

Characterization of structure, epitope and novel activities in multi-domain Bacterial Immunoglobulin-like (Big) proteins from *Leptospira*

A Thesis

Submitted for the Degree of

Doctor of Philosophy



Pankaj Kumar

(Regd. No. 15LBPH07)

Department of Biochemistry

University of Hyderabad

Hyderabad-500046

Telangana, India

July 2022

Characterization of structure, epitope and novel activities in multi-domain Bacterial Immunoglobulin-like (Big) proteins from *Leptospira*

A Thesis

Submitted for the Degree of

Doctor of Philosophy

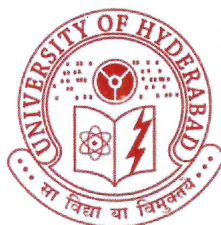
By

Pankaj Kumar

(Regd. No. 15LBPH07)



Department of Biochemistry
School of Life Sciences, University of Hyderabad
Hyderabad-500046
Telangana, India
July 2022



University of Hyderabad
Hyderabad -500046, India

Certificate

This is to certify that this thesis entitled “**Characterization of structure, epitope and novel activities in multi-domain Bacterial Immunoglobulin-like (Big) proteins from *Leptospira***” submitted by **Mr. Pankaj Kumar** bearing registration number **15LBPH07** in partial fulfilment of the requirement for the award of Doctor of Philosophy in the Department of Biochemistry, School of Life Sciences, is a bonafide work carried out by him under my supervision and guidance.

This thesis is free from plagiarism and has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma.

Part of this thesis has been presented in the following conferences:

- **Pankaj kumar** and Mohd. Akif, Identification of Epitopes on Lig Proteins: Strategy to Generate a Better Recombinant Vaccine against Leptospirosis. Presented Poster at International CCP4 workshop, 22-26 Oct, 2018, IMTech., Chandigarh
- **Pankaj Kumar** and Mohd. Akif, Stability of Lig Domains Provides a Rational for Engineering an Improved Scaffold for Recombinant Vaccine against Leptospirosis. Presented poster at Bioquest, 12-13 Oct, 2017 University of Hyderabad, Hyderabad

Published in the following journals:

- **Kumar P.**, Lata S., Shankar U.N., and Akif M., (2021) Immunoinformatics Based Designing of a Multi-Epitope Chimeric Vaccine From Multi-Domain Outer Surface Antigens of *Leptospira*. *Frontiers in Immunology*. 12:735373. (Chapter 4)
- **Kumar P.**, Chang YF, and Akif M. (2022). Characterization of Novel nuclease and protease activities among Leptospiral immunoglobulin-like proteins. *Arch Biochem Biophys*. 2022.109349 (Chapter 5) [In Press]
- Khan MA, **Kumar P.**, Akif M., and Miyoshi H (2021) Phosphorylation of eukaryotic initiation factor eIF5a4E enhances the binding rates to VPg of turnip mosaic virus. *PLoS ONE* 16(11): e0259688.
- Shiraz M, Lata S, **Kumar P.**, Shankar UN, and Akif M. (2021) Immunoinformatics analysis of antigenic epitopes and designing of a multi-epitope peptide vaccine from putative nitro-reductases of Mycobacterium tuberculosis DosR. *Infect Genet Evol.* Oct; 94:105017.
- **Kumar P.**, Shiraz M., and Akif M., (2022) Multi-epitope-based vaccine design by exploring antigenic potential among leptospiral lipoproteins using comprehensive immunoinformatics and structure-based approaches. *Biotechnology and Applied Biochemistry* (Under revision).

Under communication:

- **Kumar P.**, Chang YF, and Akif M. (2022). Crystal structure of two domains segment from the variable region of *Leptospira* host-interacting outer surface protein, LigA. *International Journal of Biological Macromolecules* (Chapter 3).

Further, the student has passed the following courses towards fulfilment of the coursework requirement for Ph.D.

Course code	Name	Credits	Pass/Fail
BC 801	Analytical Techniques	4	Passed
BC 802	Research ethics, Data analysis and Biostatistics	3	Passed
BC 803	Lab seminar and Records	5	Passed


Supervisor
MOHD. AKIF Ph.D.

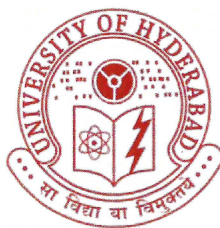
Assistant Professor
DEPARTMENT OF BIOCHEMISTRY
SCHOOL OF LIFE SCIENCES
UNIVERSITY OF HYDERABAD
HYDERABAD-500 046. (INDIA)


Head, Dept. of Biochemistry
HEAD

Dept. of Biochemistry
SCHOOL OF LIFE SCIENCES
UNIVERSITY OF HYDERABAD
HYDERABAD-500 046.


Dean, School of Life Sciences
DEAN

School of Life Sciences
University of Hyderabad
Hyderabad-500 046.



University of Hyderabad
Hyderabad -500046, India

Declaration

I, **Pankaj Kumar**, hereby declare that this thesis entitled “**Characterization of structure, epitope and novel activities in multi-domain Bacterial Immunoglobulin-like (Big) proteins from *Leptospira***” submitted by me under the guidance and supervision of **Dr. Mohd. Akif**, is original and independent research work. I also declare that it has not been submitted previously in part or in full to this University or any other University or Institution for the award of any degree or diploma.

Pankaj K.

Date: 18/07/22

Signature of the Student

Mohd. Akif

Signature of the Supervisor
MOHD. AKIF Ph.D.
Assistant Professor
DEPARTMENT OF BIOCHEMISTRY
SCHOOL OF LIFE SCIENCES
UNIVERSITY OF HYDERABAD
HYDERABAD-500 046. (INDIA)

Trust the process.

*Your time is coming, just do the work
and the results will handle themselves!*

Dedicated to My Maa and Papa

ACKNOWLEDGMENTS

Countless people have supported my efforts in completing this thesis, and I extend my sincerest thanks to all of them.

First and foremost, I am incredibly grateful to my supervisor, **Dr. Mohd. Akif** for his continuous support and invaluable advice. I am highly grateful to him for allowing me to work under his supervision and showing continuous faith in me over the years. He has been a constant source of inspiration throughout my academic research and daily life. He gave me the freedom to design and plan experiments, which gave me a lot of confidence to develop my scientific skills. He was always available for the scientific discussions related to my experiments, manuscript, and thesis writing despite his busy schedule. I always remember his encouraging words that doing a Ph.D. is like being in a sea; you have to learn to swim and come out with hidden gems. Thank you, sir, for your great patience, motivation, and enthusiasm all throughout these years.

I extend my sincere thanks to my doctoral committee members, **Dr. Santosh Kumar Padhi** and **Dr. Insaf Ahmed Qureshi**, for their critical suggestions and important comments during the DRC meetings that helped to shape my thesis.

I would also like to thank **Dr. Irfan Ahmed Ghazi** and **Dr. Insaf Ahmed Qureshi** for allowing me to use their lab facilities for my experiments.

I want to thank the present and former Head of the Biochemistry Department, **Prof. Krishnaveni Mishra**, **Prof. Mrinal Kanti Bhattacharyya**, and **Prof. N. Siva Kumar** for maintaining the department facilities. I am also extremely thankful to the present Dean of the School of Life Sciences, **Prof. N. Siva Kumar** and the former Deans, **Prof. S Dayananda**, **Prof. KVA Ramiah**, **Prof. P Reddanna**, and **Prof. MNV Prasad** for the school research facilities.

I sincerely thank **Mr. Prabhu**, **Mr. Chary**, and all other non-teaching staff of the Department of Biochemistry for their administrative help and cooperation.

I am also thankful to all the teaching and non-teaching staff of the School of Life Sciences for their timely help.

I would like to extend my sincere gratitude to **Dr. R Sankaranarayanan** (CCMB, Hyderabad) and **Prof. B Gopal** (IISc. Bangalore) for generously allowing me to use the X-Ray diffraction facilities.

I owe a deep sense of gratitude to our collaborators, **Dr. Yung-Fu Chang** (Cornell University, USA) and **Dr. S M Faisal** (NIAB, Hyderabad), for sharing the reagents and clones used in my thesis.

I greatly acknowledge UGC-RGNF for providing me the financial support (JRF and SRF) for five years of my Ph.D.

I also thank the funding agencies DST-SERB, IoE-Grant, and DBT for funding the Lab and DST-FIST and UGC-SAP for funding the department.

I would also extend my thanks to CMSD, University of Hyderabad, for providing their facilities to run simulations.

During my journey of Ph.D., I was fortunate to have great lab members who provided numerous assistance whenever required. I am grateful to my friends cum lab mates, **Nachiket**, **Shiraz**, and **Kousamvita**, for maintaining a friendly and peaceful work environment in the Lab. I am also thankful to present and former M.Sc. students of our lab **Anusha**, **Priyanka**, **Pranati**, **Pranita**, **Pradeep**, **Sree Priya**, **Haneesh**, **Yash**, **Nikita**, and **Aayushi** for their cooperation throughout my work.

Many thanks to **Nachiket** and **Shiraz** for the endless scientific discussions and their help in my professional and personal life. I had the best tea breaks in my life with them. I thank **Nachiket** for helping me out with the CD experiments. I also extend my heartfelt thanks to **Shiraz** for his constant support, countless food trips, and late-night bike rides to Charminar. I would also like to thank him for helping me to carry the crystals to IISc. Bangalore. They were great colleagues to work with.

I would like to thank my friend, **Pratul** for being such an amazing friend and for helping me during my visit to Bangalore for X-ray data collection.

I am thankful to **Vivek Bhaiya** (NIAB) for sharing clones and providing the genomic DNA of *Leptospira*.

I would like to thank all my Ph.D. colleagues, with whom I have shared moments of anxiety but also big excitement **Vishnu Priya**, **Srinivas**, **Kiran Anna**, **Dr. Arshad**, **Dr. Nisha**, **Rutuparna**, **Kesaban**, and **Nivedita**.

I am also grateful to **Dr. Santhosh**, **Dr. Sayanna** and **Pranay** for clearing my experiment-related doubts during the initial days of my Ph.D. I am also thankful to **Janish** for assisting me with my CD experiment in the Biotechnology department.

I would like to thank my previous roommate, **Ashesh Baidya**, for being a good friend and for preparing tasty foods on weekends.

My sincere gratitude goes out to **Rahat ma'am** for being a family away from home and for treating us with homemade foods.

I would like to thank my cricket team **Natural Selection** that I have been a part of since my M.Sc. days. I would also like to thank my team members **Rohith Anna**, **Thirumala Anna**, **Akash Bhaiya**, and **Gopal Bhaiya** for having a great time at the cricket ground every weekend.

I would like to thank my childhood school friends **Manish**, **Akash**, and **Kundan**, who has always been a significant source of support when things got a little tricky.

I am grateful to my college friends and hostel companions **Nikhil, Nilesh, Ravi, Anil Bhaiya, Ashish, Sumit, Avash, Pramod, Ramesh** and **Mukesh** for always supporting and encouraging me.

I cannot forget my dear friends who went through hard times together, cheered me on, and celebrated each accomplishment: **Wahida, Payal, and Ramji**. They have been a tremendous support throughout and are more like a family to me now.

Wahida deserves special thanks for always being there for me in all situations. She is an exceptional friend and supporter whom I have known since my graduation days. I must acknowledge her for always motivating me, helping me with my Pre Ph.D. slides, and listening to my presentation rehearsals. I am grateful to her for all the discussions we had about troubleshooting the failed experiments.

I thank **Wahida** and **Payal** for their support and motivation. I appreciate the time they spent with me on food outings, shopping, and movies during the weekends. The festival celebrations with them are some of the best times of my life during my stay in Hyderabad.

Many thanks to **Ramji** for always encouraging me. He has been like an elder brother who always stood beside me no matter what.

This acknowledgment would not be complete without mentioning the efforts and patience of my families. I sincerely thank my Papa, **Mr. Surbdeo Manjhi**, and my Maa, **Mrs. Savitri Devi**, for their unconditional trust, love, timely encouragement, tremendous support, and endless patience.

I would like to thank my younger sister, **Chanchala** for her love, care, and support. I thank her for helping me financially during the crisis by sharing her part of pocket money. Thanks to my younger brother, **Rohit** for patiently listening to my experimental plans and taking interest in my long-term future goals. I thank them for reading my articles and proudly sharing them. I would like to thank my cousin brother, **Kissu** for always being a good son to my mother in my absence. I would like to thank my late grandparents for their teachings during my childhood.

Pankaj

Table of Contents

	Page No.
List of Figures	I
List of Tables	V
Abbreviations	VII
Unit of measurements	IX
Preface	XI
CHAPTER 1	1-33
Introduction and review of literature	
1.1. Introduction	2
1.2. History	3
1.3. <i>Leptospira</i> Biology	4
1.3.1. General characteristics	4
1.3.2. <i>Leptospira</i> Classification	6
1.3.2.1. Taxonomical classification	6
1.3.2.2. Serological classification	6
1.3.2.3. Genotypic classification	6
1.3.3. Growth condition	7
1.4. Epidemiology	7
1.4.1. Human leptospirosis	8
1.4.2. Cycle of infection	9
1.4.3. Risk factors	9
1.4.4. Clinical manifestation	10
1.5. Global burden	11
1.5.1. Leptospirosis in India	12
1.5.2. Diagnostic challenges	13
1.6. Pathogenesis	13
1.6.1. Attachment and Invasion	14

1.6.2.	Lipopolysaccharides (LPS)	16
1.6.3.	Leptospiral lipoprotein (LipL) Proteins	16
1.6.4.	<i>Leptospira</i> endostatin-like protein (Len) and Loa22	17
1.6.5.	Hemolysins and HemO	17
1.6.6.	Leptospiral Immunoglobulin-like protein (Lig)	18
1.7.	Host Immune response against leptospirosis	18
1.8.	Immune evasion strategies	19
1.9.	<i>Leptospira</i> vaccines	21
1.10.	Leptospiral Immunoglobulin-like protein (Lig)	23
1.10.1.	Domain organization of Lig	23
1.10.2.	Structure	25
1.10.3.	Calcium-binding	26
1.10.4.	Lig in leptospiral pathogenesis	26
1.10.5.	Lig in Immune evasion	28
1.10.6.	Lig as a diagnostic marker and vaccine candidate	29
1.11.	Major challenge and aim of the thesis	31
1.12.	Objectives of the study	33

CHAPTER 2 **34-52**

Cloning, overexpression and purification of Lig and its fragments

2.1.	Introduction	35
2.2.	Materials	36
2.2.1.	Chemical and Reagents used in Cloning	36
2.2.2.	Reagents used in protein purification and detection	36
2.2.3.	Recipes of reagents used	37
2.3.	Experimental Procedures	38
2.3.1.	Preparation of Competent cells	38
2.3.2.	Cloning of fragments of <i>ligA</i> and <i>ligB</i> into the expression vector	39
2.3.3.	Transformation	42
2.3.4.	Expression and purification of all Lig constructs	43

2.3.5.	Size-exclusion chromatography	44
2.4.	Results	45
2.4.1.	Cloning of LigA and LigB fragments	45
2.4.2.	Recombinant purification of LigA fragments	48
2.4.2.1.	Single Ig-like domain	48
2.4.2.2.	Multiple Ig-like domains	49
2.4.2.	Recombinant purification of LigB fragments	50
2.5.	Conclusions	52
CHAPTER 3		53-88

Crystallization and structure determination

3.1.	Introduction	54
3.2.	Materials	56
3.3	Experimental Procedures	56
3.3.1.	Protein purification	56
3.3.2.	Crystallization Screening	57
3.3.3.	Crystallization optimization	58
3.3.4.	X-ray data collection and processing	58
3.3.5.	Structure determination and model building	58
3.3.6.	Validation and deposition of LigA8-9 structure	59
3.3.7.	Molecular docking with the ECM	59
3.4.	Results	59
3.4.1.	Crystallization screening	59
3.4.2.	Crystallization optimization	66
3.4.3.	X-ray data collection and processing	66
3.4.4.	Data Processing of LigA8-9	69
3.4.5.	Structure solution and refinement of LigA8-9 data	71
3.4.6.	Overall structure	74
3.4.7.	Conserved Hydrophobic core	76
3.4.8.	Dimerization of LigA8-9	77

3.4.9.	Binding of metals with the LigA8-9 structure	79
3.4.10.	Structural comparison among two domains	82
3.4.11.	Structural comparison of LigA8-9 with other Big domains	82
3.4.12.	Domain orientation in LigA8-9	85
3.4.13.	Interaction with fibronectin	87
3.5.	Conclusions	88
CHAPTER 4		89-117

Identification of antigenic epitopes within Lig proteins

4.1.	Introduction	90
4.2.	Methodology	92
4.2.1.	Protein sequence retrieval	92
4.2.2.	Prediction of BCL epitopes	94
4.2.3.	CTL and HTL Epitope prediction	95
4.2.4.	Removal of self-peptides and selection of IFN- γ epitopes	96
4.2.5.	Epitope selection and construction of a multi-epitopes vaccine	96
4.2.6.	Immunoinformatic and physicochemical analysis of the vaccine	96
4.2.7.	Prediction of secondary and tertiary structure	97
4.2.8.	Molecular docking of vaccine construct with immune receptor	97
4.2.9.	Molecular Dynamics Simulation studies	97
4.2.10.	Immune simulation	98
4.2.11.	<i>In-silico</i> cloning	99
4.3.	Results	99
4.3.1.	Collection and primary sequence analysis	99
4.3.2.	Identification of linear and conformational BCL epitopes	99
4.3.3.	Identification of CTL and HTL Epitopes	103
4.3.4.	Generation of multi-epitope chimeric vaccine	104
4.3.5.	Physicochemical features of chimeric vaccine construct	106
4.3.6.	Secondary and Tertiary structure prediction	107
4.3.7.	Docking complex of final vaccine construct with immune receptor	109

4.3.8.	Molecular Dynamics Simulation	110
4.3.9.	Immune response profile	114
4.3.10.	<i>In-silico</i> cloning	115
4.4.	Conclusions	117
CHAPTER 5		118-149
Characterization of novel nuclease and protease activities among Lig		
5.1.	Introduction	119
5.2.	Materials	121
5.3.	Experimental Procedures	121
5.3.1.	Nuclease activity assay	121
5.3.2.	Effect of divalent ions on nuclease activity	122
5.3.3.	Intrinsic fluorescence measurement	122
5.3.4.	Site-directed mutagenesis	122
5.3.5.	Purification of the mutant protein	124
5.3.6.	Homology modeling and molecular docking	125
5.3.7.	Protease assay	125
5.3.8.	Mass spectrometry analysis of β -casein cleavage product	125
5.4.	Results	126
5.4.1.	LigA domains possess novel Metal-dependent nuclease activity	126
5.4.2.	Preference of divalent ions on the LigA nuclease activity	132
5.4.3.	LigA7 nuclease activity on single-stranded DNA and RNA	134
5.4.4.	Binding of Mg^{2+} ion with LigA7	137
5.4.5.	Effect of LigA7 point mutations on the nuclease activity	138
5.4.6.	Molecular docking study reveals potential DNA binding surface	140
5.4.7.	LigA7 also possesses novel protease activity	141
5.5.	Proposed model of the novel activity	148
5.6.	Conclusions	149

CHAPTER 6	149-159
Discussion	150
BIBLIOGRAPHY	159
APPENDIX	194
PUBLICATIONS	214
ANTI-PLAGIARISM REPORT	219

List of Figures

	Page No.
Figure 1.1	Brief history related to Leptospirosis disease and <i>Leptospira</i> 4
Figure 1.2	Architecture of <i>Leptospira interrogans</i> 5
Figure 1.3	Phylogenetic analysis of <i>Leptospira</i> strains 7
Figure 1.4	Transmission cycle of leptospirosis 10
Figure 1.5	Global burden of leptospirosis in the context of DALY index 11
Figure 1.6	Schematic representation of the <i>Leptospira</i> membrane architecture 15
Figure 1.7	Complement evasion strategies and the associated Proteins of pathogenic <i>Leptospira</i> 20
Figure 1.8	Schematic representation of the domain organization of Leptospiral immunoglobulin-like protein (Lig) 24
Figure 1.9	Solution structure of LigB12 and LigA4 25
Figure 1.10	Lig plays multiple roles during <i>Leptospira</i> pathogenesis 29
Figure 1.11	Summary of Lig proteins and their association at different stages during leptospirosis 30
Figure 2.1	Protein purification strategy 44
Figure 2.2	Schematic representation of the gene fragments selected for the generation of truncated proteins 45
Figure 2.3	1% Agarose gel showing PCR amplification of different Lig fragments 46
Figure 2.4	1% Agarose gel showing double digestion of different <i>LigA</i> fragments from the expression vector 47
Figure 2.5	1% Agarose gel showing double digestion of different LigB fragments from the expression vector 47
Figure 2.6	SDS-PAGE showing the purification LigA fragments 48
Figure 2.7	SDS-PAGE showing the purification of multiple Ig-like fragments of LigA 49
Figure 2.8	SDS-PAGE profile showing the purification of Ig-domains of LigB 50
Figure 3.1	The experimental approach followed for protein crystallization and structure determination 57
Figure 3.2	Crystallization condition obtained for LigA8-13 fragment 60
Figure 3.3	Crystallization condition obtained for LigB7-12 fragment of LigB 61

