OXFORD

Briefings in Bioinformatics, 18(3), 2017, 467–478

doi: 10.1093/bib/bbw025
Advance Access Publication Date: 25 March 2016
Software Review

Novel *in silico* tools for designing peptide-based subunit vaccines and immunotherapeutics

Sandeep Kumar Dhanda, Salman Sadullah Usmani, Piyush Agrawal, Gandharva Nagpal, Ankur Gautam and Gajendra P.S. Raghava

Corresponding author. Gajendra P.S. Raghava, Bioinformatics Centre, Institute of Microbial Technology, Sector 39A, Chandigarh 160036, India. Tel.: +91-172-2690557; Fax: +91-172-2690632; E-mail: raghava@imtech.res.in

Abstract

The conventional approach for designing vaccine against a particular disease involves stimulation of the immune system using the whole pathogen responsible for the disease. In the post-genomic era, a major challenge is to identify antigenic regions or epitopes that can stimulate different arms of the immune system. In the past two decades, numerous methods and databases have been developed for designing vaccine or immunotherapy against various pathogen-causing diseases. This review describes various computational resources important for designing subunit vaccines or epitope-based immunotherapy. First, different immunological databases are described that maintain epitopes, antigens and vaccine targets. This is followed by *in silico* tools used for predicting linear and conformational B-cell epitopes required for activating humoral immunity. Finally, information on T-cell epitope prediction methods is provided that includes indirect methods like prediction of Major Histocompatibility Complex and transporter-associated protein binders. Different studies for validating the predicted epitopes are also examined critically. This review enlists novel *in silico* resources and tools available for predicting humoral and cell-mediated immune potential. These predicted epitopes could be used for designing epitope-based vaccines or immunotherapy as they may activate the adaptive immunity. Authors emphasized the need to develop tools for the prediction of adjuvants to activate innate and adaptive immune system simultaneously. In addition, attention has also been given to novel prediction methods to predict general therapeutic properties of peptides like half-life, cytotoxicity and immune toxicity.

Key words: immunoinformatics; subunit vaccine; peptide therapeutics; in silico tools; epitope prediction algorithms

Introduction

The traditional approach to design a vaccine is a long and tedious process that involves the attenuation of the pathogen through sub-culturing followed by its administration. The administration of whole pathogen raises several safety issues and toxicity owing to the unwanted biomaterial. It has been shown in the past that a small peptide or antigenic regions can also activate different arms of the immune system [1]. Thus, peptide can also serve as a vaccine candidate, as well as can be used for designing immunotherapeutic. However, identification of peptides (or epitopes or antigenic regions) that can be used for immunization against a pathogen is one of the major challenges. Numerous tools have been developed in the past for identification of antigenic regions in pathogenic proteins or antigens.

Sandeep Kumar Dhanda has submitted his PhD thesis in the area of immunoinformatics. He has developed useful tools for the prediction of cytokine response for any peptide.
Salman Saddullah Usmani is a thesis graduate and is interested in understanding the nature of therapeutic peptides for designing better vaccine candidates.
Piyush Agrawal has joined PhD in 2014 and is currently working on the peptide proteins interactions.
Gandharva Nagpal is a PhD scholar interested in designing algorithms to predict the immunostimulation using oligonucleotides.
Ankur Gautam is a postdoc fellow in the laboratory, and he is working on the peptide-based therapeutics.
Gajendra P.S. Raghava is a chief scientist and head of the bioinformatics centre. He has developed several open-source tools to address the diverse problems in biology.
Submitted: 30 November 2015; **Received (in revised form):** 9 February 2016

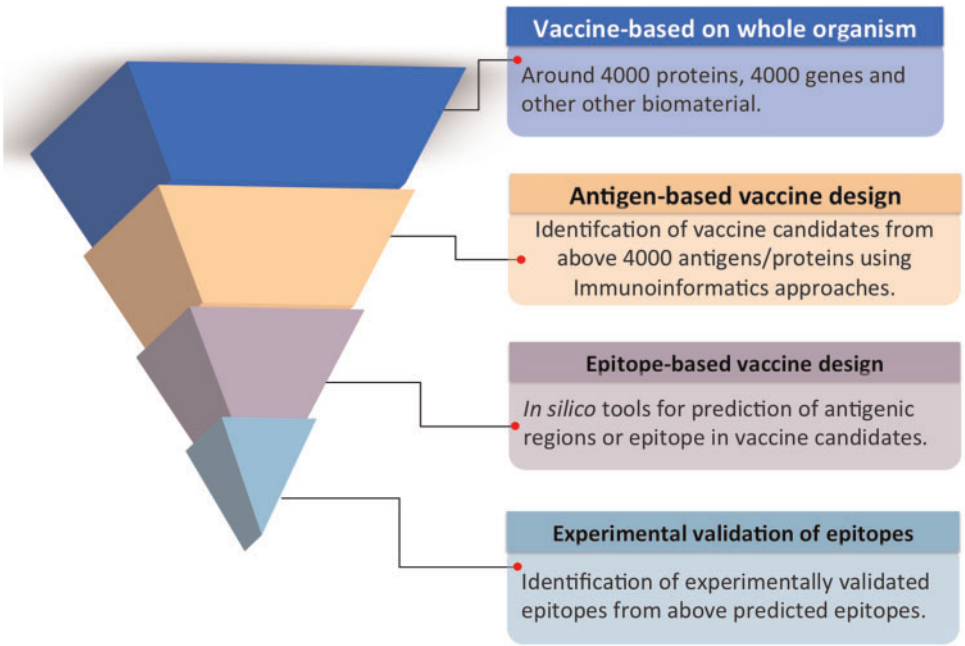

Figure 1. Computational approach for the identification of potential epitopes for designing peptide-based immune therapy or vaccine (e.g. *Mycobacterium tuberculosis*). A colour version of this figure is available at BIB online: https://academic.oup.com/bib.

These computational tools are playing an important role in the field of immunoinformatics by virtue of their assistance in the screening of potential epitopes from a large number of possible epitopes. Figure 1 demonstrates the application of immunoinformatics in virtual screening of vaccine candidates in a pathogen. The numbers, shown in Figure 1, have been estimated from the whole proteome of *Mycobacterium tuberculosis* that contain nearly 4000 proteins with average length ~350 amino acids. It has been shown that all proteins are not vaccine candidates; only around 200 proteins are potential vaccine candidates. These selected vaccine candidates may have some possible peptides, which may be further filtered based on their immunogenic and antigenic potential. Only potential epitopes, which can activate desired arm of the immune system, can be validated using experimental techniques.

This review summarizes the peptide-based novel *in silico* tools as well as tools developed in the past for the prediction of epitopes that can activate different arms of immune system. For better understanding of the field, these tools are categorized based on their function that includes immunological databases and epitope prediction tools. Finally, the review is concluded with an expert opinion where we have figured out the most commonly used tools in each category.

Repositories of immunological resources

In the past few decades, different research groups have created several databases across the world to serve the scientific community. These databases store different types of information related to the human immune system and pathogenic components responsible for generating the immune responses (Table 1). AntigenDB [7] maintains comprehensive information about experimentally validated antigens that includes structural and functional annotation of antigens. This database will be useful for designing protein- or antigen-based vaccine. PRRDb [12] contains pattern recognition receptors and their ligands; it is important for designing vaccine adjuvants. One of the largest repository of immunological epitopes is Immune Epitope Database (IEDB) [5], which maintains B-cell and T-cell epitopes including non-epitopes. In addition, there are databases that maintain interaction-related information, like antigen–antibody interactions (e.g. AgAbDb [13]) and Major Histocompatibility Complex–peptide (MHC–peptide) interactions (e.g. MPID-T2 [15]). In addition to generalized resources, disease-specific resources have also been developed in the past. These resources provide a comprehensive information required to design subunit vaccine against a specific disease, like MalVac [19], MycobacRV [8] and FungalRV [9]. One of the challenges in the field of immunological databases is to maintain and update information; numerous databases have becomes non-functional and outdated in the past two decades. Thus, it is imperative to use the most updated resources in research to increase the reliability of experiments.

