Research Article

APPLICATION OF REVERSE VACCINOLOGY IN DESIGNING A COMMON VACCINE FOR BACTERIAL ENDOCARDITIS

Menakha .M *and C. Manoharan.
PG and Research Department of Botany and Microbiology A.V.V.M Sri PushPam College Poondi, Thanjavur, Tamil Nadu, India.
*Corresponding author's E-mail: menagamuniraj@yahoo.com

Accepted on: 26-07-2011; **Finalized on:** 30-10-2011.

ABSTRACT

Bacterial endocarditis is an infection of the inner surface of the heart or the heart valves caused by a group of bacteria usually found in the mouth. Even though many drugs are available, currently effective preventive medicines and vaccines are not available for the life-threatening endocarditis. Due to the availability of proteomes of several hosts and pathogens and tools for bioinformatics prediction for effective B-cell and T-cell epitopes, recently vaccine designing is shifted from the entire pathogen or whole antigen to peptide or epitope based vaccines that are more specific, safe and easy to produce. In the present study, the process of vaccine discovery starts with *Insilco* using the protein information of whole antigen rather than the pathogen itself. This novel process is called as reverse vaccinology. In this study the basic epitope prediction strategy is based on to get a common antigenic epitope which will be able to produce both the B-cell and T-cell mediated immunity. In order to design and develop such a peptide vaccine against the pathogens of bacterial endocarditis, antigens such as FimA and EfaA of *Streptococcus parasanguinis* and *Enterococcus faecalis* respectively were selected. The predicted T- epitope of *E. faecaliswas* was not able to act as potential vaccine candidate. However the potential T-epitope (LKVVTTNSI) derived from antigenic B-cell epitope (SKGSSSGASGKLKVVTTNSI) of FimA from *S.parasanguis* was identified as potential vaccine candidate. The selected FimA T-epitope (LKVVTTNSI) can be a good target not only for *S. parasanguis* but also to act as a common vaccine candidate for oral pathogens like *S. sanguis, S. gordonii, S. pneumoniae* and *E. faecalis* of LraI family (lipoprotein receptor antigen I) causing endocarditis. Selected T-epitope (LKVVTTNSI) was antigenic and has much potential to interact with most common human HLA alleles (A*0201, A*0204, B*2705, DRB1*0101 and DRB1*0401). The suitability of the vaccine candidate (LKVVTTNSI) for antibody production was confirmed by Peptide Design Tool.

Keywords: Bacterial endocarditis, *Streptococcus parasanguis, Enterococcus faecalis.* LraI family, B-epitope, T-epitope, antigenicity.

INTRODUCTION

Bacterial endocarditis is an inflammation of the inner surface of the heart and heart valves caused by a group of oral bacteria, usually found in the mouth. Infective endocarditis is a life-threatening endovascular infection believed to occur when bacteria in the blood stream adhere to the damaged heart valves. Conditions that cause heart valve damage are risk factors for native valve endocarditis and include congenital heart disease, chronic rheumatic heart disease, and mitral valve prolapsed. This infection results in a serious illness which requires prolonged treatment and on occasion produces injury to the heart or even death. In a healthy individual, bacteraemia (where bacteria get into the blood stream through a minor cut or wound) would normally be cleared quickly with no adverse consequences. If a heart valve is damaged and covered with a piece of blood clot, this provides a place for the bacteria to attach themselves, and an infection is established[1]. The bacteraemia is often caused by minor dental procedures, such as a tooth removal, bleeding from the gums, scaling etc, by which a high number of bacteria getting into the bloodstream. Periodontal disease and heart disease are both inflammatory conditions. There appears to be an association between periodontal disease and heart disease. Patients who have heart disease should have regular exams to check for signs of periodontal disease.

A group of streptococcal genes encoding related proteins that function as oral adhesions have been identified. These adhesion enzymes include FimA from *Streptococcus parasanguis*, SsaB from *Streptococcus sanguis*, ScaA from *Streptococcus gordonii* and PsaA from *Streptococcus pneumoniae*. It is possible that this group of streptococcal adhesions also plays a role in endocarditis by promoting adherence to vegetations. Normal oral flora can gain entry into the blood stream and establish infections in individuals with cardiac disease or preexisting valvular damage. The involvement of a protein related to these oral adhesions in endocarditis has been suggested in the case of *Enterococcus faecalis.* EfaA has been hypothesized to be an adhesion in endocarditis, and it appears to be an immunodominant antigen in patients with endocarditis.[2]

Antibiotic treatment is frequently initiated with penicillin, vancomycin and gentamycin, but antibiotic-resistant enterococci or staphylococci are not only resistant to penicillin, but also to vancomycin and amino glycosides. Currently, effective preventive medicines and defensive immune mechanisms are not available. Hence in the present work reverse vaccinology has been applied to develop potential vaccine candidate. The vaccine produced in the conventional broad based method is not functioning properly. Only the vaccine, based on the specific antigen alone will function properly, for which reverse vaccinology, a new method for identification of

vaccine candidate, has been proposed. This is a web-based vaccine design system that predicts vaccine candidates from the whole antigen. In the present work, in order to develop potential vaccine candidate for oral pathogens such as *Streptococcus parasanguinis* and *Enterococcus faecalis* causing endocarditis, reverse vaccinology has been applied. In this technique, instead of whole antigen, a part of the antigen known as T-epitope, which is recognized by both B-cell and T-cell mediated immunity, was predicted.

