Original Article

JYP

An immunoinformatics approach toward epitope-based vaccine design through computational tools from *Bungarus caeruleus*'s neurotoxin

Kutub Uddin Muhammad Ashraf*, Prosenjit Barua, Ayan Saha, Nur Mahammad, Jannatul Ferdoush, Dipesh Das, Md. Hamed Hussain, Md. Jibran Alam

Department of Genetic Engineering and Biotechnology, Faculty of Biological Science, University of Chittagong, Chittagong, Bangladesh

ABSTRACT

Objective: This study aims to analyze and predict the possibility of designing a vaccine that could make humans immune to krait toxin. **Materials and Methods:** *Bungarus caeruleus* or common Indian krait is a member of the venomous big four snake species. Its venom contains a neurotoxic protein alpha-delta-bungarotoxin-4 and is found to be responsible for human death 4-8 h after the snake bite. Antigenicity of this protein was determined by Hopp and Woods and Kolaskar and Tangaonkar method. We predicted major histocompatibility complex (MHC) Class I and MHC Class II binding peptides of antigenic protein from alpha-deltabungarotoxin-4, which are an important determinant for protection of host from snake bite. Fragments selected through this study revealed higher efficiency binders. **Result:** Higher percentages of their atoms are directly involved in binding in comparison with larger molecules. These potential fragments, therefore can be a novel tool in the arena of cross protection to develop host specific antibodies in different objectives. We operated AllerHunter for predicting allergenicity based on the structural and physiochemical properties of whole alpha-delta-bungarotoin-4, and it was found to be nonallergen. The potential epitopes of alpha-delta-bungarotoxin-4 were found to be located at sequences "GENLCYTKM" and "FCSSRGKVI" and these were found to be sufficient for eliciting the desired immune response. In this study, a hypothetical immunization is developed, which demands more validation and study. It can be emphasized that such predictive *in silico* study requires an *in vivo* experiments comprehensibly, which must be assured to validate such approaches. Hence, our goal was to identify a conformationally biased epitope sequence, which aims to provide a new paradigm to design epitope-based peptide vaccines in order to alleviate immunological infections from Krait neurotoxin. **Conclusion:** Computational techniques manifest the attention of Krait neurotoxin as crucial immunodiagnostic tool for fatal venom proved that most snake venoms are in poorly characterized although they are biologically important proteins with therapeutic potentialities.

Key words: Alpha-delta-bungarotoxin-4, antigen protein, bungarotoxin, *Bungarus caeruleus,* nonamers, peptide vaccine

Access this article online	
Journal Sponsor	Website: www.jyoungpharm.org
www.phcog.net	DOI: 10.5530/jyp.2014.2.6

INTRODUCTION

The common krait (*Bungarus caeruleus*, generally known as Indian krait or Blue krait) is a species of genus *Bungarus*. To note about its habitat, it is found in the jungles of the Indian subcontinent and its local name in Bangladesh is Shan-

Address for correspondence:
Mr. Kutub Uddin Muhammad Ashraf, Department of Genetic Engineering & Biotechnology, Faculty of Biological Science, University of Chittagong, Chittagong - 4331, Bangladesh. E-mail: kutubuddinashraf@yahoo.com

Khani. Being a member big four species this venomous snake imposes a greater range of snakebites in India.[1] Its venom is extremely neurotoxic and quickly induces muscle paralysis. Study revealed that after 1 to 6 h of the Krait bite, death from respiratory failure may happen. Situation can be worsened as Indian krait bites may be 100% fatal if it is not immediately treated.[2] Clinically, its venom contains pre- and post-synaptic neurotoxins. These neurotoxins generally hamper the nerve endings vicinity to the synaptic cleft of the brain. Thus, it creates a peripheral paralysis after blocking neuromuscular transmission at the postsynaptic site which precludes long-term side-effects. Recent findings indicated that the venom of the common krait (*B. caeruleus*) produces a significant depression of vital centers in the brainstem. It contains an adequate amount of hemolysins and hemorrhagins. Right now, it is known that bungarotoxins can block the transmission at the neuromuscular junction. Moreover, anticoagulant and neurotoxic activities were identified in the protein isolated from common krait venom, which was later subjected to sequence and a crystal structure was determined. Besides, phospholipase A2 enzyme is shared by many snakes, which possesses a wide array of effects including vasodilatation, rhabdomyolysis, hemolysis, and release of endogenous autacoids that can cause neuromuscular blocking significantly.[3] In addition, the neurotoxins perhaps contribute to ultrastructural damage to motor nerve endings. Hence, polyvalent antivenom found no significant development ($t = 0.5$) in reversing respiratory paralysis and preventing delayed neurological complications. Study revealed that sixteen (7.6%) patients had to die, and a submucosal hemorrhage in the stomach was experienced at necropsy in three distinct case studies. Mortality could be reduced with an early treatment and free access to mechanical ventilation was emphasized.[3] Krait bites (*B. caeruleus*) is routinely attributed throughout South Asia based on clinical symptoms and the greater superficial similarities of *B. caeruleus*, *Bungarus walli* and *Bungarus sindanus* was noticed. However, envenoming by krait species other than *B. caeruleus* didn't respond to available antivenoms as observed in Bangladesh.[4] It is understood that lifesaving antivenoms possesses an immunoglobulin pool of known redundancy and unknown antigen specificity, which requires the transmission of huge volumes of heterologous immunoglobulin to the affected victim. Consequently, it raises the possibility of anaphylactoid and serum sickness which render a strong detrimental effect.[5] Meanwhile, recent progresses in computational tools have eased to predict and further analysis of T-cell and B-cell epitopes from antigenic proteins in specialized experiments. This has directed to peptide-based vaccines design planning that is more unique, optimized and secured to predict the peptide binding to human leukocyte antigen (HLA) alleles applying structural and modeling techniques. Surprisingly, such methodologies strengthened in recent years in order to alleviate some acute immunological infections. Until date, it is the first immunoinformatics study on *B. caeruleus* toxin and snake venom in Bangladesh perspective.

