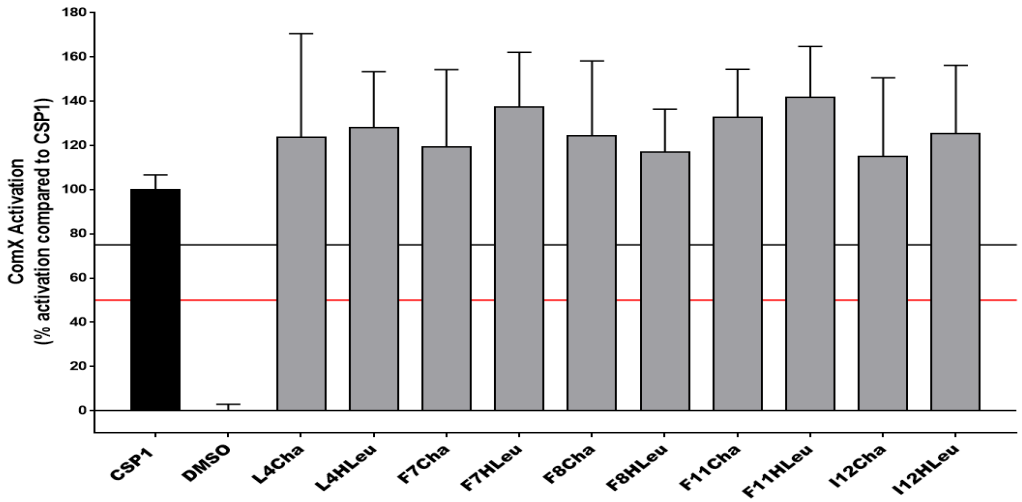


Figure S-1. Primary agonism screening assay data for the CSP1 single mutant analogs. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} values.

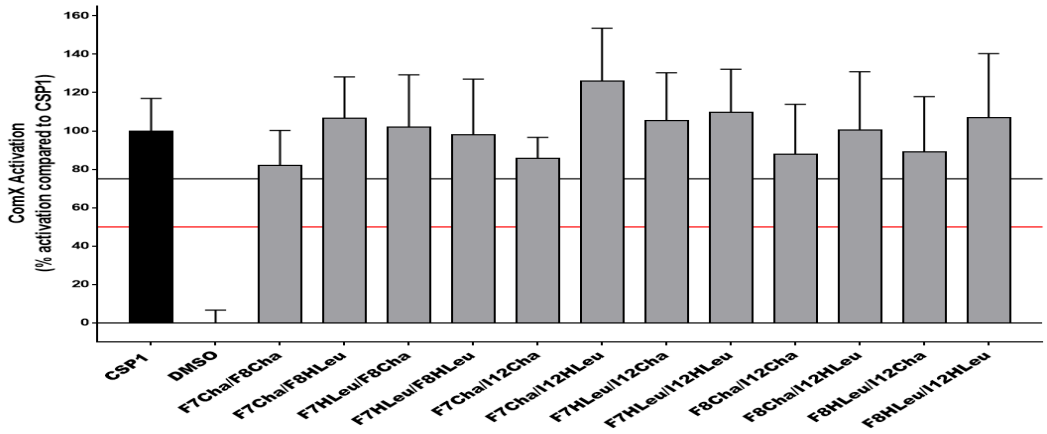


Figure S-2. Primary agonism screening assay data for the CSP1 double mutant analogs. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} values.

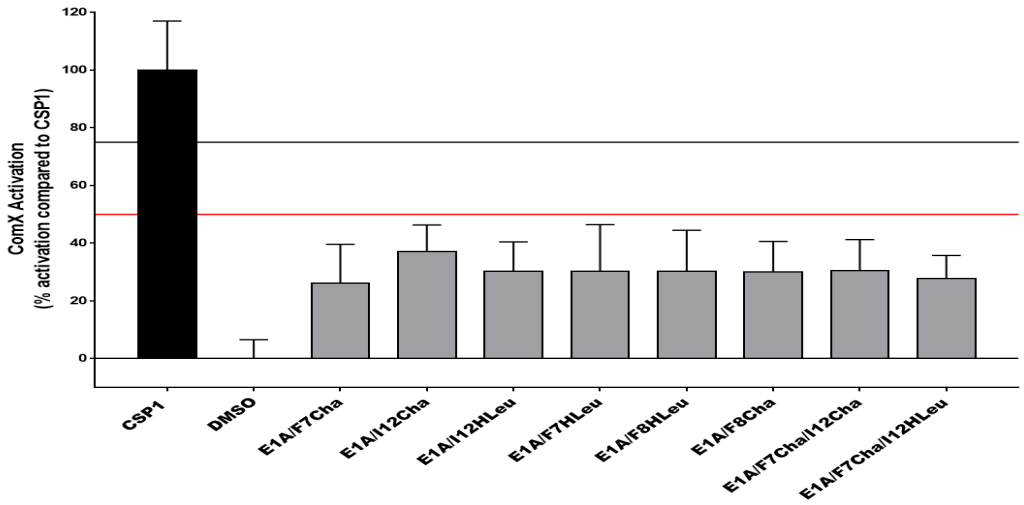


Figure S-3. Primary agonism screening assay data for the CSP1-E1A mutant analogs. None of the peptides exhibited activation of the ComD1 receptor and peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

Antagonism assays were performed at 10 μM concentration of peptides against 50 nM concentration of CSP1. CSP1 (50 nM) was used as the positive control (100%) while DMSO as the negative control (0%). Percent (%) *comX* activation was measured by normalizing the Miller units obtained for each peptide to that of CSP1. All peptides were screened in triplicate over three separate trials. Error bars indicate standard error of the mean of nine values.

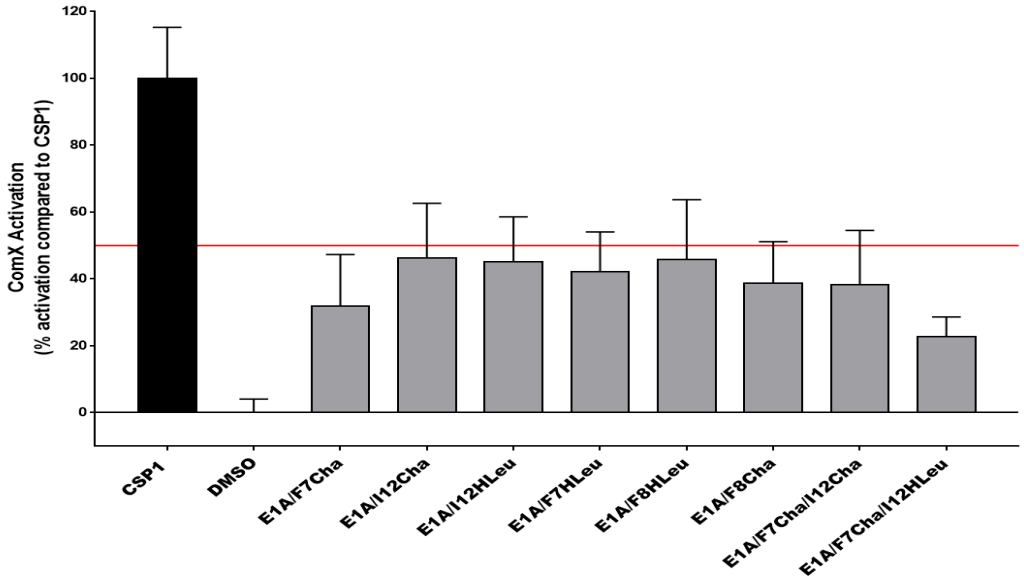


Figure S-4. Primary antagonism screening assay data for the CSP1-E1A mutant analogs. Peptides that exhibited less than 50% activation were further evaluated to determine their IC_{50} values.

***S. pneumoniae* TIGR4 pcomX::lacZ (ComD2)**

Agonism assays were performed at 10 μM concentration. CSP2 was used as the positive control (100%) while DMSO as the negative control (0%). Percent (%) ComD2 activation was measured by normalizing the Miller units obtained for each peptide to that of the native CSP2. All peptides were screened in triplicate over three separate trials. Error bars indicate standard error of the mean of nine values.

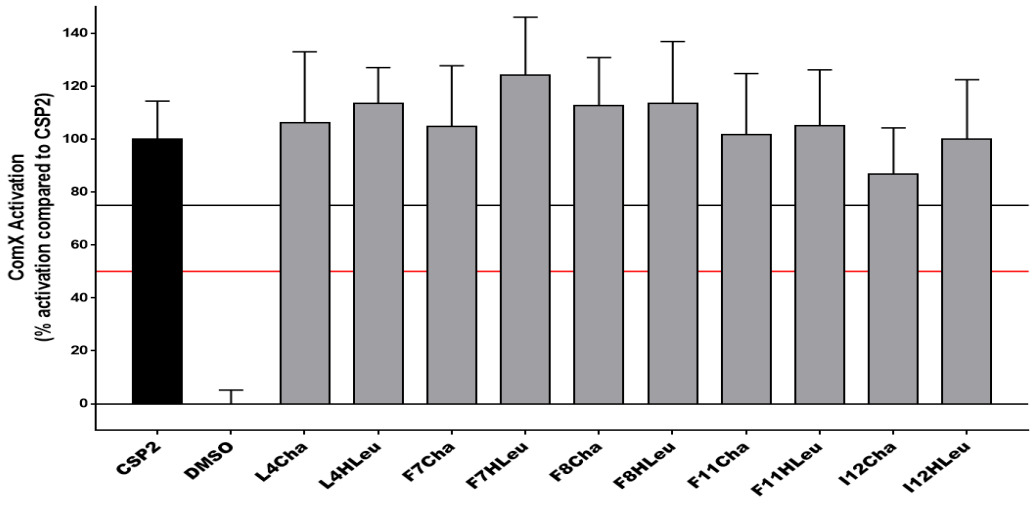


Figure S-5. Primary agonism screening assay data for the CSP1 single mutant analogs. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} values.

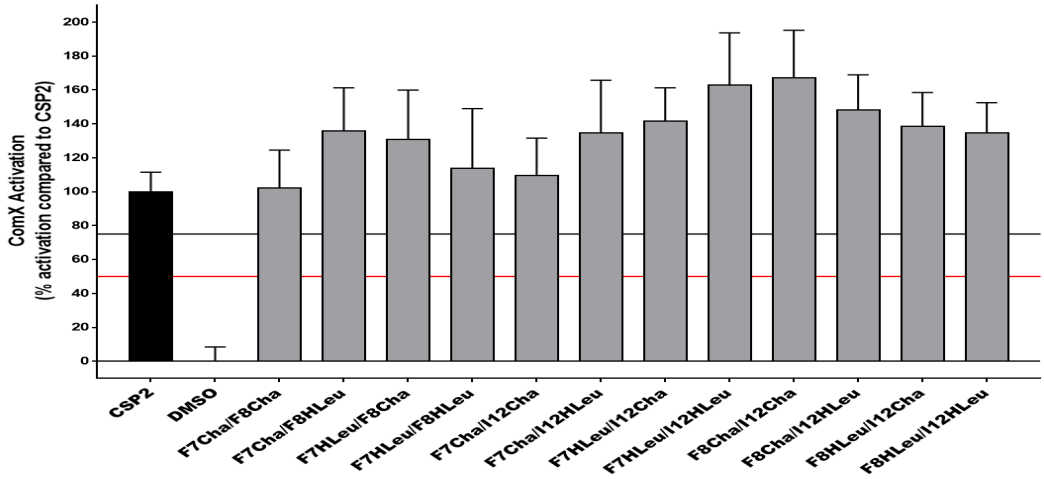


Figure S-6. Primary agonism screening assay data for the CSP1 double mutant analogs. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} values.

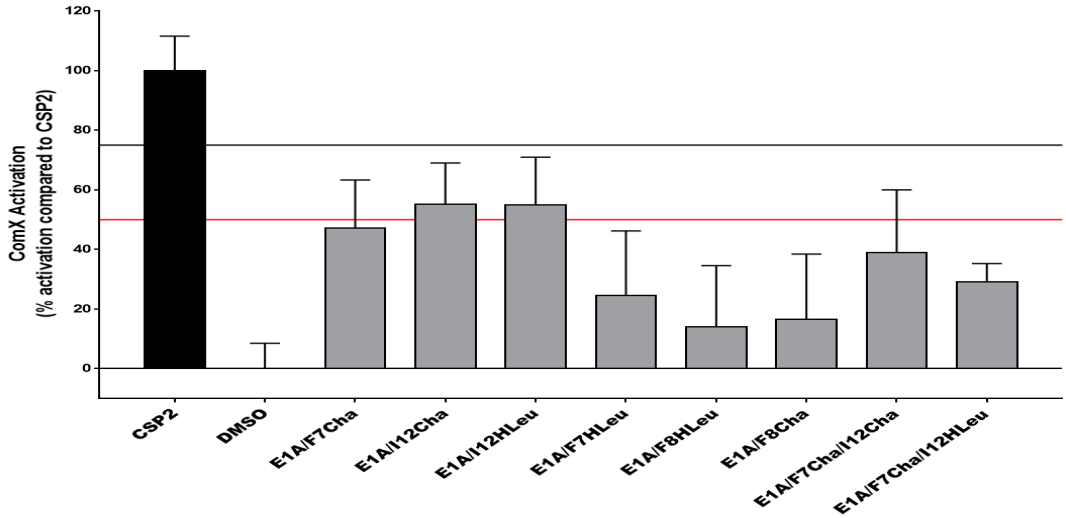


Figure S-7. Primary agonism screening assay data for the CSP1-E1A mutant analogs. None of the peptides exhibited activation of the ComD2 receptor and peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

Antagonism assays were performed at 10 μM concentration of peptides against 250 nM concentration of CSP2. CSP2 (250 nM) was used as the positive control (100%) while DMSO as the negative control (0%). Percent (%) *comX* activation was measured by normalizing the Miller units obtained for each peptide to that of CSP2. All peptides were screened in triplicate over three separate trials. Error bars indicate standard error of the mean of nine values.

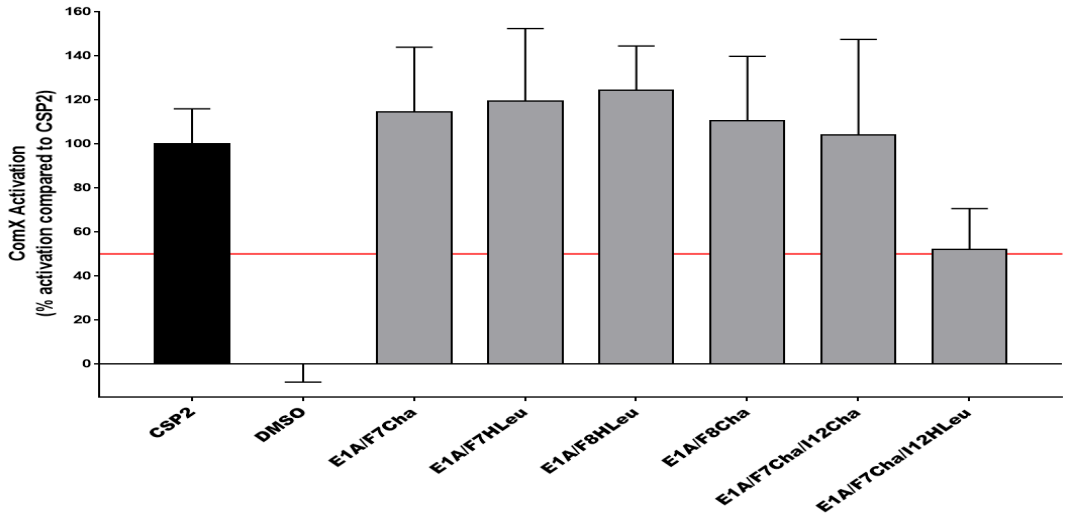


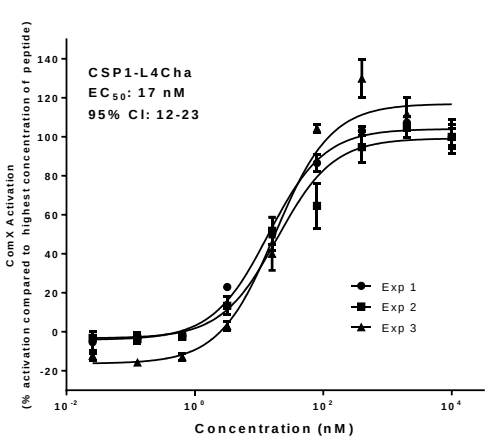
Figure S-8. Primary antagonism screening assay data for the CSP1-E1A mutant analogs. None of the peptides exhibited inhibition of the ComD2 receptor.

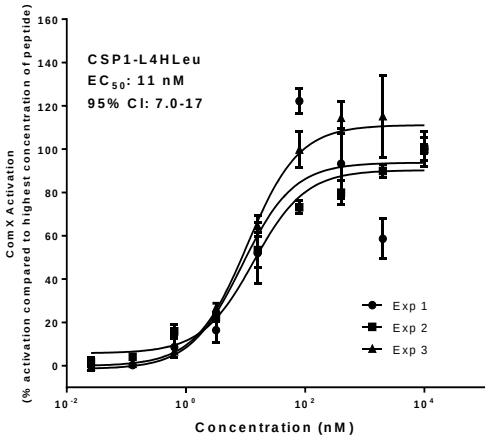
Agonism and Antagonism Dose Response Curves

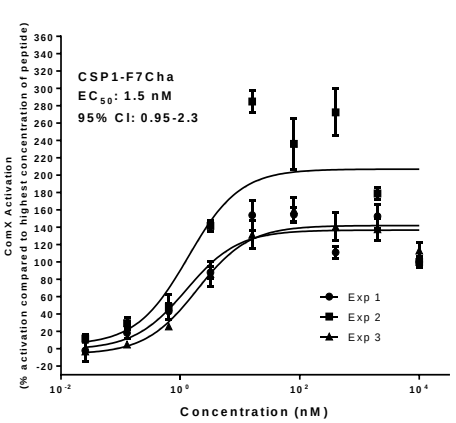
CSP1 analogs were tested to determine their EC_{50} or IC_{50} values over varying concentrations in the two indicated *S. pneumoniae* β-galactosidase reporter strains. Each dose response curve is representative of three independent experiments performed on three separate occasions (i.e., experiments (Exp.) #1-3; shown for each peptide below). Error bars represent standard error of the mean of triplicate values. In each plot, the peptide, as well as its EC_{50} or IC_{50} value (in nM) and 95% confidence interval (95% CI) values (in nM), are indicated at top left or top right.

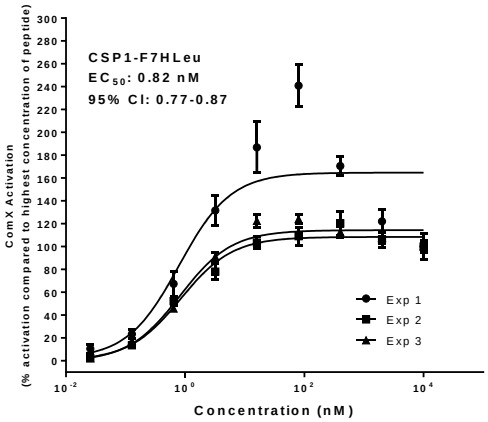
S. pneumoniae D39 pcomX::lacZ (ComD1)

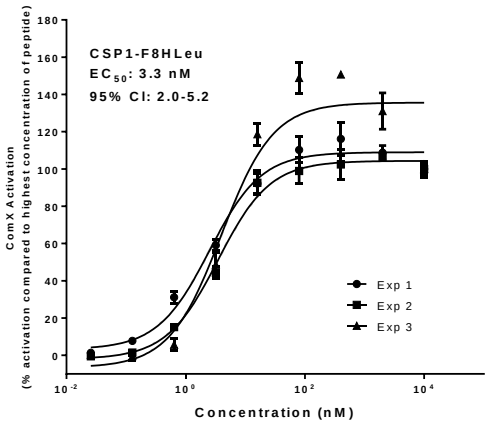
Activation dose response curves

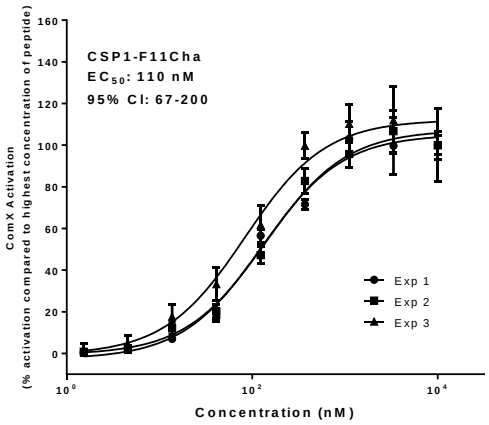


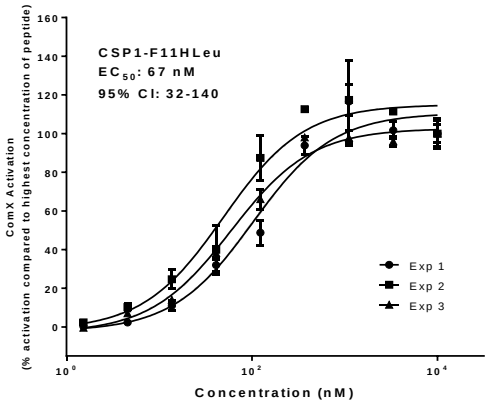


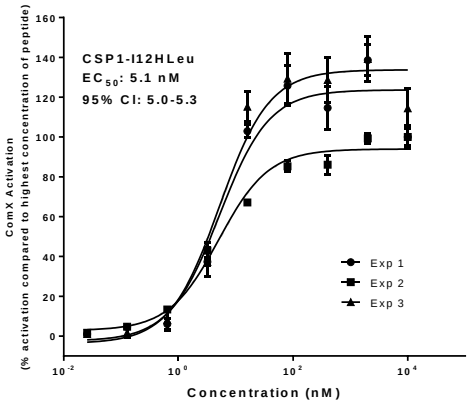


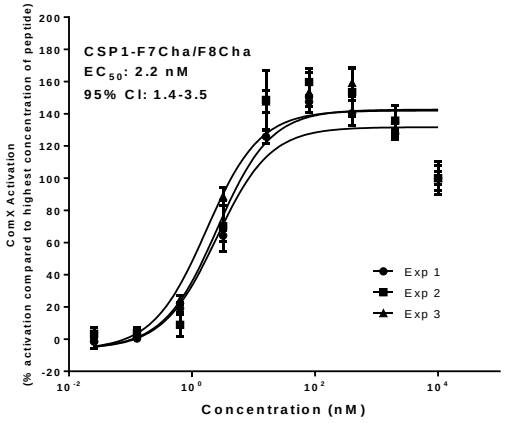


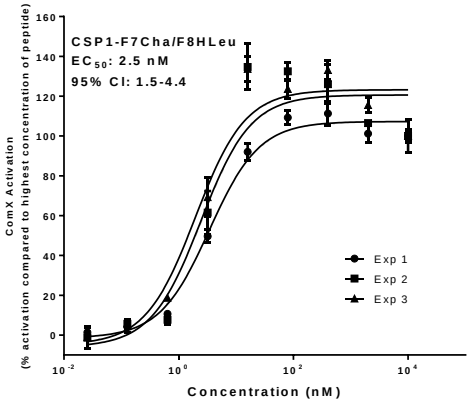

CSP1-F8HLeu
EC50: 3.3 nM
95% CI: 2.0-5.2
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

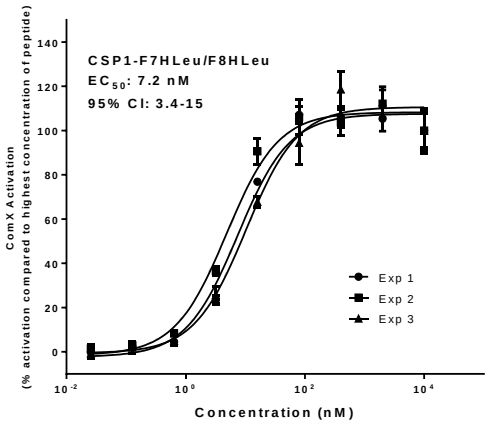

CSP1-F11Cha
EC50: 110 nM
95% CI: 67-200
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

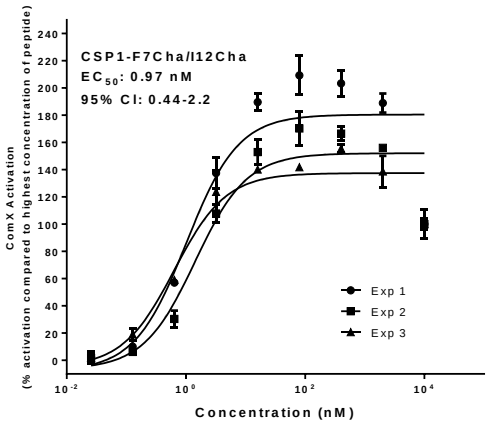

CSP1-F11HLeu
EC50: 67 nM
95% CI: 32-140
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

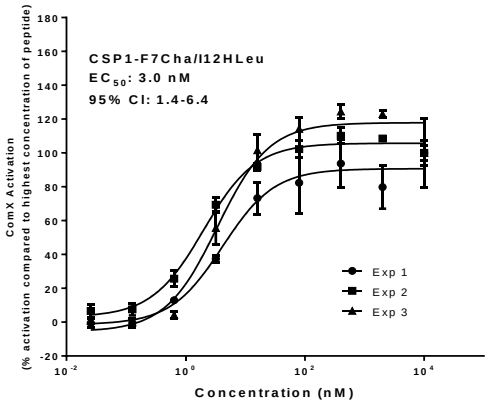

CSP1-I12HLeu
EC50: 5.1 nM
95% CI: 5.0-5.3
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

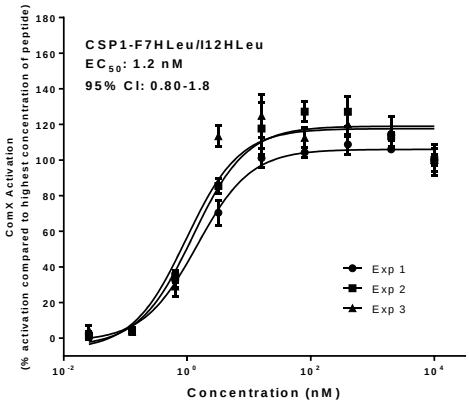

CSP1-F7Cha/F8Cha
EC50: 2.2 nM
95% CI: 1.4-3.5
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

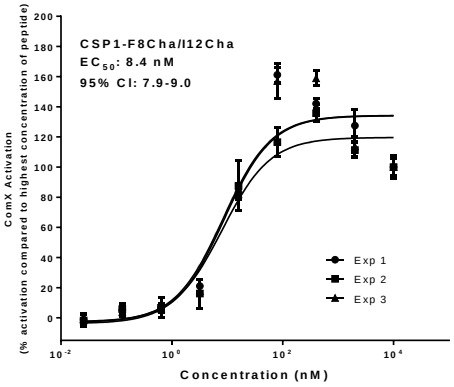

CSP1-F7Cha/F8HLeu
EC50: 2.5 nM
95% CI: 1.5-4.4
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

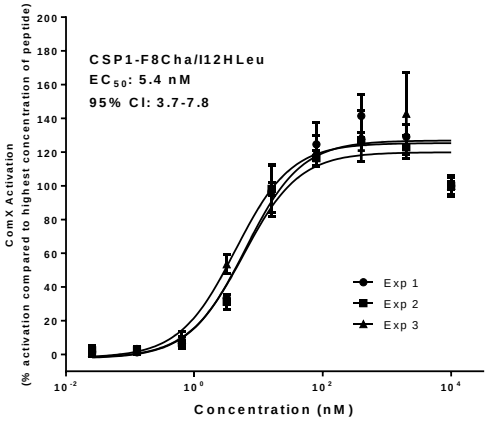

CSP1-F7HLeu/F8HLeu
EC50: 7.2 nM
95% CI: 3.4-15
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

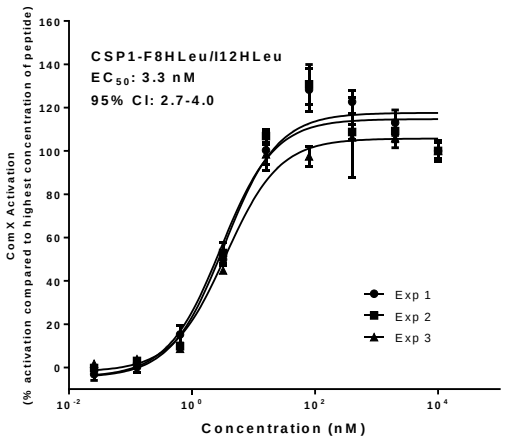

CSP1-F7Cha/I12Cha
EC50: 0.97 nM
95% CI: 0.44-2.2
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3


CSP1-F7Cha/I12HLeu
EC50: 3.0 nM
95% CI: 1.4-6.4
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

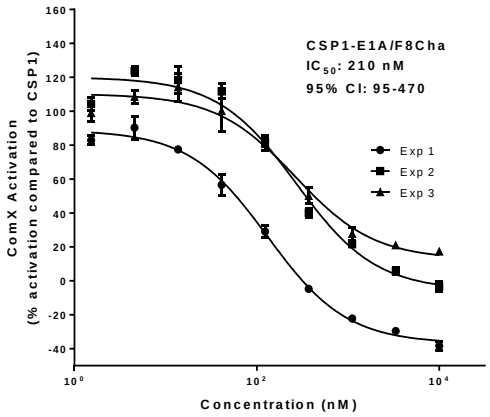

CSP1-F7HLeu/I12HLeu
EC50: 1.2 nM
95% CI: 0.80-1.8
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

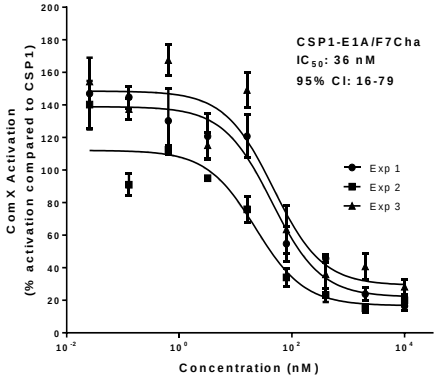

CSP1-F8Cha/I12Cha
EC50: 8.4 nM
95% CI: 7.9-9.0
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

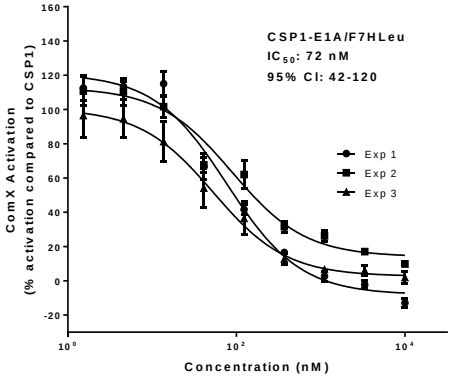

CSP1-F8Cha/I12HLeu
EC50: 5.4 nM
95% CI: 3.7-7.8
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

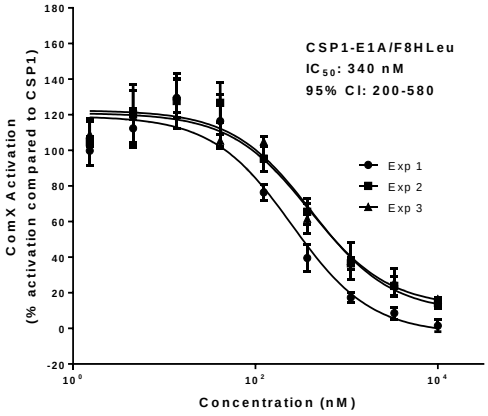


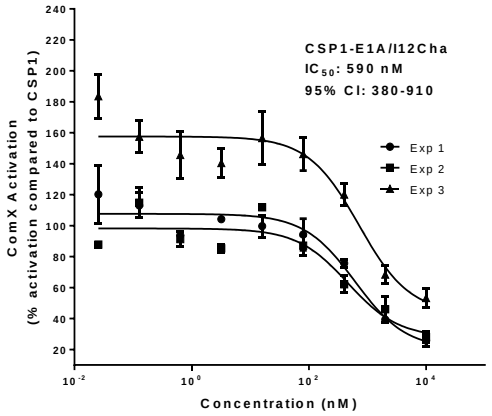
Inhibition dose response curves

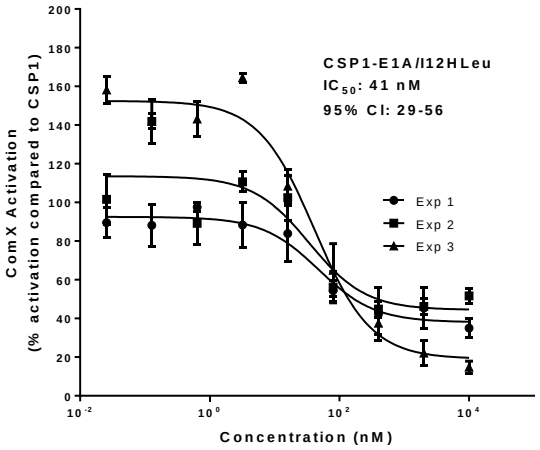


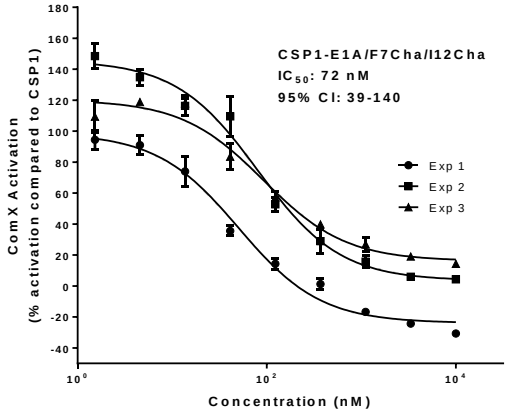


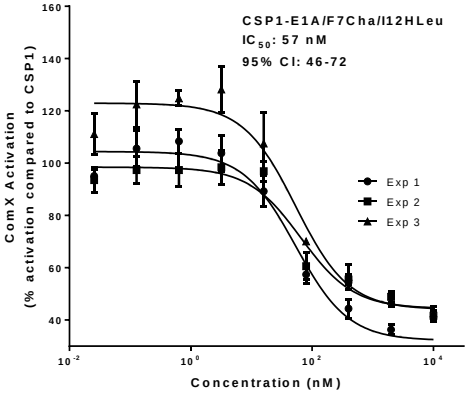






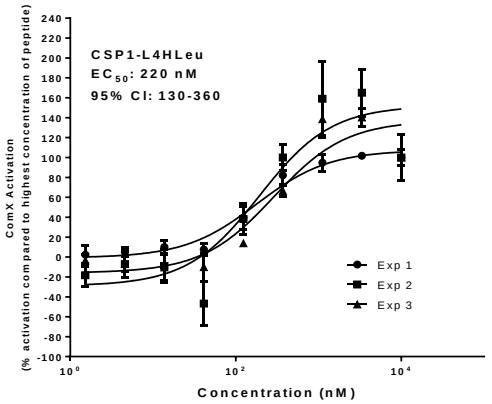


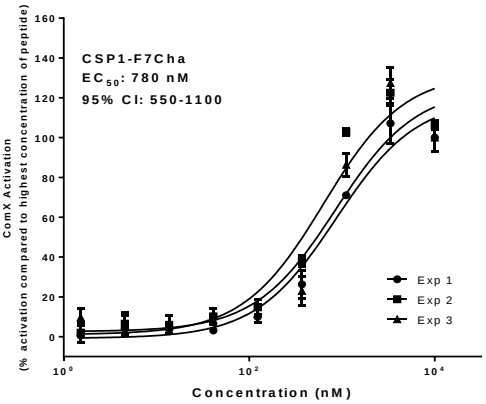


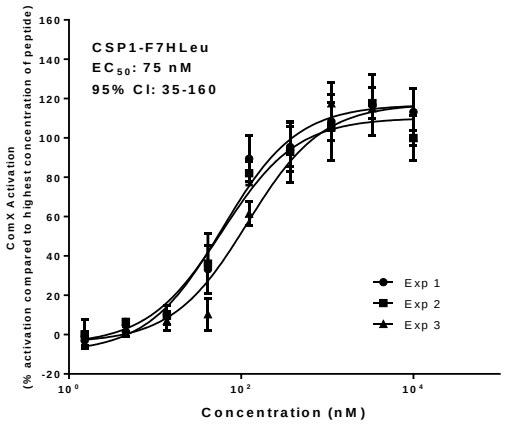


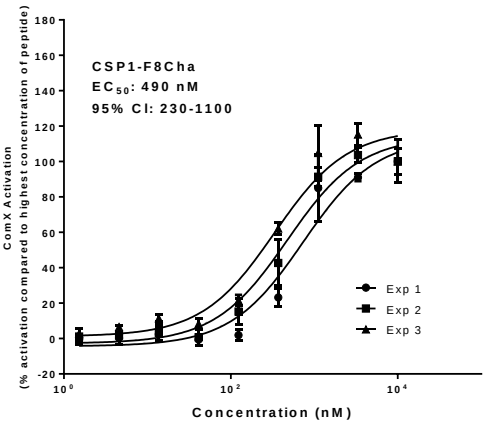
S. pneumoniae TIGR4 pcomX::lacZ (ComD2)

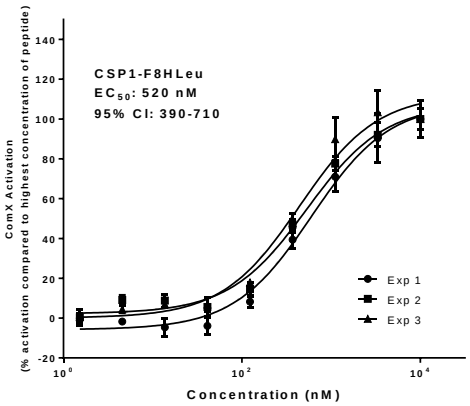
Activation dose response curves

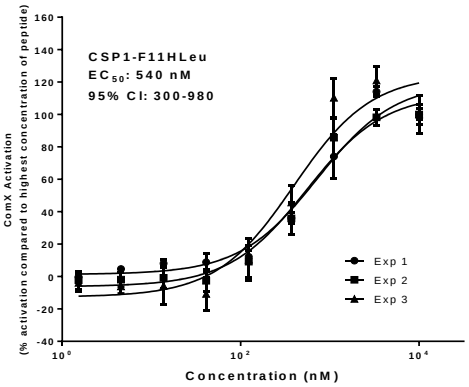


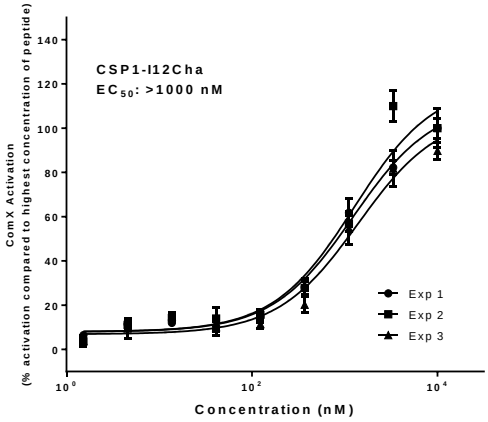


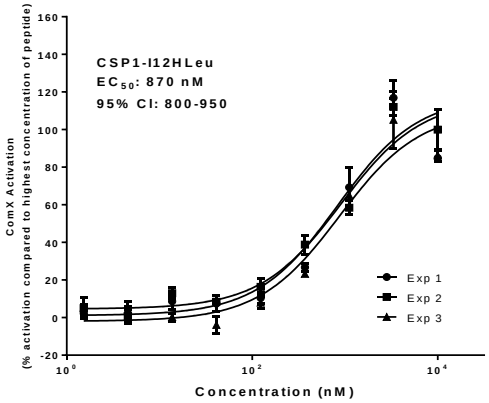

CSP1-F7HLeu
EC50: 75 nM
95% CI: 35-160
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

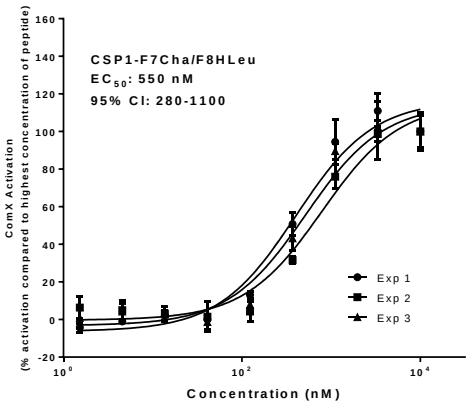

CSP1-F8Cha
EC50: 490 nM
95% CI: 230-1100
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

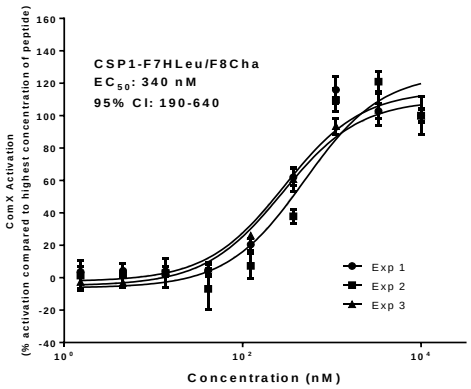

CSP1-F8HLeu
EC50: 520 nM
95% CI: 390-710
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

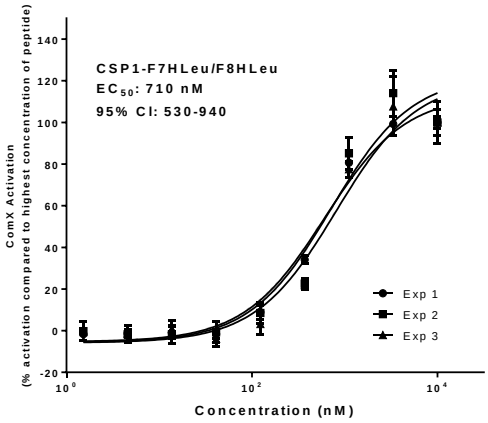

CSP1-F11HLeu
EC50: 540 nM
95% CI: 300-980
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

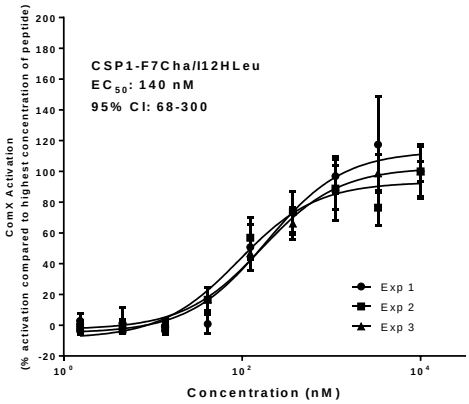

CSP1-I12Cha
EC50: >1000 nM
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

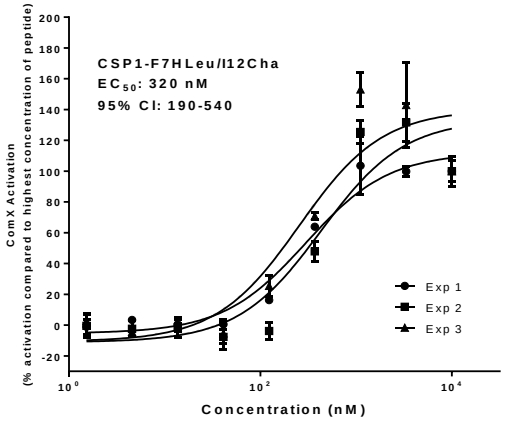

CSP1-I12HLeu
EC50: 870 nM
95% CI: 800-950
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

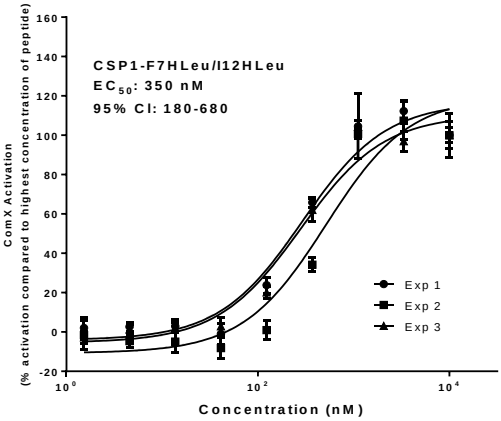

CSP1-F7Cha/F8HLeu
EC50: 550 nM
95% CI: 280-1100
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

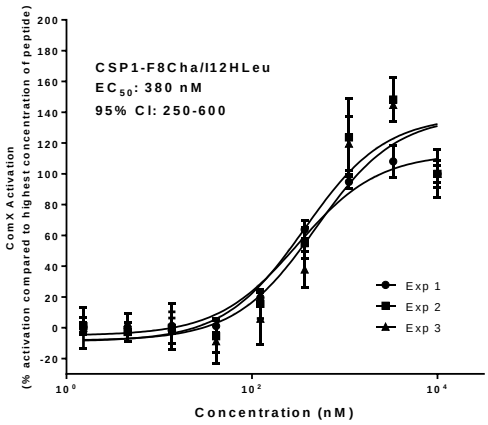

CSP1-F7HLeu/F8Cha
EC50: 340 nM
95% CI: 190-640
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

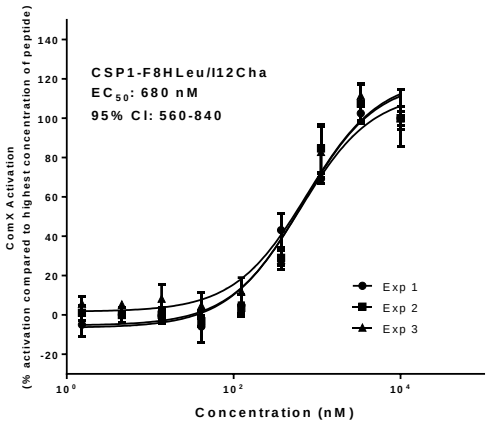

CSP1-F7HLeu/F8HLeu
EC50: 710 nM
95% CI: 530-940
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

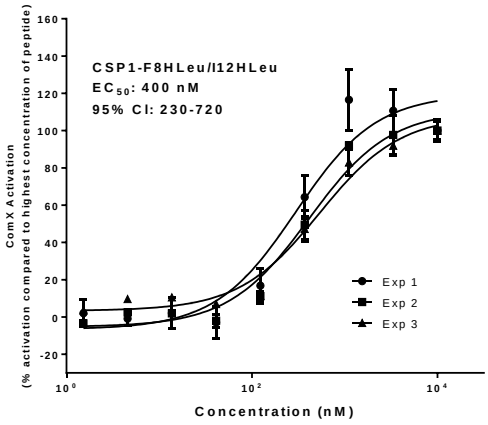

CSP1-F7Cha/I12HLeu
EC50: 140 nM
95% CI: 68-300
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3


CSP1-F7HLeu/I12Cha
EC50: 320 nM
95% CI: 190-540
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3


CSP1-F7HLeu/I12HLeu
EC50: 350 nM
95% CI: 180-680
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3


CSP1-F8Cha/I12HLeu
EC50: 380 nM
95% CI: 250-600
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3


CSP1-F8HLeu/I12Cha
EC50: 680 nM
95% CI: 560-840
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3


CSP1-F8HLeu/I12HLeu
EC50: 400 nM
95% CI: 230-720
ComX Activation
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

Metabolic Stability Analysis of Lead CSP1 Analogs

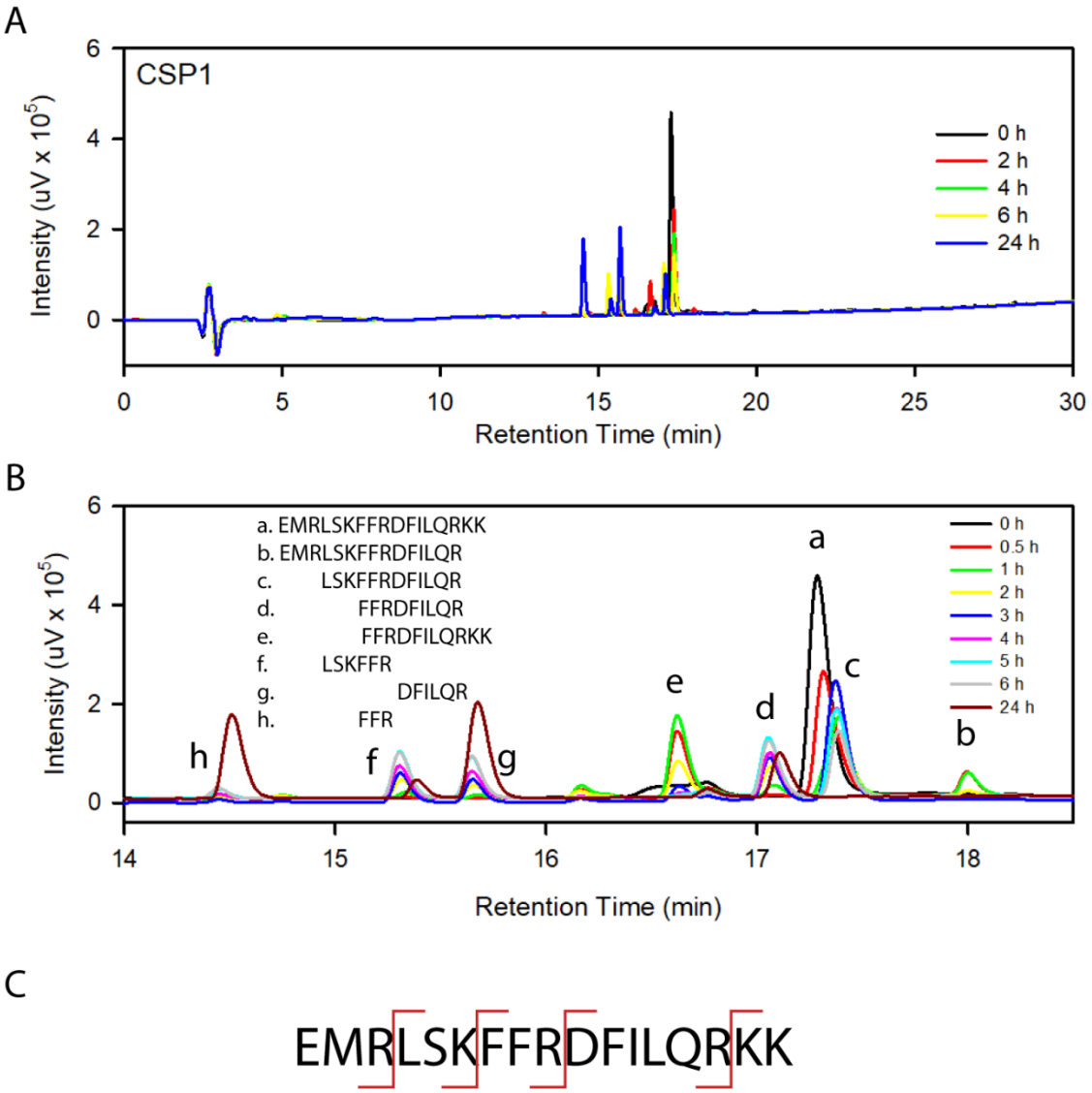


Figure S-9. Degradation pattern analysis of CSP1 in the presence of trypsin. A) Comparison of analytical HPLC chromatograms (220 nm) across the time points taken at 0, 2, 4, 6 and 24 h. B) Zoomed view of panel A with all the time points (0, 0.5, 1, 2, 3, 4, 5, 6 and 24 h). All peaks were collected and analyzed via mass spectrometry (MALDI-TOF MS and ESI-TOF LC-MS). The degradation products corresponding to each peak are listed on the left. C) A summary of observed trypsin cleavage sites annotated on the CSP1 sequence.

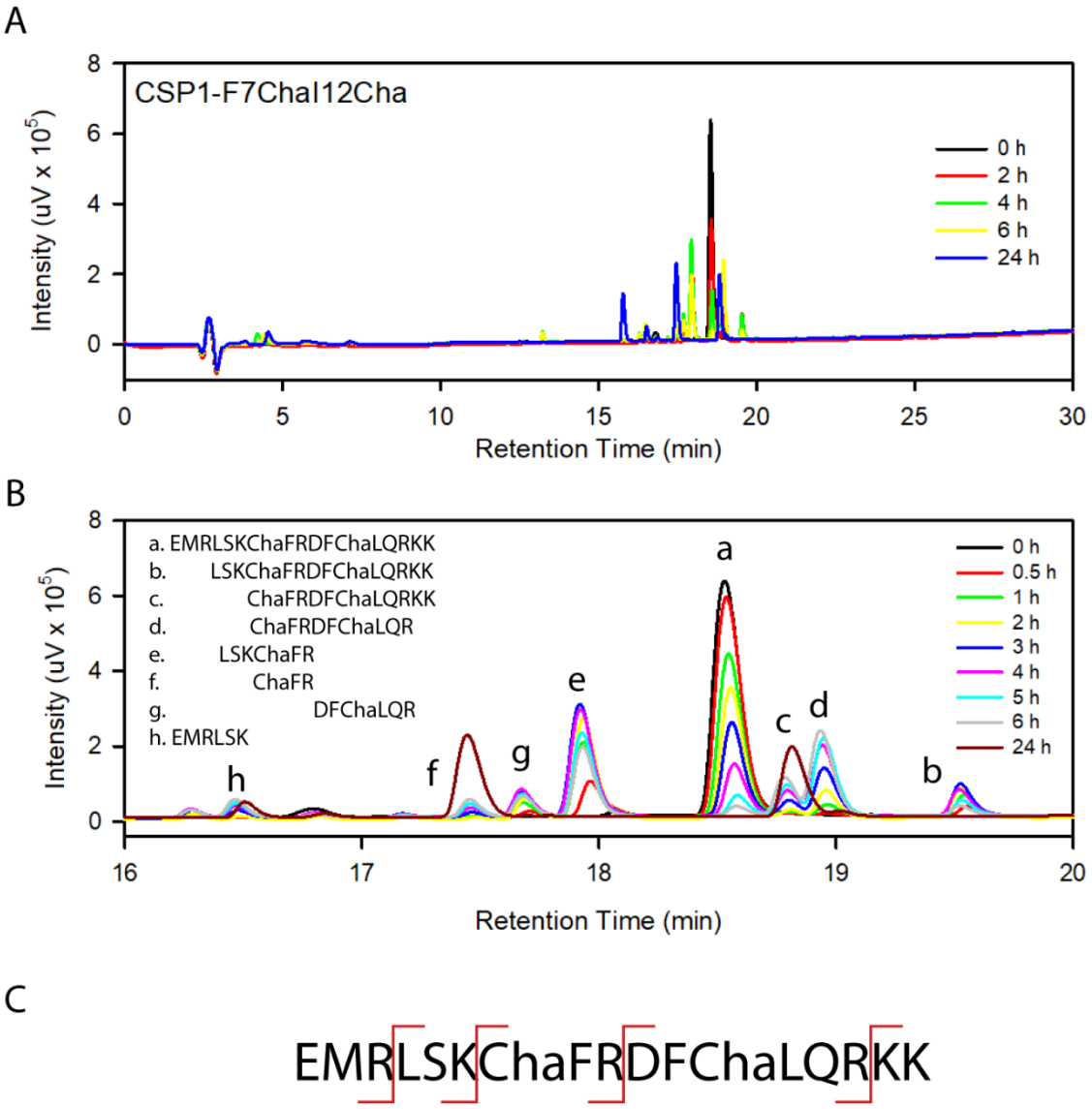


Figure S-10. Degradation pattern analysis of CSP1-F7ChaI12Cha in the presence of trypsin. A) Comparison of analytical HPLC chromatograms (220 nm) across the time points taken at 0, 2, 4, 6 and 24 h. B) Zoomed view of panel A with all the time points (0, 0.5, 1, 2, 3, 4, 5, 6 and 24 h). All peaks were collected and analyzed via mass spectrometry (MALDI-TOF MS and ESI-TOF LC-MS). The degradation products corresponding to each peak are listed on the left. C) A summary of observed trypsin cleavage sites annotated on the CSP1-F7ChaI12Cha sequence.

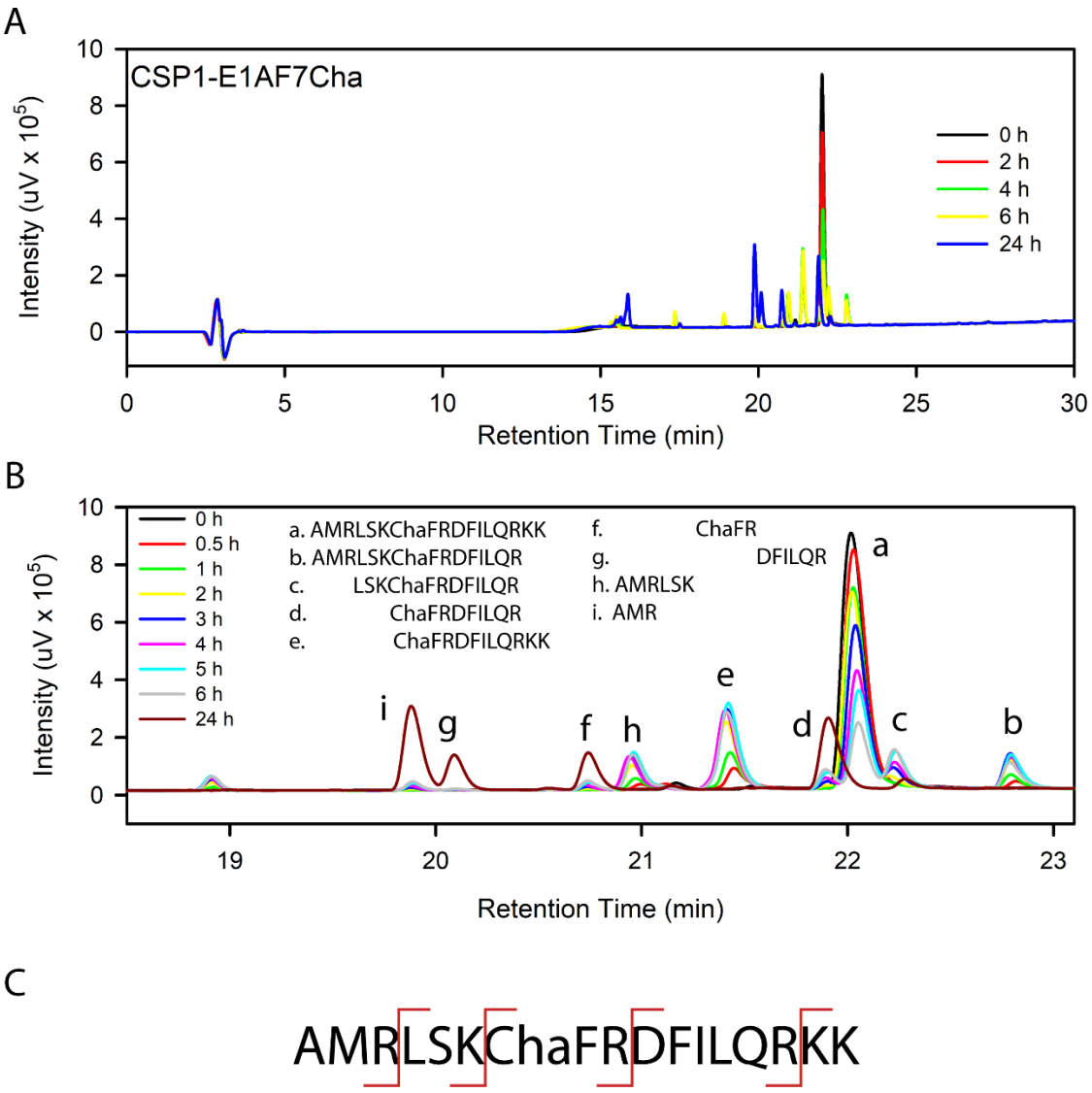


Figure S-11. Degradation pattern analysis of CSP1-E1AF7Cha in the presence of trypsin. A) Comparison of analytical HPLC chromatograms (220 nm) across the time points taken at 0, 2, 4, 6 and 24 h. B) Zoomed view of panel A with all the time points (0, 0.5, 1, 2, 3, 4, 5, 6 and 24 h). All peaks were collected and analyzed via mass spectrometry (MALDI-TOF MS and ESI-TOF LC-MS). The degradation products corresponding to each peak are listed on the left. C) A summary of observed trypsin cleavage sites annotated on the CSP1-E1AF7Cha sequence.

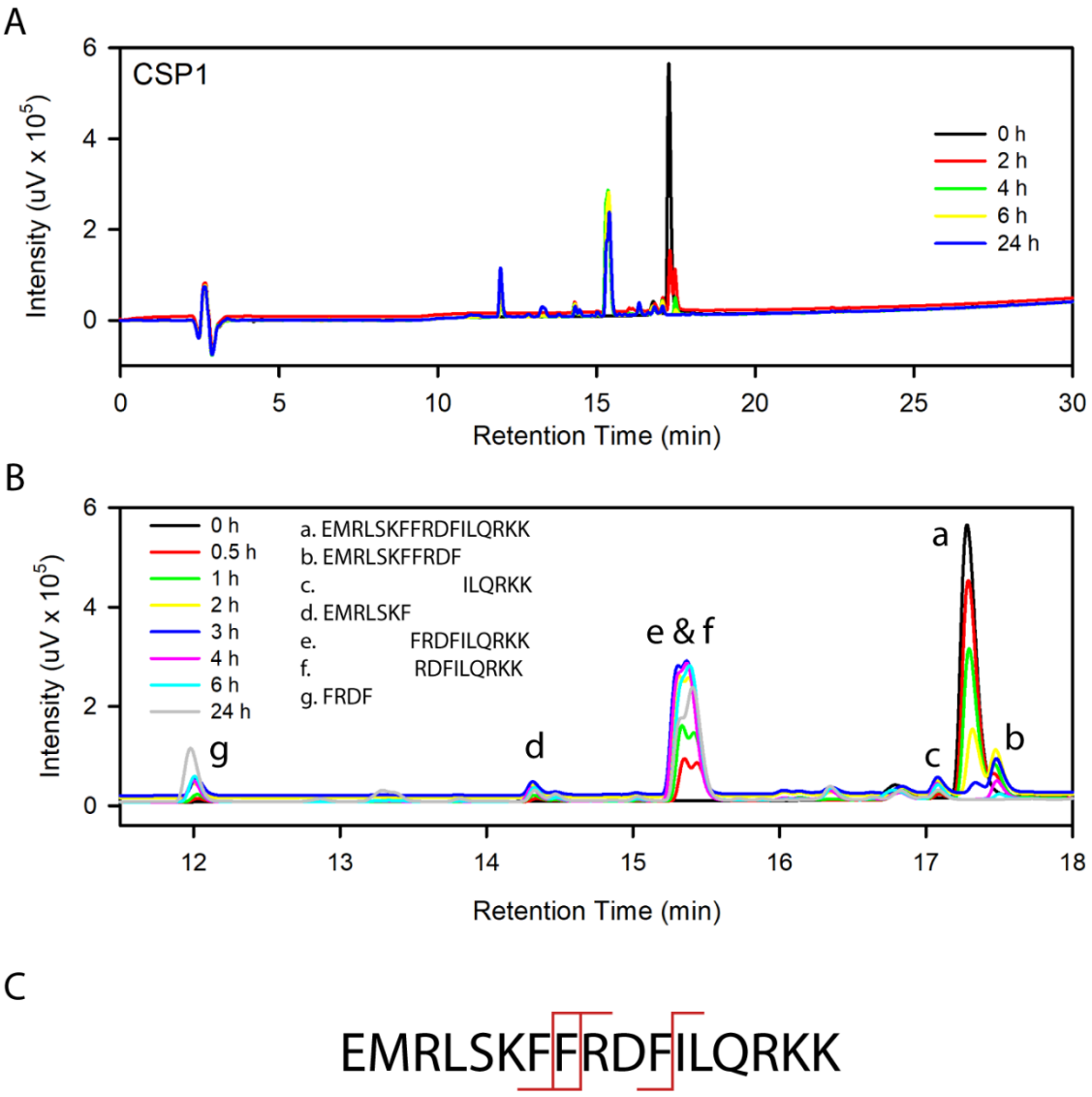


Figure S-12. Degradation pattern analysis of CSP1 in the presence of chymotrypsin. A) Comparison of analytical HPLC chromatograms (220 nm) across the time points taken at 0, 2, 4, 6 and 24 h. B) Zoomed view of panel A with all the time points (0, 0.5, 1, 2, 3, 4, 6 and 24 h). All peaks were collected and analyzed via mass spectrometry (MALDI-TOF MS and ESI-TOF LC-MS). The degradation products corresponding to each peak are listed on the left. C) A summary of observed chymotrypsin cleavage sites annotated on the CSP1 sequence.

