Ahmad, S., Ranaghan, K., & Azam, S. S. (2019). Combating tigecycline resistant *Acinetobacter baumannii*: A leap forward towards multiple-epitope based vaccine discovery. *European Journal of Pharmaceutical Sciences*, 132, 1-17.
<https://doi.org/10.1016/j.ejps.2019.02.023>

Peer reviewed version

License (if available):
CC BY-NC-ND

Link to published version (if available):
[10.1016/j.ejps.2019.02.023](https://doi.org/10.1016/j.ejps.2019.02.023)

[Link to publication record in Explore Bristol Research](#)
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Elsevier at <https://www.sciencedirect.com/science/article/pii/S0928098719300843>. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
<http://www.bristol.ac.uk/red/research-policy/pure/user-guides/ebr-terms/>

Accepted Manuscript

Combating tigecycline resistant *Acinetobacter baumannii*: A leap forward towards multi-epitope based vaccine discovery

Sajjad Ahmad, Kara E. Ranaghan, Syed Sikander Azam

PII: S0928-0987(19)30084-3

DOI: <https://doi.org/10.1016/j.ejps.2019.02.023>

Reference: PHASCI 4853

To appear in: *European Journal of Pharmaceutical Sciences*

Received date: 3 December 2018

Revised date: 31 January 2019

Accepted date: 16 February 2019

Please cite this article as: S. Ahmad, K.E. Ranaghan and S.S. Azam, Combating tigecycline resistant *Acinetobacter baumannii*: A leap forward towards multi-epitope based vaccine discovery, *European Journal of Pharmaceutical Sciences*, <https://doi.org/10.1016/j.ejps.2019.02.023>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Title Page

Manuscript title: Combating Tigecycline Resistant *Acinetobacter baumannii*: A Leap Forward towards Multi-epitope based Vaccine Discovery

Authors: Sajjad Ahmad¹, Kara E. Ranaghan², Syed Sikander Azam^{1,✉}

Institutional address:

¹Computational Biology Lab, National Center for Bioinformatics, Quaid-i-Azam University, Islamabad, Pakistan.

²Centre for Computational Chemistry, University of Bristol, Bristol, United Kingdom.

Email addresses:

Sajjad Ahmad, Email: sajjademaan8@gmail.com

Kara E. Ranaghan, Email: Kara.Ranaghan@bristol.ac.uk

Syed Sikander Azam, Email: syedazam2008@gmail.com, ssazam@qau.edu.pk

Corresponding Author:

✉ Syed Sikander Azam

E-mail: syedazam2008@gmail.com, ssazam@qau.edu.pk

Phone: 0092-51-906 44130

Computational Biology Lab, National Center for Bioinformatics, Quaid-i-Azam University, Islamabad 45320, Pakistan

Abstract

Global emergence of Tigecycline resistant *Acinetobacter baumannii* (TRAB) is on the horizon and poses a very serious threat to human health. There is a pressing demand for suitable therapeutics against this pathogen, particularly a vaccine to protect against TRAB infections. We present a comprehensive investigation of the complete proteome of a TRAB AB031 strain to predict promiscuous antigenic, non-allergenic, virulent B-cell derived T-cell epitopes and formulate a multi-epitope vaccine against the pathogen. We identified epitopes from three proteins: outer membrane protein assembly factor (BamA), fimbrial biogenesis outer membrane usher protein (FimD) and type IV secretion protein (Rhs) that are appropriate for vaccine design. These proteins constitute the core proteome of the pathogen, are essential, localized at the pathogen surface, non-homologous to humans, mice and to the beneficial probiotic bacteria that reside the human gut. Moreover, these proteins are ideal candidates for experimental investigation as they have favorable physicochemical properties and have strong cellular interacting networks. The predicted epitopes: FPLNDKPGD (BamA), FVHAE EAAA (FimD) and YVVAGTAAA (Rhs) have exo-membrane topology for efficient recognition of the host immune system and high affinity for the most prevalent allele in human populations, the DRB*0101. These epitopes were linked and attached to an adjuvant to enhance its antigenicity. The multi-epitope vaccine-construct was docked with the TLR4 receptor to assess its affinity for the protein and thus its presentation to the host immune system. Docking results were validated through molecular dynamics simulations and binding free energies were calculated using the molecular mechanics/generalized Born (MM-GBSA) method. In conclusion, we expect the designed chimeric vaccine is highly likely to be effective against infections caused by TRAB.

Keywords: Tigecycline resistant *Acinetobacter baumannii*; Vaccine; B-cell derived T-cell epitopes; Molecular docking; Molecular dynamics simulation.

1. Introduction

Acinetobacter baumannii is a Gram-negative, aerobic and non-motile bacillus that has emerged as the most troublesome pathogen for many health care institutions globally [1], [2], [3]. *A. baumannii* is commonly found in blood stream, lung, urinary tract and wound infections and it especially effects immune-compromised populations [4]. The antibiotic resistance in five bacterial pathogens: *A. baumannii*, *Escherichia coli*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Staphylococcus aureus* cost \$2.9 billion to the United States health care industry. The cost for consuming one standard unit (SU) of antibiotics ranged from \$0.6 for broad spectrum penicillins, cephalosporins, and quinolones to \$0.1 for 0.1 SU of carbapenems [5]. Multi-drug resistant *A. baumannii* (MDAB) is becoming a serious health concern and compromising healthcare outcomes due to its remarkable ability of acquiring resistant genes and surviving exposure to surface microcides [2], [6]. It has been reported that MDAB infected patients are at an increased risk of morbidity and mortality which contribute significantly to high health costs. According to the reports of Center for Disease Control and Prevention (CDC), MDAB is a serious threat that causes approximately 7,000 infections and 500 deaths each year in the United States [7]. The outbreaks of MDAB are generally pervasive and cases have been documented from countries including Brazil, China, the UK, the USA, India, Spain, Germany, Iran, Turkey and Iraq etc. [8]. The colonization rate of *A. baumannii* in hospitalized patients is 75% while in healthy individuals it is 42.5% [9]. The recommended therapy for MDAB includes carbapenems such as imipenem and meropenem; however, a significant rise of carbapenem-resistant *A. baumannii* (CRAB) has been reported worldwide [10]. The crude mortality rate observed for the hospital acquired CRAB infections in intensive care units ranged between 26% to 76% [8]. The last treatment resort for MDAB is colistin; however, severe side effects of nephrotoxicity and neurotoxicity have limited its use [11]. The first identified and US Food and Drug Administration (FDA) approved glycylcycline antibiotic, Tigecycline, mimics the tetracycline structure and is effective against carbapenem-resistant bacteria [12]. Tigecycline is recommended for the treatment of serious complications including: community acquired pneumonia, skin, and intra-abdominal infections [12], [13]. Tigecycline is considered a last resort because of its bacteriostatic activity against MDR *Acinetobacter* [14]. It disrupts protein synthesis by blocking the entry of aminoacyl-tRNA into the Aminoacyl (A) site of the 30S ribosome in prokaryotic translation [15]. Unfortunately, an increasing number of cases have been reported pointing to the development of Tigecycline-resistant *A. baumannii* (TRAB) since its first use in clinics [16], [17], [18]. TRAB is considered as a superbug because of its resistance to the majority of the antibiotics that are in use today [1], [2], [3].

Due to several limitations of the current therapy regime, such as the lack of a vaccine and incomplete efficacy of the drugs, there is a desperate need to develop novel approaches that speed up the identification of novel vaccines and drug targets to curtail serious infections of TRAB [19]. Engineering a safe vaccine is of particular importance since preventive measures are always better than curative approaches [20]. Similarly, vaccine resistance rarely emerged compared to antibiotics. It is well-established that vaccines are far more robust against evolution if compared to drugs. This can be explained in two ways. First, vaccines work prophylactically in contrast to drugs that function therapeutically. Secondly, vaccines induce immunity against

multiple targets while drugs target only one in the majority of the cases. Therefore, pathogen populations create lesser mutation for vaccine resistance in contrast to drugs [21]. Conventional vaccines are developed from live attenuated or inactivated organisms, which elicit strong humoral and cellular immunity resulting in long lasting immunity [22], [23]. However, the use of such vaccines has been questioned due to crucial safety concerns as the used pathogen may become reactivated [24], [25]. Similarly, subunit-based vaccines containing a single whole protein are also not a good choice because of non-specific immune responses, which may occur due to the presence of multiple antigenic determinants [26]. In addition, the formulation of such vaccines often leads to reactogenic and/or allergenic responses that are often not desired [26]. The design of epitope-driven or peptide-based vaccines is more attractive; they are comparatively easier to produce and construct, they lack any infectious potential and offer chemical stability [26], [27]. However, peptide-based vaccines still present many obstacles including the low intrinsic antigenicity of using individual peptides, and the shortage of suitable delivery routes. For proper engagement of cellular and humoral immunity, the combining task of different peptides and poor population coverage of T-cell epitopes is also challenging [26]. In addition, the unreliable prediction of B-cell epitopes [26]. Circumventing these limitations, the construction of multi-epitope vaccines by linking the safest and B-cell derived T-cell antigenic epitope offers many significant benefits [28], [29]. These include broader intrinsic immunogenicity, both humoral and cellular immunity can properly be engaged and better T-cell epitopes population coverage [30]. The novel combination of immuno-informative approaches, together with the knowledge of host immune responses and the exponential increase in complete genome sequencing of pathogen strains, now makes it possible to select the most antigenic epitopes [30]. As such, it is conceived that cocktails of selected and defined epitopes may provide foundations for the rational design of vaccines capable of eliciting convenient cellular and humoral immunity [31]. The majority of vaccines are based on humoral immunity; however, cellular immunity-based vaccines are preferred as the host generates strong CD8+ T-cell immune responses against the pathogen infected cells [32], [23]. T-cell immunity is vital because of the antigenic drift phenomena may aid in foreign particle escape from B-cell immunity [33] and T-cell immunity is long lasting. Thus, T-cell immunity could provide better foundations for the design of multi-epitope effective vaccines that can trigger both arms of immunity [34]. Intracellular antigen processing pathways with linear epitopes as predominant targets lead to cytotoxic T-cell responses. Therefore, it is essential to ensure that the selected epitopes of T-cell have binding to more than one major histocompatible allele and cover a major human population [23], [35], [36], [37].

In this study, we have used bioinformatics techniques to predict promiscuous antigenic epitopes, which bind to multiple HLA alleles within a heterogeneous population, therefore, applicable to a wide range of individuals [38]. Similarly, the epitopes are B-cell derived T-cell that can elicit both humoral and cellular immunity against TRAB and were formulated into a multi-epitope vaccine [35], [36], [37]. Furthermore, docking and dynamics simulations were performed to unveil vaccine-construct binding and stability with the TLR4. TLR4 is a member of toll-like receptor family that spot and mediate cytokine production against molecular patterns expressed on infectious agent. The findings of this study could be useful in the design of effective vaccines for the TRAB superbug.

