Title: A new design and epitopes analysis for recombinant vaccine against Salmonella typhi by In silico analysis

Author: Tahereh Bidmeshki Barzoki

Institute/Affiliation: &mdash;

Received: 2018-06-11

Accepted: 2018-09-05

Process: 1、First trial(Field and check) ✓
         2、Peer review ☐
         3、Editing and three trials ☐
         4、Published online ☐
A new design and epitopes analysis for recombinant vaccine against

Salmonella typhi by In silico analysis

Tahereh Bidmeshki Barzoki¹, Ali Mohammad Ahadi¹,* Hoda Ayat¹

1. Department of Genetics, Faculty of Science, Shahrekord University,

Shahrekord, Iran

* Correspond to: Dr. Ali Mohammad Ahadi, Department of Genetics, Faculty of Science, Shahrekord University, Shahrekord, Iran.

E-Mail: Ahadi_al@sci.sku.ac.ir

Mobile: +989122094047

Fax: +983832324419

Running title: Epitope mapping of a new vaccine against S. typhi
Abstract

Nowadays, foodborne diseases are one of the main problems in the world that infect humans due to consumption of contaminated water or food. Typhoid fever is one of the major causes of illness and death in the world caused by *Salmonella typhi*. Vaccination is one of the most effective approaches in order to reducing the risk of disease. The main goal of this study is designing and characterization of antigenic determinants of a fusion protein originated from *S. typhi* usable as an effective vaccine. In this study, the outer membrane proteins of *salmonella* have been considered as candidates conferring protection against typhoid. Considering the evidence, OmpA, OmpF and OmpC proteins of salmonella applied in a multivalent vaccine design. Conserved motives of these proteins were selected using the CLC software and then their extracellular regions of these peptides were identified with PRED-TMBB server. Appropriate motives were combined for design of final fusion protein. Finally epitops of designed protein with high antigenic properties were identified using BCPREDs, Ellipro, ABCpred, EpiJen, NetCTL-1.2, CTLpred, TAPpred, ProPred and VaxiJen servers. Predicted designed protein in this study reached a very high scores for antigenic indexes. Encoding Genetic construction of this fusion protein could be applied for production of the recombinant OmpA.OmpF.OmpC derived fusion protein with effective antigenic properties as a new vaccine against *S. typhi*. Laboratory experiments and animal challenging analyses is ongoing.

**Key words:** *Salmonella typhi*, recombinant multiepitopic vaccine
Introduction

Foodborne diseases are dispersed in the world in humans due to consumption of contaminated water or food [1]. Typhoid fever is one of the major etiologic factors involved in illness and death in the world caused by *S. Typhi* and it is the most common food poisoning [1, 2]. Typhoid fever is an enteric bacterial infection [3], also *S. paratyphi* A and *S. paratyphi* B cause paratyphoid fever that is less common but clinically similar to the typhoid fever [4]. The disease is transmit and spread in a fecal-oral manner [3].

Nowadays in developed countries, salmonella is commonly associated with acute non-systemic gastroenteritis. But certain serovars of Salmonella, such as *S. typhi* and *S. paratyphi* are important as the causative agent of typhoid fever that is a common diseases in developing countries [5]. Typhoid fever is a public health problem that is endemic in many developed countries, such as Africa, Asia and south America regions [6, 7]. The high incidence of the disease have been reported in South and Southeast Asia [8]. It is reported over 20 million cases of typhoid fever annually that led to the deaths of approximately 200,000 people [9, 10]. Travelers to developing countries, civilians living in endemic areas, children and technicians in microbiology laboratories are at high risk of disease [4, 6]. Typhoid fever is considered as a disease of school-aged children; and it is noteworthy economic and social impact on the community [8, 11]. Salmonella spp. are highly pathogenic members of the *Enterobacteriaceae* family in both humans and animals [12]. *S. typhi* is a human specific non-Sporulating gram-negative anaerobic bacillus and facultative intracellular pathogen [13, 14]. Salmonella is able to adapt with different conditions, such as low pH and high temperature [15]. The most important surface antigens of *Salmonella typhi* include lipopolysaccharide (O antigen), Flagella (H antigen) and capsular polysaccharide (VI antigen). More than 55% of patients who are infected with typhoid fever, ingest at least 10^5 organisms [4].

