Review Paper:

Application of *in silico* Approach in Prediction of Epitopes

Mishra Sonu* and Gomase Virendra S.
Department of Biotechnology, Mewar University, Chittorgarh, Rajasthan, 312901, INDIA
*sonumishra1014@gmail.com

Abstract

The identification of most accurate epitopes from the in silico methodology is the commonest strategy to develop novel immunotherapy; the essential part is the functional T-cell epitope discovery which has arisen importantly in recent years. The full-length protein scans through high-throughput experimental methods made it possible for epitopes bounded to a confined count of MHC alleles. The highest interest of the in silico approach is due to the high cost and limitation associated with the number of proteins and MHC alleles that are feasibly handled via such experimental methods. The available prediction methodology for MHC binding peptides is of high quality and accuracy which can predict the large grade of MHC alleles and relevant length of binding peptides.

The prediction is straightforward to execute for complete proteomes irrespective of its size and can be effectively utilized for rational epitope discovery. In the near future, we prospect to look for great development in prediction methodology and experimental validation methods and expect to see the clinical trial of products whose expansion has been headed by prediction methods.

Keywords: CTL, Epitope, HLA, MHC, Epitope prediction, T cell, Vaccines.

Introduction

The natural defense system in almost entire multicellular organisms is an immune system that protects the host from several acute and chronic illnesses. The immune system in totality acts against a larger number of factors, out of which the several are not of protein origin. The innate immunity acts against various non-self compounds at a faster rate without harming healthy and self-cells whereas the adaptive system performs its action against several specific proteins and peptides that are particular for pathogenic cells and foreign organisms[49].

Originally, epitopes describe as a portion of an antigen that specifies the linking to an immunoglobulin[21]. Antibodies also bind to native protein straightly and well-defined manner and interactions without any contributed interaction with other proteins. In general, the epitopes comprised of part of processed protein and form complex with part of the host-expressed MHC protein. Being a component of the protein (peptide), it triggers immune responses efficiently and thus referred to as epitopes and the native protein from which the epitopes belong to the antigen. Henceforth an antigen refers to a protein that illustrates immune response within a host and the T-cell epitopes refer to a sub-peptide from the antigen which forms complex with MHC and capable of recognized by antigen-specific TCRs.

In the arena of the traditional protective vaccine research, epitopes discovery has its major interest. The detection of specific epitopes can contribute to the electing of essential antigens and further helps in defining the selected antigens essential parts[58].

Currently, the designing of the minimal vaccines is in practice with the knowledgeable guidance of the specific antigen and epitopes which also comprised of the relevant antigen[44]. In the near future, this can lead to discovering vaccines of minimum unreal polypeptides designed to contain many strong relevant epitopes for the peculiar illness[59].

To fight against the established diseases like cancer and other chronic illnesses is the core focus of the researcher to utilize immune-related techniques for its cure and this approach interest is increasing continuously. Conceptually two different approaches are usually applied clinically in the immunotherapy. First, utilization of the monoclonal antibodies for a peculiar disease target and secondly through manipulation of the host's immune system keeping the focus on specific illness-relevant targets are or else being ignored by host immunity. Here, cautiously selected epitopes might be utilized to trigger a more powerful and specific immune response[22].

In course of therapeutic drug targets, considering protein and peptides with the strong knowledge of the effective strong epitopes, a protein bears further importance. The epitopes identification can also be utilized for de-selection or de-immunization and it is of ecumenical pursuit in basic immunology. The investigation of the immunogenic responses, the know-how of the epitopes and the specific signaling pathways which coincide with peculiar epitopes is an important part of immunological research that can contribute to getting depth knowledge of the behavior and evolution of immune system[5].

Processing and Displaying of T-Cell Epitopes: Based on the several different approaches of the prediction model, the important events occur in cellular responses which are

summarized and described in the following details except for other details events that are not currently predictable through mathematical models or computational algorithms.

Once the complex is formed between T-cell epitopes with MHC proteins, it identified through TCRs on the surface of T cells with different functionality, T-helper (Th) cells and cytotoxic T lymphocytes (CTL). Apart from having dissimilar co-receptors, respectively CD4 and CD8, the MHC molecules introducing epitopes of Th or CTL have significant differences that appear to be ecumenical for jawed vertebrates, with one latterly found out an exception in cod fish that deficient of the genes that are particular for Th reactions[60]. Furthermore, only a limited panel of animals has been looked into in-depth concerning to their immune systems, therefore we only consider primarily the knowledge obtained by studying the human immune system.

The CTL triggered at a point when a TCR on a CD8+ T cell binds to an MHC class-I peptide complex introduced via a nucleated cell. The limited length of the binding peptide cause of its binding pattern to the MHC class I in a specified binding groove at the endpoint and preferable length of amino acids are nine. In the availability of the free peptides in the solution, the MHC class I protein capable to interchange bound peptides with them but in general, on the cell surface presented peptides originate within the cell from the long chain of polypeptide and protein.

In current reviews, the detail description of the presentation of pathways of MHC class I peptide is elaborated.[14,15,37] The protein tagged with regulatory protein ubiquitin is digested via the proteasome complex and generates short peptides. This process occurred in normal circumstances through the constitutive proteasome that has strong stochastic elements in cleavage preferences.[31,61] The generated peptides, some of them bind to the transporter associated with antigen processing (TAP) and transported into the endoplasmic reticulum (ER). After getting into the ER, the peptide is held securely at the inner side of the ER membrane will encounter semi-folded MHC class I molecules. Upon complete peptides entering in the ER, the specific peptides trigger encouraged folding of MHC and later on altered through the protein tapasin that has a chaperone-like function.[54,55] Complete folded MHC-peptide complexes will be carried to the cell surface where they will be present for binding by a desirable TCR.

