Open Access Bioinformatics

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

# Antigenic epitopes prediction and MHC binder of a paralytic insecticidal toxin (ITX-1) of *Tegenaria agrestis* (hobo spider)

This article was published in the following Dove Press journal:
Open Access Bioinformatics
2 August 2010
Number of times this article has been viewed

AG Ingale

Department of Biotechnology, School of Life Sciences, North Maharashtra University, Jalgaon, India



**Abstract:** Spider peptide toxins with nanomolar affinities for their receptors are promising pharmacological tools for understanding the physiological role of ion channels and as leads for the development of novel therapeutic agents and strategies for ion channel-related diseases. Paralytic insecticidal toxin (*Tegenaria agrestis*) involved multiple antigenic components to direct and empower the immune system to protect the host from infection. MHC molecules are cell surface proteins, which take active part in host immune reactions and involvement of MHC class in response to almost all antigens, and it affects specific sites. Predicted MHC binding regions act like red flags for specific antigens and generate an immune response against the parent antigen. So a small fragment of antigen can induce an immune response against whole antigen. This theme is implemented in designing subunit and synthetic peptide vaccines. In this study, we analyzed secondary structure and antigenic determinants, which form antibodies against infection. The method integrates prediction of peptide MHC class binding and solvent accessible regions. Antigenic epitopes of paralytic insecticidal toxin are important antigenic determinants against the various toxic reactions and infections. There are 3 antigenic determinants in sequence. The results show highest pick at position 4–25 (QLMICLVLLPCFFCEPDEICRA) amino acid residue and 34–51 (YKSNVCNGCGDQVAACEA) amino acid residue.

**Keywords:** antigen prediction, modeling, ITX-1

## Introduction

The hobo spider, *Tegenaria agrestis*, is a member of the family of spiders known as the Agelenidae or funnel web weavers. Approximately 500 species of funnel web weavers occur worldwide; about 300 of these are found in North America, and about 100 species are native to Europe. The hobo spider is a member of the genus of spiders known colloquially as funnel web spiders. It is one of a small number of spiders in North America whose bites are generally considered to be medically significant. Although this species of spider has a reputation for aggressiveness, it will normally avoid contact with humans. Most bites occur when the spider is accidentally crushed or squeezed by a human. The spider's venom is strong enough to cause considerable local pain and necrosis.[1] Its nickname "aggressive house spider" comes from a misinterpretation of the Latin name "agrestis", which literally translates to "of the fields", but mistranslated as "aggressive". If a hobo spider is tending an egg sac, it may become aggressive if it perceives the egg sac to be threatened. However, they generally do not bite unless forced to protect themselves, and in the majority of cases, the hobo spider does not actually inject venom when it does bite. In the United States, the hobo spider has been considered to be a dangerous species based on a toxicology study on

Correspondence: AG Ingale
Department of Biotechnology, School of Life Sciences, North Maharashtra University, Jalgaon, India
Tel +910257-2257423
Fax +910257-2257406
Email agibiotech@gmail.com

rabbits where lesions appeared after spiders were induced to bite the rabbits,[2] although attempts to replicate the study (by injecting venom to ensure envenomation) have failed to produce necrotic lesions.[3] Many peptide toxins from spider venoms share structural features, amino acid composition and consensus sequences that allow them to interact with related classes of cellular receptors. They have become increasingly useful agents for the study of voltage-sensitive and ligand-gated ion channels and the discrimination of their cellular subtypes. Spider peptide toxins have also been recognized as useful agents for their antimicrobial properties and the study of pore formation in cell membranes. Their high insecticidal potency can also make them useful probes for the discovery of novel insecticide targets in the insect nervous system or for the development of genetically engineered microbial pesticides.