2. Materials and Methods

The complete workflow of the current study is illustrated in **Fig. 1**

2.1. Proteome Retrieval and Subtractive Proteomics

The complete proteome of TRAB strain AB031 [39] was retrieved from the NCBI [40] and used in subtractive proteomics hierarchy to extract proteins appropriate for peptide vaccine designing [41]. These proteins include those that are part of pathogen core proteome, host non-homologous, and essential for pathogen survival [42]. Pathogen core proteins were identified using CD-HIT [43] that removed proteins sharing sequence identity of 60% [27], [37]. Moving ahead, the non-paralogous proteins of TRAB were used in a BLASTp search against *H. sapiens* (taxid: 9606) with the inclusion of the E-value threshold set to 0.005, bit score ≥ 100 and sequence identity of $\leq 30\%$ [43]. Proteins that were adequately similar were filtered, while the remaining non-similar proteins were removed and considered as host non-homologous proteins. Essential proteins of TRAB were identified by comparing the non-similar proteins against the DEG [44]. The BLASTp search was performed with an E-value cut-off > 0.0001 , minimum bit score of 100 and sequence identity of $\geq 35\%$ [45].

2.2. Exo-proteome and Secretome Prediction

The exo-proteome and secretome of TRAB consist of those proteins that are localized in the extracellular and outer membrane, respectively [42]. A comparative subcellular localization strategy was applied for the prediction of the TRAB exo-proteome and secretome [46], [47]. First, PSORTb v 3.0 [48] was used to predict cytoplasmic, inner membrane, outer membrane, extracellular and unknown proteins. Those proteins, which were categorized as extracellular and outer membrane, were cross-validated by CELLO v 2.5 [49] and CELLO2GO [50]. The proteins of unknown localization from PSORTb, CELLO and CELLO2GO analysis were discarded to achieve consistency in the results. At the end, those predicted to be localized in extracellular and outer membrane regions by all three tools were selected and analyzed further in subsequent steps [42], [47].

2.3. Virulent Proteins Evaluation

Virulent proteins in the pool of essential exo-proteome and secretome proteins were identified using a BLASTp search of the VFDB [51] with a threshold of bit score ≥ 100 and sequence identity cut-off 30% [37].

2.4. Screening of Non-probiotic Proteins

Probiotic bacteria are friendly bacteria and inhabit the gut of humans and animals [52]. Selection of TRAB proteins that show significant homology with probiotic bacteria can lead to inaccurate and non-targeted immune responses and ultimately result in accidental inhibition of probiotic bacteria [53]. In order to avoid this, a BLASTp search of NCBI against three reference strains of intestinal lactic acid bacterial species was performed: *Lactobacillus rhamnosus* (taxid: 47715), *Lactobacillus casei* (taxid: 1582), and *Lactobacillus johnsonii* (taxid: 33959). The

identity sequence cut-off was set to $\leq 30\%$, E-value threshold set to 0.005 and minimum bit score of 100.

2.5. Screening of Non-similar Mouse Proteins

The proteins passed forward by the previous filter were further used in a BLASTp search against the mouse proteome (taxid: 10088) and those with threshold E-value of 0.005 and sequence similarity of $\leq 30\%$ were categorized as mouse non-similar proteins [54].

2.6. Physicochemical Characterization

To prioritize proteins that are most suitable for experimental validations, the proteins of the previous filter were characterized based on several physicochemical parameters [27], [36], [37], [42]. These parameters include: molecular weight, number of transmembrane helices, antigenicity and allergenicity. An online computational tool ProtParm [55] was used to estimate molecular weight, aliphatic index, extinction coefficient, instability index, atomic composition, amino acid composition, estimated half-life and grand average of hydropathicity (GRAVY). The number of transmembrane helices contained within the protein was computed through TMHMM [56] and HMMTOP [57] with the threshold set to 2 [37].

2.7. Antigenicity Prediction

The antigenicity of the proteins was determined through VaxiJen [58]. Proteins with an antigenicity score ≥ 0.4 were considered as likely to be antigenic.

2.8. Prediction of B-cell derived T-cell Epitopes

The prioritized proteins then underwent B-cell epitope prediction using BCpred [59] with epitope length set to 20. B-cell epitopes that have score > 0.8 were selected for analysis of exo-membrane topology using TMHMM. Only surface exposed B-cell epitopes were considered in the subsequent T-cell epitope prediction phase. B-cell derived T-cell epitopes contain both B-cell and T-cell epitopes and can mediate both arms of immunity i.e. humoral and cellular immunity [42]. T-cell epitopes prediction involves evaluation of B-cell epitopes for their effective binding with molecules of both classes of major histocompatibility complex (MHC) i.e. MHC class I and MHC class II [42]. Common binding T-cell epitopes interacting with the maximum number of MHC alleles were pooled out. Binding alleles of MHC class I were predicted by Propred1 [60], while those alleles of MHC class II were identified by Propred [61]. Antigenicity and virulence of the candidate B-cell derived T-cell epitopes were determined by VaxiJen [53] and VirulentPred [62]. The allergenicity of the epitopes was predicted by Allertop [63]. Lastly, the chemical stability and resistant sequence of the epitopes were evaluated through ProtParam program (<http://web.expasy.org/protparam/>) [55] and Comprehensive Antibiotic Resistance Database (CARD) (<https://card.mcmaster.ca/>), respectively.

Fig.1. The complete hierarchy of steps applied in the current study.

2.9. Targeted Proteins Structure Prediction

The antigenic, virulent and non-allergenic epitopes were then subjected to an exo-membrane topology phase, where the surface exposure of the epitopes on targeted proteins was visualized for immunological applications. This was achieved through use of the Pepitope server [64]. Prior to Pepitope analysis, the protein structures were modelled through a comparative structure prediction approach [35]. In this approach, the structure of the target proteins are first predicted by different servers: Phyre2 [65], I-Tasser [66], Modweb [67], RaptorX [68] and Swiss-Model [69]. The best model from each server then underwent a structure evaluation phase. In this phase, the proteins with most of their residues mapped in the favored region of the Ramachandran plot were selected for energy minimization [8]. Energy minimization was essential to relax the protein structures and remove any steric clashes. Minimization was carried out for total of 1500 steps: 750 steps using the steepest descent algorithm and using 750 the conjugate gradient method with a step size of 0.02 Å under the TFF [70].

2.10. Predicting Proteins with a Strong Interactome

The work was extended further to select therapeutic proteins that have strong cellular interactions [71]. Interaction analysis of physiochemically suitable proteins was executed through STRING [72]. Only interactions that have very high confidence score (>0.9) were considered.

2.11. Multi-epitope Vaccine Sequence Construction

The multi-epitope sequence was constructed by linking the B-cell derived T-cell epitopes predicted in previous steps through a GPGPG linker, while the Cholera toxin B subunit (UniProtKB - E9RIX3 (E9RIX3_VIBCL)) as adjuvant was linked to the construct using an EAAAK linker [28]. The antigenicity and allergenicity of the entire construct were re-validated through VaxiJen and Allertope, respectively. The secondary and tertiary structure of the construct were predicted through PSIPRED [73] and RaptorX [68], respectively. Intrinsic disorders in the vaccine-construct were predicted using protein disorder prediction system (PrDOS) [74]. The tertiary structure of the multi-epitope vaccine-construct was further refined through use of the Galaxy Refine server [75]. Finally, to further increase the stability of the multi-epitope sequence construct, it was subjected to Disulfide engineering utilizing Design 2 server [76].

2.12. Molecular Docking

Molecular docking of the multi-epitope vaccine sequence was done with the receptor TRL4 [77]. The three-dimensional (3D) structure of TRL4 was retrieved from the PDB using PDB ID of 4G8A [78]. Molecular docking was performed with PatchDock [79] and subsequently refined using Fast Interaction Refinement in Molecular Docking (FireDock) [80]. The top complex was then minimized using the same parameters discussed above.

2.13. MD Simulations and Binding Free Energy Calculations

The stability of multi-epitope vaccine-construct with the TLR4 receptor was assessed by subjecting the complex to 100-ns of MD simulation carried out using the AMBER 16 package [81]. System preparation was initiated with the tleap tool while energies were calculated using the MPI version of SANDER [82] using the ff03.r1 and General Amber Force Field (GAFF) forcefields [83]. The system was solvated in a water box made up of TIP3P water molecules, with a padding distance of 12 Å between the protein and the edge of the box. Complex neutralization was achieved by adding 28 Na⁺ ions. The complex was then minimized in several stages. Firstly, minimization was carried out with a restraint of force constant $k = 200$ kcal/mol Å² applied to all hydrogen atoms and water molecules for 500 and 1000 cycles, respectively, to remove steric clashes. Next, 1000 steps of minimization were carried out with a restraint of force constant $k = 5$ kcal/mol Å² applied to C-alpha atoms. In the final stage of minimization, 300 steps of minimization were performed with a restraint of force constant $k = 100$ kcal/mol Å² applied to non-heavy atoms. With restraints of 5 kcal/mol Å² on C-alpha atoms, the system was then heated to 300 K for 20 picoseconds (ps). Langevin dynamics [84] was used to maintain system temperature and SHAKE algorithm [85] was applied to bonds involving hydrogen. The constant-temperature, constant-pressure ensemble (NPT) [86] was applied for 50-ps with restraints of 5 kcal/mol Å² on C-alpha atoms to allow the pressure of the system to equilibrate. The system was then equilibrated for 1-ns. A production run of 100-ns was lastly performed using NVT ensemble [87] in combination with Berendsen's algorithm [88] and a time step of 2-fs. Deviations and fluctuations in the protein backbone were monitored throughout the simulation e.g. RMSD [89], RMSF [89], and Rg [89]. The complex was visualized at different stages of the simulation using Visual Molecular Dynamics (VMD) [90] and UCSF Chimera [91]. The strength of interactions between the vaccine-construct and the TLR4 receptor was assessed using Amber16 MM-GBSA method [92]. The overall objective of MM-GBSA is to calculate the free energy difference between the vaccine-construct-TLR4 complex and TLR4 (the receptor) and vaccine-construct (the ligand). Mathematically, it can be represented as:

$$\Delta G_{\text{Binding energy}} = \Delta G_{\text{TLR4-vaccine-construct complex}} - [\Delta G_{\text{TLR4 receptor}} - \Delta G_{\text{vaccine-construct}}]$$

In the equation, ΔG is Gibb's free energy and is calculated by the MM-GBSA method for the following terms,

$$\Delta G = E_{\text{gas}} + \Delta G_{\text{solvation}} - TS_{\text{solute}}$$

In the equation, T stands for temperature and S is the entropy contribution to vaccine binding. The gas energy is the MM energy and involves internal, electrostatic and van der Waals energies,

$$E_{\text{gas}} = E_{\text{Internal energy}} + E_{\text{Electrostatic energy}} + E_{\text{Van der Waals energy}}$$

The term ΔG solvation is calculated using an implicit system and can be split into electrostatic and non-polar contributions,

$$\Delta G_{\text{solvation}} = \Delta G_{\text{electrostatic}} + \Delta G_{\text{np}}$$

ΔG_{ele} involves electrostatic and polar solvation component and can be presented as,

$$\Delta G_{\text{electrostatic}} = E_{\text{electrostatic}} + \Delta G_{\text{GB}}$$

ΔG_{np} is directly proportional to the molecule solvent-accessible surface area,

$$\Delta G_{np} = \gamma SAS + \beta$$

Estimation of ΔG_{np} is achieved using the LCPO [93] approach with the water probe radius set to 1.4 Å. In MM-GBSA calculations, the standard values of γ and β were 0.0072 kcal/mol·Å² and 0 kcal/mol, respectively.

