In-Silico Proteomic Exploratory Quest: Crafting T-Cell Epitope Vaccine Against Whipple's Disease

Amit Joshi[1] · Vikas Kaushik[1]

Abstract

Whipple's disease is one of the rare maladies in terms of spread but very fatal one as it is linked with many disorders (like Gastroenteritis, Endocarditis etc.). Also, current regimens include less effective drugs which require long duration follows up. This exploratory study was conducted to commence the investigation for crafting multi target epitope vaccine against its bacterial pathogen *Tropheryma whipplei*. The modern bioinformatics tools like VaxiJen, NETMHCII PAN 3.2, ALLERGEN-FP, PATCH-DOCK, TOXIC-PRED, MHCPRED and IEDB were deployed, which makes the study more intensive in analyzing proteome of *T. whipplei* as these methods are based on robust result generating statistical algorithms ANN, HMM, and ML. This Immuno-Informatics approach leads us in the prediction of two epitopes: VLMVSAFPL and IRYLAALHL interacting with 4 and 6 HLA DRB1 alleles of MHC Class II respectively. VLMVSAFPL epitope is a part of DNA-directed RNA polymerase subunit beta, and IRYLAALHL epitope is a part of membranous protein insertase YidC of this bacterium. Molecular-Docking and Molecular-Simulation analysis yields the perfect interaction based on Atomic contact energy, binding scores along with RMSD values (0 to 1.5 Å) in selection zone. The IEDB (Immune epitope database) population coverage analysis exhibits satisfactory relevance with respect to world population.

Keywords Epitopes · *Tropheryma whipplei* · Population coverage · Alleles · Simulation · Vaccine

Introduction

George Hoyt Whipple in 1907 explains Whipple's disease, as a multisystemic chronic infectious disease. He identified silver stained rod shaped bacterium in vacuoles associated with macrophages of patients, he initially did not think of them as the cause for the disease rather he think that intestinal lipodystrophy (Whipple's disease) was caused due to some novel disturbances in fat metabolic schemes (Whipple 1907). When the first successful treatment started by using antibiotics in 1952, determined that this bacterium might be the major causative agent of this disease (Paulley 1952). An electron microscopic study in 1960′s provided additional support for this hypothesis (Cohen et al. 1960;

Yardley and Hendrix 1961). Whipple's disease occurs uncommonly, as a multisystemic disorder (inexact annual frequency less than 1 per 1,000,000 populace) that specially affects middle-aged Caucasian men (Fenollar et al. 2007; Ramharter et al. 2014, Dobbins et al. 1981). This bacterium was found to mostly affect small children (Keita et al. 2015) and sewage workers (Schöniger-Hekele et al. 2007). Since, its first portrayal by Whipple in very beginning of first decade in twentieth century (Whipple 1907), a limited progresses with in pathogenesis, prognosis, and treatment of the malady have been made. The bacterium gets internalized in to lamina propria of intestine and then make its way to mucosal macrophages, as this bacterium induces the decreased expression of CD11b in such macrophages (CD11b on macrophages frequently mediates the intracellular degradation of bacteria) causes flip in the scenario (inappropriate antigen presentation by such macrophages and dendritic cells). This specially reasons the boom in IL-10, TGF-β and CCL-18 expression and decrease in IFN-γ, which in turn causes destroy in maturation of phagosomes and decrease in thioredoxin expression, lead them unable to kill bacterium and antigen presentation (Moss et al. 2006, 2010).

✉ Vikas Kaushik
vikas.14664@lpu.co.in

Amit Joshi
amit34655@gmail.com

1 Domain of Bioinformatics, School of Bio-Engineering and Bio-Sciences, Lovely Professional University, Punjab, India

An unseemly development of proficient antigen-presenting cells caused by the presence of interleukin 10 and interleukin 16, and the non appearance of interferon γ and interleukin 12 might lead to inadequate antigen-presentation and hinder the incitement of antigen-specific T-helper 1 cells enhancing growth and systemic spread of *Tropheryma whipple.* The nearby generation of provocative cytokines through macrophages and endothelial cells within the fringe might actuate lymphocyte invasion through a defective endothelial obstruction taken after by central aggravation, indeed in immunologically ensured tissues such as joints or the neuronal domain (Schneider et al. 2008).Currently hydroxychloroquine (600 mg/day) and doxycycline (200 mg/day) used for treatment of whipple's disease for 12–18 months, but life time follow up is required (Lagier et al. 2014), so it is time consuming treatment process and only few handful trials were conducted in earlier studies(Feurle et al. 2013). Nowadays epitope based vaccines provide better options in search of good treatment strategy for such type of harmful and rare malady, even if the individuals are genetically predisposed as in case of classical Whipple's disease(Trotta et al. 2017). This modern approach of putative vaccine determination which involves the use of proteomic databases is very handy and easy to use method not only for rare bacterial pathogens, but also very effective in case of harmful viruses like Nipah (Kaushik 2019). *Tropheryma whipplei* was found to be associated with major ailments like gastroenteritis and endocarditis (Fenollar et al. 2013).

In this research work, five proteins from proteomic data of *T. whipplei* were analyzed for allergenicity. Non-allergenic proteins were deployed for predicting epitopes. Predicted epitopes were subjected for immunogenic properties, structural modeling and the docking with corresponding MHC II alleles to investigate the strong binding affinity. Method is more economic, time efficient, and harmless when compared to the vaccine designing and testing in wet lab and animal

testing strategies (Kumar et al. 2015). Reverse vaccinology is the suitable approach as well as novel science method that use the genomic data with the utilization of computer for the arrangement of antibodies without culturing bacterium species (Kanampalliwar et al. 2013; Tang et al. 2012). It allow the choice in hands of human interface for selecting antigens from pathogenic set of DNA and most antigenic areas could be used to synthesize potential immunization to initiate defensive responses against such pathogenic species (Ada et al. 2018). Epitopes based antibodies selection and production is explicitly less time consuming, economical and considered safest approach in vaccine designing. Earlier computational methods were found to be successful in analyzing genome and prediction of putative drugs for *T. whipplei* (Palanisamy 2018), such studies provide motivation to craft vaccine targets by deploying in-silico approach. T-cell epitopes were screened out in this study may effectively elicit immune responses against this bacterium, and also similar type of recent study was found to be successful in determining epitope based vaccine agents for SARS-Cov2 (Joshi et al. 2020). Brief flow chart of the study used to determine putative epitope based vaccine candidates against *T. whipplei* is presented in Fig. 1.

