BIOENGINEERED, 2018
VOL. 9, NO. 1, 170–177
https://doi.org/10.1080/21655979.2017.1373535

Taylor & Francis
Taylor & Francis Group

RESEARCH PAPER

OPEN ACCESS Check for updates

In silico design of a novel chimeric *shigella* IpaB fused to C terminal of *clostridium perfringens* enterotoxin as a vaccine candidate

Sina Arabshahi[a], Bahar Nayeri Fasaei[b], Abdollah Derakhshandeh[a] and Aytak Novinrooz[a]

[a]Department of Pathobiology, School of Veterinary Medicine, Shiraz University, Shiraz, Iran; [b]Department of Microbiology and Immunology, Faculty of Veterinary Medicine, University of Tehran, Tehran, Iran

ABSTRACT

This study aimed to design a novel chimeric protein *in silico* to serve as a serotype-independent vaccine candidate against *Shigella*. The chimera contains amino acid residues 240–460 of *Shigella* invasion plasmid antigen B (IpaB) and the C-terminus of *Clostridium perfringens* enterotoxin (C-CPE). Amino acid sequences of 537 peptide linkers were obtained from two protein linker databases. 3D structures of IpaB-$CPE_{290–319}$, IpaB-$CPE_{184–319}$, IpaB-$CPE_{194–319}$ and 537 newly designed IpaB-linker-$CPE_{290–319}$ constructs with varying linker regions were predicted. These predicted 3D structures were merged with the 3D structures of native $IpaB_{240–460}$, $CPE_{194–319}$, $CPE_{184–319}$ and $CPE_{290–319}$ to select the structure most similar to native IpaB and C-CPE. Several *in silico* tools were used to determine the suitability of the selected IpaB-C-CPE structure as a vaccine candidate. None of the 537 linkers was capable of preserving the native structure of $CPE_{290–319}$ within the IpaB-linker-$CPE_{290–319}$ structure. *In silico* analysis determined that the IpaB-$CPE_{194–319}$ 3D structure was the most similar to the 3D structure of the respective native CPE domain and that it was a stable chimeric protein exposing multiple B-cell epitopes. IpaB-$CPE_{194–319}$ was designed for its capability to bind to human intestinal epithelial and M cells and to accumulate on these cells. The predicted B-cell epitopes are likely to be capable of inducing a mucosal antibody response in the human intestine against *Shigella* IpaB. This study also showed that the higher binding affinities of $CPE_{184–319}$ and $CPE_{194–319}$ to claudin molecules than those of $CPE_{290–319}$ is the result of preserving the 3D structures of $CPE_{184–319}$ and $CPE_{194–319}$ when they are linked to the C-termini of other proteins.

ARTICLE HISTORY
Received 19 May 2017
Revised 28 June 2017
Accepted 30 June 2017

KEYWORDS
Shigella; CPE; IpaB; Chimeric protein; Vaccine candidate

Introduction

Shigellosis, a disease caused by *Shigella*, is endemic in developing countries due to a lack of hygienic drinking water sources, malnutrition and the high cost of antibiotic therapy. There are more than 40 *Shigella* serotypes, which are classified based on O antigen polysaccharide diversity. *S. flexneri* and *S. dysentery* are the causative agents of endemic and epidemic dysentery, particularly in children under 5 y old in developing countries. Shigellosis often causes severe symptoms and may lead to death.[1–4] An estimated 80–160 million annual cases of dysentery occur worldwide, which lead to approximately one million deaths annually.[5] To date, no effective vaccine prevents the disease caused by *Shigella*. In addition, the rapid and broad global spread of antibiotic resistance has been reported among different serotypes of *Shigella*. Thus, the World Health Organization (WHO) has prioritized the development and production of an efficient vaccine to globally protect people against *Shigella*. Immunogenic *Shigella* surface proteins that are conserved among different serotypes are thought to be capable of inducing broad-spectrum protection against multiple *Shigella* serotypes.[4,6,7] Invasion plasmid antigen B (IpaB) is a *Shigella* surface protein located at the very top of its type three secretory system (TTSS), and it is a key virulence factor of *Shigella*.[8] This protein alone may not be capable of inducing a protective immune response against *Shigella*, but it has been recently used, along with other *Shigella* proteins, to generate effective vaccine candidates.[4,9–11]

CONTACT Bahar Nayeri Fasaei nayerib@ut.ac.ir Azadi Avenue, Faculty of Veterinary Medicine, Department of Microbiology and Immunology, University of Tehran, Post Box: 14155-6453, Tehran, Iran.
Supplemental data for this article can be accessed on the publisher's website.