Figure 3.4	Crystallization condition obtained for Ig-like domains LigA8-12	62
Figure 3.5	Crystallization condition obtained for Ig-like domains	62
Figure 3.6	Crystallization condition obtained for LigA8-9 fragment	63
Figure 3.7	Crystallization condition obtained for LigA12-13 fragment	64
Figure 3.8	Crystallization condition obtained for LigB11-12 fragment	64
Figure 3.9	Crystallization condition obtained for LigA9 single Ig-like domain of LigA	65
Figure 3.10	Crystallization condition obtained for Lig A13 fragment	65
Figure 3.11	Crystals obtained after optimization	66
Figure 3.12	X-Ray diffraction pattern observed for LigB7-12	67
Figure 3.13	Diffraction pattern of one of the LigA8-9 crystals covering 1° oscillation range and diffracted up to 2.33 Å	68
Figure 3.14	Diffraction pattern of one of the LigA8-9 crystals covering 1° oscillation and diffracted up to 1.86Å	68
Figure 3.15	Cumulative distribution function of the centric and acentric reflections observed in the LigA8-9 data	70
Figure 3.16	Wilson plot for LigA8-9 data extending to 1.86Å	71
Figure 3.17	A schematic picture of LigA protein domains organization, indicating constant and variable domains	72
Figure 3.18	Ramachandran plot for the final structure of LigA8-9 fragment	73
Figure 3.19	The crystal structure LigA8-9 subunit	75
Figure 3.20	Hydrophobic core in Ig-like domain	76
Figure 3.21	Structure of LigA8-9 dimer	78
Figure 3.22	Representation of dimer structure of LigA8-9 with potassium ions	80
Figure 3.23	Representation of dimer structure of LigA8-9 with calcium ions	81
Figure 3.24	Comparison among LigA8 and LigA9 domains	82
Figure 3.25	Sequence comparison between LigA8 and its closest structural homolog	83
Figure 3.26	Sequence comparison between LigA9 and its closest structural homolog	83
Figure 3.27	Structural comparison with Ig-like domain-containing proteins	84
Figure 3.28	Size and orientation of LigA8-9	85
Figure 3.29	Proposed orientation of LigA variable domain	86
Figure 3.30	Docking GBD domain of fibronectin with the structure of LigA8-9	87

Figure 4.1	Flowchart of the methodology used in this study	92
Figure 4.2	Homologous models of single variable LigA Ig-like domains and all individual LigB Ig-like domains	101
Figure 4.3	Schematic representation of the multi-epitope vaccine construct	106
Figure 4.4	Secondary structure content of the vaccine construct	108
Figure 4.5	(A) 3-D model of vaccine construct. (B) Plot showing validation of 3-D model by Ramachandran and ProSA server. (C) Promiscuous epitopes used in the vaccine construct are mapped with same colour code on the Ig-like Domains of Lig proteins	109
Figure 4.6	(A) Represent a docked complex of vaccine construct and TLR4 and shown as marine blue cartoon and magenta surface, respectively. (B) 2-D representation of interactions between vaccine construct and TLR4. Red arcs represent salt bridges and green dotted lines represent hydrogen bonds	111
Figure 4.7	MDS plots of vaccine construct and TLR4 complex	113
Figure 4.8	(A) Separate RMSD plot for the vaccine construct and TLR4 at 100ns (B) RMSF plot for vaccine construct only at 50ns of MDS (C) Separate Rg plot for TLR4 and vaccine construct at 100ns of MDS	114
Figure 4.9	Prediction of immune response against the generated chimeric construct	115
Figure 4.10	Representation of In-silico cloning of multi-epitope vaccine construct	116
Figure 5.1	DNA cleavage activity in LigA7 on supercoiled double-standard DNA (pET28a)	126
Figure 5.2	Agarose gel showing metal-dependent DNA cleavage activity of LigA7 in the presence of both Mg^{2+} ions and EDTA	127
Figure 5.3	Agarose gel showing metal-dependent DNA cleavage activity of boiled LigA7	128
Figure 5.4	Agarose gel showing metal-dependent DNA cleavage activity of boiled LigA7 in the presence of SDS	129
Figure 5.5	Agarose gel showing metal-dependent DNA cleavage activity of different variable single domains of LigA	129
Figure 5.6	Agarose gel showing metal-dependent DNA cleavage activity of LigA12-13	130
Figure 5.7	Agarose gel showing metal-dependent DNA cleavage activity LigA11-13	130
Figure 5.8	Agarose gel showing metal-dependent DNA cleavage activity LigA10-13	131
Figure 5.9	Agarose gel showing metal-dependent DNA cleavage activity LigA9-13	131
Figure 5.10	Agarose gel showing metal-dependent DNA cleavage activity LigA8-13	132
Figure 5.11	1% agarose gel showing DNA cleavage activity of LigA7 on circular double-strand DNA in the presence of different divalent ions	133
Figure 5.12	Representation of DNase activity in LigA7 with pH variation	134

Figure 5.13	Agarose gel showing metal ion depended on DNA degradation of ssDNA by LigA7	135
Figure 5.14	Agarose gel showing metal ion-dependent DNA degradation of Un-methylated DNA by LigA7	135
Figure 5.15	Metal depended DNA cleavage activity of LigA7 on RNA monitored on 2% agarose gel	136
Figure 5.16	Fluorescence measurements of LigA	137
Figure 5.17	Multiple sequence alignment of all single Ig-like domains from the variable region of LigA	138
Figure 5.18	12% SDS-PAGE showing purification of all the generate mutants	139
Figure 5.19	Effect of Mg ²⁺ ions on LigA7 mutant proteins for the DNA cleavage activity	139
Figure 5.20	Validation of homologous model of LigA7 (A) Ramachandran plot (B) Overall Z-score	140
Figure 5.21	Representation of LigA7 and DNA complex	141
Figure 5.22	Protease activity. Degradation of β -casein in the presence of LigA7 incubated at the different time points and monitored on 12% SDS-PAGE	142
Figure 5.23	(A) MALDI-MS spectra of the degraded casein band. (B) A sequence of the degraded product	143
Figure 5.24	(A) MALDI-MS spectra of the degraded casein band. (B) A sequence of the degraded product	144
Figure 5.25	SDS-PAGE showing degradation pattern of β -Casein incubated for 12hr in the presence of different concentrations of protease inhibitor cocktail	145
Figure 5.26	SDS-PAGE showing degradation pattern of β -Casein incubated for 24hr in the presence of different concentrations of protease inhibitor cocktail	146
Figure 5.27	SDS-PAGE showing degradation pattern of β -Casein incubated for 12hr in the presence of boiled and native LigA7	147
Figure 5.28	SDS-PAGE showing degradation pattern of β -Casein incubated for 24hr in the presence of boiled and native LigA7	147
Figure 5.29	Proposed model of this novel activity in the NET escape	148

List of Tables

		Page No
Table 1.1	List of pathogenic <i>Leptospira</i> serovars and their host	8
Table 1.2	Immunological studies using the Leptospiral recombinant proteins	22
Table 2.1	Composition of culture medium	37
Table 2.2	Composition of antibiotic stocks	37
Table 2.3	Composition of solutions/reagents used for agarose gel electrophoresis	37
Table 2.4	Composition of solution used for protein purification	38
Table 2.5	Composition of solutions used for SDS-PAGE	38
Table 2.6	List of primers used for generation of LigA and LigB fragments	40
Table 2.7	Optimized concentration of PCR reactions	40
Table 2.8	Optimized concentration of PCR reactions	41
Table 2.9	Bacteria Strains/plasmids used in this study	42
Table 2.10	Summary of purification yield of the LigA and LigB fragments in this Study	51
Table 3.1	Data collection and processing statistics	69
Table 3.2	Model building and refinement statistics	72
Table 3.3	Summary of the PROCHECK results of the final refined structure of the LigA8-9 fragment	74
Table 3.4	Interface Summary analyzed by PDBePISA server	79
Table 4.1	List of tools and servers used in this study	93
Table 4.2	Ramachandran statistics and ProSA Z score	102
Table 4.3	List of selected common epitopes for vaccine construction	105
Table 4.4	Physicochemical features of chimeric vaccine construct	107
Table 4.5	List of residues involved in forming polar interaction and salt bridges	112
Table 5.1	List of primers used to generate substitution mutants in <i>ligA7</i>	123
Table 5.2	List of substitution mutants generated in <i>ligA7</i>	124
Table 6.1	Summary of the results obtained from the study	159

Appendix Table 1	List of predicted conformation BCL epitopes	194
Appendix Table 2	List of predicted linear BCL epitopes	200
Appendix Table 3	List of predicted CTL epitopes	205
Appendix Table 4	List of predicted HTL epitopes	210

ABBREVIATIONS

ACC	Auto cross-covariance
APS	Ammonium perdulfate
AHF	Andaman Hemorrhagic fever
ANN	Artificial neural network
BCL	B-cell lymphocytes
b.p	Base pair
BME	Beta mercapto ethanol
BLAST	Basic local alignment search tool
BSA	Bovine Serum Albumin
CCP4	Collaborative Computational Project Number 4
CCMB	Centre for Cellular and Molecular Biology
CD	Circular Dichroism
CTD	C-terminal domain
CTL	Cytotoxic T-lymphocyte
DALY	Disability Adjusted Life Years
DNA	Deoxyribose Nucleic Acid
DFM	Dark-field microscopy
ECM	Extracellular matrix
EDTA	Ethylene diamine tetraacetic acid
EtBr	Ethedium Bromide
EMJH	Ellinghausen – McCullough-Johnson – Harris
Fn	Fibronectin
Fg	Fibrinogen
GBD	Gelatin Binding Domain
GRAVY	Grand average hydropathicity
HCl	Hydrochloric acid
HLA	Human Leukocyte Antigen
HTL	Helper T-lymphocytes
IEDB	Immune Epitope Database
IPTG	Isopropyl thio- β -D-galactoside
IISc.	Indian Institute of Science
IgSF	Immunoglobulin superfamily
LB	Luria Bertani broth
LPS	Lipopolysaccharide
LERG	Leptospirosis Burden Epidemiology Reference Group
MAT	Microscopic Agglutination Test
MHC	Major Histocompatibility complex
MSCRAMMs	Microbial surface components recognizing adhesive matrix molecules
MCS	Multiple cloning site
MME	Monomethyl ether

MPD	2-Methyl-2, 4-pentanediol
MQ	Milli-Q
NCBI	National Center for Biotechnology Information
NET	Neutrophil extracellular Traps
NMR	Nuclear Magnetic Resonance
NTA	Nitrilotriacetic acid
OD	Optical Density
PAGE	Poly acrylamide gel electrophoresis
PAMP	Pathogen-associated molecular patterns
PBS	Phosphate Buffer Saline
PCR	Polymerase Chain Reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PME	Particle Mesh Ewald
PISA	Protein Interfaces, Surfaces and Assemblies Server
PMSF	Phenylmethylsulfonyl fluoride
Rg	Radius of gyration
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square fluctuation
RNA	Ribonucleic Acid
SAXS	Small-angle X-ray scattering
SDS	Sodium Dodecyl Sulfate
SPC	Simple point charge
SUMO	Small Ubiquitin-like Modifier
TAE	Tris base, acetic acid and EDTA
TE	Tris EDTA
TEMED	Tetramethylethylenediamine
TLR	Toll like Receptor

UNIT AND MEASUREMENT

Unit of Length

nm	nanometre
μm	micrometre
$\text{\AA}$	Angstrom

Unit of Concentration

mg/ml	milligram/millilitre
M	Molar
μM	micromolar

Unit of weight

gm	grams
kDa	Kilo Dalton
μg	microgram

Unit of Volume

l	litre
ml	millilitre
μl	microliter

Unit of Temperature

$^{\circ}\text{C}$	Degree Celsius
K	Kelvin

Preface

The outer membrane and surface-associated proteins of microorganisms perform various functions, such as protecting harsh environments, signal transduction, and protein/solute translocations (Koebnik et al., 2000). They also promote pathogenicity by enhancing pathogens' ability to attach and invade the host's immunity (Prasadarao et al., 1996). The expression of many surface proteins by pathogens such as *Staphylococcus aureus*, *Pneumococcus*, *Streptococcus*, and *Listeria monocytogenes* and their relationships in pathogenesis have been described (Andre et al., 2017; Cabanes et al., 2002; Jan-Roblero et al., 2017; Lindahl et al., 2005; Liu, 2009). Many of the surface-associated proteins contain Immunoglobulin (Ig) fold.

The Ig-like domains are widely distributed domains among the proteins and are ubiquitously present across the different phyla (Bork et al., 1994). The domain consists mainly of a β -barrel composed of seven to nine antiparallel β -strands with the typical Greek-key fold and β sandwich topology that displays a structurally conserved core (Halaby et al., 1999). The length and regularity of the β -strands, as well as the length and structure of the connecting loops, are extremely variable. The Ig-like domains are usually grouped into Ig superfamily with having a sequence homology. The presence of Ig-like domains has been reported in a large number of proteins with diverse biological functions (Hüttener et al., 2022). They are found in proteins involved in cell-cell recognition, cell-surface receptors, muscle structure protein, and the immune system (Chothia, 1953). Ig-like proteins are also reported in bacterial species known as Bacterial Ig-(Big) domain frequently found on cell surface proteins. The Big domain plays a vital role in the adhesion of bacteria to the host cell and helps in an invasion of pathogenic strains. Intimin and invasin family of outer membrane (OM) adhesin from *E. coli* and *Yersinia*, respectively, containing Ig-like domains have been studied in detail (Fairman et al., 2012; Seo et al., 2012). Apart from the structural component of adhesion, Ig-like domains in PapD protein play a chaperone function in the bacterial periplasm to help in pilus assembly. Ig-like domains

are also found in a few oxidoreductases family of proteins, sugar-binding proteins, and a few transcription factors. Some Ig domain containing proteins are reported to have hydrolysing function (Bork et al., 1994; Holmgren and Bränden, 1989).

Chapter 1 reviews the *Leptospira* and its causing agent, *Leptospira*. It also elaborates on the pathogenesis of *Leptospira*. It also discusses the outer surface proteins involved in virulence and pathogenesis. Multiple Ig-like domains containing proteins have been reported in the pathogenic *Leptospira* as well (James Matsunaga, 2003). These proteins in *Leptospira* are named Leptospiral immunoglobulin-like (Lig) proteins. These are mostly located on the surface of pathogenic *Leptospira* and help to establish host-pathogen interactions (Silva et al., 2007a). Pathogenic *Leptospira* are gram-negative spirochetes of the genus *Leptospira*, which causes a neglected tropical zoonotic disease, Leptospirosis (Adler and de la Peña Moctezuma, 2010). Each year, approximately 60,000 deaths are reported and affect almost a million population every year globally (Costa et al., 2015). The pathogenic *Leptospira* successfully establishes infection in the host mainly by using various surface proteins, including Lig proteins. Lig, including LigA and LigB, belong to the bacterial immunoglobulin-like domain (BIg) family and has been reported to be exclusively present across the pathogenic *Leptospira* (Haake and Matsunaga, 2021). LigA and LigB have 13, and 12 tandem repeated Ig-like domains, respectively, with the first six and a half domains nearly identical between the two Lig proteins. The remaining C- terminal domains, variable regions, are known to involve in interaction with the host factors. LigB, unlike LigA, has a unique carboxyl-terminal non-repeat domain (CTD) and is found in all pathogenic strains of *Leptospira*. Lig proteins have shown to be a promising vaccine candidate and provided 40-80% protection against lethality in a hamster challenge model (Haake and Matsunaga, 2021; Koizumi and Watanabe, 2004). In addition, Lig proteins have been shown to contribute to immune evasion by binding with host complement system inhibitors Factor H, FH-like 1, and C4b-binding protein (C4BP) (Castiblanco-Valencia

et al., 2012; Factor et al., 2012; Haake and Matsunaga, 2021). Many previous studies have shown the importance of Lig protein and its involvement in multiple functions. Moreover, it plays a crucial role in *Leptospira* pathogenesis and has been studied as the best vaccine candidate so far. The family of Lig proteins is a big multi-domain protein whose size ranges from 130-220 KDa, and its immunogenic region is not well characterized (Haake and Matsunaga, 2021). Recently, it was shown that antigenic motifs in a single-domain chimeric immunoglobulin-like fold generate enhanced leptospiral protection compared to whole Lig protein (Hsieh et al., 2017). Hence, it is thought that the knowledge of the immunogenic region/epitope will provide opportunities for further improvement in the Lig family vaccine efficacy through rational design. The three-dimensional structure of protein usually facilitates this direction. Unfortunately, the complete 3-D structural details of Lig proteins are largely unknown. However, recently only NMR structures of single domain LigB12 and LigA4 have been reported (Mei et al., 2015; Ptak et al., 2014). Several efforts are being put to fine map the antigenic regions among the multi-domain Lig proteins. Several reports suggest that pathogenic *Leptospira* evades host complement, as well as a neutrophil extracellular trap (NET), mediated killing (Scharrig et al., 2015). Few other pathogenic organisms evade host NET by secreting nucleases that degrade the DNA content of the NET (Andre et al., 2017; Buchanan et al., 2006; Storisteanu et al., 2017). Surprisingly, no secreted nucleases have been reported from the *Leptospira* genome so far. In the absence of secreted nucleases, we hypothesize that other surface or secreted proteins from *Leptospira* may have the ability to cleave DNA. A report suggests that proteins with Ig-like folds having His-ME finger domain may act as a nuclease (Jablonska et al., 2017). Based on this background information, we raised three questions: First, what is the full-length structure of Lig family proteins, and what is their domain orientation? Second, what are potential antigenic epitopes present on Lig proteins? Third, whether Lig proteins possess any nuclease/DNase activity.

Chapter 2 illustrates about cloning, overexpression, and purification of Lig proteins and its various smaller fragments. Both LigA and LigB are large proteins with molecular weights of 130 and 220 kDa. The full-length Lig protein showed weak expression and was inadequate for crystallographic studies. To improve the yield and protein solubility, different Lig truncated fragments were generated. The truncation is chosen to cover the complete protein in many smaller fragments. The DNA fragment containing single as well as multiple domains was cloned in the pET28a-SUMO vector at BamHI and HindIII restriction sites. The confirmed clones of LigA and LigB full-length proteins and their smaller fragments were overexpressed and purified using various chromatographic techniques. The yield of purified smaller fragments of LigA and LigB was comparatively better and was sufficient enough to carry out crystallization studies.

Chapter 3 explains the crystallization screening of all purified protein fragments using a crystallization robot. The commercially available screens from Hampton and Molecular Dimensions were used to screen the conditions in vapor diffusion sitting drop plates. The resulting crystallization conditions for many fragments were further optimized by the hanging drop vapor diffusion method. Though crystallization hits were screened for many LigA and LigB fragments, after repeated optimizations, diffraction quality crystals were obtained for LigB7-12, LigA8-13, LigA12-13, and LigA8-9. Crystals from these were exposed in X-ray in-house sources at CCMB Hyderabad. Unfortunately, the crystals of fragments LigB7-12, LigA8-13, and LigA12-13 were poorly diffracted. Only a single crystal of LigA8-9 was diffracted and diffraction data were collected up to 1.8 Å resolution. The LigA8-9 fragment contains two Ig-like domains of LigA. The diffracted crystals showed the space group P212121 with unit cell dimensions $a=34.22\text{\AA}$, $b=63.90\text{\AA}$ and $c=171.84\text{\AA}$. The calculated solvent content was 49.8%, suggesting two molecules of LigA8-9 in an asymmetric unit. The structure phase was calculated using molecular replacement using the coordinates of the LigB12 (single Ig-

like domain) NMR structure and refined refinement using REFMAC5. The final R_{free} and R_{factor} were 24 and 21%, respectively. The final model contains 185 amino acids, 33 ions, and 247 water molecules. The crystal structure of two domains of LigA showed a classical Ig-like fold. The overall structure of LigA8-9 is composed of 21 β -strands which can be divided into six sheets (three sheets in each domain) which form a β -sandwich-like structure. The structure showed some remarkably distinctive aspects even compared to the most closely related IgSF family members. Further, the structure details the relative orientation of two domains and highlights the role of the linker region in the domain orientation. The final structure also showed a clear electron density of Ca^{+2} ions, and modeled Ca^{2+} ions form a proper interacting geometry within the protein. Based on the orientation of two domain structures, the overall arrangement of Ig-like domains in LigA protein was proposed. Moreover, docking of the Gelatine binding domain (GBD) of fibronectin with the crystal structure highlighted detailed atomic interactions between the two.

Chapter 4 reports an investigation of antigenic regions present in LigA and LigB. The investigation utilizes various bioinformatics tools to identify B-cell (BCL) and T-cell (CTL) antigenic epitopes, and this information were utilized to design a multi-epitope chimeric vaccine. The potential predicted epitopes among the LigA and LigB were verified with multiple tools/servers. The epitopes which were common/overlapping to at least two servers were included in the final vaccine design. The final vaccine construct includes both linear and structural BCL epitopes. A range of linear BCL epitopes (71-97) was generated among the LigA and LigB by the three servers. Individual domains from the two proteins generated various numbers and lengths of BCL epitopes. LigA8 and LigA11 shared maximum linear BCL epitopes. In the case of structural BCL, the three-dimensional structure of each domain of LigA and LigB was modeled using the I-TASSER server. Best models were selected based on the Ramachandran analysis and Z-scores value and subjected to conformational BCL

prediction. Three different servers- DiscoTope, ElliPro, and BPro, were used. The three servers yielded a different number of discontinuous epitopes, and common epitopes were selected. As in the case of LigA7, DiscoTope and BPro predicted 13 residues that are involved in the formation of a conformational B-cell epitope, while the ElliPro for the same yielded three epitopes. Hence, common residues involved in conformational B-cell epitopes generated by three were selected. Similar strategy was applied for all other domains. For the prediction of CTL, we used IEDB, ProPred-1, and NetMHC servers. The IEDB server-generated 400 potent MHC-1 binding epitopes covering 35 HLA alleles for LigA and 36 HLA alleles for LigB. The ProPred-I generated 1032 epitopes covering all 47 HLA alleles and showed maximum binding with the HLA-B*58:01 allele. The NetMHC 4.0 generated 327 epitopes covering 72 alleles for LigA and 76 alleles for LigB.

