International Journal of Systems Biology
ISSN: 0975-2900 & E-ISSN: 0975-9204, Volume 3, Issue 1, 2012, pp.-34-44.
Available online at http://www.bioinfo.in/contents.php?id=29

MICROBIAL PROTEOMICS BASED IDENTIFICATION OF QUANTITATIVE IMMUNOGENIC SITES FROM *Toxoplasma gondii* MAJOR SURFACE ANTIGEN P30

GOMASE V.S.[1]*, CHITLANGE N.R.[1], SHERKHANE A.S.[1], CHANGBHALE S.S.[1] AND KALE K.V.[2]

[1]Jagdishprasad Jhabarmal Tibrewala University, Jhunjhunu-333001, Rajasthan, India.
[2]Department of Computer Science and IT, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad- 431001, MS, India.
*Corresponding Author: Email- gomase.viren@gmail.com

Received: July 02, 2012; Accepted: July 30, 2012

Abstract- One-third to half of the world's human population has been estimated to be infected with toxoplasma infection, Toxoplasmosis, caused by *Toxoplasma gondii*. It is found contributory factor in various psychiatric disorders including depression, schizophrenia. In this analysis we have predicted antigenic peptides from *Toxoplasma gondii* major surface antigen P30 protein for synthetic peptide vaccine design against toxoplasmosis because with a single protein subunit, immune response can be generated in large population. Analysis shows predicted epitopes of *Toxoplasma gondii* major surface antigen P30 protein are important determinant for protection from toxoplasmosis. In this assay we have analysed the binding affinity of *Toxoplasma gondii* major surface antigen P30 protein having 336 amino acids, which shows 328 nonamers. In this research work, we predicted CTL-epitopes by two different methods namely SVM (Support Vector Machine) and ANN (Artificial Neural Network), SVM based prediction shown forty valid epitopes having optimal score of 1.378 at cut off 0.36 whereas ANN based prediction shown thirty-one valid epitopes having optimal score of 1.000 at cut off 0.51. We also predicted cascade SVM based TAP binders in addition to fourteen potential antigenic epitopes as, 4-SLHHFIISSGF-14, 20-PKAVRRAVTAGVFAAPTLMSFLRCGVMASDPPLVANQVVTCPH KKSTAAVILT-72, 79-TLKCPKT-85, 88-TEPPTLAYS-96, 98-NRQICPAG-105, 107-TSSCTSKAVTLSSLIPE-123, 136-LDTAGIKLTVPIEKFPVT TQTFVVGCI-162, 167-AQSCMVTETVQARASSVVNNVARCSYG-193, 197-TLGPVKVSAE-206, 211-MTLVCGKDGVKVP-223, 227-NQYCSG T-233, 244-KDILPKL-250, 276-AESKSVIIGCTG-287, 292-KHHCTVKLEFA-302 (average propensity 1.044) predicted on the basis of highest local hydrophilicity; This approach can be implemented in designing subunit and synthetic peptide vaccine against Toxoplasmosis.
Keywords- Toxoplasmosis, parasitic disease, Antigenic peptides, MHC, SVM, ANN, CTL, Nonamers, synthetic peptide vaccines
Abbrevations- MHC: Major Histocompatibility Complex, CTL: Cytotoxic T lymphocytes, TAP: Transporter associated with Antigen Processing, SVM: Support Vector Machine, ANN: Artificial Neural Network

Citation: Gomase V.S., et al. (2012) Microbial Proteomics based Identification of Quantitative Immunogenic Sites from *Toxoplasma gondii* major surface antigen P30. International Journal of Systems Biology, ISSN: 0975-2900 & E-ISSN: 0975-9204, Volume 3, Issue 1, pp.-34-44.

Introduction

Toxoplasmosis is a parasitic disease caused by *Toxoplasma gondii*, the parasite infects most of warm blooded animals, including humans, but the primary host are of felid family. The parasite spreads by the ingestion of infected meat or the excrement of an infected cat or by vertical transmission from mother to foetus. One-third to half of the world's human population is estimated to carry a Toxoplasma infection; one may don't even know it because infection may be with no symptoms or are similar with other illnesses. Toxoplasmosis disease is most serious when one's immune system is weak. Studies show the toxoplasmosis parasite may affect behaviour and may present as or be a causative or contributory factor in various psychiatric disorders such as anxiety, depression and schizophrenia [1,2].

Pathogen Transmission

Toxoplasmosis is transmitted from animal to humans; infected cats excrete the pathogen in their faeces for a number of weeks after undertaking the disease, generally after eating an infected rodent. Even then, cat faeces are not generally infectious for the first one or two days after excretion, after which the cyst becomes potentially pathogenic. It is sometimes also found to be transmitted through ingestion of raw or partly cooked meat.

Strategy

This approach is based on the phenomenon of cross-protection [3] hereby an individual infected with a mild strain of pathogen possess immunity against more severe strain of the same pathogen. Body proteins are necessary for production of immunity in or on all food commodities. Relief from the requirement of a tolerance is established for residues of the drugs or chemicals.

MHC Class Binding Peptides

The new paradigm in vaccine design is emerging; following essential discoveries in immunology and development of new CTL binding peptides prediction tools [4]. MHC molecules are cell surface

glycoproteins, which play important role in host immune reactions. The involvement of MHC class-I in response to almost all antigens and considering its binding affinity with peptides of 9 amino acids, we have predicted the 9 amino acid long epitopes from extracted data MHCBN Comprehensive data base of MHC-I and MHC-II binding as well as non-binding epitopes. MHC molecules have been well characterized in terms of their role in immune reactions. They bind to several peptide fragments generated after proteolytic cleavage of antigen [5]. This binding act like red flags for specific antigen and generates immune response against the parent antigen. So an antigen from *Toxoplasma gondii* major surface antigen P30 subunit can induce immune response against Toxoplasmosis. CTL epitopes are most suitable for subunit vaccine development because a single epitope can generate adequate immune response. MHC peptide complexes will be translocated on the surface of antigen presenting cells (APCs). This theme is implemented in developing subunit and synthetic peptide vaccines [6-9]. One of the important problems in subunit vaccine design is to search for antigenic regions in an antigen protein [10] can stimulate T-cells called T-cell epitopes. Fortunately, a large amount of data about such peptides is available in literature. Pastly and presently, a number of databases have been designed and maintained to provide comprehensive information related to T-cell epitopes [11-16].

Materials and Methods

Protein Sequence Analysis

Toxoplasma gondii major surface antigen P30 protein is a diagnostic antigen against toxoplasmosis. The antigenic protein sequence of *Toxoplasma gondii* major surface antigen P30 protein was analyzed to study the antigenicity [17]; MHC class binding peptide and solvent accessible regions, which allows potential drug targets to predict active sites against toxoplasmosis.

A *Toxoplasma gondii* major surface antigen P30 (gi- 7387469) protein sequence is 336 amino acids long as-

MSVSLHHFIISSGFLTSMFPKAVRRAVTAGVFAAPTLMSFLRCGV
MASDPPLVANQVVTCPHKKSTAAVILTPTENHFTLKCPKTALTEP
PTLAYSPNRQICPAGTTSSCTSKAVTLSSLIPEAEDSWWTGD-
SASLDTAGIKLTVPIEKFPVTTQTFVVGCIKGDDAQSCMVTETVQ
ARASSVVNNVARCSYGADSTLGPVKVSAEEPTTMTLVCGKDGV
KVPQDNNQYCSGTTLTGCNEKSFKDILPKLTENPWQGNASSDK
GATLTIKKEAFPAESKSVIIGCTGGSPEKHHCTVKLEFAGAAGSAK
SAAGTASHVSIFAMVIGLIGSIAACVA

Antigenicity Prediction

Antigenicity Prediction program results those segments from *Toxoplasma gondii* major surface antigen P30 protein that are likely to be antigenic by eliciting an antibody response. Antigenic epitopes are determined using the Gomase method (2007), Hopp and Woods, Welling, Parker, B-EpiPred Server and Kolaskar and Tongaonkar antigenicity methods [18-23].

Protein Secondary Structure Prediction

The important concepts in secondary structure prediction are identified as: residue conformational propensities, sequence edge effects, hydrophobicity moments, insertions and Deletions positions in aligned homologous sequence, moments of conservation, autocorrelation, residue ratios, secondary structure feedback effects and filtering [24,25].

Finding the Location in Solvent Accessible Regions

Location finding in solvent accessible regions in protein, plot type determines the hydrophobic and hydrophilic scales and it is utilized for prediction. This may be applicable in predicting membrane-spanning domains, potential antigenic sites and regions that are likely exposed on the protein surface [26-47].

