Gomase and Chitlange, 1:3
http://dx.doi.org/10.4172/scientificreports.191

Open Access Scientific Reports

Research Article Open Access

# Prediction of MHC Class Antigen Peptides from Echinococcus Multilocularis: Application of Computer Intelligence

**Gomase V.S* and Chitlange N.R**

*Department of Bioinformatics, JJT University, Jhunjhunu Rajasthan, 333001, India*

## Abstract

*Echinococcus multilocularis* is a cyclophyllid cestode that causes hydatid disease in many mammals, including rodents and humans and is becoming an increasing problem in urban areas. Peptide fragments of antigen protein can be used to select nonamers for use in rational vaccine design and to increase the understanding of roles of the immune system in infectious diseases. Analysis shows MHC class II binding peptides of antigen protein from *Echinococcus multilocularis* are important determinant for protection of host form parasitic infection. In this assay, we used PSSM and SVM algorithms for antigen design and predicted the binding affinity of antigen protein having 426 amino acids, which shows 417 nonamers. Binding ability prediction of antigen peptides to major histocompatibility complex (MHC) class I & II molecules is important in vaccine development from *Echinococcus multilocularis*.

**Keywords**: Antigen protein; Epitope; PSSM; SVM; MHC; Peptide vaccine

**Abbreviations:** GES: Goldman, Engelberg and Steitz; MHC: major histocompatibility complex; PSSMs: Position Specific Scoring Matrices; SVM: Support Vector Machine

## Introduction

*Echinococcus multilocularis* is a cyclophyllid cestode that causes hydatid disease in many mammals, including rodents and humans and is becoming an increasing problem in urban areas [1,2]. Echinococcus multilocularis antigen peptides are most suitable for subunit vaccine development because with single epitope, the immune response can be generated in large population. This approach is based on the phenomenon of cross-protection, whereby a plant infected with a mild strain of virus is protected against a more severe strain of the same virus. The phenotype of the resistant transgenic hosts includes fewer centers of initial virus infection, a delay in symptom development, and low virus accumulation. Antigen protein from *Echinococcus multilocularis* is necessary for new paradigm of synthetic vaccine development and target validation [3-5].

## Methodology

In this research work antigenic epitopes of antigen protein from *Echinococcus multilocularis* is determined using the Gomase in 2007, Bull & Breeze, Eisenberg, Rao & ArgosChou & Fasman and Levitt antigenicity [6-8]. The major histocompatibility complex (MHC) peptide binding of antigen protein is predicted using neural networks trained on C terminals of known epitopes. In analysis predicted MHC/peptide binding of antigen protein is a log-transformed value related to the IC50 values in nM units. MHC2Pred predicts peptide binders to MHCI and MHCII molecules from protein sequences or sequence alignments using Position Specific Scoring Matrices (PSSMs). Support Vector Machine (SVM) based method for prediction of promiscuous MHC class II binding peptides. SVM has been trained on the binary input of single amino acid sequence [9-14]. In addition, we predict those MHC ligands from whose C-terminal end is likely to be the result of proteosomal cleavage [15].

## Results and Interpretations

We found binding of peptides to a number of different alleles using Position Specific Scoring Matrix. An antigen protein sequence is 426 residues long, having antigenic MHC binding peptides. MHC molecules are cell surface glycoproteins, which take active part in host immune reactions and involvement of MHC class-I and MHC II in response to almost all antigens. PSSM based server predict the peptide binders to MHCI molecules of antigen protein sequence are as 11mer_H2_Db, 10mer_H2_Db, 9mer_H2_Db, 8mer_H2_Db and also peptide binders to MHCII molecules of antigen protein sequence as I_Ab.p, I_Ad.p, analysis found antigenic epitopes region in putative antigen protein (Table 1). We also found the SVM based MHCII-IAb peptide regions; MHCII-IAd peptide regions; MHCII-IAg7 peptide regions and MHCII- RT1.B peptide regions, which represented predicted binders from bacterial antigen protein (Table 2). The predicted binding affinity is normalized by the 1% fractil. We describe an improved method for predicting linear epitopes (Table 2). The region of maximal hydrophilicity is likely to be an antigenic site, having hydrophobic characteristics, because terminal regions of antigen protein is solvent accessible and unstructured, antibodies against those regions are also likely to recognize the native protein (Figure1, 2, 3). It was shown that a antigen protein is hydrophobic in nature and contains segments of low complexity and high-predicted flexibility (Figure 4, 5). Predicted antigenic fragments can bind to MHC molecule is the first bottlenecks in vaccine design.

