INTERNATIONAL JOURNAL OF RESEARCHES IN BIOSCIENCES, AGRICULTURE AND TECHNOLOGY
© VISHWASHANTI MULTIPURPOSE SOCIETY (Global Peace Multipurpose Society) R. No. MH-659/13(N)
www.vmsindia.org

Computational prediction of potential vaccine target of leprosy by reverse vaccinology

A. N. Mahalley
Govt. Institute of Science, Nagpur
anitamahalley@gmail.com

Abstract

Leprosy, a communicable disease, is still prevalent in India. Though leprosy does not result into death, it can lead to numbness , paralysis, deformity and blindness, due to severe nerve damage . Vaccine Bacillus Calmette Guerin is given for leprosy as well as tuberculosis. B.C.G. provides incomplete protection against tuberculosis and leprosy. Increase in new cases of leprosy reported during door-to-door leprosy detection campaign in India, indicating the possible emergence of drug resistance in *M.leprae* (21).From the complete genome *M.leprae* TN , promiscuous epitope of penicillin binding protein with its potential to stimulate both the B-cells and T-cells by binding with MHC I and MHC II alleles, is predicted as potential vaccine target.

Keywords:

T-cell epitopes, B-cell epitopes, major histocompatibility proteins, immunogenicity ,Reverse vaccinology

Introduction

Mortality due to infectious diseases, is one of the major causes of death in developing countries. Newly emerged multidrug resistant, non cultivable microorganisms and pathogens with antigenic variations are challenges to the researchers (17). Diseases like small pox, polio and measles are successfully eradicated by vaccine but malaria, AIDS & tuberculosis and many noninfectious diseases including cancer, autoimmune diseases, and allergy still do not have effective and safe vaccines (27).

Leprosy is a neglected tropical disease. Increase in leprosy cases proves that it is difficult to eradicate the same (19). Leprae patient suffers from discrimination, stigmatization, isolation and destitution. *Mycobacterium leprae* the causative agent of leprosy is noncultivable, intracellular parasite that survive inside macrophages. The vaccine strain of BCG currently used to prevent mycobacterial infections is non-pathogenic (18).The most common BCG vaccine is less effective in MHC Class I mediated antigen presentation, which is a prerequisite in CD8+ T cell induction (20). A novel vaccine on leprosy is required to eradicate leprosy completely.

Empirical vaccine preparation is slow, time consuming and cost effective process. It fails to develop protective vaccines against nonculturable microbes and against pathogens with antigenic hypervariability . Vaccine containing killed or attenuated microorganisms may induce undesirable inflammatory response .There is a possibility that microorganisms may also revert virulent status through mutation. It ignores both pathogen and host variability (14).

To eliminate the potential undesirable effects **,** *In silico* methods are the practical alternatives (22)**.** Reverse Vaccinology is a vaccine development strategy that overcome the limits of the conventional vaccinology approaches (11). It enables to overcome the genetic variability of both pathogen and host. This approach is particularly useful for noncultivable as well as mutant microorganisms. The complete genome is screened computationally in short duration and at much lower cost. The unique , most conserved protective antigens, proteins not expressed in vitro, virulence proteins, surface proteins and immunome proteins can be searched , so as to avoid potential undesirable effects. Epitope based vaccines are better tolerated. Such predicted epitopes are important in immunodiagnosis and in antibody production (6).

In reverse vaccinology , the complete protein sequence database can be screened for every potential $CD4^+$ T cell and $CD8^+$ T cell epitope that can replace, an antigen in the process of either antibody production or detection.(27) Such a molecule can be synthesized or, in case of a protein, its gene can be cloned into an expression vector. A synthetic peptide may correspond to a short continuous stretch from a protein sequence and can bind an antibody raised against a protein. Designed synthetic peptide are inexpensive and noninfectious in contrast to viruses or bacteria.

When any foreign pathogen or antigen enters in body, it is engulfed by macrophage cells. These cells process the pathogen and present it's antigens to lymphocytes-T_C and T_H cells with the help of HLA receptors. CTLs recognize intracellular peptides presented by MHC class I molecules (CTL epitopes) and HTL recognize peptides from the extracellular space that are displayed by MHC class II molecules (HTLs epitopes). T_H cells stimulate B cells. Antibodies formed by Plasma B cells are responsible for humoral immunity. Memory B cells protect our body from further infection caused by same pathogen. In this

work, the promiscuous binders epitopes, responsible for cellular as well as humoral immunity are computationally predicted. In some cases, an identified T-cell epitope may also contain a B-cell epitope and also CTL-inducing epitope(s) (3). Many predicted T cell epitopes and B cell epitopes are possible vaccine candidate, can stimulate specific immune responses (15).

