Journal of
Proteomics & Bioinformatics

Victor et al., J Proteomics Bioinform 2014, 7:8
DOI: 10.4172/0974-276X.1000323

Research Article Open Access

Phylogenetic Characterization and Threading Based-Epitope Mapping of Leptospiral Outer Membrane Lipoprotein LipL41

Angela Asir Ramasamy Victor[2], Sunil Abraham[1], Manimaran Nellai Chinniah[1], Madanan Gopalakrishnan Madathiparambil[3] and Jebasingh Tennyson[1]*

[1]*School of Biological Sciences, Madurai Kamaraj University, Madurai-625021, Tamil Nadu, India*
[2]*St. Michael college of Engineering and Technology, St. Santhiagappar Nagar, Kalayarkoil-630 551, Sivagangai district, Tamil Nadu, India*
[3]*Regional Medical Research Centre (ICMR), Post Bag No-13, Port Blair-744 101, Andaman and Nicobar Islands, India*

Abstract

Leptospiral outer membrane Lipoprotein LipL41 is one of the key virulence factor expressed during leptospiral infection on mammalian host. Phylogenetic analysis of LipL41 from 87 pathogenic species has shown the comprising lineages of LipL41 with largely varying rates of evolution. The species of *L. borgpetersenii* are clustered together and form a group with highest bootstrap value. Relationships between the species Lai and Copenhageni are resolved by 20 and 8 sequences separately with highest and lowest bootstrap values of 99 and 50, respectively. Molecular model of LipL41 was predicted in Raptor X server to map eight sequential and conformational B cell epitopes of LipL41 in Ellipro server. All these epitopes were found to be conserved among *Leptospira* and these could be used for the development of vaccine and diagnostic kit to detect *Leptospira*.

Keywords: *Leptospira*; LipL41; Phylogenetic analysis; B-cell epitope

Abbreviations: SAVES: Structural Analysis and Verification Server; MEGA: Molecular Evolutionary Genetic Analysis; LBP: Local Bootstrap values; ACC: Auto Cross Covariance; PI: Protrusion Index; NEFF: Number of Effective sequence homologs

Introduction

Leptospirosis is one of the most widespread zoonotic diseases in the world and is caused by the pathogen *Leptospira* [1]. Susceptible animals, including humans are infected by direct contact with urine from rodents, or indirectly through contaminated water. Transmission occurs via dermal abrasions or inoculation of the mucous or conjunctival membranes [2]. It is also known as Weil's syndrome and the clinical manifestations of leptospirosis include high fever, bleeding, and renal failure. The major target of *Leptospira* is the renal proximal tubular cells of kidney. Mortality ranges from 10-15% in cases of the traditional Weil's disease and can be more than 70% in cases of severe pulmonary haemorrhage syndrome (SPHS) [3-5]. Leptospirosis is endemic in most of the southern states of India like Kerala, Tamil Nadu and certain parts of Andhra Pradesh [6-8].

The *Leptospira* genus is sub-classified into 18 genomospecies that included both saprophytic and pathogeneic species [1,9]. Based on serologic methods, approximately 300 serovars have been identified; of which more than 200 are pathogenic [1,2,10]. The availability of genome sequence data for different *Leptospira* strains drives the discovery of new diagnostic tools and vaccines for Leptospirosis [11]. The major problem associated with Leptospirosis is the diagnosis of the disease, as it shows multiple symptoms; very often it has been confused with other diseases. The diagnosis of the leptospiral infection is very much complicated when compared with other ailments.

A number of leptospiral outer membrane proteins (OMPs) have been characterized including OmpL1 [12], LipL41 [13], LipL36 [14], LipL32 [14], LipL21 [15], LipL46 [16], LenA [17], Loa22 [16] and Omp52 [12]. However their performance in diagnostic assays for acute leptospirosis or as vaccine candidates has been problematic [12,18]. Among these OMPs, LipL41 appears to be a great interest since it is one of the key virulence determinants involved in host-pathogen interactions and it is being formed only in the host during the infection [13,15,18]. LipL41 is essential for virulence of *L. interrogans* in the animals [19]. Recent studies have revealed that LipL41 may play an important role in the infection and produces immunological responses in the host during the infection of *Leptospira* [12].

Considering the large number of pathogenic leptospiral serovars and broad distribution of leptospiral host reservoirs, the potential effect of selective pressure on the evolutionary mapping of the LipL41 proteins was not studied so far. The availability of genomic sequences of various serovars and strains opened up opportunities to identify evolutionary relationships among different pathogenic strains of *L. interrogans* and others representing various kinds of serotypes (serogroups and serovars). Given the potential of the LipL41 proteins as diagnostic antigens and vaccine candidates, we examined the evolutionary relationship of LipL41 with 87 sequences from various serovars and strains followed by protein threading and mapping of B-cell epitopes from the conserved region of the alignment to develop vaccine for Leptospirosis.

Materials and Methods

Sequence retrieval, sequence alignment and dendrogram construction

Amino acid sequences used in this study were retrieved from protein knowledgebase (UniProt KB) (http://www.uniprot.org/uniprot/). A total of 87 sequences of LipL41 from different serovars and strains (Table 1) were collected from UniProtKB. Sequences with

***Corresponding author:** Jebasingh Tennyson, School of Biological Sciences, Madurai Kamaraj University, Madurai-625021, Tamil Nadu, India, Tel: +91-9043568671; Fax: +91-452-2459873; E-mail: jebasinghs@gmail.com

Received June 05, 2014; **Accepted** July 21, 2014; **Published** July 25, 2014

Citation: Victor AAR, Abraham S, Chinniah MN, Gopalakrishnan Madathiparambil M, Tennyson J (2014) Phylogenetic Characterization and Threading Based-Epitope Mapping of Leptospiral Outer Membrane Lipoprotein LipL41. J Proteomics Bioinform 7: 222-231. doi:10.4172/0974-276X.1000323

