3.1 MATERIALS

3.1.1 Antimicrobial peptides

The antimicrobial peptides used for the studies of our various objectives is as follows-

1. Four antimicrobial peptides of length less than 20 amino acids from plant origin chosen for interaction with the *Staphylococcus aureus* toxins were JCpep7, Sesquin, Snakin-2 and Ib-AMP1 are shown in TABLE 3.1.

TABLE 3.1: Names of selected antimicrobial peptides (for interaction with *S. aureus* toxins), their source of origin and their sequences

Name of Antimicrobial Peptides (AMP)	**Source**	**Sequence**	**References**
JCpep7	*Jatropha curcas*	KVFLGLK	Xiao et. al. 2011
Sesquin	*Vigna unguiculata subsp. sesquipedalis*	KTCENLADTY	Wong and Ng 2005
Snakin-2	*Solanum tuberosum*	SYKKIDCGGACAARC	Berrocal-Lobo et al. 2002
Ib-AMP1	*Impatiens balsamina*	QWGRRCCGWGPGRRYCVRWC	Tailor et. al. 1997

2. The four antimicrobial peptides Ib-AMP1, Ib-AMP2, Ib-AMP3 and Ib-AMP4, of length 20 amino acids from *Impatiens balsamina* were used for interaction study with ToxT are shown in TABLE 3.2. The sequences of these peptides were taken from CAMP database.

TABLE 3.2: Sequences of the antimicrobial peptides of *Impatiens balsamina* [Tailor et. al. 1997] used in the interaction studies with ToxT.

Name of Antimicrobial peptide	**Sequence**
Ib-AMP1	QWGRRCCGWGPGRRYCVRWC
Ib-AMP2	QYGRRCCNWGPGRRYCKRWC
Ib-AMP3	QYRHRCCAWGPGRKYCKRWC
Ib-AMP4	QWGRRCCGWGPGRRYCRRWC

3. The antimicrobial peptides of length less than 10 amino acids from microorganism source were selected from CAMP database used for the interaction studies with catalytic domain of diphtheria toxin are shown in TABLE 3.3.

TABLE 3.3: Names, source and sequence of antimicrobial peptides selected from microorganism source

Name of Antimicrobial Peptides (AMP)	**Source**	**Sequence**	**References**
Anionic Peptide SAAP	*Pasteurella haemolytics*	DDDDDD	Brodgen et. al. 1996

Microcin C7	*Escherichia coli*	MRTGNAN	García-Bustos et. al. 1984; García-Bustos et. al. 1985
Bacteriocin	*Rhizobium leguminosarum bv. viciae*	ASILTNAS	Naeem et. al. 2005
Curvalicin-28c	*Lactobacillus curvatus*	NIPQLTPTP	Galfi et. al. 2009
NRWC	*Bacillus subtilis*	NRWCFAGDD	Teixeira et. al. 2013

4. The antimicrobial peptides of length less than 10 amino acids from plant source were selected from CAMP database used for the interaction studies with catalytic domain of diphtheria toxin are shown in TABLE 3.4.

TABLE 3.4: Names, source and sequence of antimicrobial peptides selected from plant source

Name of Antimicrobial Peptide (AMP)	**Source**	**Sequence**	**References**
JC pep7	*Jatropha curcas*	KVFLGLK	Xiao et. al. 2011
Antimicrobial	*Cocos nucifera*	VAGRAQGM	Mandal et. al.

peptide 1			2009
Cr- ACP 1	*Cycas revoluta*	AWKLFDDGV	Mandal et. al. 2012
Sesquin	*Vigna unguiculata subs. sesquipedalis*	KTCENLADTY	Wong and Ng 2005
Alliumin	*Allium sativum*	DDFLCAGGCL	Xia and Ng 2005

3.1.2 Toxins

ETA, ETB and PVL (pdb id: 1DUA, 1DT2 and 1PVL respectively) [Papageorgiou et al 2000; Pédelacq et al 1999] toxins of Staphylococcus aureus; Transcription activator ToxT (pdb id: 3GBG) [Lowden et al 2010] of Vibrio cholerae; Catalytic domain of diphtheria toxin (pdb id: 1DTP) [Weiss et al 1995] of Corynebacterium diphtheriae and New Delhi Metallo-β-lactamase (pdb id: 3Q6X) [Zhang and Quan 2011] superbug protein were used in the present studies.