In silico tools for humoral immunity

The humoral immune system recognizes antigenic determinants in pathogenic proteins to activate and generate memory B-cells. These antigenic regions are called B-cell epitopes and can be used to develop immunotherapy or vaccine against deadly pathogens [20–29]. Thus, prediction of these B-cell epitopes with high accuracy is one of the important challenges in the field of epitope-based immunotherapy. Broadly, B-cell epitopes are divided in two categories: continuous (or linear) and discontinuous (or conformational) [21]. Earlier methods were developed on a small data set and were based on various physiochemical properties of amino acids like Bcepred [22, 23]. In 2006, for the first time, a machine learning technique was exploited for developing a computational method to predict continuous B-cell epitopes [24]. Since then, a number of methods were developed using various machine learning techniques like ABCpred [25], BepiPred [26], LBtope [27], LBEEP [28] and APCpred [29]. In 2013, LBtope [27] has been developed to predict linear B-cell epitopes using a large number of experimentally validated

Table 1. List of databases or repositories developed for maintaining immunological resources

Name	Description
Databases for immunogenic/antigenic peptides and proteins	
SYFPEITHI [2]	http://www.syfpeithi.de Databases for MHC ligands and peptide motifs.
MHCBN [3]	http://www.imtech.res.in/raghava/mhcbn MHC binding, non-binding and TAP-binding peptides
Bcipep [4]	http://www.imtech.res.in/raghava/bcipep/ A database of B-cell epitopes
IEDB [5]	http://www.iedb.org/ IEDB
EPIMHC [6]	http://bio.dfci.harvard.edu/epimhc/ A database of MHC ligands
AntigenDB [7]	http://www.imtech.res.in/raghava/antigendb/ A database of pathogenic antigens
Database of potential vaccine candidates for pathogenic organism	
MycobacRV [8]	http://mycobacteriarv.igib.res.in/ A database of Mycobacterial vaccine candidates.
FungalRv [9]	http://fungalrv.igib.res.in/ Immunoinformatics portal and adhesion prediction for fungal pathogens
FluKB	http://research4.dfci.harvard.edu/cvc/flukb/ Focusing on data and analytical tools for influenza vaccine discovery.
HPV [10]	http://cvc.dfci.harvard.edu/hpv/Antigen database of Human Papillomavirus (HPV)
Other immunological repositories	
IIID [11]	http://bordensteinlab.vanderbilt.edu/IIID/test_immunity.php Maintain data set of insect innate immunity genes
PRRDB [12]	http://crdd.osdd.net/raghava/prrdb/ A repository of pattern-recognition receptors and their ligands
AgAbDb [13]	http://196.1.114.46:8080/agabdb2/home.jsp Structure of antibody–antigen complexes
IMGT/HLA [14]	http://www.ebi.ac.uk/ipd/imgt/ Sequence database of MHC
MPID-T2 [15]	http://biolinfo.org/mpid-t2/ The MHC–peptide interaction database
InnateDB [16]	http://www.innatedb.com/ Comprehensive information on innate immunity
Vaxjo [17]	http://www.violinet.org/vaxjo/ Information on vaccine adjuvants
LRRSearch [18]	http://www.lrrsearch.com/ Maintain leucine-rich repeats and NOD-like receptors

Table 2. *in silico* tools for predicting linear and conformational B-cell epitopes

Resource	Description
Linear or continuous B-cell epitopes	
ABCpred [24]	http://www.imtech.res.in/raghava/abcpred/ ANN-based prediction of linear B-cell epitopes
Bcepred [25]	http://www.imtech.res.in/raghava/bcepred/ Prediction of B-cell epitope using physicochemical properties.
BepiPred [26]	http://www.cbs.dtu.dk/services/BepiPred/ HMM and propensity-based B-cell epitope prediction
SVMTriP [35]	http://sysbio.unl.edu/SVMTriP/ Linear B-epitope prediction using tripeptide
COBEpro [36]	http://scratch.proteomics.ics.uci.edu Continuous B-cell epitope prediction
EPMLR [37]	http://www.bioinfo.tsinghua.edu.cn/epitope/EPMLR/ A server for prediction of linear B-cell epitope
Pep-3D-Search [38]	http://kyc.nenu.edu.cn/Pep3DSearch/ B-cell epitope prediction based on mimotope
IgPred [39]	http://www.imtech.res.in/raghava/igpred/ Prediction of B-cell epitope for class-specific antibodies
Lbtope [23]	http://crdd.osdd.net/raghava/lbtope/ Improved method for predicting linear B-cell epitope
Conformational or discontinuous B-cell Epitopes	
BEpro [35]	http://pepito.proteomics.ics.uci.edu/ Discontinuous B-cell epitope prediction
CBTOPE [27]	http://crdd.osdd.net/raghava/cbtope/ Conformational B-cell epitopes prediction from sequence
CEP [40]	http://196.1.114.49/cgi-bin/cep.pl Conformational B-cell epitope prediction server
PEASE [41]	http://www.ofranlab.org/PEASE/ Predicting antibody-specific B-cell epitope
DiscoTope [42]	http://www.cbs.dtu.dk/services/DiscoTope-2.0 Predicts discontinuous B-epitope in structure.
Epitopia [29]	http://epitopia.tau.ac.il/ Conformation B-epitope from structure and sequence
SEPPA [43]	http://badd.tongji.edu.cn/seppa/ Discontinuous B-epitope prediction in antigen structure

B-cell epitopes and non-epitopes. Linear B cell exact epitope predictor (LBEEP), published in 2015, used dipeptide deviation from expected mean as an input feature for the identification of linear B-cell epitopes [28]. APCpred is another support vector machines (SVM)-based method, which was developed using the amino acid anchoring pair composition [29]. In addition to generalized methods, recently an organism-specific method B-cell-hepatitis C virus (HCV) has been developed to predict linear B-cell epitope in HCV [30].

Similarly, methods have also been developed to predict conformational B-cell epitopes in an antigen from its sequence like CBtope [31], or from antigen structure like DiscoTope [32], BEpro [33], etc. Epitopia can use either structure or sequence information for predicting conformation epitopes [34]. Table 2 summarizes different *in silico* tools developed for predicting linear and conformational B-cell epitopes.

Various research groups have used the B-cell epitope prediction tools for identifying epitopes in the organism for developing vaccine candidates. For example, Menezes-Souza *et al.* used the BepiPred 1.0 program for predicting the linear B-cell epitopes and found these to be a potential target for diagnosing tegumentary leishmaniasis as well as visceral leishmaniasis [44]. Likewise, Soni *et al.* used Ellipro and DiscoTope for predicting the linear and conformation B-cell epitopes in the virulence proteins LLO, ILO and SLO, coded by *Listeria monocytogenes*. They also predicted the antibody-specific epitopes using the server IgPred [45]. In one of the study done by Yang and Zhang group, ABCpred server and BepiPred server were used to predict and screen potential B-cell

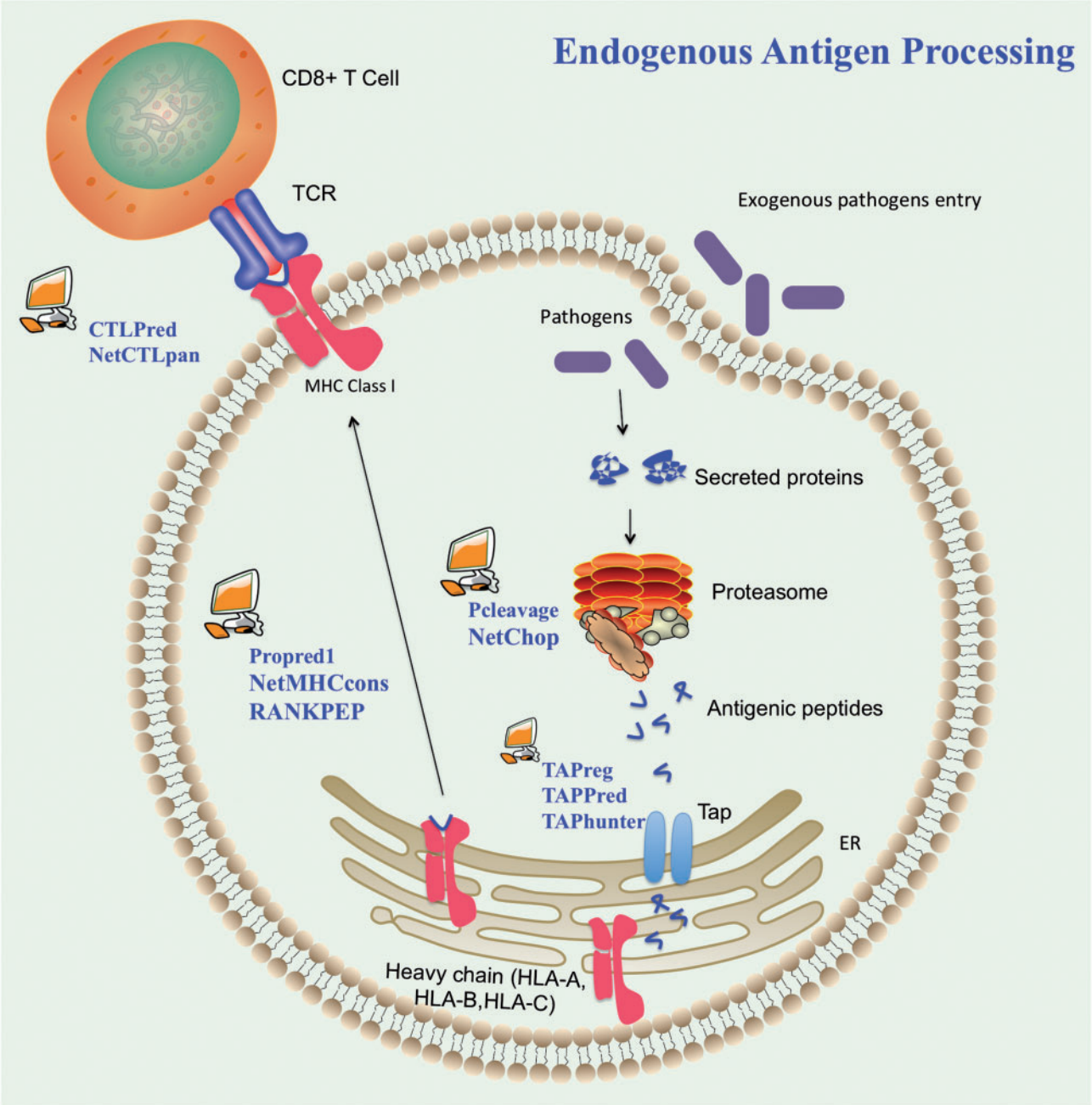

Figure 2. Diagrammatic presentation of endogenous processing of antigens. A colour version of this figure is available at BIB online: https://academic.oup.com/bib.

epitopes on spike glycoprotein of infectious bronchitis virus. They concluded with three potential epitopes on the basis of their antigenicity [46]. CBtope server was used to predict conformational B-cell epitope on collagen VII in the study done by Licarete and colleagues [47]. Experimental studies showed that predicted epitopes were localized within NC1 and NC2 domains and possibly in the hinge region of the antigen [47].

