We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

5,300
Open access books available

130,000
International authors and editors

155M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Computational Identification of the Plausible Molecular Vaccine Candidates of Multidrug-Resistant *Salmonella enterica*

Rohit Mishra, Yong Chiang Tan, Amr Adel Ahmed Abd El-Aal and Chandrajit Lahiri

Abstract

Salmonella enterica serovars are responsible for the life-threatening, fatal, invasive diseases that are common in children and young adults. According to the most recent estimates, globally, there are approximately 11–20 million cases of morbidity and between 128,000 and 161,000 mortality per year. The high incidence rates of diseases like typhoid, caused by the serovars Typhi and Paratyphi, and gastroenteritis, caused by the non-typhoidal Salmonellae, have become worse, with the ever-increasing pathogenic strains being resistant to fluoroquinolones or almost even the third generation cephalosporins, such as ciprofloxacin and ceftriaxone. With vaccination still being one of the chosen methods of eradicating this disease, identification of candidate proteins, to be utilized for effective molecular vaccines, has probably remained a challenging issue. In our study here, we portray the usage of computational tools to analyze and predict potential vaccine candidate(s) for the multi-drug resistant serovars of *S. enterica*.

Keywords: typhoid, *Salmonella* Typhi, multidrug resistance, computational identification, vaccine candidates

1. Introduction

With a current worldwide prevalence of around twenty-seven million cases [1, 2] and hundreds of thousands of deaths every year [2, 3], salmonellosis remains the second most common food/water-borne illness. It constitutes a disease caused due to the systemic infection of human and animal hosts by the facultatively anaerobic, Gram-negative rod-shaped bacterial species of *Salmonella enterica* from the family *Enterobacteriaceae*. Clinically, several serologic variants (serovars) of *S. enterica* exist, which differ with respect to their different antigenic variation in lipopolysaccharide and flagella [4, 5]. They include Typhi and Paratyphi A, besides the non-typhoidal serotypes like Typhimurium and Enteriditis [4]. Among these, the enteric fever termed typhoid, caused by *S.* Typhi and Paratyphi, is typically a more severe illness than those caused by other non-typhoidal serovars [5].

Being contagious in nature, salmonellosis, like typhoid, can spread through feces, water and the hands of those caring for the sick while, for non-typhoidal serovars, through the consumption of raw or undercooked contaminated food of animal origin such as meat, poultry, eggs and milk by humans [1, 6, 7]. Salmonellosis begins with ingestion of a dose for the bacterium enough to broach the first-line host defenses and colonize the gastrointestinal tract. The onset symptoms for typhoid are usually accompanied with fever, headache, myalgia, anorexia and sometimes diarrhea or constipation [6, 7], moving onto remittent fever, with a stepwise increment in the daily peak temperature, reaching 40°C by the end of the first week [6]. Slow recovery after 3–4 weeks is the normal case, though, for untreated patients with complications, major fatalities occur due to intestinal hemorrhage or perforation [6, 7].

Drugs available for the treatments are mostly ineffective due to the resistance developed with the emergence of multidrug-resistance (MDR) *Salmonella* strains [8]. These new strains are ineffective to the older generations of drugs including ampicillin, chloramphenicol, ciprofloxacin, trimethoprim as well as co-trimoxazole and their derivatives, thereby necessitating the newer classes of cephalosporins and quinolone derivatives to be greatly explored to combat such MDR threats [1, 8]. Moreover, dating as early as the 1890s, whole-cell vaccines with parenteral administration of killed suspensions of *S.* Typhi [9] has several problems having: a) high-reactivity with 20–25% fever and 40–50% local reactions, b) moderate efficacy with protection rates of 51–88% insufficient to halt disease transmission in endemic area and c) logistical and safety problems having the need for needles and two doses. Approaches with recent vaccines, like, single-dose Typhim Vi® containing purified Vi capsular polysaccharide, or, the live attenuated vaccine *S.* Typhi Ty21a (Vivotif®), confer around 50% protection in adults, and very poor immunogenicity among young children, without any license for under two years old, besides being considered to be expensive for low-middle income areas [10, 11]. Thus, the urgency, for new and specific vaccines and/or drugs to combat the disease, is evident and indeed, proteins of the pathogen-specific biochemical and biosynthetic pathways, involved in the virulence of *S.* Typhi, has already begun to be targeted with a view to developing novel vaccines/drugs.

While the two afore-mentioned vaccines are for *S.* Typhi, those for other serovars including Paratyphi, Typhimurium and Enteritidis were largely unavailable until some few years back [11]. Of late, efforts to confer protective immunity for serovars of Typhimurium has been reported with the *lppA* and *lppB* Braun lipoprotein genes with and without the *msbB* gene, encoding an acetyltransferase enzyme required for modification of the lipid A of lipopolysaccharide [12]. Other candidate genes proposed for effective vaccines for different serovars include *rpoS*, *phoPQ, ssaV*, *htrA* [13], besides the proteins of SseBI, OmpACDFL and SopB being used as antigens in other vaccination studies [14]. Such recombinant attenuated *Salmonella* vaccines (RASV) are considered to be same or more effective than the whole wild-type strains [15]. RASV can persistently colonize internal lymphoid tissues to produce recombinant antigens having their maximum abilities to elicit mucosal and systemic antibody along with those of the cell mediated immune responses [15]. Thus, development of such recombinant vaccines is considered to be the cost-effective and most promising strategy against the pressing antibiotic resistance threats. In this regard, several strategies have been adopted in other drug resistant bacteria including reverse vaccinology through comparative genome analysis and *in vitro* proteomics [16, 17]. These become especially effective keeping in mind the new and emerging threats of multidrug resistance strains of *Salmonella*. Such strains might possibly arise form immune selection leading to antigen

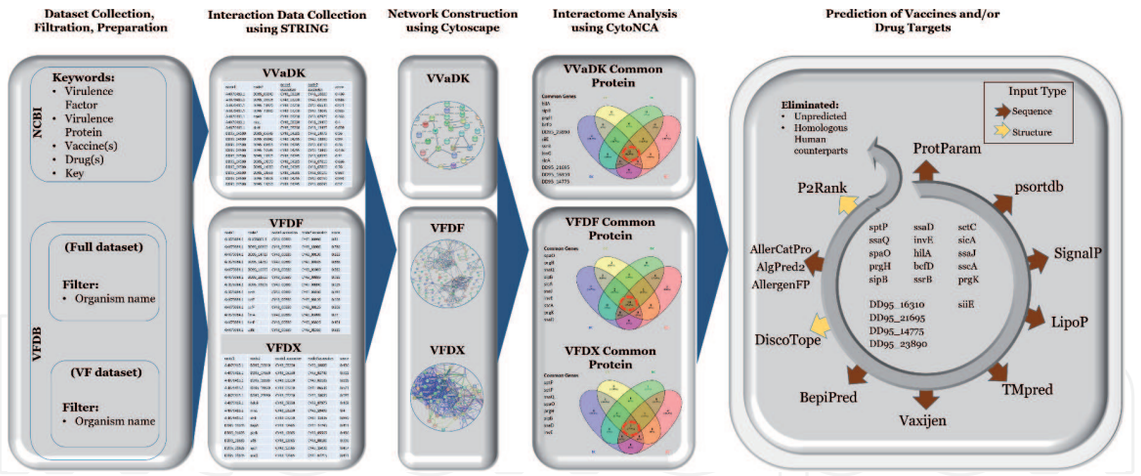


Figure 1.
Graphical summary of the methods adopted in vaccine candidates and druggability prediction. This comprises a network-based approach to identify the key players in Salmonella virulent proteome coupled with downstream predictions of vaccine candidates and druggable pockets among the top rankers.

sequence variability followed by a down-regulation of the target antigens, thereby conferring poor "cross-protective efficacy" as reported for MDR *Acinetobacter baumannii* [18]. Therefore, identification of new and effective vaccine candidates is, probably, the current need of the hour.

With an availability of different virulent proteins, reported from different experimental verification and predictive databases, selection of the most plausible vaccine candidates can be confusing. To cater to the need of simplifying this complex problem of selection, graph theoretical analysis of the interacting networks of such virulent proteins, involved in the disease scenario, might be poised to be quite useful. Such virulent protein interaction networks (PIN) can be utilized to find out the most central or sought-after proteins for such cases [19]. Ideally, the centrality of any biological networks is efficiently analyzed through global parameters like betweenness, closeness, degree and eigen-vector centralities, referred to as the BC, CC, DC and EC, respectively [19–21]. Among them, BC has been regarded to be efficient enough to impart central character of a network above CC and DC for long until EC gained some prominence and can be quite effective as reported through recent studies [22–25].

In this study, we proposed the vaccine candidates for *Salmonella* serovars (**Figure 1**) as explained in the next section. Essentially, we utilized the four different centrality measures for analyzing three different virulent PINs denoted as VVaDK, VFDF and VFDX. Among the top 20 rankers of each of the different centralities, the unanimously present unique candidates were finally collected for further downstream analyses. These shortlisted candidate virulent proteins were rigorously analyzed through different bioinformatic tools to determine their antigenic and allergenic potential besides revealing the epitopes for efficient vaccines or molecular crevices for good drug targets.