3. Results and Discussion

3.1. Proteome Retrieval and Subtractive Proteomics

Vaccination is considered an alternative and effective stratagem to prevent infection caused by drug resistant bacterial pathogens [94]. In the last few years, a substantial amount of progress has been made in the development of effective vaccines against multiple and extensive drug resistant *A. baumannii* isolates [36], [47], [95], [96], [97], [98], [99]. For active and passive immunization, different vaccines based on bacterial cell or protein antigens have been evaluated. In experimental approaches, promising antigenic-specific antibody responses were observed with varying degree of protection in animal models. However, technical and scientific hurdles remain in the development of an effective vaccine for *A. baumannii* [36].

Recent studies showed that pure antigens can elicit a faster immune response [26], [100] and thus the genome of *A. baumannii* was rigorously explored using *in silico* tools to reveal more potent ones. Subunit vaccine proteins are thought to have potential antigenic determinants: however, their role in immune protection is still questioned and further evidence is required. The appropriate combination of these shortlisted antigens is another consideration as individual antigenic peptides have several limitations; mainly they do not always stimulate strong and accurate immune responses [26]. Additionally, it has been observed that most of the identified antigens are immune-dominant and often overshadow the more effective one present in the pathogen proteome [26]. Generally, an ideal immunotherapy antigen against TRAB should offer several important qualities. Among these, the most critical is being protective to the host [26], [36], [101]. Surface topology, high reproducibility and cost are the other important factors [42]. Proteins that remain stable and soluble in recombinant form after preparation are desirable [44]. Proteins that create a balance between pro-inflammatory and anti-inflammatory cytokines (essential for the development of a well-regulated immune response) and reduce the incidence of heat shock are the most appropriate candidates [102]. Similarly, proteins that induce strong humoral immune response while maintaining moderate cellular immunity are also preferable [26]. The protection of immune-compromised patients also needs to be ensured [101] and the antigen candidate must be suitable for active and passive immunization [101]. Since limited antigen candidates are available for a TRAB vaccine compared to other bacterial pathogens [36], the search for novel antigens is imperative.

The complete proteome of TRAB strain AB031 encompasses 3,438 proteins. The CD-hit analysis revealed 3,354 proteins as non-redundant proteins, while 84 proteins were classified redundant based on the sequence similarity of 60%. Redundant proteins are paralogous in function and arise due to duplication during evolution. Because of multiplication, a defect or

mutation in a redundant protein will have a reduced adverse effect on overall fitness of the organism and as such, they are not considered attractive targets for a vaccine design [103]. Furthermore, redundant proteins are not well conserved across the bacterial species and strains and hence cannot be targeted by broad-spectrum therapeutics [104]. In contrast, orthologous proteins are non-redundant and are considered more conserved across bacterial species and strains, therefore, could be attractive candidates for vaccine design [104]. In the next step of the framework, the non-redundant proteins were aligned with the reference human protein. It was revealed that 556 proteins are human homologs, while 2798 proteins are non-human homologs. Eliminating host homologous proteins is essential as targeting such proteins could result in strong immune responses of the host. In addition, host homologous proteins can integrate and recombine in the host genome, and therefore, are avoided for vaccine development [37], [104]. The identification of essential, non-homologous proteins was crucial because the majority of vaccines target essential cellular process of the pathogenic organism. The essential proteome is a set of proteins prerequisite to maintain cellular life and as such has wider therapeutic potential [27]. In TRAB, essential proteins are involved in major biological pathways and aid survival in extreme environmental stress conditions, infection and persistence in the host, GTP, ATP and nucleotide binding, ligase, phosphatase and transferase activity. The use of essential proteins in vaccine development will ensure the continuous expression of antigenic determinants and thus will evoke the host immune response when it encounters the pathogen [36]. DEG analysis revealed 800 proteins as essential and removed 1998 proteins from further analysis as they were classed as non-essential.

3.2. TRAB Exo-proteome and Secretome

Knowledge of protein subcellular localization is important and can significantly improve the identification of potential vaccine candidates [36]. Secreted and membrane proteins are easily accessible to drug molecules because of their localization in the extra-cellular space or on the organism's cell surface [105] and are of great interest due to their potential as vaccine or diagnostic targets [42]. Similarly, targeting such proteins is also useful as they aid pathogen adherence, invasion, survival and proliferation within the host cell [36], [42]. In a comparative subcellular localization approach, proteins located in the extracellular matrix and outer membrane of the pathogen were identified first through PSORTb v 3.0. The majority of proteins were cytoplasmic (577), 181 were located in the inner membrane, 13 in the outer membrane, 9 were periplasmic, 1 extracellular and 19 could not be classified. The 13 identified outer membrane proteins and 1 extracellular protein from the search by PSORTb were used in CELLO and CELL2GO to cross-validate the protein's localization. According to CELLO, 2 of the proteins were categorized as extracellular while 12 were outer membrane. CELL2GO identified 6 proteins as outer membrane while the remaining 8 were predicted to have dual localization of extracellular and outer membrane. These shortlisted 14 proteins are now non-redundant, host non-homologous, pathogen essential and outer membrane and extracellular proteins and illustrated in **Table 1**. The Venn diagram for number of proteins shortlisted so far is shown in **Fig.2**.

Fig.2. The shortlisted 14 proteins for virulent protein analysis. **C. P.**, Complete proteome, **N. R. P.**, Non-redundant proteome, **N. H. P.**, Non-homologous proteome, **E. P.**, Essential proteome, **O. M.**, Outer membrane, **E. C.**, Extracellular.

Table 1. Fourteen proteins with subcellular localization of outer membrane and extracellular matrix.

Protein ID	Protein	PSORTb	CELLO	CELL2GO
WP_000505931.1	Methyl-accepting chemotaxis protein (MACP)	Outer Membrane	Outer Membrane	Extracellular, Outer Membrane
WP_000557460.1	M23 family peptidase	Outer Membrane	Extracellular	Extracellular, Outer Membrane
WP_001109856.1	Hypothetical protein	Outer Membrane	Outer Membrane	Extracellular, Outer Membrane

WP_001174793.1	Multidrug efflux RND transporter AdeIJK outer membrane channel subunit (AdeK)	Outer Membrane	Outer Membrane	Outer Membrane
WP_004838311.1	Hypothetical protein	Extracellular	Outer Membrane	Outer Membrane
WP_005128115.1	Peptidoglycan-associated lipoprotein (Pal)	Outer Membrane	Extracellular	Extracellular Outer Membrane
WP_025464534.1	Outer membrane protein assembly factor (BamA)	Outer Membrane	Outer Membrane	Outer Membrane
WP_031999435.1	Outer membrane protein assembly factor (BamD)	Outer Membrane	Outer Membrane	Extracellular Outer Membrane
WP_031999452.1	Insulinase family protein	Outer Membrane	Outer Membrane	Outer Membrane
WP_031999870.1	Outer membrane protein assembly factor (FimD)	Outer Membrane	Outer Membrane	Extracellular Outer Membrane
WP_032000062.1	Hypothetical protein	Outer Membrane	Outer Membrane	Extracellular Outer Membrane
WP_032000856.1	Fimbrial biogenesis outer membrane usher protein (FimD)	Outer Membrane	Outer Membrane	Outer Membrane
WP_038405564.1	TonB-dependent siderophore receptor	Outer Membrane	Outer Membrane	Outer Membrane
WP_079378209.1	Type IV secretion protein (Rhs)	Outer Membrane	Outer Membrane	Extracellular Outer Membrane

3.3. TRAB Virulent Proteins

Virulence is a key factor of bacterial pathogens that facilitates organism pathogenesis as well as survival in adverse environmental conditions [106]. Compared to other bacterial pathogens, very little is known about TRAB's pathogenic potential and virulence determinants [36]. In order to get insights about the virulent potential of TRAB, virulent proteins in the essential exo-proteome and secretome were explored. In total, six proteins were identified as virulent. These include: methyl-accepting chemotaxis protein (MACP) (Bit score: 562), outer membrane protein assembly factor (BamA) (Bit score: 241), multi-drug efflux RND transporter (RND) (Bit score: 285), TonB-dependent siderophore receptor (TonB) (Bit score: 219), type IV secretion protein Rhs (Rhs) (Bit score: 259) and fimbrial biogenesis outer

membrane usher protein (FimD) (Bit score: 591). MACP is a sensor protein that allows bacteria to sense molecule concentration in the extracellular matrix and mediate smooth swim or tumble accordingly [107]. In the case of rising levels of nutrients or declining levels of toxins, the bacteria will continue to swim smooth or swim forward [97]. On the contrary, if the concentration of the nutrients is declined or the level of toxins is on the rise, the bacteria reorient themselves or will tumble [107], [108]. In this way, the bacteria may swim away from toxins or towards nutrients [107], [108]. BamA is an important part of the outer membrane and plays a vital role in assembly and insertion of beta-barrel proteins into the outer membrane of Gram-negative bacteria [109]. RND is Resistance-Nodulation-Division family transporter, widespread among Gram-negative bacteria to actively efflux chemotherapeutic agents including antibiotics [110]. TonB is a bacterial outer membrane protein that binds and transports ferric chelates called siderophores, as well as carbohydrates, nickel complexes and vitamin B₁₂ [111]. Similarly, this protein provides immunity to bacteria against antibiotics, detergents, digestive enzymes and immune surveillance [111]. Rhs is a member of the Type IV secretion proteins and plays role in the transportation of molecules across the membrane [112]. Lastly, FimD protein is involved in pilus biogenesis in Gram-negative bacteria [113].