# Materials and methods

## Protein sequence analysis

The protein sequence of the paralytic insecticidal toxin ITX-1 of *Tegenaria agrestis* was retrieved from Swissprot sequence database UniProtKB/Swiss-prot (http://www.expasy.org/uniprot)[4] database release 54.0 NCBI gi|2920713|emb|CAA11839.1|. Here we have predicted secondary-structure MHC binding sites and solvent accessible regions in the protein sequence of ITX-1, having taxonomic information: family, Agelenidae; genus, Tegenaria Latreille; species, *Tegenaria agrestis*.[5]

## Prediction of antigenicity

Antigenic epitopes is determined using B-EpiPred Server, Kolaskar and Tongaonkar antigenicity methods. This program predicts those segments from within the protein sequence of ITX-1 that are likely to be antigenic by eliciting an antibody response. Predictions are based on a table that reflects the occurrence of amino acid residues in experimentally known segmental epitopes.[6]

## Prediction of protein secondary structure

The Jpred3 method predicted the secondary structure of the ITX-1 protein. Each residue is assigned values for alpha helix, beta sheet, turns and coils using a window of 7 residues. Using these information parameters, the likelihood of a given residue assuming each of the 4 possible conformations, alpha, beta, reverse turn, or coils calculated, and the conformation with the largest likelihood is assigned to the residue. The important concepts in secondary structure prediction are identified as: residue conformational propensities, sequence edge effects, moments of hydrophobicity, position of insertions and deletions in aligned homologous sequence, moments of conservation, autocorrelation, residue ratios, secondary structure feedback effects, and filtering.[7] The YASPIN secondary structure predictions method for the ITX-1 sequence were also used.

## Finding the location in solvent accessible regions

For setting the solvent accessible regions in protein, type of plot determine the hydrophobic scale and it is utilized for prediction. NetSurfP prediction server was used to predict the solvent accessible regions in protein. This may be useful in predicting membrane-spanning domains, potential antigenic sites, and regions that are likely exposed on the protein surface.[8–10]

## Prediction of MHC binding peptide (http://www.imtech.res.in/raghava/)

MHC2Pred predicts peptide binders to MHC class I and MHC class II molecules from protein sequences or sequence alignments using position-specific scoring matrices (PSSMs). This is a support vector machine (SVM) based method for prediction of promiscuous MHC class II-binding peptides. The average accuracy of SVM based method for 42 alleles is ~80%.[11,12] This method will be useful in cellular immunology, vaccine design, immunodiagnostics, immunotherapeutics, and molecular understanding of autoimmune susceptibility. For development of MHC binder, an elegant machine learning technique SVM has been used. SVM has been trained on the binary input of single amino acid sequence. In addition, we predict those MHC ligands from which the C-terminal end is likely to be the result of proteosomal cleavage. The identification of peptides that can stimulate cytotoxic T lymphocytes (CTLs) is one of the major challenges in subunit vaccines design. The existing epitope prediction methods are based on identification of MHC binding peptides.[13] It is not necessary that all MHC binders can act as T cell epitopes. There is a need to develop a highly accurate prediction method for CTL epitopes instead of MHC binders. The use of artificial neural network and SVM on the recent and high quality CTL epitopes and non-epitopes data is explored as a means to meet these challenges.

