In silico analysis and selection of Clostridium perfringens type D epsilon toxin multiepitope fragment

Mohsen Mehrvarz[1], Mohsen Fathi Najafi[1],T.Zahraei Salehi[2],Behjat Majidi[1]

1Education and Extension Organization (AREEO), Agricultural Research, Razi Vaccine and Serum Research Institute, Mashhad Branch, Mashhad, Iran, 2-Department of Microbiology and Immunology, Faculty of Veterinary MedicineUniversity of TehranTehran Iran

Epitope based vaccination plays an important role in the protective immunity. To construct an effective universal vaccine against domestic infections caused by *C.perfrigenes* types B and D, the B cell epitopic regions of the epsilon toxin were analyzed with different algorithms. Using different algorithms and bioinformatics tools to select the best candidate fragment with the highest number of antigenic determinant regions is a powerful tool in immunoinformatic. Sequence primary information was obtained from national center of biotechnology (NCBI) and Protein Data Bank (PDB). B-cell epitope prediction algorithms performed on amino acid sequence allowed the identification of several regions containing putative conserved B-cell epitopes and at least 3 fragments with overlapping regions were selected. Sequence analysis was performed by CLC Main Work bench software, I-TASSERand Raptrox (for tertiary structure).Immunoinformatic analysis was accomplished by IEDB, ABCpred, Seppa servers. According to all outputs received from these algorithms fragment 3 has the highest score for linear epitope regions included and fragment 1 has the highest score among three fragments in the case of having discontiniuousepitopic regions. The additional experimental work is required to validate these multiepitope fragments and confirm results obtained here.

Key words: *Clostridium perfringens*,epsilon toxin, recombinant, vaccine, B cell epitope, fragment.

Highlights

- Available Information about epsilon toxin sequence, features and structure were collected from databases.
- A collection of different algorithms were used for epsilon toxin analysis.
- The best epitopic sites on the basis of their scores were selected for experimental researches.

1. Introduction

Clostridium perfringens gram positive bacteria is an anaerobic phathogen from clostridium genus. According to production of 4 different toxin types (α-, β-, ε- and ι-toxins) , it classified into toxinotypes; A, B , C, D and E. ε-toxin, the causative agent of infections in lamb, goat and claves (in the case of rich diets) is produced by toxinotypes A and B. Disease initiate

Disruption of microbial balance caused by overeating followed by the passage of undigested carbohydrates from rumen into the intestine lead to large amount proliferation of *C. perfringens* and production of ε-toxin. By proteases secreted from the gut the inactive

form of toxin become an active agent which can severely create vascular damage and lesions in various major organs viz. brain, heart, lung and kidney[1] . vaccination is the best strategy to prevent rapid disease progression. Common problem with conventional vaccines (heat inactivated method) is their leading to non-specific and uncontrolled immune response. convenient alternative to the conventional vaccines are recombinant vaccines . A protective immune response against the pathogen is largely based on B-cell mediated immune pathways.B lymphocytesrecognize native protein, polysaccharide and glycolipids antigens on the basis of linear highlyspecific 3-dimensional conformational epitopes. Continuous linear B cell epitopic regions can be mapped experimentally bypeptidescanning technique while overlapping peptides spanning the entire sequence are specially tested for antibody interacting residues. B cell epitopes(Conformational epitopes)are influenced by the structural and physiochemical properties of spatially adjacent residues complicating their identification[2]So, the major task during recombinant vaccine preparation progress is to identify the surface antigens containing antibody binding epitopes (B-cell epitopes). Subunit vaccines includeantigens which best stimulate the immune system. They only have essential parts of complete toxin, so; the probability of non-specific responses and damaging effects is very lower. Experimental identification of epitopes is a relatively time consuming and laborious work. Computational epitope prediction method helps us to detect promising epitopes with higher precision rapidly. B cell epitope prediction softwares utilize different algorithms on the basis of amino acids physicochemical properties such as flexibility,hydrophilicity,surface accessibility andbeta-turns. Since an epitope are small entities and not highly immunogenic by themselves, multi epitopic fragments can easier stimulate immune responses and do not need to a carrier (compared with single epitopes)[3]. Vaccines development based on B cell epitope identification method has been great success in many cases of infectious diseases.These vaccines compared to the usual vaccines (like killed vaccines and attenuated vaccines) are easy to produce and safe [4]. So the present study was designed for prediction and identification of both linear and conformational b-cell epitopes and selection of the best candidate fragment in the process of generating potential subunit vaccine against clostridium perfringens type D epsilon toxin.