MATERIALS AND METHODS

Protein sequence retrieval

The protein sequence of alpha-delta-bungarotoxin-4 was retrieved from National Center for Biotechnology Information's (NCBI's) protein database (NCBI, http://www.ncbi.nlm.nih.gov/protein/) by GenBank accession no. CAM11302.1. This sequence was executed and studied in order to identify the immunologically relevant regions, B cell epitope regions and major histocompatibility complex (MHC) Classes I and II binding sites with significant scores.

Identification of conserved domain

To identify the conserved domain of venom protein sequence was aligned with protein superfamily members, cd00206[6] and pfam00087[7] by using the Conserved Domain Database of NCBI server.

Secondary structure prediction and hydrophilicity estimation

ExPASy's secondary structure prediction server (http://web.expasy.org/protparam/)[8] was used to get an idea about the secondary structure of the venom protein. Several physicochemical parameters given by protparam tool were studied, e.g. molecular weight, amino acid and atomic composition, instability and aliphatic index, theoretical pI, estimated half-life, extinction coefficient, and grand average of hydropathicity.[9-11] To calculate hydrophilicity, Hopp and Woods hydrophilicity scale was analysed.[12]

Prediction of MHC binding peptide

To predict the MHC binding peptides for venom protein, two options were used that were provided by Immune Epitope Database (IEDB) analysis resource. For MHC Class I, peptide prediction, proteasomal cleavage/TAP transporter/MHC Class I combined prediction server (http://tools.immuneepitope.org/processing/) and for MHC Class II peptide, MHC II binding prediction (http://tools.immuneepitope.org/mhcii) were used. We used the Artificial Neural Network prediction methods in these servers to predict the potential nonamers that may significantly bind to the binding grooves of the MHC molecules.

B-cell epitope prediction

Kolaskar and Tangaonkar antigenicity scale[13] available at IEDB analysis resource (http://tools.immuneepitope.org/tools/bcell/iedb_input) was analyzed to predict the B-call epitopes.

Allergenicity assessment

In order to assay the degrees of allergenicity, we operated AllerHunter (http://tiger.dbs.nus.edu.sg/AllerHunter/index.html). A combinational prediction by using both support vector machine and pair-wise sequence similarity makes AllerHunter a very useful program for cross-reactive allergen prediction. Cross-reactivity is a phenomenon, which is based on similarity among proteins and allergens, whereas allergenicity means the ability of an allergen to induce immunoglobulin E antibody production. AllerHunter predicts allergens as well as non-allergens with high specificity. Moreover, it does not compromise its efficiency, while classifying proteins with similar sequence to known allergens.

Epitope conservancy and population coverage analysis

Epitope conservancy and populations covered by epitopes when used as vaccine were analyzed by using epitope analysis tools of IEDB analysis resource server (http://tools.immuneepitope.org/main/html/analysis_tools.html). Predicted epitopes and the protein sequences that were used to predict the conserved region were used to find out the accuracy of the prediction.

Docking simulation

In silico docking simulation was done to find out whether or not these peptides will bind to the MHC molecules when will be applied for further *in vivo* experiments. For docking simulation study, we used AutoDock Vina[14] developed by The Scripps Research Institute. To carry out the docking simulations, three MHC II molecules (PDB ID: 1 H15, 1 AQD and 1 DLH) and three MHC I molecules (PDB ID: 1 A1O, 1 JHT and 3 LKN) were taken into consideration. PDB files for the predicted epitopes were prepared by using HHPred to use them as ligands. AutoDock tools were used for preparation of receptor and ligand molecules for docking simulations at the binding groove of the MHC molecules. To reduce calculation time, search exhaustiveness was set to four.

RESULTS

Identification of conserved domain

Two conserved domains were found in conserved domain search; one of which was snake toxin domain and other was a disulfide rich snake toxin. In both cases, amino acids 3-68 were shown conserved (Figure 1).

Secondary structure analysis

The secondary structural features of alpha-delta-bungarotoxin-4 protein are summarized in Table 1.

Hydrophilicity estimation

Hopp and Woods hydrophilicity scale was used to determine the hydrophilicity of the conserved region. The venom protein was found to be hydrophilic in nature as the average value of the scale was 0.185 with the minimum value of 1.171 and maximum of 0.986 (Figure 2).

Table 1: Secondary structural analysis of alpha-delta-bungarotoxin-4 by ProtParam tool

Criteria	Assessment
Number of amino acids	76
Molecular weight	8305.5 Da
Isoelectric pH	7.65
No. of negatively charged residues (Asp+Glu)	6
No. of positively charged residues (Arg+lys)	7
Formula	$C_{353}H_{556}N_{96}O_{113}S_{11}$
Extinction coefficient	10595
Instability index	45.57
Aliphatic index	48.68