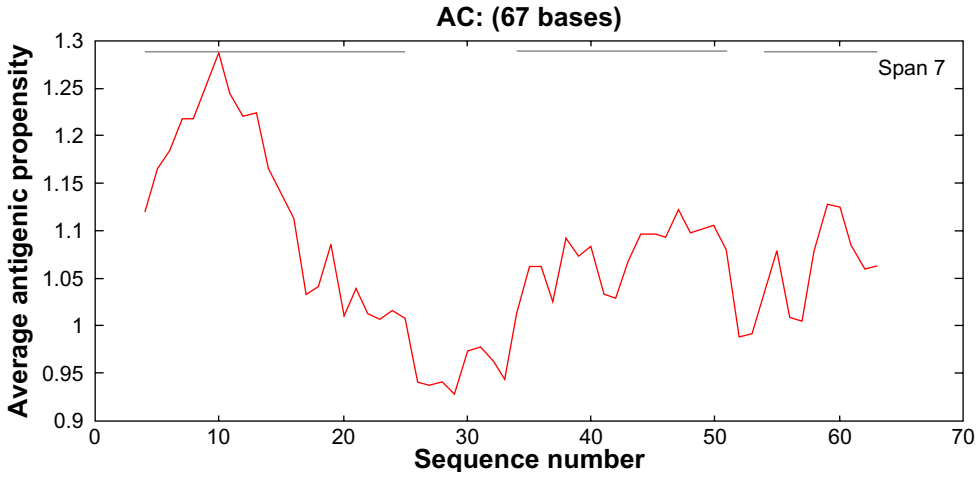


**Figure 1** Prediction of antigenic sites of the paralytic insecticidal toxin $ITX_1$ using Kolaskar and Tongaonkar.

# Results and discussions

## Prediction of antigenic peptides

The ITX-1 protein sequence is 68 residues long: MKLQLMICLV LLPCFFCEPD EICRARMTHK EFNYKSNVCN GCGDQVAACE AECFRNDVYT ACHEAQKG. As our knowledge of the immune responses to a protein antigen progressed, it became clear that the whole protein was not necessary for raising the immune response, but small segments of protein called the antigenic determinants or the epitopes is sufficient for eliciting the desired immune response. In these methods we found the antigenic determinants by finding the area of greatest local hydrophilicity. The Hopp–Woods[14] scale was designed to predict the locations of antigenic determinants in a protein, assuming that the antigenic determinants would be exposed on the surface of the protein and thus would be located in hydrophilic regions. Its values are derived from the transfer-free energies for amino acid side chains between ethanol and water. We also studied B-EpiPred Server, Kolaskar and Tongaonkar antigenicity methods, and the predicted antigenic fragments that bind to the MHC molecule is the first bottleneck in vaccine design. Figure 1 shows the antigenic determinant plot; x-axis shows sequence number and y-axis shows average antigenic propensity. Average antigenic propensity for this protein is 1.0636. There are 3 antigenic determinants in sequence. The highest pick at start position 4–25 amino acid residue and 34–51 amino acid residue. The sequence is "QLMICLV-LLPCFFCEPDEICRA", "YKSNVCNGCGDQVAACEA". The average for the whole protein is above 1.0; all residues above 1.0 are potentially antigenic. The highest pick sequence of antigenic determinant site is used for insertion. Highest pick in antigenic determinants plot indicate antigenic site for attachments. The ability of an individual antibody-combining site to react with only one antigenic determinant and the

```
aITX-1          : MKLQLMICLVLLPCFFCEPDEICRARMTHKEFNYKSNVCNGCGDQVAACEAECFRNDVYTACHEAQKG :
UniRef90_O46166 : MKLQLMICLVLLPCFFCEPDEICRARMTHKEFNYKSNVCNGCGDQVAACEAECFRNDVYTACHEAQKG :

                : 1---------11--------21--------31--------41--------51--------61------ :
OrigSeq         : MKLQLMICLVLLPCFFCEPDEICRARMTHKEFNYKSNVCNGCGDQVAACEAECFRNDVYTACHEAQKG :
Jnet            : --HHHHHHHHHHH------HHHHHH------------------EE----------HHHHHHHH---- :
jhmm            : --HHHHHHHHHHH-------HHHHHH-------------------------------HHHH-HHH---- :
jpssm           : --HHHHHHHHHHHHH-----HHHHHH------------------EEE---------HHHHHHH----- :

Lupas 14        : ------------------------------------------------------------------- :
Lupas 21        : ------------------------------------------------------------------- :
Lupas 28        : ------------------------------------------------------------------- :

Jnet_25         : --B-BBBBBBBBBBBB----BB-B-B-B--B-B---BB-BB---BB-B-B-BB---BB-BB------ :
Jnet_5          : ----B-BBBBBB-B-BB-----B-------------BB--B---B------B----B---B------ :
Jnet_0          : ------------------------------------B-------------------------------- :
Jnet Rel        : 840135788988800577643324321456777777777654410003555434433555033038 9 :
```