	Organism	Serovar	Serogroup/Name/ uncharacterized protein	Strain	Uniprot ID
1.	*Leptospira interrogans*	*Lai*	Icterohaemorrhagiae	56601	Q8F8E1
2.	*Leptospira interrogans*	Copenhageni	Icterohaemorrhagiae	Fiocruz L1-130	Q72N71
3.	*Leptospira interrogans*	Icterohaemorrhagiae	Uncharacterized protein	Verdun	M6R807
4.	*Leptospira weilii*	Manhao	Major Outer Membrane protein	-	Q6GXB6
5.	*Leptospira interrogans*	-	Outer Membrane protein	-	P71435
6.	*Leptospira interrogans*	Valbuzzi	Uncharacterized protein	Duyster	N6XR14
7.	*Leptospira interrogans*	Pomona	Uncharacterized protein	Kennewicki_LC82-25	J4T638
8.	*Leptospira interrogans*	Canicola	Major outer membrane protein		Q6GXC6
9.	*Leptospira interrogans*	Lai	Major outer membrane protein		Q6GXC8
10.	*Leptospira interrogans*	Australis	Major outer membrane protein		Q6GXC2
11.	*Leptospira interrogans*	Autumnalis			Q33BN1
12.	*Leptospira interrogans*	Hebdomadis			Q33BN0
13.	*Leptospira interrogans*	Manilae			Q33BM7
14.	*Leptospira kirschneri*	Bulgarica		Nikolaevo	M6F5M3
15.	*Leptospira kirschneri*			H1	K6FCL0
16.	*Leptospira interrogans*	Autumnalis	Major outer membrane protein		Q6GXC3
17.	*Leptospira interrogans*	Wolffi	Major outer membrane protein		Q6GXB5
18.	*Leptospira kirschneri*	Cynopteri	Uncharacterized protein	3522_CT	S3UFG2
19.	*Leptospira kirschneri*		Uncharacterized protein		Q48587
20.	*Leptospira kirschneri*		Uncharacterized protein	200801774	M6XE01
21.	*Leptospira kirschneri*	Sokoine	Uncharacterized protein	RM1	M6JZZ1
22.	*Leptospira kirschneri*	Bim	Uncharacterized protein	1051	M6I8Y4
23.	*Leptospira kirschneri*	Bim	Uncharacterized protein	str._PUO_1247	M6ETW2
24.	*Leptospira kirschneri*		Uncharacterized protein	str._MMD1493	M6DUL6
25.	*Leptospira kirschneri*	Valbuzzi	Uncharacterized protein	str._200702274	K8I2W0
26.	*Leptospira kirschneri*	Grippotyphosa		str._Moskva	K8HDP3
27.	*Leptospira kirschneri*	Grippotyphosa		str._RM52	J4SXV5
28.	*Leptospira interrogans*		Uncharacterized protein	str._L1207	M6M715
29.	*Leptospira noguchii*	Panama		str._CZ214	T0FJ19
30.	*Leptospira noguchii*			str._1993005606	S3HU63
31.	*Leptospira interrogans*		Uncharacterized protein	str._HAI1536	M6V6T1
32.	*Leptospira noguchii*	Autumnalis	Uncharacterized protein	str._ZUN142	M6U3T0
33.	*Leptospira noguchii*		Uncharacterized protein	str._2007001578	M6I0J7
34.	*Leptospira noguchii*		Uncharacterized protein	str._2006001870	K8KY78
35.	*Leptospira interrogans*	Australis	Lipoprotein		Q33BM9
36.	*Leptospira interrogans*	Icterohaemorrhagiae	Lipoprotein		Q33BM8
37.	*Leptospira kirschneri*		Uncharacterized protein	str._200801925	M6XK47
38.	*Leptospira kirschneri*		Uncharacterized protein	str._200803703	M6W9C5
39.	*Leptospira kirschneri*		Uncharacterized protein	str._200802841	K6GUX5
40.	*Leptospira interrogans*	Grippotyphosa			I0AYM4
41.	*Leptospira interrogans*	Icterohaemorrhagiae	Outer membrane lipoprotein		C9EH90
42.	*Leptospira borgpetersenii*	Hardjo-bovis	Lipoprotein	strain_JB197	Q04VM6
43.	*Leptospira borgpetersenii*	Javanica	Uncharacterized protein	str._UI_09931	S3URD2
44.	*Leptospira borgpetersenii*	Ballum	Major outer membrane protein		Q6GXC5
45.	*Leptospira borgpetersenii*	Pomona	Uncharacterized protein	str._200901868	M6W4K3
46.	*Leptospira borgpetersenii*	Mini	Uncharacterized protein	str._201000851	N6XFT9
47.	*Leptospira borgpetersenii*		Uncharacterized protein	str._Noumea	M6S0C6
48.	*Leptospira sp.*	Kenya	Uncharacterized protein	str._Sh9	M6EBK5
49.	*Leptospira borgpetersenii*		Uncharacterized protein	str._200701203	M3HTU4
50.	*Leptospira borgpetersenii*	Castellonis	Uncharacterized protein	str._200801910	K8I5J9
51.	*Leptospira borgpetersenii*		Uncharacterized protein	str._UI_09149	K8HKM7
52.	*Leptospira borgpetersenii*		Uncharacterized protein	str._200801926	K6IQU5
53.	*Leptospira borgpetersenii*	Javanica	Uncharacterized protein	str._MK146	M6MSW7
54.	*Leptospira borgpetersenii*		Uncharacterized protein	str._Brem_328	M6J8U8
55.	*Leptospira borgpetersenii*		Uncharacterized protein	str._Brem_307	M6J296
56.	*Leptospira borgpetersenii*	Javanica	Major outer membrane protein		Q6GXC7
57.	*Leptospira borgpetersenii*	Tarassovi	Major outer membrane protein		Q6GXB7
58.	*Leptospira borgpetersenii*	Hardjo-bovis	Lipoprotein	strain_L550	Q04XU7
59.	*Leptospira borgpetersenii*	Hardjo-bovis	Uncharacterized protein	str._Sponselee	M6BLU3
60.	*Leptospira santarosai*	Shermani	Uncharacterized protein	str._1342KT	S3WCD7

