

Immunoinformatics Prediction of Epitope Based Peptide Vaccine Against *Mycobacterium Tuberculosis* PPE65 family Protein

Mustafa Elhag^{1,2}, Anfal Osama Mohamed Sati², Moaaz Mohammed Saadaldin^{3,2}, Mohammed A.Hassan⁴

¹Faculty of Medicine, University of Seychelles-American Institute of Medicine, Seychelles

²Africa City of Technology, Sudan

³Department of Biotechnology, Faculty of Science and Technology, Omdurman Islamic University, Sudan

⁴Department of Bioinformatics, DETAGEN Genetics Diagnostic Center, Kayseri, Turkey

Abstract

Introduction: Tuberculosis (TB) is a serious disease with varying rates of mortality and morbidity among infected individuals which estimates for approximately two million deaths/year. The number of deaths could increase by 60% if left untreated. It mainly affects immune-compromised individuals and people of third world, due to poverty, low health standards, and inadequate medical care. It has varying range of manifestations that is affected by the host immune system response, the strain causing the infection, its virulence, and transmissibility.

Materials and methods: A total of 1750 *Mycobacterium Tuberculosis* PPE65 family protein strains were retrieved from National Center for Biotechnology Information (NCBI) database on March 2019 and several tools were used for the analysis of the T- and B-cell peptides and homology modelling.

Results and conclusion: Four strong epitope candidates had been predicted in this study for having good binding affinity to HLA alleles, good global population coverage percentages. These peptides are YAGPGSGPM, AELDASVAM, GRAFNFAAPRYGFK and a single B-cell peptide YAGP.

This study uses immunoinformatics approach for the design of peptide based vaccines for *M. tuberculosis*. Peptide based vaccines are safer, more stable and less hazardous/allergenic when compared to conventional vaccines. In addition, peptide vaccines are less labouring, time consuming and cost efficient. The only weakness is the need to introduce an adjuvant to increase immunogenic stimulation of the vaccine recipient.

Keywords: *Immunoinformatics, Mycobacterium Tuberculosis, PPE65 family protein, Peptide vaccine, Epitope.*

INTRODUCTION

Tuberculosis (TB) is a serious disease with varying rates of mortality and morbidity among infected individuals which estimates for approximately two million deaths/year. The number of deaths could increase by 60% if left untreated. [1][2][3] It mainly affects immune-compromised individuals and people of third world, due to poverty, low health standards, and inadequate medical care. [4][5][6] It has varying range of manifestations that's affected by the host immune system response, the strain causing the infection, its virulence, and transmissibility. [7][8] These manifestations include but are not limited to: Constitutional symptoms for all forms of TB include fever, night sweats, and failure to thrive (children) or weight loss (adults). [9] Millitary tuberculosis, skeletal deformities, central nervous system involvement (tuberculosis meningitis TBM) and other organs involvement i.e., cervical lymphadenitis. [10][11][12][13]The main site of infection which follows the inhalation of aerosols or droplets containing *Mycobacterium Tuberculosis* is the alveoli in lungs where they enter mononuclear cells and interact with them following phagocytosis, causing damage, notably necrosis, cavitation, and coughing of blood. [5][14][15][16][17]

Mycobacterium Tuberculosis is a distinctive Acid fast bacilli that is obligate pathogens. [18]They differ from other bacteria in that a very large portion of its coding region is devoted to the production of lipogenesis and lipolysis enzymes, and to two new families of glycine-rich proteins with a repetitive structure that may represent a source of antigenic variation. [19]*M. tuberculosis* contains around 4,000 genes, one of which is the Proline Glutamic acid PPE gene, which comes from PE/PPE/PE-PGRS gene family which is present

exclusively in genus *Mycobacterium*, accounts for approximately 10% of the *M. tuberculosis* genome and plays a role in the makeup of cell surface markers thus, helping in genetic variations, adhesion and invasion of host defense cells. [20][21][22][23][24] Studies suggest that variations in these genes may underlie the basis of virulence attenuation. [24][25]

One of the *M. tuberculosis* PPE proteins, the PPE65 (*Rv3621c*) harbored in the Region of Difference (RD8) is suggested to be potential virulence determinant. [24] It's found to be overexpressed in macrophages infected with various clinical isolates of *M. tuberculosis* and it's particularly, a good B-cell antigen with a higher antibody response. [23][26][27][28] PPE65 is thought to be a suppressor of pro-inflammatory cytokines such as TNF- α and IL-6 while it also induces high expression of anti-inflammatory IL-10 co-operated by PE32 protein causing inhibition of T-helper 1 cells. [29]

Morbidity and mortality rates of TB steadily dropped due to, improved public health practices and widespread use of the Bacille Calmette–Guérin (BCG) vaccine, as well as the development of antibiotics. [7] Then numbers of new cases started increasing. Hence, homelessness and poverty increased and the AIDS emerged, with its destruction of the cell-mediated immune response in co-infected persons. [12][30][31][32] This global crisis was followed by the emergence of multidrug resistance and extensively drug resistant strains in countries like the former Soviet Union, South Africa, and India, where some antibiotics are available in lower quality or are not used for a sufficient time to control the disease according to recommended regimens [5][33][34][35][36][37][38][39][40] In addition, BCG, the widely administered live attenuated vaccine against TB, is inconsistently effective in adults. [10][38][41][42] Therefore, development of new specific vaccine preventive against *M. tuberculosis*, is becoming a major priority of *M. tuberculosis* investigators.

This study aims to suggest new possible peptides for a peptide- based vaccine for *M. tuberculosis*, targeting PPE65 protein as an immunogen using an immunoinformatics approach. Reverse vaccinology (peptide based vaccines) is becoming more favoured over conventional vaccine methodologies. This is owes to the fact that reverse vaccine techniques are much less costing, more time and labour saving and makes patients less prone to hazards. [43][44][45] peptide vaccines are based on the chemical approach to synthesize the computationally predicted suitable B-cell and T-cell epitopes that are immunodominant and can induce specific immune responses. [43] this study differs from other studies for being the first study to use PPE65 protein as an immunogenic target for a reverse vaccine-design based study.

MATERIALS AND METHODS

Protein Sequence Retrieval

A total of 1750 *Mycobacterium Tuberculosis* PPE65 family protein strains were retrieved from National Center for Biotechnology Information (NCBI) database on March 2019 in FASTA format. These strains were submitted from different parts of the world and we collected them for immunoinformatics analysis. The retrieved protein strains had length of 284 base pairs with name “PPE65 family protein”

Determination of conserved regions

The retrieved sequences of *Mycobacterium Tuberculosis* PPE65 family protein were subjected to multiple sequence alignment (MSA) using ClustalW tool of BioEdit Sequence Alignment Editor Software version 7.2.5 to determine the conserved regions. Molecular weight and amino acid composition of the protein were also obtained using this tool. [46][47]

Sequenced-Based Method

The reference sequence (WP_096833455.1) of *Mycobacterium Tuberculosis* PPE65 family protein was submitted to different prediction tools at the Immune Epitope Database (IEDB) analysis resource (<http://www.iedb.org/>) to predict various B and T cell epitopes. Conserved epitopes would be considered as candidate epitopes for B and T cells and were subjected for further analysis. [48]

