

Research Article

Computational Identification and Characterization of a Promiscuous T-Cell Epitope on the Extracellular Protein 85B of *Mycobacterium* spp. for Peptide-Based Subunit Vaccine Design

Md. Saddam Hossain,^{1,2} Abul Kalam Azad,²
Parveen Afroz Chowdhury,³ and Mamoru Wakayama¹

¹Department of Biotechnology, Faculty of Life Sciences, Ritsumeikan University, Kusatsu, Shiga 525-8577, Japan

²Department of Genetic Engineering & Biotechnology, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh

³Department of Dermatology, Sylhet Women's Medical College, Sylhet 3114, Bangladesh

Correspondence should be addressed to Abul Kalam Azad; dakazad-btc@sust.edu

Received 25 October 2016; Revised 25 January 2017; Accepted 26 February 2017; Published 16 March 2017

Academic Editor: Marcela I. Henao-Tamayo

Copyright © 2017 Md. Saddam Hossain et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Tuberculosis (TB) is a reemerging disease that remains as a leading cause of morbidity and mortality in humans. To identify and characterize a T-cell epitope suitable for vaccine design, we have utilized the Vaxign server to assess all antigenic proteins of *Mycobacterium* spp. recorded to date in the Protegen database. We found that the extracellular protein 85B displayed the most robust antigenicity among the proteins identified. Computational tools for identifying T-cell epitopes predicted an epitope, 181-QQFIYAGSLSALLDP-195, that could bind to at least 13 major histocompatibility complexes, revealing the promiscuous nature of the epitope. Molecular docking simulation demonstrated that the epitope could bind to the binding groove of MHC II and MHC I molecules by several hydrogen bonds. Molecular docking analysis further revealed that the epitope had a distinctive binding pattern to all DRB1 and A and B series of MHC molecules and presented almost no polymorphism in its binding site. Moreover, using “Allele Frequency Database,” we checked the frequency of HLA alleles in the worldwide population and found a higher frequency of both class I and II HLA alleles in individuals living in TB-endemic regions. Our results indicate that the identified peptide might be a universal candidate to produce an efficient epitope-based vaccine for TB.

1. Introduction

Tuberculosis (TB) is a reemerging disease that remains a leading cause of morbidity and mortality in humans and represents a major public health problem in many countries [1]. Although the Bacille Calmette–Guérin (BCG) vaccine has been available for over 75 years, approximately one-third of the world's population was infected with *Mycobacterium tuberculosis* in 2011, resulting in an estimated 1.4 million people dying from the disease [2]. TB is caused by a group of phylogenetically related bacteria, the *M. tuberculosis* complex (MTBC) [3, 4], which is characterized by low overall genetic diversity and a largely clonal population structure [5–7]. MTBC has been classified into seven phylogenetic lineages, which are found in different geographic regions [6–12].

Although in most patients TB can be treated effectively with multidrug combinations of antibiotics, resistance to anti-TB drugs is increasing. TB treatment is undoubtedly inadequate for preventing disease transmission in highly endemic populations because an active TB patient usually infects approximately 10–15 people per year [2]. The current BCG vaccine has been extensively evaluated and has demonstrated variable protective efficacies ranging from 0% to 85% in different field trials [13]. The protective efficiency of BCG against pulmonary TB in adults, which represents the contagious form of this disease, is discrepant and incomplete, with BCG vaccination campaigns having only a modest impact on the incidence of pulmonary TB. One more issue of concern that compromises BCG's utility is that infants with HIV have an increased risk of developing disseminated

BCG-osis [14]. This stresses the importance of developing TB vaccines that not only have a greater capability to provoke protective immunity against TB but also have a better safety outline than BCG [14, 15].

The first attempts to create better BCG vaccines were made by Tullius et al. [17] who developed a recombinant strain of *Mycobacterium* called rBCG30, which overexpressed antigen 85B (Ag85B). This vaccine resulted in improved protection against TB in guinea pigs and seemed to be immunogenic in humans [18]. A phase I clinical safety trial investigating rBCG30 has concluded but is currently on hold [19]. Since 2009, more than a dozen TB vaccine candidates have entered clinical trials, and many more are in the preclinical pipeline for testing in phase I clinical trials [20]. However, it should be noted that success in these studies and trials may not predict a vaccine's protective coverage on the diverse global stage [21]. Therefore, scientists are tasked with determining the global coverage of novel vaccine candidates through interdisciplinary preclinical approaches that integrate comparative genomics and bioinformatics [22, 23]. Though previous studies have suggested that *M. tuberculosis* has a comparatively stable genome in contrast with other bacteria [24, 25], genomic studies have exposed biologically significant variation among clinical strains [26]. Hebert and colleagues [27], for example, revealed considerable genetic variation in the *pepA* and *PPE18* genes of the clinical *M. tuberculosis* strain. Besides, the diversity of the most polymorphic regions of the human genome, the human leukocyte antigen (HLA) loci is thought to be a response to pathogen escape variants [28–30]. For this reason, to incorporate information of host genetic diversity in vaccine design, improvement, and assaying has been strongly emphasized [31].

Of the diverse vaccine candidates anticipated, subunit vaccines have received significant attention, particularly those composed of antigenic proteins Ag85B, ESAT-6, Ag85A, *pepA*, and *PPE18* [32–34]. As these candidates move forward into clinical practices, it will be critically important to appraise their defensive potential as global vaccines via bioinformatics approaches to increase the understanding on the comparative genomics of the pathogen population and the immunomics of the host population. With the dawn of computers and informatics, novel approaches have been devised that focus on the development and applications of computational approaches to advance vaccine research and improve immunization programs. Epitope-based vaccines have attracted considerable concentration lately as a potential, inexpensive, and easy means of vaccine development for treating infectious disease [35, 36]. It was shown experimentally that a synthetic peptide can mimic the structure of epitopes and induce an immune response against the native proteins of bacteria causing Salmonellosis and *Chlamydia* infection [37, 38].

Surface and secreted proteins of any pathogen are mostly antigenic [39, 40] and are responsible for pathogenicity. The Ag85B is a secretory antigenic target and is highly conserved in MTBC [41–43]. The present study reports the identification and characterization of T-cell epitopes on Ag85B of *M. tuberculosis* and their interactions with the MHC molecules using different bioinformatics tools as well as

molecular docking approaches. Furthermore, using the Allele Frequency Database, we investigated the frequency of the HLA alleles in individuals living in TB-endemic regions.

2. Materials and Methods

2.1. Identification of T-Cell Epitopes. Protegen (<http://www.violinet.org/protegen/>) [44], an antigen database and analysis system, was used to assess the antigenicity of Ag85B. For prediction of protective antigens and subunit vaccines, the VaxiJen (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>) [45] system was utilized. Lastly, Vaxign (<http://www.violinet.org/vaxign/>) [36] was used for vaccine target prediction and analysis. The threshold value in VaxiJen was kept at 0.5. After confirmation of the highest antigenicity of Ag85B, all amino acid sequences of Ag85B of the pathogenic *Mycobacterium* spp. were retrieved from the protein database. The nonidentical sequences were subjected to multiple sequence alignment analysis with the T-Coffee server [46]. For identification of T-cell epitopes from Ag85B with strong MHC II binding affinity, artificial neural network based computer program NetMHCII 2.2 (<http://www.cbs.dtu.dk/services/NetMHCII/>) [47], motif ma trices based SYFPEITHI (<http://www.syfpeithi.de/>) [48], and quantitative matrices based ProPred (<http://www.imtech.res.in/raghava/propred/>) [49] servers were used. To predict MHC I binding ability, the epitope sequences were subjected to the NetCTL 1.2 Server (<http://www.cbs.dtu.dk/services/NetCTL/>) [50]. Binding ability of the epitopes was analyzed against all 12 available MHC I HLA supertypes. Threshold values for “weight on C terminal cleavage,” “weight on TAP (transport associated antigen processing) transport efficiency,” and “epitope identification” were kept as 0.15, 0.05, and 0.75, respectively. Predicted MHC alleles that represented strong binding affinity with the epitope were studied in the “Allele Frequency Worldwide Populations database” [51]. Additionally, 22 countries with high TB burden as identified by the World Health Organization (WHO) [2] were also analyzed for MHC alleles. Physicochemical properties of the epitope were analyzed with ProtParam computer program [52].

2.2. Retrieval of 3D Structure. The 3D structure of epitopes was built using PyMol molecule builder, as described previously [53]. There were no experimental structures available in the Protein Data Bank (PDB) (<http://www.rcsb.org/pdb/home/home.do>) [54] for the HLA-DRB1*0701, HLA-DRB1*0901, HLA-B39, and HLA-B58 molecules. The 3D structures of these four proteins were obtained using homology modeling with Molecular Operating Environment (MOE), as described previously [16]. The homology models of HLA-DRB1*0701 [Uniprot:P13761], HLA-DRB1*0901 [Uniprot:Q0R326], HLA-B39 [Uniprot:P30475], and HLA-B58 [Uniprot:P10319] were constructed based on the experimental structures of HLA-DR1 beta chain [PDB:1AQD-b], DRB1-1 beta chain [PDB:3PGD-b], HLA-B*1402 alpha chain [PDB:3BXN-a], and HLA-B*5701 alpha chain [PDB:3VRI-a], respectively. The amino acid sequence homology of HLA-DRB1*0701, HLA-DRB1*0901, HLA-B39, and HLA-B58

TABLE 1: Ramachandran plot statistics computed with PROCHECK program.

Plot statistics	DRB1*0701	DRB1*0901	HLA-B39	HLA-B58
% residues in favourable regions	89.5	89.2	92.8	92.0
% residues in additional residue regions	9.9	10.8	6.6	7.7
% residues in generously regions	0.0	0.0	0.6	0.3
% residues in disallowed regions	0.6	0.0	0.0	0.0
% of nonproline and nonglycine residues	100.00	100.00	100.00	100.00

with HLA-DR1, DRB1-1, HLA-B*1402, and HLA-B*5701 is 89.31%, 87.90%, 97%, and 98%, correspondingly. To evaluate how much the structural models deviate from the template, the root-mean-square deviation (RMSD) value between the template and the homology model superposed was determined by the Internal Coordinate Mechanics method based computer program Molsoft Internal Coordinate Mechanics- (ICM-) pro 3.5 [55]. The stereochemical quality of the homology models was assessed by the PROCHECK program through the PDBsum server [56] (Table 1). The PDB ID for experimental structures of HLA-DRA [Uniprot:P01903], HLA-DRB1*0101 [Uniprot:P04229], HLA-DRB1*0404 [Uniprot:P13760], HLA-DPBI*0101 [Uniprot:P04440], and HLA-DQBI*0301 [Uniprot:P01920] were 1A6A-a, 1SJE-b, 1D5 M-b, 3LQZ-b, and 1JK8-b, respectively. MHC I HLA supertypes A2, A24, B8, B27, and B44 were retrieved from PDB. PDB ID for experimental structures of HLA supertype A2 [Uniprot:P01892], A24 [Uniprot:P05534], B8 [Uniprot:P30460], B27 [Uniprot:P03989], and B44 [Uniprot:P30481] were 1HHI-a, 3I6L-a, 1AGB-a, 10GT-a, and 1SYS-a, respectively.