Although this disease is successfully treated with antibiotics but the increased emergence of the antimicrobial resistance phenotype in bacteria becomes a major problem that leads to difficulties in the treatment of the disease [4, 13, 16]. *Also in endemic areas, early and accurate diagnosis of typhoid fever is difficult and important because, in addition typhoid fever, many other factors are involved in fever emergence* [17]. *On the other hand*, the treatment of typhoid fever is significantly associated with direct and indirect costs which impose a burden on national health care facilities. Therefore, in order to reduce the impact of the disease, we are required to control the disease by preventive strategies, such as improved sanitation and application of effective vaccines [4]. Despite the improved quality of water and...
promoting health, vaccination is one of the most effective recommended ways by the World Health Organization for reduction of the disease risk [18]. Many studies have been done to develop the effective vaccines for protection against diseases caused by Salmonella species [16]. The first inactivated whole-cell vaccines was licensed in the 1970s in America and Europe. However due to the associated effects, it is not suitable for use [7]. The two typhoid vaccines that are currently available are the injectable VI polysaccharide vaccine and the oral live-attenuated Ty21a vaccine. Currently available vaccines are not satisfactory because of undesirable side effects or a lack of sustained effectiveness [19]. These vaccines have some limitations and cannot been used for children under two years of age and older age groups [16]. These reports show the need to development of new generation of vaccine against S. typhi.

Outer membrane of gram-negative bacteria has large number of proteins that are known as OMPs. A set of these proteins are called porins, which form channels in membrane to transport the small hydrophobic molecules [20]. Expression of OMPs is regulated according to the environmental conditions [21]. Porins are not only involved in a wide range of pathogenic process but also they are resistant to high temperatures and denaturant agent. OmpA, OmpF and OmpC are synthesized in salmonella. pH and osmolality conditions affects the expression of porins. OmpF is expressed under low osmolality conditions, but OmpC is expressed under both low and high osmolality conditions [9,21,22]. Recent studies have shown the importance of the role of porins in bacterial pathogenesis and ability of them in promoting of humoral and cellular immune system in host [9]. Nowadays there are many approaches to design vaccines, but we need new methods to access vaccines with better quality, less complications and more effectiveness. Due to the severity of the disease, relapse and spread through asymptomatic carriers, the incidence of drug resistance and use as a biological weapon, Salmonella has become a serious threat. Therefore in this study we try to design a new multiepitopic recombinant protein as a candidate model for new generation of vaccine against salmonella typhi.

Material and methods

In this study salmonella enterica, Serovar typhi was selected as a model for epitope studies. Protein sequences of OmpF, OmpA and OmpC proteins (GeneBank accession number AAO69550.1, AAO69468.1 and AAO68302.1 respectively) were collected from NCBI genome database (www.ncbi.nlm.nih.gov). The sequence of OmpF, OmpA and OmpC proteins belonging to S. bongori, S. choleraesuis, S. enteritidis, S. gallinarum, S. typhimurium, S. sofia, S. Dublin and S. minnesota species collected from NCBI genome database, are involved in a homology study for identification of
conserved motives. Using CLC main workbench 7.5.1 software \cite{23}, conserved blocks were identified for construction of recombinant fusion protein. In the next step, extracellular region of OmpF, OmpA and OmpC proteins sequences were considered by PRED-TMBB server (http://bioinformatics.biol.uoa.gr/PRED-TMBB) \cite{24}. Finally some parts of three proteins that were conserved and belonged to the extracellular region were selected to design fusion protein. Physical and chemical properties of the fusion pepti

de, such as number and percentage of constitutive amino acid residues were analyzed by ProtParam (web.expasy.org.protparam). Secondary structure of fusion protein was predicted by several different servers, such as Phyre2 (http://www.sbg.bio.ic.ac.uk.phyre2.html.page.cgi?id=index) \cite{25}, predict protein (https://www.predictprotein.org) \cite{26} and PSIPRED (http://bioinf.cs.ucl.ac.uk.psipred) \cite{27}. Tertiary structure of proposal fusion protein was predicted by Phyre2 and was analyzed by YASARA software \cite{28}.