B- Cell Epitope Prediction

B cell epitopes is the portion of the vaccine that interacts with B lymphocytes. Candidate epitopes were analysed using several B cell prediction methods from IEDB (<http://tools.iedb.org/bcell/>), to identify the surface accessibility, antigenicity and hydrophilicity with the aid of random forest algorithm, a form of unsupervised learning. The Bepipred Linear Epitope Prediction 2 was used to predict linear B cell epitope with default threshold value 0.533 (<http://tools.iedb.org/bcell/result/>). The Emini Surface Accessibility Prediction tool was used to detect the surface accessibility with default threshold value 1.000 (<http://tools.iedb.org/bcell/result/>). The Kolaskar and Tongaonker Antigenicity method was used to identify the antigenicity sites of candidate epitope with default threshold value 1.032 (<http://tools.iedb.org/bcell/result/>). The Parker Hydrophilicity Prediction tool was used to identify the hydrophilic, accessible, or mobile regions with default threshold value 1.695.[49-53]

T- Cell Epitope Prediction MHC Class I Binding

T cell epitopes is the portion of the vaccine that interacts with T lymphocytes. Analysis of peptide binding to the MHC (Major Histocompatibility complex) class I molecule was assessed by the IEDB MHC I prediction tool (<http://tools.iedb.org/mhci/>) to predict cytotoxic T cell epitopes. The presentation of peptide complex to T lymphocyte undergoes several steps. Artificial Neural Network (ANN) 4.0 prediction method was used to predict the binding affinity. Before the prediction, all human allele lengths were selected and set to 9 amino acids. The half-maximal inhibitory concentration (IC50) value required for all conserved epitopes to bind at score less than 100 were selected.[43, 54-59]

T- Cell Epitope Prediction MHC Class II Binding

Prediction of T cell epitopes interacting with MHC Class II was assessed by the IEDB MHC II prediction tool (<http://tools.iedb.org/mhcii/>) for helper T cells. Human allele references set were used to determine the interaction potentials of T cell epitopes and MHC Class II allele (HLA DR, DP and DQ). NN-align method was used to predict the binding affinity. IC50 values at score less than 500 were selected.[60-63]

Population Coverage

In IEDB, the population coverage link was selected to analyse the epitopes. This tool calculates the fraction of individuals predicted to respond to a given set of epitopes with known MHC restrictions (<http://tools.iedb.org/population/iedbinput>). The appropriate checkbox for calculation was checked based on MHC- I, MHC- II separately and combination of both.[64]

Homology Modelling

The 3D structure requires PDB ID which was obtained using raptorX (<http://raptorx.uchicago.edu>) i.e. a protein structure prediction server developed by Xu group, excelling at predicting 3D structures for protein sequences without close homologs in the Protein Data Bank (PDB). USCF chimera (version 1.13.1rc) was the program used for visualization and analysis of molecular structure of the promising epitopes (<http://www.cgl.uscf.edu/chimera>).[65, 66]

RESULTS

Multiple Sequence Alignment

The conserved regions and amino acid composition for the reference sequence of *Mycobacterium Tuberculosis* PPE65 family protein are illustrated in figure 1 and 2 respectively. Glycine and alanine were the most frequent amino acids (Table 1).

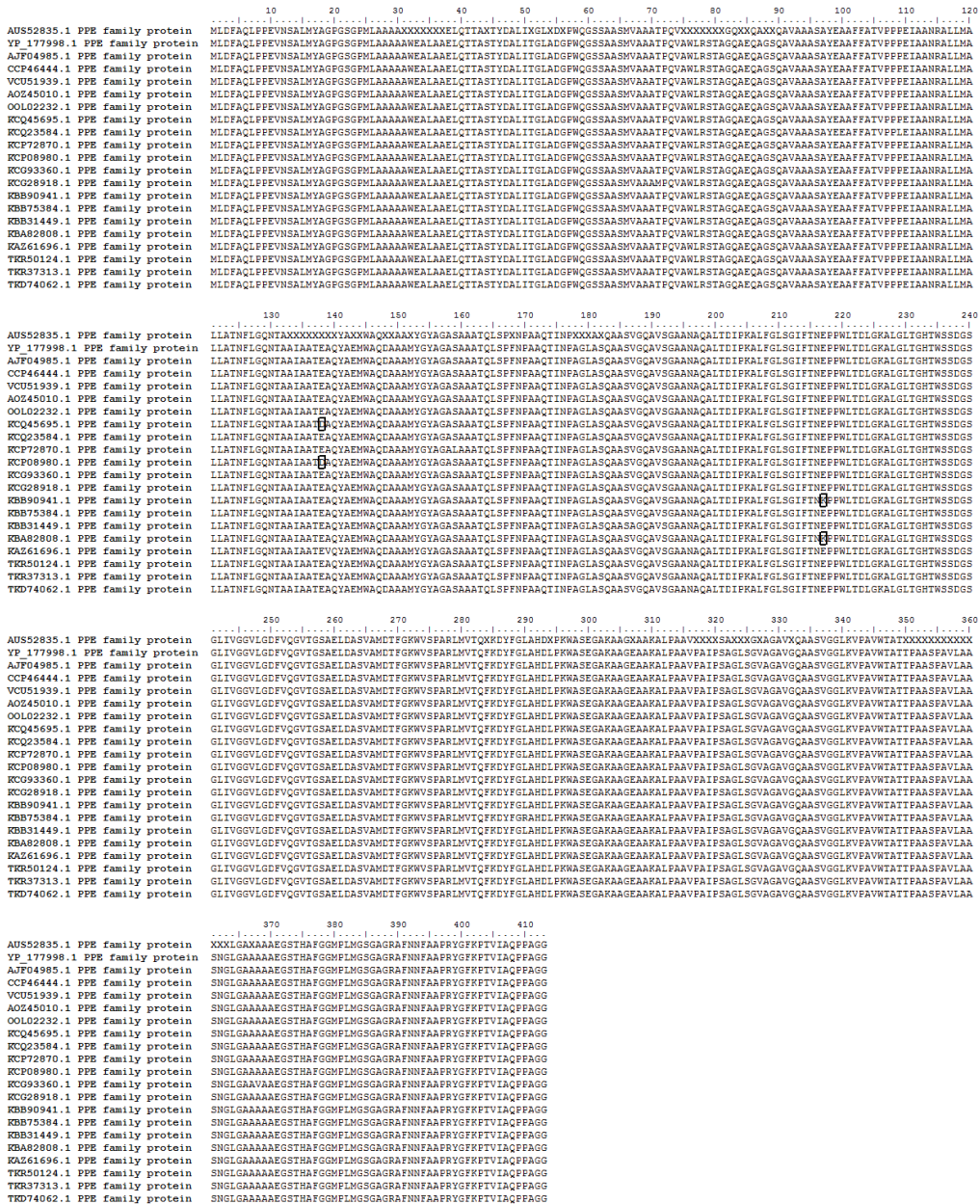


Figure 1: Multiple Sequence Alignment for partial sequence of the PPE65 protein in *M. tuberculosis*, using BioEdit software.

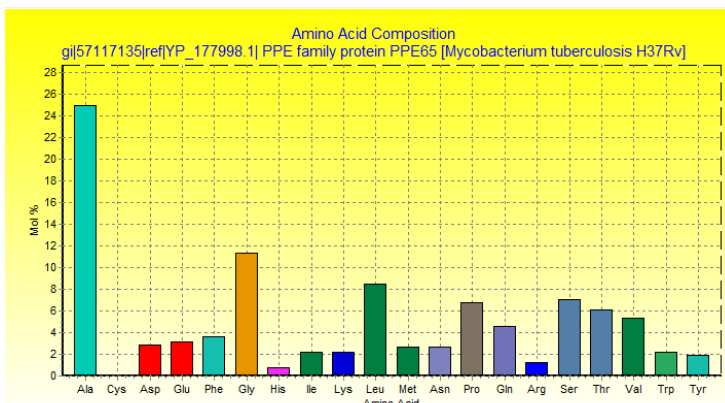


Figure 2: Amino acid composition for *Mycobacterium Tuberculosis* PPE65 family protein using BioEdit software.

