Structural insight into PNAG binding and partial De-*N*-acetylation modification site of IcaB protein in *Staphylococcus epidermidis*

Ramachandira Prabu[1*], Mohanty Amaresh Kumar[2], Raziya Shaik[3], Sarada R.[4] and Veeramani T.[1]
1. Department of Biotechnology, Center for Research and Development, PRIST university, Vallam, Thanjavur. Tamil Nadu – 613403, INDIA
2. Department of Bioinformatics, Centre for Bioinformatics, School of Life Sciences, Pondicherry University, Puducherry - 605 014, INDIA
3. Department of Chemistry, Visakha Institute of Engg. and Technology, Visakhapatnam 530027, INDIA
4. Department of Chemistry, Anil Neerukonda Institute of Technology and sciences, Visakhapatnam 531162, INDIA
*rcprabu86@gmail.com

Abstract

De-N-acetylation of Poly-N-acetylglucosamine (dPNAG) is an extracellular exopolysaccharide for biofilm development. Staphylococcus epidermidis, IcaB is poly-beta-1,6-N-acetyl-D-glucosamine N-deacetylase belonging to CE4s family of carbohydrate esterase. It overlaps with IcaC gene at 3'-end. IcaB present in periplasmic region of cell membrane required for partial de-N-acetylation of PNAG, yet details of crystal structure are unknown. We have characterized IcaB gene sequence first which may act as promoter for IcaC gene regulation by enhancing or silencing in Staphylococcus epidermidis. Protein threading was used for construction of truncated IcaB three-dimensional structure. Post-translational modification site as acetylation, ubiquitination, phosphorylation except glycosylation was commonly found in IcaB protein sequence.

DXD, DXH, NXS, HD, RXXR signature sequences were required for IcaB metal binding and enzyme catalytic activity. Binding free energy was calculated for IcaB-GluNAc ligand docking to form hydrogen bond with critical amino acids showing ΔG score −11.95 KJ/mol using Autodock 4.2.6. Ligand docking and molecular dynamics simulation (MDS) were performed to evaluate the stability of IcaB, IcaB-GlcNAc binding complexes packed within active site amino acids throughout trajectories captured with time scale 100 ns period using GROMACS 4.5. Different binding energies were calculated for IcaB-GlcNAc complex using GROMACS tools MM-PBSA with perceptive to van der Waal energy, electrostatic energy and polar solvation energy, SASA energy.

Keywords: IcaB, Biofilm, CE4s family, De-N-acetyalase, dPNAG, Protein threading; Molecular dynamics simulation, *Staphylococcus epidermidis*.

Introduction

Extracellular polysaccharides had wide range of structural functions and biofilm properties producing inside bacterial cells[40]. IcaB is De-N- deacetylase, whose function removes 15 to 20 % of *N*-acetyl group in random pattern of PNAG converting to dPNAG and acting as protective exopolymer supporting anionic interaction within cells and other surface essential for biofilm development and evasion of host immune system[6,57]. *S. epidermidis* IcaB deacetylation activity was demonstrated by using size exclusion chromatography-mass spectrometry for dPNAG modification[45]. 80% of polysaccharide are in dPNAG form which is glucosaminoglycan, linear in structure whereas other form of 6% polysaccharides was found to be anionic constituting succinyl groups (*O*-succinlyation modification)[45]. IcaB protein was performed for overexpression and purification for understating molecular mechanism and metal-dependent de-N-acetylase activity of *S. epidermidis*[33].

IcaB steady state enzyme kinetics displays a distinct low to high catalytic efficiency with respect to different oligomeric substrates such as trisaccharides, tetrasaccharides pentasaccharides, hexasaccharides. In pentasacccharide substrate, IcaB specifically acts on second residue of reducing terminus, whereas in hexasaccharides substrate, IcaB specifically acts on second and third residue of reducing terminus[33]. In *Escherichia coli*, PgaB protein function similar like IcaB is metal-dependent enzyme involved in catalytic activity with divalent cation for partial De-*N*-acetylation of β-1,6-GlcNAc oligomers[27]. PgaB structure comprises of two (β/α) barrel domains belonging to CE4 family of carbohydrate esterases at N-terminal site and conformation, identical to glycoside hydrolases[27].

PgaB is involved in β-1,6-GlcNAc oligomers substrate specificity of de-*N*-acetylase activity in length dependent manner and never acts on other non-specific substrates such as β-1,4-$(GlcNAc)_4$ oligomer chitotetraose[27]. Gene deletion of *pgaB* leads to prevention of biofilm formation and retention of full synthesis of PNAG without any partial de-*N*-acetylated form. PgaB contains lipidation domain responsible for partial de-*N*-acetylation of PNAG and N-terminal domain for CE4s function[12]. CE4s is essential for deacetylation of *O*- and *N*-acetylated polysaccharides coordinated with divalent metal ions.

PgaBC-terminal region is showing TIM barrel domain of glycoside hydrolases homology with COG1649 domain for PNAG hydrolysis or binding in *Yersinia pestis*[13]. Overexpressing *icaB* strains produces highly acetylated form of PNAG than deacetylated form of PNAG, had an

enhanced survival advantage, more susceptible to opsonophagocytosis in murine bacteremia model. This highly acetylated form of PNAG acts as opsonin by preventing opsonic killing of human monoclonal antibody (mAb)[43]. Isogenic *icaB* mutant strains, which results in non-deacetylated poly-acetylglucosamine was devoid of bacterial surface attachment due to its loss of cationic character[45]. Murine model was taken for persistence of virulence activity with perceptive to device-related infection, which impaired biofilm formation significantly due to *icaB* mutant strain[45].

O-linked binding of GlcNAc to threonine and serine amino acids residues controls transcription factors and intracellular proteins such as p53, c-myc, NFκB[26]. GlcNAc provokes virulence gene expression and induces changes in morphogenesis in candida albicans, a fungal pathogen of human [32]. Production of antibiotics and changes in cell process in soil bacterial was stimulated by GlcNAc[32]. GlcNAc binds to residues of fucose extracellular domains leading to ligand specific interaction which gets altered for notch family receptors[32]. When compared to wild type CSF41498, IcaB substitution mutant O7 showed higher IcaA transcription, increased IcaB activity which facilitates phenotype and enhanced PIA synthesis[39].

Addition role of IcaB protein revealed that extracellular DNA (eDNA) and functional poly-*N*-acetylglucosamine (PIA) polysaccharide work together for biofilm mediated infection[28]. IcaB protein was detected in bacterial culture medium revealed from Immunofluorescence microscopy. IcaA, IcaD, IcaC are present in the fraction of cell membrane, whereas IcaB protein is present in the cytosol culture supernatant and continues deacetylatase activity after membrane localization[45]. GBSPII-1, new polysaccharide of 34Kda from *G. biloba* acts against biofilm formation identified *in vitro* in *S. areus* through inhibiting icaB gene expression[31].

