

Identification of Antigenic Determinants, Solvent Accessibility and MHC Binders of Antigen NADH Dehydrogenase Subunit 5 from a Little Dragon from Medina (*Dracunculus medinensis*)

Sonu Mishra and Virendra S Gomase

Department of Biotechnology, Mewar University, Chittorgarh, India

Abstract

Dracunculiasis caused by the pathogen *Dracunculus medinensis* (a little dragon from Medina). The Cyclops is the intermediate host that ingests the larvae of parasite (*D. medinensis*), that later on ingested by the human from the contaminated stagnant unfiltered water from the water source. After an incubation period of a year these mature female worm comes towards the skin and start formation of a small round bulge on the skin by secreting an irritating chemical which causes severe pain. In this study we will summarize the potency of NADH dehydrogenase subunit 5 (mitochondrion) from *Dracunculus medinensis* with 527 amino acids. Antigenic peptide of NADH dehydrogenase subunit 5 protein is most suitable for subunit vaccine development because with single epitope, the immune response can be generated in large population. In this research, we used PSSM and SVM algorithms for the prediction of MHC class I and II binding peptide, antigenicity, Solvent accessibility, polar and nonpolar residue to analyse the regions that are likely exposed on the surface of proteins which are potentially antigenic that allows potential drug targets to identify active sites against infection as well as to design effective drug to treat it.

Keywords: Dracunculiasis, Antigenic peptides; MHC-Binders; TapPred; PSSM; SVM; Nonamers; NADH; Dehydrogenase subunit 5

Abbreviations: MHC I: Major Histocompatibility Complex-Class I; MHC II: Major Histocompatibility Complex-Class II; PSSM: Position Specific Scoring Matrices; SVM: Support Vector Machine; GWD: Guinea worm disease; UniProt: The Universal Protein Resource; NCBI: National Center for Biotechnology Information; TAP: Transporter Associated with Antigen Processing; HPLC: High Performance Liquid Chromatography

Introduction

NADH dehydrogenase 5 (ND5) protein contribute to be a part of the large enzyme complex i.e Complex I, which is active in mitochondria. Complex I is an one of the essential enzyme complexes necessary for oxidative phosphorylation and thereafter, mitochondrial enzyme complexes execute chemical reactions that contribute to the ATP production. Complex I is responsible for the first step in the electron transport process. The NADH dehydrogenase complex (or mitochondrial respiratory complex I) plays an important role in the oxidation of NADH by ubiquinone by catalysis. This chemical reaction is linked to proton transfer across mitochondrial membranes. Study reveals that there are multiple humans disorders are consociated with mutations of the mitochondrial ND genes [1]. Mutations in mitochondrial DNA of ND5 have been found to be the most important cause of several disorder like hypertrophic cardiomyopathy (homoplasmic ND5 12338T>C variant) [2], Leber's hereditary optic neuropathy (LHON) which is a maternally inherited eye disease (homoplasmic ND4 G11696A and ND5 T12338C mutation) [3], Light syndrome (ND5*13513 G to A mutation) [4], Leber's hereditary optic neuropathy is associated with the T12338C mutation in mitochondrial ND5 gene in six Han Chinese families [5]. The heteroplasmic T>C transition at position 13271 in MTND5 affects a highly conserved base and segregates with the disease. Investigation suggest that, this is the 15th mutation affecting the MTND5 subunit of respiratory chain complex I and affirms its importance site for disease with phenotypes ranging from MELAS and infantile encephalopathies to isolated syndromes affecting a single tissue such as Leber hereditary optic neuropathy and now skeletal muscle [6]. Looking at its important we have used this protein to study the antigenicity of protein, its solvent

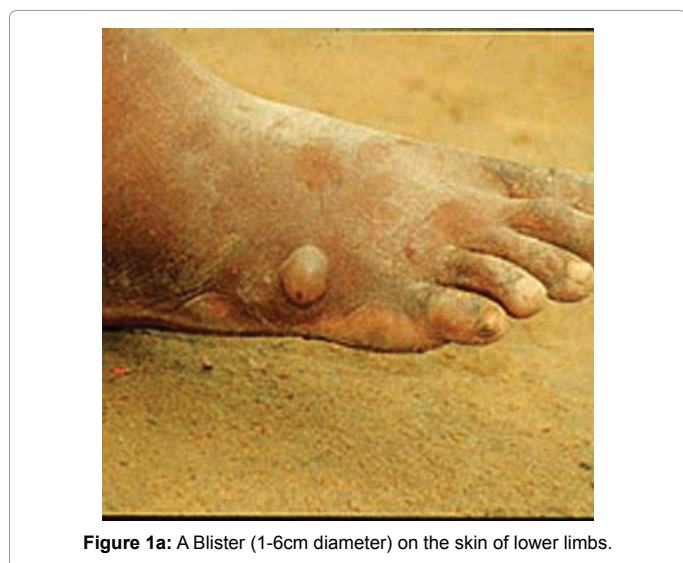
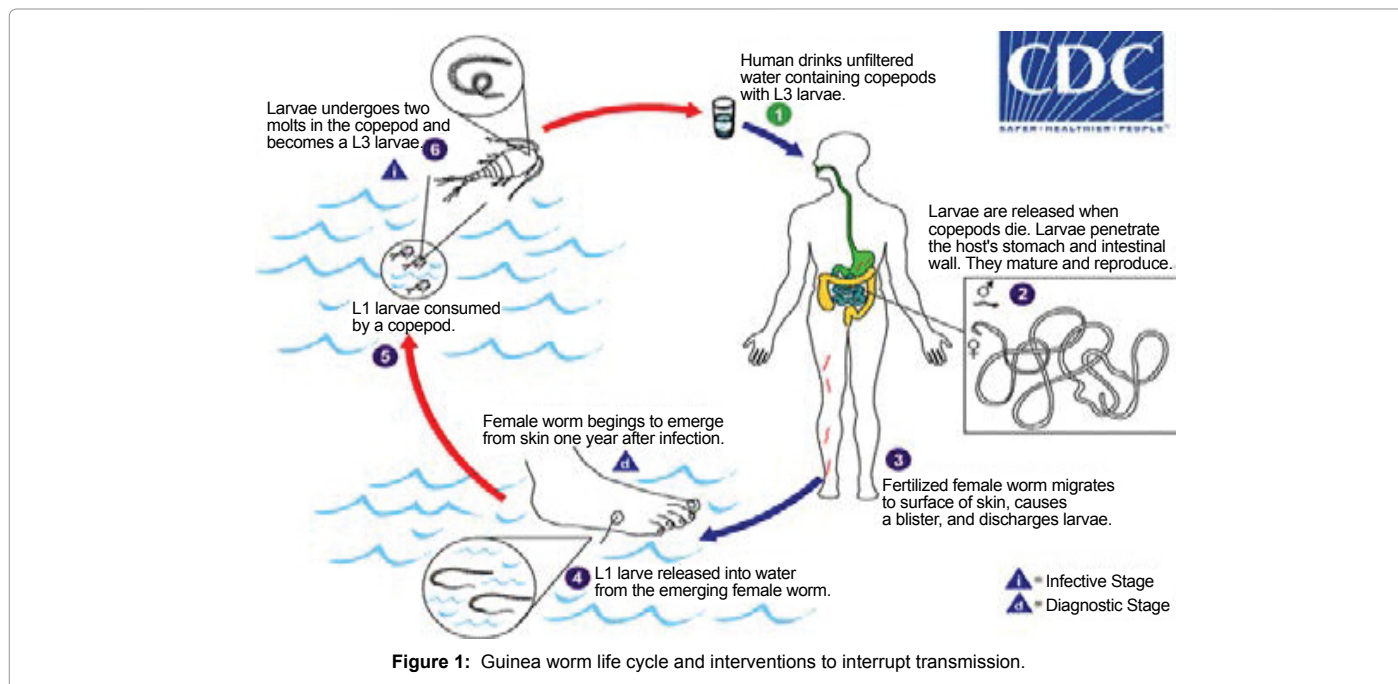
accessibility, polar and nonpolar residue to analyse the regions that are likely exposed on the surface of proteins which could be the potentially antigenic that allows potential drug targets to identify active sites against infection as well as to design effective drug to treat *Dracunculiasis*. *D. medinensis* (a little dragon from Medina) is the causative agent and it is the only species from the 12 species of *Dracunculus* [7-10] which infects humans, commonly known as "Guinea Worm Disease (GWD)". The other *Dracunculus* species generally resides in the internal tissues and body cavities of non-human mammals and reptiles (snake and turtles) [11]. This little dragon undergo a very unusual life cycle of six developmental stages with incubation period last for 1 to one and half years approximately (Figure 1) [12]. This is one of the most neglected tropical parasites which bears clinical importance and needs to be eradicated after small pox [13]. After reaching to the maturation stage, these worms copulate and an adult female produces millions of eggs in its uterus whereas male dies. Later on, the female worm release the larvae which induces a painful blister (1-6 cm diameter) on the skin of lower limbs (predominantly localized in the lower extremities (80-90%) in most of the reported cases) (Figure 1a). The infected person develops slight fever, local skin redness, swelling and severe pruritus around the blister. Other symptoms include: diarrhea, nausea, vomiting and dizziness. The blister burst within three days and female worms one or more slowly comes out from the wounds which causes an excoriating burning sensation and pain [14] (Figure 1b). Immersing or pouring water over the blister provides pain reliever. But this the moment that adult female is exposed to the external environment [15] During emergence of the limbs in open water sources it recognizes

*Corresponding author: Sonu Mishra, Department of Biotechnology, Mewar University, Chittorgarh, India, Tel: 01471 220 881, E-mail: sonumishra1014@gmail.com

Received January 04, 2016; Accepted January 28, 2016; Published February 05, 2016

Citation: Mishra S, Gomase VS (2016) Identification of Antigenic Determinants, Solvent Accessibility and MHC Binders of Antigen NADH Dehydrogenase Subunit 5 from a Little Dragon from Medina (*Dracunculus medinensis*). Drug Des 5: 126. doi: 10.4172/2169-0138.1000126

Copyright: © 2016 Mishra S, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



the temperature difference and releases the milky white liquid in the water which contains millions of immature larvae, when larvae released in water are ingested by copepods where they moult twice and become infective larvae within two weeks [16].