Methodology

Retrieval of Proteins for *T. whipplei*

Proteomes were retrieved in fasta format from NCBI-Genbank and UniProtKB databases. Five proteins of different functionality were selected with following accession no's: WP_042507409.1 DNA-directed RNA polymerase subunit beta (RPO-B), WP_033800049.1 co-chaperone GroES, WP_038104819.1 TerC/Alx family metal homeostasis membrane protein, WP_042505650.1 membrane protein

Fig. 1 Flow chart of the study used to determine putative epitope
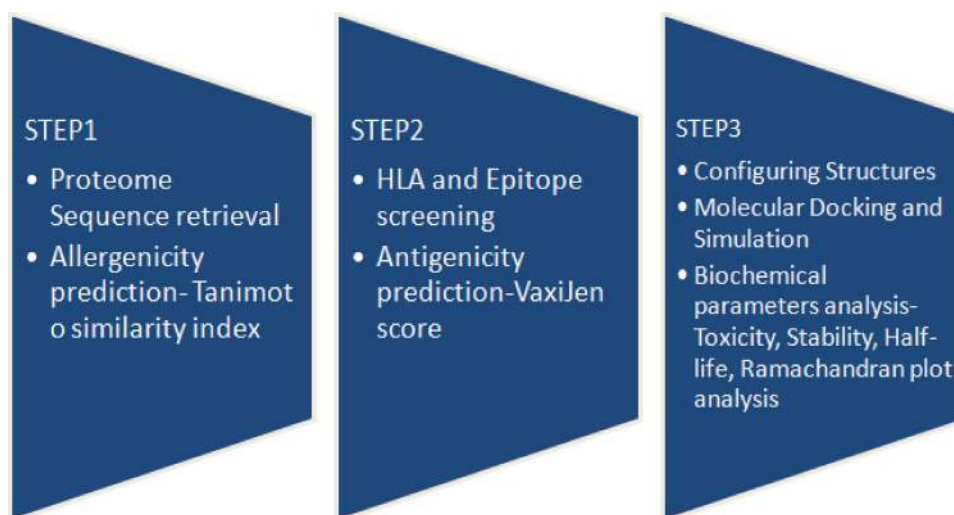
insertase YidC, WP_042505746.1 murein biosynthesis integral membrane protein. This selection depicts the variability to include crucial proteins of pathogenic domain (Table 1).

Allergenicity Prediction for Proteins

The protein sequences were then deployed for further analysis based on Allergen FP V 1.0 for predicting allergenicity (Dimitrov et al. 2014).

T-Cell Epitope Prediction

Net MHCII PAN 3.2 server is used to find and screen out HLA alleles which have good interaction with selected non-allergens of pathogenic origin (Jensen et al. 2018). To bring higher confidence in selecting epitope, VaxiJen server is deployed to determine antigenicity with threshold ≥ 0.7 for selected rare bacterium (Doytchinova et al. 2007). By subjecting proteomic sequences to Net MHCII PAN 3.2 server we obtained 1147 epitopes for WP_042507409.1, 90 epitopes for WP_033800049.1, 309 epitopes for WP_038104819.1, 302 epitopes for WP_042505650.1, and 510 epitopes for WP_011096746.1, this server was used because of its neural networking algorithm based approaches for fine predictions. 1-log50k (affinity score) ≤ 0.6 is used to screen out possible epitopes presented in Table 2. These epitopes were further subjected for antigenicity analysis based on VaxiJen scores.

Prediction of Epitope's Toxic Behavior

Toxicity for putative peptides was designated by using SVM scores from Toxin Pred web server (Gupta et al. 2013). Non toxic peptides were finalized for further analysis.

Table 1 Allergenicity results for analyzed proteins of *T. whipplei*

S. no	UniProtKB Accession No.	GenBank Accession No.	Allergen FP Score (Tanimoto similarity index)
1	Q76L83	WP_042507409.1	0.85 NON ALLERGEN
2	Q96P47	WP_033800049.1	0.81 NON ALLERGEN
3	P00846	WP_038104819.1	0.91 NON ALLERGEN
4	O14569	WP_042505650.1	0.83 NON ALLERGEN
5	Q7XBS0	WP_011096746.1	0.89 NON ALLERGEN

Molecular Structural Modeling for Epitopes and Probable HLA Alleles

The tertiary structure or 3D structure for epitope is determined by using PEP-FOLD 3 web server (Lamiable et al. 2016; Shen et al. 2014; Thévenet et al. 2012). And predicted Human leukocyte antigen alleles 3D structure was obtained from RCSB PDB database (Berman et al. 2000). Also Ramachandran plot analysis was conducted for verification of results by using Molprobity server (Williams et al. 2018).

Epitopes 3D Interaction or Molecular Docking with HLA Alleles

The docking experiments was conducted by using Patch-Dock tool (Schneidman-Duhovny et al. 2005), The predicted docked models of putative epitope and HLA alleles was selected on the basis of score, which relies on highest geometric shape complementarities and Atomic contact energy (Zhang et al. 1997). This allows the best selection of epitope and HLA allele interaction. This tool is easy to deploy for all life science domains.

Analysis of Population Coverage

Immune Epitope Database (IEDB) analysis Resource tool of population coverage was used to predict population coverage of the putative epitopes that are exhibiting interaction to HLA alleles and based on MHC-II restriction data (Bui et al. 2006). MHCPred tool was deployed for quantitative prediction of selected epitopes interacting to major Histocompatibility complexes (Guan et al. 2003).

Molecular Docking Simulations

Epitope-HLA allele docked sets were then used for simulation and dynamics analysis by deploying NAMD (Phillips et al. 2005) associated with VMD (Visual Molecular Dynamics) tool (Humphrey et al. 1996).

Results

Non-allergen Determination

Total 5 protein sequences were analyzed for allergenicity and depicted as non-allergen in Table 1 by using AllergenFP tool.

T-Cell Epitope Prediction

Net MHCII PAN 3.2 server is deployed to identify promiscuous epitopes and probable HLA alleles of MHC Class II

Table 2 List of predicted epitopes based on NetMHCII 3.2 server and VaxiJen score (threshold value of 0.7 and above was selected)