2-Materials and methods:

2.1. Sequence primary information

Etx nucleotide (AJ426474.1) and amino acid (accession no. 209947607) were obtained from www.ncbi.nlm.nih.gov . We utilizedthis information as query for in silico analysis. Also Tertiary structure with PDB format (1UYJ) and additional details were obtained from NCBI and Protein Data Bank (www.rcsb.org).

2.2. Sequence analysis

To collect and have maximum information about toxin features, we used CLC main work bench and online software at http://expasy.org/tools/rabtorax orhttp://raptorx.uchicago.edu/.

The output data from CLC was pI, amino acid composition, total number of negatively and positively charged residues, half life, instability index and aliphatic index.

The tertiary structure was obtained from raptroxhttp://raptorx.uchicago.edu and I-TASSER (http://zhanglab.ccmb.med.umich.edu/I-TASSER).

2.3. Immunoinformatic analysis

Secondary structure of toxin was acquired from CLC and then the antigenicity of toxin was plotted on the basis of it. Linear and conformational epitope prediction was performed with different algorithms including; ABCpred (www.imtech.res.in), Immune Epitope Database (:IEDB-Discotope-Ellipro)www.iedb.org.SEPPA (Spatial Epitope Prediction of Protein Antigens) server at the Life Science and Technology School, Tongji University, Shanghai China, (http://lifecenter.sgst.cn/seppa/webcite) was used to predict conformational B-cell epitope.

3. Results and discussion:

3.1.Sequence primary information and analysis:

Amino acid sequence with accession number YP_002291114.1 was compared with amino acid sequence obtained from conversion of nucleotide sequence of toxin (AJ426474.1) and two mentioned sequences were completely identical. A conserved domain related to ETX-MTX2 superfamily was identified in Blast search. The PDB ID of sequence was 1UYJ. The blast results were shown that full length sequences of epsilon toxin with different accession codes are not significantly different. We preferred that the selected fragment overlapped well with conserved domain.

The primary analysis of sequence was accomplished with CLC MAIN Work Bench. Molecular weight and calculated pI were 36.22 kDa and 8.89 respectively and aliphatic index was 82.226.Half-life was estimated to be 30 h (mammalian reticulocytes, in vitro), >20 h (yeast, in vivo), >10 h (Escherichia coli, in vivo). Hydrophobic, hydrophilic and other amino acids frequency were 40.5, 37.5 and 22 percent respectively. 32 first amino acids belong to a signal peptide and from 32 to 334 are considered as α chain and from 46 to 328 are as β chain. Tertiary structure model of α chain was received from raptorx. The obtained results are presented in figure 3. I-TASSER 3D model with the highest confidence score (a score for estimating the quality of model) wasgot. It presents in following figure.

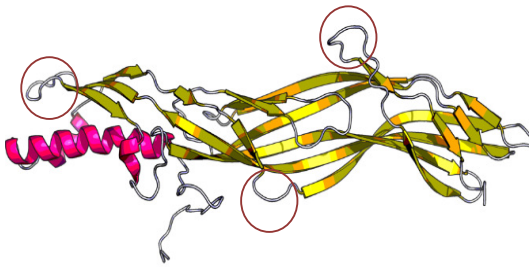


Figure 1. 3D model of epsilon toxin received from rabtorax (left) and I-TASSER. Because of similarity between two figures here we present only one from raptorx.regions surrounded by circules are turns with the higher epitope probability.