Crystal structure of IcaB has been undermined from *Ammonifex degensii* (IcaBAd), Gram negative bacteria and C-terminal domain catalytic pocket is required for efficient deacetylation activity[48]. Protein threading methodology has been assigned for structural and functional features of IcaB protein of *S. epidermidis*. Predicted binding patterns of GlcNAc ligand molecule with docking binding energy score were manually verified using molecular visualisation tools. Following MD simulation with IcaB and IcaB-GlcNAc complexes in GROMACS 4.5, trajectory period of molecular simulation revealed complete structural stability of IcaB and IcaB-GlcNAc complex protein. To cross-check, binding affinity of IcaB-GlcNAc complexes was investigated with prodigy server.

Material and Methods

IcaB protein characterization: *S. epidermidis* RP62A strain was chosen and full length *IcaB* gene sequence was retrieved from GenBank (http://www.ncbi.nlm.nih.gov/ BLAST). RXXR, NXS, DXH, invariant signature sequence was identified from IcaB protein from literature. IcaB predicted protein-protein interaction done with STRING database (https://string-db.org/). Scratch protein predictor was utilized for *in silico* bacterial protein overexpression, protein solubility, protein disorder region (http:// scratch.proteomics.ics.uci.edu/). MusiteDeep tools was utilized for protein posttranslational modifications site prediction (https://www.musite.net/). Database of transcriptional regulation in *Bacillus subtilis* (DBTBS) was utilized for identification of promoter cis-acting sequence and trans-acting elements (https://dbtbs.hgc.jp/). Metal ion-binding site prediction and docking Server (MIB) were performed for divalent ion binding site for IcaB protein (http://bioinfo.cmu.edu.tw/MIB/).

IcaB protein 3D structure prediction: IcaB truncated region taken for three-dimensional protein structure prediction was built by threading method on I-tasser, excluding N-terminal transmembrane peptide region (https://zhanggroup.org/ I-TASSER/). The predicted structure was evaluated for loop refinement by ModBase tool (http://modbase.compbio.ucsf.edu/modloop/)[15]. The predicted model was performed with SAVES [http://nihserver.mbi.ucla.edu/SAVES/]. ProFunc database was utilized for secondary structure prediction by ProMotif program and biochemical function of IcaB protein [http:// www.ebi.ac.uk/thornton-srv/databases/ProFunc/]. The transmembrane model quality score was analysed from QMEAN and QMEANBrane online server[7]. PyMOL was utilized for 3D structure visualization[46].

IcaB- protein-ligand docking: IcaB-GlcNAc blind and focused docking was performed using Schrodinger software[37]. GlcNAc ligand structure (N-Acetyl-beta-D-glucosamine) and molecular formula ($C_8H_{15}NO_6$) were obtained from PubChem site[24,39] in SDF file format (http://pubchem.ncbi.nlm.nih.gov/). IcaB and GlcNAc conversion were prepared with protein preparation and LigPrep module of schrodinger suite[37]. Grid-box was generated to cover entire active site of IcaB protein structure. UCFC chimera molecular visualization tool was utilized to get insights into docking pattern and LIGPLOT was utilized to obtain 2D interaction plots (https://www.ebi.ac.uk/ thornton-srv/software/LIGPLOT/).

Molecular dynamics simulations: Protein stability of IcaB and IcaB-GlcNAc ligand complex by molecular dynamics simulation was performed with Gromacs 5.1.4[42,58]. Automated topology builder (ATB) was utilised to create topology files for GlcNAc. IcaB and IcaB-GlcNAc complex solvated with SPC water molecule in cubic shape box placed at the center[24]. 1 nm distance was given between simulation box edge and IcaB-GlcNAc complex, to move freely in the immersed state completely. Electrostatic energy calculation was evaluated by Particle Mesh Ewald (PME) algorithms. Ewald summation and covalent bond constraints used to solve linear constraint solver (LINCS) algorithms[53]. Four chloride ions (Cl-) were added to neutralize the system.

Energy minimization of IcaB and IcaB-GlcNAc complex proceeded with steepest descent approach at 100 ps. Constant pressure, volume, temperature (300 K) were maintained to system equilibration at 100 ps with IcaB-GlcNAc complex. 100 ns time scale was set for final MD simulation run and trajectories were saved for further analysis. Root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration, hydrogen bond graph, solvent accessible surface was generated using Xmgrace tools[42]. Different energy calculation was performed using MMPBSA from protein-ligand complex file.

Results

Characterization of IcaB intercellular adhesion protein: IcaB gene accession no. U43366.1 and protein ID. AAC06118.1 were derived from NCBI. IcaB gene sequence encodes 870 nucleotides. icaB gene may behave as promoter for IcaC gene expression for transcription +1 start site (TSS) positioning at 857th nucleotide i.e. 14 nt overlapped sequence of IcaB gene at 3'-end (Supplementary data). IcaC protein as O-succinyltransferase family of protein and gene sequence overlap with *icaB* at 3'-end. PUR, AhrC, SigD, SigE, invariant N-Box cis-regulatory region are present in *icaB* gene sequence. This may act as promoter region for *icaC* gene regulation with stringent spatial-temporal manner under proximal promoter range of transcription start site (TSS). SigD transcription factor binding site identified at –86 to -113 nt (5'-AAAAAGCTGTCACACCAGATGCC GATAA–3') from upstream region of *icaC* TSS. AhrC transcription factor binding site was at -142 to -153 nt (5'-ATGAAAAAAAAT–3').

PurR transcription factor binding site was at –233 to -246 nt (5'-AAAAGCGAACACTA-3'). SigE transcription factor binding site was at –338 to -367 nt (5'- AATGTATGAA ACAGGCTTATGGGACTTTGA–3'). Enhancer box (N-Box) transcription factor binding site was at –770 to -775 nt (5'-CACCAG –3'). Predicted disorder amino acid residues as GNKSK position were at Gly190th to Lys194th and as EEK position at Glu287th to Lys289th of IcaB protein with score ≥ 0.5. DAD, DMD belonged to aspartate–any residue–aspartate (DXD) identified at different position as DAD (Asp259 – Asp261), DMD (Asp121–Asp123).

Deduced sumoylation is member of small ubiquitin-like modifier (SUMO) predicted in IcaB lysine position at 40 and 242. Acetylation PTM acetyl group may be attached to ε-amino group of lysine residues at position 104, 242. Lysine ubiquitination PTM position at lysine 114 and 123 may regulate protein stability and membrane trafficking[20]. Protein phosphorylation reversible PTM is occurring at tyrosine positon 95, threonine at position 61,137, 282,285, 401, serine amino acid at position 231, 356, 401,403, 404, lysine at position 286.