Protein ID	Allele (HLA)	POS	Peptide	1-log50k(aff)	VaxiJen	Antigen/non antigen
Q76L83	DRB1_0101	328	IRYLAALHL	0.581	0.9461	Antigen
	DRB1_0103	328	IRYLAALHL	0.311	0.9461	Antigen
	DRB1_0301	801	LSAEERLLR	0.271	0.2947	Non antigen
	DRB1_0401	225	FLRAIGMTD	0.292	− 0.3463	Non antigen
	DRB1_0404	328	IRYLAALHL	0.375	0.9461	Antigen
	DRB1_0405	328	IRYLAALHL	0.336	0.9461	Antigen
	DRB1_0406	328	IRYLAALHL	0.255	0.9461	Antigen
	DRB1_0701	328	IRYLAALHL	0.496	0.9461	Antigen
	DRB1_0802	324	IIATIRYLA	0.284	− 0.5053	Non antigen
	DRB1_1101	1010	YMYVLKLHH	0.451	1.2693	Antigen
	DRB1_1302	154	FVINGTERV	0.445	− 0.7514	Non antigen
Q96P47	DRB1_0101	84	YTILASRDV	0.410	0.2401	Non antigen
	DRB1_0406	86	ILASRDVLA	0.200	− 0.1277	Non antigen
P00846	DRB1_0101	222	FFSLTGLRQ	0.478	− 0.0649	Non antigen
	DRB1_0103	244	YMKFGVAAL	0.194	0.3077	Non antigen
	DRB1_0301	152	GLLDKVMIR	0.249	0.6173	Non antigen
	DRB1_0401	198	MFALDSIPA	0.343	0.4086	Non antigen
	DRB1_0404	295	IIALSVALS	0.339	0.8054	Antigen
	DRB1_0405	222	FFSLTGLRQ	0.350	− 0.0649	Non antigen
	DRB1_0406	295	IIALSVALS	0.237	0.8054	Antigen
	DRB1_0701	87	FRFAVPEIF	0.420	1.2093	Antigen
	DRB1_0802	18	MLVTVRRPA	0.314	− 0.5928	Non antigen
	DRB1_0901	87	FRFAVPEIF	0.407	1.2093	Antigen
	DRB1_0901	244	YMKFGVAAL	0.386	0.3077	Non antigen
	DRB1_1001	222	FFSLTGLRQ	0.477	− 0.0649	Non antigen
	DRB1_1101	18	MLVTVRRPA	0.402	− 0.5928	Non antigen
	DRB1_1302	159	IRMNVSKNY	0.420	0.7822	Antigen
	DRB1_1602	222	FFSLTGLRQ	0.335	− 0.0649	Non antigen
O14569	DRB1_0101	175	FYALQAGQA	0.539	0.6085	Non antigen
	DRB1_0301	278	LAFELRRKR	0.220	0.7910	Antigen
	DRB1_0401	175	FYALQAGQA	0.284	0.6085	Non antigen
	DRB1_0404	8	FLQNILLPI	0.299	0.1551	Non antigen
	DRB1_0405	8	FLQNILLPI	0.341	0.1551	Non antigen
	DRB1_1001	175	FYALQAGQA	0.467	0.6085	Non antigen
	DRB1_1101	63	FLKQIRAQR	0.438	− 0.0522	Non antigen
	DRB1_1302	8	FLQNILLPI	0.478	0.1551	Non antigen
Q7XBS0	DRB1_0101	374	YILQKAFYA	0.556	0.3071	Non antigen
	DRB1_0103	374	YILQKAFYA	0.274	0.3071	Non antigen
	DRB1_0401	374	YILQKAFYA	0.323	0.3071	Non antigen
	DRB1_0404	334	VLMVSAFPL	0.350	1.2114	Antigen
	DRB1_0405	374	YILQKAFYA	0.312	0.3071	Non antigen
	DRB1_0406	334	VLMVSAFPL	0.250	1.2114	Antigen
	DRB1_0701	334	VLMVSAFPL	0.539	1.2114	Antigen
	DRB1_0802	435	FLAIRVKLG	0.274	1.1141	Antigen
	DRB1_0901	334	VLMVSAFPL	0.436	1.2114	Antigen
	DRB1_1001	374	YILQKAFYA	0.473	0.3071	Non antigen
	DRB1_1101	503	YFLVITRCR	0.416	− 0.2632	Non antigen
	DRB1_1302	334	VLMVSAFPL	0.412	1.2114	Antigen
	DRB1_1602	374	VLMVSAFPL	0.366	0.3071	Non antigen

that interacts together by analyzing their 1-log50k values and binding affinities, then the VaxiJen scores were used with threshold of ≥ 0.7 with all informative details are presented in Table 2.

Molecular 3D Modeling of Epitopes and HLA Alleles

3D structural models of selected epitopes were designed by using PEP-FOLD 3 web server and than most common HLA

Table 3 HLA template model based on Pdb Id derived for MHC Class II alleles structure from RCSB-PDB

Allele name	Template structure (PDB ID)
HLA-DRB1*01_01	4AH2
HLA-DRB1*01_03	3PDO
HLA-DRB1*03_01	1A6A
HLA-DRB1*04_01	5LAX
HLA-DRB1*04_04	4IS6
HLA-DRB1*04_05	4IS6
HLA-DRB1*04_06	4IS6
HLA-DRB1*07_01	3C5J
HLA-DRB1*08_02	3PDO
HLA-DRB1*09_01	1BX2
HLA-DRB1*10_01	3PDO
HLA-DRB1*11_01	6CPL
HLA-DRB1*13_02	1FV1
HLA-DRB1*16_02	6CPO

DRB1 proteins structural models were derived by using RCSB-PDB database. In Table 3 PDB Id along with HLA alleles is exhibited. Molprobity results of Ramachandran plot analysis results shows satisfactory structural prediction (> 85% residues in favorable region) of epitopes that were finalized at last in Fig. 8.

Molecular Docking of Epitopes and HLA Alleles

PatchDock tool was deployed for interaction between selected structures of epitopes and HLA DRB1 proteins. Then interaction data produced by docked molecules include ACE (Atomic contact energy) and best model score that leads to the final selection in the way of prediction for each pair. In Table 4 the selected models and rejected models both were included to enhance the comparative analysis. The two selected epitopes were VLMVSAFPL and IRYLAALHL interacting with 4 and 6 HLA DRB1 alleles respectively. VLMVSAFPL epitope is a part of DNA-directed RNA polymerase subunit beta and IRYLAALHL epitope is a part of murein biosynthesis integral membrane protein of *T. whipplei* and are major identifiers of this bacterium. Figure 2 clearly depicts the good interaction between epitopes and HLA Alleles in docked results. In Fig. 2a Docked result of IRYLAALHL with HLA-DRB1* 01:01 exhibits perfect hydrogen bond due to presence of tyrosine residue in epitope at 3rd position, while most of the other non polar amino acids of this epitope are depicts vander waals interactions with in the HLA model and in Fig. 2c Docked result of VLMVSAFPL with HLA-DRB1* 04:04 exhibits perfect
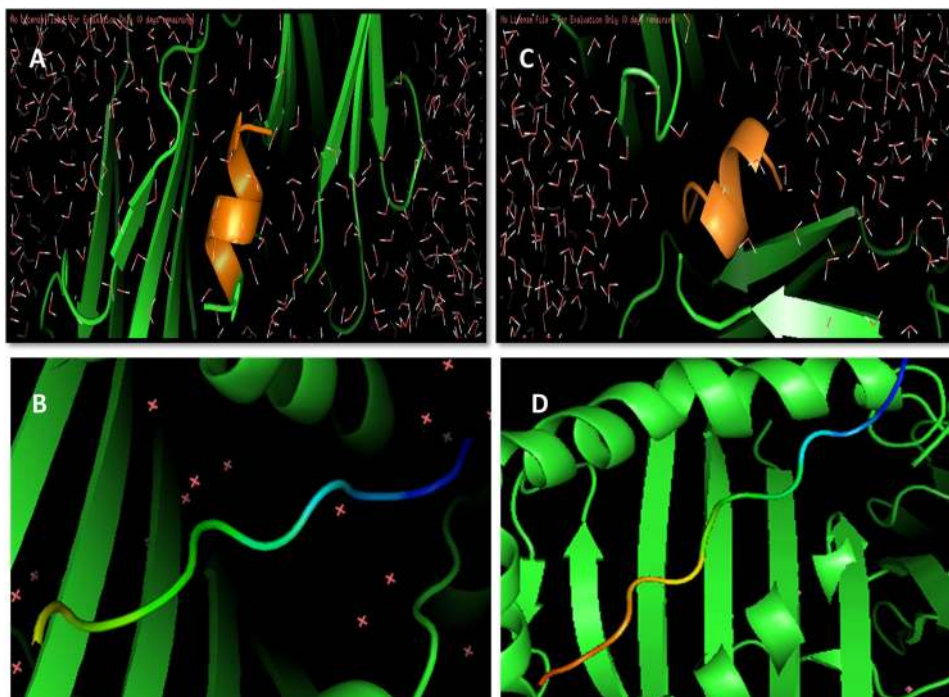
Table 4 Molecular docking results of screened epitopes with HLA alleles

