

Insilico Prediction of B-Cell and T-Cell Epitopes for Chaperones from *Mycobacterium tuberculosis*

Madhu Pearl R¹, Shanmughavel Piramanayagam²

Computational Biology lab, Department of Bioinformatics, Bharathiar University, Coimbatore-46, India

Abstract: With *Mycobacterium tuberculosis* control becoming a global necessity and history of causing more death than HIV in 2015 crucial consideration is to be given to create an effective vaccine to cope up with the multi-drug resistant forms of *M. tuberculosis*. The only licensed TB vaccine *Bacillus Calmette-Guérin* (BCG) has drawbacks that limit its efficacy and applicability for which alternate approaches like peptide or epitope based vaccine that can offer safe and specific immune response without any cross reactivity also offering easy production except for an obstacle that for designing peptide vaccines a potential candidate selection from the hundreds or thousands of potential candidates in which many number of epitopes might be present in an antigenic protein but all may not antigenic and does not involve in immune response too. In the present study, two protein *fimC* and *papD* were selected. Antigenicity, B-cell and T-cell epitopes of these proteins were identified using Vaxijen, BCPred, ProPred I, ProPred. Binding affinity of the predicted epitope with MHC/HLA alleles were calculated using MHC Pred. Surface exposed epitope may be effectively involved in immune response which is predicted by PSORTb tool. Good binding epitopes were selected based on IC₅₀ value against DRB1*0101 and DRB1*0401 allele. The finally selected potential epitopes are IRVNNTSPY, YRPAGLSDR from *fimC* and one epitope LGSAPVLGY from *papD*. This study suggested that the predicted epitopes can elicit both B-Cell and T-cell mediated immune response, however experimental validation are required.

Keywords: Antibiotic, Drugs, MDR, Vaccine, Epitope, MHC/HLA allele, prediction

1. Introduction

Tuberculosis one of the important health hazards worldwide possess a major role in India as identified by WHO. The drug susceptibility and molecular characterization of drug resistant forms need to be explored. *M. tuberculosis* follows mutations in genes that encodes drug targets or drug activating enzymes which are responsible for drug resistance. The multi-drug resistance forms of Tuberculosis acts as the major threat to all TB control programs. To obstruct the multi drug resistance (MDR) and extensive drug resistance (XDR) of tuberculosis, Peptide-based epitope vaccines offer great advantage over whole cell vaccines; that avoids non-protective responses and unwanted side effects. Peptide based vaccines provides hope for treatment and prevention of chronic infectious disease that are caused by virus and cancer cells (Hailemichael, Y and Wang, C.Y., 2013)

Humoral and adaptive immune responses are mediated by B-cell and T-cell epitopes, of which the former are linear (Continuous) and conformational (Discontinuous). Conformational epitope identification is solely based on X-Ray structure of antigen-antibody complex which is a time taking step. Computational algorithms are available for linear B-cell epitope prediction (Barlow et al., 1986). MHC binding peptide is a prerequisite for immune response, the MHC Class I (recognize CD8 + T-cells), MHC class II (recognize CD4+ T-cells). For the recognition of peptide bounded with MHC on the Antigen Presenting cells (APCs) both Helper T-Lymphocytes (HTL) and Cytotoxic T-Lymphocytes (CTL) are required (Pancre et al., 1996). The epitope functional study aid in the immunopathogenesis study understanding. It also helps in the better designing of protective epitopes and better improved epitope or peptide designing. (PMID 24715888 – Functions and potentials of *M. tuberculosis* epitope). Epitope based designs takes

understanding a broader area of the multidrug resistance mechanism accelerating the vaccine productions.

In the present approach we have selected two chaperone protein *fimC* and *papD* from the *Mycobacterium tuberculosis* which can play role in Multi Drug resistance. Computational techniques were followed to predict novel peptide –epitope which can elicit both humoral and adaptive immune response.

2. Methodology

Antigenic properties of Target protein

The amino acid sequences of *Mycobacterium tuberculosis* chaperone protein *fimC* and *papD* were retrieved from UniprotKB database. Immunogenic properties of the retrieved sequences were identified by using Vaxijen 2.0 server with default value ≥ 0.4 to increase the prediction accuracy and to minimize the false positives. The sequences with vaxijen score ≥ 0.4 and were selected.