3.4. Screening of Non-probiotic Bacterial Proteins

Probiotic bacteria are useful organisms that line the digestive tract and perform a significant role in promoting host health and disease prevention [114]. In the gut, these microorganisms play an important role in preventing the growth of harmful bacterial species by creating hydrolytic enzymes, training the host immune system by stimulating the secretion of IgA and regulatory T-cells, producing hormones that store body fats, producing vitamins for the host such as biotin, butyrate and vitamin K2, and regulating the development of the gut [52]. Accidental inhibition of these microbes could lead to severe side effects; therefore, it is imperative to discard proteins of such microbes from the design framework [53]. All the proteins were reported to possess no homology with the shortlisted proteins and hence are safe to be used in vaccine for targeted immune responses.

3.5. Screening of Non-similar Mouse Proteins

In this stage, proteins that were identical to mouse proteins were removed. This was essential as evaluating immune protection efficacy of proteins similar between the bacteria and mouse can lead to autoimmune diseases thus limiting experimental evaluations and producing false positive hits [54]. In addition, the proteins are less likely to be immunogenic because of epitope mimicry. All the six proteins were found non-homologous to mouse proteome and thus analyzed further in the framework.

3.6. Physiochemical Prioritization of Vaccine Proteins

The most crucial factor during this phase was the estimation of the protein's molecular weight [36], [37]. Proteins having molecular weight < 110 kDa are considered more appropriate as they can easily be purified and successfully subjected to vaccine development [27]. Molecular weights of the proteins were estimated as: Rhs (180.35 kDa), FimD (94.7 kDa), BamA (92.27 kDa), TonB (80.84 kDa), MACP (75.38 kDa) RND (52.82 kDa). Despite its molecular weight

falling outside the desired range, the Rhs protein was not discarded at this stage as this protein has not previously been explored and it contains potential antigenic epitopes. The theoretical pI for MACP, BamA, FimD, Rhs and TonB was found in the acidic category with range 4.93 to 5.0. The RND was predicted to be acidic with a theoretical pI of 9.02 [46]. The Grand average of hydropathicity (GRAVY) index score for all the six proteins was negative indicating their hydrophilic nature [115]. The stability index computed for the proteins was found to be < 40 , thus considered as stable [116]. The aliphatic index was used to indicate the thermostability of the proteins. The average aliphatic index value for the proteins is 83.96, indicating high thermostability. The extinction coefficient i.e. the light amount observed by a protein at wavelength of 280 nm in water [117], is a useful parameter to indicate the ease of protein purification. Two kinds of values were noted: one considered all cysteine values as half cystines, while the other considered no cysteine as half cystines. Cystine is an amino acid formed as a result of two cysteine joined together by a disulphide bond. The extinction coefficient for the proteins is shown in **S-Table.1**. Further, all the six proteins were analyzed for number of trans-membrane helices. Proteins having trans-membrane helices < 2 were selected as they are easy to clone and express during experimental evaluation [27], [36], [37]. All the proteins were found to have < 2 trans-membrane helices: MACP (2), BamA (0), RND (0), TonB (0), Rhs (0) and FimD (0).

3.7. Antigenicity Prediction

All the six proteins were found antigenic: MACP (0.44), BamA (0.61), RND (0.63), TonB (0.60), Rhs (0.53) and FimD (0.57).

3.8. Prediction of B-cell derived T-cell Epitopes

Filtration of protein candidates suitable for the production of antigenic peptides could not only help in minimizing the time, resource and labor but also at the same time optimize the chances of success of getting the most suitable vaccine for the pathogen [36]. The first step in this process was to predict B-cell epitopes for each protein. The predictions were: 31 for Rhs, 23 for FimD, 15 for BamA, 14 for TonB, 6 for MACP, and 6 for RND. Only B-cell epitopes with a B-cell epitope prediction score of > 0.8 were investigated for surface exposure as epitopes located near the surface are easily recognized by immune cells and can be efficiently targeted [26], [36]. A single exo-membrane B-cell epitope was identified for MACP, but more were predicted for the others: 4 for RND, 6 for BamA, 9 for FimD, 9 for TonB and 13 for Rhs. Surface exposed B-cell epitopes were then analyzed for binding alleles of MHC-I and MHC-II. Only those epitopes were considered for which common binding alleles in both classes were predicted [47]. MACP was excluded from further analysis at this stage as only MHC-I epitopes were predicted. The other epitopes contained several epitopes common to both MHC classes [BamA 3, RND 4, TonB 5, FimD 6, and Rhs 7]. The next stage of screening involved the prediction of the antigenicity of the epitopes. All 4 epitopes identified for the RND transporter protein were found to be non-antigenic, excluding this target from further analysis. For BamA all 3 epitopes were found to be antigenic. For FimD 3 out of 6 of the epitopes were antigenic and 4 out of 5 were for TonB. Rhs had the most antigenic epitopes with 4 out of 7 found to be antigenic. The affinity of the screened epitopes for the DRB*0101 (highly

prevalent allele in the human population) was determined. This was vital as epitopes having higher affinity for DRB*0101 produce robust and accurate immune responses. Only epitopes with an affinity < 100 nM were considered. For outer membrane protein assembly factor BamA, out of 3 antigenic epitopes only 1 was found with an IC50 value in the correct range (FPLNDKPGD IC50 = 91.62 nM). For FimD, 2 antigenic B-cell derived T-cell epitopes: FVHAEEAAA and MVLVHAPDA, were found to have higher affinity for the DRB*0101 allele with IC50 values of 3.52 and 17.42 nM, respectively. Two of the antigenic epitopes for TonB had acceptable IC50 values: YQLLDGQEV (IC50 = 26.36 nM) and YNSDGQSR (IC50 = 10.16 nM). For Rhs, three epitopes: YVVAGTAAA (IC50 = 4.90 nM), YYYFDLNGF (IC50 = 0.92 nM) and YEGITIQPL (IC50 = 11.02 nM) were selected. The virulent potential of these shortlisted epitopes was another significant consideration as relates to ability to evoke a response from the host immune system [42]. All the epitopes of the remaining vaccine candidates: BamA, FimD, TonB and Rhs had a virulent potential above the threshold (> 0.5). Lastly, in this phase, only non-allergenic epitopes were considered to avoid any side effects of the epitopes that will be used for multi-epitope vaccine construction [118]. For BamA, “FPLNDKPGD” was found non-allergenic therefore considered. For FimD protein, epitope “FVHAEEAAA” was found non-allergenic, while all the screened epitopes of TonB were allergenic, removing it from further processing in the pipeline. For Rhs, only epitope “YVVAGTAAA” was a suitable candidate. Lastly, we found all the predicted epitopes as stable and do not have any resistant sequence. The predicted B-cell derived T-cell epitopes for the potential three vaccine proteins are illustrated in **Table 2**.

Table 2. Predicted B-cell derived T-cell epitopes for potential three vaccine proteins.

Protein	B-cell epitope	B-cell derived T-cell epitope	Antigenicity	MHCpred IC50 score	Virulent Pred	Allergenicity
BamA	AFPLNDKPGDETKEIQFEIG	FPLNDKPGD	1.4	91.6	1.0	Non-allergen
FimD	TMPVVFVHAEEAAASAPVEAE	FVHAEEAAA	1.0	3.5	1.0	Non-allergen
Rhs	AASVAGYGPYVVAGTAAAGS	YVVAGTAAA	0.6	4.9	1.0	Non-allergen

3.9. Structure Prediction and Evaluation of Shortlisted Epitopes Proteins

The exo-membrane topology of the shortlisted epitopes on their respective protein surfaces was visualized. Before this, structures of the shortlisted proteins were predicted and optimal structures selected. Generally, structures predicted by RaptorX were selected. For BamA, the structure modelled has majority of its residues (92.0%) are plotted in the most favorable region. The residues in the additionally allowed region, generously allowed region, and disallowed region are 7.2%, 0.7% and 0.1%, respectively. The overall G-factor for the protein is -0.12. Additionally, the Errat, Verify-3D and Prosa Z-score for the protein are 53.01, 68.42 and -8.45, respectively. Similarly, structure of FimD was analyzed to have 88.1% residues in

the most favorable region of the Ramachandran plot, while residues in the disallowed region were 0.8%. The overall G-factor score for the protein is -0.15 . The Prosa Z-score for the model was highly acceptable i.e. -9.49 , while the Errat and Verify-3D scores were 61.84 and 77.70, respectively. In case of Rhs, the structure has 83.6% of residues in most favorable region and 1.8% in the disallowed. The G-factor for the structure is -0.21 . Errat, Verify-3D and Prosa Z-score for the structure are 77.01, 52.06 and -1.44 , respectively. The predicted structures for the three proteins are shown in **Fig. 3**, while the exo-membrane topology of the predicted epitopes are shown in **Fig. 4**.

Fig. 3. Tertiary structure of the shortlisted potential vaccine proteins. From left to right, BamA, FimD and Rhs.

Fig. 4. Exo-membrane topology of the shortlisted antigenic epitopes on their respective protein surface.

3.10. Interacting Network Analysis

Interacting network analysis was carried out for BamA, FimD and Rhs that contain desired antigenic B-cell derived T-cell epitopes. This interacting network analysis allows the interpretation of direct and indirect interactions between the mentioned proteins and other proteins of the bacteria that mediate cellular process and metabolic/signaling pathways [26]. No direct interactions among the proteins were observed. Individual clusters were constructed for BamA (**Fig. 5A**) and FimD (**Fig. 5B**), no cluster could be produced for Rhs as no similar sequences were found. For BamA using stringent interaction criteria of 0.9, a cluster of 8 proteins directly interacting with BamA protein was identified. Of these 8 proteins, BamB, BamD and BamE together with BamA form the BAM complex and catalyze folding and insertion of β -barrel proteins into the Gram-negative bacteria outer membrane [119]. SurA is a chaperone involved in maturation and efficient production of several outer membrane proteins [120]. RseP is a zinc metalloprotease enzyme [121]. FabZ is an enzyme of unsaturated fatty acids biosynthesis and catalyzes the dehydration of long and short chain, saturated and unsaturated beta-hydroxyacyl-ACPs [122]. LpxD is an enzyme of the Lpx family and is involved in the synthesis of lipid A molecules that make up part of the Gram-negative outer membrane [123]. Peptidoglycan associated protein (Pal) is an outer membrane protein essential for survival and pathogenesis of bacteria [124]. Three proteins were identified in the interacting network analysis of FimD: F911_01973 (fimbrial protein), HMPREF0010_00598 (Gram-negative pili assembly chaperone domain protein) and F911_02226 (Gram-negative pili assembly protein) [113].

Fig.5. Interacting network of (A) BamA and (B) FimD. Empty balls are proteins with unknown 3D structure while filled balls are proteins with known 3D structure.

3.11. Secondary and Tertiary Structure of Construct

The secondary structure of the vaccine-construct contains 38% of alpha helix, 19% of beta strand and 44% of coil structure as shown in **Fig. 6**. RaptorX was used to predict 5 tertiary structures for the vaccine-constructs which were then assessed by Ramachandran analysis to select the best model (**Table 3**). Model 5 was selected as it had the most favorable positioning of residues in allowed regions. The G value of 0.74 and the ProA Z-score of -5.76 also indicate that this structure is energetically stable. The Errat quality factor and Verify-3D score for the model were 84.81 and 77.71, respectively. The tertiary structure, ProSA Z-score and Ramachandran plot of the most optimal multi-epitope peptide structure are illustrated in **Fig. 7**.