The *D. medinensis* antigen peptides can be most desirable segment for the subunit vaccine development because with the single epitope, the immune response can be generated in large population. This approach is usually based on the phenomenon of cross-protection, whereby infected with the mild strain and is protected against a more severe strain of the same. The phenotype of the resistant transgenic hosts includes fewer centers of initial infection, a delay in symptom development and low accumulation. The World Health Assembly called in the 1986 in order to dracunculiasis elimination. The global Guinea Worm Eradication Program, supported by The Carter Center, World Health Organization (WHO), UNICEF, CDC, and other partners, began assisting ministries of health of countries in which dracunculiasis is endemic in meeting this goal. At that time, an estimated 3.5 million cases occurred each year in 20 countries in Africa and Asia. Dracunculiasis remains endemic in four countries in 2014 (South Sudan, Chad, Mali, and Ethiopia), but the number of overall reported incidence is decreases in 2013 by 73% and in 2014 by 71% compared with 2012. Failures in surveillance and containment, lack of clean drinking water, insecurity in Mali and parts of South Sudan, and an unusual epidemiologic pattern in Chad are the main remaining challenges to dracunculiasis eradication [17]. A case of *Onchocerca volvulus* has been reported in the Cameroon which is mimicking *Dracunculus medinensis* [18]. More than two decades after the International Drinking Water Supply and Sanitation Decade (IDWSSD) implemented by the United Nations (1981-1990) [19], the disease still lingers, underscoring the daunting challenge of disease control, as has been the case of the failure of previous attempts to eradicate diseases like malaria, hookworm and yaws [20]. Till date there is no accurate and efficient curative drug of vaccine is available against dracunculiasis [21]. The investigation suggests that the immunity is not developed by the infected individual [22]. *Dracunculus medinensis*-specific antibodies (total, IgG1 and IgG4) during the time of patency, which were significantly higher than the levels measured in the same individuals eight months later, except for a few individuals who had developed a new patent infection [23]. It was observed that the mean level of specific IgG1 and IgG4 is higher during the month of potency of the infected individual whereas variation in the IgE value is relatively

negotiable and constant before, during the infection and after the recovery. There is possibility that variation in antibody production is regulated by infected larvae (i.e by transmission) and / or by adult worms (i.e by patency) is still need to be clear out. It is possibility that increased production of IgG1,IgG4 during the time of patency plays a role in blocking or protecting immune responses otherwise it could have killed ingested infected larvae [24]. The *D. medinensis* antigen peptides can be most desirable segment for the subunit vaccine development because with the single epitope, the immune response can be generated in large population. This approach is usually based on the phenomenon of cross-protection, whereby infected with the mild strain and is protected against a more severe strain of the same. The resistant transgenic host's phenotype includes of fewer centers of initial infection, following a delay development in symptom with low accumulation. Antigenic peptides from *D. medinensis* is most suitable for the development of peptide vaccine [25] because a single protein subunit can generate sufficient immune response. In this research work we have used the phenomenon of cross-protection, whereby an individual undertaken by a mild toxin can have immunity to survive against similar strong toxic effects. MHC molecules are cell surface protein that binds to the peptides derived from host or antigenic proteins and present them to cell surface for recognition by T-cells. T cell recognition is an important mechanism of the adaptive immune system by which the host identifies and responds to foreign antigens [25,26]. There are two types of MHC molecule and are extremely polymorphic. MHC class I molecules present peptides from proteins synthesized within the cell, whereas, MHC class II molecule present peptides derived from endocytosed extracellular proteins. MHC molecules have been well characterized due to their role in immune reactions and they take active part in host immune reactions and involvement of MHC class molecule in response to almost all antigens and it give impacts on specific sites. The involvement of MHC class-I molecule in response to almost all antigens make the study very interesting. They bind to some of the peptide fragments generated after proteolytic cleavage of antigen

[27]. Identification of MHC-binding peptides and T-cell epitopes helps improve our understanding of specificity of immune responses [28]. Antigenic peptides are most suitable for peptide vaccine development because single epitope can generate large the immune response [29].

Methodology

Database searching

The antigenic protein sequence of NADH dehydrogenase subunit5 from *Dracunculus medinensis* was retrieved from www.ncbi.nlm.nih.gov, UniProt databases are initially the most important [30-32].

Prediction of antigenicity

Prediction of antigenicity program predicts those segments from Antigen NADHdehydrogenasesubunit5 protein that are likely to be antigenic by eliciting an antibody response. In this research work antigenic epitopes of *Dracunculus medinensis* antigen NADHdehydrogenasesubunit5 are determined by using the Hopp and Woods, Welling, Parker, Bepipred ,Kolaskar and Tongaonkar antigenicity methods [33-38].

Prediction of MHC binding peptide

The major histocompatibility complex (MHC) peptide binding of *Dracunculus medinensis* predicted using neural networks trained on C terminals of known epitopes. Rankpep predicts peptide binders to MHC-I ligands whose C-terminal end is likely to be the result of proteosomal cleavage using Position Specific Scoring Matrices (PSSMs). Peptides that bind to a given MHC molecule share sequence similarity. Traditionally, the sequence patterns used for the prediction of peptides binding to MHC molecules. Such sequence patterns, however, have proven to be too simple, as the complexity of the binding motif cannot be precisely represented by the few residues present in the pattern [39]. RANKPEP uses "Position Specific Scoring Matrices (PSSMs) or profiles" from set of aligned peptides known to bind to a given MHC molecule as the

MHC-I Allele	RANK	POS.	N	SEQUENCE	C	MW (Da)	SCORE	% OPT.
8mer_H2_Db	1	262	STL	SQIGFCFF	GLG	930.1	20.659	39.35%
8mer_H2_Db	7	63	GSV	VVYSGFYM	MED	947.12	13.569	25.85%
8mer_H2_Db	8	306	LCG	GQQDSRGY	VGW	891.9	12.539	23.89%
8mer_H2_Db	9	512	FSG	KEINFLFL	CLF	1005.23	12.34	23.51%
8mer_H2_Db	10	419	LWW	LNYSFVI	SSS	951.09	12.251	23.34%
9mer_H2_Db	1	418	FLW	WLNYNSFVI	SSS	1114.3	22.359	44.39%
9mer_H2_Db	2	71	FYM	MEDYNFNYF	CVV	1224.32	20.942	41.58%
9mer_H2_Db	3	95	GVV	FSNNCISML	IFW	1010.19	19.364	38.45%
9mer_H2_Db	4	332	VSL	MCLCGLFFL	GGG	1028.36	18.362	36.46%
9mer_H2_Db	6	101	NCI	SMLIFWDL	GVS	1096.39	16.749	33.26%
10mer_H2_Db	1	179	FTK	SAQYPFSSWL	PKA	1144.3	29.252	49.70%
10mer_H2_Db	4	109	WDL	LGVSYYFLVL	YYG	1079.31	12.086	20.53%
10mer_H2_Db	6	305	HLC	GGQDSRGYV	GVG	1048.08	9.329	15.85%
10mer_H2_Db	8	395	LFY	SGGSLVFSFV	SLV	981.12	8.591	14.60%
10mer_H2_Db	9	433	LLY	SDYYSLFYLF	ILG	1299.47	8.398	14.27%
11mer_H2_Db	3	263	TLS	QIGFCFFLGL	GLV	1183.44	16.546	20.81%
11mer_H2_Db	5	61	VVG	SVVYSGFYMM	EDY	1264.52	15.559	19.57%
11mer_H2_Db	6	368	LFF	FSILLTYLYCY	RLM	1380.68	13.326	16.76%
11mer_H2_Db	7	261	LST	LSQIGFCFFGL	GLG	1213.47	12.853	16.17%
11mer_H2_Db	8	93	MVG	VVFSNNCISML	IFW	1208.45	12.594	15.84%

Table 1: Promiscuous MHC ligands, having C-terminal ends are proteosomal cleavage sites of *Dracunculus medinensis*. The antigenic peptide to the MHC-I Allele i.e. 8mer_H2_Db(The binding thresholds:33.04, optimal score:52.494), 9mer_H2_Db(Optimal Score: 50.365, Binding Threshold: 17.96), 10mer_H2_Db(The Optimal Score: 58.858, BindingThreshold: 41.32),11mer_H2_Db(Optimal Score: 79.495, Binding Threshold: 56.96). (All rows highlighted in red represent predicted binders and A peptide highlighted in violet has a C-terminus predicted by the cleavage model used).

predictor of MHC-peptide binding and overcome the complexity of the binding motif limitation. RANKPEP web server is a variability masking feature to focus on the prediction of conserved epitopes, which could thus help to avoid immune evasion resulting from mutation. Support Vector Machine (SVM) based method for prediction of promiscuous MHC class II binding peptides from protein sequence; SVM has been trained on the binary input of single amino acid sequence [40-43].

Prediction of antigenic peptides by Cascade SVM based TAPPred method

In the present study, we predict cascade SVM based several TAP binders which was based on the sequence and the features of amino acids [44]. We found the MHCI binding regions (Table 1), the binding affinity of *Dracunculus medinensis*.

Solvent accessible regions

We also analyzed the solvent accessible regions of proteins having highest probability that a given protein region lies on the surface of a protein Surface Accessibility, backbone or chain flexibility by Emini et al. [45] and Karplus and Schulz [46]. By using different scale we predict the hydrophobic and hydrophilic characteristics of amino acids that are rich in charged and polar residues [47-56].

Results

The *Dracunculus medinensis* Antigen NADH dehydrogenase subunit 5, contain a long residue of 527 amino acids with 483 nonamers.

