Structural Characterization of Competence Stimulating Peptides (CSPs) Analogues Reveals Key Features for ComD1 and ComD2 Receptor Binding in *Streptococcus pneumoniae*

Yifang Yang,[†] Gabriel Cornilescu,[‡] and Yftah Tal-Gan*[,†]

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

[‡] National Magnetic Resonance Facility at Madison, University of Wisconsin-Madison, 433 Babcock Drive, Madison, Wisconsin, 53706, United States

ABSTRACT: *Streptococcus pneumoniae* is an important pathogen that utilizes quorum sensing (QS) to regulate genetic transformation, virulence and biofilm formation. The competence stimulating peptide (CSP) is a 17-amino acid signal peptide that is used by *S. pneumoniae* to trigger QS. *S. pneumoniae* strains can be divided into two main specificity groups based on the CSP signal they produce (CSP1 or CSP2) and their compatible receptors (ComD1 or ComD2 respectively). Modulation of QS in *S. pneumoniae* can be achieved by targeting the CSP:ComD interaction using synthetic CSP analogues. However, in order to rationally design CSP-based QS modulators with enhanced activities, an in-depth understanding of the structural features that are required for receptor binding is needed. Herein, we report a comprehensive in-solution three-dimensional structural characterization of eight CSP1 and CSP2 analogues with varied biological activities using NMR spectroscopy. Analysis of these structures revealed two distinct hydrophobic patches required for effective ComD1 and ComD2 binding.

Introduction

Streptococcus pneumoniae is an opportunistic, Gram-positive, commensal bacterium that predominately colonizes in the nasopharynx of many humans.[1] It is a major cause of pneumonia, meningitis, and otitis media, leading to over 22,000 deaths annually in the United States.[2] Pneumococcal strains that are resistant to numerous antimicrobial agents such as vancomycin, linezolid, and beta-lactams have been reported worldwide.[3] The use of pneumococcal conjugate vaccine (PCV) decreased the frequency of invasive diseases caused by antibiotic resistant pneumococcus. However, the frequent emergence of diseases caused by non-vaccine serotype pneumococcus suggests that PCV may not be the final answer to the problem of drug resistance.[4] Quorum sensing (QS) has attracted significant attention as a promising target for antibacterial drug design due to its critical role in infectivity.[5-15] Moreover, for most bacteria QS is not essential for survival.[5-7] Therefore, inhibiting QS has the potential to attenuate infectivity and at the same time cause minimal selective pressure for drug-resistant strains.

S. pneumoniae utilizes the competence stimulating peptide (CSP) to trigger QS and the corresponding phenotypes including competence (the ability of bacteria to acquire exogenous DNA and incorporate it in their genome), biofilm formation and virulence.[16-21] The concentration of CSP increases as the bacteria grow. When the CSP reaches a cell density-dependent threshold concentration, it can effectively bind to and activate a transmembrane histidine kinase receptor, ComD that, on activation, phosphorylates a response regulator, ComE. Active ComE then triggers the expression of numerous genes including *ComX*, which is responsible for the development of competence (**Figure 1**).[22-24] The ability of pneumococcus to become competent is critical to the development of antibiotic resistance because competent bacteria are able to lyse other non-competent bacteria and take up exogenous DNA that may contain antibiotic-resistance

genes.[16] Moreover, Lau *et al.* reported that an insertional mutation in the *comB* and *comD* genes, both of which are essential to the CSP-based QS circuitry, resulted in attenuation of *S. pneumoniae*-mediated bacteremia and pneumonia in a mouse model of infection, suggesting that inhibiting QS has therapeutic potential.[18, 21] *S. pneumoniae* strains can be divided into two main specificity groups based on the CSP signal they produce (CSP1 or CSP2) and their compatible transmembrane histidine kinase receptor (ComD1 or ComD2, respectively), and both groups are prevalent in clinical settings.[25] Thus, pan-group QS inhibitors are needed to effectively attenuate pneumococcus pathogenicity.

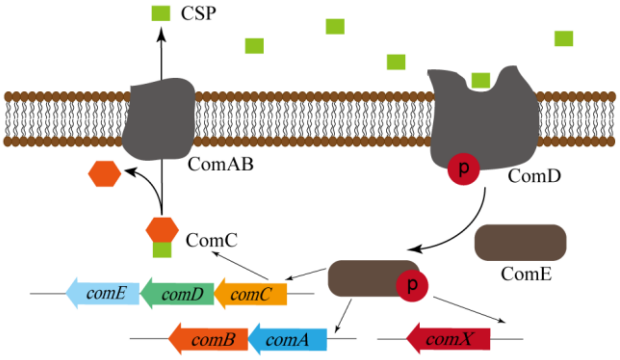

Figure 1. *S. pneumoniae* CSP-mediated QS circuit. The *ComC* gene encodes a pre-CSP peptide, which is processed and secreted by the ABC transporter (ComAB). As the bacteria grow, the concentration of CSP increases until it reaches a threshold. Upon reaching the threshold concentration, CSP activates a transmembrane histidine kinase receptor (ComD), which, after being activated, transfers a phosphate group to its cognate response regulator (ComE). After phosphorylation ComE triggers the transcription of numerous genes including *comX,* the effector molecule of the circuitry that regulates QS-mediated phenotypes.

Pneumococcal QS can be modulated through interfering with CSP:ComD interaction by using CSP analogues. In our previous study we conducted an extensive structure-activity relationship (SAR) study of CSP1 and CSP2 against both ComD1 and ComD2 where we identified residues in CSP1 and CSP2 that were important for receptor binding and activation.[26] Our analysis also revealed the first pan-group ComD activator, CSP1-K6A, as well as led to the development of the most potent ComD2 inhibitor to date, CSP2-E1Ad10. Additionally, structural analysis using

circular dichroism (CD) of CSP1, CSP2 and their analogues suggested that a helical structure is needed for CSP1:ComD1 binding and that an α-helix is also the bioactive conformation of CSP2. However, our previous study also left several unanswered questions. Broadly, the CD analysis only provided globular structural trends of the CSP analogues. However, in-depth structural analysis on the atomic level is required to fully understand the different activity trends and design potent QS modulators.[27] Specifically, we wanted to determine the reason why the Arg3 residue is more critical to ComD1 binding than any other residue in CSP1. Moreover, our previous study revealed that D-amino acid replacements in CSP1 resulted in lowered helicity, however, the specific structural changes that led to the lowered helicity were not known. We hypothesized that the orientation change of the side chain due to the D-amino acid replacement introduced steric hindrance that destabilized the helix, however, additional structural information is required to test this hypothesis. Additionally, both CSP1-K6A and CSP2-d10 adopt an α-helix conformation and have increased potency towards ComD2 compared to the natural signal CSP2, which adopts a β-sheet conformation. We hypothesized that both analogues share common structural features that are responsible for the improved ComD2 binding, but again needed detailed structural information to compare them. Lastly, as CSP1-E1A is the only potent ComD1 inhibitor to date, we wanted to probe this analogue and evaluate the structural effects of the E1A mutation on the overall CSP1 conformation.