Protein: gi|57117135|ref|YP_177998.1| PPE family protein PPE65 [Mycobacterium tuberculosis H37Rv]
 Length = 413 amino acids
 Molecular Weight = 40677.26 Daltons

Amino Acid	Number	Mol%
Ala	103	24.94
Cys	0	0
Asp	12	2.91
Glu	13	3.15
Phe	15	3.63
Gly	47	11.38
His	3	0.73
Ile	9	2.18
Lys	9	2.18
Leu	35	8.47
Met	11	2.66
Asn	11	2.66
Pro	28	6.78
Gln	19	4.6
Arg	5	1.21
Ser	29	7.02
Thr	25	6.05
Val	22	5.33
Trp	9	2.18
Tyr	8	1.94

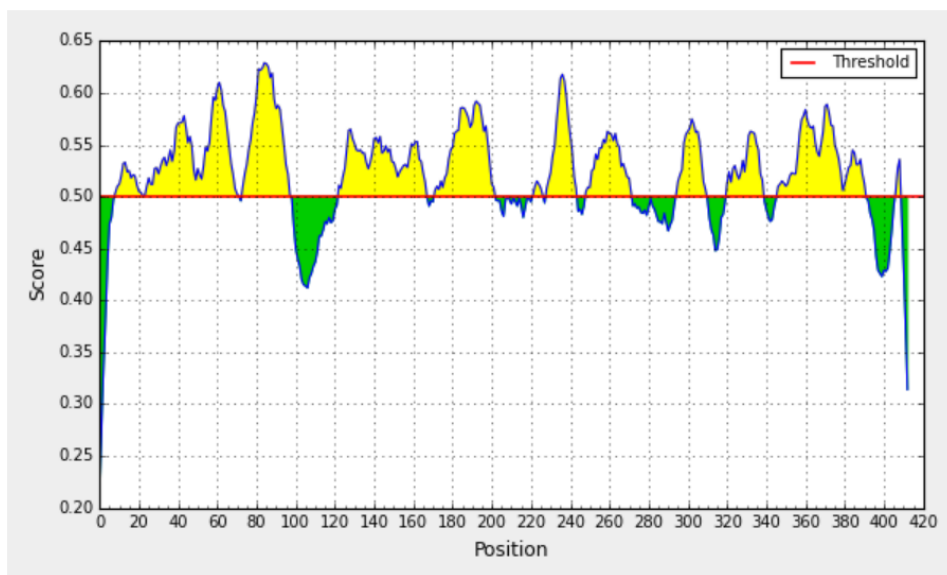
Table 1: Molecular weight and amino acid frequency distribution of the protein

B-cell epitope prediction

The reference sequence of *M. Tuberculosis* PPE65 was subjected to Bepipred linear epitope 2, Emini surface accessibility, Kolaskar & Tongaonkar antigenicity and Parker hydrophilicity prediction methods to test for various immunogenicity parameters (Table 2 and Figures 3-8). Two epitopes have successfully passed the three tests. Tertiary structure of the proposed B cell epitopes is illustrated (Figure 7 &8).

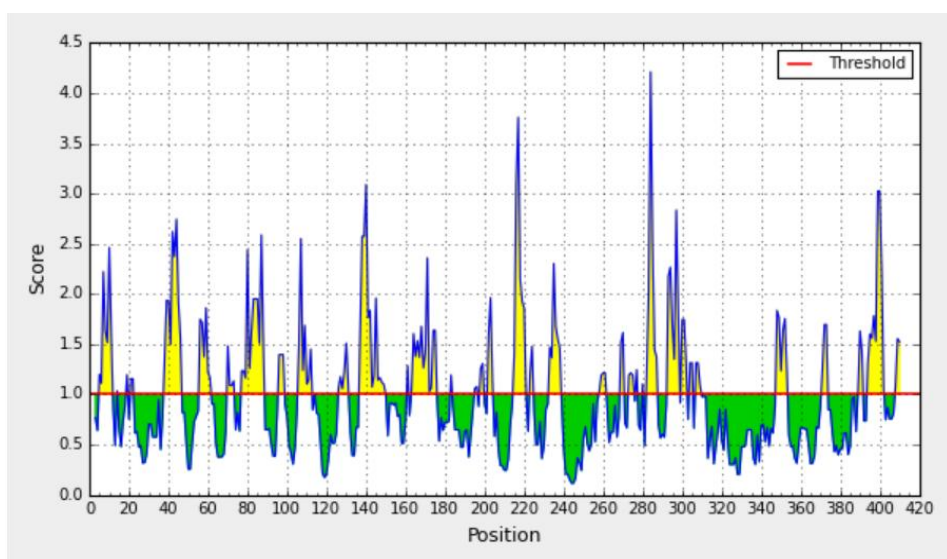
Peptide	Start	End	Length	Kolaskar & Tongaonkar antigenicity score (TH: 1.03)	Emini surface accessibility score (TH: 1)	Parker Hydrophilicity prediction score (TH: 1.43)	ALLERTOP
YAGP	17	20	4	1.041	1.24	2	Non-allergen
QAAS	183	186	4	1.039	1.212	4.175	Non-allergen
QAAS	333	336	4	1.039	1.212	4.175	Non-allergen

Table 2: Conserved B cell epitopes that had successfully passed the tests.



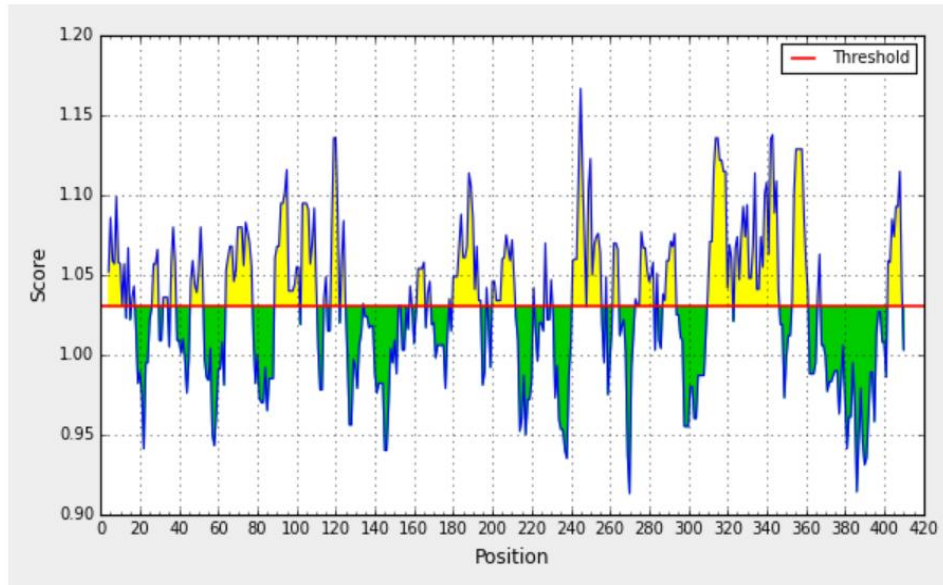
Average: 0.522 Minimum: 0.220 Maximum: 0.629

Figure 3: Bepipred Linear Epitope Prediction 2; Yellow areas above threshold (red line) are proposed to linearB- cell epitopes while the green areas are not.