2. Approach

2.1 Dataset collection

We have initiated our study with the proteins collected for *Salmonella enterica* serovar Typhimurium str. LT2 (NCBI txid: 99287) on the 19th of December 2020.

They were retrieved from two different sources namely, the National Center for Biotechnology Information (NCBI) and the Virulence Factor Database (VFDB) [26]. From NCBI, protein datasets were collected through literature search using various keywords such as Virulence, Virulence Factor, Virulence Protein, Drug(s), Vaccine(s) and Key. Some of these keywords, having essentially the same meaning, were used to get more hits and to avoid missing of any possible candidates thereby reducing the false-negative hits. Finally, all the candidates of the lists were merged, and duplicates were removed to yield 120 proteins to be considered for further analysis. They were termed as VVaDK for easy reference, where V stands for Virulence, Va represents Vaccine(s), D means Drug(s) and K denotes Key. Moreover, two types of candidates' lists were retrieved from VFDB. They comprised the Full dataset which covers all the proteins (261) related to unknown and predicted VFs of *S.* Typhimurium and were referred as VFDF. Additionally, 117 experimentally verified candidates were retrieved for *S.* Typhimurium and termed as VFDX.

All the afore-mentioned proteins for the different categories of VVaDK, VFDF and VFDX were fed as queries to the biological meta-database of protein interaction, STRING version 11.0 [27] to retrieve all the possible interactions of a particular protein [date and time of access: Dec 22, 2020, from 17 hours IST onwards]. Detailed protein links file under the accession number 90371 in STRING v11 was used to collect all the interactions of the whole genome proteins of *S.* Typhimurium. In each case, a database dictated default medium confidence value of 0.4, for the combined scores from different parameters of interaction, was used. Accordingly, the total number of protein interactions obtained were 138, 3501 and 2464 for VVaDK, VFDF and VFDX listed candidates, respectively.

2.2 Interactome construction

The protein interaction data for all individual sets for VVaDK, VFDF and VFDX, having medium confidence values, were imported into Cytoscape version 3.8.2 [28] to integrate and build the respective interactomes of protein interactions. Care was taken to remove duplicate and bidirectional interactions from each dataset. In essence, such interactome of proteins or the protein interaction network (PIN) has been constructed as an undirected graph, G = (V, E), consisting of E edges and a finite set of V vertices (or nodes) where, edge, e = (u, v), is connected to two vertices u and v. Each vertex/node in our PIN represents a protein. The number of connections/interactions/associations/links, a protein has with other proteins, reflects its degree, d [29].

2.3 Network analysis

All the constructed 3 PINs have been viewed by Cytoscape v 3.8.2 in the form of interactomes of aforementioned interconnected proteins. They were subsequently analyzed through the integrated java plugin CytoNCA version 2.1.6 [30] to compute values for BC, CC, DC and EC as the four different global network centrality parameters. The different parametric combined scores from STRING were considered as edge weights for computing the CytoNCA scores of the 4 centrality parameters. Upon sorting these 4 measures from largest to smallest, top 20 proteins for each of the categories of centrality were picked to create Venn diagrams using Venny 2.0 [31] for finding the common proteins from each of the measures. This resulted in 12, 10, 7 proteins from VVaDK, VFDF and VFDX, respectively. Among these 29 candidates, 9 duplicates were removed to yield a total of 20 proteins. Through a

BLASTp alignment, these Typhimurium proteins were unanimously found in the serovars of Typhi and Paratyphi, and thus, considered for further analyses.

2.4 Vaccine and/or drug candidature prediction

2.4.1 Basic analysis

The 20 shortlisted protein candidates from VVaDK, VFDF and VFDX PIN analysis were subjected to further analyses for predicting the plausible vaccine and/or drug candidates. All such proteins were explored for their molecular weight calculation, cellular localization, signal peptide prediction followed by antigenicity prediction. ProtParam was used to find the molecular weight and number of amino acids [32] and cellular localization was analyzed by PSORTb v3.0.2 [33]. Location of signal peptides was predicted using the server called SignalP 4.1 [34]. Lipoprotein signal peptides were predicted using the LipoP 1.0 [35]. Finally, Vaxijen was used to predict the possible antigenicity of the proteins [36].

2.4.2 Mapping of available 3D structures in PDB

For the top ranked proteins, the respective crystallized protein 3D structures available in Protein Data Bank (PDB) were retrieved (**Table 1**). The seleno-methionine in PDB structures were changed back into methionine using Dock Prep in Chimera [37].

Protein	**Structural Information**			
	PDB ID	**Chain ID**	**Structure Coverage**	**Resolution**
SptP	1G4W	R	161–543	2.20
	1JYO	E	35–139	1.90
SpaO	4YX1	A	232–297	1.35
	4YX7	A	145–213	2.00
	4YX7	B	232–297	2.00
PrgH	4G1I	A	170–392	1.85
	4G2S	A	11–119	1.86
	6UOT	A	1–392	3.30
SipB	3TUL	A	81–237	2.79
SsrB	2JPC	A	133–193 (First 19 N-terminal amino acids missing)	NMR
SctC	4G08	A	22–178	1.80
	6PEE	A	1–562	3.42
PrgK	6UOT	Y	1–252	3.30
	4OYC	A	96–200	2.60
SiiE	2YN5	A	5078–5365	1.85

SsaQ, SsaD, InvE, HilA, BcfD, SicA, SsaJ, SscA, DD95_23890, DD95_21695, DD95_16310, and DD95_14775 have no structures available in PDB.

Table 1.
PDB structure availability among top rankers.

2.4.3 B-cell epitope prediction

Unlike viral pathogens, most bacterial pathogens are not intracellular parasites, especially *Salmonella*. Thus, the humoral immune response, which involves B cells and antibodies, will be of great focus in this study. Herein, BepiPred v2.0 and DiscoTope v2.0 were utilized in predicting linear and discontinuous B-cell epitopes, respectively [38, 39]. For BepiPred, the default threshold score of 0.5 was applied for epitope recognition. For DiscoTope, the propensity score radius was 22 Angstrom, upper half sphere radius was 14 Angstrom, window size was 1, and alpha was 0.115. An in-house script (DiscoTope2ChimeraAttr) has been utilized to convert DiscoTope result into Chimera attributes for visualization in 3D, with a default threshold DiscoTope score of −3.7 [40]. These analyses were done to pinpoint the specific immunogenic regions within the full-length proteins. Thus, the immunogenically insignificant regions can be trimmed out, resulting in shorter peptides which can confer higher specificity and ease the peptide synthesis process.

2.4.4 Allergenicity prediction

The ability of proposed immunogen to potentially evoke allergic reactions can usually fail clinical trials due to the severe adverse effects arising upon vaccination. Herein, we utilized AllerCatPro, AlgPred2, and AllergenFP v1.0 to predict possible allergic reactions raised by the query proteins, which were the top rankers in this case. For AlgPred2, the hybrid algorithm was selected and the default threshold value of 0.3 was selected. AllerCatPro predicts allergenicity by comparing the protein structural and sequential information to known allergens [41]. Besides, the hybrid algorithm of AlgPred2.0 utilizes the random forest, BLAST, and MERCI algorithms to predict the allergenicity of the query proteins [42]. Moreover, the allergenicity prediction of AllergenFP v1.0 utilizes an alignment-independent fingerprint-based approach [43].

2.4.5 Druggable pocket prediction

P2Rank was being utilized to predict the presence of druggable pockets in the available 3D structures of proteins [44]. P2Rank utilizes a template-independent machine learning algorithm in predicting potential ligand-binding sites on the query proteins. Herein, the topmost ranked predicted pockets were selected for further analyses. Thus, besides being utilized in vaccination, the potential drug-gability of the top rankers can be discovered.

2.4.6 Detecting human counterparts

Peptide vaccines that contain regions of high sequence similarity to human proteome counterparts can lead to ineffective vaccination due to recognition as "self" by the immune system, which can result in low antigenicity or adverse effects that arise from potential self-reactivity. Thus, the top rankers were screened for human counterparts via sequence alignment approach using BLASTp against non-redundant proteins (nr) database with *Homo sapiens* as the specified organism [45].

3. Interactome analyses of three virulent PINs

Three different interactomes of virulent proteins of *Salmonella* were built using the method described above. The first of them comprised those available through

literature search using different keywords comprising Virulence, Virulence Factor, Virulence Protein, Drug(s), Vaccine(s) and Key. This was named as VVaDK. The other two PINs were made of the full and experimentally verified datasets of virulent proteins from *Salmonella*, listed in VFDB and were named as VFDF and VFDX, respectively. The four centrality measures were applied for analyzing each of these PINs and twenty top rankers from each of the measures were initially segregated. Among them, the proteins present unanimously for all the measures were noted as 12, 10 and 7 for VVaDK, VFDF and VFDX, respectively, and a removal of duplicates from them finally yielded 20 candidates for further downstream analysis.