The induction of the responses occurs when TCR on a CD4+ T cell binds to an MHC class II–peptide complex, which is commonly explicit in experts APCs[72]. The binding of peptides of any length is permission due to the open end groove of the MHC class II. The substantial processing occurs to the peptide before presenting by MHC class II at the cell surface and several times this process has been reviewed at various places in details.[8,32,38,76] The native protein partly degrades protein or protein complexes from extracellular space are the origin of the presented peptide origination. Through the procedure of the phagocytosis the protein gets entry into the APC and completes the whole pathways of endosomic, the process includes denaturation and processing of protein via cathepsins like proteases.[7,80] The fully folded invariant chain (Ii) in the binding groove of MHC undergoes proteolysis then CLIP (MHC class II-associated invariant chain peptide) exchanged with free peptides which are the outcomes of the proteolysis of the phagocytosis process of antigen facilitated through the activity of HLA-DM and then the fully loaded MHC class II translocate to the cell surface, which later on recognized via CD4+ T cells expressing a pertinent TCR.[1,6,29,33,65,73]

Predictions of T-cell epitopes

Prediction of peptide binding to MHC class I: Currently, in epitope discovery, *In Silico* method of T cell prediction is of point of interest for a researcher as of the number of resources required for entire epitope scans. The binding of the peptide to the MH molecule is not only important to become an epitope but it is the most restrictive phase in the pathway from native protein to an immune response. Study describes that on average, out of 200 only one of the possible peptides that originates from native protein binds to MHC class I protein[78] whereas this estimate is somewhat unpredictable in case of MHC class I protein. In both cases, the binding of the peptide position and affinity is an interesting part for the inventing of prediction techniques.[31,35,40,45]

For evaluation, any prediction methods depend on data, but in some techniques like structure-based techniques utilize knowledge of docking and molecular interactions.[19,28,62] were also developed. Hence, it's very much clear that from observation that the more data led to a more accurate prediction technique. The information about the binders is much more substantial than the non-binders and this binder's detection improves the success and improved prediction approach. In the recent, accelerated process can be observed in developing the accurate prediction method through the beneficial scientific collaboration among experimentalists, assay developers and bioinformatics prediction developers[26].

Commonly, the chances of binding to the same peptide increases if the sequence distance among two proteins is smaller, but this is not constantly true of all, exception has been reported and even for two HLAs with highest homology degree, a significant peptides number binding to one protein might not get bind to other. To get the folded or re-foldable MHC protein expression from a bigger number of distinct alleles is one of the most arduous and expensive phases in MHC-peptide binding assays and in the peptide synthesis.

However, in case of the synthesized peptide this task becomes more effortless and economically feasible to test the binding affinity and capability to several available MHC proteins and further thus synergy encourages the quantity of

presentable affinity data for each HLA allele significantly. The obtained data kind can contribute significantly to prediction system development and also for the iteration loops.[85,86] The pan-allele prediction approach is the most successful predictor which is trained to predict the affinity of a given MHC-peptide combination.[24,35,39]

Predictive algorithms used are: binding matrices[2,12,25,42,46-48,51,56,57,63,64], artificial neural networks-ANN[3,4,41,46,82], hidden Markov models-HMM[82], support vector machines[9,11,77,83], structure-based model[28,62], partial least square function[19] and peptide-peptide distance function[26].

Currently, there are several advanced methods that can predict HLA-I binding peptide and can be easily accessible through the internet [Table 1].