## Conclusion

An antigen protein from *Echinococcus multilocularis* peptide nonamers are from a set of aligned peptides known to bind to a given MHC molecule as the predictor of MHC-peptide binding. MHCII molecules bind peptides in similar yet different modes and alignments

***Corresponding author:** Virendra S Gomase, Department of Bioinformatics, JJT University, Jhunjhunu Rajasthan, 333001, India, Tel: +91-9987770696; E-mail: gomase.viren@gmail.com

**Received** April 30, 2012; **Published** July 27, 2012

**Citation:** Gomase VS, Chitlange NR (2012) Prediction of MHC Class Antigen Peptides from Echinococcus Multilocularis: Application of Computer Intelligence. 1: 191. doi:10.4172/scientificreports.191

| MHC-I | POS. | N | Sequence | C | MW (Da) | Score | % OPT. |
|---|---|---|---|---|---|---|---|
| 8mer_H2_Db | 174 | SSL | SDVTPRYI | EDM | 932.05 | 19.64 | 37.41 % |
| 8mer_H2_Db | 300 | DRG | GPIPGTTD | SGS | 738.79 | 13.544 | 25.80 % |
| 8mer_H2_Db | 280 | LAD | GSSGGVTL | TSL | 658.7 | 11.871 | 22.61 % |
| 8mer_H2_Db | 219 | SAK | PNLSQIFV | GLR | 899.06 | 10.214 | 19.46 % |
| 8mer_H2_Db | 295 | TIT | SPDRGGPI | PGT | 779.86 | 9.066 | 17.27 % |
| 8mer_H2_Db | 60 | IAA | GLINPVQL | GIK | 835.01 | 7.939 | 15.12 % |
| 8mer_H2_Db | 374 | RPG | VPIRALYD | YVG | 928.11 | 6.218 | 11.85 % |
| 8mer_H2_Db | 343 | TSS | VNGAAAAI | SKE | 667.76 | 6.033 | 11.49 % |
| 8mer_H2_Db | 341 | DYT | SSVNGAAA | AIS | 657.68 | 5.954 | 11.34 % |
| 9mer_H2_Db | 216 | LDI | SAKPNLSQI | FVG | 939.08 | 15.774 | 31.32 % |
| 9mer_H2_Db | 107 | YKH | YKNVNRCKK | EYF | 1134.35 | 13.694 | 27.19 % |
| 9mer_H2_Db | 251 | YGV | DMAPNFPVF | QEY | 1019.19 | 13.637 | 27.08 % |
| 9mer_H2_Db | 59 | NIA | AGLINPVQL | GIK | 906.09 | 12.02 | 23.87 % |
| 9mer_H2_Db | 94 | NFD | SEFENAQKT | WYK | 1035.08 | 9.414 | 18.69 % |
| 9mer_H2_Db | 87 | TSI | KEVKNFDSE | FEN | 1077.16 | 7.325 | 14.54 % |
| 9mer_H2_Db | 19 | LRK | FAARLEMFL | RTG | 1079.34 | 7.156 | 14.21 % |
| 9mer_H2_Db | 390 | ADE | LSFNSGDLF | EKL | 981.08 | 7.076 | 14.05 % |
| 9mer_H2_Db | 342 | YTS | SVNGAAAAI | SKE | 754.84 | 5.826 | 11.57 % |
| 10mer_H2_Db | 185 | EDM | TQVFNKAQAF | ERE | 1135.28 | 16.784 | 28.52 % |
| 10mer_H2_Db | 87 | TSI | KEVKNFDSEF | ENA | 1224.34 | 14.513 | 24.66 % |
| 10mer_H2_Db | 14 | AYA | SNLRKFAARL | EMF | 1157.39 | 13.843 | 23.52 % |
| 10mer_H2_Db | 275 | KGR | SALADGSSGG | VTL | 802.8 | 13.092 | 22.24 % |
| 10mer_H2_Db | 310 | DSG | SNISTSPVHT | TAY | 1024.09 | 11.937 | 20.28 % |
| 10mer_H2_Db | 260 | PVF | QEYSPEMSAL | GKK | 1136.26 | 10.912 | 18.54 % |
| 10mer_H2_Db | 190 | VFN | KAQAFERERI | IYF | 1229.42 | 10.639 | 18.08 % |
| 10mer_H2_Db | 324 | AYG | SNSYDHGSEG | ATP | 1033.97 | 10.439 | 17.74 % |
| 10mer_H2_Db | 348 | GAA | AAISKEKQRV | EDT | 1111.31 | 10.178 | 17.29 % |
| 11mer_H2_Db | 113 | VNR | CKKEYFHACKT | VRS | 1339.59 | 25.822 | 32.48 % |
| 11mer_H2_Db | 410 | QGW | CKGRKDGRVGL | YPQ | 1170.39 | 17.044 | 21.44 % |
| 11mer_H2_Db | 59 | NIA | AGLINPVQLGI | KNW | 1076.3 | 16.853 | 21.20 % |
| 11mer_H2_Db | 340 | SDY | TSSVNGAAAAI | SKE | 943.02 | 16.171 | 20.34 % |
| 11mer_H2_Db | 347 | NGA | AAAISKEKQRV | EDT | 1182.39 | 14.568 | 18.33 % |
| 11mer_H2_Db | 87 | TSI | KEVKNFDSEFE | NAQ | 1353.46 | 9.581 | 12.05 % |
| 11mer_H2_Db | 146 | EQL | RKIEDKLRKGI | MEE | 1337.63 | 8.196 | 10.31 % |
| 11mer_H2_Db | 282 | DGS | SGGVTLTSLKT | ITS | 1045.18 | 6.949 | 8.74 % |
| 11mer_H2_Db | 185 | EDM | TQVFNKAQAFE | RER | 1264.4 | 6.609 | 8.31 % |