MATERIALS AND METHODS

S. parasanguinis and *E. faecalis* of LraI family were selected for the present work. **Pub Med** (www.ncbi.nlm.nih.gov/pubmed/) was used to access literature information. **NCBI** (www.ncbi.nlm.nih.gov) was used to retrieve the gene id of endocarditis causing oral pathogens (gi|153834|, gi|153826|, gi|7920464|, gi|310633|andgi|29376585|).

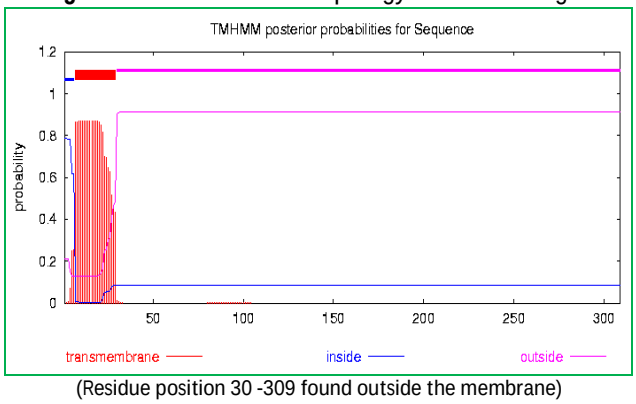
PDB (http://www.rcsb.org/) & SWISSPROT (http://www.expasy.org/sprot/) were used to retrieve the 3D structures of proteins (P31305 and Q832Z2) of FimA and EfaA. **VaxiJen v2.0** (www.ddg-pharmfac.net/vaxijen/) was used as an antigen prediction server. Proteins having antigenic score>0.4 were selected. **TMHMM** (www.cbs.dtu.dk ›), a server used to identify exo-membrane topology of amino acid sequences of protein. **BCPreds** (**http://ailab.cs.iastate.edu/bcpreds/),** a computational tool was used to predict linear B-cell epitopes in protein sequences. **Propred-1** (www.imtech.res.in/raghava/propred1/), was used to identify the MHC Class-I (47 alleles) binding regions in antigens. The server **Propred** (www.imtech.res.in/raghava/propred/) was used to predict MHC Class-II binding regions in an antigen sequence, that are useful in selecting vaccine candidates (51 MHC Class-II alleles). **MHCPred** was used to predict the binding affinity of MHC class I and II molecules and the value was given in terms of inhibitory concentration (IC50). **T-epitope Designer** was used to predict HLA binding peptides. **Peptide Design Tool (**http://www.innovagen.se/custom-peptide-polyclonalantisera/peptide-design-info pa.asp) was used to confirm the suitability of predicted T-cell epitopes for antibody production. Pepitope server was used for fold level characterization of T-epitope.

RESULTS AND DISCUSSION

Reverse vaccinology is an emerging vaccine development approach. It starts with the prediction of vaccine candidates through bioinformatics analysis from microbial proteome sequences. Since the process of vaccine discovery starts with *Insilco* using the protein information rather than the pathogen itself, this novel process can be named as reverse vaccinology. The major advantage of reverse vaccinology is finding vaccine targets quickly and efficiently. Reverse vaccinology is an improvement in vaccine development, pioneered by Rappuoli[3] and first used against meningococcal disease. Recently few works have been reported in this field[4-7]. Many vaccines impossible to develop so far will become a reality. Reverse Vaccinology is one of the best examples of how bioinformatics can boost molecular immunology. This allows scientists to save time and money while facing pathogens for which cell culture is difficult or impossible.

Bacterial endocarditis is caused by virulent group of streptococci and enterococci of which *S. parasanguinis* and *E. faecalis* respectively are the common species. FimA and EfaA, the two virulent enzymes causing heart disease, are found in *S. parasanguinis* and *E. faecalis* respectively. The above two enzymes formed the whole antigens in disease causing processes. Full length of the antigens of *S.parasanguinis* and *E. faecalis* was 309 and 308 amino acid residues in respectively. The present investigation is interested in studying these enzymes for epitope prediction in novel vaccine development. As a first step, the two enzymes were screened for eligible antigens. An eligible antigen must have a vaxijen score more than 0.4 and should be exo-membrane in topology based on TMHMM results as stated by Barh *et al*.,[8]. In the present study, the two antigens, FimA and EfaA were exo-membrane in topology (Fig.1) and the VaxiJen scores were 0.4801 and 0.4164 (Table.1) respectively. Therefore both of them have satisfied the above two conditions and selected as two eligible antigens, from which B-cell epitopes and T-cell epitopes were derived.