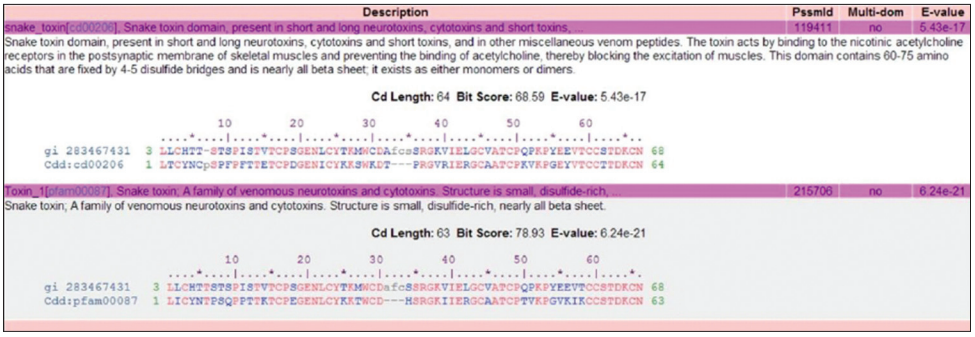

Figure 1: Conserved domain search for alpha delta bungarotoxin-4

Prediction of MHC binding peptide and B-cell epitopes

A total of 48 alleles were analyzed for MHC Class I peptide prediction by using artificial neural network method.[15,16] Again, 26 MHC Class II alleles were analyzed for prediction of MHC II binding peptides from the venom protein. Nonamers that showed high prediction results were selected in this study. Interaction among different alleles with these peptides is summarized in Tables 2 and 3. In case of MHC Class II prediction, artificial neural network alignment method was used.[17] For selection of all the MHC binding peptides, MHC IC50 score was below 250 nM. The B cell epitopes that were situated in the conserved domain region of alpha-delta-bungarotoxin-4 were selected by analyzing Kolaskar and Tangaonkar antigenicity scale. The B cell epitope regions that match the same regions as the predicted MHC binding peptides are summarized in Table 4. Two nonamers were found that fall into B cell epitope region.

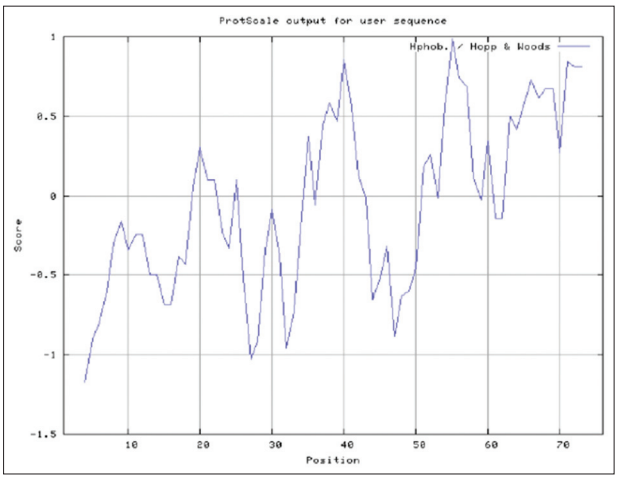

Figure 2: Hopp and Woods hydrophilicity results for alpha delta bungarotoxin-4

Table 3: Prediction of MHC II peptides

Allele	Start	End	Peptide	IC_{50}
HLA-DQA1*01:02/DQB1*06:02	24	32	LCYTKMWCD	47.20
HLA-DRB1*07:01	24	32	LCYTKMWCD	62.20
HLA-DQA1*05:01/DQB1*03:01	34	42	FCSSRGKVI	98.60
HLA-DRB1*01:01	34	42	FCSSRGKVI	34.40
HLA-DRB1*07:01	34	42	FCSSRGKVI	3.70
HLA-DRB1*09:01	34	42	FCSSRGKVI	194.80
HLA-DRB1*11:01	34	42	FCSSRGKVI	42.50
HLA-DRB3*01:01	34	42	FCSSRGKVI	244.80
HLA-DRB5*01:01	34	42	FCSSRGKVI	10.40
HLA-DQA1*05:01/DQB1*03:01	42	50	IELGCVATC	209.50
HLA-DRB1*04:04	42	50	IELGCVATC	145.40
HLA-DRB1*01:01	5	13	CHTTSTSPI	75
HLA-DRB1*07:01	5	13	CHTTSTSPI	11.10
HLA-DRB1*01:01	1	9	YTLLCHTTS	116.90
HLA-DRB1*04:01	1	9	YTLLCHTTS	42.30
HLA-DRB1*04:04	1	9	YTLLCHTTS	22.70
HLA-DRB1*04:01	4	12	LCHTTSTSP	104.80
HLA-DRB1*04:05	4	12	LCHTTSTSP	102.20
HLA-DRB1*04:01	3	11	LLCHTTSTS	179.60
HLA-DRB1*04:04	3	11	LLCHTTSTS	182.30

Table 4: Prediction of B cell epitopes by Kolaskar and Tangaonkar method

Start	End	Peptide	Peptide length
21	28	GENLCYTK	9
31	54	CDAFCSSRGKVIELGCVATCPQPK	24

Table 2: Prediction of MHC I peptides

Allele	Start	End	Length	Peptide	Proteasome score	TAP score	MHC score	Processing score	MHC IC_{50}
HLA-A*03:01	46	54	9	CVATCPQPK	0.49	0.24	−2.04	0.73	110
HLA-A*11:01	46	54	9	CVATCPQPK	0.49	0.24	−1.76	0.73	57
HLA-A*68:02	8	16	9	TSTSPISTV	0.96	0.13	−1.18	1.09	15
HLA-C*12:03	8	16	9	TSTSPISTV	0.96	0.13	−1.96	1.09	91
HLA-B*15:01	26	34	9	YTKMWCDAF	1.15	1.08	−1.61	2.22	41
HLA-C*14:02	26	34	9	YTKMWCDAF	1.15	1.08	−2.08	2.22	121
HLA-B*35:01	18	26	9	CPSGENLCY	1.31	1.12	−0.70	2.43	5
HLA-B*53:01	18	26	9	CPSGENLCY	1.31	1.12	−1.18	2.43	15
HLA-B*39:01	5	13	9	CHTTSTSPI	0.85	0.22	−2.23	1.08	170
HLA-C*14:02	5	13	9	CHTTSTSPI	0.85	0.22	−2.04	1.08	109
HLA-B*40:01	21	29	9	GENLCYTKM	0.99	-0.02	−1.85	0.96	71
HLA-B*40:02	21	29	9	GENLCYTKM	0.99	-0.02	−1.61	0.96	41
HLA-B*44:02	21	29	9	GENLCYTKM	0.99	-0.02	−2.14	0.96	138
HLA-B*44:03	21	29	9	GENLCYTKM	0.99	-0.02	−2.17	0.96	147
HLA-C*03:03	34	42	9	FCSSRGKVI	1.43	0.26	−2.13	1.69	135
HLA-C*08:02	34	42	9	FCSSRGKVI	1.43	0.26	−2.10	1.69	125
HLA-C*12:03	34	42	9	FCSSRGKVI	1.43	0.26	−1.40	1.69	25