2.3. Peptide Docking. The 3D structure of the Ag85B T-cell epitope built by the PyMol molecule builder was subjected to molecular docking simulation with MHC I and II molecules by Molsoft ICM-pro 3.5. The Ag85B T-cell epitope could bind to MHC II (HLA-DRB1*0701, HLA-DRB1*0101, HLA-DRB1*0404, HLA-DRB1*0901, HLA-DPBI*0101, and HLA-DQBI*0301) and MHC I (HLA-A2, HLA-A24, HLA-B27, HLA-B39, and HLA-B44) molecules. Molecular docking of the epitope was performed against the aforementioned 6 MHC II and 5 MHC I molecules. Ionizable group of cationic residues in neutral or mildly acidic compartment were transformed into the protonated states and ICM default partial atomic charges were set. ICM pocket finder macro was used for pocket selection, and after receptor positioning, the ICM method based molecular docking was done, as described previously [57]. The best docking orientation was selected based on binding free energy and hydrogen bond distance.

To validate the identification and characterization of the epitope with the computer-aided programs, the in silico approaches we used in this study were applied to identify the epitope on a positive control protein, an influenza viral antigen nucleoprotein (IvAgNP). Although IvAgNP is a viral antigenic protein, we used it as a positive control to cross check the in silico approaches used here and determine whether they can identify the epitopes previously reported

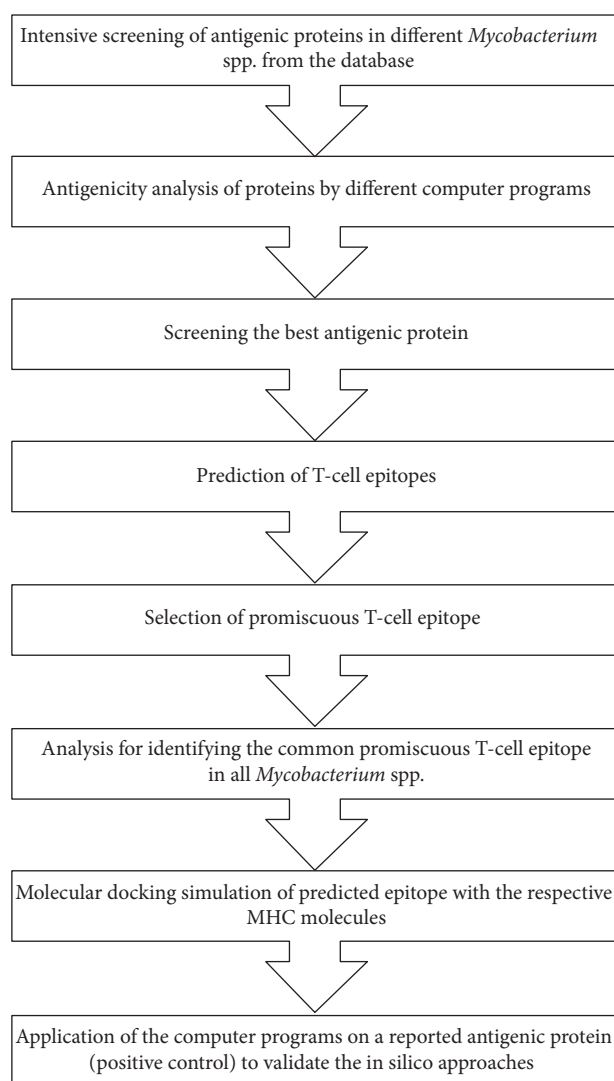


FIGURE 1: In silico approaches for T-cell epitope identification.

[58]. The epitopes of IvAgNP were used as peptide vaccines [58, 59], and the epitope NP 380-ELRSRYWAIRTRSG-393 had been experimentally characterized [60, 61] and was known as an inducer of cellular response. A summary of in silico approaches used in this study is shown in Figure 1.

TABLE 2: Comparison of antigenicity of antigenic proteins of *Mycobacterium* spp. analyzed with Vaxign computer program.

Protective antigens	Accession number	Localization probability	Vaxign analysis	
			Adhesin probability	Number of similar human proteins
Ag85B	NP_216402.1	Periplasmic (Prob. = 0.984)	0.618	0
Ag85A	AAO62004.1	Periplasmic (Prob. = 0.984)	0.558	0
ESAT-6	YP_178023.1	Unknown (Prob. = 0.2)	0.659	0
pepA	NP_214639.1	Periplasmic (Prob. = 0.976)	0.456	4
PPE18	YP_177795.1	Cytoplasmic membrane (Prob. = 0.973)	0.650	0
IvAgNP*	P03466	Periplasmic (Prob. = 0.983)	0.076	0

* Positive control.

3. Results

3.1. Identification and Characterization of T-Cell Epitopes. Proteogen analysis identified 24 experimentally verified protective antigens that had been reported to date in *M. tuberculosis*. Among the 24 antigenic proteins, 11 are extracellular, five are cytoplasmic, and one is a cell wall component. The localization of the remaining seven proteins is unknown. Five proteins, namely, Ag85B, Ag85A, ESAT-6, pepA, and PPE18, have been extensively studied and used as subunit vaccines in animal models [32–34]. Antigenicity analysis of these proteins with Vaxign (Table 2) showed that Ag85B, ESAT-6, and PPE18 had comparatively higher adhesion capability. However, it was concluded that it would be unwise to design a vaccine based on PPE18 due to polymorphism among clinical strains of *M. tuberculosis* [27]. The antigen probability was further analyzed with the VaxiJen server. The antigen probability for Ag85B was 0.5926, whereas that of ESAT-6 was 0.5657, indicating that Ag85B had higher antigenicity than ESAT-6. Therefore, Ag85B was considered for further analysis. Notably, the antigen probability of the positive control, IvAgNP, was only 0.5112. Ag85B was then subjected to the NetMHCII 2.2 web server and analyzed against 26 MHC II molecules available in the database. Ag85B showed strong binding capacity with 22 MHC II molecules (Table 3). The epitope with the strongest affinity was 181-QQFIYAGSLSALLDP-195. This epitope was found against HLA-DRB1*0101 with an affinity of 4.5 nM. Notably, an affinity below 50 nM in the NetMHCII 2.2 server is considered as strong binding. Upon further analysis, this epitope showed high affinity to a large number of MHC alleles (Table 4), revealing its promiscuous nature [57]. Four additional epitopes showed high affinity to HLA-DRB1*0701, DPB1*0201, DPB1*0401, and DQB1*0301 molecules compared to the epitope 181-QQFIYAGSLSALLDP-195 (Table 3). After being subjected to analysis through the Net MHCII 2.2 server, these epitopes showed an ability to bind only with the same MHC II molecule as shown in Table 3 with affinities below 50 nM, thereby showing a lack of promiscuous binding. The same approaches were applied to find the strongest promiscuous epitope on IvAgNP. The control IvAgNP also bound against 22 MHC II molecules (see Additional File 1, Table S1 in Supplementary Material available online at <https://doi.org/10.1155/2017/4826030>). The most promiscuous epitope, 301-IDPFRLQNSQVYSL-315, was found against HLA-DRB1*0101 with an affinity of

3.4 nM (Additional File 2, Table S2). A previously identified epitope, 380-ELRSRYWAIRTRSGG-393 was determined to be not promiscuous [58], though it did bind against HLA-DRB1*0101 with a significant affinity of 9 nM (data not shown). The binding capability of Ag85B with MHC II molecules was further analyzed with ProPred and SYFPEI-THI web servers. Each of these approaches predicted several epitopes that showed an ability to bind to many MHC II molecules. The 9-mer of the strongest predicted epitope mentioned above was identified as one of the top five epitopes in this analysis (Table 5). Data presented here were only for HLA-DRB1*0701.

To predict MHC I binding ability, epitope sequences were subjected to the NetCTL 1.2 Server. The binding ability of the epitopes was analyzed against all 12 available MHC I HLA supertypes. Analysis showed that the 9-mer of the predicted epitope, 181-QQFIYAGSLSALLDP-195, could bind significantly with five MHC I HLA supertypes, namely, A2, A24, B27, B44, and B39. The 9-mer epitope 181-QQFIYAGSL-189 could bind to B27 and B44 MHC I HLA supertypes. However, the 9-mer epitope 184-IYAGSLSAL-192 could bind to both HLA-A24 and B39 supertypes. Likewise, HLA-A2 supertypes could bind to the epitope 183-FIYAGSLSA-191. Analysis with the NetCTL 1.2 server revealed that the 9-mer of the positive control epitope, 380-ELRSRYWAIRTRSG-393 of IvAgNP, could bind to only two MHC I HLA supertypes, namely, B8 and B27. However, the strongest epitope, 301-IDPFRLQNSQVYSL-315, could bind significantly against five MHC I allele supertypes, namely, A3, A2, B27, B39, and B58.

To predict the universal effectiveness of the epitope, we retrieved all 150 Ag85B sequences (Additional File 3, Table S3) of pathogenic *Mycobacterium* spp. from the protein database and found 15 nonidentical Ag85B sequences. These nonidentical sequences were then subjected to the T-Coffee server for multiple sequence alignment to determine the identity and the universality of the predicted epitopes among pathogenic *Mycobacterium* spp. (Figure 2). This analysis revealed that the percentage of identity of the predicted strongest epitope was 100, 93, and ~87 in 9, 3, and 3 *Mycobacterium* spp., respectively. These results revealed that the predicted strongest epitope had at least 90% of amino acid identity in all *Mycobacterium* spp. indicating the universality of the epitope. Likewise, we retrieved 1300 IvAgNP sequences (data not shown) of pathogenic influenza virus and 73 nonidentical sequences were aligned to find the conservation

TABLE 3: Binding of epitopes of Ag85B from *M. tuberculosis* against MHC II molecules^a.