The prediction of HTL epitopes was completed by using the online tools ProPred, IEDB, and NetMHC-II 2.3. In the case of HTL epitopes, the IEDB server predicted a total of 185 potent MHC-II binding epitopes, the ProPred server predicted 177 epitopes, and NetMHCII 2.3 server predicted 665 epitopes. From all the predicted epitopes, the promiscuous peptides that showed binding to three or more HLA alleles were selected. The common/overlapping epitopes with higher antigenic scores were selected and fused using linkers to construct a multi-epitope vaccine. The most antigenic region identified from the prediction study is mapped on Lig proteins,

To enhance the immunogenicity of the vaccine construct, an adjuvant was also fused at the N-terminal of the construct using an appropriate linker. The 3-D structure model contains 99.7% residues in allowed regions of the Ramachandran plot. The model structure of the vaccine construct possesses 16.96% α -helix, 41.38% extended strand, and 41.66% random coil. Many previous studies signify the importance of host TLR4 in immunological mechanisms against leptospiral antigens. The multi-epitope vaccine construct also interacts with the TLR4 through

several hydrogen bonds and salt bridges, as demonstrated by a molecular docking study. The docking studies revealed a strong interaction between the vaccine construct and TLR4 and ensured the induction of immune response. The stability of the interaction was also validated by performing molecular dynamics simulations for 100ns. Molecular dynamic simulation generated RMSDs emphasized the stable interaction between vaccine construct and TLR4 receptor. Moreover, the interaction and compactness of the complex between the two are also analyzed with RMSF and RG plots. Overall, this suggests that the vaccine construct may follow TLR4 mediated immune response. The multi-epitope chimeric vaccine generated from this study potential could serve as a promising vaccine against leptospirosis. In-vivo validation and efficacy of the vaccine construct are being investigated by our collaborator.

Chapter 5 demonstrates a novel in-vitro nuclease and protease activity in the Ig-like domain of the Lig proteins. All the Lig constructs containing single, and multiple domains showed metal-dependent nuclease activity. Among various metals such as Mg^{2+} , Mn^{2+} , Ca^{2+} , Fe^{2+} , Co^{2+} , Ni^{2+} , Cu^{2+} , and Zn^{2+} , maximum nuclease activity was displayed by LigA7 in the presence of Mn^{2+} , Co^{2+} , and Mg^{2+} . The interaction of the magnesium ion with the protein was monitored using fluorescence spectroscopy. Site-directed mutagenesis revealed Mg^{2+} binding residues in the Ig-like domain of LigA7. The mutagenesis result suggests Asparagine 49 as an Mg^{2+} coordinating residue. The basis of novel nuclease activity may be associated with protein adopting different conformation in the presence of divalent ions and substrate as investigated by change of intrinsic fluorescence. The interaction of the protein with DNA for the cleavage activity was analyzed by molecular docking of LigA7 and 19bp DNA stretch. The docking results revealed the binding of DNA to positively charged residues such as Lysine. Interestingly, no activity was observed with the RNA, indicating that the protein bears DNase activity and not RNase. In addition, the Lig proteins also showed metal-independent protease activity. The cleavage of β -casein was further confirmed by MALDI-TOF MSMS analysis of the cleaved product. This

part of the study establishes that apart from adhesin-like properties, the Lig proteins also possess a metal-dependent DNase activity and metal-independent protease activity. To our surprise, though there were no His-ME motifs found in the Lig sequence, still these novel functions highlight moonlighting function in the Lig proteins. This finding proposes that the novel functions may be associated with the evasion of neutrophil extracellular Traps (NET). Overall this study enhances the basic knowledge of non-nuclease proteins involved in the NET evasion phenomenon in *Leptospira* and makes the foundation to explore other pathogens. Moreover, this information may be utilized to develop preventive strategies to interfere with *Leptospira* immune evasion.

Chapter 1

Introduction and Review of Literature

1.1. Introduction

Leptospirosis is a zoonotic disease caused by a spirochete, pathogenic *Leptospira*. The disease is categorized as an emerging and neglected tropical zoonotic disease. It is considered a public health problem globally, with an estimated ~1 million cases reported each year, causing deaths of around 60,000 (Costa et al., 2015; Putz and Nally, 2020; Rodrigues de Oliveira et al., 2021a). The disease has a broad geographical distribution occurring in rural and urban areas. The infection is more persistent in tropical regions, especially in developing countries where adequate sanitation is not provided (Faine et al., 1999). Leptospirosis usually has symptoms such as headache, chills, illness, and muscle aches while the more severe form of the disease is associated with multi-organs failure known as Weil's disease (David A. Haake, 2015). Cattles and other domestic animals are primarily affected by *Leptospira*, which causes a massive loss in livestock industries each year due to abortion, uveitis, pancreatitis, and other respiratory diseases (Adler and de la Peña Moctezuma, 2010a). Humans are the accidental hosts who get infected through the contaminated water, soil, and food (David A. Haake, 2015). The individuals associated with farming, sewage work, veterinarians, and other recreational activities such as swimming and kayaking are at higher risk (Faine et al., 1999). The cases of leptospirosis drastically increase after heavy rain or flood conditions (Faine et al., 1999). *Leptospira* enters the host body through cuts or abrasions in the skin or mucous membranes, from where it enters the bloodstream and eventually invades many organs (David A. Haake, 2015). *Leptospira* establishes infection majorly through the surface-expressed antigens. Leptospirosis is difficult to diagnose because of the nonspecific symptoms that overlap with common infectious diseases such as malaria, dengue, and flu. Available screening tests include microscopic agglutination (MAT), an antibody-based serological test that has proven to be the gold standard in diagnosing leptospirosis (Faine et al., 1999). The molecular diagnosis techniques that use Polymerase chain reaction (PCR) against the unique target genes (eg,

LipL32, *ligA*, *ligB*) to identify *Leptospira* DNA in the initial phase of infection, have proven to be advantageous over serological tests (Deneke et al., 2014; Vedhagiri et al., 2013).

1.2. History

Leptospirosis was first brought into the picture in 1886 by Adolph Weil (Weil, 1886). He described it as a type of jaundice, accompanied by renal dysfunction, skin rashes, and conjunctivitis. Later, it was given a name, Weil's disease, to honour him. Although the disease was identified in 1886, based on the documented records, it is thought to have existed in nature much earlier. Various names have described the disease in ancient literature, such as rice field jaundice, autumn fever, seven-day fever, cane cutter's disease, swineherd's disease, yellow fever, and mud fever (Faine et al., 1999). In 1907, Arthur Stimson discovered the presence of spirochete for the first time in the kidney tubules of a patient by silver staining. He named it *Spirochaeta interrogans* because of their shape (Stimson, 1907). However, this discovery was not recognized until 1915, when saprophytic forms were discovered in freshwater (Inada, R.I., 1915). The transmission of *Leptospira* to humans through rats was first discovered in 1916. The first case of human leptospirosis was observed in the year 1951 (Coggins, 1962). The family of *Leptospira* was proposed sometime in 1979 (Hovind-Hougen, 1979). After three years, in 1982, based on the antigenic variations, 200 different serovars of *Leptospira* were identified (Faine and Stallman, 1982). The complete genome sequence of *Leptospira interrogans* Serovar Lai was done in 2003 (Ren et al., 2003). Subsequently, in the year 2005, seven new species of *Leptospira* were reported (Sulzer and Rogers, 2005). The diagrammatic representation of the historical milestone is given in **Figure 1.1**.

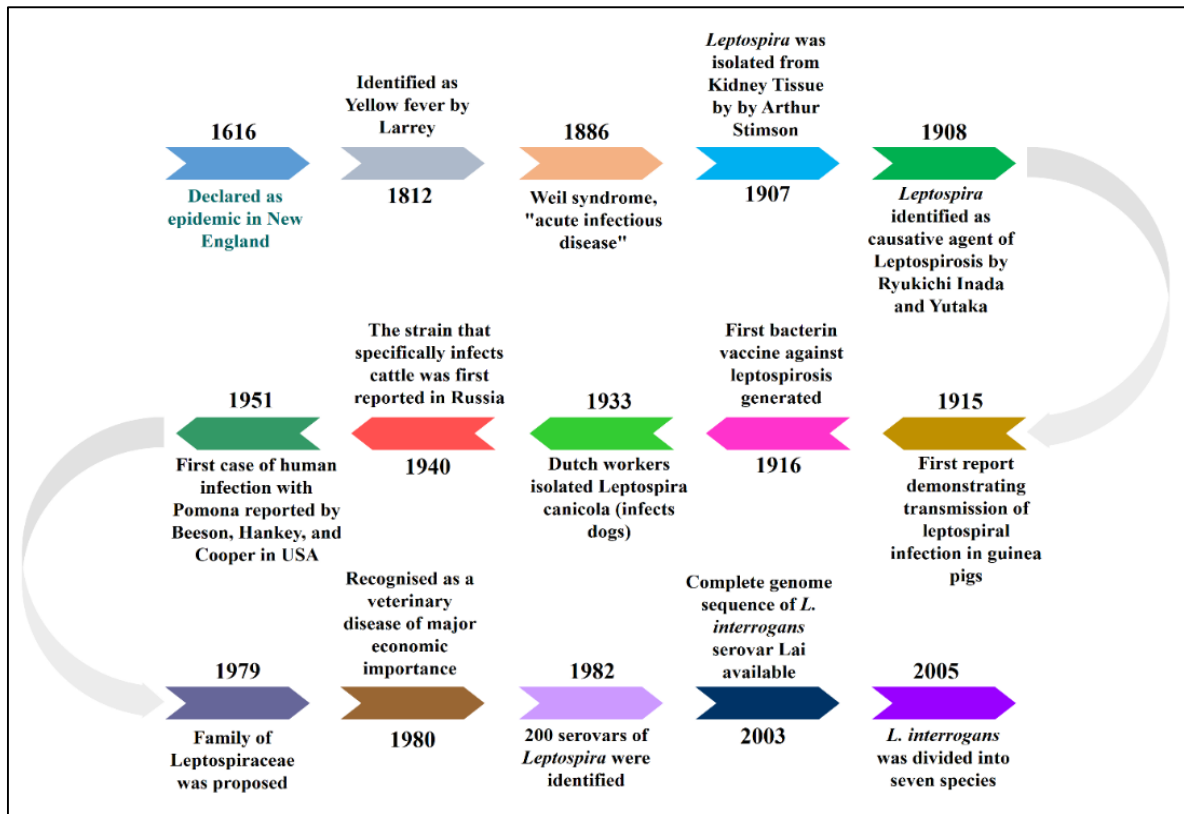


Figure 1.1: Brief history related to Leptospirosis disease and *Leptospira*

1.3. *Leptospira* Biology

1.3.1. General characteristics

The members of the leptospiral family are thin, long, and highly motile bacteria comprising both pathogenic and saprophytic species. They are microscopic and spiral-shaped bacteria with an average diameter of 0.1 μm and 6–20 μm in length (Ko et al., 2009). Motility is provided by the endoflagellum. *Leptospira* have a closely associated double-membrane structure like other spirochetes (Picardeau, 2017). The membrane thickness usually varies from 27 to 30nm, consisting of the outer membrane, peptidoglycan layer, and inner membrane (Levett, 2001). The outer membrane is attached with Lipopolysaccharides (LPS) (Haake et al., 2010). The bacterium is not visible under the light microscope due to its thin structure but can be seen

under phase-contrast or dark field microscopy (Faine et al., 1999). The images showing the morphology of *Leptospira* are given in **Figure 1.2**.

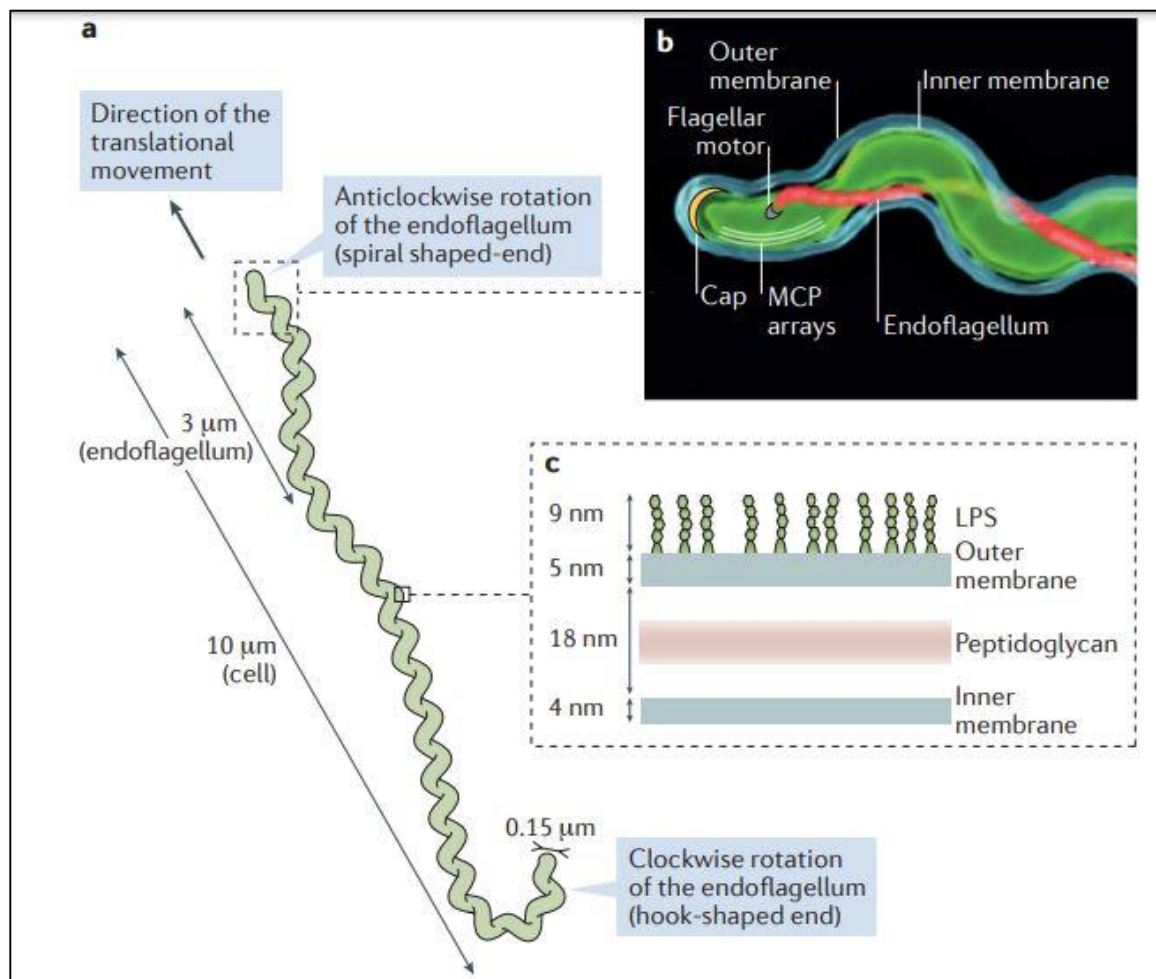


Figure 1.2: Architecture of *Leptospira interrogans* (A) general characteristics of the bacterium (B) Cryo-electron tomography shows the cap-like structure of *Leptospira interrogans* (C) Gram-negative like structure possessed by the bacterium (**Adapted from Picardeau, 2017, *Nat Rev Microbiol***)

1.3.2. *Leptospira* Classification

1.3.2.1. Taxonomical classification

Domain: - **Bacteria**

Kingdom: - **Eubacteria**

Phylum: - **Spirochaetes**

Division: - **Gracillicutes**

Class: - **Scotobacteria**

Order: - **Spirochaetales**

Family: - **Leptospiraceae**

Genus: - ***Leptospira***

1.3.2.2. Serological classification

Leptospira is grouped into pathogenic and non-pathogenic based on their antigenic relationships (Picardeau, 2017). All pathogenic serovars belong to *Leptospira interrogans* species, and they fail to grow in the presence of 8-azaguanine at 13° C (Terpstra, 1992). In contrast, the non-pathogens belonging to *Leptospira biflexa* usually grow. Another prominent feature observed, unlike non-pathogens, pathogenic species form spherical shapes in the presence of salt (Levett, 2001). In addition, the *Leptospira* species are subdivided into serovars based on the structural heterogeneity of their lipopolysaccharide (LPS) components. To date, 300 pathogenic and 60 non-pathogenic serovars have been described (Picardeau, 2017).

1.3.2.3. Genotypic classification

Based on 16S rRNA phylogenetic analysis, the genus *Leptospira* is also classified into three clades; pathogenic, non-pathogenic, and intermediates (Lehmann et al., 2014). The 22 species of *Leptospira* are divided based on their genetic diversity, as shown in **Figure 1.3**. In this classification, ten *Leptospira* species are classified as pathogenic, seven species as intermediate, and five species are classified as non-pathogenic (Picardeau, 2017).

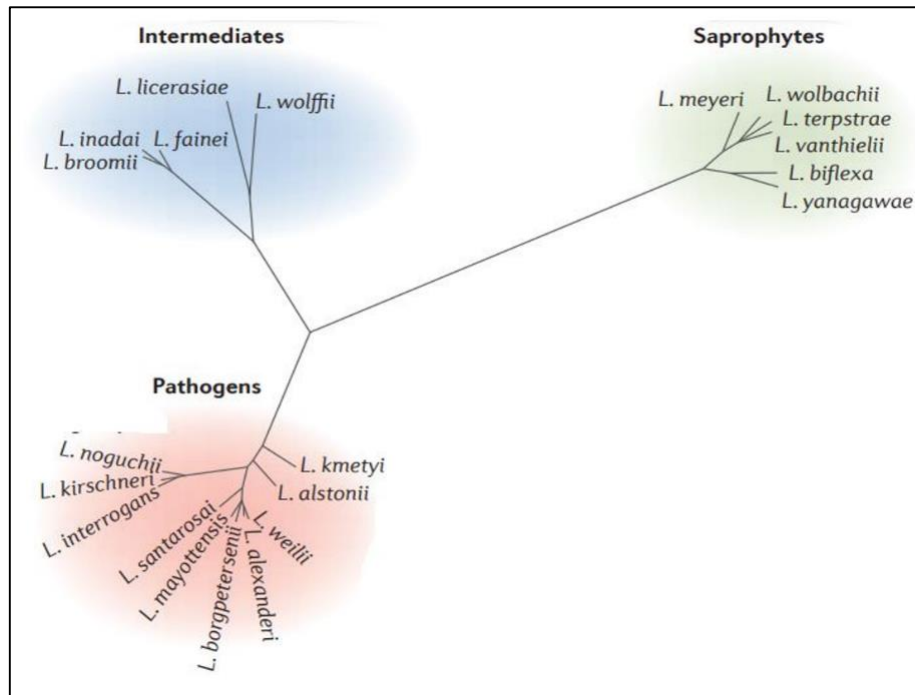


Figure 1.3: Phylogenetic analysis of *Leptospira* strains (Adapted from Picardeau, 2017, *Nat Rev Microbiol*)

1.3.3. Growth condition

Leptospira can grow with a doubling time of 6-8h on both solid and liquid media supplemented with ammonium salts, vitamins B1 and B12, and long-chain fatty acids over a temperature range of 28-30 °C (Ellis and Michno, 1976). The medium required for culturing *Leptospira* under laboratory conditions is Ellinghausen – McCullough-Johnson – Harris (EMJH) supplemented with Tween20 and 1% bovine serum albumin (BSA) (Ellinghausen and McCullough, 1965; Ellinghausen H C & McCullough W G, 1965).

1.4. Epidemiology

Leptospirosis is ubiquitous in animals and found in almost all regions across the globe. The causative agent of leptospirosis, pathogenic *Leptospira* infects various animal hosts such as rodents, cattle, sheep, and dogs (Stein, 2007). Leptospirosis is more chronic in animals and causes infertility, miscarriage, stillbirth, and pulmonary hemorrhage syndrome in dogs, sheep,

and goats (Libonati et al., 2018; Lilenbaum and Martins, 2014). The disease has shown high livestock mortality, directly contributing to economic loss (Gendron et al., 2014; Petrakovsky et al., 2014). Rodents are majorly responsible for the transmission of the disease. A list of pathogenic serovars of *Leptospira* and the organisms affected is given in **Table 1.1**.

Table 1.1: List of pathogenic *Leptospira* serovars and their host

Affected organism	Causative organism	Reference
Humans	<i>L. interrogans</i>	(Inada et al., 1916)
Cattles	<i>L. interrogans</i> serovar Canicola / <i>Copenhageni</i> , <i>L. kirshneri</i> serovar Grippotyphosa	(Michna and Campbell, 1969)
Deer	<i>Leptospira interrogans</i> serovar Hardjo	(FERRIS et al., 1960)
Sheep	<i>L. noguchi</i> serovar Autumnalis	(Ellis et al., 1983)
Dogs	<i>L. noguchi</i> , <i>L. interrogans</i> serovar Canicola, <i>L. interrogans</i> serovar Copenhageni	(Claus et al., 2008)
Swine	<i>L. interrogans</i> serovar Canicola	(Ellis, 2015)
Bats	<i>L. interrogans</i> , <i>L. kirshneri</i> , <i>L. borgpetersenii</i> and <i>L. fainei</i>	(Cox et al., 2005)
Squirrels	<i>L. interrogans</i> serovars Icterohaemorrhagiae and Canicola	(Ellis et al., 1983)
Horses	<i>L. interrogans</i> serovar Pomona type kennewicki and serovar Grippotyphosa	(Verma et al., 2013)
Goats	<i>L. kirshneri</i> serovar Grippotyphosa	(Taylor et al., 2011)

1.4.1. Human leptospirosis

Humans are accidental hosts that get infected when they come into contact with contaminated water, soil, and food (Haake and Levett, 2015). Humans living in rural areas with inadequate sanitization, poor housing, and close contact with domestic animals are at high risk of leptospirosis. Even individuals dwelling in the urban slums characterized by substandard hygiene conditions are also at risk of the disease (Dias et al., 2007). Leptospirosis infection in

humans has many nonspecific symptoms which are difficult to identify clinically. The Leptospirosis Burden Epidemiology Reference Group (LERG) provides an estimated burden of leptospirosis worldwide and promotes awareness of leptospirosis interventions and controls (Stein, 2007). According to LERG, the rate of infectivity of leptospirosis increases during rainfall, flood conditions, by animal contact, and in poor sanitation conditions (Costa et al., 2015).