MDVYIYWFGVILLCFLLLFVFFYDDFVFSLDFSSLELLQFQ-FRLDWFSFSFFCLLMVVGSVVVVYSGFYM

MEDYNFNFCVLSIFVFSMVG VVFSNNCISMLIFWDLG-VSSYFLVLYGNWDSCSGSMNTVMMNRVGDVCFVLFVFCGLFFMGIDFISLELVVSVALFFFILSTFTKSAQYPFSSWLPKAMSAPTVPVLSALVHSSSTLVTAGLFLGMCFSVEMFLDFVLDFFVGLFTMFSSGLMAYFEFDIKKLVALSTLSQIGFCFFGLGLGLVYFSFIHMLSHAVFKSCLFMQMGYIIHLCGGQQDSRQYVGVGGLSSVVYIQTFVSLMCLCGLFFLGGVSKEILLEHYFFCNWSLFLVFLFFSILLTYLYCYRLMKGFYYCSSLFYSGGSLVFSFVSLVLVVSIVFLWLN

YNSFVISSLLYSDYYSLFYLFILGLVLCVVFVKFGSFD-VKYKFGDLPKVIIRGNVVKWSDCMVDYSIMKFGDFS-FYVSKIFVMGFSGKEINFLFLCLLLMI

Prediction of antigenic peptides

In this study, we found the antigenic determinants by finding the area of greatest local hydrophilicity. The Hopp-Woods scale Hydrophilicity Prediction Result Data found high in position between 300-320 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 (Figure 2). Welling antigenicity plot gives value as the log of the quotient between percentage in a sample of known antigenic regions and percentage in average proteins and Prediction Result Data found high in position 250-270 (Figure 3). We also study Hydrophobicity plot of HPLC / Parker Hydrophilicity Prediction Result Data found between 300-320 (Maximum Score-6.3) i.e., the maximum predicted residues at the position 308(Residue is Q) is 305- GGQQDSR -311 and 309 (Residue D) is 306- GQQDSRG-312 (Figure 4), BepiPred predicts the location of linear B-cell epitopes Result found that between 300-310 is 301-MQMGYIIHLC-310 and the maximum score (1.369) is found at the position 308,309 (Figure 5), Kolaskar and Tongaonkar antigenicity methods (Figure 6) Predicted peptides result found i.e.,

4-YIYWFGVILLCFLLLFVFFYDDFVFSLDFSSLELLQFQFRLD-WFSFSFFCLLMVVGSVVVVYS-66

76-FNYFCVLSIFVFSMVG VVFSN-97

99-CISMLIFWDLG VSSYFLVLY-119

138-VGDVCFVLFVFCGLFF-152

156-DFISLELVVSVALFFFILST-175

178-KSAQYPFSSWLPKA-191

193-SAPTVPVLSALVHSSSTLVTAGLFLGMCFSVEMFLDFVLD-FMFFVGLFT-238

251-DIKKLVALSTLSQIGFCFFGLGLGLVYFSFIHMLSHAVFK-SCLFM-295

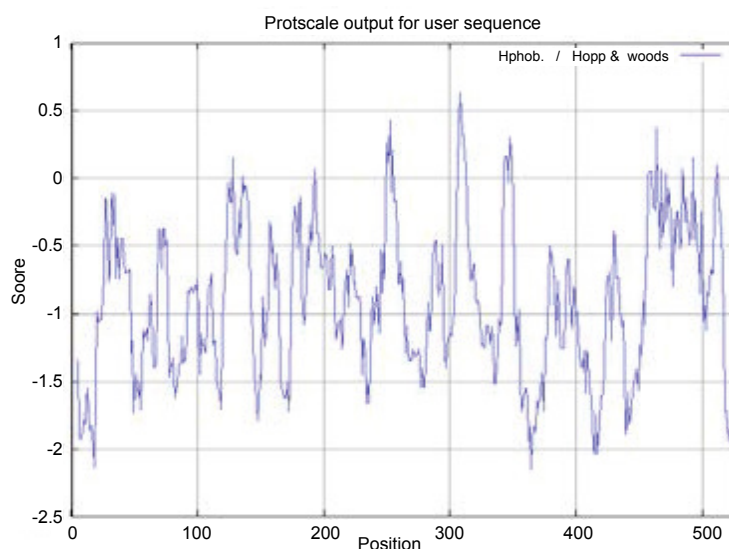


Figure 2: Hydrophobicity plot of Hopp and Woods.

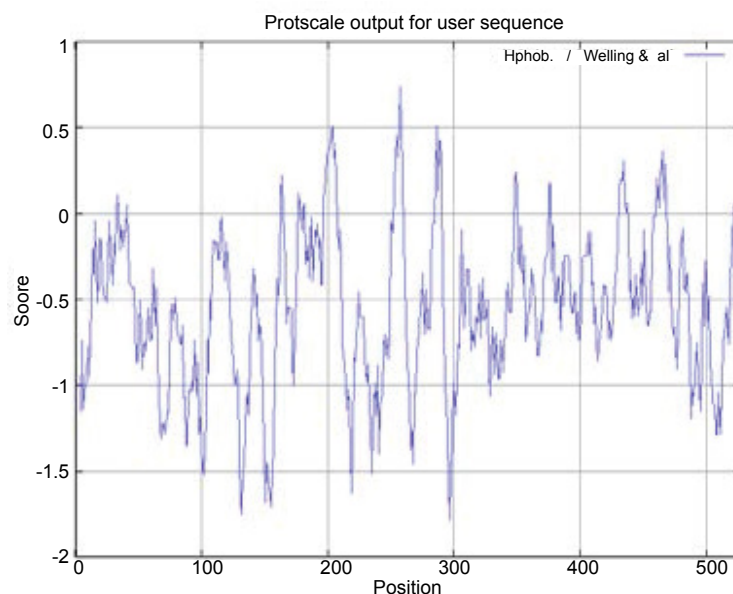


Figure 3: Hydrophobicity plot of Welling et al.

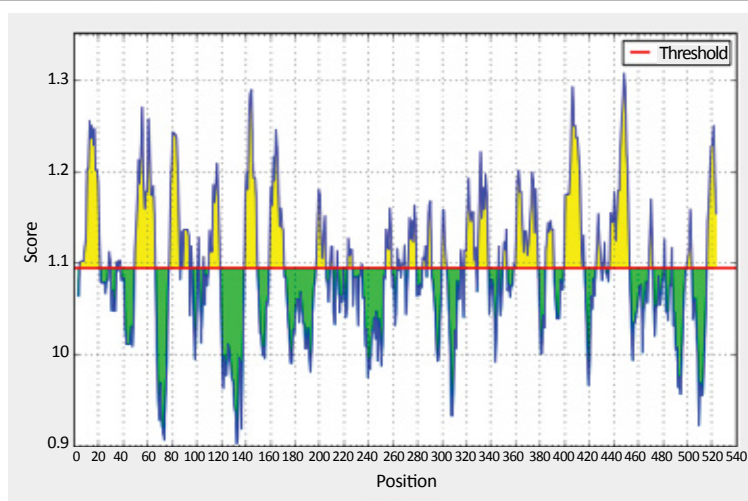


Figure 4: Hydrophobicity plot of HPLC/Parker et al.

297-MGYIIHLCGG-306
 310-SRGYVGVGGLSSVVIYIQTfVSLMCLCGLFFLG-342
 344-VSKEILLEHYFFCNWSLFLVFLFFSILLTYLYCYRLMKG-
 FYYYCSSSLFYSGSLVFSVSLVLVFSIVFLW-417
 421-YNSFVISSLLYSDYYSFLYLFILGLVLCVVFVK-454
 456-GSFDVKYKfyGDLLPKVIIRGNYVVKWSDCMVDYSI-491
 496-DFSFYVSKIFVM-507
 514-INFLFLCLFL-523 and the predicted antigenic fragments can bind to MHC molecule is the first bottlenecks in vaccine design.

Solvent accessible regions

We also predict solvent accessible regions in proteins; different measurement was performed for the prediction of antigenic activity, surface region of peptides. Emini et al. (Figure 7) predicts the

highest probability i.e. found 306-GQQDSR-311 (maximum score 7.194), 307-QQDSRG-312 (maximum score 7.194), 308-QDSRGY-313 (score-6.508), 432-YSDYYS-437 (6.379), 178-KSAQYP-183 (score 6.281), that a given protein region lies on the surface of a protein and are used to identify antigenic determinants on the surface of proteins. Karplus and Schulz (Figure 8) High score is found i.e. found 1.122 maximum in 305-GGQQDSR-311. Predict backbone or chain flexibility on the basis of the known temperature B factors of the α -carbons. The hydrophobicity and hydrophilic characteristics of amino acids is determined by using different scales that are rich in charged and polar residues i.e. Sweet et al. hydrophobicity prediction Result Data found high in position 350-370 (Figure 9), Kyte and Doolittle result high in position 10-20 (Figure 10), Abraham and Leo result high in position 350-370 (Figure 11), Bull and Breese result high in position 300-320 (Figure 12), Miyazawa result high in position 350-370 (Figure 13), Guy result high in position 300-320 (Figure 14), Wolfenden result high in position 440-460 (Figure 15), Roseman result high in position 350-380

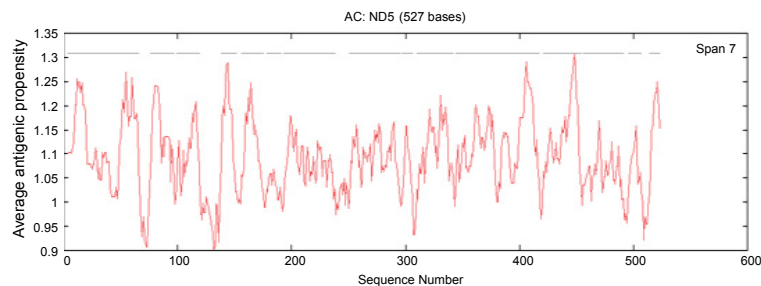


Figure 5: Bepipred linear epitope prediction plot.