Prediction algorithms for CD8+ T-cell

Over the past three decades, a large number of methods have been developed to predict T-cell epitope. In 1987, based on sequential and structural analysis of T-cell epitopes, direct prediction method was developed [48]. One of the T-cell epitope predicting tools, AMPHI, is based on the hypothesis that T-cell epitopes form stable amphipathic structure [48–50]. Similarly, SOHHA is based on the hypothesis that a helix of 3–5 turn and a strip of hydrophobic residues at one side make T-cell epitopes. Above-mentioned methods are called direct methods, as they directly predict T-cell epitopes instead of MHC binders. In 1990s, structure of MHC–peptide complex was solved by X-ray crystallography, which changed the old concept of structure-based epitope prediction. It is clear now that the peptide bound to MHC has extended conformation [51] and cannot accommodate helix in its groove. This led to development of indirect methods, where MHC binder is predicted instead of T-cell epitope. It includes motif-based methods like EpiMer and Optimer [52]. The

EpiMer method's performance was evaluated with the standard method of synthesizing short, overlapping peptides and testing them for their immunogenic reactions. The method was found to be better than AMPHI or other overlapping methods, as it requires synthesis of fewer peptides compared with other methods [53]. The prediction accuracy of both EpiMer and Optimer was found to be in between 60% and 70%. In the past two decades, instead of the T-cell epitope prediction, the scientific community was more curious toward developing methods to predict MHC binders because it provides more specificity. These methods can be classified in two categories: (i) MHC Class I binders (MHC-I) or CD8 T-cell epitopes for endogenous antigen processing and (ii) MHC Class II binders (MHC-II) or helper T-cell epitopes for exogenous antigen processing.

In the endogenous antigen processing, proteasome recognizes the foreign proteins or antigens and cleaves them into small fragments (peptides). Transporter-associated protein (TAP) then transports these fragments or peptides to the endoplasmic reticulum where the association between these peptides and MHC Class I molecules takes place. These antigenic peptides are presented to cytotoxic T lymphocyte (CTL) by MHC Class I molecules (Figure 2). *In silico* methods have been developed to predict proteasomal cleavage sites in a protein, for example, Pcleavage [54] and NetChop [55]. Pcleavage was developed using SVM technique, and its Matthews correlation coefficient value was found to be 0.54 and 0.43 on *in vitro* and MHC-ligand data, respectively. The SVM classifier in Pcleavage was able to recognize ~82% of cleavage sites and ~45% of non-cleavage sites. This performance was comparable with the NetChop, which was developed using neural network and was able to determine 65% of the cleavage site and 85% of the non-cleavage site correctly. To understand whether proteasomal fragments or peptides will bind to TAP, methods have been developed for predicting TAP binders [56, 57]. Similarly, numerous methods have been developed to predict MHC Class I binders that include ProPred-I, RANKPEP, etc. [58–63]. Propred-I server showed 80% accuracy in predicting MHC allele HLA-A*0201 and H2-Kb binders as well as non-binders. RANKPEP, which was developed using position-specific scoring matrix (PSSM) profile, was able to predict >80% of the epitopes among the top 2% of scoring peptides. NetMHCpan 2.0 server allows users to predict binding peptides for human leukocyte antigen A and B (HLA-A and HLA-B), as well as for chimpanzee, rhesus macaque, gorilla and mouse MHC Class I molecules [64]. NetMHC-4.0 is a sequence alignment-based MHC Class I prediction method, which incorporates the gapped sequence alignment and INDEL events, i.e. insertion and deletion information for predicting peptide-MHC Class I binding affinity [65]. Finally, the MHC-bound peptides are presented to the cell surface where selected peptides are recognized by CD8+ T-cells. To predict whether MHC-bound peptide will be recognized by CD8+ T-cells, direct methods, such as CTLpred, have been developed [66]. Different computational tools developed for predicting CTL epitopes have been listed in Table 3.

Experimental researchers have used the above-mentioned servers in their experiments and found them to be useful in carrying out their research and also validated them. For example, Simoneau *et al.* used Pcleavage server to predict the proteasomal cleavage site in the SHP-1 domain, which is expressed in the nuclei of intestinal epithelial cells and found that both the $Tyr^{208}$and Ser^{591} residues are important for the SHP-1 proteolysis [70]. Similarly, Frahm *et al.* predicted the proteasomal cleavage site using NetChop method, which has been found to be highly predictive of C termini in HIV-1 epitope [71]. RANKPEP server was used by Mingozzi *et al.* for predicting T-cell epitopes and was found to be accurate [72]. Nuzzaci et al. used CTLpred for predicting CTL epitopes within a series of HCV-derived antigenic sequences [73]. They predicted three epitopes with high accuracy and found that the predicted epitopes were able to accurately discriminate between T-cell epitopes and non-epitope MHC binders and further validated them experimentally. In a similar study, Geluk et al. identified four peptides from *Mycobacterium leprae* having HLA binding motifs using ProPred as well as ProPred-I servers [74]. So these studies have shown that bioinformatics servers contribute significantly in developing potential vaccine candidates.

Table 3. List of significant computational methods available for predicting MHC Class I binders and cytotoxic T-cell epitopes

Resource	Description
Prediction MHC Class I binders	
ProPred1 [52]	http://www.imtech.res.in/raghava/propred1/ Server for predicting promiscuous MHC-I binders
NetMHCcons [54]	http://www.cbs.dtu.dk/services/NetMHCcons/ A consensus method for predicting MHC-I binders
nHLAPred [56]	http://www.imtech.res.in/raghava/nhlapred/ Prediction server for promiscuous MHC Class I binders
MMBPred [57]	http://www.imtech.res.in/raghava/mmbpred/ Prediction of high-affinity mutated promiscuous MHC binders.
NetCTLpan [55]	http://www.cbs.dtu.dk/services/NetCTLpan/ CTL epitope and MHC Class I binder prediction predictions
RANKPEP [53]	http://bio.dfci.harvard.edu/RANKPEP/ Profile motifs-based prediction of MHC Class I binding peptides
Important tools for endogenous antigen processing	
CTLPred [58]	http://www.imtech.res.in/raghava/ctlpred/ Direct method for predicting CTL epitopes
NetTepi [67]	http://www.cbs.dtu.dk/services/NetTepi/ Integrated method of T-cell epitopes prediction
FRED [68]	http://www-bs.informatik.uni-tuebingen.de/Software/FRED FRED—a framework for T-cell epitope detection.
TAPPred [50]	http://www.imtech.res.in/raghava/tappred/ Prediction of TAP binding peptides
TAPhunter [69]	http://datam.i2r.a-star.edu.sg/taphunter/ Prediction of TAP ligands using local description of sequence
TAPreg [51]	http://imed.med.ucm.es/Tools/tapreg/ Affinity of TAP-binding ligands
Pcleavage [48]	http://www.imtech.res.in/raghava/pcleavage/ Proteasome/immunoproteasome cleavage sites prediction
NetChop [49]	http://www.cbs.dtu.dk/services/NetChop/ Predicting cleavage site for human proteasome

In silico tools for CD4+ T-cells

Exogenous antigen processing pathway has been shown in Figure 3, CD4+ T-cells or helper T-cell recognize antigenic regions or peptide in the form of MHC-II–peptide complex [75]. There is a fundamental difference between MHC-I and MHC-II binding groves; in case of MHC-I, the groove is closed from both sides, whereas it is open in case of MHC-II. Thus, prediction of MHC-II binders is much more difficult than the prediction of MHC-I binders [76]. During the past decade, efforts have been