Table 1
Prediction servers list of HLA-I binding peptides

ID	Abbreviation	Servers	Prediction algorithm	URLs
1	BIMAS	BIMAS	Matrix	http://www-bimas.cit.nih.gov/molbio/hla_bind/
2	HLA_LI	HLA Ligand	Matrix	http://hlaligand.ouhsc.edu/prediction.htm
3	IEDB_ANN	IEDB (ANN)	ANN	http://tools.immuneepitope.org/analyze/html/mhc_binding.html
4	IEDB_ARB	IEDB (ARB)	Matrix	http://tools.immuneepitope.org/analyze/html/mhc_binding.html
5	IEDB_SMM	IEDB (SMM)	Matrix	http://tools.immuneepitope.org/analyze/html/mhc_binding.html
6	MAPPP_B	MAPPP (Bimas)	Matrix	http://www.mpiib-berlin.mpg.de/MAPPP/binding.html
7	MAPPP_S	MAPPP (SYFPEITHI)	Matrix	http://www.mpiib-berlin.mpg.de/MAPPP/binding.html
8	MHC_BP	MHC Binder Prediction	Matrix	http://www.vaccinedesign.com/
9	MHC_BPS	MHC-BPS	SVM	MHC-BPS http://bidd.cz3.nus.edu.sg/mhc/
10	MHCI_MM	MHC-I (Multiple matrix)	Structure-based model	http://atom.research.microsoft.com/hlabinding/hlabinding.aspx
11	MHCI_SM	MHC-I (Single matrix)	Structure-based model	http://atom.research.microsoft.com/hlabinding/hlabinding.aspx
12	MHCP_I	MHCPred (Interactions)	Partial least square	http://www.jenner.ac.uk/MHCPred/
13	MHCP_AA	MHCPred (Amino Acids)	Partial least square	http://www.jenner.ac.uk/MHCPred/
14	MULTI_ANN	MULTIPRED (ANN)	ANN	http://antigen.i2r.a-star.edu.sg/multipred1/
15	MULTI_HMM	MULTIPRED (HMM)	HMM	http://antigen.i2r.a-star.edu.sg/multipred1/
16	MULTI_SVM	MULTIPRED (SVM)	SVM	http://antigen.i2r.a-star.edu.sg/multipred1/
17	NETM_ANN	NetMHC (ANN)	ANN	http://www.cbs.dtu.dk/services/NetMHC/
18	NETM_WM	NetMHC (Weight Matrix)	Matrix	http://www.cbs.dtu.dk/services/NetMHC/
19	NHP_ANN	nHLAPred (ANNPred)	ANN	http://www.imtech.res.in/raghava/nhlapred/neural.html
20	NHP_CP	nHLAPred (ComPred)	ANN and Matrix	http://www.imtech.res.in/raghava/nhlapred/comp.html
21	PEPDIST	PepDist	distance function	http://www.pepdist.cs.huji.ac.il/
22	PEPC_M	PeptideCheck (Matrix)	Matrix	http://www.peptidecheck.org/
23	PREDEP	Predep	Structure-based model	http://margalit.huji.ac.il/Teppred/mhc-bind/index.html
24	PROPRED	ProPred1	Matrix	http://www.imtech.res.in/raghava/propred1
25	RANKPEP	Rankpep	Matrix	http://bio.dfci.harvard.edu/Tools/rankpep.html
26	SMM	SMM	Matrix	http://zlab.bu.edu/SMM/
27	SVMHC_M	SVMHC (MHCPEP)	SVM	http://www.sbc.su.se/~pierre/svmhc/new.cgi
28	SVMHC_S	SVMHC (SYFPEITHI)	SVM	http://www.sbc.su.se/~pierre/svmhc/new.cgi
29	SVRMHC	SVRMHC	SVM	http://SVRMHC.umn.edu/SVRMHCdb
30	SYFPEITHI	SYFPEITHI	Matrix	http://www.syfpeithi.de/Scripts/MHCServer.dll/EpitopePrediction.htm

Prediction of Th epitopes: The defining of the binding core motif from the peptide binding data is a difficult theoretical and experimental issue in the case of MHC class II binding because of its continuous nine amino acids residue stretch and occurrence in a large stretch of peptide length of 12 to 20 residue ranges. Moreover, the protein of the MHC class II comprised of the two chains (HLA-DQ and HLA-DP), both chains actively play a role in binding of peptide and both are polymorphic residues.

Henceforth, a similar degree of the HTP biochemical assays is not present in comparison to class I MHC protein for the peptide-binding because locating the binding core accurately is a challenging task. To this limitation, the binding data availability is lesser in the case of MHC class I protein which restricts the development of improved prediction approach. Currently, there are few validated accurate prediction approaches available for peptide binding to many HLA-DR alleles and other HLA-DQ and HLA-DP[75].

The prediction of the MHC of non-human mammal's validation is performed for quite limited allele's number because of the availability of the lesser-known data[43]. There are prediction approaches like pan-MHC prediction which have been developed to predict any mammalian MHC class II protein. The possibility to produce class II multimers technically it is complicated in case of producing re-foldable class II recombinantly[27,74].

Computational approaches to predict potential T cell binding-epitopes

Discriminating the immunoprotective epitope from the non-one is the essential step in vaccine development. The interface among T-cells and ligands with accuracy can be modeled because of the boundation of the T- cell epitopes to MHCs in a linear form. Based on the information that epitopes are linked together into binding groove of class I and class II MHCs molecules via interaction among their side chain R group and pocket located on the surface of the MHC, several T cell epitopes- mapping algorithms were established and this has been used to develop efficient and rapid tools to detect T- cell epitopes.[10,13,18,36,52,81]

Currently, the tools have been designed with wide allelic coverage with an accuracy level of 90-95% in range for prediction of the MHC-I binding peptides[23,24,34]. RANKPEP, out of several servers for MHC alleles, predicts binders from the sequence of protein or alignments of sequence to MHC-I and II molecules through PSSMs (Position Specific Scoring Matrices) and also predicts the MHC-I ligands whose C- terminal end is resultant of the proteosomal cleavage[53].

Since 2011, the NetMHCpan prediction method is utilized by the IEDB Analysis Resource database, this produces a quantitative prediction of the affinity of any peptide-MHC class I interaction with coverage of HLA-A and HLA-B for humans, chimpanzee, macaque, gorilla, cow, pig and mouse. This makes up one of the elite databases that admits this form of organisms[24].

The next comprehensive approach is nHLAPred that utilizes sixty-seven MHC alleles prediction of MHC I binding peptides, which is based on ANNs (Artificial Neural Networks) and QM (Quantitative Matrices). This server facilitates twin options of Compred and ANNPred but the most wide-ranging one is 'Compred' that is based on the hybrid technique of ANN and QM[4]. NetMHC server anticipates the binding peptides to enumerate of dissimilar HLA alleles employing ANNs. ANNs have been trained for seventy-eight different human MHC (HLA) alleles standing for all twelve HLA-A and -B Supertypes. Moreover, predictions for forty-one animals including a monkey, cattle, pig and mouse alleles are also accessible[34].

