Mini Review

Epitope Based Vaccine Designing- A mini review

Aryandra Arya and Sunil K Arora*
Department of Immunopathology, Post Graduate Institute of Medical Education & Research, Chandigarh-160 012, India

Introduction

A vaccine is an antigen derived from pathogen. In its crudest form vaccine contains either attenuated pathogen, or an antigen molecule as in case of subunit vaccine, yet what interacts with immune system are few amino acids in the form of epitopes [1]. The idea to form a vaccine from selective few epitopes has emerged as a more logical approach owing to the fact that the conventional approaches are slow and selection of antigen is more or less random. In the last 5 years a lot of new vaccine candidates have been proposed which are based on B Cell Epitopes (BCE) and T Cell Epitopes (TCE) [2]. This approach of rapid identification of immuno epitopes is centered on computational predictions, which utilize advance algorithms and increasing epitope data base. Epitope prediction is one of the most important corner stones of in-silico vaccine designing, however it depends on antigen identification, and most crucially epitope selection for an effective immune response.

The in-silico vaccine designing is known as computational vaccinology. The advantage of computational vaccinology is, utilization of high through put data analysis methods for rapid antigen identification, molecular docking and simulation models to test immunological responses [3]. This method can analyze multiple antigen candidates and whole proteomes for antigenicity and efficacy in a relatively short time. Epitope search has an additional advantage to further narrow down the antigen screening for very short specific regions, thereby providing a possibility where protein-based manipulation can be used to synergies and select the appropriate immune response type (Figure 1).

Identification of antigen and computational vaccinology

With the advent of high through put proteomics, vaccinologists now have access to multiple tools which can analyze a protein sequence for identification and functioning in an organism along with its interaction and possible evolutionary conservation (4–7). This approach is helpful in providing novel vaccine candidates in a relatively shorter time. One of the main reasons for the failure of a vaccine is its inability to generate specific immune response. By identifying and selecting the most potent antigens, this shortcoming can be avoided. In a recent attempt for a vaccine candidate for visceral leishmaniasis, the authors have applied extensive immune-informatics approach to identify the most potent antigens based on KEGG (Kyoto Encyclopedia of Genes and Genomes) analysis of proteins involved in Protein–Protein Interaction (PPI) networks and metabolic pathways [8,9]. VaxiJen2.0 is an alignment independent antigen prediction server, based on physico chemical properties of protein sequences, utilizing Auto Cross Covariance (ACC) [10,11]. Another online serve Jenner-Predict predicts antigen from functional domains of proteins involved in host pathogen interaction [12]. Similarly VacSol can predict potential therapeutic targets using subtractive reverse vaccinology [13]. AntigenDB is a database of previously validated antigens, the dataset includes data from various other databases like Swiss–Prot, MHCDB, AntiJen, IEDB, and BCPEP [14]. These tools can help in identification of most probable antigens which are able to generate desired immune response.

Epitope mapping and selection

B-cells recognize discontinuous conformational epitopes and continuous linear epitopes [15–17]. BCE arise due to protein folding as these epitopes are recognized by antibodies. The role of hydrophobic vs hydrophilic regions is open to discussion as it is now known that surface regions of protein contain same number of hydrophilic and hydrophobic residues [18]. Amino acid propensity scales applicable for B-cell epitope prediction are generally based on flexibility [19], β-turn propensity [20], and surface accessibility [21]. Prediction of 3D conformational epitope is more difficult than T cell epitopes owing to the
uncertainty in prediction models of protein folding. The existing prediction models of conformational B cell epitopes require antigen 3D structure or homology-based model of the amino acid sequence. So far, no method is available which can predict conformational B cell epitope using antigen primary sequence in the absence of any homology with the known structures. The conformational B cell epitopes tend to be longer than 17 amino-acid (aa) sequence, since shorter aa sequences generally do not form conformational epitopes. ABCPred and BCPRD are B cell epitope prediction web servers which are based on ANN (Artificial Neural Network) and SVM (Support Vector Machine) [22,23]. Bcepred, predict B-cell epitopes on the basis of the physico-chemical properties (hydrophilicity, flexibility/mobility, accessibility, polarity, exposed surface and turns) [24]. The accuracy for ABCPred and Bcepred is 65% and 58% respectively. One of the major constraints of accurate B cell epitope prediction is the small size of dataset used for model training, and use of random peptides as non B cell epitopes. CBtope is another webserver which can predict conformational B cell epitopes using SVM mechanics up to a prediction accuracy of more than 85% and Area Under Curve (AUC) 0.9 [25]. LBTope is a linear B cell epitope prediction model which is based on SVM and uses a larger dataset of validated B-cell epitopes and non-epitopes (12063 epitopes and 20589 non epitopes obtained from IEDB database) [26].