1.4.2. Cycle of infection

Pathogenic leptospires are usually maintained in the kidneys of infected individuals. Human-to-human transmission is infrequent but can occur by sexual intercourse (Harrison and Fitzgerald, 1988) or breastfeeding (Carlisle et al., 1988). The schematic representation of the *Leptospira* transmission cycle in humans is given in **Figure 1.4**. The *Leptospira* can survive in the environment without any host for months. *Leptospira* enters humans body through abraded skin, cuts, and mucous membranes from the contaminated environment (Dias et al., 2007).

1.4.3. Risk factors

In addition to poor sanitation and sub-standard dwelling conditions, individuals associated with occupations such as farming, veterinary physician, animal husbandry, hunters, dairy farming, and scientists handling animals are at higher risk of the disease (Lau et al., 2016; Mwachui et al., 2015). Indirect contact with infected water is more common and associated with water sports such as kayaking, swimming, rafting, and other recreational activities (Levett, 2001). Areas with heavy rainfall and flooding are at risk of having epidemic of the disease.

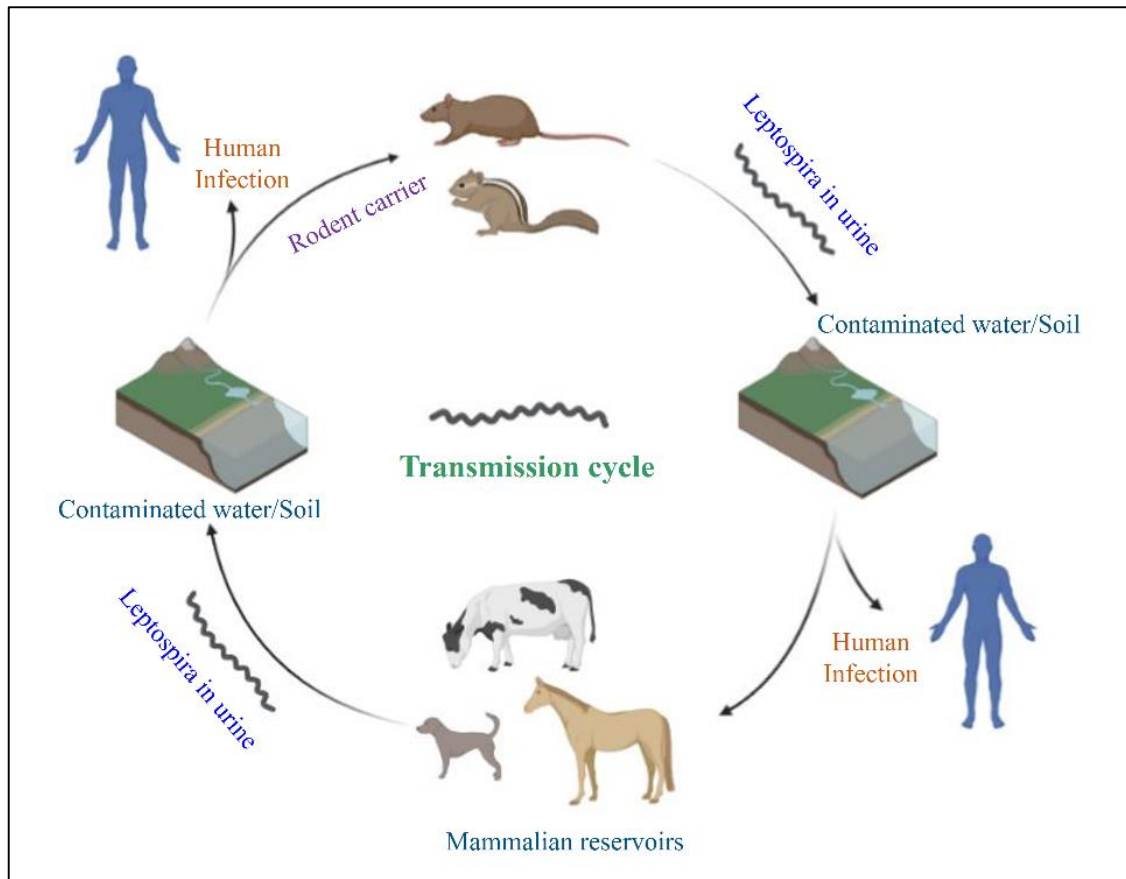


Figure 1.4: Transmission cycle of leptospirosis

1.4.4. Clinical manifestation

Leptospirosis in humans is associated with a wide range of symptoms, from headaches, fever, chills, and muscle ache to the more severe form known as Weil's syndrome, which affects multiple organs such as the liver, kidneys, lungs, and brain. Most symptoms are nonspecific in the early stages of the disease, leading to misdiagnosis and remain unnoticed by the patient. In a Weil's syndrome condition, *Leptospira* harm multiple organs leading to kidney lesions, jaundice, vascular injury, oliguria, meningitis, and uremia (Ajay R Bharti, 2003; Hartskeerl et al., 2011; Samrot et al., 2021). *Leptospira* colonizes the kidney's proximal tubules leading to renal dysfunction, oliguria, hypokalemia, and acute renal failure, which have been observed in 16-40% of cases (De Brito et al., 2018; Medicina et al., 1999).

1.5. Global burden

Leptospirosis occurs worldwide, especially in areas with tropical and sub-tropical climates. Disability Adjusted Life Years (DALY) index is a parameter used by the world health organization (WHO) to estimate the global disease burden. One DALY is equal to the loss of one year of healthy life of an individual. The leptospiral burden in terms of DALY /100,000 per year is given in **Figure 1.5**. Leptospirosis is a global burden, with the DALY index of approximately 2.90 million per year, majorly occurring in low-income tropical countries (Torgerson et al., 2015). While leptospirosis occurs worldwide, cases are more likely in remote areas where the disease is endemic or in tropical regions with humid environments, such as Latin America, South Asia, China, and Africa (Zenebe and Abdi, 2013). Areas that generally receive heavy rainfall are also affected. Fewer cases of human leptospirosis are also reported in Japan, Taiwan, South Korea, and Hong Kong (Victoriano et al., 2009).

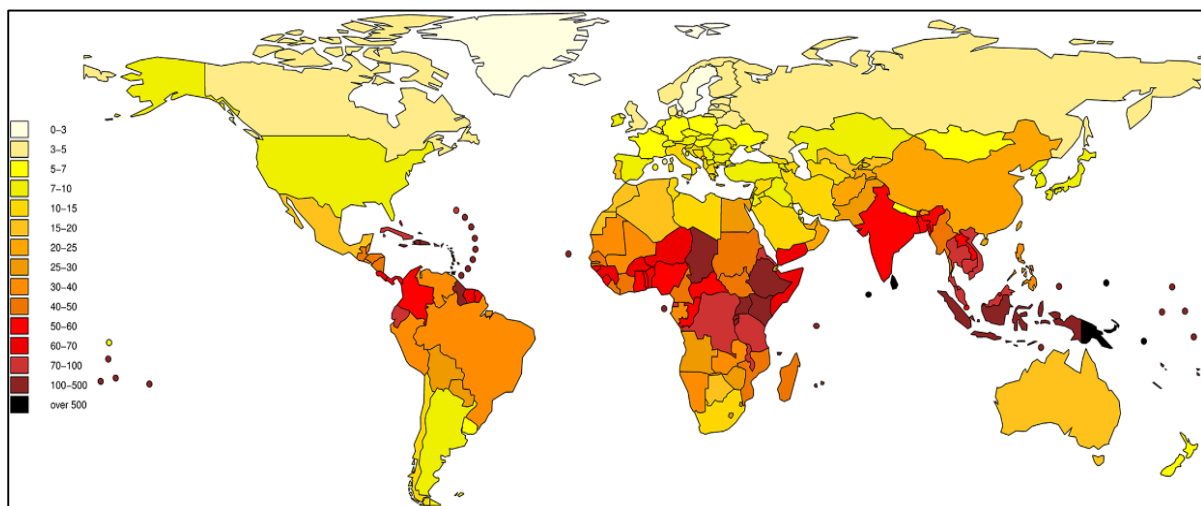


Figure 1.5: Global burden of leptospirosis in the context of DALY index (Adapted from Torgerson PR, *PLoS Negl Trop Dis*, 2015)

Sri Lanka is one of the most endemic countries for leptospirosis, with more than 7,000 reported cases of leptospirosis in 2008 (Agampodi et al., 2011). In Malaysia, cases of leptospirosis have

been reported since 1920, and since then, more than 41,736 cases and 502 deaths have been reported in Malaysia alone (Philip et al., 2020).

1.5.1. Leptospirosis in India

Since the 20th century, India has been considered as a hub for leptospirosis because of its socio-economic, geographical, and environmental factors (Sethi et al., 2010). The leptospirosis cases were reported mainly from coastal areas and rainy states like Maharashtra, Tamil Nadu, Kerala, Gujarat, and Andaman-Nicobar islands (Izurieta et al., 2008; Sambasiva and Naveen, 2003; Sethi et al., 2010). In 1960, *Leptospira icterohaemorrhagiae* and *Leptospira canicola* antigens were isolated from five patients with jaundice. In 1967, the sera of several patients with hepatitis showed evidence of infection with *L. pyrogenes* in Bombay. Sometimes, in 1984-85, 19 patients with acute renal failure due to leptospirosis were reported in Madras. Currently, the highest incidence of leptospirosis is being reported from Andaman island (50- 65/ 100,000 per year) (Shivakumar, 2008). Several outbreaks of leptospirosis occurred in the Andaman-Nicobar islands in 1932, where two leptospiral serovars *Leptospira andamans* and *Leptospira grippotyphosa* were isolated. Andaman Hemorrhagic fever (AHF) outbreak was first reported in 1988 and was further confirmed as leptospirosis in 1994. In total, 524 cases were reported from 1988 to 1997, and 524 cases were reported between 2000-2004 (Shivakumar, 2008). In 2004, 322 confirmed cases of leptospirosis were reported, most of which were associated with occupations such as sewage workers, butchers, forest workers, zookeepers, and farmers (Shivakumar, 2008). In 2005, 14 out of 58 reported cases of leptospirosis died due to pulmonary hemorrhage within 48hrs.

In the areas of south Gujrat, such as Valsad, Navsari, and Surat, leptospirosis has been endemic since 1994 (Clerke et al., 2002). In 2005, 392 cases of leptospirosis were reported in Gujrat, including 81 deaths due to renal failure and hemorrhagic Pneumonitis. The cases of

leptospirosis in Maharashtra have been reported regularly since 1998. A major outbreak was reported, following the monsoon in the year 2005, with 2355 cases and 167 deaths. Kerala is also one of the most affected states by pathogenic *Leptospira*. Such cases have been reported since 1987. A case study of 976 confirmed cases of leptospirosis with a mortality rate of 5.32% (Kuriakose et al., 1997). In another study with 282 leptospirosis patients showed a mortality rate of 6.03% (Pappachan et al., 2004). Cases of leptospirosis have been reported from Chennai, Tamil Nadu since 1980s. The number of cases increased significantly in 2006 with 2765 cases reported in a year.

1.5.2. Diagnostic challenges

Patients do not receive a direct leptospirosis diagnosis because the clinical manifestations of the early stages of infection are confused with the symptoms of other illnesses. Diagnosis is most often made when the patient shows clinical signs such as renal failure, pulmonary hemorrhage, and jaundice. The gold standard diagnostic tool currently being used is the serological test, the Microscopic Agglutination Test (MAT). Conventional clinical staining is not applicable to *Leptospira* because they are too fine to be stained by Gram staining. However, they are visible under dark-field microscopy (DFM) due to screw movement, but this method cannot be used for clinical diagnosis because it can give false-positive results for artefacts such as fibres. Other available diagnosis tests include ELISA, PCR, and commercial kits such as Leptocheck dipstick assay, LEPTO IgM MICROLISA ELISA kit, and Leptospirosis IHA test.

1.6. Pathogenesis

The complete molecular mechanism of pathogenesis has not yet been fully elucidated. However, in recent times, the availability of whole genome sequences of pathogenic leptospires such as *Leptospira interrogans* serovar Copenhageni and Lai has made significant advances in demonstrating pathogenesis. The use of various animal models in laboratories, such as guinea

pigs and hamsters, has helped to identify key factors involved in establishing host-pathogen interactions (Budihal and Perwez, 2014). In addition, pathogen-associated molecular patterns (PAMP) also play a crucial role in establishing pathogenesis (Zuerner, 2015).

1.6.1. Attachment and Invasion

The pathogenic *Leptospira* exhibits direct attachment with the host Extra Cellular Matrix (ECM) proteins such as fibronectin, laminin, collagen, elastin, and integrin to establish host-pathogen interaction. These attachments are usually mediated by various surface-exposed proteins from *Leptospira* (**Figure 1.6**). These proteins have been classified as microbial surface components recognizing adhesive matrix molecules (MSCRAMMs). Leptospiral surface-expressed proteins also help *Leptospira*, to attach host renal epithelial cells, monocytes, fibroblasts, and endothelial cells (Liu et al., 2007; Thomas and Higbie, 1990). It has been observed that pathogenic leptospires bind to host cells more efficiently than non-pathogenic ones (Liu et al., 2007). Various studies have identified MSCRAMMs in *Leptospira* and established their interaction with host proteins. Such an example is briefed in a schematic diagram showing MSCRAMMs and their host interacting partners is shown in **Figure 1.6B**.

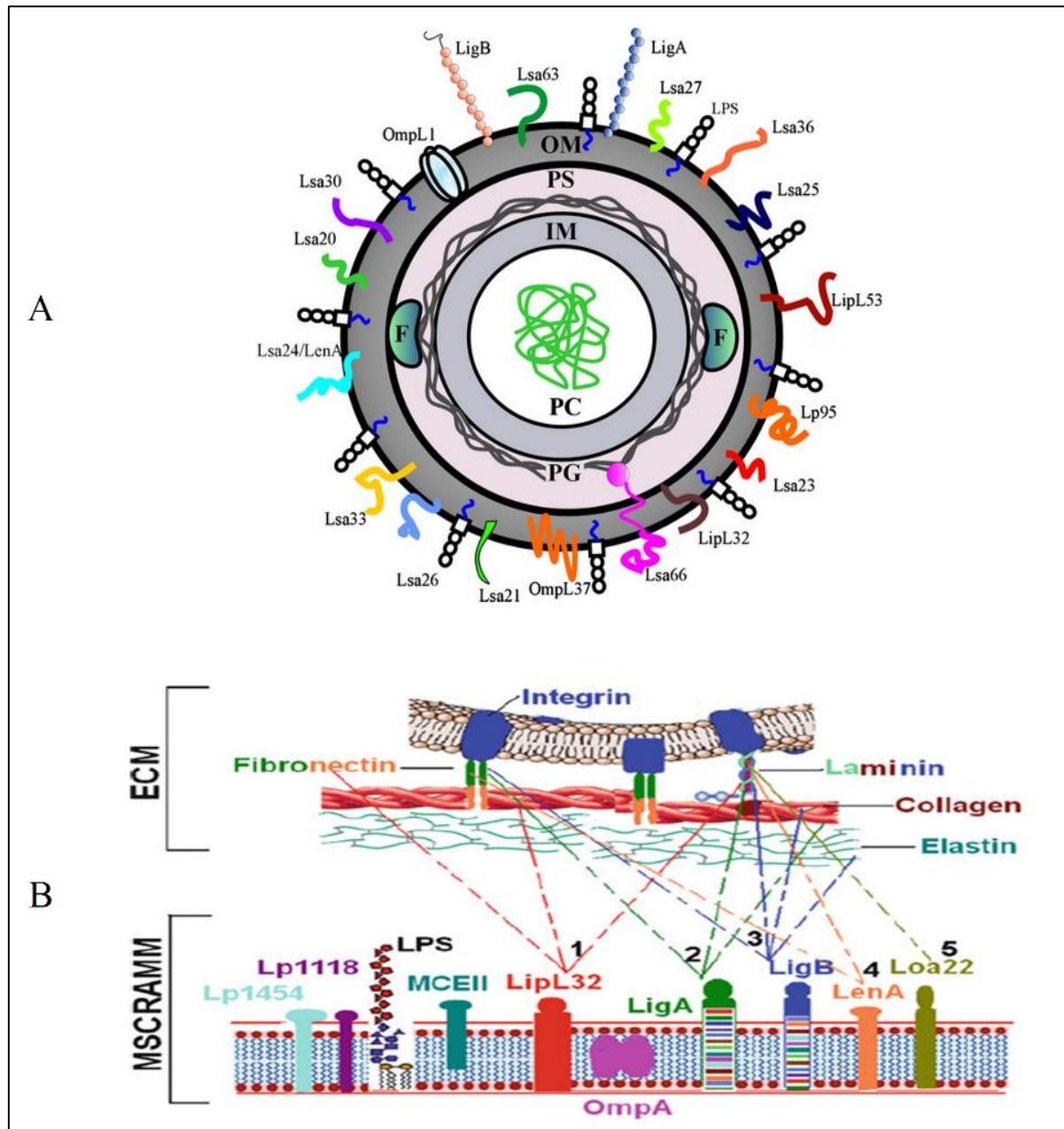


Figure 1.6: (A) Schematic representation of the *Leptospira* membrane architecture showing surface-exposed proteins (B) Interaction of *Leptospira* MSCRAMM with the host ECMs (Adapted from Vieira, M.L. *FEMS Microbiol Lett*, 2014 & Faisal, S.M. *Springer*, 2012).

After attachment, *Leptospira* migrates through the proximal tubule to target organs such as the liver, lungs, and kidneys (Ajay R Bharti, 2003; David A. Haake, 2015). The detailed roles of *Leptospira* PAMPs and surface-exposed proteins involved in pathogenesis are described below.

1.6.2. Lipopolysaccharides (LPS)

Lipopolysaccharides (LPS) are key cell wall components characteristic of Gram-negative bacteria, a representative of PAMP, and play an important role in pathogenesis. The polysaccharide components of the LPS primarily participate in protective functions for bacteria by camouflaging with common host carbohydrate residues. The Lipid component of LPS incorporates methylated phosphate, which helps the microorganism to be undetected through the human TLR4 (Werts et al., 2001). Various studies have already established that the LPS expression is essential for the survival of *Leptospira*, both inside and outside of the mammalian host (David A Haake, 2015). In addition, the LPS provides an antigenic diversity in different species of *Leptospira*. The LPS of *Leptospira* is chemically composed of arabinose, galactose, xylose, and rhamnose (Faine et al., 1974; Microbiology and Britain, 1986). It has been described in many studies as a diagnostic target and a strong vaccine candidate (Chapman et al., 1991; Matsuo, 2000; Wang et al., 2007; Widiyanti et al., 2013)

1.6.3. Leptospiral lipoprotein (LipL) Proteins

Leptospiral Lipoproteins such as LipL32, LipL53, LipL36, LipL21, and LipL46 are reported to be highly expressed during infection (Fernandes and Hartwig, 2007; Haake et al., 1998; Koizumi and Watanabe, 2003; Shang and Summers, 1996). Many of these LipL are known to interact with the host extracellular components and are involved in pathogenesis. The LipL46 shows binding with the human plasminogen, and LipL53 is reported to be involved in binding with laminin, collagen IV, and fibronectin, which highlights the importance of these proteins during *Leptospira* infection (Oliveira et al., 2010). In addition, few are reported to be antigenic in nature. LipL32 is the most abundant surface lipoprotein and a well-known immune-dominant antigen in human leptospirosis (Haake et al., 2000). The LipLs are mostly conserved in the pathogenic species of *Leptospira*. Few have been extensively evaluated as vaccine candidates

in combination with different strategies such as subunit vaccines, DNA vaccines, and with BCG (Rino Rappuoli, 2000; Serruto et al., 2012; Yang et al., 2002). The recombinant vaccine developed, consists of surface-exposed LipL41, and OmpL1 proteins which has been reported to demonstrate approximately 71% survival against *Leptospira* infection in the hamster model (Haake et al., 1999). Moreover, in a recent report, LipL21 and its truncated N-terminal domain were shown to induce pro-inflammatory cytokines by activating the host's innate immune response (Kumari et al., 2017). A microarray and *in-silico* analysis have been used to categorize LipL45, LipL21, LipL41, and LipL32 to be a potential vaccine candidates against *Leptospira interrogans* (Yang et al., 2006).