To address the aforementioned questions, we report herein a comprehensive 3D structure characterization of eight CSP1 and CSP2 analogues using 2D-NMR spectroscopy. Previously, the structure of CSP1 was reported by Johnsborg et al. using 2D-NMR spectroscopy.[28] However, this structure alone is not sufficient to identify structural features that are important for receptor binding, activation and specificity. To this end, a detailed analysis of several analogues with

various activity profiles is needed. Ideally, the peptides should be analyzed in the presence of the ComD receptors. However, since these are transmembrane receptors that cannot be isolated and purified while retaining their active form, we chose to test the CSP analogues in conditions that best mimic the ComD receptor environment. In our previous study, we used 20% TFE (trifluoroethanol) in PBS (phosphate-buffered saline) as the membrane mimicking condition to conduct the CD experiments. Thus, for consistency, we first applied the same conditions to the 2D-NMR characterization of CSP1. For comparison, we also applied the same membrane mimicking condition as used by Johnsborg et al.: the micelle-forming reagent DPC (Dodecylphosphocholine) in PBS.[28] The results showed that CSP1 was more structured in the DPC solution (20-structure ensemble backbone and heavy atom RMSD values of 0.29 and 1.10 Å for CSP1 in DPC vs. 1.38 and 2.48 Å for CSP1 in TFE, respectively; **Figure S-1A and S-1B, and Table 2**). Moreover, micelles are better cell membrane mimetics than TFE. Therefore, we chose the DPC solution as the membrane mimicking condition for the other eight analogues we evaluated. The analogues we chose for the NMR study represent a variety of properties against the ComD receptors, such as reduced activity, enhanced activity, inhibition, and pan-group activation (**Table 1**). Although caution must be taken when trying to draw conclusions about the bioactive conformations of flexible linear peptides in solution, the relative rigidity of the structure ensembles obtained from the NMR restraints, as can be seen from the relatively low RMSD values for the 20-structure ensembles of the different analogs (**Table 2**), along with our ability to correlate the structural data with the biological activity suggests that the structures have biological relevance. Thus, comparison of the structures provides valuable mechanistic insights about the structural features that are responsible for the diverse biological activities. This detailed

structural knowledge will, in turn, guide the rational design of novel CSP-based analogues with desired activity profiles.

Table 1. Biological activity and percent helicity of CSP1, CSP2 and select analogues analyzed in this study*

Name	Sequence	EC_{50} / IC_{50} (nM)		Helicity (%)
		ComD1	ComD2	
CSP1	EMRLSKFFRDFILQRKK	10.3 (6.27 - 16.8)	526 (498 - 556)	20.1 %
CSP2	EMRISRIILDFLFLRKK	1650 (1190 - 2300)	50.7 (40.6 - 63.2)	β-sheet
CSP1-E1A	**A**MRLSKFFRDFILQRKK	85.7 (50.8 - 145)	--	18.8 %
CSP1-R3A	EM**A**LSKFFRDFILQRKK	--	--	18.5 %
CSP1-K6A	EMRLS**A**FFRDFILQRKK	51.0 (37.9 - 68.6)	24.0 (14.7 - 39.3)	22.2 %
CSP1-F11A	EMRLSKFFRD**A**ILQRKK	--	--	21.5 %
CSP1-f11	EMRLSKFFRD**f**ILQRKK	--	--	8.8 %
CSP2-d10	EMRISRIIL**d**FLFLRKK	513 (437 - 602)	2.86 (1.91 - 4.31)	17.2 %
CSP2-l14	EMRISRIILDFLF**l**RKK	>1000	54.2 (52.9 - 55.6)	19.8 %
CSP2-E1Ad10	**A**MRISRIIL**d**FLFLRKK	>1000	56.5 (53.5 - 59.6)	18.0 %

* Data from reference 26.

Experimental Section

Peptide synthesis and Biological Assays

Peptides were synthesized, purified to homogeneity (>95% purity, as determined by RP-HPLC), characterized, and evaluated for their ability to modulate QS in *S. pneumoniae* using the same protocols we previously described.[26] For full details of methods, see Supporting Information.

NMR sample preparation

For NMR structure elucidation, two solution conditions were applied: 1) the peptide was dissolved in 250 mM deuterated dodecyl phosphocholine (DPC-d_{38}; CDN Isotopes) in a PBS buffer solution with 10% D_2O (Cambridge Isotope Laboratories), or 2) the peptide was dissolved in 20% (vol/vol) deuterated trifluoroethanol (TFE-d_2; Sigma-Aldrich) in a PBS buffer solution. For both conditions, the concentration of the peptide was 1.9 mM. PBS buffer solution was a water solution that contained NaCl (137 mM), KCl (2.7 mM), Na_2HPO_4 (10 mM), and KH_2PO_4 (1.8 mM), and the pH was adjusted to 7.4.

NMR spectroscopy

All NMR spectra were recorded on a Bruker 900 MHz spectrometer at 298 K. Spectra were processed using NMR Pipe software. Chemical shifts were referenced to water at 4.771 ppm. We acquired the following two-dimensional (2-D) homonuclear experiments: gradient selection COSY with presaturation, TOCSY with DIPSI spinlock and the 3-9-19 or excitation sculpting with gradient water suppression schemes, ^{1}H-^{15}N fast HSQC, sensitivity enhanced ^{1}H-^{13}C HSQC with selective 180° ^{13}C pulses, ROESY (rotating frame NOESY) with continuous wave spin lock and 3-9-19 water suppression and a 200 ms mixing time, and NOESY with flip-back and Watergate water suppression and a 200 ms mixing time. The COSY experiments were collected with 1024 and 2048 complex and real data points in the direct and indirect dimensions, respectively, with 4 scans per data point. The TOCSY and ROESY experiments were acquired with 1024 direct and 512 indirect complex data points, with 80 ms and 200 ms spin lock durations, and with 8 and 16 scans per data point, respectively. The ^{1}H-^{15}N and ^{1}H-^{13}C HSQC experiments were collected with 1024 complex data points and 13 ppm spectral width in the direct dimensions and 128 complex data points with 28 ppm spectral width and 256 complex data points and 150 ppm spectral width in the indirect dimension, with 8 and 16 scans per data point, respectively. A 1 s relaxation delay was used in all experiments except TOCSY (relaxation delay of 1.2 s). Excitation sculpting with gradient water suppression was used in the ^{1}H 1-D experiment, and 16384 real data points were acquired, with 8 scans per data point.

Spectra assignment and structure calculation

All spectra were analyzed with NMRFAM-SPARKY.[29] Assignment of resonances for each peptide was achieved using the standard sequential assignment methodology. The volumes of the NOE peaks were calculated by SPARKY and converted into a continuous distribution of interproton distance restraints, with a uniform 35% distance error applied to take into account

spin diffusion (See Table 2 for total number of NOEs divided into backbone vs. side-chain for each peptide). The 2-D ^{1}H-^{15}N and ^{1}H-^{13}C HSQC experiments allowed assigning all backbone (and C^{β}) atoms, which were subsequently used as input in the TALOS-N program to generate backbone dihedral angle restraints (ϕ/ψ) and side chain chi1 angle restraints (See Supporting Information for complete restraint tables used for structural calculations).[30] Three-dimensional structure calculations and refinements made use of the torsion angle molecular dynamics and the internal variable dynamics modules of Xplor-NIH (v. 2.42).[31] Backbone and Heavy Atom RMSD values were calculated for the 20-structure ensembles of all peptides using the entire peptide sequence (**Table 2**). For CSP2-l14, a second overlay and RMSD calculation was conducted using residues 4-14 to eliminate the effects of the highly flexible *C*- and *N*-terminal residues (residues 1-3 and 15-17). PyMOL was used for visual analysis and presentation of the peptide structure.