Structural and functional insights of IcaB protein: IcaB three-dimensional structure was retrieved from I-TASSER sever for threading based structure prediction, due to lack of specific PDB structure from PDB-PSI BLAST. 259 aa truncated regions at position Asn31 to Lys289 has been taken for protein structure prediction, excluding N-terminal transmembrane sequence position at Met1 to Ala30 (Figure 1a). I-Tasser generated 4 out of 10 PDB-ID belonging to 4f9dA hits. 4f9dA PDB-ID is orientated with respective to PgaB. 3 out of 10 PDB-ID belongs to 4wcj hits with normalized Z-score as 4.78. 4wcj is orientated with respective to structure of IcaB from *A. degensii*. Out of five model, fourth model was selected from c-score -1.06. Ramachandran plot shows that 96.1% amino acid residues were found in the favored region, 2.3% amino acids in allowed regions and 1.5% in the outlier region (Figure 1b).

IcaB PDB secondary structure were predicted as 20 β turns, 2 ϒ turns, 7 Helix-Helix interacs, 9 helices, 10 strands, 2 β-α-β units, 4 sheets. IcaB structure was visualized with QMembrane tools revealing Barel like shaped in the periplasmic region anchored with the inner membrane (Figure 1c).

Overall quality of IcaB protein three dimensional model assessment from QMEAN score was 0.55 ± 0.05. IcaB cleft, tunnel, pore analysis reveals that interior surface of funnel was found to be higher. IcaB predicted solubility upon overexpression in bacteria host with range is 0.532130 and is found to be insoluble in nature.

Active site prediction and protein-ligand docking of IcaB protein: Predicted IcaB protein-protein interaction envisages that IcaA, IcaC, IcaD, IcaR, bhp showing significant overall score are 0.999, 0.987, 0.959, 0.897,0.708 and least neighbourhood partners are atlE, aap, SERP0563, SERP2137, SERP2281. Co-occurence partners are IcaA, IcaC showing significant expression, whereas Co-expression partners are IcaA showing very least expression on IcaC, bhp, SERP0563, SERP2137 (Figure 2). The predicted IcaB active site amino acids were Asp121, Asp123, Thr125, Phe144, Ile146 hits with blind docking coherent with predicted active site amino acid. Asp121 and Asp123 are key catalytic residues which may involve in de-*N*-acetylation reported from other CE4s domain. IcaB-GlcNAc docking with significant binding energy found as -11.95 kcal/mol (Figure 3). Docking pose analysis revealed 7 hydrogen bonds (H_bonds) interactions between UDP-GlcNAc and active site amino acids residues of protein.

Here, we observed single H_bond with Asp123, Thr125, Phe144, Ile146 residues of bond length 3.14 Å, 2.46 Å,2.87 Å, 2.85 Å, while three H_bond formation were observed with Asp121 of bond length 2.52 Å, 3.15 Å, 2.87 Å. Asp123 functional carboxylic acid group form hydrogen bonding with 6th carbon atom of hydroxyl group of pyronase ring of β-N-acetyl-D-glucosamine. Ile146 nitrogen group of α carbon bond was attached with 6th carbon atom of hydroxyl group. Phe144 nitrogen group of α carbon bond was attached with 2nd carbon atom of hydroxyl group.

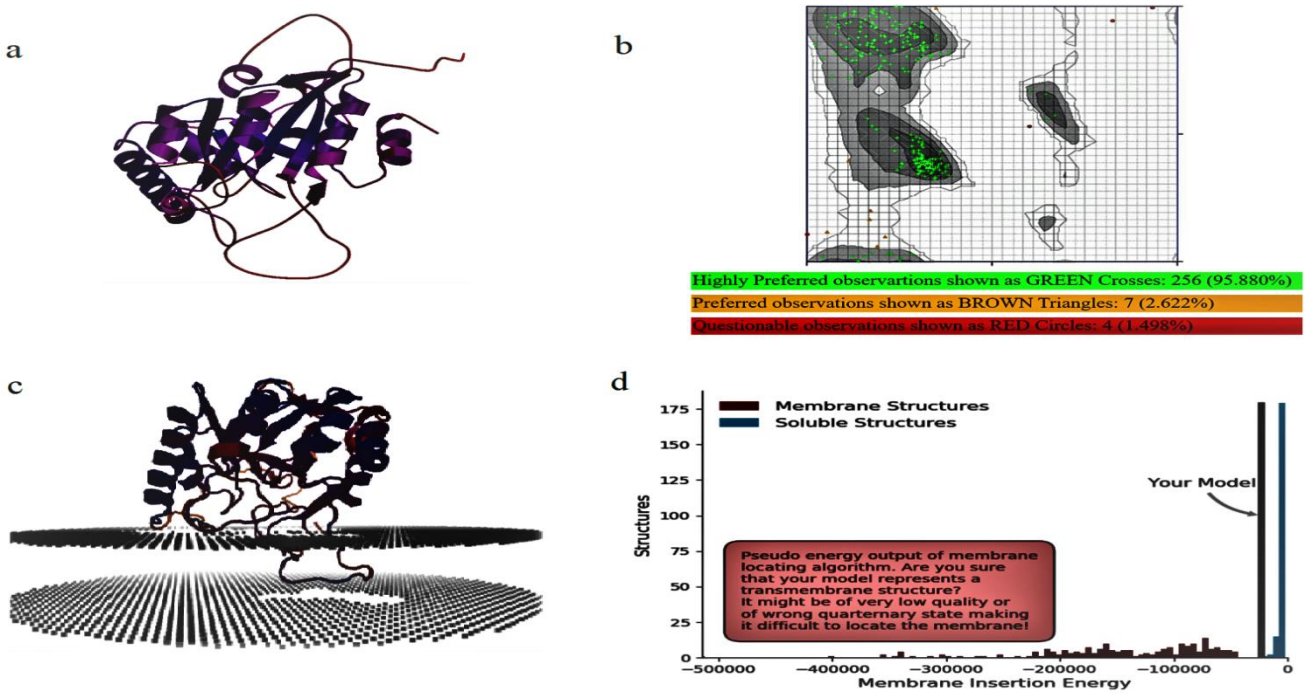

Figure 1: a. Predicted three-dimensional structure of IcaB protein, b. Ramachandran plot, c. Transmembrane model quality assessment with cytoplasmic membrane, d. QMEANBane model scoring