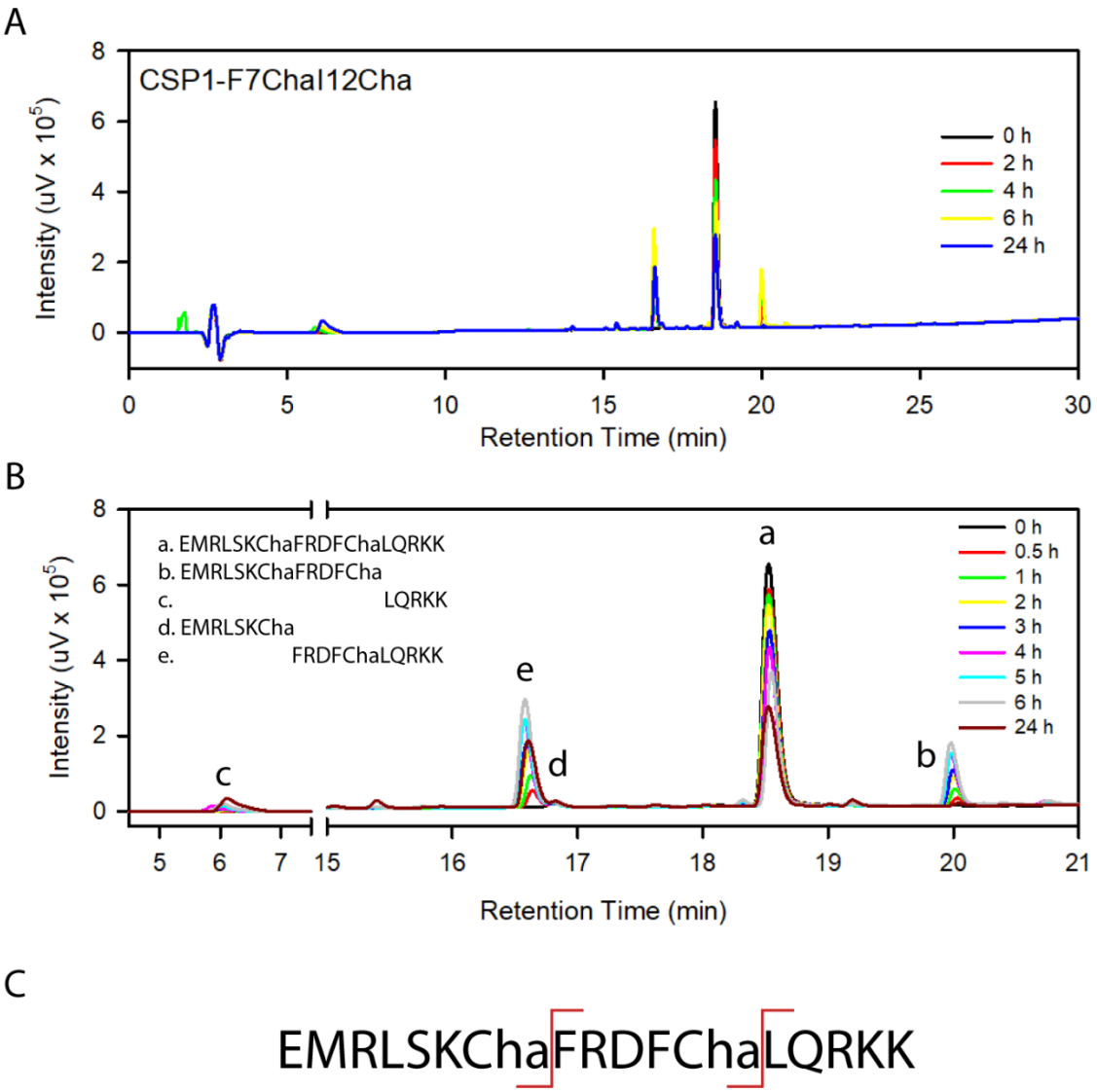


Figure S-13. Degradation pattern analysis of CSP1-F7ChaI12Cha in the presence of chymotrypsin. A) Comparison of analytical HPLC chromatograms (220 nm) across the time points taken at 0, 2, 4, 6 and 24 h. B) Zoomed view of panel A with all the time points (0, 0.5, 1, 2, 3, 4, 5, 6 and 24 h). All peaks were collected and analyzed via mass spectrometry (MALDI-TOF MS and ESI-TOF LC-MS). The degradation products corresponding to each peak are listed on the left. C) A summary of observed chymotrypsin cleavage sites annotated on the CSP1-F7ChaI12Cha sequence.

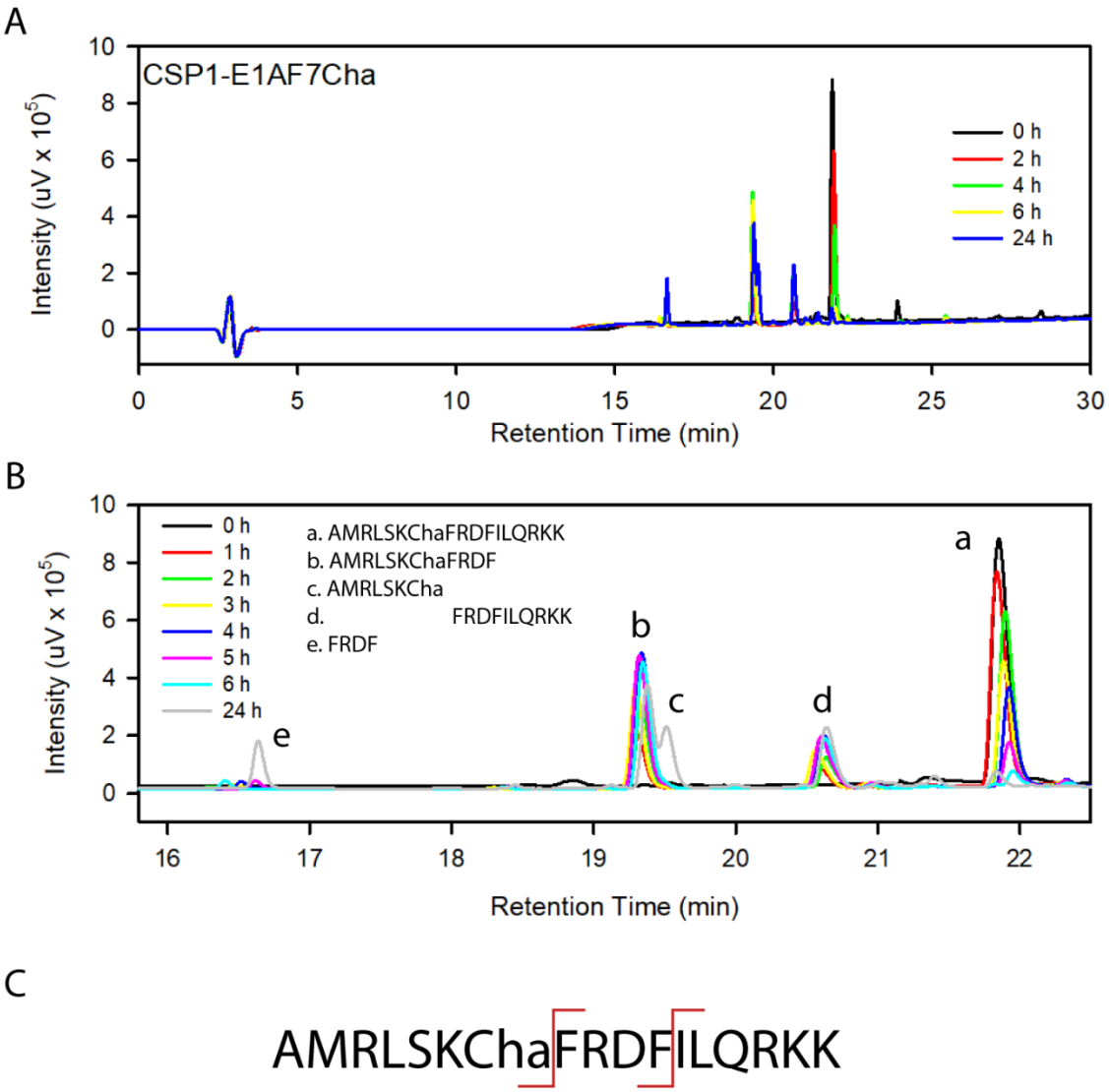


Figure S-14. Degradation pattern analysis of CSP1-E1AF7Cha in the presence of chymotrypsin. A) Comparison of analytical HPLC chromatograms (220 nm) across the time points taken at 0, 2, 4, 6 and 24 h. B) Zoomed view of panel A with all the time points (0, 1, 2, 3, 4, 5, 6 and 24 h). All peaks were collected and analyzed via mass spectrometry (MALDI-TOF MS and ESI-TOF LC-MS). The degradation products corresponding to each peak are listed on the left. C) A summary of observed chymotrypsin cleavage sites annotated on the CSP1-E1AF7Cha sequence.

Appendix 3: Biological Evaluation of Native Streptococcal Competence Stimulating Peptides Reveals Potential Crosstalk between *Streptococcus mitis* and *Streptococcus pneumoniae* and a New Scaffold for the Development of *S. pneumoniae* Quorum Sensing Modulators

[a]Reprinted with permission from Milly, T. A.; Tal-Gan, Y. Biological evaluation of native streptococcal competence stimulating peptides reveals potential crosstalk between *Streptococcus mitis* and *Streptococcus pneumoniae* and a new scaffold for the development of *S. pneumoniae* quorum sensing modulators. *RSC Chem. Biol.* **2020**, 1, 60-67.

Biological Evaluation of Native Streptococcal Competence Stimulating Peptides Reveal Potential Crosstalk Between *Streptococcus mitis* and *Streptococcus pneumoniae* and a New Scaffold for the Development of *S. pneumoniae* Quorum Sensing Modulators

Supporting Information

Tahmina Ahmed Milly and Yftah Tal-Gan*

Department of Chemistry, University of Nevada, Reno, 1664 N. Virginia Street, Reno, Nevada 89557

* To whom correspondence should be addressed. ytalgan@unr.edu

Additional Experimental Details

Peptide Synthesis: 0.1 g of resin was first swelled by suspension in DMF for 15 min at room temperature and then drained. Deprotection of the Fmoc group was performed first using 5 mL of 20% piperidine in DMF (90 sec, 90 °C) followed by another 5 mL of 20% piperidine in DMF (90 sec, 90 °C). The resin was washed with DMF (3 x 5 mL) after each deprotection cycle. Coupling reactions were performed using 2.5 mL solution containing Fmoc-protected amino acid (5 equiv.), 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU; 5 equiv.) and diisopropylethylamine (DIPEA; 5 equiv.). All amino acids were coupled for 20 min (30 W, 75 °C), except His. His was coupled for 10 min (0 W, 25 °C) then for 40 min (20 W, 50 °C). After the synthesis was completed the resin was washed with DMF (2 x 5 mL).

Cleavage: Upon completion of peptide synthesis, the resin containing the final peptide product was washed with diethyl ether (2 mL) and dried under nitrogen stream for 3 min before it was transferred into a 15 mL falcon tube. The peptide was cleaved from the resin, along with all the protecting groups, by mixing the resin with 3 mL of 2.5% de-ionized water and 2.5% triisopropylsilane (TIPS) in trifluoroacetic acid (TFA) for 3 h with agitation. The cleaved peptide was separated from the resin by filtration and the filtrate was transferred into a new 50 mL falcon tube. A cooled solution of diethyl ether:hexane (1:1, 45 mL, 0 °C) was added to the filtrate, and the peptide was allowed to precipitate in a freezer at -20 °C for 10 min. The pellet of the crude peptide was obtained by centrifuging the mixture at 3000 RPM for 5 min and the supernatant was removed to yield crude peptide, which was dissolved in 10 mL acetonitrile (ACN):water (1:1) and lyophilized before HPLC purification.

Peptide Verification with Mass Spectrometry: During purification, MALDI-TOF MS was used to verify the peaks containing the desired peptide. Samples were prepared using α-Cyano-4-hydroxycinammic acid as matrix and aliquots were taken directly from the preparative HPLC fractions. The exact masses of the peptides were obtained with a high resolution ESI-TOF MS for the final verification of the peptides. 8 - 30 μM stock solutions were prepared in acetonitrile (ACN):water (1:1). The instrument was calibrated before each run and an internal reference mass standard was used.

Beta-Galactosidase Assays: The ability of synthetic CSP analogues to activate the expression of *S. pneumoniae comX* was determined using indicated reporter strains grown in THY (pH 7.3). An initial activation screening was performed at a high concentration (10 μM) for all CSP analogues. 198 μL bacterial culture were placed in triplicate to a 96-well plate containing 2 μL of CSP peptides and incubated at 37 °C for 30 min. A total of 2 μL of 20 μM solution of *S. pneumoniae* CSP1 (200 nM final concentration) were added in triplicate and served as the positive control for the *S. pneumoniae* group I strain (D39pcomX::lacZ), while 2 μL of 100 μM solution of *S. pneumoniae* CSP2 (1000 nM final concentration) were added as the positive control for the *S. pneumoniae* group II strain (TIGR4pcomX::lacZ). These concentrations were chosen to afford full activation of the QS circuit, as determined from the dose-dependent curves created for the native *S. pneumoniae* CSPs.[47] Two μL dimethyl sulfoxide (DMSO) were added in triplicate and served as the negative control. After the incubation time (30 min) had elapsed, the absorbance at 600_{nm} was read. The cells were then lysed by incubating the culture for 30 min at 37 °C with 20 μL 0.1% Triton X-100 in water. In a new plate, 100 μL Z-buffer solution (60.2 mM Na_2HPO_4, 45.8 mM NaH_2PO_4, 10 mM KCl, and 1.0 mM $MgSO_4$ in 18

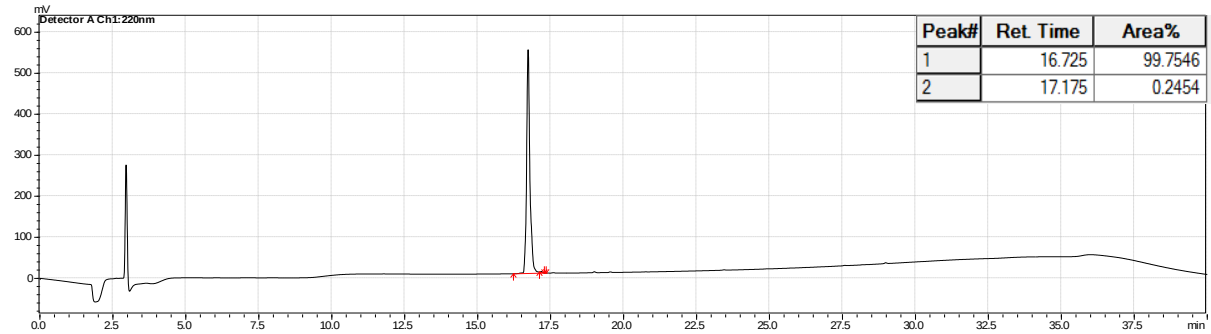
MΩ H_2O; pH was adjusted to 7.0 and the buffer was sterilized before use) containing 2-Nitrophenyl-Beta-D-galactopyranoside (ONPG) at a final concentration of 0.4 mg/mL were added, followed by 100 µL lysate, and the plate was incubated for 3 hours at 37 °C. After the incubation, the reaction was stopped by adding 50 µL of 1 M sodium carbonate solution, and the OD 420_{nm} and OD 550_{nm} were measured using a plate reader, allowing for the calculation of the activity in Miller units (see below). The results were reported as percent activation, which is the ratio between the Miller units of the analogue and that of the positive control.

$$Miller\ Unit = 1000 \times \frac{[Abs_{420} - (1.75 \times Abs_{550})]}{(t \times v \times Abs_{600})}$$

Abs_{420} is the absorbance of o-nitrophenol (ONP). Abs_{550} is the scatter from cell debris, which, when multiplied by 1.75 approximates the scatter observed at 420 nm. *t* is the duration of incubation with ONPG in minutes, *v* is volume of lysate in milliliters, and Abs_{600} reflects cell density.

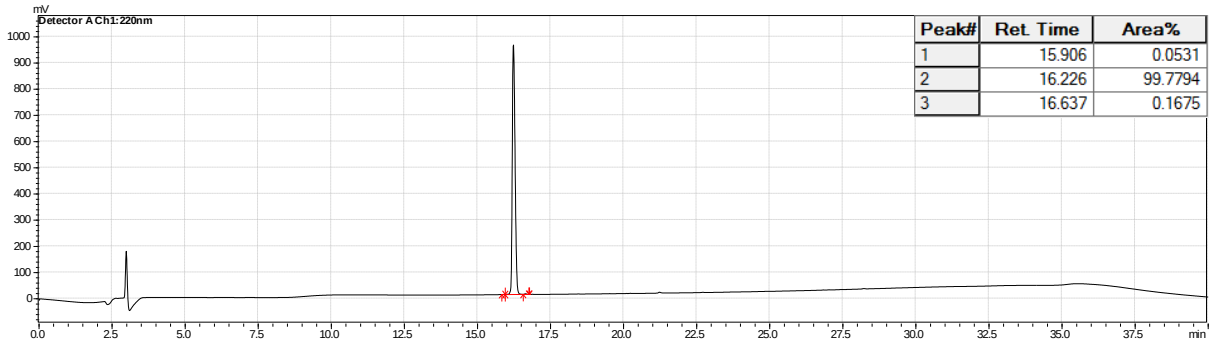
HPLC traces for CSP analogues

S. mitis-CSP-1



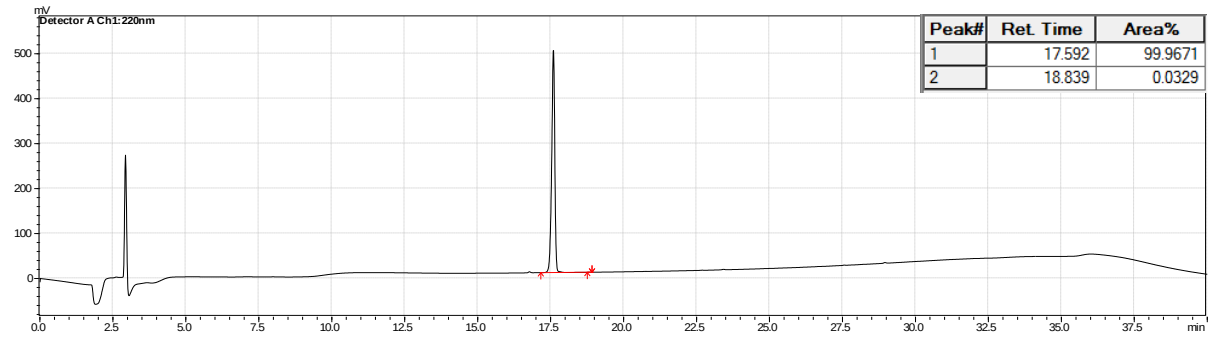
Peak#	Ret. Time	Area%
1	16.725	99.7546
2	17.175	0.2454

S. mitis-CSP-2



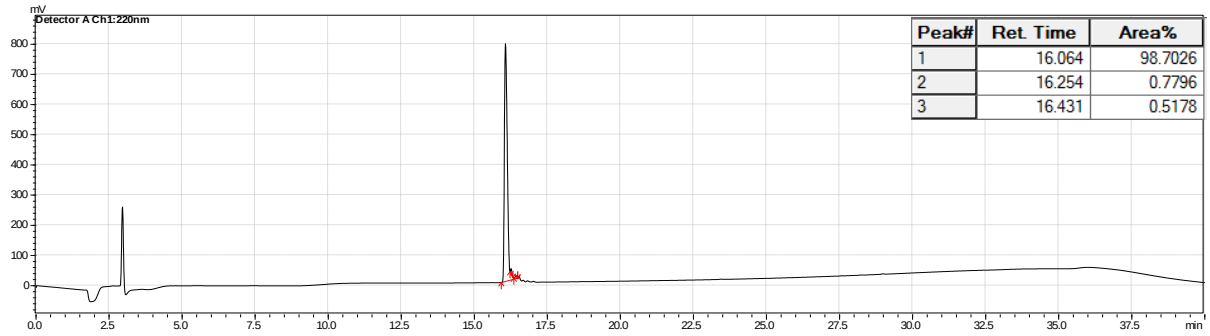
Peak#	Ret. Time	Area%
1	15.906	0.0531
2	16.226	99.7794
3	16.637	0.1675

S. oralis-CSP



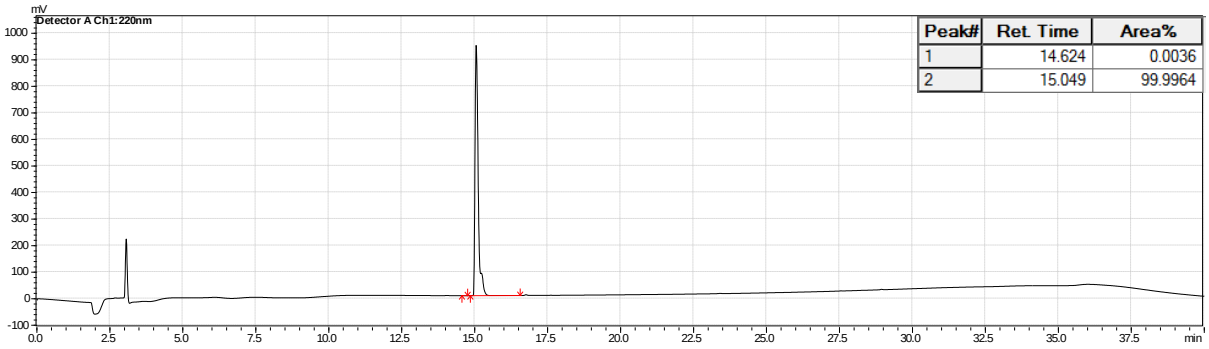
Peak#	Ret. Time	Area%
1	17.592	99.9671
2	18.839	0.0329

S. cristatus-CSP



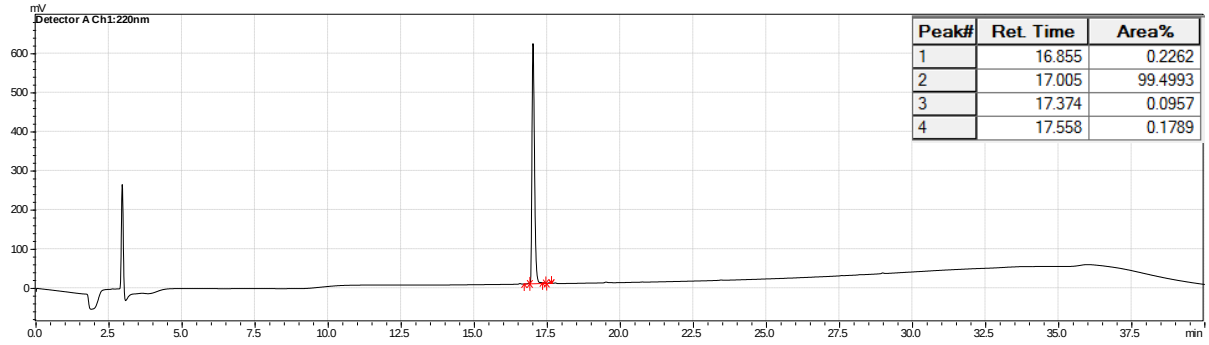
Peak#	Ret. Time	Area%
1	16.064	98.7026
2	16.254	0.7796
3	16.431	0.5178

S. oligofermentans-CSP



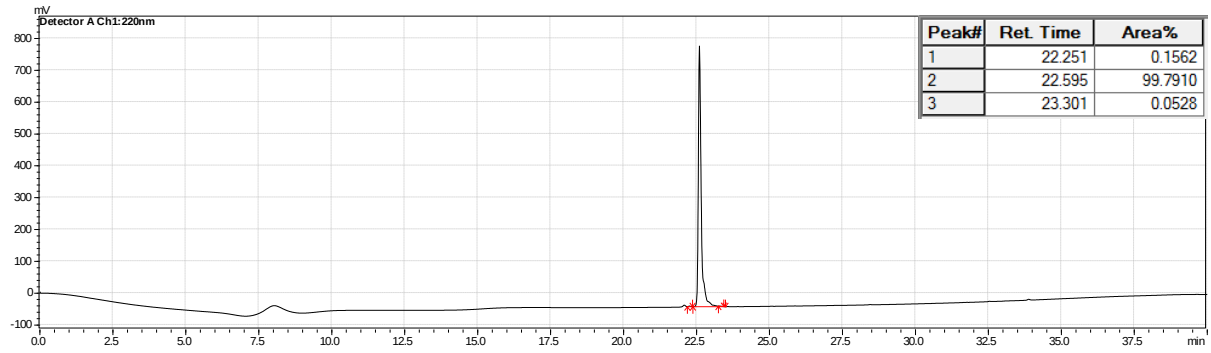
Peak#	Ret. Time	Area%
1	14.624	0.0036
2	15.049	99.9964

S. intermedius-CSP



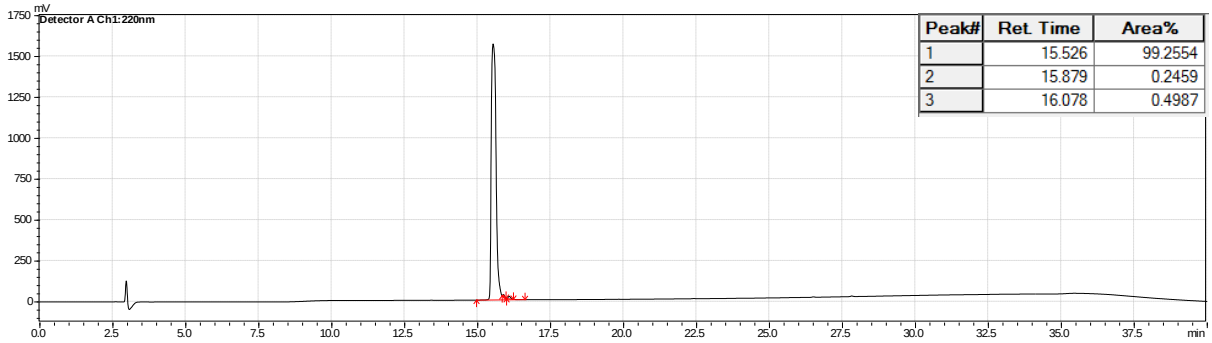
Peak#	Ret. Time	Area%
1	16.855	0.2262
2	17.005	99.4993
3	17.374	0.0957
4	17.558	0.1789

S. gordonii-CSP-1



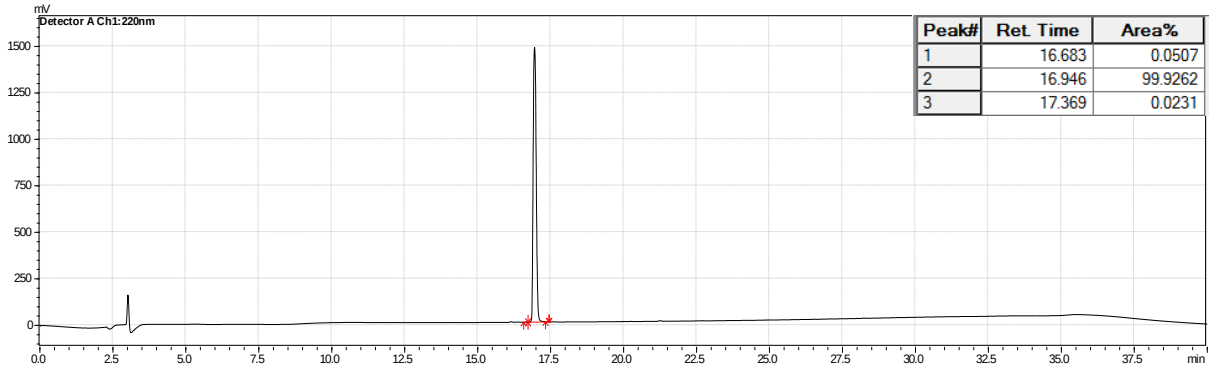
Peak#	Ret. Time	Area%
1	22.251	0.1562
2	22.595	99.7910
3	23.301	0.0528

S. gordonii-CSP-2



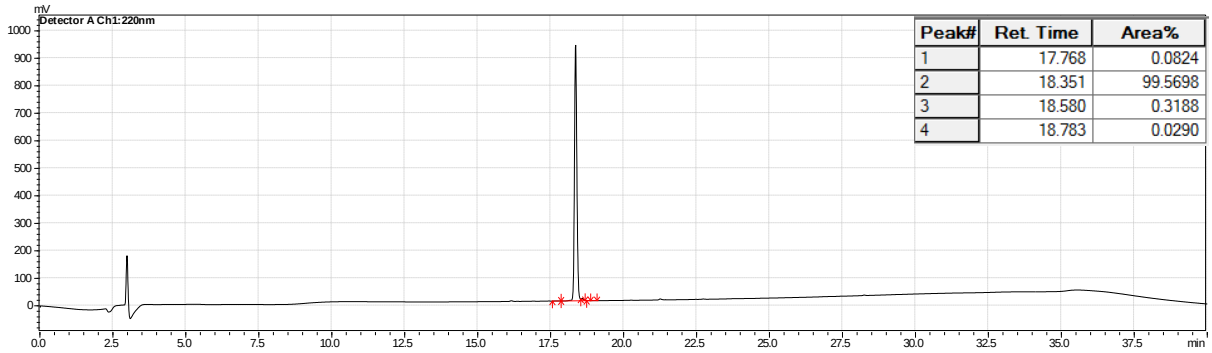
Peak#	Ret. Time	Area%
1	15.526	99.2554
2	15.879	0.2459
3	16.078	0.4987

S. gordonii challis-CSP



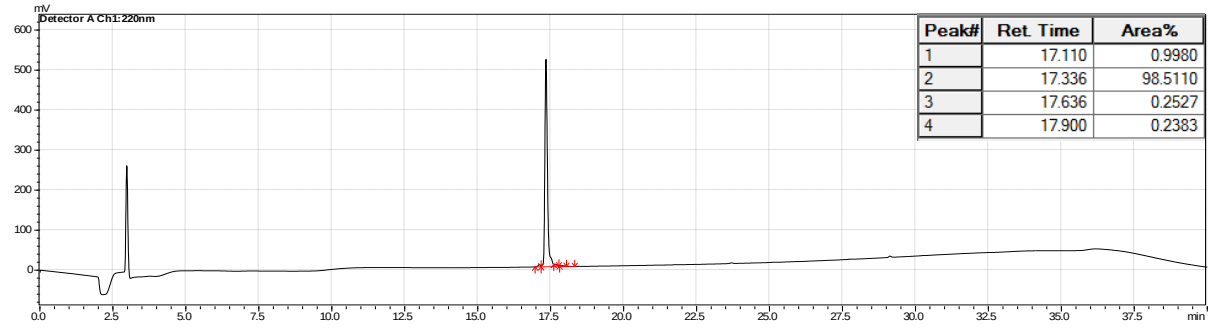
Peak#	Ret. Time	Area%
1	16.683	0.0507
2	16.946	99.9262
3	17.369	0.0231

S. sanguinis-CSP



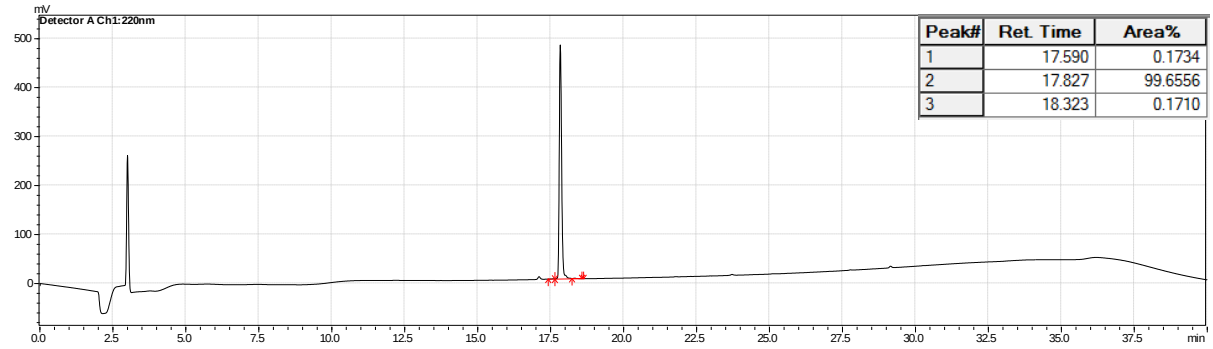
Peak#	Ret. Time	Area%
1	17.768	0.0824
2	18.351	99.5698
3	18.580	0.3188
4	18.783	0.0290

S. mitis-CSP-2-I2M



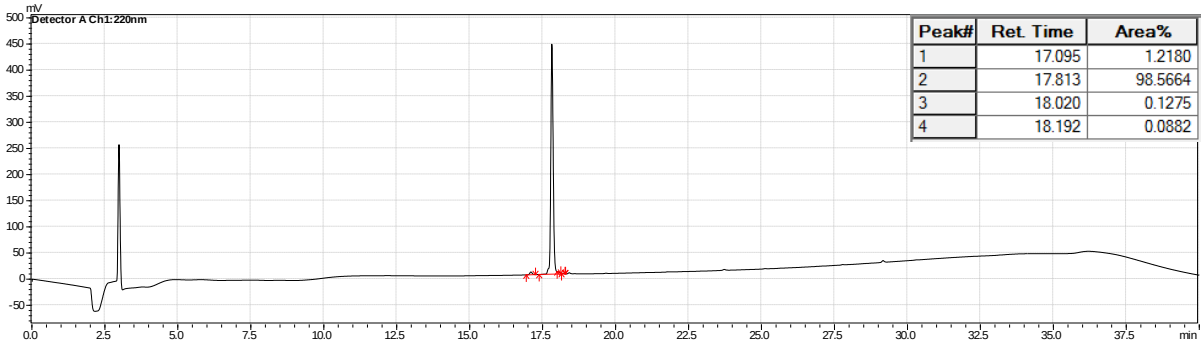
Peak#	Ret. Time	Area%
1	17.110	0.9980
2	17.336	98.5110
3	17.636	0.2527
4	17.900	0.2383

S. mitis-CSP-2-Q4L



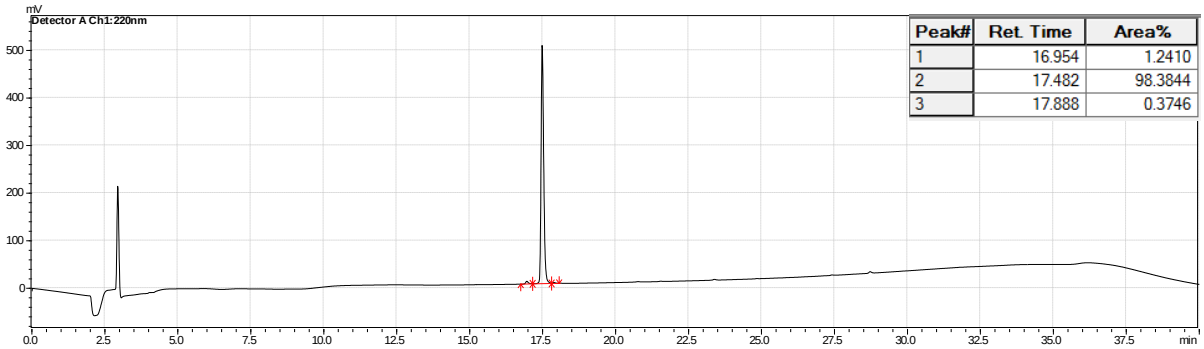
Peak#	Ret. Time	Area%
1	17.590	0.1734
2	17.827	99.6556
3	18.323	0.1710

S. mitis-CSP-2-N7F



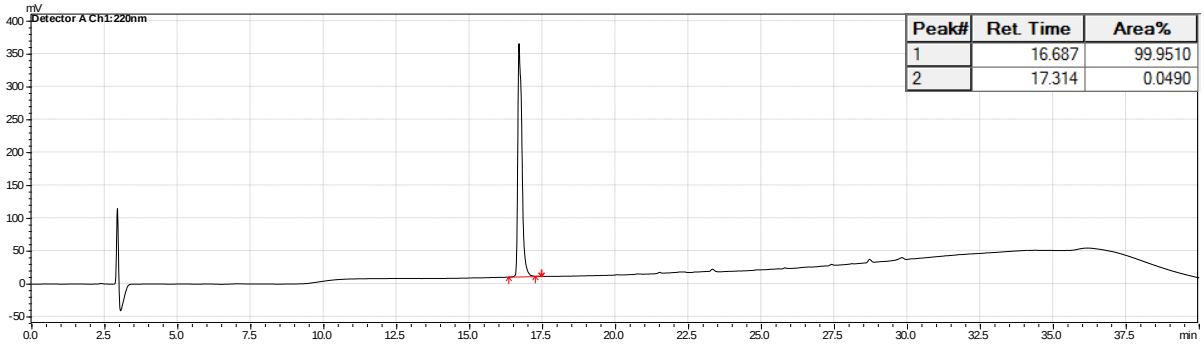
Peak#	Ret. Time	Area%
1	17.095	1.2180
2	17.813	98.5664
3	18.020	0.1275
4	18.192	0.0882

S. mitis-CSP-2-I8F



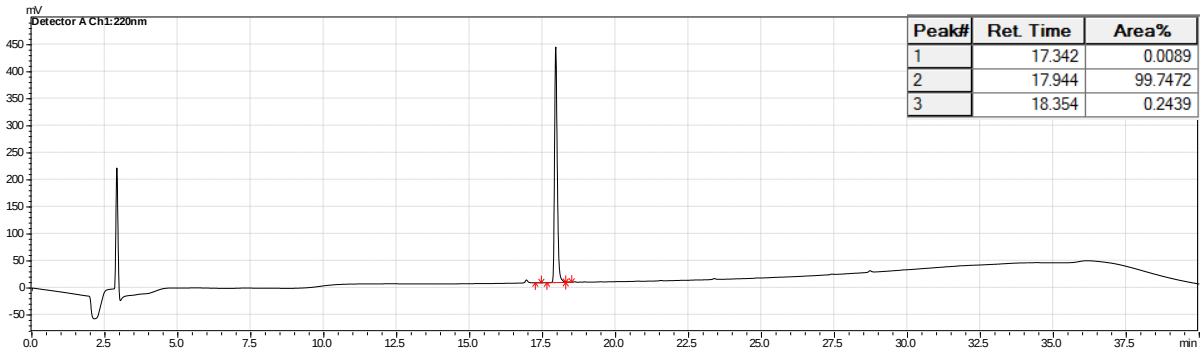
Peak#	Ret. Time	Area%
1	16.954	1.2410
2	17.482	98.3844
3	17.888	0.3746

S. mitis-CSP-2-F10D



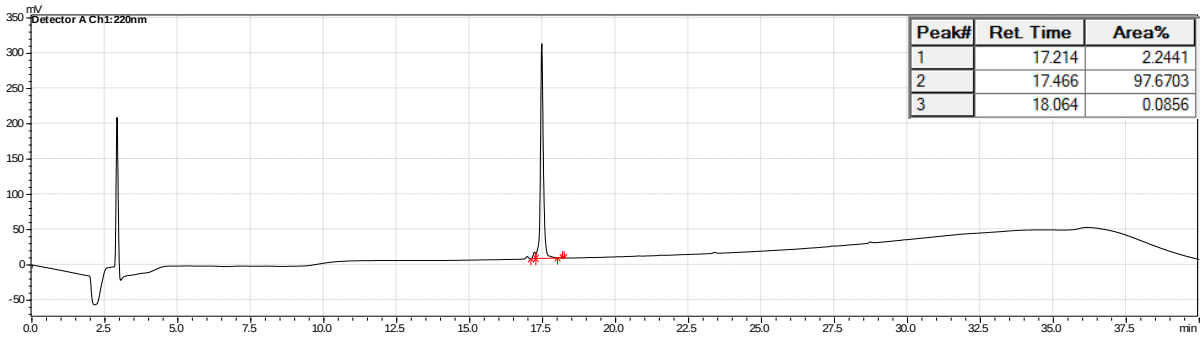
Peak#	Ret. Time	Area%
1	16.687	99.9510
2	17.314	0.0490

S. mitis-CSP-2-N11F



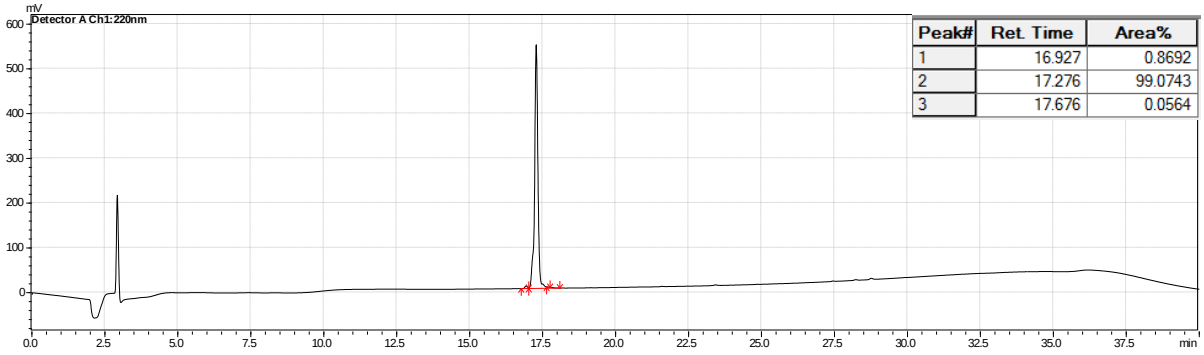
Peak#	Ret. Time	Area%
1	17.342	0.0089
2	17.944	99.7472
3	18.354	0.2439

S. mitis-CSP-2-F12I



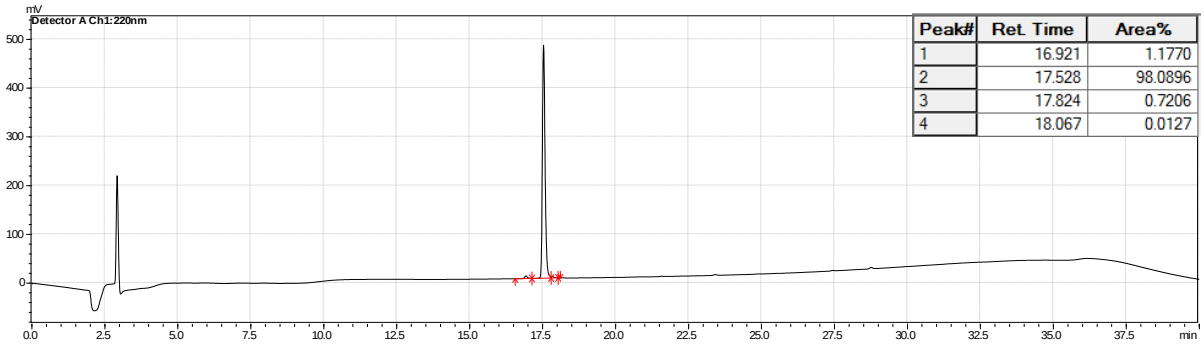
Peak#	Ret. Time	Area%
1	17.214	2.2441
2	17.466	97.6703
3	18.064	0.0856

S. mitis-CSP-2-F12L



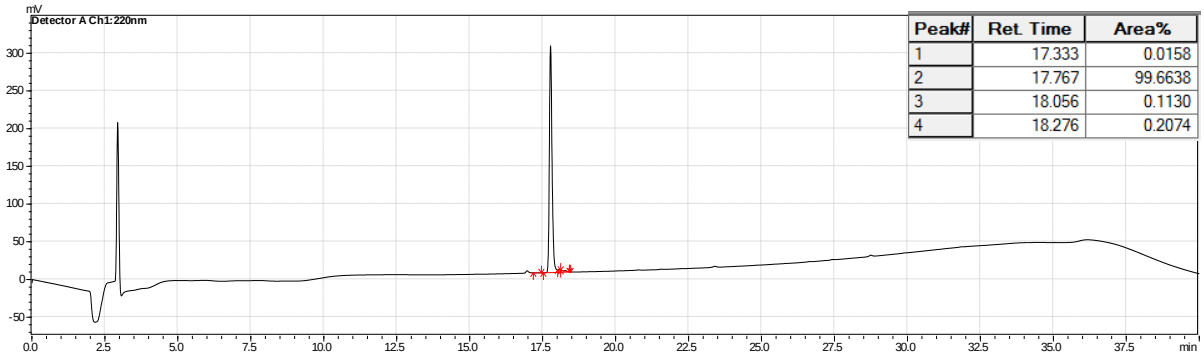
Peak#	Ret. Time	Area%
1	16.927	0.8692
2	17.276	99.0743
3	17.676	0.0564

S. mitis-CSP-2-N7I



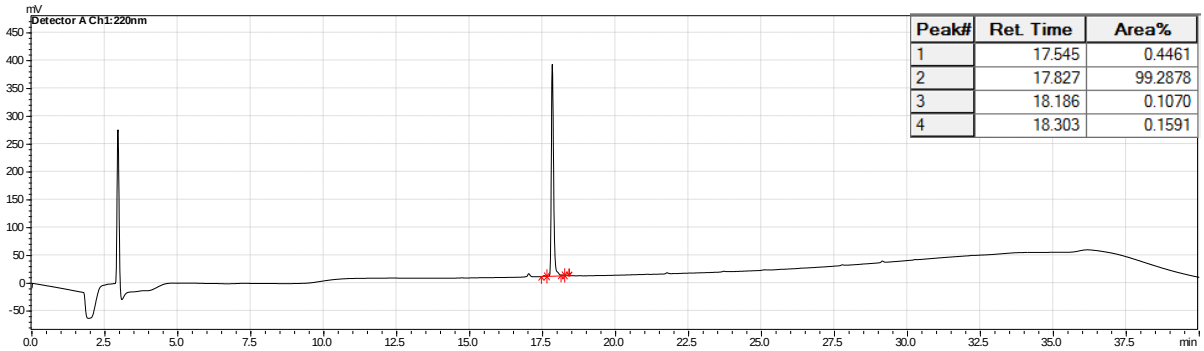
Peak#	Ret. Time	Area%
1	16.921	1.1770
2	17.528	98.0896
3	17.824	0.7206
4	18.067	0.0127

S. mitis-CSP-2-Q4I



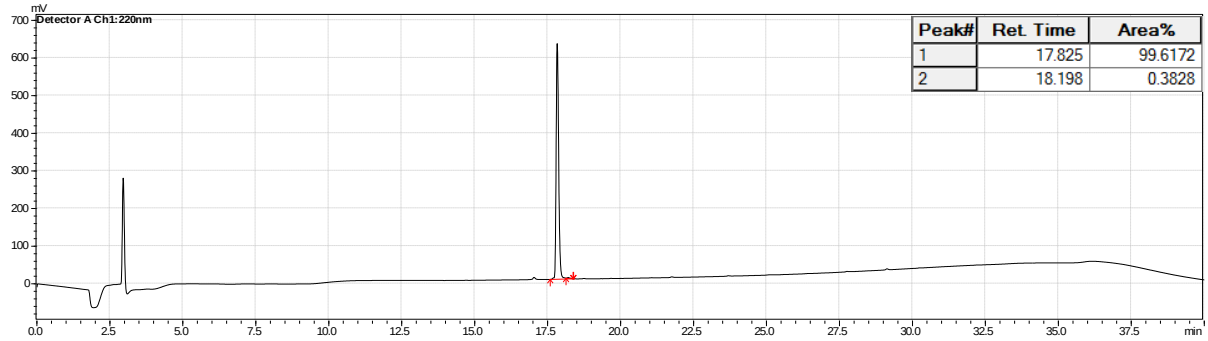
Peak#	Ret. Time	Area%
1	17.333	0.0158
2	17.767	99.6638
3	18.056	0.1130
4	18.276	0.2074

S. mitis-CSP-2-I2MQ4L



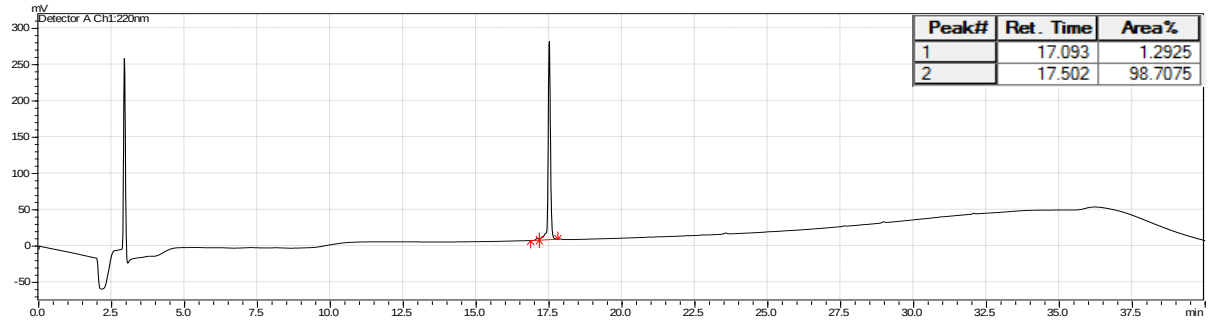
Peak#	Ret. Time	Area%
1	17.545	0.4461
2	17.827	99.2878
3	18.186	0.1070
4	18.303	0.1591

S. mitis-CSP-2-I2MN7F



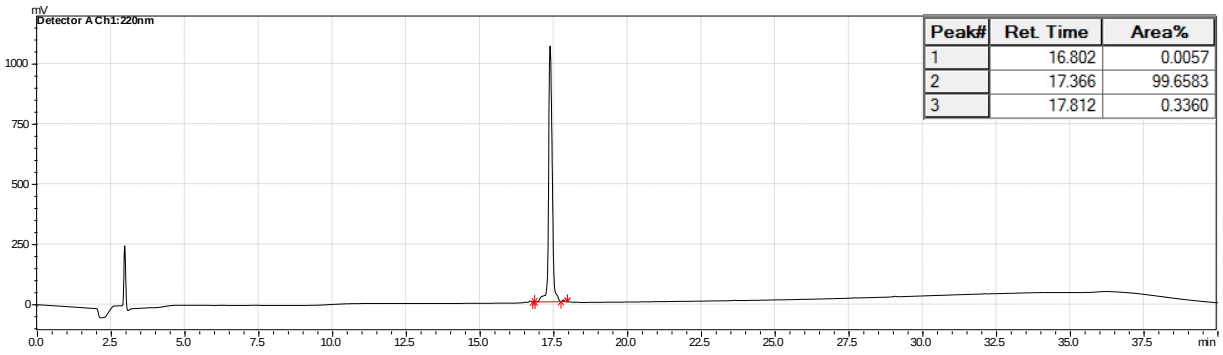
Peak#	Ret. Time	Area%
1	17.825	99.6172
2	18.198	0.3828

S. mitis-CSP-2-I2MN7I



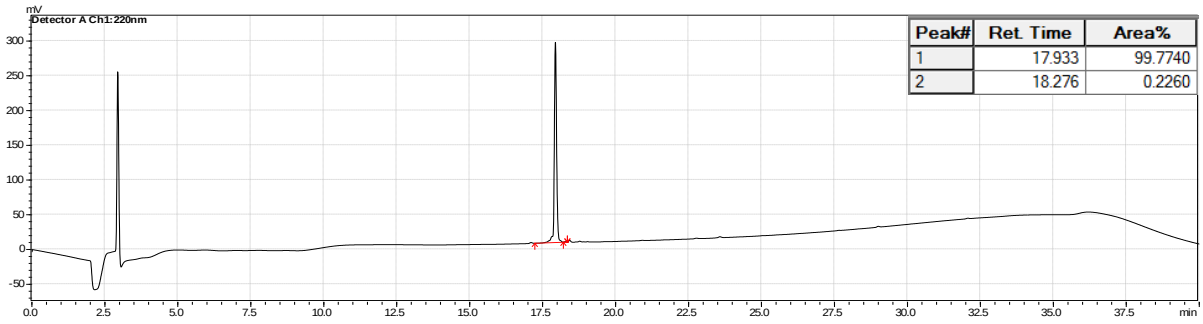
Peak#	Ret. Time	Area%
1	17.093	1.2925
2	17.502	98.7075

S. mitis-CSP-2-I2MI8F



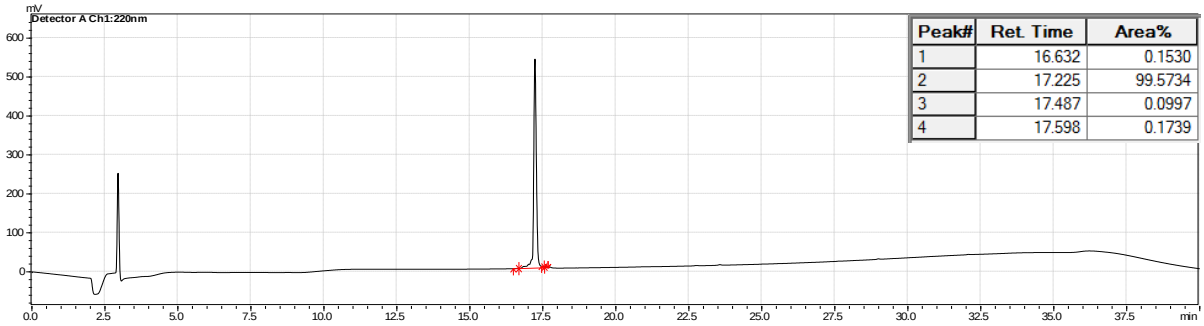
Peak#	Ret. Time	Area%
1	16.802	0.0057
2	17.366	99.6583
3	17.812	0.3360

S. mitis-CSP-2-I2MN11F



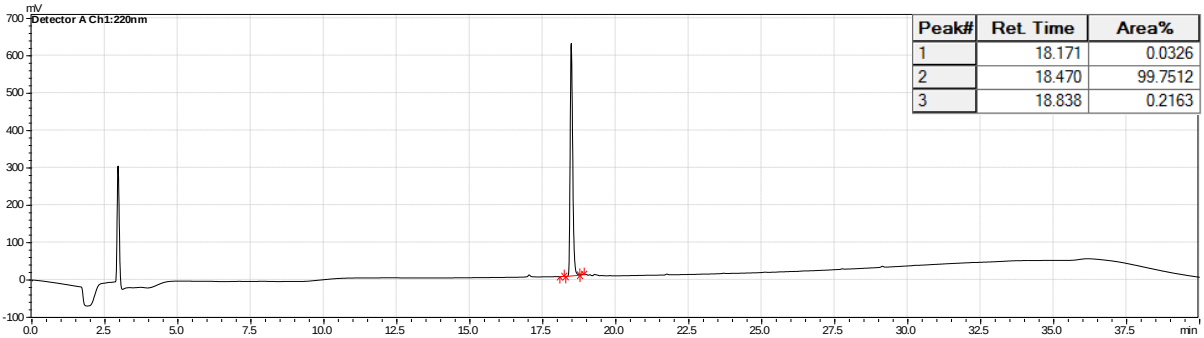
Peak#	Ret. Time	Area%
1	17.933	99.7740
2	18.276	0.2260

S. mitis-CSP-2-I2MF12L



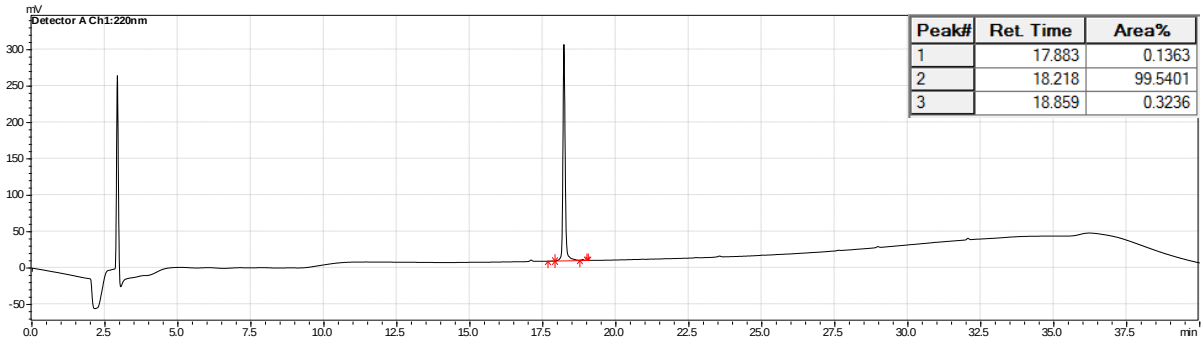
Peak#	Ret. Time	Area%
1	16.632	0.1530
2	17.225	99.5734
3	17.487	0.0997
4	17.598	0.1739

S. mitis-CSP-2-Q4LN7F



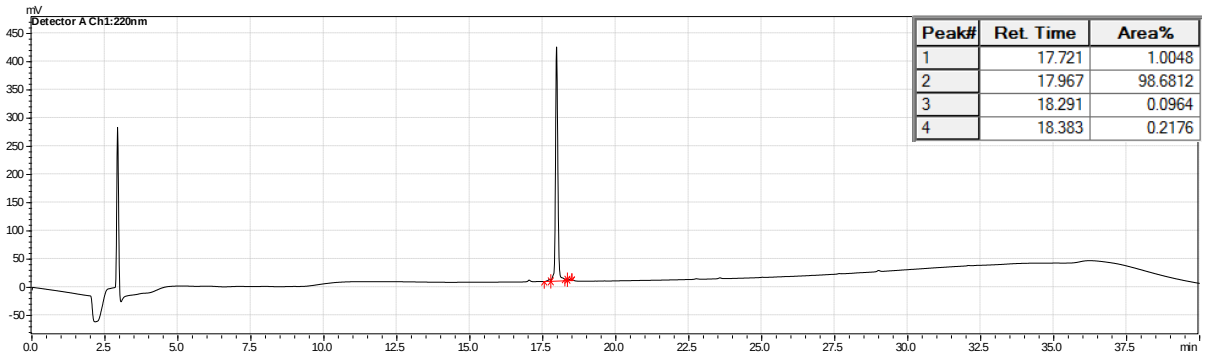
Peak#	Ret. Time	Area%
1	18.171	0.0326
2	18.470	99.7512
3	18.838	0.2163

S. mitis-CSP-2-Q4LN7I



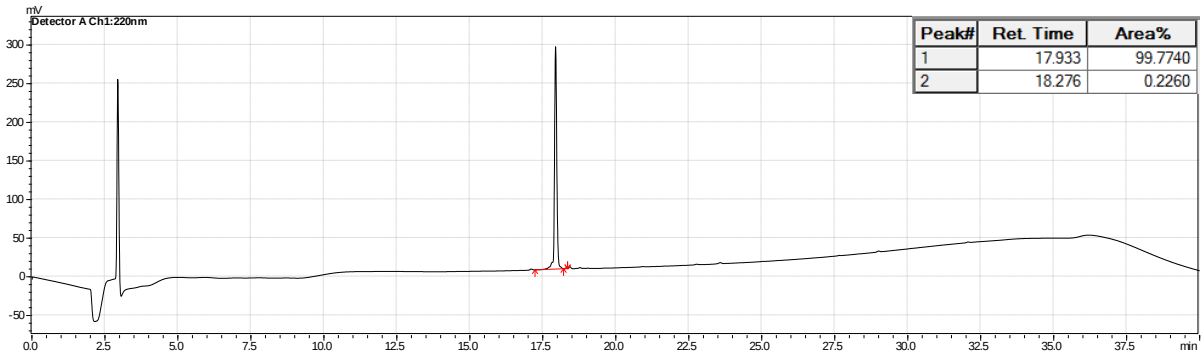
Peak#	Ret. Time	Area%
1	17.883	0.1363
2	18.218	99.5401
3	18.859	0.3236

S. mitis-CSP-2-Q4LI8F



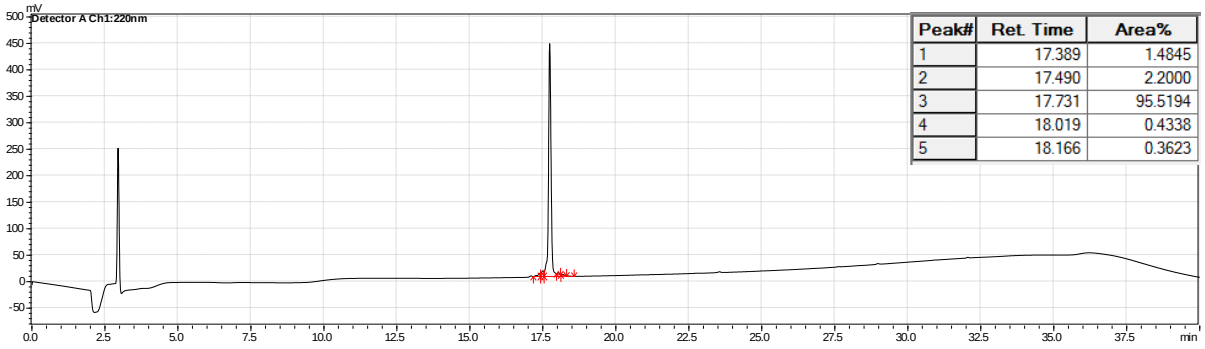
Peak#	Ret. Time	Area%
1	17.721	1.0048
2	17.967	98.6812
3	18.291	0.0964
4	18.383	0.2176

S. mitis-CSP-2-Q4LN11F



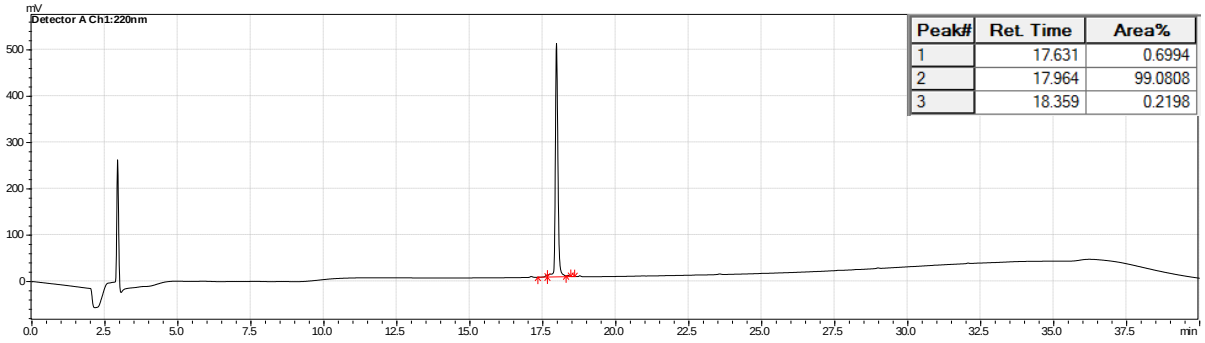
Peak#	Ret. Time	Area%
1	17.933	99.7740
2	18.276	0.2260

S. mitis-CSP-2-Q4LF12L



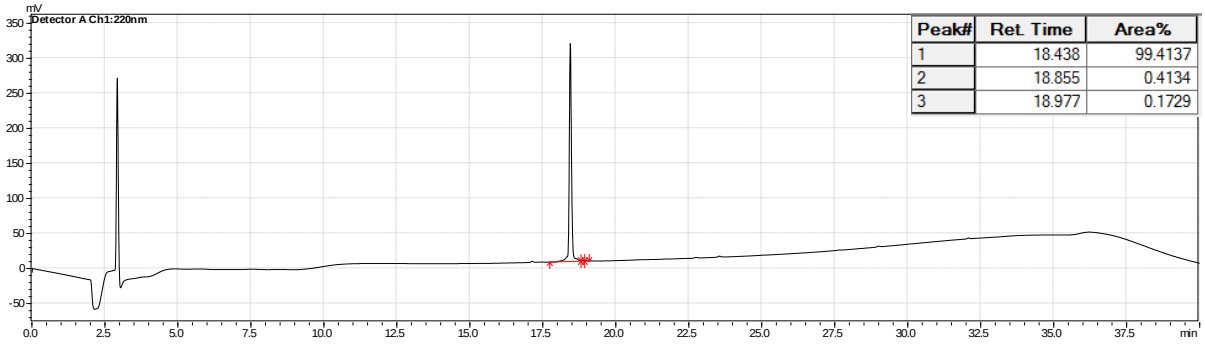
Peak#	Ret. Time	Area%
1	17.389	1.4845
2	17.490	2.2000
3	17.731	95.5194
4	18.019	0.4338
5	18.166	0.3623

S. mitis-CSP-2-N7FI8F



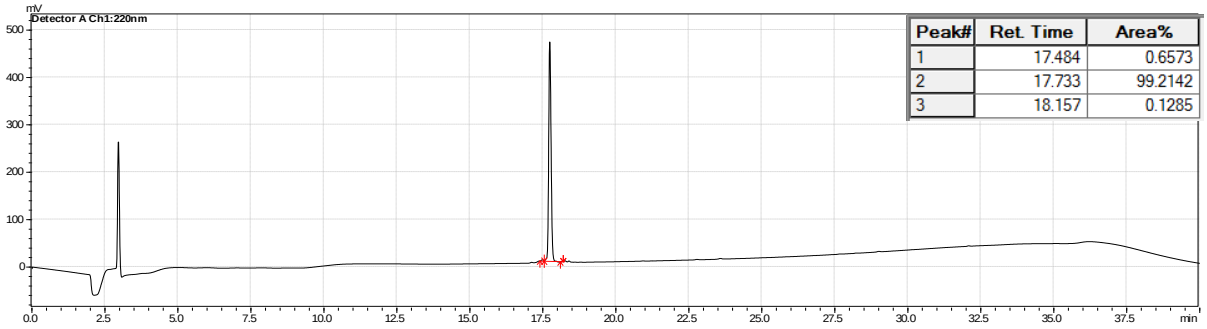
Peak#	Ret. Time	Area%
1	17.631	0.6994
2	17.964	99.0808
3	18.359	0.2198

