Molecular Interaction Studies for Inhibition of the Streptococcus pneumoniae Competence Stimulating Peptide (CSP1) by Potent Plant-Derived Compounds.

Mekhala Nagabushan Chitagudigi¹, Govindaraju Shruthi² and Chandan Shivamallu*¹

¹ Department of Biotechnology and Bioinformatics, School of Life Sciences, JSS Academy of Higher Education & Research, Mysuru- 570015, India.
² Glimetomics Bioresolve Private Limited, Hootagalli Industrial Area, Mysuru-570018, India.

Abstract: Antimicrobial resistance (AMR) is an immense medical concern; it is among the top causes of fatality worldwide. The development of resistance occurs most frequently with nosocomial and community-acquired infections, among which Streptococcus pneumoniae is a common causative organism. This study is focused on inhibiting the quorum-sensing (QS) mechanism using plant compounds as an alternate strategy to avert AMR. A major factor in the development of AMR is genetic variability. S. pneumoniae genetic variability is enabled by the natural competence and transformation of the organism, a trait historically most notable. It is regulated by the expression of com loci genes. The com loci regulation and the regulation of other subsequent signaling pathways are a QS mediated system, for which the competence stimulating peptide (CSP1) is the autoinducer. CSP1 was selected as the target for inhibition studies due to its significant role in driving bacterial communication via QS, leading to competence, virulence, and resistance. Plant-derived compounds present a vast scope for developing antimicrobials; in this study, we have proposed using the plant compounds to avert the development of AMR by inhibiting the factor directly responsible, i.e. CSP1. Five natural plant compounds, selected based on the ADMET profile, were studied for inhibition of CSP1; these compounds were curcumin, ellagic acid, eugenol, kaempferol, tinosporinone. These five compounds had credible drug likeliness with no acute toxicity and satisfactory bioavailability score. The molecular docking studies between CSP1 and the selected five compounds revealed a satisfactory interaction with the binding pocket of CSP1 and act as potential inhibitors.

Keywords: Antimicrobial resistance, Quorum Sensing, S. pneumoniae, Molecular interaction, plant compounds.
1. INTRODUCTION

Antimicrobial resistance (AMR) against the currently available medicines plays a pivotal role in treating infectious diseases. The development of resistance properties by the microorganisms renders the drug useless, increasing the morbidity and fatality due to the ensuing illness. AMR is of particular concern with the rampant transmission of the hospital and community-acquired infections. The antibiotic abuse and frequent misdiagnosis further augment the AMR crisis. Antimicrobials have a universal application in various fields such as animal husbandry, agriculture, scientific research, pharma, and the food industry; thus, its unscrupulous exploitation is not limited to the medical forum. The presently available antimicrobial drugs are structurally large and complex molecules; designed to null an infection by inhibiting critical cellular processes such as the synthesis of DNA, RNA, and proteins. However, these cellular processes are quintessential for life; hence, AMR has risen as a strategy for survival. Alternatively, targeting a non-essential microbial component is scope for consideration, allowing for the deterrence of resistance development rather than killing the organism. In this manner, the development of virulence and resistance can be stalled since the plant molecule functions in the extracellular environment and does not become the targeted antagonist of the bacteria; and meanwhile, also allowing the natural antibiotic defense mechanisms to kill the microorganisms. Furthermore, since the natural defense mechanisms are a part of normal physiology and commensal biology, the bacteria’s drastic agonistic activity is less likely. This study focused on small molecular weight plant molecules’ activity as inhibitors of the non-essential secondary molecule CSP1 produced by Streptococcus pneumoniae, a peptide essential for inducing genetic variability leading to AMR. S. pneumoniae is a gram-positive bacteria most noteworthy for the principle of genetic transformation and competence; the expression of these traits concurs with the development of resistance. The steps that bring about the development of resistance are a process precisely coordinated between the activities of autolysis and fratricide, the development of competence, and followed by the uptake of genetic material. The expression of competence increases the cell membrane permeability of S. pneumoniae, facilitating the uptake of external DNA. Competence is synchronized with S. pneumoniae autolysis and fratricide activities, carried out by the release of autolysin and bacteriocins, respectively. Autolysis and fratricide expel DNA into the extracellular environment, contributing to the genetic pool available for uptake by the surviving fraction of S. pneumoniae in the microenvironment. The development of resistance from this point onwards is ultimately a chance factor, directly proportional to the size of the gene pool; the larger the gene pool, the higher the chance of developing resistance. The microorganism’s inherent faculties enhance the development of AMR; these primarily include group behavior and communication. The QS mechanism, a form of bacterial communication prominent during the stationary phase of growth, drives the exhibition of competence and transformation via the com loci. It is directly involved in eliciting the competence phenotype, for which the competence stimulating peptide (CSP1) is the QS autoinducer. The production of autolysin and bacteriocins is also regulated by the com loci, specifically by the comE response regulator. The QS mediated upregulation of com loci by CSP1 also upregulates the autolytic and fratricidal activity. The induction of competence by CSP1 contributes to the overall development of virulence and resistance by S. pneumoniae, thus making CSP1 the primary target molecule. The doubling time of S. pneumoniae-D39 is 54 minutes, and it remains active for 12 hours before the commencement of cell death; this offers a window of opportunity to dissipate genetic transformation and development of AMR. The use of natural plant-derived compounds is an extremely conducive option, considering the extensive breadth complexity of antimicrobials and AMR. The smaller size of plant compounds is particularly advantageous in subduing the development of resistance traits because small molecules quickly diffuse across cell barriers, such that it can directly carry out its function. Therefore, this research focuses on assessing potent antimicrobial plant-derived small molecules as inhibitors of CSP1.