CTL Epitope Prediction

The CTL Epitopes of *Toxoplasma gondii* major surface antigen P30 are obtained from MHCBN comprehensive database of MHC binding and non-binding peptides using two different methods firstly with Support Vector Machine based method (Cut off is 0.36) and then by Artificial Neural Network based method (Cut off is 0.51). In this work predicted MHC-Peptide binding is a log-transformed value related to the IC50 values in nM units.

The average accuracy of Support Vector Machine (SVM) based epitope prediction method is ~76% at cut off 0.36. SVM has been trained on the binary input of single amino acid sequence. In Case of Artificial Neural Network ANN based epitope prediction method the average accuracy is ~74% at the cut off score 0.51 [4].

TAP Binding Prediction

TAP (Transporter associated with Antigen Peptide) play an important role in transportation of MHC-Peptide complexes, which elicits the immune response for clearing various intracellular infections. The prediction of TAP binding peptides is crucial in identifying the MHC class-I restricted T cell epitopes. The cascade SVM based prediction, using properties of amino acid sequence at correlation coefficient of 0.88 as per Jack-Knife validation test [48].

Result and Interpretation

Antigenic Peptides Prediction

In this assay we predicted the antigenic determinants by finding the area of highest local hydrophilicity. The Hopp & Woods scale predicts the locations of antigenic determinants in linear antigen protein sequence, assuming that the antigenic determinants would be exposed on the protein surface and thus would be located in hydrophilic regions [Fig-1]. Its values are derived from the transfer-free energies for amino acid side chains between ethanol and water. 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 [Fig-2].

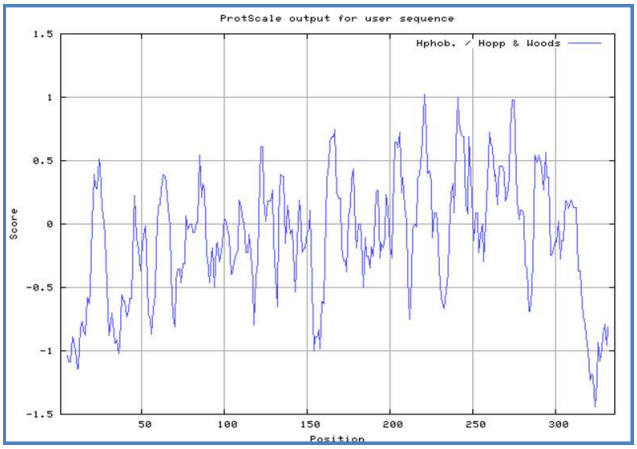

Fig. 1- Hopp and Woods (1981) Hydrophobicity plot of *Toxoplasma gondii* major surface antigen P30

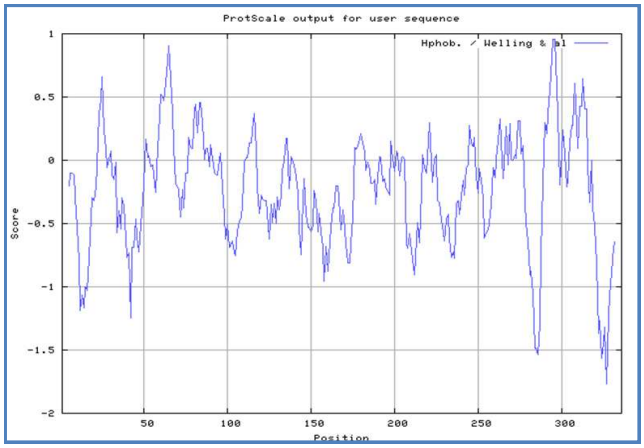

Fig. 2- Welling et al. (1985) Hydrophobicity plot of *Toxoplasma gondii* major surface antigen P30

We also studied B-EpiPred Server, Parker, Kolaskar and Tongaonkar antigenicity methods and the predicted antigenic fragments can bind to MHC molecule is the first bottleneck in vaccine design [Fig-3], [Fig-4], [Fig-5], [Fig-6].

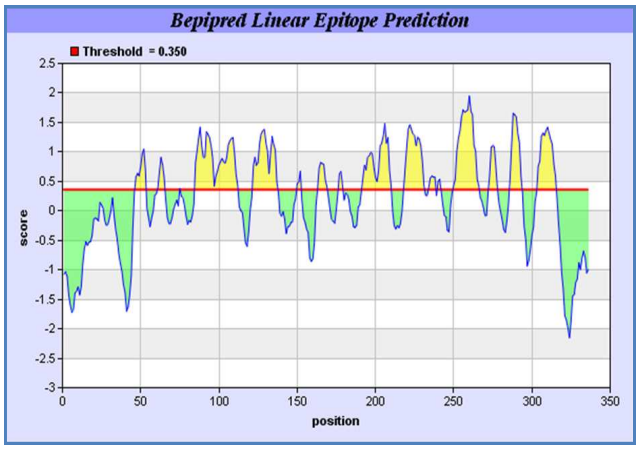

Fig. 3- B-cell epitopes are the sites of molecules that are recognized by antibodies of the immune system of the *Toxoplasma gondii* major surface antigen P30

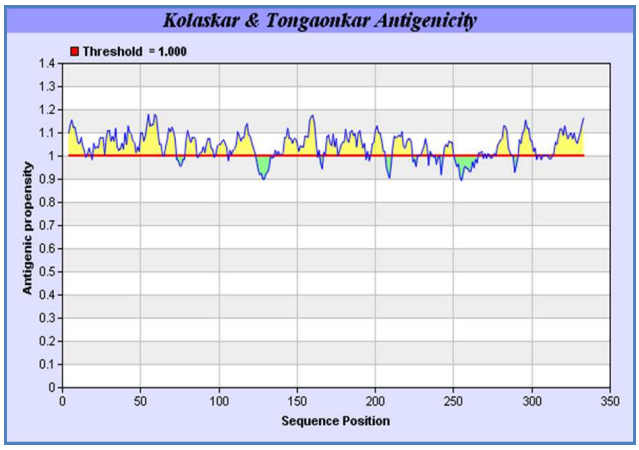

Fig. 4- Kolaskar and Tongaonkar antigenicity are the sites of molecules that are recognized by antibodies of the immune system for the *Toxoplasma gondii* major surface antigen P30

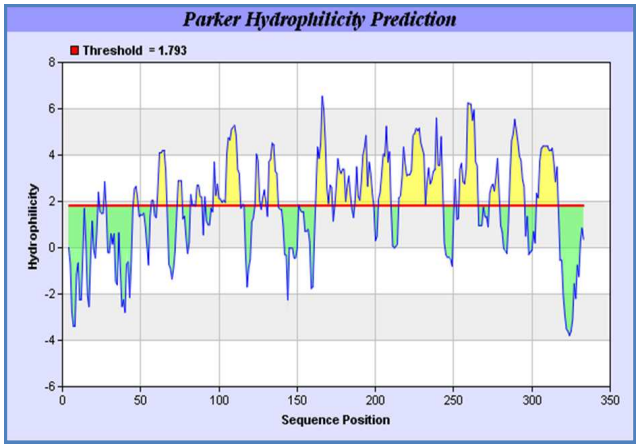

Fig. 5- Hydrophobicity plot of HPLC/Parker et al. (1986) of *Toxoplasma gondii* major surface antigen P30

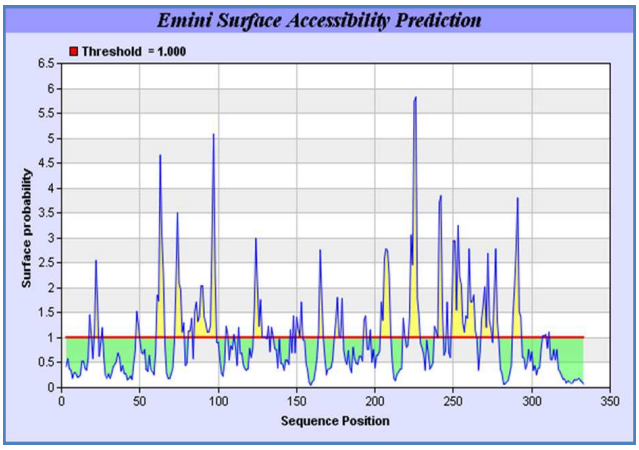

Fig. 6- Hydrophobicity plot of Emini Surface Accessibility Prediction (1990) of *Toxoplasma gondii* major surface antigen P30

Secondary Alignment

The Robson and Garnier method applied for secondary structure prediction of the *Toxoplasma gondii major surface antigen P30 protein*. Each residue have specific assigned value for alpha helix (*Shown in Red*), beta sheet (*Shown in Blue*) and coil (*Shown in Pink*) using a window of 7 residues [Fig-7]. Using these information parameters, the similar to a given residue assuming each of the four possible conformations alpha, beta, reverse turn or coils calculated and the conformation with the largest likelihood is assigned to the residue.