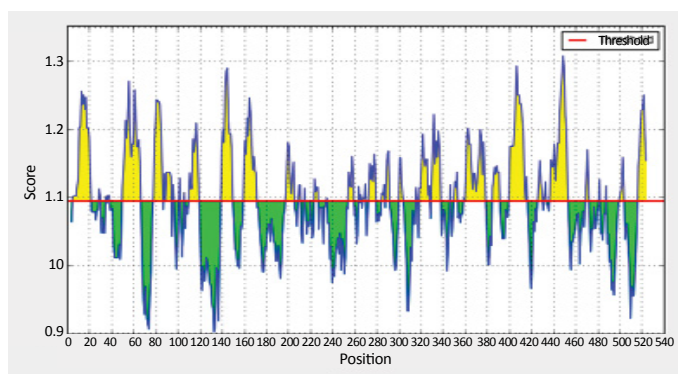


Figure 6: Kolaskar and Tongaonkar antigenicity plot.

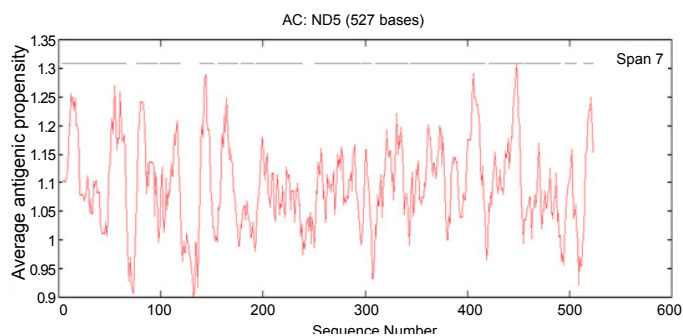


Figure 6a: Kolaskar and Tongaonkar antigenicity plot, the average antigenic propensity for protein is 1.0938.

(Figure 16), Wilson et 350-380 (Figure 17), Cowan 350-380 (Figure 18), Chothia 400-430 (Figure 19).

Prediction of MHC binding peptide

We found the binding of peptides to a number of different alleles using Position Specific Scoring Matrix. NADH dehydrogenase subunit 5 of *Dracunculus medinensis* antigen, with sequence 527 amino acid residues long, is having 519 nonamers. MHC molecules are cell surface proteins, which actively participate in host immune reactions and involvement of MHC-I and MHC-II in response to almost all antigens. We have predicted MHC-I peptide binders of NADH dehydrogenase subunit 5 from *Dracunculus medinensis* was tested with on a set of 4 different alleles i.e. H2-Db (mouse) 8mer, H2-Db (mouse) 9mer, H2-Db (mouse) 10mer and H2-Db (mouse) 11mer (Table 2) and MHC-II peptide binders for I_Ab.p, I_Ad.p, I_Ag7.p alleles highlighted in red represent predicted binders (Table 3). Here RANKPEP report PSSM-specific binding threshold and is obtained by scoring all the antigenic

peptide sequences included in the alignment from which a profile is derived, and is defined as the score value that includes 85% of the peptides within the set. Peptides whose score is above the binding threshold will appear highlighted in red and peptides produced by the cleavage prediction model are highlighted in violet. We also use a cascade SVM based TAPPred method which found 80 High affinity TAP Transporter peptide regions which represents predicted TAP binders residues which occur at N and C termini from *Dracunculus medinensis* (NADHdehydrogenasesubunit5) (Table 3).

Discussion

In this study, we found the antigenic determinants by finding the area of greatest local hydrophilicity. Hopp and Woods hydrophobicity scale is used to identify of potentially antigenic sites in proteins by analyzing amino acid sequences in order to find the point of greatest hydrophilic. Hydrophilicity Prediction result data found high in sequence position at 300-320 in a protein this scale is basically a

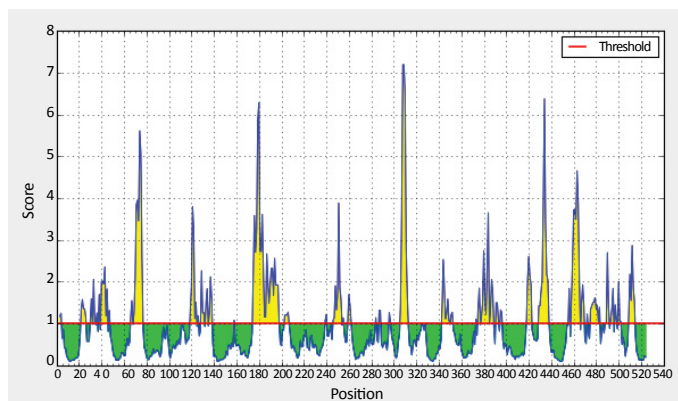


Figure 7: Emini surface accessibility prediction plot.

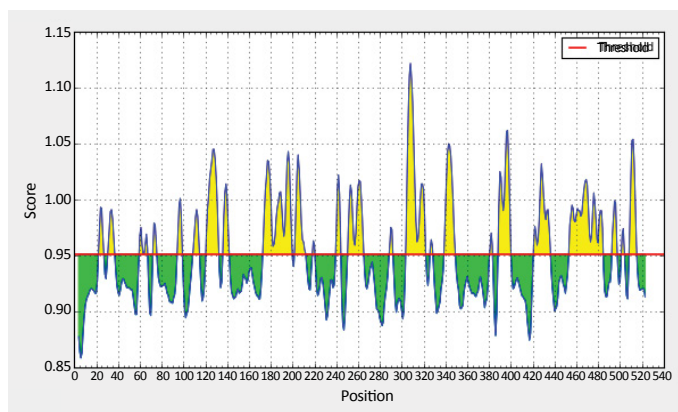


Figure 8: Karplus and Schulz flexibility prediction.

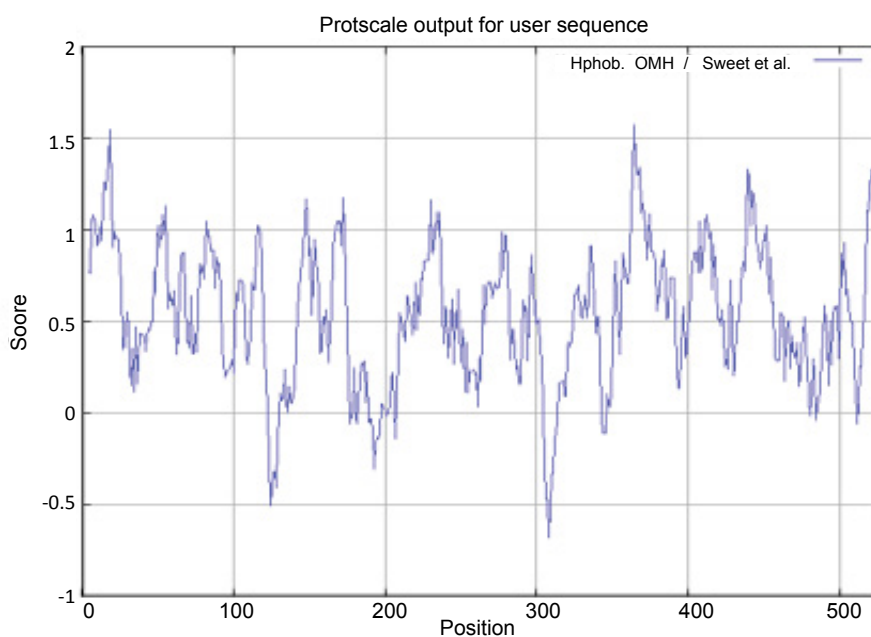


Figure 9: Hydrophobicity plot of Sweet et al.

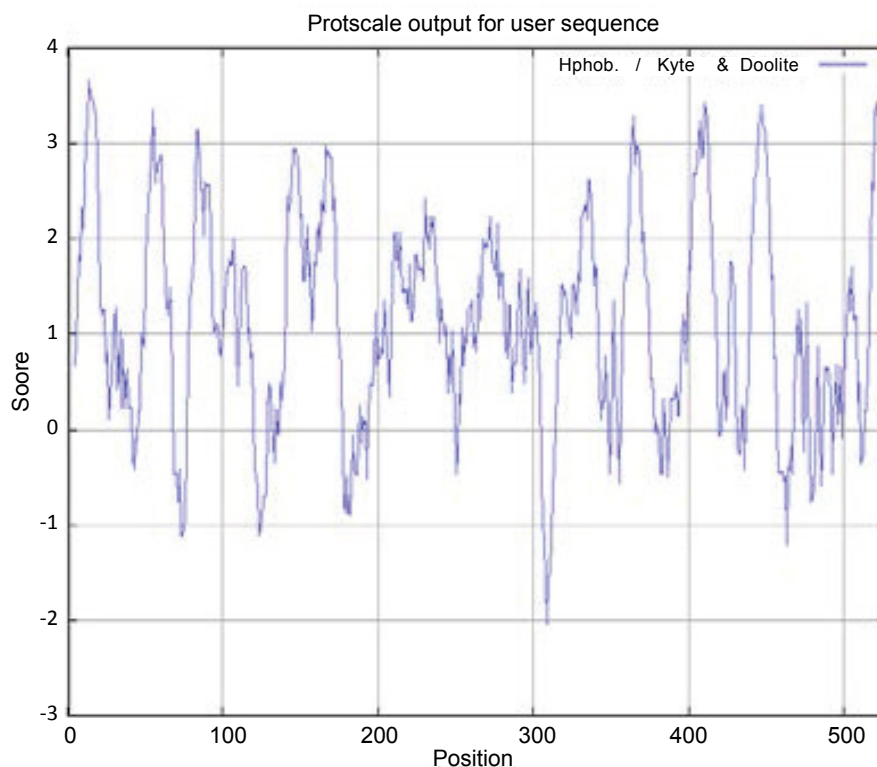


Figure 10: Kyte and Doolittle hydrophobicity plot.