Fig. 6. Secondary structure of the multi-epitope peptide vaccine.

Table 3. Structural evaluation of predicted five models of the multi-epitope peptide vaccine.

Model	Most favored regions [A,B,L]	Additional allowed regions [a,b,l,p]	Generously allowed regions [~a,~b,~l,~p]	Disallowed regions [XX]	OVERALL G-Factor (AVERAGE -0.71*)
Model-1	93 (65.5%)	41 (28.9%)	5 (3.5%)	3 (2.1%)	-0.71
Model-2	85(59.9%)	36 (25.4%)	15 (10.6%)	6 (4.2%)	-1.52
Model-3	85 (59.9%)	39 (27.5%)	9 (6.3%)	9 (6.3%)	-1.32
Model-4	84 (59.2%)	47 (33.1%)	8 (5.6%)	3 (2.1%)	-0.98
Model-5	94 (66.2%)	38 (26.8%)	8 (5.6%)	2 (1.4%)	-0.74

Fig.7. (A) Tertiary structure (B) ProsA Z-score and (C) Ramachandran plot of the multi-epitope peptide vaccine.

3.12. Intrinsic Disorder Regions Prediction

Intrinsically disordered regions of proteins are those regions that lack an ordered 3D structure, commonly found in linker or loop regions [125]. It has been recognized that these disordered regions allow the vaccine-construct the flexibility to adopt different conformations required for binding to the catalytic binding pocket of the receptor protein. **Fig. 8** infers that majority of the vaccine-construct residues are well below the threshold, indicating that the majority of the vaccine-construct residues lie in ordered regions. Two regions: one from the adjuvant (Met1-Leu4) and the second from the B-cell derived T-cell epitope region (Pro131-Ala166) are present in disordered regions. These regions are important for the proper adjustment of the construct in the receptor pocket.

Fig. 8. Intrinsic disorder graph predicting that the majority of the multi-epitope peptide regions are below the threshold value.

3.13. Tertiary Structure Refinement and Disulfide Engineering

Refinement of the construct structure was vital to reconstruct side chains, followed by repacking MD simulations to relax the modelled structure [75]. The Galaxy refine server is one of the best tools for the refinement of both the global and local structural quality of protein models [77]. Disulfide engineering helps to increase protein stability by reducing the conformational entropy of the unfolded state of the protein [76], [77]. The GalaxyRefine server generated five refined models for the vaccine-construct. Model-1 was the best model according to several parameters including: GDT-HA (0.9292), MolProbity (2.538), RMSD (0.473), Poor rotamers (1.6), Clash score (25.1), and Ramachandran favored plot (91.5). In Disulphide engineering, 8 pairs of residues in the construct were observed that violate the recommended values of energy (cut-off <

2.2) and Chi3 should be in between -87 to $+97$. These residues include: Lys5, Gly7-Val8, Thr11, Ser15-Ala17, His34, Tyr48, Ala59, Gln70, His78, Leu98, Cys107, Trp109 and Ala151.

3.14. Codon Optimization of the Vaccine-construct

The codon usage of the final vaccine-construct in *E.coli* K 12 strain was optimized to assure maximum protein expression [126]. The optimal GC content should lie in the range 30% to 70% [126]. The average GC content of the improved sequence was 50.80 %, indicating that good expression of the vaccine protein was likely in the *E.coli* host. The Codon Adaptation Index (CAI) of the optimized sequence is 0.92. The recombinant clone with inserted adopted codon sequence is shown in **Fig. 9**.

Fig. 9. *In silico* prediction of the cloning of the final multi-epitope peptide vaccine-construct (red) into pET28a expression vector.

3.15. Molecular Docking of the Vaccine-construct with the TLR4 Immune Receptor

In total 189 binding pockets were identified spanning the amino acids from all the four chains: A, B, C and D. Details of the best binding pocket is given in **Fig.10** and **S-Table 2**.

Fig. 10. The best predicted binding side of the TLR4 receptor protein. The red spheres indicate the active site region of the protein.

The surface volume of the pocket was $13287.162 \text{ \AA}^3$, surface area $8770.908 \text{ \AA}^2$, molecular area of the cavity mouth was $11028.235 \text{ \AA}^2$ and molecular circumference was $2563 \text{ \AA}^2$. Molecular docking generated 10 models ranked according to electrostatic complementarity and geometry of the protein surface (**Table 4**). The molecular docking solutions were further refined (**Table 5**) based on the Global energy, the lower the Global binding energy the more stable the complex and vice versa. Two solutions were found that have lower Global binding energies: Solution 4 and 6. Solution 6 was ranked top complex based on the lowest Global binding energy of -8.16 kcal/mol . The attractive and repulsive van der Waals interactions for the complex were -18.30 kcal/mol and 22.83 kcal/mol , respectively. The atomic contact energy for the complex is 5.96 kcal/mol and hydrogen bonding energy of -2.42 kcal/mol . Visualization of the top complex

showed binding of the vaccine-construct within the binding pocket formed by chain A, B and C. The vaccine-construct was seen docked deeply inside the pocket with helical region (Pro1509-Tyr1519) from the Cholera toxin B subunit sequence facing towards the wall of chain A and B of TLR4. The epitope from FimD was used to support the construct on chain C. The binding mode and interactions of the multi-epitope vaccine-construct at the binding site of TLR4 receptor are illustrated in **Fig.11**.

Table 4. Top 10 generated models by PatchDock.

Solution Number	Score	Area	Atomic Contact Energy (ACE) (kcal/mol)
1	21294	3038.8	37.0
2	18344	2662.0	-137.0
3	18042	3391.8	252.7
4	17776	2436.2	403.8
5	17760	2925.7	406.9
6	17456	3335.0	-20.0
7	17290	2123.2	234.9
8	16930	2561.4	380.8
9	16874	2104.0	415.4
10	16834	2471.8	-390.7

Table 5. FireDock refinement of PatchDock models.

Rank	Solution Number	Global Energy (kcal/mol)	Attractive VDW	Repulsive VDW	ACE (kcal/mol)	HB
1	6	-8.1	-18.3	22.8	5.9	-2.4
2	4	-3.2	-39.4	20.9	19.0	-2.2
3	8	6.9	-26.8	21.8	8.1	-2.9
4	2	11.0	-0.5	0.0	0.0	0.0
5	7	21.3	-56.5	77.5	5.9	-5.9
6	9	69.5	-55.2	158.3	14.0	-4.5
7	10	266.8	-54.8	442.1	-1.2	-6.5
8	1	695.9	-63.8	924.1	-0.2	-7.3
9	5	4056.5	-95.4	5145.7	18.6	-10.2
10	3	9406.4	-77.5	11829.9	7.4	-12.9

VDW, van der Waals forces, ACE, Atomic contact energy, HB, Hydrogen bonding

3.16. MD Simulations and MMGBSA based Binding Free Energy Calculations

The dynamic behavior of the vaccine-construct and TLR4 complex was studied through MD simulations. Three physical parameters were evaluated shown in **S-Fig. 1**. The atomic displacement of the C α atoms for the complex was explored first by calculating the RMSD as a function of time [89]. An average RMSD value calculated for the complex is 4.85 Å with maximum RMSD observed at time scale from 90-97 ns. The trajectory was visually inspected to explain the pattern of RMSD variations. Snapshots at different nanoseconds: 8th, 18th, 29th, 43th, 49th, 72th, 81th and 90th ns were extracted from the trajectories. All these snapshots were superimposed in Chimera to investigate the structural variations and decide whether these variations are vaccine-construct induced or part of the catalytic mechanism of TLR4. Generally, all the variations observed were due to the orientation of the vaccine-construct in the pocket of the TLR4 resulting in the rise of RMSD. The rise of RMSD throughout the simulation can be explained by the relaxation of the TLR4 chains in response to the presence of the vaccine-construct in the binding pocket. The vaccine-construct orientations in the binding cavity of TLR4 are shown in **Fig. 12**. The structural mobility of the complex was investigated by plotting the RMSF as a function of residue number. The average RMSF calculated for the system is 2.48 Å, with the first 594 residues having RMSF values around 3 Å. Higher fluctuations are seen in the other residues. Rg in general measures the overall packing quality and density of a macromolecular system [89], [127]. Variation in the pattern of Rg generally indicates changes in the compactness of the protein structure. Higher the Rg indicate loose packing of the system and low values indicate compact structures [89], [127]. The mean Rg value estimated for the system is 41.51 Å with highest Rg value observed at 47th ns with value of 42.23 Å. It was noticed that variation in the Rg value is due to the multi-epitope vaccine-construct induced variations in the protein structure. The protein structure adjustments mainly were noticed at the loop regions and seem necessary to properly recognize the construct. Hence, the variations observed in the Rg analysis complements that of RMSD.

Fig. 12. Vaccine-construct adjustments at the active site of TLR4 receptor during the simulation period. The multi-epitope peptide is in gold, while the TLR4 receptor is in dark magenta.

Molecular docking can be employed to predict the binding mode of ligands or to discriminate between binders and non-binders [128], [129], but they cannot really discriminate between ligands of very similar binding affinity e.g. < 6 kJ/mol [130]. More accurate methods, known as alchemical perturbation (AP) methods, exist that are based on statistical mechanics [128], but they require extensive Monte Carlo (MC) or MD simulations to sample conformational space, limiting their widespread application [128]. Between these two extremes, there are methods with intermediate performance that do sampling at the end states, therefore, referred as end point methods [130]. End point methods are generally more accurate than scoring functions and are less computationally expensive than AP methods [131]. The arguably most popular end point method is MM-GBSA [129], [130]. This method was developed by Kollman *et al.* in the late 90s and has been used successfully in wide range of settings including conformer stability, protein-protein interactions, protein design and re-scoring [132]. The total MM-GBSA binding energy of the complex was estimated to be -17.48 kcal/mol. The van der Waals energy (ΔE_{VDW}) reported for the complex is -98.33 kcal/mol and dominates the overall energies of the complex. Non-favorable contributions for the complex were observed for the electrostatic energy (ΔE_{ele}) that was 6.01 kcal/mol. The gas phase energy (ΔE_{gas}) for the complex was revealed favorable with value -92.31 kcal/mol. The polar solvation energy contribution towards the total binding energy is non-favorable with score of 74.83 kcal/mol. The energy for polar component of solvation energy (ΔG_{GB}) is 87 kcal/mol, which illustrates their non-favorable contribution in vaccine-construct binding to TLR4 receptor binding site.