Allele	Epitope	Binding score	Ace	Epitope selection
DRB1_0101 (4AH2)	IRYLAALHL	7704	− 367.28	Selected
DRB1_0103 (3PDO)	IRYLAALHL	7452	− 321.11	Selected
DRB1_0404 (4IS6)	IRYLAALHL	8064	− 353.51	Selected
DRB1_0405 (4IS6)	IRYLAALHL	8064	− 353.51	Selected
DRB1_0406 (4IS6)	IRYLAALHL	8064	− 353.51	Selected
DRB1_0701 (3C5J)	IRYLAALHL	7544	− 349.18	Selected
DRB1_1101 (6CPL)	YMYVLKLHH	7880	− 288.29	Rejected
DRB1_0404 (4IS6)	IIALSVALS	6562	− 136.21	Rejected
DRB1_0406 (4IS6)	IIALSVALS	6562	− 136.21	Rejected
DRB1_0701 (3C5J)	FRFAVPEIF	8260	− 223.61	Rejected
DRB1_0901 (1BX2)	FRFAVPEIF	9042	− 127.68	Rejected
DRB1_1302 (1FV1)	IRMNVSKNY	8564	344.23	Rejected
DRB1_0301 (1A6A)	LAFELRRKR	7746	− 93.04	Rejected
DRB1_0404 (4IS6)	VLMVSAFPL	6788	− 443.31	Selected
DRB1_0406 (4IS6)	VLMVSAFPL	6788	− 443.31	Selected
DRB1_0701 (3C5J)	VLMVSAFPL	6868	− 226.35	Selected
DRB1_0802 (3PDO)	FLAIRVKLG	7812	− 298.60	Rejected
DRB1_0901 (1BX2)	VLMVSAFPL	7960	− 297.99	Selected
DRB1_1302 (1FV1)	VLMVSAFPL	8500	104.94	Rejected

Fig. 2 **a** Docked result of IRYLAALHL with HLA-DRB1* 01:01 exhibits perfect hydrogen bond due to presence of tyrosine residue in epitope at 3rd position, while most of the other non polar amino acids of this epitope are depicting Vander waals interactions with in the HLA model. **b** MKMRMATPLLMQAL interacting with HLA-DRB1*01:01 (4AH2) structure considered as reference. **c** Docked result of VLMVSAFPL with HLA-DRB1* 04:04 exhibits perfect hydrogen bond due to presence of serine residue in epitope at 5th position, while most of the other non polar amino acids of this epitope are depicting Vander waals interactions with in the HLA model. **d** RQLYPEWTEAQRL epitope interacting with HLA-DRB* 04:04 (4IS6) structure considered as reference

hydrogen bond due to presence of serine residue in epitope at 5th position, while most of the other non polar amino acids of this epitope are depicts Vander waals interactions with in the HLA model. In Fig. 2b, d both RCSB-PDB structures for selected HLA DRB1 alleles (i.e. 4AH2 and 4IS6) with predicted peptide considered as reference structure (Schlundt et al. 2012; Chen et al. 2013). The reference docked peptides have great difference in amino acid sequence in comparison to our screened epitopes but exhibits some resemblance alike of our epitopes in interaction towards antigen binding pocket. Figure 3a, b represents the free undocked HLA-DRB1 receptors (4AH2, 4IS6 respectively), while Fig. 3c, d represents free unbound putative epitopes (IRYLAALHL, VLMVSAFPL respectively) and their side chains. Figure 4 graphically represents the selected epitopes and HLA alleles of MHC II on the basis of ACE values.

Toxicity, Half Life, and Stability Analysis of Putative Epitopes

Predicted epitopes VLMVSAFPL and IRYLAALHL have VaxiJen scores 0.9461 and 1.2114 respectively, they are also of non toxic nature as per the study of Toxin Pred tool and its toxicity scores (SVM score) represented in Table 5. In

Table 6 quantitative estimation of best interaction of epitope with HLADRB1 alleles were achieved with upright IC_{50} values by using MHCPred tool, this allows confidence of prediction. Table 7 shows Half-life and instability index for putative epitopes by deploying ProtParam expasy tool.

Population Coverage Analysis of Epitopes

VLMVSAFPL and IRYLAALHL manifest 28.82% and 37.06% elicitation of immune responsiveness by world population by availing IEDB tool. The epitopes VLMVS-AFPL and IRYLAALHL shows greater effect in European population by 29.63% and 42.68% respectively, and correspondingly similar results with North American population coverage analysis. This indicates its greater relevance in treatment of Whipple's disease as it is mostly seen in Caucasoid population. In Figs. 5 and 6 it is clearly represented in a graphical representation.
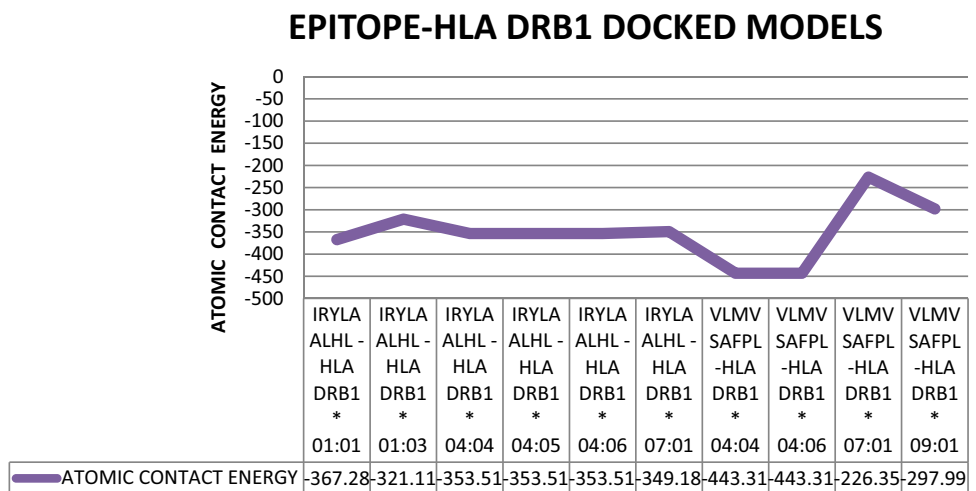
Molecular Dynamic Simulation Studies