Kernel-based Inter-allele peptide binding prediction SyStem (KISS) identifies whether nine-mer peptides will bind to MHC-I molecule for sixty-four alleles with applying SVM multitask Kernel, to leverage the usable training selective information across the alleles, which meliorates its veracity especially for the alleles with few experienced epitopes. The predictor is trained on databases that contain known epitopes from SYFPEITHI, MHCBN, LANL and IEDB databases. Although there are other servers available to identify MHC-I binding predictors, the servers' names are listed in table 2.

***In silico* Approach Accuracy:** In the past few years, a number of the prediction tools have been designed and their accuracy has been checked thoroughly and their accuracy prediction ranges from 86-93%[67]. In an open competition, some of the designed methods are analyzed[84] and it was found that all latest prediction methods are more improved and accurate compared to the two pioneer methods BIMAS[47] and SYFPEITHI[56], but still very much in practice in the discovery of epitopes. The general recommendation of the tools is listed in table 3 based on accuracy, speed and consistent availability.

Comparative study of the tools is yet a difficult task because of the following aspects:

(i) The short in the documentation of a set of data and prediction approaches,
(ii) The inaccessibility of standard dataset to assess the available approaches,
(iii) The code inaccessibility that implements the techniques,
(iv) The deficient of an amalgamated resultant format, which complicates the procedure of aggregating the results of respective servers to get consensus predictions[16,17].

Therefore, it is indispensable to build standardized data delegacies that enable the examination of distinct prediction approaches on a standardized dataset to liken the technique and originate meta-servers that are going to be blending of multiple predictions tools[68].

Table 2
List of T-cell epitope prediction servers

Server name	Predictive server for		Predictive method	Link
	MHC I	MHC II		
EpiJen	24		Multi-step algorithm	http://www.ddg-pharmfac.net/epijen/EpiJen/EpiJen.htm
SYFPEITHI	42	7	Published motifs	http://www.syfpeithi.de/bin/MHCServer.dll/EpitopePrediction.htm
ANNPRED	30		ANN-regression	http://www.imtech.res.in/raghava/nhlapred/neural.html
BIMAS	41		Published coefficient tables	http://www-bimas.cit.nih.gov/molbio/hla_bind/
ProPred I	47		Quantitative matrix	http://www.imtech.res.in/raghava/propred1/
ProPred		51	Quantitative matrix	http://www.imtech.res.in/raghava/propred/
MHCPred	14	11	Additive method	http://www.ddg-pharmfac.net/mhcpred/MHCPred/
MHC2Pred		42	SVM-based method	http://www.imtech.res.in/raghava/mhc2pred/
NetMHC	57		ANN based method	http://www.cbs.dtu.dk/services/NetMHC/
PREDEP	13		Published coefficient tables	http://margalit.huji.ac.il/Teppred/mhc-bind/index.html
RANKPEP	118	62	PSSM	http://bio.dfci.harvard.edu/RANKPEP/
SVMHC	33	51	SVM-based method	http://abi.inf.uni-tuebingen.de/Services/SVMHC
IEDB binding	77		ANN and SMM method	http://tools.immuneepitope.org/analyze/html/mhc_processing.html
EpiVax	6	8	Epimatrix algorithm	http://www.epivax.com/
MMBPred	46		Quantitative matrix	http://www.imtech.res.in/raghava/mmbpred/
NetCTL	12		ANN-regression	http://www.cbs.dtu.dk/services/NetCTL
nHLAPred	67		Artificial Neural Networks	http://www.imtech.res.in/raghava/nhlapred/
KISS	64		SVM based method	http://cbio.ensmp.fr/kiss/
SVRMHC	36	6	SVM-basedmethod	http://svrmhc.biolead.org/
IMTECH		3	Quantitative matrix	http://www.imtech.res.in/raghava/mhc

Table 3
Recommended MHC binding prediction tools

Name	Prediction	Immune Epitope Database analysis tool	CBS prediction servers
NetMHCpan	MHC class I binding	NetMHCpan	NetMHCpan
NetMHCCons	MHC class I binding		NetMHCCons
SMM	MHC class I binding	smm	
NetMHCIIpan	MHC class II binding	NetMHCIIpan	NetMHCIIpan

Many distinct databases are present currently to detect T and B-cell epitopes. However, further developments are still required.

Bioinformatics tools for predicting potential B cell binding epitopes: B cell receptor or antibodies identify B-cell epitopes in their native structure. The prediction of the B-cell is very much similar to the prediction of T cell. The continuous B cell prediction methods are based on properties of the amino acids like charge, hydrophilicity, hydrophobicity, area of exposed surface and secondary structure whereas discontinuous B-cell prediction of epitopes necessitates the antigen structure in 3D form.[19,68,79] Some comprehensive continuous and discontinuous B-cell epitopes prediction resources are listed in table 4. The Bcepred tool for prediction of B-cell epitopes is based on properties of physiochemical like hydrophilicity, flexibility, polarity and exposed surface on a non-redundant dataset.

The dataset comprises of 1029 B-cell epitopes received from the Bcipep database and an equal number of non-epitopes obtained randomly from the Swiss-Prot database. The accuracy of this prediction approach based on above-mentioned properties varies between 52.92% to 57.53%[69]. The ABCpred server based on neural networks and accuracy estimates of 65.93%[68]. The site of linear B-cell epitopes is predicted through another server BepiPred that is a union of a hidden Markov model and a propensity scale method[50].

Disco Topes tools predict discontinuous B cell epitopes from the 3D structure of proteins. The net scores are figured by aggregating the propensity scores of residues in spatial

proximity and the contact numbers. This server also identifies epitopes in complexes of multiple chains[30].

