International Journal of Advanced Research and Review

www.ijarr.in

MOLECULAR STRUCTURE AND INTERACTION STUDIES OF MECA AND SIGX IN *STREPTOCOCCUS PYOGENES*

P.C. Kunthavai, Preethi Ragunathan*

Centre of Advanced Study in Crystallography and Biophysics, University of Madras, Chennai (India)
Email: preeth_182003@yahoo.co.in/ preethiragu@gmail.com

ABSTRACT

Streptococcus pyogenes is a stern human pathogen causing deadly infections humans. Natural competence and the ability of *S. pyogenes* to grow under varied environmental conditions is mediated by alternate sigma factor SigX. However, SigX has been reported to be negatively regulated by MecA, an adapter protein in *S. mutans*. In order to understand the interaction of MecA and SigX at the molecular level in *S. pyogenes*, the structures of both these proteins are imperative. This study aims at modelling the structures of SigX and MecA and understand their mechanism of interaction by protein-protein docking studies.

KEYWORDS: Streptococcus, alternate sigma factor, gene regulation, *ab initio* modelling, protein-protein docking

INTRODUCTION

The natural competence pathway or acquisition of exogenous DNA through conjugation, transduction and transformation in bacteria has contributed to their evolution. In Gram positive bacteria, this genetic transformation is well studied in *Bacillus subtilis* and *Streptococcus pneumoniae*. Competence development in *S. pneumoniae* is signalled by a quorum sensing mechanism that consists of three components: a signal peptide and a two-component regulatory system. This two- component system consists of a membrane-bound histidine kinase sensor and an intracellular response regulator [1].

Alternate sigma factors of RNA polymerase regulate bacterial gene expression under varied environmental conditions. In *S. pneumoniae*, ComX regulates competence for gene transformation [2]. In *S. mutants*, comX is involved in stress tolerance and bacterial biofilm formation [3]. In Streptococci, the early genes of competence development include a quorum-sensing system that activates the alternative sigma factor *sigX*, the key regulator of competence. SigX, homolog of ComX recognizes a conserved DNA sequence, the CIN box present upstream of the late competence genes necessary for DNA uptake and recombination. Together with the core subunits of RNA polymerase, SigX binds to the CIN box and promotes transcription of the late genes. It is also well established that natural competence is usually regulated by a set of regulators that form a cascade of reactions in the early to late stage of competence development[4].

SigX, besides the transcriptional control of comX, is also the target of post-translational regulation of Streptococci. *In S. pneumoniae*, SigX is positively regulated by ComW while it is negatively regulated by ClpE/-P, because ComW is capable of simulating the activity of SigX and also protect it against the proteolytic activity of ClpE/-P [5]. In *S. mutans*, lot of studies on the regulatory cascade leading to transcriptional activation of *comX* had been carried out. In this study, genetic approaches have shown that MecA, an adapter protein acts as a negative regulator of genetic competence in *S. mutans* grown in a complex medium. The negative regulation of competence by MecA involves a mechanism that requires the presence of the

functional SigX [6]. Recently, a similar regulatory mechanism involving acquisition of SigX by a SigX-MecA-ClpC complex has been reported in *Streptococcus thermophilus*, in which the type I ComRS system primarily regulates the expression of *comX* [Figure 1] for genetic competence[6]. However, it is still unclear whether MecA in other organisms, particularly in *Streptococcus pyogenes* (GAS) plays a functional role in the regulation of SigX activity and competence. *S. pyogenes* is a pathogenic beta-hemolytic bacterium known to cause acute pharyngitis and other life-threatening complications such as toxic shock syndrome (TSS), glomerular nephritis, rheumatic fever and necrotizing fasciitis [7]. This study aims to understand the role of MecA and its interaction of SigX in transcriptional regulation by *in silico* methods.