Our unique way of streamlining the candidates is based upon the following facts. Under pathological conditions, the virulent proteins are expected to be working in unison to render the final disease phenotype. Thus, their connectivity could be perceived in terms of the said PINs. Among these proteins, some can be master regulators and connecting to others more frequently thereby having higher order of connectivity. This renders them degree centrality (DC). Alternatively, there could be different types of such regulators for carrying out different sub-functions of the main disease phenotype and they form the bridge between the other proteins. These could impart the betweenness centrality (BC) of such proteins. Moreover, among such conglomerate of different proteins, certain numbers could connect to others faster to sequentially carry out their function, leading to a concept of closeness for them and having higher closeness centrality (CC). Furthermore, certain proteins could be more important to render the final disease phenotype and they are only connected to other important proteins to carry out their functions. These could bring out their character of eigen vector centrality (EC). Finally, from the top-ranking proteins of all these centrality measures, those, appearing unanimously, are expected to play a major role in virulence and could be segregated to scan for further analysis. These are 20 unique virulent proteins, mostly belonging to the *Salmonella* Pathogenicity Islands (SPI) from three different PIN analyses and reflected in **Figure 1** and **Table 2**. These are discussed in the next section.

4. Features of the twenty virulent proteins

All the virulent proteins from different serovars of *Salmonella* are discussed here, with their characteristic features along with a note on their existing vaccine potential.

SptP is one of the most important SPI-1 Type III Secretion System (T3SS) effector proteins which facilitates the bacterial translocation and survival into the host non-phagocytic cells by inhibition of the extracellular-regulated kinase (ERK) mitogen-activated protein kinase (MAP) pathways [46]. It requires SicP as a chaperone protein for its secretion and stabilization [46]. Moreover, SptP is directly responsible for the reversal of the actin cytoskeletal changes in the host cells by acting as a GTPase-activating protein (GAP) for Rac-1 and Cdc42. In fact, the efficacy of *sptP* deletion mutation of *S.* Enteriditis has been shown to be effective for live attenuated vaccine (LAV) in chickens [47].

SsaQ is a member of FliN/YscQ/Spa33/HrcQ family of both T3SS and flagellum proteins [48]. The gene *ssaQ* is encoded in the *ssaMVNOPQ* operon within the SPI-2 and transcribes to two products namely, SsaQL of 322 residues and SsaQS of 106 residues. SsaQS acts as a chaperone-like protein for SsaQL and optimize its function. SsaQ interact with SsaK and SsaN to form the C-ring complex, which have a crucial role in secretion by acting as a cytoplasmic sorting platform at the base of T3SS as well as rotation and direction switching of the flagella [49].

Protein	ProtParam		PsortB	SignalP & LipoP	TMpred			Vaxijen	
	# amino acids	Molecular Weight	Localization	Prediction: CSPosition	Prediction score	Status (Predicted)	# of TM Helices	Position	Score (Orientation)
SptP	543	60047.68	E	—	0.5192	A	1	477–496	570 (o-i)
SsaQ	322	36009.35	C	—	0.3857	NA	1	186–209	611 (i-o)
SpaO	303	33793.74	C	—	0.5073	A	1	62–86	600 (o-i)
PrgH	392	44459.53	C	—	0.5122	A	1	142–163	2551 (o-i)
SipB	593	62450.71	E	—	0.4855	A	2	320–343 409–428	2293 (i-o) 2875 (o-i)
SsaD	403	44849.66	CM	—	0.4319	A	1	119–135	2978 (o-i)
InvE	372	42421.49	CM	—	0.3335	N	2	255–273 317–337	634 (o-i) 612 (i-o)
HilA	553	63040.96	C	—	0.3985	N	1	340–361	523 (o-i)
BcfD	335	35928.51	U	SP(SPI): 21–22	0.6728	A	2	1–21 204–222	1342 (o-i) 582 (i-o)
SsrB	212	24354.49	C	—	0.4053	A	0	—	—
SctC	562	61765.81	OM	SP(SPI): 24–25	0.371	N	1	5–25	1725 (i-0)
SicA	165	19220.72	C	—	0.5954	A	0	—	—
SsaJ	249	28521.47	U	LIPO(SPII): 18–19	0.529	A	2	6–24 225–245	1371 (i-o) 2963 (o-i)
SscA	157	18134.88	C	—	0.3731	N	0	—	—
PrgK	252	28210.3	OM	LIPO(SPII): 17–18	0.5132	A	1	208–225	2812 (i-o)

Protein	ProtParam		PsortB	SignalP & LipoP	TMpred			Vaxijen	
	# amino acids	Molecular Weight	Localization	Prediction: CSPosition	Prediction score	Status (Predicted)	# of TM Helices	Position	Score (Orientation)
DD95_23890 (BigA)	1721	176526.46	U	—	0.8358	A	5	193–213 466–487 800–820 952–971 1141–1165	641 (i-o) 610 (o-i) 513 (i-o) 558 (o-i) 504 (i-o)
SiiE	5559	594451.38	U	—	0.76	A	3	39–57 4349–4367 5013–5036	692 (o-I) 512 (i-o) 1827 (o-i)
DD95_21695 (SspH2)	788	87222.97	E	—	0.382	N	0	—	—
DD95_16310 (TorS)	911	100273.65	CM	—	0.4384	A	3	10–29 332–350 622–641	2340 (i-o) 2738 (o-i) 1048 (i-o)
DD95_14775 (MarT_1)	147	16757.61	U	—	0.4706	A	1	127–144	1622 (o-i)

TM: Transmembrane. For Localization, E: Extracellular, C: Cytoplasmic, CM: Cytoplasmic Membrane, OM: Outer Membrane, U: Unknown. For TMpred status, A: Antigen, N: Non-antigen.

Table 2.
Basic screening of plausible vaccine candidates.

SpaO is a major invasion factor of *S. enterica spp.* and the core component of the sorting platform in *S.* Typhimurium. SpaO is comprised of 303 residues of two translated products with SpaOS (the shorter product) encompassing the last 101 amino acids of SpaOL (full length protein) [50]. It is a highly conserved element in T3SS that shares similarity with limited residues with flagellar C-ring substructure [51]. In fact, SpaO, along with H1a, has been suggested to be promising new vaccine candidates to prevent typhoid fever caused by *S.* Paratyphi A infection [52].

PrgH is a 55 kDa protein encoded within *prgHIJK* operon in the SPI-1. All the genes of *prg* operon are essential for the formation of T3SS needle complex (NC) and known to share sequence similarity with the flagellar protein, FliF [53]. PrgH inserts in the inner membrane by its hydrophobic domain where it forms the MS-ring of the flagellar basal body as well as provides the structural foundation required for *prgK* oligomerization for further assembly of the NC [53].

SicA is a wide acting chaperone protein (18 KDa) which aids in the secretion process of all T3SS proteins through the invasion of host cells. Accordingly, it is encoded upstream to the *Sip/SspABCD* operon in SPI-1. SipB and SipC proteins are responsible for the translocon formation in the host cell membrane to facilitate the injection of Type III effector proteins into the host cell to manipulate it [54]. Moreover, SicA is essential for the expression of the most virulence genes that encode T3SS effector proteins and is identified as a co-regulator with InvF for *SigDE* and *SptP* [55].

HilA is a member of the OmpR/ToxR regulator protein family and the central activator of SPI-1 genes, belonging to T3SS. The *hilA* gene is encoded within SPI-1 and is the key factor in SPI-1-T3SS regulation, starting from the expression of downstream genes *sicA* and *invF* to ultimate regulation of the effector genes *sipA* and *sipB* [56]. The upregulation of *hilA* results in the high expression of all genes encoded within the SPI-1 which are necessary for the invasion of epithelial cells. Moreover, the expression of *hilA* is controlled by many different activators and suppressors in response to specific environmental changes during invasion of the host cells, such as, temperature, bile, fatty acids, osmolarity, pH, oxygen concentrations and growth state [57]. Additionally, certain studies considered HilA as a promising drug target to inhibit the activity of T3SS without affecting the growth of *Salmonella* [58].

SiiE is the largest protein in *Salmonella* proteome, with the size of 595 kDa. It consists of 53 repetitive bacterial immunoglobulin domains, each containing several conserved residues [59]. The protein helps to contact the host cell membrane and positions the SPI T3SS, to initiate the translocation of effector proteins. A study states that *Salmonella* SiiE-mediated entry of enterocytes via the apical route requires transmembrane mucin MUC1 [60]. Moreover, it is shown that, *siiE* is required for the prevention of efficient humoral immune response against the pathogen and it induces the high tires of specific *Salmonella*-specific IgG [61].

PrgK is a component from the inner membrane of *Salmonella* SPI-1 T3SS basal body, in its N-terminus. It possess the canonical lipoproteins which acts as anchor for the hydrophilic proteins onto the surface of the bacterial cell membranes [62]. In addition, C-terminus of PrgK is found in the cytoplasm which confirms that the protein traverses the inner membrane. A study observed reduced fever in swine which were vaccinated with *prgK* gene attenuated *S.* Typhimurium in comparison with mock-vaccinated swine [63].