3.1.3 Cultures used

Cultures of Gram negative bacteria *Escherichia coli* (MTCC No. 443) and a Gram positive bacteria *Staphylococcus aureus* (MTCC No. 7443) were used for antibacterial assay in the present study [https://www.mtccindia.res.in/].

3.1.4 Databases used

3.1.4.1 CAMP Database

This is a database for collection of antimicrobial peptides. Details of antimicrobial peptides from various sources are collected here. CAMP also has a collection of tools for prediction of

various properties of antimicrobial peptides [Thmas et al 2009; Waghu et al 2014; Waghu et al 2015].

3.1.4.2 RCSB

This database consists of 3-D structures of proteins and nucleic acids of all the organisms. The structural files can be searched using PDB identification number or description of macromolecules and the relevant structures can be downloaded from the database [Berman et al 2000].

3.1.5 Web servers used

3.1.5.1 Pepfold3

This server was used for the prediction of 3-dimensional structures of the antimicrobial peptides. The sequences of antimicrobial peptides were given as input and the predicted structures of these antimicrobial peptides were downloaded from the server. This server is used for the prediction of native conformations of linear peptides from 5 – 50 amino acids length [Thevenet et al 2012; Shen et al 2014; Lamiable et al 2016].

3.1.5.2 FireDock (Fast Interaction Refinement in molecular docking) server

The first web server for protein-protein docking and refinement was Firedock. Two steps are followed in this web server- first step is global search for best conformation of protein-protein complex using PatchDock [Duhovny et al 2002; Schneidman-Duhovny et al 2005; Schneidman-Duhovny et al 2003; Mashiach et al 2010] and second step is refinement of the conformations using FireDock [Andrusier et al 2007; Mashiach et al 2008].

3.1.5.3 CABS-dock server

This server was used for docking studies of ToxT with antimicrobial peptides. This server is specific for protein-peptide docking interaction study. The 3D structure of protein and the peptide sequence are given as input to this server. Flexible docking of proteins and peptides is carried out by the CABS-dock web server. Here the docking interactions between proteins and peptides is modelled and hence can be analysed. The protocol constitutes of the following steps –

1. Generation of random structures: Several peptide structures are randomly generate. These structures are then placed at random positions on the sphere's surface whose centre is at the geometrical centre of the receptor.

2. Simulation of docking and binding. This web server uses Replica Exchange Monte Carlo dynamics, having uniform spreading of replicas on the temperature scale. The number of replicas is 10. As the time of the simulation continues the temperature of the replicas decreases till the end which is at the bottom of the energy minima.output produces 10 trajectories.

3. A two-step process is followed for Selecting the final representative models:

- The first filtering process filters out all the unbound states from all the 10 trajectories. 100 models with lowest binding energy are considered for the next step.
- Next filtering step is of K-medoids clustering. The models selected in the above filtering process are clustered in this process. Various initial medoids are used for perform clustering 100 times. The final models selected are ten consensus medoids.

4. Final models are reconstructed: Since the models are in C- alpha form, these have to be finally reconstructed to all-atom model. [Kuecinski et al 2015; Blaszczyk et al 2016; Ciemny et al 2017]

3.1.5.4 I-TASSER server

The I-TASSER server is used for the prediction of 3D structure of proteins. The output of the I-TASSER server gives up to five models. These models are further quantitatively analysed based upon the standard scoring systems for the quantitative assessment of the I-TASSER models. Each model is predicted with a C-score, TM-score and RMSD, based upon these the best model is predicted for further studies.