Figure 1: Exomembrane topology of whole antigen



(Residue position 30 -309 found outside the membrane)

Epitope prediction

An epitope, also known as antigenic determinant, is the part of an antigen that is recognized by the immune system. Earlier studies have reported either T-cell or B-cell based epitope designing for a given pathogen[9-11]. Similarly, some vaccines can only activate helper T-lymphocytes (HTL)/CD4+ / MHC II. But activation of CD8+ cytotoxic T- lymphocytes (CTLs)/ MHC I is also required in many cases[12]. Therefore, an epitope that can produce both the B-cell and T-cell (MHC I and MHC II) mediated immunity is highly useful in developing peptide-based vaccines. Recently, Barh & Misra[13] and Barh *et al*.,[8] while studying the epitope prediction in *Neisseria gonorrhoeae* have used this new strategy to identify the epitopes which have high potentiality to produce both the B- and T- cell mediated immunity.

Table 1: Prediction of B-epitopes from full length antigens using BCPred(BCPred and AAP). (Selected B-cell epitopes are in bold letters)

Name of antigen and B-cell Epitope sequences (BCPreds)	Position of Amino acids	BCPreds Score*	VaxiJen Scores for B-cell epitope**	Topology of antigen / B-cell Epitope ***	Antigenicity
FimA antigen	**1-309**	**-**	**0.4801**	**Outside**	**Antigen**
BCPred method					
PSAYIWEINTEEEGTPEQIK	218-237	1	1.3147	inside	Non-Antigen
LHSIVPVGKDPHEYEPLPED	56-75	0.973	-0.3097	outside	Non-antigen
SKGSSSGASGKLKVVTTNSI	**22-41**	**0.946**	**1.6625**	outside	**Antigen**
LEGQNQAGKEDPHAWLNLEN	**127-146**	**0.942**	**0.9825**	outside	**Antigen**
KQAFKNIPEDKKMIVTSEGC	188-207	0.894	-0.3380	inside	Non-antigen
VESSVDERPMKTVAKDTNIP	253-272	0.778	0.9726	outside	Antigen
AAP method					
NIAGDKIELHSIVPVGKDPH	**48-67**	**1**	**0.4830**	**outside**	**Antigen**
CSKGSSSGASGKLKVVTTNS	**21-40**	**1**	**1.8323**	**outside**	**Antigen**
WEINTEEEGTPEQIKTLVEK	223-242	1	0.8281	inside	Non-Antigen
TGGNAWFTKLVKNANKVENK	94-113	0.996	0.4263	inside	Non-Antigen
IPIYAKIFTDSIAKEGEKGD	**271-190**	**0.995**	**1.0293**	**outside**	**Antigen**
EGQNQAGKEDPHAWLNLENG	128-147	0.446	0.8826	inside	Non-Antigen
IPEDKKMIVTSEGCFKYFSK	194-213	0.11	-0.4501	inside	Non-antigen
EfaA antigen BCPred method	**1-308**	-	**0.4164**	**Outside**	**Antigen**
AKEGTEGDTYYSMMNWNLTK	281-300	0.982	1.4186	inside	Non-**Antigen**
HSIVPIGTDPHEYEPLPEDI	55-74	0.967	-0.0410	outside	Non-antigen
YIWEINTESQGTPEQMTTII	219-238	0.937	0.7970	inside	Non-**Antigen**
VEKDPKNKDFYTENAKNYTE	157-176	0.873	0.4994	inside	Non-**Antigen**
AAP method					
NTESQGTPEQMTTIIDTIKK	224-243	1	0.2544	inside	Non-Antigen
APVLFVETSVDKRSMERVSK	**246-265**	**1**	**0.6601**	**outside**	**Antigen**
TSKNVTPQYLTSAGQEQTED	116-135	1	0.9347	inside	Non-Antigen
DKIELHSIVPIGTDPHEYEP	**50-69**	**1**	**0.5263**	**outside**	**Antigen**
DPKNKDFYTENAKNYTEKLS	160-179	1	0.2612	inside	Non-Antigen
TLFTDSLAKEGTEGDTYYSM	**274-293**	**1**	**1.0494**	**outside**	**Antigen**
KAKFADIPDDKKLLVTSEGA	186-205	0.981	-0.1811	inside	Non-Antigen
AWLDIENGIKYVENIRDVLV	138-157	0.05	-0.2984	outside	Non-**Antigen**

***Eligible B-epitope should be >0.8 BCPred score; ****Eligible B-epitope should be >0.4 VaxiJen score; *****B-epitope on outer surface is eligible B-epitope.