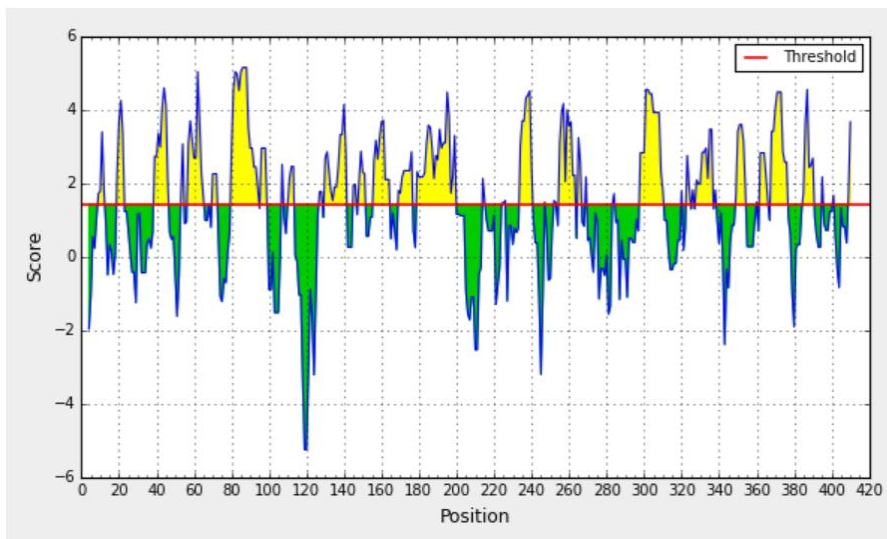
Average: 1.000 Minimum: 0.116 Maximum: 4.212

Figure 4: EMINI surface accessibility prediction; Yellow areas above the threshold (red line) are proposed to be a part of B cell epitopes and the green areas are not.



Average: 1.030 Minimum: 0.913 Maximum: 1.167

Figure 5: Kolaskar and Tonganokar antigenicity prediction; Yellow areas above the threshold (red line) are proposed to be antigenic B-cell epitopes and green areas are not.



Average: 1.430 Minimum: -5.257 Maximum: 5.171

Figure 6: Parker Hydrophilicity prediction; Yellow areas above the threshold (red line) are proposed to be hydrophilic B-cell epitopes and green areas are not.

Prediction of Cytotoxic T-lymphocyte epitopes

The reference sequence was analyzed using (IEDB) MHC-1 binding prediction tool to predict T cell epitopes interacting with different types of MHC Class I alleles. 61 peptides were predicted to interact with different MHC-I alleles. The most promising epitopes with their corresponding MHC-1 alleles and IC50 scores are shown in (Table 3) followed by the tertiary structure of the candidate T cell epitope (Figure 7).

Peptide	MHC 1 alleles	IC50
AELDASVAM*	HLA-B*40:01, HLA-B*44:02, HLA-B*40:02, HLA-B*44:03, HLA-B*18:01	7.85

FAAPRYGFK	HLA-A*68:01, HLA-C*03:03	12.86
GGLKVPVAVW	HLA-B*58:01, HLA-B*57:01	21.45
HAFGGMPLM	HLA-B*35:01, HLA-C*03:03, HLA-C*12:03, HLA-A*26:01	16.48
NRALLMALL	HLA-B*39:01, HLA-B*27:05	87.63
SVAMDTFGK	HLA-A*11:01, HLA-A*68:01	20.16
VAMDTFGKW	HLA-B*58:01, HLA-B*53:01	15.18
VLGDFVQGV	HLA-A*02:01, HLA-A*02:06	10.64
WVSPARLMV	HLA-A*68:02, HLA-A*02:06	20.77
YAGPGSGPM*	HLA-C*03:03, HLA-C*12:03, HLA-B*35:01, HLA-C*14:02, HLA-B*15:01	1.99

Table 3: The most promising T cell epitopes and their corresponding MHC-1 alleles.

Prediction of the Helper T-lymphocyte epitopes

Reference sequence was analysed using (IEDB) MHC-II binding prediction tool there were 85 predicted epitopes found to interact with MHC-II alleles. The most promising epitopes with their corresponding alleles and IC50 scores are shown in (Table 4) along with the 3D structure of the proposed epitope (Figure 10).

Peptide	MHC 2 alleles	IC50
AGRAFNNFAAPRYGF	HLA-DRB1*01:01, HLA-DRB1*09:01, HLA-DRB5*01:01, HLA-DRB1*07:01, HLA-DRB1*04:01, HLA-DRB1*15:01, HLA-DQA1*05:01/DQB1*03:01, HLA-DRB1*04:04, HLA-DRB1*04:05, HLA-DRB1*11:01, HLA-DQA1*01:02/DQB1*06:02	16.4
DTFGKWVSPARLMVT	HLA-DRB1*07:01, HLA-DRB1*09:01, HLA-DRB1*01:01, HLA-DQA1*05:01/DQB1*03:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*04:04, HLA-DRB1*11:01, HLA-DRB1*15:01, HLA-DRB5*01:01, HLA-DRB1*13:02, HLA-DPA1*02:01/DPB1*01:01	16
GRAFNNFAAPRYGFK*	HLA-DRB1*01:01, HLA-DRB5*01:01, HLA-DRB1*09:01, HLA-DRB1*04:01, HLA-DPA1*01:03/DPB1*02:01, HLA-DRB1*15:01, HLA-DRB1*07:01, HLA-DQA1*05:01/DQB1*03:01, HLA-DRB1*11:01, HLA-DRB1*04:04, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*04:05	10.1
RAFNNFAAPRYGFKP	HLA-DRB1*01:01, HLA-DRB5*01:01, HLA-DRB1*09:01, HLA-DPA1*01:03/DPB1*02:01, HLA-DQA1*05:01/DQB1*03:01, HLA-DRB1*04:01, HLA-DRB1*15:01, HLA-DRB1*07:01, HLA-DRB1*11:01, HLA-DRB1*04:04, HLA-DQA1*01:02/DQB1*06:02	14.1
TFGKWVSPARLMVTQ	HLA-DRB1*07:01, HLA-DRB1*09:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*01:01, HLA-DQA1*05:01/DQB1*03:01, HLA-DRB1*11:01, HLA-DRB5*01:01, HLA-DRB1*15:01, HLA-DRB1*04:04, HLA-DRB1*13:02, HLA-DPA1*03:01/DPB1*04:02, HLA-DPA1*02:01/DPB1*01:01	21.2

Table 4: The most promising T cell epitopes and their corresponding MHC-2 alleles.