Bioinformatics tools that can accurately model the MHC:epitope interface and predict MHC binding vaccine targets**.** Epitope prediction software filters out potential B-cell and T-cell epitopes from the vast array of possibilities .

Reverse vaccinology which has application of bioinformatics at its first stage, was first used by Rino Rappuoli in the development of a vaccine against serogroup B *Neisseria meningitidis* (MenB) (7) .It was later on successfully applied to other pathogens, including *Bacillus anthracis* (2), *Porphyromonas gingivalis* (21) *Chlamydia pneumoniae* (22), *Streptococcus pneumoniae* (25) *and Mycobacterium tuberculosis* (10). Many epitope-based vaccines developed by reverse vaccinology , are in clinical trials. (9).

In this study promiscuous epitopes of *M.leprae* TN are predicted computationally.

Material and Methods.

1) Refseq protein sequences of M. *leprae* TN, are collected from NCBI Entrez protein database (ftp://ftp.ncbi.nlm.nih.gov/genomes/genbank/bacteria).
2) Complete Protein sequences of *M.leprae* TN are screened for their potential epitopes for vaccine by Vaxign .(8,11)
3) Antigenicity of proteins is predicted by antigenicity prediction server ,VaxiJen *v* 2.0.
4) Immunogenicity of predicted antigenic proteins with their immunogenic 20 mers peptide is predicted by Ig Pred server.(8)
5) From common antigenic proteins as well as immunogenic proteins the Linear B cell epitope 15 mers are predicted by a web server EPMLR (23).
6) Binding potential of EPMLR epitopes with HLA alleles is predicted by Immune Epitope Database (IEDB) and NetMHC server using ANNs**.**

(Web site-*http://www.cbs.dtu.dk/services/NetMHC (12)*

Result and Discussion.

Leprosy is one of the least understood human diseases . The treatment given in leprosy is multi-drug therapy (MDT). Both MB- and PB-MDT treatments are of long duration. Though MDT remains effective in the majority of cases, relapse (or possibly re-infection) can occur. (21). MDT efficacy will also disappear with the emergence of drug resistance. According to a report (published on 18 October 2016 |Geneva | New Delhi), on door-to-door leprosy detection campaign spearheaded by the National Leprosy Elimination Programme, (which involved volunteers from the Accredited Social Health Activists (ASHA) project, covering 149 districts across 19 states in 2015) a total of 127 326 new cases were detected (accounting for 60% of the global total of new cases), as compared with 125 785 new cases in 2014. Highest number of cases was detected in Bihar (4400 cases). Such widespread emergence of drug resistant leprosy demands a new effective leprosy vaccine. To improve leprosy control, an attempt is made to identify promiscuous multi epitopes for leprosy vaccine by using bioinformatics .

The criteria of ideal vaccine candidate includes , the localization of antigenic protein either exposed on the surface of the pathogen (especially adhesins) or secreted into the extracellular milieu, adhesion probability, transmembrane helices , nonhomology to human proteins , highly conserved antigens and Pfam domain**(**30). From the predictions of Vaxign Vaccine Design Program(28), six proteins are selected as vaccine target proteins for the study as given in following Tables.

Hypothetical proteins are not selected for further study.

Immunogenicity is the ability of a particular foreign substance, protein or its epitope, to stimulate an immune response in the body of host. When foreign substance or antigen is recognized by the body, the antigen presenting cells engulf and process it .Later on, present it to Cytotoxic T cells (Tc) on one hand and to T helper cells on the other hand with the help of HLA receptors. Activated T_H cells multiply and differentiates into plasma B cells and memory B cells. Plasma B cells secrets antibodies to attack the foreign cell and the protein. Memory B cells gives protection to the body during subsequent attack of same pathogen.

Immunogenicity of antigenic proteins of *M.leprae* TN is tested computationally by using **IgPred** server.