B-cell epitopes

For effective antigenicity, the antigen should not be a lengthy one. An ideal antigen should have 8 to 20 amino acid residues in length. Such short peptide segment is described as an epitope. An epitope with 20 amino acid residues is termed as B-epitope and an epitope with 9 amino acid residues as T-epitope. In the present study, antigenic B-cell epitopes were predicted from whole antigen using vaxijen, TMHMM 2.0 and BCPreds. This method generated 20 mer sequences of B-epitopes from FimA and EfaA enzymes that meet all the selection criteria for the antigenic B-cell epitopes (Table 1). From the above study, 5 B-epitopes out of 13 from FimA and 3 out of 12 B-epitopes from EfaA were identified as qualified B-epitopes from which T-epitope 9mers were derived. Such a kind of study was also carried out by Barh et al.,[8].

T- cell epitopes

The T-epitopes were derived from qualified B-epitopes. T-epitope prediction was not as simple as in B-cell epitope prediction. The T-epitopes were predicted through two screening methods. In the first level of screening, each selected B-cell epitope was subjected to T-cell epitopes prediction. The qualified B-epitopes were submitted to ProPred 1 (for 47 alleles of MHC I) and ProPred (for 51 alleles of MHC II) with default parameters. Based on MHC-I binding, 38 T-epitopes from FimA and 20 from EfaA were obtained. On the other hand MHC-II binding T-epitopes were lesser with 9 each from FimA and EfaA were predicted (Table.2). The common epitopes, that can bind both the MHC classes and covers maximum (more than 15) MHC alleles, were selected by using the above two servers such as ProPred1 and ProPred. In this process, 5 T-epitopes from FimA and 3 from EfaA were selected. The selected epitopes were further analyzed for vaxijen score and MHCPred IC50 value.

MHCPred (DRB1*0101 allele) was used to identify common T-cell epitopes which can interact with both the MHC classes with highest number, and specifically interact with DRB1*0101 (DRB1*0101 is the commonest bound allele, therefore the interaction epitope should produce better antigenic response) The T-epitope must interact with HLA DRB*0101 alleles and secure IC50 value not more than 50 which will indicate as good binders. In the present study, out of 5 T-epitopes of FimA selected, only one, 'LKVVTTNSI' was found to be a good binder with HLA DRB*0101 alleles and also showed a good antigenic determinant by having vaxijen score more than 0.4. Since this T-epitope secured antigenic score value of 0.4354, it was finally selected in the first level screening. Out of 3 T-epitopes selected from EfaA, only one attained the antigenic score more than the required one (0.4392).

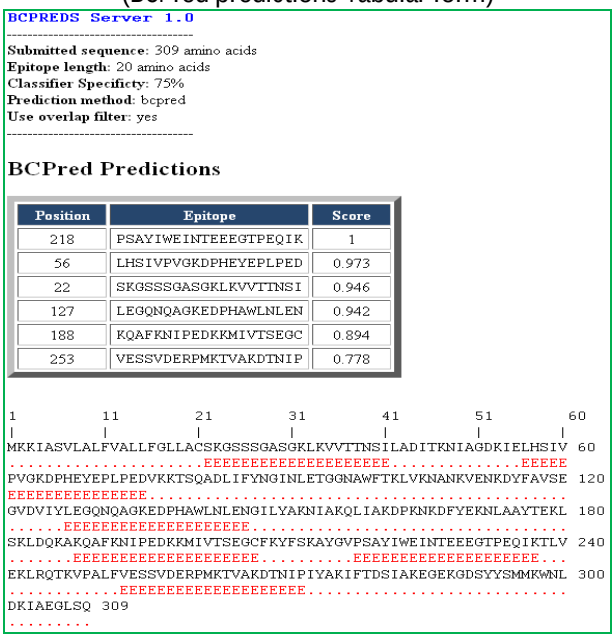
Table 2: Prediction and selection of T-epitopes from B-epitopes