B-Cell and T-cell epitope prediction

The first screening steps were carried out in BCPred server with a cut-off score of 0.8 for 20-mer with vaxijen score ≥ 0.4 . The predicted antigenic B-cell epitopes were subjected to T-cell epitope prediction based on binding affinity with MHC class I and II molecules using Propred I (Singh, H. and Raghava, G. P. S. 2003) and Propred (Singh, H. and Raghava, G. P. S. 2001). MHC class II binding peptides were predicted from 51 alleles and MHC class I binding peptides were predicted from 47 alleles. Epitopes binded with MHC I and MHC II were selected as promiscuous 9-mer epitope. A second screening was based on epitope localization on target protein which is carried out by Pepitope server (Mayrose et al., 2007) a platform based on PepSurf (Mayrose et al., 2007) and Mapitope (Bublil et al., 2007) algorithms used for best cluster prediction based on

the structure based QSAR simulation. For the best epitope cluster, IC₅₀ values were calculated using MHCpred (Guan et al., 2003) to know its binding efficiency to the specific human alleles DRB1*0101 and DRB1*0401.

MDR protein modeling and Epitope characterization:

The protein structures were modeled using Phyre2 server and validated using PROCHECK. The modeled structure was then used in Pepitope server to predict the epitope localization on target protein.

3. Results

Protein localization & Antigenic character identification:

The *Mycobacterium tuberculosis* chaperone proteins fimC and papD were identified from Vaxijen score which is ≥0.4 that indicates the proteins are antigenic. The accession number, sequence length and Vaxijen score are shown in Table 1.

Table 1: Sequence length and antigenic properties of the fimC and papD protein

S.No	Protein Name	Uniprot ID	Sequence length	Vaxijen Score
1	fimC	A0A1K3GKW8	234	0.67
2	papD	A0A1K3CIH2	252	0.46

Linear B-cell epitope prediction

The number of identified B-cell epitopes from fimC and papD are 12 and 9. Among the identified B-cell epitopes, 10 were from fimC and 8 from papD were found to be antigenic based on Vaxijen score ≥ 0.4. The predicted antigenic B-cell epitopes of fimC and papD are represented in Table 2.

Table 2: Predicted B-cell epitopes of fimC and papD

Protein Name	Sequence position	B-cell epitopes	BCPred score	Vaxijen score
fimC	157	SFQSGSHSIRVNNTSPYHIT	0.992	1.2076
	68	PVTPFELTPPISRINPGANQ	0.973	0.7975
	112	LDIPPVSETSSSLKEKNSGQL	0.971	0.7899
	30	YPKGAHDITARINNESATAA	0.848	1.1629
	211	RGDHLTVSNINDQGGNVTF	0.793	1.2667
	57	PGDSEALPAKLPVTPFELTP	1	1.3646
	115	PPVSETSSSLKEKNSGQLAVR	1	1.1028
	139	VFYRPAGLSDRNTASQRVSF	1	0.8093
	30	YPKGAHDITARINNESATAA	1	1.1629
	163	HSIRVNNTSPYHITITEITL	1	1.1373
papD	164	TRQGDKYTVNNPTPYVVTIS	0.992	1.0702
	40	LNGSERAESVEISNQNKELP	0.963	0.9094
	216	SLGSAPVLGYINDYGGRPKM	0.88	0.4570
	157	WQESITLTRQGDKYTVNNPT	1	0.8977
	203	APNSSEVLGGNAASLGSAPV	1	0.9558
	34	DRTRVILNGSERAESVEISN	1	0.1780
	119	REIPPKTKDKANTLQIALQTR	1	0.6700
	180	VTISSASAAAYKKGDISGF	0.997	0.8034

T-cell epitope prediction

8 T-cell epitopes of fimC and 7 T-cell epitopes of papD were commonly binding with Class I and Class II MHC molecules. Among the identified commonly binding T-cell epitopes, 7 epitopes of fimC and 5 epitopes of papD were found to be antigenic based on Vaxijen score ≥ 0.4. The predicted antigenic T-cell epitope from fimC and papD are shown in Table 3. The predicted T-cell epitope binding with

the number of class I and Class II MHC molecules are represented in Fig 1, 2.