S. mitis-CSP-2-N7FN11F



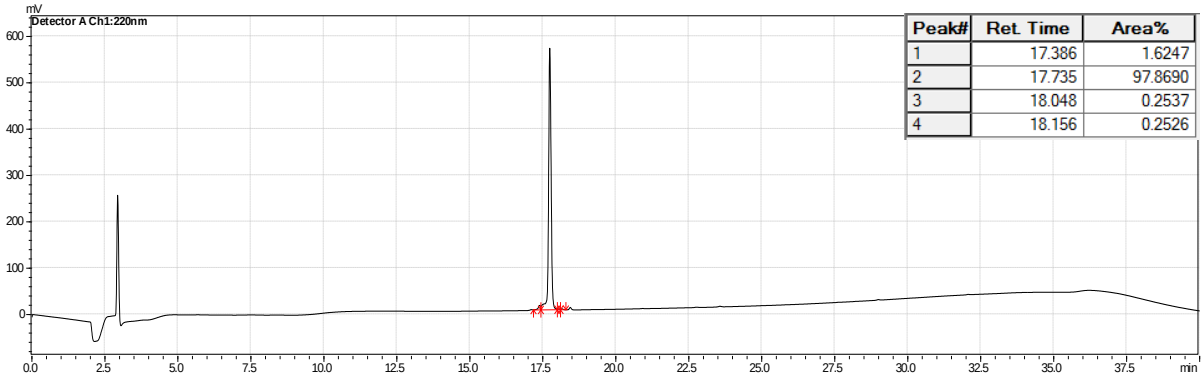
Peak#	Ret. Time	Area%
1	18.438	99.4137
2	18.855	0.4134
3	18.977	0.1729

S. mitis-CSP-2-N7FF12L



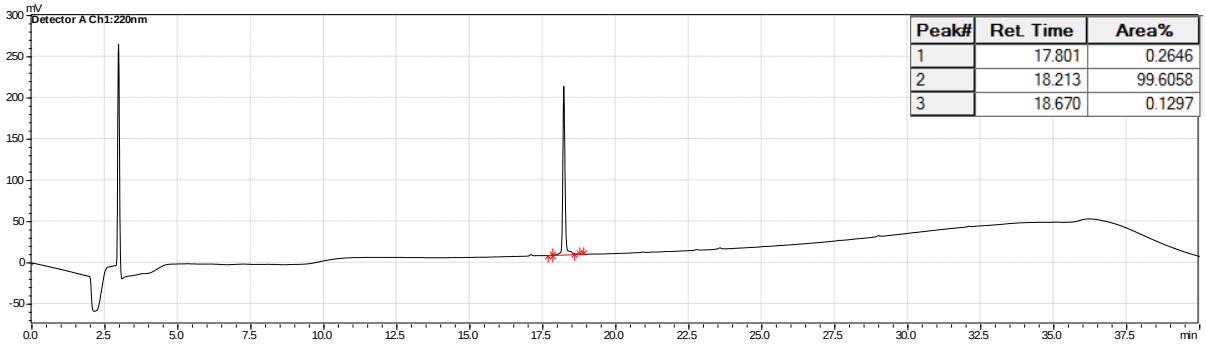
Peak#	Ret. Time	Area%
1	17.484	0.6573
2	17.733	99.2142
3	18.157	0.1285

S. mitis-CSP-2-N7II8F



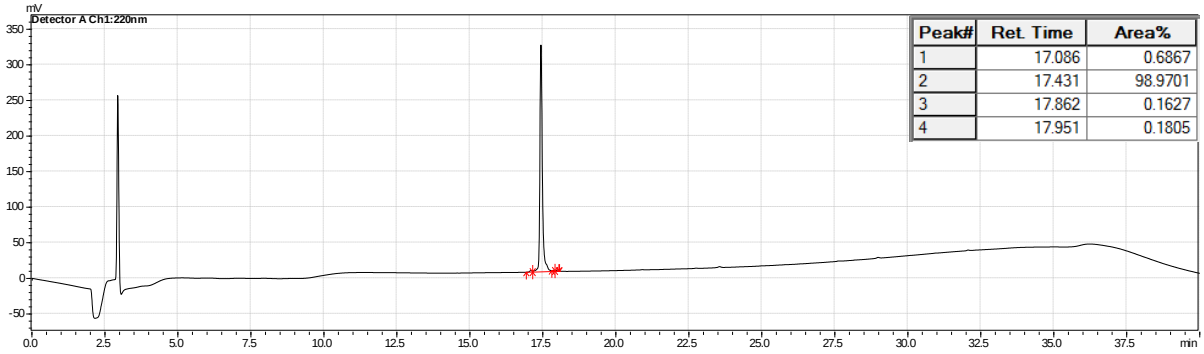
Peak#	Ret. Time	Area%
1	17.386	1.6247
2	17.735	97.8690
3	18.048	0.2537
4	18.156	0.2526

S. mitis-CSP-2-N7IN11F



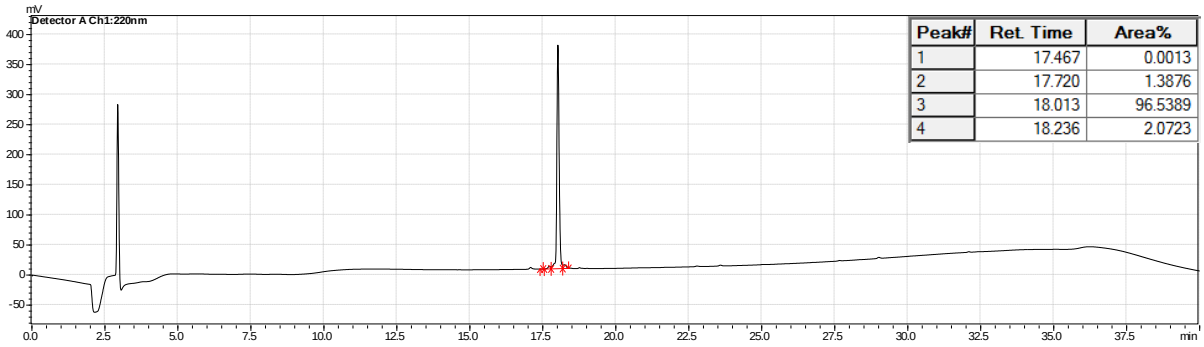
Peak#	Ret. Time	Area%
1	17.801	0.2646
2	18.213	99.6058
3	18.670	0.1297

S. mitis-CSP-2-N7IF12L



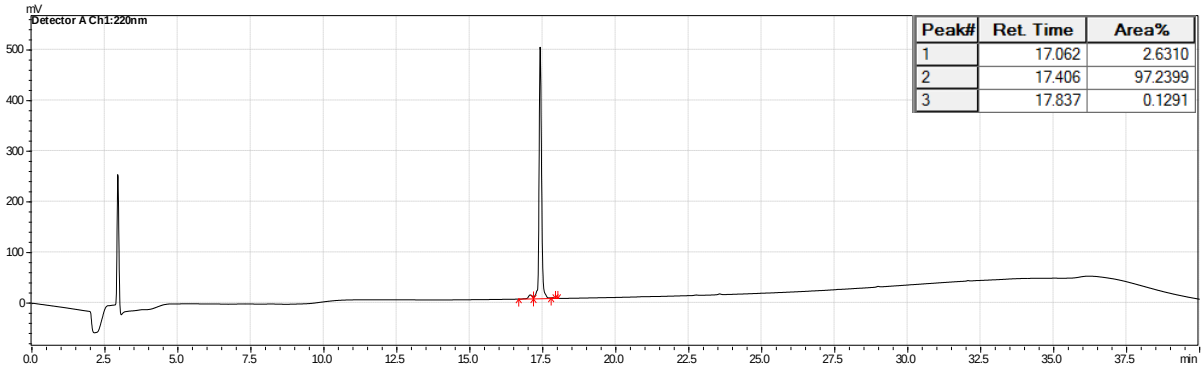
Peak#	Ret. Time	Area%
1	17.086	0.6867
2	17.431	98.9701
3	17.862	0.1627
4	17.951	0.1805

S. mitis-CSP-2-I8FN11F



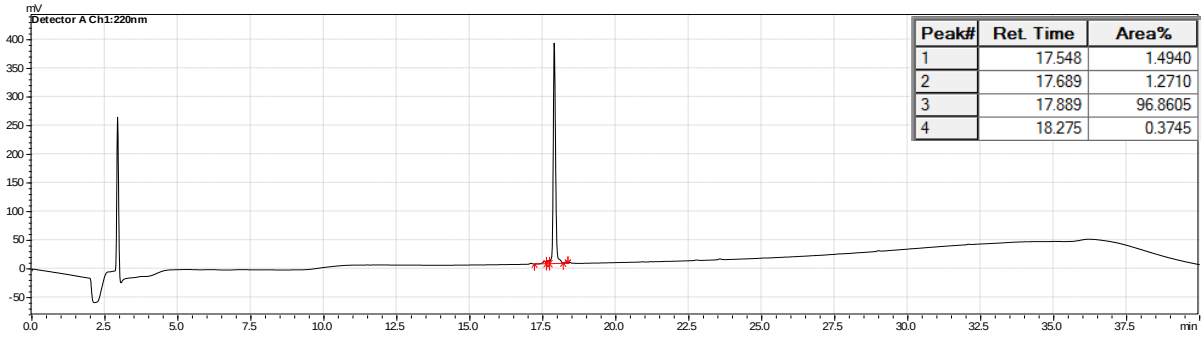
Peak#	Ret. Time	Area%
1	17.467	0.0013
2	17.720	1.3876
3	18.013	96.5389
4	18.236	2.0723

S. mitis-CSP-2-I8FF12L



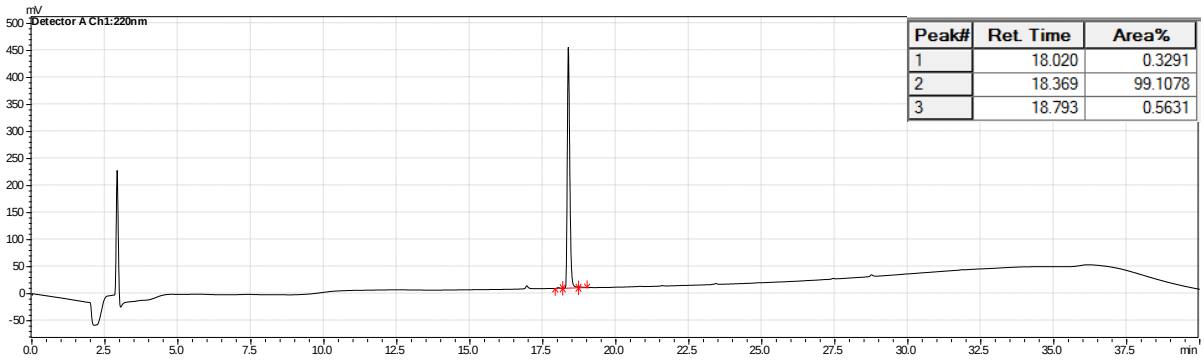
Peak#	Ret. Time	Area%
1	17.062	2.6310
2	17.406	97.2399
3	17.837	0.1291

S. mitis-CSP-2-N11FF12L



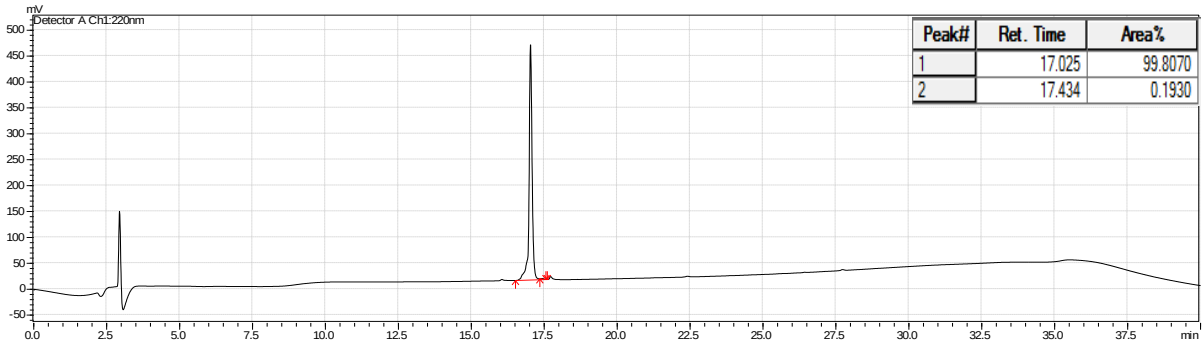
Peak#	Ret. Time	Area%
1	17.548	1.4940
2	17.689	1.2710
3	17.889	96.8605
4	18.275	0.3745

S. mitis-CSP-2-I2MQ4LN7F



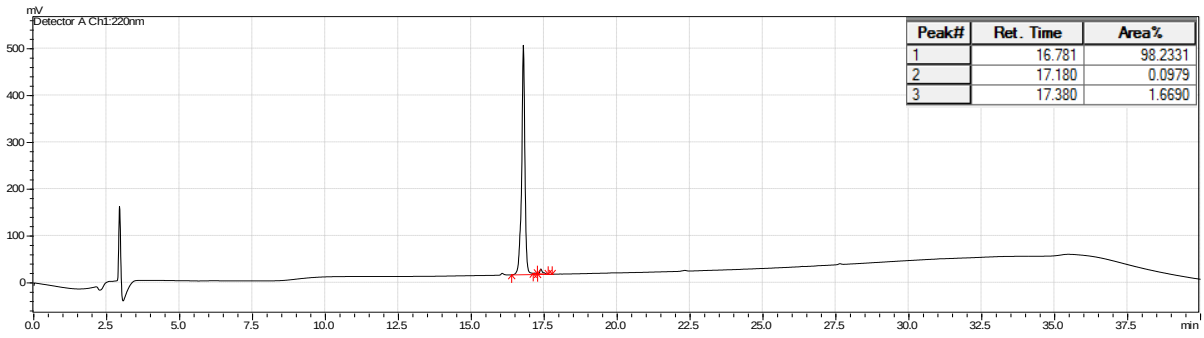
Peak#	Ret. Time	Area%
1	18.020	0.3291
2	18.369	99.1078
3	18.793	0.5631

S. mitis-CSP-2-I2MI8FN11F



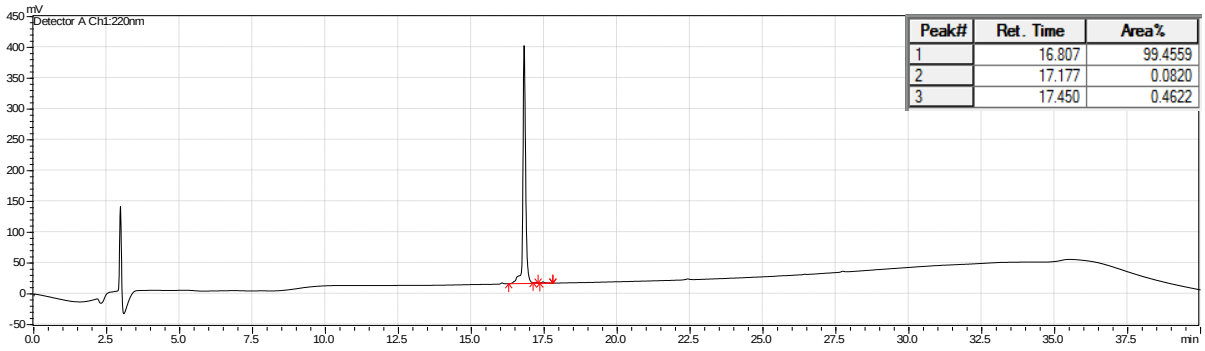
Peak#	Ret. Time	Area%
1	17.025	99.8070
2	17.434	0.1930

S. mitis-CSP-2-I2MN7FF12L

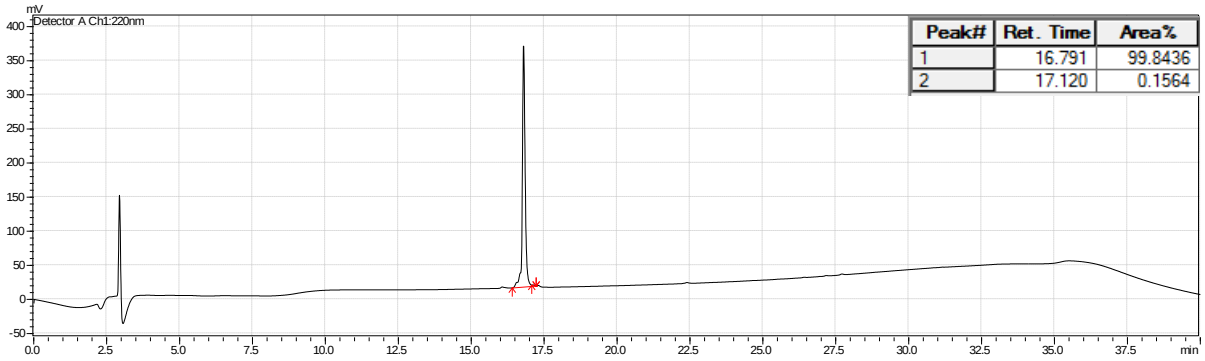


Peak#	Ret. Time	Area%
1	16.781	98.2331
2	17.180	0.0979
3	17.380	1.6690

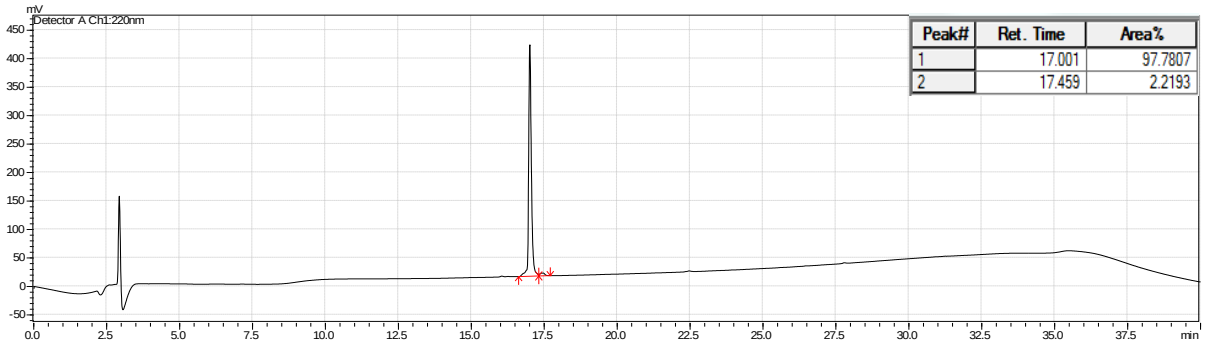
S. mitis-CSP-2-I2MN7II8F



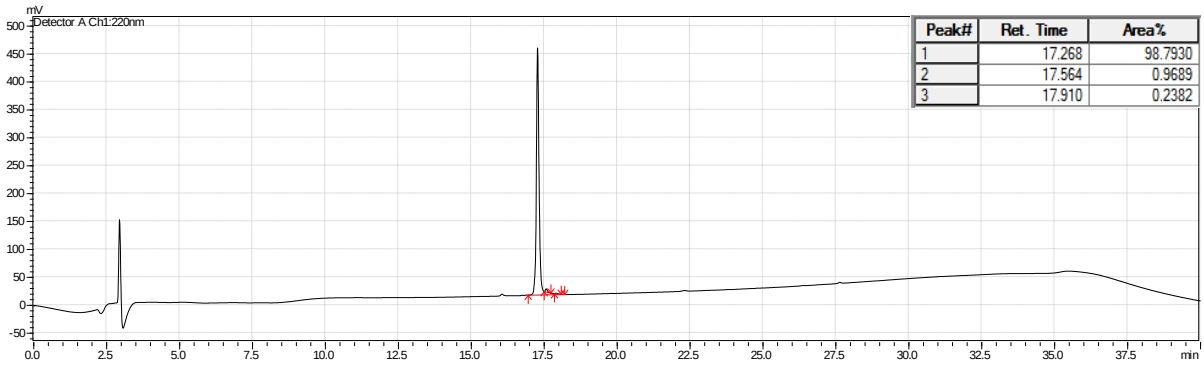
S. mitis-CSP-2-I2MQ4LF12L



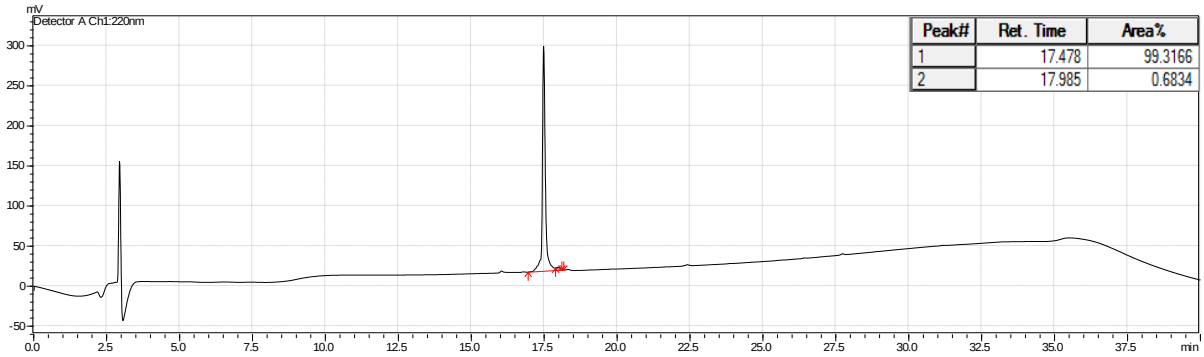
S. mitis-CSP-2-I2MQ4LI8F



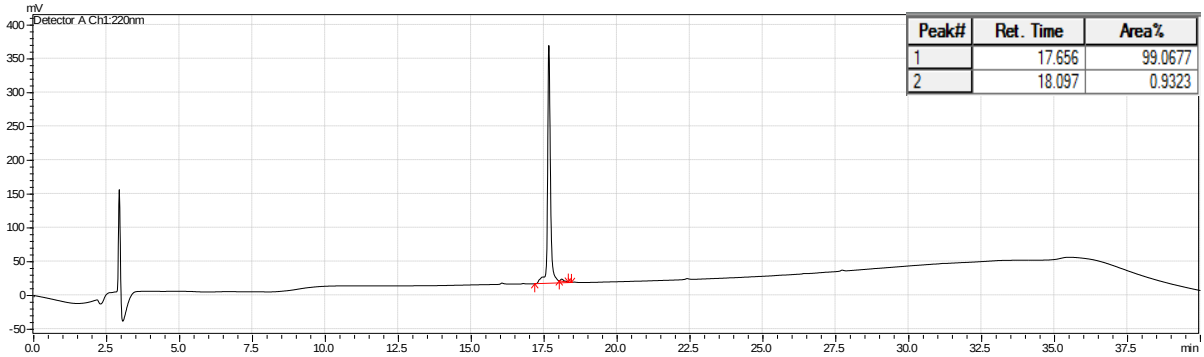
S. mitis-CSP-2-I2MQ4LN7I



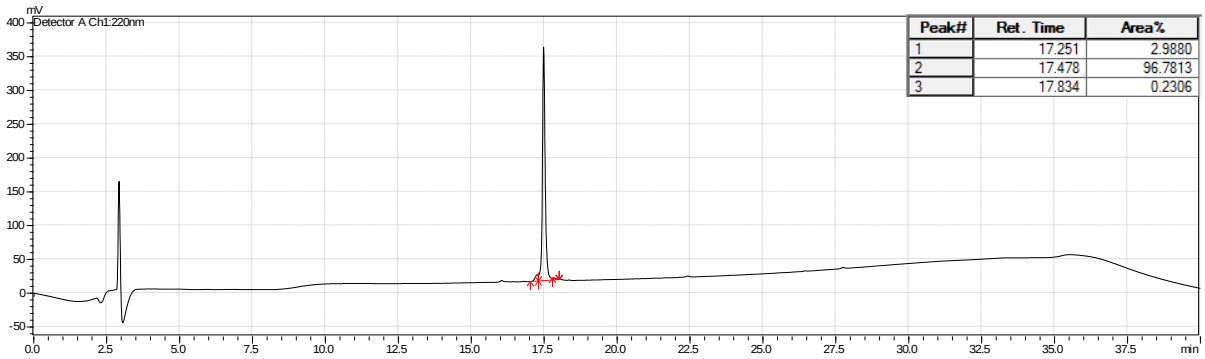
S. mitis-CSP-2-I2MQ4LN11F



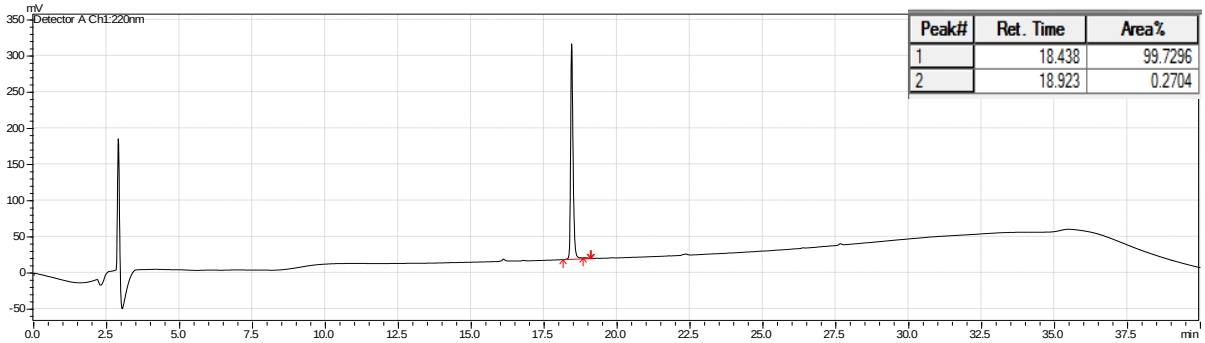
S. mitis-CSP-2-Q4LN7FI8F



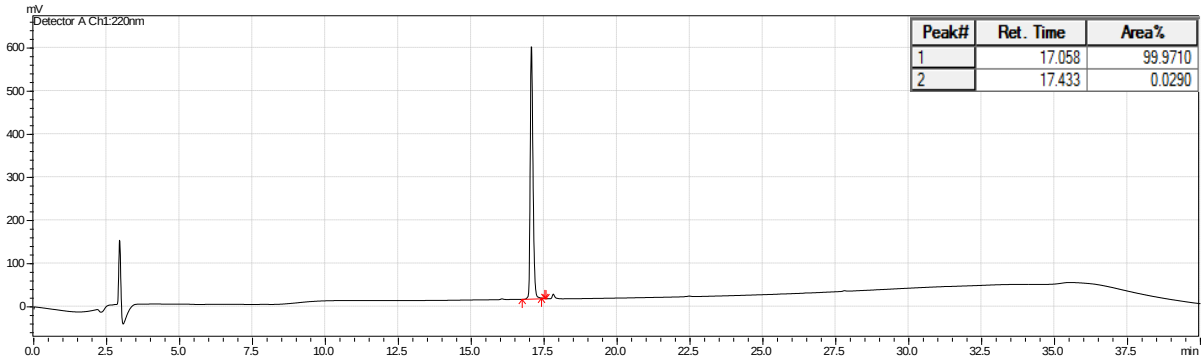
S. mitis-CSP-2-Q4LN7II8F



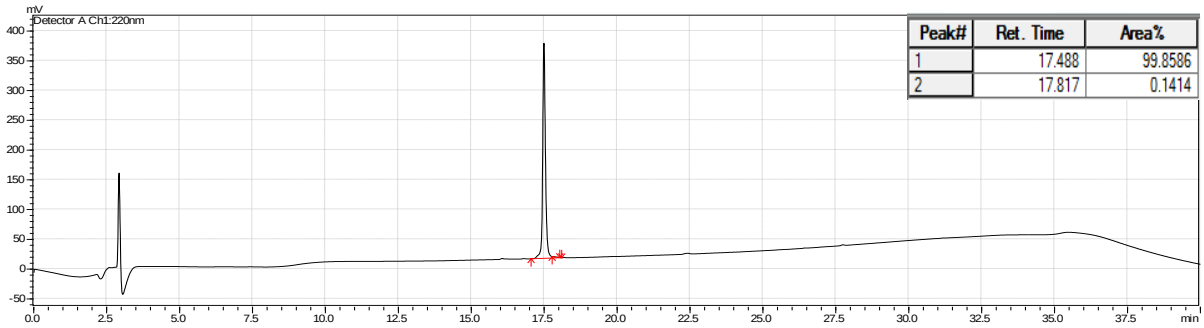
S. mitis-CSP-2-Q4LN7FN11F



S. mitis-CSP-2-I2MN7FI8F

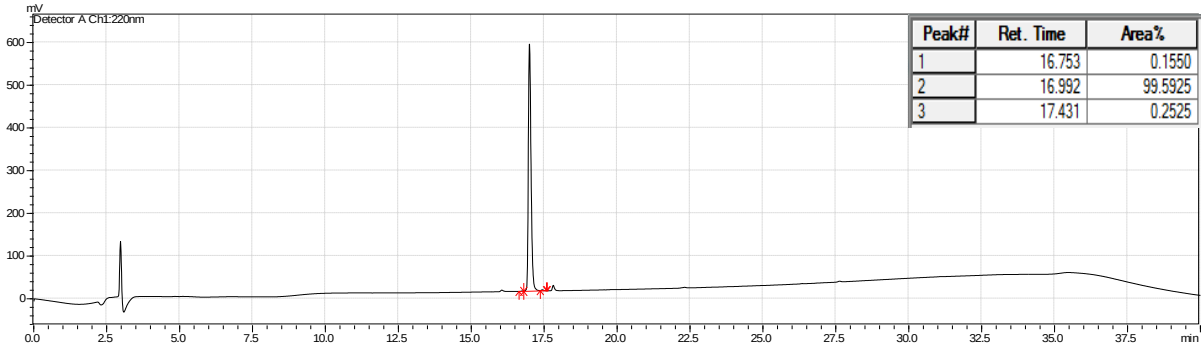


S. mitis-CSP-2-Q4LN7FF12L



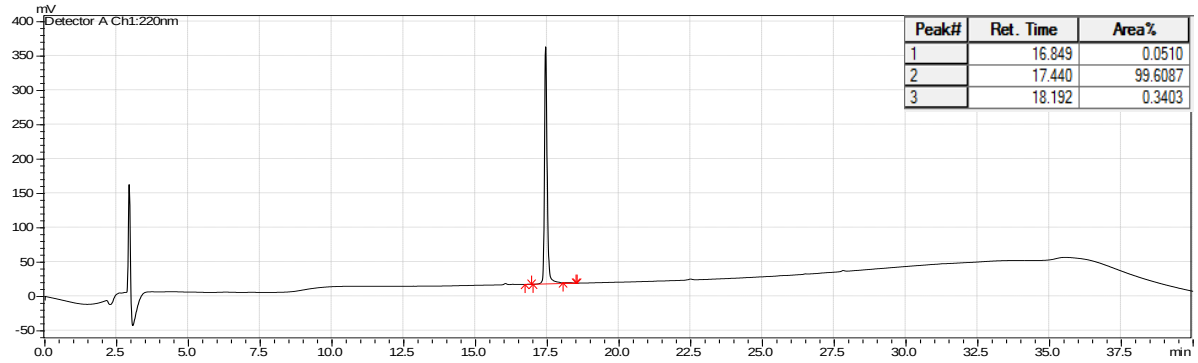
Peak#	Ret. Time	Area%
1	17.488	99.8586
2	17.817	0.1414

S. mitis-CSP-2-N7FI8FF12L



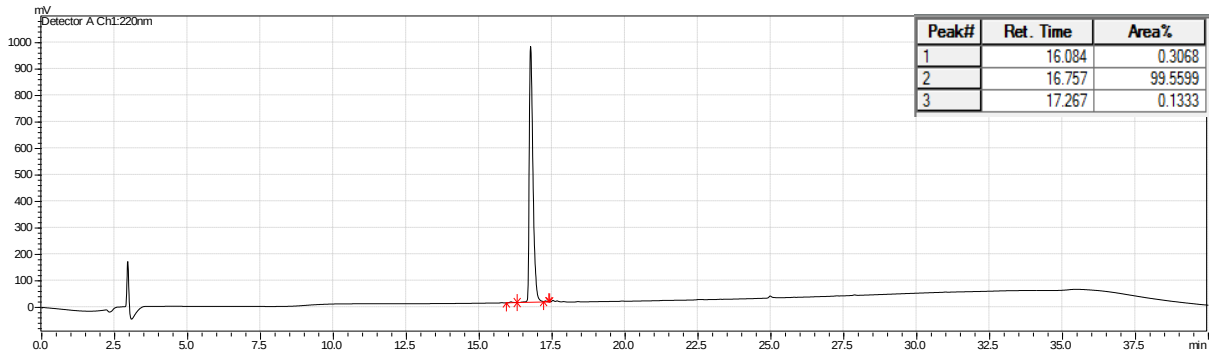
Peak#	Ret. Time	Area%
1	16.753	0.1550
2	16.992	99.5925
3	17.431	0.2525

S. mitis-CSP-2-N7II8FN11F



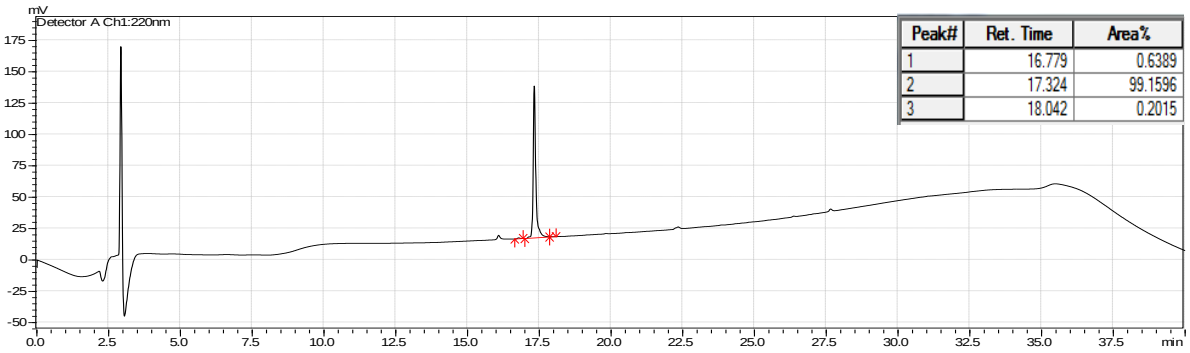
Peak#	Ret. Time	Area%
1	16.849	0.0510
2	17.440	99.6087
3	18.192	0.3403

S. mitis-CSP-2-N7II8FF12L



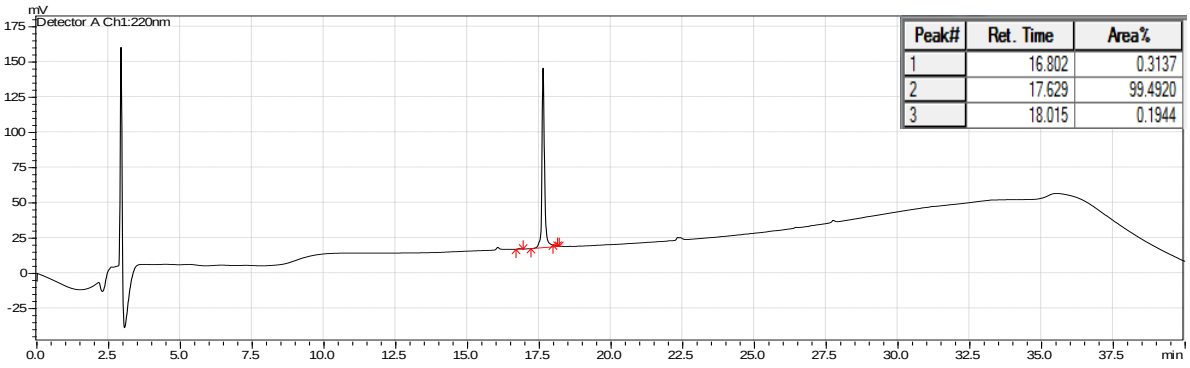
Peak#	Ret. Time	Area%
1	16.084	0.3068
2	16.757	99.5599
3	17.267	0.1333

S. mitis-CSP-2-I2MN7IN11F



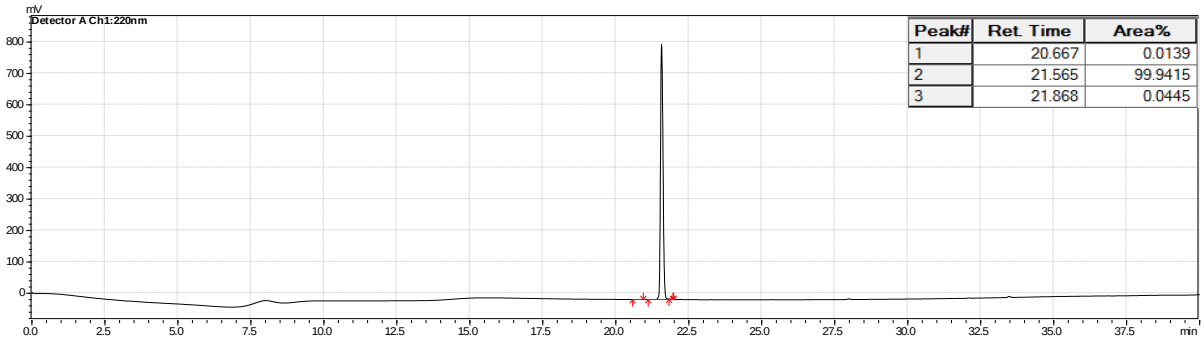
Peak#	Ret. Time	Area%
1	16.779	0.6389
2	17.324	99.1596
3	18.042	0.2015

S. mitis-CSP-2-N7FI8FN11F



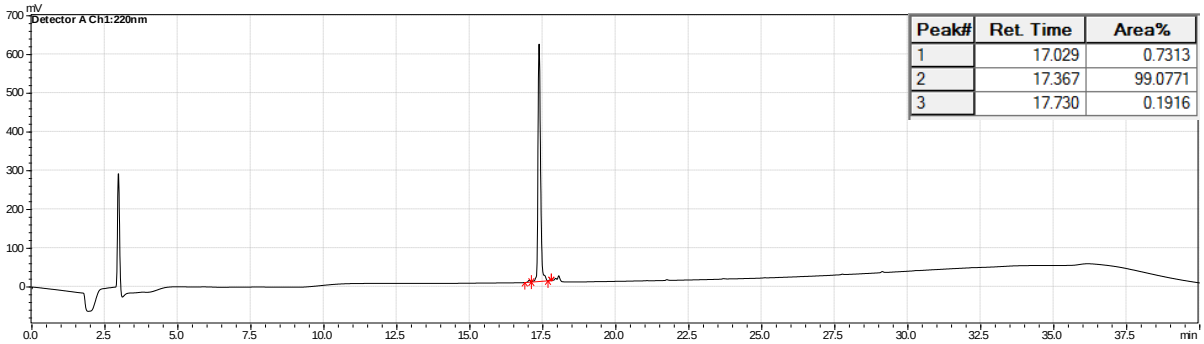
Peak#	Ret. Time	Area%
1	16.802	0.3137
2	17.629	99.4920
3	18.015	0.1944

S. mitis-CSP-2-E1A



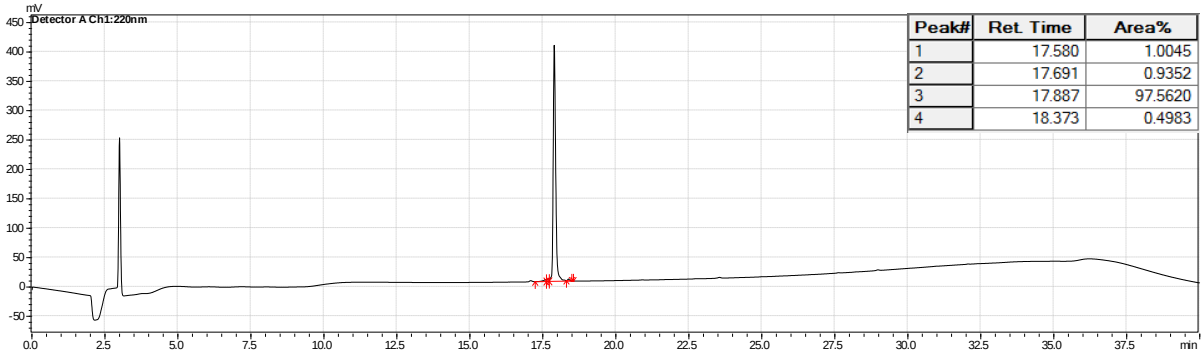
Peak#	Ret. Time	Area%
1	20.667	0.0139
2	21.565	99.9415
3	21.868	0.0445

S. mitis-CSP-2-E1AI2M



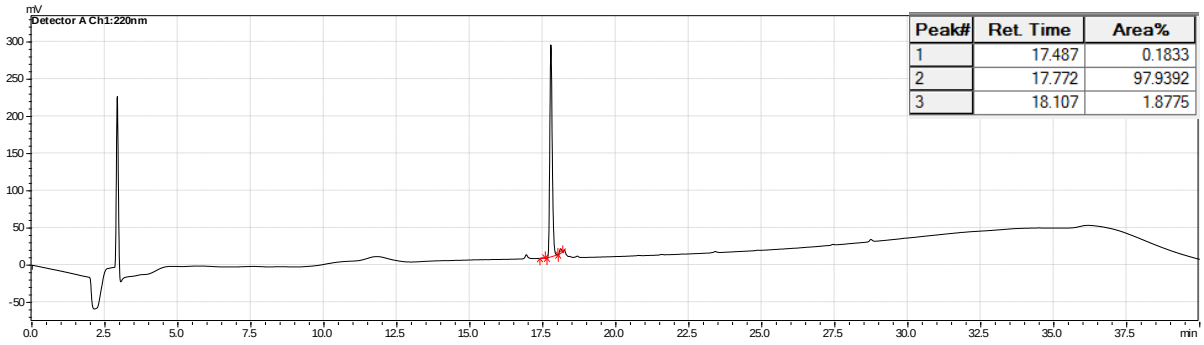
Peak#	Ret. Time	Area%
1	17.029	0.7313
2	17.367	99.0771
3	17.730	0.1916

S. mitis-CSP-2-E1AN11F



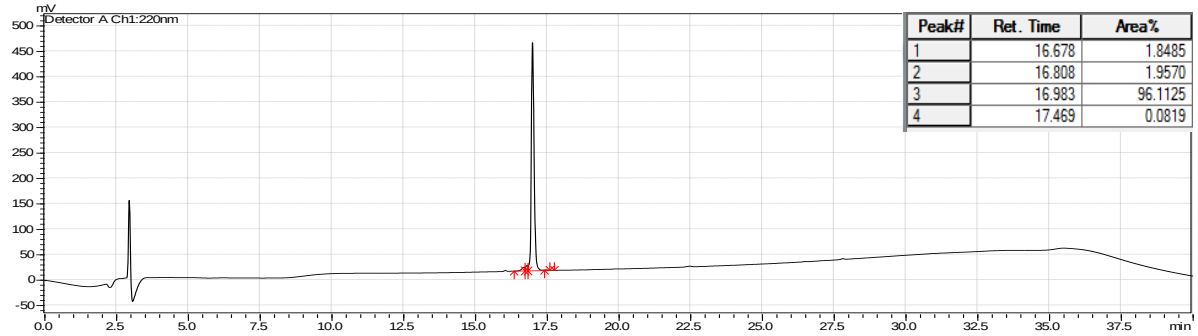
Peak#	Ret. Time	Area%
1	17.580	1.0045
2	17.691	0.9352
3	17.887	97.5620
4	18.373	0.4983

S. mitis-CSP-2-E1AI2MQ4L



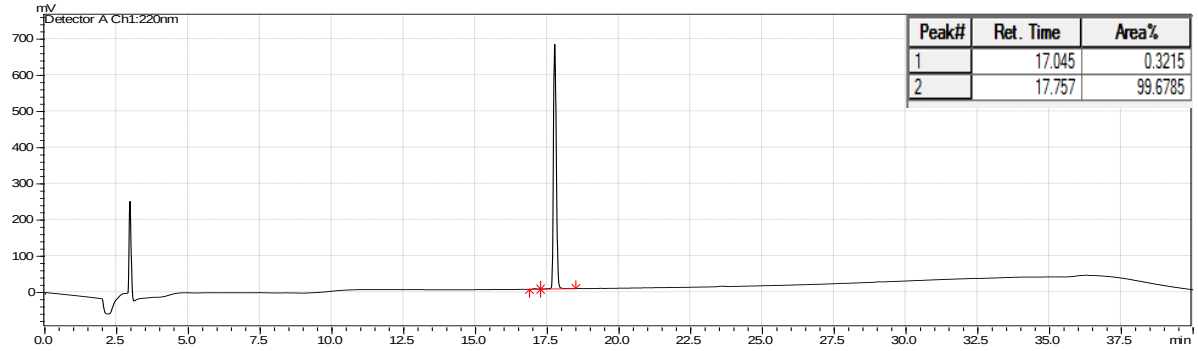
Peak#	Ret. Time	Area%
1	17.487	0.1833
2	17.772	97.9392
3	18.107	1.8775

S. mitis-CSP-2-E1AN11FF12L



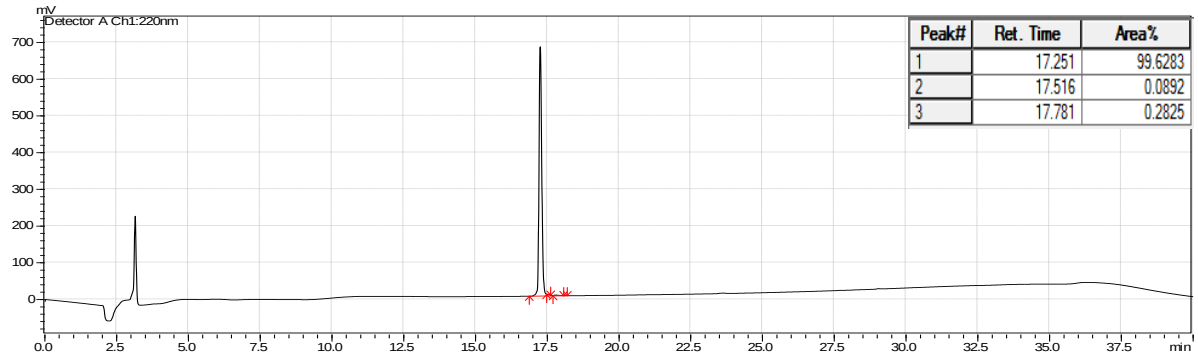
Peak#	Ret. Time	Area%
1	16.678	1.8485
2	16.808	1.9570
3	16.983	96.1125
4	17.469	0.0819

S. mitis-CSP-2-E1AN7II8F



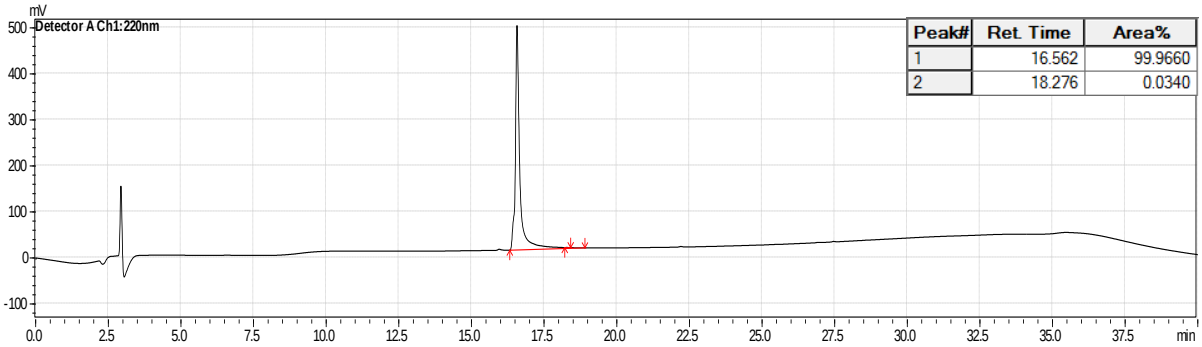
Peak#	Ret. Time	Area%
1	17.045	0.3215
2	17.757	99.6785

S. mitis-CSP-2-E1AI2MF12L



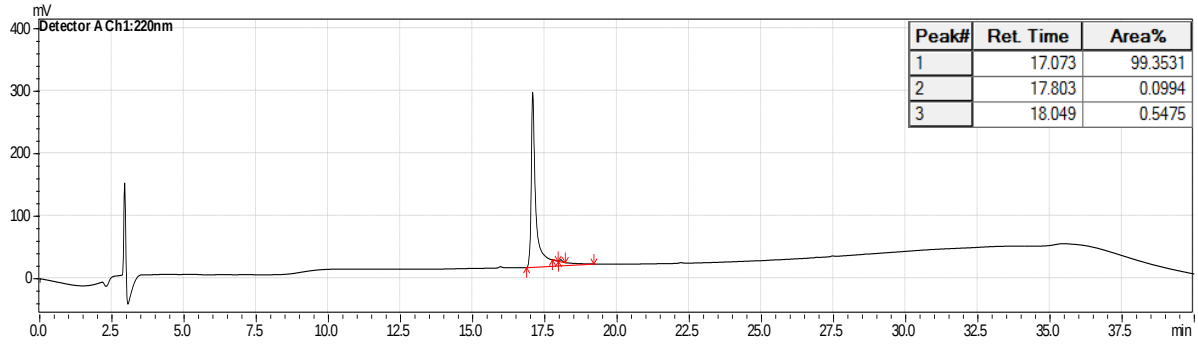
Peak#	Ret. Time	Area%
1	17.251	99.6283
2	17.516	0.0892
3	17.781	0.2825

S. mitis-CSP-2-E1AI2MI8F



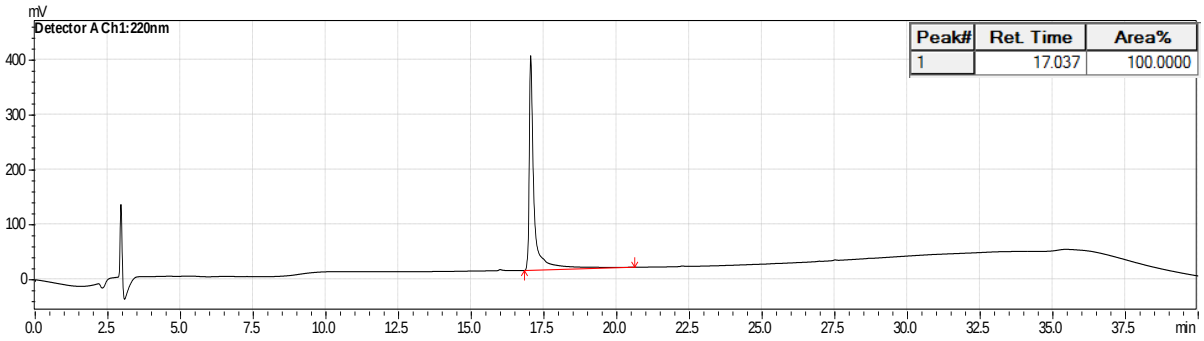
Peak#	Ret. Time	Area%
1	16.562	99.9660
2	18.276	0.0340

S. mitis-CSP-2-E1AI2MI8FN11F

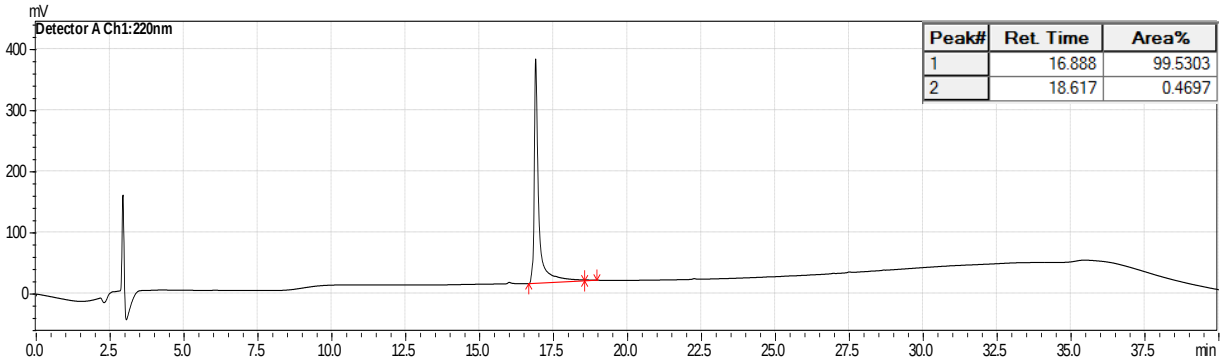


Peak#	Ret. Time	Area%
1	17.073	99.3531
2	17.803	0.0994
3	18.049	0.5475

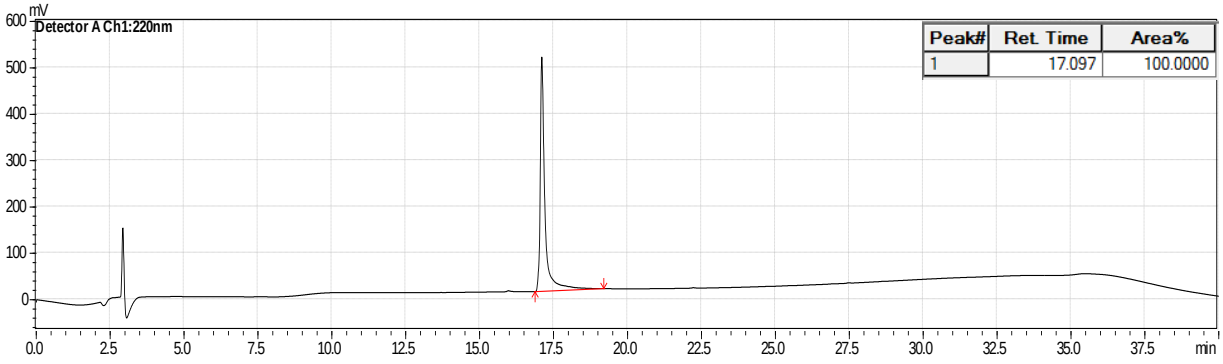
S. mitis-CSP-2-E1AI2MQ4LI8F



S. mitis-CSP-2-E1AI2MN7II8F



S. mitis-CSP-2-E1AI2MN7FI8F



MS and HPLC data for CSP analogues

Table S-1. MS and HPLC data for the synthetic native CSPs.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
S. mitis-CSP-1	1089.6404	1089.6442	≥ 99
S. mitis-CSP-2	1077.5808	1077.5850	≥ 99
S. cristatus-CSP	896.0720	896.0753	≥ 98
S. intermedius-CSP	952.9974	953.0012	≥ 99
S. oralis-CSP	1101.1281	1101.1232	≥ 99
S. oligofermentans-CSP	883.5474	883.5482	≥ 99
S. gordonii-CSP-1	1395.2448	1395.2455	≥ 99
S. gordonii-CSP-2	1241.2105	1241.2152	≥ 99
S. gordonii challis-CSP	1247.6808	1247.6844	≥ 99
S. sanguinis-CSP	885.9668	885.9653	≥ 99

EM = Exact Mass. See methods above.

Table S-2. MS and HPLC data for *S. mitis*-CSP-2-point modification analogues.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
S. mitis-CSP-2-I2M	1086.5590	1086.5602	≥ 98
S. mitis-CSP-2-Q4L	1070.0936	1070.0942	≥ 99
S. mitis-CSP-2-N7F	729.7315*	729.7332*	≥ 98
S. mitis-CSP-2-I8F	1094.5730	1094.5764	≥ 98
S. mitis-CSP-2-F10D	1061.5601	1061.5640	≥ 99
S. mitis-CSP-2-N11F	1094.0936	1094.0977	≥ 99
S. mitis-CSP-2-F12I	1060.5886	1060.5867	≥ 97
S. mitis-CSP-2-F12L	1060.5886	1060.5921	≥ 99
S. mitis-CSP-2-N7I	1077.1014	1077.0982	≥ 98
S. mitis-CSP-2-Q4I	1070.0936	1070.0964	≥ 99

EM = Exact Mass. See methods above, *MH_3^{3+}.

Table S-3. MS and HPLC data for *S. mitis*-CSP-2-multiple modification analogues.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
S. mitis-CSP-2-I2MQ4L	1079.0718	1079.0687	≥ 99
S. mitis-CSP-2-I2MN7F	1103.0718	1103.0717	≥ 99
S. mitis-CSP-2-I2MN7I	1086.0796	1086.0803	≥ 99
S. mitis-CSP-2-I2MI8F	1103.5512	1103.5481	≥ 99
S. mitis-CSP-2-I2MN11F	1103.0718	1103.0748	≥ 99
S. mitis-CSP-2-I2MF12L	1069.5668	1069.5640	≥ 99
S. mitis-CSP-2-Q4LN7F	724.7400*	724.7381*	≥ 99
S. mitis-CSP-2-Q4LN7I	1069.6142	1069.6135	≥ 99
S. mitis-CSP-2-Q4LI8F	1087.0858	1087.0833	≥ 98
S. mitis-CSP-2-Q4LN11F	1086.6063	1086.6083	≥ 99
S. mitis-CSP-2-Q4LF12L	1053.1014	1053.0997	≥ 95
S. mitis-CSP-2-N7FI8F	1111.0858	1111.0843	≥ 99
S. mitis-CSP-2-N7FN11F	1110.6063	1110.6092	≥ 99
S. mitis-CSP-2-N7FF12L	1077.1014	1077.0977	≥ 99
S. mitis-CSP-2-N7II8F	1094.0936	1094.0988	≥ 97
S. mitis-CSP-2-N7IN11F	1093.6142	1093.6111	≥ 99
S. mitis-CSP-2-N7IF12L	1060.1092	1060.1133	≥ 98
S. mitis-CSP-2-I8FN11F	1111.0858	1111.0847	≥ 96
S. mitis-CSP-2-I8FF12L	1077.5808	1077.5810	≥ 97
S. mitis-CSP-2-N11FF12L	1077.1014	1077.1060	≥ 96
S. mitis-CSP-2-I2MQ4LN7F	1095.5845	1095.5869	≥ 99
S. mitis-CSP-2-I2MI8FN11F	1120.0640	1120.0617	≥ 99
S. mitis-CSP-2-I2MN7FF12L	1086.0796	1086.0753	≥ 98
S. mitis-CSP-2-I2MN7II8F	1103.0718	1103.0679	≥ 99
S. mitis-CSP-2-I2MQ4LF12L	1062.0796	1062.0760	≥ 99
S. mitis-CSP-2-I2MQ4LI8F	1096.0639	1096.0586	≥ 97
S. mitis-CSP-2-I2MQ4LN7I	1078.5923	1078.5886	≥ 98
S. mitis-CSP-2-I2MQ4LN11F	1095.5845	1095.5807	≥ 99
S. mitis-CSP-2-Q4LN7FI8F	1103.5985	1103.5936	≥ 99
S. mitis-CSP-2-Q4LN7II8F	1086.6063	1086.6027	≥ 96
S. mitis-CSP-2-Q4LN7FN11F	1103.1191	1103.1189	≥ 99
S. mitis-CSP-2-I2MN7FI8F	1120.0640	1120.0586	≥ 99
S. mitis-CSP-2-Q4LN7FF12L	1069.6142	1069.6119	≥ 99
S. mitis-CSP-2-N7FI8FF12L	1094.0936	1094.0899	≥ 99
S. mitis-CSP-2-N7II8FN11F	1110.6063	1110.6049	≥ 99
S. mitis-CSP-2-N7II8FF12L	1077.1014	1077.0976	≥ 98
S. mitis-CSP-2-I2MN7IN11F	1102.5924	1102.5881	≥ 99
S. mitis 2 CSP-N7FI8FN11F	1127.5985	1127.6010	≥ 99

EM = Exact Mass. See methods above, $^{*}MH_3^{3+}$.

Table S-4. MS and HPLC data for *S. mitis*-CSP-2-E1A modification analogues.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
S. mitis-CSP-2-E1A	1048.5781	1048.5754	≥ 99
S. mitis-CSP-2-E1AI2M	705.3733*	705.3763*	≥ 99
S. mitis-CSP-2-E1AN11F	1065.0908	1065.0892	≥ 97
S. mitis-CSP-2-E1AI2MQ4L	1050.0690	1050.0645	≥ 97
S. mitis-CSP-2-E1AN11FF12L	1048.0987	1048.0947	≥ 96
S. mitis-CSP-2-E1AN7II8F	1065.0908	1065.0857	≥ 99
S. mitis-CSP-2-E1AI2MF12L	1040.5641	1040.5611	≥ 99
S. mitis-CSP-2-E1AI2MI8F	1074.5484	1074.5499	≥ 99
S. mitis-CSP-2-E1AI2MI8FN11F	1091.0612	1091.0647	≥ 99
S. mitis-CSP-2-E1AI2MQ4LI8F	1067.0612	1067.0613	≥ 99
S. mitis-CSP-2-E1AI2MN7II8F	1074.069	1074.0735	≥ 99
S. mitis-CSP-2-E1AI2MN7FI8F	1091.0612	1091.0608	≥ 99

EM = Exact Mass. See methods above, $^{*}\mathbf{MH_3^{3+}}$.

Primary reporter gene assay data

***S. pneumoniae* D39pcomX::lacZ (ComD1)**

Agonism assays were performed at 10 µM concentration of synthetic CSP. *S. pneumoniae* CSP1 was used as the positive control (100%) while DMSO as the negative control (0%). Percent (%) ComD1 activation was measured by normalizing the Miller units obtained for each peptide to that of the native CSP1. All peptides were screened in triplicate over three separate trials. Error bars indicate standard error of the mean of nine values.

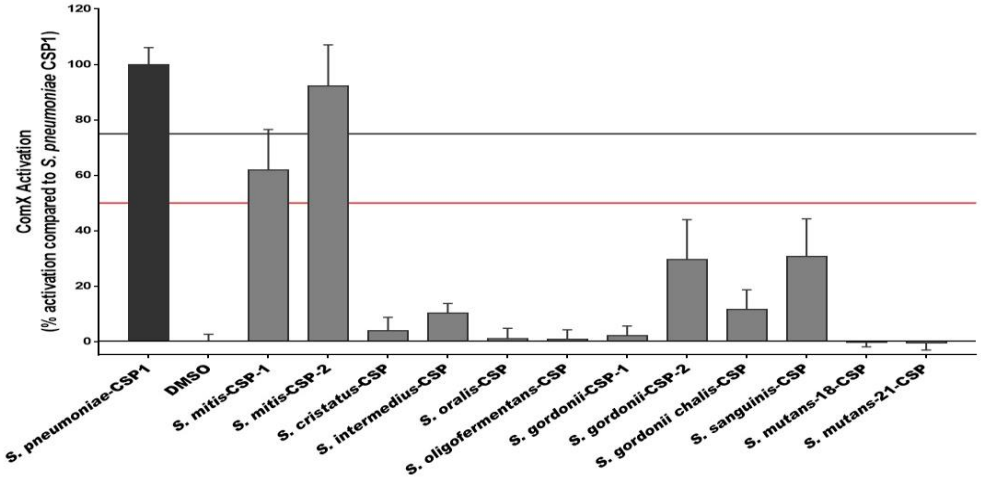


Figure S-1. Primary agonism screening assay data for the synthetic *Streptococci* native CSP pheromones. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

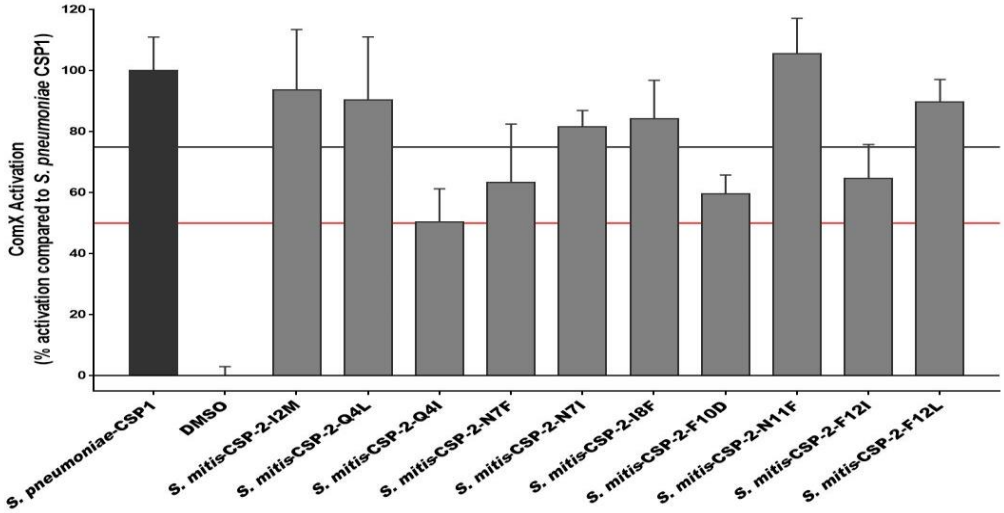


Figure S-2. Primary agonism screening assay data for the *S. mitis*-CSP-2-point modification analogues. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} values.