MATERIALS AND METHODS

Structure modelling of MecA and SigX

Structural information is of great aid in the study of protein function, dynamics, interactions with ligands and other proteins. Homology modelling is the best tool for modelling the 3-D structures of proteins from known templates. However, in cases where homology modelling fails to yield good results, *ab initio* or *de novo* modelling of protein structures are used to build three dimensional protein models from the scratch. The structures of MecA and SigX of *S. pyogenes* were built using *ab intio* modelling technique. The sequences of MecA and SigX were retrieved from EXPASY (Swiss Institute of Bioinformatics, Switzerland) proteomics server[8]. A large ensemble of structural conformations(decoys) were generated by simulations in I-Tasser [9]. The final models were selected by clustering the decoys based on the pair-wise structure similarity, and five best models which corresponds to the five largest structure clusters were chosen based on their confidence score or C-score. The best models obtained were further refined and energy minimized using the online server YASARA [9] to remove bad contacts and steric clashes.

Validation of 3D structures

The structural validation and protein geometry of the modelled structures were carried out using RAMPAGE, by analysing the phi-psi torsion angles of amino acids with Ramachandran Plot and predicting the stereochemical quality of the proteins. The native fold of the proteins, surface energy, pairwise energy and z scores were analysed using ProSAII protein structure analysis server [10].

Active site prediction and docking studies

The Active sites of the proteins were predicted by CASTp [11]. Protein-protein docking studies were carried out between MecA and SigX using Z-Dock 3.0.2 [12]. The amino acids outside the active site region were blocked using block.pl. Docked poses were created using create.pl. A total of 1000 docked structures of MecA-SigX complex were generated and ranked based on their z-score. The best ranked complexes were analysed further. The proteins and their interactions were visualized using PYMOL software [13].

RESULTS AND DISCUSSION

The crystal structures of MecA and SigX are not available in the PDB. The three dimensional structures of SigX and MecA, their active site and the protein-protein interactions are studied here.

3D model generation and validation of MecA and SigX

The MecA adapter protein (Uniprot id:P60187) belongs to the MecA superfamily of proteins and is 253 residues long. The conserved domains in MecA was analysed using the Conserved

Domain Database (CDD) of NCBI and indicates the entire protein (253 amino acids) belongs to the PRK02315 family of adaptor proteins. A BLAST search of MecA against PDB structures failed to yield a good template for homology modelling and was not successful.

SigX, an RNA polymerase sigma factor of S. pyogenes is 141 amino acid residues long and belongs to the Sigma70-ECF family. The conserved domains of SigX, identified from CDD (NCBI) are shown in Fig. The first domain (Sigma70-ECF family) lies between Arg48 and Tyr141. The second domain (DNA-directed RNA polymerase specialized sigma subunit, sigma24 family) constitutes the residues Asp66 to Tyr141; the third domain (DNA-directed RNA polymerase specialized sigma subunit, sigma24 family [Transcription]) ranges between Ser46 to Lys139. The fourth domain (RNA polymerase factor sigma-70; validated) constitutes residues between Glu78 to Lys139.

A BLAST search of the sequence of SigX retrieved several homologs with very high sequence identity, including hypothetical proteins from *S. suis*, *Tyzerella nexilis*, Blautia sp. etc. However, none of the structures of the identical proteins were solved. Hence, *ab initio* modelling was carried out for both MecA and SigX using I-Tasser. The models with the best confidence score were chosen.

The obtained models were further energy minimized and stereochemical corrections wre carried out using Yasara energy minimization server. Ramachandran plots were used to evaluate the stereochemical quality of the predicted structures. The graph of residue by residue torsion angles (phi-psi plot) for both the proteins is shown in Figure 2. The plot statistics shows that 98% and 98.6% of the amino acid residues fall in the allowed regions, for MecA and SigX respectively. The statistics of Ramachandran plot suggests that the generated protein structure is of good stereochemical quality.

The ProSA server tool gives pairwise energy and surface energy values of proteins based on the mean potential as a function of the sequence and folding energies. The overall z-score for MecA was -5.36 and SigX was -5.63. The z-score indicates the model quality of the overall fold. The z-scores and the negative energies of the models (Figure 3), suggested by ProSA is indicative that the generated models are of good quality and could be used for further studies.