B-epitopes	Epitopes binding with MHC-I alleles	No. of alleles bound	Epitopes binding with MHC-II alleles	No. of alleles bound	Total alleles (MHC-I+II)	Vaxijen score	MHCpred Score Ic50 value)	Cluster score	Selected T-epitopes
FimA	SSSGASGKL	14							
SKGSSSGASGKLKVVTTNSI	SSGASGKLK	2							
	GSSSGASGK	3							
	SKGSSSGAS	1							
	GKLKVVTTN	1							
	KLKVVTTNS	2							
	GASGKLKVV	7							
	LKVVTTNSI	4	LKVVTTNSI	42	46	0.4354	34.83	7.8225	LKVVTTNSI
	SGKLKVVTT	2							
	SGASGKLKV	7							
	ASGKLKVVT	2							
	LEGQNQAGK	1							
	NQAGKEDPH	1							
	KEDPHAWLN	3	Nil						
LEGQNQAGKEDPHAWLNLEN	GKEDPHAWL	5							
	QAGKEDPHA	3							
	AGKEDPHAW	2							
	EDPHAWLNL	11							
	DPHAWLNLE	1							
	IVPVGKDPH	**1**	**IVPVGKDPH**	**26**	**27**	0.0432			
	DKIELHSIV	5	IAGDKIELH	10					
NIAGDKIELHSIVPVGKDPH	NIAGDKIEL	11	LHSIVPVGK	9					
	GDKIELHSI	4							
	IELHSIVPV	**11**	**IELHSIVPV**	**18**	**29**	0.2825			
	GSSSGASGK	3							
	KLKVVTTNS	2							
	SSGASGKLK	3	Nil						
CSKGSSSGASGKLKVVTTNS	SSSGASGKL	14							
	GASGKLKVV	7							
	SGASGKLKV	7							
	ASGKLKVVT	2							
	SGKLKVVTT	2							
	KIFTDSIAK	5							
	DSIAKEGEK	3							
IPIYAKIFTDSIAKEGEKGD	**YAKIFTDSI**	**11**	**YAKIFTDSI**	**8**	**19**	-0.3683			
	IYAKIFTDS	3	IYAKIFTDS	1	4				
	IFTDSIAKE	1	IFTDSIAKE	12	13				
	IPIYAKIFT	6	IPIYAKIFT	23	29	0.4890	891.25		
	FVETSVDKR	6	FVETSVDKR	4	10				
	SVDKRSMER	5	VLFVETSVD	13					
	KRSMERVSK	3							
EfaA	VETSVDKRS	3							
APVLFVETSVDKRSMERVSK	VDKRSMERV	3							
	ETSVDKRSM	3							
	LFVETSVDK	1	LFVETSVDK	5	6				
	PVLFVETSV	1							
	APVLFVETS	5							
	IVPIGTDPH	1	IVPIGTDPH	25	26	0.4392	131.22	8.6559	IVPIGTDPH
	IELHSIVPI	**12**	**IELHSIVPI**	**27**	**39**	0.1382			
DKIELHSIVPIGTDPHEYEP	DKIELHSIV	6	IVPIGTDPH	8					
	VPIGTDPHE	1							
	LHSIVPIGT	**2**	**LHSIVPIGT**	**20**	**22**	-1.1749			
	TLFTDSLAK	4							
	KEGTEGDTY	8							
TLFTDSLAKEGTEGDTYYSM	TEGDTYYSM	9							
	EGTEGDTYY	6							
	LFTDSLAKE	1	LFTDSLAKE	9	10				
	TDSLAKEGT	2	LAKEGTEGD	4					

Table 3: Selection of vaccine candidate from T-epitopes through T-Epitope Designer and MHCPred (IC50 Value)

Antigens (Proteins)	Selected T-epitopes	T-Epitope Designer A*0201	A*0204	B*2705	MHCPred (IC50 Value) DRB1*0101	DRB1*0401	T-Epitope Designer (No. of binders in %)	Selected Vaccine candidate
FimA	LKVVTTNSI	224.03	172.40	1196.10	34.83	1180.32	98.74% (mostly positive for almost all A*,B* & C* alleles)	LKVVTTNSI
EfaA	IVPIGTDPH	-259.76	-756.95	1192.67	131.22	228.03		Nil

Figure 2: Prediction of B-epitopes of FimA (BCPred predictions-Tabular form)



BCPREDS Server 1.0

Submitted sequence: 309 amino acids
Epitope length: 20 amino acids
Classifier Specificty: 75%
Prediction method: bcpred
Use overlap filter: yes

