Available online www.jocpr.com

Journal of Chemical and Pharmaceutical Research, 2016, 8(7):520-524

Research Article

ISSN : 0975-7384
CODEN(USA) : JCPRC5

Computational analysis of coagulase protein from *Staphylococcus Sp. Cobs2Tis23*

Dibyajyoti Ozah[1], Jatin Kalita[1], Prasenjit Manna[1], Tapan Dey[1], Robin Das[1], Robin Kumar Sharma[2] and Balagopalan Unni[3]

[1]Biotechnology Division, CSIR- North East Institute of Science & Technology, Jorhat – 785006, Assam, India
[2]School of Biological Sciences, University of Science & Technology Meghalaya, 9th Mile - 793101, Meghalaya, India
[3]Biological Sciences, Assam Downtown University, Guwahati- 781026, Assam, India

ABSTRACT

Some of the Staphylococcus species were shown to have Coagulase properties. Earlier we reported a strain from Staphylococcus family named Staphylococcus sp. Cobs2Tis23 which possesses Coagulase properties. Since structure prediction by the bioinformatics tools depending upon its amino acid sequence is very important to have a preliminary knowledge about the characteristics of the protein before the experimental evidences by structural studies, we tried to predict the structural conformation as well as some biochemical characteristics of the coagulase extracellular protein present in Staphylococcus sp. Cobs2Tis23 strain. A range of bioinformatics programmes like Expasy, TranSeq ToxinPred, PROCHECK, ERRAT, Discovery Studio 4.0 were used for the structure prediction as well as the biochemical characterisation of the protein. A Ramachandran plot validated structure is predicted out for the Coagulase protein which is found to be nontoxic and have the molecular weight of 50049.7 Da. Prediction of the structural conformation of the protein also shows the protein have the isoelectric pH of 9.68 and has three binding sites.

Keywords: Staphylocoagulase, in Silico Characterization, Structure, Toxicity Prediction.

INTRODUCTION

Staphylocoagulase is an extracellular protein produced by several *Staphylococcus* strains, a potent human pathogenic bacterium responsible for a wide range of acute and pyogenic infections, including abscesses, central nervous system infections, endocarditis, osteomyelitis, pneumonia, urinary tract infections and several syndromes caused by exotoxins and endotoxins, besides food poisoning, scalded skin and toxic shock syndromes [1]. A few Staphylococcal strains are also found as a main cause of hospital-acquired (nosocomial) infections of surgical wounds and infections related to indwelling medical devices [2]. The study about these *Staphylococcus* strains is also important due to the spread of antibiotic resistance [3]**.** Moreover, coagulase can act as a systemic hemostatic agent [4]. The ability of certain bacterial species to coagulate the plasmas of certain animals was first described by Loeb in 1903 [5]. Coagulase is an extracellular protein that reacts with prothrombin in the blood and the interaction enables the conversion of fibrinogen to fibrin that results in clotting of blood [6]. Evidence of two forms of *Staphylococcal* Coagulase was reported so far. One is bound to the cell wall which is responsible for the slide test coagulation and the other is liberated as free coagulase into the culture medium and is responsible for the tube test [7]. Study about coagulase enzyme has become important owing to have more understanding in pathogenicity of *staphylococcus*. This will also give better understanding about the contribution of the enzyme in development of *staphylococcal* infection.

Nowadays, in the field of biological research, the theoretical models are the best representative of experimental protein structures to reduce cost and time for a getting 3D structure of a protein. The protein homology modelling or comparative modelling refers to constructing a three-dimensional model of a target protein sequence based on its similarity to one or more known experimental three-dimensional structure of related homologous proteins. A high-quality structural model can be generated through homology modelling when the targets and templates are closely related which has inspired the formation of structural genomics consortium dedicated to the production of representative experimental structures for all classes of protein folds. In this present study, we tried to predict the staphylocoagulase structure as well as some biochemical characteristics through *in silico* studies.