The overall structural model of MecA consists of 13 helices, 2 beta sheets, 1 beta hairpin, 15 helix-helix interactions and 75 beta turns and 12 gamma turns while the structure of SigX is composed of 6 helices, 17 helix-helix turns and 5 beta turns. The cartoon representation of MecA and SigX are shown in Figure 4 and 5.

Active site prediction

The active sites of MecA and SigX were predicted using CASTp server. 29 cavities were predicted for MecA while SigX did not show a binding pocket in CASTp. The largest cavity in MecA with a volume of 7104 $Å^3$ is formed between Glu2 to Thr14, Met59 to Ser96, Phe103 to Tyr194, and Ala215 to Glu251. Since structural information highlighting the binding sites are not available from any known structures, the predicted pockets could be the possible binding sites with its cognate sigma factors.

MecA-Sigx protein protein docking studies

Experimental evidence suggests that MecA is a negative regulator of SigX in *S. mutans*. However, the exact mechanism of interaction between them is not known. To study the same in *S. pyogenes*, protein-protein docking studies between MecA and SigX were carried out using ZDOCK 3.0.2. This helped us to predict the amino acid residues involved in complex formation. The docking results were further validated with solvent accessible surface area calculations (SASA), computed using VADAR [14]. The ASA of MecA was 14045 $Å^2$ and that of SigX was 8109 $Å^2$. However, the ASA of the complex was 19264 $Å^2$, and much lower. An electrostatic surface representation of the complex (Figure 6) shows that the positive and

negative charge distribution complements each other. The protein-protein interactions are listed in Table 1 and the docked complex is shown in Figure 7. The key interactions at the binding interface of the complementary regions are the hydrogen bonds, salt bridges, possibly cation-pi interactions and a large number of non-bonded interactions. The residues Val64, Lys189, Tyr194, Ile250, Gly253, Phe62, Ile250, Asp56, Glu38 and Glu35 interact of MecA interact with Tyr89, Tyr60, Glu86, Asp56, Trp53, Tyr89, Gln90, Lys101, Gln108 and Tyr141 of SigX. Interestingly, these residues fall within the binding pockets predicted by CastP in MecA. The protein-protein docking analyses have pointed out the residues involved in both MecA and SigX. These residues could further be useful for identifying novel lead compounds against GAS.

CONCLUSIONS

The structure of MecA and SigX have been obtained using computational tools and validated. The domains of both proteins have been identified and the key residues involved in protein-protein interaction have been analysed using our docking results. Although, the exact mechanism of negative regulation of SigX by MecA is yet to understood experimentally, our studies have thrown insights into the possible mechanism of MecA-SigX interactions in *S. pyogenes.*

CONFLICT OF INTEREST: None

REFERENCES

1. Kowalko JE, Sebert ME. The *Streptococcus pneumoniae* Competence Regulatory System Influences Respiratory Tract Colonization. Infection and Immunity.2008; 76: 3131-3140.
2. Luo P, Morrison DA. Transient association of an alternative sigma factor, ComX, with RNA polymerase during the period of competence for genetic transformation in *Streptococcus pneumoniae*. Journal of Bacterilogy. 2003; 185:349-358.
3. Aspiras MB, Ellen RP, Cvitkovitch DG. ComX activity of *Streptococcus mutans* growing in biofilms. FEMS Microbiology Letters. 2004; 238**:**167-174.
4. Reck M, Tomasch J, Wagner-Döbler I. The alternative sigma factor SigX controls bacteriocin synthesis and competence, the two quorum sensing regulated traits in *Streptococcus mutans*. PLoS Genetics. 2015; 11(7):e1005353.
5. Chastanet A, Prudhomme M, Claverys J P& amp; Msadek T. Regulation of *Streptococcus pneumoniae* clp genes and their role in competence development and stress survival. Journal of Bacteriology. 2001; 183: 7295-7307.
6. Tian XL, Dong G, Liu T, Gomez ZA, Wahl A, Hols P, Li YH. MecA protein acts as a negative regulator of genetic competence in *Streptococcus mutans*. Journal of Bacteriology. 2013; 195: 5196-5206.
7. Carapetis, Jonathan R, Andrew C Steer, Kim Mulholland E, Martin Weber. The global burden of group A streptococcal diseases. The Lancet: Infectious Diseases. 2005; 5: 685-694.
8. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Research. 2003; 31:3784-3788.
9. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. 2010; Nature protocols. 5: 725-738.
10. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. 2007. Nucleic Acids Research. 2007; 35: W407-410.