61.	*Leptospira santarosai*		Uncharacterized protein	str._AIM	M6YHV6
62.	*Leptospira sp.*		Uncharacterized protein	Fiocruz LV4135	M5V8R7
63.	*Leptospira santarosai*	Shermani	Lipoprotein	str._LT_821	K8Y7P9
64.	*Leptospira santarosai*		Uncharacterized protein	str._JET	K8MNX0
65.	*Leptospira sp*		Uncharacterized protein	Fiocruz_LV3954	K6HER9
66.	*Leptospira borgpetersenii*	Mini	Uncharacterized protein	str._200901116	M6UV01
67.	*Leptospira santarosai*		Uncharacterized protein	str._2000027870	M6GU93
68.	*Leptospira borgpetersenii*		Uncharacterized protein	str._200901122	K8LWU9
69.	*Leptospira santarosai*	Arenal	Uncharacterized protein	str._MAVJ_401	M6JWP5
70.	*Leptospira santarosai*		Uncharacterized protein	str._MOR084	K6FVF0
71.	*Leptospira santarosai*		Uncharacterized protein	str._200403458	M6X740
72.	*Leptospira interrogans*	Hebdomadis	Uncharacterized protein	str._R499	K8JJ55
73.	*Leptospira weilii*		Uncharacterized protein	str._UI_13098	M6Q9M2
74.	*Leptospira sp*		Uncharacterized protein	P2653	M6A0W1
75.	*Leptospira kirschneri*		Uncharacterized protein	str._H2	K6GP23
76.	*Leptospira weilii*		Uncharacterized protein	Topaz_str._LT2116	M3GXL7
77.	*Leptospira borgpetersenii*	Mini	Major outer membrane protein		Q6GXB4
78.	*Leptospira alstoni*	Pingchang	Uncharacterized protein	str._80-412	T0FYD6
79.	*Leptospira alstoni*	Sichuan	Uncharacterized protein	str._79601	M6CS57
80.	*Leptospira weilii*		Uncharacterized protein	str._LNT_1234	M6LQ40
81.	*Leptospira interrogans*	Paidjan	Outer_membrane_lipoprotein_LipL41_(Fragment)		Q5MJS6
82.	*Leptospira weilii*	Ecochallenge	Uncharacterized protein		N1U2R2
83.	*Leptospira interrogans*	Pomona	Outer_membrane_lipoprotein_LipL41_(Fragment)		Q5MJS8
84.	*Leptospira interrogans*	Bataviae	Uncharacterized protein	str._HAI135	M6TFS7
85.	*Leptospira weilii*		Uncharacterized protein	str._2006001855	M6FKQ9
86.	*Leptospira interrogans*	Pomona	Uncharacterized protein	str._CSL4002	M5ZV00
87.	*Leptospira interrogans*	Valbuzzi	Uncharacterized protein	str._Duyster	M5ZDN6

Table 1: List of LipL41 sequences of *Leptospira* species used for phylogenetic analysis in figure 1.

	Organism	Serovar	Serogroup/Name/ uncharacterized protein	Strain	Uniprot ID
1.	*Leptospira interrogans*	Hebdomadis			Q33BN0
2.	*Leptospira interrogans*	Manilae			Q33BM7
3.	*Leptospira interrogans*	Autumnalis			Q33BN1
4.	*Leptospira interrogans*	Copenhageni	Icterohaemorrhagiae	Fiocruz L1-130	Q72N71
5.	*Leptospira interrogans*	*Lai*	Icterohaemorrhagiae	56601	Q8F8E1
6.	*Leptospira interrogans*	Australis	Major outer membrane protein		Q6GXC2
7.	*Leptospira interrogans*	Wolffi	Major outer membrane protein		Q6GXB5
8.	*Leptospira interrogans*	Canicola	Major outer membrane protein		Q6GXC6
9.	*Leptospira_kirschneri*	Bulgarica		Nikolaevo	M6F5M3
10.	*Leptospira_kirschneri*			H1	K6FCL0
11.	*Leptospira interrogans*	Lai	Major outer membrane protein		Q6GXC8
12.	*Leptospira interrogans*	Autumnalis	Major outer membrane protein		Q6GXC3
13.	*Leptospira_weilii*	Manhao	Major outer membrane protein	-	Q6GXB6
14.	*Leptospira interrogans*	Australis	Lipoprotein		Q33BM9
15.	*Leptospira interrogans*	Icterohaemorrhagiae	Lipoprotein		Q33BM8
16.	*Leptospira_interrogans*		Uncharacterized protein	str._L1207	M6M715
17.	*Leptospira_interrogans*	Icterohaemorrhagiae	Outer membrane lipoprotein		C9EH90
18.	*Leptospira_interrogans*	Icterohaemorrhagiae	Uncharacterized protein	Verdun	M6R807
19.	*Leptospira_interrogans*	Hebdomadis	Uncharacterized protein	str._R499	K8JJ55
20.	*Leptospira_kirschneri*		Uncharacterized protein	str._H2	K6GP23

Table 2: List of LipL41 sequences of *Leptospira species* used for phylogenetic analysis in figure 2.

significant identity were aligned with ClustalW algorithm implemented in Molecular Evolutionary Genetic Analysis (MEGA 5.2.2) (http://www.megasoftware.net) by using distance matrix and then it was trimmed to consensus. Neighbour Joining (NJ) trees were constructed with 1000 bootstraps at uniform divergence rates with distance 'p' as the evolutionary model and with a data subset to use with gaps/ missing data treatment as complete deletion [20]. Posterior probability and conserved regions among the closely related sequences were done with MEGA 5.2.2.

Modelling, energy minimization and validation of the model

The amino acid sequence of LipL41 of *Leptospira interrogans* serogroup Icterohaemorrhagiae serovar Lai (strain 56601) was retrieved from UniProt KB (Uniprot Id: Q8F8E1). The sequence was submitted into the RaptorX server [21] (http://raptorx.uchicago.edu/) to derive the 3-dimensional structure. The modelled protein structures were viewed in Swiss-PdbViewer (http://www.expasy.org/spdbv/) and the individual residues were collected from 100 cycles of steepest

descent algorithm carried out in GROMOS96 [22] until the side chain interactions in the vicinity is readjusted and brings up lower potential energy and becomes more stable. Energy minimized models were assessed by PROCHECK [23] to analyse the stereo chemical quality and residual geometry of the model by submitting the co-ordinate file in Structural Analysis and Verification Server (SAVES) (http://nihserver.mbi.ucla.edu/SAVES/). The value of the predicted LipL41 model was analyzed by using PYMOL [24,25].

Computational mapping of epitopes

Linear and discontinuous B-Cell epitopes of LipL41 were mapped from the generated three dimensional structure of LipL41. Linear B-cell epitopes were chosen with two different algorithms: ABCPred and BepiPred. ABCPred uses a recurrent neural network to predict B-cell epitopes (http://www.imtech.res.in/raghava/abcpred/) [26,27]. The amino acid length of 16 and the scoring threshold of 0.8 were set to predict B-cell epitopes in ABCPred. The epitope prediction in BepiPred is based on hidden Markov model (http://www.cbs.dtu.dk/services/BepiPred/) and propensity scale method [28,29]. The value, 0.35 was set as the threshold value, because at this value, the sensitivity/ specificity of predictions are maximized in BepiPred. BepiPred analyzes each amino acid independently and does not have a minimum or maximum number of amino acids to predict an epitope.