*TAP: Transporter of antigenic peptide

Epitope conservancy and population coverage prediction

The epitopes 21-29 and 34-42 were found to be 100% conserved in a scale of sequence match <100%. The minimum identity for both epitopes was more than 20%, and the maximum was 88.89% for both epitopes. Population coverage analysis yielded significant results using both epitopes (Table 5).

Allergenicity evaluation

The query sequence didn't meet the criteria set by the Food and Agriculture Organization (FAO)/World Health Organization (WHO) evaluation scheme for cross-reactive allergen prediction. Hence, the query sequences were classified as a nonallergen by the FAO/WHO evaluation scheme. Alpha-delta-bungarotoxin-4 protein was predicted as a non-allergen with a prediction score of 0.0 (sensitivity = 93.0%, specificity = 79.4%).

DOCKING SIMULATION RESULTS

The area that were selected on the receptor molecules for docking with the epitopes are summarized in Table 6. One angstrom spacing was used to select the binding site. The center box area was positioned carefully to make the docking of ligands at the binding groove of the receptors. The predicted peptides showed significant binding affinity to the MHC receptors (Table 7). The binding energy of the predicted epitopes were compared with the binding energy of the Ls6 peptide (Sequence: KPIVQYDNF) from malaria parasite with HLA-B*5301 (−6.0 kCal/mol). Strong binding affinity gives a clear idea that peptide vaccine designed by using these epitopes may efficiently work *in vivo* to elicit humoral and cell mediated immunity (Figures 3 and 4).

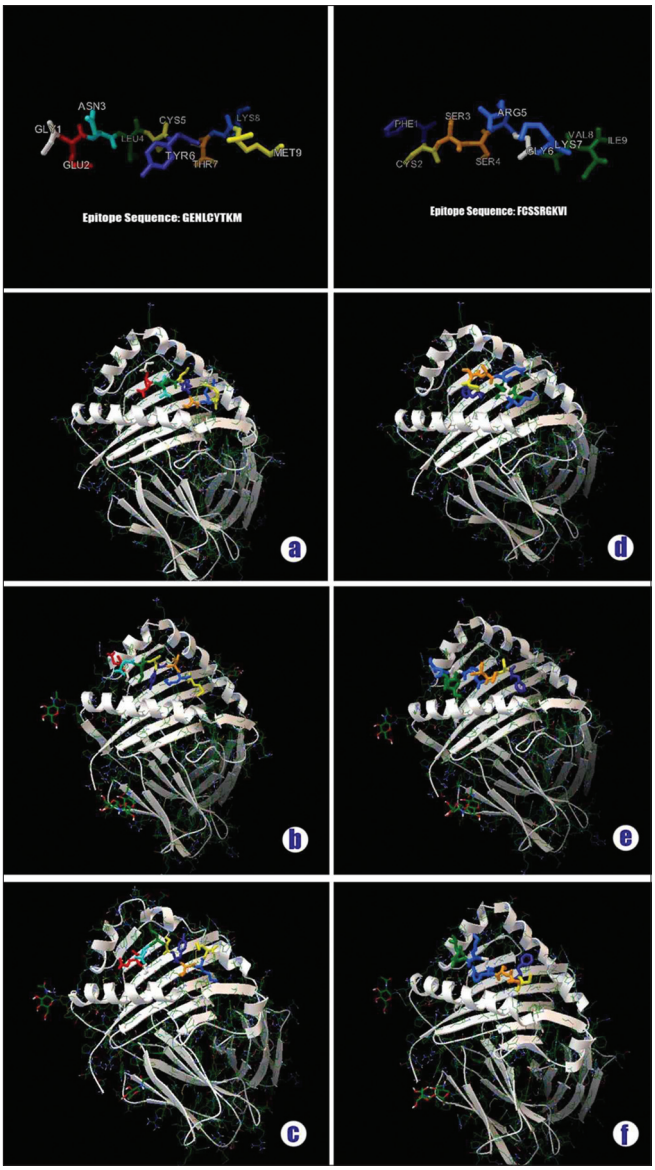

Figure 3: Visualization of docking results for predicted peptides with major histocompatibility complex Class I receptors by using AutoDock Tools. (a-c) docking images of "GENLCYTKM" with 3LKN, 1A1O, 1JHT respectively; (d-f) represents docking images of "FCSSRGKVI" with 3LKN, 1A1O, 1JHT respectively

DISCUSSION

Prediction of epitope and mapping these on the protein surface is a vital step for epitope based vaccine design. A number of ways were attempted in earlier studies, but here we tried to predict the epitopes more accurately by starting from the very basic step like finding the conserved and hydrophilic regions of protein and ending by docking of epitopes to HLA receptors.