3.2.Immunoinformatic assay results

common existing methods are used to predict B-cell epitopes use protein sequences as input, and are best convenient to predict epitopes composed of a continuous stretch of amino acids (linear epitopes).This was accomplished using CLC, ABC pred ,BCE pred and Ellipro algorithms .in order to predict linear B cell epitopes,a combination of the most common algorithms such as plots and secondary structure, loops and helixes ,location of turns was used. The antigenicity was plotted against location of alpha-helixes, beta sheets, loops and turns (figure 3).

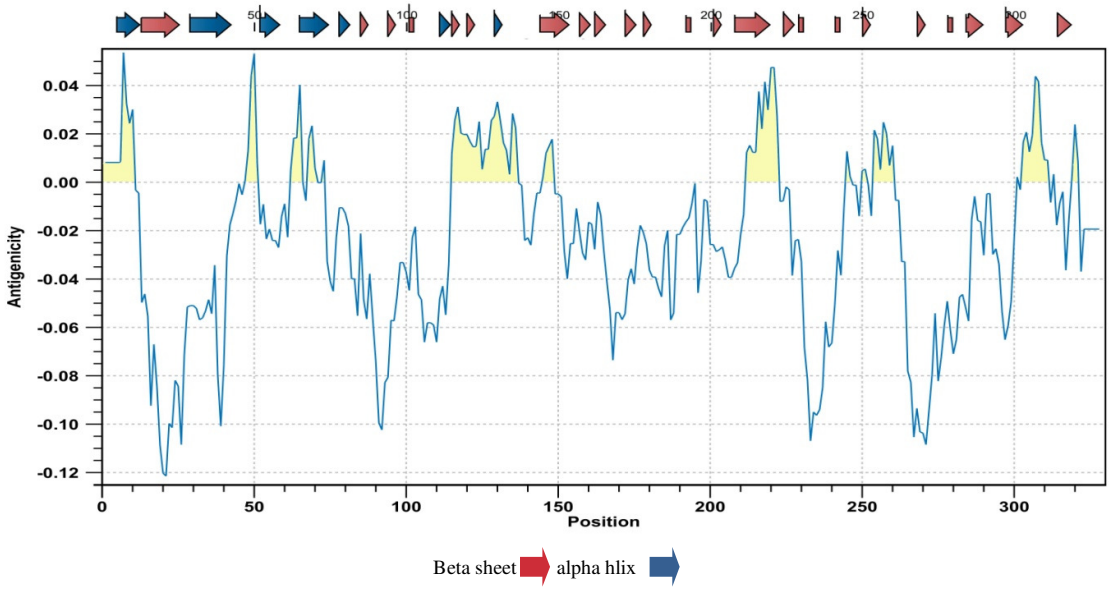


Figure3. Antigenicity plot against secondary structure of epsilon toxin (CLC-Main workbench. An antigenicity plot measures hedrophilicity of protein. The greater hydrophilicity the greater exposure and epitope probability.

Parameters such as flexibility, hydrophilicity,hydrophobicity, turns and their location,accessibility, exposed surface, polarity and antigenic propensity of polypeptides chains are related to the sites that continuous epitopes are located in. The existence of some secondary structures like alpha helix and beta sheets affects the exposure amount of amino acids to surface and as shown in figure 3,it identifies antigenicity properties of different parts of protein. Division of protein to fragments was carried out according to this plot and 3 overlapping fragments have been considered. The location and length of mentioned fragments have been shown in table 1.

Beta-turns , the most common type of non-repetitive structures, on average constitute 25% of the residues in all protein chains. Beta-turn's formation plays an important role in protein folding, molecular recognition processesandprotein stability.Therev is a high correlation between the tendency for a sequence to form a Beta-turn and the proteins reactivity to antipeptide antibodies, and furthermore beta turns are highly presented in B-cell epitopes[5]. In protein-protein interactions, the formation of Beta-turns has been indicated to be essential for high-affinity binding of antigen to antibody. IEDB Chou and Fasman Beta-Tern prediction online software was used to predict the location of beta turns. In fragments 1 and 3 there are more beta turn probability , therefore; fragment 2 has lower probability for containing epitopic regions(figure4).