C-score

The C-score of the I-TASSER models is defined as

$$C - score = \ln\{\frac{M}{Mtot} . \frac{1}{RMSD} . (\frac{\prod_{i=1}^{4} Z(i))}{(\prod_{i=1}^{4} Z0(i))}\}$$ Equation-1

Where, M is the multiplicity of structures in the SPICKER cluster; M_{tot} is the total number of the I-TASSER structure decoys used in the clustering; RMSD is the average RMSD of the decoys to the cluster centroid; Z(i) is the highest Zscore (the energy to mean in the unit of standard deviation) of the templates by the ith PPA threading program and Z0(i) is a program-specified Z-score cutoff for distinguishing between good and bad templates.

TM-score

TM-score is defined to assess the topological similarity of two protein structures

$$TM - score = \frac{1}{L}(\sum_{i=1}^{L} \frac{1}{1+\frac{di2}{d02}})$$ Equation-2

where di is the distance of the ith pair of residues between two structures after an optimal superposition, $d0 = 1.24 \sqrt[3]{L - 15}$ -1.8, and L is the protein length. TM-score stays in [0, 1] with higher values indicating better models. Statistically, a TM-score ≤ 0.17 corresponds to a

similarity between two randomly selected structures from the PDB library; a TM-score > 0.5 corresponds approximately to two structures of the similar topology. One advantage of the TM-score is that the meaning of the TM score cutoffs is independent of the size of proteins. [Zhang 2008; Roy et al 2010; Yang et al 2015].

3.1.5.5 ProtParam

ProtParam is used for the *in silico* calculation of various chemical and physical parameters of the protein entered by user or protein stored in TrEMBL or Swiss-Prot. The parameters that can be predicted here are grand average of hydropathy (GRAVY), aliphatic index, instability index, estimated half-life, extinction coefficient, atomic composition, amino acid composition, theoretical pI and molecular weight [Gasteiger et al 2005].

3.1.5.6 ToxinPred

ToxinPred web server is used for the design and prediction of Toxic and non-toxic peptides. The sequence of the peptides is given as input to this web server [Gupta et al 2013].

3.1.5.7 HemoPred webserver

The haemolytic activity of peptides is predicted and analysed by the HemoPred web server. The sequence of the peptide is given as the input for prediction [Win et al 2017].

3.1.6 Software used

3.1.6.1 Gromacs (GROningen MAchine for Chemical Simulations)

The simulation of nucleic acids, lipids and proteins can be pergormed using GROMACS, which is a molecular dynamics process. GROMACS uses command-line interface. The pdb files of the structures are converted to the GROMACS format which GROMSCS uses. GROMACS supports various forcefields because of which it is flexible in performance.

Various commands are used by GROMACS for the analysis of various outputs generated after various simulation runs [Berendsen et al 1995; Lindahl et al 2001].

3.1.7 Visualizers used

3.1.7.1 Discovery Studio Visualizer

Discovery studio visualizer is a free viewer which offers an interactive environment editing and viewing structural data of the biomolecules. The Discovery Studio visualizer was used for visualization, analysis and sharing of macromolecular data. The sequence, structure and other chemical properties of the macromolecules can be viewed and analysed using Discovery Studio Visualizer. It is used for modelling and analysis of sequences, molecular structures and other data for use of researchers in the field of life science. This application can run both on Linux and Windows environment [Visualizer Discovery Studio 2013].

3.1.7.2 PyMOL visualizer

PyMOL is used for production of 3D images of biological macromolecules and small molecules. It is an open source software for generation of high quality images to be used for structural biology [The PyMOL Molecular Graphics System, Version 2.0 Schrodinger LLC].

3.1.8 USV (P) Ltd., Mumbai

They synthesize custom peptides and supply for industrial and research purposes [http://www.usvindia.com/].

Methodology used by USV custom peptide for peptide synthesis:

- Solid phase using Fmoc and t-Boc methodologies.
- Solution phase chemistry.

3.1.9 Synthesized Peptide

Antimicrobial peptide Alliumin having sequence DDFLCAGGCL was synthesized by USV (P) Ltd.