Discontinuous epitopes were predicted using Ellipro which is an Antibody Epitope Prediction server (http://tools.immuneepitope.org/tools/ElliPro/iedbinput) [30,31]. Ellipro, with the best algorithm to predict discontinuous epitopes from 3-D structures when compared to six other software programs that predict discontinuous epitopes [30]. The default threshold value was set at 0.8. The predicted epitopes were additionally verified in VaxiJen server to predict the probability of an antigen (http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html/) [32], with a threshold of 0.4. VaxiJen uses an alignment free approach for antigen prediction and works on an auto cross covariance (ACC) transformation of protein sequences into uniform vectors of principal amino acid properties. Sixteen sequences of LipL41 were taken randomly in T-coffee programme to identify the conserved amino acids residues.

Results and Discussion

Phylogenetic analysis for genetic relatedness

The phylogenetic tree was performed by using 87 sequences of LipL41 from various serovars and strains which were retrieved from UniProt Knowledgebase (Uniprot KB) (Table 1). The phylogenetic tree of LipL41 is evidenced that the isolates are clustered with different serovars and strains (Figure 1) and diverged to form different branches in the phylogenetic tree. The following mutations are observed for Borgpetersenii strains: 33S->T, 39M->I, 40F->Y, 125A->I, 126I->L, 130S->T, 139N->S, 191D->E, 336T->V; Weilii strains: 33S->A, 247I->V; Fiocruz strains: 80A->P, 177L->I, 183A->V, 186M->A, 197E->D; Santorasai strains: 125A->I, 126I->L, 130S->T, 139N->S, 336T->I; Krishneri strains: 176I->V, 336T->A; Noguchi strains: 269I->M, 274R->K, 336T->A (data not shown). The closest neighbouring clusters include strains of *L. weilli*, *L. kirschneri* and *L. interrogans* with 100% bootstrap confidence values. Based on the phylogenetic analysis, the cluster of LipL41 of *L. borgpetersenii* together forms a clade showing the evolutionary relationship of same serovars and strains with the highest bootstrap value of 98 which indicates that it has uniform support. The reliability of a branch length in MEGA 5 is based on confidence probability (CP). The branch length is high when the CP is high, thus the branch length is considered to be statistically significant. MEGA 5 inferred the evolutionary tree by a Neighbour-Joining (NJ) algorithm by using a matrix of pairwise distances. In order to resolve the relationships of the sequences within each group of the constructed phylogenetic tree (Figure 1), a separate phylogenetic tree was constructed with 20 sequences from different serovars (Table 2) (Figure 2). It shows the highest bootstrap value of 99, indicate that the clade is close to 100%, which reveals that all the characters in a group believed to comprise all the evolutionary descendants of a common ancestor which is rooted with different serovars and strains as the ancestral group. In order to resolve the polytomies and to make the evolutionary relationship into dichotomies, 8 sequences of different serovars and strains (Table 3) (Figure 3) was used to make a separate branch of tree reflected Polytomies with Local bootstrap probability (LBP) values below 50% in *L. interrogans* Icterohaemorrhagiae serovar Lai and Copenhageni serovar (Figure 3). The phylogenetic tree evidenced that *L. interrogans* Icterohaemorrhagiae serovar Lai and Copenhageni serovar were closely clustered with other different serovars. The neighbouring groups also include serovars of *L. autumnalis* and *L. hebdomadis* which clustered with different bootstrap confidence values. Even though LipL41 is highly analogous protein present in all pathogenic *Leptospira* but the phylogenetic pattern of the present study exhibited the clonality of the sequences of the serovar Lai and Copenhageni used to analyze species separation. The evolutionary relationship was confirmed among the 87 sequences and on further confirmation with serovar Lai can be used for serodiagnosis of pathogenic leptospiral species.

Structure prediction of LipL41 by threading method

Protein threading method is a fold recognition method of protein modelling which is based on the predicted structure properties, such as predicted secondary structures and predicted residue burial status [21]. Threading based prediction for LipL41 was done in RaptorX server which uses a non-linear scoring function to combine homologous information with structural information for the given template-sequence alignment [21] . The amino acid sequence of LipL41 of *L. interrogans* Icterohaemorrhagiae serovar Lai (Q8F8E1) was taken for modelling by protein threading. Given an input sequence, RaptorX predicted its tertiary structure as well as solvent accessibility and disordered regions (Figure 4A). The RaptorX assigned the confidence score which is based on P-value and uGDT (unnormalized Global Distance Test). P-value measures the relative quality and uGDT measures the absolute quality of protein model. uGDT has greater value of 50 is a indicator for a good model. The input of 355 amino acid residues of LipL41 were completely modelled with 100% and showed 2 domains (Figure 4A). In the model, 7 positions among 355 residues were predicted as disordered regions which are 1%. The model shows P-value with 3.52 e-03 and uGDT (GDT) with 137(38). The modelled LipL41 has two domains (Figure 4A); Domain 1 (254 to 355) showed p value of 5.32 e-3 and Domain 2 (1 to 253) showed p value 3.52 e-3.

LipL41 is a haem binding protein with Cys-Ser (CS) and Cys-Pro (CP) domains [9] (Figure 4B). The CS and CP are conserved domains of pathogenic *Leptospira* which are responsible for immunoprotection [9]. Mutation of these domains fails to cause immunoprotection in mice [12]. The motifs Cys-Pro or Cys-Ser has been determined in diverse proteins binding to heme (Fe2+)/hemin (Fe3+) [33,34]. It has been reported that the cysteine containing dipeptide: CS or CP is necessary for heme binding in HRM [35-37]. These conserved residues are found in LipL41 at 140 Cys-Ser and 220 Cys-Pro are located on the surface of the predicted structure (Figure 4B), and the thiol of cysteine may be a ligand for iron on heme [36,37].