EXPERIMENTAL SECTION

Few bioinformatics tools and database were used to search out some insight and structure prediction of the coagulase protein extracted from the *Staphylococcal* strain. The preliminary characterization including the amino acid composition, molecular weight, theoretical pI, atomic composition, instability index, aliphatic index and grand average of hydropathicity (GRAVY) of the protein can be calculated through theoretical computation from available amino-acid sequences. In the present study, the primary characterization of Staphylocoagulase was done using the expasy proteomics tools program at http://au.expasy.ch/tools.

The sequence of coagulase in *Staphylococcus* strain was obtained from NCBI having the accession number "EU246837.1_4" of *Staphylococcus Sp. Cobs2Tis23* 16S ribosomal RNA gene, partial sequence; in FASTA format and used the "Transeq" tool *(http://www.ebi.ac.uk/Tools/st/emboss_transeq/)* for conversion to amino acid sequence. We have chosen all the six open reading frames and selected standard codon table during launching of the tool, keeping other parameters by default.

The query sequence of coagulase from *Staphylococcus* was searched to find out the related protein structure to be used as a template (s) by the BLAST program *[www.ncbi.nlm.nih.gov/BLAST]* against Protein Data Bank database with default parameters. As in our template search, we found the templates having less than 35% identity, so we preferred the multiple template based model generation approach of MODELER 9 version8 software. MODELLER implements comparative protein structure modelling by satisfaction of spatial restraints and can perform many additional tasks, including de novo modelling of loops in protein structures [8]. The best templates found during PDB-BLAST were used as multiple templates.

Different software tools and Meta servers are available for validation and evaluation of a generated model. A structural evaluation of the generated structure of the Staphylcoagulase was performed by using two programs called PROCHECK and ERRAT. The predicted model was submitted to the structure evaluation at Meta server "SAVES" [http://nihserver.mbi.ucla.edu/SAVES/].

There are numerous *in silico* tools available, which can predict the immunogenicity and toxicity of the peptides. Here we have used the *ToxinPred* tool, a unique *in silico* method useful in predicting toxicity of peptides/proteins developed by CSIR-IMT, Chandigarh. In addition, it will also be useful in designing least toxic peptides and discovering toxic regions within the proteins [9]. Here, we have used the three modules of designing peptides, batch submission and protein scanning for toxicity analysis.

The binding sites of the Staphylocoagulase were predicted though Discovery Studio 4.0 Molecule Viewer which predicts the possible binding sites in receptor protein.

RESULTS AND DISCUSSION

The nucleotide sequence from NCBI database having accession number "EU246837.1_4" of *Staphylococcus* sp. *Cobs2Tis23* 16S ribosomal RNA gene, partial sequence; is converted to its corresponding amino acid sequence by using the "Transeq" tool *(http://www.ebi.ac.uk/ Tools/st /emboss_transeq/)* under EMBOSS. The number of amino acids was found to be 460.

From the preliminary prediction, the coagulase protein showed the Instability Index around 50.42. The isoelectric point (pI) and theoretical molecular weight of Staphylocoagulase was determined to be 9.68 and 50049.7 Da, respectively. The result of Protparam programme also shows the protein has the aliphatic index of 75.7 and grand average of hydropathicity (GRAVY) of -0.031.

While doing PDB- BLAST of the sequence, three templates as 1BXC (Chain A, Xylose isomerase from *Thermus caldophilus*), 1N5Y (Chain L, Hiv-1 Reverse Transcriptase Crosslinked to Post- Translocation Aztmp-Terminated Dna (Complex P) [*Mus musculus*]) and 1SZJ of (Chain G, Structure of Holo-Glyceraldehyde-3-Phosphate-Dehydrogenase from *Palinurus versicolor*) were selected having identity (32%,26% and 30%) and E value (0.067, 3.3, 9.2) are chosen for the model building.