To the best of our knowledge, the present study reports the first work that describes the *in silico* design of a multi-epitope vaccine for TRAB. The vaccine-construct is composed of antigenic epitopes from three surface exposed proteins: BamA, FimD and Rhs. These proteins were prioritized based on several important parameters that are vital in vaccine design [36], [37], [42]. The BamA protein plays an essential role in outer membrane protein biogenesis [109]. Outer membrane protein biogenesis is a vital process to maintain the indispensable functions of bacterial cell envelope, which facilitates the transmembrane movement of nutrients and signaling molecules and at the same time act as a protective physical barrier [109]. Recently, in a combined study of both *in silico* and experimental immunoprotective analysis revealed that compared to previous studies, BamA as a potential candidate for vaccine designing against infections caused by *A. baumannii* [133]. Immunization with BamA conferred 80% of protection against lethal doses of the bacterium and produces high titer of macrophages depended opsonophagocytic antibodies. FimD allows the assembly and secretion of P pili fragments [113]. Pili proteins have been recognized previously to form the core proteome of the *A. baumannii* by *in silico* studies and as such can be potential vaccine candidates against the bacterium. A putative pilus assembly protein, FliF was first identified as the most promising protein using *in silico* studies. Following experimental follow up, the protein was found to elicit antibody titer of $> 64,000$ in mouse model with 50% survival rate from the *A. baumannii* challenge [96]. No previous studies were found regarding the use of Rhs protein as vaccine candidate against *A. baumannii* and therefore can be considered as a novel vaccine protein against the pathogen. The pure antigens of these shortlisted proteins were formulated in multi-

epitope vaccine in the present study that may provide a best-adopted strategy against this rogue pathogen.

4. Conclusions

In the present study, using bioinformatics and immunoinformatics tools, we designed a multi-epitope vaccine for the most troublesome bacterial pathogen, TRAB. The vaccine-construct includes the safest and antigenic epitopes from three proteins: BamA, FimD and a novel Rhs vaccine candidate. These proteins fulfilled all the requirements for potential subunit vaccine candidates. The predicted B-cell derived T-cell epitopes for these proteins were used to design a multi-epitope peptide vaccine to engage properly both cellular and humoral immunity and enhance antigenicity of the individual peptides. Molecular docking and MD simulations show the high stability of the complex of the construct with the active site of the TLR4 receptor. The affinity of the construct for the TLR4 receptor was further validated through MM-GBSA calculations. This multi-epitope peptide vaccine can potentially be used against TRAB infections in both therapeutic and prophylactic situations.

Supplementary Files

S-Table 1. Extinction coefficients for the shortlisted proteins

S-Table 2. Active site residues of TLR4.

S-Fig.1. A. RMSD, B. RMSF, C. Rg,

Conflict of Interest

The authors declare that they have no conflict of interest.

Acknowledgments

Authors are highly grateful to the Pakistan-United States Science and Technology Cooperation Program (Grant No. Pak-US/2017/360) and the Higher Education Commission (HEC) for granting the financial assistance.

References

- [1] A. Howard, M. O'Donoghue, A. Feeney, R.D. Sleator, *Acinetobacter baumannii*: an emerging opportunistic pathogen, *Virulence*. 3 (2012) 243–250.
- [2] M.K. Almaghrabi, M.R.P. Joseph, M.M. Assiry, M.E. Hamid, Multidrug-Resistant *Acinetobacter baumannii*: An Emerging Health Threat in Aseer Region, Kingdom of Saudi Arabia, *Can. J. Infect. Dis. Med. Microbiol.* 2018 (2018).
- [3] A.Y. Peleg, H. Seifert, D.L. Paterson, *Acinetobacter baumannii*: emergence of a successful pathogen, *Clin. Microbiol. Rev.* 21 (2008) 538–582.
- [4] C.-R. Lee, J.H. Lee, M. Park, K.S. Park, I.K. Bae, Y.B. Kim, C.-J. Cha, B.C. Jeong, S.H. Lee, Biology of *Acinetobacter baumannii*: pathogenesis, antibiotic resistance mechanisms, and prospective treatment options, *Front. Cell. Infect. Microbiol.* 7 (2017) 55.
- [5] P. Shrestha, B.S. Cooper, J. Coast, R. Oppong, N.D.T. Thuy, T. Phodha, O. Celhay, P.J. Guerin, H. Wertheim, Y. Lubell, Enumerating the economic cost of antimicrobial resistance per antibiotic consumed to inform the evaluation of interventions affecting their use, *Antimicrob. Resist. Infect. Control.* 7 (2018) 98.
- [6] V. Manchanda, S. Sanchaita, N.P. Singh, Multidrug resistant acinetobacter, *J. Glob. Infect. Dis.* 2 (2010) 291.
- [7] Center for Disease Control, Prevention, others, Antibiotic/antimicrobial resistance: biggest threats, (2016).
- [8] S. Skariyachan, M. Manjunath, N. Bachappanavar, Screening of potential lead molecules against prioritised targets of multi-drug-resistant-*Acinetobacter baumannii*-insights from molecular docking, molecular dynamic simulations and in vitro assays, *J. Biomol. Struct. Dyn.* (2018) 1–24.
- [9] R. Mohd, T. Nesam, L. Kamaruzaman, R.A. Cader, R. Mustafar, W.-Y. Kong, Community acquired multi drug resistant (MDR) *Acinetobacter baumannii* pneumonia in Malaysia--A case report, *Respir. Med. Case Reports*. 24 (2018) 147–149.
- [10] J.M. Pogue, T. Mann, K.E. Barber, K.S. Kaye, Carbapenem-resistant *Acinetobacter baumannii*: epidemiology, surveillance and management, *Expert Rev. Anti. Infect. Ther.* 11 (2013) 383–393.
- [11] H. Spapen, R. Jacobs, V. Van Gorp, J. Troubleyn, P.M. Honoré, Renal and neurological side effects of colistin in critically ill patients, *Ann. Intensive Care.* 1 (2011) 14.
- [12] D.E. Karageorgopoulos, T. Kelesidis, I. Kelesidis, M.E. Falagas, Tigecycline for the treatment of multidrug-resistant (including carbapenem-resistant) *Acinetobacter* infections: a review of the scientific evidence, *J. Antimicrob. Chemother.* 62 (2008) 45–55.
- [13] D. Curcio, F. Fernandez, Tigecycline for *Acinetobacter baumannii* infection: other considerations, *Clin. Infect. Dis.* 46 (2008) 1797–1798.
- [14] J.A. Viehman, M.H. Nguyen, Y. Doi, Treatment options for carbapenem-resistant and

- extensively drug-resistant *Acinetobacter baumannii* infections, *Drugs*. 74 (2014) 1315–1333.
- [15] D. Dixit, R.P. Madduri, R. Sharma, The role of tigecycline in the treatment of infections in light of the new black box warning, (2014).
- [16] S. Navon-Venezia, A. Leavitt, Y. Carmeli, High tigecycline resistance in multidrug-resistant *Acinetobacter baumannii*, *J. Antimicrob. Chemother.* 59 (2007) 772–774.
- [17] C. Kulah, G. Celebi, E. Aktas, Z. Mengeloglu, F. Comert, H. Ankarali, Unexpected tigecycline resistance among *Acinetobacter baumannii* Isolates: high minor error rate by Etest, *J. Chemother.* 21 (2009) 390–395.
- [18] N. Taneja, G. Singh, M. Singh, M. Sharma, Emergence of tigecycline & colistin resistant *Acinetobacter baumannii* in patients with complicated urinary tract infections in north India, *Indian J. Med. Res.* 133 (2011) 681.
- [19] R. Hashemzahi, A. Doosti, M. Kargar, M. Jaafarinia, Cloning and expression of nlpA gene as DNA vaccine candidate against *Acinetobacter baumannii*, *Mol. Biol. Rep.* (2018) 1–7.
- [20] F.E. Andre, R. Booy, H.L. Bock, J. Clemens, S.K. Datta, T.J. John, B.W. Lee, S. Lolekha, H. Peltola, T.A. Ruff, others, Vaccination greatly reduces disease, disability, death and inequity worldwide, *Bull. World Health Organ.* 86 (2008) 140–146.
- [21] D.A. Kennedy, A.F. Read, Why does drug resistance readily evolve but vaccine resistance does not?, *Proc. R. Soc. B Biol. Sci.* 284 (2017) 20162562.
- [22] I.P. Nascimento, L.C.C. Leite, Recombinant vaccines and the development of new vaccine strategies, *Brazilian J. Med. Biol. Res.* 45 (2012) 1102–1111.
- [23] P.D. Minor, Live attenuated vaccines: Historical successes and current challenges, *Virology*. 479 (2015) 379–392.
- [24] A. Vitrinel, G.Ç. Erdaug, Principles of immunization, *Turkish J. Fam. Pract.* 11 (2008) 149–153.
- [25] S. Chang, J.R. Dunn, M. Heidari, L.F. Lee, J. Song, C.W. Ernst, Z. Ding, L.D. Bacon, H. Zhang, Genetics and vaccine efficacy: Host genetic variation affecting Marek's disease vaccine efficacy in White Leghorn chickens, *Poult. Sci.* 89 (2010) 2083–2091.
- [26] W. Li, M.D. Joshi, S. Singhanian, K.H. Ramsey, A.K. Murthy, Peptide vaccine: progress and challenges, *Vaccines*. 2 (2014) 515–536.
- [27] S. Baseer, S. Ahmad, K.E. Ranaghan, S.S. Azam, Towards a peptide-based vaccine against *Shigella sonnei*: A subtractive reverse vaccinology based approach, *Biologicals*. (2017).
- [28] M. Saadi, A. Karkhah, H.R. Nouri, Development of a multi-epitope peptide vaccine inducing robust T cell responses against *brucellosis* using immunoinformatics based approaches, *Infect. Genet. Evol.* 51 (2017) 227–234.
- [29] N. Nezafat, M. Eslami, M. Negahdaripour, M.R. Rahbar, Y. Ghasemi, Designing an