2. METHODOLOGY

2.1 Screening of Biochemical Pathway

The virulent encapsulated D39 strain of S. pneumoniae was considered for the study. The QS pathway of S. pneumoniae D39 was analyzed using the KEGG pathway database.

2.2 Screening of Small Molecules

Plant-derived natural molecules with molecular mass less than 500, known to possess antimicrobial properties, were screened through the PubChem Database. The screened molecules were further scrutinized through the Drugbank based on Lipinski’s rule and the ADME (absorption, distribution, metabolism, and excretion) and toxicity profile. ADME toxicity of all the screened compounds was estimated through the swiss ADME tool.

2.3 Ligand Preparation

The 2D structure of the molecules selected as ligands were sketched using the Chemsketch tool. These structures were further cleaned and subsequently converted into the corresponding 3D structures by incorporating the appropriate 3D coordinates and hydrogens using the OpenBabel tool. Next, the 3D structures were processed to obtain a clean geometry of the structures by removing any deviations concerning its stereochemical properties using the ArgusLab tool. Finally, the cleaned 3D structures were saved in the PDB format.

2.4 Protein Structure Retrieval

The functional structure of CSP1, having no mutations and without the presence of any ligands, was screened for in the RCSB PDB database, and the appropriate form was retrieved in the PDB format.

2.5 Binding Site Prediction

The ligand-binding site in CSP1 was predicted using multiple approaches involving the in silico screening via binding site prediction server Pankweb. In silico screening by data mining through the RCSB PDB database, multiple structures of CSP1 bound to a ligand were analyzed to check the site of ligand interaction. This analysis was followed by a thorough
literature review concerning reports on in vitro studies of the interaction and binding sites of CSP1 was conducted 28–31.

2.6 Molecular Interaction Studies

Molecular docking of the five selected natural compounds against CSP1 was carried out using Autodock 4.0 48–50. The grid was generated wherein CSP1 was considered a macromolecule, and its binding site residues were comprised in the gridbox. The grid log file was generated using Autogrid v.4.0. The water molecules were deleted, followed by non-polar hydrogen merging, and the Gasteiger charges were conferred to the protein moiety. The rigid docking program was run using a genetic algorithm; the search parameters were set as follows: the population size of 150 at maximum, the number of evaluations at medium, and the maximum number of generations set to 27000. The docking parameter file was used to generate a docking log file through AutoDock 4.0. 48

2.7 Interaction Analysis

Molecular interaction between CSP1 and each of the five natural compounds was individually studied and analyzed using the PMV and chimera applications 51. The binding energy between CSP1 and the five selected natural compounds were analyzed using the PMV of MGL tools 50.