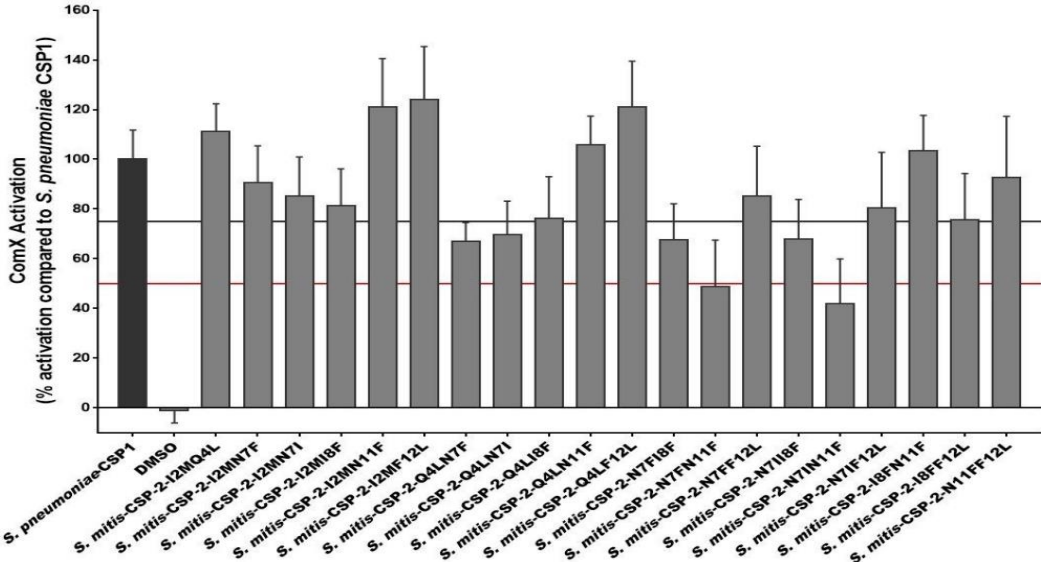


Figure S-3. Primary agonism screening assay data for the *S. mitis*-CSP-2 double modification analogues. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

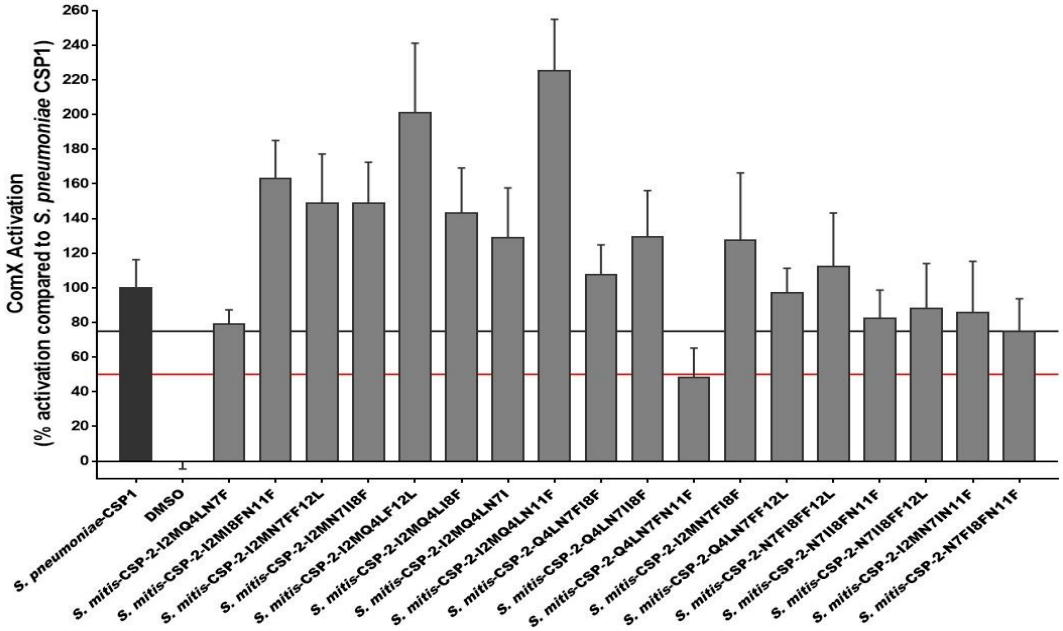


Figure S-4. Primary agonism screening assay data for the *S. mitis*-CSP-2 triple modification analogues. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

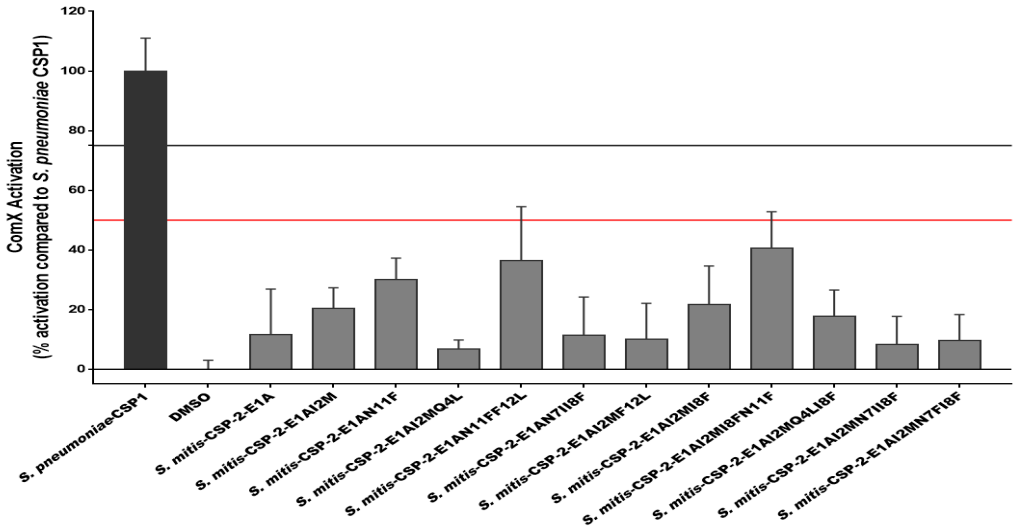


Figure S-5. Primary agonism screening assay data for the *S. mitis*-CSP-2-E1A modification analogues. None of the peptides exhibited activation of the *S. pneumoniae* ComD1 receptor and peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

Antagonism assays were performed at 10 μM concentration of peptides against 50 nM concentration of *S. pneumoniae* CSP1. *S. pneumoniae* CSP1 (50 nM) was used as the positive control (100%) while DMSO as the negative control (0%). Percent (%) *comX* activation was measured by normalizing the Miller units obtained for each peptide to that of CSP1. All peptides were screened in triplicate over three separate trials. Error bars indicate standard error of the mean of nine values.

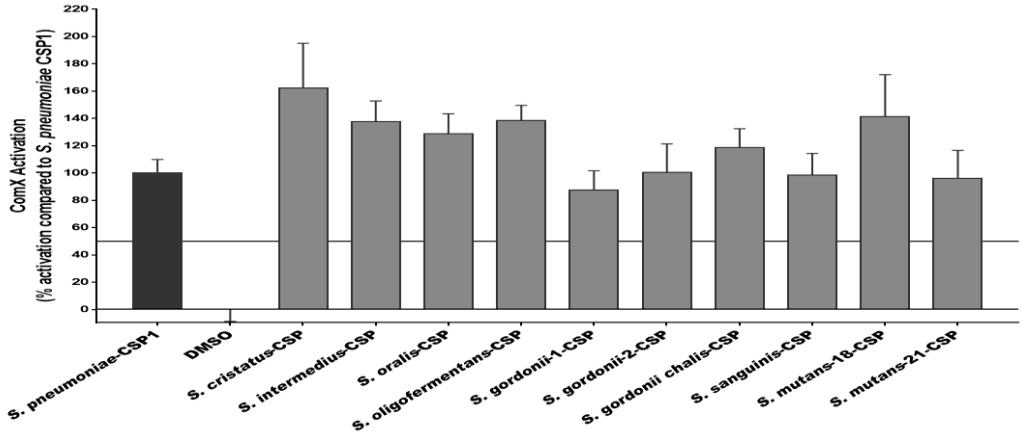


Figure S-6. Primary antagonism screening assay data for the synthetic *Streptococci* native CSP pheromones. None of the peptides exhibited inhibition of the *S. pneumoniae* ComD1 receptor.

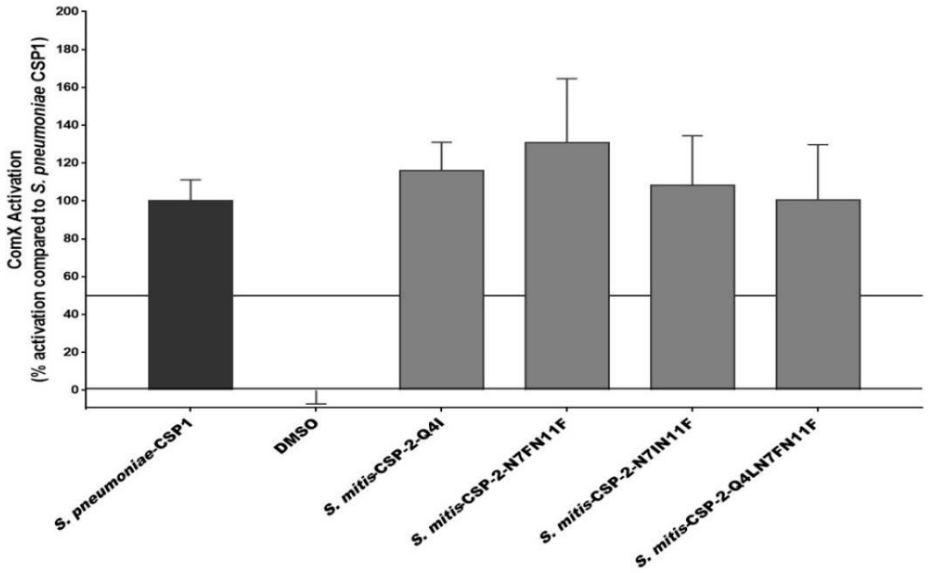


Figure S-7. Primary antagonism screening assay data for the *S. mitis*-CSP-2 point and multiple modification analogues. None of the peptides exhibited inhibition of the *S. pneumoniae* ComD1 receptor.

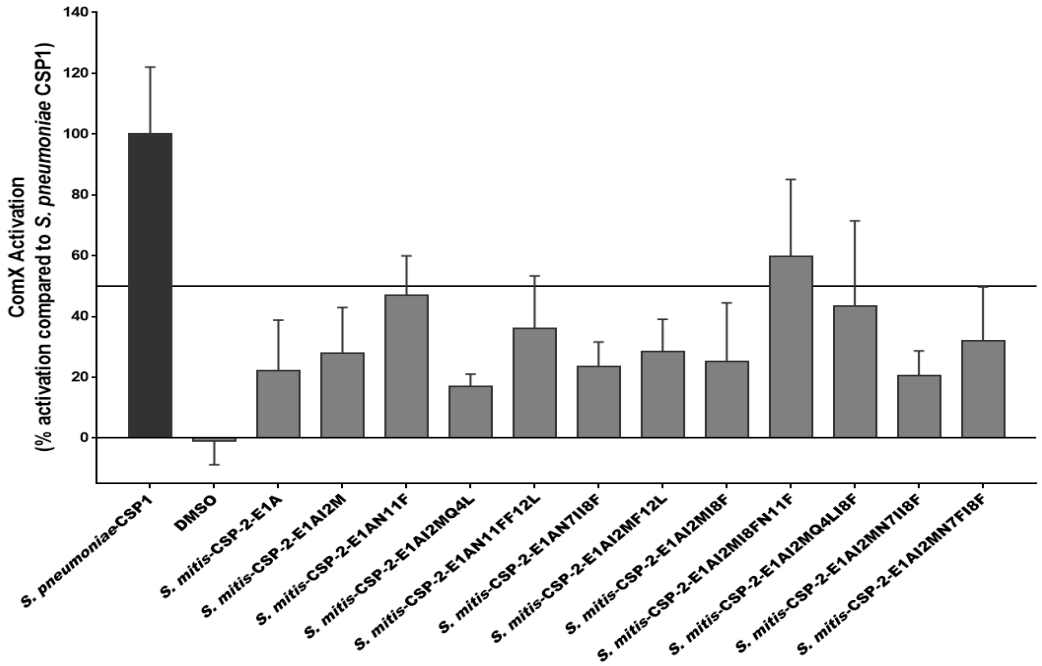


Figure S-8. Primary antagonism screening assay data for the *S. mitis*-CSP-2-E1A modification analogues. Peptides that exhibited less than 50% activation were further evaluated to determine their IC_{50}.

***S. pneumoniae* TIGR4pcomX::lacZ (ComD2)**

Agonism assays were performed at 10 μM concentration. *S. pneumoniae* CSP2 was used as the positive control (100%) while DMSO as the negative control (0%). Percent (%) *comX* activation was measured by normalizing the Miller units obtained for each peptide to that of CSP2. All peptides were screened in triplicate over three separate trials. Error bars indicate standard error of the mean of nine values.

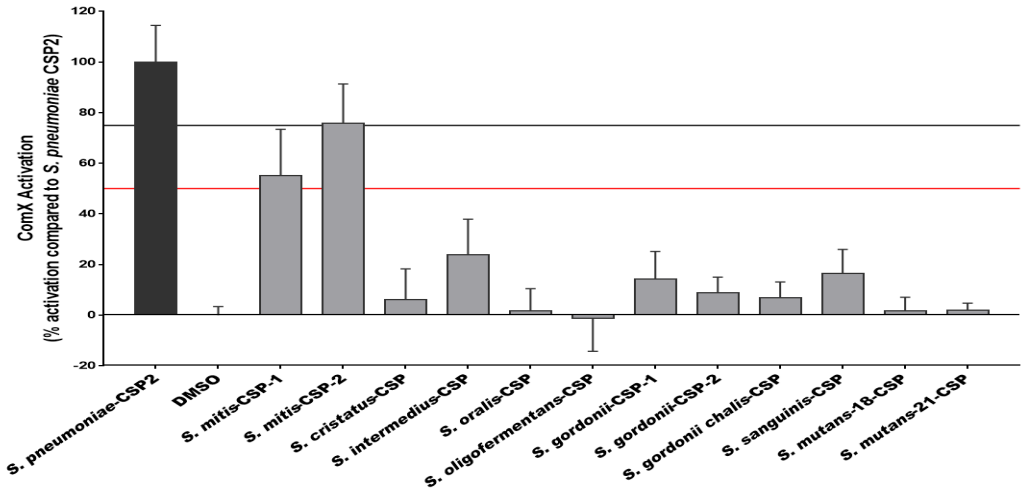


Figure S-9. Primary agonism screening assay data for the synthetic *Streptococci* native CSP pheromones. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

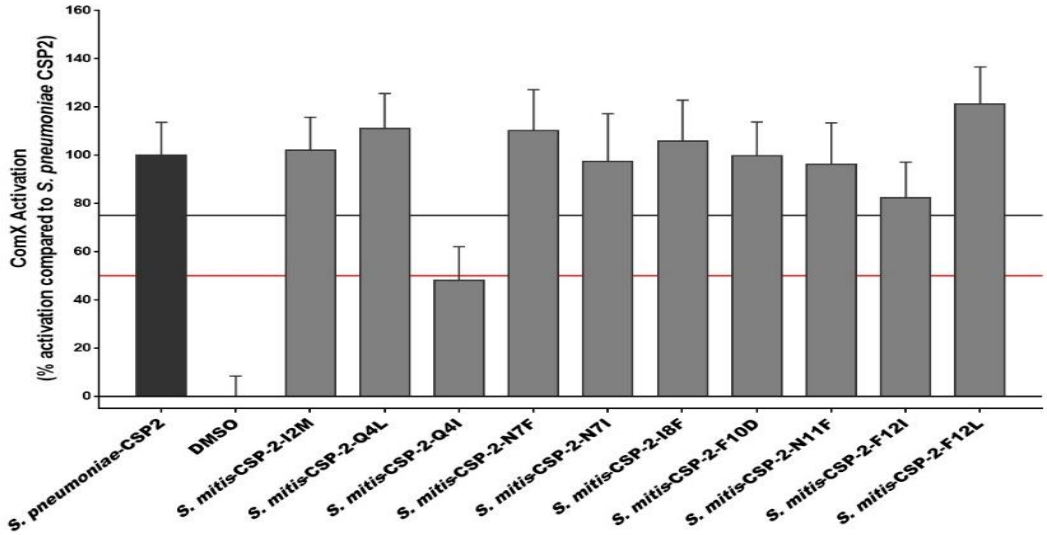


Figure S-10. Primary agonism screening assay data for the *S. mitis*-CSP-2-point modification analogues. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

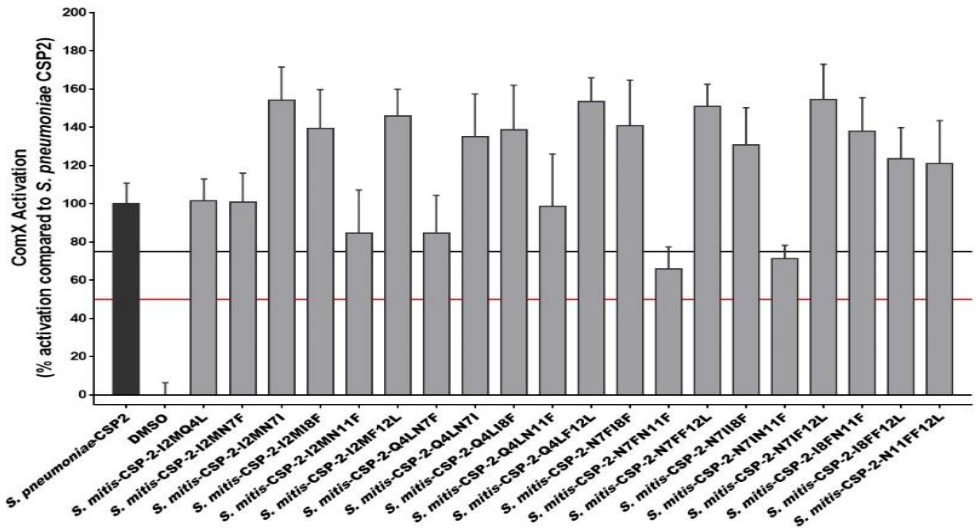


Figure S-11. Primary agonism screening assay data for the *S. mitis*-CSP-2 double modification analogues. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} values.

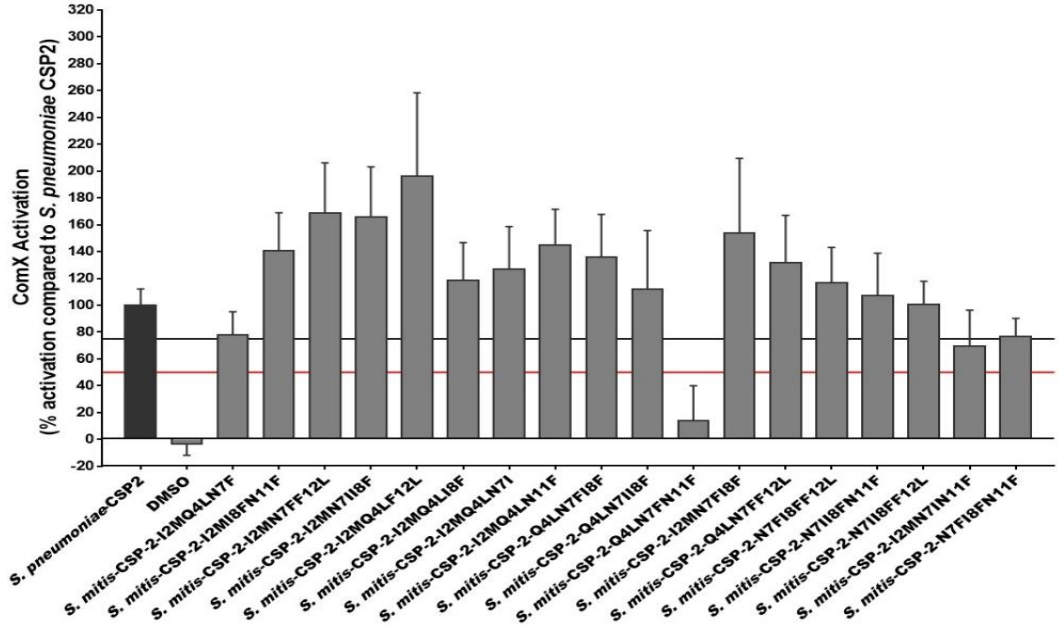


Figure S-12. Primary agonism screening assay data for the *S. mitis*-CSP-2 triple modification analogues. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

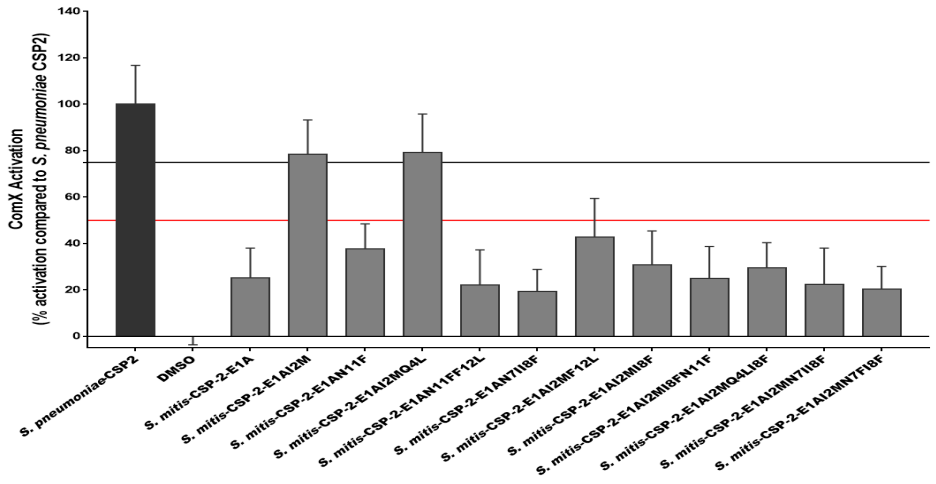


Figure S-13. Primary agonism screening assay data for the *S. mitis*-CSP-2-E1A modification analogues. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

Antagonism assays were performed at 10 μM concentration of peptides against 250 nM concentration of *S. pneumoniae* CSP2. *S. pneumoniae* CSP2 (250 nM) was used as the positive control (100%) while DMSO as the negative control (0%). Percent (%) *comX* activation was measured by normalizing the Miller units obtained for each peptide to that of CSP2. All peptides were screened in triplicate over three separate trials. Error bars indicate standard error of the mean of nine values.

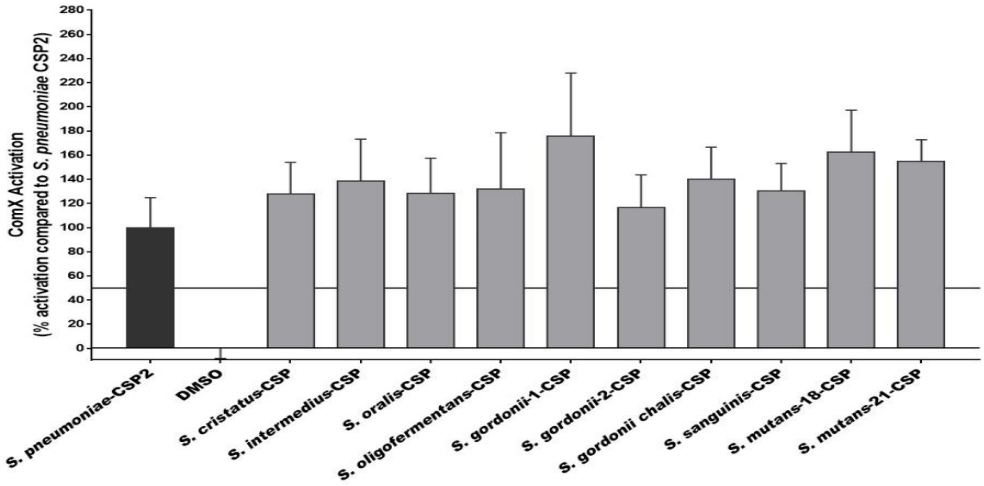


Figure S-14. Primary antagonism screening assay data for the synthetic *Streptococci* native CSP pheromones. None of the peptides exhibited inhibition of the *S. pneumoniae* ComD2 receptor.

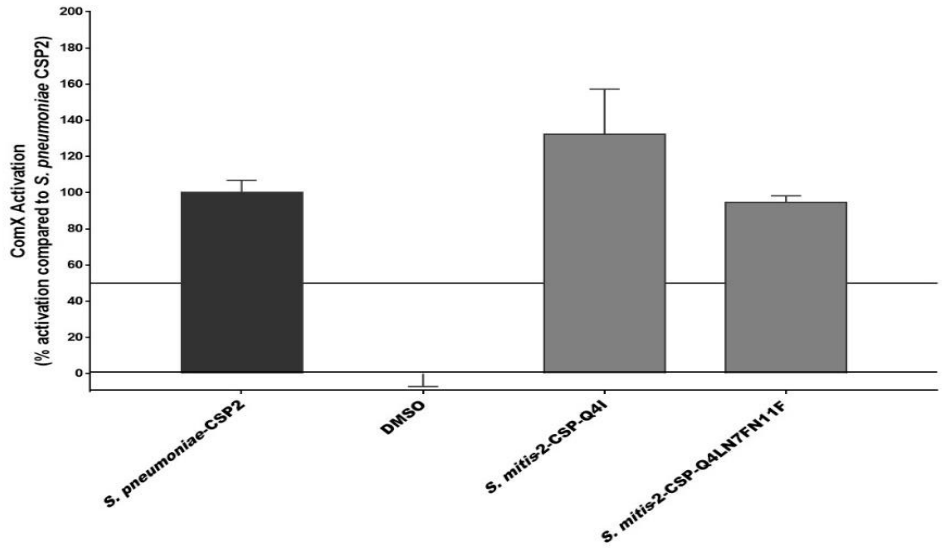


Figure S-15. Primary antagonism screening assay data for the *S. mitis*-CSP-2 point and multiple modification analogues. None of the peptides exhibited inhibition of the *S. pneumoniae* ComD2 receptor.

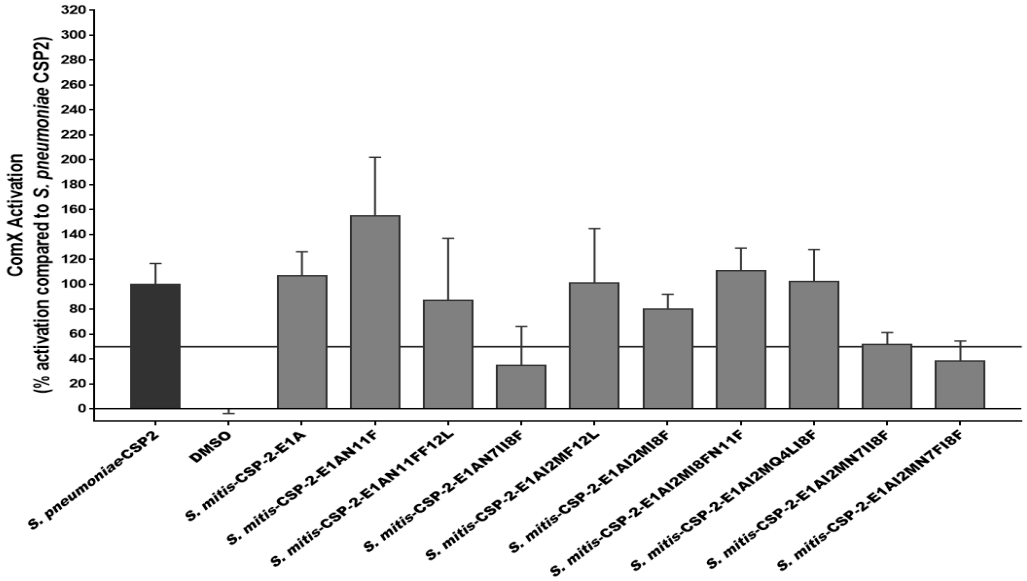


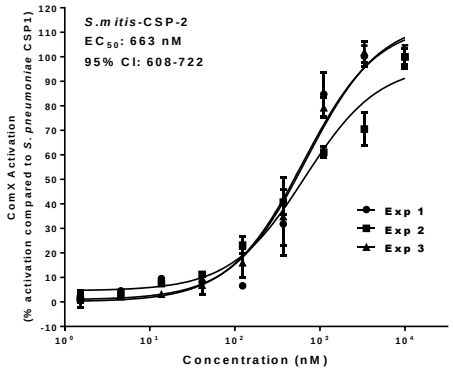
Figure S-16. Primary antagonism screening assay data for the *S. mitis*-CSP-2-E1A modification analogues. Peptides that exhibited less than 50% activation were further evaluated to determine their $IC_{50.}$

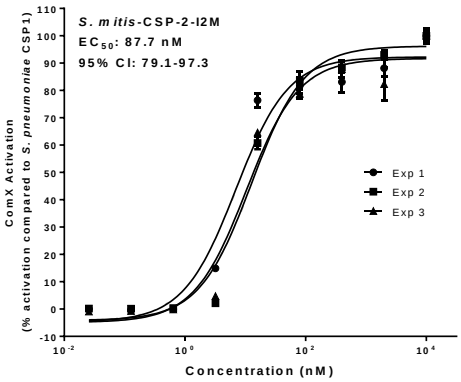
Agonism and antagonism dose response curves

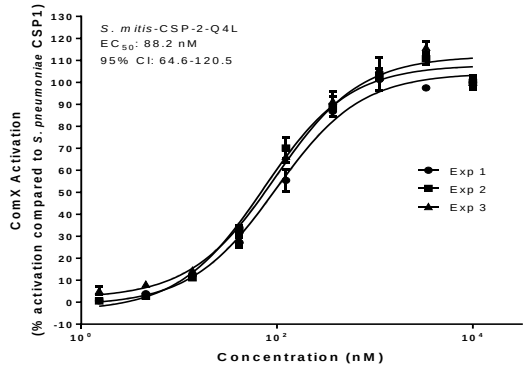
S. mitis-CSP-2 analogues were tested to determine their EC_{50} or IC_{50} values over varying concentrations in the two indicated *S. pneumoniae* beta-galactosidase reporter strains. Each dose response experiment was performed in triplicate on three separate occasions (i.e., experiments (Exp.) #1-3; shown for each peptide below). Error bars indicate standard error of the mean of triplicate values. In each plot, the peptide, as well as its EC_{50} or IC_{50} value (in nM) and 95% confidence interval (95% CI) values (in nM), are indicated at top left.

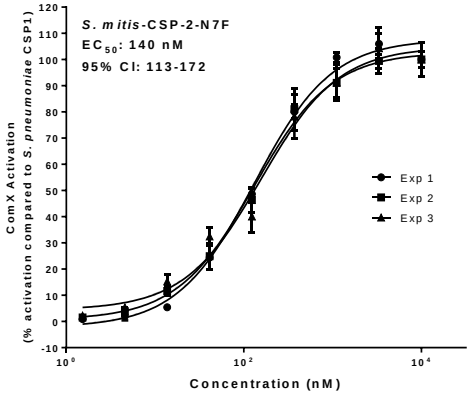
S. pneumoniae D39 pcomX::lacZ (ComD1)

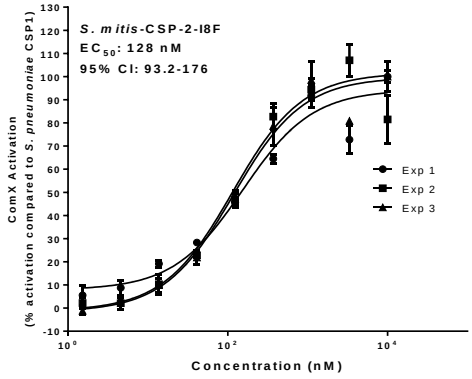
Activation dose response curves

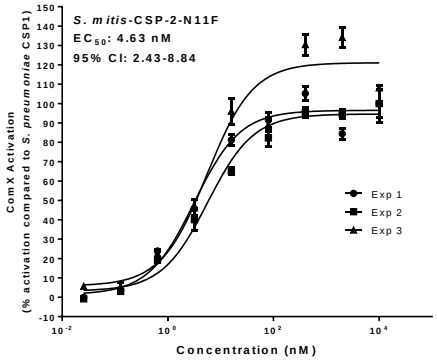


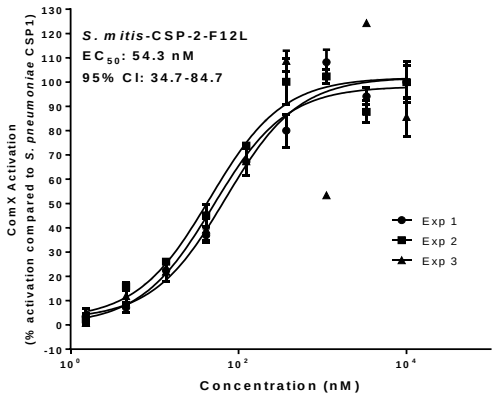


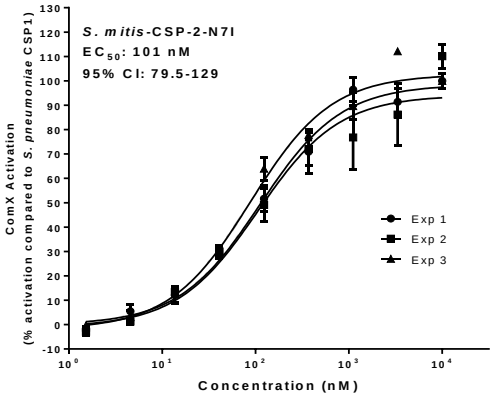


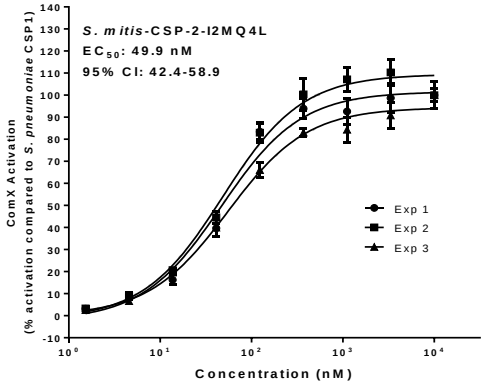


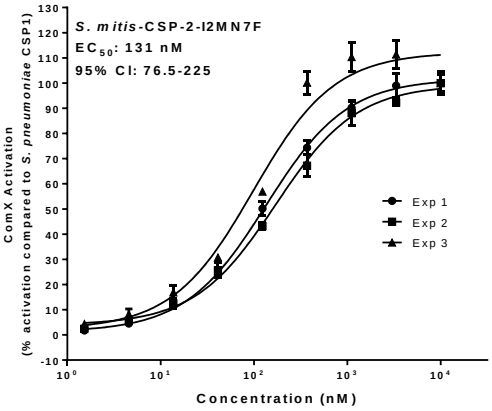


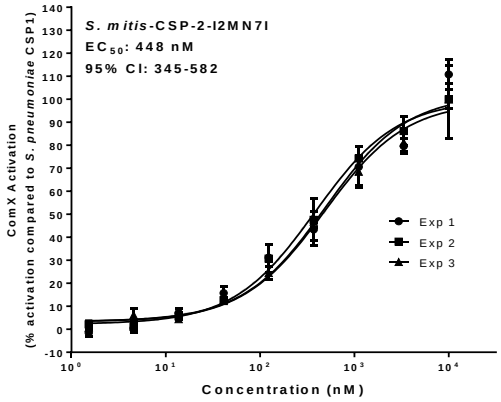


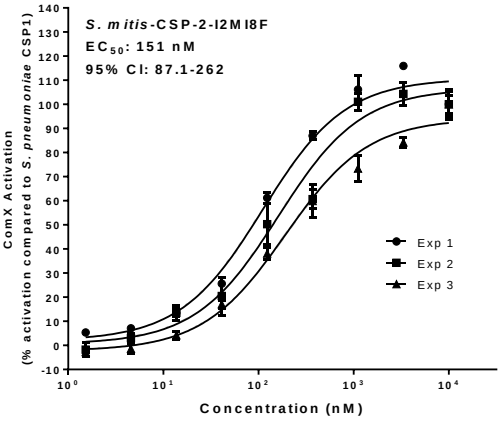

S. mitis-CSP-2-F12L
EC50: 54.3 nM
95% CI: 34.7-84.7
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

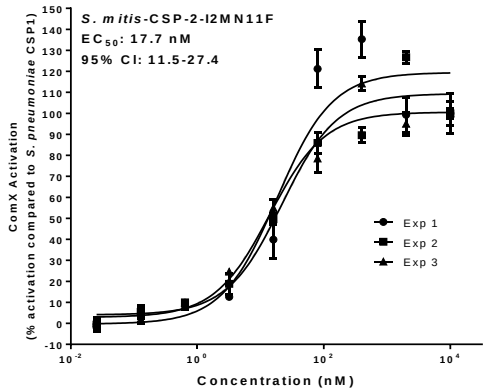

S. mitis-CSP-2-N7I
EC50: 101 nM
95% CI: 79.5-129
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

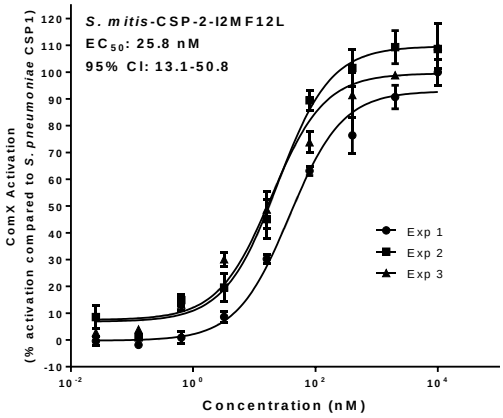

S. mitis-CSP-2-I2MQ4L
EC50: 49.9 nM
95% CI: 42.4-58.9
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

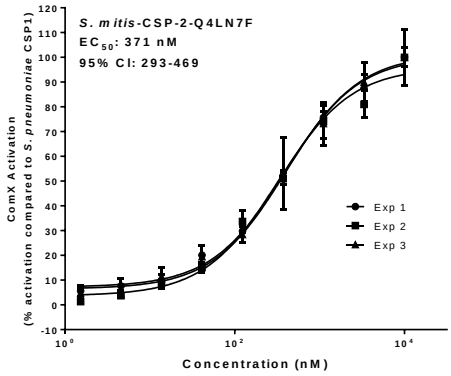

S. mitis-CSP-2-I2MN7F
EC50: 131 nM
95% CI: 76.5-225
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

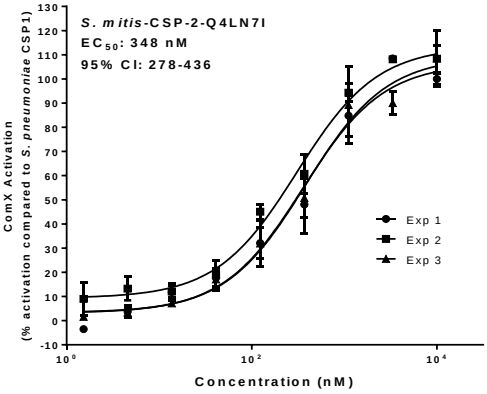

S. mitis-CSP-2-I2MN7I
EC50: 448 nM
95% CI: 345-582
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

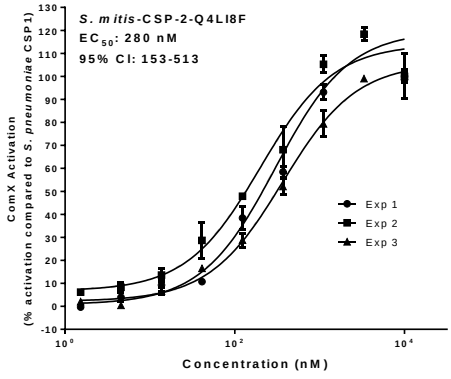

S. mitis-CSP-2-I2MI8F
EC50: 151 nM
95% CI: 87.1-262
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

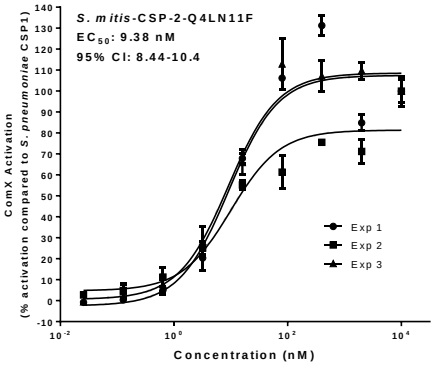

S. mitis-CSP-2-I2MN11F
EC50: 17.7 nM
95% CI: 11.5-27.4
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

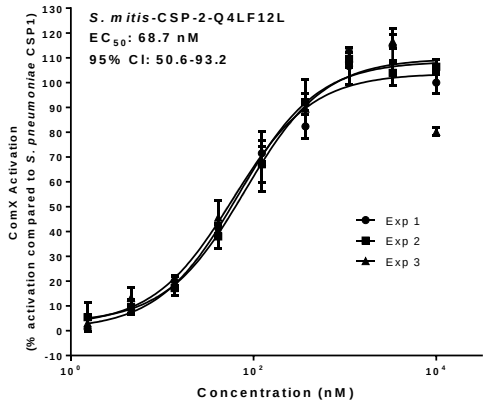

S. mitis-CSP-2-I2MF12L
EC50: 25.8 nM
95% CI: 13.1-50.8
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

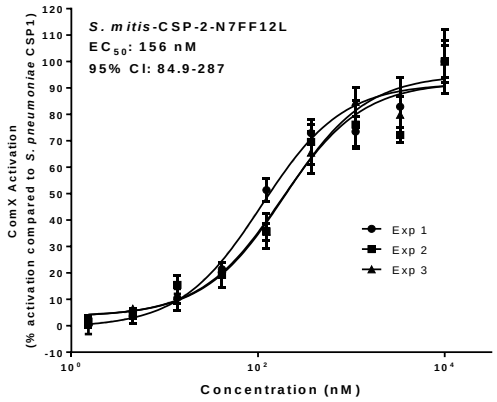

S. mitis-CSP-2-Q4LN7F
EC50: 371 nM
95% CI: 293-469
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

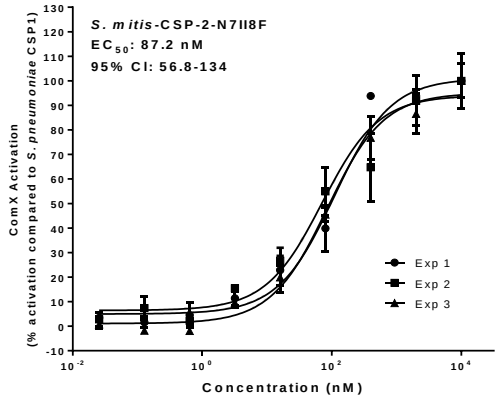

S. mitis-CSP-2-Q4LN7I
EC50: 348 nM
95% CI: 278-436
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

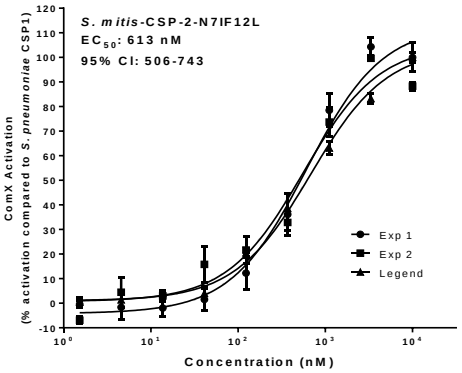

S. mitis-CSP-2-Q4LI8F
EC50: 280 nM
95% CI: 153-513
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

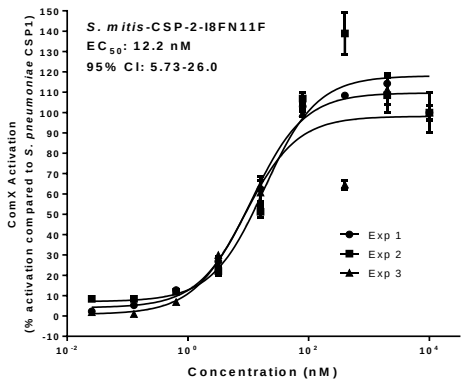

S. mitis-CSP-2-Q4LN11F
EC50: 9.38 nM
95% CI: 8.44-10.4
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

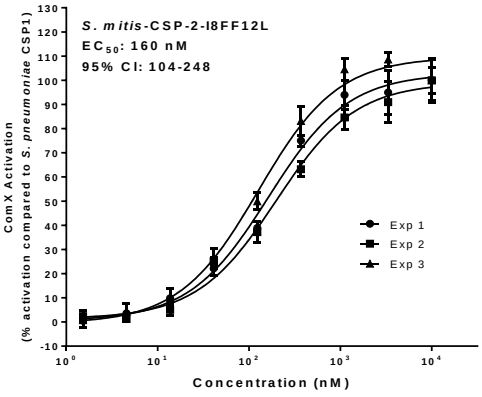

S. mitis-CSP-2-Q4LF12L
EC50: 68.7 nM
95% CI: 50.6-93.2
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

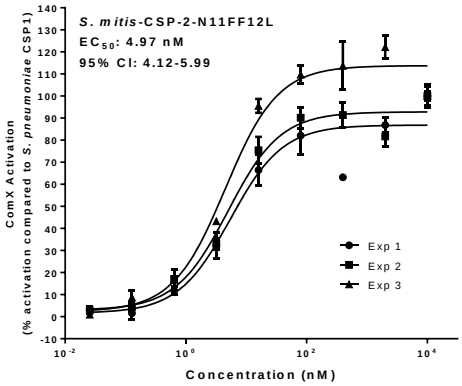

S. mitis-CSP-2-N7FF12L
EC50: 156 nM
95% CI: 84.9-287
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

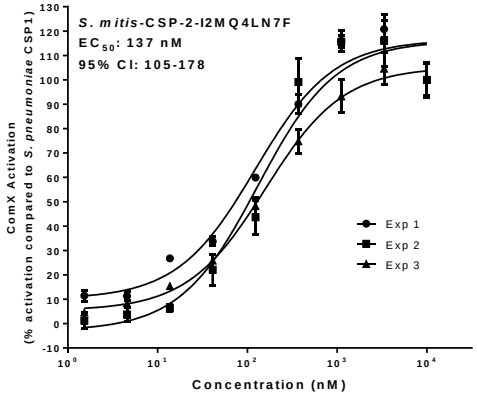

S. mitis-CSP-2-N7I8F
EC50: 87.2 nM
95% CI: 56.8-134
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

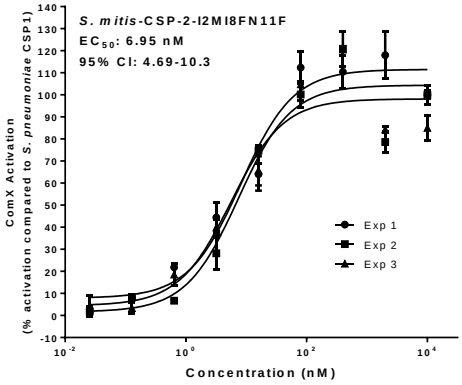

S. mitis-CSP-2-N7IF12L
EC50: 613 nM
95% CI: 506-743
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Legend

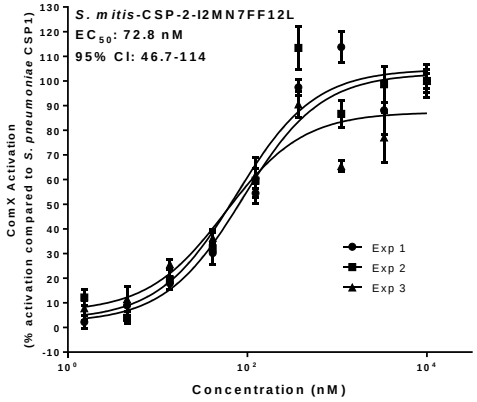

S. mitis-CSP-2-I8FN11F
EC50: 12.2 nM
95% CI: 5.73-26.0
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

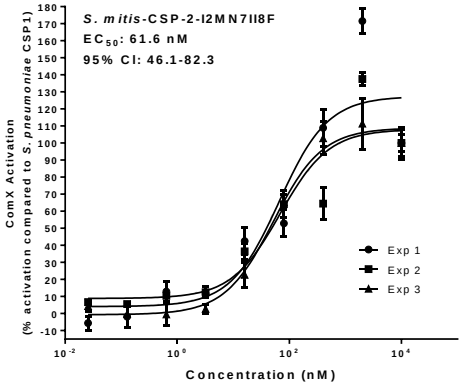

S. mitis-CSP-2-I8FF12L
EC50: 160 nM
95% CI: 104-248
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

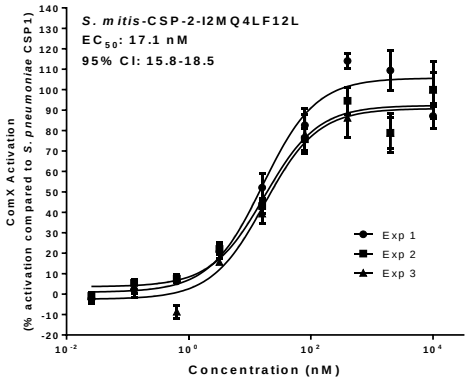

S. mitis-CSP-2-N11FF12L
EC50: 4.97 nM
95% CI: 4.12-5.99
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

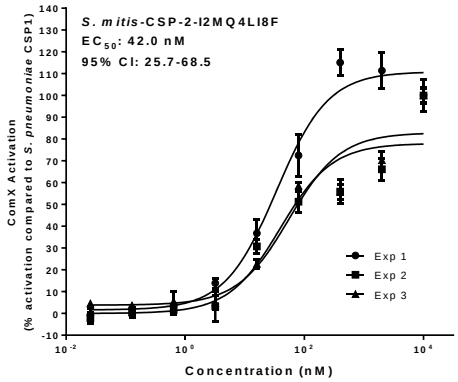

S. mitis-CSP-2-I2MQ4LN7F
EC50: 137 nM
95% CI: 105-178
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

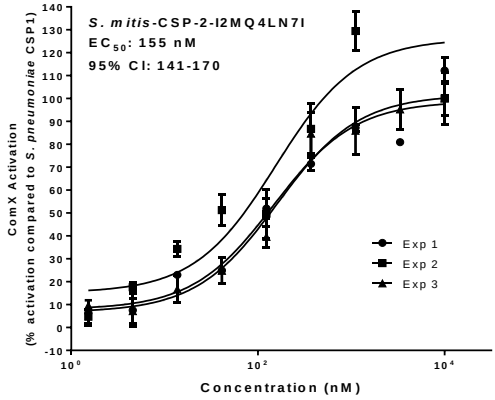

S. mitis-CSP-2-I2MI8FN11F
EC50: 6.95 nM
95% CI: 4.69-10.3
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

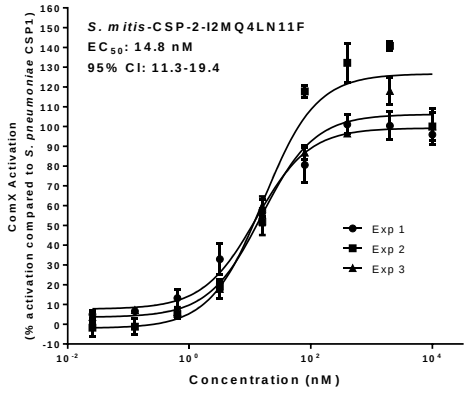

S. mitis-CSP-2-I2MN7FF12L
EC50: 72.8 nM
95% CI: 46.7-114
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

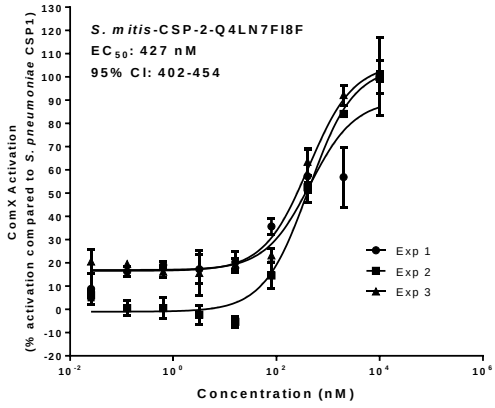

S. mitis-CSP-2-I2MN7II8F
EC50: 61.6 nM
95% CI: 46.1-82.3
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

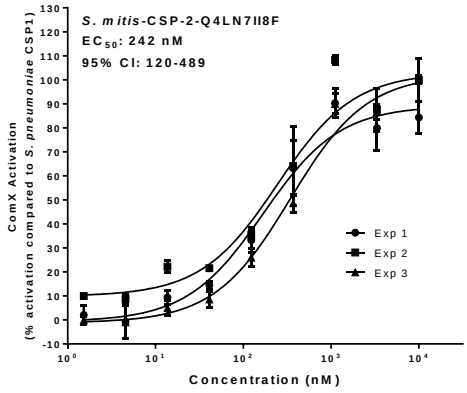

S. mitis-CSP-2-I2MQ4LF12L
EC50: 17.1 nM
95% CI: 15.8-18.5
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

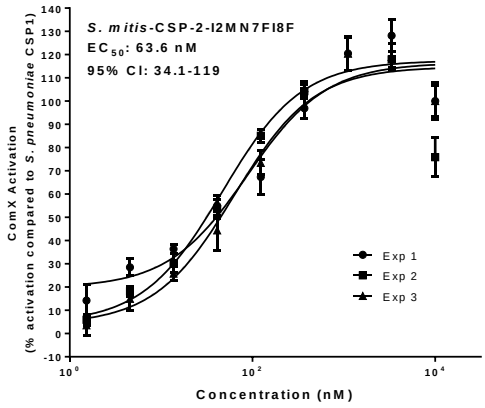

S. mitis-CSP-2-I2MQ4LI8F
EC50: 42.0 nM
95% CI: 25.7-68.5
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

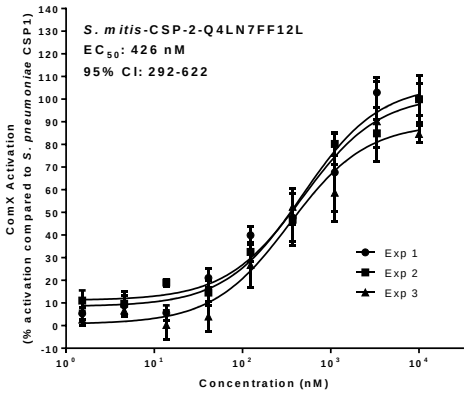

S. mitis-CSP-2-I2MQ4LN7I
EC50: 155 nM
95% CI: 141-170
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

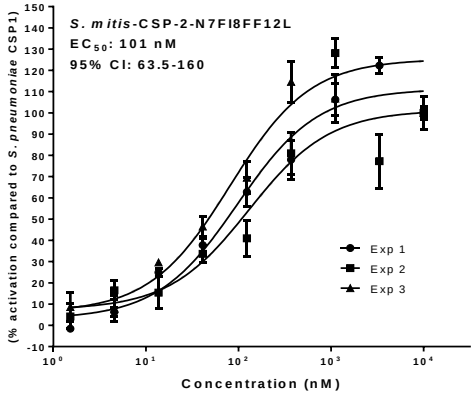

S. mitis-CSP-2-I2MQ4LN11F
EC50: 14.8 nM
95% CI: 11.3-19.4
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

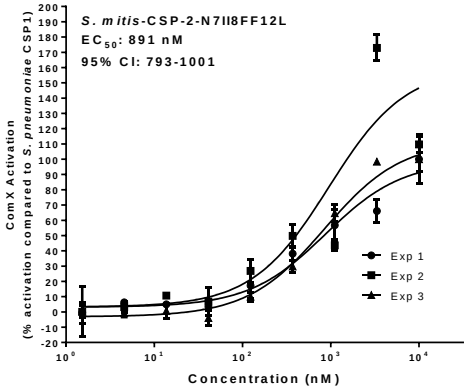

S. mitis-CSP-2-Q4LN7FI8F
EC50: 427 nM
95% CI: 402-454
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

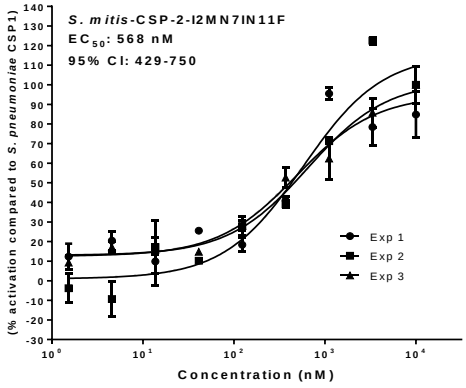

S. mitis-CSP-2-Q4LN7II8F
EC50: 242 nM
95% CI: 120-489
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3


S. mitis-CSP-2-I2MN7FI8F
EC50: 63.6 nM
95% CI: 34.1-119
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)
Exp 1
Exp 2
Exp 3

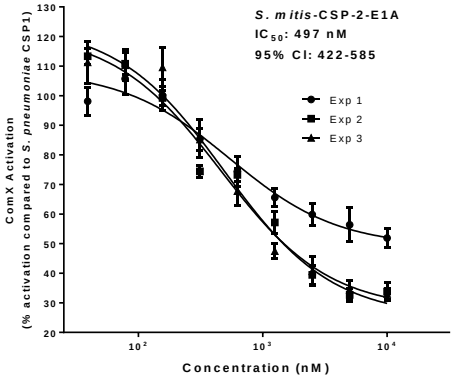


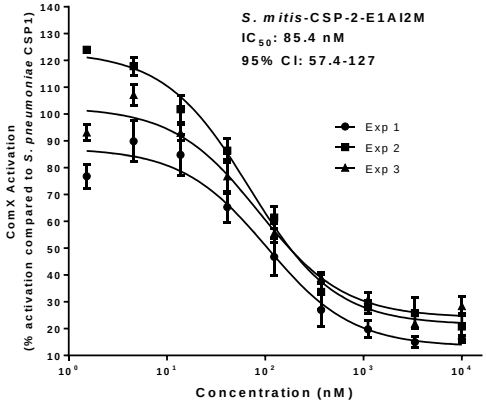


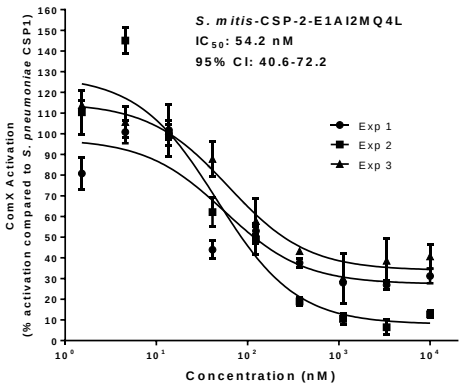


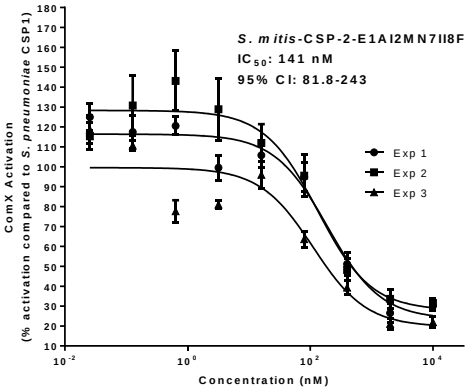


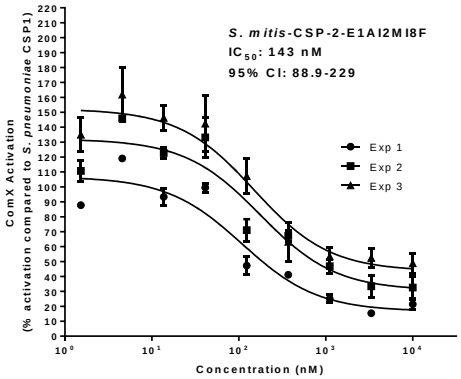
Inhibition dose response curves

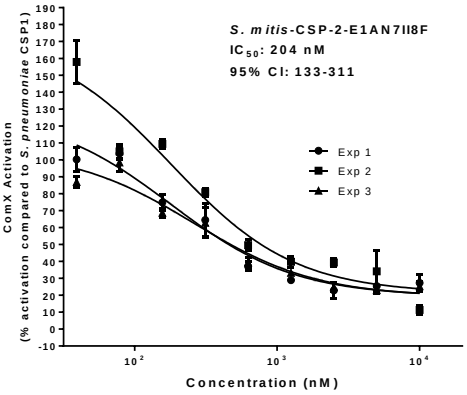


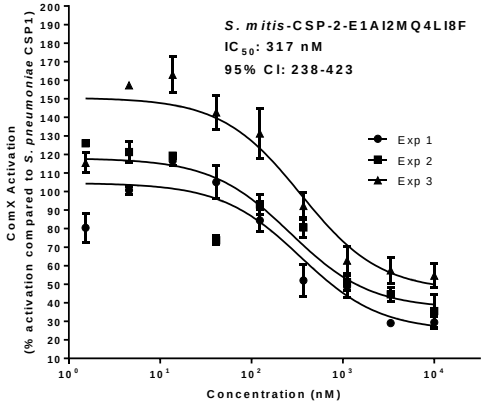


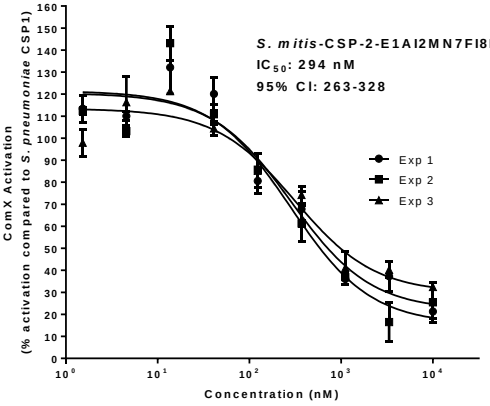

S. mitis-CSP-2-E1AI2MQ4L
IC50: 54.2 nM
95% CI: 40.6-72.2
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)


S. mitis-CSP-2-E1AI2MN7II8F
IC50: 141 nM
95% CI: 81.8-243
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)


S. mitis-CSP-2-E1AI2MI8F
IC50: 143 nM
95% CI: 88.9-229
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)

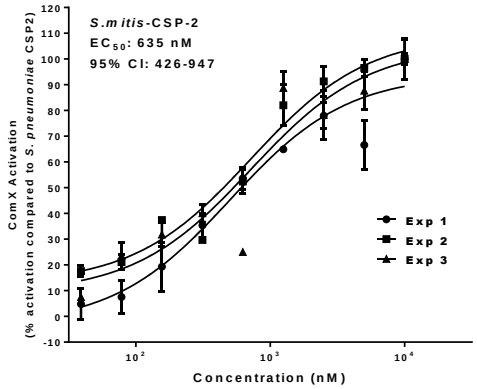

S. mitis-CSP-2-E1AN7II8F
IC50: 204 nM
95% CI: 133-311
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)

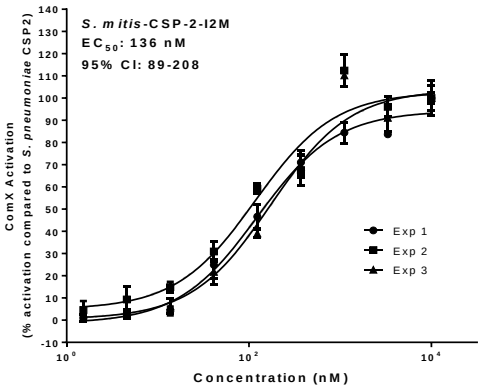

S. mitis-CSP-2-E1AI2MQ4LI8F
IC50: 317 nM
95% CI: 238-423
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)

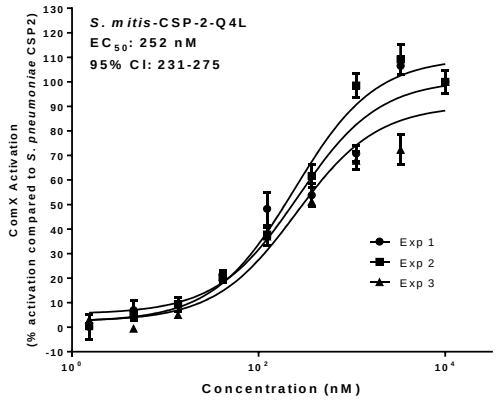

S. mitis-CSP-2-E1AI2MN7FI8F
IC50: 294 nM
95% CI: 263-328
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to S. pneumoniae CSP1)
Concentration (nM)

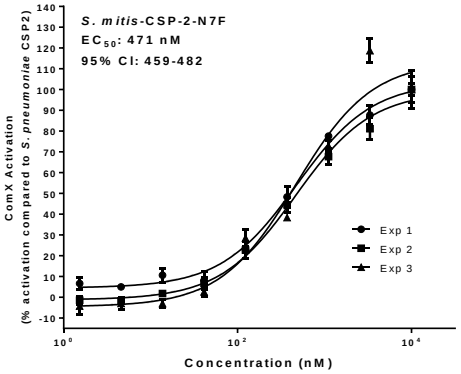
S. pneumoniae TIGR4 pcomX::lacZ (ComD2)