Since 2012, the combinations of IpaB/IpaD and IpaB/GroEL have been reported aspotential vaccine candidates capable of inducing protection against experimental *Shigella* infections in mice.[7,11,12]

Clostridium perfringens enterotoxin (CPE) consists of two functional domains: a cytotoxic N-terminal domain and a non-toxic C-terminal domain (C-CPE). The C-terminal domain of CPE is known to be capable of binding to claudin-3 and claudin-4 molecules at the surface of some mammalian cells. The strong attachment of the C-CPE to claudin-3 and -4 on human intestinal epithelial cells causes this peptide to aggregate on the luminal sides of intestinal epithelial cells, allowing it to act as a biological and mucosal adjuvant for orally administered drugs and vaccines. Previous studies have indicated that these attachments eventually lead to the transfer of the C-CPE, and likely any protein fused to it, into the cytoplasm of intestinal epithelial cells through receptor mediated endocytosis.[13]

Claudin-3 and -4 molecules belong to the claudin receptor molecule family.[14,15] Claudin-3 is localized on the apical surfaces of villi and cryptic epithelial cells of the human small and large intestines. Moreover, these molecules are localized on the entire apical surfaces of follicle associated epithelia (FAEs) in the human intestinal tract. However, claudin-4, another receptor for the C-CPE, is found more frequently on the basolateral surfaces of epithelial cells in the intestinal villi. Claudin-4 is also found on the apical surfaces of FAEs in Peyer's patches and on the apical surfaces of cecal patches in the human large intestine.[16]

The aim of the current study is to design a chimeric IpaB-C-CPE protein for use in the development of a vaccine to combat *Shigella* infections. To the best of our knowledge, this report is the first to design a chimeric IpaB-C-CPE for any biological or medical purpose.

Results

The amino acid sequences of 537 chimeric IpaB-linker-$CPE_{290–319}$ variants generated using 537 different peptide linkers to connect $IpaB_{240–460}$ to $CPE_{290–319}$ were obtained using an in silico sequence conversion tool. Furthermore, the amino acid sequences of IpaB-$CPE_{184–319}$ and IpaB-$CPE_{194–319}$ were obtained successfully using the abovementioned software. All the obtained amino acid sequences were analyzed by phyre2 software, and 3D structures of the recombinant proteins were downloaded as 539 PBD files. The 3D structures of 537 chimeric proteins were compared with the 3D structures of native $IpaB_{240–460}$ and $CPE_{290–319}$ using TM-align software, and the rate of similarity index between these proteins and recombinant IpaB-linker-$CPE_{290–319}$ was calculated by measuring the TM-score. The obtained results showed that all 537 linkers had impaired $CPE_{290–319}$ and $IpaB_{240–460}$ 3D structures. Using TM-align software from the Zhang lab, the 3D structures of $CPE_{184–319}$ and $CPE_{194–319}$ were found to exhibit considerable similarity to their corresponding domains in IpaB-$CPE_{184–319}$ (TM-score: 0.97) and IpaB-$CPE_{194–319}$ (TM-score: 0.99), respectively (Fig. 1).

IpaB-$CPE_{194–319}$ was predicted to be a stable 38.5 KDa protein with an estimated half-life of more than 10 hours in *E. coli*. The aliphatic and grand average of hydropathicity (GRAVY) indices of IpaB-$CPE_{194–319}$ were predicted by ExPASy ProtParam online software to be 109.66 and 0.117, respectively. The results obtained from IpaB-$CPE_{194–319}$ secondary structure analysis revealed a high percentage of alpha helices and low percentages of beta strands and disordered regions. In addition to these secondary structures, a moderate percentage of transmembrane helices was predicted (residues 72–184).

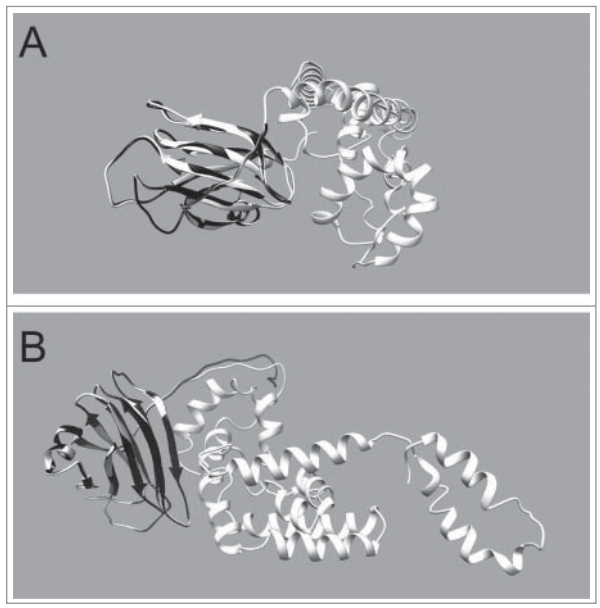

Figure 1. Protein 3D structure alignments generated by the UCSF Chimera package.[38] (A) Structural alignment of IpaB-$CPE_{194–319}$ (in white) with native $CPE_{194–319}$ (in black). (B) 3D structure of IpaB-$CPE_{184–319}$ (in white) aligned with the 3D structure of native $CPE_{184–319}$ (in black).

Table 1. IpaB-CPE$_{194–319}$ B-cell conformational epitopes predicted by the DiscoTope server. In total, 28 epitope residues were identified within IpaB-CPE$_{194–319}$.