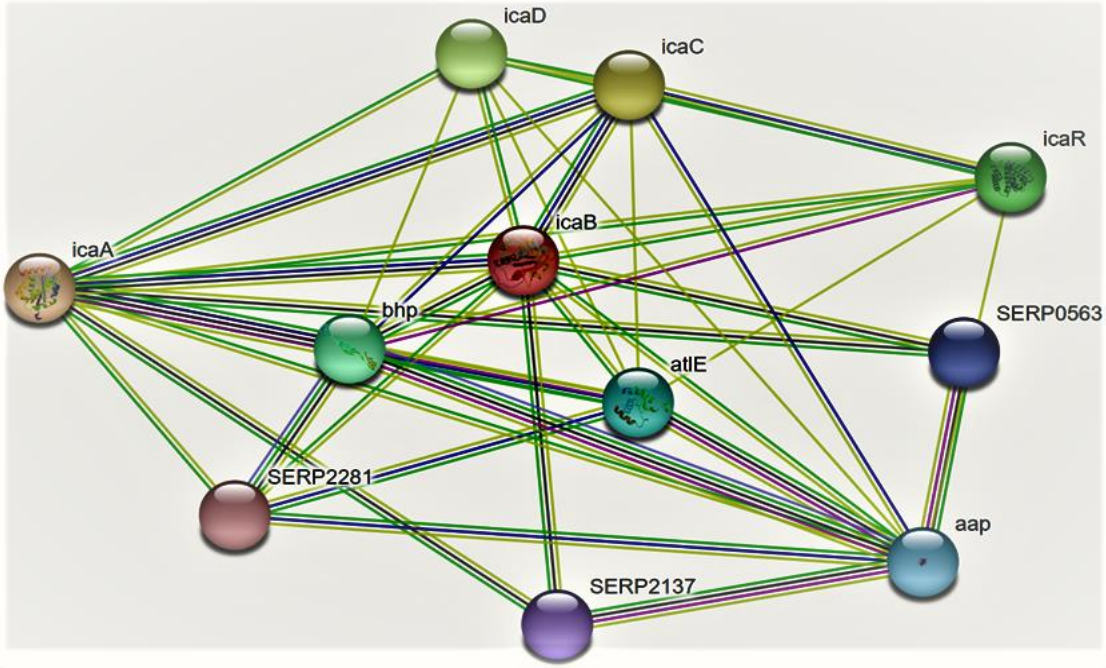

Figure 2: Screenshot from STRING protein-protein interaction of IcaB protein, results showing a set of 11 proteins involved in biofilm formation. The protein nodes corresponding to this pathway are highlighted in color.

Asp121 nitrogen group of α carbon bond was attached with 4th carbon atom of hydroxyl group. Asp121 functional carboxylic acid group of α carbon bond was attached with 4th carbon atom of hydroxyl group. Thr125 oxygen atom at ϒ1 position bond was attached with 3th carbon atom of hydroxyl group. Hydrophobic interaction was observed with Phe119, Gln124, Gly143, Leu145, Leu166, IIe 150 amino acids showing an arc with radiating spokes towards GlcNAc atom with contacts showing spokes radiating back (Figure 3).

Molecular dynamics simulations of IcaB, IcaB-GlcNAc: Root mean square deviations (RMSD) was calculated for IcaB, IcaB-GlcNAc complex. The graphs were generated for protein backbone flexibility. Throughout simulation period, protein backbone stability was not disturbed and strong interaction with IcaB-GlcNAc complex at 15 ns and IcaB at 14 ns, with no significant fluctuations was seen. Root Mean Square Fluctuation (RMSF) was evaluated for the residual mobility of IcaB and IcaB-GlcNAc molecule against residue number and graph was generated. High fluctuation was

observed for IcaB protein from 25 to 50 ns and IcaB-GlcNAc complex from 150 to 175 ns during trajectory period of MD simulation (Figure 4). The radius of gyration (Rg) was calculated as high secondary structure compactness and shown decreasing from 10,000 ps to 100,000 ps throughout 100 ns trajectory period of MD simulation.

Rg graph starts at Y-axis 1.95 nm of IcaB and IcaB-GlcNAc complex, eventually reaching to the lowest range to 1.89 and 1.93 nm (Figure 4). To determine hydrogen bond stability of IcaB-GlcNAc, binding sites were monitored throughout trajectory time scale 100 ns. The analysis revealed that H_bonds graph starts at Y-axis with the number of hydrogen bond from 0 to 6 and Y-axis with 0–100 ns period. Three hydrogen bond profiles were shown with 0–1, 1–2, 2–3, 3-4, 4-5, 5-6. H_bond profile 0–1 range was calculated with no significant fluctuation and breaking was observed. Profile 1–2 range was calculated with no significant fluctuation, however, least H_bond breakage was observed.

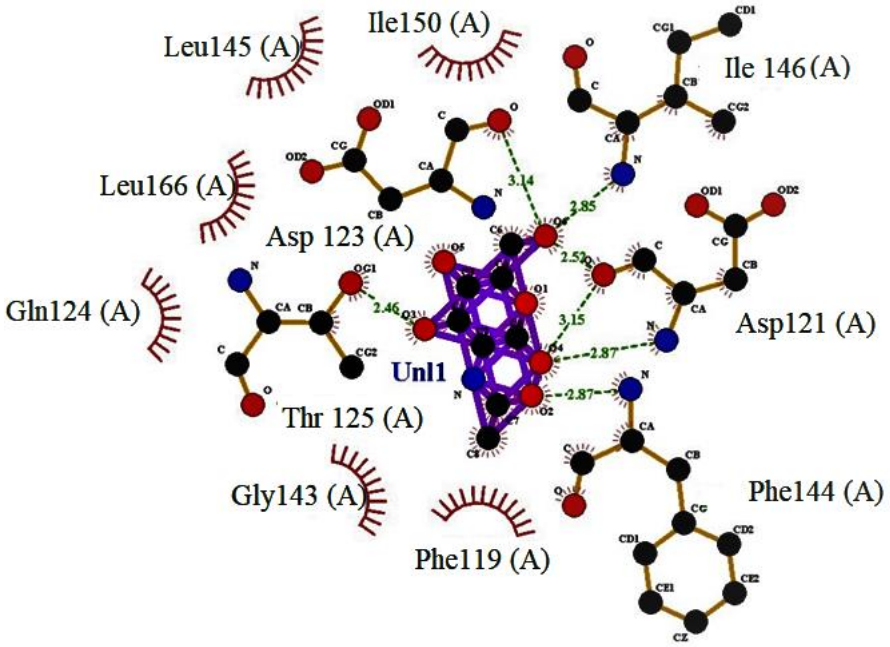

Figure 3: IcaB-GlcNAc docking.