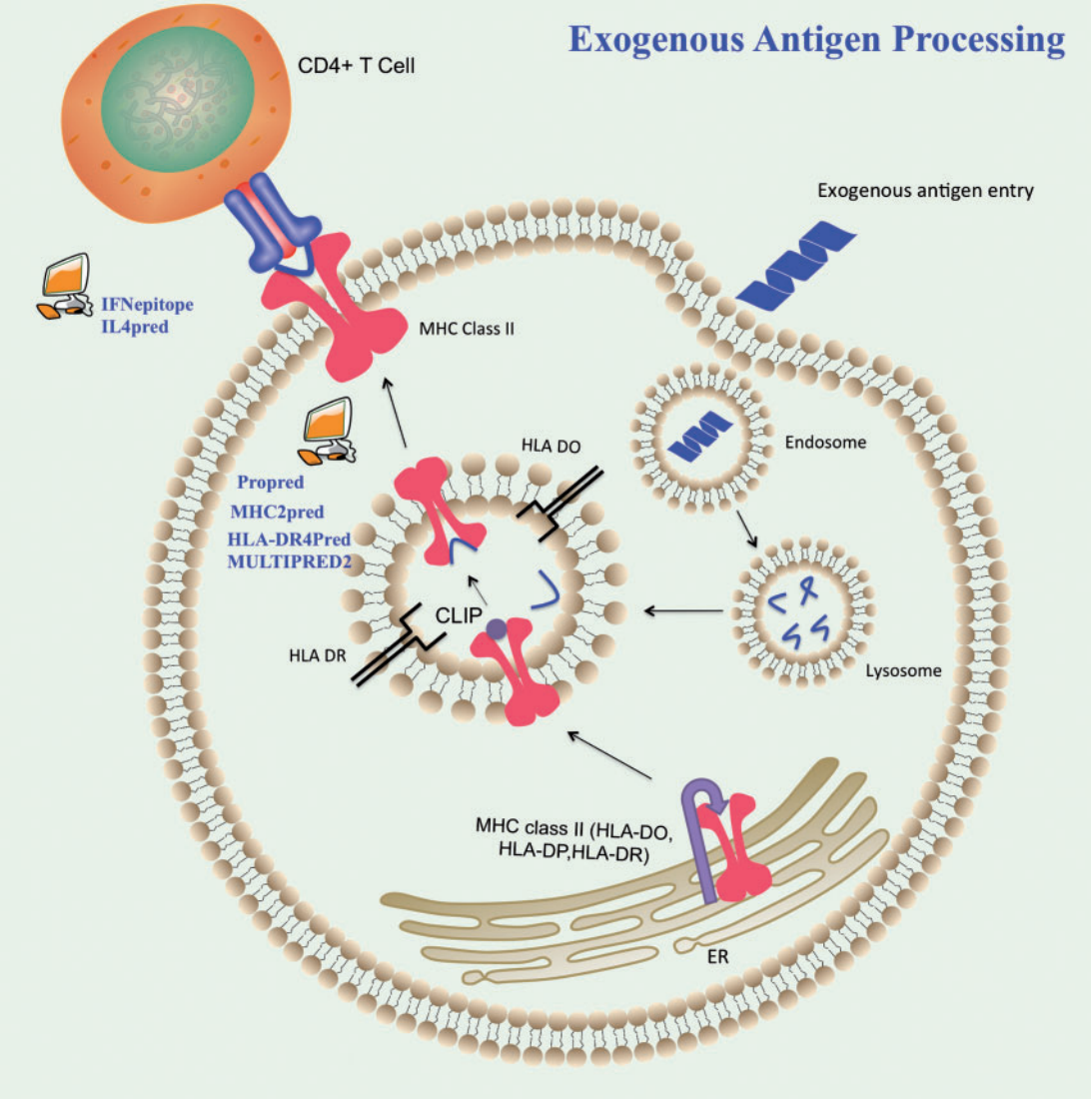

Figure 3. Diagrammatic presentation of exogenous processing of antigens. A colour version of this figure is available at BIB online: https://academic.oup.com/bib.

made in developing data-driven methods to predict MHC-II binders. In 1998, Brusic *et al.* have used an evolutionary algorithm and artificial neural network to predict MHC-II binders [77]. Later, many other algorithms like ant colony [78], hidden Markov models [79], SVM [80], as well as consensus methods integrating the output from two or more methods [81, 82], were proposed for MHC-II binding prediction. In the past, methods have also been developed to predict promiscuous MHC-II binders (which may bind more than one allele) like EpiDOCK, EpiTOP, MHC2Pred, etc. [83–85]. EpiDOCK method is structure based, and models are represented by docking score-based quantitative matrices. EpiTOP is the proteochemometrics-based server for T-cell epitope prediction. Proteochemometrics is a quantitative structure activity relationship (QSAR)-based approach developed by Lapinsh *et al.* [86]. ProPred is another method developed using quantitative matrix [87]. HLA-DR4Pred is a method for predicting MHC-II binding peptides with accuracy ~86% and ~78% in case of SVM and artificial neural network (ANN) models, respectively [88]. Recently, updated NetMHCIIpan 3.1 is a version of a quantitative method that predicts 9 amino acid long binding core of a peptide likely to bind human and mouse MHC Class II molecules [89].

All the above MHC-II binder prediction methods are indirect methods for predicting T-helper epitopes. Recently, a few direct methods have been developed to predict T-helper (e.g. Th1, Th2) epitopes like IFNepitope, IL4Pred [90, 91], etc. Another recent CD4+ T Cell epitope prediction method is PREDIVAC, which made a significant advancement over the previous methods, as its prediction accounts for 95% of the available MHC Class II protein variants [92]. All these methods are summarized in Table 4.

In human system, immune response initiation and regulation are achieved through recognition of HLA-II presented peptides. These peptides (HLA-II T-cell epitopes) can be targeted for the development of the vaccine. Complete viral proteome and individual antigen have to be studied in multiple spanning patterns to get HLA-II binding peptides. Because of the excessive polymorphism of HLA and pathogenic antigen variation, the time and cost of experimental screening methods would be high. *In silico* prediction tools have proven to be a complement to experimental studies in epitope mapping process. *Mycobacterium. tuberculosis* secretes an antigen MPT64 (Rv1980c), which is recognized by Th1 [94], and previous studies have shown that when mycobacterial protein/peptide antigen is associated with HLA-II, it is recognized by human Th1 [95, 96].

Table 4. List of major computational tools commonly used for predicting MHC Class II binders and T-helper epitopes

Resource	Description
MHC Class II binder prediction servers	
Propred [79]	http://www.imtech.res.in/raghava/propred/ Prediction of promiscuous HLA-DR binders
Consensus [74]	http://tools.immuneepitope.org/mhcii/ IEDB tool for predicting MHC Class II binders
EpiDOCK [75]	http://epidock.ddg-pharmfac.net/ Docking-based method for predicting MHC-II binders
EpiTOP [76]	http://www.pharmfac.net/EpiTOP/ An *in silico* tool for MHC Class II binding prediction
MHC2Pred [77]	http://crdd.osdd.net/raghava/mhc2pred/ Prediction of promiscuous binders for MHC Class II alleles.
HLA-DR4Pred [80]	http://www.imtech.res.in/raghava/hladr4pred/ Prediction of HLA-DRB1*0401 binding peptides in an antigen sequence.
MULTIPRED2 [93]	http://cvc.dfci.harvard.edu/multipred2/ A web server for identification of peptides predicted to bind to HLA supertypes and alleles
Direct method for predicting T-helper epitopes	
IFNepitope [81]	http://crdd.osdd.net/raghava/ifnepitope/ Prediction of interferon-gamma inducing T-helper (Th1) epitopes
IL4pred [82]	http://crdd.osdd.net/raghava/il4pred/ Prediction of IL4 inducing T-helper (Th2) epitopes

Table 5. Miscellaneous *in silico* tools and resources for designing immunotherapy

Tool	Description
Prediction of toxic effect of peptides	
Toxinpred [107]	http://crdd.osdd.net/raghava/toxinpred/ Prediction of toxicity of peptides and proteins
Algpred [108]	http://www.imtech.res.in/raghava/algpred/ Prediction of allergenicity for peptide and proteins
AllerHunter [109]	http://tiger.dbs.nus.edu.sg/AllerHunter Allergen prediction program based on SVM
PREAL [110]	http://gmobl.sjtu.edu.cn/PREAL/index.php Prediction of allergenic protein
Hemolytik [111]	http://crdd.osdd.net/raghava/hemolytik/ Database of experimentally verified hemolytic and non-hemolytic peptides
AHTPDB [112]	http://crdd.osdd.net/raghava/ahtpdb/ Database of antihypertensive peptides
Pipeline and meta-servers	
Vacceed [117]	https://github.com/sgoodswe/vacceed/releases *In silico* vaccine candidate for eukaryotic pathogens discovery pipeline
EpiToolKit [119]	http://www.epitoolkit.de/ EpiToolKit offers an array of immunoinformatics tools
NERVE [118]	http://www.bio.unipd.it/molbinfo New enhanced reverse vaccinology environment
Mosaic Vaccine Tool Suite [126, 127]	http://www.hiv.lanl.gov/content/sequence/MOSAIC/ Generation of multiple T-cell epitopes for covering all variants of HIV

In one study, Rv1980 sequence was analyzed for binding to molecules expressed from 51 HLA-DR in ProPred and then predicted regions were experimentally evaluated for HLA-promiscuous Th1 cell reactivity [97]. The study concluded that MPT64 deserves consideration as a candidate vaccine and could have useful diagnostic applications. Similarly, Mustafa used EpiTOP for identifying HLA-promiscous peptide in *Mycobacterium bovis* BCG-vaccinated healthy subjects [98]. MULTIPRED2 has been used in several experimental studies [99–101] as for the identification of an HPV-16 tumor antigen from cancer biopsy specimen [100] and in the detection of influenza A epitope [101]. Consensus prediction approach, a service provided by IEDB, has been used in various experimental studies [102–106].