- efficient multi-epitope oral vaccine against *Helicobacter pylori* using immunoinformatics and structural vaccinology approaches, *Mol. Biosyst.* 13 (2017) 699–713.
- [30] D.S. Rosa, S.P. Ribeiro, S.G. Fonseca, R.R. Almeida, V.C. Santana, J. de S. Apostólico, J. Kalil, E. Cunha-Neto, Multiple approaches for increasing the immunogenicity of an epitope-based anti-HIV vaccine, *AIDS Res. Hum. Retroviruses*. 31 (2015) 1077–1088.
- [31] A. Sette, R. Rappuoli, Reverse vaccinology: developing vaccines in the era of genomics, *Immunity*. 33 (2010) 530–541.
- [32] R.E. Soria-Guerra, R. Nieto-Gomez, D.O. Govea-Alonso, S. Rosales-Mendoza, An overview of bioinformatics tools for epitope prediction: implications on vaccine development, *J. Biomed. Inform.* 53 (2015) 405–414.
- [33] A.S. Clem, Fundamentals of vaccine immunology, *J. Glob. Infect. Dis.* 3 (2011) 73.
- [34] I.J. Amanna, M.K. Slifka, Contributions of humoral and cellular immunity to vaccine-induced protection in humans, *Virology*. 411 (2011) 206–215.
- [35] I.R. Cohen, H.L. Weiner, T-cell vaccination, *Immunol. Today*. 9 (1988) 332–335.
- [36] A. Hassan, A. Naz, A. Obaid, R.Z. Paracha, K. Naz, F.M. Awan, S.A. Muhammad, H.A. Janjua, J. Ahmad, A. Ali, Pangenome and immuno-proteomics analysis of *Acinetobacter baumannii* strains revealed the core peptide vaccine targets, *BMC Genomics*. 17 (2016) 732.
- [37] Y. Asad, S. Ahmad, T. Rungtongmongkol, K.E. Ranaghan, S.S. Azam, Immuno-informatics driven proteome-wide investigation revealed novel peptide-based vaccine targets against emerging multiple drug resistant *Providencia stuartii*, *J. Mol. Graph. Model.* (2018).
- [38] H. Dar, T. Zaheer, M.T. Rehman, A. Ali, A. Javed, G.A. Khan, M.M. Babar, Y. Waheed, Prediction of promiscuous T-cell epitopes in the Zika virus polyprotein: An in silico approach, *Asian Pac. J. Trop. Med.* 9 (2016) 844–850.
- [39] P.C. Loewen, Y. Alsaadi, D. Fernando, A. Kumar, Genome sequence of a tigecycline-resistant clinical isolate of *Acinetobacter baumannii* strain AB031 obtained from a bloodstream infection, *Genome Announc.* 2 (2014).
- [40] K.D. Pruitt, T. Tatusova, D.R. Maglott, NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins, *Nucleic Acids Res.* 33 (2005) D501–D504.
- [41] S. Ahmad, S. Raza, R. Uddin, S.S. Azam, Comparative subtractive proteomics based ranking for antibiotic targets against the dirtiest superbug: *Acinetobacter baumannii*, *J. Mol. Graph. Model.* 82 (2018) 74–92.
- [42] A. Naz, F.M. Awan, A. Obaid, S.A. Muhammad, R.Z. Paracha, J. Ahmad, A. Ali, Identification of putative vaccine candidates against *Helicobacter pylori*, *Helicobacter*. 3 (2015) 1.
- [43] W. Li, A. Godzik, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences, *Bioinformatics*. 22 (2006) 1658–1659.

- [44] D. Barh, N. Barve, K. Gupta, S. Chandra, N. Jain, S. Tiwari, N. Leon-Sicairos, A. Canizalez-Roman, A.R. dos Santos, S.S. Hassan, others, Exoproteome and secretome derived broad spectrum novel drug and vaccine candidates in *Vibrio cholerae* targeted by Piper betel derived compounds, PLoS One. 8 (2013) e52773.
- [45] R. Zhang, H.-Y. Ou, C.-T. Zhang, DEG: a database of essential genes, Nucleic Acids Res. 32 (2004) D271--D272.
- [46] S.S. Azam, A. Shamim, An insight into the exploration of druggable genome of *Streptococcus gordonii* for the identification of novel therapeutic candidates, Genomics. 104 (2014) 203–214.
- [47] S. Ahmad, S.S. Azam, A novel approach of virulome based reverse vaccinology for exploring and validating peptide-based vaccine candidates against the most troublesome nosocomial pathogen: *Acinetobacter baumannii*, J. Mol. Graph. Model. 83 (2018) 1–11.
- [48] N.Y. Yu, J.R. Wagner, M.R. Laird, G. Melli, S. Rey, R. Lo, P. Dao, S.C. Sahinalp, M. Ester, L.J. Foster, others, PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes, Bioinformatics. 26 (2010) 1608–1615.
- [49] C.-S. Yu, Y.-C. Chen, C.-H. Lu, J.-K. Hwang, Prediction of protein subcellular localization, Proteins Struct. Funct. Bioinforma. 64 (2006) 643–651.
- [50] C.-S. Yu, C.-W. Cheng, W.-C. Su, K.-C. Chang, S.-W. Huang, J.-K. Hwang, C.-H. Lu, CELLO2GO: a web server for protein subcellular localization prediction with functional gene ontology annotation, PLoS One. 9 (2014) e99368.
- [51] L. Chen, J. Yang, J. Yu, Z. Yao, L. Sun, Y. Shen, Q. Jin, VFDB: a reference database for bacterial virulence factors, Nucleic Acids Res. 33 (2005) D325--D328.
- [52] M. Kechagia, D. Basoulis, S. Konstantopoulou, D. Dimitriadi, K. Gyftopoulou, N. Skarmoutsou, E.M. Fakiri, Health benefits of probiotics: a review, ISRN Nutr. 2013 (2013).
- [53] A. Wadood, A. Jamal, M. Riaz, A. Khan, R. Uddin, M. Jelani, S.S. Azam, Subtractive genome analysis for in silico identification and characterization of novel drug targets in *Streptococcus pneumonia* strain JJA, Microb. Pathog. 115 (2018) 194–198.
- [54] Y. He, Z. Xiang, H.L.T. Mobley, Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development, Biomed Res. Int. 2010 (2010).
- [55] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein identification and analysis tools on the ExPASy server, Springer, 2005.
- [56] A. Krogh, B. Larsson, G. Von Heijne, E.L.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305 (2001) 567–580.
- [57] G.E. Tusnady, I. Simon, The HMMTOP transmembrane topology prediction server, Bioinformatics. 17 (2001) 849–850.

- [58] I.A. Doytchinova, D.R. Flower, VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines, *BMC Bioinformatics*. 8 (2007) 4.
- [59] Y. EL-Manzalawy, D. Dobbs, V. Honavar, Predicting linear B-cell epitopes using string kernels, *J. Mol. Recognit. An Interdiscip. J.* 21 (2008) 243–255.
- [60] H. Singh, G.P.S. Raghava, ProPred1: prediction of promiscuous MHC Class-I binding sites, *Bioinformatics*. 19 (2003) 1009–1014.
- [61] H. Singh, G.P.S. Raghava, ProPred: prediction of HLA-DR binding sites, *Bioinformatics*. 17 (2001) 1236–1237.
- [62] A. Garg, D. Gupta, VirulentPred: a SVM based prediction method for virulent proteins in bacterial pathogens, *BMC Bioinformatics*. 9 (2008) 62.
- [63] I. Dimitrov, I. Bangov, D.R. Flower, I. Doytchinova, AllerTOP v. 2—a server for in silico prediction of allergens, *J. Mol. Model.* 20 (2014) 2278.
- [64] I. Mayrose, O. Penn, E. Erez, N.D. Rubinstein, T. Shlomi, N.T. Freund, E.M. Bublil, E. Rupp, R. Sharan, J.M. Gershoni, others, Pepitope: epitope mapping from affinity-selected peptides, *Bioinformatics*. 23 (2007) 3244–3246.
- [65] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis, *Nat. Protoc.* 10 (2015) 845–858.
- [66] Y. Zhang, I-TASSER server for protein 3D structure prediction, *BMC Bioinformatics*. 9 (2008) 40.
- [67] U. Pieper, N. Eswar, F.P. Davis, H. Braberg, M.S. Madhusudhan, A. Rossi, M. Marti-Renom, R. Karchin, B.M. Webb, D. Eramian, others, MODBASE: a database of annotated comparative protein structure models and associated resources, *Nucleic Acids Res.* 34 (2006) D291–D295.
- [68] M. Källberg, H. Wang, S. Wang, J. Peng, Z. Wang, H. Lu, J. Xu, Template-based protein structure modeling using the RaptorX web server, *Nat. Protoc.* 7 (2012) 1511.
- [69] T. Schwede, J. Kopp, N. Guex, M.C. Peitsch, SWISS-MODEL: an automated protein homology-modeling server, *Nucleic Acids Res.* 31 (2003) 3381–3385.
- [70] S. Ahmad, S. Raza, R. Uddin, S.S. Azam, Binding mode analysis, dynamic simulation and binding free energy calculations of the MurF ligase from *Acinetobacter baumannii*, *J. Mol. Graph. Model.* (2017).
- [71] M.I. Rashid, A. Naz, A. Ali, S. Andleeb, Prediction of vaccine candidates against *Pseudomonas aeruginosa*: An integrated genomics and proteomics approach, *Genomics*. 109 (2017) 274–283.
- [72] D. Szklarczyk, A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K.P. Tsafou, others, STRING v10: protein–protein interaction networks, integrated over the tree of life, *Nucleic Acids Res.* 43 (2014) D447–D452.
- [73] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server,

- Bioinformatics. 16 (2000) 404–405.
- [74] T. Ishida, K. Kinoshita, PrDOS: prediction of disordered protein regions from amino acid sequence, *Nucleic Acids Res.* 35 (2007) W460--W464.
- [75] L. Heo, H. Park, C. Seok, GalaxyRefine: protein structure refinement driven by side-chain repacking, *Nucleic Acids Res.* 41 (2013) W384--W388.
- [76] D.B. Craig, A.A. Dombkowski, Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins, *BMC Bioinformatics.* 14 (2013) 346.
- [77] R.K. Pandey, T.K. Bhatt, V.K. Prajapati, Novel immunoinformatics approaches to design multi-epitope subunit vaccine for malaria by investigating anopheles salivary protein, *Sci. Rep.* 8 (2018) 1125.
- [78] U. Ohto, N. Yamakawa, S. Akashi-Takamura, K. Miyake, T. Shimizu, Structural analyses of human Toll-like receptor 4 polymorphisms D299G and T399I, *J. Biol. Chem.* 287 (2012) 40611–40617.
- [79] D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, H.J. Wolfson, PatchDock and SymmDock: servers for rigid and symmetric docking, *Nucleic Acids Res.* 33 (2005) W363--W367.
- [80] E. Mashiach, D. Schneidman-Duhovny, N. Andrusier, R. Nussinov, H.J. Wolfson, FireDock: a web server for fast interaction refinement in molecular docking, *Nucleic Acids Res.* 36 (2008) W229--W232.
- [81] D.A. Pearlman, D.A. Case, J.W. Caldwell, W.S. Ross, T.E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, P. Kollman, AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules, *Comput. Phys. Commun.* 91 (1995) 1–41.
- [82] J. Wang, R.M. Wolf, J.W. Caldwell, P.A. Kollman, D.A. Case, Development and testing of a general amber force field, *J. Comput. Chem.* 25 (2004) 1157–1174.
- [83] G.A. Özpınar, W. Peukert, T. Clark, An improved generalized AMBER force field (GAFF) for urea, *J. Mol. Model.* 16 (2010) 1427–1440.
- [84] M.G. Paterlini, D.M. Ferguson, Constant temperature simulations using the Langevin equation with velocity Verlet integration, *Chem. Phys.* 236 (1998) 243–252.
- [85] V. Kräutler, W.F. Van Gunsteren, P.H. Hünenberger, A fast SHAKE algorithm to solve distance constraint equations for small molecules in molecular dynamics simulations, *J. Comput. Chem.* 22 (2001) 501–508.
- [86] M.J. Uline, D.S. Corti, Molecular dynamics at constant pressure: allowing the system to control volume fluctuations via a “shell” particle, *Entropy.* 15 (2013) 3941–3969.
- [87] S. Ahmad, S. Raza, S.W. Abbasi, S.S. Azam, Identification of natural inhibitors against *Acinetobacter baumannii* D-alanine-D-alanine ligase enzyme: A multi-spectrum in silico approach, *J. Mol. Liq.* 262 (2018) 460–475.