Table 2. Number of assigned NOEs and RMSD values for each peptide

Peptide name	Total NOEs[c]	Intra-residue NOEs	Inter-residue NOEs						Backbone RMSD (Å)[a]	Heavy atom RMSD (Å)[a]
			i - i + 1	*i - i + 2*	*i - i + 3*	*i - i +4*	*i - i + 5*	other		
CSP1(DPC)	577 (87,318,172)	200 (15,111,74)	131 (28,80,23)	93 (23,52,18)	71 (12,41,18)	46 (8,21,17)	27 (1,11,15)	9 (0,2,7)	0.29	1.10
CSP1(TFE)	306 (58,188,60)	121 (14,88,19)	90 (23,48,19)	27 (9,14,4)	41 (9,25,7)	26 (3,12,11)	1 (0,1,0)	0	1.38	2.48
CSP1-E1A	664 (88,361,215)	198 (15,120,63)	186 (33,107,46)	69 (16,38,15)	105 (16,61,28)	78 (8,28,42)	14 (0,3,11)	14 (0,4,10)	0.24	0.72
CSP1-R3A	476 (62,266,148)	183 (16,107,60)	135 (26,87,22)	33 (7,10,16)	73 (9,46,18)	44 (4,16,24)	6 (0,0,6)	2 (0,0,2)	0.22	0.70
CSP1-K6A	806 (101,441,264)	202 (17,121,64)	208 (33,117,58)	107 (19,61,27)	143 (13,77,53)	94 (12,46,36)	31 (4,10,17)	21 (3,9,9)	0.10	0.60
CSP1-F11A	550 (84,311,155)	190 (14,116,60)	161 (38,100,23)	41 (17,18,6)	82 (9,47,26)	67 (5,26,36)	6 (1,1,4)	3 (0,3,0)	0.25	0.94
CSP1-f11	592 (64,305,223)	162 (13,97,52)	164 (22,93,49)	83 (12,38,33)	96 (7,48,41)	59 (6,23,30)	20 (4,5,11)	8 (0,1,7)	0.75	1.44
CSP2-d10	548 (101,298,149)	183 (15,111,57)	174 (42,97,35)	71 (20,34,17)	70 (16,34,20)	43 (7,19,17)	5 (1,3,1)	2 (0,0,2)	0.76	1.74
CSP2-l14	564 (94,309,161)	191 (15,117,59)	202 (43,113,46)	67 (17,26,24)	65 (12,34,19)	27 (5,13,9)	10 (1,5,4)	2 (1,1,0)	1.12 (0.14[b])	2.22 (1.05[b])
CSP2-E1Ad10	399 (79,221,99)	163 (15,105,43)	129 (36,66,27)	38 (11,19,8)	43 (10,22,11)	25 (7,8,10)	0	1 (0,1,0)	0.39	0.96

[a] RMSD values calculated for the 20-structure ensemble of each peptide. [b] RMSD values calculated for the 20-structure ensemble without the flexible end (E1-R3 and R15-K17). [c] Total # of NOEs (#of Backbone-to-Backbone NOEs, # of Backbone-to-Side-Chain NOEs, # of Side-Chain-to-Side-Chain NOEs).

Bacteria strain and growth conditions

S. pneumoniae D39pcomX::lacZ (group I) and TIGR4pcomX::lacZ (group II) reporter strains were used to examine the ability of the synthesized CSP analogs to modulate the ComD receptors, and thus the QS circuit in *S. pneumoniae*. Freezer stocks 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 ºC. For experiments, bacteria from the freezer stocks were streaked onto a THY agar plate containing 5% serum and chloramphenicol at a final concentration of 4 μg/mL. The plate was incubated for 8-9 hours in a CO_2 incubator (37 ºC with 5% CO_2). Fresh colonies were transferred to 5 mL THY broth supplemented with chloramphenicol at a final concentration of 4 μg/mL and the culture was incubated in a CO_2 incubator overnight (15 hours). 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 (0.30-0.35 for D39pcomX::lacZ; 0.20-0.25 for TIGR4pcomX::lacZ) as determined by using a plate reader.

Beta-Galactosidase dose response assay

CSP analog stock solutions were diluted with dimethyl sulfoxide (DMSO) in serial dilutions (either 1:2, 1:3, or 1:5). 2 μL of solution with each concentration was added in triplicate to a clear 96-well microtiter plate. 2 μL of 20 μM solution of CSP1 were added in triplicate and served as the positive control for the group I strain (D39pcomX::lacZ), while 2μL of 100 μM solution of CSP2 were added as the positive control for the group II strain (TIGR4pcomX::lacZ). 2 μL DMSO were added in triplicate and served as the negative control. Then, 198 μL bacterial culture were added to each well with CSP and analogs. The plate was incubated at 37 ºC for 30 minutes, and the OD 600_{nm} was measured. In order to measure the beta-galactosidase activity in

the pneumococcal culture, the cells were lysed by incubating the culture for 30 minutes at 37 °C with 20 µL 0.1% Triton X-100. 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 is 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 ºC. 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. The final 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. GraphPad Prism 5 was used to calculate the EC_{50} values, which are the concentrations of an analog where 50% of the maximum activity is reached.

Data analyses

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.

Circular Dichroism (CD) spectroscopy measurement

CD experiments in this study were performed as we previously described.[26] Briefly, an Aviv Biomedical CD spectrometer (model 202-01) was used to record all CD spectra. Samples were prepared by dissolving pure peptide in 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) mixed with 20% trifluoroethanol (TFE) to

achieve a final concentration of 20 μM. Measurements were performed at 25 ºC with a quartz cuvette with a path length of 1 cm. Samples were scanned one time at 3 nm min^{-1} with a bandwidth of 1 nm and a response time of 20 s over a wavelength range of 195 to 260 nm.