3.2 METHODOLOGY

The methodologies used in each sub-section of work are described below -

3.2.1 Selection of potential Antimicrobial peptide against *Staphylococcus aureus* toxins

- Four antimicrobial peptide of length less than or equal 20 amino acids were chosen from CAMP database. The four chosen antimicrobial peptides were JCpep7, Sesquin, Snakin-2 and Ib-AMP1.
- The 3D structures of all the four chosen antimicrobial peptides were predicted using Pepfold3 server.
- The 3D structures of exfoliative toxin A (ETA), exfoliative toxin B (ETB) and Panton-Valentine leukocidin (PVL) were downloaded from RCSB database. The PDB structures of toxins were checked for the presence of water molecules and inhibitors. The water molecules and the inhibitors attached with the PDB structures were removed and hydrogens were added to these toxin structures. Then these structures were energy minimized and used for docking.
- All the four selected antimicrobial peptides were docked with the three toxins of *S. aureus* using FireDock protein- protein docking web server. The dockings were performed first with Patchdock server. The toxins were given as the receptor input and the antimicrobial peptides were given as the ligand input. Clustering RMSD is kept at 4 Å for the docking. The solutions of these dockings were again given as input to the Firedock server for refinement.
- The best docked conformations of all the sets of dockings were selected based on global energies and non-bond interactions between the toxins and peptides.

- The antimicrobial peptide which gave most significant result with all the three toxins was found to be most effective and selected for further studies.

3.2.2 Studies of interaction of antimicrobial peptides with *Vibrio cholerae* transcription activator ToxT

- The 3D structure of ToxT (pdb id 3GBG) was downloaded from RCSB database.
- The structure was checked for missing residues and the structure was modelled using I-TASSER server. The sequence of ToxT in FASTA format was given as input to the server and 3GBG was given as Template.
- The modelled structure of ToxT and the four antimicrobial peptides were given to CABS-dock server for docking.
- The results of all the four sets of dockings were analysed and the best docked complexes of all the sets based upon the protein peptide non-bond interactions were predicted.
- The best docked complexes were analysed and the suitable complexes were taken for molecular dynamics simulations.
- The MD simulations of the docked complexes and ToxT protein were carried out for 10nsusing GROMACS 4.5.3. The respective charges were added to neutralize the protein systems if they had some net positive or negative charges on them. The root mean square deviations (RMSD), root mean square fluctuations (RMSF) and radius of gyration (Rg) were calculated for all the simulations [Lobanov et al 2008; Elengoe et al 2014; Singh et al 2018; Thangapandian et al 2012]. The results for the complexes were analysed. Thus predicting the most suitable AMP of *Impatiens balsamina* against ToxT.

- The selected antimicrobial peptide was given for in silico characterization using ProtParam and ToxinPred.

3.2.3 Selection of potential antimicrobial peptide against Diphtheria toxin

- Five antimicrobial peptides of length =<10 amino acids from both plant and microorganism sources were selected from CAMP database.
- The 3D structures of Anionic peptide SAAP, Bacteriocin, Curvalicin-28c, NRWC, JC-pep7, Antimicrobial peptide1, Cr-ACP1, Sesquin and Allliumin were predicted using Pepfold3 server. The Microcin C7 structure was downloaded from RCSB database. The G-chain of the structure with pdb id :3H9J was the structure of Microcin C7, this was isolated and used for our studies.
- The catalytic domain of the diphtheria toxin was taken for the present studies. The pdb structure of isolated catalytic domain of diphtheria toxin (PDB id: 1DTP) was downloaded from RCSB database (Berman et al., 2000). This pdb structure of the catalytic domain of diphtheria toxin is bound to inhibitor adenylyl (3′–5′) uridine 3′-monophosphate (ApUp) (Weiss et al., 1995). This inhibitor bound with this structure was removed and the hydrogens were added to the structure.
- All the selected antimicrobial peptides were docked with the 3D structure of catalytic domain of diphtheria toxin using Firedock web server.
- The docked complexes were analysed for the interactions between catalytic domain of diphtheria toxin and antimicrobial peptides.
- The best docked complex each from five sets of dockings for plant source antimicrobial peptides and five sets of dockings for microorganism source antimicrobial peptides were selected and taken for molecular dynamics studies.
- The molecular dynamics simulation of both the selected docked complexes was carried out for 10ns using GROMACS 4.5.3. The respective charges were added to