The conserved region was found by aligning with protein family member cd00206 and pfam00087. The actual alignment for conserved region was done with the sequences, chain B-acetylcholinesterase (E.C. 3.1.1.7) complexed with fasciculin-II (PDB: 1FSS_B), chain A-crystal structure of kappa-bungarotoxin At 2.3-angstrom resolution (PDB: 1KBA_A), chain A-erabutoxin (PDB: 1QKD_A), Chain A-nuclear magnetic resonance structure of the complex between A-bungarotoxin and A mimotope of the nicotinic acetylcholine receptor (PDB: 1HOY_A), toxin S4C6 (UniProtKB/Swiss-Prot: P25670.1), alpha-elapitoxin-Ast2a (UniProtKB/Swiss-Prot: P01380.1), long neurotoxin homolog Pa ID (UniProtKB/Swiss-Prot: P14612.2), short neurotoxin one (UniProtKB/Swiss-Prot: P10808.1), weak toxin CM-1c (UniProtKB/Swiss-Prot: P25676.1), alpha-neurotoxin (Pseudonaja textilis) (GenBank: AAF75223.1)

Table 5: Population coverage by epitopes "GENLCYTKM" and "FCSSRGKVI"

Population/area	Class I and II		
	Coverage[a]	Average hit[b]	PC90[c]
Australia	76.50%	2.44	0.85
Australia : Cape York	68.30%	2.06	0.63
Australia : Groote Eylandt	71.53%	1.93	0.70
Australia : Kimberley	90.26%	3.31	2.01
Australia : Yuendumu	48.51%	1.10	0.39
Europe	91.41%	4.02	2.11
Europe : Bulgarian	31.70%	0.71	0.29
Europe : Croatian	21.97%	0.46	0.26
Europe : Cuban (Eu)	25.30%	0.53	0.27
Europe : Czech	91.86%	4.04	2.14
Europe : Finn 90	77.36%	2.61	0.88
Europe : Georgian	37.37%	0.85	0.32
Europe : Irish	77.51%	2.49	0.89
Europe : North America (Eu)	58.43%	1.51	0.48
Europe : Slovenian	81.71%	2.67	1.09
North Africa	85.65%	3.13	1.39
North Africa : Algerian 99	63.06%	1.63	0.54
North Africa : Chaouya	52.87%	1.30	0.42
North Africa : Metalsa	59.83%	1.56	0.50
North Africa : Moroccan 98	83.17%	2.71	1.19
North Africa : Moroccan 99	74.21%	2.17	0.78
North America	92.18%	3.73	2.15
North America : Amerindian	41.30%	0.97	0.34
North America : Lacandon	58.15%	1.42	0.48
North America : Seri	72.67%	2.18	0.73
North America : Yupik	98.02%	4.73	2.94
North-East Asia	72.53%	2.22	0.73
North-East Asia : Buriat	0.00%	0.00	0.00
North-East Asia : Korean 200	72.83%	2.21	0.74
North-East Asia : Tuva	72.48%	2.22	0.73
Oceania	71.53%	2.15	0.70
Oceania : American Samoa	58.27%	1.31	0.48
Oceania : Filipino	66.22%	1.90	0.59
Oceania : Ivatan	62.84%	1.61	0.54
Other	88.33%	3.50	1.71
Other : Brazilian	25.33%	0.53	0.27
Other : Brazilian (Af Eu)	62.84%	1.76	0.54
Other : Cuban (Af Eu)	24.48%	0.51	0.26
Other : Mexican	88.23%	3.44	1.70
Other : North America (Hi)	44.83%	1.08	0.36
South America	98.32%	4.35	2.97
South America : Bari	63.73%	1.27	0.55
South America : Guarani-Kaiowa	97.47%	3.95	2.68
South America : Guarani-Nandewa	99.24%	4.59	3.56
South-East Asia	81.52%	2.78	1.08
South-East Asia : Ami 97	71.26%	1.85	0.70
South-East Asia : Atayal	81.18%	2.49	1.06
South-East Asia : Bunun	61.89%	1.62	0.52
South-East Asia : Chinese	40.85%	0.92	0.34
South-East Asia : Hakka	71.17%	2.02	0.69
South-East Asia : Han-Chinese 149	35.14%	0.72	0.31
South-East Asia : Han-Chinese 572	30.60%	0.62	0.29
South-East Asia : Kinh	74.35%	1.95	0.78
South-East Asia : Malay	79.78%	2.46	0.99
South-East Asia : Minnan	71.90%	2.08	0.71
South-East Asia : Muong	31.62%	0.70	0.29
South-East Asia : North America (As)	40.96%	0.94	0.34
South-East Asia : Okinawan	58.78%	1.47	0.49
South-East Asia : Paiwan 51	87.01%	2.87	1.54
South-East Asia : Pazeh	73.97%	2.18	0.77
South-East Asia : Puyuma 49	70.69%	2.05	0.68
South-East Asia : Rukai	82.85%	2.71	1.17
South-East Asia : Ryukuan	0.00%	0.00	0.00
South-East Asia : Saisiat	68.10%	1.74	0.63
South-East Asia : Singapore (Chinese)	24.96%	0.50	0.27
South-East Asia : Siraya	81.52%	2.70	1.08
South-East Asia : Thai	20.03%	0.42	0.25
South-East Asia : Thao	68.36%	1.79	0.63
South-East Asia : Toroko	84.51%	2.70	1.29
South-East Asia : Tsou	66.51%	1.84	0.60
South-East Asia : Yami	58.17%	1.43	0.48
South-West Asia	90.19%	3.63	2.01
South-West Asia : Arab Druze	36.60%	0.84	0.32
South-West Asia : Israeli Jews	34.79%	0.79	0.31
South-West Asia : Kurdish	35.58%	0.80	0.31
South-West Asia : Omani	2.47%	0.05	0.21
South-West Asia : Turk	85.89%	2.94	1.42
Sub-Saharan Africa	78.06%	2.63	0.91
Sub-Saharan Africa : Doggon	29.30%	0.66	0.28
Sub-Saharan Africa : Kenyan 142	71.37%	1.93	0.70
Sub-Saharan Africa : Kenyan Highlander	4.91%	0.10	0.21
Sub-Saharan Africa : Kenyan Lowlander	15.04%	0.31	0.24
Sub-Saharan Africa : Mandenka	1.06%	0.02	0.20
Sub-Saharan Africa : North America (Af)	25.06%	0.55	0.27
Sub-Saharan Africa : Rwandan	53.89%	1.32	0.43
Sub-Saharan Africa : Shona	76.84%	2.51	0.86
Sub-Saharan Africa : Ugandan	40.97%	0.97	0.34
Sub-Saharan Africa : Zambian	24.36%	0.53	0.26
Sub-Saharan Africa : Zulu	85.68%	3.12	1.40
Average (standard deviation)	59.66% (26.09%)	1.84 (1.14)	0.81 (0.69)