Results and Discussion

All NMR spectra were obtained on a Bruker Avance III 900 MHz spectrometer with Bruker 5 mm cryoprobe at 298 K. Two-dimensional COSY, TOCSY, NOESY, ^{1}H-^{13}C and ^{1}H-^{15}N HSQC experiments were performed to assign each peptide. All spectra were analyzed with NMRFAM-SPARKY and interproton distance restraints were obtained from converting the volume of NOE peaks.[29] Dihedral angle restraints and chi1 angle restraints were obtained by using TALOS-N with chemical shift values.[30] Three-dimensional structure calculations and refinements were performed with Xplor-NIH.[31]

Our CSP1 structure revealed an amphiphilic α-helix region where hydrophobic side chains such as L4, F7, F8, F11, and I12 are located on one side of the helix (**Figure 2A**), while hydrophilic side chains such as K6, R9 and D10 are located on the other side of the helix. A similar amphiphilic feature was also observed by Johnsborg et al.[28] The main difference between the two structures is that in our structure, the helix spans from L4 to K16, while in Johnsborg's structure a shorter helix (from K6 to L12) was observed. This difference can be explained by the much larger number of NOE restraints (577 versus 278) we obtained from the NOESY spectrum, potentially due to the higher resolution NMR spectrometer we used and the improved version of the TALOS software. In fact, Johnsborg et al. also found certain NOEs that suggested that a longer helix is present, however they did not have enough NOE restraints to reflect that in the final structure.[28] Interestingly, the amphiphilic characteristic was found in all the CSP1 analogues we analyzed, suggesting that this feature is critical for receptor binding. Moreover, our

previous bioassay results indicated that alanine replacements in positions L4, F7, F8, F11 or I12 caused between 20- to over 100-fold decreases in activity against ComD1, while replacements in K6, R9 and D10 led to only 3- to 5-fold decreases in activity against ComD1.[26] Combined, these results suggest that the hydrophobic side of the helix plays a major role in the binding of CSP1 to ComD1. We therefore hypothesized that L4, F7, F8, F11 and I12 in CSP1 form a hydrophobic patch that spans across two full helical turns and provides optimal binding to ComD1, and that any structural change affecting this hydrophobic patch will weaken the binding. Importantly, the side-chain rotameric conformations of the key hydrophobic residues forming the hydrophobic patch were found to be well-defined in all the CSP analogues we studied, allowing clear distinction between different hydrophobic patches (**Figures S-1 – S-3**).

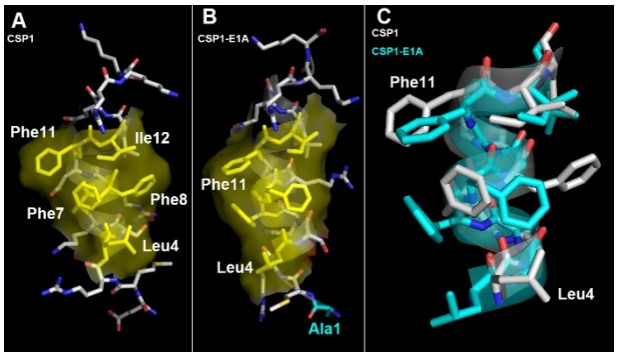

Figure 2. (A) Heavy atom average structure of CSP1 (BMRB accession ID: 30416, and reference 25). (B) Heavy atom average structure of CSP1-E1A (BMRB accession ID: 30427). (C) Overlay of CSP1 (silver) and CSP1-E1A (cyan) structures. Residues E1-R3 and Q14-K17, as well as the side chains of S5, K6, R9 and D10 residues are hidden in the structures in (C) for clarity. Yellow surfaces in (A) and (B) represent the hydrophobic patch formed by the critical residues. See SI for overlay of the full structures.

CSP1-E1A is a potent ComD1 inhibitor (IC_{50} = 85.7 nM against ComD1; **Table 1**). Since this peptide was capable of binding to the ComD1 receptor, we expected this analogue to contain a similar hydrophobic patch to that of CSP1. Indeed, structural comparison of CSP1 and CSP1-E1A revealed that, with the exception of the loss of E1 in CSP1-E1A, both peptides adopt a very similar conformation (**Figure 2A-B**). Furthermore, the hydrophobic patch of CSP1-E1A aligned

very well to that of CSP1, with only minor perturbations in the orientation of the side chains in residues L4, F8 and F11 (**Figure 2C**). Overall, the structural data confirmed that the E1 residue is critical to ComD1 activation, presumably through direct interaction with the receptor in the form of a salt-bridge or metal-mediated bond, and that the presence of the hydrophobic patch in CSP1-E1A is the reason this analogue binds strongly to ComD1, resulting in a potent competitive inhibitor. Unfortunately, this analysis could not provide a more definitive structural explanation to the role of Glu1 in receptor activation.

CSP1-K6A was found to be the only pan-group activator (EC_{50} = 51.0 nM against ComD1 and 24.0 nM against ComD2; **Table 1**). In the context of ComD1, this analogue was also capable of binding the ComD1 receptor. Thus, we expected this analogue to also contain a similar hydrophobic patch. Similarly to CSP1-E1A, structural analysis of CSP1-K6A revealed a hydrophobic patch that aligned well with that of CSP1, with the exception of the F11 residue (**Figure 3A-B**). The modified orientation of the F11 side chain may be the reason that CSP1-K6A has a slightly weaker binding affinity to ComD1 compared to CSP1 (5-fold decrease in potency). An overlay of CSP1-E1A and CSP2-K6A revealed that the hydrophobic patches of these two analogues also align well. Specifically, the F11 residue overlays better in CSP1-E1A and CSP1-K6A (**Figure 3C**). In our previous study, a dual modified analogue, CSP1-E1AK6A, was found to be a potent ComD1 inhibitor with a similar IC_{50} value to that of CSP1-E1A (104 nM vs. 85.7 nM, respectively).[26] The similarity in structures of CSP1-E1A and CSP1-K6A can thus explain the similar binding affinities of these analogues to the ComD1 receptor and further emphasize the importance of the hydrophobic patch to ComD1 binding.

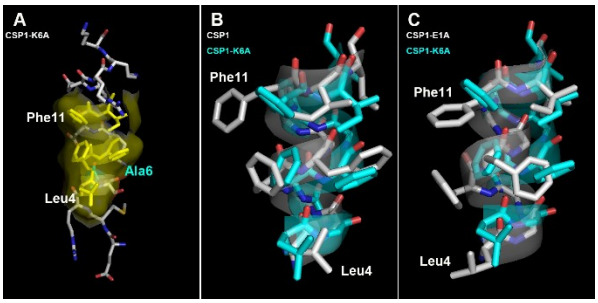

Figure 3. (A) Heavy atom average structure of CSP1-K6A (BMRB accession ID: 30429). (B) Overlay of CSP1 (silver) and CSP1-K6A (cyan) structures. (C) Overlay of CSP1-E1A (silver) and CSP1-K6A (cyan) structures. Residues E1-R3 and Q14-K17, as well as the side chains of S5, K6 (A6 for CSP1-K6A), R9 and D10 residues are hidden in the structures in (B) and (C) for clarity. Yellow surface in (A) represents the hydrophobic patch formed by the critical residues. See SI for overlay of the full structures.

CSP1-R3A is a CSP1 analogue that exhibited complete loss of activity against the ComD1 receptor (**Table 1**). Since we hypothesized that a hydrophobic patch is vital for activity, we expected to observe a significantly disrupted hydrophobic patch for this analogue. In our previous CD analysis this analogue exhibited similar α-helical propensity to that of CSP1 (18.5% vs. 20.1%, respectively; **Table 1**), suggesting that the two peptides assume a similar conformation. However, as expected, the detailed NMR analysis revealed that the side chains of three of the five key hydrophobic residues in CSP1-R3A (F7, F11 and I12) dramatically deviated from their original position in CSP1 (**Figure 4A and 4D**), resulting in a significant change to the overall shape of the hydrophobic patch. These deviations can both lead to the loss of binding interactions with the ComD1 receptor as well as introduce steric clashes that may potentially interfere with the interactions of other residues with the receptor binding site. Overall, the analysis of CSP1-R3A highlights the importance of detailed structural determination at the atomic level (NMR) in addition to an overall conformational analysis that is focused on the peptide backbone (CD).

