Vaccine xxx (2018) xxx–xxx

Contents lists available at ScienceDirect

Vaccine

journal homepage: www.elsevier.com/locate/vaccine

Multi-epitope DnaK peptide vaccine against S.Typhi: An in silico approach

Shivani Verma [a], Ragumani Sugadev [a], Ashish Kumar [b], Sudhir Chandna [b], Lilly Ganju [a], Anju Bansal [a,*]

[a] Defence Institute of Physiology and Allied Sciences, Lucknow Road, Timarpur, Delhi 110054, India
[b] Institute of Nuclear Medicine and Allied Sciences, Lucknow Road, Timarpur, Delhi 110054, India

ARTICLE INFO

Article history:
Received 12 February 2018
Received in revised form 25 May 2018
Accepted 26 May 2018
Available online xxxx

Keywords:
Peptide vaccine
DnaK protein
Salmonella
Immunoinformatics

ABSTRACT

Salmonella is one of the key global causes of food and water borne enteric infections, responsible for significant morbidity and mortality worldwide especially in developing countries. Currently available vaccines against typhoid are moderately effective with several side effects and not efficacious against all *Salmonella* serovars. Due to limitations of these vaccines and emerging threats of multidrug resistance, developing an effective vaccine against these infections has increasingly become a priority. Heat shock proteins (Hsps), being evolutionarily conserved, represent dominant antigens in the host immune response. In continuation of our earlier studies on the development of S. Typhi DnaK and GroEL vaccine candidates, highly efficacious against *Salmonella* and multiple pathogens, in the present study, we have designed multi-epitope vaccine candidates common to multiple serovars of *Salmonella* using bioinformatics approach. Implementing various immunoinformatics tools such as IEDB, EpiJen, BCPRED, ElliPro and VaxiJen, led to the identification of many immunogenic B and T cell epitopes. The 3-D structure model of DnaK was generated to predict conformational B-cell epitopes using ElliPro server. Most promising T cell epitopes (29 CTLs, 18 T-helper cells) were selected based on their binding efficiency with commonly occurring MHC alleles. Finally we narrowed down to 5 protective antigenic peptides (PAPs), comprising highly conserved, antigenic and immunogenic B /T cell epitopes, least homologous with human host. These PAPs were predicted to be non-allergenic by allergenicity prediction tools (SORTALLER and AllerHunter). Hence, these immunogenic epitopes can be used for prophylactic or therapeutic usages specifically to defeat antibiotic-resistant *Salmonella.* These antigens have been reported for the first time and their conserved nature endow them as potential future vaccine candidates against other multiple pathogens as well.

1. Introduction

Inspite of the advances in food safety surveillance and regulations, the burden of foodborne diseases is considerable. Almost 550 million people fall ill each year due to diarrhoeal diseases, including 220 million children under the age of 5 years [1]. Increased burden of food and water borne infections has necessitated the development of medical countermeasures to protect against virulent pathogens that produce disease and death in humans, livestock, and crops [2]. *Salmonella* is 1 of the 4 key global causes of diarrhoeal diseases, ubiquitous and hardy bacteria that can survive several weeks in a dry environment and several months in water. *Salmonella* infection ranging from mild, self-limiting diarrhoea to serious gastrointestinal, septicemic disease and enteric fever (typhoid) is a global health problem for humans /animals and has been attributed to one of the most important bacterial etiologies for enteric infections worldwide. The major causes of the prevalence of this disease are related to poor food hygiene and inadequate sanitation. In developing countries, it is difficult to maintain good living standard by most of the population, accentuating the need for vaccine to prevent this infection.

A vaccine would be highly useful for tourists, pilgrims, military personnel and in general, due to their exposure to diverse environmental conditions. Increasing reports on the occurrence of *Salmonella* resistant to oral antibiotics like ampicillin, chloramphenicol, co-trimoxazole, quinolones and fluoroquinolones,

Abbreviations: PDB, Protein Data Bank; SSE, secondary structure element; ACC, accessibility.

* Corresponding author at: Experimental Biology Division, Defence Institute of Physiology and Allied Sciences, Lucknow Road, Timarpur, India.
E-mail address: anjubansal_dipas@rediffmail.com (A. Bansal).

https://doi.org/10.1016/j.vaccine.2018.05.106
0264-410X/

decreased susceptibility to ciprofloxacin [3] and non-availability of relevant drugs are of significant public health concern, underscoring the urgent need to develop an effective vaccine against *Salmonella*.

Pathogen-derived Heat Shock Proteins (Hsps) serve as important antigens in defense against numerous infectious agents [4]. Hsps are highly conserved family of proteins, act as a sort of link between innate and adaptive immune responses [5]. These proteins undertake crucial functions in maintaining basic cellular processes to preserve cell viability and homeostasis, since they act as chaperons and mediate a range of cytoprotective and housekeeping functions [6]. DnaK, bacterial homolog of human Hsp70 has been characterized in several pathogenic bacteria and seems to have important functions in stress resistance and pathogenicity in multiple-drug-resistant bacteria [7]. Being a conserved protein and having a pivotal role in pathogenic microbes, bacterial DnaK protein is of particular interest in vaccine discovery. Conventional vaccines comprising of whole organism or large proteins incorporate unnecessary antigenic load and complicate the situation by inducing allergenic response. To circumvent this, the peptide based vaccine is an attractive alternative strategy as it relies on the usage of short immunogenic peptide fragments that induce highly targeted immune responses, consequently avoiding allergenic response. The explosion in knowledge of the sequence and structure of proteins from pathogenic organisms and integration of this information with informatics tools, aid in the development of effective vaccines [8–10]. Peptide vaccine design based on computational paradigm facilitates the identification of highly immunogenic epitopes within protein antigen. Prediction of antigenic epitopes has long been the focus of immunoinformatics and several tools have been developed given the potential translational implications [11,12]. Immunoinformatics approach overcomes the hurdles of cost, time duration, and accuracy associated with traditional vaccinology and has been applied successfully in vaccine designing [13,14].