Table 1: Characterization of Predicted Vaccine Target of *Mycobacterium leprae TN* by Vaxign: Vaccine Target Prediction

S.N.	Protein Accession	Protein Note	Localization (Probability)	Adhesin Probability	Trans-membrane helices	pFAM Domain	Protein Length
1	**NP_302395.1**	PstS component of phosphate uptake	Unknown (Prob.=0.333)	0.651	0	PF01547.18	**369**
2	**NP_301391.1**	serine-rich protein	Unknown (Prob.=0.25)	0.607	0	PF12484.1	**408**
3	NP_302435.1	phosphate-binding protein 3	Cytoplasmic Membrane - (Prob.=0.955)	0.603	0	PF01547.18	**429**
4	NP_302091.1	penicillin binding protein	Cytoplasmic Membrane (Prob.=0.878)	0.296	1	PF0090.15	**608**
5	**NP_301163.1**	hypothetical protein ML0050	Extracellular (Prob.=0.972)	0.572	0	PF06013.5	**100**
6	NP_301670.1	hypothetical protein ML0885	Cytoplasmic Membrane (Prob.=0.88)	0.630	1	PF00877.12	**374**

Antigenicity property is important in selection of the best potential vaccine the candidates.(5)

Table 2: **Overall Prediction for the Antigenicity of target proteins VaxiJen**

	protein	**score**	
1	PstS	**0.7478**	Probable **ANTIGEN**
2	phosphate-binding protein 3	**0.6499**	Probable **ANTIGEN**
3	penicillin binding protein	**0.5351**	Probable **ANTIGEN**
4	serine-rich protein	**0.5033**	Probable **ANTIGEN**
5	hypothetical protein ML2312	**0.5302**	Probable **ANTIGEN**
6	periplasmic solute-binding proteins	**0.4546**	Probable **ANTIGEN**

Table3: Prediction of Immunogenicity of antigenic proteins of *M.leprae TN* by IgPred server (**www.imtech.res.in/raghava/igpred/**) for Prediction of Antibody-specific B-cell epitopes

	IgG Score	IgE Score	IgA Score	Prediction
PstS component of phosphate uptake	1.097 *	-0.130	0.709	IgG Epitope
serine-rich protein	1.329 *	-0.675	0.360	IgG Epitope
periplasmic solute-binding proteins	1.354 *	-0.517	0.364	IgG Epitope
phosphate-binding protein 3	1.246 *	-0.298	0.355	IgG Epitope
penicillin binding protein	1.381 *	-0.502	0.510	IgG Epitope

Among the tested five antigenic proteins, penicillin binding protein has the highest IgG score. IgG epitopes (above the threshold value 1) of these antigenic proteins are given in following table.

Table4: Immunogenic epitopes of antigenic proteins of *M.leprae TN* predicted by Ig Pred server.