This tool on with BEpro (formerly known as PEPITO) and SEPPA (Spatial Epitope Prediction of Protein Antigens) needs a 3-D structure as input specifically, in PDB format[70]. Employing SEPPA, every residue in the inquiry protein will be given a score concording to information from its vicinity residues. The higher score symbolizes to the higher probability of the residue to be involved in an epitope[71]. ElliPro, the most demanding tool in this field predicts linear and discontinuous epitopes based on a protein antigen's 3-D structure. This tool links each predicted epitope with a score defined as a PI (Protrusion Index) value.

Homology modeling based 3-D structure prediction is performed based on the user-selected structure template for the input protein sequence, then compute the linear and discontinuous epitopes based on the predicted protein structure. Some other bioinformatics tools to predict continuous and discontinuous B cell epitopes are included in table 4. All of these consolidative tools constitute a possibility for the expansion of a new vaccine in special those that aim at the elicitation of humoral responses.

Conclusion

In this review, we discussed several T-cell epitope prediction tools discovery, its accuracy. The feasibility and sensitivity of the HTP ex vivo method are not up to the mark for elaborated assay types for an extensive number of HLA alleles and pathogens to make a complete proteome scan. Also, the application of the pure assay-based methods detecting the subdominant epitopes is a challenge but can take an essential part in treating certain illnesses whereas the *in silico* models contribute significantly to identify the potent sub-dominant epitopes. In the end, we believe that a conjunctive attempt in a non-hostile arms race of the exploitation of assay procedures and computational modeling will lead to reciprocal welfares.

The process of selecting the potential epitopes is enabled through bioinformatics tools without any hazardous risk. The application of this approach represents an upper hand advantage with speedy outputs and lowers cost over conventional vaccinology techniques. Regardless of numerous epitope prediction techniques availability, building up a systematic appraisal of dissimilar approaches on the standard norm datasets is required. It is expected that in the next few years, the development will be seen in the T-cell prediction method significantly. The development will also be likely seen in inaccurate predictions not only for all HLA-DR alleles, but also for HLA-DQ and -DP alleles and probably for MHC class II proteins from several other mammals as well.

Peptides spotted or synthesized in microarrays can be used to measure a very high number of different peptides in one go. The micro-array peptide chip technology is growing and seems to be appropriate for multi peptide assessment hereafter; especially for MHC class II because of the presence of the binding groove open end of the peptide allowing MHC to link with the peptide despite peptide one end immobilization. The development of computational methods in combination with general obtained biological knowledge will lead to the possibility of highly personalized treatments. Such treatments might be seen in trials against untreatable serious diseases within 5 years period and most certainly within the next 10 years.

Table 4
Comprehensive list of B cell epitope prediction servers.

Server name	Type of Prediction	Link
Bcepred	Prediction of continuous B-cell epitopes	http://www.imtech.res.in/raghava/bcepred/
BepiPred	Prediction of continuous B-cell epitopes	http://www.cbs.dtu.dk/services/BepiPred/
ABCPred	Prediction of continuous B-cell epitopes	http://www.imtech.res.in/raghava/abcpred/
BEST	Prediction of continuous B-cell epitopes	http://biomine.ece.ualberta.ca/BEST/
EPCES	Prediction of discontinuous B-cell epitopes	http://sysbio.unl.edu/services/EPCES/
Discotope	Prediction of discontinuous B-cell epitopes	http://www.cbs.dtu.dk/services/DiscoTope/
BEPro (PEPITO)	Prediction of discontinuous B-cell epitopes	http://pepito.proteomics.ics.uci.edu/
SEPPA	Prediction of discontinuous B-cell epitopes	http://lifecenter.sgst.cn/seppa/index.php
EpiSearch	Prediction of discontinuous B-cell epitopes	http://curie.utmb.edu/episearch.html
MimoPro	Prediction of discontinuous B-cell epitopes	http://informatics.nenu.edu.cn/MimoPro
MIMOX	Prediction of discontinuous B-cell epitopes	http://immunet.cn/mimox/
Pep-3D-Search	Prediction of discontinuous B-cell epitopes	http://kyc.nenu.edu.cn/Pep3DSearch
Epitopia	Prediction of continuous and discontinuous B-cell epitopes	http://epitopia.tau.ac.il/
PepSurf	Prediction of continuous and discontinuous B-cell epitopes	http://pepitope.tau.ac.il
ElliPro	Prediction of continuous and discontinuous B-cell epitopes	http://tools.immuneepitope.org/tools/ElliPro/iedb_input

References

1. Amigorena S. and Savina A., Intracellular mechanisms of antigen cross presentation in dendritic cells, *Curr Opin Immunol.*, **22**,109–117 **(2010)**

2. Bui H.H., Sidney J., Peters B., Sathiamurthy M., Sinichi A., Purton K.A., Mothe B.R., Chisari F.V., Watkins D.I. and Sette A., Automated generation and evaluation of specific MHC binding predictive tools: ARB matrix applications, *Immunogenetics*, **57**, 304–314 **(2005)**

3. Buus S., Lauemoller S.L., Worning P., Kesmir C., Frimurer T., Corbet S., Fomsgaard A., Hilden J., Holm A. and Brunak S., Sensitive quantitative predictions of peptide-MHC binding by a 'Query by Committee' artificial neural network approach, *Tissue Antigens*, **62**, 378–384 **(2003)**

4. Bhasin M. and Raghava G.P., A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes, *J Biosci.*, **32,** 31–42 **(2007)**