Earlier we have developed recombinant Hsp60/70 based candidate vaccine molecules and found these to be highly effective against S. Typhi and *S* Typhimurium in mouse model [15]. As peptide based vaccines are more specific and easy to produce, present study aims to identify major immunodominant peptides of *S*. Typhi DnaK protein, sharing both B cell and T cell epitopes, non-allergenic, showing least sequence similarity with human proteome, and covering worldwide population using in silico studies for the development of vaccine candidates against *Salmonella* infection/typhoid. We strongly believe that the outcome of study will provide potential candidate vaccine molecules that provide protection against multiple enteric infections.

2. Material and methods

The designed framework used in the current study for peptide-based candidate vaccine molecules mining within the DnaK protein of S. Typhi is shown in Fig. 1.

2.1. Protein sequence retrieval

The sequence of DnaK (Reference sequence NP_454622.1, UniProt Accession no. Q8Z9R1) of *S*. Typhi was retrieved from the Universal Protein Resource (UniProt) at http://www.uniprot.org (Consortium, 2015) saved in FASTA format.

2.2. Determination of conserved region

A preliminary analysis of DnaK protein sequences of *S*.Typhi was performed to evaluate the reliability of this protein sequence as representative for all the possible serovars of bacterium *S*.Typhi using protein Basic Local Alignment Search Tool (BLASTp). BLAST search was performed within non redundant protein sequences (nr) database of bacteria using blosum62 matrix. The retrieved sequences were aligned by multiple sequence alignment (MSA) using BLAST software [16], to obtain the conserved regions.

2.3. Sequence based B-cell epitope prediction

B-cell candidate epitopes were analyzed by several B-cell prediction methods from IEDB including Chou & Fasman Beta-Turn Prediction [17], Emini Surface Accessibility Prediction [18], Karplus & Schulz Flexibility Prediction [19], Kolaskar & Tongaonkar Antigenicity [20] and Parker Hydrophilicity Prediction [21] with default window and threshold values. The linear predicted epitopes were also obtained by using BCPRED server (http://ailab.ist.psu.edu/bcpred/predict.html) using three methods namely fixed length AAP [22], BCPred and Flexible length method (FBCPred) [23] with the threshold >1, >0.9 and >0.9 respectively.

2.4. Structure based B-cell epitope prediction

Discontinuous B-cell epitope prediction was carried out using the 3D structure of DnaK protein sequence accession no. NP_454622.1. As 3D structure of *S*. Typhi DnaK protein is not available in protein data bank and is prerequisite for prediction of conformational epitopes, the same was generated by SWISS homology MODEL [24]. DnaK protein of E. coli PDB ID 2kho.1 was used as the template to build our model. The validation of the predicted structure was done using QMEAN analysis [25].The structure based epitopes were then obtained using ElliPro [26] with the threshold score >0.5.

2.5. Cytotoxic T lymphocyte epitope predictions

For T-cell epitope predictions, predominantly occurring MHC (major histocompatibility complex) alleles in human population for HLA class I (A*01:01, A*02:01, A*03:01, A*11:01, A*24:02, HLA-A*33:03B*07:02, B*08:01 HLA-B*35:0, HLA-B*40, HLA-B*44) and HLA class II (DRB1*01:01, DRB1*03:01, DRB1*04:01, DRB1*07:01, DRB1*08:02, DRB1*11:01, DRB1*13:02, DRB1*15:01, HLA-DQA1*05:01|DQB1*02:01, HLA-DPA1*02:01|DPB1*01:01, HLA-DQA1*01:02|DQB1*06:02, HLA-DQA1*03:01|DQB1*03:02) were selected [27,28]. The peptide length was set at 9 amino acids prior to the prediction by MHC Class I epitope binding analysis tools of IEDB. The promising MHC I T-cell epitopes were predicted on the basis of IC50 values (<500 nM) of peptides. EpiJen v1.0 tool was also employed to predict T-cell epitopes which consider several important stages of the MHC degradation pathway: proteasome cleavage, TAP binding, MHC binding and epitope selection.

2.6. Helper T lymphocyte epitope predictions

To predict MHC II binding epitopes by IEDB prediction tool (http://tools.immuneepitope.org/mhcii/), peptide length was set at 15 amino acids with above mentioned selected alleles. Among the different MHCII prediction methods, we used IC50 value (<500 nM)) to narrow down to the most promising MHC II epitopes.

2.7. Population coverage of selected epitopes

Population coverage analysis was performed by IEDB population coverage tool of the selected T cell epitopes with their respective HLA alleles, to predict the coverage within South Asian and Worldwide population.

Please cite this article in press as: Verma S et al. Multi-epitope DnaK peptide vaccine against S.Typhi: An in silico approach. Vaccine (2018), https://doi.org/10.1016/j.vaccine.2018.05.106