Epitope position	Epitope sequence	Contact number	DiscoTope score
1–4	MQKS	1, 8, 6, 0	−0.257, −1.600, −2.288, −2.433
25	N	0	−2.840
45	K	11	−3.585
47	E	11	−2.467
152	G	0	−2.651
154	D	5	−3.627
217–218	SQ	5, 5	−2.878, −3.606
220–224	DDLDI	6, 4, 9, 0, 3	−2.641, −1.886, −3.211, −1.327, −2.379
248–249	PA	0, 9	−1.307, −2.740
262	P	5	−3.416
309–312	DGVK	5, 4, 0, 4	−2.702, −1.851, −1.469, −2.795
334–338	SYSGN	1, 20, 0, 5, 9	−1.462, −2.922, 0.607, −0.292, −1.866

In total, 150 B-cell lineal epitopes (5–20 amino acids) exhibiting FBCpred scores greater than 0.7 were predicted within the IpaB-CPE$_{194–319}$ structure (Supplementary Table 1). Among all 150 lineal epitopes within the chimeric protein, 62 B-cell epitopes fit within residues 1–221, which correspond to the IpaB$_{240–460}$ structure. Moreover, 28 B-cell conformational epitope residues were predicted by the DiscoTope server within full-length IpaB-CPE$_{194–319}$ (Table 1). Among all predicted conformational epitope residues, 13 epitope residues were predicted within residues 1–221.

Codon optimization of *ipaB-c-cpe* (encoding IpaB-CPE$_{194–319}$) was performed using OPTIMIZER according to an *E. coli* codon usage table provided by the Kazusa codon usage database. The codon adaptation indices (CAIs) of the codon-optimized and non-optimized recombinant genes were analyzed by both the OPTIMIZER and GenScript rare codon analysis tools. The results showed a significant increase in the CAI following codon optimization. The CAI value determined by OPTIMIZER following codon optimization was 1, whereas the CAI values calculated by the GenScript rare codon analysis tool were 0.62 and 0.94 prior to and following codon optimization, respectively (Fig. 2).

mRNA secondary structure prediction of ipaB-c-cpe showed that the most stable mRNA structure had a ΔG equal to −311.2 kcal/mol, and no stable hairpins or pseudo knots were observed in this predicted structure (Fig. 3).

Discussion

Currently, there is no available licensed vaccine against *Shigella*; however, several ongoing studies are focusing on *Shigella* surface proteins that are common among different serotypes of this microorganism as potential vaccine candidates.[4,5,7,12] IpaB, one of these surface proteins, has been shown to be capable of inducing an antibody response in humans. Moreover, IpaB-specific antibodies have been shown to be negatively correlated with shigellosis severity in humans.[17] Recently, a *Lactococcus lactis* bacterium-like particles (BLP) vaccine candidate harboring IpaB and IpaD was successfully designed and used in adult and newborn mouse studies. Protective immune responses against *S. flexneri* and *S. sonnei* were generated in 90% and 80% of the adult mice and 90% and 44% of newborn mice, respectively, following immunization with BLP-IpaB/D.[18] IpaB fused to GroEL (IpaB-GroEL) has been investigated in a more recent study as a vaccine candidate against shigellosis. Interestingly, the immunization of mice with chimeric IpaB-GroEL resulted in the protection of 90–95% of mice following lethal challenges with *S. flexneri, S. boydii* and *S. sonnei*.[19] The present

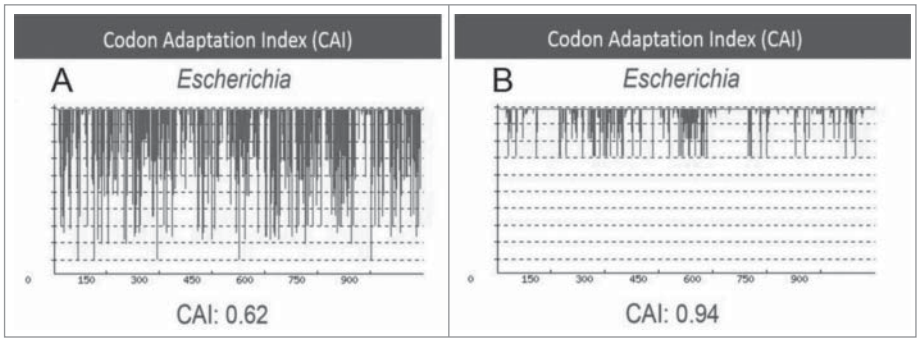

Figure 2. Codon adaptation index (CAI) values of *ipaB-c-cpe*. (A) A CAI of 0.62 was calculated by the GenScript rare codon analysis tool before codon optimization. (B) A CAI of 0.94 was calculated by the GenScript rare codon analysis tool after codon optimization.

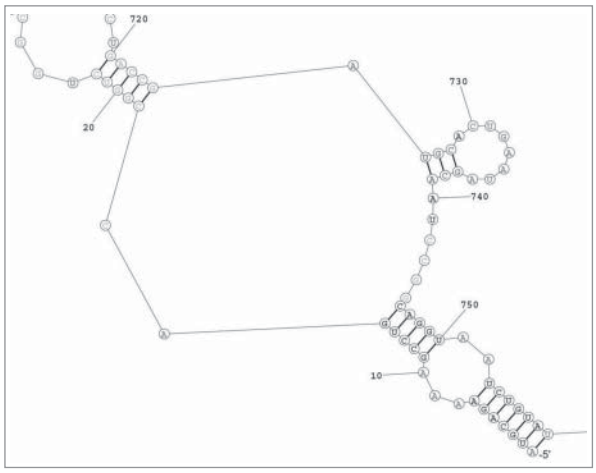

Figure 3. Prediction of *ipaB-c-cpe* mRNA secondary structure. No long stable hairpins or pseudo knots were observed at the 5′ of the predicted structure.

study focuses on the design of a novel chimeric protein consisting of a variant of IpaB, $IpaB_{240–460}$, which is an immunogenic surface protein conserved among different serotypes of *Shigella*, fused to the C-terminus of the *Clostridium perfringens* enterotoxin.