11. Binkowski T, Naghibzadeh S, Liang J. CASTp: computed atlas of surface topography of proteins. Nucleic Acids Research. 2003; 31:3352-3355.
12. Pierce BG, Hourai Y, Weng Z. Accelerating protein docking in ZDOCK using an advanced 3D Convolution Library. PLoS ONE. 2011;6:e24657.
13. Lill MA, Danielson ML. Computer-aided drug design platform using PyMOL. Journal of Computer Aided Molecular Design. 2011; 25:13-19.

Table 1: List of hydrogen bonds and salt bridges observed in the docked complex of MecA and SigX

Hydrogen bonds	
S. no	**Residue name and number (MecA: SigX)**
1.	Val 64 [N]: Tyr 89[OH]
2.	Lys 189 [NZ]: Tyr60 [OH]
3.	Lys 189 [NZ]: Glu 86 [OE2]
4.	Tyr 194 [OH]: Glu86 [OE]
5.	Ile 250[N]: Asp56 [OD1]
6.	Gly 253 [O]: Trp53 [NE1]
7.	Phe 62 [O]: Tyr89 [OH]
8.	Ile 250[O]: Gln90[NE2]
9.	Asp 56 [OD2]: Lys101[NZ]
10.	Glu 38 [O]: Gln108 [NE2]
11.	Glu 35 [OE2]: Tyr141 [OH]
Salt bridges	
1.	Lys 189 [NZ]: Tyr60 [OH]
2.	Asp 56 [OD2]: Lys101 [NZ]

Figures

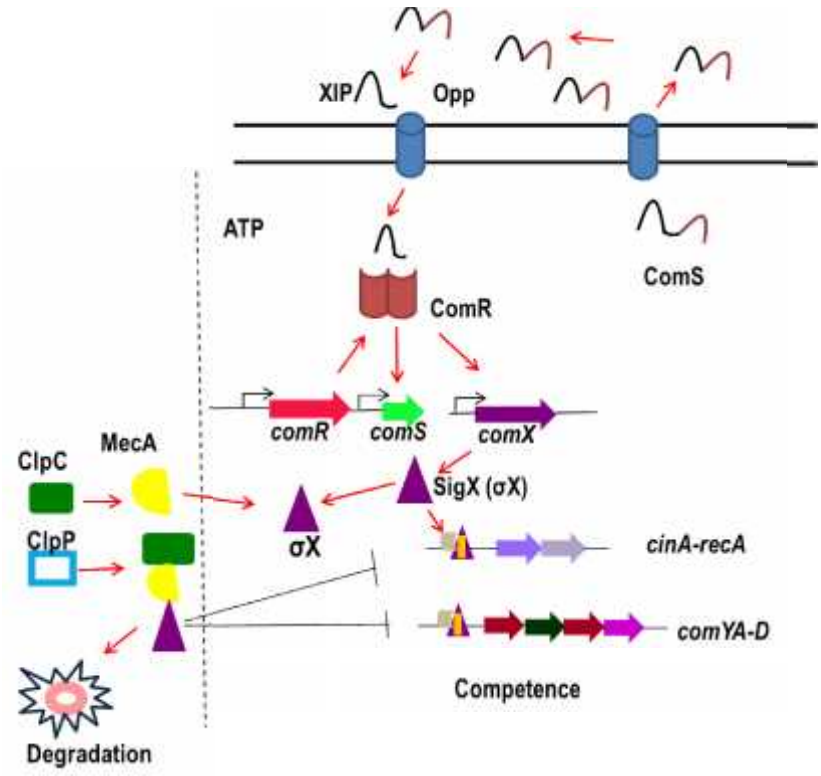

Figure 1. Diagrammatic representation showing regulatory mechanism of SigX by SigX-MecA-ClpP complex in *Streptococcus thermophilus*