SscA is a chaperone protein of about 18 KDa size. It is an independent α-helical protein, that consists of eight α-helices and repeated large tetratricopeptide domain from 36 to 137 amino acids. SscA is a virulence factor which encodes the chaperonin of SseC and the translocon is involved during the adaptation and survival to

desiccation [64]. A huge effect of the gene expression level of *sscA,* has been noted on treatment of the samples with ciprofloxacin [65].

SsaJ is a core encoding component of the T3SS. It is required for SpvB, in-order to induce the actin depolymerization, especially inside the human macrophages. *Salmonella* depends on SsaJ effector protein as it prevents the interaction of NADPH oxidase subunit Cytb558 with the *Salmonella* containing vesicle (SCV) thereby helping to avoid the oxidative burst [66]. An *in vivo* study, conducted with the peptide of SsaJ, however, showed its inability to provide antigen specific immunity when compared with the other chosen peptides [67].

SctC is a layer of outer membrane anchor forming two distinct outer rings namely, OR1 and OR2. It is homologous to a protein of Type II Secretion System (T2SS) which requires pilotin lipoprotein for its optimal assembly and localization [68]. SctC serves as a midline between the inner and outer membrane, with evidence showing that the translocation of foreign antigens can induce potent immune response against pathogens [69].

SsrB is responsible for the survival and replication of *Salmonella* in the host cell and plays an important role in the transcription of multiple genes of SPI-2. SsrB has been claimed as one of the most important factors for *Salmonella*'s virulence by the fact that, a mutated *ssrB*, resulted in reduced ability of colonization on comparing with the wild type [70]. Moreover, one alteration in the gene *ssrB*, preferentially silencing the acquired DNA, can have a high contribution towards low transcription in the virulence factors of *Salmonella* [71].

BcfD is a fimbrial protein and part of the operon Bcf [72]. BcfD is a surface molecule, which helps in the adherence through specific receptors on the host cell. This step of adhesion is considered to be an important course during infection as it allows bacteria to initiate the colonization [73]. A research shows that the knockout of this gene influenced in the low adhesion capacity of *Salmonella* to the host cell [74].

InvE, encoded within SPI-1, is a protein located in the cell membrane and said to be essential for the translocation of *Salmonella* proteins into the host cells by regulating the functions of the Sip protein translocases [75]. An investigation of finding the region of InvE, as the T3SS regulator protein, indicates that it may have two functional domains which are responsible for regulating the secretion of translocases as N-terminal secretion signal and C-terminal regulatory domain [76]. An *in-vivo* study conducted with the BALB/c mice, showed less pathogenicity when it is injected with the mutated *invE* gene *Salmonella* on comparing with the wild strain [77].

SipB is one of the effector proteins of SPI-1 T3SS which facilitates the entry of *Salmonella* into the host cell. It is also called as an invasion protein as it initiates the bacterial entry process. It forms a complex along with the SipC to assemble into plasma membrane-integral structure which mediates the effectors delivery [78]. It also affects the membrane fluidity and bacterial osmotolerance and hence a small alteration of this gene will pave a huge way to prevent *Salmonella* entry into the host cell [79]. In fact, a study evaluating the effect of *sipB* deleted mutants, showed significant decrease in the virulence of *sipB* mutants when compared with the wild-type strains [80].

SsaD is an important cellular component which is responsible for the virulence of *Salmonella.* It is found to be in the transmembrane of the bacteria. The gene *ssaD* encodes for the proteins related to the basal body, cytoplasmic rings and export apparatus and it is also involved in the ATPase complex, regulation and translocation of T3SS [81]. A study shows that there is an important defect in the intercellular survival with the mutant *ssaD* strains on comparing with the wild-type *Salmonella* [82].

DD95_23890 refers to the computationally predicted protein, mapping to the autotransporter adhesin BigA protein. The BigA protein in *Salmonella* has recently

been identified via automated genome annotation in 2015. Thus, studies on this protein has been scarce. Inferring from its homolog in *Brucella*, the cell surface BigA protein promotes adhesion of bacteria on host epithelial cells [83, 84]. The adhesive properties of the BigA protein can be established by binding onto the cell adhesion molecules on the host epithelial cellular surface [85].

DD95_21695 maps to the RING-type E3 ubiquitin transferase (SspH2) protein. The SspH2 protein aids in *Salmonella* pathogenicity by conferring anti-inflammatory properties, hence delaying the host immune response in reaction to bacterial invasion [86]. Moreover, the ability of SspH2 to ubiquitinate host NOD1 protein, through an essential interaction with host SGT1 protein, can result in NOD1-mediated IL-8 secretion in host [87].

DD95_16310 maps to the *Salmonella* TorS histidine kinase sensor. The TorS protein comprises the two-component systems along with the TorT response regulator [88]. Upon stimulation by Trimethylamine-N-oxide, TorS, along with TorT, carry out osmoregulation and protect the cellular proteins against low-pH induced denaturation in urea [88].

DD95_14775 refers to the putative transcriptional regulator *marT_1* in *Salmonella*. The MarT protein mainly regulates the expression of MisL autotransporter protein, which is a fibronectin-binding protein that is involved in the cell adhesive properties of *Salmonella* [89]. Moreover, MarT has also been reported to regulate the expression of genes related to bacterial biofilm formation [90].

5. Initial screening of the candidate proteins

All the twenty proteins were screened to ascertain their potential for plausible candidatures as vaccines (**Table 2**). Proteins were localized in extracellular matrix (3), cytoplasm (7), cytoplasmic membrane (3) and outer membranes (2), besides some of them being predicted with unknown cellular location (5). Of these, surface/outer membrane proteins and vesicles have been deployed for prospective vaccinations against bacterial pathogens [91–94]. Again, extracellular proteins have been potentiated as drugs for prospects against disease management, *albeit,* in a different scenario [95, 96]. Our results predict the proteins namely, SptP, SipB, SsaD, PrgK and TorS to be potentially antigenic except InvE, SctC and SspH2. Notably, the five proteins of unknown location, namely, BcfD, SsaJ, BigA, SiiE, and MarT_1 are all potentially antigenic. Of the two signal peptides BcfD and SctC, the latter was predicted to be non-antigenic while SsaJ and PrgK belongs to another category of signal peptides (lipoproteins) with good antigenic potential. Of these, SsaJ has been predicted with two transmembrane (TM) spanning helices and poses itself a good candidate for vaccines. Other candidates with more TM helices are BigA (5), SiiE (3) and TorS (3). Furthermore, a BLASTp alignment of these 20 proteins revealed SptP and SspH2 to have 40–50% similarity for 101 and 106 hits, respectively, against human counterparts, thereby completely ruling out their candidature as potential vaccines.

6. Selection of potential vaccine candidates

The 20 top ranked proteins were further screened for B cell epitopes. Therein, InvE, SsrB, SicA, and SscA were omitted from being considered as vaccine candidates due to the absence of predicted epitopes that fall within the normal range of peptide length (**Table 3**). Moreover, in allergenicity prediction, HilA, BcfD, SicA, BigA, SiiE, and MarT_1 were predicted to be potential allergens (**Table 4**), and thus, were excluded from consideration as well. Hence, we report SptP, SsaQ, SpaO, PrgH, SipB, SsaD, SctC, SsaJ, PrgK, SspH2, and TorS to be