Residues 240–280 of IpaB are involved in protein stability and folding, while residues 313–333 and 399–419 form two large globular domains located at the Tip Complex (TC) of the TTSS that are likely responsible for any contact of IpaB with the host cell as well as the penetration of this protein into the host cell membrane.[20,21] Our designed chimeric protein contains amino acid residues from IpaB, as well as their respective domains, that are mostly in contact with host cells and the environment and are essential for the initiation of *Shigella* invasion into epithelial cells. Therefore, the neutralization of these structures and domains by the induction of mucosal immunity likely leads to the inhibition of *Shigella* invasion and the prevention of the disease caused by these bacteria.

The presence of the C-CPE in the structure of our designed chimeric protein makes this chimeric protein better able to induce a mucosal immune response, as the C-CPE exhibits adjuvant properties when administered on mucosal surfaces.[13,22] Several C-CPE variants, such as $CPE_{168–319}$, $CPE_{171–319}$, $CPE_{184–319}$, $CPE_{194–319}$ and $CPE_{290–319}$, have been shown to be capable of binding to claudin family proteins.[15,23] Among these variants, $CPE_{290–319}$ is the most frequently studied. This variant is sufficient for binding to some claudin proteins with an affinity comparable to that of native CPE.[24] Thus, the first chimeric protein designed in this study was IpaB-$CPE_{290–319}$.

To preserve the native conformations and functions of $IpaB_{240–460}$ and $CPE_{290–319}$ in the structure of our first designed chimeric protein, different peptide linkers were placed between $IpaB_{240–460}$ and $CPE_{290–319}$. Then, the 3D structures of 537 recombinant IpaB-linker-$CPE_{290–319}$ proteins differing in their peptide linkers were analyzed by TM-align software to determine the linker sequence resulting in the highest TM-score. The TM-score varies between 0 and 1. A TM-score of 1 indicates perfect similarity between the 3D structures of two proteins. As TM-align software compares uploaded protein data bank (PDB) files to the 3D structures of different proteins and peptides in several protein databases, and there is no 3D structure related to full-length IpaB or the C-terminus of IpaB, TM-align software was unable to determine any valid TM-scores for the comparisons of the IpaB-linker-$CPE_{290–319}$ variants to $IpaB_{240–460}$. However, we successfully used this software to compare our designed protein with $CPE_{290–319}$. The similarity between IpaB-linker-$CPE_{290–319}$ and $CPE_{290–319}$ 3D structures was determined to be random, with a TM-score of 0.39, and all the analyzed linkers were determined to cause significant disorganization in the 3D structure of $CPE_{290–319}$.

In addition to the efforts described above, the 3D structures of IpaB-$CPE_{290–319}$, IpaB-$CPE_{184–319}$ and IpaB-$CPE_{194–319}$ were compared to the 3D structures of $CPE_{290–319}$, $CPE_{184–319}$ and $CPE_{194–319}$, respectively. The highest TM-score, 0.99, was achieved for the alignment of the IpaB-$CPE_{194–319}$ and $CPE_{194–319}$ 3D structures, showing a high degree of similarity between these two 3D structures. Bioinformatic analysis also showed excellent similarity between the 3D structures of IpaB-$CPE_{184–319}$ and $CPE_{184–319}$, with a TM-score of 0.97.

Both $CPE_{194–319}$ and $CPE_{184–319}$ have been shown to be capable of high-affinity specific binding to claudin-3 and -4 at the tight junctions of intestinal epithelial cells without inducing cytotoxicity.[23,25] Although $CPE_{290–319}$ is known as the claudin-binding domain of the CPE, this variant of the C-CPE is not able to bind claudin-4 without the presence of its N-terminal domain. However, $CPE_{184–319}$ and $CPE_{194–319}$, containing longer N-terminal regions than $CPE_{290–319}$, are capable of binding to claudin-4 more strongly, likely

because they contain more of the functional domains required to bind their receptor molecules and claudins.[23,26] According to previous results and the *in silico* analysis performed in the current study, the N-terminal domains of $CPE_{184–319}$ or $CPE_{194–319}$ are likely involved in the folding of these proteins, as they are necessary for the formation of a strong attachment to claudin-4.