Table 3: T-cell epitope of fimC and papD binding with class I and class II MHC

Protein Name	Position	Common in ProPred I and ProPred	Vaxijen score
fimC	167	VNNTSPYHI	0.7834
	165	IRVNNTSPY	0.6557
	158	FQSGSHSIR	1.8821
	37	ITARINNES	1.4025
	41	INNESATAA	0.8367
	141	YRPAGLSDR	0.9376
	172	YHITITEIT	0.8859
papD	224	YINDYGGRP	1.0321
	217	LGSAPVLGY	0.9953
	208	VLGGNAASL	1.3725
	188	YKKGDISGF	0.9230
	180	VTISSASAA	0.8144

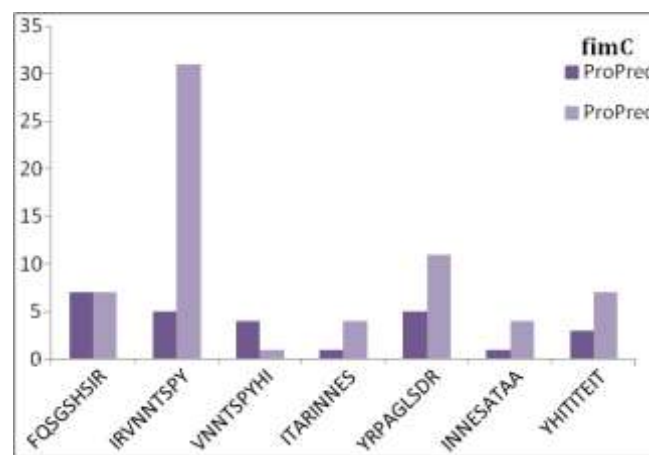


Figure 1: T-cell epitopes of fimC binding with Class I MHC and Class II MHC molecules

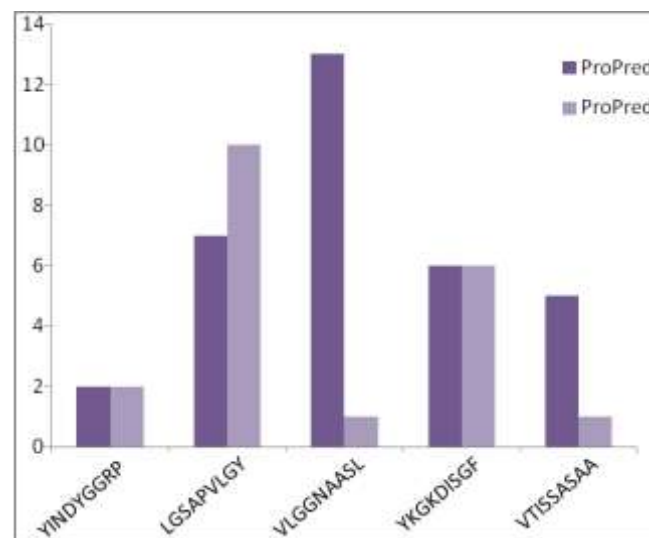


Figure 2: T-cell epitopes of papD binding with Class I MHC and Class II MHC molecules

Epitope characterization

To analyze the surface localization of the peptide to the respective target protein, 3-D structure of papD (PDB ID: 5K93) was retrieved from PDB. Three dimensional structure of fimC protein was modeled using Phyre2 server and validated using Ramachandran plot. More than 85% of the

amino acids of all the modeled protein were present in allowed region of the Ramachandran Plot.

The best epitope clusters were identified using P-epitope server. Among the identified clusters, epitope located on outer surface of the protein was selected. The number of short listed epitope from each protein includes 2 epitope from fimC whose cluster score is 272.45 and 5 epitope from papD which has a cluster score of 134.06. Epitope located with the surface of fimC and papD are shown in Fig 3.