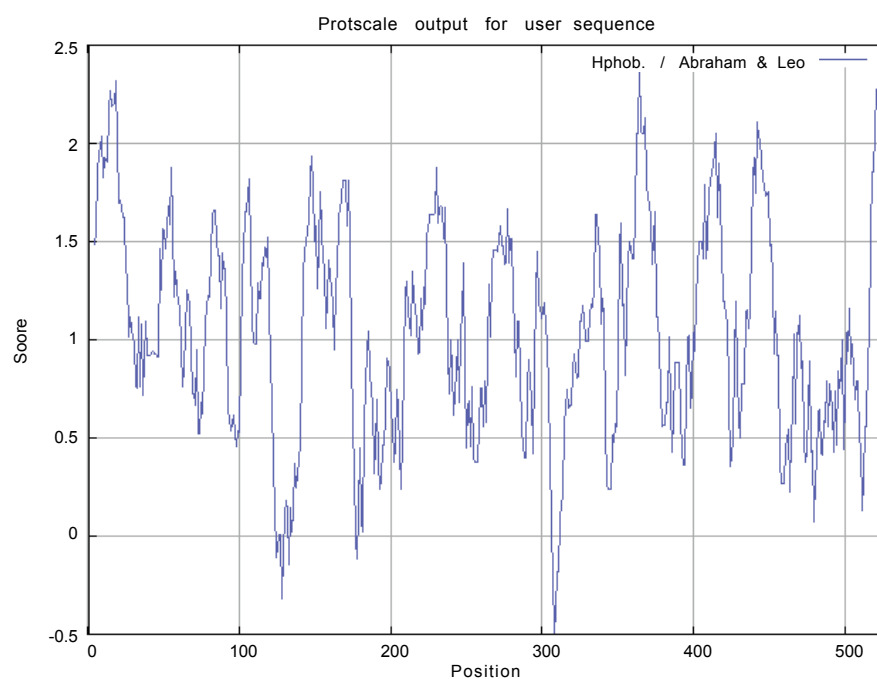


Figure 11: Abraham and Leo hydrophobicity plot.

hydrophilic index where apolar residues have been assigned negative values. The Window size of 5-7 is good for finding hydrophilic regions, greater than 0 values are consider as hydrophilic which is consider as antigenic. Welling used information on the relative occurrence of

amino acids in antigenic regions to make a scale which is useful for prediction of antigenic regions and the predicted result data found high in sequence position 250-270. Welling antigenicity plot gives value as the log of the quotient between percentage in a sample of known

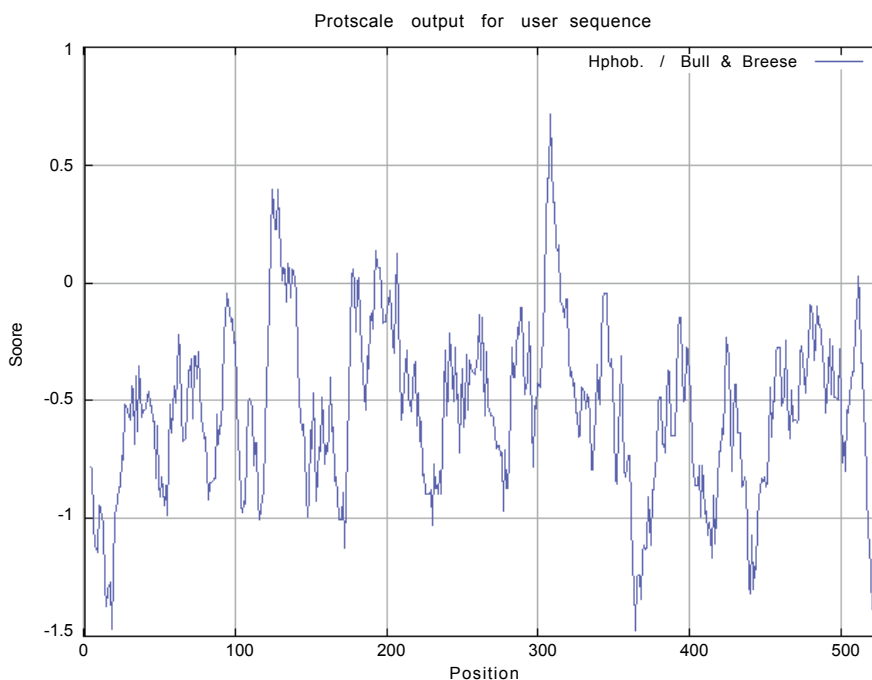


Figure 12: Bull and Breese use surface tension to measure hydrophobicity and also uses negative values to describe the hydrophobicity of antigen NADH dehydrogenase subunit 5.

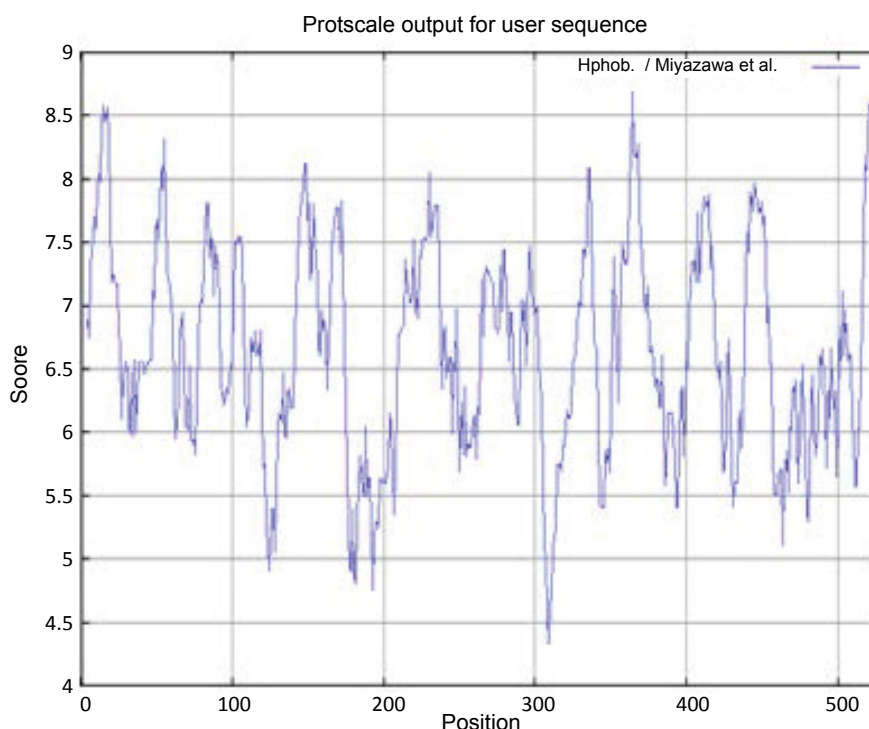


Figure 13: Hydrophobicity plot of Miyazawa et al.

antigenic regions and percentage in average proteins. We also study Hydrophobicity plot of HPLC / Parker Hydrophilicity Prediction Result Data found between 300-320 (Maximum Score-6.3) i.e., the maximum predicted residues at the position 308(Residue is Q) is 305- GGQQDSR-311 and 309(Residue D) is 306- GQQDSRG-312. BepiPred predicts

the location of linear B-cell epitopes Result found that between 300-310 is 301-MQMGYIIHLC-310 and the maximum score (1.369) is found at the position 308,309. There are 15 antigenic determinant sequences is found by Kolaskar and Tongaonkar antigenicity scales (Figures 6a and Table 4) the results show highest pick at position

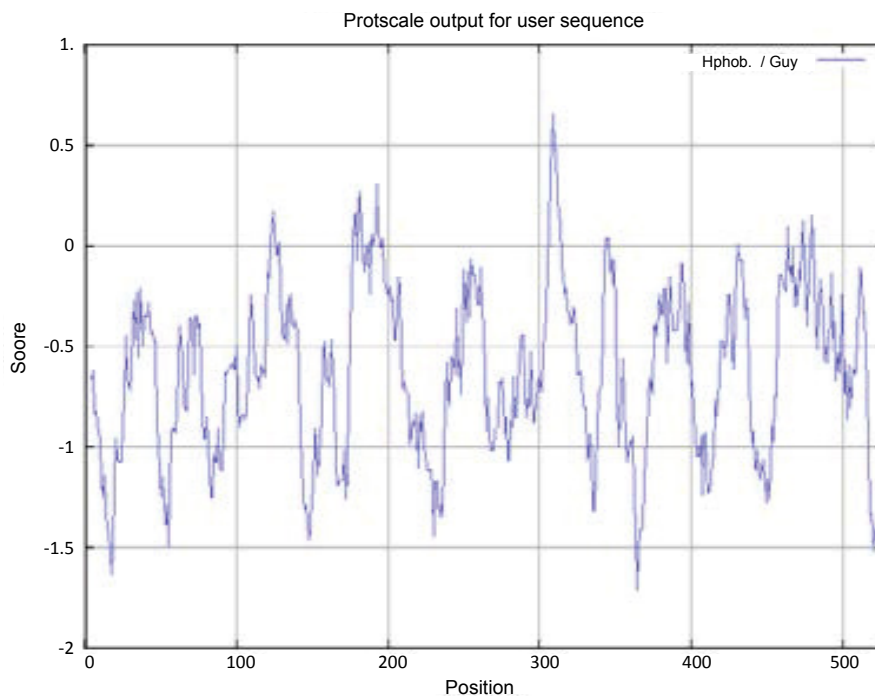


Figure 14: Hydrophobicity plot of Guy.

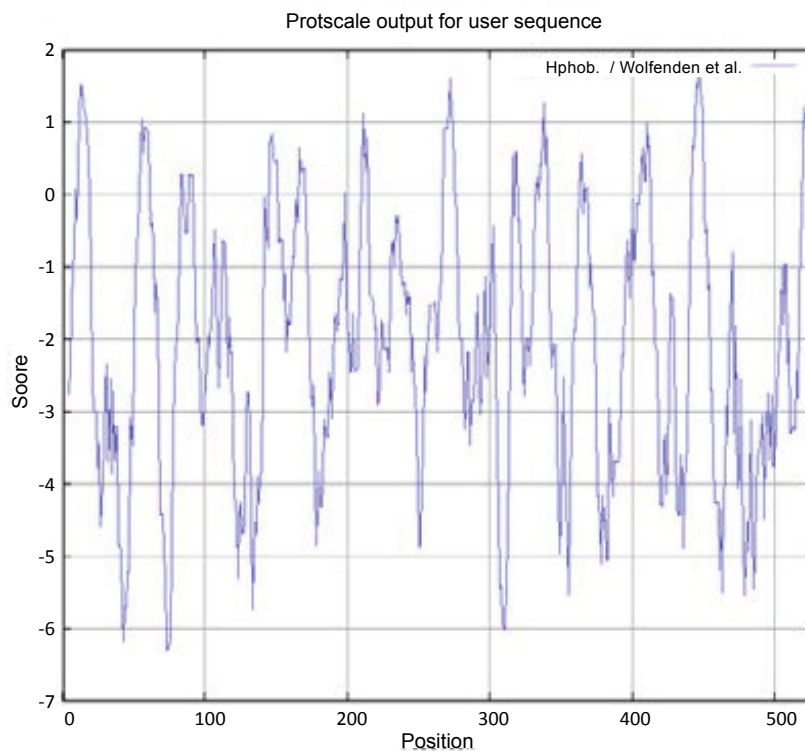


Figure 15. Hydrophobicity plot of Wolfenden et al.