**Figure 2** Secondary structure prediction of paralytic insecticidal toxin $ITX_1$ using JPred3 (jnet3 algorithm).

```
  AA: MKLQLMICLVLLPCFFCEPDEICRARMTHKEFNYKSNVCNGCGDQVAACEAECFRNDVYTACHEAQKG
Pred: --EEEEEEEEEE-------HHHHHHH---------------HHHHHHHHHHH----HHHHHHH----
Conf: 912335665611110016201541100346201122201365201221000521111153533211 39

Hconf: 00000000000000000026999998600000000001100087898961999312299999992200
Econf: 05999999999862576000000000000015330002000000001020000000000000000000
Cconf: 94000000000137423973000001399984668996899912000018000687700000007799
```

**Figure 3** Secondary structure prediction of paralytic insecticidal toxin ITX$_1$ using YASPIN prediction method.

ability of a population of antibody molecules to react with only one antigen.

## Secondary alignment

The jnet3 method predicted the secondary structure of the ITX-1 protein. Each residue is assigned values for alpha helix, beta sheet, turns and coils using a window of 7 residues. Using these information parameters, the likelihood of a given residue assuming each of the 4 possible conformations, alpha, beta, reverse turn, or coils calculated, and the conformation with the largest likelihood is assigned to the residue. The more red a position is, the higher the level of conservation of chemical properties of the amino acids. Jnet show the final secondary structure prediction for ITX-1 protein. Lupas Coil prediction (window size of 14, 21, and 28) (dashed line in Figure 2) show less than 50% probability. Jnet-25 used for prediction of burial, less than 25% solvent accessibility. Jnet-5 show prediction of burial, less than 5% exposure, Jnet show prediction of burial, 0% exposure. Jnet Rel show reliability of prediction accuracy, ranges from 0 to 9, bigger is better.

The YASPIN secondary structure predictions for the ITX-1 sequence is directly under its corresponding amino acid. The numbers under each position are the confidence values for each prediction as calculated by the HMM. The higher the number from 0–9 the more confident the prediction (Figure 3). The values are separated into overall confidence (Conf), helix prediction confidence (Hconf), strand prediction confidence (Econf) and coil prediction confidence (Cconf).

## Solvent accessible regions

Solvent accessible prediction of ITX-1 protein was done using NetSurfP. Solvent accessible scales for delineating hydrophobic and hydrophilic characteristics of amino acids and scales are developed for predicting potential antigenic sites of globular proteins, which are likely to be rich in charged and polar residues. It was shown (Table 1) that a ITX-1 protein is hydrophobic in nature and contains segments of low complexity and high-predicted flexibility. The region of maximal hydrophilicity is likely to be an antigenic site, having hydrophobic characteristics, because C-terminal regions of ITX-1 protein is solvent accessible and unstructured; antibodies against those regions are also likely to recognize the native protein.

## The promiscuous MHC binding peptide prediction

These MHC binding peptides are sufficient for eliciting the desired immune response. The prediction is based on cascade SVM, using sequence and properties of the amino acids. The correlation coefficient of 0.88 was obtained by using jack-knife validation test. In this test, we found the MHC class I- and MHC class II-binding regions (Figures 4 and 5). MHC molecules are cell surface glycoproteins, which take active part in host immune reactions and involvement of MHC class I and MHC class II in response to almost all antigens.[15] In this assay we predicted the binding affinity of ITX-1 protein having 68 amino acids, which shows 60 different nonamers. For development of MHC binder prediction method, SVM has been used. SVM has been trained on the binary input of single amino acid sequence. In this assay we predicted the binding affinity of ITX-1 protein sequence having 68 amino acids, which shows 60 nonamers. MHC2Pred predicts peptide binders to MHC class I and MHC class II molecules from