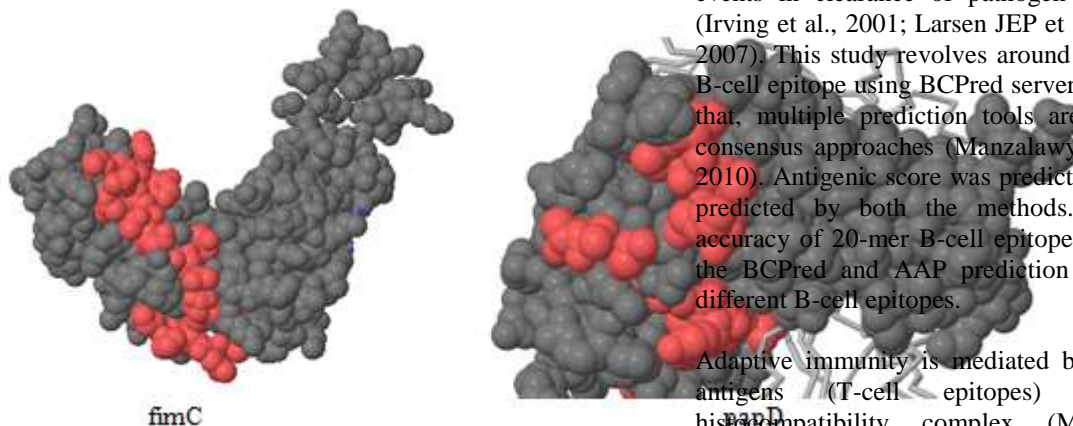


Figure 3: Surface localized epitopes of fimC and papD

Epitope validation

The potential T-cell epitopes of fimC are IRVNNTSPY , YRPAGLSDR and T-cell epitopes of papD are YINDYGGRP, LGSAPVLGY, VLGGNAASL, YKGKDISGF, VTISSASAA. The binding affinities of these potential T-cell epitopes were predicted against DRB1*0101 and DRB1*0401 allele. The final shortlisted epitopes based on IC₅₀ value are IRVNNTSPY, YRPAGLSDR from fimC and one epitope LGSAPVLGY from papD. The results are represented in Table 4.

Table 4: Binding affinity of predicted T-cell epitope with DRB1*0101 and DRB1*0401 allele

Protein Name	Sequence position	T-cell epitope	IC ₅₀ Value (nM)	
			DRB1*0101	DRB1*0401
fimC	165	IRVNNTSPY	8.11	341.98
	141	YRPAGLSDR	5.24	95.28
papD	224	YINDYGGRP	10.64	767.36
	217	LGSAPVLGY	42.95	325.09
	208	VLGGNAASL	229.09	1291.22
	188	YKGKDISGF	239.33	717.79
	180	VTISSASAA	76.03	762.08

4. Discussion

Experimental approaches of vaccine designing are costly and time consuming to determine the epitope which elicit humoral and T-cell immunity (Sakib et al., 2014). *In silico* prediction of epitopes from immunogenic protein a familiar tool that can save time from large scale screening of experimental approach (Li et al., 2005). Proteins located on the surface of pathogenic organism mainly interact with host immune system since surface proteins are highly immunogenic. (Da Dara et al., 2003; Tran et al., 2006). Identification of target antigen being the initial step in rational vaccine design, the surface exposed proteins become

more favorable for a potential vaccine (Almeida et al., 2003).

The protein sequence availed from UniprotKB database; provides information about protein localization. The proteins fimC and papD are found to be antigenic based on VaxiJen server whose prediction accuracy is nearly 70%-89 %. B-cell epitope prediction is essential, as the cells are the sites of antigen antibody interaction since antibodies are produced by B-lymphocytes. Antigen- antibody interactions are key events in clearance of pathogen by the immune system (Irving et al., 2001; Larsen JEP et al., 2006; Gershoni et al., 2007). This study revolves around the prediction of 20-mer B-cell epitope using BCPred server. Several studies reported that, multiple prediction tools are promising strategy for consensus approaches (Manzalawy et al., 2008; Barh et al 2010). Antigenic score was predicted for the B-cell epitopes predicted by both the methods. order to improve the accuracy of 20-mer B-cell epitope, two prediction methods the BCPred and AAP prediction were used which gives different B-cell epitopes.