[a]Projected population coverage. [b]Average number of epitope hits/HLA combinations recognized by the population. [c]Minimum number of epitope hits/HLA combinations recognized by 90% of the population

and chain A-crystal structure of the extracellular domain of the nicotinic acetylcholine receptor one subunit bound to alpha-bungarotoxin At 1.9 A resolution (PDB: 2QC1_A), cytotoxin 10 (UniProtKB/Swiss-Prot: P01453.1), cytotoxin V-II-2/V-II-3 (UniProtKB/Swiss-Prot: P01474.1), cytotoxin 11 (UniProtKB/Swiss-Prot: P62394.1), cytotoxin homolog S3C2 (UniProtKB/Swiss-Prot: P19003.1), cytotoxin homolog S4C8 (UniProtKB/Swiss-Prot: P19004.1), toxin C10S2C2 (UniProtKB/

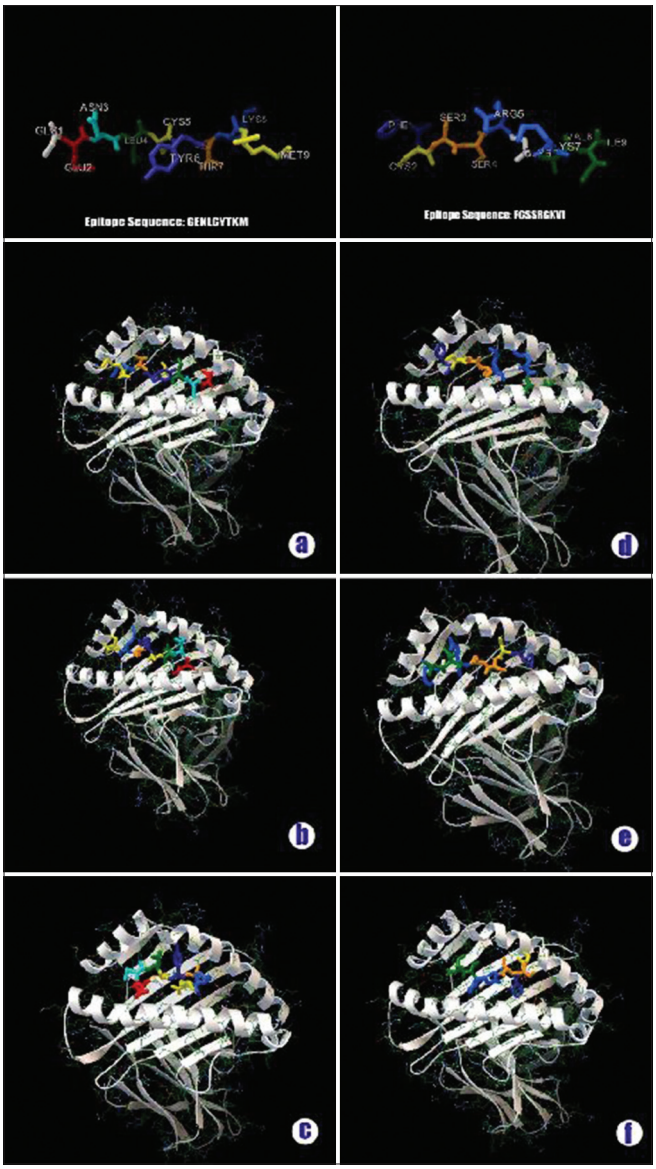

Figure 4: Visualization of docking results for predicted peptides with major histocompatibility complex Class I receptors by using AutoDock Tools. (a-c) docking images of "GENLCYTKM" with 1AQD, 1DLH, 1H15 respectively; (d-f) docking images of "FCSSRGKVI" with 1AQD, 1DLH, 1H15 respectively

Swiss-Prot: P25684.1), toxin S4C8 (UniProtKB/Swiss-Prot: P25683.10), toxin 4.9.6 (UniProtKB/Swiss-Prot: P01405.1), toxin C13S1C1 (UniProtKB/Swiss-Prot: P18329.10. cd00206 protein domain family represents a snake toxin domain that is present in short and long neurotoxins. This domain blocks the excitation of muscles by preventing binding of acetylcholine to the acetylcholine receptors at the postsynaptic membrane of skeletal muscles.[6] pfam00087 represents a family of venomous neurotoxins and cytotoxins.

The local hydrophilic region of the protein which is typically more exposed to the surface is detected as the antigenic site and the corresponding amino acids of these sites are detected as the antigenic peptides. Hopp and Woods hydrophilicity scale was used to predict the antigenic peptides of the selected venom protein. Hopp and Woods hydrophobicity scale is actually a hydrophilicity scale in which window size seven gives the ideal values for a protein hydrophilicity nature. Hopp and Woods scale assigns non-polar residues with a negative value. By analyzing the ProtScale results, the acetylcholine receptor binding domain of alpha-delta-bungarotoxin-4 was found to be hydrophilic in nature and may be this hydrophilicity favors the protein to interact the domain with the receptors.