**Table 1:** PSSM based prediction of MHC ligands, from whose C-terminal ends are proteosomal cleavage sites.

| MHC ALLELE | Rank | Sequence | Residue No. | Peptide Score |
|---|---|---|---|---|
| I-Ab | 1 | RALYDYVGV | 377 | 1.101 |
| I-Ab | 2 | GVTLTSLKT | 284 | 1.061 |
| I-Ab | 3 | KTWYKHYKN | 101 | 0.952 |
| I-Ab | 4 | KGIMEEEKT | 154 | 0.934 |
| I-Ad | 1 | LSFNSGDLF | 390 | 0.841 |
| I-Ad | 2 | TATNILSGL | 35 | 0.722 |
| I-Ad | 3 | TGVEYGTAT | 29 | 0.669 |
| I-Ad | 4 | GSSGGVTLT | 280 | 0.640 |
| I-Ag7 | 1 | VNGAAAAIS | 343 | 1.876 |
| I-Ag7 | 2 | YGVDMAPNF | 248 | 1.865 |
| I-Ag7 | 3 | YVGVEADEL | 382 | 1.804 |
| I-Ag7 | 4 | SSVNGAAAA | 341 | 1.733 |
| RT1.B | 1 | KTVRSLQVQ | 122 | 1.377 |
| RT1.B | 2 | FKEQALQMQ | 202 | 1.230 |
| RT1.B | 3 | TTAYGSNSY | 319 | 0.947 |
| RT1.B | 4 | LRKFAARLE | 16 | 0.937 |