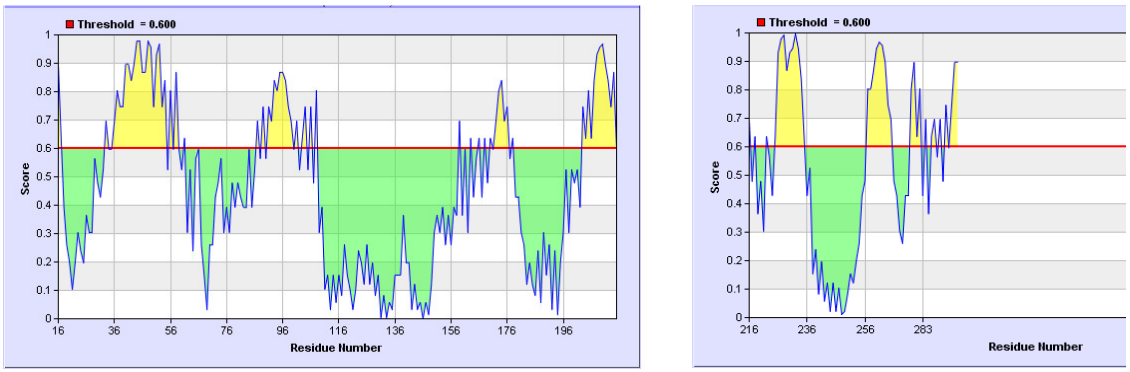


Figure4- Chou and Fasman Beta-Tern prediction

Approximately 90% of B-cell epitopes are discontinuous ,so, identification of these types of epitopes is more important than their linear forms[6].Antibody-antigen interactions are often conformation-dependent , therefore; in order to predict conformational B cell epitopes we require sturactural models of toxin which accuired from I-TASSER and RABTROX.It has used as a query for Discotope discontinuous epitope prediction server.

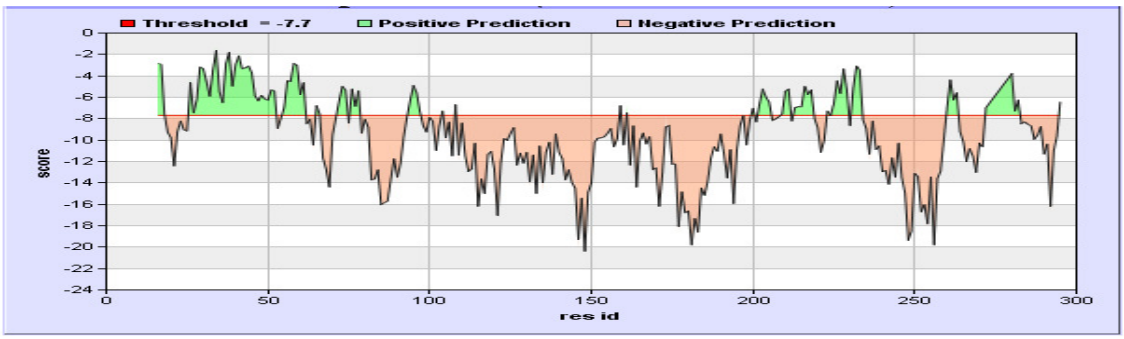


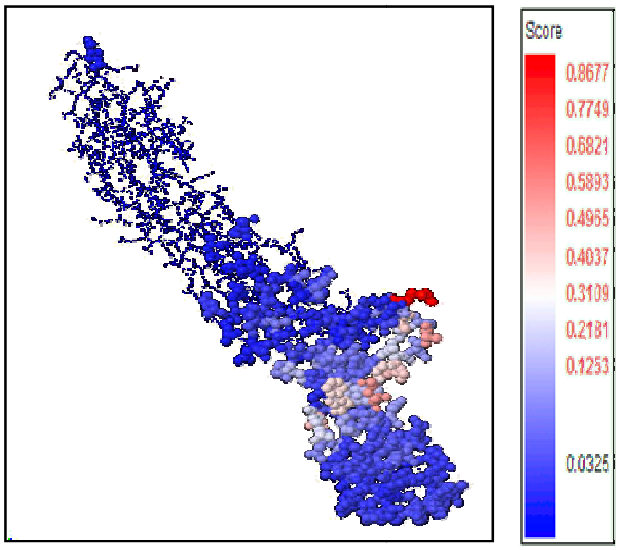
Figure 5- discotopediscontinuous epitope prediction software resulted graph.