Based on our findings, many natural and synthetic peptide linkers with different physical and chemical properties were unable to provide the physicochemical interactions necessary to correctly fold $CPE_{290–319}$, which is a requirement for it to bind claudin-4 molecules. Thus, although $CPE_{290–319}$ has been frequently used as a natural ligand for the attachment of other molecules to targeted cells,[13,27,28] The results of this study indicate that $CPE_{194–319}$ and $CPE_{184–319}$ are better suited to this purpose. Both of these C-CPE variants have been successfully tested as absorption enhancers for delivering biologically active peptides to pulmonary, nasal and jejunal tissues. $CPE_{184–319}$ is capable of enhancing the mucosal absorption of dextran, a model drug, approximately 400–fold more than sodium caprate, a common clinical absorption enhancer. $CPE_{194–319}$ was shown to have similar absorption enhancing properties, while $CPE_{289–319}$ was shown to have no enhancing effect on the absorption of dextran. However, the low solubility of $CPE_{184–319}$ limits its application as a mucosal absorbent. The solubility of $CPE_{194–319}$ has been shown to be 30–fold greater than that of $CPE_{184–319}$, making it more suitable for use as a mucosal absorbent than $CPE_{184–319}$.[23,29,30]

The physical and chemical properties of chimeric IpaB-$CPE_{194–319}$ were predicted by the ExPASy ProtParam server. The isoelectric point (pI) value was predicted to be lower than 7, which indicates the intrinsic acidity of the designed protein. The approximate low molecular weight of the designed protein decreases the probability of inclusion body formation consequent to its expression in *E. coli*. In addition, its estimated half-life of more than 10 hours in *E. coli* suggests an enhanced probability of the large-scale production and purification of this protein in *E. coli* without significant loss of its structure and activity. The high aliphatic index of chimeric IpaB-$CPE_{194–319}$ shows the high stability of this protein across a wide range of temperatures. The positive predicted GRAVY value indicates the hydrophobic nature of this designed protein.

As a vaccine candidate against *Shigella* infections, IpaB-$CPE_{194–319}$ is designed to induce mucosal antibody production to prevent IpaB-mediated *Shigella* invasion. The B-cell epitope prediction conducted in this study determined 150 lineal B-cell epitopes (5–20 amino acids in length) and 28 conformational B-cell epitope residues in the chimeric IpaB-$CPE_{194–319}$. Among these, 62 lineal B-cell epitopes and 13 conformational B-cell epitope residues fit within the IpaB domain of the IpaB-$CPE_{194–319}$ chimeric protein. According to these B-cell epitope prediction results, we assume that IpaB-$CPE_{194–319}$ is theoretically capable of inducing a mucosal antibody response in the human intestinal tract against several IpaB epitopes. The abundance of predicted IpaB epitopes enhances the probability that chimeric IpaB-$CPE_{194–319}$ can induce a protective humoral immune response.

Codon optimization of the recombinant protein was conducted by OPTIMIZER to enhance its further expression levels in *E. coli*. To calculate the CAI of the designed *ipaB-c-cpe* gene, two web-based programs, OPTIMIZER and the GenScript rare codon analysis tool, were used. The CAI is associated with the expression level of the desired protein and varies between 0 and 1. Genes with CAI values close to 1.0 and G/C contents between 30% and 70% are predicted to be ideally expressed. The CAI value of non-optimized *ipaB-c-cpe* was calculated to be 0.62 by the GenScript rare codon analysis tool. The values for the codon-optimized *ipaB-c-cpe* were 1.00 and 0.94, as calculated by OPTIMIZER and the GenScript rare codon analysis tool, respectively. The G/C contents of *ipaB-c-cpe* were also optimized to 51.61% and 42.95%, respectively, using OPTIMIZER. The CAI and G/C content values of the optimized genes indicate an enhanced probability of the high-yield expression of IpaB-$CPE_{194–319}$ in *E. coli*.

Its predicted mRNA secondary structure indicates that the *ipaB-c-cpe* gene will be transcribed into a stable mRNA. The predicted mRNA structure does not include any translation inhibitory structures such as pseudo knots or long stable hairpins. Therefore, the predicted mRNA is theoretically capable of being translated into IpaB-$CPE_{194–319}$ in *E. coli* cells.

This study describes the successful *in silico* design of a novel IpaB-$CPE_{194–319}$ chimeric protein that is a potential vaccine candidate against *Shigella* infections. The obtained results show that, in comparison to $CPE_{290–319}$, two other variants of the C-CPE, $CPE_{184–319}$ and $CPE_{194–319}$, are better able to preserve their natural 3D structures and to form stronger attachments to their natural ligands when they are linked to the C-terminal regions of other proteins. This ability may be a consequence of extra domains (184–289 and 194–289) that are involved in correct folding by CPE, $CPE_{184–319}$ and $CPE_{194–319}$. However, more in-depth studies are required to validate the results obtained in this study from the *in silico* analyses of different chimeric proteins containing different variants of the C-CPE. Furthermore, *in vitro* expression and purification of the designed IpaB-$CPE_{194–319}$, followed by *in vitro* and *in vivo* evaluation of the immune response to this protein are proposed to determine the potential of this newly designed protein as a vaccine against shigellosis.

Materials and methods

Sequence analysis

To design the IpaB-C-CPE encoding gene, nucleotides 718–1380 of the full-length *ipaB* gene (NCBI accession number NC_004851.1) and a 5′ fragment of the *cpe* gene (NCBI accession number NC_008262.1) were used. The 5′ fragment of the *cpe* gene was chosen from the sequences encoding $CPE_{290–319}$, $CPE_{194–319}$ and $CPE_{184–319}$ as is described in the construct design section. All amino acid sequences were obtained by converting their respective nucleotide sequences using an in silico sequence conversion tool (http://in-silico.net/tools/biology/sequence_conversion).