NAMD was deployed for simulation studies on docked Epitope—HLA allele sets to obtain RMSD values. Maximum value of RMSD for VLMVSAFPL and IRYLAALHL

Fig. 3 Free HLA-DRB1 receptors and putative epitopes structure: **a** 4AH2 **b** 4IS6 **c** IRYLAALHL **d** VLMVSAFPL
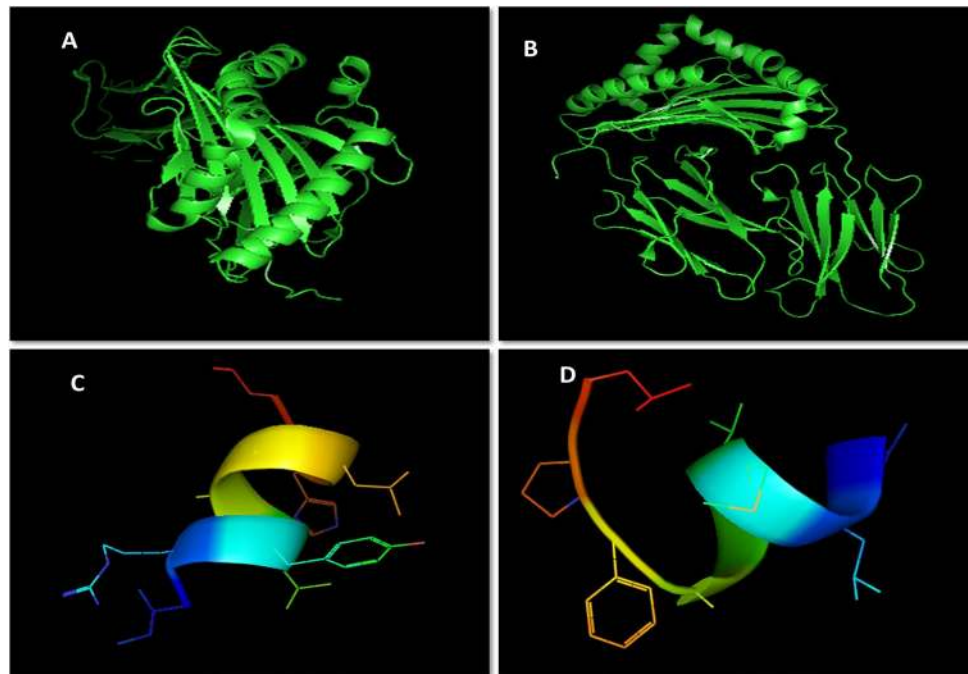
Fig. 4 Graphical representation of atomic contact energy (ACE) for docked complexes of putative epitopes and HLA DRB 1 alleles

EPITOPE-HLA DRB1 DOCKED MODELS

	IRYLA ALHL - HLA DRB1 * 01:01	IRYLA ALHL - HLA DRB1 * 01:03	IRYLA ALHL - HLA DRB1 * 04:04	IRYLA ALHL - HLA DRB1 * 04:05	IRYLA ALHL - HLA DRB1 * 04:06	IRYLA ALHL - HLA DRB1 * 07:01	VLMV SAFPL -HLA DRB1 * 04:04	VLMV SAFPL -HLA DRB1 * 04:06	VLMV SAFPL -HLA DRB1 * 07:01	VLMV SAFPL -HLA DRB1 * 09:01
ATOMIC CONTACT ENERGY	-367.28	-321.11	-353.51	-353.51	-353.51	-349.18	-443.31	-443.31	-226.35	-297.99

Table 5 Toxicity scores for putative epitopes

Epitope	No. of HLA binders	Toxicity score	Toxicity
VLMVSAFPL	4	− 1.19	Non toxin
IRYLAALHL	6	− 0.61	Non toxin

epitopes were analyzed, this gives more confidentiality in selection of vaccine candidate against *Tropheryma whipplei*. Figures 7 and 8 shows RMSD plots that indicates clear picture of selection of these two epitopes.

Discussion

Immuno-informatics is the suitable approach as well as novel science method that use the proteomic data with the utilization of computer systems for predicting epitopes without culturing bacterium species (Kanampalliwar et al. 2013; Tang et al. 2012). It allow the choice in hands of human interface for selecting antigens from pathogenic set of DNA and most antigenic areas could be used to synthesize potential immunization to initiate defensive responses against harmful pathogenic species (Ada et al. 2018). In-silico approach was earlier successful in case of *Staphylococcus aureus* (Delfani et al. 2015), *Mycobacterium*

**Table 6** MHCPRED results for interacting epitopes and HLA DRB1 alleles

Epitope	HLA allele used in test	Predicted -logIC$_{50}$ (M)	Predicted IC$_{50}$ value (nM)	Confidence of prediction (Max = 1)
IRYLAALHL	**DRB0101**	8.598	2.52	0.89
IRYLAALHL	**DRB0701**	5.935	1161.45	1.00
VLMVSAFPL	**DRB0701**	5.954	1111.73	1.00

Only two alleles were present in this database for MHC II i.e. DRB0101and DRB0701

Table 7 ProtParam tool used for predicting biochemical parameters (GRAVY, half-life, and instability index)

Epitope	Theoretical PI	Half life (for mammalian reticulocytes)	Instability index (percentage)	GRAVY
IRYLAALHL	8.75	20 h	0.51 (stable)	1.167
VLMVSAFPL	5.49	100 h	30.29 (stable)	2.233

Fig. 5 Graphical representation of Population coverage for VLMVSAFPL

POPULATION COVERAGE FOR VLMVSAFPL

Fig. 6 Graphical representation of population coverage for IRYLAALHL
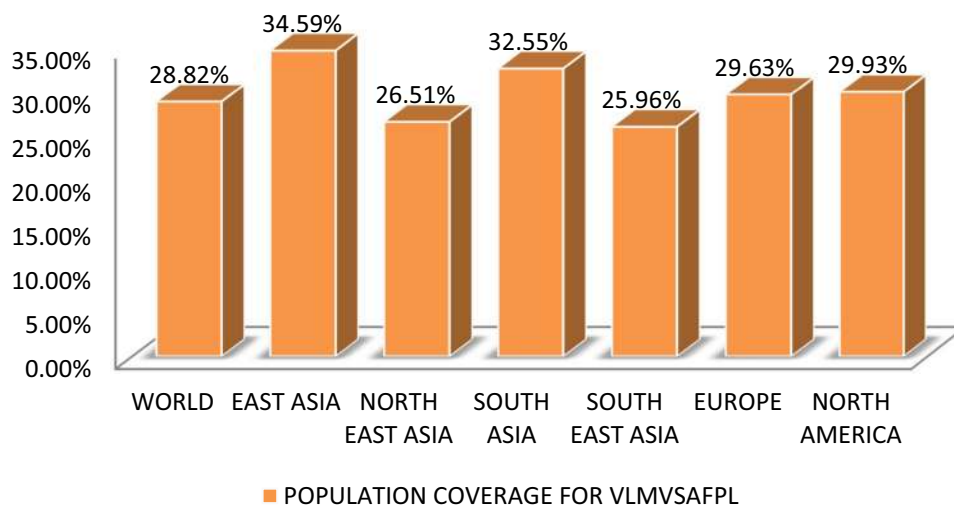