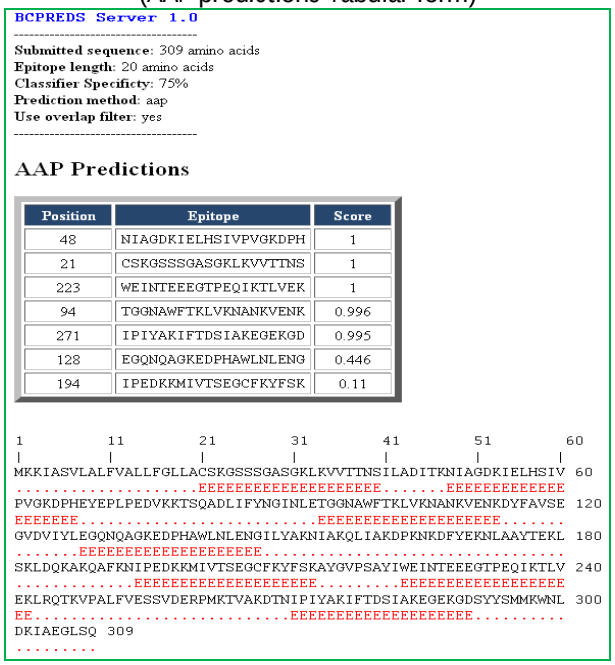
BCPred Predictions

Position	Epitope	Score
218	PSAYIWEINTEEEGTPEQIK	1
56	LHSIVPVGKDPHEYEPLPED	0.973
22	SKGSSSGASGKLKVVTTNSI	0.946
127	LEGQNQAGKEDPHAWLNLEN	0.942
188	KQAFKNIPEDKKMIVTSEGC	0.894
253	VESSVDERPMKTVAKDTNIP	0.778

```
1          11         21         31         41         51         60
|          |          |          |          |          |          |
MKKIASVLALFVALLFGLLACSKGSSSGASGKLKVVTTNSILADITKNIAGDKIELHSIV 60
.....................EEEEEEEEEEEEEEEEEEE..............EEEEE
PVGKDPHEYEPLPEDVKKTSQADLIFYNGINLETGGNAWFTKLVKNANKVENKDYFAVSE 120
EEEEEEEEEEEEEE.....................................................
GVDVIYLEGQNQAGKEDPHAWLNLENGILYAKNIAKQLIAKDPKNKDFYEKNLAAYTEKL 180
......EEEEEEEEEEEEEEEEEEE.....................................
SKLDQKAKQAFKNIPEDKKMIVTSEGCFKYFSKAYGVPSAYIWEINTEEEGTPEQIKTLV 240
.......EEEEEEEEEEEEEEEEEEE..........EEEEEEEEEEEEEEEEEEE...
EKLRQTKVPALFVESSVDERPMKTVAKDTNIPIYAKIFTDSIAKEGEKGDSYYSMMKWNL 300
............EEEEEEEEEEEEEEEEEEE.....................
DKIAEGLSQ 309
.........
```

Visualization of B-cell epitopes on protein.

Figure 3: Prediction of B-epitopes of FimA (AAP predictions-Tabular form)



BCPREDS Server 1.0

Submitted sequence: 309 amino acids
Epitope length: 20 amino acids
Classifier Specificty: 75%
Prediction method: aap
Use overlap filter: yes

AAP Predictions

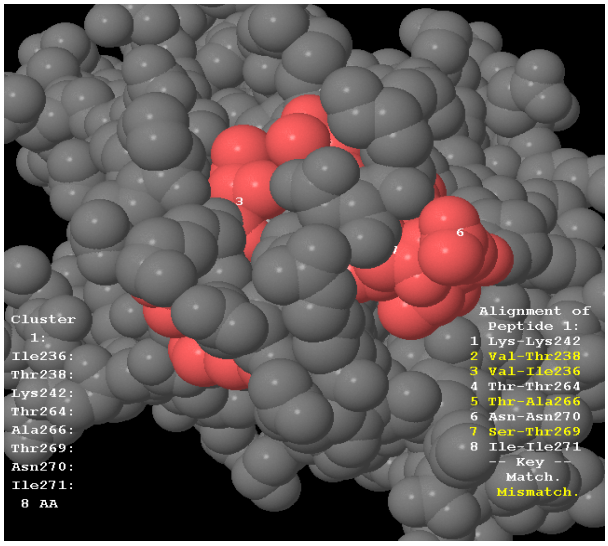
Position	Epitope	Score
48	NIAGDKIELHSIVPVGKDPH	1
21	CSKGSSSGASGKLKVVTTNS	1
223	WEINTEEEGTPEQIKTLVEK	1
94	TGGNAWFTKLVKNANKVENK	0.996
271	IPIYAKIFTDSIAKEGEKGD	0.995
128	EGQNQAGKEDPHAWLNLENG	0.446
194	IPEDKKMIVTSEGCFKYFSK	0.11

```
1          11         21         31         41         51         60
|          |          |          |          |          |          |
MKKIASVLALFVALLFGLLACSKGSSSGASGKLKVVTTNSILADITKNIAGDKIELHSIV 60
.....................EEEEEEEEEEEEEEEEEEE.......EEEEEEEEEEEEE
PVGKDPHEYEPLPEDVKKTSQADLIFYNGINLETGGNAWFTKLVKNANKVENKDYFAVSE 120
EEEEEEE.............................EEEEEEEEEEEEEEEEEEE.......
GVDVIYLEGQNQAGKEDPHAWLNLENGILYAKNIAKQLIAKDPKNKDFYEKNLAAYTEKL 180
.......EEEEEEEEEEEEEEEEEEE.....................................
SKLDQKAKQAFKNIPEDKKMIVTSEGCFKYFSKAYGVPSAYIWEINTEEEGTPEQIKTLV 240
.............EEEEEEEEEEEEEEEEEEE.........EEEEEEEEEEEEEEEEEEE
EKLRQTKVPALFVESSVDERPMKTVAKDTNIPIYAKIFTDSIAKEGEKGDSYYSMMKWNL 300
EE.............................EEEEEEEEEEEEEEEEEEE.........
DKIAEGLSQ 309
.........
```

Visualization of B-cell epitopes on protein.