MHC II molecule	Start position	Strongest epitope		Affinity (nM)	Number of strong binding epitopes	Number of weak binding epitopes
		Full peptide (15 mer)	Core peptide (9 mer)			
<i>HLA-DRB1*0101</i>	181	<i>QQFIYAGSLSALLDP</i>	<i>FIYAGSLSA</i>	4.5	106	105
HLA-DRB1*0301	269	LENFVRSSNLKFQDA	FVRSSNLKF	17.1	6	26
HLA-DRB1*0401	46	LPVEYLQVPSMGR	YLQVPSM	18.1	9	95
HLA-DRB1*0404	46	LPVEYLQVPSMGR	VEYLQVSP	22.3	10	99
HLA-DRB1*0405	180	PQQFIYAGSLSALLD	FIYAGSLSA	18.5	18	65
HLA-DRB1*0701	179	HPQQFIYAGSLSALL	FIYAGSLSA	4.3	49	36
HLA-DRB1*0901	181	QQFIYAGSLSALLDP	YAGSLSALL	9.2	25	91
HLA-DRB1*1101	5	SRKIRAWGRRMIGT	IRAWGRRM	17.3	12	52
HLA-DRB1*1302	269	LENFVRSSNLKFQDA	FVRSSNLKF	13.5	8	35
HLA-DRB1*1501	3	DVSRKIRAWGRRMI	KIRAWGRR	13.9	9	90
HLA-DRB3*0101	74	PAVYLLDGLRAQDDY	YLLDGLRAQ	18	13	40
HLA-DRB4*0101	59	GRDIKVQFQSGGNS	IKVQFQSGG	13.9	6	48
HLA-DRB5*0101	269	LENFVRSSNLKFQDA	FVRSSNLKF	9.2	21	67
HLA-DPA1*0103-DPB1*0401	132	CQTYKWETFLTSELP	YKWETFLTS	3.9	12	31
HLA-DPA1*0103-DPB1*0201	131	GCQTYKWETFLTSEL	YKWETFLTS	2.3	17	55
HLA-DPA1*0201-DPB1*0101	132	CQTYKWETFLTSELP	YKWETFLTS	7.7	17	40
HLA-DPA1*0201-DPB1*0501	132	CQTYKWETFLTSELP	YKWETFLTS	39.7	2	18
HLA-DPA1*0301-DPB1*0402	132	CQTYKWETFLTSELP	YKWETFLTS	21.7	10	44
HLA-DQA1*0101-DQB1*0501	114	GQSSFYSDWYSPACG	QSSFYSDWY	19.4	5	21
HLA-DQA1*0102-DQB1*0602	165	LSMAGSSAMILAAYH	GSSAMILAA	7.1	24	86
HLA-DQA1*0501-DQB1*0201	259	LGGANIPAEFLENFV	IPAEFLENF	34.5	3	24
HLA-DQA1*0501-DQB1*0301	27	GLVGLAGGAATAGAF	GLAGGAATA	3.7	82	94

^aBinding of epitopes of Ag85B from *M. tuberculosis* was predicted using NetMHCII 2.2 server in accordance with the instruction of the server. 22 MHC II molecules against which bindings were strong are shown. Binding affinity below 50 nM is considered strong. The strongest epitope bound to HLA-DRB1*0101 MHC II molecule is italicized.

TABLE 4: Binding of 181-QQFIYAGSLSALLDP-195 epitope of Ag85B to respective allele predicted by NetMHCII 2.2 servers^a.

MHC II alleles	Promiscuous peptide sequence		Affinity (nM)
	Full peptide (15 mer)	Core peptide (9 mer)	
HLA-DRB1*0101	QQFIYAGSLSALLDP	FIYAGSLSA	4.5
HLA-DRB1*0404	QQFIYAGSLSALLDP	FIYAGSLSA	49.1
HLA-DRB1*0405	QQFIYAGSLSALLDP	FIYAGSLSA	20.7
HLA-DRB1*0701	QQFIYAGSLSALLDP	FIYAGSLSA	6.1
HLA-DRB1*0901	QQFIYAGSLSALLDP	FIYAGSLSA	9.2
HLA-DRB5*0101	QQFIYAGSLSALLDP	FIYAGSLSA	33.7
HLA-DPA1*0201-DPB1*0101	QQFIYAGSLSALLDP	FIYAGSLSA	38.1
HLA-DQA1*0501-DQB1*0301	QQFIYAGSLSALLDP	FIYAGSLSA	5.7

^aBinding affinity below 50 nM is considered as strong immunogenicity.

<u><i>M. tuberculosis_1 H37Rv</i></u>	141	LTSELPQWLSANRAVKPTGSAAIGLSMAGSSAMILAAYHP	<u>QQFIYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	216
<i>M. tuberculosis_2 CDC1551</i>	190	LTSELPQWLSANRAVKPTGSAAIGLSMAGSSAMILAAYHP	<u>QQFIYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	265
<i>M. bovis_1 BCG str.</i>	141	LTSELPQWLSANRAVKPTGSAAIGLSMAGSSAMILAAYHP	<u>QQFIYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	216
<i>M. tuberculosis_3</i>	141	LTSELPQWLSANRAVKPTGSAAIGLSMAGSSAMILAAYHP	<u>QQFIYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	216
<i>M. tuberculosis_4 SUMu003</i>	138	LTSELPQWLSANRAVKPTGSAAIGLSMAGSSAMILAAYHP	<u>QQFIYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	213
<i>M. tuberculosis_5 H37Ra</i>	126	LTSELPQWLSANRAVKPTGSAAIGLSMAGSSAMILAAYHP	<u>QQFIYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	201
<i>M. bovis_2</i>	141	LTSELPQWLSANRAVKPTGSAAIGLSMAGSSAMILAAYHP	<u>QQFIYAGSLSALLDP</u>	SQGMG--LIGLAMGDAGGYKA	214
<i>M. tuberculosis_6</i>	101	LTSELPQWLSANRAVKPTGSAAIGLSMAGSSAMILAAYHP	<u>QQFIYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	176
<i>M. tuberculosis_7</i>	101	LTSELPQWLSANRAVKPTGSAAIGLSMAGSSAMILAAYHP	<u>QQFIYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	176
<i>M. kansasii ATCC 12478</i>	141	<u>LTSELPQWLSANRSVKPTGSAAVGISMAGSSALILSVYHP</u>	<u>QQFIYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	216
<i>M. avium_1</i>	141	<u>LTSELPQYLASNKGVKRTGNAAVGISMSGSSAMILAVNHP</u>	<u>QQFIYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	216
<i>M. intracellulare_1 ATCC</i>	141	<u>LTSELPQYLASNKGVKSTGSAAVGISMSGSSAMILAVNHP</u>	<u>QQFVYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	216
<i>M. avium_2</i>	141	<u>LTSELPQYLASNKGVKRTGSAAVGISMSGSSAMILAVNHP</u>	<u>QQFVYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	216
<i>M. intracellulare_2 ATCC</i>	126	<u>LTSELPQYLASNKGVKSTGSAAVGISMSGSSAMILAVNHP</u>	<u>QQFVYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	201
<i>M. intracellulare_3</i>	95	<u>LTSELPQYLASNKGVKSTGSAAVGISMSGSSAMILAVNHP</u>	<u>QQFVYAGSLSALLDP</u>	SQGMGPSLIGLAMGDAGGYKA	170

*****.:*:*:* **.*:*:*:*:*:*:*:*:*:* * *****

FIGURE 2: Multiple alignment of protein sequences of Ag85B from *Mycobacterium* spp. Accession numbers of Ag85B from *M. tuberculosis* were NP_216402 (*M. tuberculosis_1 H37Rv*), NP_336393 (*M. tuberculosis_2 CDC1551*), IF0PA (*M. tuberculosis_3*), ZP_07422952 (*M. tuberculosis_4 SUMu003*), ZP_02552646 (*M. tuberculosis_5 H37Ra*), IF0N_A (*M. tuberculosis_6*), and AAO62005 (*M. tuberculosis_7*). Accession numbers of Ag85B from *M. bovis* were YP_978013 (*M. bovis_1 BCG str.*) and AAA25359 (*M. bovis_2*). Accession numbers of Ag85B from *M. intracellulare* were YP_005338175 (*M. intracellulare_1 ATCC*), BAA03981 (*M. intracellulare_2 ATCC*), and BAA03243 (*M. intracellulare_3*). Accession numbers of Ag85B from *M. avium* were Q06947 (*M. avium_1*) and AAM21939 (*M. avium_2*), and those of Ag85B from *M. kansasii* were ZP_04749192 (*M. kansasii ATCC 12478*). The best epitope based on conservation and NetMHCII 2.2 analysis is shown in the box. The epitope sequence from *M. tuberculosis_1 H37Rv* is underlined. Nonidentical epitope sequences are shaded and the amino acid residues for which the epitope varies are bold, italic, and underlined.

TABLE 5: Binding of epitopes of Ag85B from *M. tuberculosis* against HLA-DRB1*0701 molecules^a.

Epitope rank	Start position	Sequence	Score
<i>Predicted with ProPed</i>			
1	271	FVRSSNLKF	8
2	166	MAGSSAMIL	5.9
3	154	VKPTGSAAI	5.8
4	184	YAGSLSALL	5.7
5	101	YQSGLSIVM	5.3
<i>Predicted with SYFPEITHI</i>			
1	152	NRAVKPTGSAAIGLS	32
2	182	QFIYAGSLSALLDS	32
3	162	AIGLSMAGSSAMILA	30
4	180	PQQFIYAGSLSALLD	28
5	217	ADMWGPSSDPAWERN	28

^aBinding of epitopes of Ag85B from *Mycobacterium* spp. was predicted by analyzing with ProPred and SYFPEITHI servers. The higher the score of an epitope, the greater the probability of binding to a given MHC molecule. The core sequence (9-mer) of the epitope 181-QQFIYAGSLSALLDP-195 is italicized. Five epitopes with the top scores are shown.

of the control epitope and the strongest promiscuous epitope (Additional File 4, Figure S1). Figure S1 shows the percentage of identity in amino acid level of the predicted strongest epitope was 100, 93, and 86 in 6, 42, and 25 influenza A viruses, respectively. The level of identity was 86.75% in all influenza A virus.