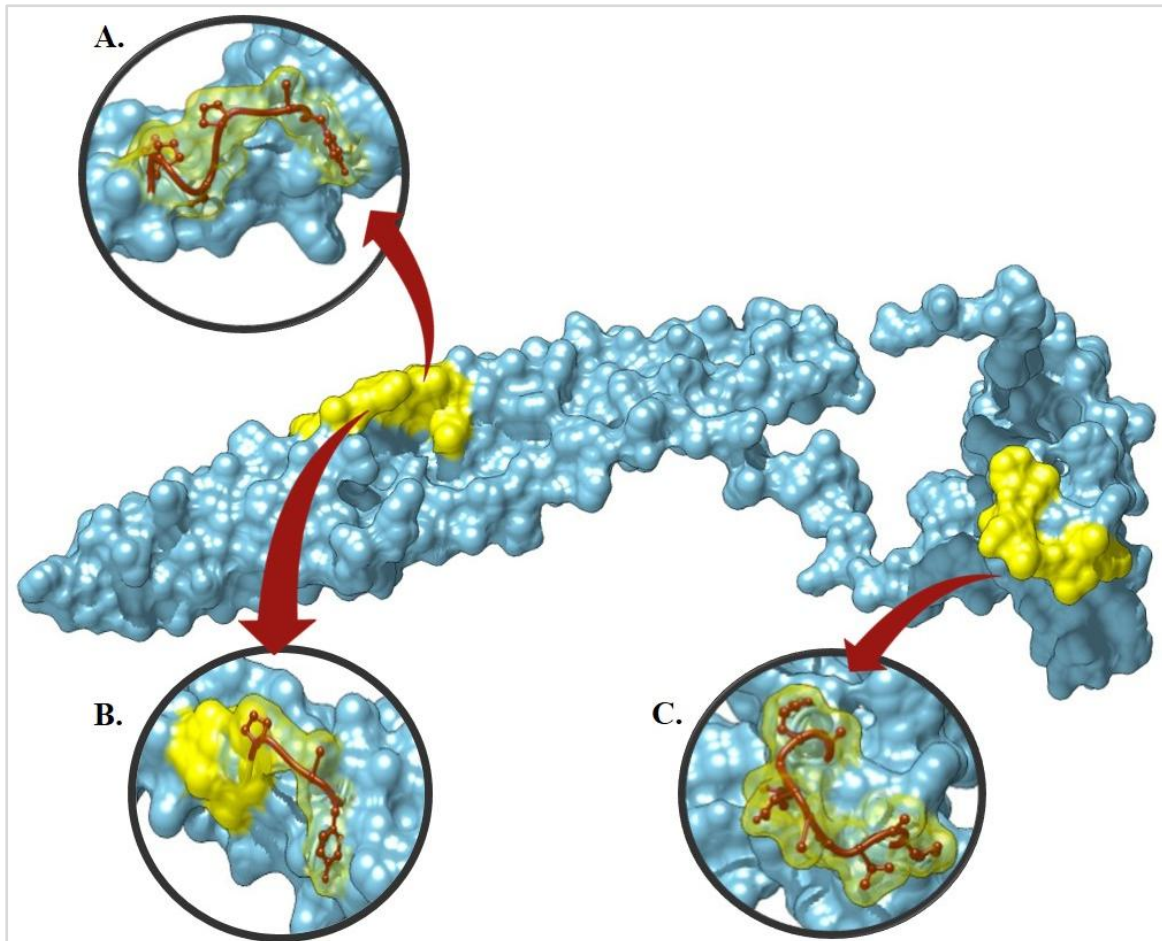


Figure 7. Three-dimensional visualisation of the most promising peptides in the study using Chimera (version 1.13.1rc) A. MHC-I binding T-cell peptide **YAGAPGSGPM** B. **YAGP** B-cell peptide which is a part of YAGPGSGPM and C. MHC-I binding T-cell peptide **AELDASVAM**.

Population Coverage Analysis:

All MHC I and MHC II epitopes were assessed for population coverage against the whole world using IEDB population coverage tool. For MHC I, epitopes with highest population coverage were VLGDFVQGV (40.6%) and YAGPGSGPM (33.81) (Figure 8 and Table 5). For MHC class II, the epitopes that showed highest population coverage were GRAFNNFAAPRYGFK (68.15%) and AGRAFNNFAAPRYGF (68.15%) (Figure 9 and Table 6). When combined together, the epitopes that showed highest population coverage were GRAFNNFAAPRYGFK (68.15) and AGRAFNNFAAPRYGF (68.15) (Figure10 and Table 7).

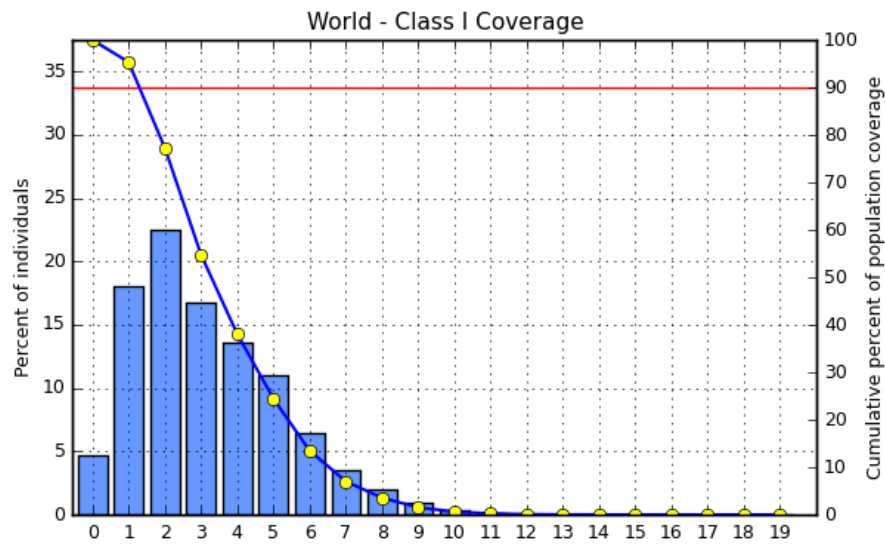


Figure 8. Population coverage for MHC class I epitopes.

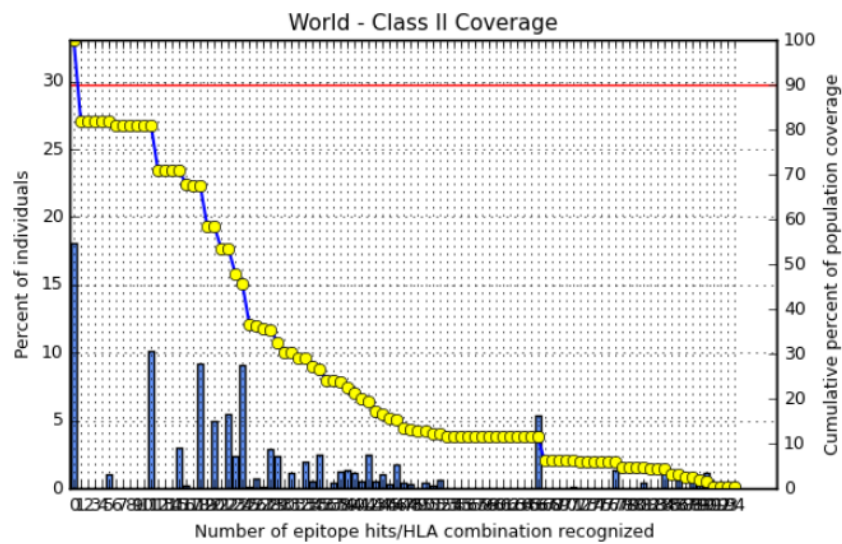


Figure 9. Population coverage for MHC class II epitopes.

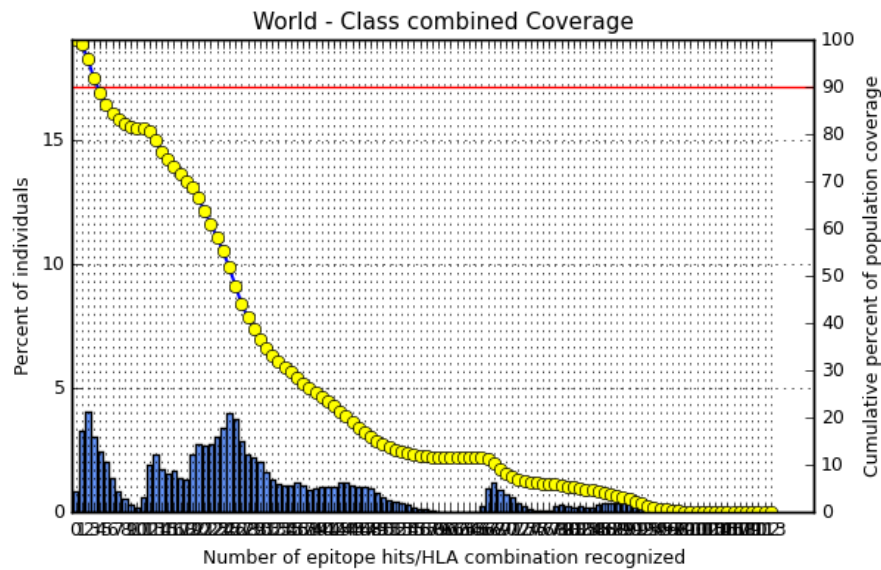


Figure 10. Population coverage for combined MHC I and II epitopes.