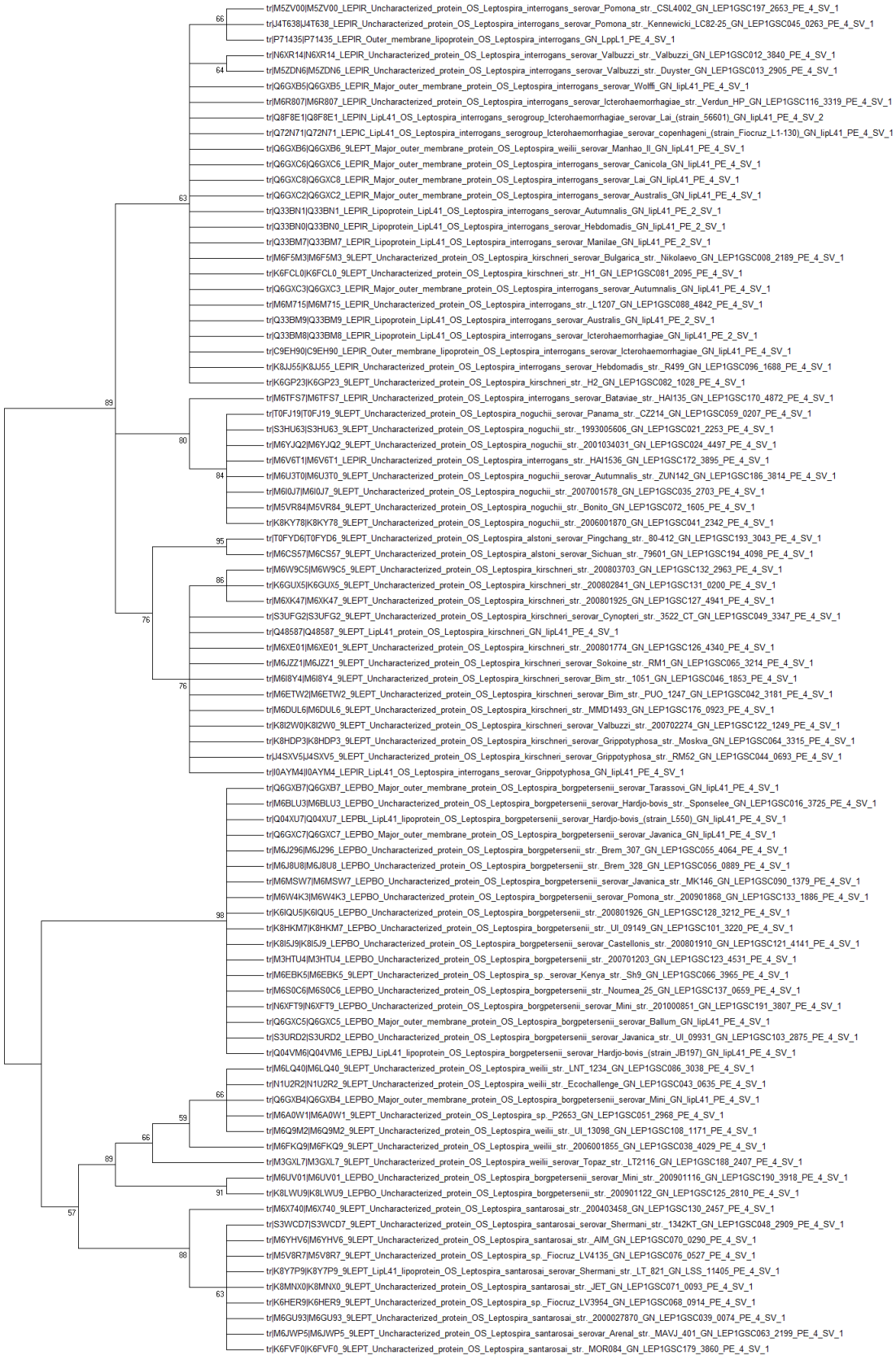

Figure 1: Phylogenetic tree with 87 sequences of LipL 41. Bootstrap confidence levels are shown as percentages on nodes and confidence values are shown in branches. The phylogenetic tree showed the divergence of Leptospiral strains and serovar with LipL41.

Figure 2: A phylogram of 20 selected sequences of LipL41. The bootstrap values calculate the frequency for each taxon bipartition during replication and boot strapping denotes measures how consistently the data support given taxon bipartitions. The scale bar also represents branch length (number of amino acid substitutions/100 residues). The high bootstrap value 99 shows the uniform support and bootstrap values close to 100% which indicate that the clade is a group which means that all the characters in a group believed to comprise all the evolutionary descendants of a common ancestor which is rooted with different serovars and strains as the ancestral group.

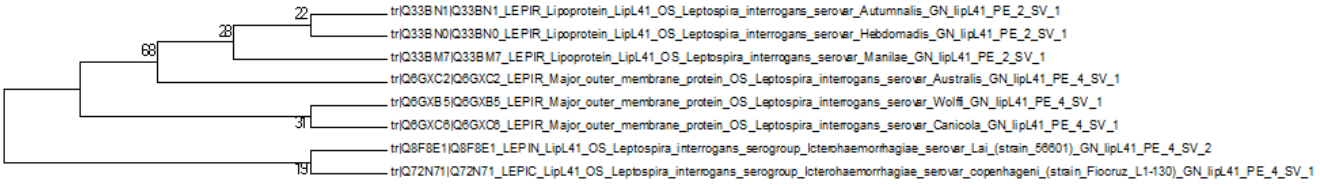

Figure 3: A phylogram of 8 selected sequences of LipL 41. Polytomies reflect nodes with LBP values below 50% for Lai and Copenhageni serovar of *L. interrogans* Icterohaemorrhagiae. This tree resolves the relationship among different LipL 41 groups and the sequences within each group and serovars.

	Organism	Serovar	Serogroup/Name/ uncharacterized protein	Strain	Uniprot ID
1.	*Leptospira interrogans*	Wolffi	Major outer membrane protein		Q6GXB5
2	*Leptospira interrogans*	Canicola	Major outer membrane protein		Q6GXC6
3.	*Leptospira interrogans*	Australis	Major outer membrane protein		Q6GXC2
4.	*Leptospira interrogans*	*Lai*	Icterohaemorrhagiae	56601	Q8F8E1
5.	*Leptospira interrogans*	Copenhageni	Icterohaemorrhagiae	Fiocruz L1-130	Q72N71
6.	*Leptospira interrogans*	Autumnalis			Q33BN1
7.	*Leptospira interrogans*	Hebdomadis			Q33BN0
8.	*Leptospira interrogans*	Manilae			Q33BM7

Table 3: List of LipL41 sequences of *Leptospira interrogans* used for phylogenetic analysis in figure 3.