On the other hand, a T-cell recognizes epitopes in context of MHC-I and MHC-II molecules. The MHC-II molecules are presented on the surface of professional Antigen Presenting Cells (APC), these cells process and present antigens to CD4+ T-cells through exogenous pathway. The MHC-I is expressed on the surface of all nucleated cells, it presents endogenous antigens (peptides generally associated with cancer and infecting viruses) to CD8+ T-cells [27]. The peptide length for MHC-I associated epitopes are 9–15 amino acids long, while for MHC-II it can vary up to 22 amino acids. The huge diversity of MHC alleles and T-cell receptors (TCR) presents a challenge in prediction of TSE along with antigen processing during which large protein molecules are chopped into smaller peptides and loaded onto MHC molecule [28]. A large number of prediction software have come up in last 20 years with prediction accuracy levels up to 75%. NetMHCCcons 1.1 has integrated three prediction tools NetMHC, NetMHCPan and PickPocket for more accurate prediction than its competitors [29]. NetChop can predict the proteasomal cleavage sites for MHC-I epitopes, while TAPPred can predict the binding affinity towards TAP (Transporter associated with antigen processing) [30,31]. The TCE prediction software are basically divided in two groups viz. for MHC-I binders and for MHC-II binders. The large number of alleles and sub alleles make it difficult
to optimize the epitope selection, and for that purpose the AlleleFreq.Net software can be used to narrow down the allele requirements based on population selection criteria. Prediction of epitopes for MHC-II binders is comparatively more difficult than for MHC-I owing to two factors: a) the large number of alleles and sub allele frequency in MHC-II loci, and b) due to the physicochemical property of MHC-II grooves, which are open ended and can fit a larger peptide molecule. For MHC-I and MHC-II epitope selection a number of web servers have emerged due to a bigger data sets available for machine learning, majority of servers are based on ANN, SVM, and HMM (Hidden Markov model). SYFPEITHI is an online server, which is based on previous publications of T-cell epitopes and MHC ligands [32]. Similarly, NET MHC-I and NET MHC-II are able to predict TCE for humans as well as mouse [33]. The Propred is an MHC-II prediction server based on quantitative matrix [34].

For a long time, the idea of TCE prediction has been in context of either CD4+ T-cell or CD8+ T-cell activation, but none of the methods can predict the nature of T-cell response. The T-cells, after activation may generate different type of responses e.g. either Th1 or Th2. The signature cytokine for Th1 is IFN-γ and for Th2 is IL-4, IL-5 and IL-10 [35]. Th1 type of responses are important in context of intracellular infections such as viral, parasitic or bacterial infections. The nature of this response depends on multiple different factors like secondary signal for activation, TLR and PAMPs activation, but so far it has not been shown to be dependent on the nature of epitope, although the alteration of single amino acid in epitope has been shown to completely alter the immune response type [36,37]. In this context so far, it has been impossible to predict the epitopes for Th1 or Th2 cell responses. One online server, the IFNepiotpe can predict IFN-γ inducing epitopes, thereby predicting the immunological response type [38]. Similar to IFNepiotpe CTLpred is based on direct methods of prediction where the information or patterns of T cell epitopes, instead of MHC binders, are used for the development of methods [39]. The method is based on Artificial Neural network and support vector machine, which allows the consensus and combined prediction based on these two approaches.