Epitope	Coverage Class 1 (%)	Total HLA hits
VLGDFVQGV	40.60%	2
YAGPGSGPM	33.81%	5
AELDASVAM	30.31%	5
HAFGGMPLM	29.26%	4
RYGFKPTVI	21.38%	1
SVAMDTFGK	20.88%	2
FAAPRYGFK	13.48%	2
APRYGFKPT	12.78%	1
ALFGLSGIF	8.44%	1
WQGSSAASM	8.44%	1

Table 5: Population coverage of proposed peptides interaction with MHC class I

Epitope	Coverage Class 2 (%)	Total HLA hits
GRAFNNFAAPRYGFK	68.15%	12
TFGKWVSPARLMVTQ	63.61%	12
AGRAFNNFAAPRYGF	68.15%	11
DTFGKWVSPARLMVT	66.41%	11
RAFNNFAAPRYGFKP	63.61%	11
AFNNFAAPRYGFKPT	63.61%	10
APRYGFKPTVIAQPP	63.56%	10
MDTFGKWVSPARLMV	57.83%	10
AAPRYGFKPTVIAQP	61.74%	9
AELDASVAMDTFGKW	61.74%	9

Table 6: Population coverage of proposed peptides interaction with MHC class II

Epitope	Coverage Class 1&2 combined (%)	Total HLA hits
AELDASVAM	68.15%	12
AFNNFAAPR	68.15%	11
AGRAFNNFA	66.41%	11
ALFGLSGIF	63.61%	12
APRYGFKPT	63.61%	11
EIAANRALL	63.61%	10
FAAPRYGFK	63.56%	10
FNNFAAPRY	61.74%	9
FVQGVTSAS	61.74%	9
GGLKVPVAVW	57.83%	10

Table 7: Population coverage of proposed peptide interaction with MHC class I & II combined

Country	MHC I	MHC II	MHC I&II combined
World	95.34 %	81.94 % *	99.16% *

Table 8: The population coverage of whole world for the epitope set for MHC I, MHC II and MHC I&II combined.

DISCUSSION

This study suggests four most promising peptide candidates for a reverse vaccine design of *M. tuberculosis* targeting PPE65 protein, including three T-cell peptides (YAGPGSGPM, AELDASVAM, GRAFNNFAAPRYGFK) and a single B-cell peptide YAGP. These peptides make the strongest candidates due to their relatively high global population coverage, scoring the lowest rates of IC50 with their corresponding HLA alleles, indicating strong interaction between the peptide and allele.

This study also has a major focus on the analysis of T-cell peptides. Hence, tuberculosis is mainly maintained from dissemination by cell-mediated immunity.[67] CD4+ T cells are the principal antigen-specific cells responsible for containment of *M. tuberculosis* infection, although they can also be major contributors to disease during *M. tuberculosis* infection in several cases.[68]

For MHC-I, the length of peptides is strictly nine amino acids thus all peptides are nine amino acid long. Peptide YAGPGSGPM, is the most promising peptide in this study. It binds to five MHC-I alleles, (HLA-C*03:03, HLA-C*12:03, HLA-B*35:01, HLA-C*14:02, and HLA-B*15:01). Besides, scoring the lowest IC50 of 1.99 when binding to HLA-C*03:03 indicating very strong interaction. Moreover, the peptide's global population coverage was the highest among the candidate peptides and was estimated for 33.81%. Another good finding which strengthens the point of selection of this particular peptide is that, a part of it YAGP peptide during the analysis of linear B-cell peptides was predicted to elicit a good immune response according to antigenicity, surface accessibility, hydrophilicity and linearity tests. This indicates that YAGPGSGPM is at least partially located on the surface of the protein and that it's immunogenic stimulating both B and T-cell responses.

Peptide AELDASVAM is also a strong possible peptide for MHC-I related immune responses. Not only does the peptide bind to five HLA alleles (HLA-B*40:01, HLA-B*44:02, HLA-B*40:02, HLA-B*44:03, HLA-B*18:01), but it has a very low IC50 of 7.85 to HLA-B*40:01 allele. The peptide is predicted to cover 30.31% of the total global population.

As for MHC-II alleles, the proposed alleles were longer than nine peptides to give more space for manipulation during in vivo and in vitro analysis. GRAFNNFAAPRYGFK peptide is the most promising peptide among the suggested peptides. It binds to a vast number of MHC-II alleles, approximately 12 alleles. In addition, it scored lowest IC50 of 10.1 to HLA-DRB1*01:01 and the highest population coverage with a score of 68.15%.

In this study, twenty-three peptides were shared in eliciting both MHC-I and MHC-II immune responses. Among these peptides were YAGPGSGPM, and AELDASVAM which proves the remarkable efficiency of the selected epitopes. This is advantageous when designing a peptide-based vaccine as it reduces the vast number of selected epitopes.

Different peptide based vaccine approaches for TB have been reported previously. A study by Rai PK et al. 2017 has reported that a lipidated peptide was introduced (L91) and engaged in an animal model trials to detect the efficiency of the peptide in activating T- helper cells. The result then confirmed that L91 peptide led to enhanced immune response to BCG vaccine.[69] A study by Coppola et al. 2015, suggests that a long synthetic peptide derived from Latency antigen Rv1733c of *M. tuberculosis* which is highly expressed in patients with dormant TB, can efficiently elicit an immune response, particularly T-cell response. The result was validated by using animal models for in vivo response.[70]

In this study, and besides being strictly computational, another limitation was found in population coverage analysis of MHC II; a total of twelve alleles did not give predictions. This includes (HLA-DQA1*05:01/DQB1*03:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DPA1*01/DPB1*04:01, HLA-DPA1*01:03/DPB1*02:01, HLA-DQA1*03:01/DQB1*03:02, HLA-DRB3*01:01, HLA-DRB4*01:01, HLA-DRB5*01:01, HLA-DQA1*05:01/DQB1*02:01, HLA-DPA1*03:01/DPB1*04:02, HLA-DQA1*04:01/DQB1*04:02, and HLA-DPA1*02:01/DPB1*01:01). In addition, the tertiary structure of the PPE65 protein was incomplete and this resulted in failure of visualization of the promising MHC-II peptide GRAFNNFAAPRYGFK. We recommend future researches conducting an in vivo and in vitro analysis of the proposed peptides and a population coverage analysis of the HLA alleles that did not give result along with homology modelling for PPE65 protein.

Peptide vaccines are an attractive alternative to conventional vaccines, as their concept relies on usage of short peptide fragments to induce a highly targeted immune responses, consequently avoiding allergenic and/or reactogenic sequences. [44]Reverse vaccine technology has proved its efficiency as many of the designed vaccines have made it into clinical trials. For example, a multiple peptides vaccine derived from tumor-associated antigens (TAAs) made it to phase II clinical trial for head and neck squamous cell cancer (HNSCC). A phase I trial was registered with University Hospital Medical Information Network (UMIN) number UMIN000002022, with a fixed 2-mg dose of KIF20A and VEGFR1 peptides for patients with advanced unresectable pancreatic cancer targeting nineteen patients sub grouped into the HLA-A*2402-positive group and HLA-A*2402-negative group respectively. Results of this study showed that vaccination with KIF20A and VEGFR1 peptides was a safe treatment, and might also be a promising treatment as HLA-A*2402-positive group had an increased survival rate in comparison to the negative group.[71]

In general, peptide based vaccines are safer, more stable with a minimal exposure to hazardous or allergenic substances compared to conventional vaccines. Moreover, these peptide vaccines are less labouring, time consuming and cost efficient. The only weakness regarding them is the need to introduce an adjuvant to increase immunogenic stimulation of the vaccine recipient. [44][72][73][74]

CONCLUSION

Since the need for a proper vaccine for M. tuberculosis is increasing with time, along with the reduced response to BCG vaccine, and growing population rates, poverty and pollution. This study suggests four peptides suitable for further analysis using reverse vaccinology techniques. YAGPGSGPM, AELDASVAM, and GRAFNNFAAPRYGFK peptides have acceptable global population coverage percentages, good binding affinity to MHC alleles.