No.	Start Position	End Position	Peptide	Number of Residues	Score
1	22	42	ATVDVEYPVFPKDKEGRALQK	21	0.758
2	68	97	EGSSFIDQMPSKVFEAFDKESYYKLTDLSK	30	0.821
3	132	150	GYQKPYTECSTENKIDAVA	19	0.726
4	167	182	DVNTGNEPVSKPTGVR	16	0.608
5	185	189	LIPLD	5	0.521
6	195	213	VETGEVKKAVVSSPAKIFN	19	0.606
7	267	284	QEGYEEIVGETPSFKKAK	18	0.712
8	303	355	ANLATYYFSTGDFEKSIKLYEEAMKLKDADKSYLRELRKRVEATFAVDESNAK	53	0.774

Table 4: List of epitopes predicted based on amino acid sequence and structure.

Validation and evaluation of LipL41 model

The NEFF (Number of Effective sequence homologs) score which was ranging from 1 to 20 for the predicted structure was estimated by PROCHECK. The results obtained from PROCHECK [23,38] was evaluated for protein backbone conformations by Ramachandran Plot [39,40]. The phi-psi torsion angle for 92.7% of residues of LipL41 are in the most favourable region (A, B and L); 6.6%, 0.3% and 0.2% in additionally allowed, generously allowed and disallowed regions, respectively (a, b, l, p), indicate that LipL41 model is stereo chemically good (Figure 5) and the model derived from RaptorX was of higher quality in terms of protein folding.

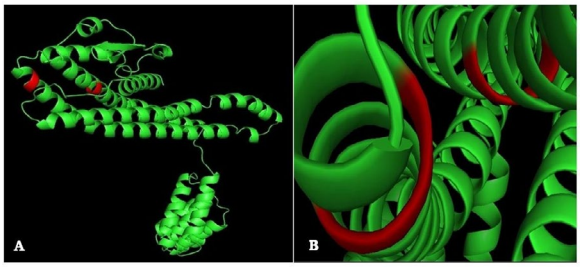

Figure 4: A: Three dimensional structure of LipL41 of *Leptospira interrogans* serogroup Icterohaemorrhagiae serovar Lai (strain 56601 (Uniprot Id: Q8F8E1). **B:** LipL41 model showing haem binding regions (CS and CP) marked in red.

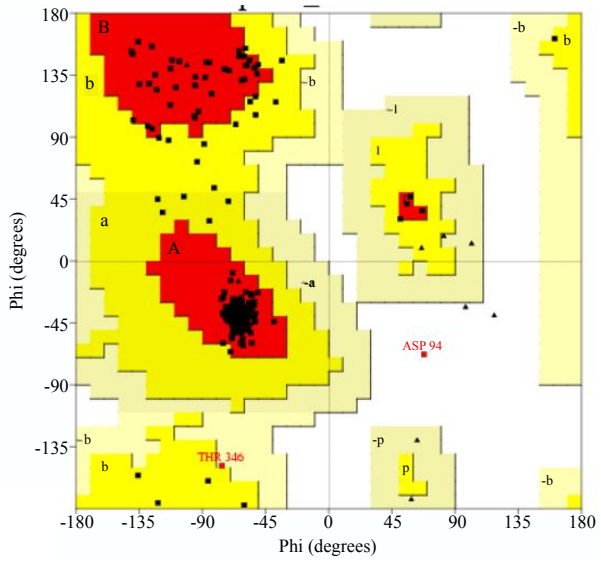

Figure 5: Ramachandran plot of the LipL41 model. The most favoured regions are colored red, additional allowed, generously allowed and disallowed regions are indicated as yellow, light yellow and white fields, respectively.

Prediction and immunoinformatic analysis of antigenic peptides

Two different epitope prediction software programs (ABCPred and BepiPred) were utilized to predict the most immunogenic linear B-cell epitopes on the surface of the leptospiral OMP LipL41. ABCPred and BepiPred predicted 9 different overlapping and potentially immunogenic regions within LipL41, respectively (data not shown). ABCPred is able to predict epitopes with approximately 66% accuracy using the recurrent neural network [26]. ABCPred assigns scores between 0 and 1 for each epitope it predicts. If prediction shows score closer to 1, the particular prediction can be taken as epitope and prediction closer to 0 is not suitable for an epitope. Eight B- cell discontinuous epitopes of LipL41 were mapped from the predicted 3-D structure of LipL41 by using Ellipro (Figure 6). These epitopes spans (EP1: 22-42), (EP2: 68-97), (EP3: 132-150), (EP4: 167-182), (EP5: 185-189), (EP6: 195-213), (EP7: 267-284) and (EP8: 303-355) positions of LipL41 (Table 4). Ellipro predicts epitope with Protrusion Index (PI) value which is percentage of the protein atoms enclosed in the ellipsoid, at which the residue first becomes lying outside the ellipsoid; whereas all the residues which were lying 90% outside the ellipsoid had the PI value 9, i.e., 0.9 in Ellipro. This gives information of amino acids lying outside the ellipsoid.

The prediction of peptides is vital not only for diagnostics but also for vaccines. It became clear that the small segments of protein called the antigenic determinants or the epitopes are sufficient for eliciting the desired immune response. Based on the threshold value, all the predicted epitopes are antigenic nature. All these epitopes predicted can be used for the development of Monoclonal antibody or epitope based diagnostic kit for the Leptospirosis.