During model assessment the stereo chemical and geometrical parameter of the generated models (Figure 1) were evaluated using different online and offline bioinformatics tools and meta servers. The QMEAN, dDFIRE & ERRAT values for the generated structure are recorded as 0.097, 734.35 KJ/mol and 26.386, respectively. In Ramachandran analysis (Figure 2), a good distribution of 460 amino acid residues of the generated model of Staphylocoagulase was found where about 95.1% residues were obtained in most favored and additionally allowed region, and only 2.3 % residues are in generously allowed region. During Disorder prediction study the protein did not showed any disordered region for the given sequence.

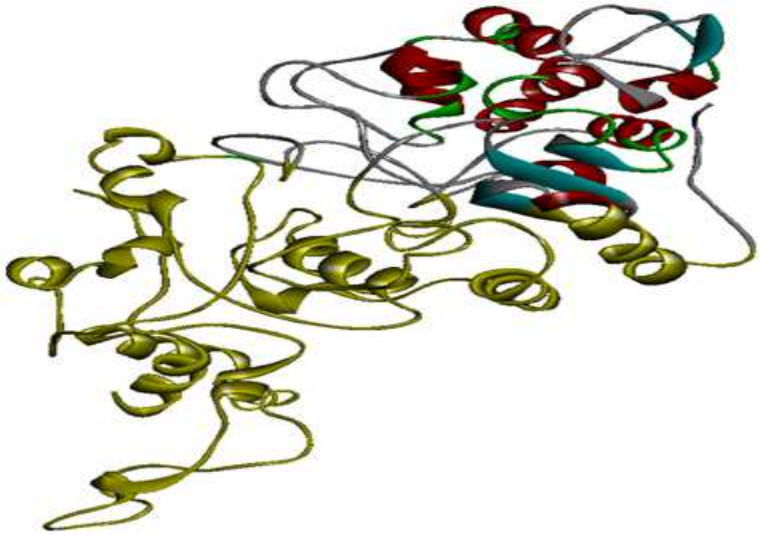

Figure 1: 3D Structure of the coagulase protein generated by MODELER

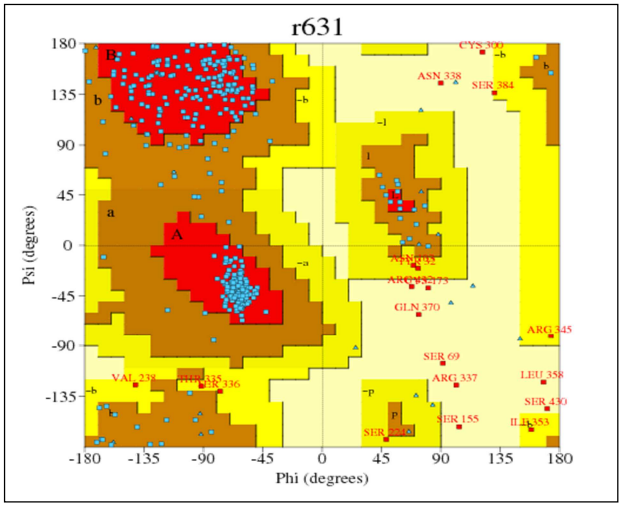

Figure 2: Ramachandran plot of the generated model of the coagulase protein

From the present *in silico* toxicity prediction studies, the concerned protein did not shows any mutation or disorder region (Figure 3) and found to be non-toxic (Figure 4) .