References:

1. Kaye, K. and T.R. Frieden, Tuberculosis Control: the Relevance of Classic Principles in an Era of Acquired Immunodeficiency Syndrome and Multidrug Resistance. *Epidemiologic Reviews*, 1996. 18(1): p. 52-63.
2. Dolin, P.J., M.C. Raviglione, and A. Kochi, Global tuberculosis incidence and mortality during 1990-2000. *Bulletin of the World Health Organization*, 1994. 72(2): p. 213-20.
3. Dye, C., et al., Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. *Jama*, 1999. 282(7): p. 677-686.
4. Waaier, H.T., Tuberculosis and poverty. *The International Journal of Tuberculosis and Lung Disease*, 2002. 6(9): p. 745-746.
5. Marais, B.J. and H.S. Schaaf, Tuberculosis in children. *Cold Spring Harbor perspectives in medicine*, 2014. 4(9): p. a017855-a017855.
6. Colditz, G.A., et al., Efficacy of BCG vaccine in the prevention of tuberculosis: meta-analysis of the published literature. *Jama*, 1994. 271(9): p. 698-702.
7. Smith, I., Mycobacterium tuberculosis pathogenesis and molecular determinants of virulence. *Clinical microbiology reviews*, 2003. 16(3): p. 463-96.
8. Hesselting, A.C., et al., A critical review of diagnostic approaches used in the diagnosis of childhood tuberculosis. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*, 2002. 6(12): p. 1038-45.
9. Marais, B.J., et al., The natural history of childhood intra-thoracic tuberculosis: a critical review of literature from the pre-chemotherapy era. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*, 2004. 8(4): p. 392-402.
10. Thwaites, G.E., R. van Toorn, and J. Schoeman, Tuberculous meningitis: more questions, still too few answers. *The Lancet Neurology*, 2013. 12(10): p. 999-1010.
11. Awasthi, S., European Journal of Molecular Biology and Biochemistry PATHOGENESIS AND EFFECT OF BCG VACCINATION AGAINST MYCO BACTERIUM TUBERCULOSIS IN ABDOMEN. *European Journal of Molecular Biology and Biochemistry*, 2016. 3(3): p. 124-128.
12. Marais, B.J., et al., The spectrum of disease in children treated for tuberculosis in a highly endemic area. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*, 2006. 10(7): p. 732-8.
13. Marais, B.J., et al., Tuberculous Lymphadenitis as a Cause of Persistent Cervical Lymphadenopathy in Children From a Tuberculosis-Endemic Area. *The Pediatric Infectious Disease Journal*, 2006. 25(2): p. 142-146.
14. Marais, B.J., et al., A proposed radiological classification of childhood intra-thoracic tuberculosis. *Pediatric Radiology*, 2004. 34(11): p. 886-894.
15. Ernst, J.D., Macrophage receptors for Mycobacterium tuberculosis. *Infection and immunity*, 1998. 66(4): p. 1277-1281.
16. Bermudez, L.E. and J. Goodman, Mycobacterium tuberculosis invades and replicates within type II alveolar cells. *Infection and immunity*, 1996. 64(4): p. 1400-6.
17. Mehta, P.K., et al., Comparison of in vitro models for the study of Mycobacterium tuberculosis invasion and intracellular replication. *Infection and immunity*, 1996. 64(7): p. 2673-9.
18. Weiner, J. and S.H.E. Kaufmann, Recent advances towards tuberculosis control: vaccines and biomarkers. *Journal of Internal Medicine*, 2014. 275(5): p. 467-480.
19. Cole, S.T., et al., Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence. *Nature*, 1998. 393(6685): p. 537-544.
20. Kohli, S., et al., Comparative genomic and proteomic analyses of PE/PPE multigene family of Mycobacterium tuberculosis H37Rv and H37Ra reveal novel and interesting differences with implications in virulence. *Nucleic acids research*, 2012. 40(15): p. 7113-7122.
21. Tundup, S., et al., The co-operonic PE25/PPE41 protein complex of Mycobacterium tuberculosis elicits increased humoral and cell mediated immune response. *PloS one*, 2008. 3(10): p. e3586.
22. Strong, M., et al., Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from Mycobacterium tuberculosis. *Proceedings of the National Academy of Sciences*, 2006. 103(21): p. 8060-8065.
23. Choudhary, R.K., et al., PPE antigen Rv2430c of Mycobacterium tuberculosis induces a strong B-cell response. *Infection and Immunity*, 2003. 71(11): p. 6338-6343.
24. Mohareer, K., S. Tundup, and S.E. Hasnain, Transcriptional regulation of Mycobacterium tuberculosis PE/PPE genes: a molecular switch to virulence. *Journal of molecular microbiology and biotechnology*, 2011. 21(3-4): p. 97-109.
25. Zheng, H., et al., Genetic Basis of Virulence Attenuation Revealed by Comparative Genomic Analysis of Mycobacterium tuberculosis Strain H37Ra versus H37Rv. *PLoS ONE*, 2008. 3(6): p. e2375-e2375.
26. Qureshi, R., et al., PPE65 of M. tuberculosis regulate pro-inflammatory signalling through LRR domains of Toll like receptor-2. *Biochemical and Biophysical Research Communications*, 2019. 508(1): p. 152-158.
27. Homolka, S., et al., Functional Genetic Diversity among Mycobacterium tuberculosis Complex Clinical Isolates: Delineation of Conserved Core and Lineage-Specific Transcriptomes during Intracellular Survival. *PLoS Pathogens*, 2010. 6(7): p. e1000988-e1000988.