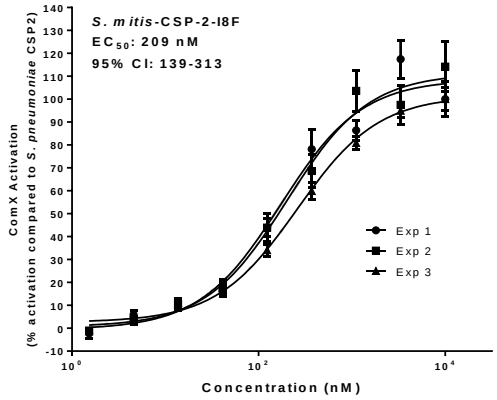
Activation dose response curves

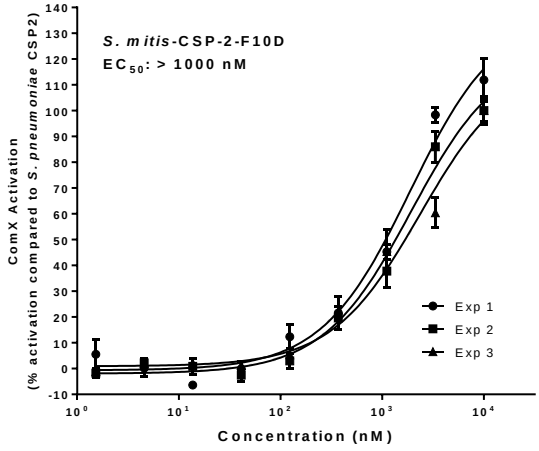


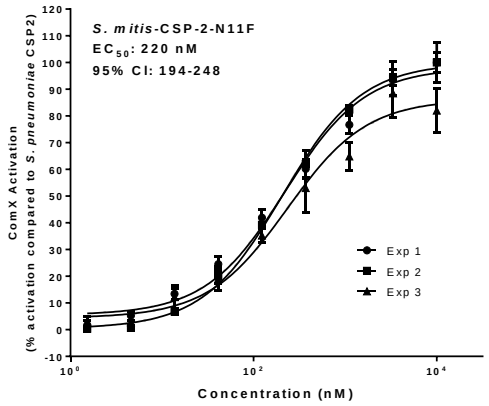


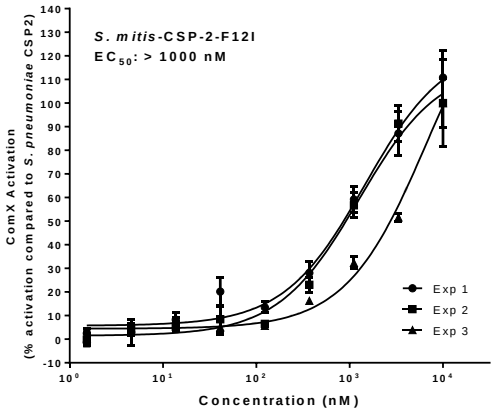


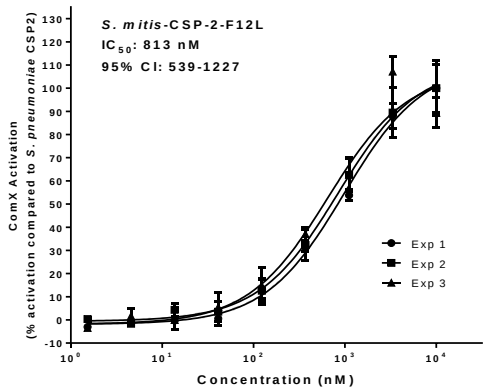


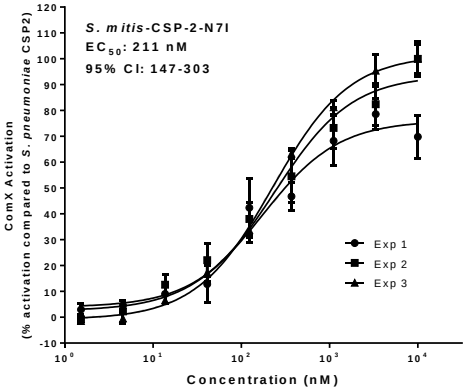


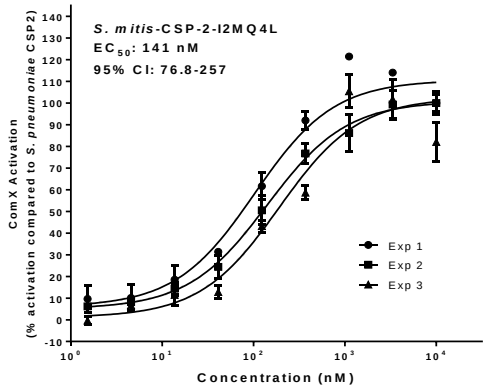


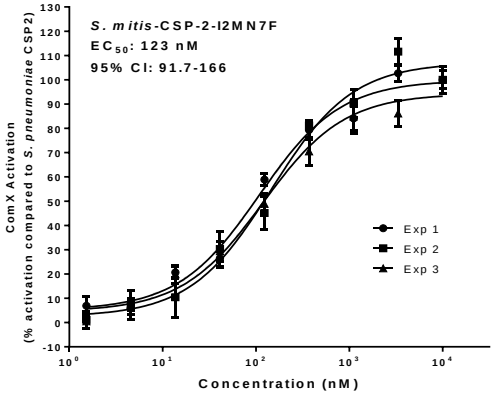

S. mitis-CSP-2-N11F
EC50: 220 nM
95% CI: 194-248
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

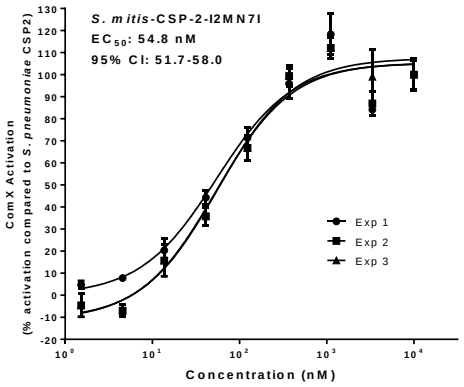

S. mitis-CSP-2-F12I
EC50: > 1000 nM
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

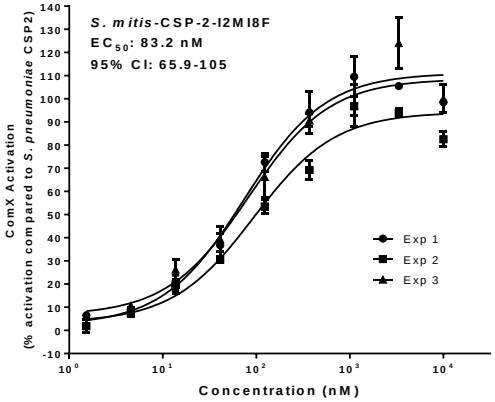

S. mitis-CSP-2-F12L
IC50: 813 nM
95% CI: 539-1227
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

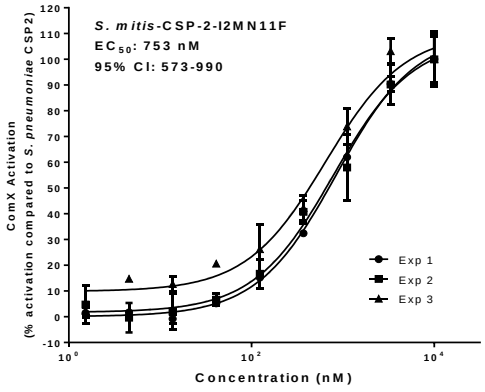

S. mitis-CSP-2-N7I
EC50: 211 nM
95% CI: 147-303
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

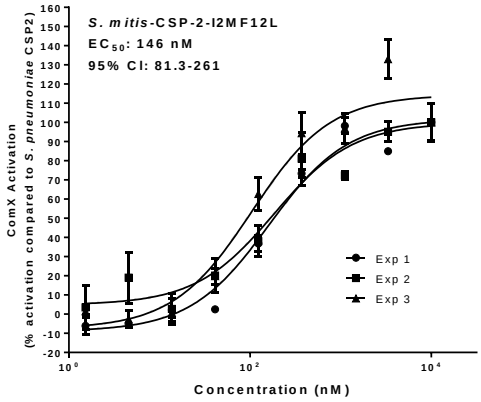

S. mitis-CSP-2-I2MQ4L
EC50: 141 nM
95% CI: 76.8-257
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

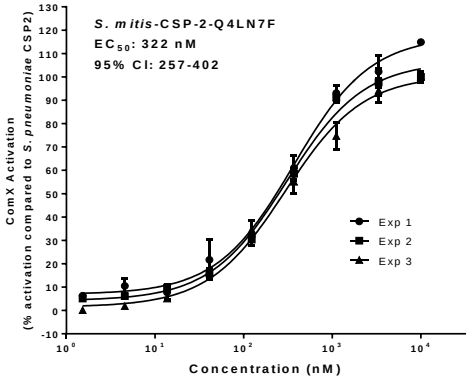

S. mitis-CSP-2-I2MN7F
EC50: 123 nM
95% CI: 91.7-166
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

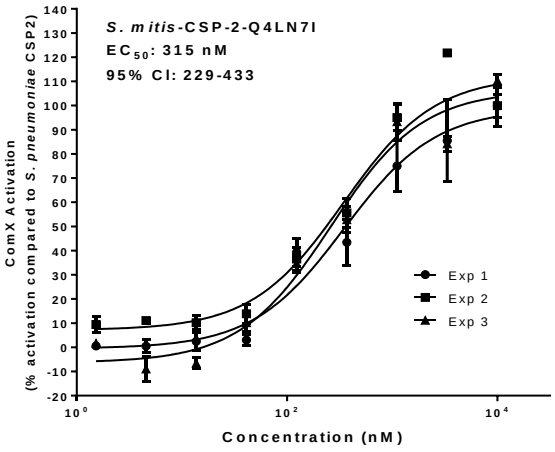

S. mitis-CSP-2-I2MN7I
EC50: 54.8 nM
95% CI: 51.7-58.0
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

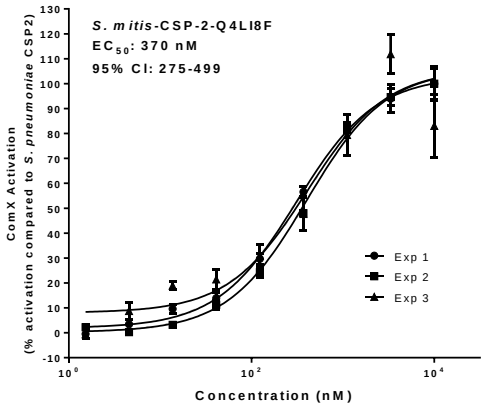

S. mitis-CSP-2-I2MI8F
EC50: 83.2 nM
95% CI: 65.9-105
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

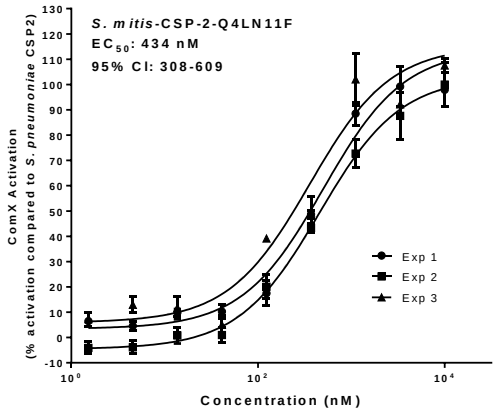

S. mitis-CSP-2-I2MN11F
EC50: 753 nM
95% CI: 573-990
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

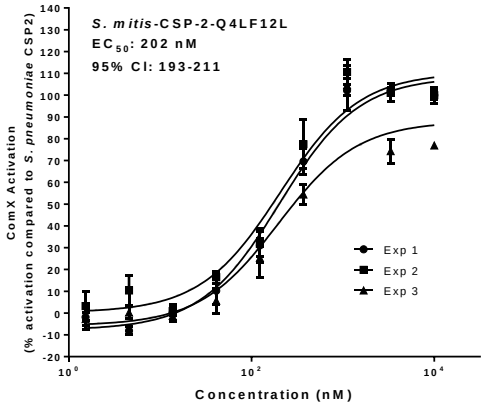

S. mitis-CSP-2-I2MF12L
EC50: 146 nM
95% CI: 81.3-261
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

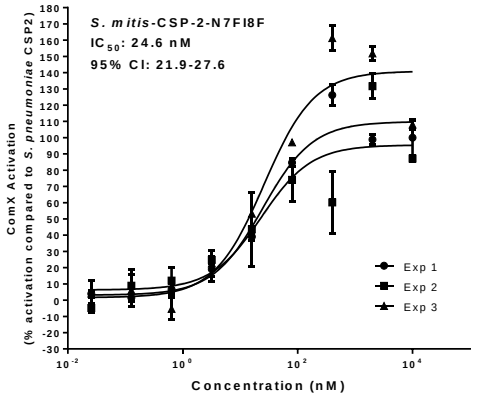

S. mitis-CSP-2-Q4LN7F
EC50: 322 nM
95% CI: 257-402
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

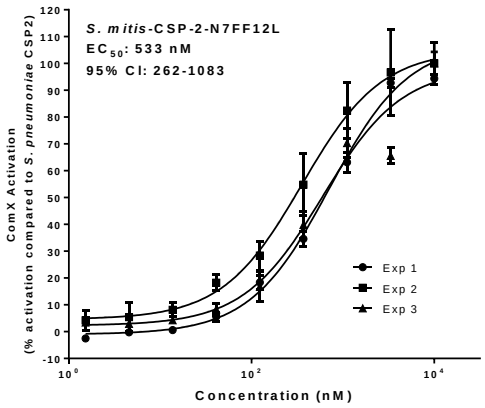

S. mitis-CSP-2-Q4LN7I
EC50: 315 nM
95% CI: 229-433
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

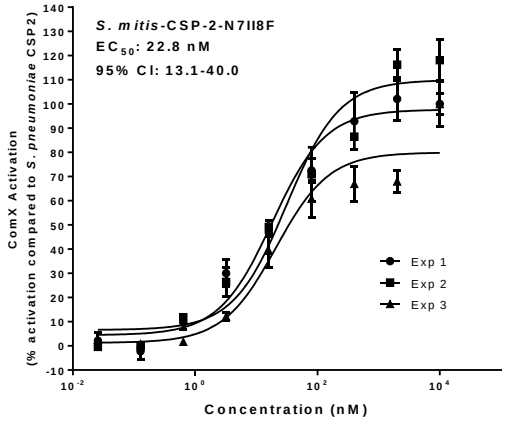

S. mitis-CSP-2-Q4LI8F
EC50: 370 nM
95% CI: 275-499
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

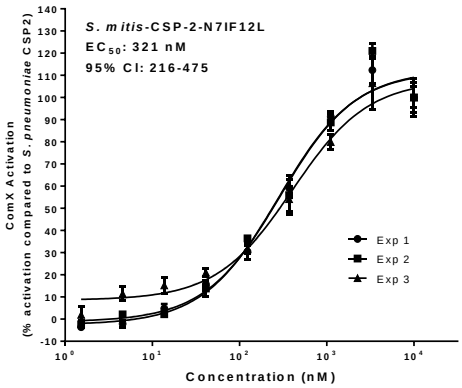

S. mitis-CSP-2-Q4LN11F
EC50: 434 nM
95% CI: 308-609
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

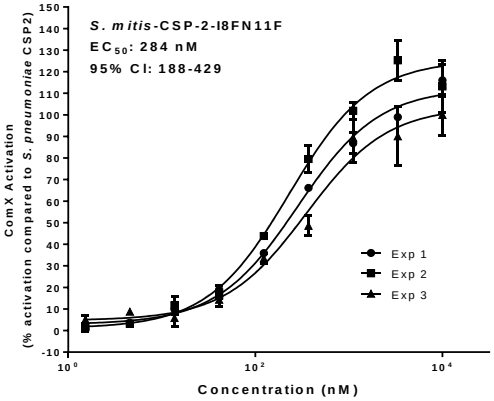

S. mitis-CSP-2-Q4LF12L
EC50: 202 nM
95% CI: 193-211
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

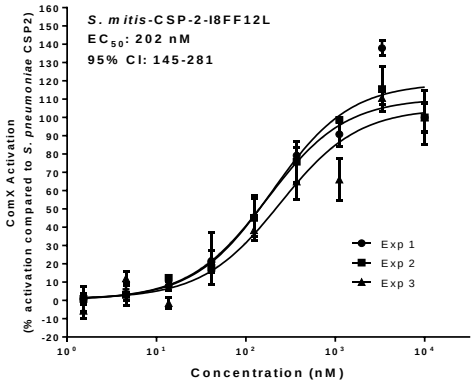

S. mitis-CSP-2-N7FI8F
IC50: 24.6 nM
95% CI: 21.9-27.6
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

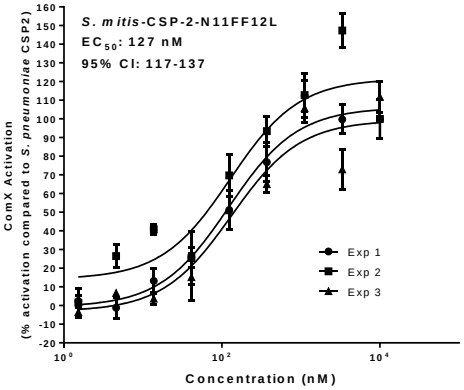

S. mitis-CSP-2-N7FF12L
EC50: 533 nM
95% CI: 262-1083
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

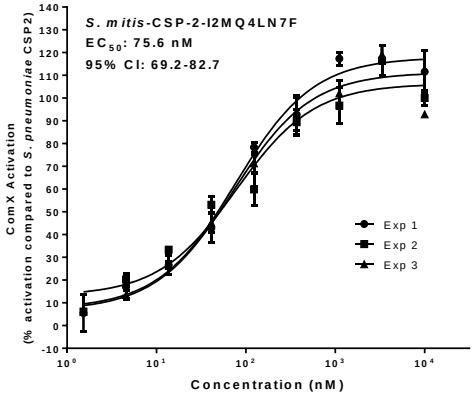

S. mitis-CSP-2-N7II8F
EC50: 22.8 nM
95% CI: 13.1-40.0
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

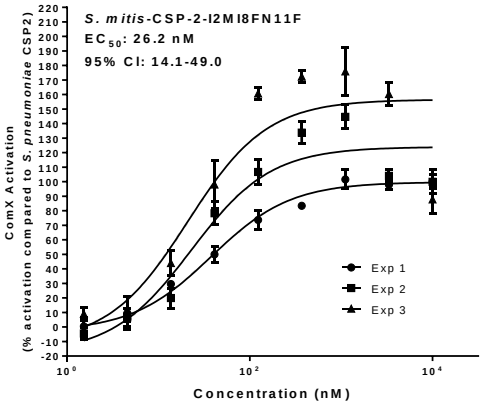

S. mitis-CSP-2-N7IF12L
EC50: 321 nM
95% CI: 216-475
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

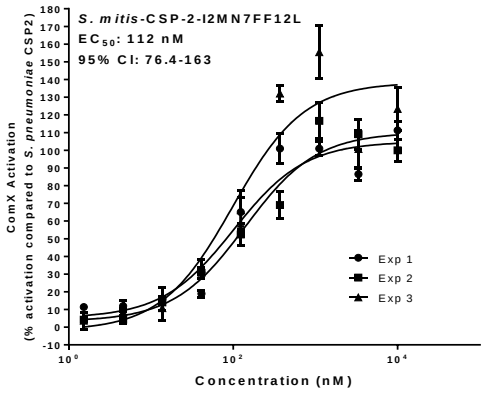

S. mitis-CSP-2-I8FN11F
EC50: 284 nM
95% CI: 188-429
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

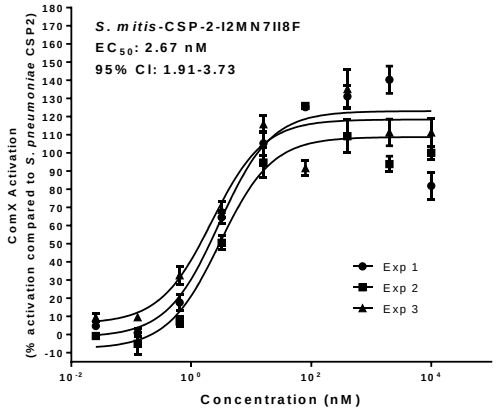

S. mitis-CSP-2-I8FF12L
EC50: 202 nM
95% CI: 145-281
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

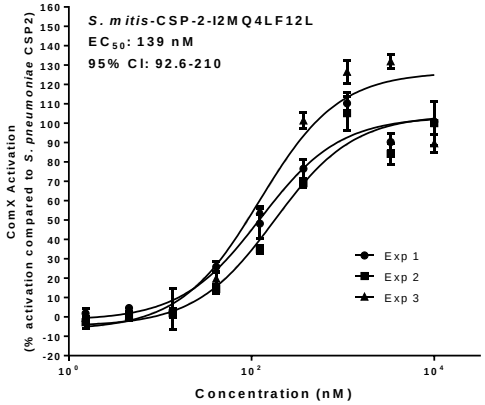

S. mitis-CSP-2-N11FF12L
EC50: 127 nM
95% CI: 117-137
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

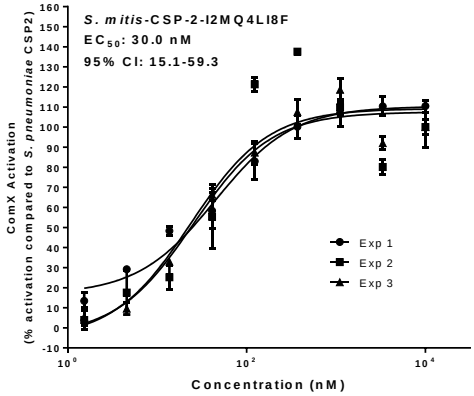

S. mitis-CSP-2-I2MQ4LN7F
EC50: 75.6 nM
95% CI: 69.2-82.7
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

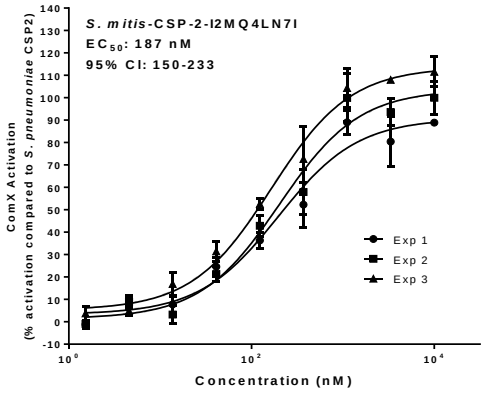

S. mitis-CSP-2-I2MI8FN11F
EC50: 26.2 nM
95% CI: 14.1-49.0
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

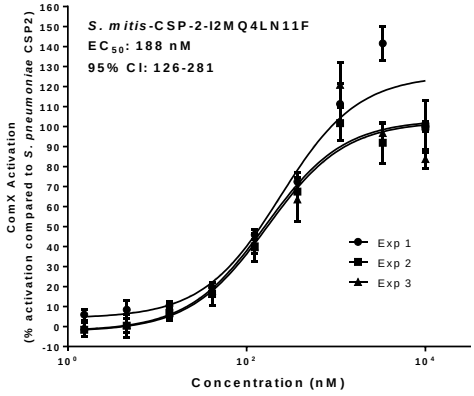

S. mitis-CSP-2-I2MN7FF12L
EC50: 112 nM
95% CI: 76.4-163
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

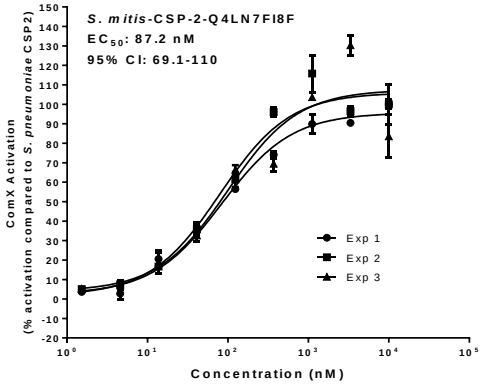

S. mitis-CSP-2-I2MN7II8F
EC50: 2.67 nM
95% CI: 1.91-3.73
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

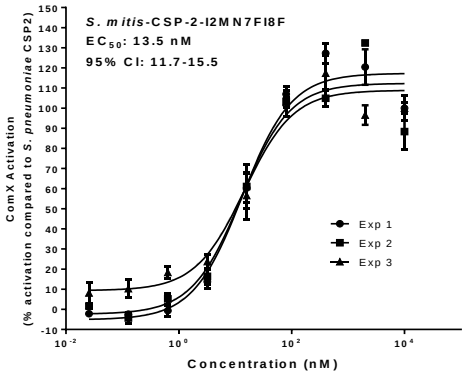

S. mitis-CSP-2-I2MQ4LF12L
EC50: 139 nM
95% CI: 92.6-210
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

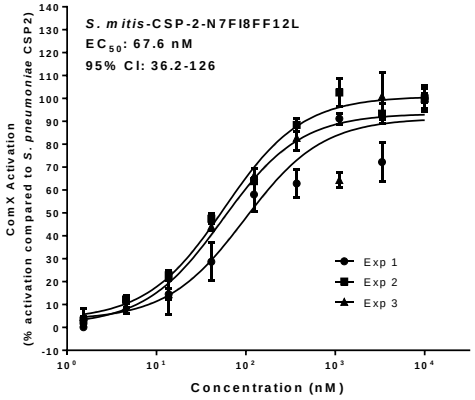

S. mitis-CSP-2-I2MQ4LI8F
EC50: 30.0 nM
95% CI: 15.1-59.3
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3


S. mitis-CSP-2-I2MQ4LN7I
EC50: 187 nM
95% CI: 150-233
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

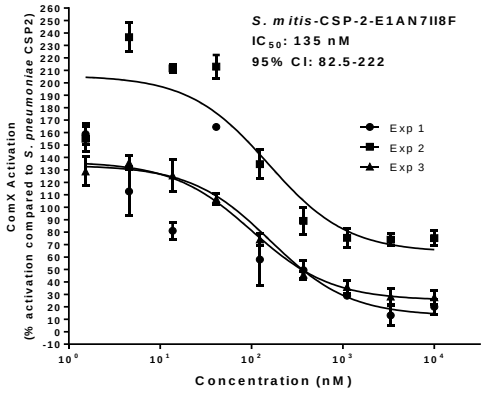

S. mitis-CSP-2-I2MQ4LN11F
EC50: 188 nM
95% CI: 126-281
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)
Exp 1
Exp 2
Exp 3

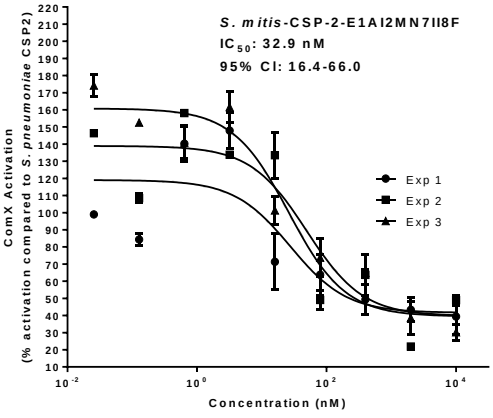


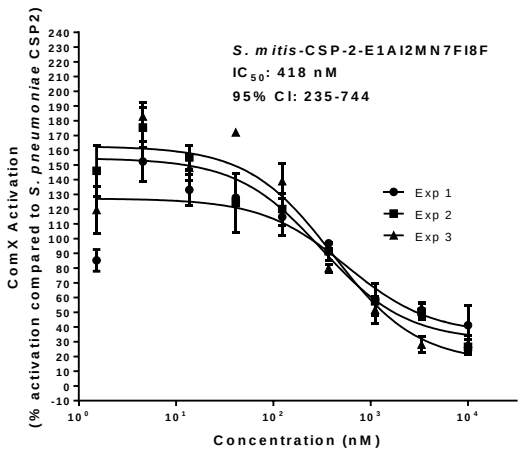




Inhibition dose response curves






S. mitis-CSP-2-E1AI2MN7FI8F
IC50: 418 nM
95% CI: 235-744
Exp 1
Exp 2
Exp 3
ComX Activation
(% activation compared to S. pneumoniae CSP2)
Concentration (nM)

Appendix 4: Developing Multi-Species Quorum Sensing Modulators Based on the *Streptococcus mitis* Competence-Stimulating Peptide[a]

[a]Submitted

Developing Multi-Species Quorum Sensing Modulators Based on the *Streptococcus mitis* Competence-Stimulating Peptide

Supporting Information

Tahmina A. Milly, Clay P. Renshaw, and Yftah Tal-Gan*

Department of Chemistry, University of Nevada, Reno, 1664 North Virginia Street, Reno, Nevada, 89557, United States

* To whom correspondence should be addressed. ytalgan@unr.edu

Full Experimental Methods

Chemical Reagents and Instrumentation:

All chemical reagents and solvents were purchased from Chem-Impex or Sigma-Aldrich and used without further purification. Water (18 MΩ) was purified using a Thermo Scientific Smart2- Pure Pro UV/UF 16 LPH water purification system. Solid-Phase resin was purchased from Chem-Impex. Reversed-phase high-performance liquid chromatography (RP-HPLC) was performed using a Shimadzu system equipped with a CBM-20A communications bus module, two LC-20AT pumps, an SIL-20A auto sampler, an SPD-20A UV/VIS detector, a CTO-20A column oven, and an FRC-10A fraction collector. All RP-HPLC solvents (18 MΩ water and HPLC-grade acetonitrile (ACN)) contained 0.1% trifluoroacetic acid (TFA). Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) data were obtained on a Bruker Microflex spectrometer equipped with a 60 Hz nitrogen laser and a reflectron. In positive ion mode, the acceleration voltage on Ion Source 1 was 19.01 kV. Exact mass (EM) data were obtained on an Agilent Technologies 6230 TOF LC/MS spectrometer. The samples were sprayed with a capillary voltage of 3500 V and the electrospray ionization (ESI) source parameters were as follows: gas temperature of 325 °C at a drying gas flow rate of 8 L/min at a pressure of 35 psi.

Solid Phase Peptide Synthesis:

All the *S. mitis*-CSP-2 analogs were synthesized using standard Fluorenyl methoxycarbonyl (Fmoc)-based solid-phase peptide synthesis (SPPS) procedures on preloaded Fmoc-L-Arg (Pbf) Wang resin (0.305 mmol g^{-1}) by using a Liberty 1 automated peptide synthesizer (CEM Corporation). The resin (0.1 g) was first swelled by suspension

in *N*, *N*-dimethylformamide (DMF) for 15 min at room temperature and then drained. Fmoc removal was accomplished with treatment of the resin by 5 mL of 20% piperidine in DMF (90 s, 90 °C) followed by another 5 mL of 20% piperidine in DMF (90 s, 90 °C). The resin was then washed with DMF (3 X 5 mL) after each deprotection cycle. To couple each amino acid, Fmoc-protected amino acids (5 equiv. relative to the overall loading of the resin) were dissolved in DMF (5 mL) and mixed with 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU; 5 equiv.) and diisopropylethylamine (DIPEA; 5 equiv.). All amino acids were coupled for 5 min (30 W, 75 °C). Arginine residues present within the sequence were double coupled. Following completion of the coupling reaction, the resin was drained and washed with DMF (2 X 5 mL). This process was repeated until the desired peptide sequence was obtained.

Peptide cleavage from solid support:

Following coupling of the final residue, the resin was washed three times with 2 mL DCM with manual shaking for 1 min. The resin was washed with 2 mL diethyl ether and dried under nitrogen stream for 3 min and then transferred into a 15 mL falcon tube. A 3 mL solution of 2.5% 18 MΩ water and 2.5% triisopropylsilane (TIPS) in 95% trifluoroacetic acid (TFA) for every 0.1 g of resin was added and the tube was shaken for 3 h at 200 rpm. Following completion of the cleavage reaction, the resin was filtered using a cotton plug in a polypropylene syringe and washed with a small amount of TFA. The filtrate was collected in a 50 mL polypropylene centrifuge tube, and a cooled solution of diethyl ether:hexane (1:1, 45 mL, -20 °C) was added to the tube. To precipitate the crude peptide, the tube was kept in a freezer for 10 min at -20 °C. The pellet of the crude peptide was obtained by centrifugation of the 50 mL tube in a Beckman Coulter Allegra 6 centrifuge

equipped with a GH3.8 rotor at 3000 RPM for 5 min. Following centrifugation, the 1:1 ether:hexane solution was removed, and the solid peptide product was re-dissolved in 1:1 ACN:water, frozen in a dry ice-acetone bath, and then lyophilized for a minimum of 24 h.

Peptide Purification by HPLC:

Crude peptides were purified using RP-HPLC. The crude peptide was dissolved in ACN: H_2O (1:4; volume of ACN in water depends on the solubility of the peptide) and purified in 4 mL portions on a Phenomenex Kinetex 5 μm C18 semi-preparative column (10 mm x 250 mm, 110 Å) with a flow rate of 5 mL min^{-1}; mobile phase A = 18 MΩ water + 0.1% TFA and mobile phase B = ACN + 0.1% TFA. Initially, crude peptides were purified with a linear gradient from 5% to 40% ACN in 46 min. After identifying the relative retention time of these peptides, a second preparatory run was performed using a 10% ACN gradient centered on the ACN concentration where peptide elution was observed, for example a 15% to 25% ACN gradient for a peptide that eluted at 20% ACN during the first run. These conditions were typically sufficient to purify the peptides to ≥ 95%. Fraction purity was determined through analysis on a Phenomenex Kinetex 5 μm analytic C18 column (4.6 mm x 250 mm, 110 Å). Purities were determined by integration of peaks with UV detection at 220 nm. The gradient used for analytical analysis was from 5% to 95% ACN over 24 min. Following purification, peptides were frozen with a dry ice-acetone bath, and then lyophilized for a minimum of 24 h. Before the final masses and yields of purified peptides were determined, peptides were dissolved in 25% acetic acid in up to 1:1 ACN:water to allow removal of any residual TFA. The solution was then frozen and lyophilized for at least 24 h before peptide DMSO stocks were made for bioassays.

Peptide Verification with Mass Spectrometry:

Following purification of crude peptide, peaks were verified to contain the desired peptide mass by MALDI-TOF MS. Samples were prepared using 1 μL α-Cyano-4-hydroxycinammic acid (10 μg) in 1:1 water:ACN as a matrix and 1 μL of the desired peptide fraction. Final verification of the peptide mass was conducted by obtaining their exact masses with a high-resolution ESI-TOF MS (**Tables S-2 to S-5**). The instrument was calibrated before each run and an internal reference mass standard was used.

Isolation of Crude Peptides from Bacterial Supernatants:

Overnight *S. mitis* ATCC culture (200 mL) was centrifuged at 4,500 rpm for 10 min. The supernatants were fractionated and filtered through a sterile 0.22-μm polyether sulfone (PES) filter into sterile 50 ml centrifuge tubes. Ammonium sulfate was added to the supernatants to afford a 40% (wt/vol) concentration and the solutions were mixed by inversion until the ammonium sulfate was completely dissolved. The solution was stored at 4 °C for 1 hour, followed by centrifugation at 4,500 rpm for 15 min. The supernatants were then carefully decanted from the centrifuged tubes and the remaining pellet was dissolved in 10 mL of ddH_2O:ACN (1:1), after which all of the fractions were combined into a single centrifuge tube and freeze-dried. The lyophilized material was then purified by RP-HPLC. Following purification, the mass of the extracted peptide was verified using a high-resolution ESI-TOF MS and through MS/MS analysis.

MS/MS analysis:

In 1% Formic Acid in 1:1 H_2O:Methanol (mass spectrometry grade), either the extracted or synthetic *S. mitis*-CSP-2 peptide was dissolved to a final concentration of ~0.5-1.0 μM. To remove any fine particles, the solutions were filter sterilized using 0.45 μm syringe

filters. For MS and MS/MS analysis, a ThermoFisher™ Orbitrap Fusion™ Tribrid™ Mass Spectrometer was used. From the MS data, the two highest intensity peaks (MH^{+3} and MH^{+4} for the extracted peptide; MH^{+4} & MH^{+5} for the synthetic peptide) were isolated using quadrupole isolation and fragmented using higher energy collisionally activated dissociation (HCD). The Orbitrap™ detector was used at a resolution of 60,000 for both peptides. To determine the HCD energy for optimal peptide fragmentation, an HCD energy scan from 0-50% was performed (data not shown) using 5% intervals. For the final MS/MS fragmentation, the HCD energy was set to 43% (MH^{+3}) and 42% (MH^{+4}) for the extracted peptide and 41% (MH^{+4}) and 24% (MH^{+5}) for the synthetic peptide. All MS and MS/MS data were copied to Microsoft Excel and analyzed using Origin Pro.

Development of *S. mitis* ATCC 49456 reporter strain:

Bacterial gDNA extraction: An isolated fresh single colony of *S. mitis* ATCC 49456 was picked into a sterilized cultural tube containing 5 mL of THY media (pH 7.3) and the culture was incubated in a CO_2 incubator overnight (16 hours). Following incubation, 1.5 mL of pure culture broth was added to a sterile 1.5 mL microcentrifuge tube. Cells were pelleted by centrifugation at 10,000 rpm for 1 min. Following centrifugation, the supernatant was discarded, and 0.5 mL sterile 1x phosphate buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4; pH adjusted to 7.2 - 7.4) was added to the tube containing the bacterial cell pellet. The pellet was then resuspended using 5-sec pulses with a vortex mixer set at 4,000 rpm. The cells were pelleted again under the same centrifugation conditions as above. The supernatant was discarded, and the process was repeated twice more for a total of three times. After the PBS washes, 0.5 mL of sterile distilled water (dH_2O) was added to the bacterial pellet, and it was resuspended by pulse

vortex mixing at 4,000 rpm. Using the same conditions, the cells were pelleted. The supernatant was discarded, and 200 μL of sterile distilled water (dH_2O) was added to the tube. The pellet was again resuspended by vortex mixing at 4,000 rpm, incubated in a standard heat block set at 95 °C for 5 min, and immediately placed in a −80 °C freezer for 10 min. The tubes were then kept at room temperature to thaw completely and gently mixed with a vortex mixer set at 4,000 rpm using 5-sec pulses, followed by centrifugation at 5,000 rpm for 60 sec. The supernatant was then transferred to a new, sterile, 0.6 mL microcentrifuge tube and stored at −20 °C.

***S. mitis* ATCC 49456 *comX* promoter amplification:** The promoter region (911 bp) of the *S. mitis* ATCC 49456 *comX* gene was amplified from extracted bacteria gDNA using the primer pair SmATCCComXfwd and SmATCCComXrev (**Table S-1**) on an Eppendorf Mastercycler Gradient 5331 PCR machine. For *comX* PCR amplification, total reaction volume of 25 μL was prepared containing 12.5 μL Hot Start Taq Master Mix (VWR), 2.5 μL template DNA, 2.5 μL of 1 μM primer SmATCCComXfwd, 2.5 μL of 1 μM primer SmATCCComXrev, and 5 μL nuclease free water. PCR amplification consisted of the following steps: 3 min initial denaturation at 95 °C to activate the polymerase, followed by 35 cycles of 30 sec denaturation at 95 °C, 1 min of annealing at 59 °C, and 1 min of extension at 72 °C, followed by a final extension at 72 °C for 7 min. Following amplification, PCR products were run on a 1% agarose gel to verify the presence of the *comX* PCR amplicon containing ends homologous to the restriction sites for *BamHI* and *NheI* (911 bp). Once verified, PCR products were purified using a E.Z.N.A Cycle Pure Kit (Omega Bio-Tek).

Table S-1. Primers used for the design of the *S. mitis* ATCC 49456 luciferase reporter

Primer	**Sequence**
SmATCCComXfwd	aaagctagcAGCTGCTTTAGTCGCTGCTC[a]
SmATCCComXrev	aaaggatccCAATCCCCTGGACTTCTTCA[b]

[a]*BamHI* restriction site underlined. [b]*NheI* restriction site underlined

Restriction digestion: The plasmid pFW5-luc ($Spec^R$) was extracted and purified from *Escherichia coli* using a GenElute Plasmid Miniprep Kit (Sigma). Both pFW5-luc and the PCR amplified *comX* promoter amplicon were then restriction digested using 1 μg DNA, 5 μL 10X CutSmart Buffer (NEB), 1 μL *NheI-HF* (NEB), 1 μL *BamHI-HF* (NEB) and nuclease free water to bring the total reaction volume up to 50 μL. The restriction digestion was carried out for 15 min at 37 °C, followed by heat inactivation of the restriction enzymes at 80 °C for 20 min. Following restriction digestion, digested products were run on a 1% agarose gel to verify the presence of bands with a different size than undigested plasmid or PCR amplified *comX* promoter, respectively. After the verification of the correct amplicon, digested product was purified using a E.Z.N.A Cycle Pure Kit (Omega Bio-Tek).

Ligation and cloning: Ligation of purified restriction digested pFW5-luc ($Spec^R$) and PCR amplified *comX* promoter was carried out in a microcentrifuge tube placed on ice. Ligation was carried out at an insert to vector ratio of 3:1. Ligation mixture was prepared with the addition of 2 μL of 10X T4 DNA ligase reaction buffer (NEB), 1 μL of T4 DNA ligase (NEB), and nuclease free water to bring the total reaction volume up to 20 μL. Then the

ligation mixture was gently mixed by pipetting up and down, and the ligation reaction of the insert and vector was performed at 16 ℃ overnight. After the completion of ligation reaction, heat inactivation of the ligated construct was performed at 65 °C for 10 min. Following that, the ligated construct was transformed into competent *E. coli*. Competent *E. coli* were removed from a -80 °C freezer, thawed at room temperature, and immediately placed on ice for 10 min. Ligated construct to be transformed was added at a concentration of 10 ng per 50 μL of competent cells, with the total volume of added DNA not exceeding 5% that of the competent cells. Tubes containing the competent *E. coli* and construct DNA were gently inverted several times to mix contents to homogeneity, then the tubes were placed on ice for 30 min. Next, the tubes were placed in a 42 °C heat bath for 90 sec, and immediately transferred to an ice bath for 2 min. Following that, 400 μL of SOC media (2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, and 20 mM glucose) was added to the tube, which was then transferred to a 37 °C shaking incubator set at 200 rpm for 45 min. The entire content of the tube was then transferred to a LB agar plate containing 100 μg/mL spectinomycin. The plate was then incubated overnight at 37 °C and was checked for the presence of positive transformants the following day.

Construct transformation: A positive *E. coli* transformant was inoculated into 5 mL of LB media containing 100 μg/mL spectinomycin, which was then grown overnight at 37 °C with shaking at 200 rpm. Purified plasmid construct was extracted from the *E. coli* overnight culture using a GenElute Plasmid Miniprep Kit (Sigma). A small amount of the extracted plasmid was subjected to restriction digestion by *BamHI-HF* and *NheI-HF* using the above conditions and the presence of the correct vector and insert was verified.

Successful constructs were identified as containing two bands corresponding to the approximate size of the vector and insert when run on a 1% agarose gel. Following that, a single colony of *S. mitis* ATCC 49456 was inoculated into 5 mL THY media and grown overnight (16 h) at 37 °C with 5% CO_2. Then the overnight culture was diluted 1:100 in fresh THY media and was incubated at 37 °C in a 5% CO_2-supplemented atmosphere for 4-5 h until the culture reached an absorbance at 600 nm (optical density at 600 nm [OD_{600}]) of 0.30. Following incubation, 100 µL of culture was transferred to 900 µL of THY media and then synthetic *S. mitis*-CSP-2 at a final concentration of 200 nM and plasmid construct at a final concentration of 1 µg/mL were added to the diluted culture, and the bacteria were subjected to another 3 to 4 h growth at 37 °C. After 3 to 4 h of incubation at 37 °C, 200 µL of the culture was plated on THY agar containing 200 µg/mL spectinomycin and incubated at 37 °C with 5% CO_2 for 24 to 48 h to identify positive transformants. A single colony was then picked and grown for 16 h in 5 mL TSB media containing 200 µg/mL spectinomycin at 37 °C with 5% CO_2. Following overnight growth, the *S. mitis* culture was diluted 1:10 in fresh TSB media, and the presence of the reporter plasmid was validated by both sequencing of the *comX* promoter region and the observation of luminescence following treatment with exogenous *S. mitis*-CSP-2 (10,000 nM) and D-luciferin.

Biofilm Formation Assay:

A single colony of *S. mitis* ATCC 49456 was grown for 16 hours in 5 mL THY media (pH 7.3) at 37 °C with 5% CO_2. Following incubation, *S. mitis* ATCC 49456 was diluted 1:10 in fresh THY media containing 1% D-glucose, and 196 µL of culture was added in triplicate to a 96-well microtiter plate. Two µL of *S. mitis*-CSP-2 (at a final concentration of 200 nM), the lead activator, *S. mitis*-CSP-2-n11 (at a final concentration of 200 nM), or

S. mitis-CSP-2 (at a final concentration of 200 nM) together with the lead inhibitor, *S. mitis*-CSP-2-E1Af10r16 (at a final concentration of 500 nM) in 2 μL of DMSO were added to the culture. For background subtraction, a set of wells containing only 196 μL of THY media containing 1% D-glucose and 4 μL of DMSO (no bacteria) was included. The plate was then incubated for 24 hours at 37 °C with 5% CO_2. Following 24 h incubation, the optical density at 600 nm (OD_{600}) was recorded. The contents of all wells were carefully decanted by shaking the plate gently over a glass basin. Experimental wells were then gently washed with 100 μL 1x PBS. The 96 well microtiter plate was then incubated at 55 °C for 2.5 h to heat fix bacterial biofilms to the bottom of the well. Two hundred μL of a 0.05% crystal violet solution was then added to each well, and the solution was kept at room temperature for 5 min. The wells were then carefully decanted and washed with 200 μL distilled water, and a total of 200 μL of a 30% (v/v) acetic acid in distilled water was added to the wells. The plate was placed in shaker at 150 rpm for 15 min at 37 °C, and experimental wells were then further diluted 1:5 in distilled water. The absorbance at 595 nm (OD_{595}) was then measured for each well. Each OD_{595} value was divided by its corresponding OD_{600} values. Experiments were performed in triplicate on three separate days. Data are presented as the percent biofilm formation relative to wild type untreated with exogenous *S. mitis*-CSP-2. Results are expressed as the mean +/- the standard deviation of three independent experiments.

Hemolysis Assay:

A single colony of *S. mitis* ATCC 49456 was grown for 16 h at 37 °C with 5% CO_2 in 5 mL THY media (pH 7.3). Overnight cultures were then diluted 1:10 in fresh THY media.

Samples were prepared in clear bottom 96-well microtiter plates. Experimental samples were prepared by adding 196 μL of the 1:10 diluted culture and 2 μL of a 1 mM *S. mitis*-CSP-2 and the lead *S. mitis*-CSP-2 analogs (lead activator, *S. mitis*-CSP-2-n11, and lead inhibitor, *S. mitis*-CSP-2-E1Af10r16), or 196 μL of fresh THY media and 2 μL of a 1 mM *S. mitis*-CSP-2/ *S. mitis*-CSP-2 analog stock. Two μL of DMSO was added to each well. A positive control was prepared by adding 2 μL of a 1% Triton X solution and 2 μL of DMSO in 196 μL THY media, and negative controls were prepared by adding 4 μL DMSO in either 196 μL 1:10 diluted culture or 196 μL THY media. The 96-well microtiter plate was then incubated for 2 h at 37 °C with 5% CO_2, after which the optical density at 600 nm (OD_{600}) was recorded and hemolysis was assessed. A 1 mL portion of defibrinated rabbit red blood cells (VWR International) was aliquoted into a sterile 1.5 mL microcentrifuge tube, and centrifuged at 2,000 rpm for 2 min. Following centrifugation, the top layer was aliquoted off, and red blood cells were gently washed with sterile1 mL 1X PBS. This process was repeated for a total of three times, until following centrifugation the top layer was mostly clear. Washed red blood cells were resuspended in sterile 1 mL 1X PBS solution, and a 15 μL aliquot was added to each well of the 96-well microtiter plate. The plate was then incubated for 30 min at 37 °C. Following incubation, the plate was centrifuged for 4 min at 4 °C at 1600 rpm. Experimental wells were then diluted 1:5 in distilled water to prevent saturation of the detector by the positive control. The absorbance at 420 nm (OD_{420}) was then measured for each well. Experiments were performed in triplicate on three separate days. Data are presented as the percent hemolysis relative to the

.01% Triton X positive control. Results are expressed as the mean +/- the standard deviation of three independent experiments.

Circular Dichroism (CD) Spectroscopy:

CD spectra were recorded using an Aviv Biomedical CD spectrometer (model 202-01). All the measurements were performed with a peptide concentration of 200 μM in 1X PBS buffer (137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4; pH was adjusted to 7.4) with 0% or 20% TFE. Measurements were performed at 25 °C using a quartz cuvette (Starna Cells) with a pathlength of 0.1 cm. Samples were scanned once at 3 nm min^{-1} with a bandwidth of 1 nm and a response time of 20 sec over a wavelength range (195 to 260 nm). Single scans were acquired and corrected for their respective solvent concentrations and then converted to mean residue ellipticity (MRE) values using the following equation:

$$\mathrm{MRE} = \left(\frac{\theta}{10 \times c \times l}\right)/n$$

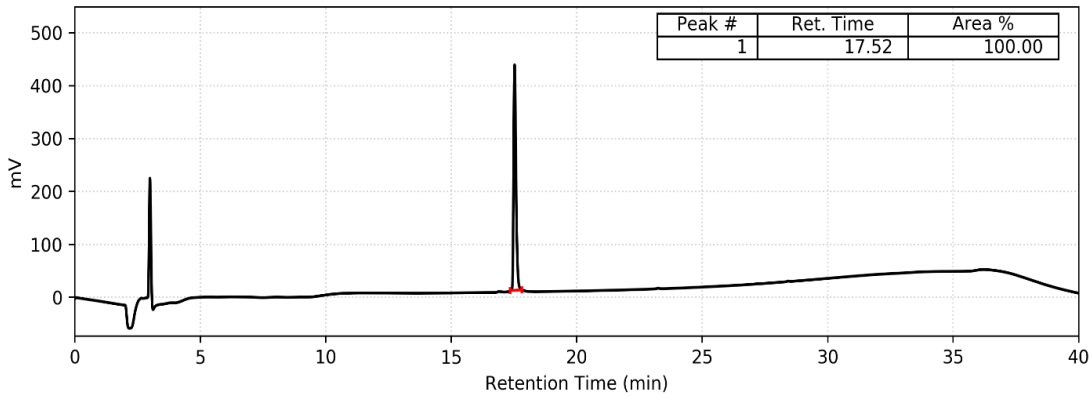
θ is the observed ellipticity in millidegrees, c is the molar peptide concentration, l is the pathlength in centimeters, and n is the number of residues in the peptide sequence. Percent helicity (f_H) was calculated for all peptide analogs using the following equation:

$$f_H = \frac{[\theta_{222}]}{[\theta_\infty]_{222}\left(1-\frac{x}{n}\right)}$$

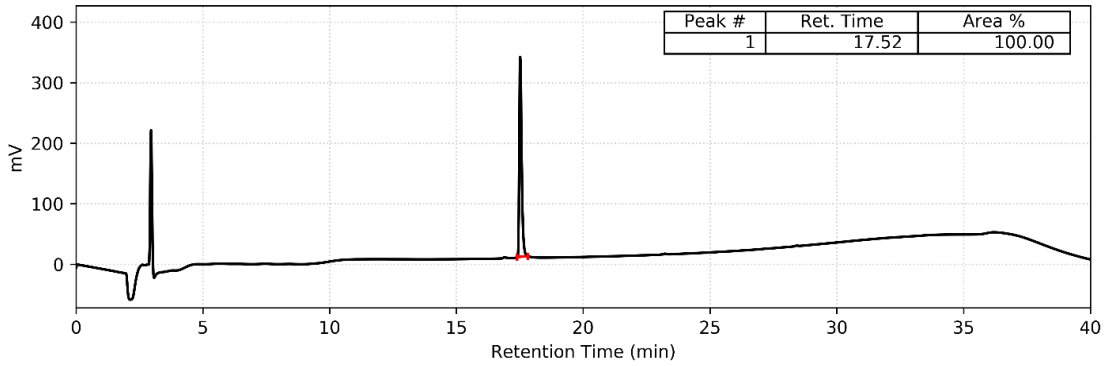
$[\theta_{222}]$ is the mean residue ellipticity of the peptide at 222 nm, $[\theta_\infty]_{222}$ is the assumed mean residue ellipticity for a peptide with 100% helicity (−44,000 deg cm^2 $dmol^{-1}$), n is the number of residues within the tested peptide sequence, and x is an empirical correction for end effects.

HPLC Traces for *S. mitis*-CSP-2 analogs

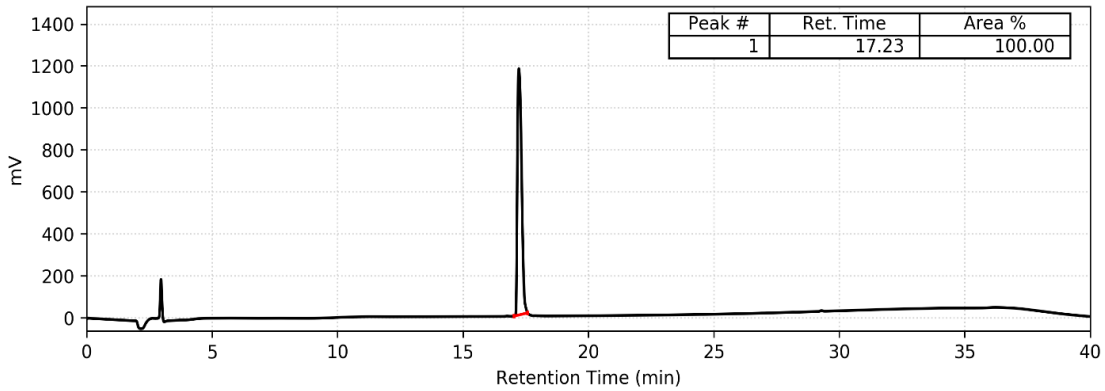
S. mitis-CSP-2 (synthesized)



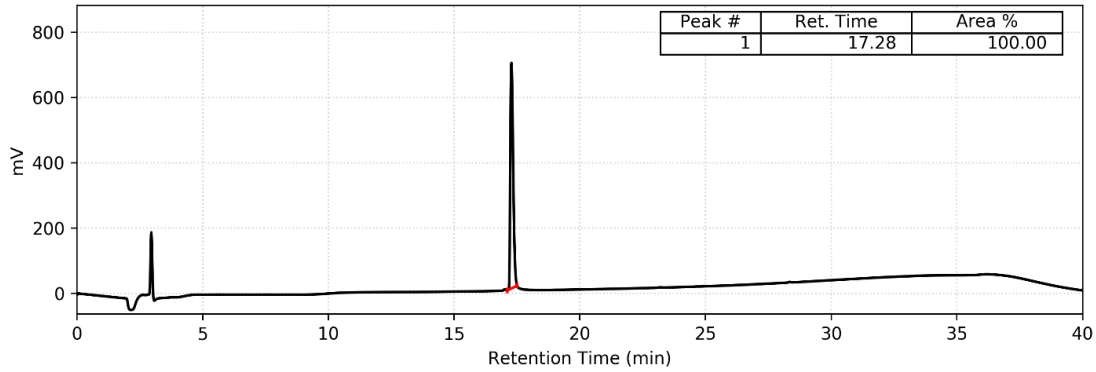
S. mitis-CSP-2 (extracted)