Construct design

To construct a chimeric protein resembling the native folding and function of the respective regions of IpaB and the C-CPE, 537 derivatives of chimeric IpaB-linker-$CPE_{290–319}$ constructs varying in their linker moieties were analyzed *in silico*. The amino acid sequences of all 537 natural and synthetic linkers were adapted from Klein et al., 2014 (for a review, see ref. 31) and two protein linker databases [(http://www.ibi.vu.nl/programs/linkerdbwww) and (http://parts.igem.org/Protein_domains/Linker)] (Supplementary Table 2). In addition to these 537 derivatives of IpaB-linker-$CPE_{290–319}$, three other variants of IpaB-C-CPE were designed: IpaB-$CPE_{290–319}$, IpaB-$CPE_{184–319}$ and IpaB-$CPE_{194–319}$.

3D structure analysis

Three dimensional structures of 537 different IpaB-linker-$CPE_{290–319}$ variants were obtained as PDB files by introducing the amino acid sequences of all 537 chimeric proteins to the Protein Homology/analogY Recognition Engine (Phyre2) (http://www.sbg.bio.ic.ac.uk/~phyre2/html/page.cgi?id=index).[32] In addition, PDB files containing 3D structures of $IpaB_{240–460}$, $CPE_{184–319}$, $CPE_{194–319}$, $CPE_{290–319}$, IpaB-$CPE_{290–319}$, IpaB-$CPE_{184–319}$ and IpaB-$CPE_{194–319}$ without any linkers were created. Afterward, residue to residue alignments of the 3D structures of the chimeric proteins with native $IpaB_{240–460}$ and $CPE_{290–319}$ were performed using TM-align to analyze the similarity of the 3D structures of the relevant proteins based on their TM-scores (http://zhanglab.ccmb.med.umich.edu/TM-align).[33] The 3D structures of IpaB-$CPE_{184–319}$ and IpaB-$CPE_{194–319}$ were also compared to $CPE_{184–319}$ and $CPE_{194–319}$, respectively. The chimeric protein most closely resembling the native structure of the C-CPE was selected for further analysis.

Physical and chemical parameters and secondary structure

The physical and chemical properties of IpaB-$CPE_{194–319}$, including its theoretical pI, estimated half-life, instability index, aliphatic index and GRAVY, were predicted using ExPASy ProtParam online software (http://web.expasy.org/protparam/). Secondary structures and probable transmembrane helices of IpaB-$CPE_{194–319}$ were predicted using Phyre2 web-based software.

B-cell epitope prediction

IpaB-$CPE_{194–319}$ continuous/lineal B-cell epitopes (5–20 amino acids) were predicted through flexible length epitope prediction (FBCpred) using the BCPREDS server (http://ailab.ist.psu.edu/bcpred/predict.html).[34] IpaB-$CPE_{194–319}$ discontinuous/conformational B-cell epitopes were analyzed by uploading the 3D structure

of chimera to the Disco Tope server (http://www.cbs.dtu.dk/services/DiscoTope/).[35]

Codon optimization

The amino acid sequence of IpaB-CPE$_{194–319}$ was then converted back to a nucleotide sequence using an in silico sequence conversion tool (http://in-silico.net/tools/biology/sequence_conversion). The nucleotide sequence of recombinant *ipaB-c-cpe* was codon optimized to be expressed in *E. coli* using OPTIMIZER (http://genomes.urv.es/OPTIMIZER).[36] An *E. coli* codon usage table was adapted from a codon usage database (www.kazusa.or.jp/codon). The CAI, which is directly related to the expression level of the desired gene, was analyzed by both OPTIMIZER and the GenScript rare codon analysis tool (http://www.genscript.com/cgibin/tools/rare_codon_analysis) to determine the expression efficiency of the recombinant gene in *E. coli*.

mRNA secondary structure

The secondary structure of *ipaB-c-cpe* messenger RNA was predicted using the RNAstructure Web Server (http://rna.urmc.rochester.edu/RNAstructureWeb), and the presence of hairpins and pseudo knots at the 5′ end of the mRNA were analyzed.[37]

Disclosure of potential conflicts of interest

The authors claim no conflict of interest to declare.

Acknowledgments

The authors would like to thank Mrs. B. Arabshahi for her helpful assistance during the project.