28. Monterrubio-López, G.P., J.A. González-Y-Merchand, and R.M. Ribas-Aparicio, Identification of Novel Potential Vaccine Candidates against Tuberculosis Based on Reverse Vaccinology. *BioMed Research International*, 2015. 2015: p. 1-16.
29. Khubaib, M., et al., Mycobacterium tuberculosis Co-operonic PE32/PPE65 Proteins Alter Host Immune Responses by Hampering Th1 Response. *Frontiers in Microbiology*, 2016. 7: p. 719-719.
30. Bruchfeld, J., M. Correia-Neves, and G. Källenius, Tuberculosis and HIV Coinfection: Table 1. *Cold Spring Harbor Perspectives in Medicine*, 2015. 5(7): p. a017871-a017871.
31. Lawn, S.D., et al., Impact of HIV Infection on the Epidemiology of Tuberculosis in a Peri-Urban Community in South Africa: The Need for Age-Specific Interventions. *Clinical Infectious Diseases*, 2006. 42(7): p. 1040-1047.
32. Hesseling, A.C., et al., High Incidence of Tuberculosis among HIV- Infected Infants: Evidence from a South African Population- Based Study Highlights the Need for Improved Tuberculosis Control Strategies. *Clinical Infectious Diseases*, 2009. 48(1): p. 108-114.
33. Iseman, M.D., Evolution of drug-resistant tuberculosis: a tale of two species. *Proceedings of the National Academy of Sciences*, 1994. 91(7): p. 2428-2429.
34. Schaaf, H.S., et al., Culture-confirmed childhood tuberculosis in Cape Town, South Africa: a review of 596 cases. *BMC Infectious Diseases*, 2007. 7(1): p. 140-140.
35. Zignol, M., et al., Multidrug-resistant tuberculosis in children: evidence from global surveillance. *European Respiratory Journal*, 2013. 42(3): p. 701-707.
36. O'BRIEN, R., Scientific Blueprint for Tuberculosis Drug Development (Global Alliance for TB Drug Development). *Tuberculosis*, 2001. 81(1): p. 1-52.
37. Smith, C.V., V. Sharma, and J.C. Sacchettini, TB drug discovery: addressing issues of persistence and resistance. *Tuberculosis*, 2004. 84(1-2): p. 45-55.
38. McKinney, J.D., W.R. Jacobs, and B.R. Bloom, 3 Persisting problems in tuberculosis. *Biomedical Research Reports*, 1998. 1: p. 51-146.
39. Abubakar, I., et al., Drug-resistant tuberculosis: time for visionary political leadership. *The Lancet Infectious Diseases*, 2013. 13(6): p. 529-539.
40. Seung, K.J., S. Keshavjee, and M.L. Rich, Multidrug-resistant tuberculosis and extensively drug-resistant tuberculosis. *Cold Spring Harbor perspectives in medicine*, 2015. 5(9): p. a017863.
41. Buddle, B.M., et al., Overview of Vaccination Trials for Control of Tuberculosis in Cattle, Wildlife and Humans. *Transboundary and Emerging Diseases*, 2013. 60: p. 136-146.
42. Trunz, B.B., P.E.M. Fine, and C. Dye, Effect of BCG vaccination on childhood tuberculous meningitis and miliary tuberculosis worldwide: a meta-analysis and assessment of cost-effectiveness. *The Lancet*, 2006. 367(9517): p. 1173-1180.
43. Patronov, A. and I. Doytchinova, T-cell epitope vaccine design by immunoinformatics. *Open biology*, 2013. 3(1): p. 120139.
44. Li, W., et al., Peptide Vaccine: Progress and Challenges. *Vaccines*, 2014. 2(3): p. 515-536.
45. Reche, P., et al., Peptide-based immunotherapeutics and vaccines 2015. *Journal of immunology research*, 2015. 2015.
46. Hall, T., I. Biosciences, and C. Carlsbad, BioEdit: an important software for molecular biology. *GERF Bull Biosci*, 2011. 2(1): p. 60-61.
47. Zhang, G. and B. Fang, A uniform design- based back propagation neural network model for amino acid composition and optimal pH in G/11 xylanase. *Journal of Chemical Technology & Biotechnology: International Research in Process, Environmental & Clean Technology*, 2006. 81(7): p. 1185-1189.
48. Vita, R., et al., The immune epitope database (IEDB) 3.0. *Nucleic acids research*, 2014. 43(D1): p. D405-D412.
49. Hall, T.A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. in *Nucleic acids symposium series*. 1999. [London]: Information Retrieval Ltd., c1979-c2000.
50. Larsen, J.E.P., O. Lund, and M. Nielsen, Improved method for predicting linear B-cell epitopes. *Immunome research*, 2006. 2(1): p. 2.
51. Emini, E.A., et al., Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. *Journal of virology*, 1985. 55(3): p. 836-839.
52. Kolaskar, A. and P.C. Tongaonkar, A semi-empirical method for prediction of antigenic determinants on protein antigens. *FEBS letters*, 1990. 276(1-2): p. 172-174.
53. Parker, J.M., D. Guo, and R.S. Hodges, New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. *Biochemistry*, 1986. 25(19): p. 5425-32.
54. Andreatta, M. and M. Nielsen, Gapped sequence alignment using artificial neural networks: application to the MHC class I system. *Bioinformatics*, 2015. 32(4): p. 511-517.
55. Buus, S., et al., Sensitive quantitative predictions of peptide- MHC binding by a 'Query by Committee' artificial neural network approach. *Tissue antigens*, 2003. 62(5): p. 378-384.
56. Nielsen, M., et al., Reliable prediction of T- cell epitopes using neural networks with novel sequence representations. *Protein Science*, 2003. 12(5): p. 1007-1017.
57. Peters, B. and A. Sette, Generating quantitative models describing the sequence specificity of biological processes with the stabilized matrix method. *BMC bioinformatics*, 2005. 6(1): p. 132.
58. Lundegaard, C., et al., NetMHC-3.0: accurate web accessible predictions of human, mouse and monkey MHC class I affinities for peptides of length 8–11. *Nucleic acids research*, 2008. 36(suppl_2): p. W509-W512.
59. Abdelbagi, M., et al., Immunoinformatics prediction of peptide-based vaccine against african horse sickness virus. *Immunome Research*, 2017. 13(135): p. 2.

60. Wang, P., et al., A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS computational biology*, 2008. 4(4): p. e1000048.
61. Wang, P., et al., Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC bioinformatics*, 2010. 11(1): p. 568.
62. Kim, Y., et al., Immune epitope database analysis resource. *Nucleic acids research*, 2012. 40(W1): p. W525-W530.
63. Nielsen, M. and O. Lund, NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction. *BMC bioinformatics*, 2009. 10(1): p. 296.
64. Bui, H.-H., et al., Predicting population coverage of T-cell epitope-based diagnostics and vaccines. *BMC bioinformatics*, 2006. 7(1): p. 153.
65. Pettersen, E.F., et al., UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*, 2004. 25(13): p. 1605-1612.
66. Peng, J. and J. Xu, RaptorX: exploiting structure information for protein alignment by statistical inference. *Proteins: Structure, Function, and Bioinformatics*, 2011. 79(S10): p. 161-171.
67. Cooper, A.M., Cell-mediated immune responses in tuberculosis. *Annual review of immunology*, 2009. 27: p. 393-422.
68. Mayer-Barber, K.D. and D.L. Barber, Innate and Adaptive Cellular Immune Responses to Mycobacterium tuberculosis Infection. *Cold Spring Harbor perspectives in medicine*, 2015. 5(12): p. a018424.
69. Rai, P.K., et al., A lipidated peptide of Mycobacterium tuberculosis resuscitates the protective efficacy of BCG vaccine by evoking memory T cell immunity. *Journal of translational medicine*, 2017. 15(1): p. 201-201.
70. Coppola, M., et al., Synthetic Long Peptide Derived from *Mycobacterium tuberculosis* Latency Antigen Rv1733c Protects against Tuberculosis. *Clinical and Vaccine Immunology*, 2015. 22(9): p. 1060.
71. Kato, J., A. Nagahara, and T. Kodani, Phase I Clinical Trial of Peptide Vaccination with KIF20A and VEGFR1 Epitope Peptides in Patients with Advanced Pancreatic Cancer. *Pancreatic Disorders and Therapy*, 2012. 02.
72. Obara, W., et al., Present status and future perspective of peptide-based vaccine therapy for urological cancer. *Cancer Science*, 2018. 109(3): p. 550-559.
73. Berger C.M., K.K.L., Salazar L.G., Kathy Schiffman PC., Disis M.L. , Peptide-Based Vaccines. In: Morse M.A., Clay T.M., Lyerly H.K. (eds) in *Handbook of Cancer Vaccines*. Cancer Drug Discovery and Development. . 2004, Humana Press, Totowa, NJ.
74. Slingsluff, C.L., Jr., The present and future of peptide vaccines for cancer: single or multiple, long or short, alone or in combination? *Cancer journal (Sudbury, Mass.)*, 2011. 17(5): p. 343-350.