Effective vaccine

In natural course of infection both cell–mediated immunity as well as humoral immunity are required to clear the infection, therefore an integral approach, combining both T-cell and B-cell epitopes, is the appropriate way to design a vaccine. The BCE and TCE can be linked by linker sequences which are amino acids with neutral charge and maximum rotational degree of freedom. Another possibility to combine TCE and BCE is to isolate those regions of antigens which are both B cell epitope positive and T cell epitope positive. For designing an epitope-based vaccine one must address the question of, how many epitopes and which epitopes to choose. There is no straight forward solution to this problem but it is evident from previous research that large size molecules generate better immune response as they are able to mimic the natural antigen and its course of immune activation, shorter peptides have high Expect value (e value) and the probability of finding similar peptides as self-antigen is relatively larger compared to proteins with larger size [40,41]. One possibility is to take it one step further in this regard by integrating multiple epitopes from different antigens and use them in a cocktail/chimeric manner forming a multiepitope construct to synergize the immune response of single epitope into a cumulative effect. One of the advantages of multiple epitope utilization is the increase in the HLA diversity. More research is required to understand the synergy of epitope integration as very few vaccine candidates with chimeric epitopes have been tested as on today [42,43].

Another significant aspect of vaccine development is the generation of memory response, since no vaccine can be effective if memory response is inadequate to respond to re-infection in a heightened manner. The memory response is linked to two different but associated phenomenon, firstly the TCR signal strength and pro-survival signals received, secondly the influence of cytokines and costimulatory signals on the transcriptomal regulation of T cell differentiation. The stability of CD4+ and CD8+ T memory cells generated differs remarkably in their need for cytokine milieu, for example memory CD8+ T-cells proliferate in response to IL-15 but CD4+ T-cells do not [44]. The generation of memory is linked to the amount and duration for which the antigen persists. Longer exposures of high antigen amount are known to induce senescence in CD8+ T-cells but might be required for CD4+ T-cells [45,46]. Memory T-Cells have Stem cell like capabilities and it appears that homeostatic signals drive self-renewal whereas antigenic signals drive effector differentiation. The formation of effector CD8+ T-cell seems to follow a developmental program, which can be triggered by a brief 2–24 hr antigen stimulation and seems to be affected largely by the extrinsic factors such as antigen exposure, its duration and cytokine milieu. Memory T-cells, both CD8+ and CD4+, follow one common pattern i.e. they go through an expansion phase upon antigen stimulation, this expansion phase drives effector memory cell and persists for a relatively shorter duration. At the end of effector expansion phase, a contraction phase kills all the effector T cells, at this stage some effector T cells survive and change into memory subset. The minimum number of memory T-cells surviving has direct correlation with effective protection against re-infection. Memory T cell number below a certain threshold is ineffective against future reinfection [47]. A vaccine candidate should be able to generate above threshold level Memory T-cells in order to be effective. The initial expansion phase of effector T-cell formation is directly related to the size of memory T-cell at the end [47,48]. In an experiment by Murli Kaza et al. 1998 only 5% of initially activated CD8+ T-cells were able to pass into memory pool, in the same set of experiments the memory pool size for CD8+ T cells was 10% of total CD8+ T-cell population [49]. Therefore, it is expected from a vaccine candidate to induce a larger effector T-cell population possible. Epitope based vaccines can enhance this initial effector burst phase by utilizing only the relevant and immunogenic epitopes. Another advantage of using epitope-based vaccine compared to conventional vaccines is in case of chronically ill patients, which have consistent high antigen levels causing T cell exhaustion. In those cases, epitope identification and use
of newer or protective epitopes can provide a better vaccine candidate. Epitope based vaccine development provides a better grip on the amount and specificity of antigen required to activate T cells and can provide better candidates which are effective in generating either Memory CD4+ T cells by using selective MHC–II epitopes or Memory CD8+ T cells by MHC–I epitopes.

Recent developments

Most important implication of epitope–based vaccines would be to address diseases for which conventional methodologies of vaccine development have been unsuccessful till date, like HIV, TB, Leishmaniasis and SARS COV–2. The designing of a vaccine depends upon the detailed knowledge of natural immune profile of the infection and precise identification of the immune–correlates of protection, as any random antigen may not provide perfect protection. The immune response generated by an antigen depends upon its interaction with TCR, PAMPs and most importantly the HLA. Not every antigen carries the epitopes which can be presented by HLA molecules to a TCR to generate a robust immune response. On the same lines, it is necessary that a strong association between HLA epitope and TCR complex will generate the immune response in a specific manner (Th1 or Th2). The strength and duration of such TCR MHC complexes along with secondary signaling molecules forming the supramolecular activation cluster or SMAC affects the future fate of T cell. Naïve T cells subjected to strong TCR stimulations favor Th1 lineage, while weak signals favor Th2 type [50,51]. Remarkably weak TCR stimulation at immuno synapse is enough to generate a robust memory response, while sustained duration of signal along with higher antigen levels generates enhanced proliferation [52]. The role of epitope alone during TCR and HLA association dictating the downstream events would need to be explored, as this will answer the question why certain antigens are able to generate good memory response and others fail.