5. Cantor J.R., Yoo T.H., Dixit A., Iverson B.L., Forsthuber T.G. and Georgiou G., Therapeutic enzyme deimmunization by combinatorial T-cell epitope removal using neutral drift, Proc Natl Acad Sci., USA, **108**, 1272–1277 **(2011)**

6. Chemali M., Radtke K., Desjardins M. and English L., Alternative pathways for MHC class 1 presentation: a new function for autophagy, *Cell Mol Life Sci*., **68**, 1533–1541 **(2011)**

7. Conus S. and Simon H.U., Cathepsins and their involvement in immune responses, *Swiss Med Wkly*, **140**, w13042 **(2010)**

8. Call M.J., Small molecule modulators of MHC class 2 antigen presentation: mechanistic insights and implications for therapeutic application, *Mol Immunol.*, **48**, 1735–1743 **(2011)**

9. Cui J., Han L.Y., Lin H.H., Tang Z.Q., Jiang L., Cao Z.W. and Chen Y.Z., MHC-BPS: MHC-binder prediction server for identifying peptides of flexible lengths from sequence-derived physicochemical properties, *Immunogenetics*, **58**, 607–613 **(2006)**

10. De Groot A.S. and Moise L., New tools, new approaches and new ideas for vaccine development, *Expert Rev Vaccine*, **6**, 125–7 **(2007)**

11. Donnes P. and Kohlbacher O., SVMHC: a server for prediction of MHC-binding peptides, *Nucleic Acids Res.*, **23**, W194–197 **(2006)**

12. De Luca D.S., Khattab B. and Blasczyk R., A modular concept of HLA for comprehensive peptide binding prediction, *Immunogenetics*, **59**, 25–35 **(2007)**

13. De Lisi C. and Berzofsky J.A., T-cell antigenic sites tend to be amphipathic structures, *PNAS*, **82**, 7048–52 **(1985)**

14. Del Val M., Iborra S., Ramos M. and Lázaro S., Generation of MHC class 1 ligands in the secretory and vesicular pathways, *Cell Mol Life Sci.*, **68**, 1543–1552 **(2011)**

15. Dragovic S.M. et al, Proteasomes, TAP and endoplasmic reticulum-associated aminopeptidase associated with antigen processing control CD4+ Th cell responses by regulating indirect presentation of MHC class 2-restricted cytoplasmic antigens, *J Immunol.*, **186(12)**, 6683–6692 **(2011)**

16. EL-Manzalawy Y. and Honavar V., Recent advances in B-cell epitope prediction methods, *Immun Res*., **6**, 1–9 **(2010)**

17. El-Manzalawy Y. and Honavar V., A framework for developing epitope prediction tools, In Proceedings of the first ACM international conference on bioinformatics and computational biology, ACM, 660–2 **(2010)**

18. Falk K., Rotzschke O., Stevanovic S., Jung G. and Rammensee H.G., Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules, *Nature*, **351,** 290–6 **(1991)**

19. Guan P., Hattotuwagama C.K., Doytchinova I.A. and Flower D.R., MHCPred 2.0: an updated quantitative T-cell epitope prediction server, *Appl Bioinformatics*, **5**, 55–61 **(2006)**

20. Greenbaum J.A. et al, Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools, *J Mol Recognit.*, **20**, 75–82 **(2007)**

21. Huang J. and Honda W., CED: a conformational epitope database, *BMC Immunol.*, **7,** 7 **(2006)**

22. Hansel T.T., Kropshofer H., Singer T., Mitchell J.A. and George A.J., The safety and side effects of monoclonal antibodies, *Nat Rev Drug Discov.*, **9**, 325–338 **(2010)**

23. He Y., Rappuoli R., De Groot A.S. and Chen R., Emerging vaccine informatics, *J Biomed Biotechnol*., **2010**, 1–26 **(2010)**

24. Hoof I. et al, NetMHCpan, a method for MHC class 1 binding prediction beyond humans, *Immunogenetics*, **61**, 1–13 **(2009)**

25. Hakenberg J., Nussbaum A.K., Schild H., Rammensee H.G., Kuttler C., Holzhutter H.G., Kloetzel P.M., Kaufmann S.H. and Mollenkopf H.J., MAPPP: MHC class I antigenic peptide processing prediction, *Appl Bioinformatics*, **2**, 155–158 **(2003)**

26. Hertz T. and Yanover C., PepDist: a new framework for protein-peptide binding prediction based on learning peptide distance functions, *BMC Bioinformatics*, **7**, S3 **(2006)**

27. James E.A., La Fond R., Durinovic-Bello I. and Kwok W., Visualizing antigen specific CD4+ T cells using MHC class 2 tetramers, *J Vis Exp.*, **6(25)**, 1167 **(2009)**

28. Jojic N., Reyes-Gomez M., Heckerman D., Kadie C. and Schueler-Furman O., Learning MHC I–peptide binding, *Bioinformatics*, **22**, e227–235 **(2006)**

29. Jorgensen K.W., Buus S. and Nielsen M., Structural properties of MHC class 2 ligands, implications for the prediction of MHC class 2 epitopes, *PLoS One*, **5**, e15877 **(2010)**

30. Kringelum J.V., Lundegaard C., Lund O. and Nielsen M., Reliable B cell epitope predictions: impacts of method development and improved benchmarking, *Plos Comput Biol.*, **8**, 12 **(2012)**

31. Lundegaard C., Lund O., Kesmir C., Brunak S. and Nielsen M., Modeling the adaptive immune system: predictions and simulations, *Bioinformatics*, **23**, 3265–3275 **(2007)**