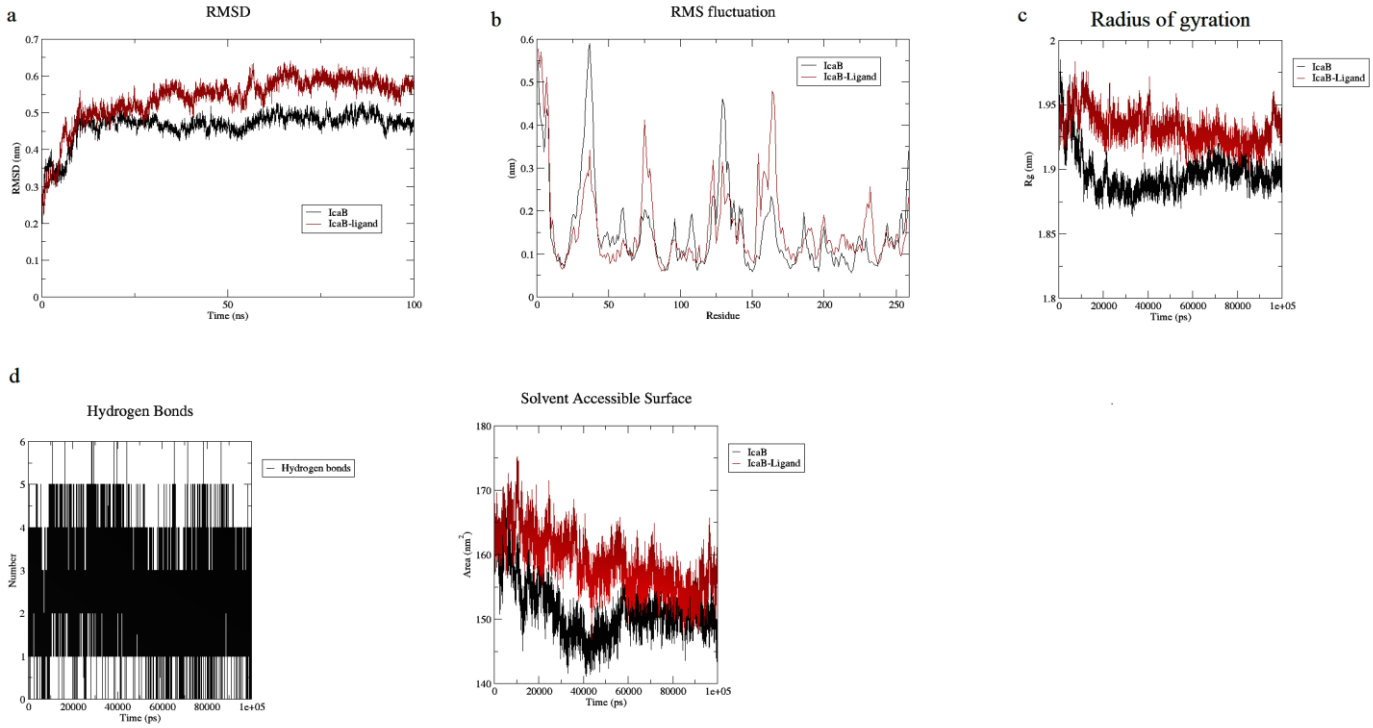

Figure 4: IcaB and IcaB-GlcNAc MD simulation a. RMSD, b. RMS fluctuation, c. Radius of gyration, d. Hydrogen bonds, e. Solvent accessible surface.

Profile 2–3 range has a significant range of fluctuation and one H_bond breakage was observed. Profile 3–4, 4-5, 5-6 range has a significant range of fluctuation and higher H_bond breakage was observed (Figure 4). Solvent accessibility surface area results were observed that IcaB and IcaB-GlcNAc reside at 165 nm2 in the simulation graph throughout trajectory time scale 100 ns IcaB and IcaB-GlcNAc complex showed a gradual decrease and reached to limit of 149 nm2 and 157 nm2 at 100 ns (Figure 4). For icaB-GlcNAc complex, van der Waal energy is -90.684 +/- 12.802 kJ/mol, electrostatic energy is -237.251 +/- 33.353 kJ/mol, polar solvation energy is 788.381 +/- 33.438 kJ/mol, SASA energy is -10.309 +/- 0.531 kJ/mol, SAV energy is 0.000 +/- 0.000 kJ/mol, WCA energy is 0.000 +/- 0.000 kJ/mol and binding energy is -450.137 +/-21.030 kJ/mol.

Discussion

In *S. epidermidis*, PNAG deacetylation modification plays key pathogenic role in bacterial colonization, protects against neutrophil phagocytosis and biofilm formation, virulence against human antibacterial peptides[45]. There is strong association with promoter regulatory sequence and level of RNA transcripts synthesis. IcaC plays a significant role in succinate addition in growing PNAG during biosynthesis[16]. σD factor is holoenzyme component of RNA polymerase by enhancing mycolate synthesis and peptidoglycan conformation modification provides resistance to environmental stress in *Corynebacterium glutamicum*[47]. *Bacillus subtilis* controlled the level of L-arginine by AhrC, transcriptional regulator. AhrC protein binds catabolic gene clusters operator region of promoter resulting in repression of L-arginine biosynthetic genes and enhancing catabolic gene clusters transcription[38].

PurR transcriptional factor comes under LacI family of metabolite-responsive regulators and plays a critical role in purine biosynthesis regulation, metabolism and virulence in *S. aureus*[19]. PurR transcription factor binds directly to promoter genes acting as virulence responsive master regulators and pathogenic factors to control their gene expression[19]. *PurR* gene inactivation leads to enhanced pathogenic traits in mice model with perceptive to bloodstream infection. Subunits of sigma factors belong to RNA polymerase of bacteria which are responsible for promoter binding region at specific site and control RNA transcription efficiently[30]. *HEY1* gene belongs to basic helix–loop–helix proteins (bHLH) with YRPW motif which binds specifically to N-Box CACNAG motifs[14]. Human IDE promoter of Notch targets genes.

One of the key active site amino acids as aspartic acid at conserved DMD motifs plays a central role in De-N-acetylation activity. Beta/alpha barrel catalytic domains involved metal-binding co-ordination for De-N-acetylase. Quantitative analysis of bacterial membrane proteins is still challenging due to dynamical entities of protein structure modification[4]. IcaB predicted protein solubility was found to be overexpressed in the *E. coli* system and insoluable in nature due to post-translational modification site. It was hypothesized that *icaB* gene silencing or expression involved in natural occurrence of insertion sequence IS256 through excision or insertion in the *ica* locus responsible for bacterial phase variation[29,41].

Histidine–aspartate (HD) domain belongs to metalloprotein superfamily and predicted phosphohydrolase activity involved in stress response, virulence, inflammation, nucleic acid metabolism, signal transduction[23,55]. HD coordinates at least one metal ions in mono-, di- and trinuclear conformation catalyse range of hydrolysis to oxygenation reaction[55]. Predicted Asn-Lys-Ser (NYS) invariant sequence may necessitate hydrophobic surrounding interaction with perceptive to membrane protein fold to packing within phospholipid bilayer of IcaB protein [25,60]. Ca^{2+}-dependent serine protease enzyme called furin is member of SPC family (subtilisin-like proprotein convertase) playing a significant part in distinct pathological response through proteolytic activation of proteins with invariant RXXR motif within reactive site loop constituting arginine at P1 site[59] .

DLH invariant signature sequence present in various phosphoesterases family of enzymes for metal ion binding and catalysis deployed phosphate esters hydrolysis [49]. Asparagine residue is vital for interaction between oppositely charged and hydrophobic amino acids within α-helical topological sector[17]. Sumoylation is strongly regulated and reversible PTM involved in covalent attachment for membrane trafficking, intracellular trafficking, signal transduction and membrane structure stabilization[50]. Acetylation PTM plays vital role in cellular process, stress factor detrimental and pathogenesis. Protein phosphorylation involved in transition from apolar state to polar state and promoted association with other molecule via kinase activity for activation which may involve in cell activation, deactivation and regulation[22].