MHC-II Allele	RANK	POS.	N	SEQUENCE	C	MW (Da)	SCORE	% OPT.
MHC-II I_Ab	1	385	KGF	YYYCSSSLF	YSG	1114.26	12.219	34.29%
MHC-II I_Ab	2	484	KWS	DCMVDYSIM	KFG	1058.25	12.095	33.94%
MHC-II I_Ab	3	384	MKG	FYYYCSSSL	FYS	1114.26	11.744	32.96%
MHC-II I_Ab	4	324	SVV	YIQTFSVSLM	CLC	1083.31	11.388	31.96%
MHC-II I_Ab	5	192	PKA	MSAPTPVSA	LVH	841.98	11.24	31.54%
MHC-II I_Ad	1	89	FVF	SMVGVVFSN	NCI	921.07	16.144	30.38%
MHC-II I_Ad	2	282	SFI	HMLSHAVFK	SCL	1051.27	11.598	21.82%
MHC-II I_Ad	3	77	YNF	NYFCVLSI	FVF	1039.26	7.574	14.25%
MHC-II I_Ad	4	387	FYY	YCSSLFYYS	GGG	1038.16	7.329	13.79%
MHC-II I_Ag7	1	332	VSL	MCLCGLFFL	GGG	1028.36	10.512	25.72%
MHC-II I_Ag7	2	33	LDF	SSLELLQFQ	FRL	1046.2	9.939	24.32%
MHC-II I_Ag7	3	373	ILL	TYLYCYRLM	KGF	1207.48	9.4	23.00%
MHC-II I_Ag7	4	457	KFG	SFDVKYKFY	GDL	1178.36	9.342	22.86%
MHC-II I_Ag7	5	71	FYM	MEDYFNIFY	CVV	1224.32	8.271	20.24%

Table 2: Prediction of MHCII ligands all rows highlighted in red represent predicted binders to the MHC-II Allele i.e. MHC-II I_Ab, MHC-II I_Ad, MHC-II I_Ag7. (All rows highlighted in red represent predicted binders).

Peptide Rank	Start Position	Sequence	Score	Predicted Affinity
1	489	YSIMKFGDF	8.651	High
2	483	SDCMVDYSI	8.648	High
3	161	ELVVSVALF	8.643	High
4	96	SNNCISMLI	8.64	High
5	337	LFFLGGSVS	8.639	High
6	305	GGQQDSRGY	8.636	High
7	262	SQIGFCFFG	8.626	High
8	502	SKIFVMGFS	8.62	High
9	417	WWLNYNSFV	8.616	High
10	217	CFSEVMFLD	8.613	High

Table 3: Cascade SVM based high affinity tap binders of *Dracunculus medinensis*.

n	Start Position	Sequence	End Position
1	4	YIYWFGVILLCFLLLFVFFYDDFVFSLDFS SLELLQFQRLDWFSFSFFCLLMVVGS VVVYS	66
2	76	FNYFCVLSIFVFSMVGVVFSN	97
3	99	CISMLIFWDLGVSYYFLVLY	119
4	138	VGDVCFVLFVFCGLFF	152
5	156	DFISLELVVSVALFFFILST	175
6	178	KSAQYPFSSWLPKA	191
7	193	SAPTPVSALVHSSTLVTAGLFLGMCFSE VM FLDFVDFM FVGLFT	238
8	251	DIKKLVALSTLSQIGFCFFGLGLVYFSF IHMLSHAVFKSCLFM	295
9	297	MGYIIHLCGG	306
10	310	SRGYVGVGGLSSVYIQTFSVSLMCLCGL FFLGG	342
11	344	VSKEILLEHYFFCNWSLFLVFLFFSILLT YLYCYRLMKGFYYYCSSSLFYGGSLV FSVSLVLFVFSIVFLW	417
12	421	YNSFVISSLLYSDYSLFYLFILGLVLCV VFFK	454
13	456	GSFDVKYKFYGDLLPKVIIRGNVVKWS DCMVDYSI	491
14	496	DFSFYVSKIFVM	507
15	514	INFLFLCLFL	523

Table 4: The 15 antigenic determinants of NADH dehydrogenase subunit 5.

4-YIYWFGVILLCFLLLFVFFYDDFVFSLDFS SLELLQFQ-
 FRLDWFSFSFFCLLMVVGSVVVYS-66, 76-FNYFCVLSIFV-
 SMVGVVFSN-97, 99-CISMLIFWDLGVSYYFLVLY-119, 138-VG-
 DVCVFLVFCGLFF-152,156-DFISLELVVSVALFFFILST-175,178-
 KSAQYPFSSWLPKA-191,193-SAPTPVSALVHSSTLVTAGLFLGM-
 CFSEVMFLDFVDFMFFVGLFT-238,251-DIKKLVALSTLSQIGF-
 CFFGLGLVYFSFIHMLSHAVFKSCLFM-295,297-MGYIIHLCGG-
 306,310-SRGYVGVGGLSSVYIQTFSVSLMCLCGLFFLGG-342,344-
 SKEILLEHYFFCNWSLFLVFLFFSILLTYLYCYRLMKGFYYYCSSSL-

FYSGGSLVFSVSLVLFVFSIVFLW-417,421-YNSFVISSLLYSDYSSL-
 FYLFILGLVLCVFFK-454,456-GSFDVKYKFYGDLLPKVIIRGNV-
 VKWSDCMVDYSI-491, 496-DFSFYVSKIFVM-507, 514-INFLFLCLFL-
 523.

Result of determined antigenic sites on proteins has revealed that the hydrophobic residues if they occur on the surface of a protein are more likely to be a part of antigenic sites. This method can predict antigenic determinants with about 75% accuracy and also gives the information of surface accessibility and flexibility. Further this region

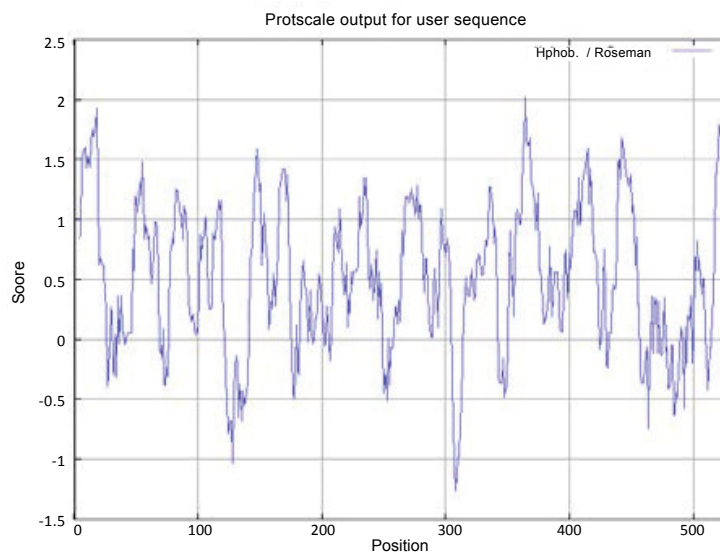


Figure 16: Hydrophobicity plot of Roseman.

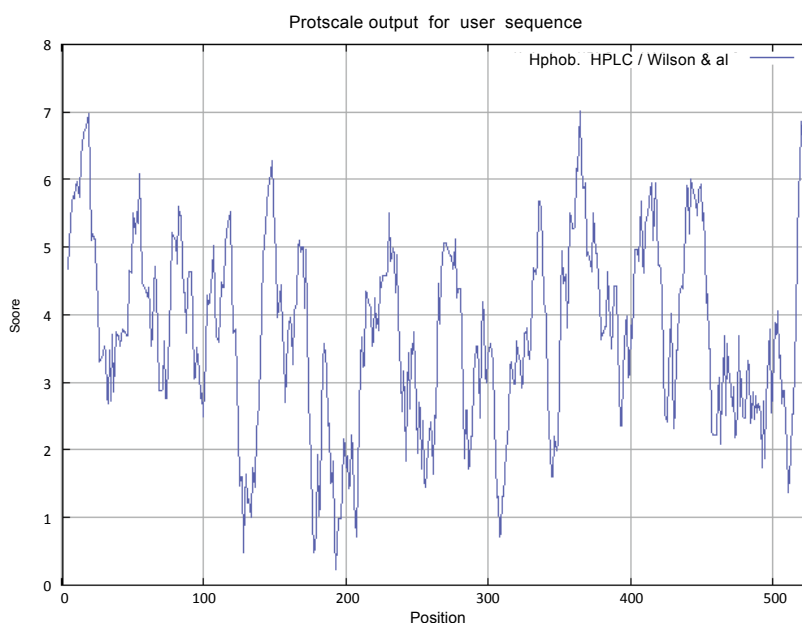


Figure 17: Hydrophobicity/HPLC plot of Wilson.