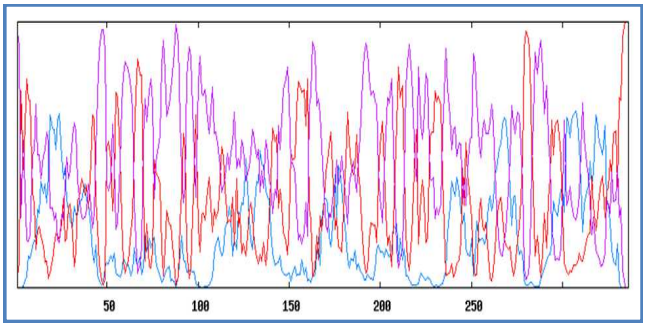

Fig. 7- Secondary structure plot of the *Toxoplasma gondii* major surface antigen P30
**Red: helix, Blue: Sheet, Pink: Coil*

Solvent Accessible Regions

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 that *Toxoplasma gondii* major surface antigen P30 *protein* is hydrophobic in nature and contains segments of low complexity and high-predicted flexibility [Fig-8], [Fig-9], [Fig-10], [Fig-11], [Fig-12], [Fig-13], [Fig-14], [Fig-15], [Fig-16], [Fig-17], [Fig-18], [Fig-19], [Fig-20], [Fig-21], [Fig-22], [Fig-23], [Fig-24], [Fig-25], [Fig-26], [Fig-27], [Fig-28].

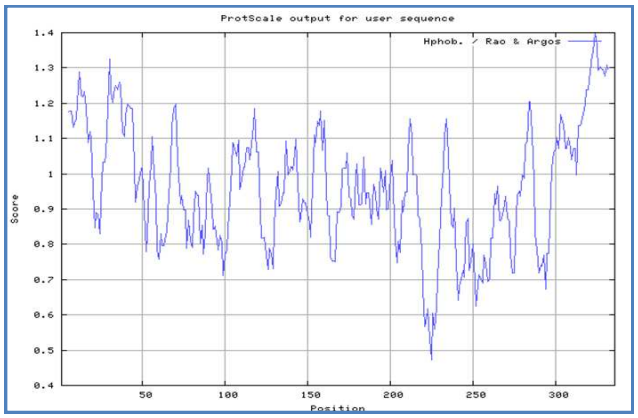

Fig. 8- Rao and Argos (1986) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

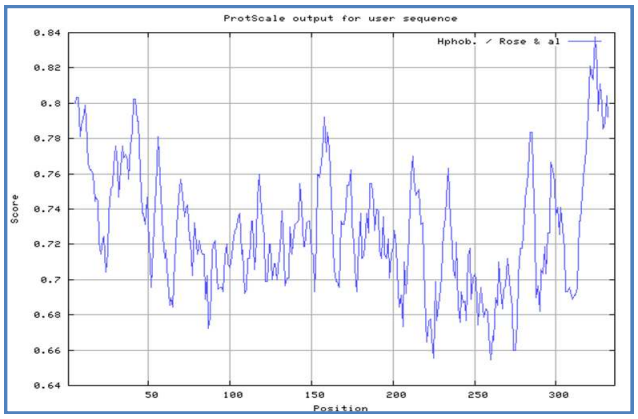

Fig. 9- Rose et al. (1985) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

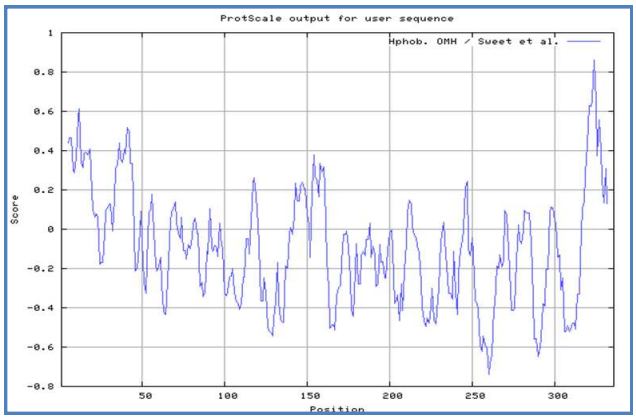

Fig. 10- Sweet Hydrophobicity plot of OMH for the *Toxoplasma gondii* major surface antigen P30

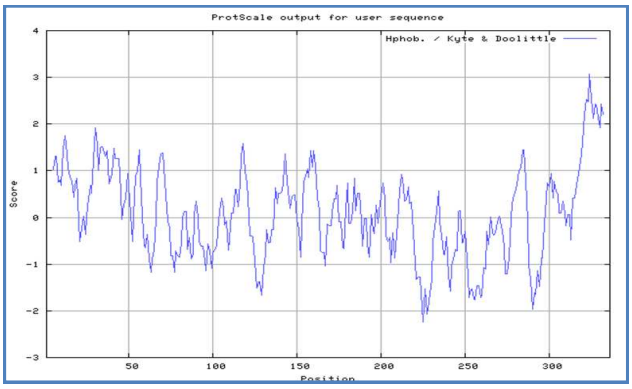

Fig. 11- Kyte and Doolittle (1982) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

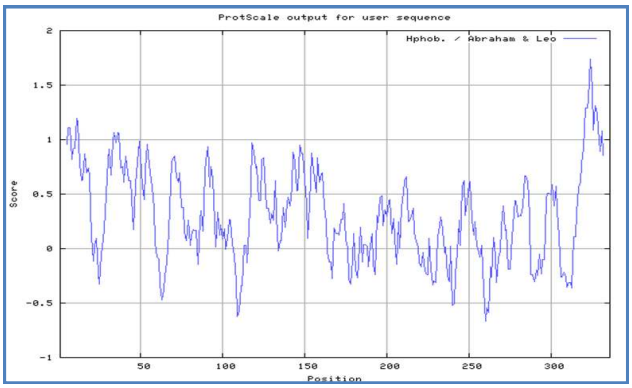

Fig. 12- Abraham and Leo (1987) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

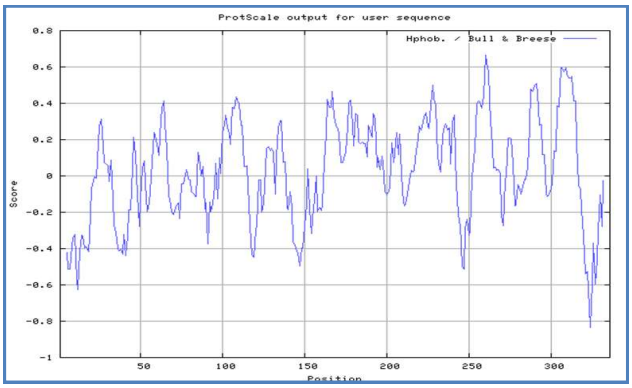

Fig. 13- Bull and Breese (1974) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

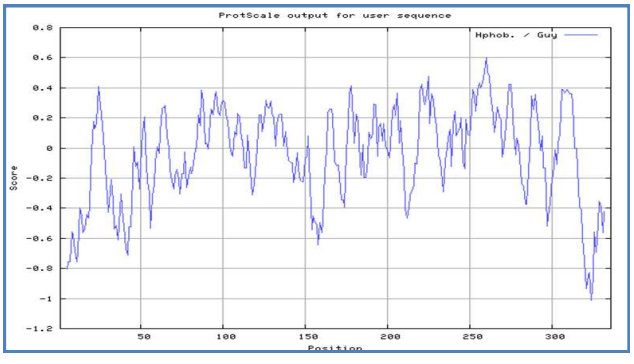

Fig. 14- Guy (1985) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

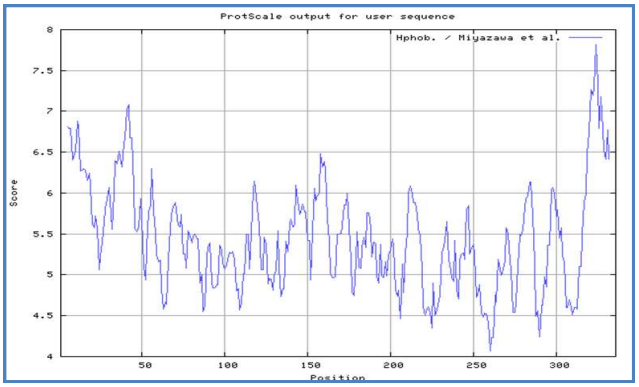

Fig. 15- Miyazawa, et al (1985) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