S.N	Protein	Peptide	IgG Score	Location
1	PstS component of phosphate uptake	IHVVFRNDESGTTDNFQRYL	1.097	181-220
2	serine-rich protein	MFDFMVYSPEVNAFLMSRGP	1.329	1-20
3	serine-rich protein	VNAFLMSRGPGSTPLWGAAE	1.079	11-30
4	serine-rich protein	QLVSRYCMDRRDSVNSFHSS	1.080	174-193
5	serine-rich protein	LVSRYCMDRRDSVNSFHSSS	1.086	175-194
6	serine-rich protein	VSRYCMDRRDSVNSFHSSSS	1.084	176-195
7	serine-rich protein	EEHGSDSMSQSYNTCGSVAQ	1.088	214-224
8	Periplasmic solute binding protein	AYSLLASRGATTEGPADEHV	1.020	110-129
9	Periplasmic solute binding protein	THENENDPSAADMAAALNLI	1.045	215-234
10	phosphate-binding protein 3	GVRRKCAEKARSVNHDKVCR	1.062	44-63
11	phosphate -binding protein 3	GKRTMTAEGSTAQQNAIALF	1.159	107-126
12	phosphate -binding protein 3	KRTMTAEGSTAQQNAIALFN	1.100	108-127
13	phosphate -binding protein 3	VCSKGYDPDTFAAIKSFLTV	1.043	377-396
14	phosphate -binding protein 3	DVLAKIFSGVITTWNDGILA	1.039	211-230
15	penicillin binding protein	VEDTGTIAYRFTWHLPKNRT	1.066	91-110
16	penicillin binding protein	EDTGTIAYRFTWHLPKNRTW	1.217	92-111
17	penicillin binding protein	DTGTIAYRFTWHLPKNRTWS	1.300	93--112
18	penicillin binding protein	TGTIAYRFTWHLPKNRTWSY	1.312	94-113
19	penicillin binding protein	GTIAYRFTWHLPKNRTWSYD	1.257	95--114
20	penicillin binding protein	TIAYRFTWHLPKNRTWSYDG	1.424	96-115
21	penicillin binding protein	IAYRFTWHLPKNRTWSYDGQ	1.649	97--116
22	penicillin binding protein	AYRFTWHLPKNRTWSYDGQL	1.534	98-117
23	penicillin binding protein	YRFTWHLPKNRTWSYDGQLK	1.383	99-118
24	penicillin binding protein	RFTWHLPKNRTWSYDGQLKM	1.377	100-119
25	penicillin binding protein	FTWHLPKNRTWSYDGQLKMV	1.491	101-120
26	penicillin binding protein	TWHLPKNRTWSYDGQLKMVR	1.576	102-121
27	penicillin binding protein	WHLPKNRTWSYDGQLKMVRY	1.454	103-122
28	penicillin binding protein	HLPKNRTWSYDGQLKMVRYE	1.218	104-123
29	penicillin binding protein	NDRVASVIGRLPGVVVTLQA	1.032	230- 249
30	penicillin binding protein	RVVSVNRNGVDVAVLHEVEP	1.018	281-300
31	penicillin binding protein	VSVNRNGVDVAVLHEVEPSP	1.228	283-302
32	penicillin binding protein	SVNRNGVDVAVLHEVEPSPA	1.040	284-303
33	penicillin binding protein	VNRNGVDVAVLHEVEPSPAS	1.064	285-304
34	penicillin binding protein	NRNGVDVAVLHEVEPSPASS	1.007	286-305

According to Prof. R. Arnon (4) a successful synthetic vaccine should contain both B and T cell epitopes , and also CTL-inducing epitope (s) . When considering the B-cell antigens, as potential subunit vaccines, it also may be important to consider their T-cell epitope content since the antibody response is dependent upon the presence of T help. (1).

B-cell epitope is a region on the surface of an antigen that binds to paratope of an antibody. Most of the experimentally validated B-cell epitopes are in the range of 5–22 amino acid length (29).

From 25 mer immunogenic peptides VEDTGTIAYRFTWHLPKNRTWSYDGQLKMVRYE, location - 91—123) of penicillin binding protein predicted by IgG Pred server , 'fifteen ' mer epitopes are searched out by using B cell epitope prediction server EPLMR .(26)

For immunological applications, a minimum conserved sequence length of nine amino acids is required, as this represents the typical length of peptides that bind to HLA molecules. (4) From IgG pred epitopes, the sequence of nine amino acids are selected to test their potential as T cell epitopes.

Table 5 : EPLMR Predicted 15 mers Epitopes of Penicillin binding protein with immunogenic IgG pred peptide binding with MHCII allele of penicillin binding protein

Penicillin binding protein		
EPLMR prediction	Prediction of MHC II binding by NetMHCII 2.2	
Epitope 15 mer	binding core	MHC-II allele
EDTGTIAYRFTWHLP 92-106 Score - 0.90915 DTGTIAYRFTWHLPK 93-107 score-1	IAYRFTWHL 97-105	HLA-DRB10701 HLA-DRB11501 HLA-DRB50101

Table 6 : The binding potential of nine mers epitopes from immunogenic peptide (VEDTGTIAYRFTWHLPKNRTWSYDGQLKMVRYE , location 91—123) of penicillin binding protein with potential to bind MHC I alleles as per IEDB analysis .