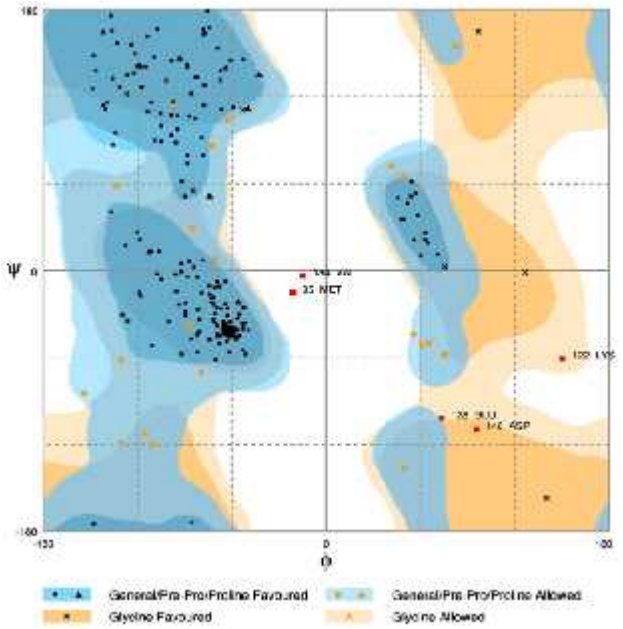

Figure 2: Conformation analysis of MecA protein using RAMPAGE. The plot shows that the model is of good stereochemical quality.

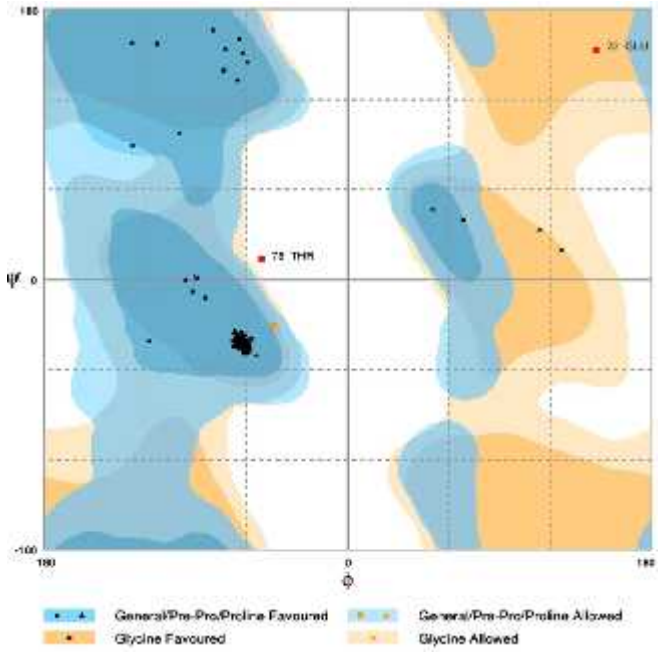

Figure 3: Conformation analysis of SigX protein using RAMPAGE. The plot shows that the model is of good stereochemical quality.