There is every possibility of the identified epitopes getting folded inside the tertiary structure of the corresponding protein[8]. Therefore, a fold level analysis of two identified epitopes was carried out to determine the position of epitopes in folded proteins and to confirm their exo-membrane topology. Analysis of the epitope clusters and positions of epitopes with in the folded protein by Pepitope server showed that the identified two epitopes are situated within clusters having acceptable scores, located at the surface of the corresponding protein. Therefore, these results confirmed the two identified epitopes as qualified epitopes at the first level screening (Fig. 4).

Figure 4: Fold level characterization of cluster and topology of T-epitope (in red) using Pepitope server.



In the second level of screening, the predicted T-epitope from first level of screening should satisfy certain criteria. They must bind with large number of alleles (798) and also bind with a minimum of >75% of total HLA molecules. In the present study, the selected T-epitope (LKVVTTNSI) of FimA from *S. parasanguinis* bound with 98.74% of HLA alleles (mostly positive for almost all A*, B* & C* alleles). The identified epitope is highly expected to bind both the classes of MHC molecules specifically, A*0201, A*0204, B*2705, DRB1*0101, and DRB1*0401, that are most frequent MHC alleles in human population[8]. In the present study also the selected T-epitope was able to bind with all kinds of above alleles through T-Epitope Designer and MHC Pred (IC50 Value). The single T-epitope (LKVVTTNSI) of FimA from *S. parasanguinis* satisfied all the above conditions and qualified to act as a vaccine candidate**.** On the other hand, the predicted T-epitope (IVPIGTDPH) of EfaA found on *E. faecalis* was not qualified because of its negative binding score with **A*0201 (**-259.76) **and A*0204 (**-756.95) HLA alleles. Its binding ability with DRB1*0101 was also very poor and therefore it was rejected to act as a vaccine candidate.

Viscount et al., [14] while studying the suitability of FimA as a vaccine for streptococcal endocarditis, examined the degree of conservation of the gene encoding FimA among several species of streptococci. FimA-like proteins were commonly expressed by species of *Streptococcus* and *Enterococcus* that cause endocarditis. Presence of FimA homology among the viridans group of streptococci strongly suggested that the FimA protein is a good target for broad-based preventive vaccine against streptococcal native valve endocarditis. In the present study also, from the *Insilco* approach, it was found that the T-epitope (LKVVTTNSI) of FimA from *S.parasanguinis* was qualified to act as a common vaccine candidate for the species of *Streptococcus* and *Enterococcus* causing endocarditis (Fig.5).

Figure 5: Pair wise alignments showing conserved vaccine candidate (LKVVTTNSI) with T-epitope sequences in other bacterial endocarditis causing enzymes of LraI protein family (lipoprotein receptor antigen I).

```
FimA    1 ------------------------------LKVVTTNSI----------  9
                                         |.:||||||
EfaA    1 MKKFSLFFLTLLAGLTLAACGNQAAEKKEKLAIVTTNSILSDLVKNVGQD  50   8/9 (88.9%)

FimA    1 ---------------------------------LKVVTTNSI---------  9
                                             |||||.||||
ssaB    1 MKKLGFLSLLLLAVCTLFACSNQKNASSDSSKLKVVATNSIIADITKNIA  50   8/9 (88.9%)

FimA    1 ---------------------------------LKVVTTNSI--------  9
                                             |.||.||||
scaA    1 MKKCRFLVLLLLAFVGLAACSSQKSSTDSSSKLNVVATNSIIADITKNI  50   7/9 (77.8%)

FimA    1 ---------------------------------LKVVTTNSI---------  9
                                             ||||.||||
PsaA    1 MKKLGTLFVLFLSVIVLVACASGKKDAASGQKLKVVATNSIIADITKNIA  50   8/9 (88.9%)
```

1. Bar indicates identical amino acids
2. Double points indicate conserved substitutions
3. Single point indicates semi-conserved substitution.

Identification of common epitope for endocarditis.