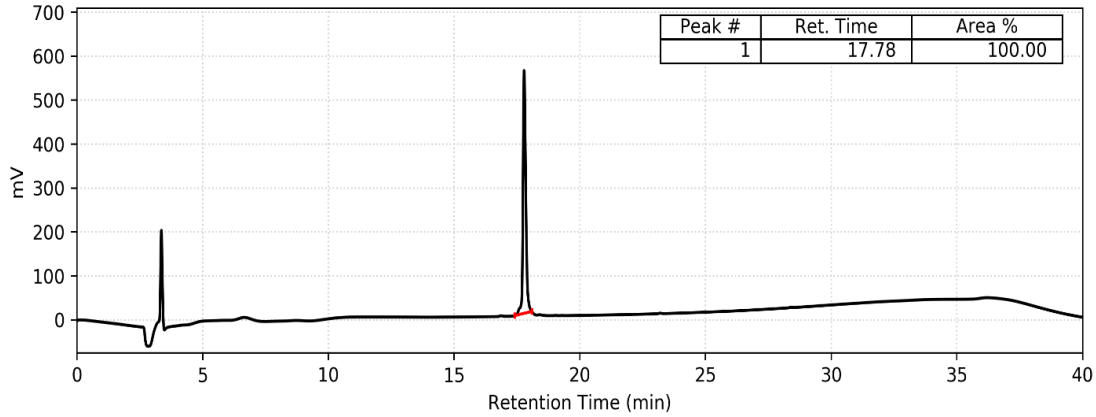
S. mitis-CSP-2-E1A



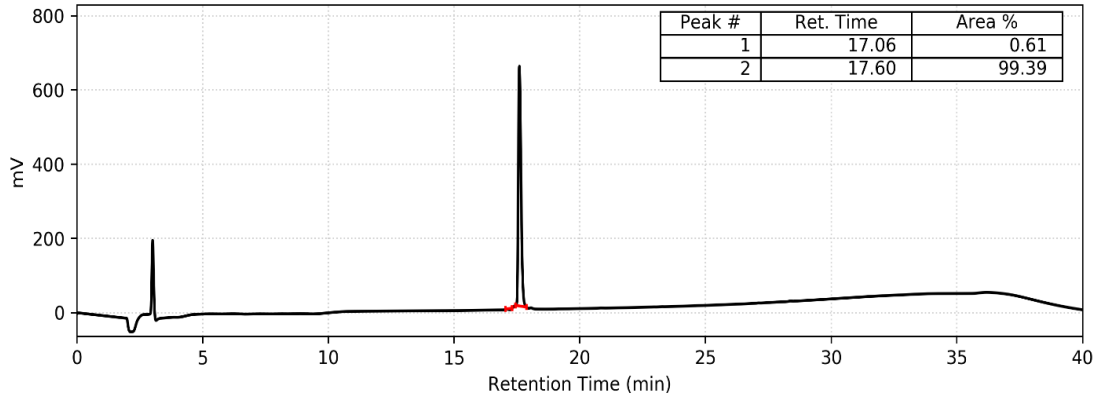
S. mitis-CSP-2-I2A



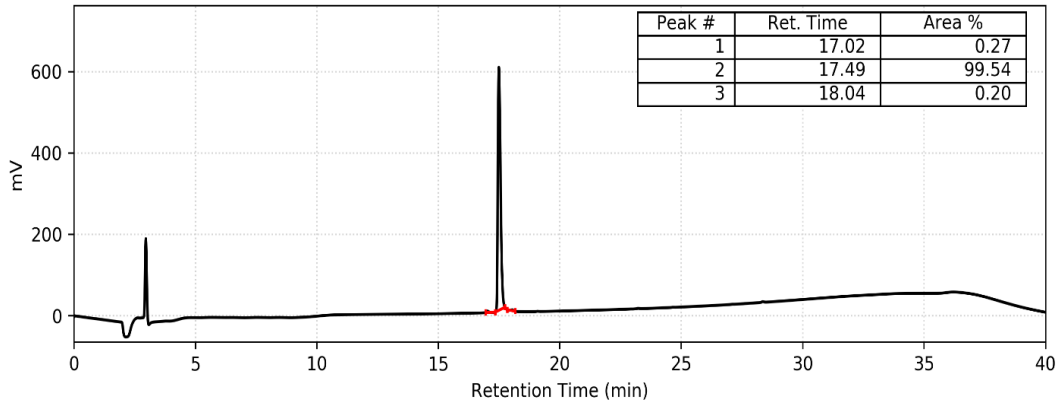
S. mitis-CSP-2-R3A



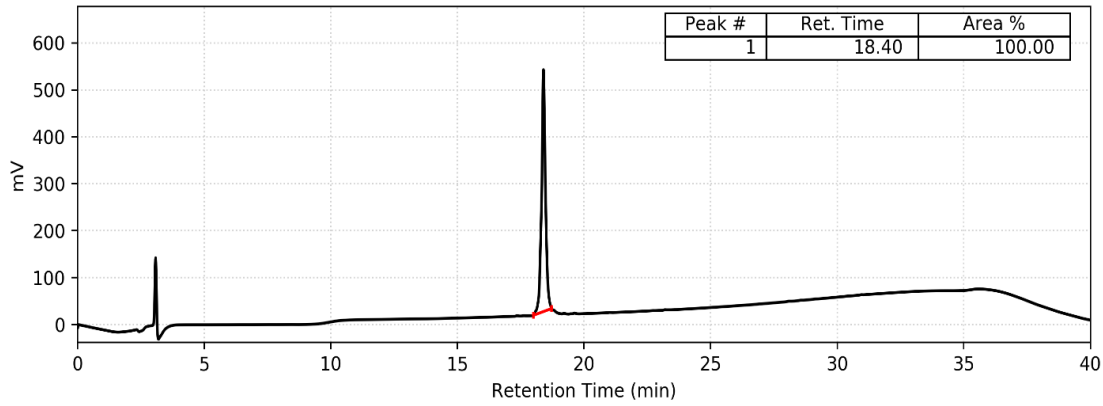
S. mitis-CSP-2-Q4A



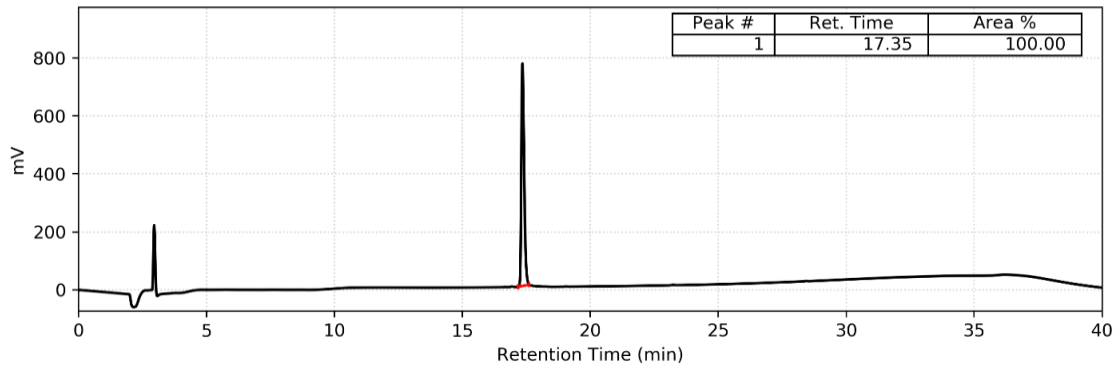
S. mitis-CSP-2-T5A



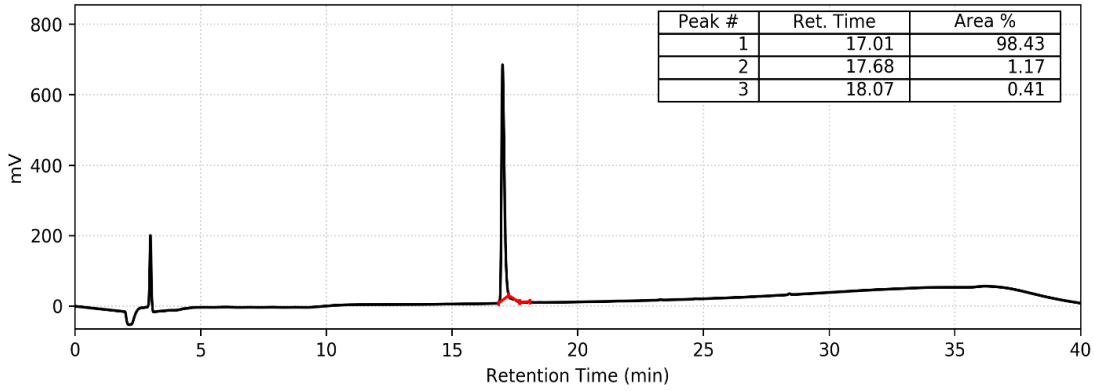
S. mitis-CSP-2-H6A



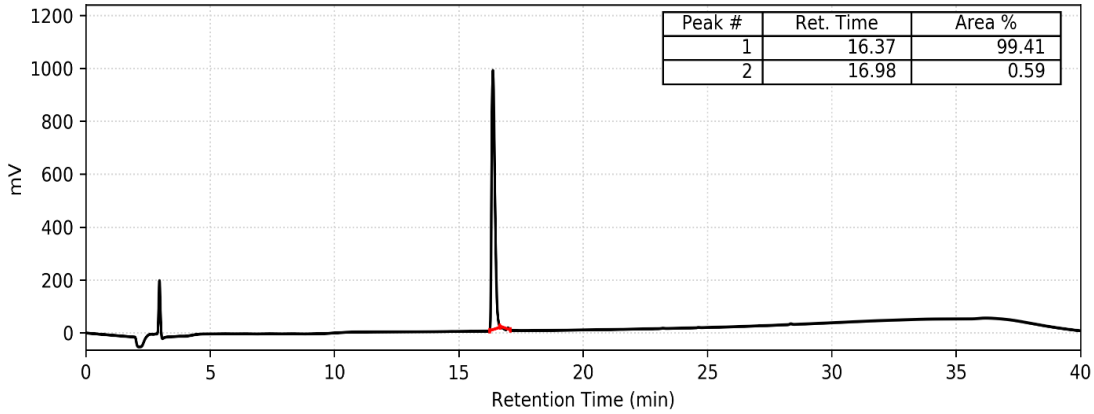
S. mitis-CSP-2-N7A



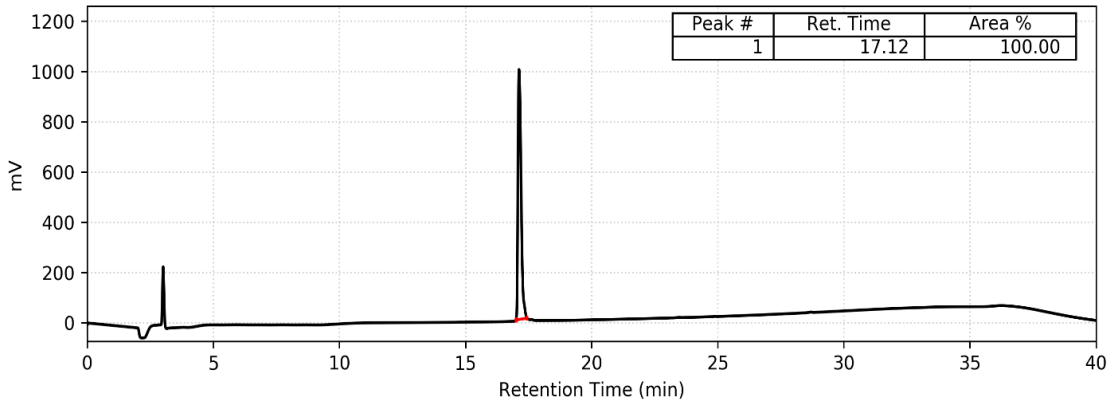
S. mitis-CSP-2-I8A



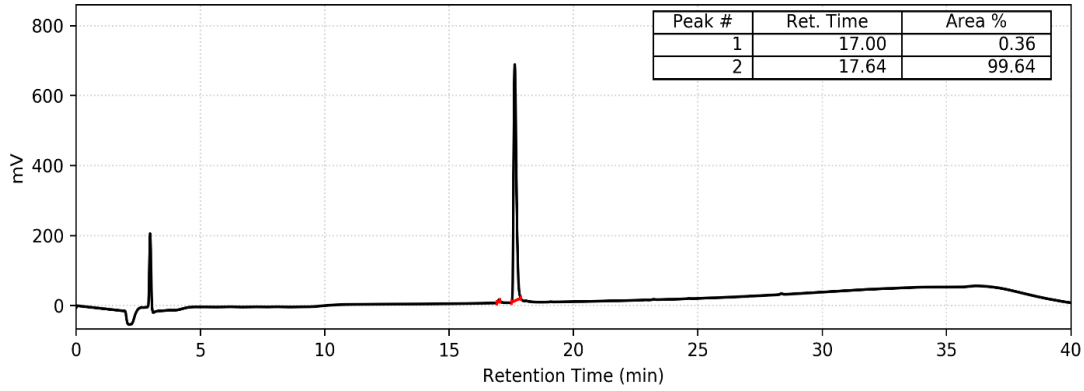
S. mitis-CSP-2-F9A



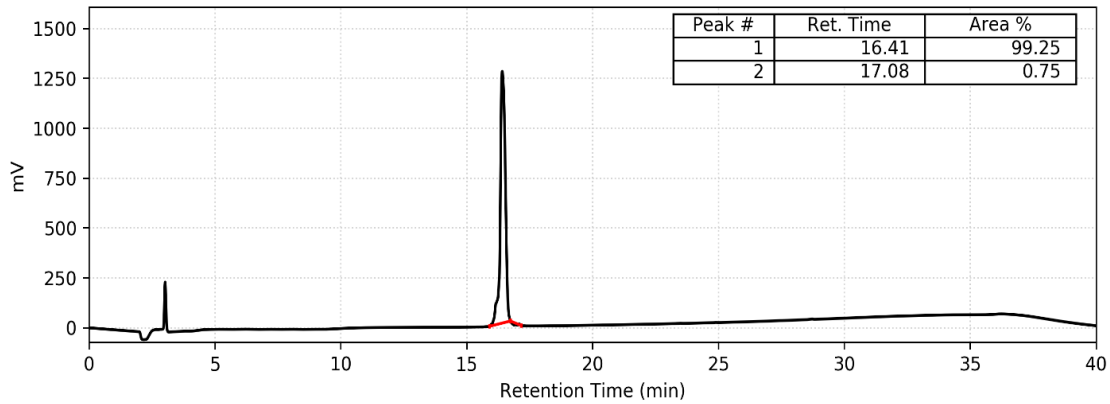
S. mitis-CSP-2-F10A



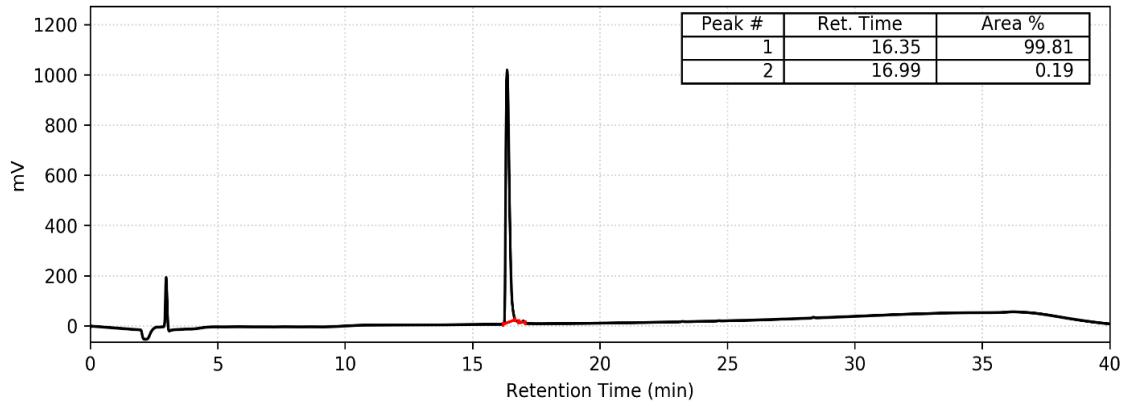
S. mitis-CSP-2-N11A



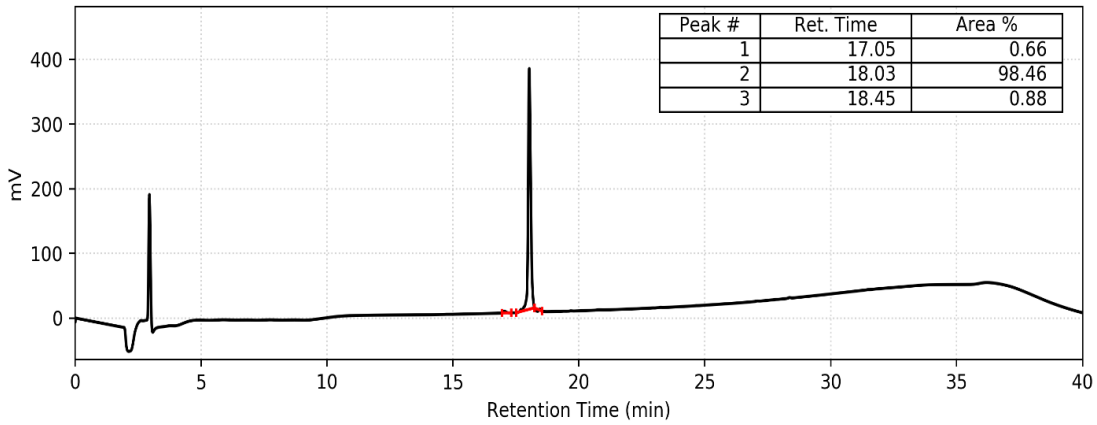
S. mitis-CSP-2-F12A



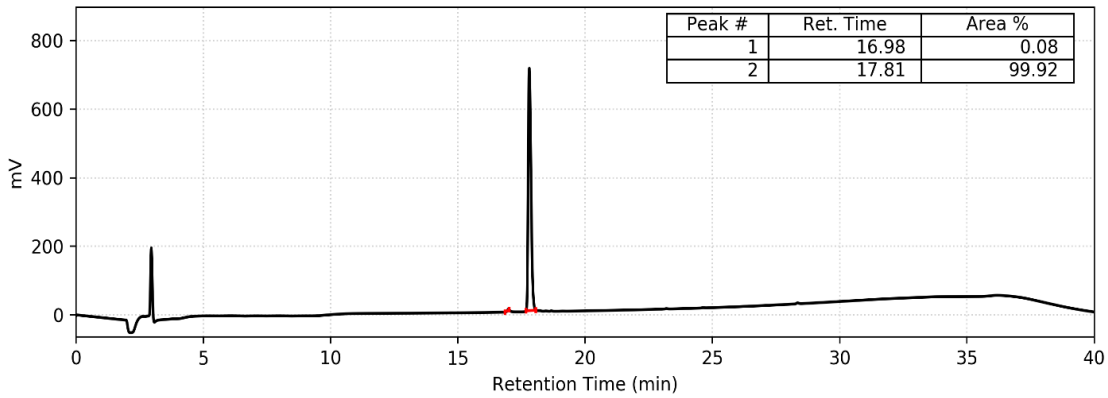
S. mitis-CSP-2-F13A



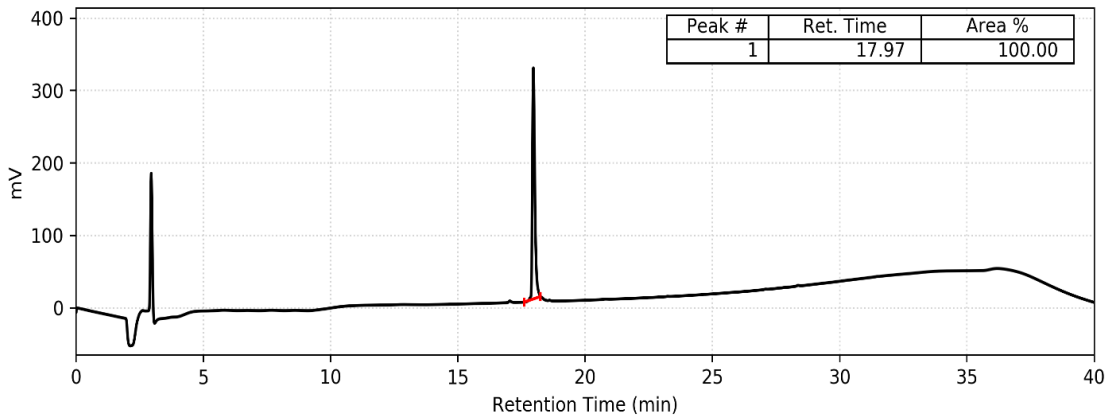
S. mitis-CSP-2-K14A



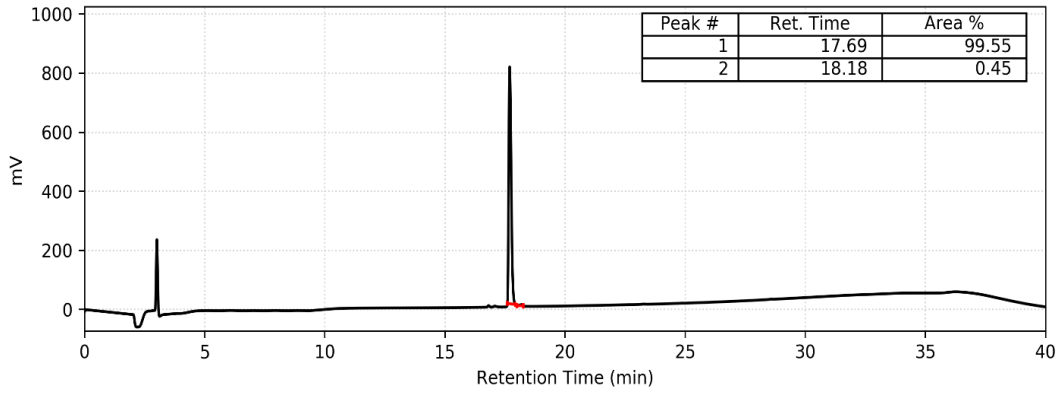
S. mitis-CSP-2-R15A



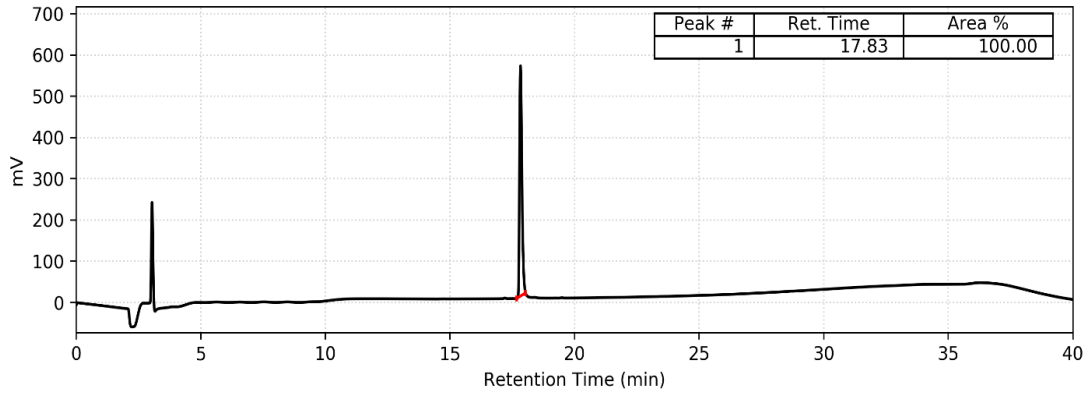
S. mitis-CSP-2-R16A



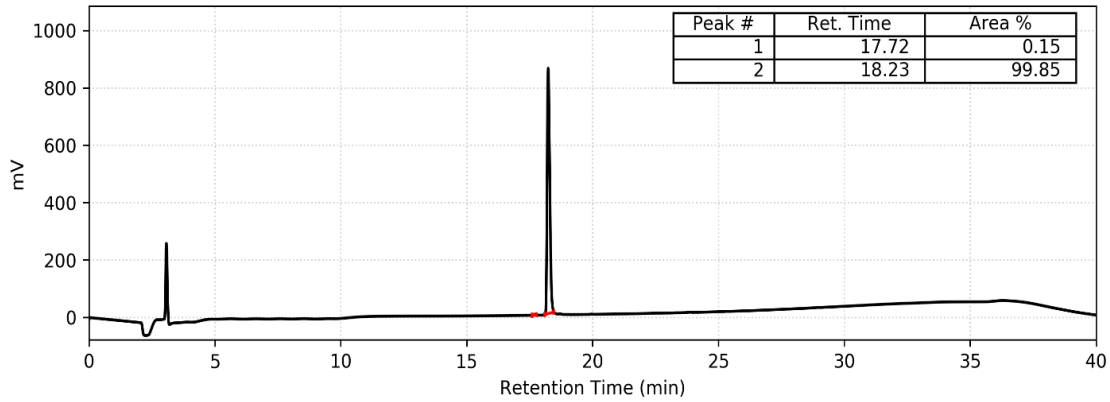
S. mitis-CSP-2-e1



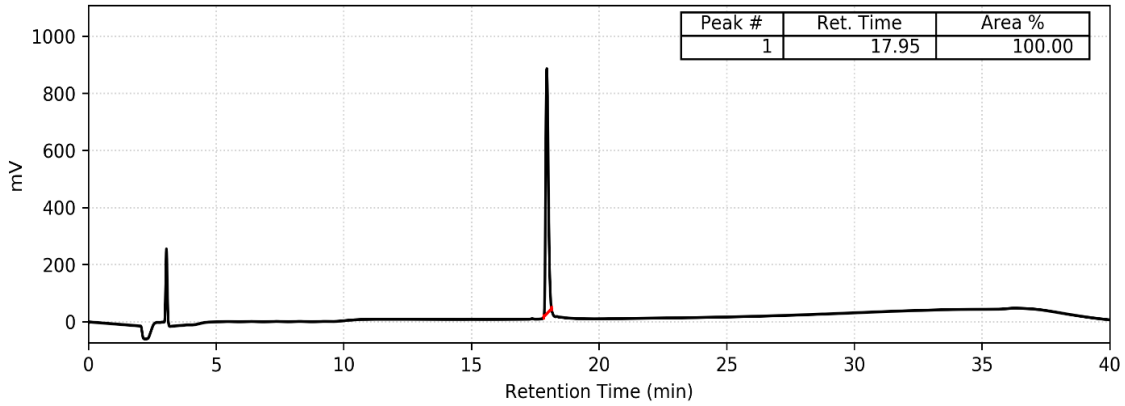
S. mitis-CSP-2-i2



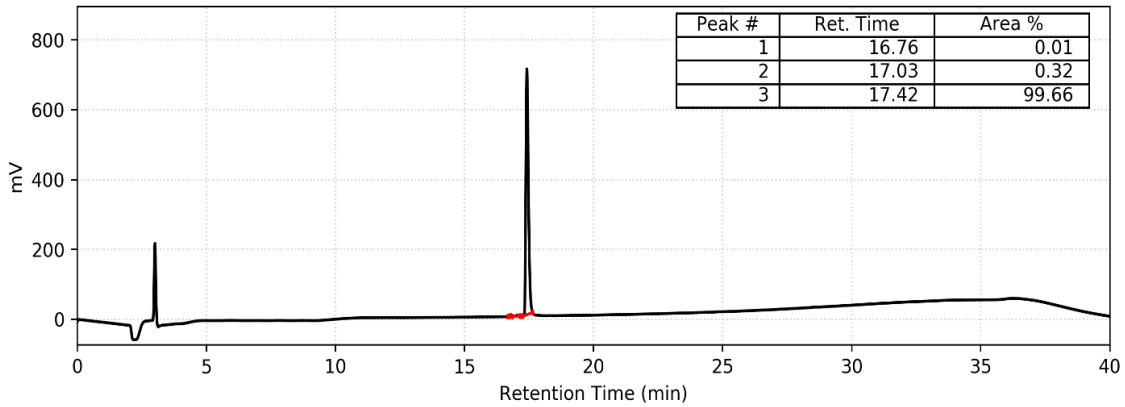
S. mitis-CSP-2-r3



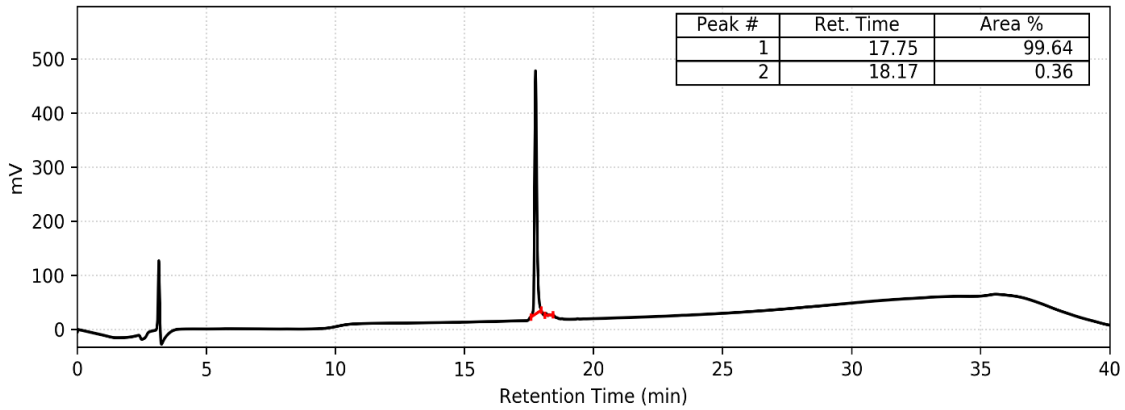
S. mitis-CSP-2-q4



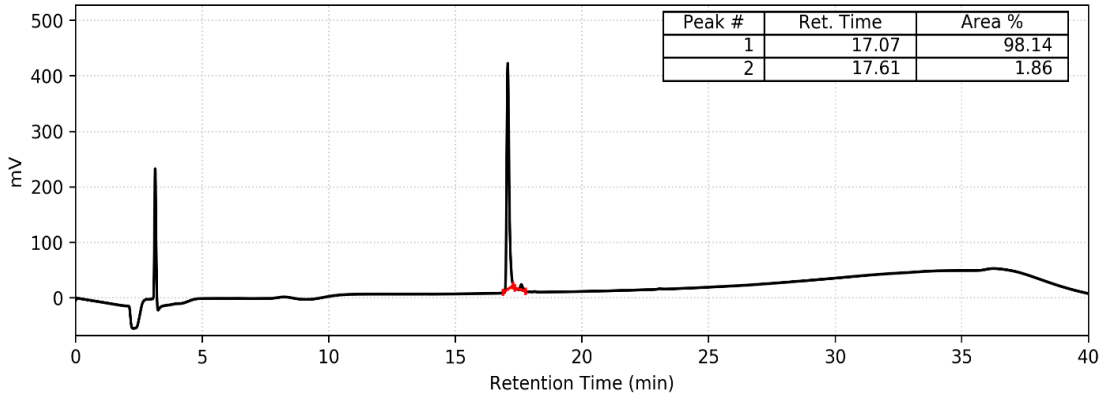
S. mitis-CSP-2-t5



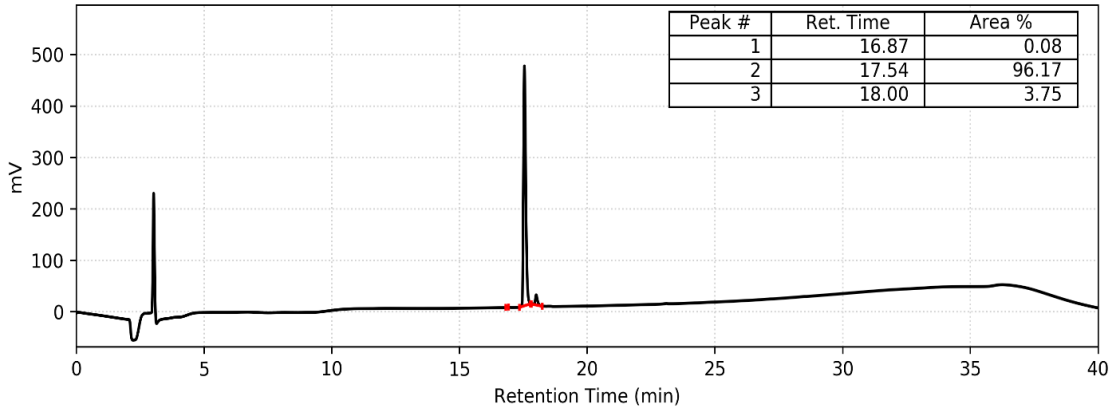
S. mitis-CSP-2-h6



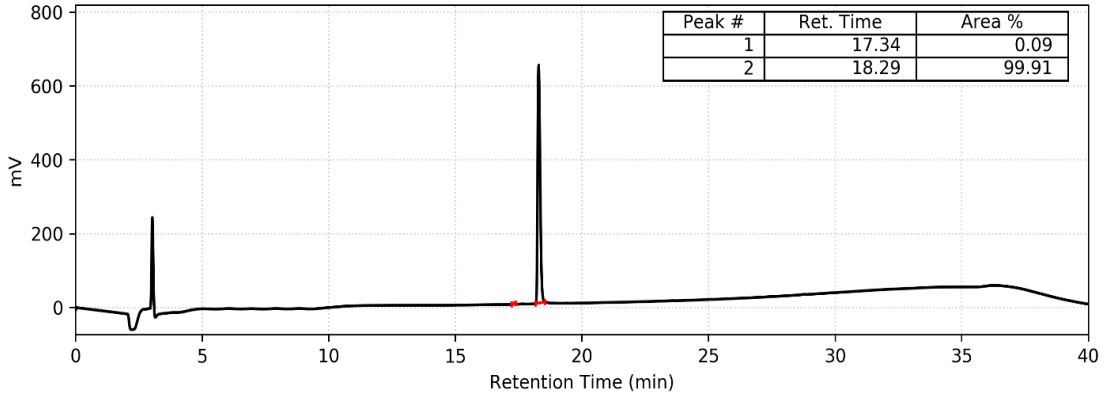
S. mitis-CSP-2-n7



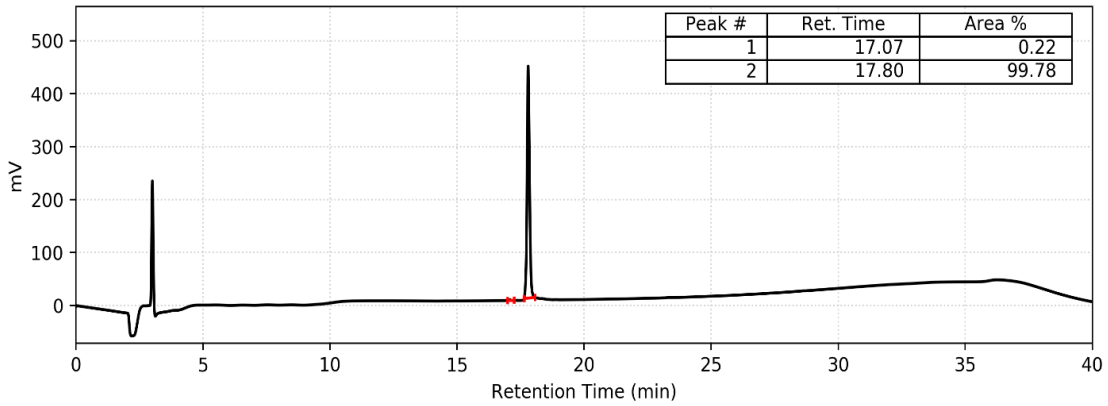
S. mitis-CSP-2-i8



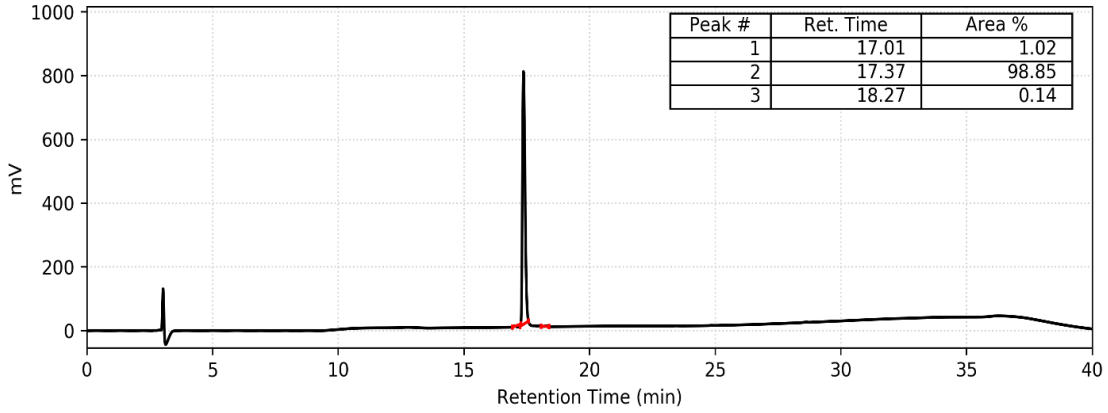
S. mitis-CSP-2-f9



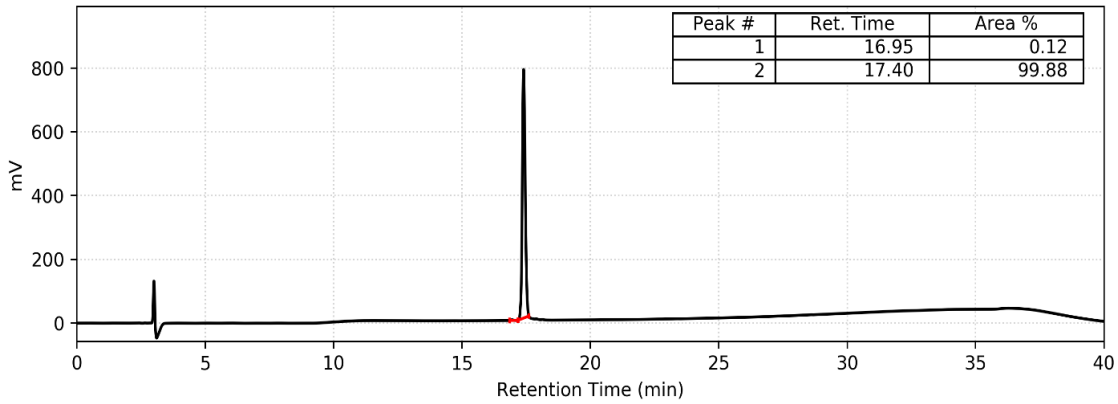
S. mitis-CSP-2-f10



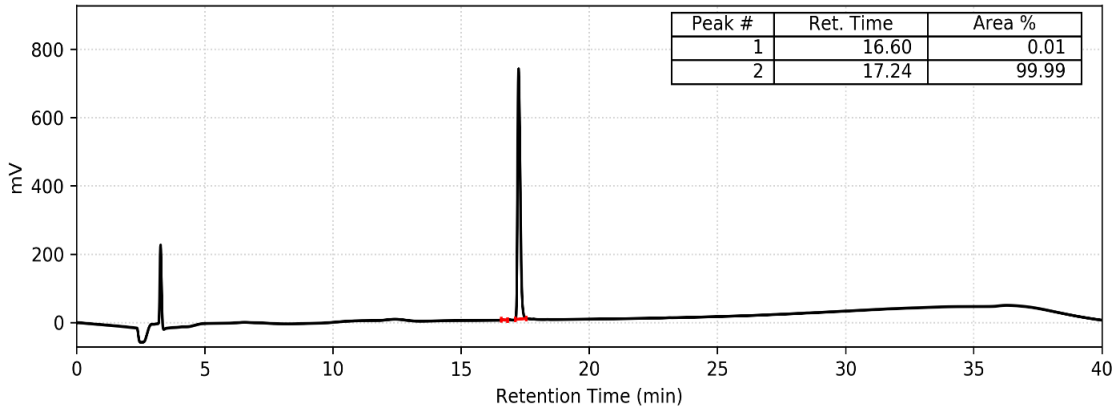
S. mitis-CSP-2-n11



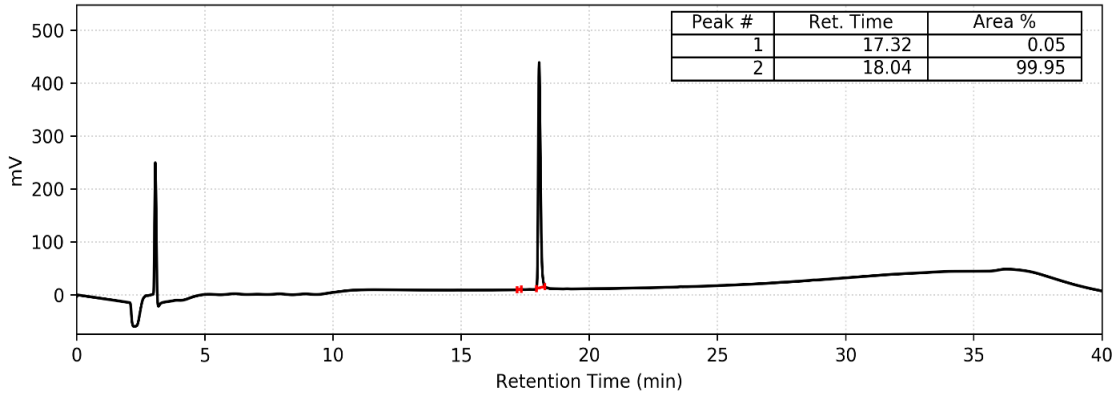
S. mitis-CSP-2-f12



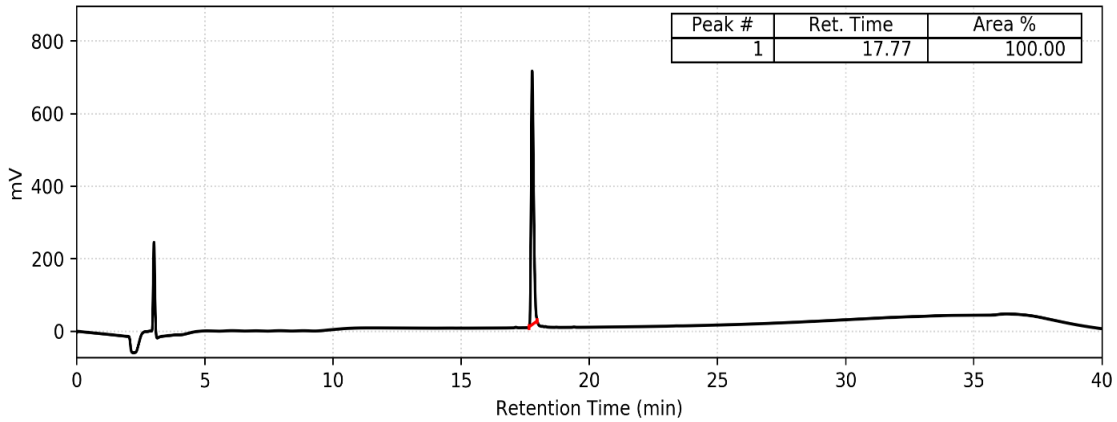
S. mitis-CSP-2-f13



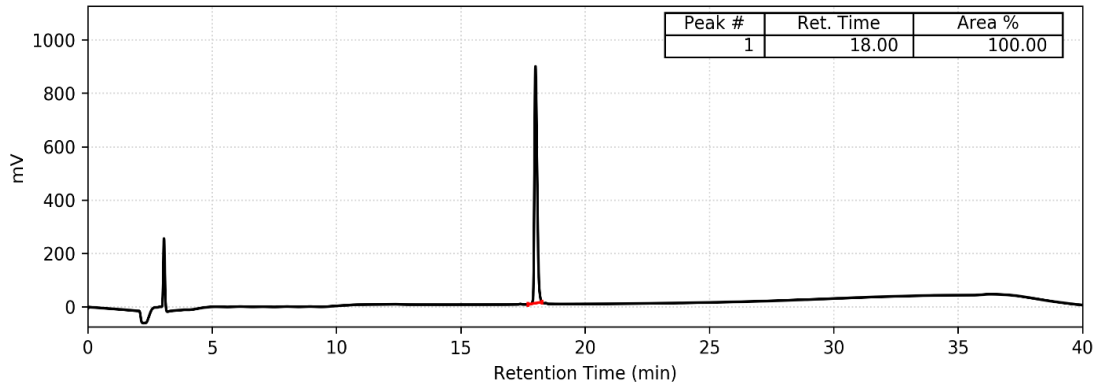
S. mitis-CSP-2-k14



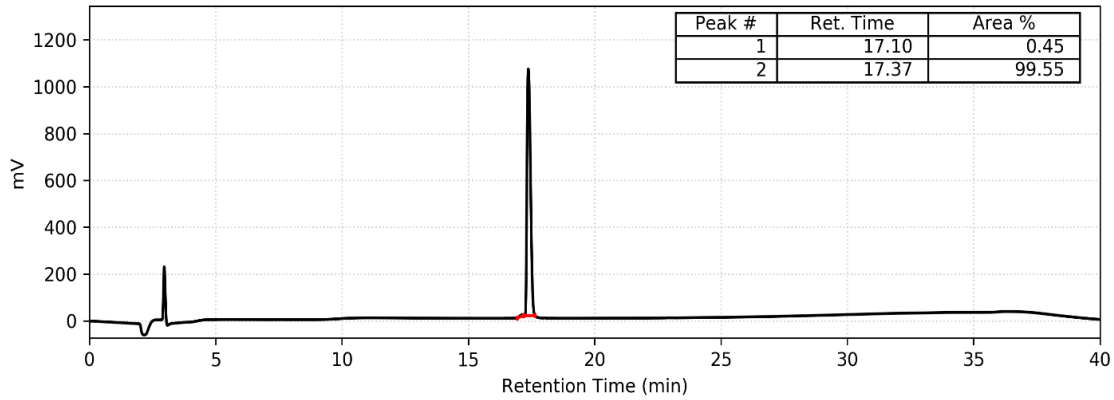
S. mitis-CSP-2-r15



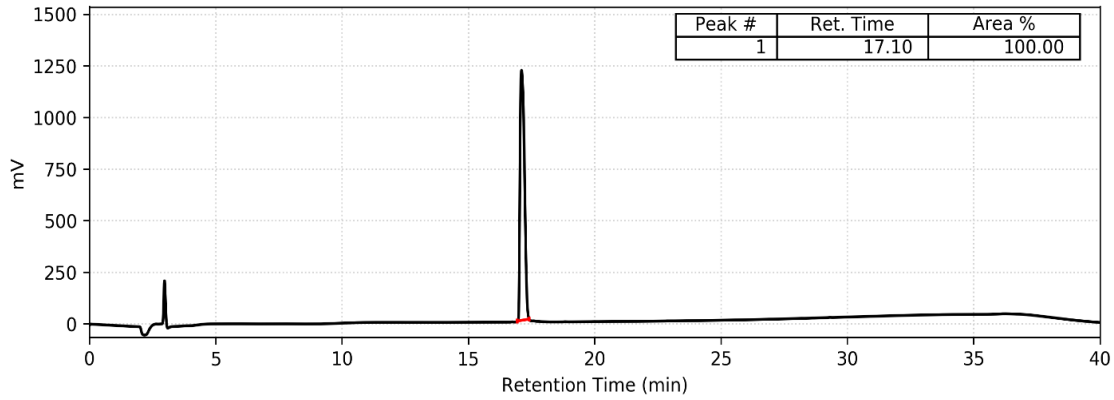
S. mitis-CSP-2-r16



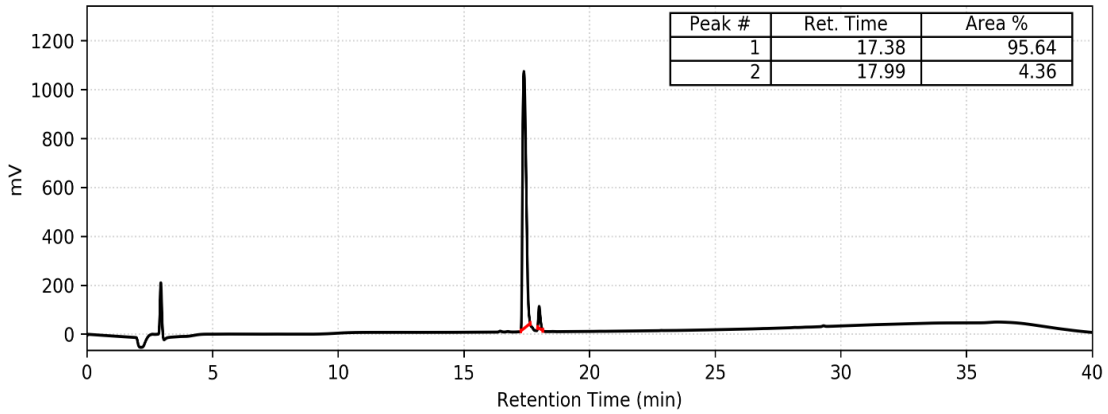
S. mitis-CSP-2-des-E1



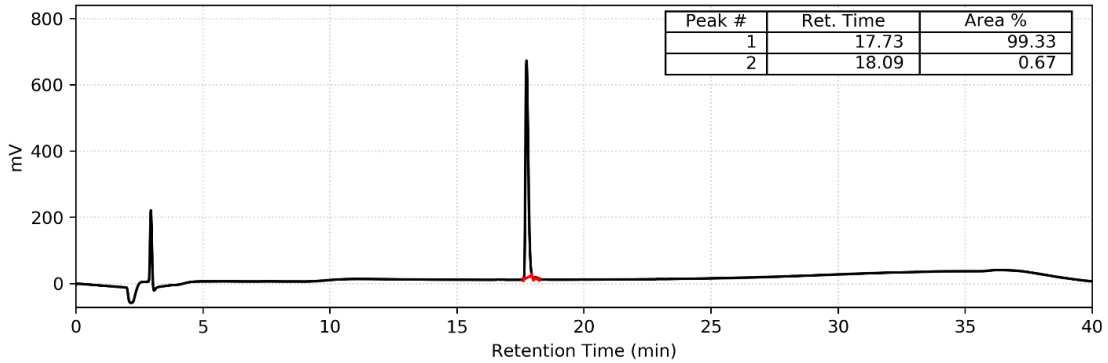
S. mitis-CSP-2-des-E1I2



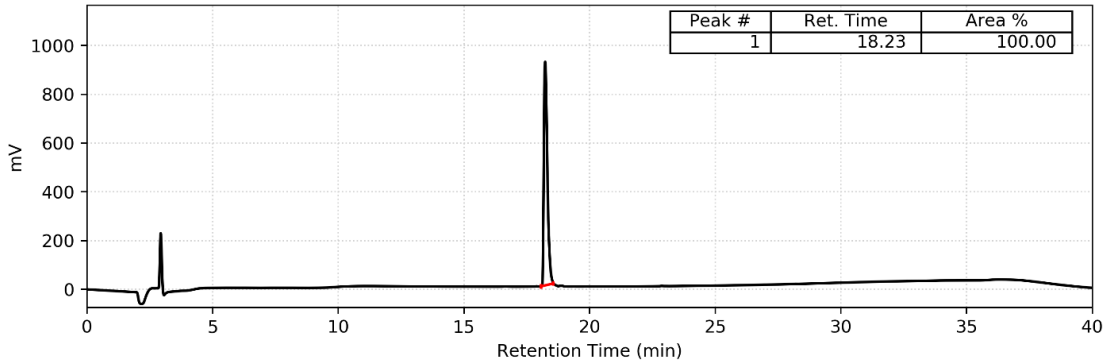
S. mitis-CSP-2-des-E1I2R3



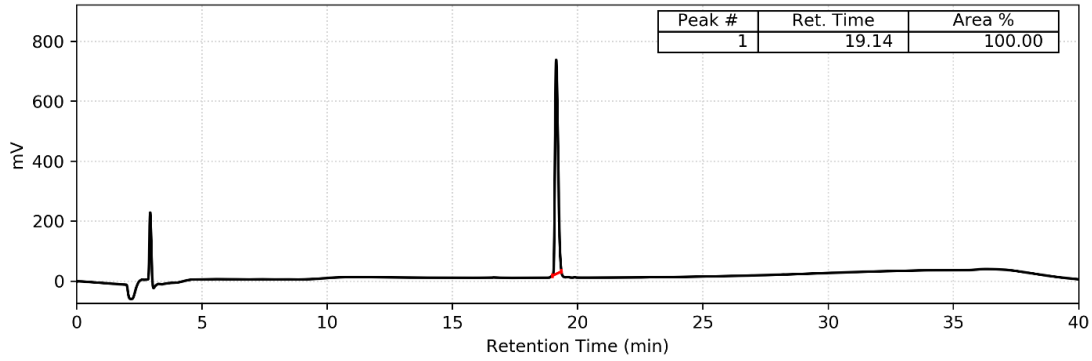
S. mitis-CSP-2-des-R16



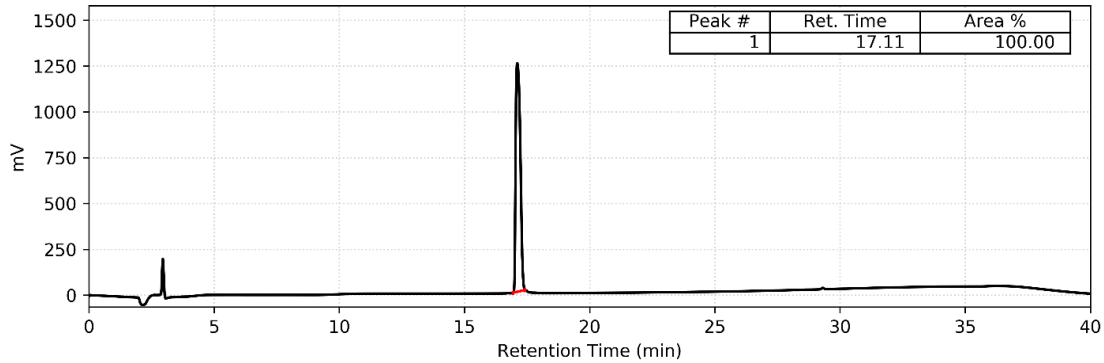
S. mitis-CSP-2-des-R15R16



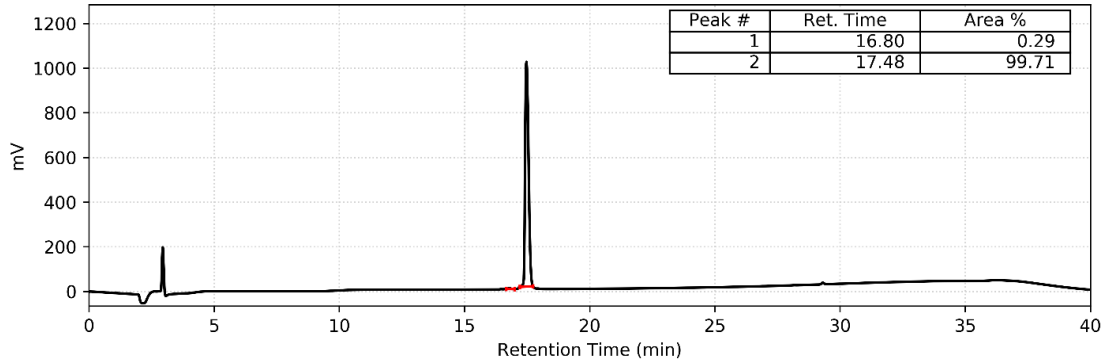
S. mitis-CSP-2-des-K14R15R16



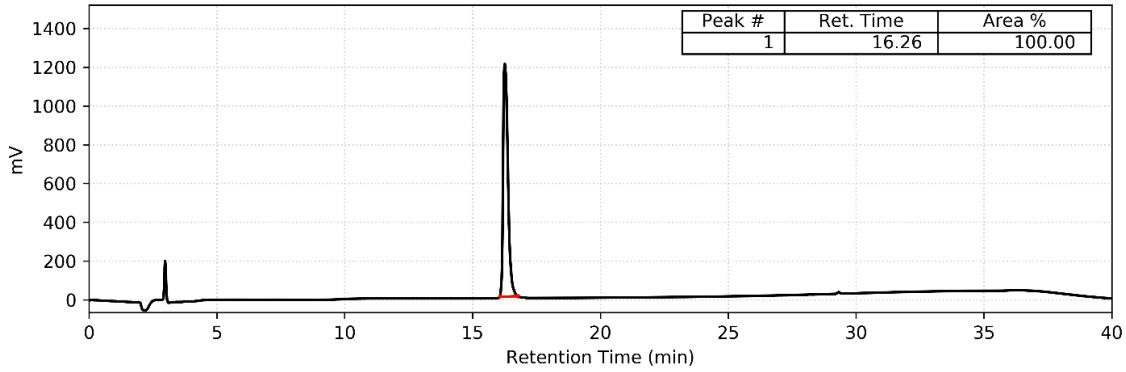
S. mitis-CSP-2-E1AN7A



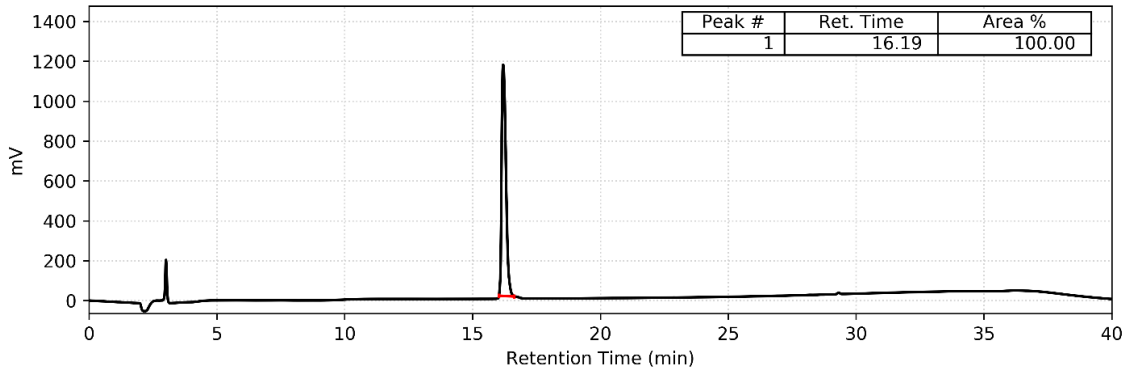
S. mitis-CSP-2-E1AN11A



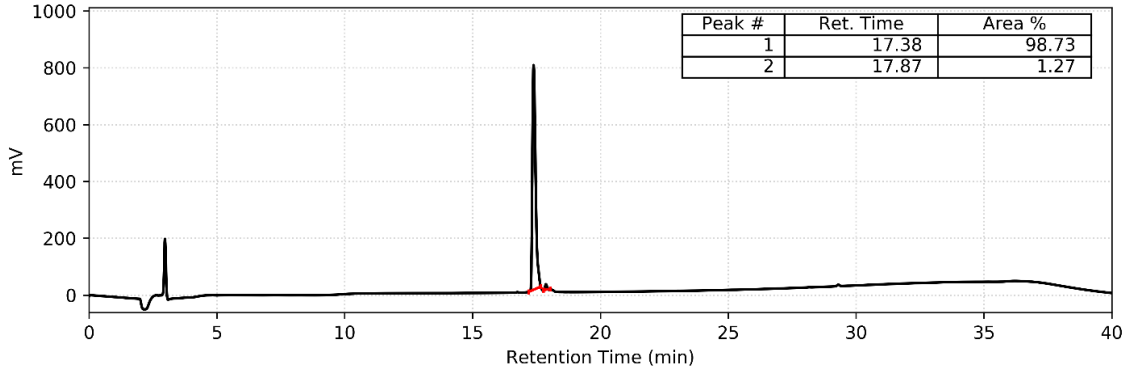
S. mitis-CSP-2-E1AF12A



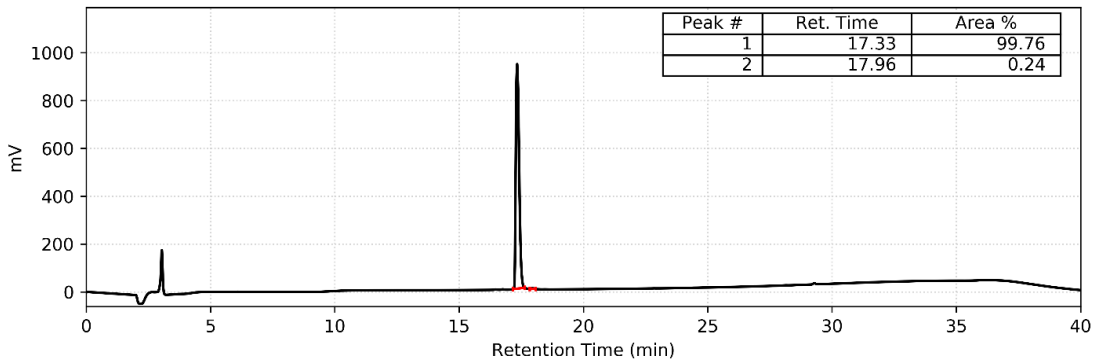
S. mitis-CSP-2-E1AF13A



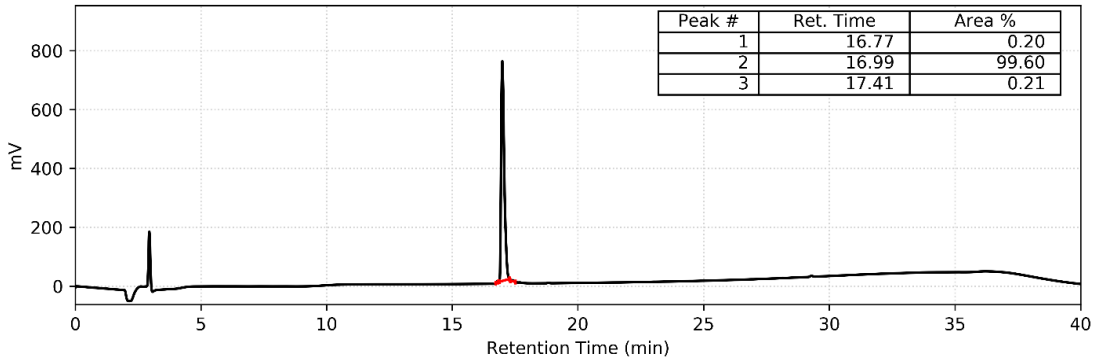
S. mitis-CSP-2-E1Ai8



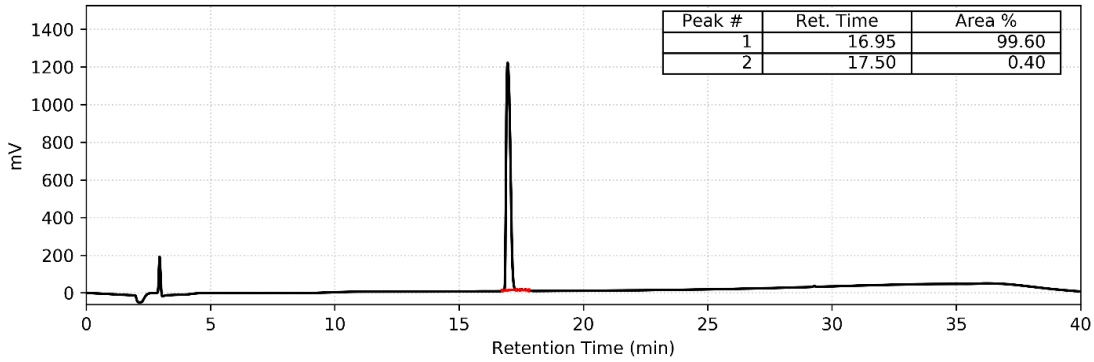
S. mitis-CSP-2-E1Af10



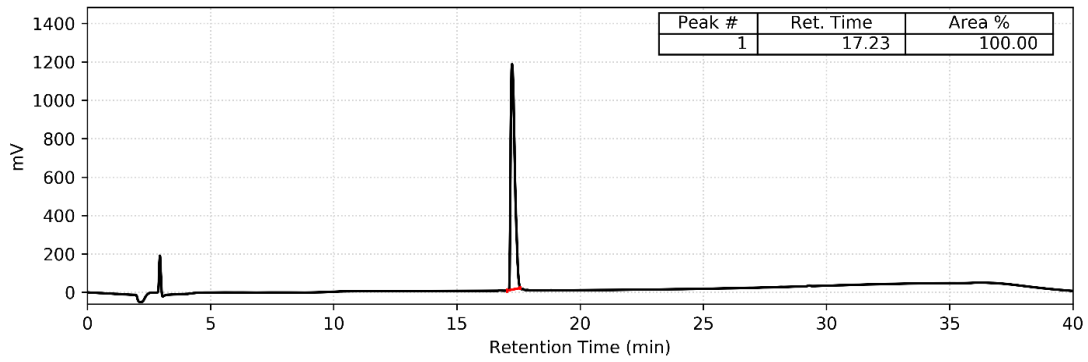
S. mitis-CSP-2-E1An11



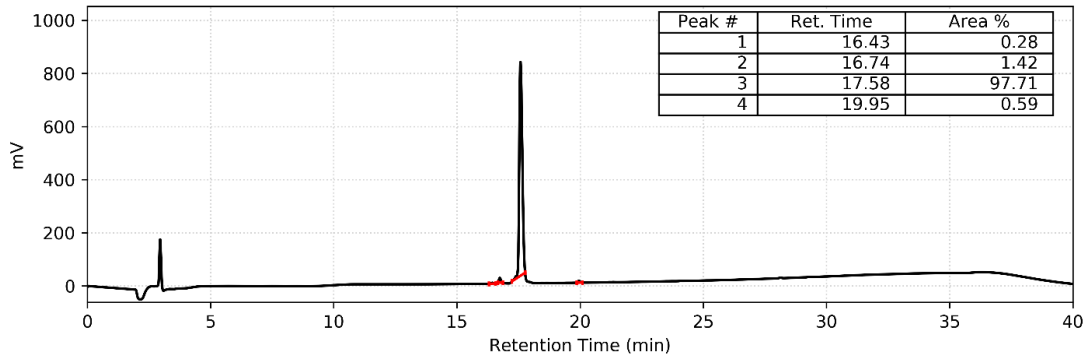
S. mitis-CSP-2-E1Ak14



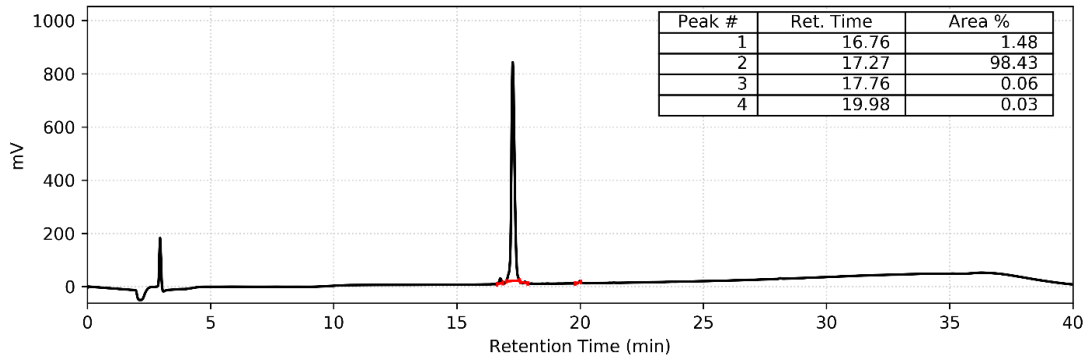
S. mitis-CSP-2-E1Ar16



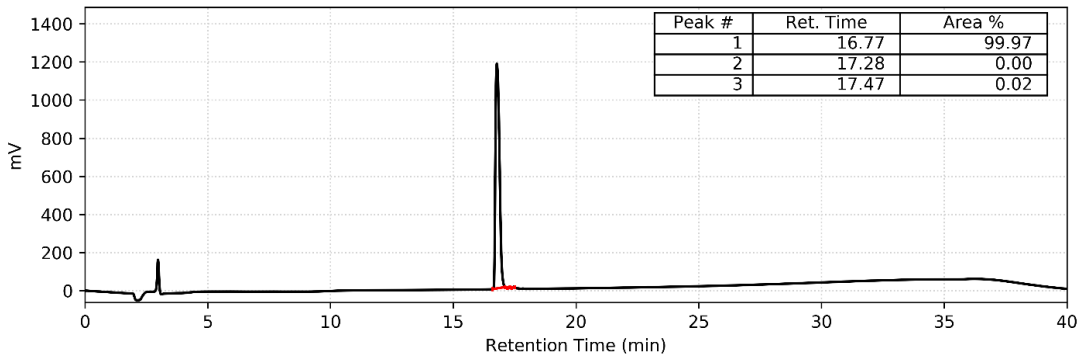
S. mitis-CSP-2-E1A-des-R16



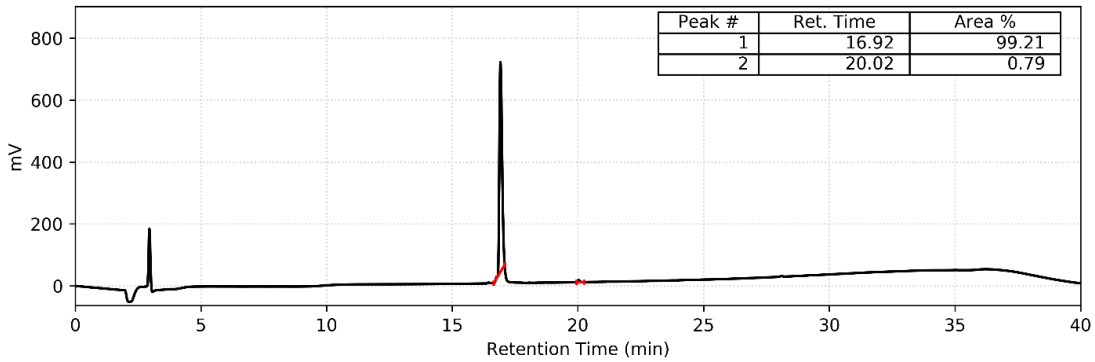
S. mitis-CSP-2-E1AN7Af10



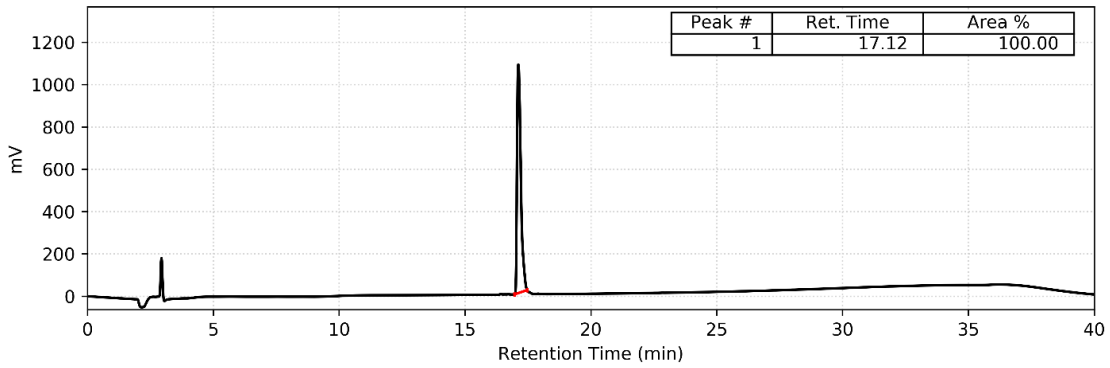
S. mitis-CSP-2-E1AN7An11



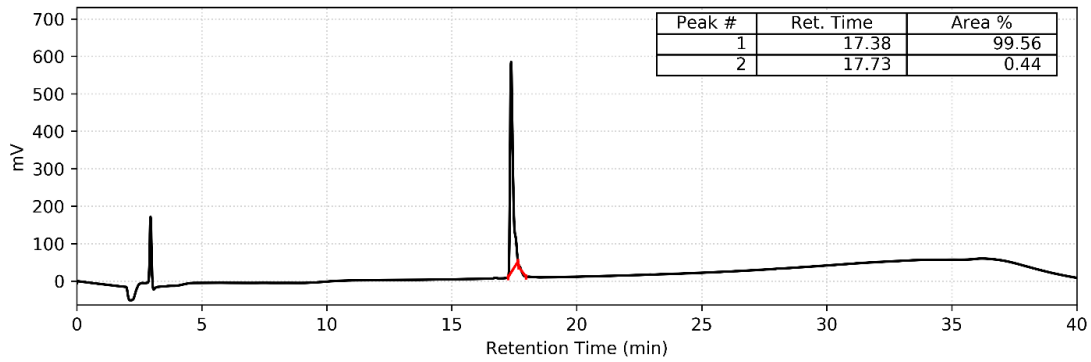
S. mitis-CSP-2-E1AN7Ak14



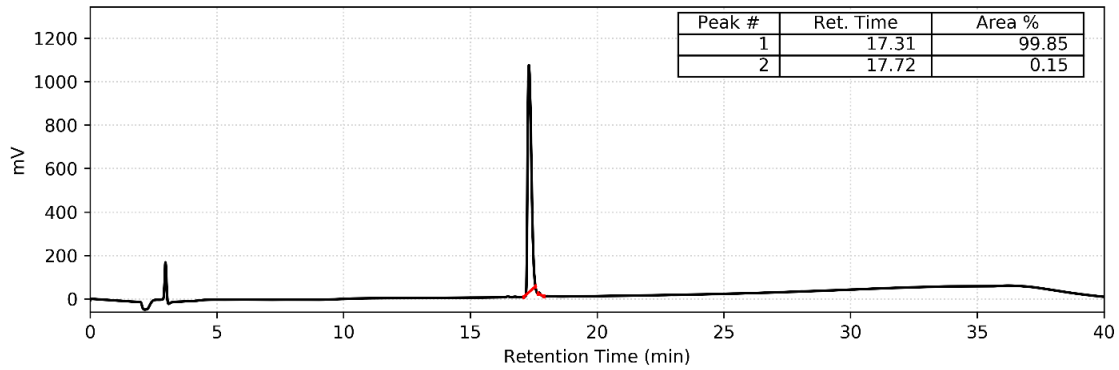
S. mitis-CSP-2-E1AN7Ar16



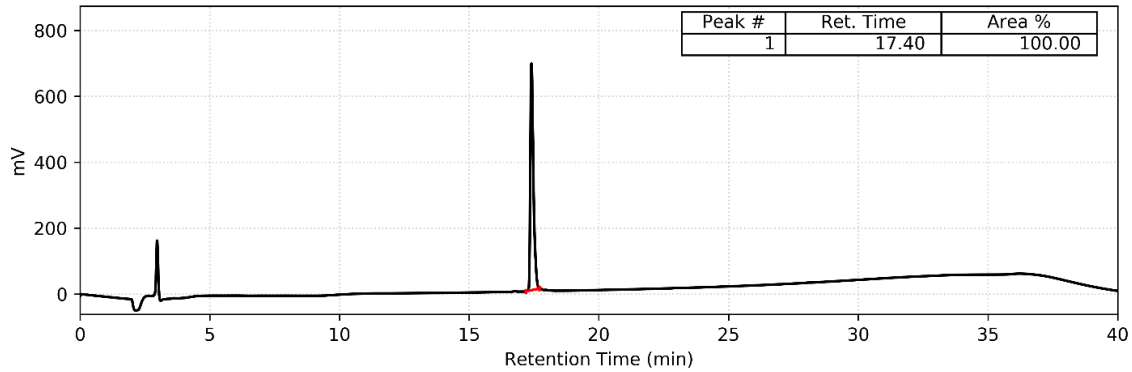
S. mitis-CSP-2-E1Af10n11



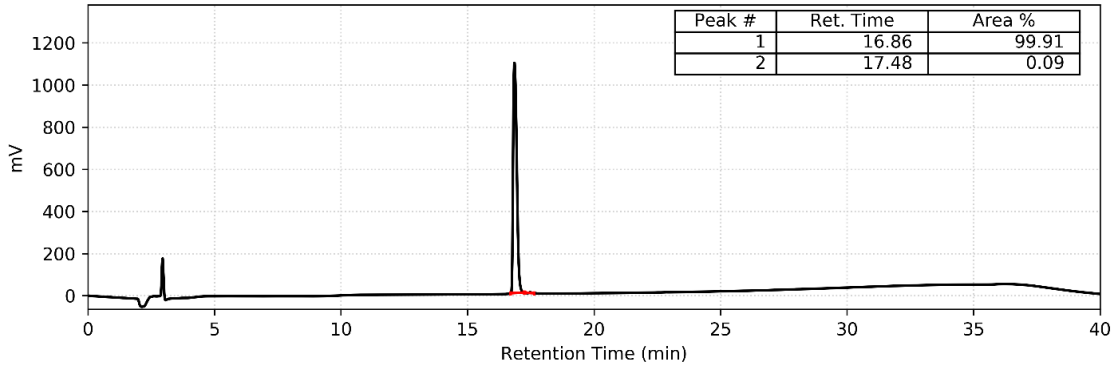
S. mitis-CSP-2-E1Af10k14



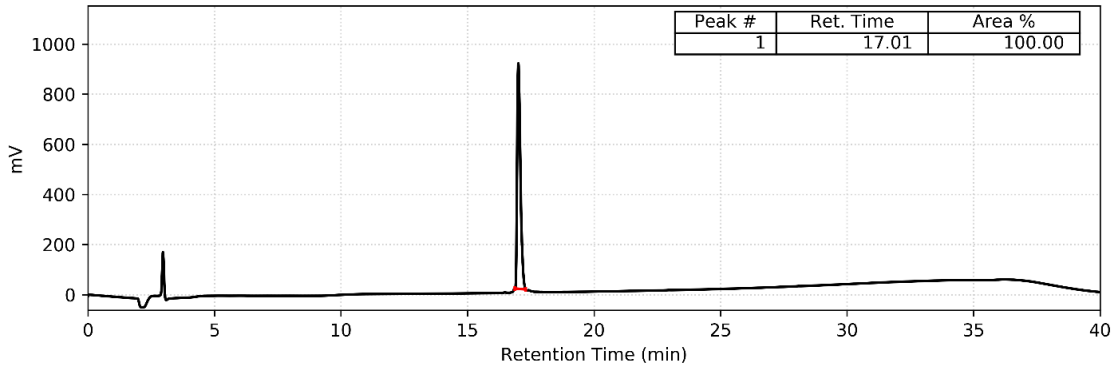
S. mitis-CSP-2-E1Af10r16



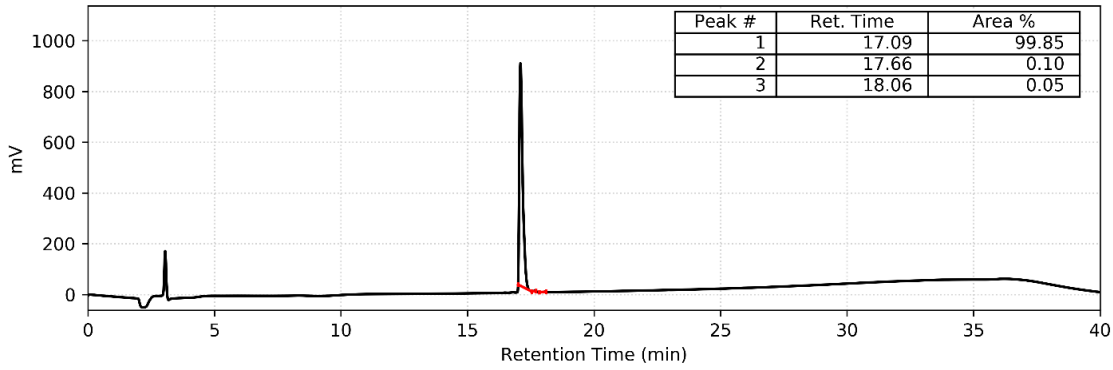
S. mitis-CSP-2-E1An11k14



S. mitis-CSP-2-E1An11r16



S. mitis-CSP-2-E1Ak14r16



MS and HPLC data for *S. mitis*-CSP-2 analogs

Table S-2. MS and HPLC data for *S. mitis*-CSP-2 alanine-screen analogs.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
S. mitis-CSP-2 (synthesized)	1077.5808	1077.5829	≥99
S. mitis-CSP-2 (extracted)	1077.5808	1077.5833	≥99
S. mitis-CSP-2-E1A	1048.5781	1048.5754	≥99
S. mitis-CSP-2-I2A	1056.5573	1056.5587	≥99
S. mitis-CSP-2-R3A	1035.0488	1035.0520	≥99
S. mitis-CSP-2-Q4A	1049.0701	1049.0711	≥99
S. mitis-CSP-2-T5A	1062.5755	1062.5787	≥99
S. mitis-CSP-2-H6A	1044.5699	1044.5712	≥99
S. mitis-CSP-2-N7A	1056.0779	1056.0822	≥99
S. mitis-CSP-2-I8A	1056.5573	1056.5610	≥98
S. mitis-CSP-2-F9A	1039.5651	1039.5666	≥99
S. mitis-CSP-2-F10A	1039.5651	1039.5665	≥99
S. mitis-CSP-2-N11A	1056.0779	1056.0797	≥99
S. mitis-CSP-2-F12A	1039.5651	1039.5674	≥99
S. mitis-CSP-2-F13A	1039.5651	1039.5674	≥99
S. mitis-CSP-2-K14A	1049.0519	1049.0558	≥98
S. mitis-CSP-2-R15A	1035.0488	1035.0519	≥99
S. mitis-CSP-2-R16A	1035.0488	1035.0507	≥99

EM = Exact Mass. See the methods above.