Conservancy of LipL41 epitopes

Universal epitope vaccine development requires conserved amino acids of a protein among the various pathogenic strains [41]. Thus 16 different strains and serovars of *Leptospira* were taken randomly. Multiple sequence alignment of LipL41 for 16 different strains and

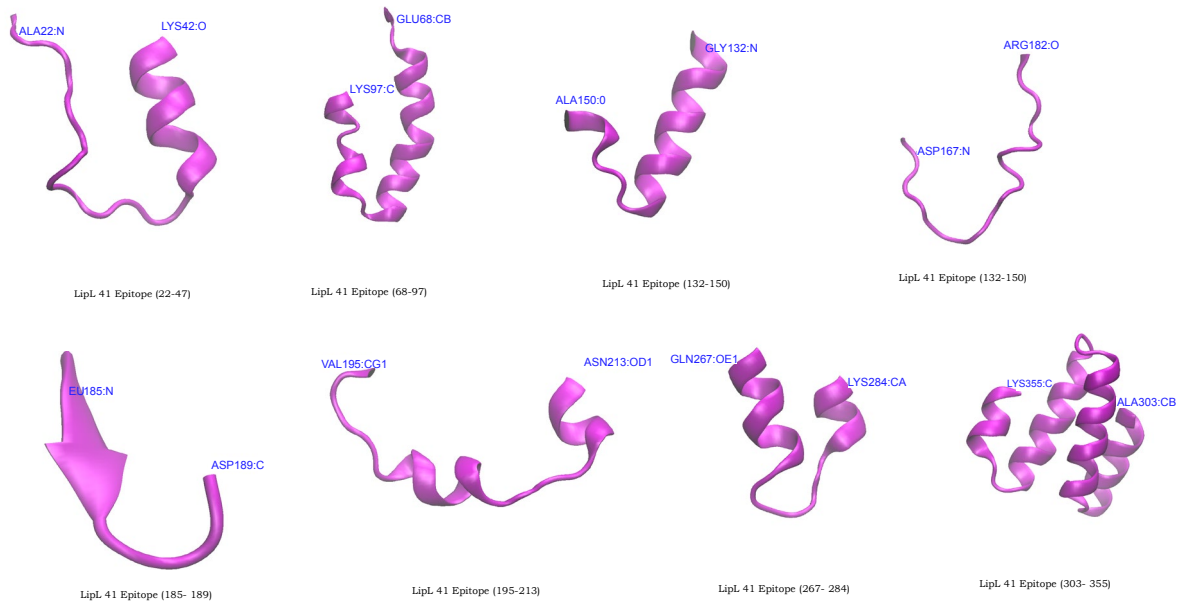

Figure 6: Structures of B-cell epitopes of LipL41.

Figure 7: Multiple amino acid sequences alignment of LipL41. Red regions in the above alignment show perfect sequence alignment among the sequences. Conserved epitopes are marked.

serovars of *Leptospira interrogans*: *Leptospira interrogans* serovar Wolffi (Q6GXB5), *Leptospira weilii* serovar Manhao II (Q6GXB6), *Leptospira interrogans* serovar Canicola (Q6GXC6), *Leptospira interrogans* serovar Australis (Q6GXC2), *Leptospira interrogans* serovar Autumnalis (Q6GXC3), *Leptospira kirschneri* serovar Bulgarica str. Nikolaevo (M6F5M3), *Leptospira interrogans* str.L1207(M6M715), *Leptospira interrogans* serogroup Icterohaemorrhagiae serovar Lai (strain_56601) (Q8F8E1), *Leptospira interrogans* serogroup Icterohaemorrhagiae serovar copenhageni (strain_Fiocruz_L1-130) (Q72N71), *Leptospira interrogans* serovar Autumnalis (Q33BN1), *Leptospira interrogans* serovar Hebdomadis (Q33BN0), *Leptospira interrogans* serovar Manilae (Q33BM7), *Leptospira interrogans* serovar Australis (Q33BM9), *Leptospira interrogans* serovar Icterohaemorrhagiae (Q33BM8), *Leptospira kirschneri* str. (K6FCL0), *Leptospira interrogans* serovar Icterohaemorrhagiae (C9EH90) were analyzed by using T-coffee program (http://www.tcoffee.org/) [42] (Figure 7) and found that all these epitopes are conserved among all these strains and serovars.

Prediction of transmembrane domain, signal sequence and topology of LipL41

In order to ensure that the epitopic region should not overlap with signal peptide or transmembrane domain of LipL41, amino acid sequence of LipL41 of *Leptospira interrogans* serogroup Icterohaemorrhagiae serovar Lai (strain_56601) was analyzed by TMHMM 2.0 server (http://www.cbs.dtu.dk/services/TMHMM/) [43]. It was found that amino acids from 1-6, 7-29, and 30-355 are located inside the plasma membrane, inside the transmembrane (TM) helix, and outside the plasma membrane of the cell, respectively. A combined trans-membrane topology and signal peptides were predicted using Phobius online server (http://phobius.sbc.su.se/) [44]. Based on the probability of occurrences of specific amino acids, the sequence from 1 to 22 was predicted to be a signal peptide and between 23 and 355 is non-cytoplasmic region of LipL41. This analysis confirms that all the eight epitopes are topologically surface exposed and do not have any signal sequences in them.

Conclusion

In this present study, we have characterized LipL41 for its genetic diversity among the *Leptospira* species. The phylogenetic relationship of *Leptospira* with LipL41 from 87 sequences of different serovars and serogroup have shown that the comprising lineages with largely varying rates of evolution. The alignment also has shown the presence of haem binding motifs are conserved in all the LipL41. The three dimensional structure of LipL41 was predicted by using RAPTOR X and was validated by Ramachandran plot. Eight B-cell epitopes were predicted from LipL41. Antibody developed against these conserved epitopic regions could be used to develop a detection kit or as a vaccine candidate for Leptospirosis.

Acknowledgement

The authors thank DBT-IPLS and UGC-Networking Resource Centre in Biological Sciences (NRCBS) for providing necessary facilities to carry out the work.