3. ADMET Studies

Five molecules found to fit the aforementioned criteria and possess potent antimicrobial and anti biofilm properties were selected as ligands for further study; these were curcumin, ellagic acid, eugenol, kaempferol, and tinosporin. 53–60. Lipinski’s rule of five was a guideline for determining the druggability of the test molecules. The rule states that for any molecule to work as a drug, it should have less than five hydrogen bond donors and ten hydrogen bond acceptors, a molecular weight less than 500, less than ten rotatable bonds, and a calculated p-log value less than five which corresponds to the octanol-water partition coefficient. Any deviation from the rules leads to lower permeability and poor absorption 61. Therefore, it is critical for a molecule to obey Lipinski’s rule of five to be considered a potential drug candidate. The Ghose filter determines the drug likeliness of a molecule via the following constraints, a p-log value must range between -0.4 and 5.6, the molecular weight must be between 160 and 480, a molecular refractivity between 40 and 130, and the total number of atoms must be between 20 to 70 62. Veber’s rule decrees that two essential criteria must be met for a molecule to be acceptable for oral bioavailability. First, the number of hydrogen bonds in the molecule should not exceed ten, and second, the polar surface area must not exceed 140Å², which subsequently corresponds to the molecule having less than twelve hydrogen donors and acceptors 63. Egan’s rule states that a molecule has good oral bioavailability if they satisfy the p-log value in the range of -1.0 and 5.8 and a topological polar surface area (TPSA) value less than or equal to 130Å² 64. Muegge’s rule edicts for a molecule’s druggability are that the molecular weight must be in the range of 200 to 600, a lipophilicity profile (xlogP3) between -2 and 5, the TPSA less than or equal to 150Å² 65. Additionally, the number of rings must be less than or equal to seven, the number of carbon atoms must be greater than or equal to four; the number of heteroatoms greater than one. The number of rotatable bonds must be lesser than or equal to fifteen. Finally, The number of hydrogen bond acceptors must be lesser than or equal to ten, and hydrogen

3.1 Screening of Biochemical Pathway Involved in QS

Upon screening of the various pathways involved in the QS regulatory mechanism, it was found that the com loci play a significant and integral role in QS of S. pneumoniae D39 strain (Figure 1) 28. Further analysis led to the understanding that the com loci comprises six essential genes, namely comABCDE and comX; these genes encode for two transmembrane proteins, the competence stimulating peptide, the two-component regulatory system, and a transcriptional regulator 18,22. Among the six proteins expressed by the com loci, CSP1 is the autoinducer of the com loci QS mechanism. It creates a QS mediated positive feedback loop involved in constituent gene expression 26. CSP1 is known to be significantly involved in the QS signaling of S. pneumoniae biofilm and also reported to cause autolysis of neighboring bacterial cells in its niche, thus, imbibing virulence and pathogenicity to an otherwise commensal organism 29,31. Hence, CSP1 was considered for further quorum quenching studies using the natural compounds selected as inhibitors.

Fig 1: KEGG pathway of the S. pneumoniae QS pathway.
bond donors must be lesser than or equal to five. Curcumin, kaempferol, and tinosporinone concur with all the discussed five criteria for drug likeliness; however, eugenol and ellagic acid were found to deviate slightly. Eugenol does not comply with Muegge’s rule, having a molecular weight lesser than 200. Ellagic acid similarly digresses from Veber’s rule and Egan’s rule, having a TPSA greater than 140Å. The A-bioavailability score (ABS) is a semiquantitative rule that predicts the degree of oral bioavailability and permeability of a molecule based on Lipinski’s rule of five, TPSA, and total molecular charge. An ABS of 0.55 is assigned to any molecule which passes Lipinski’s rule of five; as such, all five molecules tested for druggability obtained an ABS of 0.55. (Supplementary Data Figure1-5).

3.3 Ligand Preparation

The selected ligands’ 2D structures were sketched then cleaned using the Chem Sketch tool to obtain geometrically correct 2D structures. The 2D structures were converted to 3D upon addition of the 3D coordinates via the Open Babel tool; this conferred preliminary 3D structures with slight deviations in bond length and bond angles. These predicted structures were subsequently corrected using the Argus Lab tool, which yielded stereo chemically fit 3D structures of the ligands.