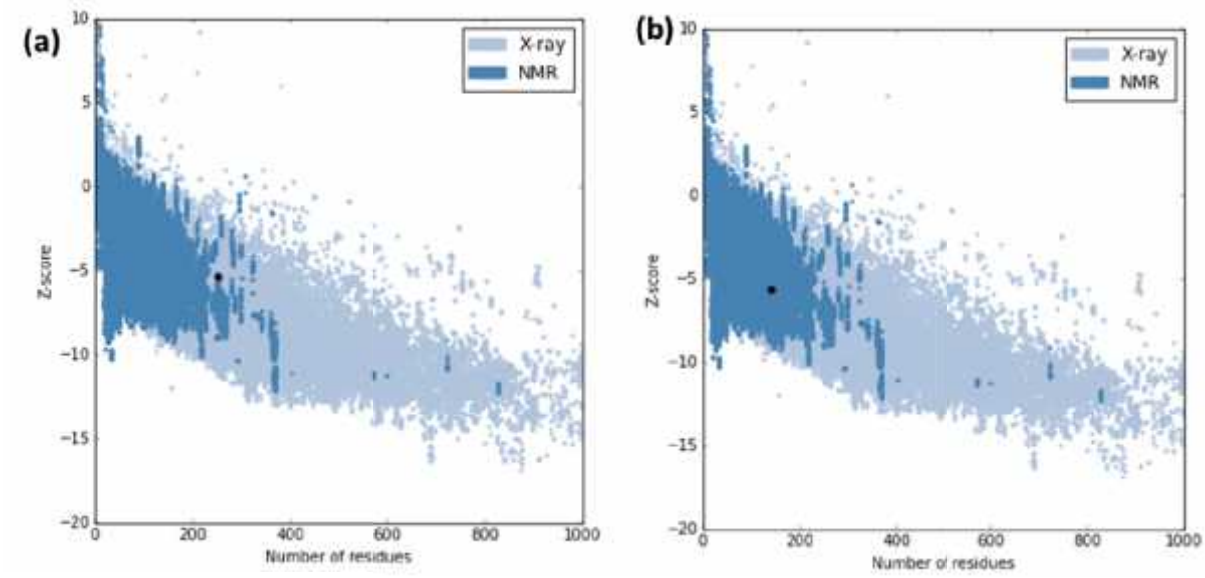

Figure 4: The overall model quality of (a) MecA and (b) SigX protein in comparison with the experimentally determined native proteins of the same size. Dark blue dots NMR structures, light blue dots X-ray structures and the black dot SigF protein with z score value of -5.36 and -5.63 respectively are represented.

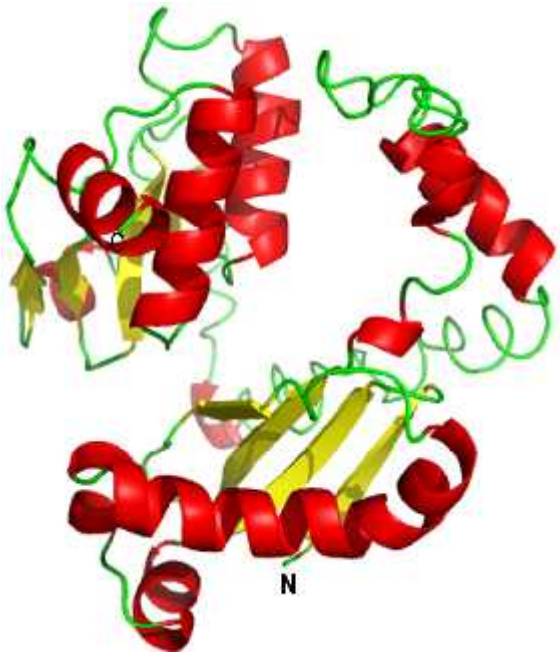

Figure 5: Three dimensional structure of MecA protein built by *ab initio* modelling. Helices are represented in red and beta strands in yellow.

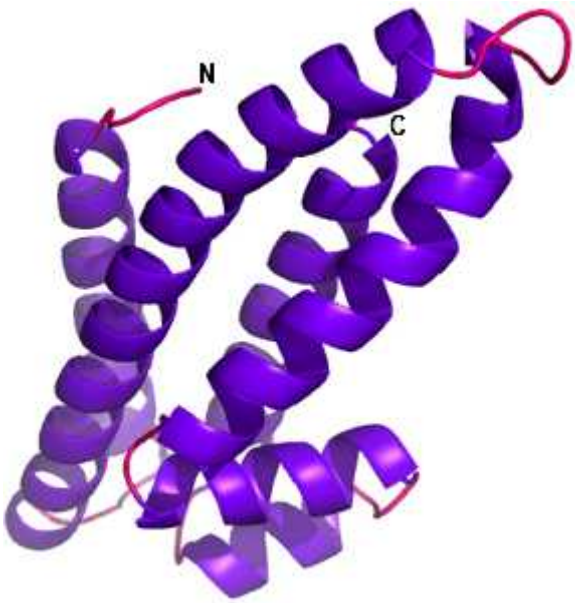

Figure 6: Three dimensional structure of SigX protein built by *ab initio* modelling. Helices are represented in purple and loops in pink.