References

1. Levett PN, Branch SL, Whittington CU, Edwards CN, Paxton H (2001) Two methods for rapid serological diagnosis of acute leptospirosis. Clin Diagn Lab Immunol 8: 349-351.
2. Faine S, Adler B, Bolin C, Perolat P (1999) Leptospira and leptospirosis. (2nd Edn0 MediSci, Australia.
3. Gouveia EL, Metcalfe J, de Carvalho AL, Aires TS, Villasboas-Bisneto JC, et al. (2008) Leptospirosis-associated severe pulmonary hemorrhagic syndrome, Salvador, Brazil. Emerg Infect Dis 14: 505-508.
4. McBride AJ, Athanazio DA, Reis MG, Ko AI (2005) Leptospirosis. Curr Opin Infect Dis 18: 376-386.
5. Segura ER, Ganoza CA, Campos K, Ricaldi JN, Torres S, et al. (2005) Clinical spectrum of pulmonary involvement in leptospirosis in a region of endemicity, with quantification of leptospiral burden. Clin Infect Dis 40: 343-351.
6. Koteeswaran AS (2006) Seroprevalence of leptospirosis in man and animals analysis of the gene encoding LipL32 of Leptospira in Tamil Nadu. Indian J Med Microbiol 24: 329-331.
7. Balakrishnan G, Govindarajan R, Meenambigai TV, Jayakumar V, Manohar MB (2008) Diagnosis of leptospirosis by recombinant antigen: Seroprevalence of animal based single serum dilution ELISA. Indian Vet J 85: 227- 228
8. Abhinay G, Joseph S, Ambily R (2012) Sero prevalence of canine leptospirosis. Indian Vet J 89: 72-73.
9. Lin MH, Chang YC2, Hsiao CD3, Huang SH, Wang MS, et al. (2013) LipL4, a hemin binding protein from Leptospira santarosai serovar Shermani. PLoS One 8: e83246.
10. Bharti AR, Nally JE, Ricaldi JN, Matthias MA, Diaz MM, et al. (2003) Leptospirosis: a zoonotic disease of global importance. Lancet Infect Dis 3: 757-771.
11. Adler B, de la Peña Moctezuma A (2010) Leptospira and leptospirosis. Vet Microbiol 140: 287-296.
12. Haake DA, Mazel MK, McCoy AM, Milward F, Chao G, et al. (1999) Leptospiral outer membrane proteins OmpL1 and LipL41 exhibit synergistic immunoprotection. Infect Immun 67: 6572-6582.
13. Shang ES, Summers TA, Haake DA (1996) Molecular cloning and sequence analysis of the gene encoding LipL4, a surface-exposed lipoprotein of pathogenic Leptospira species. Infect Immun 64: 2322-2330.
14. Haake DA, Chao G, Zuerner RL, Barnett JK, Barnett D, et al. (2000) The leptospiral major outer membrane protein LipL32 is a lipoprotein expressed during mammalian infection. Infect Immun 68: 2276-2285.
15. Cullen PA, Xu X, Matsunaga J, Sanchez Y, Ko AI, et al. (2005) Surfaceome of Leptospira spp. Infect Immun 73: 4853-4863.
16. Matsunaga J, Barocchi MA, Croda J, Young TA, Sanchez Y, et al. (2003) Pathogenic Leptospira species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily. Mol Microbiol 49: 929-945.
17. Verma A, Brissette CA, Bowman AA, Shah ST, Zipfel PF, et al. (2010) Leptospiral endostatin-like protein A is a bacterial cell surface receptor for human plasminogen. Infect Immun 78: 2053-2059.
18. Guerreiro H, Croda J, Flannery B, Mazel M, Matsunaga J, et al. (2001) Leptospiral proteins recognized during the humoral immune response to leptospirosis in humans. Infect Immun 69: 4958-4968.
19. Monahan AM, Callanan JJ, Nally JE (2008) Proteomic analysis of Leptospira interrogans shed in urine of chronically infected hosts. Infect Immun 76: 4952-4958.
20. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance and maximum parsimony methods. Mol Biol Evol 28: 2731-2739.
21. Källberg M, Wang H, Wang S, Peng J, Wang Z, et al. (2012) Template-based protein structure modeling using the RaptorX web server. Nat Protoc 7: 1511-1522.
22. Bonvin AMJJ, Mark AE, Van Gunsteren WF (2000) The GROMOS96 benchmarks for molecular simulation. Comput Phys Commun 128: 550-557.
23. Unit M, Street G (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 26: 283-291.
24. Schrodinger LLC (2010) The PyMOL Molecular graphics system, version~1.3r1.
25. DeLano WL (2002) The PyMOL Molecular graphics system.
26. Pirovano W, Heringa J (2010) Protein secondary structure prediction. Methods Mol Biol 609: 327-348.

27. Saha S, Raghava GP (2007) Prediction of neurotoxins based on their function and source. In Silico Biol 7: 369-387.

28. Larsen JE, Lund O, Nielsen M (2006) Improved method for predicting linear B-cell epitopes. Immunome Res 2: 2.

29. Kringelum JV, Lundegaard C, Lund O, Nielsen M (2012) Reliable B cell epitope predictions: impacts of method development and improved benchmarking. PLoS Comput Biol 8: e1002829.

30. Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, et al. (2008) ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics 9: 514.

31. Ponomarenko JV, Regenmortel Marc HV (2009) B-cell epitope prediction. Struct Bioinfo 35: 849-881.

32. Doytchinova IA, Flower DR (2007) VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics 8: 4.

33. Lathrop JT, Timko MP (1993) Regulation by heme of mitochondrial protein transport through a conserved amino acid motif. Science 259: 522-525.

34. Zhang L, Guarente L (1995) Heme binds to a short sequence that serves a regulatory function in diverse proteins. EMBO J 14: 313-320.

35. Huang TJ, McCoubrey WK Jr, Maines MD (2001) Heme oxygenase-2 interaction with metalloporphyrins: function of heme regulatory motifs. Antioxid Redox Signal 3: 685-696.

36. Yang F, Xia X, Lei HY, Wang ED (2010) Hemin binds to human cytoplasmic arginyl-tRNA synthetase and inhibits its catalytic activity. J Biol Chem 285: 39437-39446.

37. Westberg JA, Jiang J, Andersson LC (2011) Stanniocalcin 1 binds hemin through a partially conserved heme regulatory motif. Biochem Biophys Res Commun 409: 266-269.

38. Laskoswki RA, MacArthur MW, Moss DS, Thorton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. J Appl Crystallogr 26: 283-291.

39. Ramachandran GN, Ramakrishnan C, Sasisekharan V (1963) Stereochemistry of polypeptide chain configurations. J Mol Biol 7: 95-99.

40. Ramachandran GN, Sasisekharan V (1968) Conformation of polypeptides and proteins. Adv Protein Chem 23: 283-438.

41. Sirskyj D, Diaz-Mitoma F, Golshani A, Kumar A, Azizi A (2011) Innovative bioinformatic approaches for developing peptide-based vaccines against hypervariable viruses. Immunol Cell Biol 89: 81-89.

42. Notredame C, Higgins DG, Heringa J (2000) T-Coffee: A novel method for fast and accurate multiple sequence alignment. J Mol Biol 302: 205-217.

43. Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305: 567-580.

44. Käll L, Krogh A, Sonnhammer EL (2004) A combined transmembrane topology and signal peptide prediction method. J Mol Biol 338: 1027-1036.