Physicochemical properties of the epitope include a molecular weight of 1622.8 Da with 230 atoms (C₇₅H₁₁₅N₁₇O₂₃). Analysis with the ProtParam computer

program revealed that the instability index was 32.93. This result indicates a high stability of the epitope since a ProtParam instability index smaller than 40 is considered to be stable. However, the ProtParam instability index of the positive control epitope and the predicted strongest promiscuous epitope of IvAgNP was 76.04 and 46.69, respectively, indicating that these epitopes are unstable.

3.2. Docking of the Predicted T-Cell Epitope to MHC Molecules.

Molecular docking was performed to analyze the interaction of the epitope with MHC molecules. Docking showed that the protrusion of the side chains of the epitope bound in cavities within the groove of MHC II HLA-DRB1*0101 through nine hydrogen bonds (Figure 3(a); Additional File 5, Table S4). Docking simulation revealed that the binding energy between the epitope and HLA-DRB1*0101 was -181.93 kcal/mol. A binding energy less than -32 kcal/mol is generally considered a good score in the ICM method [55]. Molecular docking of the epitope was further performed with DRB1*0901, DRB1*0404, DRB1*0701, DPB1*0101, and DQB1*0301 (Figures 3(b)–3(f)). Docking analysis showed that the epitope side chains bound in the cavities within the grooves of HLA-DRB1*0404, DRB1*0701, DRB1*0901, DPB1*0101, and DQB1*0301 through 9, 8, 6, 5, and 9 hydrogen bonds, respectively, and the binding energy of the epitope with these MHC II molecules were -162.42, -179.42, -178.35, -159.39, and -202.60 kcal/mol, correspondingly (Additional File 5, Table S4). As HLA-DRA does not have polymorphisms in the peptide binding region, we performed molecular docking simulation between the epitope and HLA-DRA (Figure 3(g)). Docking analysis revealed that the epitope bound to the groove by six hydrogen bonds with the binding energy of

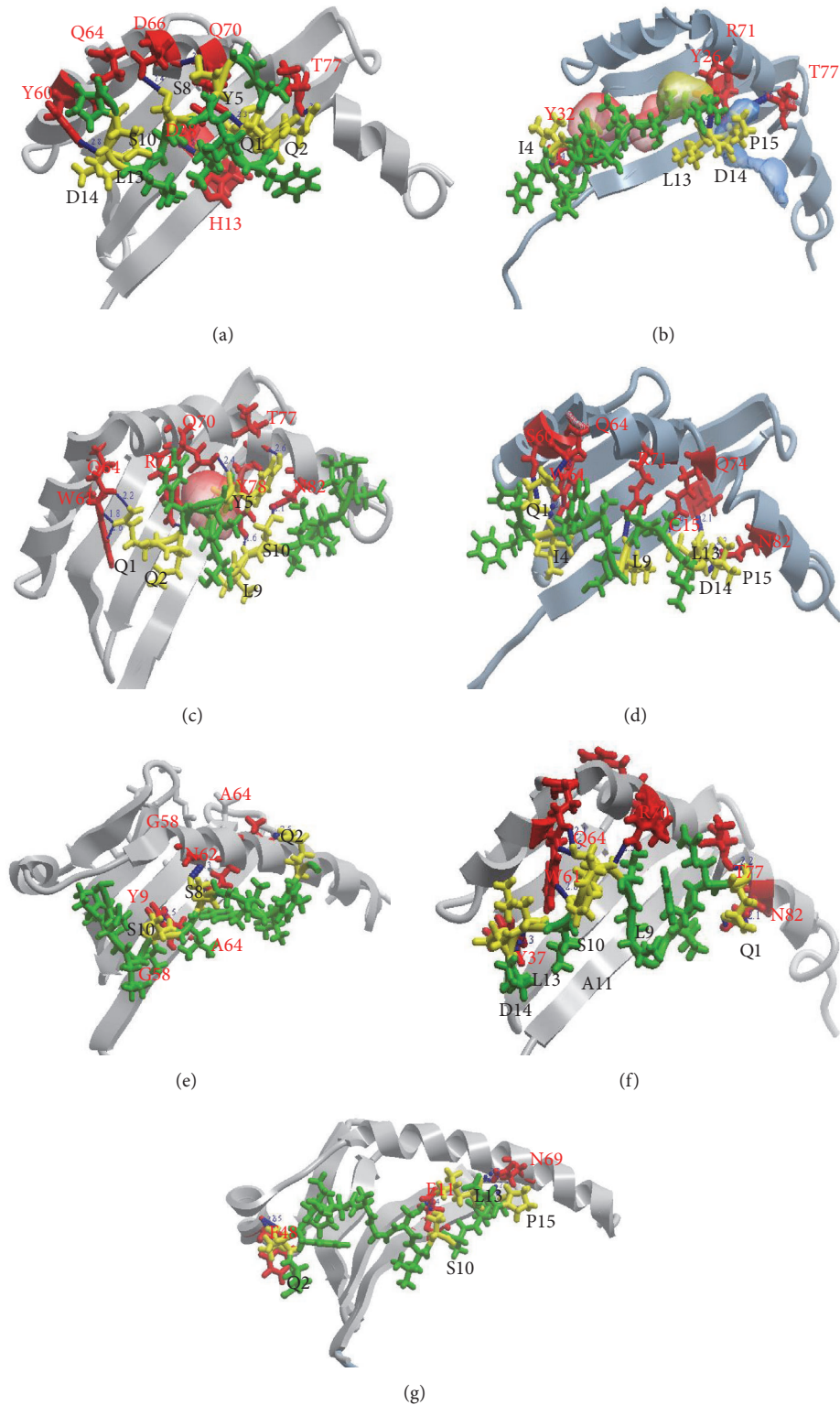


FIGURE 3: Docking of the epitope with HLA-DRB1*0101 (a), HLA-DRB1*0901 (b), HLA-DRB1*0404 (c), HLA-DRB1*0701 (d), HLA-DPB1*0101 (e), HLA-DQB1*0301 (f), and HLA-DRA (g). The 3D structure of the Ag85B T-cell epitope was made in PyMol molecule builder. Docking of receptor MHC with epitope was performed in Molsoft ICM-pro 3.5. All 3D structures of HLA proteins were obtained from the PDB server unless noted. Homology models of HLA-DRB1*0701 and HLA-DRB1*0901 were prepared in MOE as described by Azad et al. [16]. MHC II structures are shown in gray ribbon and the amino acid residues involved in the hydrogen bonding interaction are shown as red sticks and labeled red. The epitope is shown as green ribbon and the amino acid residues involved in the hydrogen bonding network are shown as yellow sticks and labeled black. MHC II binding pockets are shown as an electrostatic sphere. Hydrogen bonds are displayed as blue spheres and the hydrogen bonding distances are labeled blue. Hydrogen bonding interaction between the amino acid residues of the epitope and those of MHC II are detailed in Additional File 5, Table S4.

−204.17 kcal/mol. Docking analysis of the predicted strongest epitope, 301-IDPFRLLQNSQVYSL-315, on the positive control protein showed that the epitope side chains bound in the cavities within the grooves of HLA-DRB1*0101, DRB1*0901, DRB1*0404, and DRB1*0701 through 9, 9, 7, and 8 hydrogen bonds, respectively, and the binding energy of the epitope with these MHC II molecules were −194.87, −211.17, −189.62, and −185.42 kcal/mol, correspondingly (Additional File 6, Figure S2; Additional File 7, Table S5). However, docking analysis showed that the epitope 380-ELRSRYWAIRTRSG-393 in IvAgNP bound within the grooves of only HLA-DRB1*0101 through 10 hydrogen bonds with a binding energy of −191.05 kcal/mol (data not shown).

Docking of the 9-mers of the predicted epitope was performed with MHC I molecules. Docking analysis showed that the 9-mer epitope 183-FIYAGLSA-191 bound to HLA-A2 through eight hydrogen bonds with a binding energy of −162.01 kcal/mol (Figure 4; Additional File 8, Table S6). The epitope 184-IYAGLSAL-192 bound in the cavities within the groove of HLA-A24 and B39 through seven hydrogen bonds with a binding energy of −185.45 and −175.42 kcal/mol, respectively (Figure 4; Additional File 8, Table S6). The 9-mer epitope 181-QQFIYAGSL-189 bound to HLA-B27 and HLA-B44 through 11 hydrogen bonds with a binding energy of −193.62 and −200.13 kcal/mol, respectively (Figure 4; Additional File 8, Table S6). Docking analysis also revealed that the 9-mers of the epitopes in IvAgNP bound to MHC I supertypes with a significant number of hydrogen bonds and binding energy (Additional Files 9 and 11, Figures S3, S4; Additional Files 10 and 12, Tables S7, S8).

3.3. Receptor Binding Sites for the Epitope 181-QQFIYAGSLSALLDP-195. The MHC II binding site is formed by 35 amino acids from the N-terminal 80 and 90 residues of the HLA-DRA and HLA-DRB, respectively [52, 55, 62]. As the HLA-DRA exhibits no binding site polymorphisms [63], we used only the HLA-DRB1 chains for analysis. Our analysis revealed that DRB1 chains contain 29 polymorphic amino acids in the binding sites. The polymorphic positions are 4, 9, 10, 11, 12, 13, 14, 16, 26, 28, 30, 31, 32, 33, 37, 38, 40, 47, 57, 60, 67, 70, 71, 73, 74, 77, 78, 85, and 86 (Figure 5). Docking analysis of the receptor position showed that the epitope bound to 15 positions of HLA-DRB1 alleles. Among them, positions 15(C), 61(W), 64(Q), 66(D), and 82(N) presented no polymorphism in the receptor binding site. However, only one residue among the positions 28(D), 32(Y), 60(Y), 77(T), and 78(Y) of the DRB1 chain was found to be substituted. We also analyzed interaction sites (residues 1–179) of MHC I molecules and found that the peptide interacted at 17 positions of HLA-A and B proteins (Additional File 13, Figure S5). Among them, residues at positions 59(Y), 84(Y), 99(Y), 143(T), 147(W), 155(Q), and 159(Y) showed no polymorphism.