**Table 1** Solvent accessible regions prediction of paralytic insecticidal toxin ITX$_1$ (partial results are shown here) using NetSurfp

| Class assignment | Amino acid | Amino acid number | Relative surface accessibility | Absolute surface accessibility |
|---|---|---|---|---|
| E | M | 1 | 0.729 | 145.853 |
| E | K | 2 | 0.595 | 122.371 |
| B | L | 3 | 0.226 | 41.399 |
| E | Q | 4 | 0.359 | 64.171 |
| B | L | 5 | 0.124 | 22.778 |
| B | M | 6 | 0.123 | 24.552 |
| B | I | 7 | 0.071 | 13.172 |
| B | C | 8 | 0.029 | 0.631 |
| B | L | 9 | 0.054 | 4.044 |
| B | V | 10 | 0.050 | 9.942 |
| B | L | 11 | 0.050 | 7.716 |
| B | L | 12 | 0.046 | 8.349 |

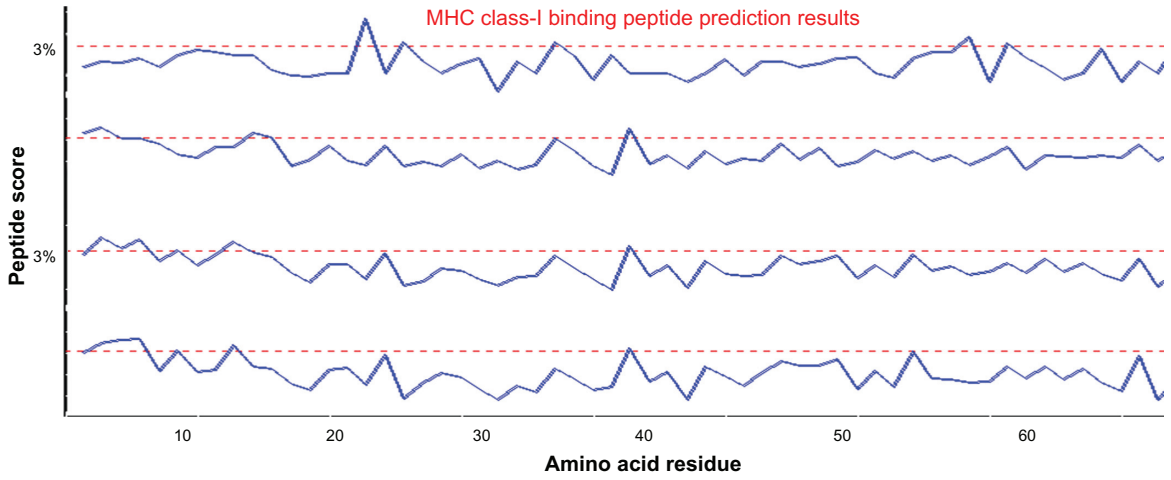


**Figure 4** The MHC class I binding peptide prediction of paralytic insecticidal toxin $ITX_1$. A peak crossing the dashed line means predicted binder starting from amino acid under peak to 9 amino acids ahead. This output is a quantitative way to look at prediction results, predicted binders, and other regions in antigen.

protein sequences or sequence alignments using PSSMs. In addition, we predict those MHC class I ligands which C-terminal end is likely to be the result of proteosomal cleavage.[16,17] The MHC peptide binding is predicted using neural networks trained on C terminals of known epitopes. Analysis predicted MHC peptide binding is a log-transformed value related to the IC50 values in nM units. These MHC binding peptides are sufficient for eliciting the desired immune response. The predicted binding affinity is normalized by the 1% fractil. The MHC peptide binding is predicted using neural networks trained on C terminals of known epitopes.[12]

B-EpiPred Server, Kolaskar and Tongaonkar antigenicity scales were designed to predict the locations of antigenic determinants in Paralytic insecticidal toxin alpha-ITX-1. ITX-1 protein shows beta sheets regions, which are high antigenic response than helical region of this peptide and shows high antigenicity.

The positions of the linear B-cell epitopes are predicted to be located at the residues with the highest scores (Figure 6). The annotation, given in the column denoted "?", is determined by the given threshold for the scores. A residue annotated with an "E" is predicted as being part of a linear B-cell epitope (score above threshold), where an "."