- [88] H.J.C. Berendsen, J.P.M. van Postma, W.F. van Gunsteren, A. DiNola, J.R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.* 81 (1984) 3684–3690.
- [89] F.U. Haq, A. Abro, S. Raza, K.R. Liedl, S.S. Azam, Molecular dynamics simulation studies of novel β -lactamase inhibitor, *J. Mol. Graph. Model.* 74 (2017) 143–152.
- [90] W. Humphrey, A. Dalke, K. Schulten, VMD-Visual Molecular Dynamics *J Mol Graph* 14: 33–38, (1996).
- [91] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera—a visualization system for exploratory research and analysis, *J. Comput. Chem.* 25 (2004) 1605–1612.
- [92] B.R. Miller III, T.D. McGee Jr, J.M. Swails, N. Homeyer, H. Gohlke, A.E. Roitberg, MMPBSA. py: an efficient program for end-state free energy calculations, *J. Chem. Theory Comput.* 8 (2012) 3314–3321.
- [93] J. Weiser, P.S. Shenkin, W.C. Still, Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO), *J. Comput. Chem.* 20 (1999) 217–230.
- [94] K.U. Jansen, C. Knirsch, A.S. Anderson, The role of vaccines in preventing bacterial antimicrobial resistance, *Nat. Med.* 24 (2018) 10.
- [95] W. Chen, Current advances and challenges in the development of *Acinetobacter* vaccines, *Hum. Vaccin. Immunother.* 11 (2015) 2495–2500.
- [96] R. Singh, N. Garg, G. Shukla, N. Capalash, P. Sharma, Immunoprotective efficacy of *Acinetobacter baumannii* outer membrane protein, FilF, predicted in silico as a potential vaccine candidate, *Front. Microbiol.* 7 (2016) 158.
- [97] D.G. Moriel, S.A. Beatson, D.J. Wurpel, J. Lipman, G.R. Nimmo, D.L. Paterson, M.A. Schembri, Identification of novel vaccine candidates against multidrug-resistant *Acinetobacter baumannii*, *PLoS One.* 8 (2013) e77631.
- [98] M.-H. Chiang, W.-C. Sung, S.-P. Lien, Y.-Z. Chen, A.F. Lo, J.-H. Huang, S.-C. Kuo, P. Chong, Identification of novel vaccine candidates against *Acinetobacter baumannii* using reverse vaccinology, *Hum. Vaccin. Immunother.* 11 (2015) 1065–1073.
- [99] Z. Ni, Y. Chen, E. Ong, Y. He, Antibiotic resistance determinant-focused *Acinetobacter baumannii* vaccine designed using reverse vaccinology, *Int. J. Mol. Sci.* 18 (2017) 458.
- [100] R. Donev, *Peptide and Protein Vaccines*, Academic Press, 2015.
- [101] T.A. Ahmad, D.M. Tawfik, S.A. Sheweita, M. Haroun, L.H. El-Sayed, Development of immunization trials against *Acinetobacter baumannii*, *Trials Vaccinol.* 5 (2016) 53–60.
- [102] S. Venkatesha, S. Dudics, B. Acharya, K. Moudgil, Cytokine-modulating strategies and newer cytokine targets for arthritis therapy, *Int. J. Mol. Sci.* 16 (2015) 887–906.
- [103] R. Xie, X.D. Zhang, Q. Zhao, B. Peng, J. Zheng, Analysis of global prevalence of antibiotic resistance in *Acinetobacter baumannii* infections disclosed a faster increase in OECD countries, *Emerg. Microbes Infect.* 7 (2018) 31.

- [104] G. Sanober, S. Ahmad, S.S. Azam, Identification of plausible drug targets by investigating the druggable genome of MDR *Staphylococcus epidermidis*, *Gene Reports*. 7 (2017) 147–153.
- [105] G. Grandi, Bacterial surface proteins and vaccines, *F1000 Biol. Rep.* 2 (2010).
- [106] P.F. Sparling, Bacterial virulence and pathogenesis: an overview, *Rev. Infect. Dis.* 5 (1983) S637–S646.
- [107] M.R. Kehry, F.W. Dahlquist, The methyl-accepting chemotaxis proteins of *Escherichia coli*. Identification of the multiple methylation sites on methyl-accepting chemotaxis protein I., *J. Biol. Chem.* 257 (1982) 10378–10386.
- [108] P. Derr, E. Boder, M. Goulian, Changing the specificity of a bacterial chemoreceptor, *J. Mol. Biol.* 355 (2006) 923–932.
- [109] R. Albrecht, M. Schütz, P. Oberhettinger, M. Faulstich, I. Bermejo, T. Rudel, K. Diederichs, K. Zeth, Structure of BamA, an essential factor in outer membrane protein biogenesis, *Acta Crystallogr. Sect. D Biol. Crystallogr.* 70 (2014) 1779–1789.
- [110] H. Nikaido, Y. Takatsuka, Mechanisms of RND multidrug efflux pumps, *Biochim. Biophys. Acta (BBA)-Proteins Proteomics*. 1794 (2009) 769–781.
- [111] N. Noinaj, M. Guillier, T.J. Barnard, S.K. Buchanan, TonB-dependent transporters: regulation, structure, and function, *Annu. Rev. Microbiol.* 64 (2010) 43–60.
- [112] S. Koskiniemi, J.G. Lamoureux, K.C. Nikolakakis, C. t’Kint de Roodenbeke, M.D. Kaplan, D.A. Low, C.S. Hayes, Rhs proteins from diverse bacteria mediate intercellular competition, *Proc. Natl. Acad. Sci.* (2013) 201300627.
- [113] S.-P. Nuccio, A.J. Bäumlér, Evolution of the chaperone/usher assembly pathway: fimbrial classification goes Greek, *Microbiol. Mol. Biol. Rev.* 71 (2007) 551–575.
- [114] R. Nagpal, A. Kumar, M. Kumar, P. V Behare, S. Jain, H. Yadav, Probiotics, their health benefits and applications for developing healthier foods: a review, *FEMS Microbiol. Lett.* 334 (2012) 1–15.
- [115] J. Kyte, R.F. Doolittle, A simple method for displaying the hydropathic character of a protein, *J. Mol. Biol.* 157 (1982) 105–132.
- [116] K. Guruprasad, B.V.B. Reddy, M.W. Pandit, Correlation between stability of a protein and its dipeptide composition: a novel approach for predicting in vivo stability of a protein from its primary sequence, *Protein Eng. Des. Sel.* 4 (1990) 155–161.
- [117] S.C. Gill, P.H. Von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem.* 182 (1989) 319–326.
- [118] E.H. Chung, Vaccine allergies, *Clin. Exp. Vaccine Res.* 3 (2014) 50–57.
- [119] L. Han, J. Zheng, Y. Wang, X. Yang, Y. Liu, C. Sun, B. Cao, H. Zhou, D. Ni, J. Lou, others, Structure of the BAM complex and its implications for biogenesis of outer-membrane proteins, *Nat. Struct. Mol. Biol.* 23 (2016) 192.

- [120] G.R. Soltes, J. Schwalm, D.P. Ricci, T.J. Silhavy, The activity of *Escherichia coli* chaperone SurA is regulated by conformational changes involving a parvulin domain, *J. Bacteriol.* (2016) JB--00889.
- [121] K. Koide, K. Ito, Y. Akiyama, Substrate recognition and binding by RseP, an *Escherichia coli* intramembrane protease, *J. Biol. Chem.* (2008).
- [122] L. Zhu, J. Cheng, B. Luo, S. Feng, J. Lin, S. Wang, J.E. Cronan, H. Wang, Functions of the *Clostridium acetobutylicum* FabF and FabZ proteins in unsaturated fatty acid biosynthesis, *BMC Microbiol.* 9 (2009) 119.
- [123] L. Buetow, T.K. Smith, A. Dawson, S. Fyffe, W.N. Hunter, Structure and reactivity of LpxD, the N-acyltransferase of lipid A biosynthesis, *Proc. Natl. Acad. Sci.* 104 (2007) 4321–4326.
- [124] L.M. Parsons, F. Lin, J. Orban, Peptidoglycan recognition by Pal, an outer membrane lipoprotein, *Biochemistry.* 45 (2006) 2122–2128.
- [125] M.M. Babu, The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease, *Biochem. Soc. Trans.* 44 (2016) 1185–1200.
- [126] N. Khatoon, R.K. Pandey, V.K. Prajapati, Exploring *Leishmania* secretory proteins to design B and T cell multi-epitope subunit vaccine using immunoinformatics approach, *Sci. Rep.* 7 (2017) 1–12. doi:10.1038/s41598-017-08842-w.
- [127] S. Abbasi, S. Raza, S.S. Azam, K.R. Liedl, J.E. Fuchs, Interaction mechanisms of a melatonergic inhibitor in the melatonin synthesis pathway, *J. Mol. Liq.* 221 (2016) 507–517.
- [128] L.G. Ferreira, R.N. dos Santos, G. Oliva, A.D. Andricopulo, Molecular docking and structure-based drug design strategies, *Molecules.* 20 (2015) 13384–13421.
- [129] S. Genheden, U. Ryde, The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities, *Expert Opin. Drug Discov.* 10 (2015) 449–461.
- [130] G. Wang, W. Zhu, Molecular docking for drug discovery and development: A widely used approach but far from perfect, (2016).
- [131] C. Wang, D. Greene, L. Xiao, R. Qi, R. Luo, Recent developments and applications of the MMPBSA method, *Front. Mol. Biosci.* 4 (2018) 87.
- [132] P.A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, others, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, *Acc. Chem. Res.* 33 (2000) 889–897.
- [133] R. Singh, N. Capalash, P. Sharma, Immunoprotective potential of BamA, the outer membrane protein assembly factor, against MDR *Acinetobacter baumannii*, *Sci. Rep.* 7 (2017) 12411.

Graphical abstract