1.6.4. *Leptospira* endostatin like protein (Len) and Loa22

LenA is a 24-kDa protein that binds to multiple host factors such as fibrinogen, complement factor H, laminin, and fibronectin. The Len family includes six different proteins LenA to LenF, and almost all are reported to interact with the host factors (Angela S Barbosa et al., 2006; Stevenson et al., 2007; Verma et al., 2006). Loa is the second most abundant outer membrane protein of *Leptospira* (Beck et al., 2009) and has a peptidoglycan-binding motif which makes them similar to outer membrane protein A (OmpA). Loa22 is one of the outer-membrane proteins which gets upregulated during *Leptospira* infection (Koizumi and Watanabe, 2003; Nally et al., 2007). The upregulation of Loa22 may be associated with pathogenesis.

1.6.5. Hemolysins and HemO

Hemolysins serve as potential virulent factors that have the ability to lyse the cell membranes. These proteins facilitate the invasion of *Leptospira* either by damaging the membrane or by forming the pores. The pathogenic *L. interrogans* Lai genome encodes for nine hemolysin genes which include sphingomyelinases, non-sphingomyelinases, and pore-forming

hemolysins (Ren et al., 2003). Although the exact role of sphingomyelinases in pathogenesis has not been identified, their absence from the saprophytic *Leptospira* highlights their role in pathogenicity and survival (Adler and de la Peña Moctezuma, 2010a; Bulach et al., 2006). The outer membrane of *Leptospira* contains many iron up-taking proteins like HemO, and the distribution of these genes has a direct effect on the virulence and pathogenesis of *Leptospira* (Murray et al., 2009).

1.6.6. Leptospiral Immunoglobulin-like protein (Lig)

The family of Leptospiral Immunoglobulin-like (Lig) proteins, present exclusively in pathogenic species, consists of LigA and LigB, with 13 and 12 homologous extracellular Ig-like repeated domains, respectively, known to play an important role in bacterial pathogenesis (James Matsunaga, 2003; Palaniappan et al., 2007; Christopher P. Ptak et al., 2014; Silva et al., 2007a). The expression of Lig proteins are found to be upregulated during mammalian infection (James Matsunaga, 2003). The family of Lig proteins binds to host extracellular matrix components (ECM) and helps pathogens to invade and in host tissue colonization (Choy et al., 2007; Palaniappan et al., 2007). In addition, Lig proteins are known to bind to the complement factors in order to evade innate immunity and establish the infection (Factor et al., 2012). Many studies have been performed on Lig proteins in terms of vaccines and pathogenesis.

1.7. Host Immune response against leptospirosis

The alternate pathways of the complement system play a crucial role in the elimination of *Leptospira* during the first hours of infection. Non-pathogenic *Leptospira biflexa* is usually cleared within a minute from the human serum, but pathogenic *Leptospira* survives within serum by using different escape mechanisms. Studies on cattle and hamsters suggest their ability to induce Th1 response and release of IFN- γ against killed *Leptospira borgpetersenii*

serovar Hardo and *Leptospira interrogans* serovar Pomona (Bolin and Alt, 2001; Brown et al., 2003; Faisal et al., 2008). The involvement of cell-mediated immunity against *Leptospira* infection is also well demonstrated, especially for CD4 and Gamma-delta-T cells ($\gamma\delta$ -T cells). These $\gamma\delta$ -T cells serve as the link between the innate and adaptive immune responses.

The innate immune system is first activated in response to the bacterial LPS and lipoproteins through TLR2 and TLR4 interactions leading to the induction of inflammatory cytokines such as IL-10, IL-1 β , MCP-1, IL-6, and TNF- α (Wang et al., 2012). In addition, the host macrophage also gets activated in response to *Leptospira* LPS and induces IFN, IL-1 β , TNF α , and IL-6 cytokines (Isogai et al., 1990). Polymorph nuclear neutrophils (PMNs) play a vital role in clearing the pathogen by producing hydrogen peroxide (H₂O₂) (Murgia et al., 2002). Furthermore, human defensin such as HNP1, HNP2, and HNP3 produced by neutrophils and epithelial cells also effectively kill *Leptospira*.

1.8. Immune evasion strategies

Pathogenic bacteria adopt various strategies to avoid host immune surveillance to survive. Complement-system is the first arm of the innate immune system and plays a vital role in the clearance of *Leptospira*. Non-pathogenic *Leptospira* are more susceptible to complement-mediated destruction than pathogenic. Many unique proteins present in the pathogenic *Leptospira*, are known to play a role in evading complement-mediated killing (Barbosa et al., 2009; Meri et al., 2005). It is reported that the pathogenic *Leptospira* usually avoids complement-mediated attacks by three main ways. 1. by acquiring complement regulator of the host, 2. By acquisition of host proteases, and 3. By secreting its proteases to cleave the complement factor. All three pathways of the complement system are being affected by pathogenic *Leptospira*. Pathogen utilizes its outer surface and secretory proteins such as LigA, LigB, and *Leptospira* complement binding protein A (LcpA) to bind the complement regulator

C4BP and increase the cleavage C3 convertase, which ultimately inhibits the activation of classical and lectin pathways (Breda et al., 2015; Factor et al., 2012). These proteins also bind to Factor H to inhibit the activation of the alternate pathway of the complement system (Factor et al., 2012). Pathogenic *Leptospira* enhance the etiology by interacting with the host plasminogen (PLG) and converting it to plasmin (Verma et al., 2010; Vieira et al., 2009). Plasmin is a serine protease that degrades various host substrates such as fibrinogen, ECM, and complement molecules C3b and C5 (Barthel et al., 2012). In addition, the pathogenic *Leptospira* secretes proteases such as Thermolysin that directly cleave complement proteins, affecting all three complement pathways (Fraga et al., 2014).

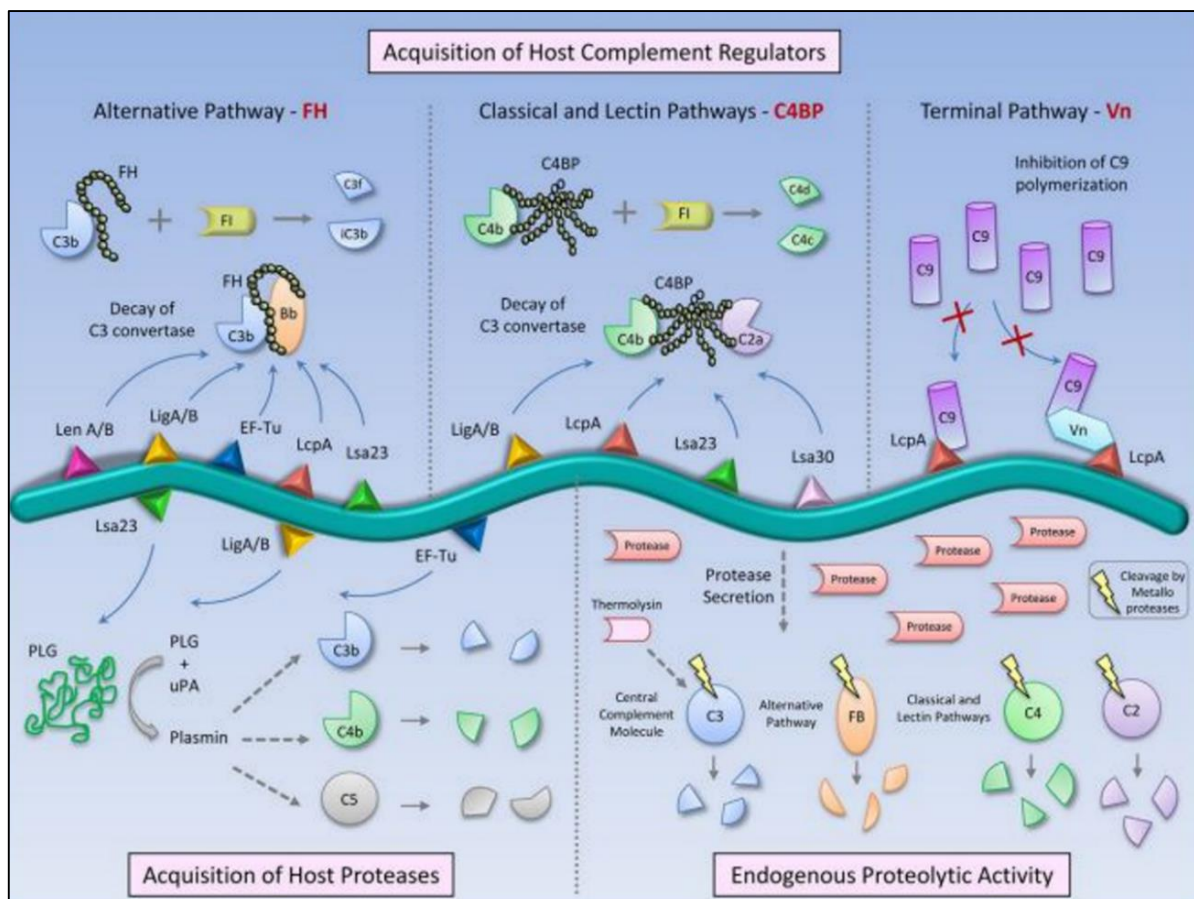


Figure 1.7: Complement evasion strategies and the associated proteins of pathogenic *Leptospira* (Adapted from Fraga TR., *Front. Immunol.*, 2016).

1.9. *Leptospira* vaccines

The first *Leptospira* vaccine was developed in 1916 using phenol-inactivated *Leptospira*. The vaccine was successfully tested in guinea pigs and showed protective immunity against leptospiral infections. Between 1919 and 1921, Japan demonstrated the first large-scale human testing of vaccines. Besides phenol, many other methods such as freezing, heat, ethanol, formalin, and irradiation have been used to kill *Leptospira* to develop an improved vaccine. However, to date, inactivated whole-cell bacteria have been used as an approved vaccine for dogs, cattle, and pigs, and due to undesirable side effects, their use in humans is not recommended. Moreover, the immunity from many known vaccines is serovar specific and doesn't provide cross-protection (B. M. Naiman et al., 2001). In many countries, combinations of up to eight different *Leptospira* serotypes have been used for vaccine purposes, but efficacy of these vaccines has been found to be very low. Live attenuated vaccines have been developed by continuous passage of serotype Pomona but do not provide long-term protection (J.D Allen, 1982). The lipopolysaccharides (LPS) derived vaccine showed generation of antibodies in hamster models (Jost et al., 1989). However, due to the presence of multiple sera, the high-cost implications and complex structure of LPS make it difficult to be used as a vaccine. The availability of genome sequences and advancements in recombinant DNA technology has accelerated the process of antigen identification. Leptospiral antigens are being explored for the development of an efficacious vaccine. A list of antigens, adjuvant, model organisms, and protection rates is given in **Table 1.2**.

Table 1.2: Immunological studies using the Leptospiral antigens

Antigen Type	Animal model	Protection rate (%)	Notes	Reference
<i>lipL32</i> adenovirus	Gerbil	86 %	50% survival in control	(Branger et al., 2001)
<i>lipL32</i> DNA	Gerbil	60 %	35% survival in control	(Branger et al., 2005)
LipL32	Guinea pig	50%	13-60% survival in control	(Luo et al., 2009)
LipL32- LTB fusion	Hamster	40–100 %	0-60% survival in control	(Enterotoxin et al., 2012)
LigA or LigB C–terminus. GST fusions	C3H/HEJ mouse	90 %	40% survival in GST control	(Koizumi and Watanabe, 2004a)
LigA C–terminus plus LigB C–terminus. GST fusion	C3H/HEJ mouse	100 %	40% survival in GST control	(Koizumi and Watanabe, 2004a)
LigA C-terminus. His tag	Hamster	63–100 %	Higher dose requirement	(Silva et al., 2007a)
<i>ligA</i> DNA	Hamster	100 %	50-75% survival in control	(Faisal et al., 2008)
LigB fragments	Hamster	25-87%	Variability observed	(Yan, 2009)
LigA C-terminus	Hamster	76-92%	Alum adjuvant showed 50% protection	(Faisal et al., 2009)
OmpL1 Plus LipL41 and OmpL1	Hamster	100%	33% survival in control	(Haake et al., 1999)
<i>ompL1</i> DNA	Hamster	33%	Only 2/6 survival rate	(Heterologous et al., 2008)
<i>lipL21</i> DNA	Guinea pig	Nil	All animal survived	(He et al., 2008)
Lp0607, Lp118, Lp1454 combined	Hamster	75%	No protection with liposomes	(Faisal et al., 2009)
Lic12720, Lic10494, Lic12922 His tag	Hamster	30-44%	Nil	(Atzingen et al., 2010)
Lsa21, Lsa66, Lic11030. His tag	Hamster	20-30%	No reproducibility	(Atzingen et al., 2012)
<i>lemA</i> DNA and LemA prime boost	Hamster	87%	Nil	(Hartwig et al., 2013)

As shown in **Table 1.2**, almost all studies performed with recombinant proteins showed problems in the control group. Experimental reproducibility is also one of the significant

concerns in vaccine production. Based on studies conducted using a variety of recombinant proteins, Lig proteins have been found to be a potential candidate for vaccine production.

1.10. Leptospiral Immunoglobulin-like protein (Lig)

Leptospiral immunoglobulin-like proteins (Lig) are the surface-expressed multi-domain protein from *Leptospira* that play an important role in leptospiral pathogenesis. The Lig protein family was identified along with GroEL, LipL41, and DnaK from *Leptospira*-infected human serum by screening bacterial expression libraries (Palaniappan et al., 2002b). The expression of Lig proteins is increased during infection in mammals, demonstrating their importance in pathogenesis (Matsunaga et al., 2013, 2007, 2005a). The family of *Leptospira* consists of three proteins LigA, LigB, and LigC. They are one of the most studied leptospiral proteins in terms of pathogenicity, diagnostic markers, and potent vaccine candidates. Several reports (mentioned above) show that vaccines developed using the Lig protein can provide more than 90% protection against leptospiral infections in hamster models. In addition, the Lig protein is exclusively conserved throughout the pathogenic serovars.

1.10.1. Domain organization of Lig

The Lig proteins are multi-domain proteins consisting of 12-13 tandem Ig-like fold domains. LigA is composed of 13 Ig-like domains, while LigB and LigC are composed of 12 Ig-like domains, and each domain is composed of 90 amino acids (Matsunaga et al., 2003b). The overall domain orientation of Lig protein is shown in **Figure 1.8**. LigB and LigC have an extra non-Ig-like domain at the C-terminal instead of the 13th domain. Each domain of Lig is connected by small 3-4 amino acid linkers, which probably provide flexibility to the protein.

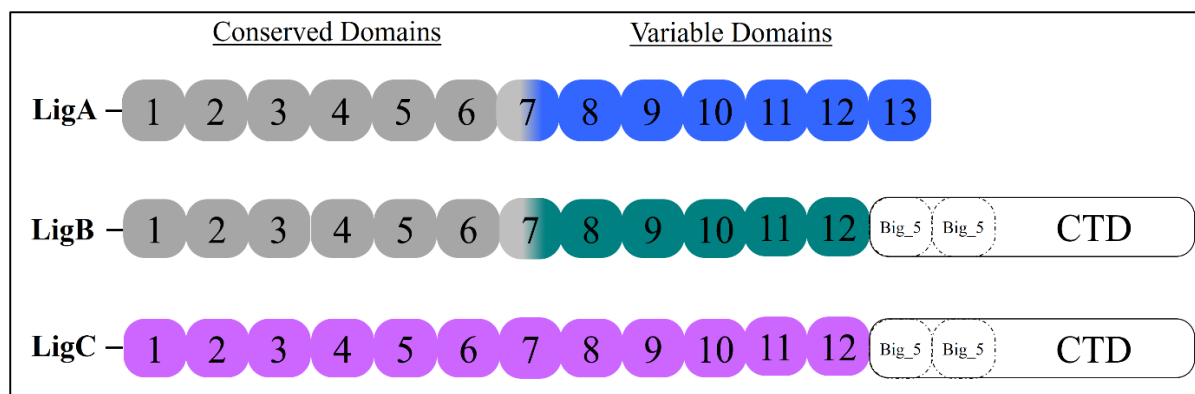


Figure 1.8: Schematic representation of the domain organization of Leptospiral immunoglobulin-like protein (Lig).

The first six and a half N-terminal domains of LigA and LigB are identical and known as constant domains. Rest other domains from 6 & 1/2 to 13 in case of LigA and up to 12 in case of LigB are unique and designated as a variable domain. LigC is a pseudogene expressed as a truncated form in some species with a domain organization similar to LigB. Both LigA and LigB variable regions have been shown to be more antigenic compared to conserved regions. The C-terminal domains (CTDs) of LigB and LigC show 51% sequence similarity among each other (Matsunaga et al., 2003a). The total molecular weight of LigA is about 130 kDa, and the molecular weight of LigB is about 212 kDa. The C-terminal domain of LigB is composed of 772 amino acids. The structure and functionality of the LigB and LigC C-terminal domains are not known, yet. The CTD of LigB and LigC may have important structure and function. The 3-dimensional structure of CTD will provide a complete functional insight of the same in the LigB and LigC. Lig proteins are usually attached to the outer membrane via LipoBox (Haake et al., 2000).

1.10.2. Structure

All the domains of LigA and LigB are predicted to have structural similarities with the immunoglobulin-like fold like a Fab fragment of human immunoglobulin (Saul, 1973). The protein folds are named Big-2 and are also commonly found in the structure of intimin of *E.coli* and invasin of *Yersinia*. Three- dimensional structure of the full length of LigA and LigB is not available. Only, NMR solution structures of single domain LigB12 (2MOG) and LigA4 (2N7S) are present in the protein database (Mei et al., 2015a, 2014; Ptak, 2014). These structures have been reported to adopt a classical β -sandwich-like fold. A cartoon representation of LigB12 and LigA4 is shown in **Figure 1.9**.

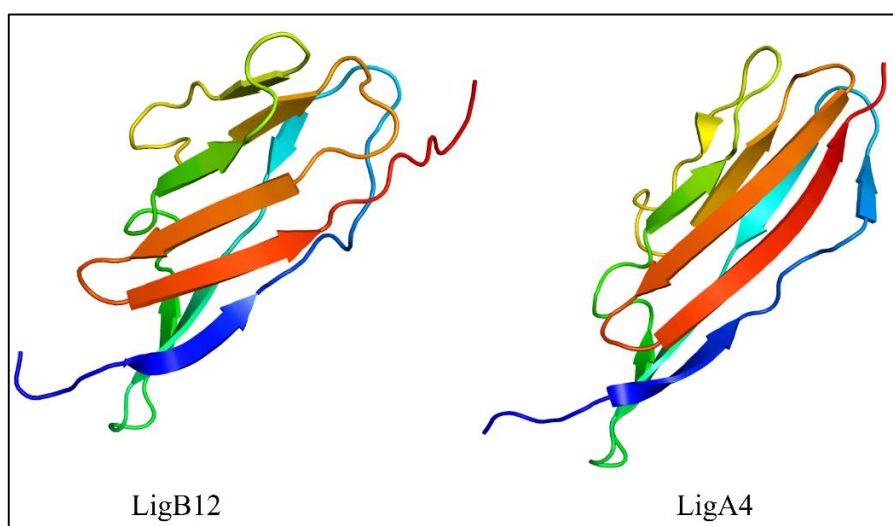


Figure 1.9: Solution structure of LigB12 and LigA4.

The first LigB and LigC CTD sequences (two domains) showed similarities to the Big_5 domain. A recent, small-angle X-ray scattering (SAXS) study on the full domain LigB1-12 showed that the protein attained elongated conformation and interdomain involved in salt bridge interactions (Hsieh et al., 2017).

1.10.3. Calcium-binding

Various *in-vitro* studies have demonstrated that LigA and LigB bind to calcium ions. Calcium-binding has been mapped in many Ig-like domains of the proteins. Individual Ig-like domains LigA9 and LigA10 showed binding to calcium. In other studies, it was demonstrated that LigBCen2, which comprised LigB12 plus extra amino acids at its amino- and carboxy termini, binds to calcium. The chemical shifts in the previous NMR experiments on LigA4, LigBCen2, and LigBCen2R (comprising LigB12 and an N-terminal extension), in the presence of Ca^{2+} ion, have suggested proteins' ability to bind to Ca^{2+} ions. Other groups did not observe a similar *in-vitro* binding. Hence, calcium-binding to Lig proteins is still debatable. Moreover, the physiological role of calcium-binding in context to the conventional function of Lig Protein is not clearly understood. Crystallographic studies on Ig-like domains or full-length Lig proteins may provide structural determinants of calcium-binding within the protein.

1.10.4. Lig in leptospiral pathogenesis

It is well established that the Lig proteins are involved in leptospiral pathogenesis. Earlier studies suggest that the expression of the LigA gene is induced during infection. The reduction of LigA and LigB expression decreases the virulence *L. interrogans*. Moreover, their expression also increases at physiological osmolarity and temperature. Targeted mutagenesis studies also have furnished evidence that expression of either *ligA* or *ligB* is required for virulence. In another study, an artificial transcriptional repressor of the lig genes using a Transcription activator-like Effector (TALE) in the L495 strain of *L. interrogans* was observed to be avirulent following intraperitoneal injection into the hamster model.