The analysis of CSP1-E1A and CSP1-K6A revealed that the conformation of the F11 side chain can be somewhat modified without significantly affecting the ability of the peptides to bind the

ComD1 receptor. We therefore decided to evaluate the two F11 modified peptides, CSP1-f11 and CSP1-F11A, both of which were very weak QS activators (60%-70% QS activation at 10 μM concentration compared to CSP1).[26] CSP1-f11 exhibited low helical propensity in the CD analysis compared to CSP1 (**Table 1**). We therefore hypothesized that the D-amino acid replacement resulted in disruption of the helical structure near the *C*-terminus leading to elimination of part of the hydrophobic patch. The NMR structure of CSP1-f11 revealed that, due to the D-amino acid replacement, the orientation of the F11 side chain changed from pointing out of the helix to pointing into the helix, resulting in the disruption of hydrogen bonding of the backbone and destabilization of the helix in this region (**Figure 4B**). Thus, the helical region of CSP1-f11 spans from F8 to D10, significantly shorter than that of CSP1 (L4 to K16). Although the F11 side chain is pointing away from the hydrophobic patch, the rest of the hydrophobic patch (residues L4, F7, F8 and I12) is maintained and is likely responsible for the weak activity of CSP1-f11 (**Figure 4E**).

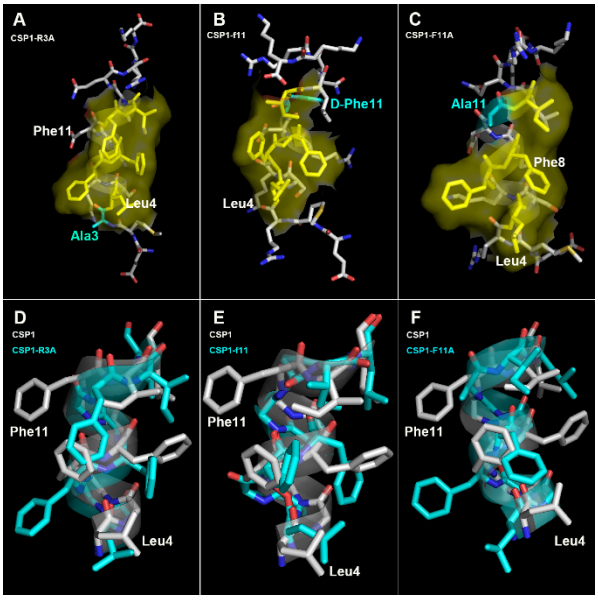

Figure 4. (A) Heavy atom average structure of CSP1-R3A (BMRB accession ID: 30428). (B) Heavy atom average structure of CSP1-f11 (BMRB accession ID: 30431). (C) Heavy atom average structure of CSP1-F11A (BMRB accession ID: 30430). (D) Overlay of CSP1 (silver) and CSP1-R3A (cyan) structures. (E) Overlay of CSP1 (silver) and CSP1-f11 (cyan) structures. (F) Overlay of CSP1 (silver) and CSP1-F11A (cyan) structures. Residues E1-R3 and Q14-K17, as well as the side chains of S5, K6, R9 and D10 residues are hidden in the structures in (D), (E) and

(F) for clarity. Yellow surfaces in (A), (B) and (C) represent the hydrophobic patch formed by the critical residues. See SI for overlay of the full structures.

CSP1-F11A exhibited helical propensity similar to CSP1 in the CD analysis (**Table 1**). We therefore hypothesized that the loss of binding affinity in this analogue is likely due to the elimination of the binding interactions provided by the F11 residue. Indeed, structural comparison of CSP1-F11A with CSP1 revealed that, although most of the hydrophobic patch residues align poorly, the overall shape of the modified patch is similar, especially the orientation of L4, F7 and F8, with only the F11 side chain missing (**Figure 4C and 4F**). Together, the structural analysis of CSP1-f11 and CSP1-F11A, along with the biological data, revealed that the F11 residue is critical for effective binding.

Our analysis thus far has focused on the ComD1 receptor, however we also wanted to gain structural insights regarding the ComD2 receptor. In our previous study, CSP2 was found to adopt a β-sheet conformation, presumably through peptide aggregates, and this conformation is likely not the bioactive conformation. For instance, a modified CSP2 analogue, CSP2-d10, was found to be ~20-fold more active than the native CSP2 and to adopt an α-helix conformation (**Table 1**). Thus, the increased activity of this analogue was hypothesized to be due to a shift in the oligomerization state, from the inactive β-sheet aggregate to the active α-helical monomeric form.[26] We therefore decided to study the structure of this analogue, the most potent ComD2 activator known. Due to the similarity between the two ComD receptors (they share 97% homology), we hypothesized that a hydrophobic patch is required for effective binding to the ComD2 receptor. Moreover, due to the specificity of the receptors (only CSP1-K6A was found to effectively activate both receptors), we hypothesized that the required shape of the hydrophobic patch differs between the two receptors. Evaluation of CSP2-d10 revealed that the helix region spans from D10 to K16 significantly shorter compared to most of the CSP1

analogues. However, closer analysis of the structure revealed that CSP2-d10 has a rigid, α-helix-like region spanning from I4 to L9 (**Figure 5A**). Moreover, TALOS-N analysis of this region indicated ϕ/ψ angles of -60/-30 degrees, close to the -64/-41 ϕ/ψ canonical angles of helical regions. Lastly, the CSP2-d10 structure exhibited an amphiphilic feature with residues I8, L9, F11, L12 and F13 forming a hydrophobic patch that spans across one helical turn.

To further evaluate the structural requirements of the hydrophobic patch needed for ComD2 binding, we compared the structures of CSP2-d10 with that of CSP1-K6A, the most potent CSP1 analogue against the ComD2 receptor (**Figure 5B**). We found that although the two peptides have different hydrophobic sequences, the hydrophobic patches formed by the two peptides are similar. Specifically, the hydrophobic patch formed by F7, F8, F11 and I12 in CSP1-K6A aligns well with that formed by I8, L9, F11, L12 and F13 in CSP2-d10. It is possible that when binding to ComD2, residues F7, F8, and F11 in CSP1-K6A take the place of residues I8, L9, and F11 in CSP2-d10, while residue I12 in CSP1-K6A may function as both residues L12 and F13 in CSP2-d10.

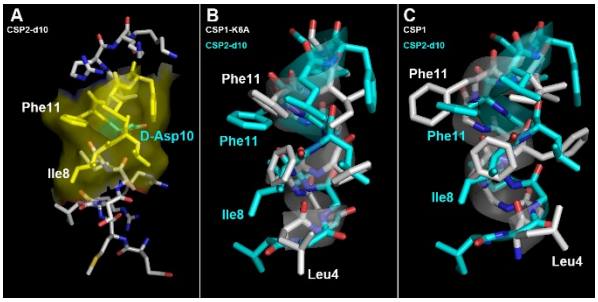

Figure 5. (A) Heavy atom average structure of CSP2-d10 (BMRB accession ID: 30432). (B) Overlay of CSP1-K6A (silver) and CSP2-d10 (cyan) structures. (C) Overlay of CSP1 (silver) and CSP2-d10 (cyan) structures. Residues E1-R3 and Q14-K17, as well as the side chains of S5, K6 (A6 for CSP1-K6A), R9 and D10 residues for CSP1-K6A and CSP1 are hidden in the structures in (B) and (C) for clarity. Residues E1-R3 and L14-K17, as well as the side chains of S5, R6 and D10 for CSP2-d10 are hidden in the structures in (B) and (C) for clarity. Yellow surface in (A) represents the hydrophobic patch formed by the critical residues. See SI for overlay of the full structures.