Discussion

Considerable progress has been made in the past few decades, as to develop novel *in silico* tools, which play an important role in designing immunotherapeutic peptide, directly or indirectly. Identification of suitable vaccine candidates, which can induce significant immune response using high-throughput screening, is a time-consuming and labor-intensive approach. The development of *in silico* tools not only expedites the drug discovery process but also saves the time and labor consumed in channelizing the whole process. One of the challenges in designing peptide-based therapeutics is their toxicity. It could be cytotoxicity or immune toxicity or hemolysis. Peptides are generally considered as good inducers of immune response, but their toxicity limit their approval from clinical trials. A few tools have been developed in the past to predict the toxicity of peptides. In 2014, Gupta *et al.* have developed a tool, Toxinpred, for predicting the toxic effects of peptides [107]. To address the issue of immune toxicity or allergenicity, researchers developed numerous tools that include Algpred [108], AllerHunter [109] and PREAL [110]. In addition, there are tools or resources that allow users to predict other side effects of peptides like Hemolytik [111] for hemolysis and AHTpin [112, 113] for hypertension. *In silico* tools like half-life prediction may help researchers in understanding stability of peptide-based vaccine candidates in body fluid [114]. Gosh et al. applied Algpred for assessment of biosafety in antifungal protein to explore its transgenic applications [115]. Similarly, capsid protein from *Tobacco mosaic* virus has been identified using computational tools and verified in wet laboratory experiments [116].

There is a pressing need to create a pipeline to run multiple algorithms using a single input point. These pipelines may join end to end or in parallel, but the final output should be provided in a single file or folder (Table 5), like Vacceed [117], NERVE [118], EpiToolKit [119]. EpIC is another pipeline that can predict immunogenecity in protein sequences [120]. Vaccine candidates against edwardsiellosis disease and tuberculosis predicted using NERVE pipeline has been validated in experimental studies [121, 122]. The use of Vacceed pipeline has been suggested for rational selection of peptides as candidates [123, 124]. Allard *et al.* described the immune escape mechanism of HIV-1 Rev epitope using EpiTookit [125].

In this review, our main focus is to cover freely available software packages for predicting immunogenic peptides that can stimulate immune system. These methods generally

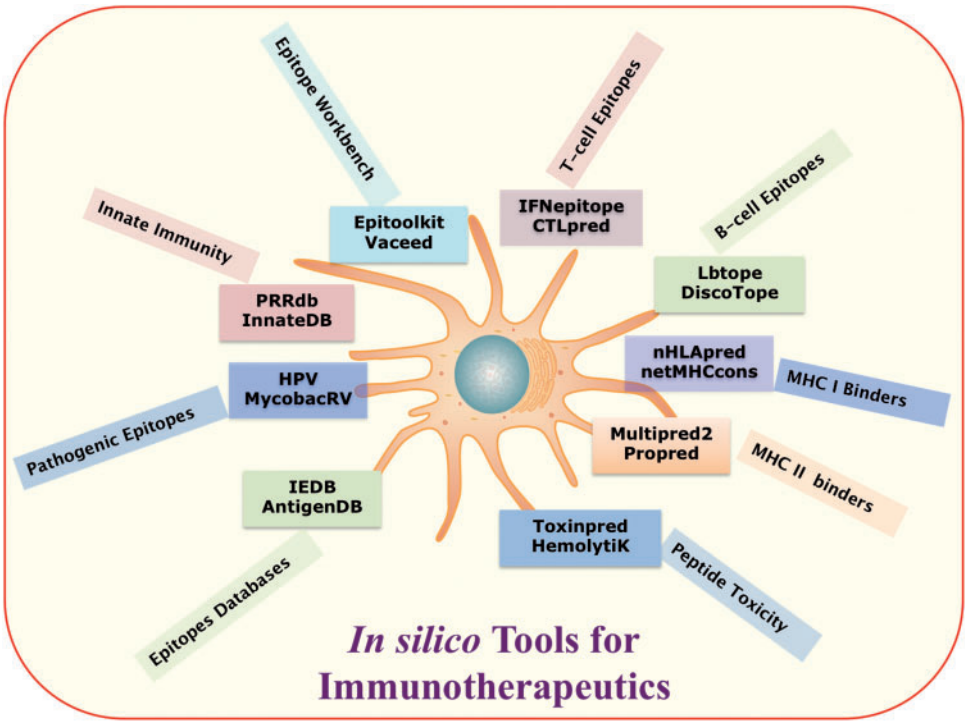

Figure 4. List of *in silico* tools recommended for researchers working in the field of immunotherapy and vaccine. A colour version of this figure is available at BIB online: https://academic.oup.com/bib.

consider static aspect of immune system for designing a vaccine candidate. There is need to understand dynamics behavior of immune system that includes immune-cell populations and their interactions. Over the years, considerable efforts have been made to investigate dynamics of immune system using mathematical modeling, simulation methods and physical concepts [128–132]. Pappalatrado *et al.* developed an *in silico* model based on six ordinary differential equation for modeling antigen-specific CTL activation and their differentiation into memory cell on vaccination with dendritic cells [133]. It is able to predict the dimensions of immune response and contemplate the persistence of antigen-specific memory cells. Arianna Palladin *et al.* showed an integrated *in vivo–in silico* approach for cancer immunoprevention [134]. They developed an antigen-based mathematical model, SimTriplex, that integrates a genetic algorithm. The model is competitive with human-designed time-taking experiments and gives vaccination protocols against cancer. The whole coverage of immune dynamics is beyond the scope of this review.

Conclusion

The past two decades have seen a tremendous progress in the field of immunoinformatics, which reflects in the form of numerous resources developed in these years. These resources cover nearly all aspects required to design peptide-based immunotherapy. One of the challenges is to identify the best resource for each aspect; for example, user may be interested to know the best method for predicting linear B-cell epitope. To serve the scientific community, particularly experimental researchers, we summarize the best resource for predicting each therapeutic property of peptides (Figure 4). The list of currently functional tools has been compiled based on certain criteria that include prediction accuracy, user-friendliness and data set used for training. Here, we suggest LBtope for predicting linear B-cell epitope because it has been trained on largest data set of experimentally proven B-cell epitopes and non-B-cell epitopes. In case of epitope databases, we suggest IEDB, as it is the largest repository for experimentally verified positive as well as negative epitopes. Similarly, AntigenDB is recommended for designing antigen-based vaccines, as it provides detailed information about an antigen.

One of the major disadvantages of epitope-based vaccine over conventional approach is its inability to activate the innate immune system. Ideally, a vaccine candidate should have ability to stimulate both adaptive and innate arms of immune system simultaneously. Unfortunately, limited attempts have been made in the past for predicting vaccine targets that may activate innate immune system as well. The rational design of adjuvants largely banks on the search for ligands of innate immunity receptors. Though some molecules like CpG oligonucleotides, lipopeptides, synthetic chemokines, etc. are in clinical trials for their development as adjuvants, the area remains relatively unexplored owing to the limited number of systematic investigations. There is a need to develop peptide-based tools for adjuvant prediction, as we have tool for DNA-based adjuvants [135]. The field of rational vaccine adjuvant design holds a vast potential with respect to host defense peptides. These have been shown to activate the innate immune system and are increasingly being discovered more in numbers.

In case of B-cell epitope prediction, we could not achieve as good accuracy as we have achieved in the case of T-cell epitope prediction. The poor performance could be attributed to the conformational nature of the B-cell epitope. Therefore, there is a need to understand structural aspects of the epitopes before developing tools for the prediction. Alternatively, B-cell epitope algorithm can be improved through the development of species-specific tools like the prediction of B-cell epitopes that can raise antibody response in different animals like mice, rabbit,

horse, human, etc. In case of prediction of T-cell epitopes, numerous tools from MHC Class I pathways have been developed, but MHC Class II pathway needs to be explored further. Only a handful of direct T-cell epitope prediction algorithms like PREDIVAC, IFNepitope and IL4pred have been developed so far. There is a great scope for further improvement and also development of new tools for predicting helper T-cell epitopes.

Immunoinformatics tools are required to cope with next-generation sequencing (NGS) data. NGS is based on parallel sequencing technology that generates millions of short reads in a short time at a low cost [136]. In the era of the NGS, size of databases consisting of whole genome of different pathogens is growing at an exponential rate. There is a need to develop computational tools or pipelines that can identify the best vaccine candidates for designing vaccine against a pathogen from its whole-genome sequence. As we are moving in the era of personalized medicine, we could not achieve our dreams as long as we do not have tools where comparison of epitopes from diseased and healthy sample can be achieved. We anticipate that with the advances in NGS technologies and *in silico* prediction, algorithms will expedite the drug discovery process and novel efficient immunotherapeutic agents will be discovered.