The AllerHunter score value is the probability that a particular sequence is a cross-reactive allergen. However, the threshold for prediction of cross-reactive allergen is

Table 6: Binding site coordinates for protein-ligand docking between MHC molecules and peptides prepared by autodock tools

MHC molecule PDB ID	Axis	Center box	Size
3LKN	X	3.659	36
	Y	−15.259	22
	Z	−36.372	18
1A1O	X	3.12	44
	Y	26.631	22
	Z	19.201	16
1JHT	X	20.706	26
	Y	37.098	36
	Z	72.438	20
1AQD	X	12.939	34
	Y	24.708	18
	Z	43.286	22
1DLH	X	4.301	40
	Y	74.656	14
	Z	19.422	22
1H15	X	95.81	22
	Y	−5.497	16
	Z	16.03	36

Table 7: Docking simulation results prepared by autodock vina for predicted epitopes

Epitope sequence/ ligand	MHC	Receptor PDB ID	Affinity (Kcal/mol)	Dist. from RMSD l.b.	Best mode RMSD u.b.
GENLCYTKM	MHC I	3LKN	−6.5	0.0	0.0
		1A1O	−6.8	0.0	0.0
		1JHT	−5.5	0.0	0.0
	MHC II	1AQD	−6.0	0.0	0.0
		1DLH	−5.5	0.0	0.0
		1H15	−6.2	0.0	0.0
FCSSRGKVI	MHC I	3LKN	−6.8	0.0	0.0
		1A1O	−6.0	0.0	0.0
		1JHT	−6.5	0.0	0.0
	MHC II	1AQD	−5.6	0.0	0.0
		1DLH	−4.6	0.0	0.0
		1H15	−5.9	0.0	0.0

adjusted such that a sequence is predicted as a cross-reactive allergen if its probability is ≥0.06. The probability threshold was determined during the fine tuning of the prediction model. AllerHunter has optimum prediction result at that particular threshold. The FAO and WHO evaluation scheme is a guideline by the FAO and WHO for sequence based allergenicity prediction. This guideline clearly states that a sequence can be a potentially allergenic if it either has an approximated identity of at least six contiguous amino acids or >35% sequence identity over a window of 80 amino acid chains when compared with known allergens.[18] Hence, if a vaccine was developed by using the venom peptides, it will not create allergic reactions.

For the prediction of MHC binding molecules in both cases (MHC I and II), artificial neural network method was used. For T-cell Class I epitope prediction, the neural network method was designed by combining sparse encoding, blosum encoding and input derived from hidden Markov models.[15] MHC Class II molecules are highly polymorphic in nature, and this polymorphism exclusively corresponds with a few differences along the peptide binding groove in antigenic fragments.[19] The binding between antigenic peptides (epitopes) and the MHC molecule is a crucial step in the cellular immune response. In this study, for MHC Class II peptide prediction, we used artificial neural network based method NN-align, which was evaluated by 26 human MHC Class II alleles.[17] IC50 is a measure of half of a compound's concentration that would be required to inhibit biological effectiveness. Lower IC50 calculation reflects a drug's effectiveness in a lower concentration. A highly potent drug should be effective *in vivo* at a lower concentration to prevent consumption of the large amount of a given drug.[20-23] In this study, the epitopes were predicted at IC50 level of lower than 250 nM. Hence, a low dose of vaccine preparation by these peptides may be potent. Kolaskar and Tangaonkar antigenicity scale is the simplest method for determining antigenic determinants. This method is based on the occurrence of amino acid residues in experimentally determined epitopes. By comparing the MHC Class I, Class II and B-cell epitope predictions two nonameric peptides "GENLCYTKM" and "FCSSRGKVI" were found to be most common. The nonameric peptide "GENLCYTKM" interacted four MHC I alleles. "FCSSRGKVI" interacted with three MHC I alleles and seven MHC II alleles. These two peptides were the most interacting nonamers than others that were found during the prediction analysis.

Epitope conservancy check and population coverage are two important analysis steps that might reflect to the possibility of an epitope to be used in designing a vaccine. High epitope conservancy score indicates a good chance of effectiveness of epitope vaccine *in vivo*. Population coverage is a limitation for which reason drugs could be limited to a specific region or population. High population coverage of vaccine compound is significant due to a lot of people can be benefited by only one vaccine preparation. It was predicted that in combination of the two epitopes, more than 90% of the population of Kimberly, Europe, Czech, Yupik, North America, South America, Guarani-Kaiowa, Guarani-Nandewa, South West Asia was covered. The maximal population coverage was 99.24% for Guarani-Nandewa and minimal was 0% for Buriat. The prediction showed promising results that these epitopes may cover a high amount of population when applied as a vaccine.

In this study, we tried to minimize the predicted promiscuous epitopes and pin-point the efficient epitope sequences that have the greatest chance for eliciting humoral and cell mediated immunity in the human body against alpha-delta-bungarotoxin-4. As it is a concern that the prediction based epitope design might not work in reality, the epitopes were subjected to *in silico* validation by protein ligand docking simulation. Docking stimulation of the predicted MHC peptides along with HLA molecules was executed to investigate whether or not the designed vaccine will elicit the sufficient immunological responses *in vivo*. Lower energy scores represent better binding between receptor and ligand.[24] Docking simulation energy scores of the predicted epitopes were found significantly low. By summing up the prediction results, we hypothesize in this study a divalent peptide vaccine for immunization against alpha-delta-bungarotoxin-4.