Next, we evaluated CSP1 in the context of the ComD2 receptor. CSP1 is a relatively weak ComD2 activator (EC_{50} = 526 nM against ComD2; **Table 1**), thus we hypothesized that its hydrophobic patch would not align well with that of CSP2-d10. As expected, comparison of the structures of CSP1 and CSP2-d10 revealed that only residues F7 and I12 in CSP1 match well with residues I8 and L12 in CSP2-d10, while residues F8 and F11 in CSP1 align poorly with residues L9 and F11 in CSP2-d10 respectively (**Figure 5C**). We then evaluated two additional CSP2 analogues capable of binding well to the ComD2 receptor: CSP2-E1Ad10, the most potent ComD2 inhibitor (IC_{50} = 56.5 nM against ComD2; **Table 1**) and CSP2-l14, a potent ComD2 activator (EC_{50} = 54.2 nM against ComD2; **Table 1**). Since both analogues are capable of binding the ComD2 receptor, we hypothesized that they both possess a functional hydrophobic patch similar to that of CSP2-d10. As expected, the results showed that all five hydrophobic residues (I8, L9, F11, L12 and F13) in CSP2-E1Ad10 align well with their counterparts in CSP2-d10 (**Figure 6A and 6C**), whereas residues I8, L9 and L12 in CSP2-l14 align well with their corresponding residues in CSP2-d10 while F11 and F13 exhibit relatively modest perturbation between the two analogues (**Figure 6B and 6D**). Combined, these results support our hypothesis that the hydrophobic patch formed by residues I8, L9, F11, L12 and F13 in the CSP2 scaffold is required for effective ComD2 binding.

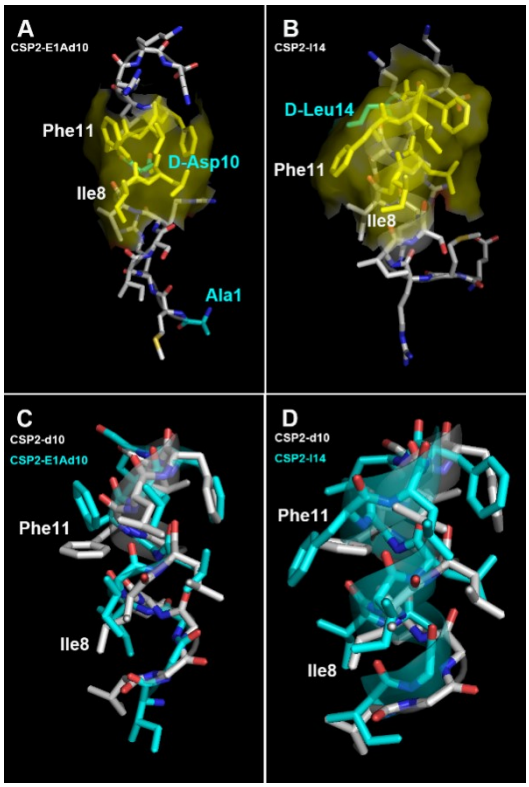

Figure 6. (A) Heavy atom average structure of CSP2-E1Ad10 (BMRB accession ID: 30433). (B) Heavy atom average structure of CSP2-l14 (BMRB accession ID: 30434). (C) Overlay of CSP2-d10 (silver) and CSP2-E1Ad10 (cyan) structures. (D) Overlay of CSP2-d10 (silver) and CSP2-l14 (cyan) structures. Residues E1-R3 and L14-K17, as well as the side chains of S5, R6 and D10 are hidden in all structures in (C) and (D) for clarity. Yellow surfaces in (A) and (B) represent the hydrophobic patch formed by the critical residues. See SI for overlay of the full structures.

Overall, our structural analysis of the CSP1 and CSP2 analogues provided a predictive model for the design of next-generation CSP analogues with minimal structural features. Based on this model we hypothesize that any CSP1 or CSP2 analogue capable of supporting an α-helix conformation and bearing the following minimal requirements, would be able to effectively bind the ComD1 or ComD2 receptors, respectively. For CSP1: E-M-R-L-X-X-F-F-X-X-F-I-X-X-R-K-K, and for CSP2: E-M-R-X-X-X-I-I-L-d-F-L-F-X-R-K-K (**Figure 7**). To test our hypothesis and validate our predictive model, we synthesized the two CSP1 and CSP2 analogues bearing our hypothesized minimally-required residues and tested their ability to activate QS using a cell-based reporter system.[26] Unfortunately, the CSP1 analogue was found to be inactive. To further evaluate the validity of our model, we synthesized a second CSP1 analogue bearing one less

mutation in the sequence (A13L): E-M-R-L-X-X-F-F-X-X-F-I-**L**-X-R-K-K, however this analogue was also found to be inactive in our reporter assay (**Table 3**). We hypothesized that the inability of the CSP1 analogues to activate the ComD1 receptor is due to an unexpected conformational change of the backbone, side-chain residues, or both. To test that, we evaluated the overall structural features of the two CSP1 analogues using circular dichroism (CD). Indeed, the two peptides were unstructured in membrane mimicking conditions (20% TFE in PBS), suggesting that several of the "dispensable" residues we identified have a role in stabilizing the bioactive conformation (**Figure 8**). Contrary to the CSP1:ComD1 system, to our satisfaction, the CSP2 analogue we designed was found to activate the ComD2 receptor with an EC_{50} value comparable to that of CSP2 (**Table 3**). Structural evaluation of this analogue using CD spectroscopy revealed that this peptide exhibits an α-helix fiber characteristics,[32] likely due to the high alanine content in the sequence, suggesting that this peptide indeed adopts an α-helix conformation. Overall, the results of the new CSP analogues reaffirms the importance of the α-helix conformation and supports the validity of the structural model as a tool to the design of novel CSP-based QS modulators, with an emphasis on the importance of stabilizing the bioactive conformation when modifying the peptide sequence.

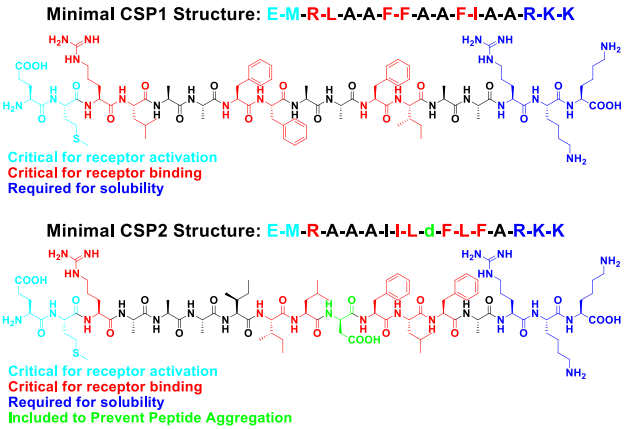

Figure 7. Predicted minimal structural requirements for CSP1 and CSP2 analogues based on the NMR analysis.