Studies suggest that *L. interrogans* adherence to the host ECM is correlated with leptospiral virulence and pathogenesis. This has been an essential phenomenon for establishing host-pathogen interactions during leptospiral infection. Many *in-vitro* reports suggested that the Lig

proteins bind to various host ECM components such as fibronectin, elastin, tropoelastin, collagen, and laminin (Ito and Yanagawa, 1987; Toshihiro ITO, 1987). The *in-vivo* investigation of adhesion to *L. interrogans* on MDCK cells is inhibited if *L. interrogans* is premixed with fibronectin, suggesting the involvement of host fibronectin in facilitating *Leptospira* adhesion. Further, the fine mapping of regions/domains on the Lig proteins interacting with the host ECM has also been demonstrated. The variable regions of LigA and LigB displayed high-affinity binding with the fibronectin (Henry A Choy et al., 2007a; Lin et al., 2010). The affinity of LigB for fibronectin was significantly higher. Moreover, an increase in affinity towards the fibronectin was observed when the first Big_5 domain was introduced, along with the variable domains of LigB (Henry A Choy et al., 2007a). However, the N-terminal Ig-like constant regions of LigA and LigB do not interact with the fibronectin. Usually, fibronectin's gelatin binding domain (Henry A Choy et al., 2007a)(GBD) interacts with the Lig proteins (Henry A Choy et al., 2007a). Several reports have shown the ability of LigB7-12 to bind collagen I and collagen IV with a moderate affinity, while LigA7-13 has a low binding affinity. The binding affinity with laminin showed similar results, the LigB7-12 variable region showed moderate binding, and the LigA8-13 binding was reported as undetectable. Heterologous expression of LigA and LigB on the surface of non-pathogenic *Leptospira* spp. *L. biflexa* displayed adherence to fibronectin and laminin. However, this expression did not show increased binding with elastin, collagen I, or collagen IV (Figueira et al., 2011).

Various groups have *in-vitro* characterized/mapped the domains of LigA and LigB involved in host protein interaction. It has been demonstrated that the binding affinity of LigB for fibronectin and fibrinogen occurs mainly through LigB9-10. Its binding affinity is enhanced by adding the 11th Ig-domain of LigB. However, studies conducted by other groups contradict these findings. In another study, the LigB12 domain and the additional 47 residues from the

non-Ig-like portion of LigB have demonstrated a higher binding affinity with the N-terminal domains of fibronectin (Y. Lin et al., 2009a).

Interestingly, both LigA and LigB variable domains have been reported to show a good affinity for fibronectin, but a single truncated Lig domain did not even show moderate affinity. It has been suggested that the interaction with host GBD requires multi-domains of LigA and LigB (Henry A. Choy et al., 2007). However, the individual domain LigB4 showed the highest binding affinity toward tropoelastin (Y. Lin et al., 2009b).

1.10.5. Lig in Immune evasion

To survive in the host during infection, the *Leptospira* must resist the host's immune attack. Various host factors, such as antibacterial peptides, neutrophils, macrophages, and the bactericidal activity of the host's complement system, are involved in the removal of pathogen load from the body. Pathogenic *Leptospira* protects itself within the host using a variety of immune escape strategies. As mentioned above, the Lig protein has demonstrated the ability to control the activation of all three pathways of the complement system by interacting with complement regulators such as C4BP, Factor H, FHL-1, and FHR-1 (**Figure 1.10**). The role of Lig protein in immune evasion and their interaction with the complement regulator proteins has been described in various reports. Through various biochemical analyses, Breda *et.al.* has performed fine mapping of Lig and C4b-Binding protein interaction (Breda et al., 2015). Their results have demonstrated that the Lig and LcpA proteins bind to complement control protein (CCP) via CCP7, CCP8, and CCP4 domains of C4BP. The Lig domains such as LigA7-8, LigB7-8, and LigB9-10 showed strong binding to CCP 7 and 8 domains of C4BP. Moreover, these interactions were found to be sensitive to ionic strength. Similarly, the LigB domains (LigB9-11) are reported to bind directly to C3b and C4b to inhibit both alternate and classical pathways of the human complement system (Choy, 2012).

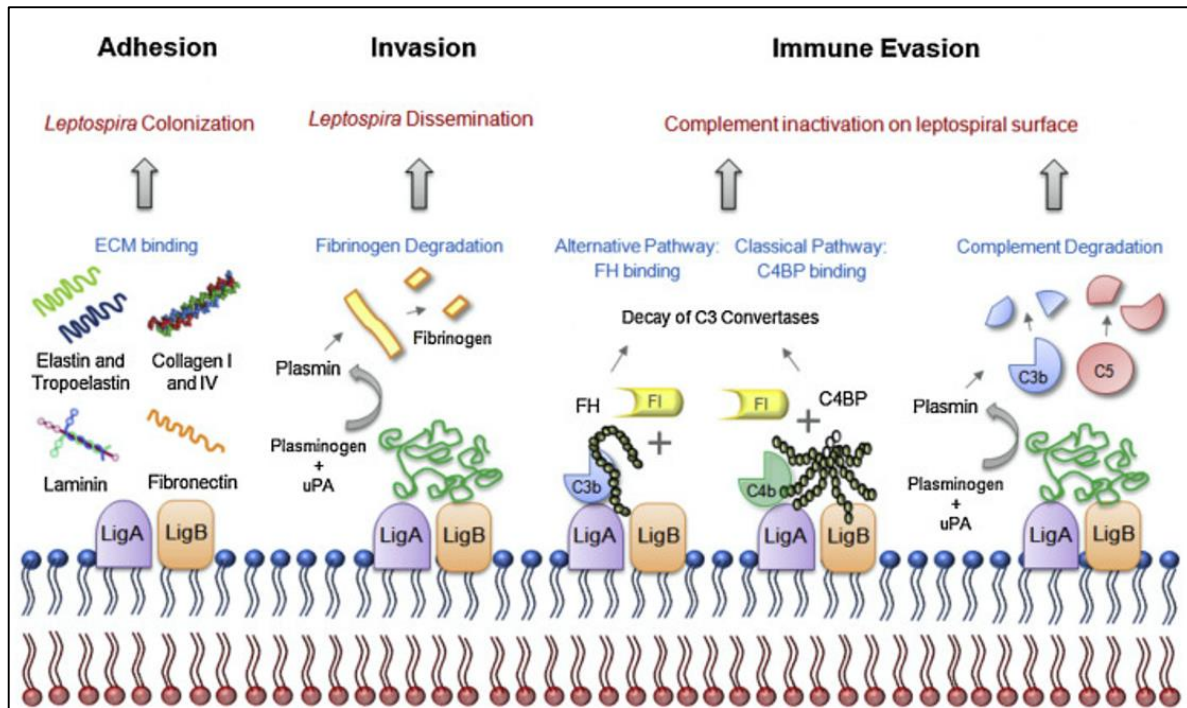


Figure 1.10: Lig plays multiple roles during *Leptospira* pathogenesis (Adapted from Mónica, *Immunobiology*, 2016).

1.10.6. Lig as a diagnostic marker and vaccine candidate

LigA and LigB are potential diagnostic markers for seroreactivity in patients with acute leptospirosis due to their expression in the early stage of infection. Their expressions have been observed during the etiology of *Leptospira* (Matsunaga et al., 2003a). Many studies have shown that the variable Ig-like domains of LigA (68-1224) and LigB (68-1191) provide high levels of protection but, at the same time, do not provide sterile immunity (Koizumi and Watanabe, 2004a). Recent reports have shown 80-100% protection and 77 -100% sterile immunity in hamster models using recombinant LigB (131-645) fragments. Animals immunized with rLigB (131-645) showed IgG antibody production, but IgG levels were lower compared to the bacterin group (Souza et al., 2017). The combination of the LigA variable region and the conserved region was used as a DNA vaccine in a study that showed increased antibody titration and 100% survival in a hamster challenge model, but the control group also

showed 62.5% survival (Faisal et al., 2008). In another study, when testing five shortened LigA and LigB DNA vaccines, a maximum survival rate of 62.5% was observed in hamsters containing LigB conserved portions (Forster et al., 2013). A chimera consisting of a conserved LigB and a viable LigA region showed 100% protection in hamsters as a DNA vaccine but was unable to confer sterile immunity (Eduardo et al., 2019). Various studies have been performed with Lig and its fragments and combination of different adjuvants for vaccine development. A study by Haake et al. showed that a fragment of LigA10-13 was sufficient to provide 100% protection in hamster challenge models (Coutinho et al., 2011).

Altogether, Lig proteins are associated with various functions such as adhesion, invasion, evasion, and many more. The role of Lig proteins at multiple stages of leptospiral infection is summarized in **Figure 1.11**.

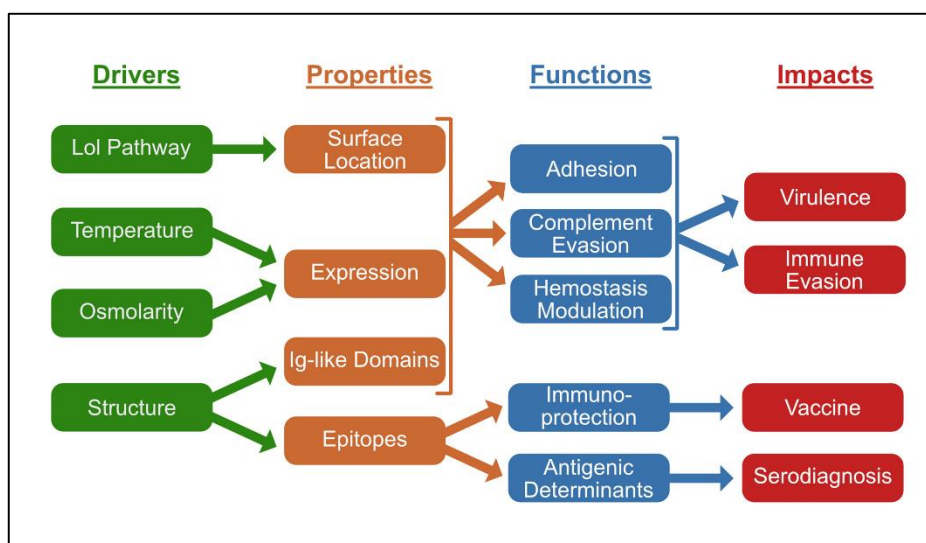


Figure 1.11: Summary of Lig proteins and their association at different stages during leptospirosis (Adapted from Haake et., *Front. in immunology*, 2021).

1.11. Major challenge and aim of the thesis

Although the pathogen was identified over 100 years ago, effective vaccines are not yet available. The presence of multiple serotypes makes vaccine development a difficult task, as the vaccines developed do not provide cross-protection. The existing vaccines have undesirable side effects, require multiple doses, and are less effective. In addition, nonspecific symptoms in the early stages of infection disrupt the diagnostic process. Several antigens such as LigA, LigB, LipL32, and LenA have been identified, and their roles in the establishment of host-pathogen interactions have been demonstrated. Still, the atomic details of those interactions are unknown. The family of Lig proteins is a big multi-domain protein whose size ranges from 130-220 KDa, and its immunogenic region is not well characterized (Haake and Matsunaga, 2021). Recently, it was shown that antigenic motifs in a single-domain chimeric immunoglobulin-like fold generate enhanced leptospiral protection compared to whole Lig protein (Hsieh et al., 2017). Hence, it is thought that the knowledge of the immunogenic region/epitope will provide opportunities for further improvement in the Lig family vaccine efficacy through rational design. The three-dimensional structure of protein usually facilitates in this direction. Unfortunately, the complete 3-D structural details of full-length Lig proteins are largely unknown. However, recently only NMR structures of single domains LigB12 and LigA4 have been reported (Mei et al., 2015a; Ptak, 2014). No X-ray crystallographic studies have been reported so far on Lig and any of its small fragments. Several efforts are being put to fine map the antigenic regions among the multi-domain Lig proteins. A recent report suggests that pathogenic *Leptospira* evades host complement, as well as a neutrophil extracellular trap (NET), mediated killing (Scharrig et al., 2015). Few other pathogenic organisms evade host NET by secreting nucleases that degrade the DNA content of the NET (Andre et al., 2017; Fairman et al., 2012; Storisteanu et al., 2017). Surprisingly, no secreted nucleases have been reported from the *Leptospira* genome so far. In the absence of secreted

nucleases, it is hypothesized that other surface or secreted proteins from *Leptospira* may have the ability to cleave DNA. A report suggests that proteins with Ig-like folds having His-ME finger domain may act as a nuclease (Jablonska et al., 2017a). Since Lig is a multifaceted protein and its role in hydrolysing DNA is not been investigated. Based on this background information, the questions warranted to address are; what is the full-length 3-dimensional structure of Lig family proteins and their various smaller fragments, and what is their domain orientation? What are potential antigenic epitopes present on Lig proteins? And whether Lig proteins possess any nuclease/DNase activity that may substitute the function of nuclease.

1.12. Objectives of the study

1. Structural characterization of leptospiral immunoglobulin-like (Lig) proteins
2. Identifying antigenic regions within the Lig proteins
3. Investigation of novel nuclease and protease activities among the Lig proteins

Chapter - 2

*Cloning, Overexpression and
Purification of Lig proteins and its
fragments*

2.1. Introduction

A family of Leptospiral Immunoglobulin-like (Lig) proteins is an important adhesin expressed on the outer surface of pathogenic *Leptospira* (Matsunaga et al., 2003b). Members of the Lig family, also known as microbial surface components recognizing adhesive matrix molecules (MSCRAMMs) and are usually localized on the outer surface of *Leptospira* (Pinne et al., 2012). Proteins of this family consist of 12-13 extracellular homologous immunoglobulin-like repeats and are anchored to the outer membrane with an N-terminal Lipo-box (Matsunaga et al., 2003b). It includes three proteins, LigA, LigB, and LigC. The LigA is a 130 KDa protein consisting of 12 Ig-like repeated domains. LigB is almost double the size of LigA and possesses 13 Ig-like domains and a non-random domain at the C-terminal. LigC is found to be a pseudogene (Pinne et al., 2012). These proteins are usually present only in pathogenic species whose expression is transcriptionally upregulated during the infection and an important in *Leptospira* pathogenesis (Matsunaga et al., 2005b). Lig proteins family has been characterized as the best antigen and a potential diagnostic marker for pathogenesis (Croda et al., 2007). The Ig-like domain of the Lig proteins, also known as bacterial immunoglobulin-like (Big) repeat domains, was initially identified in virulence determinants from *E.coli* and *Yersinia pseudotuberculosis* (Matsunaga et al., 2003c). Emerging serologic vaccine and pathogenesis studies indicate Lig proteins as crucial virulence determinants in host-pathogen interaction. As it belongs to MSCRAMMs, it was shown to bind to ECM such as laminin, collagen, fibronectin (Fn), fibrinogen, elastin, and tropoelastin of the host cells (Choy et al., 2007b; Y.-P. Lin et al., 2009; Lin et al., 2010). The variable region of LigA consists of Ig-like domains 7-13 of LigA, and 6 Ig-like domains (6-12) constitute a variable region of LigB (Palaniappan et al., 2002a). These variable regions from the LigA and LigB bind to host ECM (Choy et al., 2007b). Each Big domain consists of approximately 90-100 amino acid residues (Matsunaga et al., 2003c). Proteins or antigens with multiple repetitive domains, with large molecular sizes, like Lig

protein family, usually show poor expression and are unamenable to purifying recombinantly with a good yield. Such proteins can be challenging to crystallize for structural studies. Smaller fragments of such proteins may provide a better yield and overcome the limitation of crystallization. This chapter reports the cloning, expression, and purification of various fragments of LigA and LigB using multiple chromatographic techniques.

2.2. Materials

Chemicals and reagents used for cloning, expression, and purification of LigA and LigB and their various fragments were obtained from several commercial sources.

2.2.1. Chemical and Reagents used in cloning

All chemicals were purchased from Sigma Aldrich. Plasmid isolation and DNA gel-extraction kits were procured from Qiagen and Thermo scientific. DNA polymerases, dNTPs, T4 DNA Ligase, and Restriction endonucleases were procured from New England Bio Labs (NEB). The quick-site-directed mutagenesis kit was procured from Invitrogen. Oligonucleotides were purchased from Integrated DNA Technology, Inc.

The plasmids containing a single Ig-like fragment of *LigA* and *LigB* as well as *LigB8-12* and *LigB9-12* were gifted by Prof. Yung-Fu Chang, Cornell University, USA

2.2.2. Regents used in protein purification and detection

Chemicals including Sodium phosphate, Tris, NaCl, glycerol, imidazole, NiSO₄, DTT, IPTG, X-gal, β -casein, Ni-NTA, etc. were procured from Sigma Inc. (MERCK). All pre-packed columns such as His-Traps and size-exclusion chromatography were purchased from GE Healthcare Life Sciences.

2.2.3. Recipes of reagents used

All media, buffers, and stocks used in cloning, expression, and purification were prepared either in Milli-Q or double-distilled water with low conductivity. All stock solutions were prepared under standard procedures described in Sambrook *et al.* 1989. Recipes/compositions of stock solutions are mentioned in the following tables.

Table 2.1: Composition of culture medium

Luria Bertani Broth (LB)	10g tryptone +5g yeast extract +10g NaCl per litre in double-distilled water. pH was adjusted to 7.2 with NaOH. Sterilized by autoclaving
--------------------------	---

Table 2.2: Composition of antibiotic stocks

Antibiotic	Stock solution*	Working concentration
Ampicillin	100mg/ml in Milli-Q	100µg/ml
Kanamycin	50mg/ml in Milli-Q	50µg/ml
Chloramphenicol	34mg/ml in ethanol	34µg/ml

* Sterilized by passing through a 0.22µm filter

Table 2.3: Composition of solutions/reagents used for agarose gel electrophoresis

Reagents	Compositions
50X TAE	242g Tris base +57.1ml of glacial acetic acid + 100ml of 0.5M EDTA per litre
6X sample loading dye	0.25% Xylene, 0.25% Bromophenol Blue, 30% Glycerol
Ethidium bromide	Stock of 10mg/ml (Puregene; Cat no- PG-823)

Table 2.4: Composition of solution used for protein purification

Buffer	Composition
1X PBS Solution	NaCl:137 mM, KCl: 2.7 mM, Na ₂ HPO ₄ : 10 mM, KH ₂ PO ₄ :1.8 mM, pH 7.4
Lysis Buffer	1X PBS and 5mM Imidazole
Washing Buffer	1X PBS and 30mM Imidazole
Elution Buffer	1X PBS and 250mM Imidazole

Table 2.5: Composition of solutions used for SDS-PAGE.

Reagents	Compositions
30% acrylamide	29.2% acrylamide +0.8% Bis-acrylamide
Stacking Buffer	1.5M Tris-HCl, pH8.8 + 0.4% SDS
Resolving buffer	1M Tris-HCl, pH 6.8 + 0.4% SDS
1X Running buffer	3g Tris-HCl +14.4g glycine +1g SDS per liter
De-staining solution	Water: Acetic acid: Methanol:: 5:1:4
Staining solution	2g/L of coomassie brilliant blue R250 in de-staining solution
1X Laemmli sample buffer	10% glycerol +1% β-mercaptoethanol + 2% SDS + 0.1% bromophenol blue in 1X separating buffer

2.3. Experimental Procedures

2.3.1. Preparation of Competent cells

The competent chemical cells such as DH5 α , BL21 (DE3), and Rosetta (DE3) used for cloning and overexpression were prepared by the calcium chloride method. The inoculum of cells was added to 5ml of Luria Broth medium (Himedia) with or without antibiotics, followed by incubation at 37°C overnight in a shaker incubator. The next day, 1% culture from the primary culture was added to 100ml of secondary culture and grown till the OD₆₀₀ reached 0.4-0.6 at 37°C. The grown culture was then incubated on ice for 30min followed by centrifugation at 6000rpm for 30min while the centrifuge was maintained at 4°C. The pellet and supernatant

were separated, and the pellet was suspended in 100mM calcium chloride solution. The suspended cells in calcium chloride were incubated on ice for 3-4hrs. The cells were then centrifuged at 6000 rpm again, and the supernatant was discarded. The cell pellet was dissolved in a solution containing 100mM calcium chloride and 10% glycerol. The solution was aliquoted in 1.5ml centrifuge tubes followed by snap freezing using liquid nitrogen and stored at -80°C until use. The efficiency and the contamination in the competent cells were checked before using them for transformation.

2.3.2. Cloning of fragments of *LigA* and *LigB* into the expression vector

The fragments of *ligA* gene (*ligA8-12*, *ligA8-11*, *ligA8-10*, *ligA8-9*, *ligA9-13*, *ligA10-13*, *ligA11-13*, and *ligA12-13*) and fragments of *ligB* (*ligB7-11*, *ligB7-10*, *ligB7-9*, *ligB7-8*, *ligB10-12*, *ligB11-12* and *ligB_CTE*) were PCR amplified from the existing clones of *LigA8-13* and *ligB7-12*, respectively, using oligonucleotides primers mentioned in **Table 2.6**. The optimized PCR condition and PCR cycles for all the mentioned fragments are mentioned in **Tables 2.7** and **2.8**. The PCR amplified gene fragments were cloned in the pET28a-SUMO expression vector under BamHI and HindIII restriction enzyme sites. Confirmed clones were selected on the basis of double digestion with the restriction endonucleases. The cloning was further confirmed by DNA sequencing. The respective clones in pET28a-SUMO are listed in **Table 2.9**.