Epitope based vaccines are also emerging for cancer therapy and many candidates are being tested on both therapeutic and prophylactic basis. Since almost all proteins from a cancerous cell are similar to the normal proteins present in the body, the minute differences limit to only a few amino acid alterations, which can be utilized as epitope–based vaccine candidate. One candidate to use multiple epitopes is Survivin derived multi-epitope vaccine EMD640744 for advanced solid tumors [53]. Another candidate based on anti–FSTL1 [Follistatin-related protein 1] mAbs protein by Kudo and Kawami has shown remarkable immune response against tumors [54].

In the recent outbreak of severe acute respiratory syndrome–corona virus–2 [SARS–COV–2], many vaccine candidates have emerged based on epitope mapping. Hong–Zhi Chen, et al. 2020, have identified BCE and TCE from nucleocapsid protein of SARS–COV–2 virus as probable candidates for vaccine designing [55]. The epitopes are based on ABCpred and BepiPred servers for sequential B-cell epitope selection and discontinuous B-cell epitopes identification by DiscoTope 2.0 [56]. In the same study IEDB server has been utilized for HLA-I and HLA-II binding peptides computation. Tamalika Kar, et al. have proposed a multi-epitope vaccine using spike glycoprotein of SARS–COV–2, docking of vaccine candidate confirmed stable interactions with TLRs and MHC [57]. For all the proposed vaccines for SARS–COV–2 the immune–efficacy has been assessed in-silico by immune–simulation. Many of these candidates are hypothetical and are not validated by in-vitro analysis for immunological responses instead they have utilized molecular docking to give an early assessment. The vaccine designing for SARS–COV–2 has emphasized the importance of immune–informatics and epitope mapping to identify protein regions, which are physiologically vital for the virus and have the ability to generate immune response.

Concluding remarks

With ever increasing database of confirmed epitopes and new algorithms for computation, B and T cell epitope prediction is becoming more reliable for novel vaccine designing. After the emergence of servers like IFNepitope, it might become possible in future to identify cytokine specific epitopes. These cytokine specific epitopes could then be used as therapeutic candidates. The presence of cytokine specific epitopes could be the reason why crude antigens fail to generate a specific immune response (presence of undesirable epitopes), although there is no experimental evidence present and this idea requires more research. Nonetheless, the possibility of cytokine specific epitopes has not only provided us a tool for vaccine designing but has also emphasized on the possible role of epitope alone in T cell activation. This question seems to be centered around the role of T cell receptor and its mechanism of activation. Currently two mechanisms available to address this scenario are, Conformation Change Model and the Kinetic Segregation Model, both models fail to address how a single amino acid substitution in the epitope can results in different downstream pathway activation. Since it is an epitope, which actually holds the specificity while generating an antibody response and T cell activation, the non-epitope domains of protein become not only useless but can also sterically shield the immunogenic epitope domains. Thus, the epitope–based vaccine designing can provide us with new candidates in this regard where large antigen derived vaccines have not been successful.

References


Discover a bigger impact and Visibility of your article with Peertechz Publications

**Highlights**
- Signatory publisher of ORCID
- Signatory publisher of DORA (San Francisco Declaration on Research Assessment)
- Articles archived in world’s renowned service providers such as Portico, CNKI, AGIRIS, TDI Net, Base (Bairfield University Library), CrossRef, Scilt, J-Gate etc.
- Journals indexed in IM/IE, SHERPA/ROMEO, Google Scholar etc.
- OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting)
- Dedicated Editorial Board for every journal
- Accurate and rapid peer-review process
- Increased citations of published articles through promotions
- Reduced timeline for article publication

Submit your articles and experience a new surge in publication services (https://www.peertechz.com/submission).

Copyright: © 2020 Arya A, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.