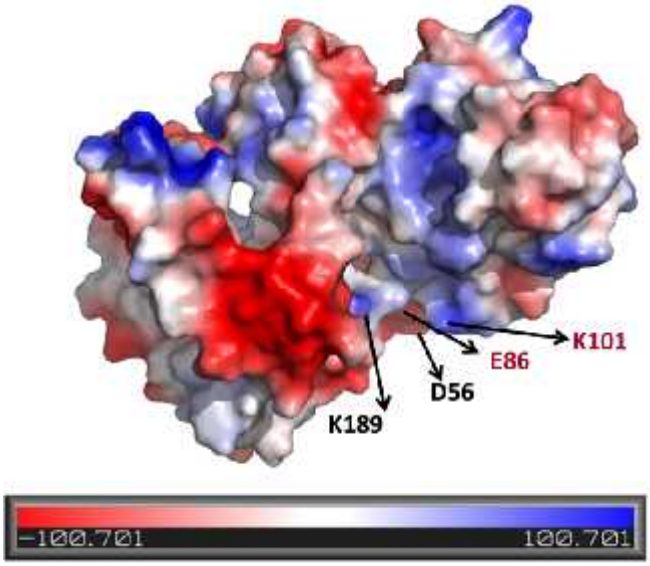

Figure 7: Electrostatic surface representation of MecA-SigX docked complex. Some of the residues at the protein-protein interface are highlighted in MecA (black) and SigX (brown).

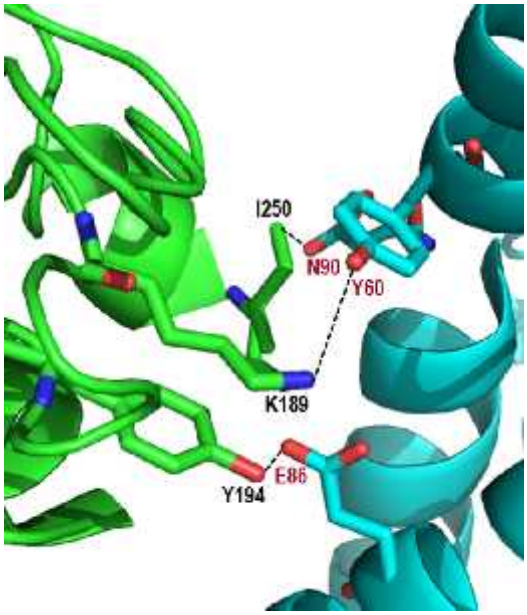

Figure 8: Close-up view of MecA-SigX complex at the interface showing some of the hydrogen bonds. Binding site residues of MecA are shown in green and SigX in blue. The hydrogen bonds are represented as dotted lines.

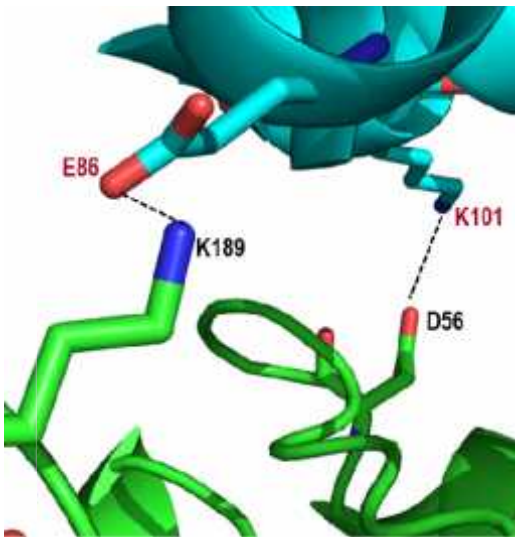

Figure 9: Close-up view of MecA-SigX complex at the interface showing the two salt bridges. Binding site residues of MecA are shown in green and SigX in blue. The salt briges are represented as dotted lines.