CONCLUSION

All of these computational techniques manifest the attention of Krait neurotoxin as crucial immunodiagnostic tool for initial research methodologies in order to disease diagnosis and future drug design against this fatal venom. However, it is proved that most snake venoms are in poorly characterized although they are biologically important proteins with therapeutic potentialities. Hence, further studies and research in this field is quite obligatory to identify and characterize novel venom proteins in order to use it as a lead or structural templates for discovering new therapeutic agents in the near future. In addition, the above immunoinformatics attempt can be a new paradigm in improving immunotherapeutics, immunodiagnostics and gaining a better understanding of molecular autoimmune

susceptibility in a broader range. *B. caeruleus* alpha-delta-bungarotoxin-4 sequence is directly involved to empower and direct the immune system to protect the individual host from the bungarotoxin. Apart these, being cell surface proteins, MHC molecules play a superficial role in host immune reactions and respond to almost all antigens which render effects on specific sites. That's why, the predicted regions of MHC binding molecules act as red flags for detecting antigenic specificity, which generate immune responses comparatively against the parent antigen. Hence, a little fragment of antigen has potentialities to induce immune responses against whole antigen precisely. As a result, this novel method accumulates the prediction of MHC class binding peptides, TAP transport efficiency and proteosomal C terminal cleavage in a well fashioned manner. Hence, this superficial concept can be implemented to design synthetic and subunit peptide vaccine against lethal Krait venom that may save thousand lives especially in India, Sri Lanka, and Bangladesh. Thus, we opine that the given information and approaches in this study will be more blissful for researchers to investigate novel human therapeutics like design of subunit and synthetic peptide vaccine from snake venoms alpha-delta-bungarotoxin-4.

REFERENCES

1. WCH Clinical Toxinology Resources. *Bungarus caeruleus*. Available from: http://www.toxinology.com/fusebox.cfm?fuseaction=main.snakes.display&id=SN0015, http://www.toxinology.com/index.cfm. [Last accessed on 2013 Jun 17].
2. Reptilechannel.com. Walls Jerry G. Deadly Snakes: What are the world's most deadly venomous snakes? Available from: http://www.reptilechannel.com/snakes/venomous-snakes/deadly-snakes.aspx, http://www.reptilechannel.com/. [Last accessed on 2013 Jun 17].
3. Kularatne SA. Common krait (*Bungarus caeruleus*) bite in Anuradhapura, Sri Lanka: A prospective clinical study, 1996-98. Postgrad Med J 2002;78:276-80.
4. Alirol E, Sharma SK, Bawaskar HS, Kuch U, Chappuis F. Snake bite in South Asia: A review. PLoS Negl Trop Dis 2010;4:e603.
5. Wagstaff SC, Laing GD, Theakston RD, Papaspyridis C, Harrison RA. Bioinformatics and multiepitope DNA immunization to design rational snake antivenom. PLoS Med 2006;3:e184.
6. Dufton MJ. Classification of elapid snake neurotoxins and cytotoxins according to chain length: Evolutionary implications. J Mol Evol 1984;20:128-34.
7. Tsetlin V. Snake venom alpha-neurotoxins and other 'three-finger' proteins. Eur J Biochem 1999;264:281-6.
8. Wilkins MR, Gasteiger E, Bairoch A, Sanchez JC, Williams KL, Appel RD, *et al.* Protein identification and analysis tools in the ExPASy server. Methods Mol Biol 1999;112:531-52.
9. Garnier J, Osguthorpe DJ, Robson B. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins. J Mol Biol 1978;120:97-120.
10. Robson B, Garnier J. Protein structure prediction. Nature 1993;361:506.
11. King RD, Sternberg MJ. Identification and application of the concepts important for accurate and reliable protein secondary structure prediction. Protein Sci 1996;5:2298-310.
12. Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. Proc Natl Acad Sci U S A 1981;78:3824-8.
13. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett 1990;276:172-4.
14. Trott O, Olson AJ. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 2010;31:455-61.
15. Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, Buus S, *et al.* Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci 2003;12:1007-17.
16. Lundegaard C, Lamberth K, Harndahl M, Buus S, Lund O, Nielsen M. NetMHC-3.0: Accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8-11. Nucleic Acids Res 2008;36:W509-12.
17. Nielsen M, Lund O. NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. BMC Bioinformatics 2009;10:296.
18. Muh HC, Tong JC, Tammi MT. AllerHunter: A SVM-pairwise system for assessment of allergenicity and allergic cross-reactivity in proteins. PLoS One 2009;4:e5861.
19. Consogno G, Manici S, Facchinetti V, Bachi A, Hammer J, Conti-Fine BM, *et al.* Identification of immunodominant regions among promiscuous HLA-DR-restricted CD4+ T-cell epitopes on the tumor antigen MAGE-3. Blood 2003;101:1038-44.
20. Cheng Y, Prusoff WH. Relationship between the inhibition constant (K1) and the concentration of inhibitor which causes 50 per cent inhibition (I50) of an enzymatic reaction. Biochem Pharmacol 1973;22:3099-108.
21. Lazareno S, Birdsall NJ. Estimation of competitive antagonist affinity from functional inhibition curves using the Gaddum, Schild and Cheng-Prusoff equations. Br J Pharmacol 1993;109:1110-9.
22. Stewart MJ, Watson ID. Standard units for expressing drug concentrations in biological fluids. Br J Clin Pharmacol 1983;16:3-7.
23. Robinson SF, Marks MJ, Collins AC. Inbred mouse strains vary in oral self-selection of nicotine. Psychopharmacology (Berl) 1996;124:332-9.
24. Islam R, Sakib MS, Zaman A. A computational assay to design an epitope-based peptide vaccine against chikungunya virus. Future Virol 2012;7:1029-42.