Table S-3. MS and HPLC data for *S. mitis*-CSP-2 D-amino acid scan analogs.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
S. mitis-CSP-2-e1	1077.5808	1077.5802	≥99
S. mitis-CSP-2-i2	1077.5808	1077.5807	≥99
S. mitis-CSP-2-r3	1077.5808	1077.5835	≥99
S. mitis-CSP-2-q4	1077.5808	1077.5834	≥99
S. mitis-CSP-2-t5	1077.5808	1077.5806	≥99
S. mitis-CSP-2-h6	1077.5808	1077.5793	≥99
S. mitis-CSP-2-n7	1077.5808	1077.5803	≥98
S. mitis-CSP-2-i8	1077.5808	1077.5785	≥96
S. mitis-CSP-2-f9	1077.5808	1077.5787	≥99
S. mitis-CSP-2-f10	1077.5808	1077.5795	≥99
S. mitis-CSP-2-n11	1077.5808	1077.5814	≥98
S. mitis-CSP-2-f12	1077.5808	1077.5800	≥99
S. mitis-CSP-2-f13	1077.5808	1077.5820	≥99
S. mitis-CSP-2-k14	1077.5808	1077.5831	≥99
S. mitis-CSP-2-r15	1077.5808	1077.5832	≥99
S. mitis-CSP-2-r16	1077.5808	1077.5845	≥99

EM = Exact Mass. See the methods above.

Table S-4. MS and HPLC data for *S. mitis*-CSP-2 truncated analogs.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
S. mitis-CSP-2-des-E1	1013.0595	1013.0624	≥99
S. mitis-CSP-2-des-E1I2	956.5175	956.5205	≥99
S. mitis-CSP-2-des-E1I2R3	878.4669	878.4683	≥95
S. mitis-CSP-2-des-R16	999.5303	999.5330	≥99
S. mitis-CSP-2-des-R15R16	921.4797	921.4823	≥99
S. mitis-CSP-2-des-K14R15R16	856.9308	856.9336	≥99

EM = Exact Mass. See the methods above.

Table S-5. MS and HPLC data for *S. mitis*-CSP-2-E1A modification analogs.

Compound Name	Calc. EM MH_2^{2+}	Obs. EM MH_2^{2+}	Purity (%)
S. mitis-CSP-2-E1AN7A	1027.0752	1027.0790	≥99
S. mitis-CSP-2-E1AN11A	1027.0752	1027.0784	≥99
S. mitis-CSP-2-E1AF12A	1010.5624	1010.5655	≥99
S. mitis-CSP-2-E1AF13A	1010.5624	1010.5651	≥99
S. mitis-CSP-2-E1Ai8	1048.5781	1048.5808	≥98
S. mitis-CSP-2-E1Af10	1048.5781	1048.5812	≥99
S. mitis-CSP-2-E1An11	1048.5781	1048.5790	≥99
S. mitis-CSP-2-E1Ak14	1048.5781	1048.5788	≥99
S. mitis-CSP-2-E1Ar16	1048.5781	1048.5790	≥99
S. mitis-CSP-2-E1A-des-R16	970.5275	970.5297	≥97
S. mitis-CSP-2-E1AN7Af10	1027.0752	1027.0769	≥98
S. mitis-CSP-2-E1AN7An11	1027.0752	1027.0755	≥99
S. mitis-CSP-2-E1AN7Ak14	1027.0752	1027.0778	≥99
S. mitis-CSP-2-E1AN7Ar16	1027.0752	1027.0767	≥99
S. mitis-CSP-2-E1Af10n11	1048.5781	1048.5782	≥99
S. mitis-CSP-2-E1Af10k14	1048.5781	1048.5780	≥99
S. mitis-CSP-2-E1Af10r16	1048.5781	1048.5780	≥99
S. mitis-CSP-2-E1An11k14	1048.5781	1048.5774	≥99
S. mitis-CSP-2-E1An11r16	1048.5781	1048.5782	≥99
S. mitis-CSP-2-E1Ak14r16	1048.5781	1048.5782	≥99

EM = Exact Mass. See the methods above.

Primary reporter gene assay data

Agonism assays were performed at a 10 µM concentration of synthetic CSP. *S. mitis*-CSP-2 was used as the positive control (100%) while DMSO as the negative control (0%). For all figures, *comX* expression is presented as relative luminescence units (RLU) divided by measured cultural optical density (OD). Percent (%) ComD activation was measured by normalizing the RLU/OD value obtained for each peptide to that of the native *S. mitis*-CSP-2. The results are averages from three replicates and representative of three independent experiments. Error bars represent the standard error of the mean of nine values.

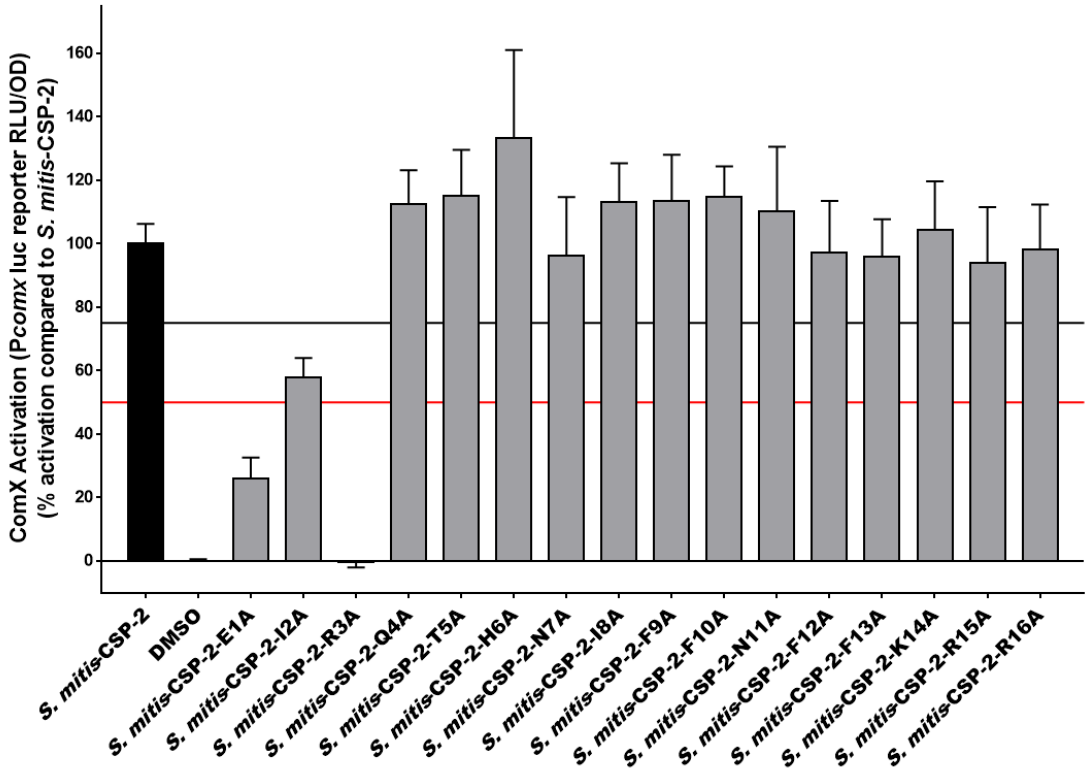


Figure S-1. Primary agonism screening assay data for the *S. mitis*-CSP-2 alanine scan analogs. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} values while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

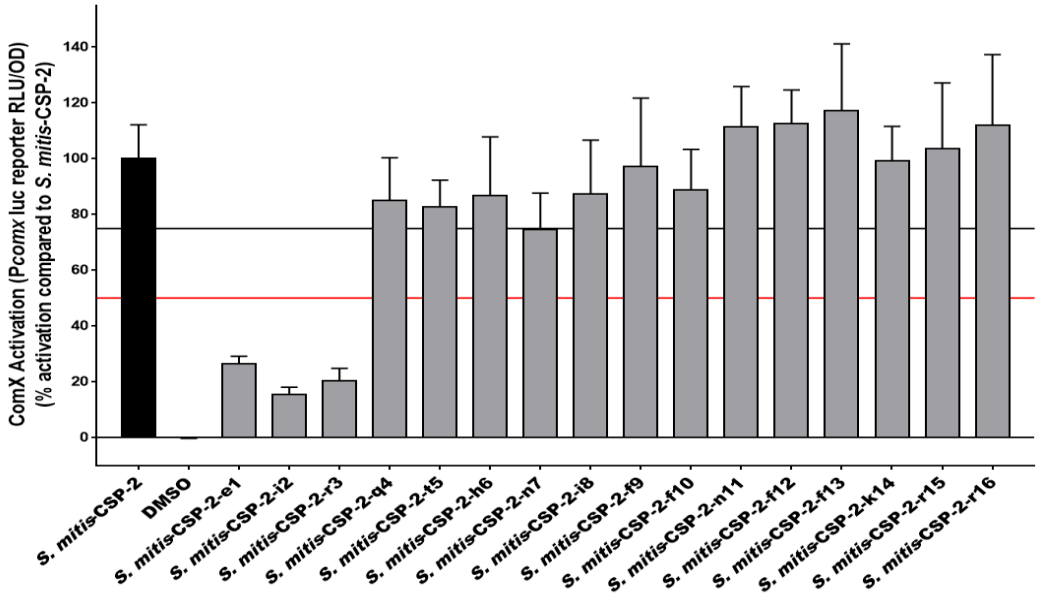


Figure S-2. Primary agonism screening assay data for the *S. mitis*-CSP-2 D-amino acid scan library. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} values while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

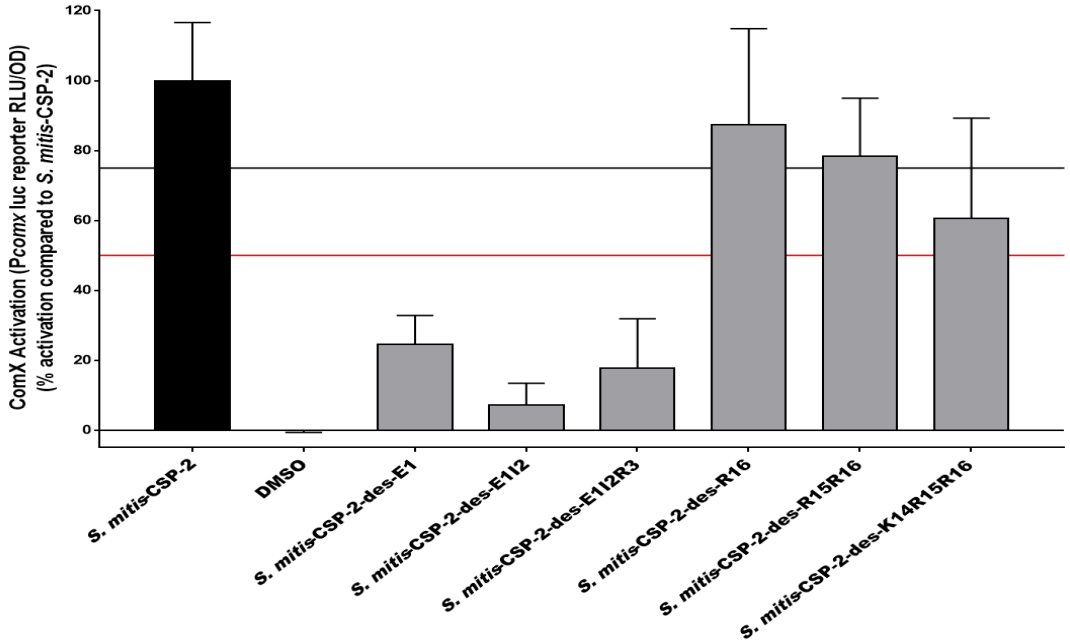


Figure S-3. Primary agonism screening assay data for the *S. mitis*-CSP-2 truncated analogs. Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} values while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

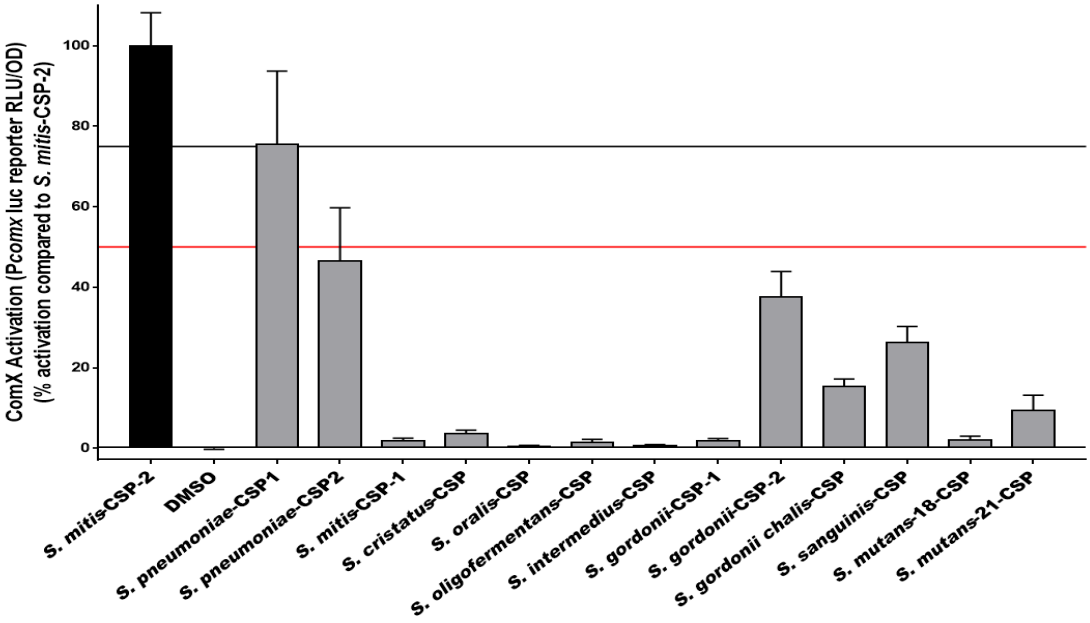


Figure S-4. Primary agonism screening assay data for the synthetic *Streptococci* native CSP pheromones.[1] *S. pneumoniae* CSP1, which exhibited over 75% activation, was further evaluated to determine its EC_{50} value, while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

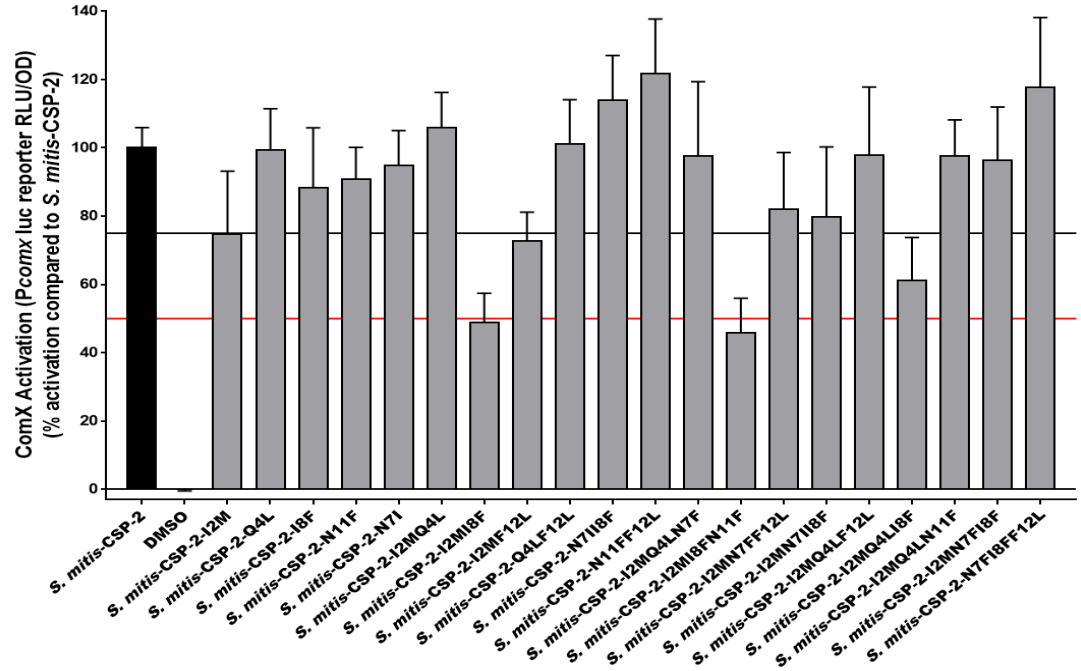


Figure S-5. Primary agonism screening assay data for the select activators from previous studies.[1] Peptides that exhibited over 75% activation were further evaluated to determine their EC_{50} values while peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

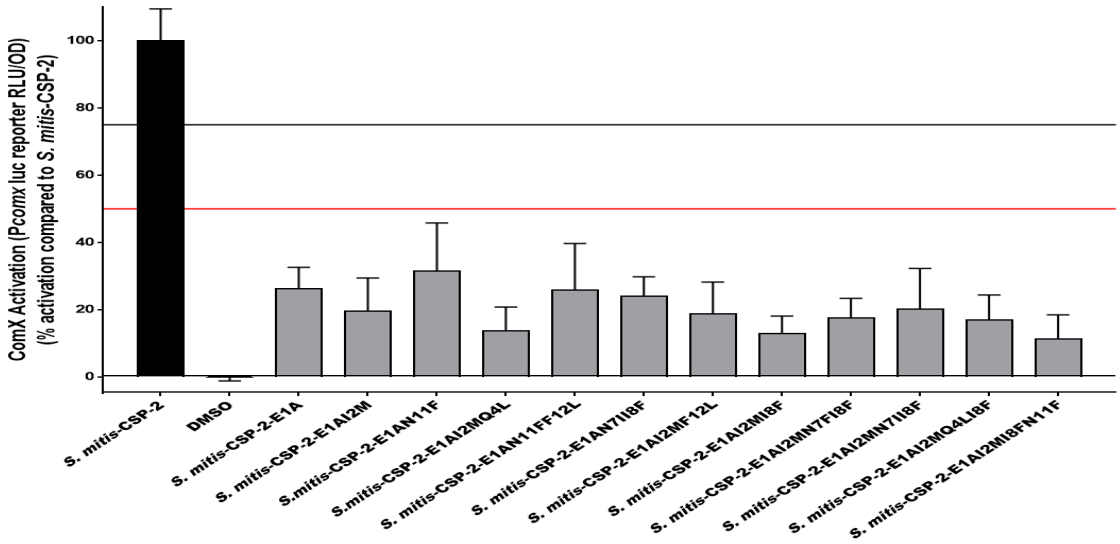


Figure S-6. Primary agonism screening assay data for the *S. mitis*-CSP-2-E1A mutant analogs from previous studies.[1] None of the peptides exhibited activation of the ComD receptor and peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

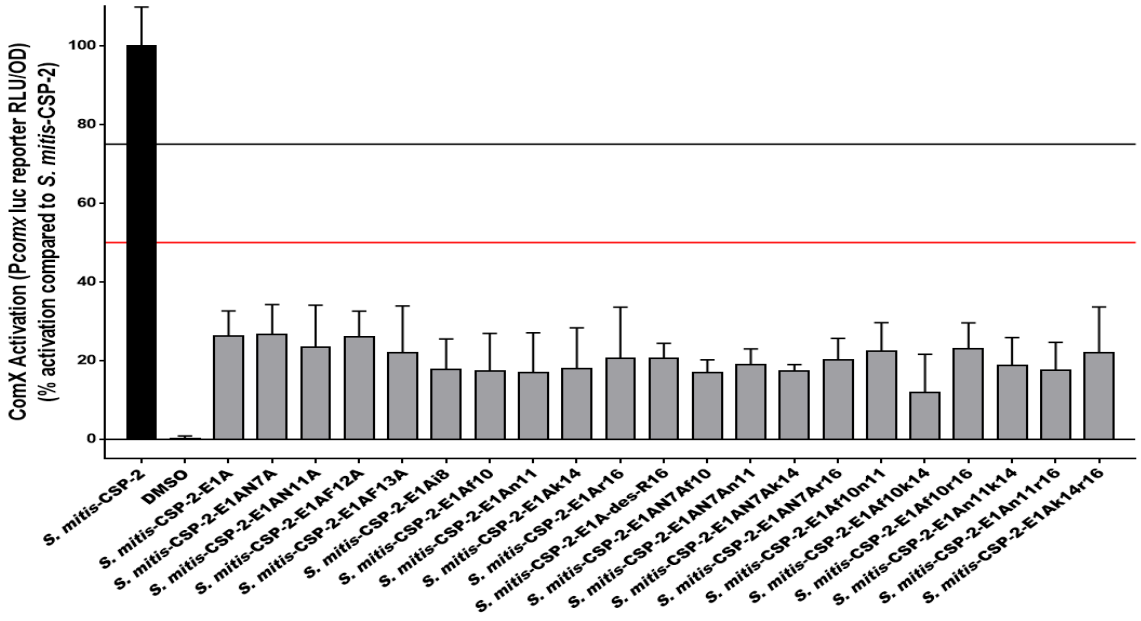


Figure S-7. Primary agonism screening assay data for the *S. mitis*-CSP-2-E1A modification analogs. None of the peptides exhibited activation of the ComD receptor and peptides that exhibited less than 50% activation were evaluated as potential competitive inhibitors.

Antagonism assays were performed at 10 μM concentration of peptides against 1 μM concentration of *S. mitis*-CSP-2. *S. mitis*-CSP-2 (1 μM) was used as the positive control (100%) while DMSO as the negative control (0%). For all figures, *comX* expression is presented as relative luminescence units (RLU) divided by measured cultural optical density (OD). Percent (%) ComD activation was measured by normalizing the RLU/OD value obtained for each peptide to that of the native *S. mitis*-CSP-2. The results are averages from three replicates and representative of three independent experiments. Error bars represent the standard error of the mean of nine values.

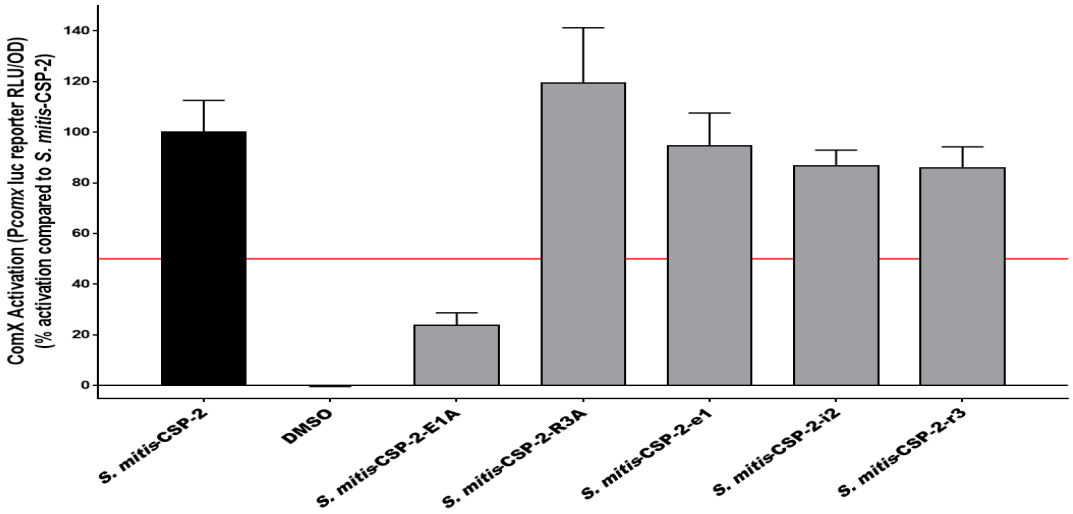


Figure S-8. Primary antagonism screening assay data for the *S. mitis*-CSP-2 alanine scan and D-amino acid scan analogs. *S. mitis*-CSP-2-E1A, which exhibited less than 50% activation, was further evaluated to determine its IC_{50} value.

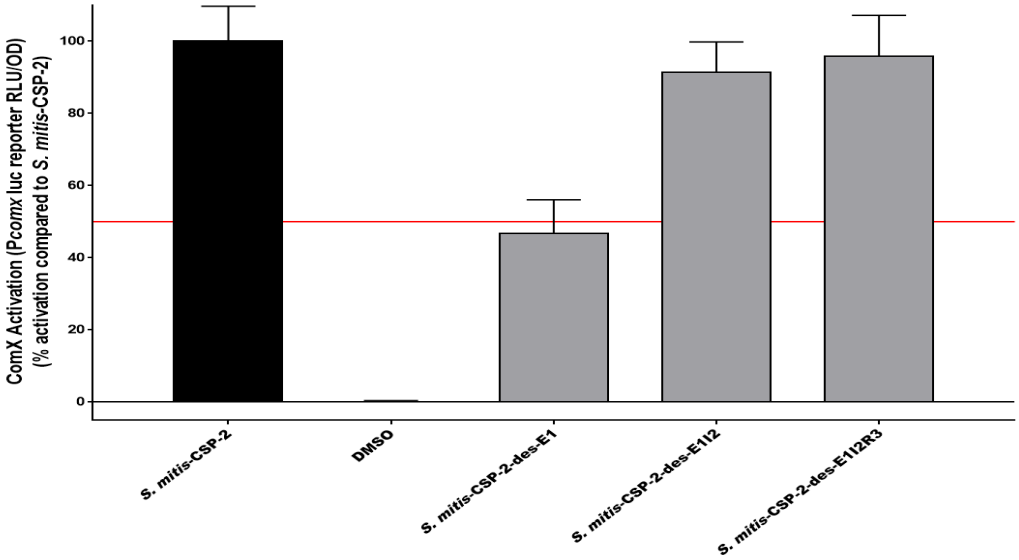


Figure S-9. Primary antagonism screening assay data for the *S. mitis*-CSP-2 truncated analogs. *S. mitis*-CSP-2-des-E1, which exhibited less than 50% activation was further evaluated to determine its IC_{50} value.

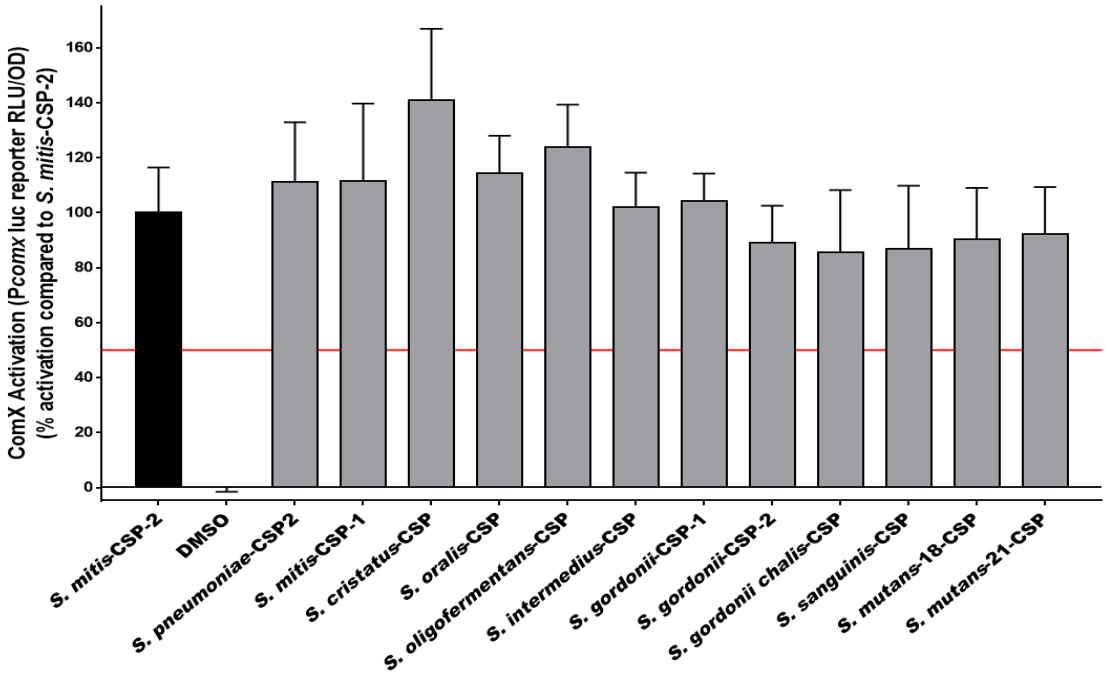


Figure S-10. Primary antagonism screening assay data for the synthetic *Streptococci* native CSP pheromones.[1] None of the peptides exhibited inhibition of the ComD receptor.

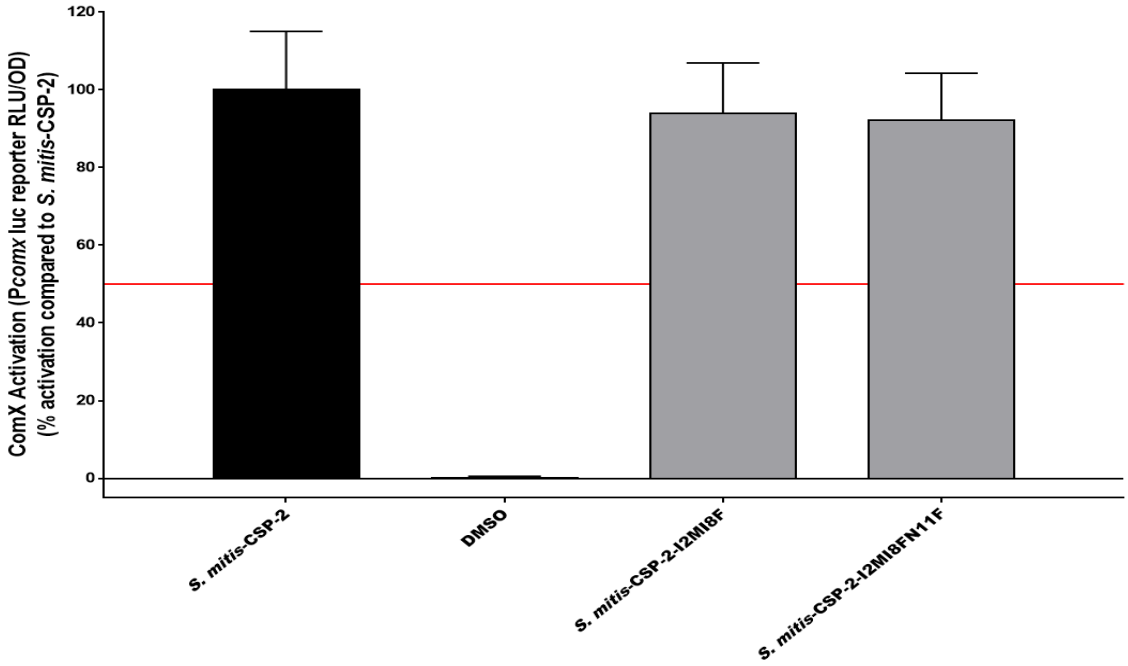


Figure S-11. Primary antagonism screening assay data for the select activators from previous studies.[1] None of the peptides exhibited inhibition of the ComD receptor.

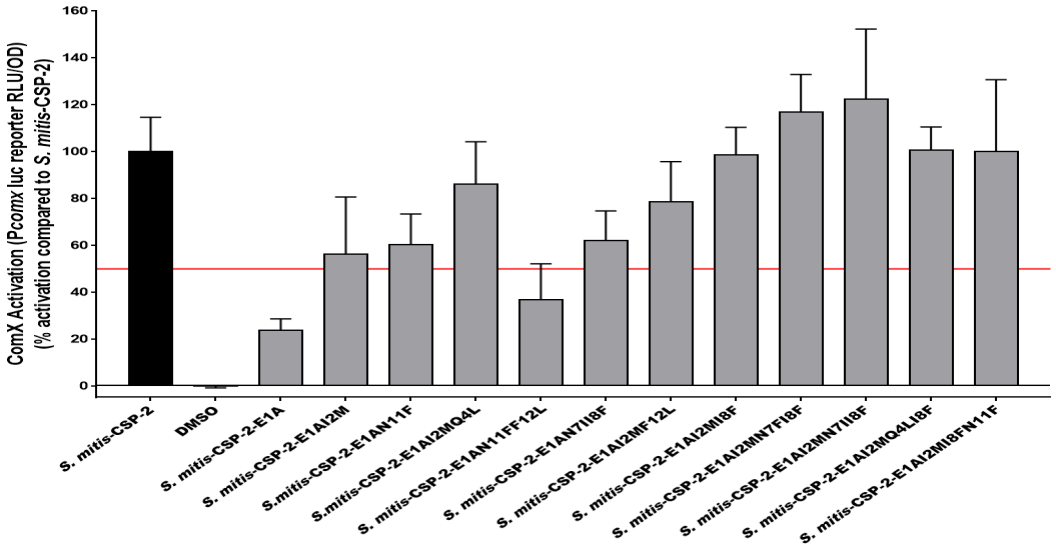


Figure S-12. Primary antagonism screening assay data for the *S. mitis*-CSP-2-E1A mutant analogs from previous studies.[1] Peptides that exhibited less than 50% activation were further evaluated to determine their IC_{50} values.

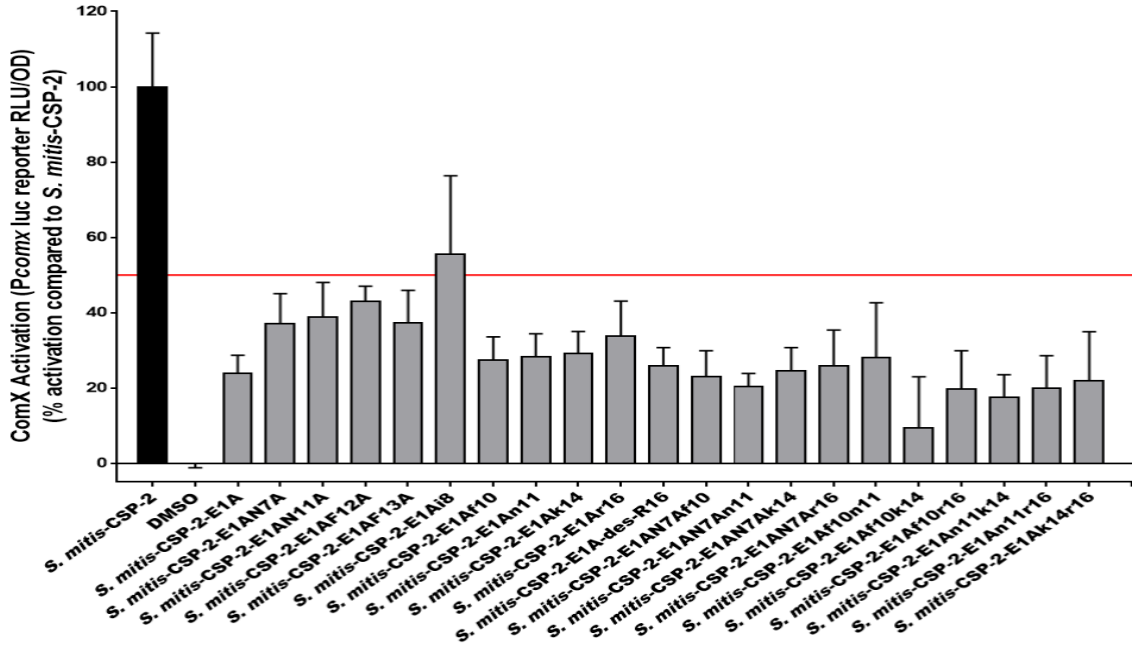
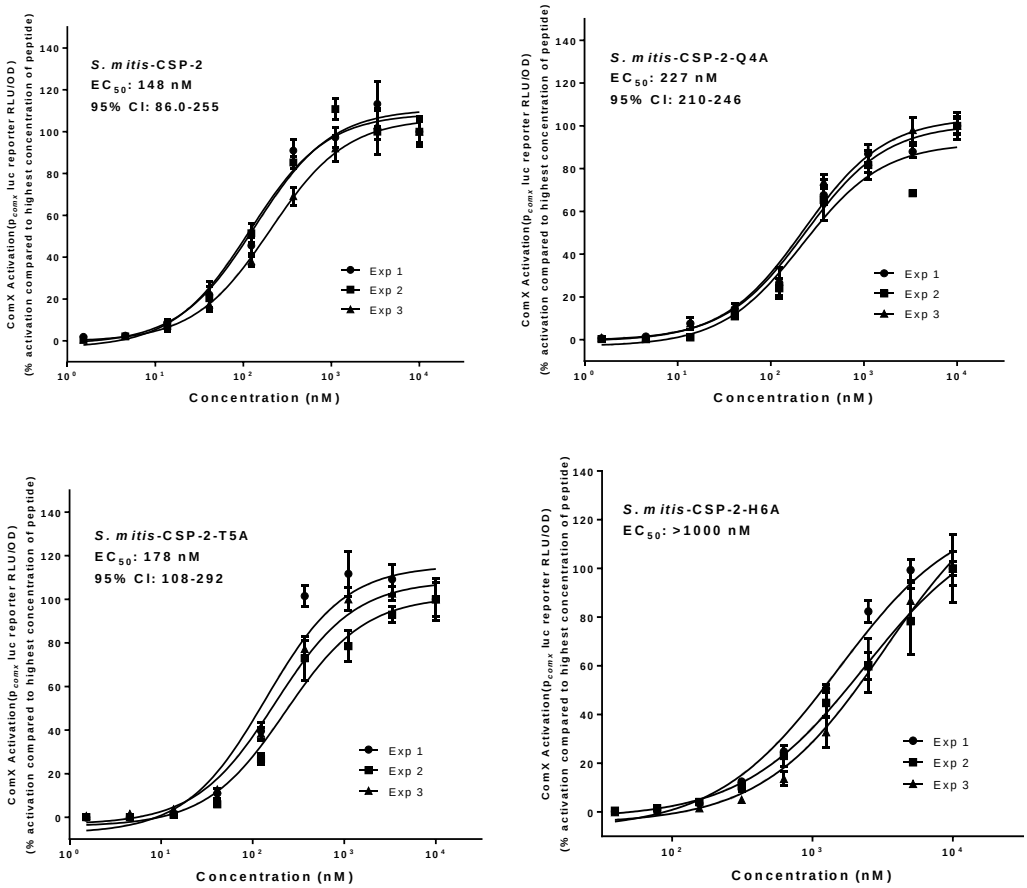


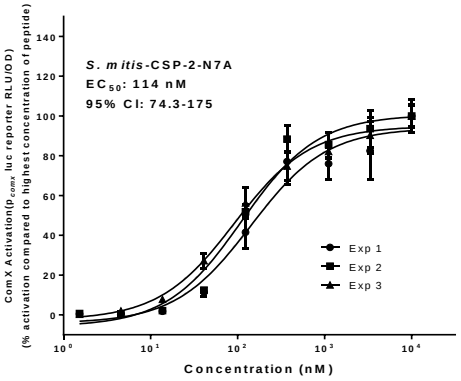
Figure S-13. Primary antagonism screening assay data for the *S. mitis*-CSP-2-E1A modification analogs. Peptides that exhibited less than 50% activation were further evaluated to determine their IC_{50} values.

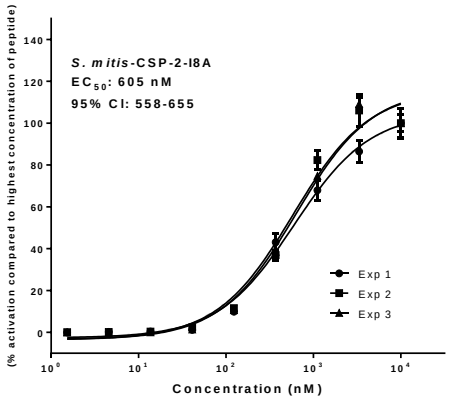
Agonism and antagonism dose response curves

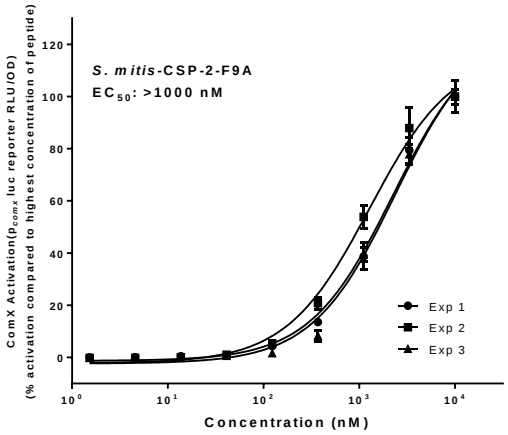
Synthetic CSP analogs were screened to determine their EC_{50} or IC_{50} values over varying concentrations in the *S. mitis comX* reporter type strain, $ATCC_{49456}$ P_{comX} luc::spc. Each dose response experiment was performed in triplicate on three separate occasions (i.e., experiments (Exp.) #1-3; shown for each peptide below). Error bars indicate standard error of the mean of triplicate values. In each plot, the peptide as well as its EC_{50} or IC_{50} value (in nM) and 95% confidence interval (95% CI) values (in nM), are indicated at top left.

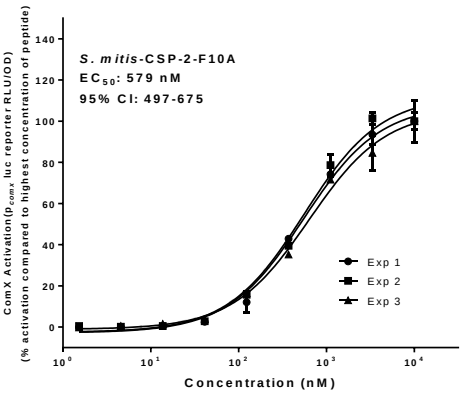
Activation dose response curves

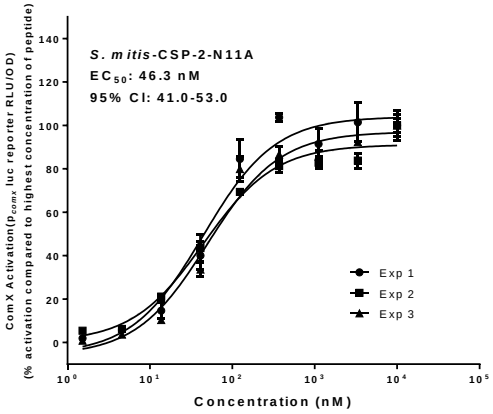


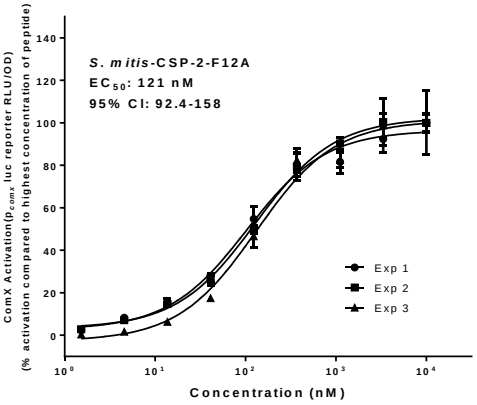

S. mitis-CSP-2-N7A
EC50: 114 nM
95% CI: 74.3-175
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

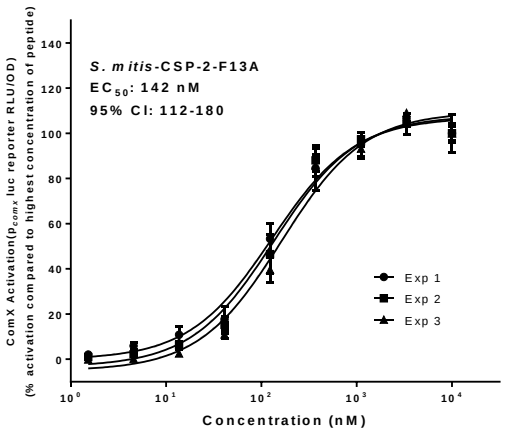

S. mitis-CSP-2-I8A
EC50: 605 nM
95% CI: 558-655
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

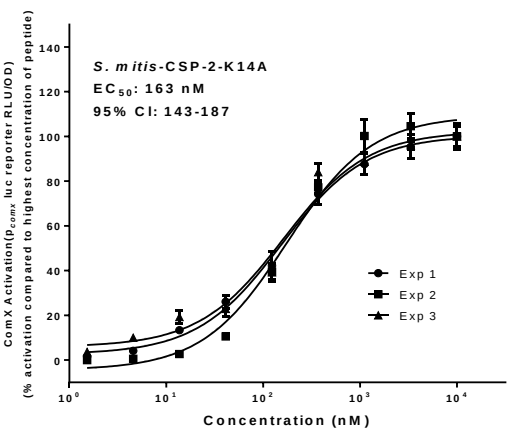

S. mitis-CSP-2-F9A
EC50: >1000 nM
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

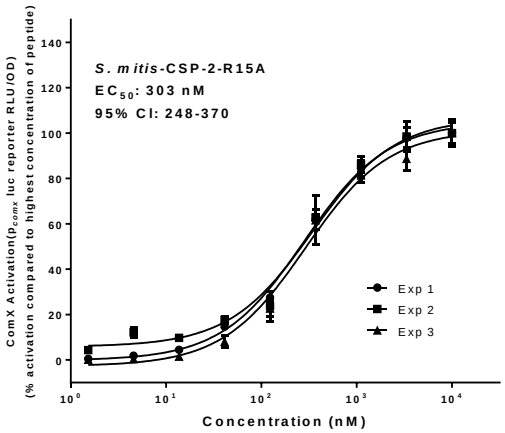

S. mitis-CSP-2-F10A
EC50: 579 nM
95% CI: 497-675
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

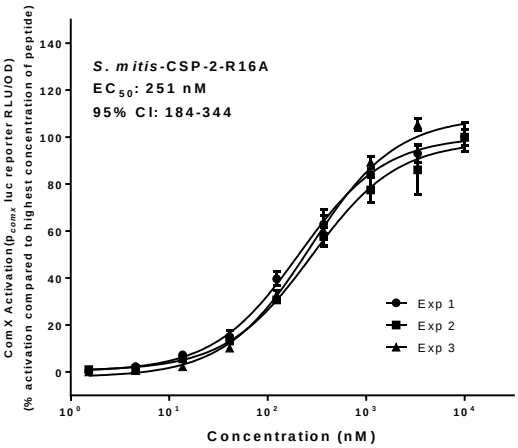

S. mitis-CSP-2-N11A
EC50: 46.3 nM
95% CI: 41.0-53.0
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

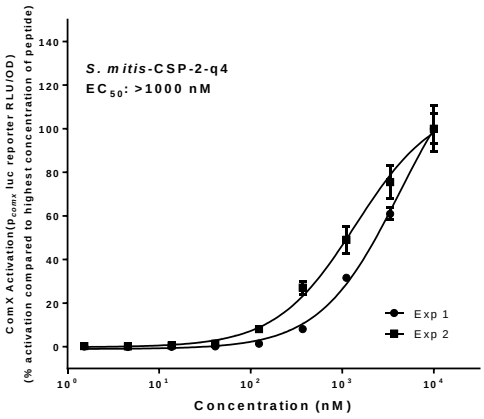

S. mitis-CSP-2-F12A
EC50: 121 nM
95% CI: 92.4-158
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

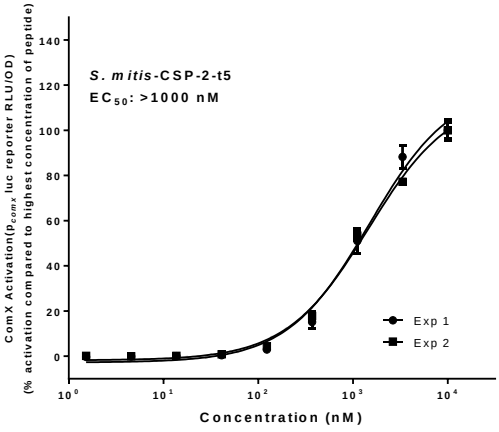

S. mitis-CSP-2-F13A
EC50: 142 nM
95% CI: 112-180
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

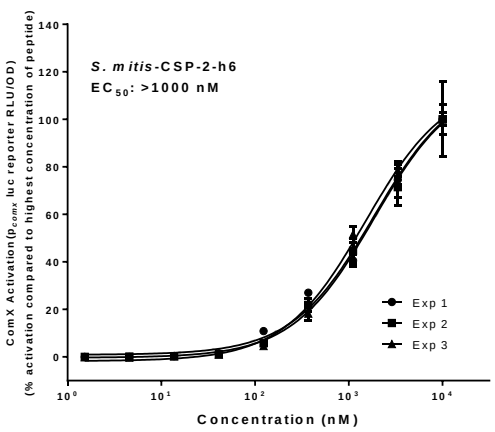

S. mitis-CSP-2-K14A
EC50: 163 nM
95% CI: 143-187
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

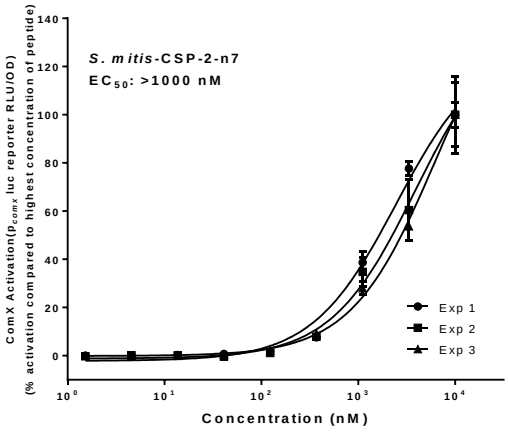

S. mitis-CSP-2-R15A
EC50: 303 nM
95% CI: 248-370
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

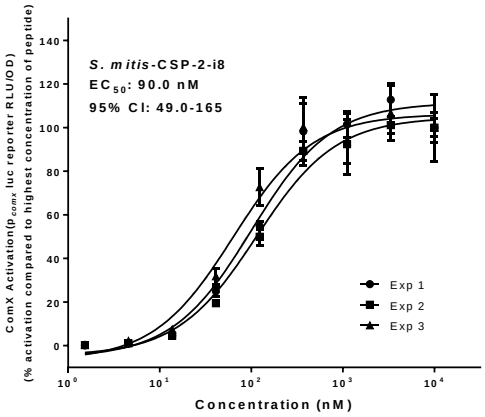

S. mitis-CSP-2-R16A
EC50: 251 nM
95% CI: 184-344
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

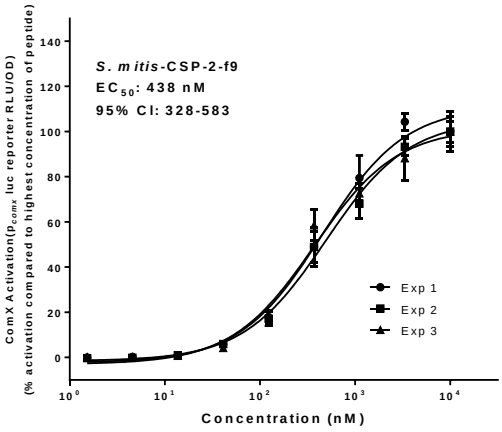

S. mitis-CSP-2-q4
EC50: >1000 nM
Exp 1
Exp 2
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

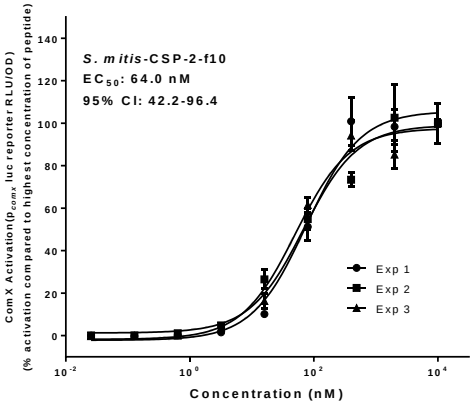

S. mitis-CSP-2-t5
EC50: >1000 nM
Exp 1
Exp 2
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

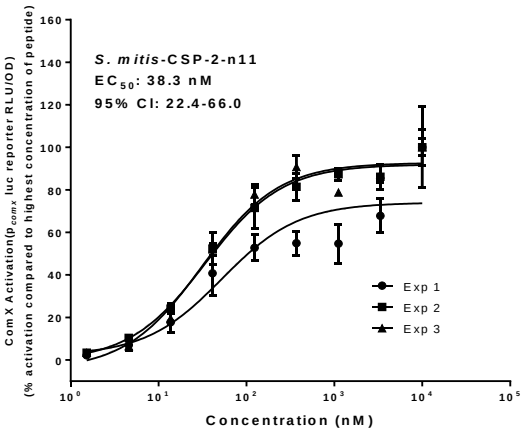

S. mitis-CSP-2-h6
EC50: >1000 nM
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

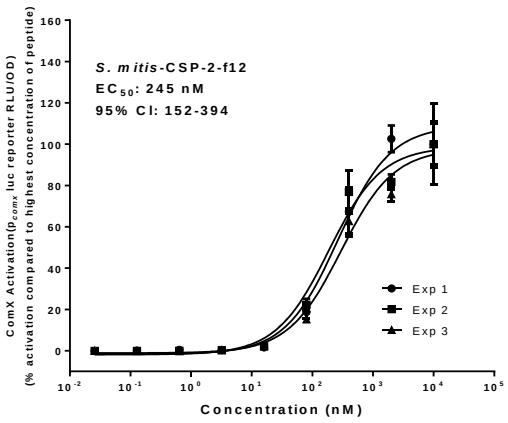

S. mitis-CSP-2-n7
EC50: >1000 nM
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

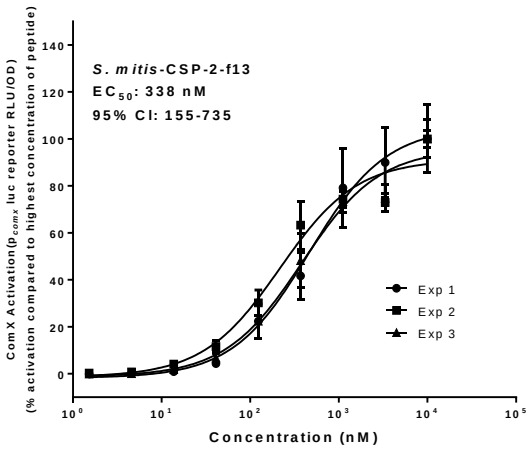

S. mitis-CSP-2-i8
EC50: 90.0 nM
95% CI: 49.0-165
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

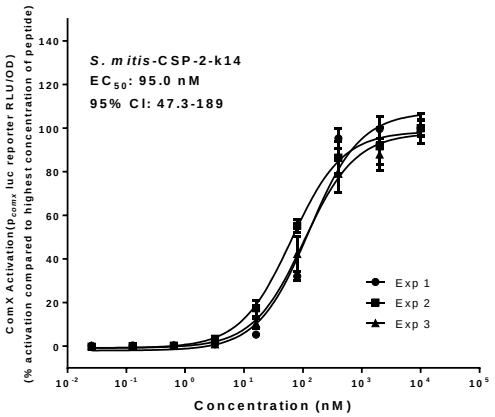

S. mitis-CSP-2-f9
EC50: 438 nM
95% CI: 328-583
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

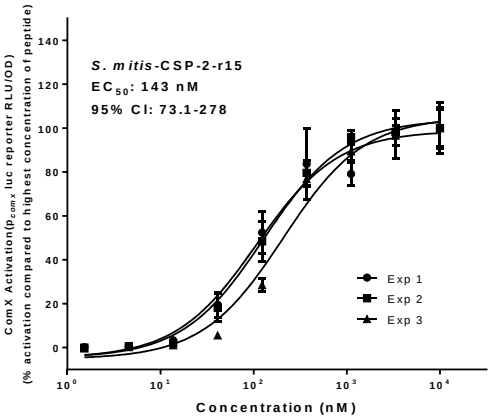

S. mitis-CSP-2-f10
EC50: 64.0 nM
95% CI: 42.2-96.4
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

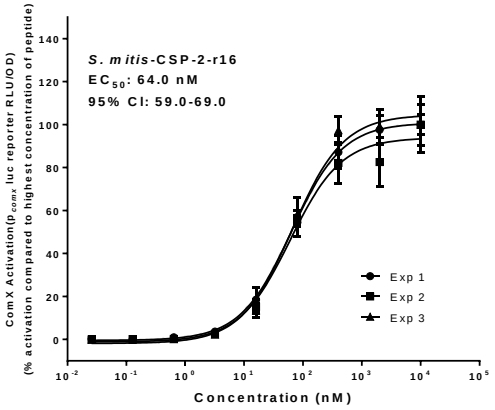

S. mitis-CSP-2-n11
EC50: 38.3 nM
95% CI: 22.4-66.0
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

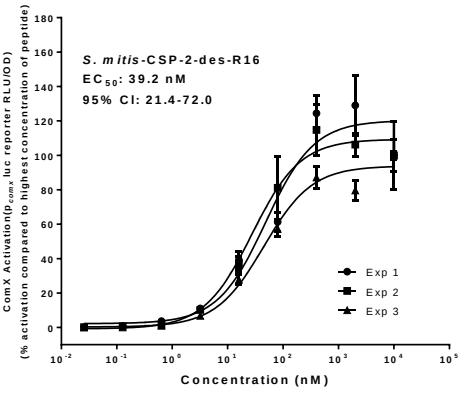

S. mitis-CSP-2-f12
EC50: 245 nM
95% CI: 152-394
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

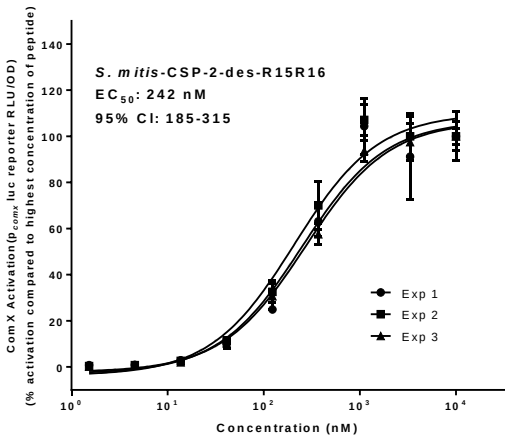

S. mitis-CSP-2-f13
EC50: 338 nM
95% CI: 155-735
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

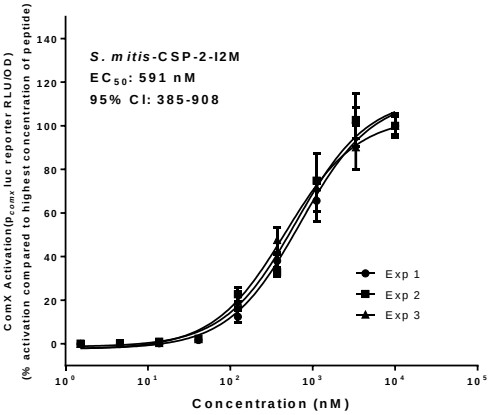

S. mitis-CSP-2-k14
EC50: 95.0 nM
95% CI: 47.3-189
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

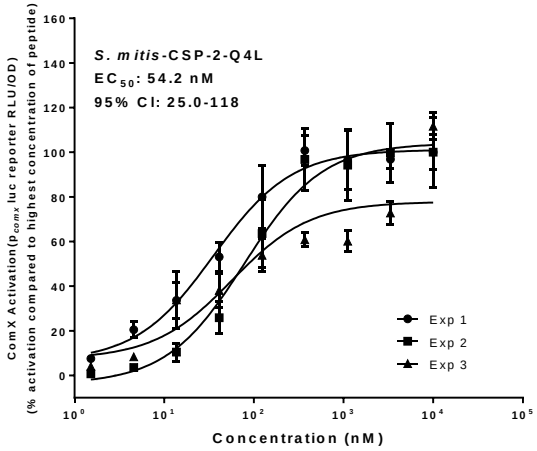

S. mitis-CSP-2-r15
EC50: 143 nM
95% CI: 73.1-278
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

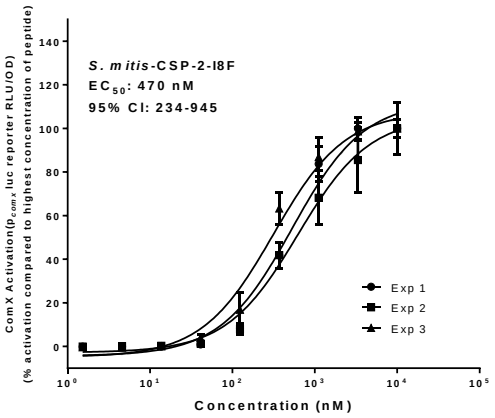

S. mitis-CSP-2-r16
EC50: 64.0 nM
95% CI: 59.0-69.0
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

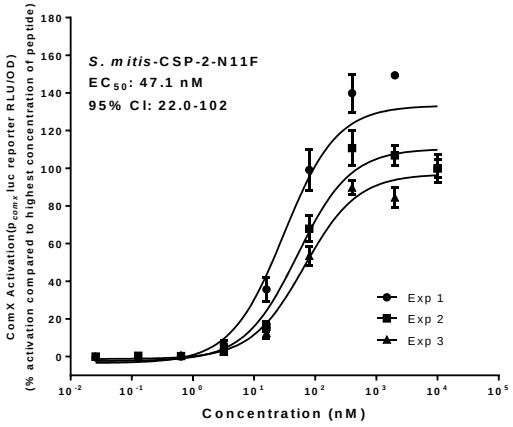

S. mitis-CSP-2-des-R16
EC50: 39.2 nM
95% CI: 21.4-72.0
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)
Exp 1
Exp 2
Exp 3