As shown in figure 5, fragments 1&3 has green marked regions with positive scores and in total discotope scores (table 1) again these results have been confirmed.

table1- scoring of different fragments based on some used algorithms.

Fragment:	location	length	ABCpred score	Discotope (discontinuous)	Ellipro (discontinuous)	Ellipro(linear)
1	1-120	120	14.8	-732.244	0.902	1.535
2	100-215	115	13.7	-1008.72	0.5589	1.407
3	200-328	128	14.3	-834.266	2.04178	3.169

SEPPA server aims to divide antigen surface into epitopic and non epitopic small patchesand analyses 3D structures on the basis of solvent accessibility and different propensity scores relying on training data series comprising resolved antigen- antibody complexes.the results obtained by this software can be visualized only by JMOL in different renderings(figure 6).In figure 6, the Highlighted epitopic residues were predicted and shown in red colored spheres. Lower results are shown in blue. The sequence of toxin and its residues are listed in figure 7. Yellow highlighted parts are the best scored epitopic regions and other non-epitopic regions haveshown by lowercase . Antigenic epitopes are recognized by antibodies that can help us in designing of immuno-diagnostic reagents andvaccine components.



Sequence analysis results by SEPPA server.

Figure6. SEPPA server analyses 3D structures. Highlighted epitope residues were predicted and shown in red sphers.

In this studythe screening for immunogenic epitopes of epsilon toxin was performed to select multiepitope fragments as convenient candidate vaccine against clostridium perfringens type D. 3 fragments with acceptable size and with considering antigenicity plot were selected and analyzed for epitopic regions. Although we used different programs and these used softwares utilize different algorithms, there was considerable overlap in the residues identified. Additionally, the fragment score obtained from themwas appromixately similar. According to all outputs received from these algorithms fragment 3 has the highest score for linear epitope regions included and fragment 1 has the highest score among three fragments in the case of having discontinuousepitopic regions. The additional experimental work is required to confirm these multiepitope fragments and confirm results obtained here.

References:

[1]. Knapp, O., et al., *Identification of the channel-forming domain of Clostridium perfringens Epsilon-toxin (ETX).* Biochim Biophys Acta, 2009. **1788**(12): p. 2584-93.

[2]. Bergmann-Leitner, E.S., et al., *Computational and experimental validation of B and T-cell epitopes of the in vivo immune response to a novel malarial antigen.* PLoS One, 2013. **8**(8): p. e71610.

[3]. Kaushik, H., et al., *Recombinant expression of in silico identified Bcell epitope of epsilon toxin of Clostridium perfringens in translational fusion with a carrier protein.* Bioinformation, 2013. **9**(12): p. 617-21.

[4]. prabhavathy, K., P. perumal, and N. sundaraBaalaji, *In silico identification of B- and T-cell epitopes on ompla and LsrC from Salmonella Typhi for peptide-based subunit vaccine design.* indian journal of biotechnology, 2011. **10**: p. 440-451.

[5]. Petersen, B., C. Lundegaard, and T. Petersen, *NetTurnp - Neural Network Prediction of Beta-turns by Use of Evolutionary Information and Predicted Protein Sequence Features.* PLoS ONE, 2010. **5**((11):e15079 doi:10.1371/journal.pone.0015079).

[6]. Andersen, P.H., M. Morten Nielsen, and O. Lund, *Prediction of residues in discontinuous B-cell epitopes using protein 3D structures.* Protein Science, 2006. **15**(11): p. 2558-2567.