Fig 2: 3D structures of curcumin, ellagic acid, eugenol, kaempferol, and tinosporinone.
3.4 Protein Structure Retrieval

Upon screening through the RCSB PDB database for CSP structures, 15 structures from the source *S. pneumoniae* were retrieved. Among the 15 structures, 12 structures were found to be of CSP1, and the remaining three were found to be structures of CSP2 protein. Out of the 12 structures of CSP1, two structures were devoid of mutations and ligands; these were PDB ID: 6COW and 6CJ8. PDB ID: 6COW had the least RMSZ value of 1.27 for bond length and 1.43 for the bond angle; hence, this structure was retrieved and used for further studies.

![Fig 3: 3D structures of competence stimulating peptide (CSP1) (PDB ID: 6COW) retrieved from Protein Data Bank.](image)

3.5 Binding Site Prediction

Previously published data provided the evidence that CSP1 acts via its cognate receptor comD1, which subsequently brings about the QS mediated gene expression and regulation. The hydrophobic path of CSP was deciphered by Johnsborg et al. wherein the CSP1 was found to have an amphiphilic \( \alpha \)-helical structure; the amino acids in positions 6-12 contributed the specificity towards the cognate receptor comD1 \(^{31}\). The amphiphilic region of CSP1 comprises the non-polar residues PHE7, PHE8, PHE11, and ILE12 on one side of the helix and LYS56, ARG9, and ASP10 on the opposite side. Mutagenesis studies involving the replacement of the amino acid residues PHE7, PHE8, PHE11, and ILE12 revealed that these particular residues play a critical role in recognition and binding to the comD1 receptor. Furthermore, a recent study by Yang et al. established that the \( \alpha \)-helix spans between LEU4 and LYS16 via high-resolution solution NMR spectroscopy \(^{28}\). The amphiphilic nature of the helix was reiterated, reemphasizing its critical role in receptor binding (supplementary data figures 15 & 16). The specific interaction between the pair of molecules occurs through hydrophobic interactions. The residues of CSP1 involved in this interaction are LEU4, PHE7, PHE8, PHE11, ILE12, LEU13, and ARG3. Since these residues form the binding site for CSP1, it is also a critical site for studying inhibition. Hence, these residues were considered for the interaction studies of CSP1 with the five ligands (Figure 2).

3.6 Grid Generation

The grid encompassing the binding site residues with a dimension of X:61, Y:42, and Z:83 points was generated using Autogrid 4.0.

3.7 Interaction Analysis

All the five compounds, i.e., curcumin, ellagic acid, eugenol, kaempferol, and tinosporinone, interact with CSP1 via hydrophobic interactions at the binding site residues. Curcumin establishes hydrophobic interactions with ILE12, LEU13, GLN14, and LYS17 residues, and a hydrogen bond interaction with LEU13 having bond length of 2.158 Å, respectively (Figures 4 & 5). Ellagic acid specifically interacts with PHE8, PHE11, ILE12, GLN14 and forms hydrogen bonds with PHE and GLN having bond length of 1.740 Å, and 1.884 Å, respectively (Figures 6 & 7). Eugenol binds to PHE11, ILE12, LEU13, GLN14, and ARG15 at the binding pocket (Figure 8 & 9). Whereas, Kaempferol interacts with the residues PHE8, PHE11, ILE12, LEU13, and ARG15 along with the formation of hydrogen bonds with ILE12, GLN14 and ARG15 having bond lengths of 2.136 Å, 2.162 Å, and 1.977 Å, respectively (Figures 10 & 11). Finally, tinosporinone interacts with ARG9, ILE12, and LEU13 residues and a hydrogen bond is formed with ARG9 having bond length of 1.910 Å of the CSP1 binding site; analyzed through UCSF chimera and Ligplot + (v.2.2) (Figure 12 & 13; Table1; Figure 14) (Supplementary Data Figures 6-14).