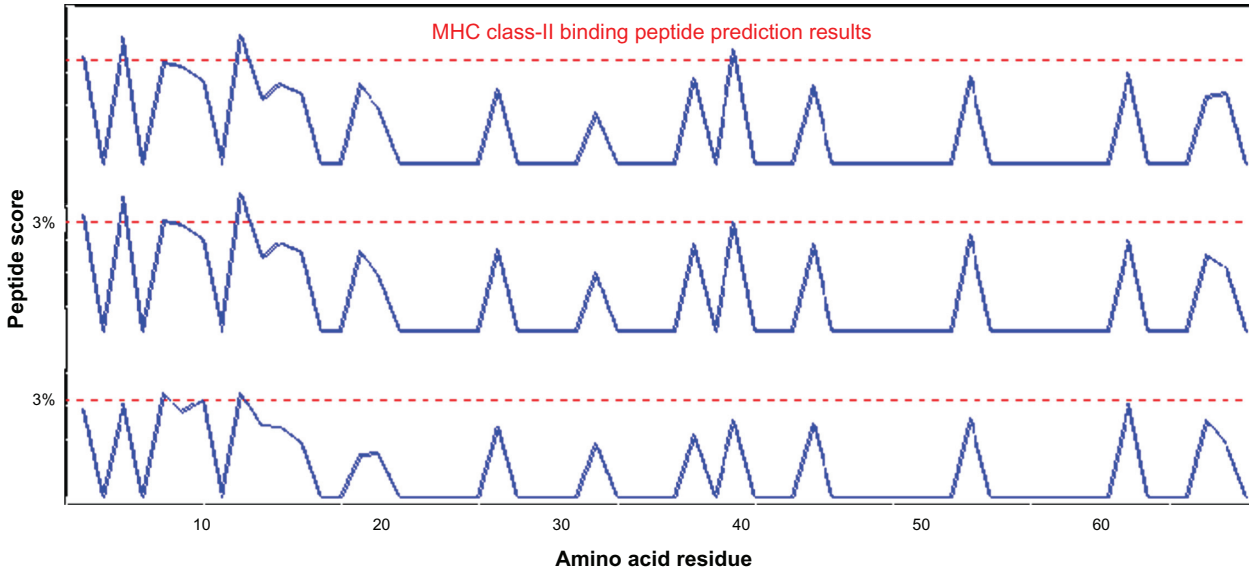


**Figure 5** The MHC class II binding peptide prediction of paralytic insecticidal toxin $ITX_1$. A peak crossing the dashed line means predicted binder starting from amino acid under peak to 9 amino acids ahead. This output is a quantitative way to look at prediction results, predicted binders, and other regions in antigen.

| # Seqname | Source | Feature | Start | End | Score | N/A | ? |
|---|---|---|---|---|---|---|---|
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 1 | 1 | -1.140 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 2 | 2 | -0.962 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 3 | 3 | -1.252 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 4 | 4 | -1.555 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 5 | 5 | -1.895 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 6 | 6 | -2.226 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 7 | 7 | -2.750 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 8 | 8 | -2.961 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 9 | 9 | -3.149 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 15 | 15 | -1.086 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 16 | 16 | -0.518 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 17 | 17 | 0.094 | . . | . |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 18 | 18 | 0.780 | . . | E |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 19 | 19 | 1.013 | . . | E |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 20 | 20 | 1.221 | . . | E |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 21 | 21 | 1.006 | . . | E |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 22 | 22 | 1.047 | . . | E |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 23 | 23 | 0.817 | . . | E |
| Sp_P01233_CGHB_HUMAN | bepipred-1.0b | epitope | 24 | 24 | 1.004 | . . | E |

**Figure 6** BEpipred prediction of paralytic insecticidal toxin ITX$_1$ shows the 'E' linear B-cell epitope, and "." represents a residue predicted not to be part of an epitope.

represents a residue predicted not to be part of an epitope (score below threshold).