neutralize the protein systems if they had some net positive or negative charges on them. The root mean square deviations (RMSD), root mean square fluctuations (RMSF) and radius of gyration (Rg) were calculated for all the simulations [Lobanov et al 2008; Elengoe et al 2014; Singh et al 2018; Thangapandian et al 2012]. The results for the complexes were analysed and compared. Afterwards this simulation was extended till 20ns for the best complexes selected among the two. Intermolecular hydrogen bond interaction was analysed. Thus, the potential antimicrobial peptide was selected and used for further studies.

3.2.4 Interaction of Alliumin with *Vibrio cholerae* transcription activator ToxT

- The modelled structure of ToxT modelled in the section 3.2.2 was used in this study. The sequence of antimicrobial peptide alliumin and modelled structure of ToxT were given to CABS-dock server for docking.
- The docking results were analysed and the best docked complex was selected based upon the protein peptide interaction non-bond interactions.
- This best docked complex was taken for molecular dynamics simulation.
- The MD simulations of the docked complex and ToxT protein were carried out for 20ns using GROMACS 5.0.7. The respective charges were added to neutralize the protein systems if they had some net positive or negative charges on them. The root mean square deviations (RMSD), root mean square fluctuations (RMSF), intermolecular hydrogen bonds and radius of gyration (Rg) were calculated for all the simulations [Lobanov et al 2008; Elengoe et al 2014; Singh et al 2018; Thangapandian et al 2012]. The results were analysed to predict if the antimicrobial peptide alliumin is a potential antimicrobial peptide against ToxT.

3.2.5 Interaction of Alliumin with New Delhi Metallo-beta-lactamase 1

- The 3D structure of NDM-1 (pdb id 3Q6X) was downloaded from RCSB database.
- The structure was checked for missing residues and the structure was modelled using I-TASSER server. The sequence of NDM-1 in FASTA format was given as input to the server and 3Q6X was given as Template.
- The modelled structure of NDM-1 and the sequence of antimicrobial peptide Alliumin were given to CABS-dock server for docking.
- The docking results were analysed and the best docked complex was selected based upon the protein peptide non-bond interactions.
- This best docked complex was taken for molecular dynamics simulation.
- The MD simulations of the docked complex and NDM-1 protein were carried out for 20ns using GROMACS 5.0.7. The respective charges were added to neutralize the protein systems if they had some net positive or negative charges on them. The root mean square deviations (RMSD), root mean square fluctuations (RMSF), intermolecular hydrogen bonds and radius of gyration (Rg) were calculated for all the simulations [Lobanov et al 2008; Elengoe et al 2014; Singh et al 2018; Thangapandian et al 2012]. The results were analysed to predict if the antimicrobial peptide alliumin is the potential antimicrobial peptide against NDM-1.

3.2.6 Synthesis of Alliumin and studies to check its antimicrobial properties

- The in silico characterization of Alliumin was done using ProtParam Tool.
- The Toxicity and haemolytic characterization of Alliumin was done using ToxinPred and HemoPred web servers respectively.
- The sequence of Alliumin (DDFLCAGGCL) was sent to USV (P) Ltd, Mumbai for custom peptide synthesis. 10mg peptide having purity >95% was procured.

- To determine the antibacterial activity of synthesized peptide Alliumin, standard disk diffusion method was carried out against *Staphylococcus aureus* and *Escherichia coli*. The agar plates having suitable nutrient media (LB media) was prepared, sterilized and allowed to solidify [Tassou *et al*. 2000]. After solidification, the agar plates were inoculated with bacterial cultures. 6mm disks were impregnated with a required quantity of peptide dissolved in deionised water. (A stock solution of 1mg Alliumin in 1ml of deionised water was prepared). These plates were incubated at 28°C for two days. The zone of inhibition was calculated and recorded for each microorganism and expressed in millimeter (mm). All the above experiments were carried out in triplicate.