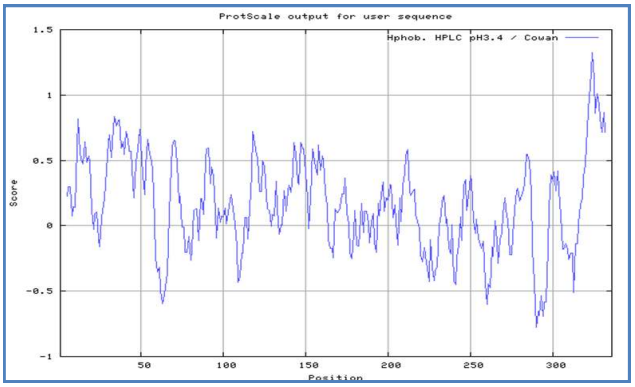

Fig. 19- Cowan (1990) Hydrophobicity plot of HPLC pH3.4 for the *Toxoplasma gondii* major surface antigen P30

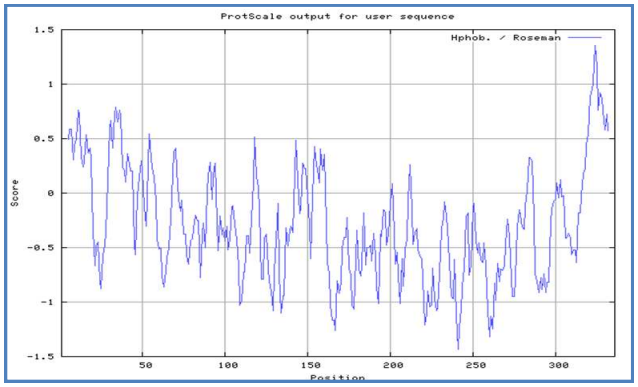

Fig. 16- Roseman (1988) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

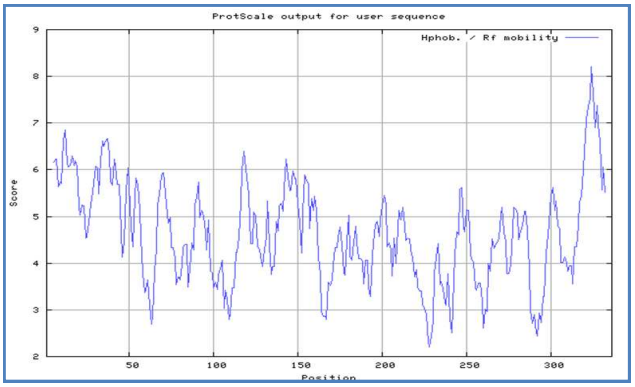

Fig. 20- RF mobility Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

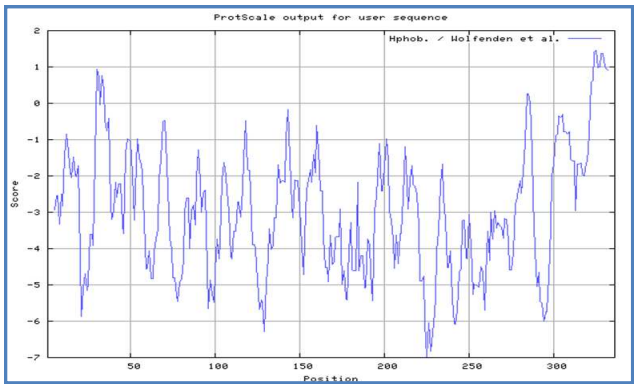

Fig. 17- Wolfenden et al. (1981) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

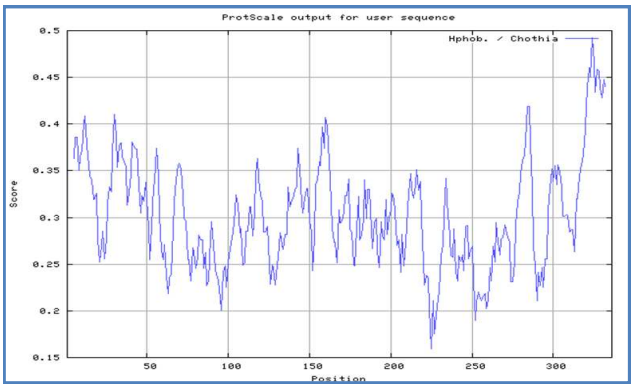

Fig. 21- Chothia (1976) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

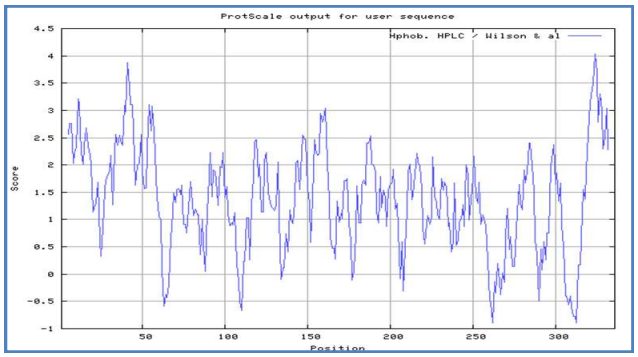

Fig. 18- Wilson et al. (1981) Hydrophobicity plot of HPLC for the *Toxoplasma gondii* major surface antigen P30

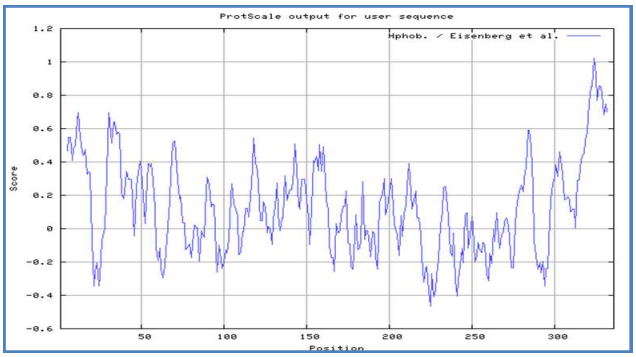

Fig. 22- Eisenberg et al. (1984) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

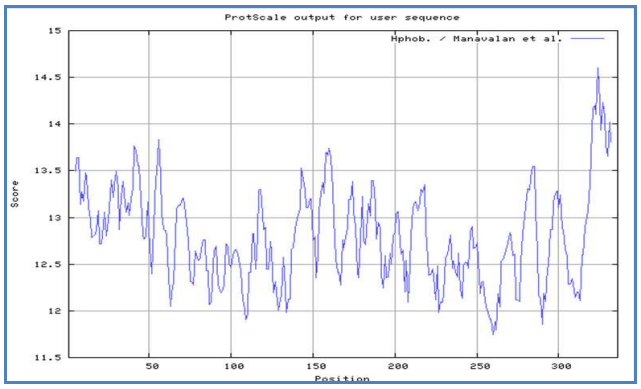

Fig. 23- Manavalan, et al (1978) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

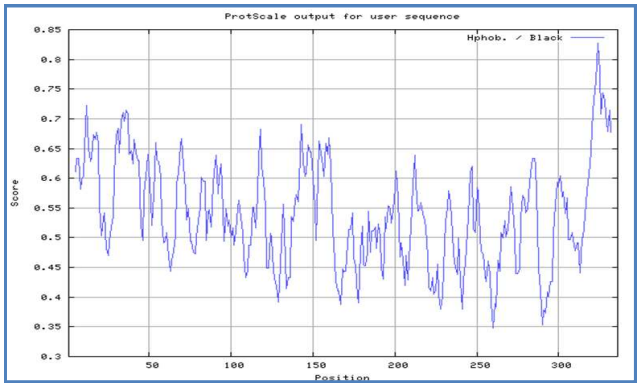

Fig. 24- Black (1991) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

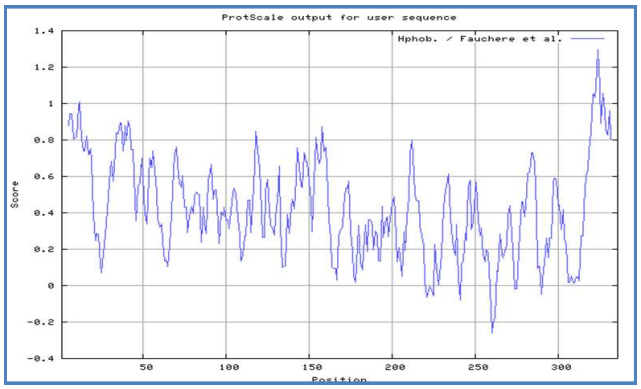

Fig. 25- Fauchere, et al (1983) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