Key Points

- Tremendous progress has been made in the field of immunology over the years that led to development of highly accurate computational tools.
- Immunological databases are playing an important role in the field of immunoinformatics.
- Sequence-based methods are available for predicting linear or conformational B-cell epitopes in antigenic proteins.
- There is a need to develop tools for predicting vaccine adjuvants and peptide toxicity for designing better subunit vaccines.
- *In silico* tools or pipelines are required for predicting vaccine candidates against a pathogen from its whole genome sequence obtained from NGS.

Acknowledgments

The authors are thankful to the Indian Council of Medical Research (ICMR), Department of Biotechnology (DBT), Department of Science and Technology (DST) and Council of Scientific and Industrial Research (CSIR), Government of India for financial support and fellowships.

Funding

Council of scientific and Industrial Research (CSIR) [project: Open Source Drug Discovery and GENESIS BSC0121]; Department of Biotechnology [project: BTISNET], Government of India.

References

1. Chaplin DD. Overview of the immune response. *J Allergy Clin Immunol* 2010;**125**:S3–23.
2. Rammensee H, Bachmann J, Emmerich NP, *et al*. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* 1999;**50**:213–19.
3. Bhasin M, Singh H. Raghava GPS. MHCBN: a comprehensive database of MHC binding and non-binding peptides. *Bioinformatics* 2003;**19**:665–6.
4. Saha S, Bhasin M, Raghava GPS. Bcipep: a database of B-cell epitopes. *BMC Genomics* 2005;6–79.
5. Vita R, Zarebski L, Greenbaum JA, *et al*. The immune epitope database 2.0. *Nucleic Acids Res* 2010;**38**:D854–62.
6. Reche PA, Zhang H, Glutting J-P, *et al*. EPIMHC: a curated database of MHC-binding peptides for customized computational vaccinology. *Bioinformatics* 2005;**21**:2140–1.
7. Ansari HR, Flower DR, Raghava GPS. AntigenDB: an immunoinformatics database of pathogen antigens. *Nucleic Acids Res* 2010;**38**:D847–53.
8. Chaudhuri R, Kulshreshtha D, Raghunandanan MV, *et al*. Integrative immunoinformatics for Mycobacterial diseases in R platform. *Syst Synth Biol* 2014;**8**:27–39.
9. Chaudhuri R, Ansari FA, Raghunandanan MV, *et al*. FungalRV: adhesin prediction and immunoinformatics portal for human fungal pathogens. *BMC Genomics* 2011;**12**:192.
10. Zhang GL, Riemer AB, Keskin DB, *et al*. HPVdb: a data mining system for knowledge discovery in human papillomavirus with applications in T cell immunology and vaccinology. *Database (Oxford)* 2014, doi: 10.1093/database/bau031.
11. Brucker RM, Funkhouser LJ, Setia S, *et al*. Insect Innate Immunity Database (IIID): an annotation tool for identifying immune genes in insect genomes. *PLoS One* 2012;**7**:e45125.
12. Lata S, Raghava GPS. PRRDB: a comprehensive database of pattern-recognition receptors and their ligands. *BMC Genomics* 2008;**9**:180.
13. Kulkarni-Kale U, Raskar-Renuse S, Natekar-Kalantre G, *et al*. Antigen-Antibody Interaction Database (AgAbDb): a compendium of antigen-antibody interactions. *Methods Mol Biol* 2014;**1184**:149–64.
14. Robinson J, Waller MJ, Parham P, *et al*. IMGT/HLA and IMGT/MHC: sequence databases for the study of the major histocompatibility complex. *Nucleic Acids Res* 2003;**31**:311–14.
15. Khan JM, Cheruku HR, Tong JC, *et al*. MPID-T2: a database for sequence-structure-function analyses of pMHC and TR/pMHC structures. *Bioinformatics* 2011;**27**:1192–3.
16. Breuer K, Foroushani AK, Laird MR, *et al*. InnateDB: systems biology of innate immunity and beyond–recent updates and continuing curation. *Nucleic Acids Res* 2013;**41**:D1228–33.
17. Sayers S, Ulysse G, Xiang Z, *et al*. Vaxjo: a web-based vaccine adjuvant database and its application for analysis of vaccine adjuvants and their uses in vaccine development. *J Biomed Biotechnol* 2012;**2012**:831486.
18. Bej A, Sahoo BR, Swain B, *et al*. LRRsearch: an asynchronous server-based application for the prediction of leucine-rich repeat motifs and an integrative database of NOD-like receptors. *Comput Biol Med* 2014;**53**:164–70.
19. Chaudhuri R, Ahmed S, Ansari FA, *et al*. MalVac: database of malarial vaccine candidates. *Malar J* 2008;**7**:184.
20. Escors D. Tumour immunogenicity, antigen presentation and immunological barriers in cancer immunotherapy. *New J Sci* 2014;**2014**:734515.
21. Zhang J, Zhao X, Sun P, *et al*. Conformational B-cell epitopes prediction from sequences using cost-sensitive ensemble classifiers and spatial clustering. *Biomed Res Int* 2014;**2014**:689219.
22. Saha S, Raghava GPS. BcePred: Prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. In: Proceedings of ICARIS 2004: Third International Conference on Artificial Immune Systems, Catania Sicily, Italy, September 13-16, 2004, pp. 197–204.

23. Saha S, Raghava GPS. Prediction methods for B-cell epitopes. *Methods Mol Biol* 2007;**409**:387–94.
24. El-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. *J Mol Recognit* 2008;**21**:243–55.
25. Saha S, Raghava GPS. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. *Proteins* 2006;**65**:40–8.
26. Larsen JEP, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. *Immunome Res* 2006;**2**:2.
27. Singh H, Ansari HR, Raghava GPS. Improved method for linear B-cell epitope prediction using antigen's primary sequence. *PLoS One* 2013;**8**:e62216.
28. Saravanan V, Gautham N. Harnessing computational biology for exact linear B-cell epitope prediction: a novel amino acid composition-based feature descriptor. *OMICS* 2015;**19**:648–58.
29. Shen W, Cao Y, Cha L, *et al*. Predicting linear B-cell epitopes using amino acid anchoring pair composition. *BioData Min* 2015;**8**:14.
30. Huang W-L, Tsai M-J, Hsu K-T, *et al*. Prediction of linear B-cell epitopes of hepatitis C virus for vaccine development. *BMC Med Genomics* 2015;**8(Suppl 4)**:S3.
31. Ansari HR, Raghava GP. Identification of conformational B-cell Epitopes in an antigen from its primary sequence. *Immunome Res* 2010;**6**:6.
32. Haste Andersen P, Nielsen M, Lund O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Sci* 2006;**15**:2558–67.
33. Sweredoski MJ, Baldi P. PEPITO: improved discontinuous B-cell epitope prediction using multiple distance thresholds and half sphere exposure. *Bioinformatics* 2008;**24**:1459–60.
34. Rubinstein ND, Mayrose I, Martz E, *et al*. Epitopia: a web-server for predicting B-cell epitopes. *BMC Bioinformatics* 2009;**10**:287.
35. Yao B, Zhang L, Liang S, *et al*. SVMTriP: a method to predict antigenic epitopes using support vector machine to integrate tri-peptide similarity and propensity. *PLoS One* 2012;**7**:e45152.
36. Sweredoski MJ, Baldi P. COBEpro: a novel system for predicting continuous B-cell epitopes. *Protein Eng Des Sel* 2009;**22**:113–20.
37. Lian Y, Ge M, Pan X-M. EPMLR: sequence-based linear B-cell epitope prediction method using multiple linear regression. *BMC Bioinformatics* 2014;**15**:414.
38. Huang YX, Bao YL, Guo SY, *et al*. Pep-3D-Search: a method for B-cell epitope prediction based on mimotope analysis. *BMC Bioinformatics* 2008;**9**:538.
39. Gupta S, Ansari HR, Gautam A, *et al*. Identification of B-cell epitopes in an antigen for inducing specific class of antibodies. *Biol Direct* 2013;**8**:27.
40. Kulkarni-Kale U, Bhosle S, Kolaskar AS. CEP: a conformational epitope prediction server. *Nucleic Acids Res* 2005;**33**:W168–71.
41. Sela-Culang I, Ashkenazi S, Peters B, *et al*. PEASE: predicting B-cell epitopes utilizing antibody sequence. *Bioinformatics* 2015;**31**:1313–15.
42. Kringelum JV, Lundegaard C, Lund O, *et al*. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. *PLoS Comput Biol* 2012;**8**:e1002829.
43. Qi T, Qiu T, Zhang Q, *et al*. SEPPA 2.0–more refined server to predict spatial epitope considering species of immune host and subcellular localization of protein antigen. *Nucleic Acids Res* 2014;**42**:W59–63.
44. Menezes-Souza D, Mendes TAdO, Gomes M, *et al*. Epitope mapping of the HSP83.1 protein of Leishmania braziliensis discloses novel targets for immunodiagnosis of tegumentary and visceral clinical forms of leishmaniasis. Clin. *Vaccine Immunol* 2014;**21**:949–59.
45. Soni D, Nath A, Dubey S. Evaluation and use of in silico structure based epitope prediction for listeriolysin O of Listeria monocytogenes. *Indian J Biotechnol* 2015;**14**:160–66.
46. Ding M-D, Wang H-N, Cao H-P, *et al*. Development of a multi-epitope antigen of S protein-based ELISA for antibodies detection against infectious bronchitis virus. *Biosci Biotechnol Biochem* 2015;**79**:1287–95.
47. Licarete E, Ganz S, Recknagel MJ, *et al*. Prevalence of collagen VII-specific autoantibodies in patients with autoimmune and inflammatory diseases. *BMC Immunol* 2012;**13**:16.
48. Stille CJ, Thomas LJ, Reyes VE, *et al*. Hydrophobic strip-of-helix algorithm for selection of T cell-presented peptides. *Mol Immunol* 1987;**24**:1021–7.
49. Spouge JL, Guy HR, Cornette JL, *et al*. Strong conformational propensities enhance T cell antigenicity. *J Immunol* 1987;**138**:204–12.
50. DeLisi C, Berzofsky JA. T-cell antigenic sites tend to be amphipathic structures. *Proc Natl Acad Sci USA* 1985;**82**:7048–52.
51. Stern LJ, Brown JH, Jardetzky TS, *et al*. Crystal structure of the human class II MHC protein HLA-DR1 complexed with an influenza virus peptide. *Nature* 1994;**368**:215–21.
52. Meister GE, Roberts CG, Berzofsky JA, *et al*. Two novel T cell epitope prediction algorithms based on MHC-binding motifs; comparison of predicted and published epitopes from Mycobacterium tuberculosis and HIV protein sequences. *Vaccine* 1995;**13**:581–91.
53. Roberts CG, Meister GE, Jesdale BM, *et al*. Prediction of HIV peptide epitopes by a novel algorithm. *AIDS Res Hum Retroviruses* 1996;**12**:593–610.
54. Bhasin M, Raghava GPS. Pcleavage: an SVM based method for prediction of constitutive proteasome and immunoproteasome cleavage sites in antigenic sequences. *Nucleic Acids Res* 2005;**33**:W202–7.
55. Nielsen M, Lundegaard C, Lund O, *et al*. The role of the proteasome in generating cytotoxic T-cell epitopes: insights obtained from improved predictions of proteasomal cleavage. *Immunogenetics* 2005;**57**:33–41.
56. Bhasin M, Lata S, Raghava GPS. TAPPred prediction of TAP-binding peptides in antigens. *Methods Mol Biol* 2007;**409**:381–6.
57. Diez-Rivero CM, Chenlo B, Zuluaga P, *et al*. Quantitative modeling of peptide binding to TAP using support vector machine. *Proteins* 2010;**78**:63–72.
58. Singh H, Raghava GPS. ProPred1: prediction of promiscuous MHC Class-I binding sites. *Bioinformatics* 2003;**19**:1009–14.
59. Reche PA, Glutting J-P, Reinherz EL. Prediction of MHC class I binding peptides using profile motifs. *Hum Immunol* 2002;**63**:701–9.
60. Karosiene E, Lundegaard C, Lund O, *et al*. NetMHCcons: a consensus method for the major histocompatibility complex class I predictions. *Immunogenetics* 2012;**64**:177–86.
61. Stranzl T, Larsen MV, Lundegaard C, *et al*. NetCTLpan: pan-specific MHC class I pathway epitope predictions. *Immunogenetics* 2010;**62**:357–68.
62. Bhasin M, Raghava GPS. A hybrid approach for predicting promiscuous MHC class I restricted T cell epitopes. *J Biosci* 2007;**32**:31–42.
63. Bhasin M, Raghava GPS. Prediction of promiscuous and high-affinity mutated MHC binders. *Hybrid Hybridomics* 2003;**22**:229–34.