form beta sheet which show high antigenic response than helical region of this peptide and shows highly antigenicity. The Structure of the *Dracunculus medinensis* antigen-NADHdehydrogenasesubunit5 is predicted by SWISS-MODEL (automated protein structure homology-modelling server) DeepView (Swiss Pdb-Viewer) program (Figure 20). The purpose of this server is to make Protein Modelling accessible to all biochemists and molecular biologists worldwide. The target structure will also serve as a detailed model for determining the structure of peptide within that protein structure. We predict Solvent accessibility by using Emini et al. the result found the highest probability i.e. found 306-GQDSR-311(maximum score 7.194), 307-QQDSRG-312(maximum score 7.194), 308- QDSRGY-313(score-6.508),432-YSDYYS-437(6.379),178- KSAQYP-183(score 6.281), that a given

protein region lies on the surface of a protein and are used to identify antigenic determinants on the surface of proteins. This algorithm also used to identify the antigenic determinants on the surface of proteins and Karplus and Schulz predict backbone or chain flexibility on the basis of the known temperature B factors of the α -carbons here we found the result with High score 1.122 maximum in 305- GGQDSR-311. We predict Solvent accessibility of *Dracunculus medinensis* antigen NADH dehydrogenase subunit 5 for delineating hydrophobic and hydrophilic characteristics of amino acids. Solvent accessibility used to identify active site of functionally important residues in membrane proteins. Solvent-accessible surface areas and backbone angles are continuously varying because proteins can move freely in a three-dimensional space. The mobility of protein segments which are located on the surface of a

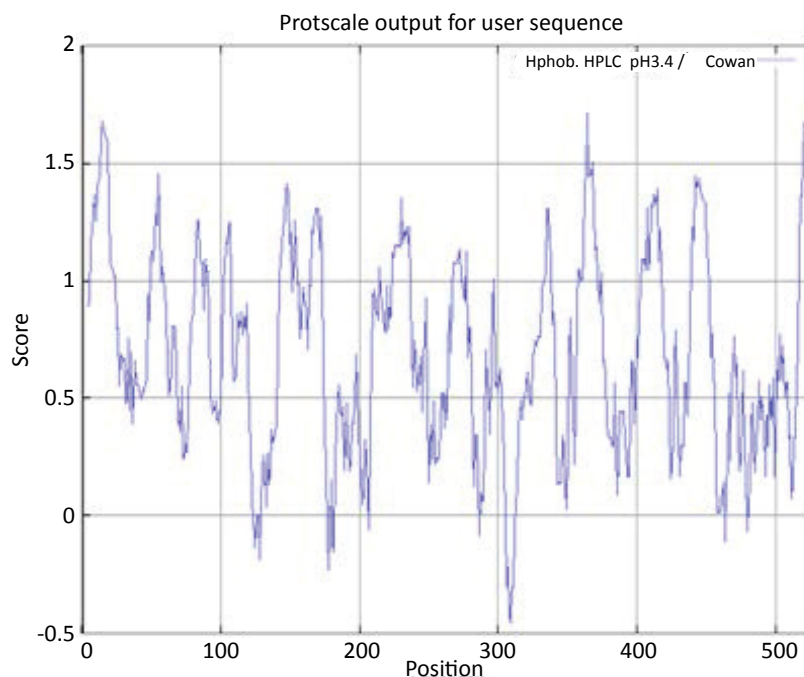


Figure 18: Hydrophobicity/HPLC pH 3.4/ plot of Cowan.

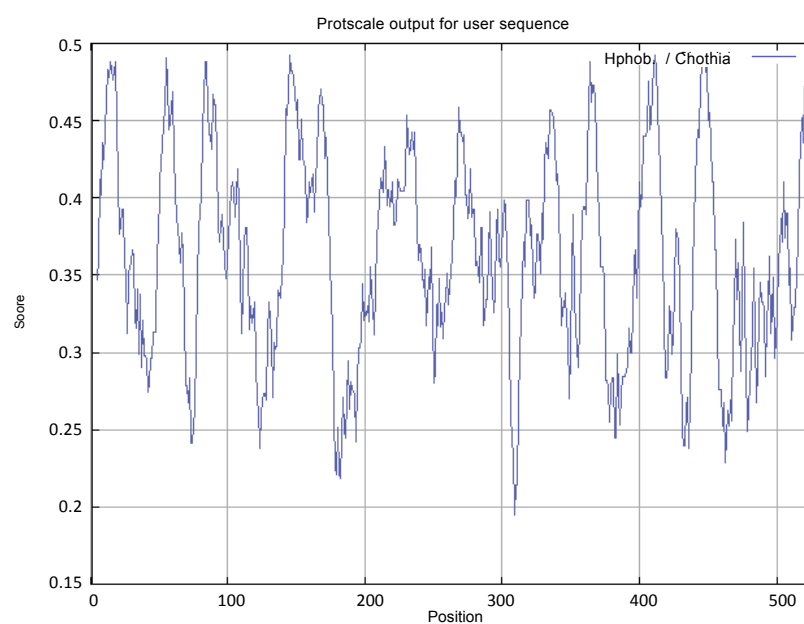
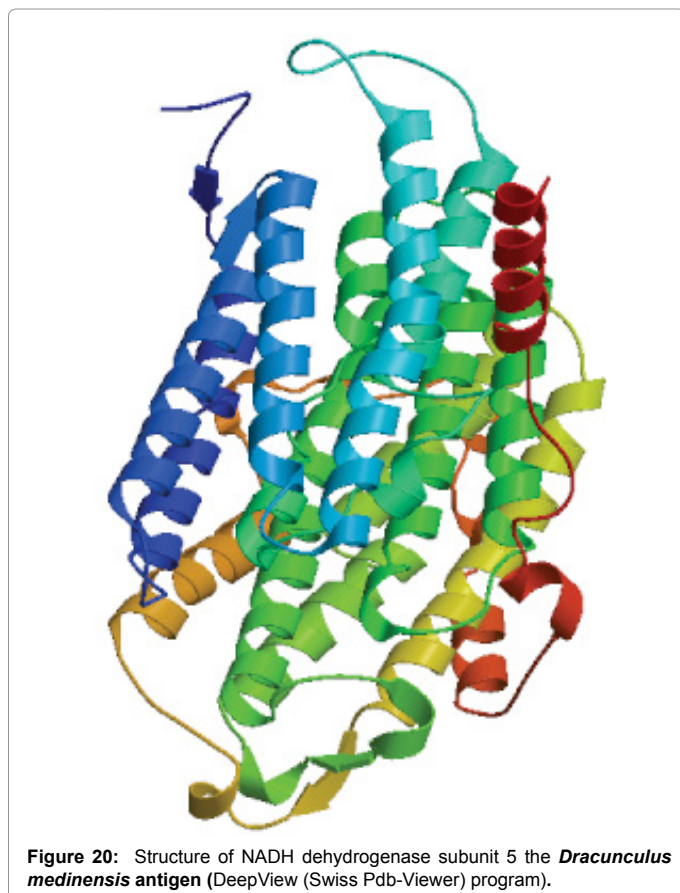


Figure 19: Hydrophobicity plot of Chothia.

protein due to an entropic energy potential and which seem to correlate well with known antigenic determinants. We also found the i.e. Sweet et al. hydrophobicity prediction result data found high in position 350-370, Kyte and Doolittle result high in position 10-20, Abraham and Leo result high in position 350-370, Bull and Breese result high in position 300-320, Guy result high in position 300-320, Miyazawa result high in position 350-370, Roseman result high in position 350-380, Wolfenden result high in position 440-460, Wilson et 350-380,

Cowan 350-380, Chothia 400-430. These scales are a hydrophilic with a polar residues assigned negative value. Because the N- and C- terminal regions of proteins are usually solvent accessible and unstructured, antibodies against those regions recognize the antigenic protein. In this study, we found predicted MHC-I peptide binders of toxin protein for 8mer_H2_Db alleles with the consensus sequence QNWNCCIT that yields the maximum score i.e. 52.494, 9mer_H2_Db with, the consensus sequence FCIHNCXYM that yields the maximum score i.e.



50.365, 10mer_H2_Db with, the consensus sequence SGYYNFFWCL that yields the maximum score i.e. 58.858, 11mer_H2_Db with, the consensus sequence CGVYNFYCCY that yields the maximum score i.e. 79.495 and I_Ab with the consensus sequence YYAPWCNNA that yields the maximum score i.e. 35.632, I_Ad with the consensus sequence QMVHAAHAE that yields the maximum score i.e. 53.145, MHC-II I_Ag7 with the consensus sequence WYAHAFKYV that yields the maximum score i.e. 40.873 for MHC II allele was tasted. We also use a cascade SVM based TAPPred method which found 160 High affinity TAP Transporter peptide regions which represents predicted TAP binders residues which occur at N and C termini from *Dracunculus medinensis* antigen NADHdehydrogenasesubunit5. TAP is an important transporter that transports antigenic peptides from cytosol to ER. TAP binds and translocate selective antigenic peptides for binding to specific MHC molecules. The efficiency of TAP-mediated translocation of antigenic peptides is directly proportional to its TAP binding affinity. Thus, by understanding the nature of peptides, that bind to TAP with high affinity, is important steps in endogenous antigen processing. The correlation coefficient of 0.88 was obtained by using jackknife validation test. In this test, we found the MHCI and MHCII binding regions. T cell immune responses are derived by antigenic epitopes hence their identification is important for design synthetic peptide vaccine. T cell epitopes are recognized by MHCI molecules producing a strong defensive immune response against antigen NADHdehydrogenasesubunit5. Therefore, the prediction of peptide binding to MHCI molecules by appropriate processing of antigen peptides occurs by their binding to the relevant MHC molecules. Because, the C-terminus of MHCI-restricted epitopes results from cleavage by the proteasome and thus, proteasome specificity

is important for determining T-cell epitopes. Consequently, RANKPEP also focus on the prediction of conserved epitopes. C-terminus of MHCI-restricted peptides is generated by the proteasome, and thus RANKPEP also determines whether the C-terminus of the predicted MHCI-peptide binders is the result of proteasomal cleavage. Moreover, these sequences are highlighted in purple in the output results. Proteasomal cleavage predictions are carried out using three optional models obtained applying statistical language models to a set of known epitopes restricted by human MHCI molecules as indicated here.