References

1. Heine SJ, Diaz-McNair J, Martinez-Becerra FJ, Choudhari SP, Clements JD, Picking WL, Pasetti MF. Evaluation of immunogenicity and protective efficacy of orally delivered shigella type iii secretion system proteins ipab and ipad. Vaccine. 2013;31(28):2919–2929. doi:10.1016/j.vaccine.2013.04.045.
2. Dharmasena MN, Hanisch BW, Wai TT, Kopecko DJ. Stable expression of shigella sonnei form i o-polysaccharide genes recombineered into the chromosome of live salmonella oral vaccine vector ty21a. Int J Med Microbiol. 2013;303(3):105–113. doi:10.1016/j.ijmm.2013.01.001.
3. Carayol N, Tran Van Nhieu G. Tips and tricks about shigella invasion of epithelial cells. Curr Opin Microbiol. 2013;16(1):32–37. doi:10.1016/j.mib.2012.11.010.
4. Carayol N, Tran Van Nhieu G. The inside story of shigella invasion of intestinal epithelial cells. Cold Spring Harb Perspect Med. 2013;3(10):a016717. doi:10.1101/cshperspect.a016717.
5. Jennison AV, Verma NK. Shigella flexneri infection: Pathogenesis and vaccine development. FEMS Microbiol Rev. 2004;28(1):43–58. doi:10.1016/j.femsre.2003.07.002.
6. Rahman MZ, Akter S, Azmuda N, Sultana M, Weill FX, Khan SI, Grimont PA, Birkeland NK. Serological cross-reaction between o-antigens of shigella dysenteriae type 4 and an environmental escherichia albertii isolate. Curr Microbiol. 2013;67(5):590–595. doi:10.1007/s00284-013-0405-7.
7. Martinez-Becerra FJ, Chen X, Dickenson NE, Choudhari SP, Harrison K, Clements JD, Picking WD, Van De Verg LL, Walker RI, Picking WL. Characterization of a novel fusion protein from ipab and ipad of shigella spp. And its potential as a pan-shigella vaccine. Infect Immun. 2013;81(12):4470–4477. doi:10.1128/IAI.00859-13.
8. Dickenson NE, Zhang L, Epler CR, Adam PR, Picking WL, Picking WD. Conformational changes in ipad from shigella flexneri upon binding bile salts provide insight into the second step of type iii secretion. Biochemistry. 2011;50(2):172–180. doi:10.1021/bi101365f.
9. Oaks EV, Turbyfill KR. Development and evaluation of a shigella flexneri 2a and s. Sonnei bivalent invasin complex (invaplex) vaccine. Vaccine. 2006;24(13):2290–2301. doi:10.1016/j.vaccine.2005.11.040.
10. Riddle MS, Kaminski RW, Williams C, Porter C, Baqar S, Kordis A, Gilliland T, Lapa J, Coughlin M, Soltis C, et al. Safety and immunogenicity of an intranasal shigella flexneri 2a invaplex 50 vaccine. Vaccine. 2011;29(40):7009–7019. doi:10.1016/j.vaccine.2011.07.033.
11. Martinez-Becerra FJ, Kissmann JM, Diaz-McNair J, Choudhari SP, Quick AM, Mellado-Sanchez G, Clements JD, Pasetti MF, Picking WL. Broadly protective shigella vaccine based on type iii secretion apparatus proteins. Infect Immun. 2012;80(3):1222–1231. doi:10.1128/IAI.06174-11.
12. Chitradevi ST, Kaur G, Uppalapati S, Yadav A, Singh D, Bansal A. Co-administration of ripab domain of shigella with rgroel of s. Typhi enhances the immune responses and protective efficacy against shigella infection. Cell Mol Immunol. 2015;12(6):757–767. doi:10.1038/cmi.2014.86.
13. Yuan X, Lin X, Manorek G, Kanatani I, Cheung LH, Rosenblum MG, Howell SB. Recombinant cpe fused to tumor necrosis factor targets human ovarian cancer cells expressing the claudin-3 and claudin-4 receptors. 2009;Mol Cancer Ther. 8(7):1906–1915. doi:10.1158/1535-7163.MCT-09-0106.
14. Ebihara C, Kondoh M, Hasuike N, Harada M, Mizuguchi H, Horiguchi Y, Fujii M, Watanabe Y. Preparation of a claudin-targeting molecule using a c-terminal fragment of clostridium perfringens enterotoxin. J Pharmacol Exp Ther. 2006;316(1):255–260. doi:10.1124/jpet.105.093351.