32. Landsverk O.J., Bakke O. and Gregers T.F., MHC II and the endocytic pathway: regulation by invariant chain, *Scand J Immunol.*, **70**, 184–193 **(2009)**

33. Larsen M.V., Nielsen M., Weinzierl A. and Lund O., TAP-independent MHC class 1 presentation, *Curr Immunol Rev.*, **2**, 233–245 **(2006)**

34. Lundegaard C., Lund O. and Nielsen M., Accurate approximation method for prediction of class I MHC affinities for peptides of length 8, 10 and 11 using prediction tools trained on 9mers, *Bioinformatics*, **24**, 1397–8 **(2008)**

35. Lundegaard C., Hoof I., Lund O. and Nielsen M., State of the art and challenges in sequence based T-cell epitope prediction, *Immunome Res.*, **6(Suppl 2)**, S3 **(2010)**

36. McMurry J., Sbai H., Gennaro M.L., Carter E.J., Martin W. and De Groot A.S., Analyzing Mycobacterium tuberculosis proteomes for candidate vaccine epitopes, *Tuberculosis*, **85**, 95–105 **(2005)**

37. Mester G., Hoffmann V. and Stevanovi S., Insights into MHC class 1 antigen processing gained from large-scale analysis of class 1 ligands, *Cell Mol Life Sci.*, **68**, 1521–1532 **(2011)**

38. Münz C., Enhancing immunity through autophagy, *Ann Rev Immunol.*, **27**, 423–449 **(2009)**

39. Nielsen M. et al, NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence, *PLoS One*, **2**, e796 **(2007)**

40. Nielsen M., Lund O., Buus S. and Lundegaard C., MHC class 2 epitope predictive algorithms, *Immunology*, **130**, 319–328 **(2010)**

41. Nielsen M., Lundegaard C., Worning P., Lauemoller S.L., Lamberth K., Buus S., Brunak S. and Lund O., Reliable prediction of T-cell epitopes using neural networks with novel sequence representations, *Protein Sci.*, **12**, 1007–1017 **(2003)**

42. Nielsen M., Lundegaard C., Worning P., Hvid C.S., Lamberth K., Buus S., Brunak S. and Lund O., Improved prediction of MHC class I and class II epitopes using a novel Gibbs sampling approach, *Bioinformatics*, **20**, 1388–1397 **(2004)**

43. Nielsen M., Justesen S., Lund O., Lundegaard C. and Buus S., NetMHCIIpan-2.0- improved pan-specific HLA-DR predictions using a novel concurrent alignment and weight optimization training procedure, *Immunome Res.*, **6**, 9 **(2010)**

44. Ottenhoff T.H. et al, First in humans: a new molecularly defined vaccine shows excellent safety and strong induction of long-lived Mycobacterium tuberculosis-specific Th1-cell like responses, *Hum Vaccin.*, **6**, 1007–1015 **(2010)**

45. Peters B. et al, A community resource benchmarking predictions of peptide binding to MHC-I molecules, *PLoS Comput Biol.*, **2**, e65 **(2006)**

46. Peters B. et al, The immune epitope database and analysis resource: from vision to blueprint, *PLoS Biol.*, **3**, e91 **(2005)**

47. Parker K.C., Bednarek M.A. and Coligan J.E., Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains, *J Immunol.*, **152**, 163–175 **(1994)**

48. Peters B., Tong W., Sidney J., Sette A. and Weng Z., Examining the independent binding assumption for binding of peptide epitopes to MHC-I molecules, *Bioinformatics*, **19**, 1765–1772 **(2003)**

49. Pulendran B. and Ahmed R., Immunological mechanisms of vaccination, *Nat Immunol.*, **12**, 509–517 **(2011)**

50. Pontoppidan-Larsen J.E., Lund O. and Nielsen M., Improved method for predicting linear B-cell epitopes, *Immun Res.*, **2**, 2 **(2006)**

51. Peters B. and Sette A., Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method, *BMC Bioinformatics*, **6**, 132 **(2005)**

52. Rötzschke O., Falk K., Stevanovic´ S., Jung G., Walden P. and Rammensee H.G., Exact prediction of a natural T cell epitope, *Eur J Immunol.*, **21**, 2891–4 **(1991)**

53. Reche P.A. and Reinherz E.L., Prediction of peptide-MHC binding using profiles, *Methods Mol Biol.*, **409**, 185–200 **(2007)**

54. Roder G. et al, The outermost N-terminal region of tapasin facilitates folding of major histocompatibility complex class 1, *Eur J Immunol.*, **39**, 2682–2694 **(2009)**

55. Roder G., Geironson L., Rasmussen M., Harndahl M., Buus S. and Paulsson K.M., Tapasin discriminates peptide-HLA-A*02:01 complexes formed with natural ligands, *The Journal of Biological Chemistry*, **286**, 20547-20557 **(2011)**

56. Rammensee H., Bachmann J., Emmerich N.P., Bachor O.A. and Stevanovic S., SYFPEITHI: database for MHC ligands and peptide motifs, *Immunogenetics*, **50**, 213–219 **(1999)**

57. Reche P.A., Glutting J.P. and Reinherz E.L., Prediction of MHC class I binding peptides using profile motifs, *Hum Immunol.*, **63**, 701–709 **(2002)**

58. Sette A. and Peters B., Immune epitope mapping in the post-genomic era: lessons for vaccine development, *Curr Opin Immunol.*, **19**, 106–110 **(2007)**

59. Söllner J. et al, Concept and application of a computational vaccinology workflow, *Immunome Res.*, **6(Suppl 2)**, S7 **(2010)**