We also found the Sweet hydrophobicity, Kyte and Doolittle hydrophobicity, Abraham and Leo. In this assay we predicted the binding affinity of ITX-1 protein having 68 amino acids, which shows 60 nonamers. Adducts of MHC and peptide complexes are the ligands for T cell receptors (TCR). MHC molecules are cell surface glycoproteins, which take active part in host immune reactions and involvement of MHC class I and MHC class II in response to almost all antigens. Kolaskar and Tongaonkar antigenicity are the sites of molecules that are recognized by antibodies of the immune system for the neurotoxin protein. Analysis shows epitopes present in the ITX-1 protein the desired immune response.

## Conclusion

Paralytic insecticidal toxin protein sequence (alpha-ITX-1) of the hobo spider involved multiple antigenic components to direct and empower the immune system to protect the host from the toxin. MHC molecules are cell surface proteins, which take active part in host immune reactions and involvement of MHC class in response to almost all antigens and it affects specific sites. Predicted MHC binding regions act like red flags for antigen specific and generate immune response against the parent antigen. So a small fragment of antigen can induce immune response against whole antigen. The method integrates prediction of peptide MHC class binding; proteosomal C terminal cleavage and TAP transport efficiency and identification of peptides that can stimulate CTLs. This theme is implemented in designing subunit and synthetic peptide vaccines.

## Disclosure

Author reports no conflict of interest in this work.

## References


1. Akre RD, Myhre EA. The great spider whodunit. *Pest Control Technology*. 1994;22(4):44–46, 50.
2. Vest DK. Envenomation by *Tegenaria* agrestis *(Walckenaer) spiders in rabbits*. *Toxicon*. 1987;25(2):221–224.
3. Vetter RS, Isbister GK. Do hobo spider bites cause dermonecrotic injuries? *Ann Emerg Med*. 2004;44:605–607.
4. Bairoch A, Boeckmann B, Ferro S, Gasteiger E. Swiss-Prot: juggling between evolution and stability. *Brief Bioinform*. 2004;5: 39–55.
5. Walckenaer CA. *Faune parisienne. Insectes. Ou histoire abrégée des insectes de environs de Paris*. Paris: Dentu; 1802:187–250.
6. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett*. 1990;276:172–174.
7. Cole C, Barber JD, Barton GJ. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res*. 2008;36:W197–W201.
8. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. *J Mol Biol*. 1982;157:105–132.
9. Sweet RM, Eisenberg D. Correlation of sequence hydrophobicities measures similarity in three-dimensional protein structure. *J Mol Biol*. 1983;171:479–488.
10. Abraham DJ, Leo AJ. Extension of the fragment method to calculate amino acid zwitterions and side chain partition coefficients. *Proteins*, 1987;2:130–152.
11. Brusic V, Rudy G, Honeyman G, Hammer J, Harrison L. Prediction of MHC class II-binding peptides using an evolutionary algorithm and artificial neural network. *Bioinformatics*. 1998;14(2):121–130.

12. Bhasin M, Raghava GPS. Pcleavage: an SVM based method for prediction of constitutive proteasome and immunoproteasome cleavage sites in antigenic sequences. *Nucleic Acids Res*. 2005;33: W202–W207.
13. Reche PA, Glutting JP, Reinherz EL. Prediction of MHC class I binding peptides using profile motifs. *Hum Immunol*. 2002;63:701–709.
14. Hopp TP, Woods KR. Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci U S A*. 1981;78: 3824–3828.
15. Sturniolo T, Bono E, Ding J, et al. Generation of tissue-specific and promiscuous HLA ligand database using DNA microarrays and virtual HLA class II matrices. *Nat Biotechnol*. 1999;17:555–561.
16. Nielsen M, Lundegaard C, Lund O, Kesmir C. The role of the proteasome in generating cytotoxic T cell epitopes: Insights obtained from improved predictions of proteasomal cleavage. *Immunogenetics*. 2005;57(1–2):33–41.
17. Kesmir C, Nussbaum A, Hansjorg S, Vincent D, Brunak S. Prediction of proteasome cleavage motifs by neural networks. *Prot Eng*. 2002; 15(4):287–296.