Table 3. Biological activity of the rationally-designed CSP1 and CSP2 analogues[a]

Name	Sequence	EC_{50}(nM)[b] (95% CI[c])	
		ComD1	ComD2
CSP1	EMRLSKFFRDFILQRKK	10.3 (6.27 - 16.8)	526 (498 - 556)
CSP2	EMRISRIILDFLFLRKK	1650 (1190 - 2300)	50.7 (40.6 - 63.2)
CSP1-S5AK6AR9AD10AL13AQ14A	EMRL**AA**FF**AA**FI**AA**RKK	>1000	>1000
CSP1-S5AK6AR9AD10AQ14A	EMRL**AA**FF**AA**FIL**A**RKK	>1000	>1000
CSP2-I4AS5AR6Ad10L14A	EMR**AAA**IIL**d**FLF**A**RKK	>1000	87.7 (79.5 - 96.6)

[a] See Experimental section for details of reporter strains and methods. See Supporting Information for plots of agonism dose response. All bioassays were performed in triplicate. [b] EC_{50} values determined by testing peptides over a range of concentrations. [c] 95% confidence interval.

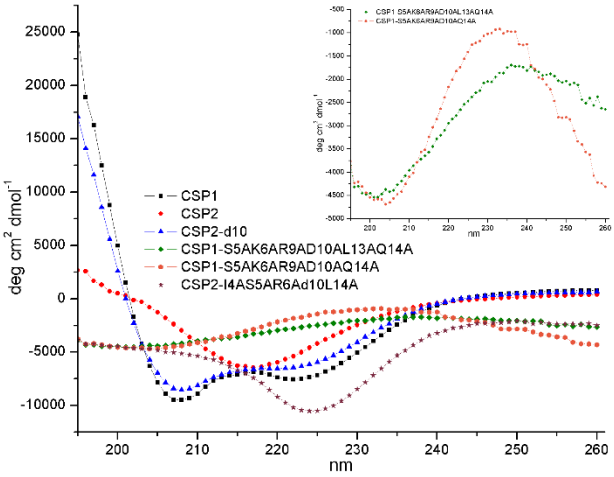

Figure 8. CD spectra of the three CSP1 and CSP2 analogues in membrane mimicking conditions. The two CSP1 analogues were found to be unstructured while the CSP2 analogue was found to exhibit an α-helix fiber characteristics. The spectra on the top right corner are the spectra of the two CSP1 analogues with adjusted y axis range to show their shapes more clearly. CD spectra of CSP1, CSP2, and CSP2-d10 were taken from reference 26. All CD spectra were acquired using the same conditions (20% TFE in PBS), with the exception that the concentration of CSP1, CSP2 and CSP2-d10 was 200 μM, while the concentration of the new CSP1 and CSP2 analogues was 20 μM.

Conclusions

In conclusion, we have performed an in-depth structural characterization of pneumococcal CSPs. Our analysis, which included CSP1, five CSP1 analogues and three CSP2 analogues, revealed distinct hydrophobic patches that are required for effective ComD1 and ComD2 binding. Specifically, CSP1 forms an elongated hydrophobic patch that spans across two helical turns,

whereas CSP2 forms a more compact hydrophobic patch spanning only one helix turn. The structures reported herein provide 3-D scaffolds for the design of new ComD1, ComD2 and pan-group CSP-based modulators with improved potencies. Lastly, we were able to validate our structural predictive model by constructing a CSP2 analogue bearing minimal structural features with activity similar to the native CSP2 signal.

Supporting Information

Full details of peptide synthesis and characterization, dose response curves for CSP analogues, tables of resonance assignments, additional structural figures, and PDB coordinates.

Corresponding Author

ytalgan@unr.edu

Acknowledgement

This work was supported by the Nevada INBRE through a grant from the National Institutes of Health (GM103440), the National Science Foundation (CHE-1808370), and the National Institutes of Health (R35GM128651). This study made use of the National Magnetic Resonance Facility at Madison, which is supported by NIH grants P41 GM103399 (NIGMS) and P41GM66326 (NIGMS). Additional equipment was purchased with funds from the University of Wisconsin, the NIH (RR02781, RR08438), the NSF (DMB-8415048, OIA-9977486, BIR-9214394), and the USDA. We would like to thank M. J. Tucker (University of Nevada, Reno) for the use of the CD spectrometer.

BMRB accession ID

CSP1 (in DPC): 30416; CSP1 (in TFE): 30435; CSP1-E1A: 30427; CSP1-R3A: 30428; CSP1-K6A: 30429; CSP1-F11A: 30430; CSP1-f11: 30431; CSP2-d10: 30432; CSP2-E1Ad10: 30433;

CSP2l14: 30434. Authors will release the atomic coordinates and experimental data upon article publication.

Abbreviations

CD, Circular dichroism; COSY, Correlation spectroscopy; CSP, Competence stimulating peptide; DIPSI, Decoupling In the Presence of Scalar Interactions; DMSO, dimethyl sulfoxide; DPC, Dodecylphosphocholine; HSQC, Heteronuclear single-quantum correlation spectroscopy; NMR, Nuclear magnetic resonance; NOE, Nuclear Overhauser effect; NOESY, Nuclear Overhauser effect spectroscopy; ONP, o-nitrophenol; ONPG, 2-Nitrophenyl-Beta-D-galactopyranoside; PBS, phosphate-buffered saline; PCV, Pneumococcal conjugate vaccine; PPM, Parts per million; QS, Quorum sensing; RMSD, Root-mean-square deviation; ROESY, Rotating-frame Overhauser Spectroscopy; RP-HPLC, Reversed-phase high performance liquid chromatography; SAR, Structure-activity relationship; TFE, trifluoroethanol; TOCSY, Total correlation spectroscopy.

References

[1] Mehr, S., and Wood, N. (2012) *Streptococcus pneumoniae*--A Review of Carriage, Infection, Serotype Replacement and Vaccination, *Paediatr Respir Rev 13*, 258-264.

[2] Huang, S. S., Johnson, K. M., Ray, G. T., Wroe, P., Lieu, T. A., Moore, M. R., Zell, E. R., Linder, J. A., Grijalva, C. G., Metlay, J. P., and Finkelstein, J. A. (2011) Healthcare Utilization and Cost of Pneumococcal Disease in the United States, *Vaccine 29*, 3398-3412.

[3] Linares, J., Ardanuy, C., Pallares, R., and Fenoll, A. (2010) Changes in Antimicrobial Resistance, Serotypes and Genotypes in *Streptococcus pneumoniae* Over a 30-Year Period, *Clin Microbiol Infect 16*, 402-410.

[4] Mitchell, P. K., Lipsitch, M., and Hanage, W. P. (2015) Carriage Burden, Multiple Colonization and Antibiotic Pressure Promote Emergence of Resistant Vaccine Escape Pneumococci, *Philos Trans R Soc Lond, B, Biol Sci 370*, 20140342.