S.N.	Peptide with location	Allele	Consensus Score
1	DTGTIAYRF 93 -101	HLA-B*51:01	31.0
		HLA-A*02:01	50.0
2	GTIAYRFTW 95 -103	HLA-B*40:01	28.0
		HLA-B*51:01	50.0
		HLA-A*02:06	17.5
3	TIAYRFTWH 96-103	HLA-B*40:01	47.0
		HLA-B*51:01	49.0
		HLA-A*02:01	50.0
4	IAYRFTWHL **97-104**	**HLA-B*51:01**	**1.9**
		HLA-A*02:06	**2.8**
		HLA-A*02:01	**4.1**
		HLA-B*40:01	**7.65**
5	AYRFTWHLP 98-106	HLA-B*40:01	49.5
6	YRFTWHLPK 99 -107	HLA-A*02:01	42.0
		HLA-B*51:01	40.0
		HLA-B*40:01	23.0
		HLA-A*02:06	54.0
7	RFTWHLPKN 100 -108	HLA-B*40:01	25.5
8	FTWHLPKNR 101 -109	HLA-B*51:01	37.0
		HLA-B*51:01	27.0

Ultimately from immunogenic sequence 'VEDTGTIAYRFTWHLPKNRTWSYDGQLKMVRYE' of peniciliin binding protein ,eight epitopes were predicted as T cell epitopes. Epitope 'IAYRFTWHL ' 97-104 has highest potential to bind with MHC I ALLELES- HLA-B*51:01, HLA-A*02:06, HLA-A*02:01 and HLA-B*40:01 . Similarly predicted by as NetMHCII 2.2, **'**IAYRFTWHL**'** has potential to bind with MHC II ALLELES- HLA-DRB10701 ,HLA-DRB11501 and HLA-DRB50101 with good binding score .

Conclusion : Reverse vaccinology technology has great potential to search most promising vaccine target in short duration. When applied to *M. leprae TN*, a nonculturable bacteria , reverse vaccinology has predicted common promiscuous B cell linear epitopes and CTL epitopes for Vaccine designing. Thus this promiscuous epitope' IAYRFTWHL' is predicted as potential vaccine target for leprosy.

References:

(1)Andras Falus (2008), Epitope based Immunome derived vaccines, Clinical Applications of Immunomics,Pp 43.

(2) Ariel N., Zvi A., Makarova K. S., Chitlaru T., Elhanany E., Velan B.,Cohen S., Friedlander A M and Shafferman A.,(2003): Genome-based bioinformatic selection of chromosomal Bacillus anthracis putative vaccine candidates coupled with proteomic identification of surface-associated antigens. *Infect. Immun.* 71:4563-4579. View this article via: Cross Ref PubMed ,Google Scholar *Neisseria meningitidis* (MenB)

(3) Arnon R .(2006) (2003) Old and new vaccine approaches,International Immunopharmacology 3 ,1195 – 1204

(4) Asif M. Khan, Olivo Miotto, A.T. Heiny, Jerome Salmon, K.N. Srinivasan, Eduardo Nascimento, Ernesto T. Marques, Vladimir Brusic, Tin Wee Tan, and J. Thomas, August (2006) , A systematic bioinformatics approach for selection of epitope-based vaccine targets, Cell Immunol.; 244(2): 141–147.

(5) Cafardi V. , Telford J, L., and Serruto D;(2012), Bacterial Genomes and Vaccine Design, Volume 5 of the series Immunomics Reviews: pp 13-37

(6) Chen J.,Liu H.,Yang J and Chou K.C. (2007):Prediction of linear B-cell epitopes using amino acid pair antigenicity scale,Amino Acids.,Sep;33(3):423-8. Epub 2007 Jan 26.

(7) Daniela R C., Telford J L., Rappuoli R, Seib K L. ,(2009), Vaccinology in the genome era, *J Clin Invest.* ;119(9):2515-2525. doi:10.1172/JCI38330.

(8) Doytchinova I A and Flower D R: (2007) VaxiJen: a serve for prediction of protective antigens,tumour antigens and subunit vaccines.BMC Bioinformatics. (2007)8:4

(9)Flower D R , Macdonald I K , Ramakrishnan K , Davies M N and Doytchinova I A (2010),Computer aided selection of candidate vaccine Antigens, Immunome Research 6 (Suppl 2):S1

(10) Gloria P. Monterrubio-López, Jorge A. González-Y-Merchand,and RosaMaría Ribas-Aparicio(2015), Identification of Novel Potential Vaccine Candidates against Tuberculosis Based on Reverse Vaccinology, BioMed Research International,Volume, Article ID 483150, 16 pages

(11) He Y, Xiang Z, Mobley HLT (2010), Vaxign: the first web-based vaccine design program for reverse vaccinology and an application for vaccine development. *Journal of Biomedicine and Biotechnology*. Volume 2010 Article ID 297505, 15 pages. [PMID: 20671958]

(12) Larsen MV, Lundergaard C, Lamberth K, Buus S, Lund O,Nietsen M. (2007),Large- scale validation of methods for cytotoxic T-lymphocyte epitope prediction.,BMC Bioinformatics. Oct 31; 8:424.