60. Star B. et al, The genome sequence of Atlantic cod reveals a unique immune system, *Nature*, **477**, 207–210 **(2011)**

61. Sijts E.J. and Kloetzel P.M., The role of the proteasome in the generation of MHC class 1 ligands and immune responses, *Cell Mol Life Sci.*, **68**, 1491–1502 **(2011)**

62. Schueler-Furman O., Altuvia Y., Sette A. and Margalit H., Structure-based prediction of binding peptides to MHC class I molecules: application to a broad range of MHC alleles, *Protein Sci.*, **9**, 1838–1846 **(2000)**

63. Singh H. and Raghava G.P., ProPred1: prediction of promiscuous MHC Class-I binding sites, *Bioinformatics*, **19**, 1009–1014 **(2003)**

64. Sathiamurthy M., Hickman H.D., Cavett J.W., Zahoor A., Prilliman K., Metcalf S., Fernandez Vina M. and Hildebrand W.H., Population of the HLA ligand database, *Tissue Antigens*, **61**,12–19 **(2003)**

65. Segura E. and Villadangos J.A., A modular and combinatorial view of the antigen cross-presentation pathway in dendritic cells, *Traffic*, **12(12)**, 1677–1685 **(2011)**

66. Schueler-Furman O., Altuvia Y., Sette A. and Margalit H., Structure-based prediction of binding peptides to MHC class I molecules: application to a broad range of MHC alleles, *Protein Sci.*, **9**, 1838–1846 **(2000)**

67. Sylvester-Hvid C. et al, SARS CTL vaccine candidates; HLA supertype-, genome-wide scanning and biochemical validation, *Tis Antigens*, **63**, 395–400 **(2004)**

68. Sun P. et al, Bioinformatics resources and tools for conformational B-cell epitope prediction, *Comput Math Methods Med*., **2013**, 943636 **(2013)**

69. Saha S. and Raghava G.P.S., Prediction of continuous B-cell epitopes in an antigen using recurrent neural network, *Proteins*, **65**, 40–8 **(2006)**

70. Sweredoski M.J. and Baldi P., PEPITO: improved discontinuous B-cell epitope prediction using multiple distance thresholds and half-sphere exposure, *Bioinformatics*, **24**, 1459–60 **(2008)**

71. Sun J. et al, SEPPA: a computational server for spatial epitope prediction of protein antigens, *Nucl Acids Res*., **37**, W612–6 **(2009)**

72. van den Hoorn T., Paul P., Jongsma M L. and Neefjes J., Routes to manipulate MHC class 2 antigen presentation, *Curr Opin Immunol.*, **23**, 88–95 **(2011)**

73. van Endert P., Post-proteasomal and proteasome-independent generation of MHC class 1 ligands, *Cell Mol Life Sci.*, **68**, 1553–1567 **(2011)**

74. Vollers S.S. and Stern L.J., Class 2 major histocompatibility complex tetramer staining: progress, problems and prospects, *Immunology*, **123**, 305–313 **(2008)**

75. Wang P. et al, Peptide binding predictions for HLA DR, DP and DQ molecules, *BMC Bioinform.*, **11**, 568 **(2010)**

76. Watts C., The endosome–lysosome pathway and information generation in the immune system, *Biochim Biophys Acta*, **1824(1)**, 14-21 **(2012)**

77. Wan J., Liu W., Xu Q., Ren Y., Flower D.R. and Li T., SVRMHC prediction server for MHC-binding peptides, *BMC Bioinformatics*, **7,** 463 **(2006)**

78. Yewdell J.W., Confronting complexity: real-world immunodominance in antiviral CD8+ T cell responses, *Immunity*, **25**, 533–543 **(2006)**

79. Yao B., Zheng D., Liang S. and Zhang C., Conformational B-cell epitope prediction on antigen protein structures: a review of current algorithms and comparison with common binding site prediction methods, *PLoS One*, **19(8)**, e62249 **(2013)**

80. Zavasnik-Bergant T. and Turk B., Cysteine proteases: destruction ability versus immunomodulation capacity in immune cells, *Biol Chem.*, **388**, 1141–1149 **(2007)**

81. Zinkernagel R.M. and Doherty P.C., The discovery of MHC restriction, *Immunol Today*, **18**, 14–7 **(1997)**

82. Zhang G.L., Khan A.M., Srinivasan K.N., August J.T. and Brusic V., MULTIPRED: a computational system for prediction of promiscuous HLA binding peptides, *Nucleic Acids Res.*, DOI: 10.1093/nar/gki452, W172–179 **(2005)**

83. Zhang G.L., Bozic I., Kwoh C.K., August J.T. and Brusic V., Prediction of supertype-specific HLA class I binding peptides using support vector machines, *J Immunol Methods*, **320**, 143–154 **(2007)**

84. Zhang G.L. et al, Machine learning competition in immunology-prediction of HLA class 1 molecules, *J Immunol Methods*, **374(1-2)**, 1–4 **(2011)**

85. Zhang H., Lundegaard C. and Nielsen M., Pan-specific MHC class 1 predictors: a benchmark of HLA class 1 pan-specific prediction methods, *Bioinformatics*, **25**, 83–89 **(2009)**

86. Zhang L., Udaka K., Mamitsuka H. and Zhu S., Toward more accurate pan-specific MHC-peptide binding prediction: a review of current methods and tools, *Briefings in Bioinformatics*, **13(3)**, 350–364 **(2012)**.

(Received 09th December 2019, accepted 08th February 2020)