Adaptive immunity is mediated by recognition of peptide antigens (T-cell epitopes) bound with major histocompatibility complex (MHC) molecules. The prediction of MHC peptide binding is lead priority for anticipating potential T-cell epitopes (Lafuente et al., 2009). The identification of both CD8 and CD4 T-cells are basic for the development of epitope based vaccines and in understanding the disease pathogenesis (Tchernev et al., 2006; Reche et al., 2006).

In the present study T-cell epitopes (9-mer) from antigenic B-cell epitope were predicted by using Propred I and Propred tool. Epitope which binds to MHC I and MHC II good antigenic score is selected as required epitopes.

Surface localization of peptide is one of the significant properties of peptide vaccine. We have predicted surface localization of T-cell epitope through P-epitope server. The modeled antigenic proteins were validated using PROCHECK. From this server, best cluster epitopes were selected and again its localization was cross checked with PSORTb 3.0.2 tool. In human, MHC molecules known as Human Leucocyte Antigen(HLA) whose polymorphic diversity in peptide binding specificity leads to effective immune response (Reche et al 2003; Kuhns et al., 1999). The Peptide binding to more than one HLA alleles is essential in vaccine development with impartial and extensive human population coverage. Selected peptide binding affinity predicted against human DRB1*0101 and DRB1*0401 HLA allele based on IC₅₀ value, in which (i) IC₅₀< 50 nM represent good binding. (ii) IC₅₀=50-100 nM represent intermediary binder. (iii) IC₅₀=500- 5000 nM represent weak binder. (iv) IC₅₀ > 5000 nM represent non binder. The predicted epitopes can induce both humoral and adaptive immune response but experimental approaches are required.

5. Conclusion

This study reported the epitopes from the chaperones of *Mycobacterium tuberculosis*. 3 potential epitopes from two

proteins were selected can be developed as peptide based vaccines in future against *Mycobacterium tuberculosis*. The epitopes predicted have the potential to elicit immune responses. Further evaluation of these in silico identified epitopes is highly warranted to confirm their immunogenicity and vaccine potential.