64. Hoof I, Peters B, Sidney J, *et al*. NetMHCpan, a method for MHC class I binding prediction beyond humans. *Immunogenetics* 2009;**61**:1–13.
65. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics* 2015;**32**:511–17.
66. Bhasin M. Raghava GPS. Prediction of CTL epitopes using QM, SVM and ANN techniques. *Vaccine* 2004;**22**:3195–204.
67. Trolle T, Nielsen M. NetTepi: an integrated method for the prediction of T cell epitopes. *Immunogenetics* 2014;**66**:449–56.
68. Feldhahn M, Dönnes P, Thiel P, *et al*. FRED–a framework for T-cell epitope detection. *Bioinformatics* 2009;**25**:2758–9.
69. Lam TH, Mamitsuka H, Ren EC, *et al*. TAP Hunter: a SVM-based system for predicting TAP ligands using local description of amino acid sequence. *Immunome Res* 2010;**6(Suppl 1)**:S6.
70. Simoneau M, Boulanger J, Coulombe G, *et al*. Activation of Cdk2 stimulates proteasome-dependent truncation of tyrosine phosphatase SHP-1 in human proliferating intestinal epithelial cells. *J Biol Chem* 2008;**283**:25544–56.
71. Frahm N, Korber BT, Adams CM, *et al*. Consistent cytotoxic-T-lymphocyte targeting of immunodominant regions in human immunodeficiency virus across multiple ethnicities. *J Virol* 2004;**78**:2187–200.
72. Mingozzi F, Meulenberg JJ, Hui DJ, *et al*. AAV-1-mediated gene transfer to skeletal muscle in humans results in dose-dependent activation of capsid-specific T cells. *Blood* 2009;**114**:2077–86.
73. Nuzzaci M, Piazzolla G, Vitti A, *et al*. Cucumber mosaic virus as a presentation system for a double hepatitis C virus-derived epitope. *Arch Virol* 2007;**152**:915–28.
74. Geluk A, van der Ploeg J, Teles ROB, *et al*. Rational combination of peptides derived from different Mycobacterium leprae proteins improves sensitivity for immunodiagnosis of *M. leprae* infection. *Clin Vaccine Immunol* 2008;**15**:522–33.
75. Robinson J, Halliwell JA, Hayhurst JD, *et al*. The IPD and IMGT/HLA database: allele variant databases. *Nucleic Acids Res* 2014;**43**:D423–31.
76. Nielsen M, Lundegaard C, Worning P, *et al*. Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci* 2003;**12**:1007–17.
77. Brusic V, Rudy G, Honeyman G, *et al*. Prediction of MHC class II-binding peptides using an evolutionary algorithm and artificial neural network. *Bioinformatics* 1998;**14**:121–30.
78. Karpenko O, Shi J, Dai Y. Prediction of MHC class II binders using the ant colony search strategy. *Artif Intell Med* **35**:147–56.
79. Noguchi H, Kato R, Hanai T, *et al*. Hidden Markov model-based prediction of antigenic peptides that interact with MHC class II molecules. *J Biosci Bioeng* 2002;**94**:264–70.
80. Wan J, Liu W, Xu Q, *et al*. SVRMHC prediction server for MHC-binding peptides. *BMC Bioinformatics* 2006;**7**:463.
81. Karpenko O, Huang L, Dai Y. A probabilistic meta-predictor for the MHC class II binding peptides. *Immunogenetics* 2008;**60**:25–36.
82. Wang P, Sidney J, Dow C, *et al*. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Comput Biol* 2008;**4**:e1000048.
83. Atanasova M, Patronov A, Dimitrov I, *et al*. EpiDOCK: a molecular docking-based tool for MHC class II binding prediction. *Protein Eng Des Sel* 2013;**26**:631–4.
84. Dimitrov I, Garnev P, Flower DR, *et al*. EpiTOP–a proteochemometric tool for MHC class II binding prediction. *Bioinformatics* 2010;**26**:2066–8.
85. Lata S, Bhasin M, Raghava GPS. Application of machine learning techniques in predicting MHC binders. *Methods Mol Biol* 2007;**409**:201–15.
86. LADEHOFF A. Clinical investigation of venous pressure in cases of venous insufficiency of the lower extremities. *Ugeskr Laeger* 1957;**119**:91–7.
87. Singh H, Raghava GP. ProPred: prediction of HLA-DR binding sites. *Bioinformatics* 2001; **17**:1236–7.
88. Bhasin M, Raghava GPS. SVM based method for predicting HLA-DRB1*0401 binding peptides in an antigen sequence. *Bioinformatics* 2004; **20**:421–3.
89. Andreatta M, Karosiene E, Rasmussen M, *et al*. Accurate pan-specific prediction of peptide-MHC class II binding affinity with improved binding core identification. *Immunogenetics* 2015;**67**:641–50.
90. Dhanda SK, Vir P, Raghava GPS. Designing of interferon-gamma inducing MHC class-II binders. *Biol Direct* 2013;**8**:30.
91. Dhanda SK, Gupta S, Vir P, *et al*. Prediction of IL4 inducing peptides. *Clin Dev Immunol* 2013;**2013**:263952.
92. Oyarzún P, Ellis JJ, Bodén M, *et al*. PREDIVAC: CD4+ T-cell epitope prediction for vaccine design that covers 95% of HLA class II DR protein diversity. *BMC Bioinformatics* 2013;**14**:52.
93. Eshete F, Bennett JL. Schistosoma mansoni: biochemical characteristics of the antischistosomal effects of Ro 15-5458. *Exp Parasitol* 1990; **71**:69–80.
94. Mustafa AS, Shaban F. Mapping of Th1-cell epitope regions of *Mycobacterium tuberculosis* protein MPT64 (Rv1980c) using synthetic peptides and T-cell lines from *M. tuberculosis*-infected healthy humans. *Med Princ Pract* 2010;**19**:122–8.
95. Mustafa AS. HLA-restricted immune response to mycobacterial antigens: relevance to vaccine design. *Hum Immunol* 2000;**61**:166–71.
96. Al-Attiyah R, Shaban FA, Wiker HG, *et al*. Synthetic peptides identify promiscuous human Th1 cell epitopes of the secreted mycobacterial antigen MPB70. *Infect Immun* 2003;**71**:1953–60.
97. Mustafa AS. In silico binding predictions for identification of HLA-DR-promiscuous regions and epitopes of *Mycobacterium tuberculosis* protein MPT64 (Rv1980c) and their recognition by human Th1 cells. *Med Princ Pract* 2010;**19**:367–72.
98. Mustafa AS. Comparative evaluation of MPT83 (Rv2873) for T helper-1 cell reactivity and identification of HLA-promiscuous peptides in *Mycobacterium bovis* BCG-vaccinated healthy subjects. *Clin Vaccine Immunol* 2011;**18**:1752–9.
99. Kong YM, Brown NK, Flynn JC, *et al*. Efficacy of HLA-DRB1*03:01 and H2E transgenic mouse strains to correlate pathogenic thyroglobulin epitopes for autoimmune thyroiditis. *J Autoimmun* 2011;**37**:63–70.
100. Keskin DB, Reinhold B, Lee SY, *et al*. Direct identification of an HPV-16 tumor antigen from cervical cancer biopsy specimens. *Front Immunol* 2011;**2**:75.
101. Keskin DB, Reinhold BB, Zhang GL, *et al*. Physical detection of influenza A epitopes identifies a stealth subset on human lung epithelium evading natural CD8 immunity. *Proc Natl Acad Sci USA* 2015;**112**:2151–6.
102. Oseroff C, Sidney J, Kotturi MF, *et al*. Molecular determinants of T cell epitope recognition to the common Timothy grass allergen. *J Immunol* 2010;**185**:943–55.
103. Cantor JR, Yoo TH, Dixit A, *et al*. Therapeutic enzyme deimmunization by combinatorial T-cell epitope removal using neutral drift. *Proc Natl Acad Sci USA* 2011;**108**:1272–7.
104. Wang B, Kuroiwa JMY, He L-Z, *et al*. The human cancer antigen mesothelin is more efficiently presented to the mouse