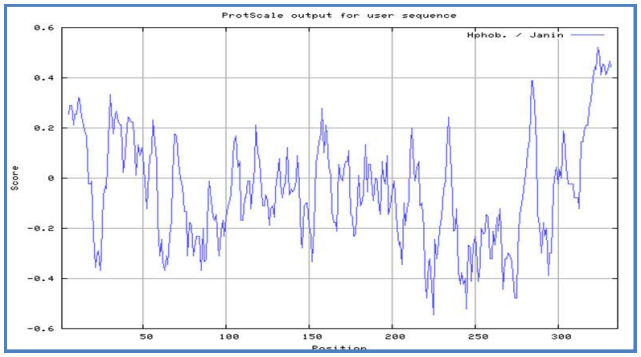

Fig. 26- Janin (1979) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

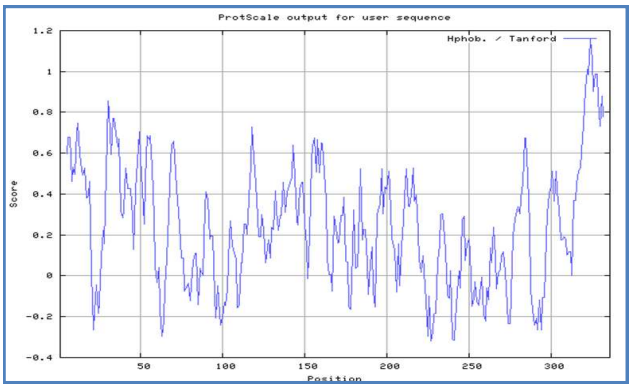

Fig. 27- Tanford (1962) Hydrophobicity plot for the *Toxoplasma gondii* major surface antigen P30

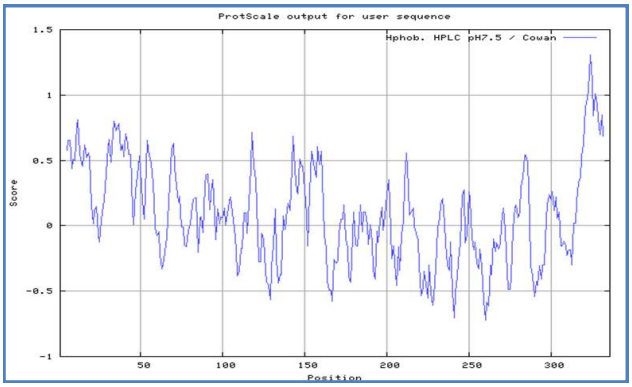

Fig. 28- Cowan (1990) Hydrophobicity plot of HPLC pH7.5 for the *Toxoplasma gondii* major surface antigen P30

Prediction of 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 so these MHC binding peptides are sufficient for eliciting the desired immune response. In analysis we determined the CTL binding regions, TAP binding regions and several potential antigenic epitopes. In this study we predicted the binding affinity of *Toxoplasma gondii* major surface antigen P30 protein having 336 amino acids, which shows number of peptides. The CTL epitope prediction is based on an elegant machine learning technique Support Vector Machine (SVM) and Artificial Neural Network (ANN). SVM has been trained on the binary input of single amino acid sequence. In this assay we have predicted the binding affinity of *Toxoplasma gondii* major surface antigen P30 protein sequence having 336 amino acids, which shows 328 nonamers.

The SVM based CTL epitopes, 179-RASSVVNNV, 79-TLKCPKTAL, 205-AEEPTTMTL, 94-AYSPNRQIC, 313-GTASHVSIF, 26-AVTAGVFAA, 242-SFKDILPKL, 259-ASSDKGATL, 319-SIFAMVIGL, 317-HVSIFAMVI, 171-MVTETVQAR, 314-TASHVSIFA, 19-FPKAVRRAV, 135-SLDTAGIKL, 17-SMFPKAVRR, 23-VRRAVTAGV, 182-SVVNNVARC, 186-NVARCSYGA, 315-ASHVSIFAM, 33-AAPTLMSFL, 251-TENPWQGNA, 316-SHVSIFAMV, 178-ARASSVVNN, 21-KAVRRAVTA, 32-FAAPTLMSF, 149-KFPVTTQTF, 22-AVRRAVTAG, 154-TQTFVVGCI, 49-DPPLVANQV, 271-KEAFPAESK, 74-TENHFTLKC, 277-

ESKSVIIGC, 175-TVQARASSV, 203-VSAEEPTTM, 181-SSVVNNVAR, 116-TLSSLIPEA, 239-NEKSFKDIL, 310-SAAGTASHV, 163-KGDDAQSCM, 72-TPTENHFTL, (optimal score is 1.378); the ANN based CTL epitopes recorded are 99-RQICPAGTT, 146-PIEKFPVTT, 35-PTLMSFLRC, 104-AGTTSSCTS, 249-LTENPWQG, 112-SKAVTLSSL, 117-LSSLIPEAE, 177-QARASSVVN, 36-TLMSFLRCG, 43-CGVMASDPP, 25-RAVTAGVFA, 79-TLKCPKTAL, 46-MASDPPLVA, 127-SWWTGDSAS, 176-VQARASSVV, 216-GKDGVKVPQ, 252-ENPWQGNAS, 114-AVTLSSLIP, 183-VVNNVARCS, 16-TSMFPKAVR, 18-MFPKAVRRA, 96-SPNRQICPA, 272-EAFPAESKS, 68-AVILTPTEN, 167-AQSCMVTET, 56-QVVTCPHKK, 243-FKDILPKLT, 294-HCTVKLEFA, 44-GVMASDPPL, 45-VMASDPPLV, 128-WWTGDSASL, 237-GCNEKSFKD, 129-WTGDSASLD, 164-GDDAQSCMV, 133-SASLDTAGI, 165-DDAQSCMVT, 287-GGSPEKHHC, 301-FAGAAGSAK, 132-DSASLDTAG, 214-VCGKDGVKV, 279-KSVIIGCTG, 286-TGGSPEKHH, 8-FIISSGFLT, 47-ASDPPLVAN, 125-EDSWWTGDS, 211-MTLVCGKDG, 145-VPIEKFPVT, 14-FLTSMFPKA, 78-FTLKCPKTA, 203-VSAEEPTTM, 285-CTGGSPEKH, 189-RCSYGADST, 24-RRAVTAGVF, 170-CMVTETVQA, 228-QYCSGTTLT, 233-TTLTGCNEK, 268-TIKKEAFPA, 251-TENPWQGNA, 312-AGTASHVSI, 134-ASLDTAGIK, 175-TVQARASSV, 253-NPWQGNASS, 9-IISSGFLTS, 292-KHHCTVKLE, 139-AGIKLTVPI, 126-DSWWTGDSA, 151-PVTTQTFVV, 215-CGKDGVKVP, 11-SSGFLTSMF, 246-ILPKLTENP, 57-VVTCPHKKS, 5-LHHFIISSG, 40-FLRCGVMAS, 313-GTASHVSIF, 106-TTSSCTSKA, 200-PVKVSAEEP, 159-VGCIKGDDA, 153-TTQTFVVGC, 244-KDILPKLTE, 274-FPAESKSVI, 238-CNEKSFKDI (optimal score is 1.000) which represented predicted binders from *Toxoplasma gondii* major surface antigen P30 protein [Table-1], [Table-2]. In addition to above CTL epitopes we also have predicted several high Affinity TAP binders as 275-PAESKSVII, 257-GNASSDKGA, 208-PTTMTLVCG, 18-MFPKAVRRA, 77-HFTLKCPKT, 156-TFVVGCIKG, 307-SAKSAAGTA, 108-SSCTSKAVT, 196-STLGPVKVS, 312-AGTASHVSI, 273-AFPAESKSV, 31-VFAAPTLMS, 197-TLGPVKVSA, 308-AKSAAGTAS, 258-NASSDKGAT, 3-VSLHHFIIS, 216-GKDGVKVPQ, 189-RCSYGADST, 177-QARASSVVN, 262-DKGATLTIK, 54-ANQVVTCPH, 236-TGCNEKSFK, 188-ARCSYGADS, 299-LEFAGAAGS, 233-TTLTGCNEK, 84-KTALTEPPT, 116-TLSSLIPEA, 5-LHHFIISSG, 155-QTFVVGCIK, 134-ASLDTAGIK, 14-FLTSMFPKA, 144-TVPIEKFPV, 192-YGADSTLGP, 316-SHVSIFAMV, 303-GAAGSAKSA, 183-VVNNVARCS, 36-TLMSFLRCG, 261-SDKGATLTI, 295-CTVKLEFAG, 194-ADSTLGPVK, 320-IFAMVIGLI, 56-QVVTCPHKK, 10-ISSGFLTSM, 72-TPTENHFTL, 133-SASLDTAGI, 123-EAEDSWWTG, 179-RASSVVNNV, 173-TETVQARAS, 214-VCGKDGVKV, 296-TVKLEFAGA, 162-IKGDDAQSC, 70-ILTPTENHF, 163-KGDDAQSCM, 204-SAEEPTTMT, 126-DSWWTGDSA, 170-CMVTETVQA, 272-EAFPAESKS, 203-VSAEEPTTM, 141-IKLTVPIEK, 279-KSVIIGCTG, 45-VMASDPPLV, 151-PVTTQTFVV, 212-TLVCGKDGV, 169-SCMVTETVQ, 74-TENHFTLKC, 107-TSSCTSKAV, 180-ASSVVNNVA, 201-VKVSAEEPT, 185-NNVARCSYG, 91-PTLAYSPNR, 280-SVIIGCTGG, 62-HKKSTAAVI, 220-VKVPQDNNQ, 105-GTTSSCTSK, 23-VRRAVTAGV, 206-EEPTTMTLV, 248-PKLTENPWQ, 314-TASHVSIFA, 243-FKDILPKLT, 59-TCPHKKSTA, 15-LTSMFPKAV, 202-KVSAEEPTT, 249-KLTENPWQG, 27-VTAGVFAAP, 291-EKHHCTVKL, 277-ESKSVIIGC, 323-MVIGLIGSI, 41-LRCGVMASD, 268-TIKKEAFPA, 263-KGATLTIKK, 260-SSDKGATLT, 311-AAGTASHVS, 4-SLHHFIISS, 288-GSPEKHHCT, 57-VVTCPHKKS, 154-TQTFVVGCI, 276-AESKSVIIG, 230-CSGTTLTGC, 42-RCGVMASDP, 120-LIPEAEDSW, 213-LVCGKDGVK, 223-PQDNNQYCS, 53-VANQVVTCP, 294-HCTVKLEFA, 148-EKFPVTTQT, 193-GADSTLGPV, 99-RQICPAGTT, 9-IISSGFLTS, 171-MVTETVQAR, 11-SSGFLTSMF, 146-PIEKFPVTT, 98-NRQICPAGT, 64-KSTAAVILT, 244-KDILPKLTE, 221-KVPQDNNQY, 228-QYCSGTTLT, 306-GSAKSAAGT, 103-PAGTTSSCT, 326-GLIGSIAAC, 80-LKCPKTALT, 83-PKTALTEPP, 226-NNQYCSGTT, 8-FIISSGFLT, 317-HVSIFAMVI, 310-SAAGTASHV, 67-AAVILTPTE, 168-QSCMVTETV, these TAP binders are important in generating immune response [Table-3]. The average propensity for the *Toxoplasma gondii* major surface antigen P30 protein found is 1.044 [Fig-4]. All residues having more than 1.0 propensity are always potentially antigenic [Table-4].