Using subtractive genomics and homology analysis, single epitope of FimA was identified within the nearly same accessible regions (30-39 AA) of LraI family of enzymes. So that the identified T-epitope (LKVVTTNSI) could be used as a common epitope against a wide range of viridans groups streptococci causing endocarditis. The identified T-epitope (vaccine candidate) from *S. parasanguis* may induce both B-cell and T-cell mediated immunity. The sequence homology analysis demonstrated that this epitope sequence of *S. parasanguis* FimA *is* conserved in SsaB from *S. sanguis*, ScaA from *S.gordonii*, PsaA from *S. pneumoniae* and EfaA from *E. faecalis.* The pair wise alignments of conserved vaccine candidate (LKVVTTNSI) with T-epitope sequences in other bacterial enzymes causing endocarditis showed more than 77.8% homology of amino acids. The T-epitope (LKVVTTNSI) of FimA (*S.parasanguis*) with EfaA *(E.faecalis.),* SsaB (*S. sanguis*) and PsaA (*S.pneumoniae)* showed 88.9% homology while it was 77.8% homology with ScaA from *S.gordonii (*Fig.5). It is well known that the above five amino acid sequences will be functionally similar since they share >77.8% sequence identity at amino acid level. Therefore, this vaccine candidate (LKVVTTNSI) could be used in designing a common vaccine against all the 5 oral pathogens causing endocarditis .Similar type of work has already been reported by Barh *et al*.,[8] in their *Insilco* Identification of dual ability of *Neisseria gonorrhea* ddl enzyme for developing drug and vaccine against pathogenic *Neisseria* and other human pathogens.

In order to find out the suitability for antibody production, the qualified vaccine candidate (LKVVTTNSI) was submitted to Peptide Station at Sweden for verification through online Peptide Design Tool. They confirmed that the submitted T-epitope was suitable for antibody production.

CONCLUSION

The selected T-epitope (LKVVTTNSI) was antigenic and has much potential to interact with most common human HLA alleles and can be a good target, not only for *S. parasanguis* but also, to act as a common vaccine candidate for endocarditis causing oral pathogens like *E. faecalis., S. sanguis*, *S. gordonii* and *S. pneumoniae.* The identified epitope requires proper design and subsequent validation for its use as peptide vaccine against the oral pathogens causing endocarditis.

Acknowledgement: The authors are thankful to Dr. K. Karthikeyan INI Centre for Research and Cognizance (ICRAC)**,** Chennai for the technical support and healthy discussion.

REFERENCES

1. Kasper DL, Brunwald E, Fauci AS, Hauser S, Longo DL, Jameson JL. 2005. Harrison's Principles of Internal Medicine. McGraw-Hill. pp. 731–40. ISBN 0-07-139140-1.
2. Dana burnette-curley, virginia wells, helen viscount, cindy munro, j. christopher fenno, p. fives-taylor, and francis. macrina. FimA, a Major Virulence Factor Associated with *Streptococcus parasanguis* Endocarditis *infection and immunity*, 63, 1995, p. 4669–4674.
3. Rappuoli R: Reverse vaccinology. *Curr Opin Microbiol* 3: 2000, 445-450.
4. Vivona .S, F. Bernante, and F. Filippini, "NERVE: new enhanced reverse vaccinology environment," *BMC Biotechnology*, 6, 2006, article 35.
5. Xiang Z, He Y. Vaxign: a web-based vaccine target design program for reverse vaccinology. Procedia in Vaccinology. 1, 2009, P.23-29.
6. He Y, Xiang Z, Mobley HLT. Vaxign: the first web-based vaccine design program for reverse vaccinology and an application for vaccine development. *J Biomed Biotechnol*. 2010; 2010:297505. Epub 2010 Jul [PMID: 20671958].
7. Menakha, M and Manoharan, C. Computer Aided Drug Design from Immunoinfomatics to Reverse vaccinology, *Advanced Biotech*. 10. 2011.P:144-147.
8. Barh D, Misra AN, Kumar A. 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: 2010, 082-090.

9. Arockiasamy A and Krishnaswamy S. Prediction of B-cell epitopes for Salmonella typhi OmpC. J Biosci 20: 1995, p.235-243.

10. Arévalo-Herrera M, Valencia AZ, Vergara J, Bonelo A, Fleischhauer K, *et al*. Identification of HLA-A2 restricted CD8(+) T-lymphocyte responses to *Plasmodium vivax* circum sporozoite protein in individuals naturally exposed to malaria. Parasite Immunol 24: 2002, p.161-169.

11. Sollner J, Grohmann R, Rapberger R, Perco P, Lukas A, *et al*. Analysis and prediction of protective continuous B-cell epitopes on pathogen proteins. *Immunome Res* 4: 2008, p.1.

12. Pancré V, Gras-Masse H, Delanoye A, Herno J, Capron A, *et al*. Induction of cytotoxic T-cell activity by the protective antigen of *Schistosoma mansoni* Sm28GST or its derived C-terminal lipopeptide. Scand J *Immunol* 44: 1996, p-485-492.

13. Barh D and Misra AN. in silico Identification of membrane associated candidate drug targets in Neisseria gonorrhoeae. *International Journal of Integrative Biology* 6: 2009, 65-67.

14. Helen b. viscount, cindy munro, dana burnette-curley, darrell. peterson and francis. Macrina. Immunization with FimA Protects against *Streptococcus parasanguis* Endocarditis in Rats *Infection and immunity*, 65: 1997, p. 994–1002.

****