4. Discussion

Peptide or epitope-based vaccines are specific and induce immunity against the selected epitope(s). Due to this specificity, these vaccines can avoid immune responses induced

by unfavorable epitopes of the antigen, thereby preventing complications [64]. Along with increased safety, peptide-based vaccines give the opportunity to engineer the epitope to increase its universality, potency, and breadth [65]. In addition, epitope-based vaccines have been shown to be successful in preventing infectious diseases like influenza [66]. Small epitopes (usually 8–15 amino acids in length) can be delivered easily by any of the following devices such as peptide-adjuvant conjugates (lipid-core peptides), peptide amphiphiles, lipid-based synthetic vesicles, endogenous exosomes containing peptide-loaded MHC molecules, noninfectious virus-like particles displaying recombinant epitopes, surface-conjugated peptides, and solid-core nanobeads with conjugated peptides [67]. The effectiveness of a peptide-based vaccine is greatly dependent on the exact identification of the immunogenic epitopes that bestow protection. In this study, different computer-aided bioinformatics tools have been used to identify the potentially promiscuous epitope 181-QQFIYAGSLSALLDP-195 from Ag85B, which may induce T-cell-mediated immunity against *Mycobacterium* spp. In the Immune Epitope Database, the Ag85B epitopes are recognized as immunodominant peptides that may react with human T-cells [68]. Computational tools used in our analysis have been validated by applying them on a reported IvAgNP epitope that was used experimentally as a vaccine following characterization of its epitope, 380-ELRSRYWAIRTRSG-393. The epitope of IvAgNP was associated with CD8+ cell-based immunity [60, 61]. Through analysis by these tools, we found the same HLA allele that was previously characterized. However, the epitope 380-ELRSRYWAIRTRSG-393 did not possess high potentiality to induce both types of cell-based immunity. Through our analysis on IvAgNP, we identified several limitations of its epitope, including (i) lack of high conservation in almost all pathogenic influenza viruses, (ii) absence of a promiscuous nature, and (iii) a ProtParam instability index greater than 40 that indicates a significantly reduced half-life. These limitations may suggest why this epitope was only investigated at the experimental level, never reaching successful epitope-based vaccination. However, the strongest epitope identified in our study, 181-QQFIYAGSLSALLDP-195 on Ag85B, had overcome such limitations.

Analysis with Vaxign revealed that Ag85B was more antigenic than other immunogenic mycobacterial proteins. Results of analyses with NetMHCII 2.2, SYFPEITHI, ProPred, and NetCTL 1.2 servers confirmed that the predicted epitope from Ag85B of *M. tuberculosis* could bind to several MHC II and MHC I molecules and thus revealed its promiscuous nature [52, 57]. Epitope promiscuity is usually assessed in the generation of epitope-based vaccines [69–71]. In addition, promiscuously binding antigenic epitopes are considered to act against a vast range of host immune systems.

It is important to recognize the inherent limitations of epitope prediction programs. No epitope prediction program is absolutely accurate, and there is substantial discrepancy among the epitopes predicted by each program [72, 73]. We increased the accuracy of our analysis by using multiple prediction programs for each DRB allele that have undergone

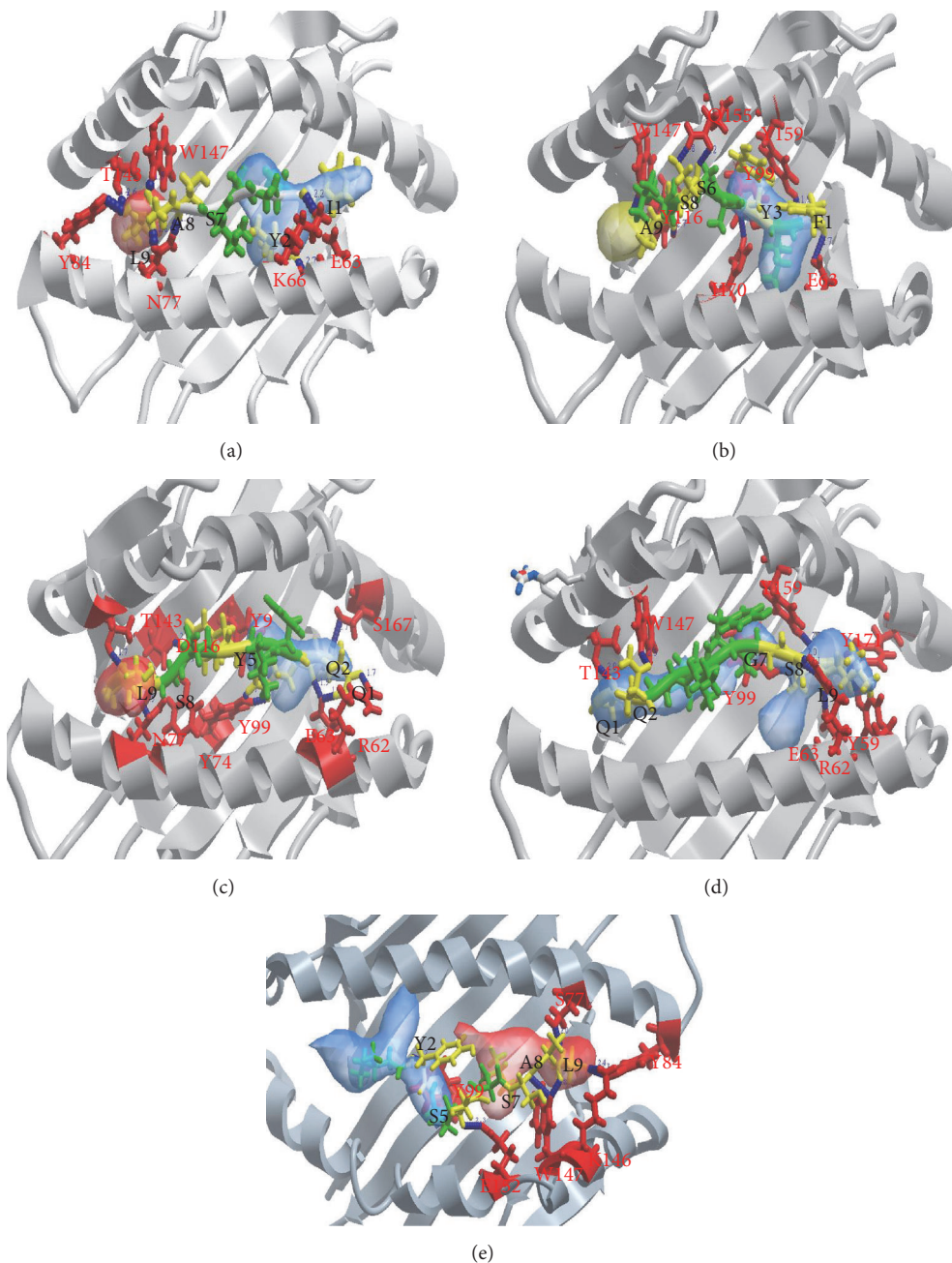


FIGURE 4: Docking of the epitope with HLA-A24 (a), HLA-A2 (b), HLA-B44 (c), HLA-B27 (d), and HLA-B39 (e). Building of the 3D structure of the Ag85B T-cell epitope and 3D models of HLA-B39 and the docking of receptor MHC with the epitope were performed as mentioned in Figure 3. MHC I structures are shown in gray ribbon and the amino acid residues involved in the H-bonding interaction are shown as red sticks and labeled red. The epitope is shown as green ribbon and the amino acid residues involved in the H-bonding network are shown as yellow sticks and labeled black. MHC I binding pockets are shown as electrostatic sphere. H-bonds are displayed as blue spheres and the H-bonding distances are labeled blue. H-bonding interaction between the amino acid residues of the epitope and those of the MHC I are detailed in Additional File 8, Table S6.

extensive validation [47, 72, 73]. Fortunately, there was agreement among programs. We focused more on the class II MHC DRB1 alleles because of the comparative abundance of data on epitopes binding to DR alleles. It has been documented for TB that more than 90% of class II *M. tuberculosis* epitopes bind to DR antigens [74]. In addition, the *DRB1* gene was extensively studied because the expression level of DRB1 proteins is

typically fivefold greater than that of *DRB3*, *DRB4*, or *DRB5* gene [75].

NetCTL 1.2 is an integrative approach to CTL epitope prediction that predicts MHC I binding, TAP transport efficiency, and proteasomal cleavage [50, 62]. Therefore, the epitope's ability to induce cytotoxic T-cell (CD8+), and natural killer cell (inhibitory receptor), mediated immunity

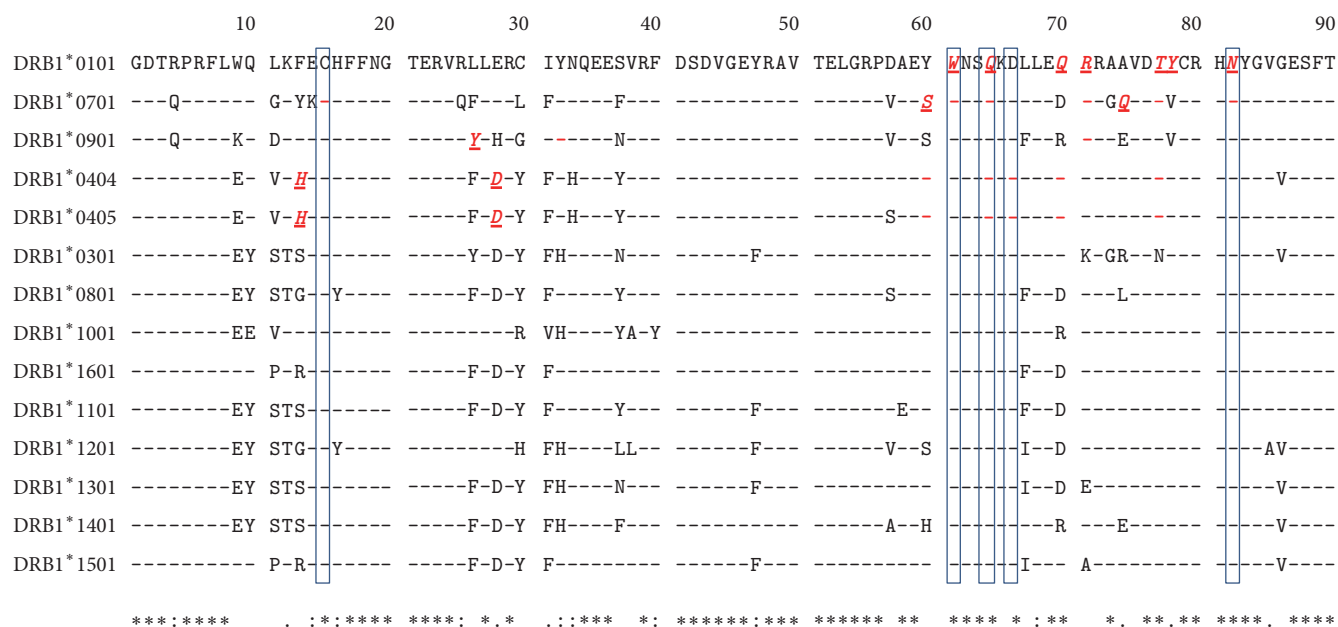


FIGURE 5: Multiple sequence alignment of amino acid sequences of HLA-DRB1 series. The T-Coffee server was used for multiple alignment of the N-terminal 90 amino acid residues of the HLA-DRB alleles. The HLA-DRB1*0101 sequence was chosen as reference. Identities of the residues are illustrated by dashes. Positions 15(C), 61(W), 64(Q), 66(D), and 82(N) are indicated as boxes to present no polymorphic binding site in the MHC protein. Binding residues are labeled red, bold, italic, and underlined. Identical binding residues in MHC proteins are indicated as red dashes.

might be satisfactory as it can bind to five MHC I HLA (two A and three B) supertypes. Immunity to TB in humans, primates, and mice depends on T lymphocytes [76, 77]. So the treatment of TB actually requires cell-mediated immunity that includes both CD4+ and CD8+ T-cells [4, 18, 19, 78]. MHC II molecules interact with CD4+ T helper cells that stimulate macrophages to kill phagocytosed pathogens. Since MTBC inhibits macrophages, CD4+ T-cell-mediated immunity is particularly important for preventing and clearing MTBC infection [18].