Table 1- SVM based Predicted CTL-Epitopes of Toxoplasma gondii major surface antigen P30 protein

Peptide Rank	Start Position	Sequence	Score
1	179	RASSVVNNV	1.378
2	79	TLKCPKTAL	1.209
3	205	AEEPTTMTL	1.108
4	94	AYSPNRQIC	1.093
5	313	GTASHVSIF	0.912
6	26	AVTAGVFAA	0.886
7	242	SFKDILPKL	0.875
8	259	ASSDKGATL	0.836
9	319	SIFAMVIGL	0.824
10	317	HVSIFAMVI	0.783
11	171	MVTETVQAR	0.778
12	314	TASHVSIFA	0.765
13	19	FPKAVRRAV	0.738
14	135	SLDTAGIKL	0.733
15	17	SMFPKAVRR	0.728
16	23	VRRAVTAGV	0.698
17	182	SVVNNVARC	0.674
18	186	NVARCSYGA	0.665
19	315	ASHVSIFAM	0.654
20	33	AAPTLMSFL	0.596
21	251	TENPWQGNA	0.562
22	316	SHVSIFAMV	0.536
23	178	ARASSVVNN	0.532
24	21	KAVRRAVTA	0.502
25	32	FAAPTLMSF	0.502
26	149	KFPVTTQTF	0.502
27	22	AVRRAVTAG	0.465
28	154	TQTFVVGCI	0.458
29	49	DPPLVANQV	0.434
30	271	KEAFPAESK	0.428
31	74	TENHFTLKC	0.419
32	277	ESKSVIIGC	0.416
33	175	TVQARASSV	0.411
34	203	VSAEEPTTM	0.408
35	181	SSVVNNVAR	0.4
36	116	TLSSLIPEA	0.399
37	239	NEKSFKDIL	0.382
38	310	SAAGTASHV	0.381
39	163	KGDDAQSCM	0.369
40	72	TPTENHFTL	0.368

Table 2- ANN based Predicted CTL-Epitopes of Toxoplasma gondii major surface antigen P30 protein

Peptide Rank	Start Position	Sequence	Score
1	99	RQICPAGTT	1
2	146	PIEKFPVTT	1
3	35	PTLMSFLRC	0.99
4	104	AGTTSSCTS	0.99
5	249	KLTENPWQG	0.99
6	112	SKAVTLSSL	0.98
7	117	LSSLIPEAE	0.98
8	177	QARASSVVN	0.98
9	36	TLMSFLRCG	0.97
10	43	CGVMASDPP	0.97
11	25	RAVTAGVFA	0.96
12	79	TLKCPKTAL	0.96
13	46	MASDPPLVA	0.95
14	127	SWWTGDSAS	0.95
15	176	VQARASSVV	0.95
16	216	GKDGVKVPQ	0.95
17	252	ENPWQGNAS	0.94
18	114	AVTLSSLIP	0.93
19	183	VVNNVARCS	0.93
20	16	TSMFPKAVR	0.92
21	18	MFPKAVRRA	0.92
22	96	SPNRQICPA	0.92
23	272	EAFPAESKS	0.92
24	68	AVILTPTEN	0.91
25	167	AQSCMVTET	0.91
26	56	QVVTCPHKK	0.89
27	243	FKDILPKLT	0.89
28	294	HCTVKLEFA	0.89
29	44	GVMASDPPL	0.88
30	45	VMASDPPLV	0.88
31	128	WWTGDSASL	0.87
32	237	GCNEKSFKD	0.87
33	129	WTGDSASLD	0.86
34	164	GDDAQSCMV	0.86
35	133	SASLDTAGI	0.85
36	165	DDAQSCMVT	0.85
37	287	GGSPEKHHC	0.85
38	301	FAGAAGSAK	0.84
39	132	DSASLDTAG	0.83
40	214	VCGKDGVKV	0.83
41	279	KSVIIGCTG	0.81
42	286	TGGSPEKHH	0.81
43	8	FIISSGFLT	0.79
44	47	ASDPPLVAN	0.79
45	125	EDSWWTGDS	0.77
46	211	MTLVCGKDG	0.77
47	145	VPIEKFPVT	0.75
48	14	FLTSMFPKA	0.74
49	78	FTLKCPKTA	0.74
50	203	VSAEEPTTM	0.74
51	285	CTGGSPEKH	0.73
52	189	RCSYGADST	0.72
53	24	RRAVTAGVF	0.71
54	170	CMVTETVQA	0.71
55	228	QYCSGTTLT	0.71
56	233	TTLTGCNEK	0.7
57	268	TIKKEAFPA	0.7
58	251	TENPWQGNA	0.69
59	312	AGTASHVSI	0.69
60	134	ASLDTAGIK	0.68
61	175	TVQARASSV	0.68
62	253	NPWQGNASS	0.67
63	9	IISSGFLTS	0.66
64	292	KHHCTVKLE	0.66
65	139	AGIKLTVPI	0.64

Table 2-Continue

Peptide Rank	Start Position	Sequence	Score
66	126	DSWWTGDSA	0.63
67	151	PVTTQTFVV	0.62
68	215	CGKDGVKVP	0.6
69	11	SSGFLTSMF	0.59
70	246	ILPKLTENP	0.59
71	57	VVTCPHKKS	0.58
72	5	LHHFIISSG	0.57
73	40	FLRCGVMAS	0.57
74	313	GTASHVSIF	0.57
75	106	TTSSCTSKA	0.55
76	200	PVKVSAEEP	0.55
77	159	VGCIKGDDA	0.54
78	153	TTQTFVVGC	0.53
79	244	KDILPKLTE	0.52
80	274	FPAESKSVI	0.52
81	238	CNEKSFKDI	0.51

Table 3- Cascade SVM based High affinity TAP epitopes of Toxoplasma gondii major surface antigen P30 protein