Conclusion

From the above result and discussion it is concluded that the ability of RANKPEP to predict MHC binding peptides, and thereby potential T-cell epitopes, antigenic peptide that binds to MHC molecule are antigenic that means hydrophilic in nature. This means the increase in affinity of MHC binding peptides may result in enhancement of immunogenicity of *Dracunculus medinensis* antigen NADH dehydrogenase subunit5 and are helpful in the designing of synthetic peptide vaccine. This approach can help reduce the time and cost of experimentation for determining functional properties of *Dracunculus medinensis* antigen NADHdehydrogenasesubunit5. Overall, the results are encouraging, both the 'sites of action' and 'physiological functions' can be predicted with very high accuracies helping minimize the number of validation experiments.

References

- Wallace DC (1992) Diseases of the mitochondrial DNA. Annu Rev Biochem 61: 1175-1212.
- Liu Z, Song Y, Gu S, He X, Zhu X, et al. (2012). Mitochondrial ND5 T12338T>C variant is associated with maternally inherited hypertrophic cardiomyopathy in a Chinese pedigree. Gene 506: 339-343.
- Ji YC, Liu XL, Zhao FX, Zhang JJ, Zhang Y, et al. (2011). The mitochondrial ND5 T12338C mutation may be associated with Leber's hereditary optic neuropathy in two Chinese families. 33: 322-328.
- Wang K, Yan CZ, Wang GX, Jiao JS, Jin M (2010) Mitochondrial ND5 as the causative gene of Leight syndrome. 27: 616-619.
- Liu XL, Zhou X, Zhou J, Zhao F, Zhang J, et al. (2011) Leber's hereditary optic neuropathy is associated with the T12338C mutation in mitochondrial ND5 gene in six Han Chinese families. Ophthalmology 118: 978-985.
- Downham E, Winterthun S, Nakkestad HL, Hirth A, Halvorsen T, et al. (2008) A novel mitochondrial ND5 (MTND5) gene mutation giving isolated exercise intolerance. Neuromuscul Disord 18: 310-314.
- Muller R (1971) *Dracunculus* and dracunculiasis. Adv Parasitol 9: 73-151.
- Muller R (1979) Guinea worm disease: epidemiology, control, and treatment. Bull World Health Organ 57: 683-689.
- Jones HI, Mulder E (2007) *Dracunculus mulbus* n. sp. (Nematoda: Spirurida) from the water python *Liasis fuscus* (Serpentes: Boidae) in northern Australia. Syst Parasitol 66: 195-205.
- Moravec F, Santos CP (2009) *Dracunculus brasiliensis* sp. n. (Nematoda: Dracunculidae) from the anaconda, *Eunectes murinus* (Ophidia: Boidae). Parasitol Res 104: 589-592.
- Birare SD, Kamble MH, Lanjewar DN, Parija SC, Girji DD, et al. (2005) Guinea worm infection of urinary bladder manifesting as obstructive uropathy in rural Maharashtra. Trop Doct 35: 242.
- <http://www.cdc.gov/parasites/guineaworm/biology.html>.
- Greenaway C (2004). Dracunculiasis (Guinea worm disease). CMAJ 170: 495-500.
- Müllner A, Helfer A, Kotlyar D, Oswald J, Efferth T (2011) Chemistry and pharmacology of neglected helminthic diseases. Curr Med Chem 18: 767-789.
- Ruiz-Tiben E, Hopkins DR (2006) Dracunculiasis (Guinea worm disease) eradication. Adv Parasitol 61: 275-309.

16. Iriemenam NC, Oyibo WA, Fagbenro-Beyioku AF (2008) Dracunculiasis--the saddle is virtually ended. *Parasitol Res* 102: 343-347.
17. Hopkins DR, Ruiz-Tiben E, Eberhard ML, Roy SL (2014) Centers for Disease Control and Prevention (CDC) Progress toward global eradication of dracunculiasis--January 2013-June 2014. *MMWR Morb Mortal Wkly Rep* 63: 1050-1054.
18. Eta N, Gerald ES, Flaubert D, Walter KK, Valeri O, et al. (2015) Not every worm wrapped around a stick is a guinea worm: a case of *Onchocerca volvulus* mimicking *Dracunculus medinensis*. *Parasit Vectors* 8: 374.
19. Richards FO, Ruiz-Tiben E, Hopkins DR (2011) Dracunculiasis eradication and the legacy of the smallpox campaign: what's new and innovative? What's old and principled? *Vaccine* 29 Suppl 4: D86-90.
20. Hopkins DR (2013) Disease eradication. *N Engl J Med* 368: 54-63.
21. Muller R (1985) Life cycle of *Dracunculus medinensis*. In workshop on opportunities for control of dracunculiasis : contaminated papers, washington, DC: National Academy Press.
22. Cairncross S, Muller R, Zagaria N (2002) Dracunculiasis (Guinea worm disease) and the eradication initiative. *Clin Microbiol Rev* 15: 223-246.
23. Bloch P, Simonsen PE, Vennervald BJ (1993) The antibody response to *Dracunculus medinensis* in an endemic human population of northern Ghana. *J Helminthol* 67: 37-48.
24. Bloch P, Simonsen PE (1998) Immunoepidemiology of *Dracunculus medinensis* infections II. Variation in antibody responses in relation to transmission season and patency. *Am J Trop Med Hyg* 59: 985-990.
25. Flower DR (2008) Vaccines: how they work in Bioinformatics for Vaccinology. Wiley-Blackwell, Oxford, UK. 73-112.
26. Batalia MA, Collins EJ (1997) Peptide binding by class I and class II MHC molecules. *Biopolymers* 43: 281-302.
27. Marrack P, Scott-Browne JP, Dai S, Gapin L, Kappler JW (2008) Evolutionarily conserved amino acids that control TCR-MHC interaction. *Annu Rev Immunol* 26: 171-203.
28. Chapman HA (1998) Endosomal proteolysis and MHC class II function. *Curr Opin Immunol* 10: 93-102.
29. Gomase VS, Chittlange NR (2012) Prediction of MHC Class Antigen Peptides from *Echinococcus Multilocularis*: Application of Computer Intelligence 1: 191.
30. <http://www.ncbi.nlm.nih.gov>.
31. Sayers EW, Acland A, Agarwala R, Barrett T, Beck J, et al. (2012). Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 40: D13-25.
32. Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, et al. (2005) The Universal Protein Resource (UniProt). *Nucleic Acids Res* 33: D154-159.
33. Hopp TP, Woods KR (1981) Prediction of protein antigenic determinants from amino acid sequences. *Proc Natl Acad Sci U S A* 78: 3824-3828.
34. Welling GW, Weijer WJ, van der Zee R, Welling-Wester S (1985) Prediction of sequential antigenic regions in proteins. *FEBS Lett* 188: 215-218.
35. Parker KC, Bednarek MA, Coligan JE (1994) Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *J Immunol* 152: 163-175.
36. Larsen JE, Lund O, Nielsen M (2006) Improved method for predicting linear B-cell epitopes. *Immunome Res* 2: 2.
37. Kolaskar AS, Tongaonkar PC (1990) A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS Lett* 276: 172-174.
38. Sonu M, Gomase VS (2015) Prediction of antigenic epitope from *D. medinensis*: new paradigm of synthetic vaccine development. International Conference on "Recent Research Development in Environment, Social Sciences and Humanities. ICRRDESH, India.
39. Ruppert J, Sidney J, Celis E, Kubo RT, Grey HM, et al. (1993) Prominent role of secondary anchor residues in peptide binding to HLA-A2.1 molecules. *Cell* 74: 929-937.
40. Reche PA, Glutting JP, Reinherz EL (2002) Prediction of MHC class I binding peptides using profile motifs. *Hum Immunol* 63: 701-709.
41. Reche PA, Reinherz EL (2003) Sequence variability analysis of human class I and class II MHC molecules: functional and structural correlates of amino acid polymorphisms. *J Mol Biol* 331: 623-641.
42. Craiu A, Akopian T (1997) Two distinct proteolytic processes in the generation of a major histocompatibility complex class I-presented peptide. *Proc Natl Acad Sci U S A* 94: 10850-10855.
43. Pieters J (2000) MHC class II-restricted antigen processing and presentation. *Adv Immunol* 75: 159-208.
44. Bhasin M, Raghava GP (2004) Analysis and prediction of affinity of TAP binding peptides using cascade SVM. *Protein Sci* 13: 596-607.
45. Emini, EA, Hughes JV, Perlow DS, Boger J (1985) Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *J Virol* 55: 836-839.
46. Karplus PA, Schulz GE (1985) Prediction of chain flexibility in proteins: a tool for the selection of peptide antigen. *Naturwissenschaften* 72: 212-213.
47. Sweet RM, Eisenberg D (1983) Correlation of sequence hydrophobicities measures similarity in three-dimensional protein structure. *J. Mol. Biol* 171: 479-488.
48. Kyte J, Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* 157: 105-132.
49. Abraham DJ, Leo AJ (1987) Extension of the fragment method to calculate amino acid zwitterion and side chain partition coefficients. *Proteins* 2: 130-152.
50. Bull HB, Breese K (1974) Surface tension of amino acid solutions: a hydrophobicity scale of the amino acid residues. *Arch Biochem Biophys* 161: 665-670.
51. Miyazawa S, Jernigen RL (1985) Estimation of Effective Interresidue Contact Energies from Protein Crystal Structures: Quasi-Chemical Approximation. *Macromolecules* 18: 534-552.
52. Roseman MA (1988) Hydrophilicity of polar amino acid side-chains is markedly reduced by flanking peptide bonds. *J Mol Biol* 200: 513-522.
53. Wolfenden R, Andersson L, Cullis PM, Southgate CC (1981) Affinities of amino acid side chains for solvent water. *Biochemistry* 20: 849-855.
54. Wilson KJ, Honegger A, Stötzel RP, Hughes GJ (1981) The behaviour of peptides on reverse-phase supports during high-pressure liquid chromatography. *Biochem J* 199: 31-41.
55. Cowan R, Whittaker RG (1990) Hydrophobicity indices for amino acid residues as determined by high-performance liquid chromatography. *Pept Res* 3: 75-80.
56. Chothia C (1976) The nature of the accessible and buried surfaces in proteins. *J Mol Biol* 105: 1-12.