Protein	Start	End	Peptide	Length	Average Score
SptP	5	25	EERKLNNLTLSSFSKVGVSND	21	0.5922
	59	78	FKNTEVVQKHTENIRVQDQK	20	0.5383
	378	396	EDQMQAKQLPPYFRGSYTF	19	0.5593
SsaQ	5	23	ANEERPWVEILPTQGATIG	19	0.5839
	58	75	WQRWCEGLIGTANRSAID	18	0.5467
	93	113	ASDATLCQNEPPTSCSNLPHQ	21	0.5675
SpaO	21	37	ECQRHGREATLEYPTRQ	17	0.5256
PrgH	84	104	LHELKEGNSESRSVQLNTPIQ	21	0.5941
	115	138	ESEPWVPEQPEKLETSAKKNEPRF	24	0.6261
	164	182	NSPQRQAAELDSLLGQEKE	19	0.5371
	261	277	SRQRNTMSKKELEVLSQ	17	0.5634
SipB	5	22	ASSISRSGYTQNPRLAEA	18	0.5777
	232	253	GTANAASQNQVSQGEQDNLSNV	22	0.5356
	545	569	MDQIQQWLKQSVEIFGENQKVTAEL	25	0.5389
SsaD	18	38	GHVLQGREVWLNEGNLSLGEK	21	0.5384
	155	170	LDKSNIHYVRAQWKED	16	0.5236
	257	274	IPGLLHWQISHSHQSQGD	18	0.5234
	331	345	QDIAPSHDESKYLPA	15	0.5794
HilA	206	227	VKGYHLLHQESIKLIEHQPASL	22	0.5358
	242	256	GLRWDTKQISELNSI	15	0.5672
BcfD	121	144	PMNNVLMGYDENVKAGQPFYVRDS	24	0.5932
	214	232	LYSGNFNHAGQKPEGVRAK	19	0.6014
	282	303	NALIPNDVQSVAPFITDSAGRA	22	0.5514
SctC	430	449	DGNDKTPQSDTTTSVDALPE	20	0.6323
SsaJ	20	34	DVDLYRSLPEDEANQ	15	0.5328
	91	111	NQLVVSPQEEQQKINFLKEQR	21	0.5596
PrgK	20	34	DKDLLKGLDQEQANE	15	0.5634
	188	207	SERSDAQLQAPGTPVKRNSF	20	0.5906
	229	249	YYKNHYARNKKGITADDKAKS	21	0.6077
BigA	1429	1448	RVLSNRFTMLADAAPQIKDG	20	0.5452
	1456	1475	KGDPRAELGNDTQYDMLALR	20	0.5380
	1689	1709	SSNDTALHLDAYQWKEDGISD	21	0.5748
SspH2	72	95	FELLRTLAYAGWEESIHSGQHGEN	24	0.5541
	431	446	RNQLTRLPESLIHLSS	16	0.5176
	574	594	TEATSSCEDRVTFFLHQMKNV	21	0.5220
	620	639	FRLGKLEQIAREKVRTLALV	20	0.5312
TorS	133	148	TLRAQQQQLSRQIAEA	16	0.5421
	172	192	AGIYDLIESGKGDQAERALDR	21	0.5473
	663	677	SKPASKSAFREPINL	15	0.5645

Protein	Start	End	Peptide	Length	Average Score
MarT_1	106	126	ITIIATDSETKGRKKQIVRQT	21	0.5981

Only predicted peptides of length between 15 to 25 amino acids were selected [97]. SiiE protein were omitted from prediction because of its overly huge sequence.

Table 3.
BepiPred v2.0 prediction of linear B-cell epitopes.

Protein	AllerCatPro	AlgPred2		AllergenFP v1.0
		Hybrid Score	Prediction	
SptP	No Hits	0.04	Non-Allergen	Probable Non-Allergen
SsaQ	No Hits	0.08	Non-Allergen	Probable Non-Allergen
SpaO	No Hits	0.03	Non-Allergen	Probable Non-Allergen
PrgH	No Hits	0.03	Non-Allergen	Probable Non-Allergen
SipB	No Hits	0.24	Non-Allergen	Probable Non-Allergen
SsaD	No Hits	0.09	Non-Allergen	Probable Non-Allergen
InvE	No Hits	0.02	Non-Allergen	Probable Non-Allergen
HilA	No Hits	0.54	**Allergen**	Probable Non-Allergen
BcfD	No Hits	0.85	**Allergen**	Probable Non-Allergen
SsrB	No Hits	−0.43	Non-Allergen	Probable Non-Allergen
SctC	No Hits	0.09	Non-Allergen	Probable Non-Allergen
SicA	No Hits	0.31	**Allergen**	**Probable Allergen**
SsaJ	No Hits	−0.45	Non-Allergen	Probable Non-Allergen
SscA	No Hits	0.18	Non-Allergen	Probable Non-Allergen
PrgK	No Hits	−0.48	Non-Allergen	Probable Non-Allergen
BigA	No Hits	0.75	**Allergen**	Probable Non-Allergen
SiiE	No Hits	0.86	**Allergen**	N/A
SspH2	No Hits	−0.48	Non-Allergen	Probable Non-Allergen
TorS	No Hits	0.02	Non-Allergen	Probable Non-Allergen
MarT_1	No Hits	0.55	**Allergen**	Probable Non-Allergen

For AllergenFP v1.0, N/A refers to Not Available because of overly large protein size.

Table 4.
Allergenicity assessment through different predictive tools. Potential allergens are in bold case.

potentially utilized as B cell epitopes. Moreover, in discontinuous B cell epitope prediction, the localizations of the highly antigenic regions were illustrated in 3D (**Figure 2**). For successful vaccination, these regions should be prioritized and retained as much as possible due to their important roles in antigenicity.

7. Potential druggable proteins

Besides potential vaccine candidates, we have conducted predictions on the druggability and druggable sites of the 20 top ranked proteins which have their 3D crystallized structures available in PDB. Eventually, the localization of the top

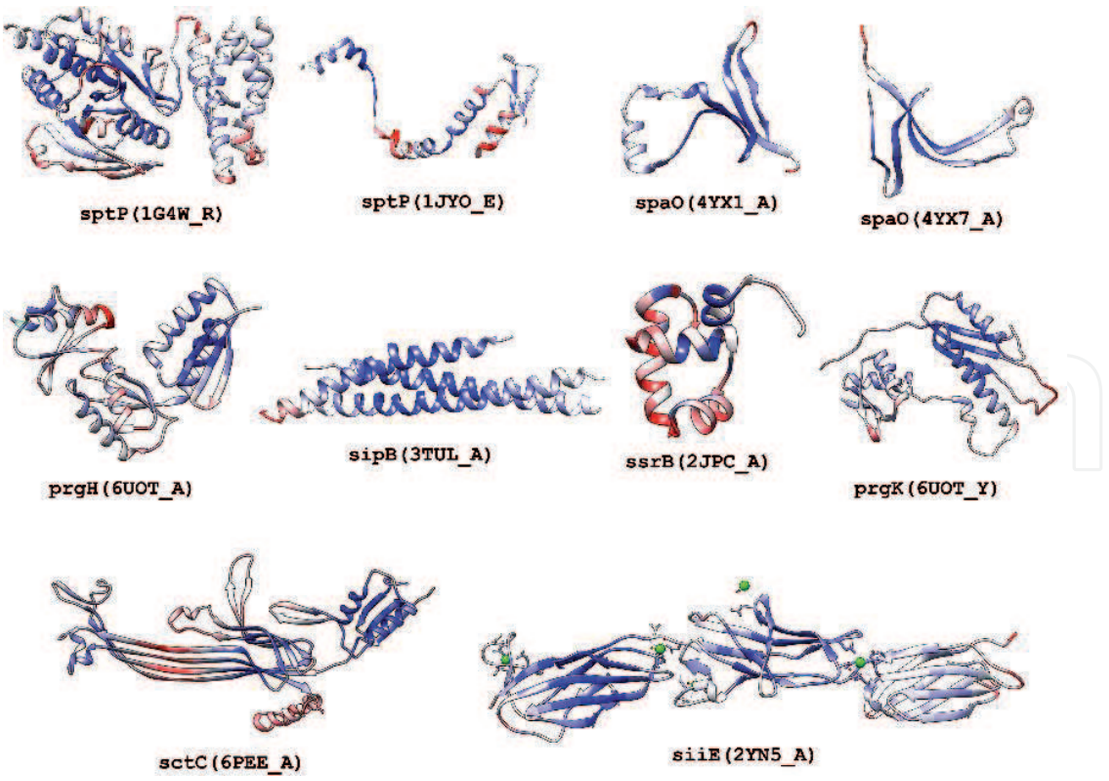


Figure 2.
DiscoTope v2.0 prediction of discontinuous B-cell epitopes. The residues are colored according to their respective DiscoTope scores (red: high, white: threshold of −3.7 and blue: low).

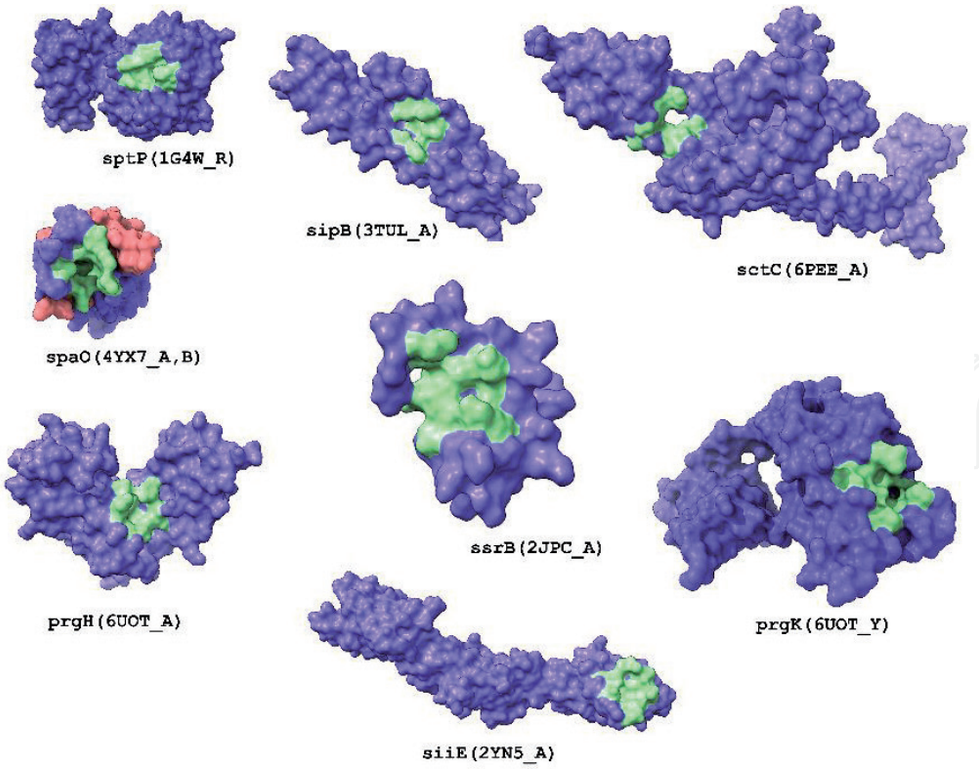
ranked druggable pockets of SptP, SipB, SctC, SpaO, SsrB, PrgK, PrgH, and SiiE were illustrated in 3D (**Figure 3**). This can help future research in structure-aided drug discovery, by designing drugs specific for the druggable pockets to suppress the virulence of *Salmonella*.