Peptide Rank	Start Position	Sequence	Score
1	275	PAESKSVII	8.643
2	257	GNASSDKGA	8.642
3	208	PTTMTLVCG	8.641
4	18	MFPKAVRRA	8.636
5	77	HFTLKCPKT	8.628
6	156	TFVVGCIKG	8.626
7	307	SAKSAAGTA	8.622
8	108	SSCTSKAVT	8.621
9	196	STLGPVKVS	8.617
10	312	AGTASHVSI	8.614
11	273	AFPAESKSV	8.61
12	31	VFAAPTLMS	8.608
13	197	TLGPVKVSA	8.605
14	308	AKSAAGTAS	8.603
15	258	NASSDKGAT	8.602
16	3	VSLHHFIIS	8.596
17	216	GKDGVKVPQ	8.591
18	189	RCSYGADST	8.58
19	177	QARASSVVN	8.579
20	262	DKGATLTIK	8.578
21	54	ANQVVTCPH	8.571
22	236	TGCNEKSFK	8.558
23	188	ARCSYGADS	8.538
24	299	LEFAGAAGS	8.536
25	233	TTLTGCNEK	8.518
26	84	KTALTEPPT	8.515
27	116	TLSSLIPEA	8.514
28	5	LHHFIISSG	8.511
29	155	QTFVVGCIK	8.503
30	134	ASLDTAGIK	8.479
31	14	FLTSMFPKA	8.469
32	144	TVPIEKFPV	8.469
33	192	YGADSTLGP	8.464
34	316	SHVSIFAMV	8.45
35	303	GAAGSAKSA	8.424
36	183	VVNNVARCS	8.417
37	36	TLMSFLRCG	8.41
38	261	SDKGATLTI	8.402
39	295	CTVKLEFAG	8.395
40	194	ADSTLGPVK	8.388
41	320	IFAMVIGLI	8.369
42	56	QVVTCPHKK	8.367
43	10	ISSGFLTSM	8.363
44	72	TPTENHFTL	8.349
45	133	SASLDTAGI	8.333

Table 3- Continue

Peptide Rank	Start Position	Sequence	Score
46	123	EAEDSWWTG	8.33
47	179	RASSVVNNV	8.318
48	173	TETVQARAS	8.313
49	214	VCGKDGVKV	8.301
50	296	TVKLEFAGA	8.245
51	162	IKGDDAQSC	8.244
52	70	ILTPTENHF	8.24
53	163	KGDDAQSCM	8.214
54	204	SAEEPTTMT	8.18
55	126	DSWWTGDSA	8.158
56	170	CMVTETVQA	8.135
57	272	EAFPAESKS	8.101
58	203	VSAEEPTTM	8.099
59	141	IKLTVPIEK	8.099
60	279	KSVIIGCTG	8.08
61	45	VMASDPPLV	8.047
62	151	PVTTQTFVV	8.023
63	212	TLVCGKDGV	7.974
64	169	SCMVTETVQ	7.936
65	74	TENHFTLKC	7.913
66	107	TSSCTSKAV	7.913
67	180	ASSVVNNVA	7.901
68	201	VKVSAEEPT	7.873
69	185	NNVARCSYG	7.853
70	91	PTLAYSPNR	7.826
71	280	SVIIGCTGG	7.797
72	62	HKKSTAAVI	7.782
73	220	VKVPQDNNQ	7.741
74	105	GTTSSCTSK	7.735
75	23	VRRAVTAGV	7.61
76	206	EEPTTMTLV	7.599
77	248	PKLTENPWQ	7.58
78	314	TASHVSIFA	7.575
79	243	FKDILPKLT	7.564
80	59	TCPHKKSTA	7.517
81	15	LTSMFPKAV	7.461
82	202	KVSAEEPTT	7.456
83	249	KLTENPWQG	7.448
84	27	VTAGVFAAP	7.421
85	291	EKHHCTVKL	7.401
86	277	ESKSVIIGC	7.396
87	323	MVIGLIGSI	7.34

Table 3- Continue

Peptide Rank	Start Position	Sequence	Score
88	41	LRCGVMASD	7.288
89	268	TIKKEAFPA	7.271
90	263	KGATLTIKK	7.25
91	260	SSDKGATLT	7.245
92	311	AAGTASHVS	7.24
93	4	SLHHFIISS	7.189
94	288	GSPEKHHCT	7.178
95	57	VVTCPHKKS	7.176
96	154	TQTFVVGCI	7.101
97	276	AESKSVIIG	7.098
98	230	CSGTTLTGC	7.079
99	42	RCGVMASDP	7.033
100	120	LIPEAEDSW	6.991
101	213	LVCGKDGVK	6.941
102	223	PQDNNQYCS	6.935
103	53	VANQVVTCP	6.922
104	294	HCTVKLEFA	6.877
105	148	EKFPVTTQT	6.867
106	193	GADSTLGPV	6.853
107	99	RQICPAGTT	6.741
108	9	IISSGFLTS	6.733
109	171	MVTETVQAR	6.704
110	11	SSGFLTSMF	6.69
111	146	PIEKFPVTT	6.683
112	98	NRQICPAGT	6.635
113	64	KSTAAVILT	6.581
114	244	KDILPKLTE	6.522
115	221	KVPQDNNQY	6.522
116	228	QYCSGTTLT	6.513
117	306	GSAKSAAGT	6.463
118	103	PAGTTSSCT	6.432
119	326	GLIGSIAAC	6.356
120	80	LKCPKTALT	6.345
121	83	PKTALTEPP	6.288
122	226	NNQYCSGTT	6.21
123	8	FIISSGFLT	6.196
124	317	HVSIFAMVI	6.181
125	310	SAAGTASHV	6.18
126	67	AAVILTPTE	6.11
127	168	QSCMVTETV	6.081

Table 4- Potential Antigenic epitopes of Toxoplasma gondii major surface antigen P30 protein

Start Position	End Position	Peptide	Peptide Length
4	14	SLHHFIISSGF	11
20	72	PKAVRRAVTAGVFAAPTLMSFLRCGVMASDPPLVANQVVTCPHKKSTAAVILT	53
79	85	TLKCPKT	7
88	96	TEPPTLAYS	9
98	105	NRQICPAG	8
107	123	TSSCTSKAVTLSSLIPE	17
136	162	LDTAGIKLTVPIEKFPVTTQTFVVGCI	27
167	193	AQSCMVTETVQARASSVVNNVARCSYG	27
197	206	TLGPVKVSAE	10
211	223	MTLVCGKDGVKVP	13
227	233	NQYCSGT	7
244	250	KDILPKL	7
276	287	AESKSVIIGCTG	12
292	302	KHHCTVKLEFA	11

The predicted segments in *Toxoplasma gondii* major surface antigen P30 protein are 4-SLHHFIISSGF-14, 20-PKAVRRAVTAGVFAAPTLMSFLRCGVMASDPPLVANQVVTCPHKKSTAAVILT-72, 79-TLKCPKT-85, 88-TEPPTLAYS-96, 98-NRQICPAG-105, 107-TSSCTSKAVTLSSLIPE-123, 136-LDTAGIKLTVPIEKFPVTTQTFVVGCI-162, 167-AQSCMVTETVQARASSVVNNVARCSYG-193, 197-

TLGPVKVSAE-206, 211-MTLVCGKDGVKVP-223, 227-NQYCSGT -233, 244-KDILPKL-250, 276-AESKSVIIGCTG-287, 292-KHHCTV KLEFA-302 Fragments identified through this approach tend to be high-efficiency binders, Which is a larger percentage of their atoms are directly involved in binding as compared to larger molecules.
These MHC binding peptides are sufficient for inducing the desired immune response. Predicted MHC binding regions in *Toxoplasma gondii* major surface antigen P30 sequence and these are actively taking part in immune reactions.