POPULATION COVERAGE FOR IRYLAALHL
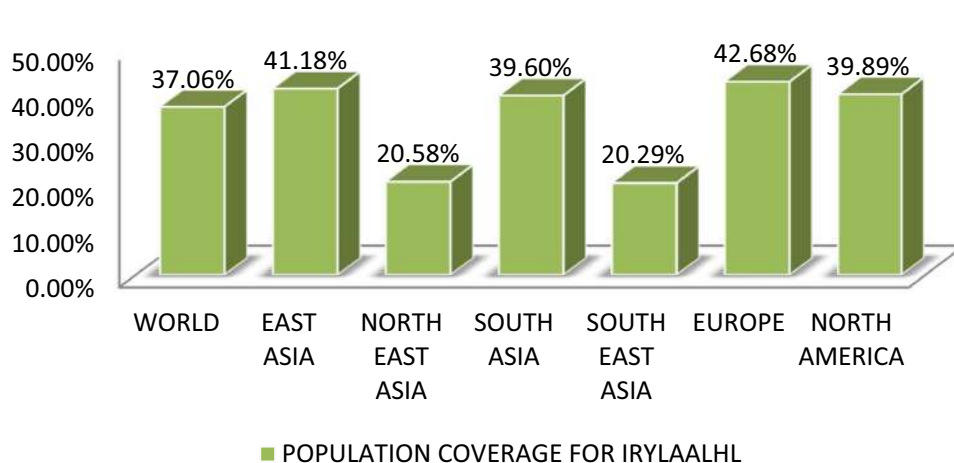
tuberculei (Mustafa 2013) and numerous bacterial species, but *T. whipplei* is still not fully explored in this domain. Current regimens include hydroxychloroquine and doxycycline for treatment of Whipple's disease for 12–18 months, but life time follow up is required (Lagier et al. 2014), so it is time consuming treatment process and
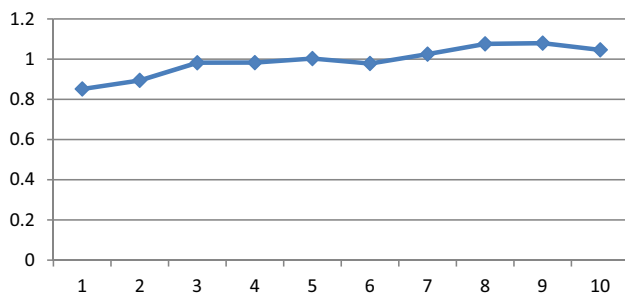
RMSD VALUE FOR IRYLAALHL
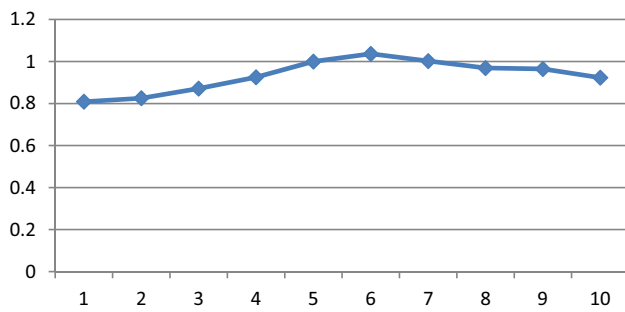
Fig. 7 Graphical representation of RMSD values for epitope IRY-LAALHL

RMSD VALUE FOR VLMVSAFPL

Fig. 8 Graphical representation of RMSD values for epitope VLM-VSFAPL

only few handful trials were conducted in earlier studies (Feurle et al. 2013).
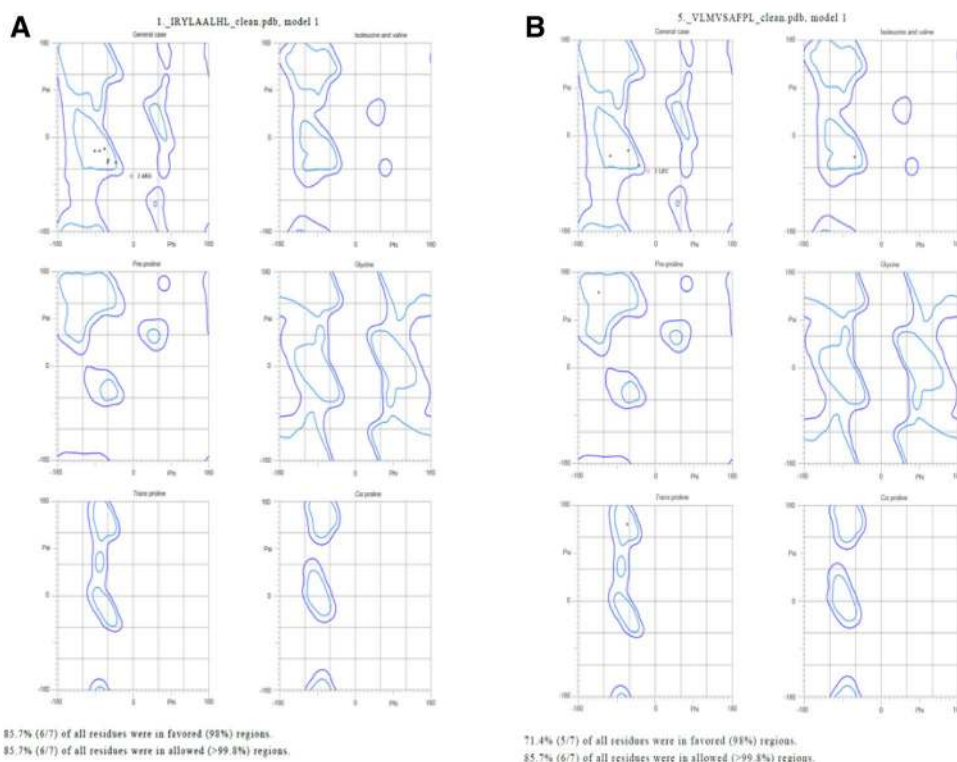
In present study we identified two possible epitopes that can interact with MHC-II alleles to elicit immune response on individuals namely VLMVSAFPL epitope (part of DNA-directed RNA polymerase subunit beta), and IRYLAALHL epitope (part of murein biosynthesis integral membrane protein). These epitopes exhibit better interaction with HLA DRB1 alleles, as confirmed by deploying Molecular-Docking and Molecular-Simulation studies (Adhikari et al. 2018). Population coverage analysis was found to be satisfactory and in earlier studies it was used in strengthening vaccine prediction aspects (Misra et al. 2011). Very similar studies were also conducted successfully for related bacterium *Mycobacterium avium* and found to be successful in predicting epitopes (Gurung et al. 2012). The epitope VLM-VSAFPL was found to interact with 4 HLA alleles of HLA-DRB1 domain (04:04, 04:06, 07:01, 09:01) with satisfactory ACE values ($-$ 443.3, $-$ 443.3, $-$ 226.3, $-$ 297.9 respectively); and the epitope IRYLAALHL found to interact with 6 alleles of HLA-DRB1 domain (01:01, 01:03, 04:04, 04:05, 04:06, 07:01) with satisfactory ACE values ($-$ 367.2, $-$ 321.1, $-$ 353.5, $-$ 353.5, $-$ 353.5, $-$ 349.1 respectively) in docking results similar type of methodology was seen in recent studies in screening epitopes for SARS-Cov-2 (Joshi et al. 2020). Both selected epitopes exhibit structural integrity as possess less than 35% instability index score, and half life greater than 20 h for mammalian reticulocytes, this makes the screening criteria more reliable. Also, more than 85% residues of selected epitopes come under favorable region in Ramachandran plot analysis (Fig. 9). Still no one has used vaccine based treatments for Whipple's disease, as it is thought to be rare and possess reduced genome but considered one of the harmful pathogen of human (Raoult et al. 2003; La Scola et al. 2001; Marth et al. 2016).The effectiveness of epitope based vaccines for treatment of endocarditis has already been claimed (Priyadarshini et al. 2014). But in our study we found the short peptides that can easily be synthesized and deployed in developing immunity in Caucasian populations against Whipple's disease.