Little is known about the amino acid variation of Ag85B proteins in *Mycobacterium* spp. If the proposed epitope sequence is highly variable among the Ag85B, the protective efficacy of Ag85B as an epitope-based vaccine might be compromised on the global stage. Our investigation showed that the proposed epitope had 100% identity among MTBC and 87–93% identity among the remaining 40% of *Mycobacterium* spp., defining it as a universal epitope (Figure 2). A seminal observation of 22 genomes of MTBC reveals that most of the experimentally identified human T-cell epitopes are hyperconserved on the evolutionary scale [4]. These findings provide evidence that the relationship between MTBC and its human host actually may differ from a classical evolutionary arms-race and informs the development of new approaches to control TB. However, recently, Copin et al. described computationally and experimentally that highly conserved MTB T-cell epitopes in fact contribute to pathology [30]. Therefore, it is noteworthy that a vaccine encoding conserved promiscuous epitopes not only shows protective immunity against wide range of pathogenic strains but also induces broad T-cell responses [79–81].

Recently, the incidence of Nontuberculous Mycobacterial (NTM) infections has surpassed the incidence of TB infections in developed countries [82–86]. Although infection may occur virtually in any organ, pulmonary infections are most common [87–89] with *M. avium* and *M. kansasii* as the causative agents of lung disease [82, 84, 90]. Difficulty eradicating NTM and its high rate of reoccurrence have made it complicated to make clinical decisions for treatment. The proposed epitope in our study had 87–93% identity among *M. avium*, *M. kansasii*, and *M. intracellulare*. Therefore, this vaccine candidate would potentially target the diversity of *Mycobacterium* spp. worldwide. Multiple alignment of the Ag85B protein sequence showed that glutamine (Q181 in *M. tuberculosis*-1 H37Rv) present in 10 Ag85B sequences is substituted by aspartic acid and asparagine in *M. avium* and *M. intracellulare*, respectively. Isoleucine (I184 in *M. tuberculosis*-1 H37Rv) present in 12 Ag85B sequences is substituted by valine in *M. intracellulare*. Leucine (L192 in *M. tuberculosis*-1 H37Rv) present in 14 Ag85B sequences is substituted by methionine in *M. kansasii*. Glutamine has strongly similar properties (roughly the same size and the same hydrophathy) as aspartic acid and asparagine, while isoleucine is similar to valine [91]. Therefore, the substituted amino acid residues in the predicted epitope of Ag85B in only ~33% of *Mycobacterium* spp. may not alter the properties of the epitope to change its universality. ProtParam analysis of the epitope with the substituted amino acid(s) to mimic the polymorphism in *M. avium*, *M. intracellulare*, and *M. kansasii* showed almost the same half-life and instability index as that of the predicted epitope (data not shown). Thus, comparative proteomics of the pathogen population from a greater range

of geographic origins stands as an additional useful tool for preclinical evaluation of new vaccine targets.

The Ag85B T-cell epitope formed a number of significant hydrogen bonds with the binding groove of each of the MHC II and MHC I molecules (Figures 3 and 4; Additional Files 5 and 8, Tables S4, S6). Hydrogen bonding distance less than 3 Å is usually considered biologically significant. The binding energy of the Ag85B T-cell epitope with each of the MHC I and II molecules is also noteworthy since a binding energy below -32 kcal/mol is biologically significant [55]. The epitope has the ability to bind to both the residues inside the groove as well as outside the binding groove (flanking residues) of the MHC molecules. This is considered as a principal determinant of epitope binding affinity [53, 57, 92–97].

It is necessary to incorporate information on host and pathogen genetic diversity in vaccine design, development, and testing. The assortment of the HLA and other genes involved in the immune system results in variable vaccine efficacy in different individuals even in the absence of pathogen variation [98]. Our analysis of the “Allele Frequency Database” revealed that the HLA alleles to which the predicted epitope bound were commonly found in TB-endemic regions. Among them, the DRB1 alleles have the highest frequency (~95%) in the TB burden countries. This result stresses the notion that the proposed epitope may confer an enhanced immunity to the patients of certain endemic region of the world. Furthermore, it might be a better candidate as an epitope-based vaccine because our data showed that the epitope had preference for both CD4+ and CD8+ T-cell-based immunity. Additionally, the binding residues of MHC had almost no polymorphism in all MHC DRB1 and MHC A and B series of proteins. Therefore, the predicted epitope would induce immunity against almost all class I and class II alleles and may have great implication for an epitope-based vaccine with the potential to impact the global population.

5. Conclusion

Here, we report (i) high percentage of identity of the epitope in all Ag85B sequences, (ii) adhesin capability and antigen probability of Ag85B from *M. tuberculosis*, (iii) strong binding affinity of the epitope with MHC I and II molecules, (iv) physicochemical properties of the epitope, (v) molecular docking simulation, and (vi) coverage of HLA alleles in TB-endemic countries, which supports the notion that the proposed epitope might be a novel universal efficient epitope to produce peptide or epitope-based vaccines against TB. While biochemical analysis remains necessary to validate the interaction between the epitope and the MHC molecules through elucidation of immunity induction, most of the approaches have been validated by applying them on a positive control protein, IvAgNP.

Disclosure

Md Saddam Hossain and Abul Kalam Azad are co-first authors.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] World Health Organization, *Global Tuberculosis Report 2013*, World Health Organization, Geneva, Switzerland, 2013.
- [2] World Health Organization, *WHO Report. Global Tuberculosis Control 2011*, World Health Organization, Geneva, Switzerland, 2013.
- [3] S. T. Cole, R. Brosch, J. Parkhill et al., “Deciphering the biology of mycobacterium tuberculosis from the complete genome sequence,” *Nature*, vol. 393, no. 6685, pp. 537–544, 1998.
- [4] Í. Comas, J. Chakravarti, P. M. Small et al., “Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyper-conserved,” *Nature Genetics*, vol. 42, no. 6, pp. 498–503, 2010.
- [5] M. Achtman, “Evolution, population structure, and phylogeography of genetically monomorphic bacterial pathogens,” *Annual Review of Microbiology*, vol. 62, pp. 53–70, 2008.
- [6] I. Comas, S. Homolka, S. Niemann, and S. Gagneux, “Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies,” *PLoS ONE*, vol. 4, no. 11, Article ID e7815, 2009.
- [7] R. Hershberg, M. Lipatov, P. M. Small et al., “High functional diversity in *Mycobacterium tuberculosis* driven by genetic drift and human demography,” *PLoS Biology*, vol. 6, no. 12, pp. 2658–2671, 2008.
- [8] I. Comas, M. Coscolla, T. Luo et al., “Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans,” *Nature Genetics*, vol. 45, no. 10, pp. 1176–1182, 2013.
- [9] R. Firdessa, S. Berg, E. Hailu et al., “Mycobacterial lineages causing pulmonary and extrapulmonary Tuberculosis, Ethiopia,” *Emerging Infectious Diseases*, vol. 19, no. 3, pp. 460–463, 2013.
- [10] S. Gagneux, K. DeRiemer, T. Van et al., “Variable host-pathogen compatibility in *Mycobacterium tuberculosis*,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 8, pp. 2869–2873, 2006.
- [11] A. E. Hirsh, A. G. Tsolaki, K. DeRiemer, M. W. Feldman, and P. M. Small, “Stable association between strains of *Mycobacterium tuberculosis* and their human host populations,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, no. 14, pp. 4871–4876, 2004.
- [12] M. B. Reed, V. K. Pichler, F. McIntosh et al., “Major *Mycobacterium tuberculosis* lineages associate with patient country of origin,” *Journal of Clinical Microbiology*, vol. 47, no. 4, pp. 1119–1128, 2009.
- [13] P. E. M. Fine, “Variation in protection by BCG: implications of and for heterologous immunity,” *The Lancet*, vol. 346, no. 8986, pp. 1339–1345, 1995.
- [14] A. C. Hesselning, B. J. Marais, R. P. Gie et al., “The risk of disseminated Bacille Calmette-Guerin (BCG) disease in HIV-infected children,” *Vaccine*, vol. 25, no. 1, pp. 14–18, 2007.
- [15] R. E. Weir, G. F. Black, B. Nazareth et al., “The influence of previous exposure to environmental mycobacteria on the interferon-gamma response to bacille Calmette-Guérin vaccination in southern England and northern Malawi,” *Clinical and Experimental Immunology*, vol. 146, no. 3, pp. 390–399, 2006.
- [16] A. K. Azad, N. Yoshikawa, T. Ishikawa, Y. Sawa, and H. Shibata, “Substitution of a single amino acid residue in the