8. Conclusions

The study depicted here essentially delineates a schematic approach of shortlisting the most probable virulent proteins as potential vaccine and/or drug candidates from the proteome of *Salmonella* spp. It starts with the building of the theoretical PIN comprising the known and predicted virulent proteins followed by the graph theoretical parametric analyses for identifying a probable set of them. These were further screened through different essential tools enabling the prediction of cellular localisation, signal peptides, transmembrane helices, antigenicity, epitopes, allergenicity and molecular crevices besides comparing with any human homologs. A thorough analysis revealed SsaJ and PrgK to come to the forefront among those already known to be virulent. PrgK even has nice druggable pocket to be targeted through potential drugs. Our approach can pave the way for screening such effective molecular vaccines and/or drug targets for such pathogens. Newer candidates, however, could be unraveled through other effective methods.

Acknowledgements

The authors wish to acknowledge the support of Sunway University, Malaysia for the provision of computational facilities.



Protein	PDB ID	Chain ID	Druggable Residues in Rank 1 Pocket
SptP	1G4W	R	310, 314, 315, 441, 442, 482, 483, 485, 486, 487, 519, 522, 526
	1JYO	E	NONE
SpaO	4YX7	A	19, 20, 21, 25, 34, 35, 36, 38
		B	7, 9
PrgH	4G2S	A	NONE
	6UOT	A	226, 229, 230, 233, 245, 246, 247, 347, 352, 353
SipB	3TUL	A	89, 92, 93, 96, 206, 209, 210
SsrB	2JPC	A	25, 26, 27, 28, 29, 30, 60, 61, 64, 68, 73
SctC	6PEE	A	187, 188, 189, 207, 297, 299, 301, 376, 378, 379
PrgK	6UOT	Y	102, 106, 109, 133, 143, 144, 146, 175, 176
	4OYC	A	NONE
SiiE	2YN5	A	5304, 5306, 5307, 5311, 5313, 5327, 5328, 5329, 5330, 5331, 5335

Figure 3.
P2Rank predicted druggable pockets colored in light green. For SpaO, chain A is in blue, while B is in red. Residues contributing to druggability are tabulated.

Conflict of interest

The authors declare no conflict of interest.

DOI: http://dx.doi.org/10.5772/intechopen.95856

Author details

Rohit Mishra[1†], Yong Chiang Tan[2†], Amr Adel Ahmed Abd El-Aal[2] and Chandrajit Lahiri[2*]

1 Advanced Centre for Treatment, Research and Education in Cancer, Tata Memorial Centre, Mumbai, India

2 Department of Biological Sciences, Sunway University, Petaling Jaya, Selangor, Malaysia

*Address all correspondence to: chandrajitl@sunway.edu.my

† These authors have contributed equally to the work.

References

[1] Miller S. Salmonella species, including S. typhi [Internet]. [place unknown]: InfectiousDiseaseAdvisor; [date unknown; cited 2021 Jan 01]. Available from: https://www.infectiousdiseaseadvisor.com/home/decision-support-in-medicine/infectious-diseases/salmonella-species-including-s-typhi/

[2] World Health Organization. Typhoid [Internet]. [place unknown]: World Health Organization; 2018 [updated 2018; cite 2021 Jan 01]. Available from: https://www.who.int/news-room/fact-sheets/detail/typhoid

[3] World Health Organization. Salmonella (non-typhoidal) [Internet]. [place unknown]: World Health Organization; 2018 [updated 2018; cite 2021 Jan 01]. Available from: https://www.who.int/news-room/fact-sheets/detail/salmonella-(non-typhoidal)

[4] Giannella RA. Salmonella. 4th edition. Galveston (TX): Medical Microbiology; 1996. Chapter 21.

[5] Ashurst JV, Truong J, Woodbury B. Salmonella Typhi. InStatPearls [Internet] 2019 Jun 11. StatPearls Publishing.

[6] Mandell GL. Mandell, Douglas, and Bennett's Principles and Practice of Infectious Diseases. 7th Edition. United States: Churchill Livingstone; 2009.

[7] Bennett JE, Dolin R, Blaser MJ. Mandell, douglas, and bennett's principles and practice of infectious diseases: 2-volume set. Elsevier Health Sciences; 2014 Aug 28.

[8] Liao YS, Chen BH, Hong YP, Teng RH, Wang YW, Liang SY, Liu YY, Tu YH, Chen YS, Chang JH, Tsao CS. Emergence of multidrug-resistant *Salmonella enterica* serovar Goldcoast strains in Taiwan and international spread of the ST358 clone. Antimicrobial agents and chemotherapy. 2019 Oct 1;63(10):e01122-19.

[9] Tacket CO, Levine MM. Human Typhoid Vaccines-Old and New. In: Ala'Aldeen AA, Hormaeche CE, eds. Molecular and Clinical Aspects of Bacterial Vaccine Development. Chichester: John Wiley & Sons, 1995. p. 155-178.

[10] Papadimitropoulos V, Vergidis PI, Bliziotis I, Falagas ME. Vaccination against typhoid fever in travellers: a cost-effectiveness approach. Clinical Microbiology and Infection. 2004 Aug;10(8):681-3.

[11] MacLennan CA, Martin LB, Micoli F. Vaccines against invasive Salmonella disease: current status and future directions. Human vaccines & immunotherapeutics. 2014 Jun 4;10(6):1478-93.

[12] Erova TE, Kirtley ML, Fitts EC, Ponnusamy D, Baze WB, Andersson JA, Cong Y, Tiner BL, Sha J, Chopra AK. Protective immunity elicited by oral immunization of mice with *Salmonella enterica* serovar Typhimurium Braun lipoprotein (Lpp) and acetyltransferase (MsbB) mutants. Frontiers in cellular and infection microbiology. 2016 Nov 10;6:148.

[13] Galen JE, Buskirk AD, Tennant SM, Pasetti MF. Live attenuated human Salmonella vaccine candidates: tracking the pathogen in natural infection and stimulation of host immunity. EcoSal Plus. 2016 Nov;7(1).

[14] Gayet R, Bioley G, Rochereau N, Paul S, Corthésy B. Vaccination against Salmonella infection: the mucosal way. Microbiology and Molecular Biology Reviews. 2017 Sep 1;81(3).

[15] Clark-Curtiss JE, Curtiss R. Salmonella vaccines: conduits

for protective antigens. The Journal of Immunology. 2018 Jan 1;200(1):39-48.

[16] Moriel DG, Beatson SA, Wurpel DJ, Lipman J, Nimmo GR, Paterson DL, Schembri MA. Identification of novel vaccine candidates against multidrug-resistant Acinetobacter baumannii. PloS one. 2013 Oct 8;8(10):e77631.

[17] Singh R, Capalash N, Sharma P. Immunoprotective potential of BamA, the outer membrane protein assembly factor, against MDR Acinetobacter baumannii. Scientific reports. 2017 Sep 29;7(1):1-1.

[18] Chiang MH, Sung WC, Lien SP, Chen YZ, Lo AF, Huang JH, Kuo SC, Chong P. Identification of novel vaccine candidates against Acinetobacter baumannii using reverse vaccinology. Human vaccines & immunotherapeutics. 2015 Apr 3;11(4):1065-73.

[19] Pan A, Lahiri C, Rajendiran A, Shanmugham B. Computational analysis of protein interaction networks for infectious diseases. Briefings in bioinformatics. 2016 May 1;17(3):517-26.

[20] Jeong H, Mason SP, Barabási AL, Oltvai ZN. Lethality and centrality in protein networks. Nature. 2001 May;411(6833):41-2.

[21] Lahiri C, Pawar S, Sabarinathan R, Ashraf MI, Chand Y, Chakravortty D. Interactome analyses of Salmonella pathogenicity islands reveal SicA indispensable for virulence. Journal of theoretical biology. 2014 Dec 21;363:188-97.