Conclusion

In this study we obtained VLMVSAFPL and IRYLAALHL as predicted epitopes for vaccine crafting. This novel approach in crafting vaccine based treatment of *T. whipplei* will open new doors in research for creating regimens to treat such harmful bacterium by developing adaptive immune response and eradicating it globally before any future escalations takes place. The predicted epitopes can be deployed in crafting vaccines against *T. whipplei* bacterium after Molecular-wet lab corroboration.

Fig. 9 Ramachandran plot for putative epitopes **a** IRYLAALHL **b** VLMVSAFPL

85.7% (6/7) of all residues were in favored (98%) regions.
85.7% (6/7) of all residues were in allowed (>99.8%) regions.

71.4% (5/7) of all residues were in favored (98%) regions.
85.7% (6/7) of all residues were in allowed (>99.8%) regions.

References

Ada K, Candy C, Abu B, Abdul M, Chiuan CH, Boon HL, Chiuan YL (2018) Current progress of immunoinformatics approach harnessed for cellular and antibody-dependent vaccine design. Pathog Glob Health 112:3

Adhikari UK, Tayebi M, Rahman MM (2018) Immunoinformatics approach for epitope-based peptide vaccine design and active site prediction against polyprotein of emerging *Oropouche virus*. J Immunol Res. https://doi.org/10.1155/2018/6718083

Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE (2000) The Protein Data Bank. Nucleic Acids Res 28:235–242

Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A (2006) Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinform 7(1):153

Chen S, Li Y, Depontieu FR, McMiller TL, English AM, Shabanowitz J, Kos F, Sidney J, Sette A, Rosenberg SA, Hunt DF, Mariuzza RA, Topalian SL (2013) Structure-based design of altered MHC class II-restricted peptide ligands with heterogeneous immunogenicity. J Immunol 191:5097–5106. https://doi.org/10.4049/jimmunol.1300467

Cohen AS, Schimmel EM, Holt PR, Isselbacher KJ (1960) Ultrastructural abnormalities in Whipple's disease. Proc Soc Exp Biol Med 105:411–414

Delfani S, Imani Fooladi AA, Mobarez AM, Emaneini M, Amani J, Sedighian H (2015) In silico analysis for identifying potential vaccine candidates against *Staphylococcus aureus*. Clin Exp Vaccin Res 4(1):99–106. https://doi.org/10.7774/cevr.2015.4.1.99

Dimitrov I, Naneva L, Doytchinova I, Bangov I (2014) AllergenFP: allergenicity prediction by descriptor fingerprints. Bioinformatics 30:846–851. https://doi.org/10.1093/bioinformatics/btt619

Dobbins WO (1981) (1981) Is there an immune deficit in Whipple's disease? Dig Dis Sci 26:247–252

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

Fenollar F, Célard M, Lagier JC, Lepidi H, Fournier PE, Raoult D (2013) *Tropheryma whipplei* endocarditis. Emerg Infect Dis 19:1721–1730. https://doi.org/10.3201/eid1911.121356

Fenollar F, Puechal X, Raoult D (2007) Whipple's disease. N Engl J Med 356:55–66

Feurle GE, Moos V, Blaker H, Loddenkemper C, Morter A, Stroux A, Marth T, Schneider T (2013) Intravenous seftriaxone, followed by 12 or three months of oral treatment with trimethoprim-sulfamethoxazole in Whipple's disease. J Infect 66:263–270

Guan P, Doytchinova IA, Zygouri C, Flower DR (2003) MHCPred: a server for quantitative prediction of peptide—MHC binding. Nucleic Acids Res 31:3621–3624

Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R (2013) In silico approach for predicting toxicity of peptides and proteins. PLoS ONE 8:e73957

Gurung RB, Purdie AC, Begg DJ, Whittington RJ (2012) In silico identification of epitopes in *Mycobacterium avium* subsp. *paratuberculosis* proteins that were upregulated under stress conditions. Clin Vaccine Immunol 19:855–864. https://doi.org/10.1128/CVI.00114-12

Humphrey W, Dalke A, Schulten K (1996) VMD—visual molecular dynamics. J Mol Graph 14:33–38

Jensen KK, Andreatta M, Marcatili P, Buus S, Greenbaum JA, Yan Z, Sette A, Peters B, Nielsen M (2018) Improved methods for predicting peptide binding affinity to MHC class II molecules. Immunology. https://doi.org/10.1111/imm.12889

Joshi A, Joshi BC, Amin-ul Mannan M, Kaushik V (2020) Epitope based vaccine prediction for SARS-COV-2 by deploying immuno-informatics approach. Inform Med Unlocked. https://doi.org/10.1016/j.imu.2020.100338

Kanampalliwar AM, Soni R, Gridhar A, Tiwari A (2013) Reverse vaccinology: basics and applications. J Vaccines Vaccin 4(6):1–5