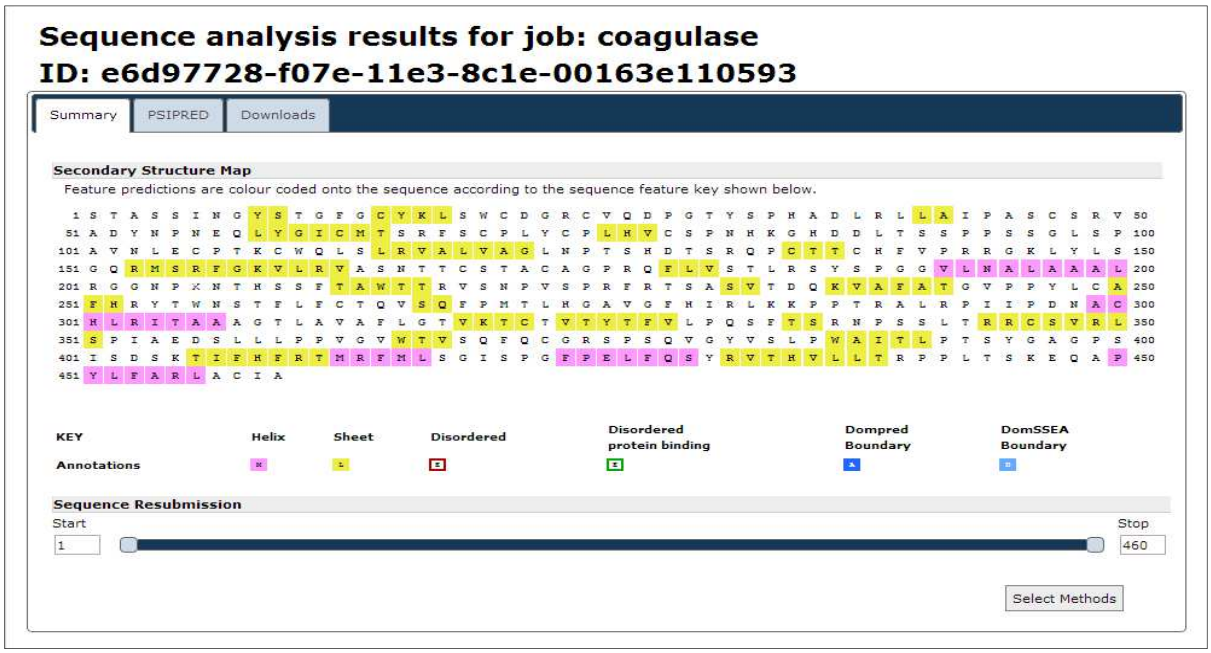

Figure 3: Results of Disorder Prediction for protein disordered region detection

Mutation Position	SVM score	Prediction	Hydrophobicity	Hydropathicity	Hydrophilicity	Charge
No Mutation	-0.80	Non-Toxin	-0.21	-0.45	0.10	49.50

Figure 4: Results of Toxin Prediction for Toxicity of the coagulase protein

The best three binding sites we found in our study were shown have following amino acids in the respective positions (Figure 5). Summary of the volume as well as the 3D direction measures were shown in Table 1.