[22] Ashraf MI, Ong SK, Mujawar S, Pawar S, More P, Paul S, Lahiri C. A side-effect free method for identifying cancer drug targets. Scientific reports. 2018 Apr 27;8(1):1-9.

[23] Pawar S, Ashraf M, Mehata K, Lahiri C. Computational identification of indispensable virulence proteins of Salmonella Typhi CT18. Curr. Top. Salmonella Salmonellosis. 2017 Apr 5.

[24] Pawar S, Ashraf M, Mujawar S, Mishra R, Lahiri C. In silico identification of the indispensable quorum sensing proteins of multidrug resistant *Proteus mirabilis*. Frontiers in cellular and infection microbiology. 2018 Aug 7;8:269.

[25] Lahiri C, Mujawar S, Mishra R, Ashraf M, Pawar S, Gatherer D. Delineating the plausible molecular vaccine candidates and drug targets of multidrug resistant Acinetobacter baumannii. Frontiers in Cellular and Infection Microbiology. 2019;9:203.

[26] Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic acids research. 2019 Jan 8;47(D1):D687-92.

[27] Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ. STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic acids research. 2019 Jan 8;47(D1):D607-13.

[28] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research. 2003 Nov 1;13(11):2498-504.

[29] Diestel R. Graph Theory. Berlin and Heidelberg: Springer-Verlag, 2000.

[30] Tang Y, Li M, Wang J, Pan Y, Wu FX. CytoNCA: a cytoscape plugin for centrality analysis and evaluation of protein interaction networks. Biosystems. 2015 Jan 1;127:67-72.

[31] Oliveros JC. Venny. An interactive tool for comparing lists with Venn's diagrams [Internet]. Venny 2.1.0. 2007 [cited 2020]. Available from: https://bioinfogp.cnb.csic.es/tools/venny/index.html

[32] Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A. Protein identification and analysis tools on the ExPASy server. InThe proteomics protocols handbook 2005 (pp. 571-607). Humana press.

[33] Yu NY, Wagner JR, Laird MR, Melli G, Rey S, Lo R, Dao P, Sahinalp SC, Ester M, Foster LJ, Brinkman FS. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. Bioinformatics. 2010 Jul 1;26(13):1608-15.

[34] Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nature biotechnology. 2019 Apr;37(4):420-3.

[35] Juncker AS, Willenbrock H, Von Heijne G, Brunak S, Nielsen H, Krogh A. Prediction of lipoprotein signal peptides in Gram-negative bacteria. Protein Science. 2003 Aug;12(8):1652-62.

[36] Doytchinova IA, Flower DR. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC bioinformatics. 2007 Dec 1;8(1):4.

[37] Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. Journal of computational chemistry. 2004 Oct;25(13):1605-12.

[38] Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic acids research. 2017 Jul 3;45(W1):W24-9.

[39] Kringelum JV, Lundegaard C, Lund O, Nielsen M. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. PLoS Comput Biol. 2012 Dec 27;8(12):e1002829.

[40] Tan YC. DiscoTope2ChimeraAttr_v1.0.0 [Internet]. Malaysia: Zenodo; 2021 [updated 2021, cited 2021 Jan 01]. Available from: https://github.com/TanYongChiang/DiscoTope2ChimeraAttr/tree/v1.0.0. DOI: 10.5281/zenodo.4408680

[41] Maurer-Stroh S, Krutz NL, Kern PS, Gunalan V, Nguyen MN, Limviphuvadh V, Eisenhaber F, Gerberick GF. AllerCatPro—prediction of protein allergenicity potential from the protein sequence. Bioinformatics. 2019 Sep 1;35(17):3020-7.

[42] Sharma N, Patiyal S, Dhall A, Pande A, Arora C, Raghava GP. AlgPred 2.0: an improved method for predicting allergenic proteins and mapping of IgE epitopes. Briefings in Bioinformatics. 2020 Nov 17.

[43] Dimitrov I, Naneva L, Doytchinova I, Bangov I. AllergenFP: allergenicity prediction by descriptor fingerprints. Bioinformatics. 2014 Mar 15;30(6):846-51.

[44] Krivák R, Hoksza D. P2Rank: machine learning based tool for rapid and accurate prediction of ligand binding sites from protein structure. Journal of cheminformatics. 2018 Dec 1;10(1):39.

[45] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. Journal

of molecular biology. 1990 Oct 5;215(3):403-10.

[46] Ehrbar K, Friebel A, Miller SI, Hardt WD. Role of the Salmonella pathogenicity island 1 (SPI-1) protein InvB in type III secretion of SopE and SopE2, two Salmonella effector proteins encoded outside of SPI-1. Journal of bacteriology. 2003 Dec 1;185(23):6950-67.

[47] Guo Y, Xu Y, Kang X, Gu D, Jiao Y, Meng C, Tang P, Wang X, Huang C, Geng S, Jiao X. Immunogenic potential and protective efficacy of a sptP deletion mutant of Salmonella Enteritidis as a live vaccine for chickens against a lethal challenge. International Journal of Medical Microbiology. 2019 Dec 1;309(8):151337.

[48] Yoshida Y, Miki T, Ono S, Haneda T, Ito M, Okada N. Functional characterization of the type III secretion ATPase SsaN encoded by Salmonella pathogenicity island 2. PLoS One. 2014 Apr 10;9(4):e94347.

[49] Sakai T, Miyata T, Terahara N, Mori K, Inoue Y, Morimoto YV, Kato T, Namba K, Minamino T. Novel insights into conformational rearrangements of the bacterial flagellar switch complex. Mbio. 2019 Apr 30;10(2).

[50] Lara-Tejero M, Qin Z, Hu B, Butan C, Liu J, Galán JE. Role of SpaO in the assembly of the sorting platform of a Salmonella type III secretion system. PLoS pathogens. 2019 Jan 22;15(1):e1007565.

[51] Notti RQ, Bhattacharya S, Lilic M, Stebbins CE. A common assembly module in injectisome and flagellar type III secretion sorting platforms. Nature communications. 2015 May 21;6(1):1-1.

[52] Ruan P, Xia XP, Sun D, Ojcius DM, Mao YF, Yue WY, Yan J. Recombinant SpaO and H1a as immunogens for protection of mice from lethal infection with Salmonella paratyphi A: Implications for rational design of typhoid fever vaccines. Vaccine. 2008 Dec 2;26(51):6639-44.

[53] Kimbrough TG, Miller SI. Contribution of Salmonella typhimurium type III secretion components to needle complex formation. Proceedings of the National Academy of Sciences. 2000 Sep 26;97(20):11008-13.

[54] Myeni SK, Wang L, Zhou D. SipB-SipC complex is essential for translocon formation. PLoS One. 2013 Mar 27;8(3):e60499.

[55] Darwin KH, Miller VL. Type III secretion chaperone-dependent regulation: activation of virulence genes by SicA and InvF in Salmonella typhimurium. The EMBO journal. 2001 Apr 17;20(8):1850-62.

[56] Lv Q, Li S, Wei H, Wen Z, Wang Y, Tang T, Wang J, Xia L, Deng X. Identification of the natural product paeonol derived from peony bark as an inhibitor of the *Salmonella enterica* serovar Typhimurium type III secretion system. Applied Microbiology and Biotechnology. 2020 Feb 1;104(4):1673-82.

[57] Queiroz MH, Madrid C, Paytubi S, Balsalobre C, Juarez A. Integration host factor alleviates H-NS silencing of the *Salmonella enterica* serovar Typhimurium master regulator of SPI1, hilA. Microbiology. 2011 Sep 1;157(9):2504-14.

[58] Zhang Y, Liu Y, Wang T, Deng X, Chu X. Natural compound sanguinarine chloride targets the type III secretion system of *Salmonella enterica* Serovar Typhimurium. Biochemistry and biophysics reports. 2018 Jul 1;14:149-54.

[59] Barlag B, Hensel M. The giant adhesin SiiE of *Salmonella enterica*. Molecules. 2015 Jan;20(1):1134-50.

[60] Li X, Bleumink-Pluym NM, Luijkx YM, Wubbolts RW, van Putten JP, Strijbis K. MUC1 is a receptor for the Salmonella SiiE adhesin that enables apical invasion into enterocytes. PLoS pathogens. 2019 Feb 4;15(2):e1007566.

[61] Männe C, Takaya A, Yamasaki Y, Mursell M, Hojyo S, Wu TY, Sarkander J, McGrath MA, Cornelis R, Hahne S, Cheng Q. Salmonella SiiE prevents an efficient humoral immune memory by interfering with IgG+ plasma cell persistence in the bone marrow. Proceedings of the National Academy of Sciences. 2019 Apr 9;116(15):7425-30.

[62] Gan Q, Lehman BP, Bobik TA, Fan C. Expanding the genetic code of Salmonella with non-canonical amino acids. Scientific reports. 2016 Dec 23;6:39920.

[63] Bearson BL, Bearson SM, Brunelle BW, Bayles DO, Lee IS, Kich JD. Salmonella DIVA vaccine reduces disease, colonization and shedding due to virulent S. Typhimurium infection in swine. Journal of medical microbiology. 2017 May;66(5):651.