Table 2.6: List of primers used for generation of LigA and LigB fragments

Construct	Reverse Primer
LigA8F	5'-GCATGGATCCACACAGGCGACTTTGACTTC-3'
LigA8-12R	5'-GGCGAAGCTTCTAACTTTCCGTAACCGTAACTG-3'
LigA8-11R	5'-GAACGAAGCTTCTAAAGAAGCGCTGGAGTGAC-3'
LigA8-10R	5'-GTCCAAGAAGCTTCTATACTTTAGCCGGAGTAACTTG-3'
LigA8-9R	5'-GTAAGAAGCTTCTAAAGTTCCGCTGCGGTAACG-3'
LigA13R	5'-GGACGGAAGCTTCTATAATATTTCTGGAGTTACTTCAAAATC-3'
LigA9-13F	5'-AAAAGGATCCACTCCCGCAATTCTTACTTCAATT-3'
LigA10-13F	5'-GAATGGATCCACCGCAGCGGAAGTTATT-3'
LigA11-13F	5'-GAATGGATCCACTCCGGCTAAAGTAGTTTCGAT-3'
LigA12-13F	5'-GAATAGGATCCACTCCAGCGCTTCTTCGTTAC-3'
LigB12R	5'-GTTAAGCAAGCTTCTACGTGTCCGTTTTGTTTACTGTG-3'
LigB10-12F	5'-GCCTTCGGATCCACTGACTTAAAACTGAAAAGTATAAC-3'
LigB11-12F	5'-GAAAAGGATCCGCTGCCACGTTAGATTCC-3'
LigB7F	5'-GAAAAGGATCCACAGCTGCAAAGCTTGTTGAA-3'
LigB7-11R	5'-AAAGAAAGCTTCTAAAGGGTTGCTGCGCTG-3'
LigB7-10R	5'-GGCAAAGAAGCTTCTATAACGTGGCAGCACTTAC-3'
LigB7-9R	5'-CCGGGCAAGCTTCTACAGTTTTAAGTCAGTGAC-3'
LigB7-8R	5'-CAAGGAAGCTTCTAAAGAAGTGCAGGAGTGACATT-3'
LigB_C-Ter F	5'-GGCCATATGATAGCTCCGACGGTTCAATCC-3'
LigB_C-Ter R	5'-AAAGCGGCCGCTATTGATTCTGTTGTCTGTAAAT-3'

* Restriction sites are underlined, F-Forward primer and R- Reverse primer

Table 2.7: Optimized concentration of PCR reactions

Components	Concentration
Template	20ng
Forward Primers	0.5 μ M
Reverse Primers	0.5 μ M
dNTPs	100 μ M
DNA Polymerase	1 unit

Table 2.8: Optimized PCR cycle parameters

Gene	Initial Denaturation	Denaturation	Annealing	Extension	Final Extension
LigA8-12	98°C (30s)	98°C (10s)	61°C (30s)	72°C (45s)	72°C (10 min)
LigA8-11	98°C (30s)	98°C (10s)	61°C (30s)	72°C (45s)	72°C (10 min)
LigA8-10	98°C (30s)	98°C (10s)	63°C (30s)	72°C (30s)	72°C (10 min)
LigA8-9	98°C (30s)	98°C (10s)	63°C (30s)	72°C (30s)	72°C (10 min)
LigA9-13	98°C (30s)	98°C (10s)	61°C (30s)	72°C (45s)	72°C (10 min)
LigA10-13	98°C (30s)	98°C (10s)	61°C (30s)	72°C (45s)	72°C (10 min)
LigA11-13	98°C (30s)	98°C (10s)	63°C (30s)	72°C (30s)	72°C (10 min)
LigA12-13	98°C (30s)	98°C (10s)	63°C (30s)	72°C (30s)	72°C (10 min)
LigB7-11	98°C (30s)	98°C (10s)	62°C (30s)	72°C (45s)	72°C (10 min)
LigB7-10	98°C (30s)	98°C (10s)	58°C (30s)	72°C (45s)	72°C (10 min)
LigB7-9	98°C (30s)	98°C (10s)	61°C (30s)	72°C (30s)	72°C (10 min)
LigB7-8	98°C (30s)	98°C (10s)	61°C (30s)	72°C (30s)	72°C (10 min)
LigB10-12	98°C (30s)	98°C (10s)	61°C (30s)	72°C (30s)	72°C (10 min)
LigB11-12	98°C (30s)	98°C (10s)	61°C (30s)	72°C (30s)	72°C (10 min)
LigB C_Ter	98°C (30s)	98°C (10s)	57°C (30s)	72°C (1.15m)	72°C (10 min)

* s- Seconds, min- Minutes

Table 2.9: Bacteria Strains/plasmids used in this study

Bacterial Strains/plasmids	Description	Source
Strains		
DH5 α	λ^- ϕ 80dlacZ Δ M15 Δ (lacZYA-argF)U169 <i>recA1 endA hsdR17</i> (rk $^-$ mk $^-$) <i>supE44 thi-1 gyrA relA1</i>	(Taylor et al., 1993)
BL21(DE3)	F $^-$ <i>ompT hsdS_B</i> (r _B $^-$ m _B $^-$) <i>gal dcm</i> (DE3)	(Jeong et al., 2009)
Rosetta (DE3)	F $^-$ <i>ompT hsdS_B</i> (r _B $^-$ m _B $^-$) <i>gal dcm</i> (DE3) pRARE (Cam ^R)	Novagen
Plasmids/Clones		
pET28a-His-SUMO	P _{T7} -based expression vector, with SUMO tag	ThermoFisher
pET28a/ligA8-13	pET28a-SUMO bearing <i>ligA8-13</i>	This study
pET28a/ligA8-12	pET28a-SUMO bearing <i>ligA8-12</i>	This study
pET28a/ligA8-11	pET28a-SUMO bearing <i>ligA8-11</i>	This study
pET28a/ligA8-9	pET28a-SUMO bearing <i>ligA8-9</i>	This study
pET28a/ligA9-13	pET28a-SUMO bearing <i>ligA9-13</i>	This study
pET28a/ligA10-13	pET28a-SUMO bearing <i>ligA10-13</i>	This study
pET28a/ligA11-13	pET28a-SUMO bearing <i>ligA11-13</i>	This study
pET28a/ligA12-13	pET28a-SUMO bearing <i>ligA12-13</i>	This study
pET28a/ligB7-11	pET28a-SUMO bearing <i>ligB7-11</i>	This study
pET28a/ligB7-10	pET28a-SUMO bearing <i>ligB7-10</i>	This study
pET28a/ligB7-9	pET28a-SUMO bearing <i>ligB10-13</i>	This study
pET28a/ligB7-8	pET28a-SUMO bearing <i>ligB7-8</i>	This study
pET28a/ligB10-12	pET28a-SUMO bearing <i>ligB7-12</i>	This study
pET28a/ligB11-12	pET28a-SUMO bearing <i>ligB11-12</i>	This study

2.3.3. Transformation

The plasmid containing gene fragments of interest were transformed into competent cells using the heat shock method. Around 1ng of plasmid DNA was added to the competent cells and incubated on ice for half an hour. Heat shock was given for 90 seconds at 42°C, followed by

incubating the cells on ice immediately. One ml of LB was added to the tube and incubated at 37°C in a shaker incubator for 1hr. The competent cells were centrifuged at 4000 rpm for five minutes, then decanted most of the supernatant. The cells were resuspended in the remaining LB medium and spread on LB agar plates containing appropriate antibiotics in the required amount.

2.3.4. Expression and purification of all Lig constructs

The vector containing genes of multiple Ig-like domains of *ligA* and *ligB* were transformed in Rosetta (DE3) competent, and transformants were grown in 500 ml Luria Bertani broth (LB) medium containing the required amount of kanamycin and chloramphenicol. Isopropyl-D-1-thiogalactopyranoside (IPTG) was added to the mid-log-phase (OD_{600} reached 0.4-0.6) culture, which was further incubated overnight at 20°C shaker-incubator maintaining at 120rpm. The cells were harvested by centrifugation at 8000 rpm for 10 min. Cell-pellet was resuspended in 1X Phosphate Buffer Saline (PBS) containing 0.1 mM phenylmethylsulfonyl fluoride (PMSF) and subjected to sonicate to break the cells. The second round of centrifugation at 13,000 rpm for 45 minutes was carried out to remove cell debris. The resulting supernatant was applied to a Ni^{2+} -nitrilotriacetic acid (Ni^{2+} -NTA) column pre-equilibrated with 1XPBS. The supernatant containing the mixture of soluble proteins was allowed to bind on the Ni^{2+} -NTA column. The column was washed (five-column bed volumes) with 1XPBS containing 30mM imidazole, followed by elution with 1XPBS containing 250mM Imidazole. To remove the SUMO tag, the eluted recombinants were treated with ulp-1 and kept overnight for dialysis in 1XPBS. The dialyzed sample was applied again to the Ni^{2+} -NTA column. The desired protein was collected in the flow-through, and the SUMO tag was retained on the column. The protocol followed for the purification is summarized in **Figure 2.1**. The flow-through containing recombinant proteins was further concentrated using an Amicon concentrator with 3,000-Da-cutoff membranes and further purified by loading them on a size-exclusion chromatography column

equilibrated with 1XPBS. The purified protein fractions were analyzed on 15% sodium dodecyl sulphate-polyacrylamide gel electrophoresis. The purified recombinant proteins were pooled and stored at -80 °C until further use.

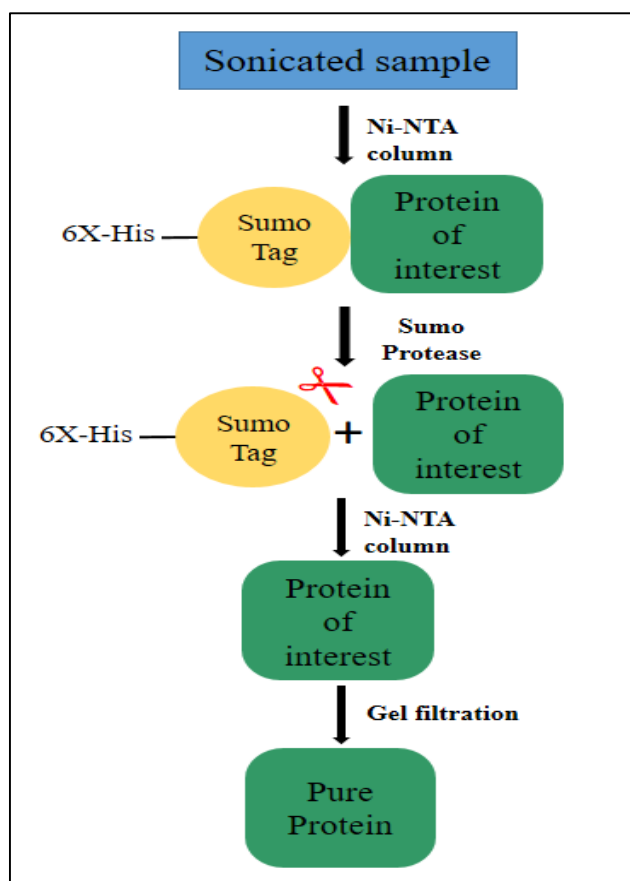


Figure 2.1: Protein purification strategy

2.3.5. Size-exclusion chromatography

The purity and molecular weights of the recombinant proteins were analyzed by size-exclusion chromatography. The chromatography experiments were carried out with Superdex 75 FPLC and 16/600 Superdex 200pg columns from the GE healthcare, using 1XPBS as the column buffer. The void volume of the column was determined using Blue Dextran 200. The elution times/volumes of all the recombinant proteins were recorded, and the molecular weights were calculated by estimating the elution volumes of standards of known molecular weights.

2.4. Results

2.4.1. Cloning of *ligA* and *ligB* fragments

Both LigA and LigB are large and multi-domain proteins with molecular weights of 130 and 220 kDa. The full-length Lig proteins showed weak expression and were inadequate. Purification of full-length protein with good yield was a challenge. Moreover, multi-domain proteins mostly provide restraint for crystallization. To improve the yield and protein solubility, different Lig truncated fragments were generated. The truncation is chosen to cover the complete protein by selecting many smaller fragments containing single as well as multiple Ig-like domains (**Figure 2.2**).

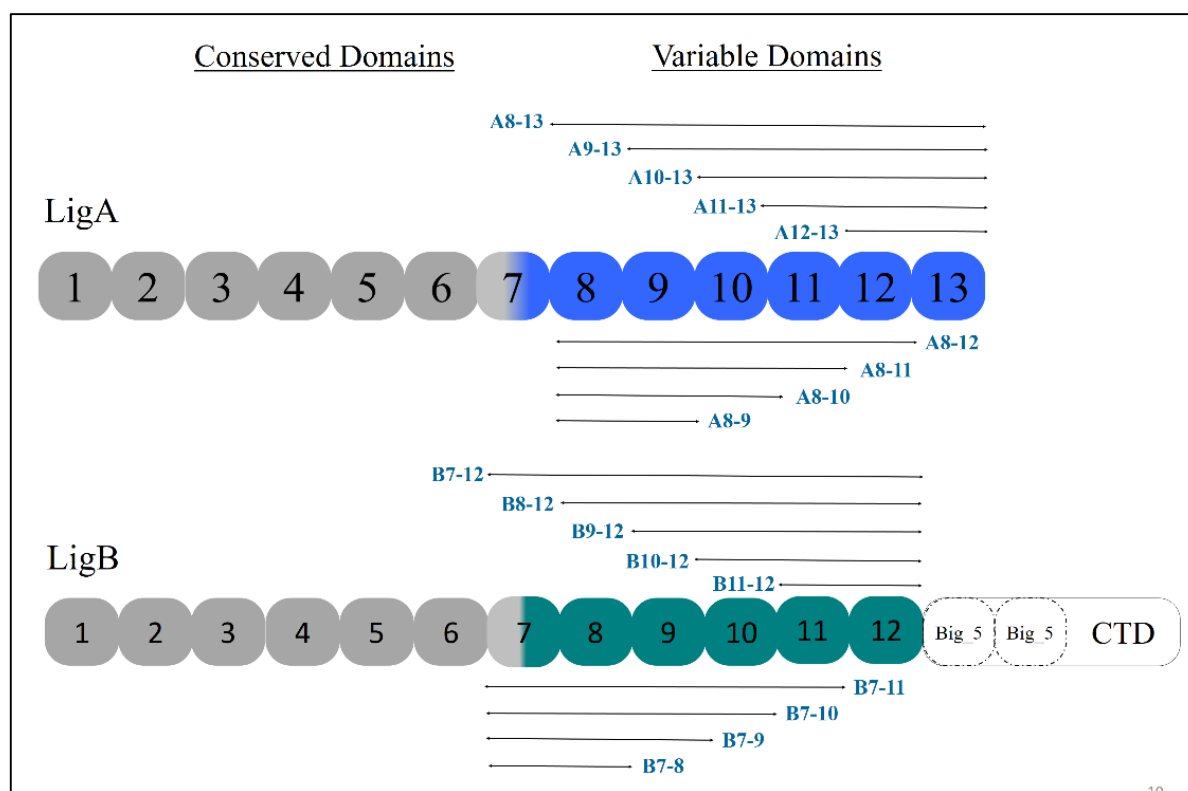


Figure 2.2: Schematic representation of the gene fragments selected for the generation of truncated proteins

The fragments of the gene encoding the particular fragments were amplified using the optimized PCR conditions and checked on 1% agarose gel. The fragments showed amplification DNA bands at their respective sizes (**Figure 2.3**).

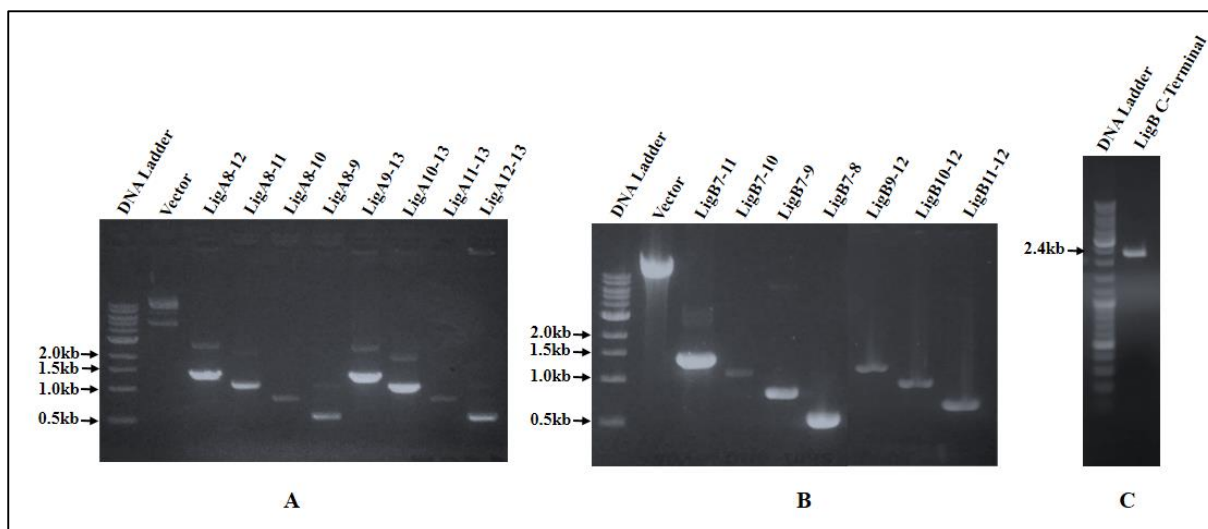


Figure 2.3: 1% Agarose gel showing PCR amplification of different Lig fragments (A) *LigA* fragments amplification (B) *LigB* fragment amplification (C) *LigB* C-terminal Non-Ig-like domain amplification.

Ligation of the PCR products of different fragments was confirmed by double digestion. The fall-out of the correct size of fragments confirms their integration in the expression vector (**Figure 2.4 & 2.5**).

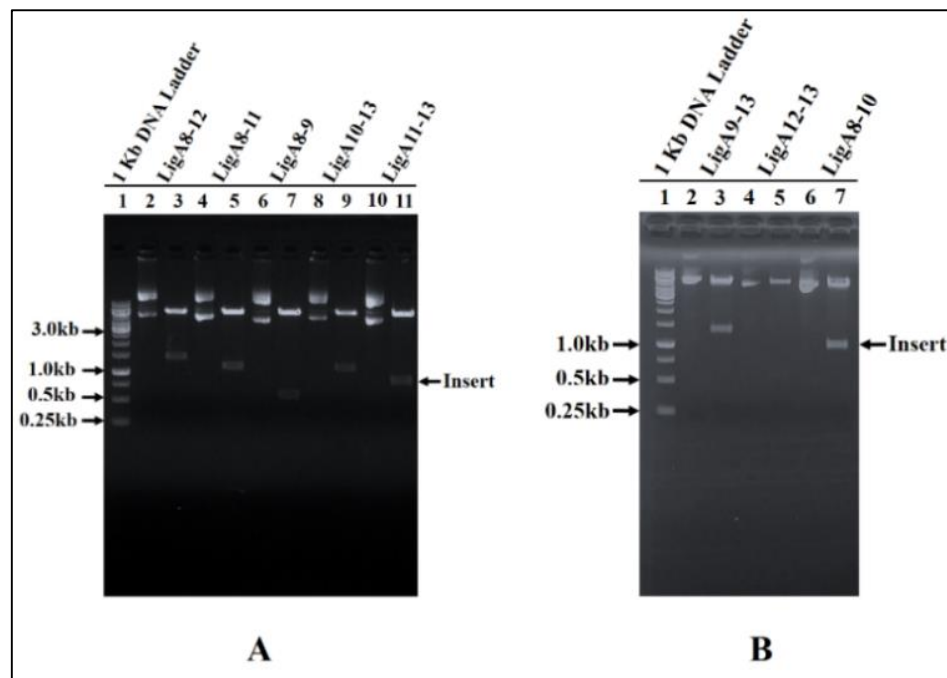


Figure 2.4: 1% Agarose gel showing double digestion of different *LigA* fragments from the expression vector. Lanes are labelled with the fragments of *LigA* and 1kb DNA ladder. Fall-out of the fragments is indicated with insert.

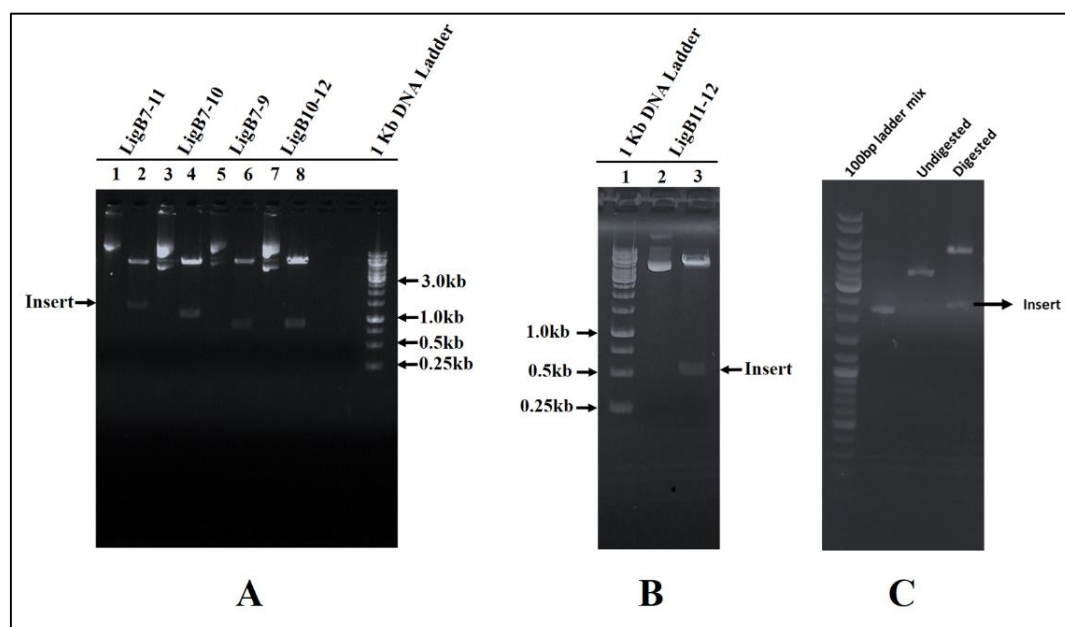


Figure 2.5: 1% Agarose gel showing double digestion of different *ligB* fragments from the expression vector. Lanes are labelled with the fragments of *LigB* and 1kb DNA ladder. Fall-out of the fragments is indicated with insert.