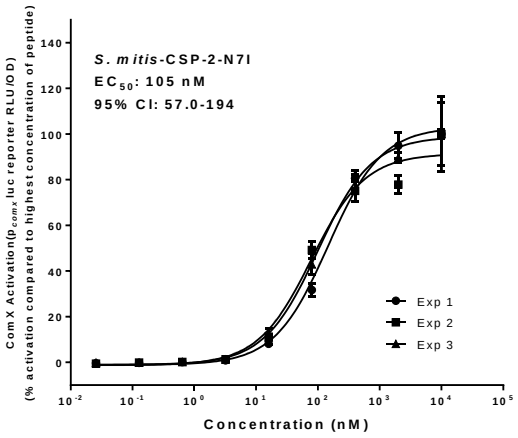

S. mitis-CSP-2-des-R15R16
EC50: 242 nM
95% CI: 185-315
Exp 1
Exp 2
Exp 3
ComX Activation(pcomX luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

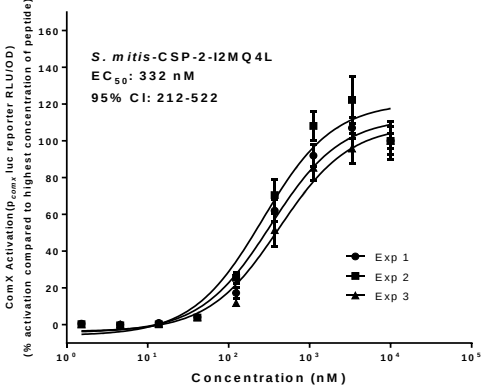

S. mitis-CSP-2-I2M
EC50: 591 nM
95% CI: 385-908
Exp 1
Exp 2
Exp 3
ComX Activation(pcomX luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

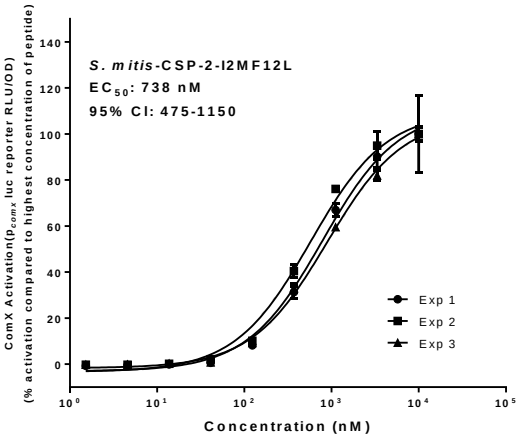

S. mitis-CSP-2-Q4L
EC50: 54.2 nM
95% CI: 25.0-118
Exp 1
Exp 2
Exp 3
ComX Activation(pcomX luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

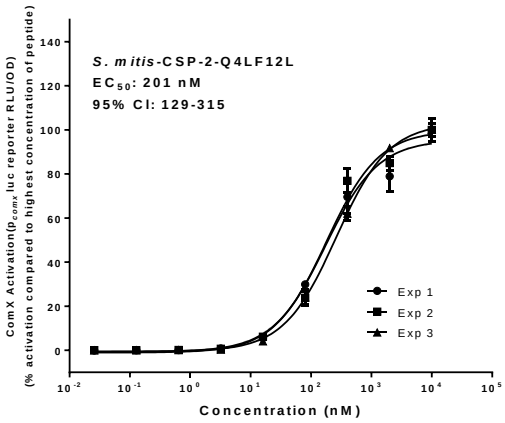

S. mitis-CSP-2-I8F
EC50: 470 nM
95% CI: 234-945
Exp 1
Exp 2
Exp 3
ComX Activation(pcomX luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

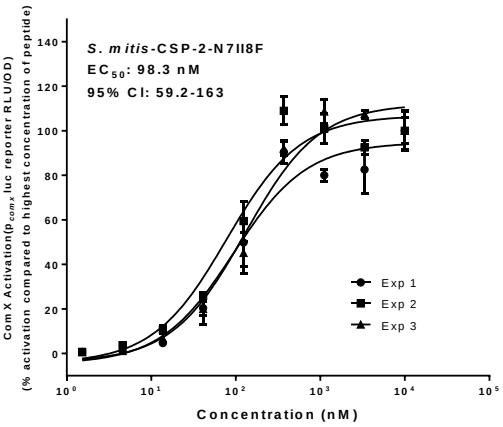

S. mitis-CSP-2-N11F
EC50: 47.1 nM
95% CI: 22.0-102
Exp 1
Exp 2
Exp 3
ComX Activation(pcomX luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

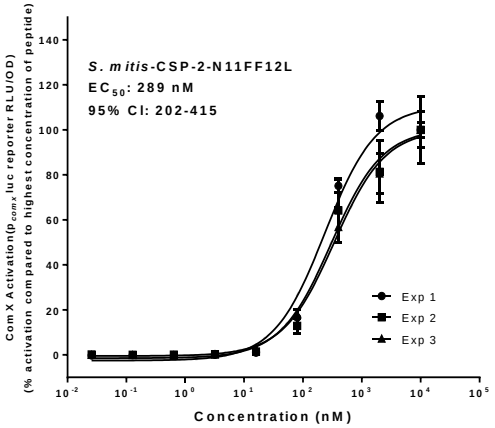

S. mitis-CSP-2-N7I
EC50: 105 nM
95% CI: 57.0-194
Exp 1
Exp 2
Exp 3
ComX Activation(pcomX luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

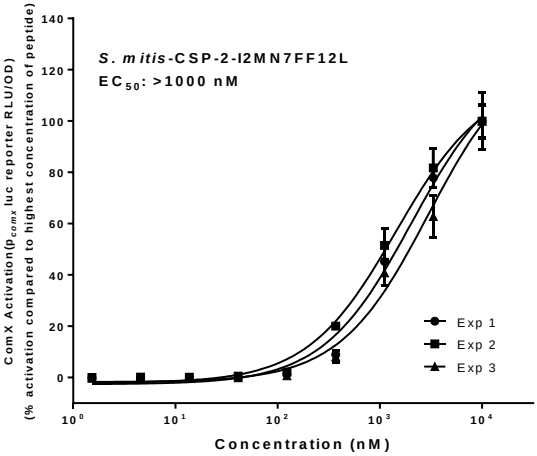

S. mitis-CSP-2-I2MQ4L
EC50: 332 nM
95% CI: 212-522
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

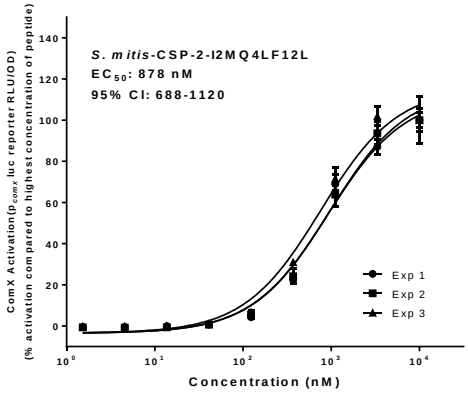

S. mitis-CSP-2-I2MF12L
EC50: 738 nM
95% CI: 475-1150
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

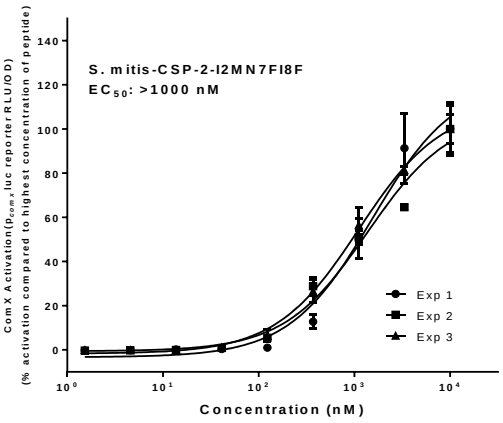

S. mitis-CSP-2-Q4LF12L
EC50: 201 nM
95% CI: 129-315
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

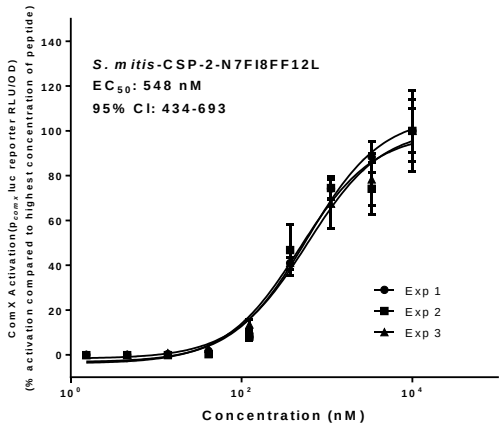

S. mitis-CSP-2-N7I8F
EC50: 98.3 nM
95% CI: 59.2-163
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

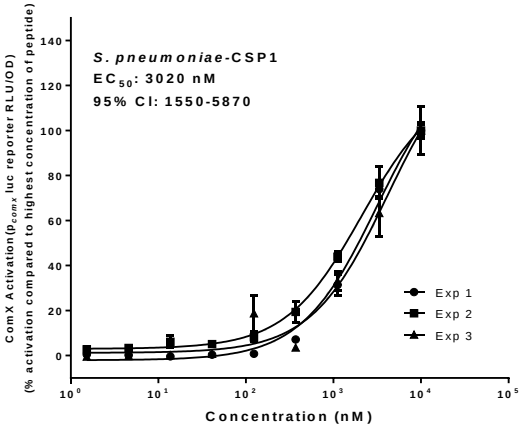

S. mitis-CSP-2-N11FF12L
EC50: 289 nM
95% CI: 202-415
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)


S. mitis-CSP-2-I2MN7FF12L
EC50: >1000 nM
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to highest concentration of peptide)
Concentration (nM)

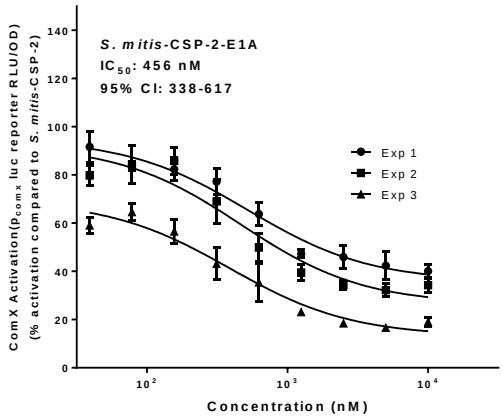


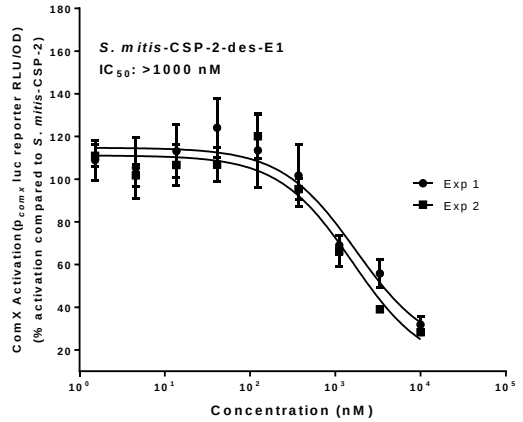


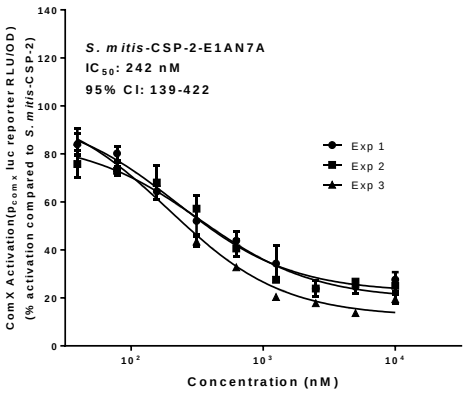


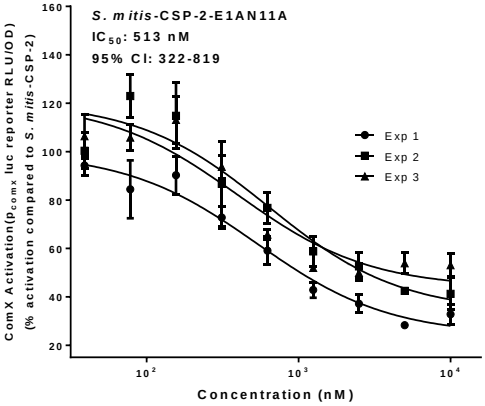


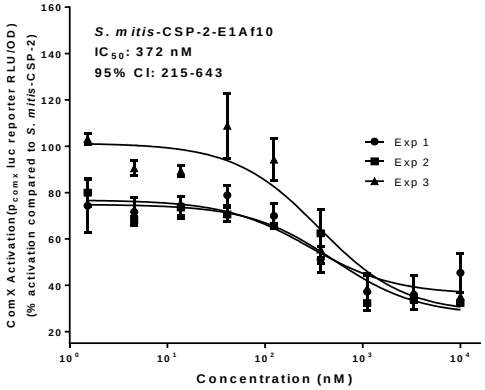
Inhibition dose response curves

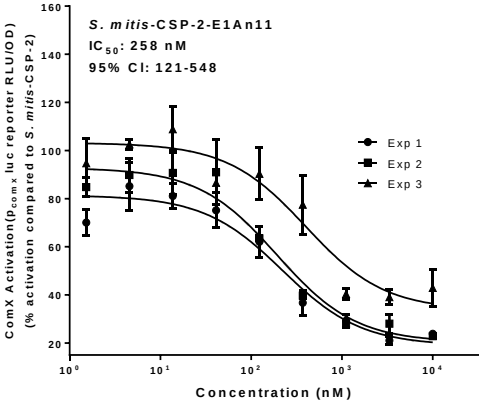


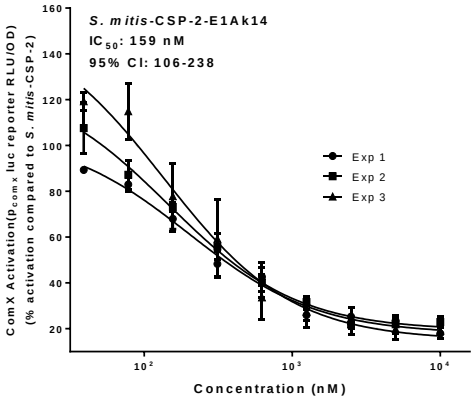


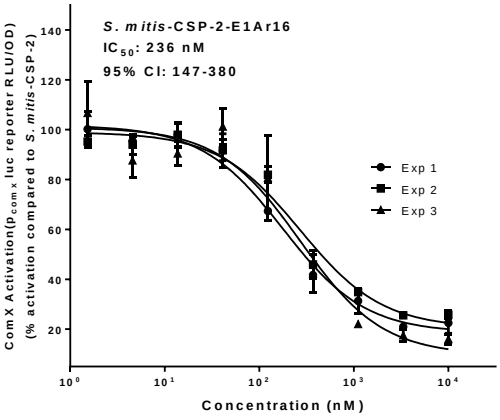

S. mitis-CSP-2-E1AN7A
IC50: 242 nM
95% CI: 139-422
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD) (% activation compared to S. mitis-CSP-2)
Concentration (nM)

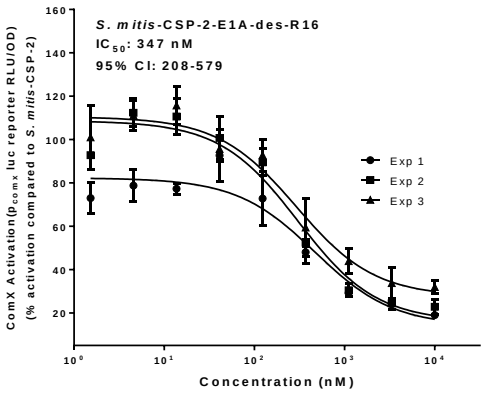

S. mitis-CSP-2-E1AN11A
IC50: 513 nM
95% CI: 322-819
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD) (% activation compared to S. mitis-CSP-2)
Concentration (nM)

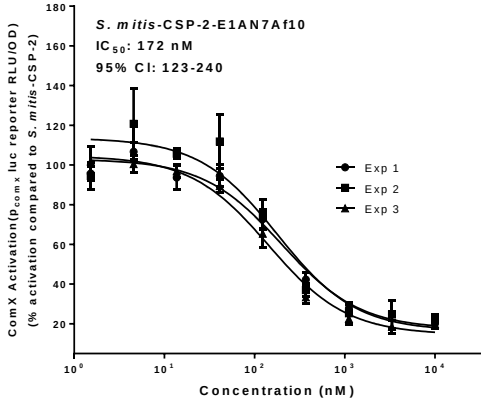

S. mitis-CSP-2-E1Af10
IC50: 372 nM
95% CI: 215-643
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD) (% activation compared to S. mitis-CSP-2)
Concentration (nM)

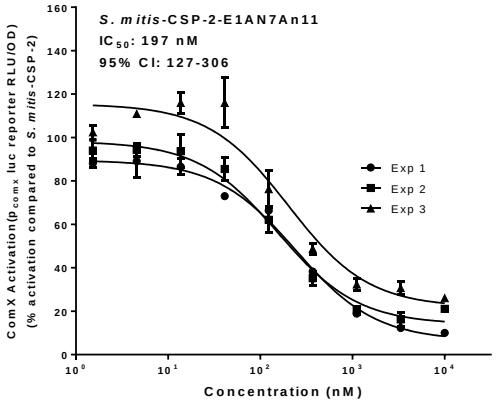

S. mitis-CSP-2-E1An11
IC50: 258 nM
95% CI: 121-548
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD) (% activation compared to S. mitis-CSP-2)
Concentration (nM)

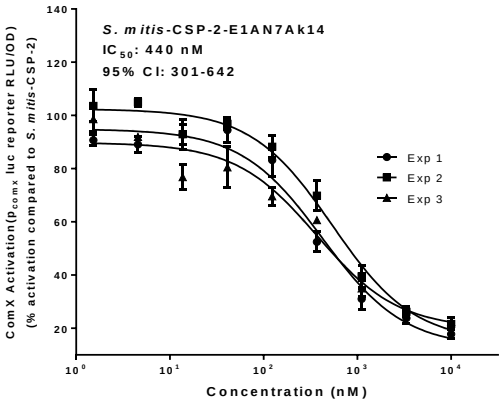

S. mitis-CSP-2-E1Ak14
IC50: 159 nM
95% CI: 106-238
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD) (% activation compared to S. mitis-CSP-2)
Concentration (nM)

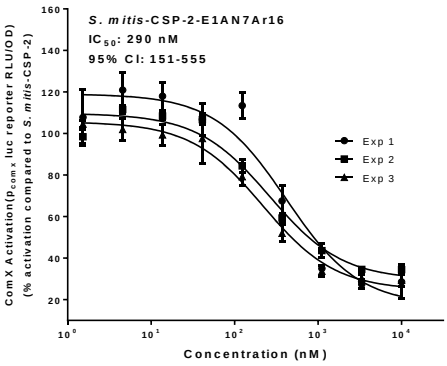

S. mitis-CSP-2-E1Ar16
IC50: 236 nM
95% CI: 147-380
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD) (% activation compared to S. mitis-CSP-2)
Concentration (nM)

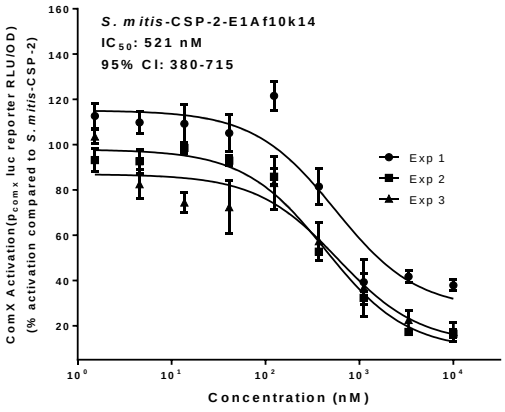

S. mitis-CSP-2-E1A-des-R16
IC50: 347 nM
95% CI: 208-579
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to S. mitis-CSP-2)
Concentration (nM)

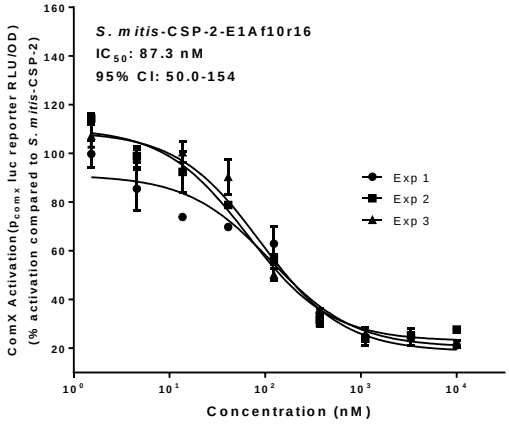

S. mitis-CSP-2-E1AN7Af10
IC50: 172 nM
95% CI: 123-240
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to S. mitis-CSP-2)
Concentration (nM)

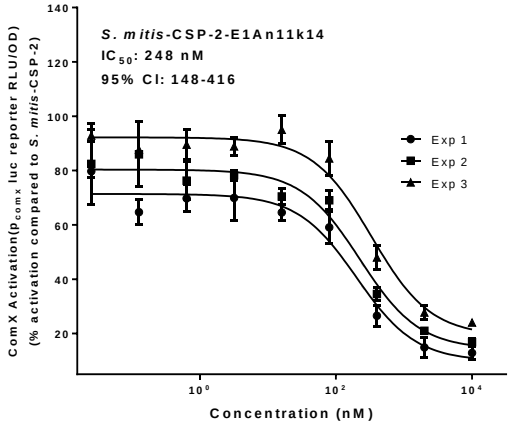

S. mitis-CSP-2-E1AN7An11
IC50: 197 nM
95% CI: 127-306
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to S. mitis-CSP-2)
Concentration (nM)

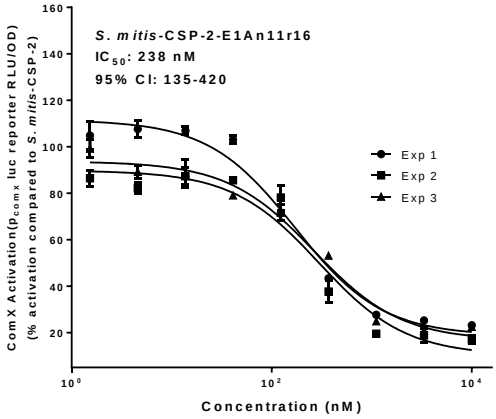

S. mitis-CSP-2-E1AN7Ak14
IC50: 440 nM
95% CI: 301-642
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to S. mitis-CSP-2)
Concentration (nM)

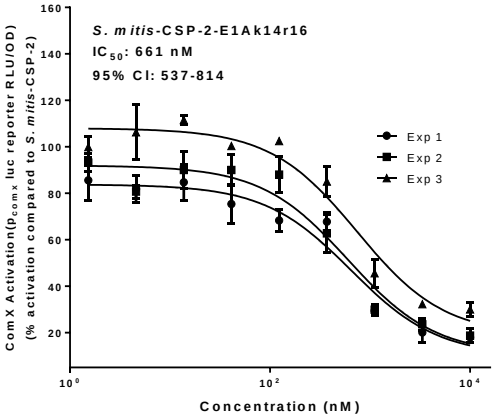

S. mitis-CSP-2-E1AN7Ar16
IC50: 290 nM
95% CI: 151-555
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to S. mitis-CSP-2)
Concentration (nM)


S. mitis-CSP-2-E1Af10k14
IC50: 521 nM
95% CI: 380-715
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to S. mitis-CSP-2)
Concentration (nM)


S. mitis-CSP-2-E1Af10r16
IC50: 87.3 nM
95% CI: 50.0-154
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to S. mitis-CSP-2)
Concentration (nM)


S. mitis-CSP-2-E1An11k14
IC50: 248 nM
95% CI: 148-416
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to S. mitis-CSP-2)
Concentration (nM)


S. mitis-CSP-2-E1An11r16
IC50: 238 nM
95% CI: 135-420
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to S. mitis-CSP-2)
Concentration (nM)


S. mitis-CSP-2-E1Ak14r16
IC50: 661 nM
95% CI: 537-814
Exp 1
Exp 2
Exp 3
ComX Activation(pcomx luc reporter RLU/OD)
(% activation compared to S. mitis-CSP-2)
Concentration (nM)

Circular dichroism (CD) spectra

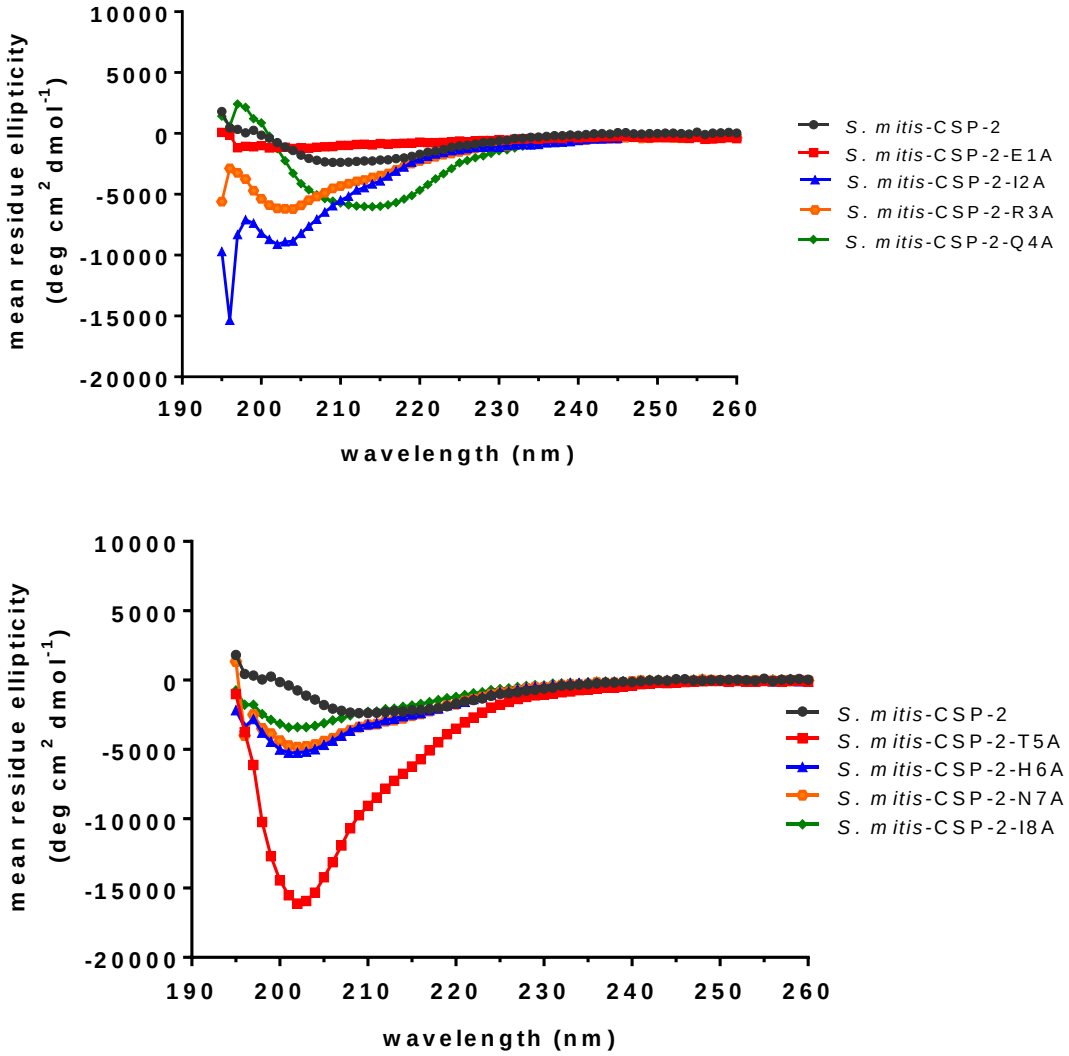


Figure S-14. CD spectra of the *S. mitis*-CSP-2 alanine scan library in aqueous solution (PBS, pH 7.4). All the measurements were performed with a peptide concentration of 200 µM. *S. mitis*-CSP-2 was added as a control. Most of the alanine screen analogs were unstructured exhibiting a random coil pattern, with the exception of *S. mitis*-CSP-2-Q4A that exhibited some β-sheet pattern.

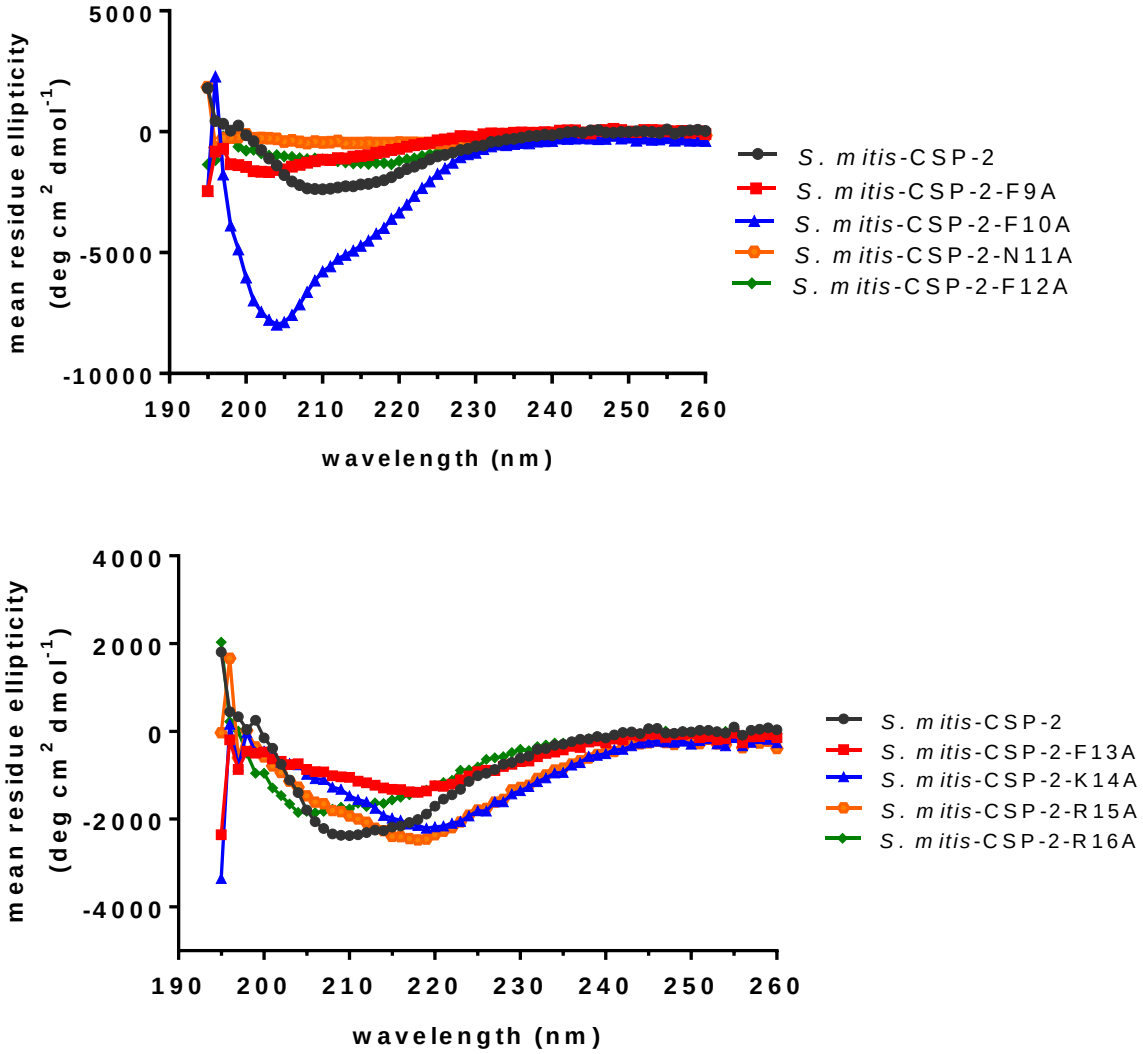


Figure S-15. CD spectra of the *S. mitis*-CSP-2 alanine scan library in aqueous solution (PBS, pH 7.4). All the measurements were performed with a peptide concentration of 200 µM. *S. mitis*-CSP-2 was added as a control. Most of the alanine screen analogs were unstructured exhibiting a random coil pattern, with the exception of *S. mitis*-CSP-2-F13A, *S. mitis*-CSP-2-K14A, and *S. mitis*-CSP-2-R15A that exhibited some β-sheet pattern.

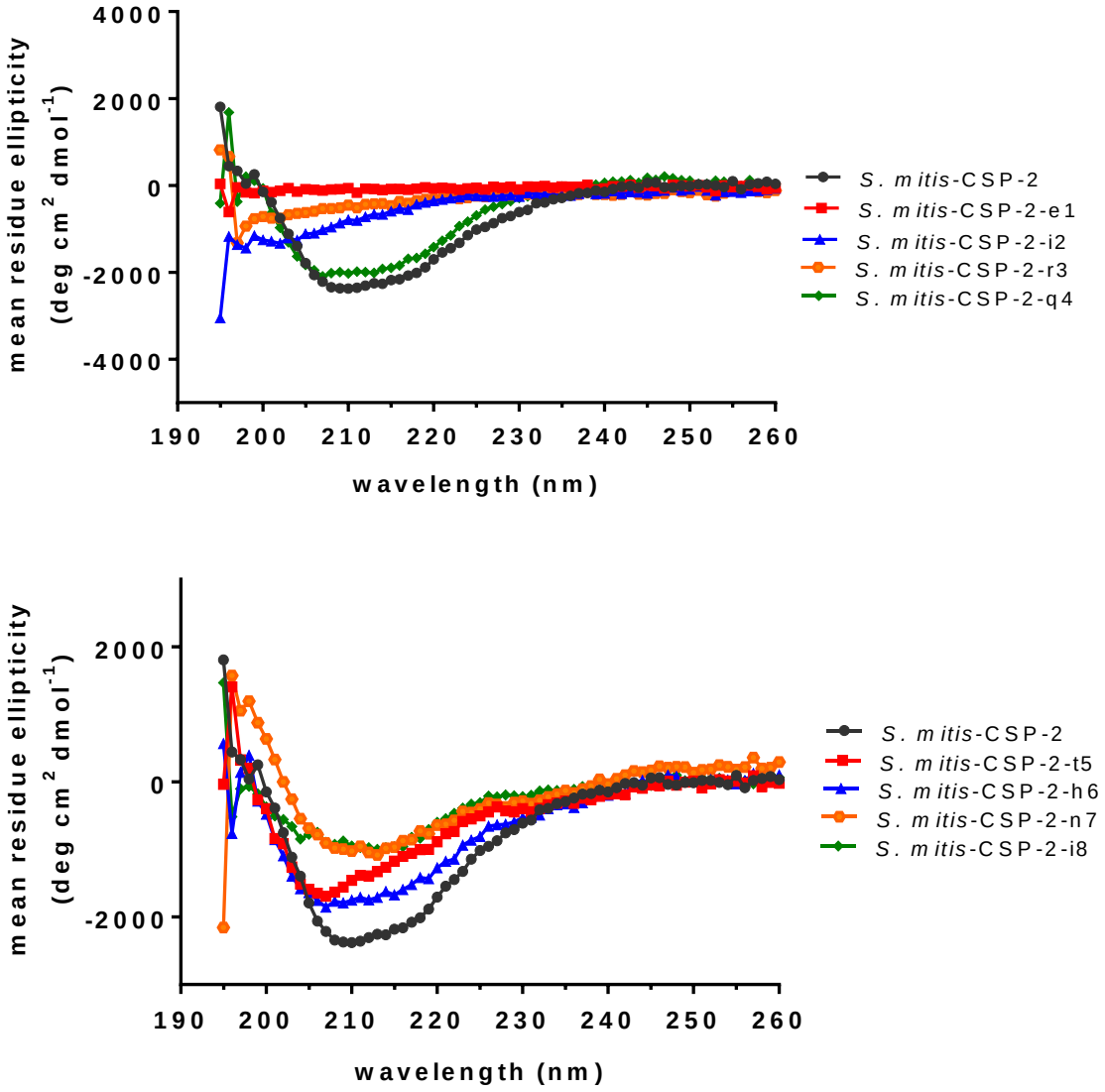


Figure S-16. CD spectra of the *S. mitis*-CSP-2 D-amino acid scan library in aqueous solution (PBS, pH 7.4). All the measurements were performed with a peptide concentration of 200 µM. *S. mitis*-CSP-2 was added as a control. All of the analogs were unstructured exhibiting a random coil pattern.

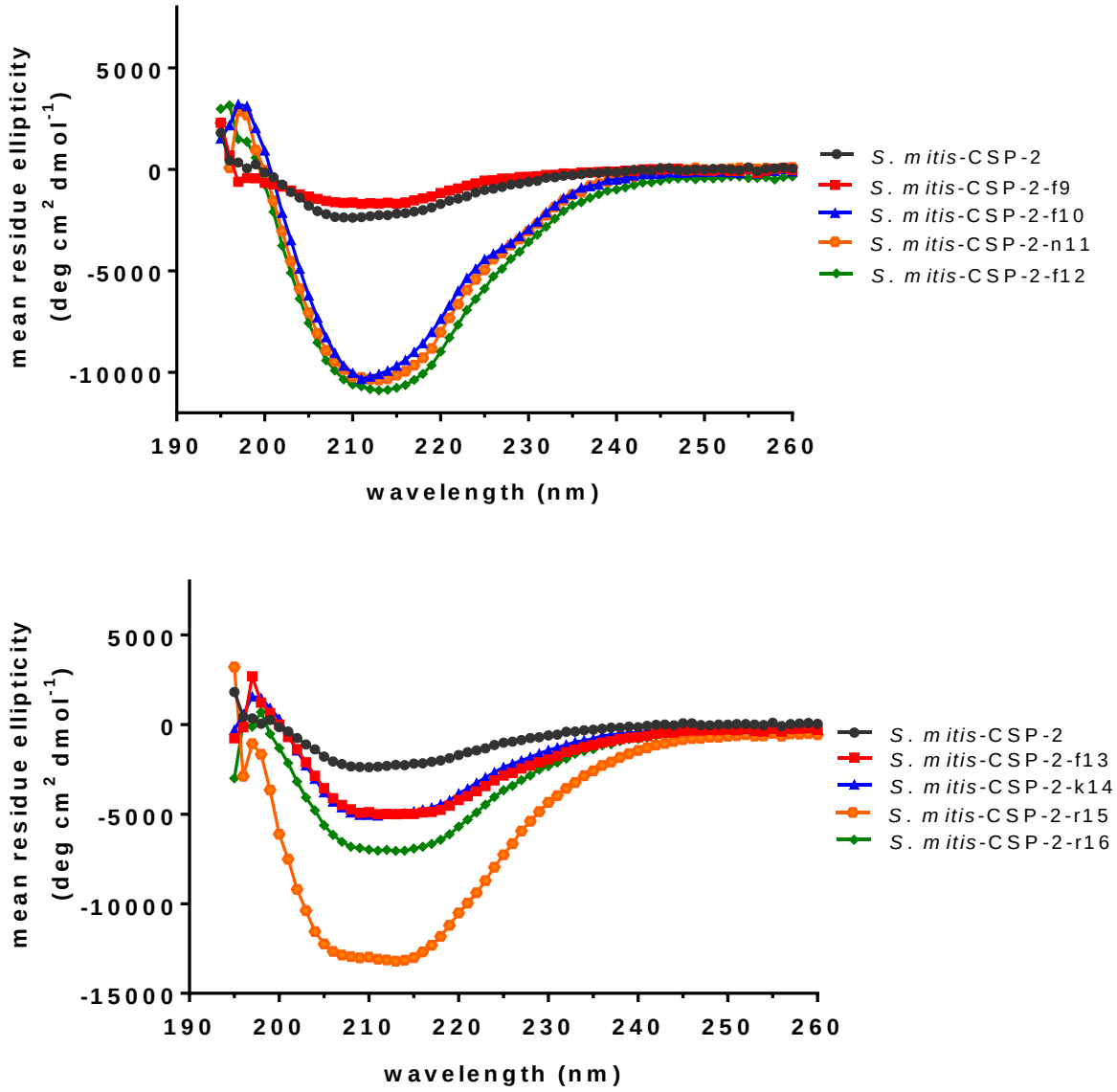


Figure S-17. CD spectra of the *S. mitis*-CSP-2 D-amino acid scan library in aqueous solution (PBS, pH 7.4). All the measurements were performed with a peptide concentration of 200 µM. *S. mitis*-CSP-2 was added as a control. Most of the D-amino acid scan analogs were unstructured exhibiting a random coil pattern, with the exception of *S. mitis*-CSP-2-f10, *S. mitis*-CSP-2-n11, and *S. mitis*-CSP-2-f12 that exhibited some β-sheet pattern.

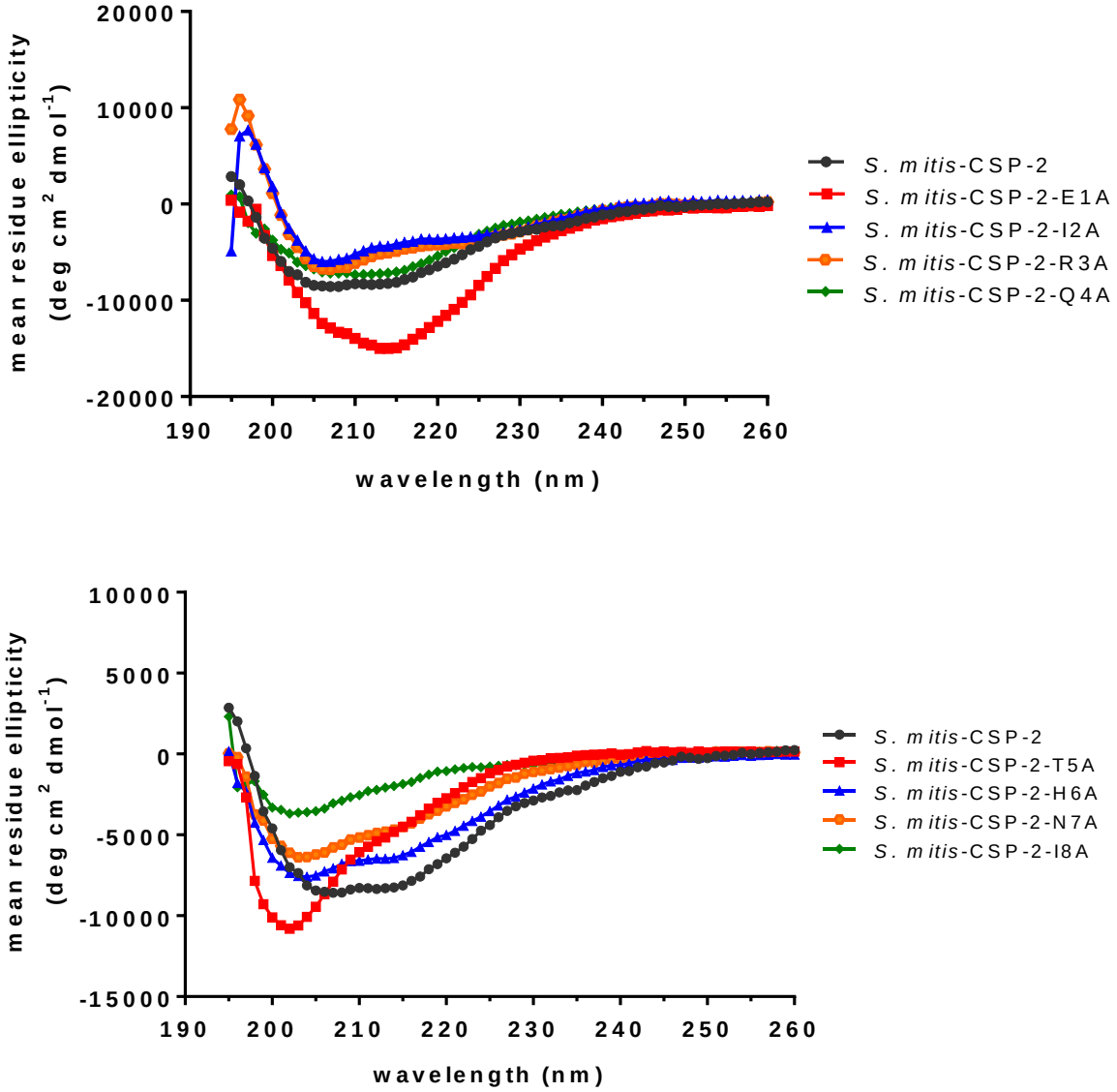


Figure S-18. CD spectra of the *S. mitis*-CSP-2 alanine scan library in membrane mimicking conditions (20% TFE: 80% PBS, pH 7.4). All the measurements were performed with a peptide concentration of 200 µM. *S. mitis*-CSP-2 was added as a control. Most of the alanine screen analogs were unstructured exhibiting a random coil pattern, with the exception of *S. mitis*-CSP-2-E1A that exhibited some β-sheet pattern.

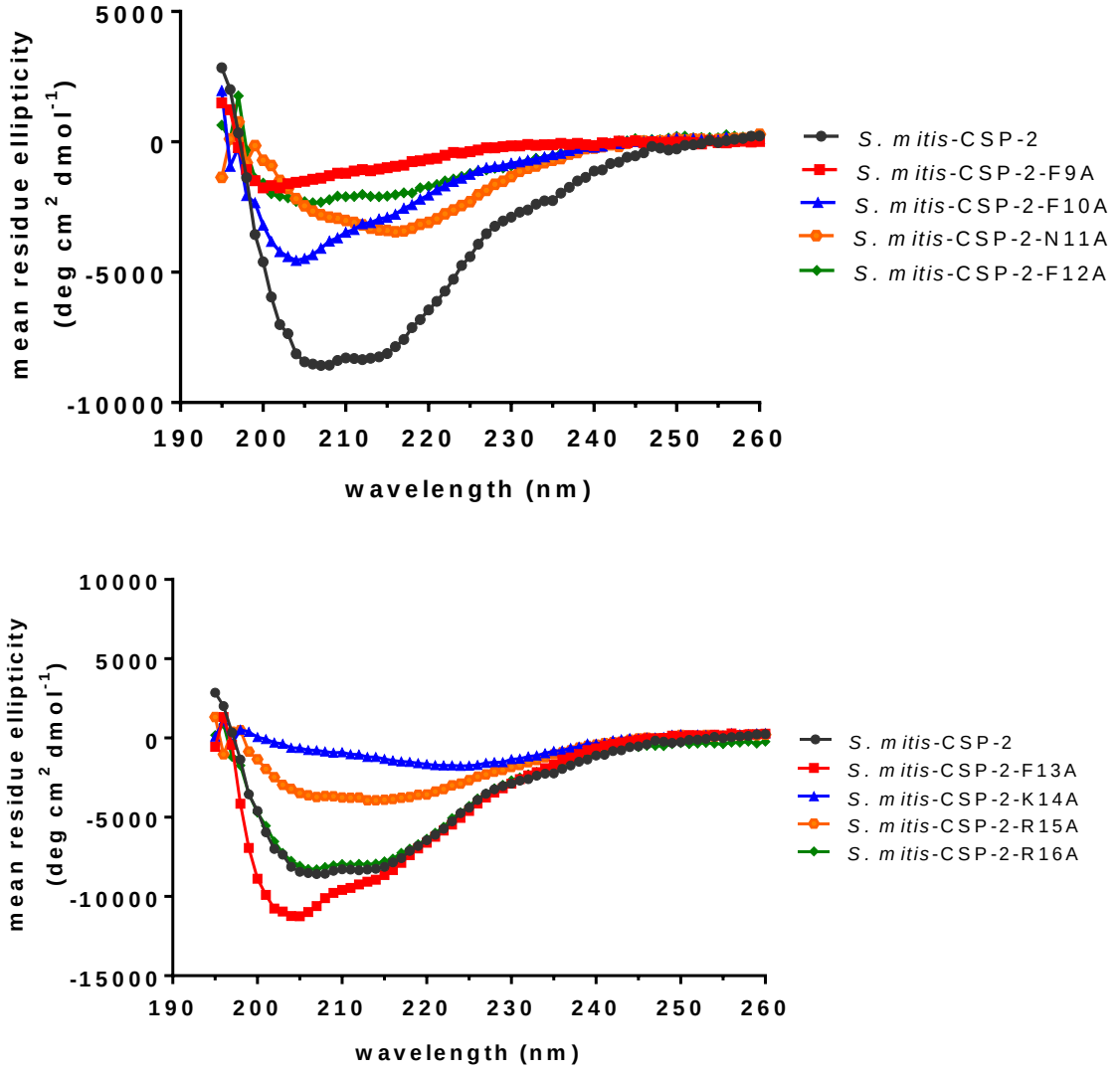


Figure S-19. CD spectra of the *S. mitis*-CSP-2 alanine scan library in membrane mimicking conditions (20% TFE: 80% PBS, pH 7.4). All the measurements were performed with a peptide concentration of 200 µM. *S. mitis*-CSP-2 was added as a control. Most of the alanine screen analogs were unstructured exhibiting a random coil pattern, with the exception of *S. mitis*-CSP-2-N11A that exhibited some β-sheet pattern.

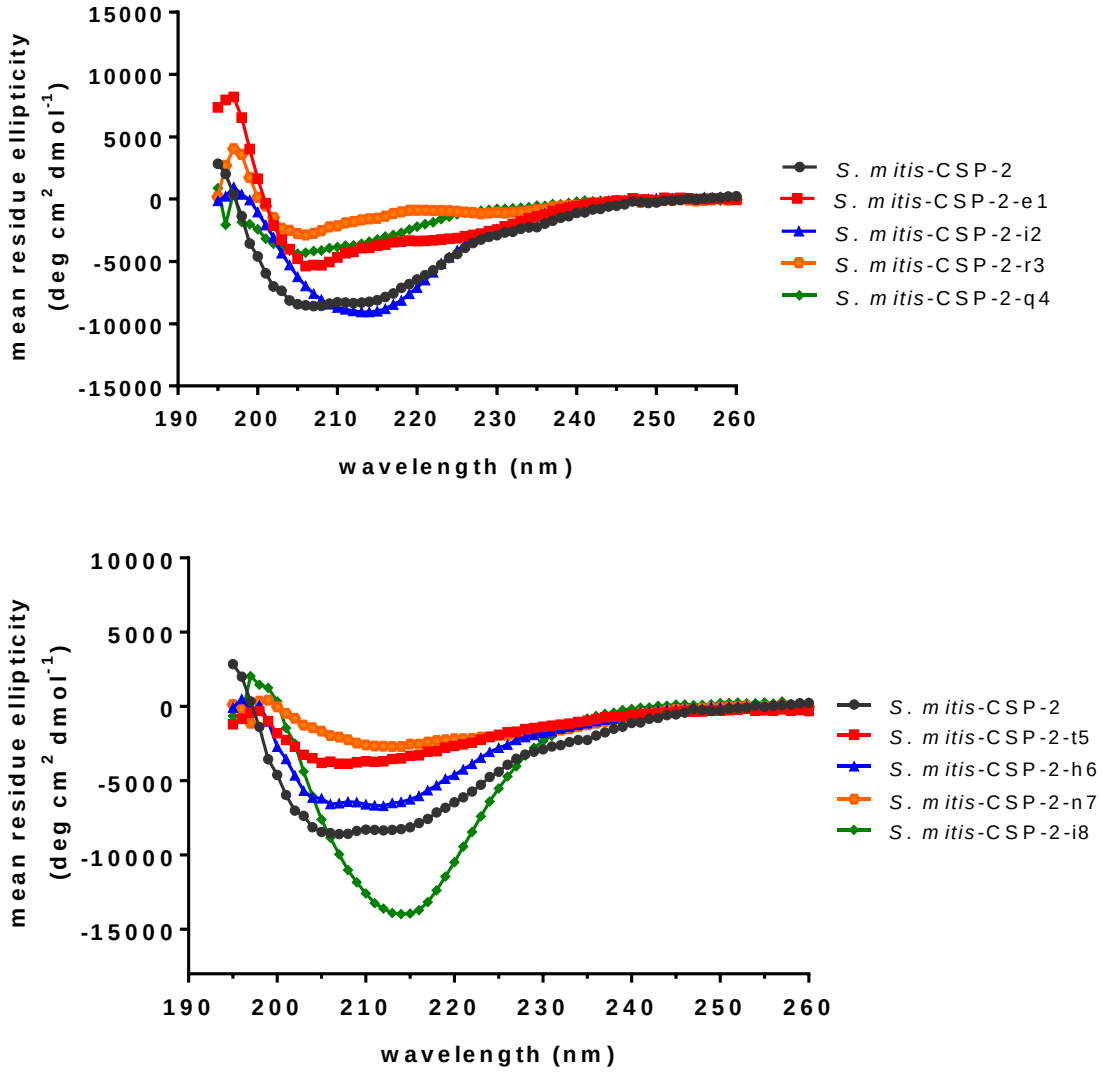


Figure S-20. CD spectra of the *S. mitis*-CSP-2 D-amino acid scan library in membrane mimicking conditions (20% TFE: 80% PBS, pH 7.4). All the measurements were performed with a peptide concentration of 200 µM. *S. mitis*-CSP-2 was added as a control. Most of the D-amino acid scan analogs were unstructured exhibiting a random coil pattern, with the exception of *S. mitis*-CSP-2-i2, and *S. mitis*-CSP-2-i8 that exhibited some β-sheet pattern.

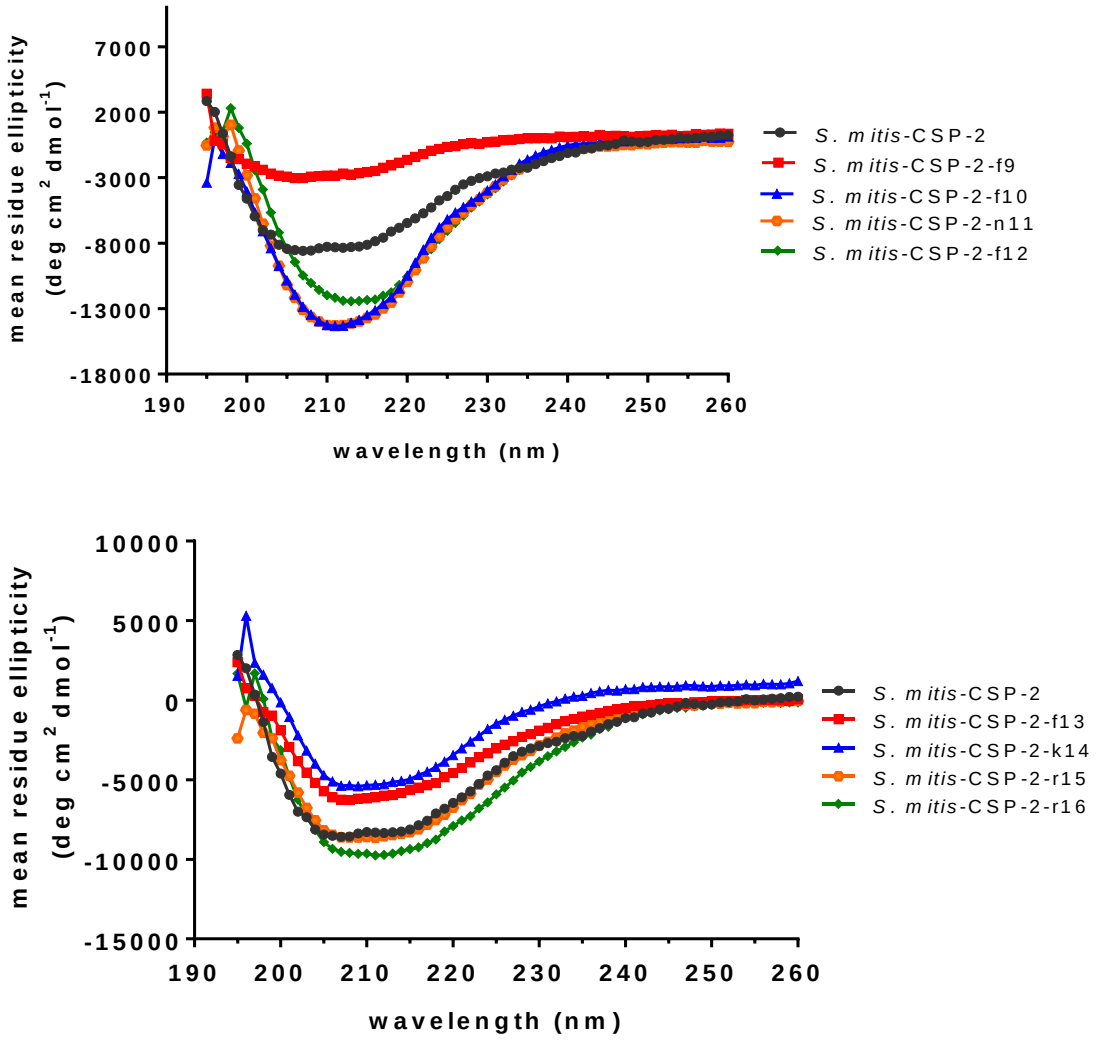


Figure S-21. CD spectra of the *S. mitis*-CSP-2 D-amino acid scan library in membrane mimicking conditions (20% TFE: 80% PBS, pH 7.4). All the measurements were performed with a peptide concentration of 200 µM. *S. mitis*-CSP-2 was added as a control. Most of the D-amino acid scan analogs were unstructured exhibiting a random coil pattern, with the exception of *S. mitis*-CSP-2-f10, *S. mitis*-CSP-2-n11, and *S. mitis*-CSP-2-f12 that exhibited some β-sheet pattern.

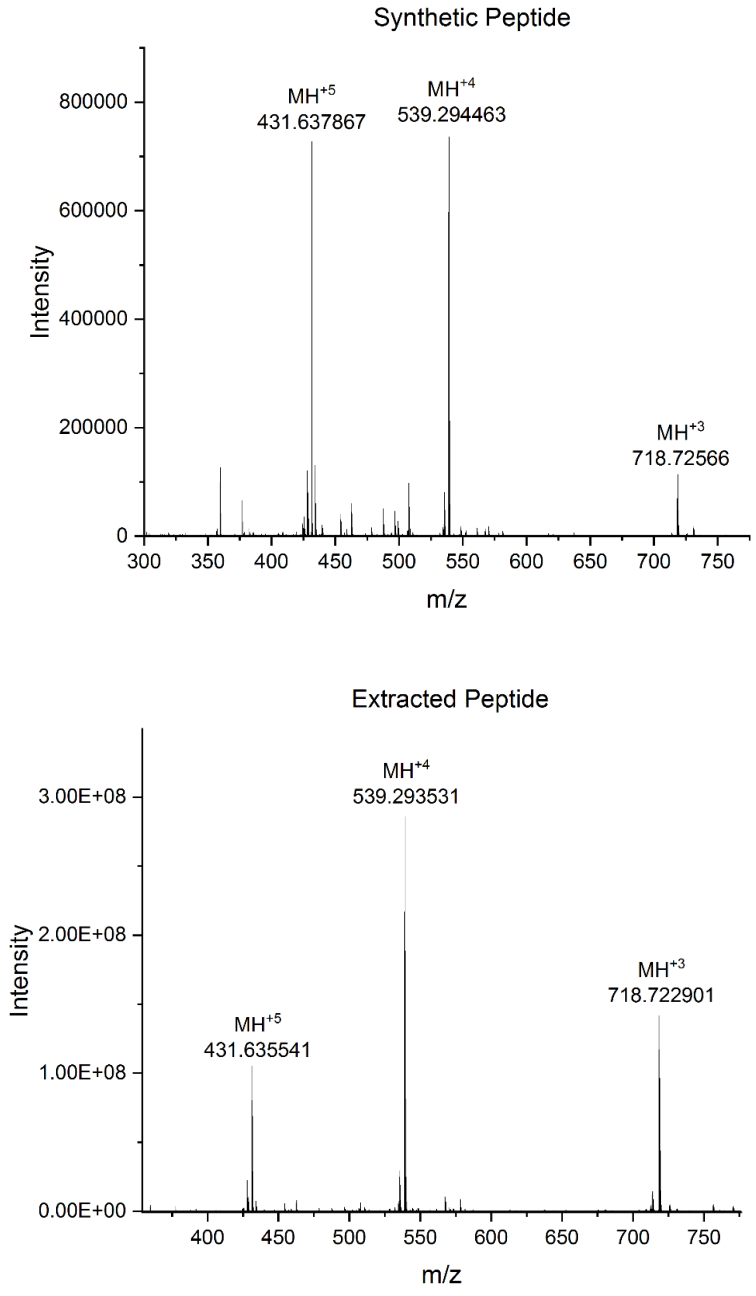
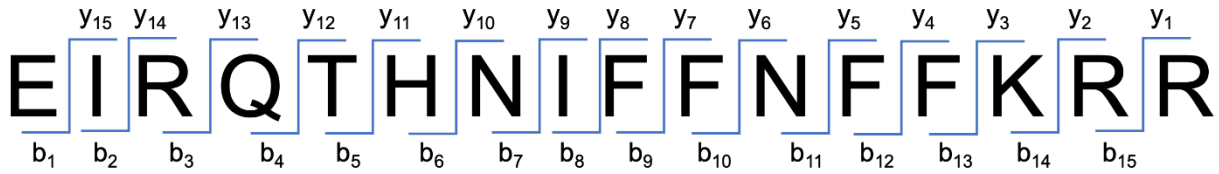


Figure S-22. Exact mass spectra of synthetic and extracted *S. mitis*-CSP-2.

A.

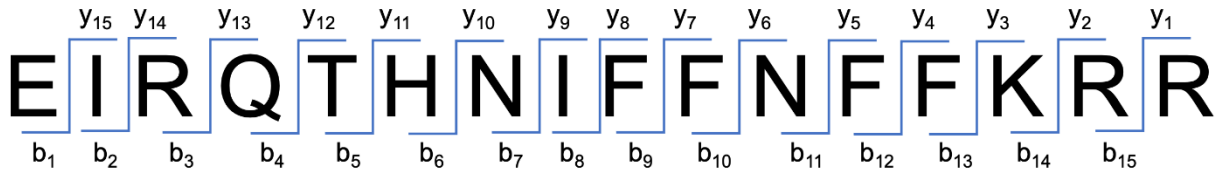


B.

Extracted Peptide			
m/z	**Y-Fragment**	**m/z**	**B-Fragment**
675.372021	y_{15}^{+3}	243.13322	b_2
637.677296	y_{14}^{+3}	399.234413	b_3
877.96467	y_{13}^{+2}	527.293264	b_4
813.933694	y_{12}^{+2}	628.341333	b_5
763.411276	y_{11}^{+2}	765.399488	b_6
694.882832	y_{10}^{+2}	879.442599	b_7
637.860666	y_9^{+2}	992.523141	b_8
581.318775	y_8^{+2}	570.301192	b_9^{+2}
507.783843	y_7^{+2}	643.834733	b_{10}^{+2}
867.494174	y_6	700.855728	b_{11}^{+2}
753.452229	y_5	774.393209	b_{12}^{+2}
606.382597	y_4	ND	ND
459.313944	y_3	452.237913	b_{14}-NH_3^{+4}
331.219297	y_2	990.023572	b_{15}^{+2}

Figure S-23. MS/MS analysis of the extracted *S. mitis*-CSP-2. (A) *S. mitis*-CSP-2 sequence showing the MS/MS fragments patterns. (B) Complete table of the fragments detected through the MS/MS analysis of the extracted *S. mitis*-CSP-2. ND, Not detected

A.



B.

Synthetic Peptide			
m/z	**Y-Fragment**	**m/z**	**B-Fragment**
675.373252	y_{15}^{+3}	225.122826	b_2-H_2O
637.678024	y_{14}^{+3}	399.233677	b_3
877.96259	y_{13}^{+2}	527.292339	b_4
813.934907	y_{12}^{+2}	628.340269	b_5
763.410042	y_{11}^{+2}	765.401093	b_6
694.881683	y_{10}^{+2}	879.441229	b_7
637.861752	y_9^{+2}	992.525846	b_8
1161.63507	y_8	570.300207	b_9^{+2}
1014.56025	y_7	643.833649	b_{10}^{+2}
867.492818	y_6	691.850776	b_{11}-H_2O^{+2}
753.451007	y_5	ND	ND
606.383567	y_4	ND	ND
459.314439	y_3	ND	ND
331.219479	y_2	491.25873	b_{15}-NH_3^{+4}
175.11883	y_1	ND	ND

Figure S-24. MS/MS analysis of the synthetic *S. mitis*-CSP-2. (A) *S. mitis*-CSP-2 sequence showing the MS/MS fragments patterns. (B) Complete table of the fragments detected through the MS/MS analysis of the synthetic *S. mitis*-CSP-2. ND, Not detected

Reference:

1 Milly, T. A. & Tal-Gan, Y. Biological evaluation of native streptococcal competence stimulating peptides reveals potential crosstalk between Streptococcus mitis and Streptococcus pneumoniae and a new scaffold for the development of S. pneumoniae quorum sensing modulators. *RSC Chem Biol* **1**, 60-67 (2020).