Lack of crystal structure for intercellular adhesion protein has been major limitation of understanding *S. epidermidis* biofilm formation and quorum sensing. 45 amino acid constitutes of lysine and arginine in which side chains positive charge attracts higher water and phosphate molecule in the cell membrane and may form predominate H-bonding to Lys/Arg-phosphate clusters[36,54]. The hydrophilic amino acids side chain of lysine, arginine and histidine also play vital function perturbation and disrupts phospholipids on transmembrane helix model contributing to protein stability at greater net charge reported in other *Staphylococcus spp.*[2,3,44,56] In *E.coli*, attenuated activity was reported with perceptive to low levels of de-*N*-acetylation of PNAG due to the absence of aspartic acid residues PgaB.

IcaB surface exposed loop is essential for anchors and membrane association involved during PIA biosynthesis[48]. However, *icaADB* produces dPNAG by unknown mechanism for survival advantage of *Staphylococcus aureus* fitness. *icaB* is most commonly expressing gene among the

other potent genes revealed from *Staphylococcus aureus* clinical isolates. In methicillin-susceptible *S. aureus* (MSSA), PCR experiment was deployed for identifying frequency of *icaADBC* genes available, it was found that *icaB* genes were majorly detected as 76% compared with *IcaA, icaD, IcaC* constituted as 71%, 41% and 72%[11].

icaB metabolic pathway recruits *icaADC* intercellular adhesion operon undergoing spatial-temporal expression during extreme environmental habitat for bacterial biofilm formation, whereas co-expression and co-occurrence partners as IcaA and IcaC cellular localization and function are distinct with IcaB protein[10,52]. Hence protein-protein interaction shows *icaB*, a principle determinant for *icaA* and *icaC* co-expression[52].

Enzyme kinetics of bacterial CE4s enzymes cleaves acetyl groups coordinating with metal cofactor through binding of catalytic water molecule[34,35]. Catalytic base atom is for abstraction of proton by the catalytic water molecule which provide a nucleophilic attack of the carbonyl carbon of the acetyl group substrate. A tetrahedral oxyanion is the resulting intermediate stabilized through metal cofactor, whereas catalytic acid molecule then protonates nitrogen atom of the intermediate and producing free amine from the acetate product bound from the metal cofactor[34]. CE4 member is characterized of prototypical NodB conserved domain constituting an Asp-His-His triad in which Asp molecule behaves as catalytic bases and His molecule behaves as the catalytic acids[34,35].

IcaB is CE4 member of beta/alpha barrel fold encompassing NodB homology domain, N-deacetylation activity and metalloenzymatic activity on PNAG substrate through acid/base catalytic mechanism interaction with Asp121. ASP123 was found to be key catalytic base molecule for dPNAG modification[21]. Hydrophobic interaction of IcaB protein is stabilizing tertiary structure and principal force behind stability and folding determination in phospholipid bilayer.

IcaB is stably folded throughout the point of time with the relative lowest Rg value considered with tight packing of IcaB protein in terms of helix-helix interaction. Mostly hydrophobic and hydrophilic amino acids are found to be exposed and buried in IcaB protein and possibility of predicting protein structure and folding characteristics from estimation of solvent accessibility surface area. Stability of IcaB and IcaB-GlcNAc was compactly folded and conformation in the native state was derived from solvent accessible surface area, hence hydrophobic exposed, hydrophobic imposed, hydrophilic exposed, hydrophilic imposed are not disturbed among non-polar residues of hydrophobic interaction.

Conclusion

We predicted cis-acting DNA sequence of PurR (metabolite-responsive), AhrC, SigD, SigE. Invariant N-Box present in *IcaB* gene may act as proximal promoter of *IcaC* gene regulation. Particulary *IcaB* gene sequence overlap with *IcaC* gene may act as promoter for enhancing *IcaC* gene expression. Our findings reveal a novel off-on switch mechanism of PIA/PNAG expression. We highlighted three dimensional structure of IcaB *S. epidermidis*, which gain insights into structure and function of CE4 type of family protein.

Acknowledgement

Authors are thankful to the Pondicherry University for providing bioinformatics facilities to carry out the work.

References

1. Ahmad F., Ali S.A., Hassan M.I. and Islam A., A review of methods available to estimate solvent-accessible surface areas of soluble proteins in the folded and unfolded states, *Curr. Protein Pept. Sci*., **15(5)**, 456–476 **(2014)**

2. Alaa S. Sadah Al-Halfi, Maher Zaki Faisal Al-Shammari and Suaad Khalil Ibrahim, Study of Staphylococcus aureus genotyping using Coagulase gene method, *Res. J. Chem. Environ.*, **28(3)**, 40-47 **(2024)**

3. Allen T.W., Li L. and Vorobyov I., The Different Interactions of Lysine and Arginine Side Chains with Lipid Membranes, *J. Phys. Chem. B.*, **117(40)**, 11906-20 **(2013)**

4. Arciola C.R., Campoccia D., Gamberini S., Rizzi S., Donati M.E., Baldassarri L. and Montanaro L., Search for the insertion element IS256 within the ica locus of Staphylococcus epidermidis clinical isolates collected from biomaterial-associated infections, *Biomaterials*, **25**, 4117–4125 **(2004)**

5. Ardito F., Giuliani M., Perrone D., Troiano G. and Muzio L.L., The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review), *Int J Mol Med.*, **40(2)**, 271-280 **(2017)**

6. Atkin K.E., MacDonald S.J., Potts J.R. and Thomas G.H., A different path: Revealing the function of staphylococcal proteins in biofilm formation, *FEBS Letters*, **588(10)**, 1869-1872 **(2014)**

7. Banks J.L., Beard H.S., Cao Y., Cho A.E., Damm W., Farid R., Felts A.K., Halgren T.A., Mainz D.T. and Maple J.R., Integrated Modelling Program, Applied Chemical Theory (IMPACT), *J. Comput. Chem.*, **26**, 1752–1780 **(2005)**

8. Barji Isfaniza and Awang Adeni Dayang Salwani, Diluted Sulphuric Acid Hydrolysis of Destarched Sago Fibre assisted with Selected Pre-treatments for Glucose and Xylose Production, *Res. J. Chem. Environ*., **28(1)**, 13-20 **(2024)**

9. Blair D.E., Hekmat O., Schüttelkopf A.W., Shrestha B., Tokuyasu K., Withers S.G. and van Aalten D.M., Structure and mechanism of chitin deacetylase from the fungal pathogen Colletotrichum lindemuthianum, *Biochemistry*, **45**, 9416–9426 **(2006)**