**Prediction of antigenic properties**

Antigen properties of OmpF, OmpA, OmpC proteins and derived fusion protein was calculated by VaxiJen v2.1 (http://www.ddg-pharmfac.net.vaxijen.VaxiJen.VaxiJen.html). VaxiJen is a server developed to determine antigenic probability based on the physicochemical properties of proteins without recourse to sequence alignment \cite{29}. Accuracy of this predictions provided by VaxiJen is 70% to 89%.

**Prediction of linear and conformational B cell epitopes**

In order to the prediction of the linear B cell epitopes, the sequence of the proposed fusion protein was submitted to Ellipro (http://tools.immuneepitope.org.ellipro) \cite{30} and ABCored servers (http://www.imtech.res.in.raghava.abcpred) \cite{31}. Conformational B cell epitopes were predicted by submission of PDB format of the proposed fusion protein to Ellipro server.

**Prediction of MHC and T-cell related epitopes**

Prediction of epitopes that interact with MHC-I and T-cells was carried out by some different servers as listed in table 1. Finally, concluded results were compared. MHC-II binding motives were predicted by ProPred server.

**Table 1.** The List of some servers applied to the prediction of epitopes that intract
with MHC and T cell receptors in this study.

<table>
<thead>
<tr>
<th>Server</th>
<th>Address</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NetCTL-1.2</td>
<td><a href="http://www.cbs.dtu.dk/services.NetCTL.">http://www.cbs.dtu.dk.services.NetCTL.</a></td>
<td>[33]</td>
</tr>
<tr>
<td>CTLpred</td>
<td><a href="http://www.imtech.res.in.raghava.ctlpred.">http://www.imtech.res.in.raghava.ctlpred.</a></td>
<td>[34]</td>
</tr>
</tbody>
</table>

**Results**

Because of laboratory research on this vaccine is being done, the sequence of our proposed multiepitopic protein dose not showed here, but it is available on request. Physical and chemical characterization showed that this fusion protein composed of 297 amino acid 3.2248 kD molecular weight. Based on ProtParam algorithms, isoelectric pH for this protein is 4.79 with hydrophobic property (GRAVY: -0.712). The secondary structure of this protein composes 3.4% α-helix, 49.8% β-sheet and 46.8 loop (Figure 1). Stability index (18.11) show that these protein is stable. Whole structure is very similar to structure of OMPs protein family.

![Figure 1: Tertiary structure of proposed protein. Beta-sheets were shown in dark ribbons. View by YASARA software.](http://www.imtech.res.in/raghava/propred/protein.png)

**Antigenic properties:**

**Linear B-cell Epitopes.** Antigenic property probability of our proposed fusion protein was calculated 0.8896. The results from the different servers showed linear B-cell epitopes with nearby points. Table 2 shows some epitope sequence extracted by two different servers.
Table 2: Linear B-cell Epitopes Predicted by ABCpred and Ellipro servers.

<table>
<thead>
<tr>
<th>ABCpred server</th>
<th>Start position</th>
<th>Score</th>
<th>Ellipro server</th>
<th>Start position</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDANTIGTRPDNGLLS</td>
<td>210</td>
<td>0.87</td>
<td>NNIGDA</td>
<td>207</td>
<td>0.828</td>
</tr>
<tr>
<td>NHIINSQNGDGVGYTM</td>
<td>55</td>
<td>0.81</td>
<td>GIQYQGKNQDNHSI</td>
<td>45</td>
<td>0.777</td>
</tr>
<tr>
<td>TSNGSNPSTSYGFANK</td>
<td>236</td>
<td>0.87</td>
<td>GSNPSTSYGFAN</td>
<td>239</td>
<td>0.726</td>
</tr>
<tr>
<td>VPGGASTKHHDGTGSP</td>
<td>168</td>
<td>0.86</td>
<td>WRADTKSNPGGASTKD</td>
<td>160</td>
<td>0.691</td>
</tr>
<tr>
<td>GLRPSVAYLQSGKDI</td>
<td>266</td>
<td>0.74</td>
<td>FEVVAQYQDFGLRPSVAYLQSGKDIS</td>
<td>255</td>
<td>0.661</td>
</tr>
<tr>
<td>PYKGDNTNGQGQLTA</td>
<td>124</td>
<td>0.64</td>
<td>GDNTNGQGV</td>
<td>127</td>
<td>0.636</td>
</tr>
<tr>
<td>SSQTNATRGTSNGS</td>
<td>225</td>
<td>0.74</td>
<td>SQFTYNAT</td>
<td>226</td>
<td>0.628</td>
</tr>
<tr>
<td>GKDISHNGAYSGYGDQD</td>
<td>278</td>
<td>0.81</td>
<td>SYGQDDIVKY</td>
<td>288</td>
<td>0.613</td>
</tr>
<tr>
<td>GETWGGAYTDNYMTSR</td>
<td>8</td>
<td>0.92</td>
<td>YTDNYMTSR</td>
<td>15</td>
<td>0.599</td>
</tr>
<tr>
<td>TNDQQRDNGDRAES</td>
<td>89</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Start point relay on amino acid residues in proposed protein. Just motives with score more than 0.6 were listed. Reference score for ABCpred and Ellipro servers is 0/51 and 0/5.