- aromatic/arginine selectivity filter alters the transport profiles of tonoplast aquaporin homologs,” *Biochimica et Biophysica Acta—Biomembranes*, vol. 1818, no. 1, pp. 1–11, 2012.
- [17] M. V. Tullius, G. Harth, S. Masleša-Galić, B. J. Dillon, and M. A. Horwitz, “A replication-limited recombinant *Mycobacterium bovis* BCG vaccine against tuberculosis designed for human immunodeficiency virus-positive persons is safer and more efficacious than BCG,” *Infection and Immunity*, vol. 76, no. 11, pp. 5200–5214, 2008.
- [18] D. F. Hoft, “Tuberculosis vaccine development: goals, immunological design, and evaluation,” *The Lancet*, vol. 372, no. 9633, pp. 164–175, 2008.
- [19] T. H. M. Ottenhoff and S. H. E. Kaufmann, “Vaccines against tuberculosis: where are we and where do we need to go?” *PLoS Pathogens*, vol. 8, no. 5, Article ID e1002607, 2012.
- [20] S. H. Kaufmann, G. Hussey, and P.-H. Lambert, “New vaccines for tuberculosis,” *The Lancet*, vol. 375, no. 9731, pp. 2110–2119, 2010.
- [21] J. Davila, L. A. McNamara, and Z. Yang, “Comparison of the predicted population coverage of tuberculosis vaccine candidates Ag85b-ESAT-6, Ag85B-TB10.4, and Mtb72f via a bioinformatics approach,” *PLoS ONE*, vol. 7, no. 7, Article ID e40882, 2012.
- [22] L. A. McNamara, Y. He, and Z. Yang, “Using epitope predictions to evaluate efficacy and population coverage of the Mtb72f vaccine for tuberculosis,” *BMC Immunology*, vol. 11, article no. 18, 2010.
- [23] M. J. Brennan and J. Thole, “Tuberculosis vaccines: a strategic blueprint for the next decade,” *Tuberculosis*, vol. 92, no. 1, pp. S6–S13, 2012.
- [24] M. Marmiesse, P. Brodin, C. Buchrieser et al., “Macro-array and bioinformatic analyses reveal mycobacterial “core” genes, variation in the ESAT-6 gene family and new phylogenetic markers for the *Mycobacterium tuberculosis* complex,” *Microbiology*, vol. 150, no. 2, pp. 483–496, 2004.
- [25] S. Sreevatsan, X. Pan, K. E. Stockbauer et al., “Restricted structural gene polymorphism in the *Mycobacterium tuberculosis* complex indicates evolutionarily recent global dissemination,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 94, no. 18, pp. 9869–9874, 1997.
- [26] I. Filliol, A. S. Motiwala, M. Cavatore et al., “Global phylogeny of *Mycobacterium tuberculosis* based on single nucleotide polymorphism (SNP) analysis: insights into tuberculosis evolution, phylogenetic accuracy of other DNA fingerprinting systems, and recommendations for a minimal standard SNP set,” *Journal of Bacteriology*, vol. 188, no. 2, pp. 759–772, 2006.
- [27] A. M. Hebert, S. Talarico, D. Yang et al., “DNA polymorphisms in the *pepA* and *PPE18* genes among clinical strains of *Mycobacterium tuberculosis*: implications for vaccine efficacy,” *Infection and Immunity*, vol. 75, no. 12, pp. 5798–5805, 2007.
- [28] A. V. S. Hill, C. E. M. Allsopp, D. Kwiatkowski et al., “Common West African HLA antigens are associated with protection from severe malaria,” *Nature*, vol. 352, no. 6336, pp. 595–600, 1991.
- [29] J. C. Howard, “Disease and evolution,” *Nature*, vol. 352, no. 6336, pp. 565–567, 1991.
- [30] R. Copin, M. Coscollá, S. N. Seiffert et al., “Sequence diversity in the *pe_pgrs* genes of *Mycobacterium tuberculosis* is independent of human T cell recognition,” *mBio*, vol. 5, no. 1, Article ID e00960-13, 2014.
- [31] J. M. Blackwell, S. E. Jamieson, and D. Burgner, “HLA and infectious diseases,” *Clinical Microbiology Reviews*, vol. 22, no. 2, pp. 370–385, 2009.
- [32] L. Brandt, M. Elhay, I. Rosenkrands, E. B. Lindblad, and P. Andersen, “ESAT-6 subunit vaccination against *Mycobacterium tuberculosis*,” *Infection and Immunity*, vol. 68, no. 2, pp. 791–795, 2000.
- [33] S. B. Sable, I. Verma, and G. K. Khuller, “Multicomponent antituberculous subunit vaccine based on immunodominant antigens of *Mycobacterium tuberculosis*,” *Vaccine*, vol. 23, no. 32, pp. 4175–4184, 2005.
- [34] A. W. Olsen, L. A. H. Van Pinxteren, L. M. Okkels, P. B. Rasmussen, and P. Andersen, “Protection of mice with a tuberculosis subunit vaccine based on a fusion protein of antigen 85B and ESAT-6,” *Infection and Immunity*, vol. 69, no. 5, pp. 2773–2778, 2001.
- [35] Y. He and Z. Xiang, “Bioinformatics analysis of *Brucella* vaccines and vaccine targets using VIOLIN,” *Immunome Research*, vol. 6, no. 1, article S5, 2010.
- [36] Y. He, Z. Xiang, and H. L. T. Mobley, “Vaxign: the first web-based vaccine design program for reverse vaccinology and applications for vaccine development,” *Journal of Biomedicine and Biotechnology*, vol. 2010, Article ID 297505, 15 pages, 2010.
- [37] J. R. Kurtz, H. E. Petersen, D. R. Frederick, L. A. Morici, and J. B. McLachlan, “Vaccination with a single CD4 T cell peptide epitope from a *Salmonella* type III-secreted effector protein provides protection against lethal infection,” *Infection and Immunity*, vol. 82, no. 6, pp. 2424–2433, 2014.
- [38] J. Tu, B. Hou, B. Wang et al., “A multi-epitope vaccine based on *Chlamydia trachomatis* major outer membrane protein induces specific immunity in mice,” *Acta Biochimica et Biophysica Sinica*, vol. 46, no. 5, pp. 401–408, 2014.
- [39] E. K. Jagusztyn-Krynicka, P. Roszczenko, and A. Grabowska, “Impact of proteomics on anti-*Mycobacterium tuberculosis* (MTB) vaccine development,” *Polish Journal of Microbiology*, vol. 58, no. 4, pp. 281–287, 2009.
- [40] Y. He, R. Rappuoli, A. S. De Groot, and R. T. Chen, “Emerging vaccine informatics,” *Journal of Biomedicine and Biotechnology*, vol. 2010, Article ID 218590, 26 pages, 2010.
- [41] Z. Yang, J. Davila, L. Zhang, C. F. Marrs, and R. Durmaz, “Assessment of the genetic diversity of *Mycobacterium tuberculosis* *esxA*, *esxH*, and *fbpB* genes among clinical isolates and its implication for the future immunization by new tuberculosis subunit vaccines Ag85B-ESAT-6 and Ag85B-TB10.4,” *Journal of Biomedicine and Biotechnology*, vol. 2010, Article ID 208371, 6 pages, 2010.
- [42] A. W. Olsen, A. Williams, L. M. Okkels, G. Hatch, and P. Andersen, “Protective effect of a tuberculosis subunit vaccine based on a fusion of antigen 85B and ESAT-6 in the aerosol guinea pig model,” *Infection and Immunity*, vol. 72, no. 10, pp. 6148–6150, 2004.
- [43] J. T. van Dissel, S. M. Arend, C. Prins et al., “Ag85B-ESAT-6 adjuvanted with IC31® promotes strong and long-lived *Mycobacterium tuberculosis* specific T cell responses in naïve human volunteers,” *Vaccine*, vol. 28, no. 20, pp. 3571–3581, 2010.
- [44] B. Yang, S. Sayers, Z. Xiang, and Y. He, “Protegen: a web-based protective antigen database and analysis system,” *Nucleic Acids Research*, vol. 39, S1, pp. D1073–D1078, 2011.
- [45] I. A. Doytchinova and D. R. Flower, “VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines,” *BMC Bioinformatics*, vol. 8, no. 1, article 4, 2007.
- [46] C. Notredame, D. G. Higgins, and J. Heringa, “T-coffee: a novel method for fast and accurate multiple sequence alignment,” *Journal of Molecular Biology*, vol. 302, no. 1, pp. 205–217, 2000.