immune system when targeted to the DEC-205/CD205 receptor on dendritic cells. *Ann NY Acad Sci* 2009;**1174**:6–17.
105. Weiskopf D, Yauch LE, Angelo MA, *et al*. Insights into HLA-restricted T cell responses in a novel mouse model of dengue virus infection point toward new implications for vaccine design. *J Immunol* 2011;**187**:4268–79.
106. Günther S, Schlundt A, Sticht J, *et al*. Bidirectional binding of invariant chain peptides to an MHC class II molecule. *Proc Natl Acad Sci USA* 2010;**107**:22219–24.
107. Gupta S, Kapoor P, Chaudhary K, *et al*. In silico approach for predicting toxicity of peptides and proteins. *PLoS One* 2013;**8**:e73957.
108. Saha S, Raghava GPS. AlgPred: prediction of allergenic proteins and mapping of IgE epitopes. *Nucleic Acids Res* 2006;**34**:W202–9.
109. Muh HC, Tong JC, Tammi MT. AllerHunter: a SVM-pairwise system for assessment of allergenicity and allergic cross-reactivity in proteins. *PLoS One* 2009;**4**:e5861.
110. Wang J, Zhang D, Li J. PREAL: prediction of allergenic protein by maximum Relevance Minimum Redundancy (mRMR) feature selection. *BMC Syst Biol* 2013;**7(Suppl 5)**:S9.
111. Gautam A, Chaudhary K, Singh S, *et al*. Hemolytik: a database of experimentally determined hemolytic and non-hemolytic peptides. *Nucleic Acids Res* 2014;**42**:D444–9.
112. Kumar R, Chaudhary K, Sharma M, *et al*. AHTPDB: a comprehensive platform for analysis and presentation of antihypertensive peptides. *Nucleic Acids Res* 2015;**43**:D956–62.
113. Kumar R, Chaudhary K, Singh Chauhan J, *et al*. An in silico platform for predicting, screening and designing of antihypertensive peptides. *Sci Rep* 2015;**5**:12512.
114. Sharma A, Singla D, Rashid M, *et al*. Designing of peptides with desired half-life in intestine-like environment. *BMC Bioinformatics* 2014;**15**:282.
115. Ghosh P, Roy A, Chakraborty J, *et al*. Biological safety assessment of mutant variant of Allium sativum leaf agglutinin (mASAL), a novel antifungal protein for future transgenic application. *J Agric Food Chem* 2013;**61**:11858–64.
116. de Lima Neto D, Stach-Machado D. Epitope mapping of tobacco mosaic virus capsid protein: prediction and experimental data from spot Synthesis. *Mol Biol* 2013;**2**:108.
117. Goodswen SJ, Kennedy PJ, Ellis JT. Vacceed: a high-throughput in silico vaccine candidate discovery pipeline for eukaryotic pathogens based on reverse vaccinology. *Bioinformatics* 2014;**30**:2381–3.
118. Vivona S, Bernante F, Filippini F. NERVE: new enhanced reverse vaccinology environment. *BMC Biotechnol* 2006;**6**:35.
119. Schubert B, Brachvogel H-P, Jürges C, *et al*. EpiToolKit-a web-based workbench for vaccine design. *Bioinformatics* 2015;**31**:2211–13.
120. Marciniuk K, Trost B, Napper S. EpIC: a rational pipeline for epitope immunogenicity characterization. *Bioinformatics* 2015;**31**:2388–90.
121. Zhang M, Wu H, Li X, *et al*. Edwardsiella tarda flagellar protein FlgD: a protective immunogen against edwardsiellosis. *Vaccine* 2012;**30**:3849–56.
122. Monterrubio-López GP, González-Y-Merchand JA, Ribas-Aparicio RM. Identification of novel potential vaccine candidates against tuberculosis based on reverse vaccinology. *Biomed Res Int* 2015;**2015**:483150.
123. Cardona NI, Moncada DM, Gómez-Marin JE. A rational approach to select immunogenic peptides that induce IFN-γ response against Toxoplasma gondii in human leukocytes. *Immunobiology* 2015;**220**:1337–42.
124. Xu Q, Ma X, Wang F, *et al*. Design and construction of a chimeric multi-epitope gene as an epitope-vaccine strategy against ALV-J. *Protein Expr Purif* 2015;**106**:18–24.
125. Allard SD, de Goede AL, De Keersmaecker B, *et al*. Sequence evolution and escape from specific immune pressure of an HIV-1 Rev epitope with extensive sequence similarity to human nucleolar protein 6. *Tissue Antigens* 2012;**79**:174–85.
126. Thurmond J, Yoon H, Kuiken C, *et al*. Web-based design and evaluation of T-cell vaccine candidates. *Bioinformatics* 2008;**24**:1639–40.
127. Fischer W, Perkins S, Theiler J, *et al*. Polyvalent vaccines for optimal coverage of potential T-cell epitopes in global HIV-1 variants. *Nat Med* 2007;**13**:100–6.
128. Germain RN, Meier-Schellersheim M, Nita-Lazar A, *et al*. Systems biology in immunology: a computational modeling perspective. *Annu Rev Immunol* 2011;**29**:527–85.
129. Motta S, Pappalardo F. Mathematical modeling of biological systems. *Brief Bioinform* 2013;**14**:411–22.
130. Pappalardo F, Flower D, Russo G, *et al*. Computational modelling approaches to vaccinology. *Pharmacol Res* 2015; **92**:40–5.
131. Belfiore M, Pennisi M, Aricò G, *et al*. In silico modeling of the immune system: cellular and molecular scale approaches. *Biomed Res Int* 2014;**2014**:371809.
132. Pappalardo F, Chiacchio F, Motta S. Cancer vaccines: state of the art of the computational modeling approaches. *Biomed Res Int* 2013;**2013**:106407.
133. Pappalardo F, Pennisi M, Ricupito A, *et al*. Induction of T-cell memory by a dendritic cell vaccine: a computational model. *Bioinformatics* 2014;**30**:1884–91.
134. Palladini A, Nicoletti G, Pappalardo F, *et al*. In silico modeling and in vivo efficacy of cancer-preventive vaccinations. *Cancer Res* 2010;**70**:7755–63.
135. Nagpal G, Gupta S, Chaudhary K, *et al*. VaccineDA: prediction, design and genome-wide screening of oligodeoxynucleotide-based vaccine adjuvants. *Sci Rep* 2015;**5**:12478.
136. Unamba CIN, Nag A, Sharma RK. Next generation sequencing technologies: the doorway to the unexplored genomics of non-model plants. *Front Plant Sci* 2015;**6**:1074.