**Table 2:** SVM based prediction of promiscuous MHC class II binding peptides from antigen protein.

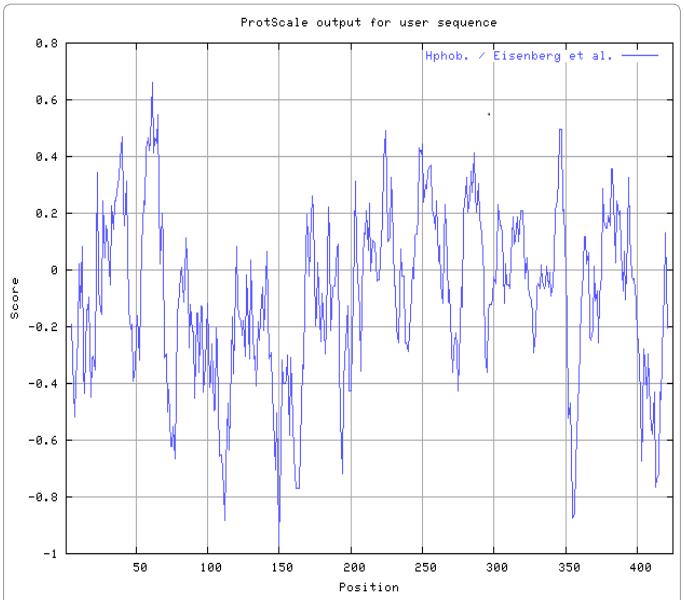


**Figure 1:** Hydrophobicity plot of antigen protein by Eisenberg, et al., scale.

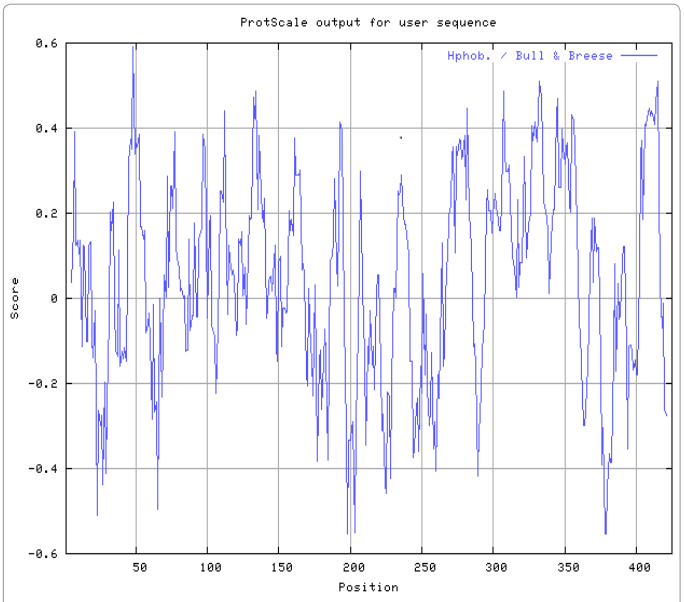


**Figure 2:** Hydrophobicity plot of antigen protein by Bull & Breese scale.

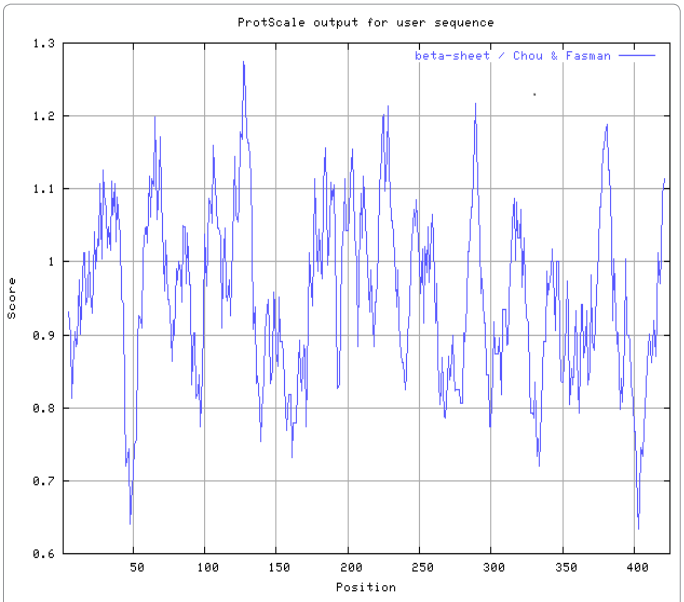


**Figure 4:** Antigenicity plot of antigen protein by beta-sheet / Chou & Fasman, scale.

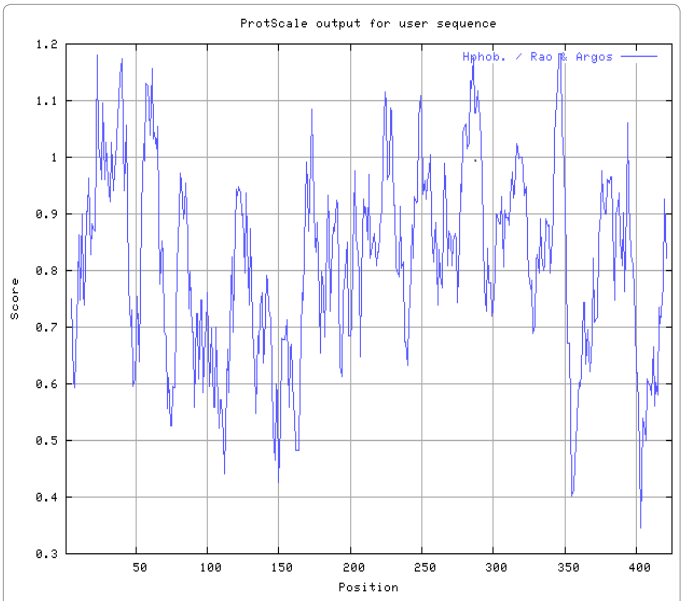


**Figure 3:** Hydrophobicity plot of antigen protein by Hphob./ Rao &Argosscale.

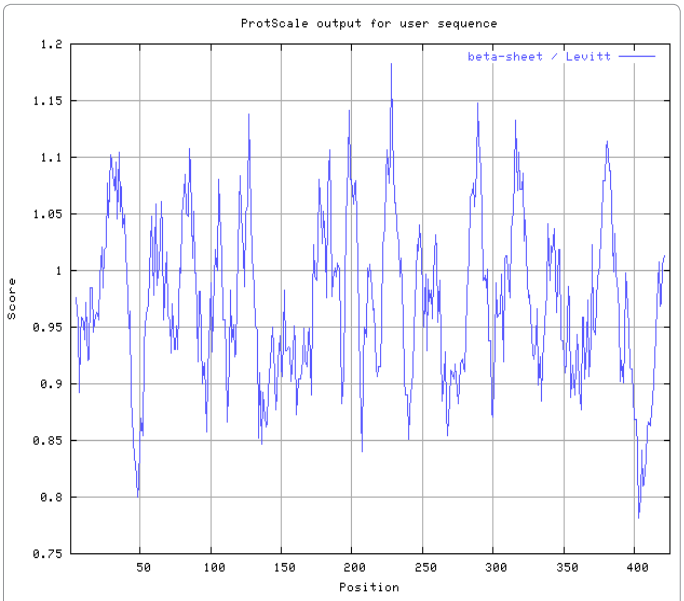


**Figure 5:** Antigenicity plot of antigen protein by Beta-Sheet / Levitt., scale.

of MHCII-ligands were obtained to be consistent with the binding mode of the peptides to their MHC class, this means the increase in affinity of MHC binding peptides may result in enhancement of immunogenicity of bacterial antigen protein. These predicted of antigen protein antigenic peptides to MHC class molecules are important in vaccine development from *Echinococcus multilocularis*.

## References

1. John DT, Petri Jr WA (2006) Markell and Voge's Medical Parasitology. (9th Edn), Elsevier, St. Louis.
2. Kayacan SM, Vatansever S, Temis S, Uslu B, Kayacan D (2008) Alveolar echinococcosis localized in the liver, lung and brain. Chin Med J 121: 90-92.
3. Joshi A, Tang J, Kuzma M, Wagner J, Mookerjee B, et al. (2009) Adenovirus DNA polymerase is recognized by human CD8+ T cells. J Gen Virol 90: 84-94.
4. McDonald D, Stockwin L, Matzow T, Blair Zajdel ME, Blair GE (1999) Coxsackie and adenovirus receptor (CAR)-dependent and major histocompatibility complex (MHC) class I-independent uptake of recombinant adenoviruses into human tumour cells. Gene Ther 6: 1512-1519.
5. Gomase VS, Kale KV, Shyamkumar K (2008) Prediction of MHC Binding Peptides and Epitopes from Groundnut Bud Necrosis Virus (GBNV). J Proteomics Bioinform 1: 188-205.