- [47] M. Nielsen and O. Lund, "NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction," *BMC Bioinformatics*, vol. 10, article 296, 2009.
- [48] M. M. Schuler, M.-D. Nastke, and S. Stevanović, "SYFPEI-THI: database for searching and T-cell epitope prediction," in *Immunoinformatics: Predicting Immunogenicity In Silico*, pp. 75–93, 2007.
- [49] H. Singh and G. P. S. Raghava, "ProPred: prediction of HLA-DR binding sites," *Bioinformatics*, vol. 17, no. 12, pp. 1236–1237, 2002.
- [50] M. V. Larsen, C. Lundegaard, K. Lamberth, S. Buus, O. Lund, and M. Nielsen, "Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction," *BMC Bioinformatics*, vol. 8, no. 1, article 424, 2007.
- [51] F. F. Gonzalez-Galarza, S. Christmas, D. Middleton, and A. R. Jones, "Allele frequency net: a database and online repository for immune gene frequencies in worldwide populations," *Nucleic Acids Research*, vol. 39, supplement 1, pp. D913–D919, 2011.
- [52] E. Gasteiger, C. Hoogland, A. Gattiker et al., *Protein Identification and Analysis Tools on the ExPASy Server*, Springer, Berlin, Germany, 2005.
- [53] A. Patronov, I. Dimitrov, D. R. Flower, and I. Doytchinova, "Peptide binding prediction for the human class II MHC allele HLA-DP2: a molecular docking approach," *BMC Structural Biology*, vol. 11, article 32, 2011.
- [54] H. M. Berman, J. Westbrook, Z. Feng et al., "The protein data bank," *Nucleic Acids Research*, vol. 28, no. 1, pp. 235–242, 2000.
- [55] M. Totrov and R. Abagyan, "Flexible protein-ligand docking by global energy optimization in internal coordinates," *Proteins: Structure, Function and Genetics*, vol. 29, no. 1, pp. 215–220, 1997.
- [56] R. A. Laskowski, "PDBsum: summaries and analyses of PDB structures," *Nucleic Acids Research*, vol. 29, no. 1, pp. 221–222, 2001.
- [57] J. M. Khan and S. Ranganathan, "PDOCK: a new technique for rapid and accurate docking of peptide ligands to Major Histocompatibility Complexes," *Immunome Research*, vol. 6, no. 1, article no. S2, 2010.
- [58] T. Ben-Yedidia, H. Marcus, Y. Reisner, and R. Arnon, "Intranasal administration of peptide vaccine protects human/mouse radiation chimera from influenza infection," *International Immunology*, vol. 11, no. 7, pp. 1043–1051, 1999.
- [59] T. Ben-Yedidia and R. Arnon, "Towards an epitope-based human vaccine for influenza," *Human Vaccines*, vol. 1, no. 3, pp. 95–101, 2005.
- [60] S. Huet, D. F. Nixon, J. B. Rothbard, A. Townsend, S. A. Ellis, and A. J. McMichael, "Structural homologies between two HLA B27-restricted peptides suggest residues important for interaction with HLA B27," *International Immunology*, vol. 2, no. 4, pp. 311–316, 1990.
- [61] A. Suhrbier, C. Schmidt, and A. Fernan, "Prediction of an HLA B8-restricted influenza epitope by motif," *Immunology*, vol. 79, no. 1, pp. 171–173, 1993.
- [62] M. V. Larsen, C. Lundegaard, K. Lamberth et al., "An integrative approach to CTL epitope prediction: a combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions," *European Journal of Immunology*, vol. 35, no. 8, pp. 2295–2303, 2005.
- [63] I. Dimitrov, P. Garnev, D. R. Flower, and I. Doytchinova, "MHC class II binding prediction—a little help from a friend," *Journal of Biomedicine and Biotechnology*, vol. 2010, Article ID 705821, 8 pages, 2010.
- [64] X. Wang, G. Xie, J. Liao et al., "Design and evaluation of a multi-epitope assembly peptide (MEAP) against herpes simplex virus type 2 infection in BALB/c mice," *Virology Journal*, vol. 8, article 232, 2011.
- [65] A. Sette and J. Fikes, "Epitope-based vaccines: an update on epitope identification, vaccine design and delivery," *Current Opinion in Immunology*, vol. 15, no. 4, pp. 461–470, 2003.
- [66] T. Ben-Yedidia and R. Arnon, "Epitope-based vaccine against influenza," *Expert Review of Vaccines*, vol. 6, no. 6, pp. 939–948, 2007.
- [67] M. Black, A. Trent, M. Tirrell, and C. Olive, "Advances in the design and delivery of peptide subunit vaccines with a focus on Toll-like receptor agonists," *Expert Review of Vaccines*, vol. 9, no. 2, pp. 157–173, 2010.
- [68] R. Vita, J. A. Overton, J. A. Greenbaum et al., "The immune epitope database (IEDB) 3.0," *Nucleic Acids Research*, vol. 43, no. D1, pp. D405–D412, 2015.
- [69] J. McMurry, H. Sbai, M. L. Gennaro, E. J. Carter, W. Martin, and A. S. De Groot, "Analyzing Mycobacterium tuberculosis proteomes for candidate vaccine epitopes," *Tuberculosis*, vol. 85, no. 1-2, pp. 95–105, 2005.
- [70] P. A. Reche, J.-P. Glutting, and E. L. Reinherz, "Prediction of MHC class I binding peptides using profile motifs," *Human Immunology*, vol. 63, no. 9, pp. 701–709, 2002.
- [71] H. Shams, P. Klucar, S. E. Weis et al., "Characterization of a *Mycobacterium tuberculosis* peptide that is recognized by human CD4+ and CD8+ T cells in the context of multiple HLA alleles," *Journal of Immunology*, vol. 173, no. 3, pp. 1966–1977, 2004.
- [72] U. Gowthaman and J. N. Agrewala, "In silico tools for predicting peptides binding to HLA-class II molecules: more confusion than conclusion," *Journal of Proteome Research*, vol. 7, no. 1, pp. 154–163, 2008.
- [73] P. Wang, J. Sidney, C. Dow, B. Mothé, A. Sette, and B. Peters, "A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach," *PLOS Computational Biology*, vol. 4, no. 4, Article ID e1000048, 2008.
- [74] M. J. Blythe, Q. Zhang, K. Vaughan et al., "An analysis of the epitope knowledge related to *Mycobacteria*," *Immunome Research*, vol. 3, no. 1, article 10, 2007.
- [75] S. Contini, M. Pallante, S. Vejbaesya et al., "A model of phenotypic susceptibility to tuberculosis: deficient in silico selection of *Mycobacterium tuberculosis* epitopes by HLA alleles," *Sarcoidosis Vasculitis and Diffuse Lung Diseases*, vol. 25, no. 1, pp. 21–28, 2008.
- [76] J. L. Flynn, "Immunology of tuberculosis and implications in vaccine development," *Tuberculosis*, vol. 84, no. 1-2, pp. 93–101, 2004.
- [77] J. L. Flynn and J. Chan, "Immunology of tuberculosis," *Annual Review of Immunology*, vol. 19, pp. 93–129, 2001.
- [78] D. A. Lewinsohn, E. Winata, G. M. Swarbrick et al., "Immuno-dominant tuberculosis CD8 antigens preferentially restricted by HLA-B," *PLoS Pathogens*, vol. 3, no. 9, article e127, 2007.
- [79] P. T. Tan, A. T. Heiny, O. Miotto et al., "Conservation and diversity of influenza A H1N1 HLA-restricted T cell epitope candidates for epitope-based vaccines," *PLOS ONE*, vol. 5, no. 1, Article ID e8754, 2010.
- [80] S. P. Ribeiro, D. S. Rosa, S. G. Fonseca et al., "A vaccine encoding conserved promiscuous HIV CD4 epitopes induces broad T cell responses in mice transgenic to multiple common HLA class II molecules," *PLoS ONE*, vol. 5, no. 6, Article ID e11072, 2010.

- [81] X. Lin, A. Sun, P. Ruan, Z. Zhang, and J. Yan, "Characterization of conserved combined T and B cell epitopes in leptospira interrogans major outer membrane proteins OmpL1 and LipL41," *BMC Microbiology*, vol. 11, article 21, 2011.
- [82] M. M. Johnson and J. A. Odell, "Nontuberculous mycobacterial pulmonary infections," *Journal of Thoracic Disease*, vol. 6, no. 3, pp. 210–220, 2014.
- [83] T. K. Marras, P. Chedore, A. M. Ying, and F. Jamieson, "Isolation prevalence of pulmonary non-tuberculous mycobacteria in Ontario, 1997–2003," *Thorax*, vol. 62, no. 8, pp. 661–666, 2007.
- [84] N. Martín-Casabona, A. R. Bahrmand, J. Bennedsen et al., "Non-tuberculous mycobacteria: patterns of isolation. A multi-country retrospective survey," *International Journal of Tuberculosis and Lung Disease*, vol. 8, no. 10, pp. 1186–1193, 2004.
- [85] J. Van Ingen, S. A. Bendien, W. C. M. De Lange et al., "Clinical relevance of non-tuberculous mycobacteria isolated in the Nijmegen-Arnhem region, The Netherlands," *Thorax*, vol. 64, no. 6, pp. 502–506, 2009.
- [86] J. E. Moore, M. E. Kruijshaar, L. P. Ormerod, F. Drobniewski, and I. Abubakar, "Increasing reports of non-tuberculous mycobacteria in England, Wales and Northern Ireland, 1995–2006," *BMC Public Health*, vol. 10, no. 1, article 612, 2010.
- [87] W.-J. Koh, O. J. Kwon, K. Jeon et al., "Clinical significance of nontuberculous mycobacteria isolated from respiratory specimens in Korea," *Chest*, vol. 129, no. 2, pp. 341–348, 2006.
- [88] V. M. Katoch, "Infections due to non-tuberculous mycobacteria (NTM)," *Indian Journal of Medical Research*, vol. 120, no. 4, pp. 290–304, 2004.
- [89] R. J. Wallace Jr., R. O'Brien, J. Glassroth, J. Raleigh, and A. Dutt, "Diagnosis and treatment of disease caused by nontuberculous mycobacteria," *American Review of Respiratory Disease*, vol. 156, pp. S1–S25, 2012.
- [90] H. C. Poyntz, E. Stylianou, K. L. Griffiths, L. Marsay, A. M. Checkley, and H. McShane, "Non-tuberculous mycobacteria have diverse effects on BCG efficacy against *Mycobacterium tuberculosis*," *Tuberculosis*, vol. 94, no. 3, pp. 226–237, 2014.
- [91] C. B. Do and K. Katoh, "Protein multiple sequence alignment," in *Functional Proteomics: Methods and Protocols*, pp. 379–413, Springer, 2008.
- [92] B. J. McFarland and C. Beeson, "Binding interactions between peptides and proteins of the class II Major histocompatibility complex," *Medicinal Research Reviews*, vol. 22, no. 2, pp. 168–203, 2002.
- [93] A. J. Sant, C. Beeson, B. McFarland et al., "Individual hydrogen bonds play a critical role in MHC class II: peptide interactions: implications for the dynamic aspects of class II trafficking and DM-mediated peptide exchange," *Immunological Reviews*, vol. 172, pp. 239–253, 1999.
- [94] V. Brusic, G. Rudy, M. Honeyman, J. Hammer, and L. Harrison, "Prediction of MHC class II-binding peptides using an evolutionary algorithm and artificial neural network," *Bioinformatics*, vol. 14, no. 2, pp. 121–130, 1998.
- [95] Z. Zavala-Ruiz, I. Strug, M. W. Anderson, J. Gorski, and L. J. Stern, "A polymorphic pocket at the P10 position contributes to peptide binding specificity in class II MHC proteins," *Chemistry & Biology*, vol. 11, no. 10, pp. 1395–1402, 2004.
- [96] P. Y. Arnold, N. L. La Gruta, T. Miller et al., "The majority of immunogenic epitopes generate CD4+ T cells that are dependent on MHC class II-bound peptide-flanking residues," *The Journal of Immunology*, vol. 169, no. 2, pp. 739–749, 2002.
- [97] A. J. Godkin, K. J. Smith, A. Willis et al., "Naturally processed HLA class II peptides reveal highly conserved immunogenic flanking region sequence preferences that reflect antigen processing rather than peptide-MHC interactions," *Journal of Immunology*, vol. 166, no. 11, pp. 6720–6727, 2001.
- [98] T. G. Kimman, R. J. Vandebriel, and B. Hoebee, "Genetic variation in the response to vaccination," *Community Genetics*, vol. 10, no. 4, pp. 201–217, 2007.