References

- [1] Almeida, S.V., DeMarco, R., Martins, E.A., 2003. Transcriptome analysis of the acoelomate human parasite *Schistosoma mansoni*. *Nature Genetics*, 35 (2), 148-157.
- [2] Barh, D., Misra, A.N., Kumar, A., 2010. In silico identification of dual ability of *N. gonorrhoeae* ddl for developing drug and vaccine against pathogenic *Neisseria* and other human pathogens. *J. Proteomics Bioinform.* 3, 82-90. doi:http://dx.doi.org/10.4172/jpb.1000125.
- [3] Barlow, D., Edwards, M., Thornton, J., 1986. Continuous and discontinuous protein antigenic determinants. *Nature* 322, 747-748.
- [4] Bublil EM, Freund NT, Mayrose I, Penn O, Roitburd-Berman A, Rubinstein ND, Pupko T, Gershoni JM. 2007. Stepwise prediction of conformational discontinuous B-cell epitopes using the Mapitope algorithm. *Proteins*. 68(1):294-304
- [5] Da Dara, A.A., Skelly, P.J., Walker, C.M., Harn, D.A., 2003. ADNA –Prime/protein boost vaccination regimen enhances Th2 immune responses but not protection following *Schistosoma mansoni* infection. *Parasite Immunol.* 25 (8-9), 429-437.
- [6] EL-Manzalawy, Y., Dobbs, D., Honavar, V., 2008. Predicting linear B-cell epitopes using string kernels. *J. Mol. Recognit.* 21, 243-255. doi:http://dx.doi.org/10.1002/jmr.893.
- [7] Gershoni, J.M., Roitburd –Berman, A., Simon-Tov, D.D., Tarnovitski Freund, N., Weiss, Y., 2007. Epitope mapping: the first step in developing epitope-based vaccines. *BioDrugs* 21, 145-156.
- [8] Guan P, Doytchinova IA, Zygori C, Flower DR. MHCpred: bringing a quantitative dimension to the online prediction of MHC binding. *Appl Bioinformatics*. 2003 2:63-66.
- [9] Hailemichael, Y., Overwijk, W.W., 2013. Peptide based anticancer vaccines: the making and unmaking of a T-cell graveyard. *Oncoimmunology* 2.
- [10] Irving, M.B., Pan, O., Scott, J.K., 2001. Random-peptide libraries and antigen-fragment libraries for epitope mapping and the development of vaccines and diagnostics. *Current Opinions in Chemical Biology* 5, 314-324.
- [11] Kuhns, J.J., Batalia, M.A., Shuquin, Y., and Collins, E.J., 1999. Poor binding of a Her-2/neu epitope (GP2) to HLA-A2.1 is due to a lack of interactions with the center of the peptide. *Journal of Biological Chemistry* 274, 36422-36127.
- [12] Lafuente, E.M., Reche, A.P., 2009. Prediction of MHC-Peptide binding: A systematic and comprehensive overview. *Current pharmaceutical design* 15, 3209-3220.
- [13] Larsen-Pontoppidan, J.E., Lund, O., Nielsen, M., 2006. Improved method for predicting linear B-cell epitopes. *Immunome research* 2, 1745-7580.
- [14] Li, G-F., Wang, Y., Zhang, Z.S., 2005. Identification of immunodominant Th1-type T cell epitopes from *Schistosoma japonicum* 28 kDa glutathione-S-transferase, a vaccine candidate. *Acta Biochimica et Biophysica Sinica*, 37 (11), 751-758.
- [15] Mayrose I, Penn O, Erez E, Rubinstein ND, Shlomi T, Tarnovitski Freund N, Bublil E, Rupin E, Sharan R, Gershoni JM, Martz E, Pupko T. 2007. Peptide: epitope mapping from affinity-selected peptides. *Bioinformatics* 23(23):3244-3246.
- [16] Mayrose I, Shlomi T, Rubinstein ND, Gershoni JM, Ruppin E, Sharan R, Pupko T. 2007. A graph-based algorithm for epitope mapping using combinatorial phage-display libraries. *Nucleic Acid Research*. 35(1):69-78
- [17] Pancre, V.H., Gras-Mass, A., Delanoye, J., Hernot, A.C., Auriault C., 1996. Induction of cytotoxic T-cell activity by the protective antigen of *Schistosoma mansoni* Sm28GST or its derived C-terminal lipopeptide. *Scand J Immunol* 44: 485-4.
- [18] Reche, P.A., Keskin, D.B., Hussey, R.E., Ancuta, P., Gabuzda, D., Reinherz, E.L., 2006. Elicitation from virus-naïve individuals of cytotoxic T lymphocytes directed against conserved HIV-1 epitopes. *Med Immunol* 5, 1.
- [19] Reche, P.A., Reinherz, E.L., 2003. Sequence variability analysis of human class I and class II MHC molecules functional and structural correlates of amino acid polymorphisms. *J Mol Biol* 331, 623-41.
- [20] Sakib S.M, Islam R.Md, Hasan A.K.M, Nabi A.H.M.N. 2014. Prediction of Epitope-based peptides for the utility of vaccine development from fusion and glycoprotein of Nipah Virus using *In Silico* approach. *Advances in Bioinformatics*. 402492
- [21] Singh, H. and Raghava, G. P. S. (2003) ProPred1: Prediction of promiscuous MHC class-I binding sites. *Bioinformatics* 22;19(8):1009-14
- [22] Singh, H. and Raghava, G.P.S. (2001) ProPred: Prediction of HLA-DR binding sites. *Bioinformatics*, 17(12), 1236-37.
- [23] Tchernev, G., Orfanos, C.E., 2006. Antigen mimicry, epitope spreading and the pathogenesis of pemphigus. *Tissue Antigens* 68, 280-6.
- [24] Tran, M.H., Pearson, M.S., Bethony, J.M., Smyth, D.J., Jones, M.K., Duke, M., Don, T.A., McManus, D.P., Correa-oliveria, R., Loukas, A., 2006. Tetraspanins on the surface of *Schistosoma mansoni* reprotective antigens against Schistosomiasis. *Microb.Infect.* 6.1125-1132.