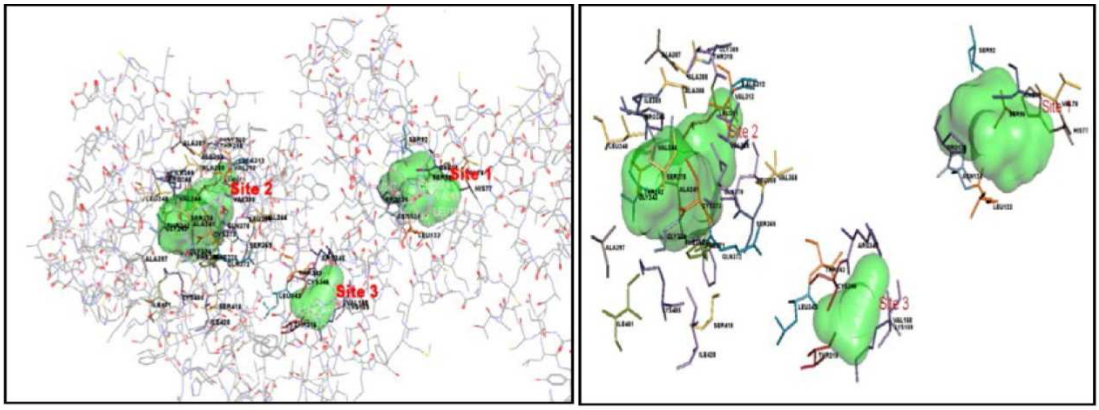

Figure 5: Active sites of generated protein model with residues view through Discovery Studio 4.0
Cavity 1:Lue 123, Asn 124,Pro 125,Ser 92,Ser 96,Cys 79,Val 78,His 77.
Cavity 2:Leu 311,Val 380,Ala 312,Val 313,Ser 378,Ala 241,Thr242,Gly 243,Lev 248,Pro 245,Val 244,Phe 371,Gly.
Cavity 3: Lys 159, Val 160, Thr 343, Arg 345, Leu 342, Cys 346, Thr 216

Table 1: Summary of the Coagulase proteins predicted 3 active binding sites

Protein Binding Sites	Position			Volume (Å)
	X- plane	Y- plane	Z- plane	
Cavity 1	24.153	-13.459	-12.97	42.375
Cavity 2	18.153	-13.209	-43.97	103.625
Cavity 3	32.153	-11.209	-30.22	6.875

In this modern era of computational biology, structure prediction is very important to predict the characteristics as well as its mechanism of action of a protein. Homology modeling is one of the most reliable approaches used nowadays to predict the structure from an amino acid sequence using known 3D structure of a protein having similar amino acid sequences. Structure of a protein not only reveals its characteristics but also helps to understand the mechanism of action of the respective proteins.

Staphylococcus is a commonly encountered organism of human body, causing serious infections. Most of the species of *Staphylococcus* strains were toxic to human beings. But as per the prediction studies the soil *Staphylococcus* strain Cobs2Tis23 from where we isolated the staphylocoagulase protein is found to be non toxic. The total numbers of amino acids were found to be 460. The structure we predicted from these 460 amino acids is found to possess three most active binding sites. The structure was validated with plotting the values of φ and ψ angles in the Ramachandran plot. 95.1% residues are found to be in most favored and additionally allowed region of the plot. The structure was further validated with different bioinformatics programs such as QMEAN, dDFIRE & ERRAT which have respective values of 0.097, 734.35 KJ/mol and 26.386. These values also validate our predicted structure of the coagulase protein.

The active site prediction is very important to know the mechanism of action of a particular protein. From our prediction studies, we found three most active binding sites of the staphylocoagulase protein.

Acknowledgements
Authors are thankful to the Director, CSIR-North-East Institute of Science & Technology, Jorhat, Assam (India), for constant support and encouragement and to the Council of Scientific & Industrial Research (CSIR) for financial support under Network Project (BSC0111).

REFERENCES

[1] Lindsay JA; Ruzin A; Ross HF; Kurepina N; Novick RP. *Mol. Microbiol.*, **1998**, 29, 527-543.
[2] Archer GL. *Clin Infect Dis.*, **1998**, 26, 1179-1181.
[3] Mohamed NA; Mohamed R; Chong TT. *Bioinfo.*, **2012**, 8, 412-414.
[4] Mojovic B; Mojovic N; Tager M; Drummond MC. *Yale J. Biol Med.*, **1969**, 42, 11–20.
[5] Loeb L. *J. Med Res.*, **1903**, 10, 407-419.
[6] Miale JB. *Am Soc of Hematology.*, **1949**, 4, 1039-1048.
[7] Duthie ES. *J. Gen Microbiol.*, **1954**, 10, 427- 436.
[8] Eswar N; Marti-Renom MA; Webb B; Madhusudhan MS; Eramian D; Shen M; Pieper U; Sali A. *Current Protocols in Bioinformatics*, John Wiley & Sons, Inc. Supplement, **2006**, 15, 5.6.1-5.6.30.
[9] Gupta S; Kapoor P; Chaudhary K; Gautam A; Kumar R; et al. *PLoS One*, **2013**, 8, e73957.