Kaushik V (2019) In Silico identification of epitope based peptide vaccine for Nipah virus. Int J Pept Res Ther. https://doi.org/10.1007/s10989-019-09917-0

Keita AK, Dubot-Pérès A, Phommasone K, Sibounheuang B, Vongsouvath M, Mayxay M, Raoult D, Newton PN, Fenollar F (2015) High prevalence of *Tropheryma whipplei* in Lao kindergarten children. PLoS Negl Trop Dis. https://doi.org/10.1371/journal.pntd.0003538

Kumar A, Hays M, Lim F, Foster LJ, Zhou M, Zhu G, Miesner T (2015) Protective enterotoxigenic *Escherichia coli* antigens in a murine intranasal challenge model. PLoS Negl Trop Dis. https://doi.org/10.1371/journal.pntd.0003924

La Scola B, Fenollar F, Fournier PE, Altwegg M, Mallet MN, Raoult D (2001) Description of *Tropheryma whipplei* gen. nov., sp. nov., the Whipple's disease bacillus. Int J Syst Evol Microbiol 51:1471–1479. https://doi.org/10.1099/00207713-51-4-1471

Lagier JC, Fenollar F, Lepidi H, Glorgi R, Million M, Raoult D (2014) Treatment of classical Whipple's disease: from in vitro results to clinical outcome. J Antimicrob Chemother 69:219–227

Lamiable A, Thévenet P, Rey J, Vavrusa M, Derreumaux P, Tufféry P (2016) PEP-FOLD3: faster de novo structure prediction for linear peptides in solution and in complex. Nucleic Acids Res 44(W1):W449–W454

Marth T, Moos V, Müller C, Biagi F, Schneider T (2016) *Tropheryma whipplei* infection and Whipple's disease. Lancet Infect Dis 16:e13–e22. https://doi.org/10.1016/S1473-3099(15)00537-X

Misra N, Panda PK, Shah K, Sukla LB, Chaubey P (2011) Population coverage analysis of T-Cell epitopes of *Neisseria meningitidis* serogroup B from iron acquisition proteins for vaccine design. Bioinformation 6(7):255–261. https://doi.org/10.6026/97320630006255

Moos V, Kunkel D, Marth T, Feurle GE, La Scola B, Ignatius R, Zeitz M, Schneider T (2006) Reduced peripheral and mucosal *Tropheryma whipplei* specific Th1 response in patients with Whipple's disease. J Immunol 177:2015–2022

Moos V, Schmidt C, Geelhaar A, Kunkel D, Allers K, Schinnerling K, Loddenkemper C, Fenollar F, Morter A, Raoult D, Ignautius R, Schneider T (2010) Impaired immune functions of monocytes and macrophages in Whipple's disease. Gastroenterology 138:210–220

Mustafa AS (2013) In silico analysis and experimental validation of *Mycobacterium tuberculosis*-specific proteins and peptides of *Mycobacterium tuberculosis* for immunological diagnosis and vaccine development. Med Princ Pract. https://doi.org/10.1159/000354206

Palanisamy N (2018) Identification of putative drug targets and annotation of unknown proteins in *Tropheryma whipplei*. Comput Biol Chem. https://doi.org/10.1016/j.compbiolchem.2018.05.024

Paulley JW (1952) A case of Whipple's disease (intestinal lipodystrophy). Gastroenterology 22:128–133

Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Christophe C, Skeel RD, Kalé L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26(16):1781–1802

Priyadarshini V, Pradhan D, Munikumar M, Swargam S, Umamaheswari A, Rajasekhar D (2014) Genome-based approaches to develop epitope-driven subunit vaccines against pathogens of infective endocarditis. J Biomol Struct Dyn 32(6):876–889

Ramharter M, Harrison N, Bühler T, Herold B, Lagler H, Lötsch F, Mombo-Ngoma G, Müller C, Adegnika AA, Kremsner PG, Makristathis A (2014) Prevalence and risk factor assessment of *Tropheryma whipplei* in a rural community in Gabon: a community based cross-sectional study. Clin Microbiol Infect 20:1189–1194. https://doi.org/10.1111/1469-0691.12724

Raoult D, Ogata H, Audic S, Robert C, Suhre K, Drancourt M, Claverie JM (2003) *Tropheryma whipplei* Twist: a human pathogenic Actinobacteria with a reduced genome. Genome Res 13:1800–1809. https://doi.org/10.1101/gr.1474603

Schlundt A, Günther S, Sticht J, Wieczorek M, Roske Y, Heinemann U, Freund C (2012) Peptide Linkage to the α-subunit of MHCII creates a stably inverted antigen presentation complex. J Mol Biol 423(3):294–302. https://doi.org/10.1016/j.jmb.2012.07.008

Schneider T, Moos V, Loddenkemper C, Marth T, Fenollar F, Raoult D (2008) Whipple's disease: new aspects of pathogenesis and treatment. Lancet Infect Dis 8:179–190

Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res 33:W363–W367. https://doi.org/10.1093/nar/gki481

Schöniger-Hekele M, Petermann D, Weber B, Müller C (2007) *Tropheryma whipplei* in the environment: survey of sewage plant influxes and sewage plant workers. Appl Environ Microbiol 73:2033–2035. https://doi.org/10.1128/AEM.02335-06

Shen Y, Maupetit J, Derreumaux P, Tufféry P (2014) Improved PEP-FOLD approach for peptide and miniprotein structure prediction. J Chem Theor Comput 10:4745–4758

Tang H, Liu XS, Fang YZ, Pan L, Zhang ZW, Zhou P, Lv JL, Jiang ST, Hu WF, Zhang P, Wang YL, Zhang YG (2012) The epitopes of foot and mouth disease. Asian J Anim Vet Adv 7:1261–1265

Thévenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tufféry P (2012) PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res 40:W288–293

Trotta L, Weigt K, Schinnerling K, Geelhaar-Karsch A, Oelkers G, Biagi F, Corazza GR, Allers K, Schneider T, Erben U, Moos V (2017) Peripheral T-cell reactivity to heat shock protein 70 and its cofactor GrpE from *Tropheryma whipplei* is reduced in patients with classical Whipple's disease. Infect Immun 85:e00363–e417

Whipple GH (1907) A hitherto undescribed disease characterized anatomically by deposits of fat and fatty acids in the intestinal and mesenteric lymphatic tissues. Bull Johns Hopkins Hosp 18:382–393

Williams CJ et al (2018) MolProbity: more and better reference data for improved all-atom structure validation. Protein Sci 27:293–315

Yardley JH, Hendrix TR (1961) Combined electron and light microscopy in Whipple's disease. Demonstration of "bacillary bodies" in the intestine. Bull Johns Hopkins Hosp 109:80–98

Zhang C, Vasmatzis G, Cornette JL, DeLisi C (1997) Determination of atomic desolvation energies from the structures of crystallized proteins. J Mol Biol 267(3):707–726

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.