Discussion and Conclusion

Gomase method (2007), B-EpiPred Server, Hopp and Woods, Welling, Parker, Kolaskar and Tongaonkar antigenicity scales were designed to predict the locations of antigenic determinants in *Toxoplasma gondii* major surface antigen P30 protein sequence. It shows beta sheets regions, which have higher antigenic response than helical region of this peptide and shows high antigenicity [Fig-1] to [Fig-5]. We also found the Sweet hydrophobicity, Kyte & Doolittle hydrophobicity, Abraham & Leo, Bull & Breese hydrophobicity, Guy, Miyazawa hydrophobicity, Roseman hydrophobicity, Cowan HPLC pH7.5 hydrophobicity, Rose hydrophobicity, Eisenberg hydrophobicity, Manavalan hydrophobicity, Black hydrophobicity, Fauchere hydrophobicity, Janin hydrophobicity, Rao & Argos hydrophobicity, Wolfenden hydrophobicity, Wilson HPLC hydrophobicity, Cowan HPLC pH3.4, Tanford hydrophobicity, RF mobility hydrophobicity and Chothia hydrophobicity scales, These scales are essentially a hydrophilic index, with a polar residues assigned negative values [Fig-8] to [Fig-28]. In this assay we predicted the binding affinity of *Toxoplasma gondii* major surface antigen P30 protein having 336 amino acids, which shows 328 nonamers. We predicted SVM and ANN based CTL epitopes [Table-1], [Table-2].
We have determined fourty CTL predicted epitopes (Optimal score is 1.378) at cut off 0.36 using SVM based method [Table-1] and eighty one CTL predicted epitopes using ANN based method at cut off 0.51 (Optimal score is 1.000) [Table-2] which represents predicted peptide binders from *Toxoplasma gondii* major surface antigen P30 protein. We also have predicted Cascade SVM based One twenty seven High affinity TAP binding epitopes (Optimal score is 8.643) [Table-3] in addition to fourteen potentially antigenic peptides recognized by antibodies of the immune system for *Toxoplasma gondii* major surface antigen P30 protein in the region of maximum local hydrophilicity [Table-4]. Kolaskar and Tongaonkar antigenicity are the sites of molecules that are recognized by antibodies of the immune system for *Toxoplasma gondii* major surface antigen P30 protein, analysis shows epitopes present in the major surface antigen P30 protein of *Toxoplasma gondii* elicites desired immune response. The region of maximal hydrophilicity is likely to be an antigenic site, having hydrophobic characteristics, because C-terminal regions of *Toxoplasma gondii* major surface antigen P30 protein is solvent accessible and unstructured, antibodies against those regions are also likely to recognize the native protein. For the prediction of antigenic determinant site of *Toxoplasma gondii* major surface antigen P30 protein, we predicted four antibody recognized antigenic determinant sites in the *Toxoplasma gondii* major surface antigen P30 sequence. The highest pick is recorded between sequence of amino acid in the region are 20-PKAVRRAVTAGVFAAPTLMSFLRCGVMASDPPLVANQVVT CP HKKSTAAVILT-72 and 136-LDTAGIKLTVPIEKFP VTTQT FVVGCI -162 [Table-4].

Future Perspectives

This method will be useful in cellular immunology, Vaccine design, immunodiagnostics, immunotherapeutics and molecular understanding of autoimmune susceptibility. *Toxoplasma gondii* major surface antigen P30 protein sequence contains multiple antigenic components to direct and empower the immune system to protect an individual from toxoplasmosis disease. MHC molecules are cell surface proteins, which actively participates in host immune reactions and involvement of MHC class in response to almost all antigens and it give effects on target sites. Predicted MHC binding regions acts like red flags for specific antigen and generate immune response against the whole antigen. So a small antigen fragment can generate immune response against entire antigen. The method integrates prediction of Peptide-MHC class binding; proteosomal C terminal cleavage and TAP transport efficiency. This approach is implemented in designing subunit and synthetic peptide vaccines.

References

[1] Henriquez S.A., Brett R., Alexander J., Pratt J., Roberts C.W. (2009) *Neuroimmunomodulation*, 16(2), 122-33.
[2] Montoya J., Liesenfeld O. (2004) *Lancet,* 363(9425), 1965-76.
[3] Valkonen J.P., Rajamäki M.L., Kekarainen T. (2002) *Mol. Plant Microbe. Interact.*, 15, 683-92.
[4] Bhasin M. and Raghava G.P.S. (2004) *Vaccine*, 22, 3195-201.
[5] Kumar M., Gromiha M.M., Raghava G.P. (2007) *BMC Bioinformatics,* 8, 463.
[6] Gomase V.S., Kale K.V., Chikhale N.J., Changbhale S.S. (2007) *Curr. Drug Discov. Technol*., 4, 117-1215.
[7] Gomase V.S. and Chitlange N.R. (2011) *Bioinfo Journal of Proteomics,* 1(1), 06-10.
[8] Gomase V.S. and Chitlange N.R. (2012) *Drug Designing*, 1(1), 101.
[9] Gomase V.S. and Chitlange N.R. (2012) *Journal of Vaccines & Vaccination* 1(1), 131.
[10] Schirle M., Weinschenk T., Stevanovic S. (2001) J. Immunol. Methods, 257, 1-16.
[11] Gomase V.S., Changbhale S.S., Kale K.V. (2008) *Advancements in Information Technology and Internet Security*, 370-378.
[12] Gomase V.S., Changbhale S.S., Patil S.A. and Kale K.V. (2008) *Current Drug Metabolism*, 9(1), 88-98.
[13] Rammensee H., Bachmann J., Emmerich N.P., Bachor O.A., Stevanovic S. (1999) *Immunogenetics,* 50, 213-9.
[14] Blythe M.J., Doytchinova I.A., Flower D.R. (2002) *Bioinformatics,* 18, 434-9.
[15] Schonbach C., Koh J.L., Flower D.R., Wong L., Brusic V. (2002) *Nucleic Acids Res.,* 30, 226-9.
[16] Korber T.M.B., Brander C., Haynes B.F., Koup R., Kuiken C., et al. (2001) *Theoretical Biology and Biophysics*, Los Alamos, New Mexico. LA-UR 02-4663.
[17] Saritha R.K., Jain R.K. (2007) *Arch. Virol*., 152(6), 1195-1200.
[18] Gomase V.S. (2006) *Curr. Drug Discov. Technol.,* 3, 225-229.
[19] Hopp T.P., Woods K.R. (1981) *Proc. Natl. Acad. Sci.* USA, 78, 3824-3828.
[20] Welling G.W., Weijer W.J., van der Zee R., Welling-Wester S. (1985) FEBS Lett., 188, 215-218.
[21] Jens E.P.L., Ole L., Morten N. (2006) *Immunome. Res.,* 2(2).
[22] Parker J.M.R., Guo D., Hodges R.S. (1986) *Biochemistry,* 25,

5425- 5431.
[23]Kolaskar A.S., Tongaonkar P.C. (1990) *FEBS Lett*., 276, 172-174.
[24]Garnier J., Osguthorpe D.J., Robson B. (1978) J. Mol. Biol., 120, 97-120.
[25]Robson B., Garnier J. (1993) *Nature,* 361(506).
[26]Sweet R.M., Eisenberg D. (1983) *J. Mol. Biol.,* 171, 479-488.
[27]Abraham D.J., Leo A.J. (1987) *Proteins,* 2, 130-152.
[28]Bull H.B., Breese K. (1974) *Arch. Biochem. Biophys.*, 161, 665-670.
[29]Guy H.R. (1985) *Biophys. J.*, 47, 61-70.
[30]Miyazawa S., Jernigen R.L. (1985) *Macromolecules,* 18, 534-552.
[31]Roseman M.A. (1988) *J. Mol. Biol*., 200, 513-522.
[32]Wolfenden R.V., Andersson L., Cullis P.M., Southgate C.C.F. (1981) *Biochemistry,* 20, 849-855.
[33]Wilson K.J., Honegger A., Stotzel R.P., Hughes G.J. (1981) *Biochem J.,* 199, 31-41.
[34]Chothia C. (1976) *J. Mol. Biol.,* 105, 1-14.
[35]Eisenberg D., Schwarz E., Komaromy M., Wall R. (1984) *J. Mol. Biol.,* 179, 125-142.
[36]Eisenberg D., Weiss R.M., Terwilliger T.C. (1984) *Proc. Natl. Acad. Sci. USA,* 81, 140-4.
[37]Manavalan P., Ponnuswamy P.K. (1978) *Nature,* 275, 673-674.
[38]Black S.D., Mould D.R. (1991) *Anal. Biochem.,* 193, 72-82.
[39]Fauchere J.L., Pliska V.E. (1983) *Eur. J. Med. Chem*., 18, 369-375.
[40]Janin J. (1979) *Nature,* 277, 491-492.
[41]Rao M.J.K., Argos P. (1986) *Biochim. Biophys. Acta.,* 869, 197-214.
[42]Tanford C. (1962) J. Am. Chem. Soc., 84, 4240-4274.
[43]Cowan R., Whittaker R.G. (1990) *Pept. Res.,* 3, 75-80.
[44]Kyte J., Doolittle R.F. (1982) *J. Mol. Biol.,* 157, 105-132.
[45]Rose G.D., Geselowitz A.R., Lesser G.J., Lee R.H., Zehfus M.H. (1985) *Science,* 229, 834-838.
[46]Wilkins M.R., Gasteiger E., Bairoch A., Sanchez J.C., Williams K.L., et al. (1999) *Methods Mol. Biol.,* 112, 531-52.
[47]Tanford C. (1962) *J. Am. Chem. Soc.*, 84, 4240-4274.
[48]Bhasin M. and Raghava G.P.S. (2004) Analysis and prediction of affinity of TAP binding peptides using cascade SVM. Protein Sci., 13 (3), 596-607.