<table>
<thead>
<tr>
<th>Sl.No.</th>
<th>Compound</th>
<th>Hydrophobic Interaction</th>
<th>Hydrogen Bond Interaction</th>
<th>Hydrogen Bond Length</th>
<th>Binding Energy (Kcal/mol)</th>
<th>Obedience of Lipinski’s rule</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Curcumin</td>
<td>ILE12, LEU13, GLN14, LYS17</td>
<td>LEU13</td>
<td>2.158 Å</td>
<td>-4.2</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>Ellagic Acid</td>
<td>GLN14, PHE8, ILE12, PHE11</td>
<td>PHE, GLN14</td>
<td>1.740 Å, 1.884 Å</td>
<td>-4.7</td>
<td>Yes</td>
</tr>
<tr>
<td>3</td>
<td>Eugenol</td>
<td>ARG15, PHE11, ILE12, LEU13, GLN14</td>
<td></td>
<td>-3.6</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Molecule</td>
<td>Residues</td>
<td>Interaction Distances</td>
<td>Lipinski's Rule</td>
<td>Obedience</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>--------------</td>
<td>----------------</td>
<td>-----------------------</td>
<td>----------------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Kaempferol</td>
<td>PHE11, ILE12, PHE8, LEU13, ARG15, GLN14</td>
<td>ARG15, GLN14, ILE12</td>
<td>2.136 Å, 2.162 Å, 1.977 Å</td>
<td>-5.1</td>
<td>Yes</td>
</tr>
<tr>
<td>5</td>
<td>Tinosporinone</td>
<td>ARG9, ILE12, LEU13, GLN14, LYS16</td>
<td>ARG9</td>
<td>1.910 Å</td>
<td>-4.2</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Obedience to Lipinski's rule indicates that the molecules have MW ≤ 500, MlogP ≤ 4.15, N or O ≤ 10, NH or OH ≤ 5.

Fig 4: Interaction of curcumin with binding site residues of CSP1 depicted in the ribbon model.

Fig 5: Curcumin bound to the binding pocket of CSP1 depicted in the hydrophobicity surface model.

Fig 6: Interaction of ellagic acid with binding site residues of CSP1 depicted in ribbon model.
Fig 7: Ellagic acid bound to the binding pocket of CSP1 depicted in the hydrophobicity surface model.

Fig 8: Interaction of eugenol with binding site residues of CSP1 depicted in the ribbon model.

Fig 9: Eugenol bound to the binding pocket of CSP1 depicted in the hydrophobicity surface model.
Fig 10: Interaction of kaempferol with binding site residues of CSP1 depicted in the ribbon model.

Fig 11: Kaempferol bound to the binding pocket of CSP1 depicted in the hydrophobicity surface model.

Fig 12: Interaction of tinosporinone with binding site residues of CSP1 depicted in the ribbon model.
Fig 13: Tinosporinone bound to the binding pocket of CSP1 depicted in the hydrophobicity surface model.

Fig 14: Ligplot exhibiting interactions of curcumin, ellagic acid, eugenol, kaempferol, and tinosporinone with the binding site residues of CSP1.
4. CONCLUSION

The inhibition of CSP1 presented to be a prospective antimicrobial approach since quorum-quenching can avert virulence and resistance. Upon screening, five plant-derived molecules were selected for study based on the ADME profile; these were curcumin, ellagic acid, eugenol, kaempferol, and tinosporinone. The compounds were found to have credible drug likeliness with no acute toxicity and satisfactory bioavailability score. The molecular docking studies between CSP1 and the selected compounds revealed a satisfactory interaction as inhibitors; the binding energy ranged between -3.6 to -5.1. Curcumin, ellagic acid, eugenol, kaempferol, and tinosporinone act as potential inhibitors of CSP1.

5. ACKNOWLEDGMENT

MCN and CS thank the JSS Academy of Higher Education & Research for funding this work under the JSSAHER Student Research Grant 2019 (REG/DIR(R)/URG/54/2011-12). MCN, CS, and GS thank Glimetomics Bioresolve Private Limited for its enduring scientific support and aid in conducting this research study.

6. FUNDING ACKNOWLEDGEMENT

MCN and CS acknowledge and thank JSS Academy of Higher Education & Research for funding this work under the JSSAHER Student Research Grant 2019 (REG/DIR(R)/URG/54/2011-12).

7. AUTHORS CONTRIBUTION STATEMENT

CS guided planned and designed the work with MCN. MCN and GS have executed the work and gathered the data. Further data processing and analysis was carried out by MCN and GS. All authors contributed to final manuscript.

8. CONFLICT OF INTEREST

Conflict of interest declared none.
9. REFERENCE