10. Blair D.E., Schüttelkopf A.W., MacRae J.I. and van Aalten D.M., Structure and metal-dependent mechanism of peptidoglycan deacetylase, a streptococcal virulence factor, Proc. Natl. Acad. Sci. U.S.A., **102**, 15429–15434 **(2005)**

11. Branda S.S., Vik S., Friedman L. and Kolter R., Biofilms: the matrix revisited, *Trends Microbiol*., **13(1)**, 20–26 **(2005)**

12. Caufrier F., Martinou A., Dupont C. and Bouriotis V., Carbohydrate esterase family 4 enzymes, Substrate specificity, *Carbohydr. Res*., **338**, 687–692 **(2003)**

13. Cerca N., Jefferson K.K., Maira-Litrán T., Pier D.B., Kelly-Quintos C., Gerald B. and Pier G.B., Molecular Basis for Preferential Protective Efficacy of Antibodies Directed to the Poorly Acetylated Form of Staphylococcal Poly-N-Acetyl-β-(1-6)-Glucosamine, *Infect Immun*., **75(7)**, 3406–3413 **(2007)**

14. Cox J. and Mann M., Is proteomics the new genomics?, *Cell*, **130(3)**, 395–398 **(2007)**

15. DeLano W.L., The PyMOL User's Manual, DeLano Scientific, San Carlos, CA, USA **(2002)**

16. Dennis C.C.A., Glykos N.M., Parsons M.R. and Phillips S.E.V., The structure of AhrC, the arginine repressor/activator protein from *Bacillus subtilis*, *Acta Crystallogr D Biol Crystallogr*., **58(Pt 3)**, 421-30 **(2002)**

17. Drazic A., Myklebust L.M., Ree R. and Arnesen T., The world of protein acetylation, *Biochim Biophys Acta*, **1864(10)**, 1372-401 **(2016)**

18. Forman S., Bobrov A.G., Kirillina O., Craig S.K., Abney J., Fetherston J.D. and Perry R.D., Identification of critical amino acid residues in the plague biofilm Hms proteins, *Microbiology*, **152**, 3399–3410 **(2006)**

19. Fukusumi T., Guo T.W., Ren S., Haft S., Liu C., Sakai A., Ando M., Saito Y., Sadat S. and Califano J.A., Reciprocal activation of HEY1 and NOTCH4 under SOX2 control promotes EMT in head and neck squamous cell carcinoma, *Int J Oncol*., **58(2)**, 226–237 **(2021)**

20. Galaktionov S., Nikiforovich G.V. and Marshall G.R., Ab initio modeling of small, medium and large loops in proteins, *Biopolymers*, **60(2)**, 153–168 **(2001)**

21. Galzitskaya O.V., Lobanov M.Y. and Bogatyreva N.S., Radius of gyration as an indicator of protein structure compactness, *Mol. Biol*., **42**, 623–628 **(2008)**

22. Gandhi N.S. and Mancera R.L., The structure of glycosaminoglycans and their interactions with proteins, *Chem. Biol. Drug Des*., **72(6)**, 455–482 **(2008)**

23. Gavel Y. and Heijne G.V., Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: implications for protein engineering, *Protein Eng*., **3(5)**, 433–442 **(1990)**

24. Guo P. and Lam S.L., Unusual structures of TTTA repeats in icaC gene of Staphylococcusaureus, *FEBS Letters*, **12**, 1296-1300 **(2015)**

25. Hada K., Isshiki K., Matsuda S., Yuasa K. and Tsuji A., Engineering of α1-antitrypsin variants with improved specificity for the proprotein convertase furin using site-directed random mutagenesis, *Protein Engineering, Design and Selection*, **26(2)**, 123–131 **(2013)**

26. Hoang T., Zhou C., Lindgren J.K., Galac M.R., Corey B., Endres J.E., Olson M.E. and Fey P.D., Transcriptional Regulation of icaADBC by both IcaR and TcaR in Staphylococcus epidermidis, *Journal of Bacteriology*, **201**, 6 **(2019)**

27. Itoh Y., Rice J.D., Goller C., Pannuri A., Taylor J., Meisner J., Beveridge T.J., Preston J.F. and Romeo T., Roles of pgaABCD genes in synthesis, modification and export of the Escherichia coli biofilm adhesin poly-β-1,6-N-acetyl-d-glucosamine, *J. Bacteriol*., **190**, 3670–3680 **(2008)**

28. Jiang H., Wu X., Xu Z., Zhou T. and Wang H., Antibacterial, Antibiofilm and Antioxidant Activity of Polysaccharides Obtained from Fresh Sarcotesta of Ginkgo biloba: Bioactive Polysaccharide that Can Be Exploited as a Novel Biocontrol Agent, *Evid Based Complement Alternat Med*., **2021**, 5518403 **(2021)**

29. Langton M., Sun S., Ueda C., Markey M., Chen J., Paddy I., Jiang P., Chin N., Milne A. and Pandelia M.E., The HD-Domain Metalloprotein Superfamily: An Apparent Common Protein Scaffold with Diverse Chemistries, *Catalysts*, **10(10)**, 1191 **(2020)**

30. Leal M.C., Surace E.I., Holgado M.P., Ferrari C.C., Tarelli R., Itossi F., Wisniewski T., Castaño E.M. and Morelli L., Notch signaling proteins HES-1 and Hey-1 bind to insulin degrading enzyme (IDE) proximal promoter and repress its transcription and activity: Implications for cellular Aβ metabolism, *Biochim Biophys Acta*, **1823(2)**, 227–235 **(2012)**

31. Little D.J., Bamford N.C., Pokrovskaya V., Robinson H., Nitz M. and Howell P.L., Structural Basis for the De-N-acetylation of Poly-β-1,6-N-acetyl-D-glucosamine in Gram-positive Bacteria, *J Biol Chem*., **289(52)**, 35907–35917 **(2014)**

32. Little D.J., Poloczek J., Whitney J.C., Robinson H., Nitz M. and Howell P.L., The Structure- and Metal-dependent Activity of Escherichia coli PgaB Provides Insight into the Partial De-N-acetylation of Poly-β-1,6-N-acetyl-D-glucosamine, *J Biol Chem*., **287(37)**, 31126-37 **(2012)**

33. Little D.J., Ganguly T., DiFrancesco B.R., Nitz M., Deora R. and Howell P.L., The Protein BpsB Is a Poly-β-1,6-N-acetyl-D-glucosamine Deacetylase Required for Biofilm Formation in Bordetella bronchiseptica, *J Biol Chem*., **290(37)**, 22827–22840 **(2015)**

34. Liu L. et al, Structural and biochemical insights into the catalytic mechanisms of two insect chitin deacetylases of the carbohydrate esterase 4 family, *J Biol Chem*., **294(15)**, 5774–5783 **(2019)**

35. Lomize A.L., Pogozheva I.D., Lomize M.A. and Mosberg H.I., The role of hydrophobic interactions in positioning of peripheral proteins in membranes, *BMC Structural Biology*, **7**, 44 **(2007)**