Conformational and discontinuous epitopes of proposed fusion protein were predicted by Ellipro. This server acts based on homology results and determining the tertiary structure of protein in order to identify conformational epitopes (Table 3). Figures show conformational position of every epitope.

Table 3: Conformational epitopes predicted by Ellipro server.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length</th>
<th>Score</th>
<th>Tertiary structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>_Y1, _F2, _S3, _G4, _E5, _T6, _W7, _G8, _F31</td>
<td>9</td>
<td>0.82</td>
<td><img src="image_url" alt="Image" /></td>
</tr>
</tbody>
</table>
Just conformational epitopes with score more than 0.7 is shown. Figures show conformational position of every epitope. Reference score for Ellipro server is 0/5.

T cell receptors and MHC-I related epitopes were predicted by different programs as listed in table 1. MHC-I binding motives were predicted by EpiJen, NetCTL, CTLpred and TAPpred servers (Table 4). MHC-II binding motives were predicted by ProPred server (Table 5).

**Table 4: MHC-I and T-cell receptors binding motives predicted by different servers.**

<table>
<thead>
<tr>
<th>Number</th>
<th>sequence</th>
<th>Start point</th>
<th>score</th>
<th>Server</th>
</tr>
</thead>
<tbody>
<tr>
<td>14</td>
<td>_K129, _A130, _Q131, _D145, _L147, _D148, _V149, _Y150, _F183, _A184, _G185, _G186, _I187, _E188</td>
<td>14</td>
<td>0.754</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>_T226, _Y227, _Y259, _Q260, _F261, _D262, _F263, _G264, _L265, _R266</td>
<td>10</td>
<td>0.739</td>
<td></td>
</tr>
</tbody>
</table>
Table 5: MHC-II binding motives predicted by ProPred server.

<table>
<thead>
<tr>
<th>number</th>
<th>sequence</th>
<th>Start point</th>
<th>score</th>
<th>allele</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>YKGDNNTNGA</td>
<td>118</td>
<td>5.2</td>
<td>HLA-DRB1*0401</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.8</td>
<td>HLA-DRB1*0405</td>
</tr>
<tr>
<td>2</td>
<td>LEYQWTNNI</td>
<td>198</td>
<td>2.9</td>
<td>HLA-DRB1*0401</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.7</td>
<td>HLA-DRB1*0402</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.4</td>
<td>HLA-DRB1*0405</td>
</tr>
<tr>
<td>3</td>
<td>LLSSQTYNA</td>
<td>220</td>
<td>1.7</td>
<td>HLA-DRB1*0401</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3.0</td>
<td>HLA-DRB1*0402</td>
</tr>
<tr>
<td>4</td>
<td>YKAQGVQLT</td>
<td>127</td>
<td>1.6</td>
<td>HLA-DRB1*0401</td>
</tr>
<tr>
<td>5</td>
<td>LRPSVAYLQ</td>
<td>264</td>
<td>1.6</td>
<td>HLA-DRB1*0401</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.9</td>
<td>HLA-DRB1*0402</td>
</tr>
<tr>
<td>6</td>
<td>YQWTNNIGD</td>
<td>200</td>
<td>1.58</td>
<td>HLA-DRB1*0401</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2.6</td>
<td>HLA-DRB1*0405</td>
</tr>
<tr>
<td>7</td>
<td>MVWRADTKS</td>
<td>155</td>
<td>2.8</td>
<td>HLA-DRB1*0402</td>
</tr>
</tbody>
</table>

*Threshold for epitope identification in NetCTL is 0/75.