[64] Maserati A, Fink RC, Lourenco A, Julius ML, Diez-Gonzalez F. General response of *Salmonella enterica* serovar Typhimurium to desiccation: a new role for the virulence factors sopD and sseD in survival. PLoS One. 2017 Nov 8;12(11):e0187692.

[65] Askoura M, Hegazy WA. Ciprofloxacin interferes with Salmonella Typhimurium intracellular survival and host virulence through repression of Salmonella pathogenicity island-2 (SPI-2) genes expression. Pathogens and Disease. 2020 Feb;78(1):ftaa011.

[66] Wang M, Qazi IH, Wang L, Zhou G, Han H. Salmonella Virulence and Immune Escape. Microorganisms. 2020 Mar;8(3):407.

[67] Singh S, Pratap Cb, Nath G. Vaccine Development Against Salmonella Typhi: A Search is Still On. Journal of Clinical & Diagnostic Research. 2019 Aug 1;13(8).

[68] Notti RQ, Stebbins CE. The structure and function of type III secretion systems. Virulence Mechanisms of Bacterial Pathogens. 2016 Jun 22:241-64.

[69] Seyrek K, Ivanisenko NV, Richter M, Hillert LK, König C, Lavrik IN. Controlling Cell Death through Post-translational Modifications of DED Proteins. Trends in Cell Biology. 2020 Mar 29.

[70] Cox CE, Wright AC, McClelland M, Teplitski M. Influence of *Salmonella enterica* Serovar Typhimurium ssrB on Colonization of Eastern Oysters (*Crassostrea virginica*) as Revealed by a Promoter Probe Screen. Applied and environmental microbiology. 2016 Jan 1;82(1):328-39.

[71] Sanapala S, Mosca L, Wang S, Curtiss R. Comparative evaluation of Salmonella Typhimurium vaccines derived from UK-1 and 14028S: Importance of inherent virulence. PloS one. 2018 Sep 7;13(9):e0203526.

[72] De Masi L, Yue M, Hu C, Rakov AV, Rankin SC, Schifferli DM. Cooperation of adhesin alleles in Salmonella-host tropism. Msphere. 2017 Apr 26;2(2).

[73] Azinheiro S, Carvalho J, Prado M, Garrido-Maestu A. Evaluation of different genetic targets for *Salmonella enterica* serovar Enteriditis and Typhimurium, using Loop-Mediated Isothermal AMPlification for detection in food samples. Frontiers in Sustainable Food Systems. 2018 Feb 21;2:5.

[74] Hu Y, Cong Y, Li S, Rao X, Wang G, Hu F. Identification of in vivo induced protein antigens of *Salmonella enterica*

serovar Typhi during human infection. Science in China Series C: Life Sciences. 2009 Oct 1;52(10):942-8.

[75] Du X, Liu Y, Sun X, Zhao J, Wang J, Li J, Liu X, Wang J, Li Y. The mRNA expression of ompF, invA and invE was associated with the ciprofloxacin-resistance in Salmonella. Archives of Microbiology. 2020 Jun 13:1-6.

[76] Kim JS, Im Jang J, Eom JS, Oh CH, Kim HG, Kim BH, Bang IS, Bang SH, Park YK. Molecular characterization of the InvE regulator in the secretion of type III secretion translocases in *Salmonella enterica* serovar Typhimurium. Microbiology. 2013 Mar 1;159(Pt_3):446-61.

[77] Velasco-Carrillo J, Araque-Granadillo M, Ayala-Serrano J, Dávila-Vera D, Peña-Contreras Z, Mendoza-Briceño R. Patogenia de mutantes de Salmonella Typhimurium en dos modelos experimentales in vivo. VacciMonitor. 2019;28(1):1-8.

[78] Gong H, Su J, Bai Y, Miao L, Kim K, Yang Y, Liu F, Lu S. Characterization of the expression of Salmonella Type III secretion system factor PrgI, SipA, SipB, SopE2, SpaO, and SptP in cultures and in mice. BMC microbiology. 2009 Dec 1;9(1):73.

[79] Birhanu BT, Park NH, Lee SJ, Hossain MA, Park SC. Inhibition of Salmonella Typhimurium adhesion, invasion, and intracellular survival via treatment with methyl gallate alone and in combination with marbofloxacin. Veterinary research. 2018 Dec 1;49(1):101.

[80] Chen S, Liao C, Zhang C, Cheng X. Roles of the crp and sipB genes of *Salmonella enterica* serovar Typhimurium in protective efficacy and immune responses to vaccination in mice. Canadian Journal of Veterinary Research. 2018 Apr 1;82(2):102-5.

[81] Seribelli AA, Cruz MF, Vilela FP, Frazão MR, Paziani MH, Almeida F, Medeiros MI, Rodrigues DD, Kress MR, Allard MW, Falcão JP. Phenotypic and genotypic characterization of Salmonella Typhimurium isolates from humans and foods in Brazil. Plos one. 2020 Aug 18;15(8):e0237886.

[82] Riquelme S, Varas M, Valenzuela C, Velozo P, Chahin N, Aguilera P, Sabag A, Labra B, Álvarez SA, Chávez FP, Santiviago CA. Relevant genes linked to virulence are required for Salmonella Typhimurium to survive intracellularly in the social amoeba Dictyostelium discoideum. Frontiers in microbiology. 2016 Aug 23;7:1305.

[83] Guard J, Cao G, Luo Y, Baugher JD, Davison S, Yao K, Hoffmann M, Zhang G, Likens N, Bell RL, Zheng J. Genome sequence analysis of 91 Salmonella Enteritidis isolates from mice caught on poultry farms in the mid 1990s. Genomics. 2020 Jan 1;112(1):528-44.

[84] Czibener C, Merwaiss F, Guaimas F, Del Giudice MG, Serantes DA, Spera JM, Ugalde JE. BigA is a novel adhesin of Brucella that mediates adhesion to epithelial cells. Cellular microbiology. 2016 Apr;18(4):500-13.

[85] Bialer MG, Sycz G, Muñoz González F, Ferrero MC, Baldi PC, Zorreguieta A. Adhesins of Brucella: Their Roles in the Interaction with the Host. Pathogens. 2020 Nov;9(11):942.

[86] Shappo MO, Li Q, Lin Z, Hu M, Ren J, Xu Z, Pan Z, Jiao X. SspH2 as anti-inflammatory candidate effector and its contribution in Salmonella Enteritidis virulence. Microbial Pathogenesis. 2020 May 1;142:104041.

[87] Bhavsar AP, Brown NF, Stoepel J, Wiermer M, Martin DD, Hsu KJ, Imami K, Ross CJ, Hayden MR, Foster LJ, Li X. The Salmonella type III

effector SspH2 specifically exploits the NLR co-chaperone activity of SGT1 to subvert immunity. PLoS Pathog. 2013 Jul 25;9(7):e1003518.

[88] Moore JO, Hendrickson WA. An asymmetry-to-symmetry switch in signal transmission by the histidine kinase receptor for TMAO. Structure. 2012 Apr 4;20(4):729-41.

[89] Tükel Ç, Akçelik M, De Jong MF, Şimşek Ö, Tsolis RM, Bäumler AJ. MarT activates expression of the MisL autotransporter protein of *Salmonella enterica* serotype Typhimurium. Journal of bacteriology. 2007 May 15;189(10):3922-6.

[90] Eran Z, Akçelik M, Özcengiz G, Akçelik N. Regulation of biofilm formation by marT in Salmonella Typhimurium. Molecular Biology Reports. 2020 Jun 11:1-0.

[91] Grandi G. Bacterial surface proteins and vaccines. F1000 biology reports. 2010;2.

[92] Baliga P, Shekar M, Venugopal MN. Potential outer membrane protein candidates for vaccine development against the pathogen Vibrio anguillarum: a reverse vaccinology based identification. Current microbiology. 2018 Mar 1;75(3):368-77.

[93] Maiti B, Dubey S, Munang'andu HM, Karunasagar I, Karunasagar I. Application of outer membrane protein-based vaccines against major bacterial fish pathogens in india. Frontiers in Immunology. 2020;11:1362.

[94] Hu R, Li J, Zhao Y, Lin H, Liang L, Wang M, Liu H, Min Y, Gao Y, Yang M. Exploiting bacterial outer membrane vesicles as a cross-protective vaccine candidate against avian pathogenic *Escherichia coli* (APEC). Microbial Cell Factories. 2020 Dec;19(1):1-7.

[95] Cudic M, Fields GB. Extracellular proteases as targets for drug development. Current Protein and Peptide Science. 2009 Aug 1;10(4):297-307.

[96] Ahmad V. Prospective of Extracellular Matrix and Drug Correlations in disease management. Asian Journal of Pharmaceutical Sciences. 2020 Aug 30.

[97] Potocnakova L, Bhide M, Pulzova LB. An introduction to B-cell epitope mapping and in silico epitope prediction. Journal of immunology research. 2016 Oct;2016.