36. Mack D., Fischer W., Krokotsch A., Leopold K., Hartmann R., Egge H. and Laufs R., The intercellular adhesion involved in biofilm accumulation of Staphylococcus epidermidis is a linear beta -1,6-linked glucosaminoglycan: purification and structural analysis, *J. Bacteriol*., **178(1)**, 175–183 **(1996)**

37. Malde A.K., Zuo L., Breeze M., Stroet M., Poger D., Pramod C. and Nair P.C., An Automated Force Field Topology Builder (ATB) and Repository: Version 1.0, *J. Chem. Theory Comput*., **7**, 4026–4037 **(2011)**

38. Mandell, Sigma factors (σ) are another major mechanism of response to environmental stimuli, Mandell, Douglas and Bennett's Principles and Practice of Infectious Diseases, Eighth Edition **(2015)**

39. Mlynek K.D., Bulock L.L., Stone C.J., Curran L.J., Sadykov M.R., Bayles K.W. and Brinsmade S.R., Genetic and Biochemical Analysis of CodY-Mediated Cell Aggregation in *Staphylococcus aureus* Reveals an Interaction between Extracellular DNA and Polysaccharide in the Extracellular Matrix, *J Bacteriol*., **202(8)**, e00593-19 **(2020)**

40. Morgan J.L.W., Strumillo J. and Zimmer J., Crystallographic snapshot of cellulose synthesis and membrane translocation, *Nature*, **493(7431)**, 181–186 **(2013)**

41. Nagata M., Kaito C. and Sekimizu K., Phosphodiesterase activity of cvfa is required for virulence in *staphylococcus aureus*, *J Biol Chem.*, **283(4)**, 2176-84 **(2008)**

42. Oostenbrink C., Villa A., Mark A.E. and Van Gunsteren W.F.A., Biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets 53A5 and 53A6, *J. Comput. Chem.*, **25**, 1656–1676 **(2004)**

43. Ozcan S., Andrali S.S. and Cantrell J.E.L., Modulation of transcription factor function by O-GlcNAc modification, *Biochim Biophys Acta*, **1799(5-6)**, 353–364 **(2010)**

44. Peerayeh S.N., Ghasemian A., Bakhshi B. and Mirzaee M., Comparison of Biofilm Formation between Methicillin-Resistant and Methicillin-Susceptible Isolates of *Staphylococcus aureus*, *Iran Biomed J*., **20(3)**, 175-81 **(2016)**

45. Pokrovskaya V., Poloczek J., Little D.J., Griffiths H., Howell P.L. and Nitz M., Functional Characterization of Staphylococcus epidermidis IcaB, a De N acetylase Important for Biofilm Formation, *Biochemistry*, **52(32)**, 5463–5471 **(2013)**

46. Pronk S., Pall S., Schulz R., Larsson P., Bjelkmar P., Apostolov R., Shirts M.R. and Smith J.C., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit, *Bioinformatics*, **29**, 845–854 **(2013)**

47. Sause W.E., Balasubramanian D., Irnov I. and Torres V.J., The purine biosynthesis regulator PurR moonlights as a virulence regulator in *Staphylococcus aureus*, *PNAS*, **116(27)**, 13563-13572 **(2019)**

48. Schwede T., Studer G. and Biasini M., Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane), *Bioinformatics*, **30(17)**, i505–i511 **(2014)**

49. Sirnes S., Bruun J., Yohannes Z., Rivedal E. and Leithe E., The gap junction channel protein connexin 43 is covalently modified and regulated by SUMOylation, *J Biol Chem.*, **287(19)**, 15851-61 **(2012)**

50. Suryadinata R., Parker M.W., Papaleo E. and Sarcevic B., Molecular and structural insight into lysine selection on substrate and ubiquitin lysine 48 by the ubiquitin-conjugating enzyme Cdc34, *Cell Cycle*, **12(11)**, 1732–1744 **(2013)**

51. Thai Ke Quan, Huynh Phuoc and Nguyen Thi Thuong Huyen, Exploring the mutation landscape of SARS-CoV-2 Variant JN.1, *Res. J. Biotech.*, **19(6)**, 76-83 **(2024)**

52. Taylor E.J., Gloster T.M., Turkenburg J.P., Vincent F., Brzozowski A.M., Dupont C., Shareck F., Centeno M.S., Prates J.A., Puchart V., Ferreira L.M., Fontes C.M., Biely P. and Davies G.J., Structure and activity of two metal ion-dependent acetylxylan esterases involved in plant cell wall degradation reveals a close similarity to peptidoglycan deacetylases, *J. Biol. Chem*., **281**, 10968–10975 **(2006)**

53. Toyoda K. and Inui M., Extracytoplasmic function sigma factor σD confers resistance to environmental stress by enhancing mycolate synthesis and modifying peptidoglycan structures in Corynebacterium glutamicum, *Molecular Microbiology*, **107(3)**, 312–329 **(2018)**

54. Tusnady G.E., Molnar J. and Szakacs G., Characterization of Disease-Associated Mutations in Human Transmembrane Proteins, *PLoS One*, **11(3)**, e0151760 **(2016)**

55. Varki A. and Sharon N., Historical background and overview. In: Essentials of Glycobiology, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, USA **(2009)**

56. Vu B., Chen M., Crawford R.J. and Ivanova E.P., Bacterial extracellular polysaccharides involved in biofilm formation, *Molecules*, **14(7)**, 2535–2554 **(2009)**

57. Vuong C., Kocianova S., Voyich J.M., Yao Y., Fischer E.R., DeLeo F.R. and Otto M., A crucial role for exopolysaccharide modification in bacterial biofilm formation, immune evasion and virulence, *J Biol Chem*., **279(52)**, 54881-6 **(2004)**

58. Wang H., Dommert F. and Holm C., Optimizing working parameters of the smooth particle mesh Ewald algorithm in terms of accuracy and efficiency, *J. Chem. Phys*., **133**, 034117 **(2010)**

59. Wolynes P.G., Kim B.L. and Schafer N.P., Predictive energy landscapes for folding α-helical transmembrane proteins, Proc. Natl. Acad. Sci. U.S.A., **111(30)**, 11031–11036 **(2014)**

60. Zhuo S., Clemens J.C., Stone R.L. and Dixon J.E., Mutational analysis of a Ser/Thr phosphatase. Identification of residues important in phosphoesterase substrate binding and catalysis, *J Biol Chem*., **269(42)**, 26234-8 **(1994)**

61. Ziebuhr W., Krimmer V., Rachid S., Lössner I., Götz F. and Hacker J., A novel mechanism of phase variation of virulence in Staphylococcus epidermidis: evidence for control of the polysaccharide intercellular adhesin synthesis by alternating insertion and excision of the insertion sequence element IS256, *Mol. Microbiol*., **32**, 345–356 **(1999)**.

(Received 17th November 2024, accepted 21st January 2025)