*Referenve score is 1.
Discussion:

Salmonella has become a serious threat in populations due to disease severity, the return and spreading manner of disease through the carriers without symptoms, the incidence of drug resistance and its application as a weapon in bioterrorism. Although the disease is rare in the developed industrial countries but is still considered as a serious problem in many developing countries, especially Southeast Asian countries, Latin America and Africa. In order to fight the disease and reduction of resulted mortality, many preventive measures are taken, the most important is vaccination\(^{[37]}\). There are two available vaccines against *Salmonella*, but these vaccines are not immunogenic in children under two years old. Due to some similar problems, design and development of new vaccines against *Salmonella* is recommended\(^{[38]}\).

In this study, we selected three outer membrane proteins (OMP A, F and C) for designing of a proposal multiepitopic protein with effective antigenic properties as we carried out in a similar work previously\(^{[39]}\). It has focused on bacterial outer membrane proteins as diagnostic markers and effective candidate antigenic protein vaccine. These proteins have a potential for subunit vaccine development against *Salmonella*. These proteins, which are abundant on the surface of the cell have been reported to induce a significant increase in antibody titers as compared with other bacteria structure proteins\(^{[22,40]}\). An immune response to OMPs has been reported in people who were vaccinated with an attenuated live strain as well as in the patients with natural typhoid fever\(^{[16]}\).

Studies on VaxiJen server showed that proposed fusion protein introduced in this study has high antigen probability compared with single candidate proteins. Antigenicity of proposed fusion protein was predicted 0.8896 which is a high value. Nowadays, the importance of in silico studies was emphasized in identifying the epitopes. B lymphocytes plays an important role in the formation of immune memory and remove pathogens by producing specific secretory antibodies. In Ellipro server, score higher than 0.5 and ABCpred, higher than 0.51 have potential in inducing humoral immunity. Ellipro and ABCpred servers were predicted 9 and 10 epitopes in designed protein with acceptable score, respectively. In Ellipro server, high score (0.828) belongs to NNIGDA and in ABCpred server, GETWGGAYTDNYMTSR had high score (0.92). However, score of NNIGDA is equivalent 0.87 in ABCpred. Also, conformational epitopes, _Y1_, _F2_, _S3_, _G4_, _E5_, _T6_, _W7_, _G8_, _F31 had higher score that is 0.82; exactly 7 amino acid residues of this motif is located on an extracellular loop of Omp F derived domain.

Due to the important role of CD4\(^+\) T cells in both innate and adaptive immune responses and involvement of CD8\(^+\) cells in destruction of viral infected cells,
identification of CD4\(^+\) and CD8\(^+\) T cells receptor related epitopes is essential for vaccine efficacy \[41\]. In this study, several servers were used to determine epitopes binding to MHC-I and II which are necessary for successful presentation of antigen to CD8\(^+\) and CD4\(^+\) cells. NetCTL server identified eight effective epitopes that higher score (1.9919) belongs to YTDNYMTSR. This motif was predicted in EpiJen server too. In ProPred server, two motives (YKAQGVQLT and LRPSVAYLQ) were predicted as the most characteristic epitopes (score: 1.6). LEYQWTNNI motif was predicted in all three alleles, including HLA-DRB1*0401, HLA-DRB1*0402 and HLA-DRB1*0405 with a score of 2.9, 2.7 and 2.4 respectively and can act as a very effective epitope for MHC-II. Our results support proposed fusion protein presented in this study as an effective multiepitopic vaccine against salmonella. However laboratorial examination and animal challenge are necessary for a confident comment on this proposed protein.

**Conclusion**

We have introduced a vaccine here which at the same time is capable of producing an immune response against some important structural proteins of S. typhi. In this paper, in addition to designing the structure and the exact identification of the relevant epitopes, the response rate of various immune cells is also measured. Laboratory studies and immunization of this vaccine are also underway in the laboratory and will be presented in the future.

**Acknowledgement:** We thanks Shahrekord University for its supports.

**Conflict of interest:** The authors declare that there is no conflict of interest regarding the publication of this manuscript.

**References**


3. Butler T, Treatment of typhoid fever in the 21st century: promises and