(13) Nielsen M, Lundegaard C, Worning P, Lauemoller SL, Lamberth K, Buus S, Brunak S, Lund O (2003), Reliable prediction of T-cell epitopes using neural networks with novel sequence representations. *Protein Sci* 12:1007-1017.

(14)Oberg AL, Kennedy RB, Li P, Ovsyannikova IG, Poland GA. (2011), "Systems Biology Approaches to New Vaccine Development." Curr Opin Immunol. Jun;23(3):436-43. doi: 10.1016/j.coi.2011.04.005. Epub 2011 May 11.

(15) Patronov A, Doytchinova I (2013) T-cell epitope vaccine design by immunoinformatics. Open Biol 2013;3:120139.

(16) Ponomarenko, J V., and Marc HV Van Regenmortel (2009): "B cell epitope prediction." *Structural bioinformatics* 849-879.

(17) RAPPUOLI R: Reverse vaccinology, a genome-based approach to vaccine development. *Vaccine (2001)*19(17-19):2688-2691.

(18) Report of the Scientific Working Group meeting on Leprosy (2002),Geneva, 26–28 November.

(19) Richardus JH , Habbema JD (2007),The impact of leprosy control on the transmission of M. leprae: is elimination being attained? Lepr Rev. Dec; 78(4):330-7.

(20) Rueckert C, Guzmán C A (2012) Vaccines: From Empirical Development to Rational Design, PLoS Pathog 8(11): e1003001. https://doi.org/10.1371/journal.ppat.1003001

(21) Ross BC, Czajkowski L, Hocking D, Margetts M, Webb E, Rothel L, Patterson M, Agius C, Camuglia S, Reynolds E, Littlejohn T, Gaeta B, Ng A, Kuczek ES, Mattick JS, Gearing D, Barr IG. (2001),Identification of vaccine candidate antigens from a genomic analysis of Porphyromonas gingivalis. *Vaccine*. 19:4135-4142.

(22) Seib KL, Dougan G, Rappuoli R. (2009),The key role of genomics in modern vaccine and drug design for emerging infectious diseases. PLoS Genet. 2009; 5:e1000612. [PMC free article] [PubMed]

(23) Singh H, Ansari HR, Raghava GPS (2013) Improved Method for Linear B-Cell Epitope Prediction Using Antigen's Primary Sequence. PLoS ONE 8(5): e62216. https://doi.org/10.1371/journal.pone.0062216.

(24) Tamara Davenne , and Helen McShane(2016)Why don't we have an effective tuberculosis vaccine yet?, Expert Rev Vaccines. Aug 2; 15(8): 1009–1013

(25) Wizemann TM, Heinrichs JH, Adamou JE, Erwin AL, Kunsch C, Choi GH, Barash SC, Rosen ,CA, Masure HR, Tuomanen E, Gayle A, Brewah YA, Walsh W, Barren P, Lathigra R, Hanson , M, Langermann S, Johnson

S, Koenig S.(2001), Use of a whole genome approach to identify vaccine molecules affording protection against Streptococcus pneumoniae infection. *Infect. Immun.* 69:1593-1598.

(26) Yao Lian,Meng Ge and Xian-Ming Pan(2014), EPMLR: sequence-based linear B-cell epitope prediction method using multiple linear regression, *BMC Bioinformatics* 2014 15:414

(27) Yongqun He, Rino Rappuoli, Anne S. De Groot, Robert T. Chen(2010), Emerging Vaccine Informatics, Journal of Biomedicine and Biotechnology,Volume 2010 ,Article ID 218590

(28)Yongqun He, Zuoshuang Xiang and Harry L. T. Mobley 2010 ,Vaxign: The First Web-Based Vaccine Design Program for Reverse Vaccinology and Applications for Vaccine Development. Journal of Biomedicine and Biotechnology, Volume 2010 (2010), Article ID 297505, 15 pages