6. Gomase VS, Kale KV, Chikhale NJ, Changbhale SS (2007) Prediction of MHC Binding Peptides and Epitopes from Alfalfa mosaic virus. Curr Drug Discov Technol 4: 117-215.
7. Gomase VS, Kale KV (2008) In silico prediction of epitopes: a new approach for fragment based viral peptide vaccines. Int J of Applied Computing 1: 39-46.
8. Gomase VS, Kale KV (2008) Approach of proteomics system architecture in plant virus's database Int J of Applied Computing 1: 33-38.

9. Gomase VS, Kale KV (2008) Bioinformatics based sequence analysis of Nucleoplasmin like viral coat protein. Int J of Information Retrieval 1: 11-15.

10. Gomase VS and Kale KV, Shyamkumar K, Shankar S (2008) Computer Aided Multi Parameter Antigen Design: Impact of Synthetic Peptide Vaccines from Soybean Mosaic Virus. ICETET 2008, IEEE Computer Society in IEEE Xplore, Los Alamitos, California.

11. Gomase VS, Tandale JP, Patil SA, Kale KV (2006) Automatic modeling of protein 3D structure Nucleoplasmin-like viral coat protein from Cucumber mosaic virus". 14th International Conference on Advance Computing & Communication, Published by IEEE Computer Society in IEEE Xplore USA 614-615.

12. Reche PA, Glutting JP, Reinherz EL (2002) Prediction of MHC Class I Binding Peptides Using Profile Motifs. Hum Immun 63: 701-709.

13. Buus S, Lauemøller SL, Worning P, Kesmir C, Frimurer T, et al. (2003) Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach. Tissue Antigens 62: 378-384.

14. Nielsen M, Lundegaard C, Worning P, Lauemøller SL, Lamberth K, et al. (2003) Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. Protein Sci 12: 1007-1017.

15. Bhasin M, Raghava GP (2005) Pcleavage: an SVM based method for prediction of constitutive proteasome and immunoproteasome cleavage sites in antigenic sequences. Nucleic Acids Research 33: W202-207.