15. Van Itallie CM, Betts L, Smedley JG 3rd, McClane BA, Anderson JM. Structure of the claudin-binding domain of clostridium perfringens enterotoxin. J Biol Chem. 2008;283(1):268–274. doi:10.1074/jbc.M708066200.
16. Tamagawa H, Takahashi I, Furuse M, Yoshitake-Kitano Y, Tsukita S, Ito T, Matsuda H, Kiyono H. Characteristics of claudin expression in follicle-associated epithelium of peyer's patches: Preferential localization of claudin-4 at the apex of the dome region. Lab Invest. 2003;83(7):1045–1053. doi:10.1097/01.LAB.0000078741.55670.6E.
17. Shimanovich AA, Buskirk AD, Heine SJ, Blackwelder WC, Wahid R, Kotloff KL, Pasetti MF. Functional and antigen-specific serum antibody levels as correlates of protection against shigellosis in a controlled human challenge study. Clin Vaccine Immunol. 2017;24(2). doi:10.1128/CVI.00412-16.
18. Heine SJ, Franco-Mahecha OL, Chen X, Choudhari S, Blackwelder WC, van Roosmalen ML, Leenhouts K, Picking WL, Pasetti MF. Shigella ipab and ipad displayed on l. Lactis bacterium-like particles induce protective immunity in adult and infant mice. Immunol Cell Biol. 2015;93(7):641–652. doi:10.1038/icb.2015.24.
19. Chitradevi ST, Kaur G, Sivaramakrishna U, Singh D, Bansal A. Development of recombinant vaccine candidate molecule against shigella infection. Vaccine. 2016;34(44):5376–5383. doi:10.1016/j.vaccine.2016.08.034.
20. Dickenson NE, Arizmendi O, Patil MK, Toth RTt, Middaugh CR, Picking WD, Picking WL. N-terminus of ipab provides a potential anchor to the shigella type iii secretion system tip complex protein ipad. Biochemistry. 2013;52(49):8790–8799. doi:10.1021/bi400755f.
21. Shen DK, Saurya S, Wagner C, Nishioka H, Blocker AJ. Domains of the shigella flexneri type iii secretion system ipab protein involved in secretion regulation. Infect Immun. 2010;78(12):4999–5010. doi:10.1128/IAI.00470-10.
22. Suzuki H, Kakutani H, Kondoh M, Watari A, Yagi K. The safety of a mucosal vaccine using the c-terminal fragment of clostridium perfringens enterotoxin. Pharmazie. 2010;65(10):766–769.
23. Uchida H, Kondoh M, Hanada T, Takahashi A, Hamakubo T, Yagi K. A claudin-4 modulator enhances the mucosal absorption of a biologically active peptide. Biochem Pharmacol. 2010;79(10):1437–1444. doi:10.1016/j.bcp.2010.01.010.
24. Mitchell LA, Koval M. Specificity of interaction between clostridium perfringens enterotoxin and claudin-family tight junction proteins. Toxins. 2010;2(7):1595–1611. doi:10.3390/toxins2071595.
25. Hashimoto Y, Yagi K, Kondoh M. Roles of the first-generation claudin binder, clostridium perfringens enterotoxin, in the diagnosis and claudin-targeted treatment of epithelium-derived cancers. Pflugers Arch. 2017;469(1):45–53. doi:10.1007/s00424-016-1878-6.
26. Veshnyakova A, Protze J, Rossa J, Blasig IE, Krause G, Piontek J. On the interaction of clostridium perfringens enterotoxin with claudins. Toxins. 2010;2(6):1336–1356. doi:10.3390/toxins2061336.
27. Yuan X, Lin X, Manorek G, Howell SB. Challenges associated with the targeted delivery of gelonin to claudin-expressing cancer cells with the use of activatable cell penetrating peptides to enhance potency. BMC cancer. 2011;11:61. doi:10.1186/1471-2407-11-61.
28. Cocco E, Casagrande F, Bellone S, Richter CE, Bellone M, Todeschini P, Holmberg JC, Fu HH, Montagna MK, Mor G, et al. Clostridium perfringens enterotoxin carboxy-terminal fragment is a novel tumor-homing peptide for human ovarian cancer. BMC cancer. 2010;10:349. doi:10.1186/1471-2407-10-349.
29. Masuyama A, Kondoh M, Seguchi H, Takahashi A, Harada M, Fujii M, Mizuguchi H, Horiguchi Y, Watanabe Y. Role of n-terminal amino acids in the absorption-enhancing effects of the c-terminal fragment of clostridium perfringens enterotoxin. J Pharmacol Exp Ther. 2005;314(2):789–795. doi:10.1124/jpet.105.085399.
30. Takahashi A, Kondoh M, Uchida H, Kakamu Y, Hamakubo T, Yagi K. Mutated c-terminal fragments of clostridium perfringens enterotoxin have increased affinity to claudin-4 and reversibly modulate tight junctions in vitro. Biochem Biophys Res Commun. 2011;410(3):466–470. doi:10.1016/j.bbrc.2011.05.161.
31. Klein JS, Jiang S, Galimidi RP, Keeffe JR, Bjorkman PJ. Design and characterization of structured protein linkers with differing flexibilities. Protein Eng Des Sel. 2014;27(10):325–330. doi:10.1093/protein/gzu043.
32. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The phyre2 web portal for protein modeling, prediction and analysis. Nat Protoc. 2015;10(6):845–858. doi:10.1038/nprot.2015.053.
33. Zhang Y, Skolnick J. Tm-align: A protein structure alignment algorithm based on the tm-score. Nucleic Acids Res. 2005;33(7):2302–2309. doi:10.1093/nar/gki524.
34. El-Manzalawy Y, Dobbs D, Honavar V. Predicting flexible length linear b-cell epitopes. Comput Syst Bioinformatics Conf. 2008;7:121–132.
35. Kringelum JV, Lundegaard C, Lund O, Nielsen M. Reliable b cell epitope predictions: Impacts of method development and improved benchmarking. PLoS Comput Biol. 2012;8(12):e1002829. doi:10.1371/journal.pcbi.1002829.
36. Puigbo P, Guzman E, Romeu A, Garcia-Vallve S. Optimizer: A web server for optimizing the codon usage of DNA sequences. Nucleic Acids Res. 2007;35(Web Server issue):W126–131. doi:10.1093/nar/gkm219.
37. Reuter JS, Mathews DH. Rnastructure: Software for rna secondary structure prediction and analysis. BMC bioinformatics. 2010;11:129. doi:10.1186/1471-2105-11-129.
38. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. Ucsf chimera–a visualization system for exploratory research and analysis. J Comput Chem. 2004;25(13):1605–1612. doi:10.1002/jcc.20084.