[5] Rutherford, S. T., and Bassler, B. L. (2012) Bacterial Quorum Sensing: Its Role in Virulence and Possibilities for Its Control, *Cold Spring Harb Perspect Med 2*.

[6] LaSarre, B., and Federle, M. J. (2013) Exploiting Quorum Sensing to Confuse Bacterial Pathogens, *Microbiol Mol Biol Rev 77*, 73-111.

[7] Defoirdt, T. (2018) Quorum-Sensing Systems as Targets for Antivirulence Therapy, *Trends Microbiol 26*, 313-328.

[8] McBrayer, D. N., Gantman, B. K., Cameron, C. D., and Tal-Gan, Y. (2017) An Entirely Solid Phase Peptide Synthesis-Based Strategy for Synthesis of Gelatinase Biosynthesis-Activating Pheromone (GBAP) Analogue Libraries: Investigating the Structure-Activity Relationships of the Enterococcus faecalis Quorum Sensing Signal, *Org Lett 19*, 3295-3298.

[9] Tal-Gan, Y., Stacy, D. M., Foegen, M. K., Koenig, D. W., and Blackwell, H. E. (2013) 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.

[10] Tal-Gan, Y., Ivancic, M., Cornilescu, G., Yang, T., and Blackwell, H. E. (2016) Highly Stable, Amide-Bridged Autoinducing Peptide Analogues that Strongly Inhibit the AgrC Quorum Sensing Receptor in Staphylococcus aureus, *Angew Chem Int Ed Engl 55*, 8913-8917.

[11] Wang, B., and Muir, T. W. (2016) Regulation of Virulence in Staphylococcus aureus: Molecular Mechanisms and Remaining Puzzles, *Cell Chem Biol 23*, 214-224.

[12] Harrington, A., and Tal-Gan, Y. (2018) Identification of *Streptococcus gallolyticus* subsp. *gallolyticus* (biotype I) Competence Stimulating Peptide Pheromone, *J Bacteriol 200*, JB.00709-00717.

[13] Mull, R. W., Harrington, A., Sanchez, L. A., and Tal-Gan, Y. (2018) Cyclic Peptides that Govern Signal Transduction Pathways: From Prokaryotes to Multi-Cellular Organisms, *Curr Top Med Chem 18*, 625-644.

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

[15] Bikash, C. R., Hamry, S. R., and Tal-Gan, Y. (2018) Structure-Activity Relationships of the Competence Stimulating Peptide in Streptococcus mutans Reveal Motifs Critical for Membrane Protease SepM Recognition and ComD Receptor Activation, *ACS Infect Dis*, In press.

[16] Ween, O., Gaustad, P., and Havarstein, L. S. (1999) Identification of DNA Binding Sites for ComE, A Key Regulator of Natural Competence in *Streptococcus pneumoniae*, *Mol Microbiol 33*, 817-827.

[17] Claverys, J. P., and Havarstein, L. S. (2002) Extracellular-Peptide Control of Competence for Genetic Transformation in *Streptococcus pneumoniae*, *Front Biosci 7*, d1798-1814.

[18] Lau, G. W., Haataja, S., Lonetto, M., Kensit, S. E., Marra, A., Bryant, A. P., McDevitt, D., Morrison, D. A., and Holden, D. W. (2001) A Functional Genomic Analysis of Type 3 *Streptococcus pneumoniae* Virulence, *Mol Microbiol 40*, 555-571.

[19] Hava, D. L., and Camilli, A. (2002) Large-Scale Identification of Serotype 4 *Streptococcus pneumoniae* Virulence Factors, *Mol Microbiol 45*, 1389-1406.

[20] Oggioni, M. R., Trappetti, C., Kadioglu, A., Cassone, M., Iannelli, F., Ricci, S., Andrew, P. W., and Pozzi, G. (2006) Switch from Planktonic to Sessile Life: A Major Event in Pneumococcal Pathogenesis, *Mol Microbiol 61*, 1196-1210.

[21] Zhu, L., Lin, J., Kuang, Z., Vidal, J. E., and Lau, G. W. (2015) Deletion Analysis of *Streptococcus pneumoniae* Late Competence Genes Distinguishes Virulence Determinants that are Dependent or Independent of Competence Induction, *Mol Microbiol 97*, 151-165.

[22] Pestova, E. V., Havarstein, L. S., and Morrison, D. A. (1996) Regulation of Competence for Genetic Transformation in *Streptococcus pneumoniae* by an Auto-Induced Peptide Pheromone and a Two-Component Regulatory System, *Mol Microbiol 21*, 853-862.

[23] Hui, F. M., and Morrison, D. A. (1991) Genetic Transformation in *Streptococcus pneumoniae*: Nucleotide Sequence Analysis Shows *comA*, a Gene Required for Competence Induction, to be a Member of the Bacterial ATP-Dependent Transport Protein Family, *J Bacteriol 173*, 372-381.

[24] Cvitkovitch, D. G., Li, Y. H., and Ellen, R. P. (2003) Quorum Sensing and Biofilm Formation in Streptococcal Infections, *J Clin Invest 112*, 1626-1632.

[25] Pozzi, G., Masala, L., Iannelli, F., Manganelli, R., Havarstein, L. S., Piccoli, L., Simon, D., and Morrison, D. A. (1996) Competence for Genetic Transformation in Encapsulated Strains of *Streptococcus pneumoniae*: Two Allelic Variants of the Peptide Pheromone, *J Bacteriol 178*, 6087-6090.

[26] Yang, Y., Koirala, B., Sanchez, L. A., Phillips, N. R., Hamry, S. R., and Tal-Gan, Y. (2017) 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.

[27] Tal-Gan, Y., Ivancic, M., Cornilescu, G., Cornilescu, C. C., and Blackwell, H. E. (2013) Structural Characterization of Native Autoinducing Peptides and Abiotic Analogs Reveals Key Features Essential for Activation and Inhibition of an AgrC Quorum Sensing Receptor in *Staphylococcus aureus*, *J Am Chem Soc 135*, 18436–18444.

[28] Johnsborg, O., Kristiansen, P. E., Blomqvist, T., and Havarstein, L. S. (2006) A Hydrophobic Patch in the Competence-Stimulating Peptide, a Pneumococcal Competence Pheromone, is Essential for Specificity and Biological Activity, *J Bacteriol 188*, 1744-1749.

[29] Lee, W., Tonelli, M., and Markley, J. L. (2015) NMRFAM-SPARKY: Enhanced Software for Biomolecular NMR Spectroscopy, *Bioinformatics 31*, 1325-1327.

[30] Shen, Y., and Bax, A. (2013) Protein Backbone and Sidechain Torsion Angles Predicted from NMR Chemical Shifts Using Artificial Neural Networks, *J Biomol NMR 56*, 227-241.

[31] Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) The Xplor-NIH NMR Molecular Structure Determination Package, *J Magn Reson 160*, 65-73.

[32] Bromley, E. H., Channon, K. J., King, P. J., Mahmoud, Z. N., Banwell, E. F., Butler, M. F., Crump, M. P., Dafforn, T. R., Hicks, M. R., Hirst, J. D., Rodger, A., and Woolfson, D. N. (2010) Assembly pathway of a designed alpha-helical protein fiber, *Biophys J 98*, 1668-1676.

For Table of Contents Use Only