2.4.2. Recombinant purification of LigA fragments

2.4.2.1. Single Ig-like domain

All single Ig-like domain fragments of LigA were purified at homogeneity. Their elution from size-exclusion chromatography and comparison with known molecular weight markers suggested that protein fragments exist as monomers in the solution. The single domains of Lig displayed a protomer mass of ~10,000 Da. The purified protein showed more than 90% purity on the SDS-PAGE (Figure 2.6).

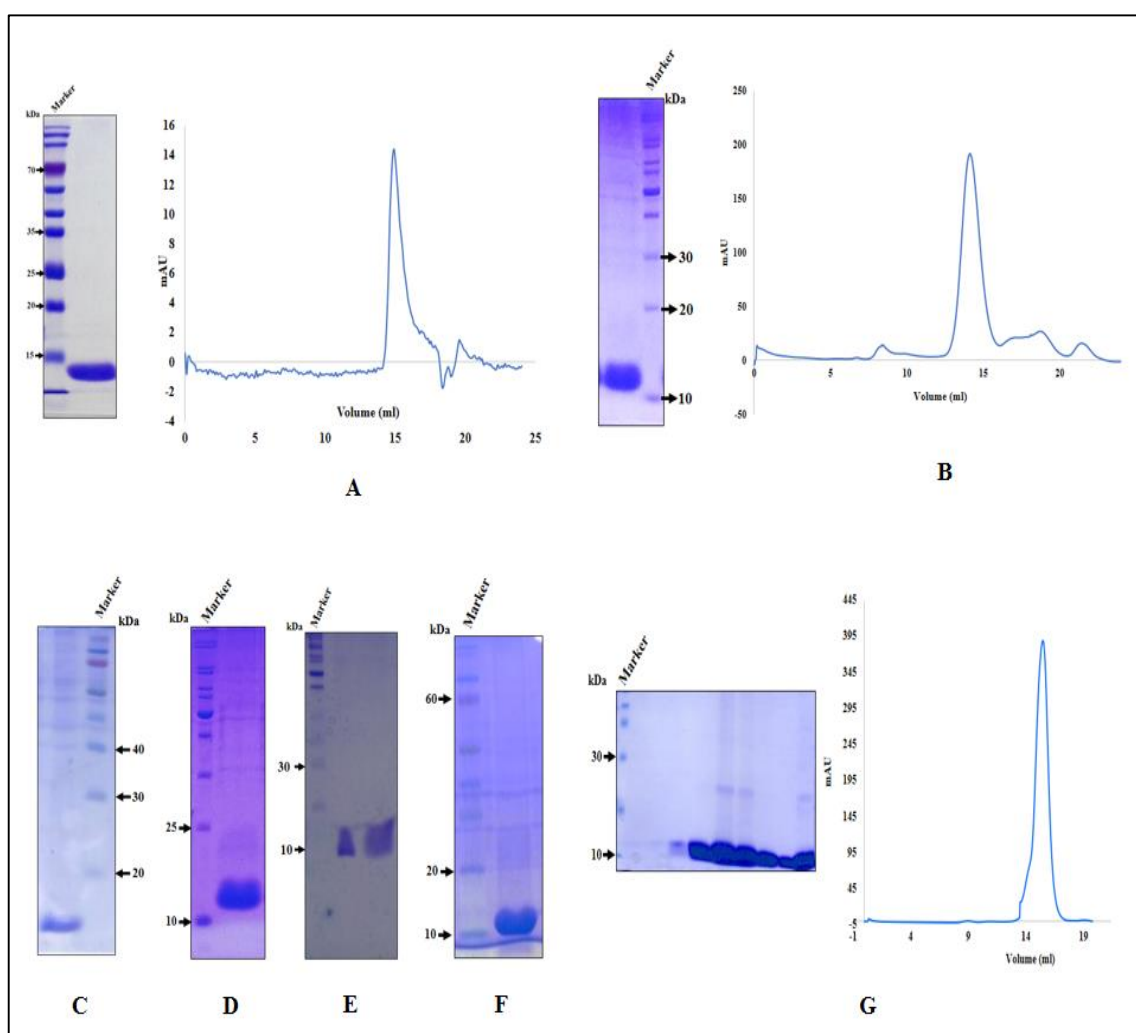


Figure 2.6: SDS-PAGE showing the purification LigA fragments. (A) LigA7 and its size-exclusion chromatogram, (B) LigA8 and its size-exclusion chromatogram, (C) LigA9, (D) LigA10, (E) LigA11, (F) LigA12, and (G) LigA13, and its size-exclusion chromatogram.

2.4.2.2. Multiple Ig-like domains

Various multiple Ig-like domain fragments of LigA were also purified at homogeneity. Their elution from size-exclusion chromatography and comparison with known molecular weight markers suggested that these fragments also exist as monomers in the solution. The purified proteins showed more than 90% purity on the SDS-PAGE (**Figure 2.7**).

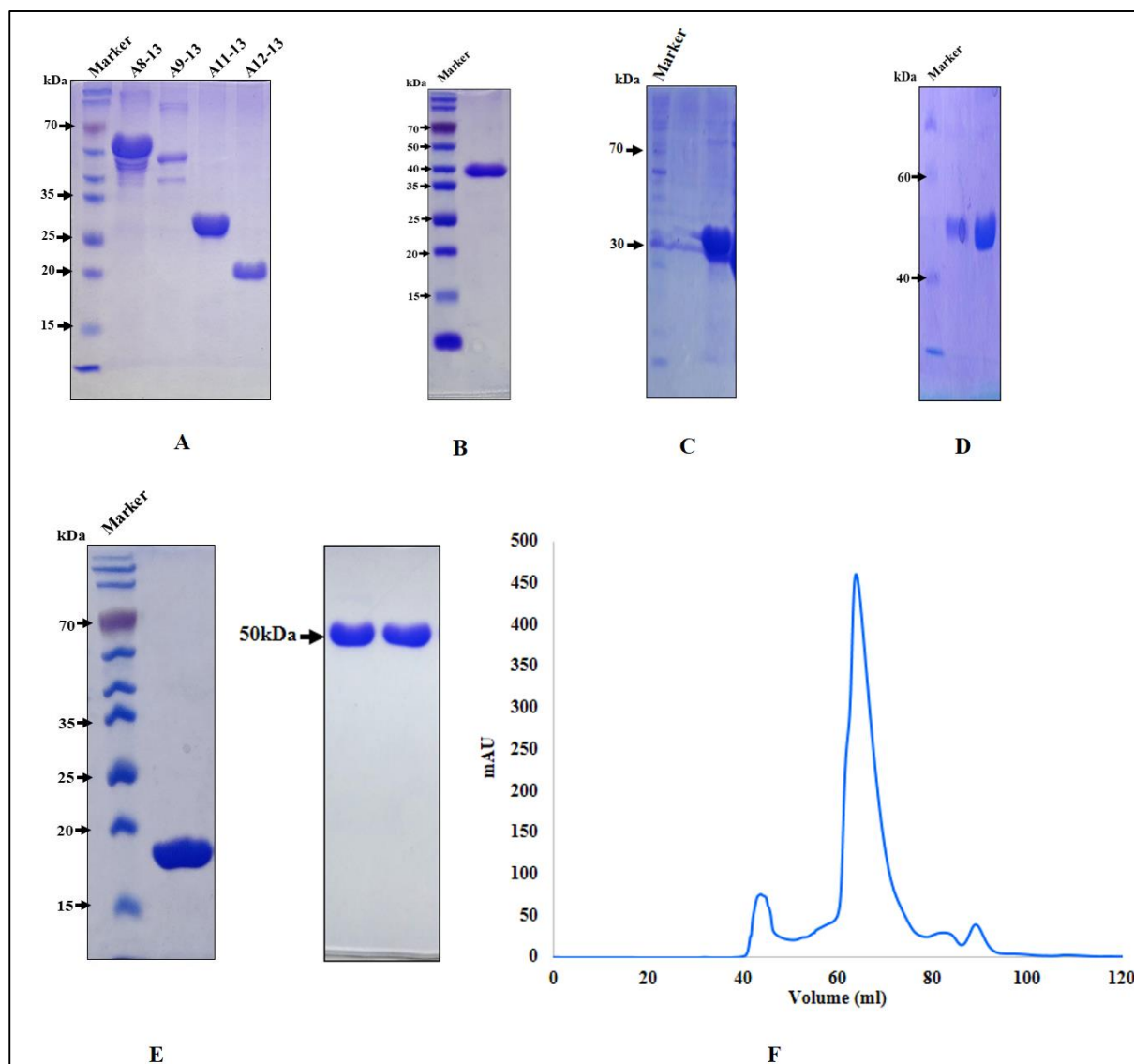


Figure 2.7: SDS-PAGE showing the purification of multiple Ig-like fragments of LigA (A) LigA8-13, LigA9-13, LigA11-13 and LigA12-13, (B) LigA10-13, (C) LigA8-10, (D) LigA8-11, (E) LigA8-9, (F) LigA8-12 and its size-exclusion chromatogram.

2.4.2. Recombinant purification of LigB fragments

Similarly, single Ig-like domains and various multiple Ig-like domain fragments of LigB were also purified (**Figure 2.8**).

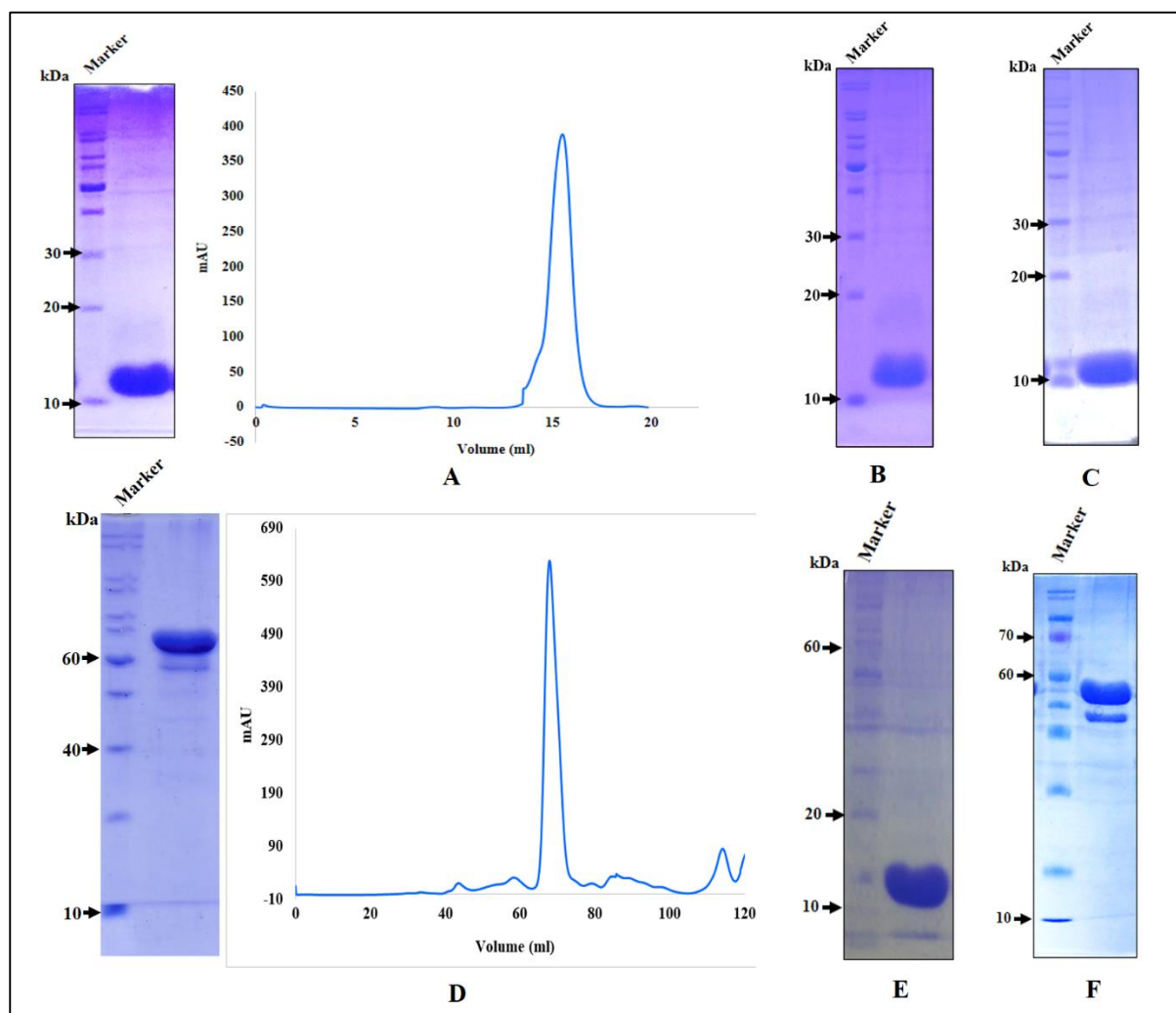


Figure 2.8: SDS-PAGE profile showing the purification of Ig-domains of LigB. (A) LigB2 and its size-exclusion chromatogram, **(B)** LigB4, **(C)** LigB5, **(D)** LigB7-12 and its size-exclusion chromatogram, **(E)** LigB10 **(F)** LigB8-12.

Interestingly, all the purified single and multiple Ig-like domains of LigA and LigB showed very good yield and seemed sufficient for crystallographic studies. The yield obtained after two rounds of Ni-NTA affinity chromatography and size exclusion chromatography of each protein is given in **Table 2.10**.

Table 2.10: Summary of purification yield of the LigA and LigB fragments in this study.

Lig fragments	Yield (mg/ml)	Lig fragments	Yield (mg/ml)
LigA7	14	LigA8-13	10
LigA8	8.5	LigA9-13	9.0
LigA9	5.3	LigA10-13	5.2
LigA10	10.5	LigA11-13	8
LigA11	2.5	LigA12-13	11
LigA12	3.2	LigA8-9	8.2
LigA13	10.2	LigA8-10	2.5
LigB1	5.2	LigA8-11	3.0
LigB2	10.3	LigA8-12	10.5
LigB3	8.5	LigB7-12	5.3
LigB4	9.5	LigB1-7	1.5
LigB5	10	LigB8-12	3.5
LigB6	9.7	LigB10-12	2.5
LigB7	5.4	LigB11-12	9.5
LigB8	1.5	LigB7-11	4.5
LigB9	2.4	LigB7-10	2.3
LigB10	3.5	LigB7-9	5.0
LigB11	8.5	LigB11-NR	2.4
LigB12	2.6	LigB12-NR	5.6

2.5. Conclusions

The Lig proteins consist of multiple Ig-like domains that showed poor expression when complete protein was tried to overexpress. In order to improve the yield and solubility of the protein, various truncations were made in these proteins. All the truncated variants of LigA and LigB were cloned successfully in the pET28a-Sumo vector. The truncated proteins showed good expression and solubility. Various fragments of LigA and LigB were purified to homogeneity using affinity chromatography and size exclusion chromatography using an AKTA purifier. All the domains were purified in good amounts for crystallographic studies. The chapter deals with the preparation of proteins that is further used for setting up crystallization. The protein preparation mentioned in this chapter was also used further to perform novel protease and nuclease activity mentioned in chapter 5 of this thesis.

Chapter - 3

Crystallization and structure determination

This chapter's manuscript is under communication:

Kumar P., Chang YF and Akif M. (2022). Crystal structure of two domains segment from the variable region of *Leptospira* host-interacting outer surface protein, LigA

3.1. Introduction

The outer membrane and surface-associated proteins of microorganisms perform various functions, such as providing protection against harsh environments, signal transduction, and protein/solute translocations (Koebnik et al., 2000). They also promote pathogenicity by enhancing the ability of pathogens to attach, invade the host, and evade the host's immunity (Hughes et al., 2002). The expression of many surface proteins by pathogens such as *Staphylococcus aureus*, *Pneumococcus*, *Streptococcus*, and *Listeria monocytogenes* and their relationships in pathogenesis have been described (Kreikemeyer et al., 2004; Niemann et al., 2004; Pizarro-Cerdá and Cossart, 2006). Pathogenic *Leptospira* expresses a plethora of outer surface proteins, and many of them play a crucial role in pathogenesis. The Outer surface proteins such as LipL32, leptospiral endostatin-like protein (LenA), Loa22, *Leptospira* lipoprotein (LipL53), Lp95, and leptospiral immunoglobulin-like protein (Lig) are known to involve in colonization and pathogenesis (Atzingen et al., 2009, 2008; Angela S. Barbosa et al., 2006; Hoke et al., 2008; Oliveira et al., 2010; Stevenson et al., 2007).

Lig is a surface-anchored multi-Ig-like domain containing protein. The Ig-like domain shares a similar fold as bacterial immunoglobulin superfamily-2 (BIg_2), which is widely distributed in nature, including many bacterial species such as intimin and invasion from *E.coli* and *Yersinia*, respectively (Matsunaga et al., 2003b). The N-terminal 630 amino acids constituting 6 & 1/2 Ig-like domains of LigA and LigB are highly conserved, and the remaining Ig-like domains among the two are unique. The LigB consists of an additional C-terminal domain (CTD) comprising 772 amino acids that form mostly non-structural regions in the protein. The beginning of this region shares homology with the BIg_5 domain (Haake and Matsunaga, 2021). The Lig family proteins serve as multifaceted proteins in complement evasion and binding with ECM components such as fibronectin (Fn), collagen, elastin, tropoelastin, laminin, and fibrinogen (Fg) (Ito and Yanagawa, 1987). These functions are

essential in host cell attachment and invasion during leptospiral pathogenesis. In addition, Lig proteins are highly antigenic and have been extensively studied as vaccine candidates. Recent studies on LigB have shown that chimera of most antigenic regions provides better protection (Hsieh et al., 2017). Moreover, the immunoinformatics approaches have identified most antigenic regions of LigA and LigB proteins. A multi-epitope chimeric construct has been proposed using these antigenic regions, which could serve as an effective vaccine (P. Kumar et al., 2021). In addition to multifaceted function, the Lig proteins are reported to bind with divalent calcium ions, enhancing Lig *in-vitro* binding to the host extracellular matrix (ECM) proteins (Raman et al., 2010). The Lig protein is a vital outer surface protein, but its complete 3-D structure and domain organization are largely unknown. Moreover, the atomic details of interaction with the host proteins have also not been investigated thoroughly. So far, only NMR structures of the single domain, LigB12, and LigA4 from LigB and LigA are available (Mei et al., 2015b; Christopher P Ptak et al., 2014). Moreover, the NMR studies on LigA4 and LigBCen2 (complete LigB12 and additional amino acid stretches from adjacent regions) have demonstrated the binding to the Ca^{2+} ions with the proteins (Lin et al., 2008a). Other *In-vitro* investigations suggest the binding of Ca^{+2} ions with various Ig-like domains of Lig proteins. However, the physiological relevance of Ca^{+2} ion binding with the Lig protein is still debatable. Recent low-resolution small-angle X-ray scattering (SAXS) of LigB has been demonstrated that all 12 Ig-like domains are arranged as an elongated conformation (Hsieh et al., 2017). However, detailed structural studies in the context of LigA protein are still lacking.

In this chapter, we describe the first high-resolution crystal structure of two domains segment from the variable region (LigA8-9) of LigA to 1.87 Å resolution. The structure showed some remarkably distinctive aspects compared to the most closely related IgSF family members. Further, the structure details provide insight into the relative orientation of two domains and highlight the role of the linker region in the domain orientation. We also observed an apparent

electron density of Ca^{+2} ion and modeled Ca^{2+} ions forms a proper interacting geometry within the protein. Overall the study suggests an arrangement of Ig-like domains in LigA protein. Moreover, docking of fibronectin's Gelatin binding domain (GBD) with the crystal structure highlighted complex atomic interactions between the two.

3.2. Materials

Crystallization screens containing a combination of buffers, precipitants, and salts were commercially procured from Jena biosciences (Wizard 1&2, JGSG++, and Wizard 3 & 4), Hampton (crystal screen 1&2), Qiagen (JCSG++), and Molecular dimensions (Morpheus, structure screen 1 &2, PACT Premier and JGSC Plus). Siliconized glass coverslips and micro-bridges were purchased from Hampton research. Additionally, coverslips were also purchased locally and siliconized manually using silicon oil from thermo scientific. Grease was procured locally and from Hampton research. The vapor diffusion crystallization screening sitting drop plates (96 wells, two drops) were purchased from Hampton research. Crystallization optimization was done in 24 well-hanging drop vapor diffusion plates purchased from Hampton research as well as Thermo scientific. All fine chemicals used for preparing the crystallization buffers were purchased from Sigma-Aldrich. All the reagents were prepared in autoclave double distilled water and solutions were filtered with 0.22-micron or 0.45-micron filters.

3.3. Experimental Procedures

3.3.1. Protein purification

The LigA and its various fragments were purified to homogeneity using different chromatographic techniques as explained in the experimental methods of the previous chapter 2, 33-50.

3.3.2. Crystallization Screening

The purified homogeneous recombinant proteins were used to set up crystallization screening with the commercially available screens mentioned above. A few nano liters of purified proteins were blended with the crystallization screens using crystallization robotics in a sitting drop plate followed by sealing the plates with clear tape and incubating it at 22°C and 4°C respectively. The plates were observed weekly under stereo zoom microscope until they dried up. The crystallization hits of each protein were monitored separately and the screened conditions were further optimized in hanging drop plates in order to get crystals suitable for diffraction. The overall experimental approach followed for crystallization and structure determination is shown in **Figure 3.1**.

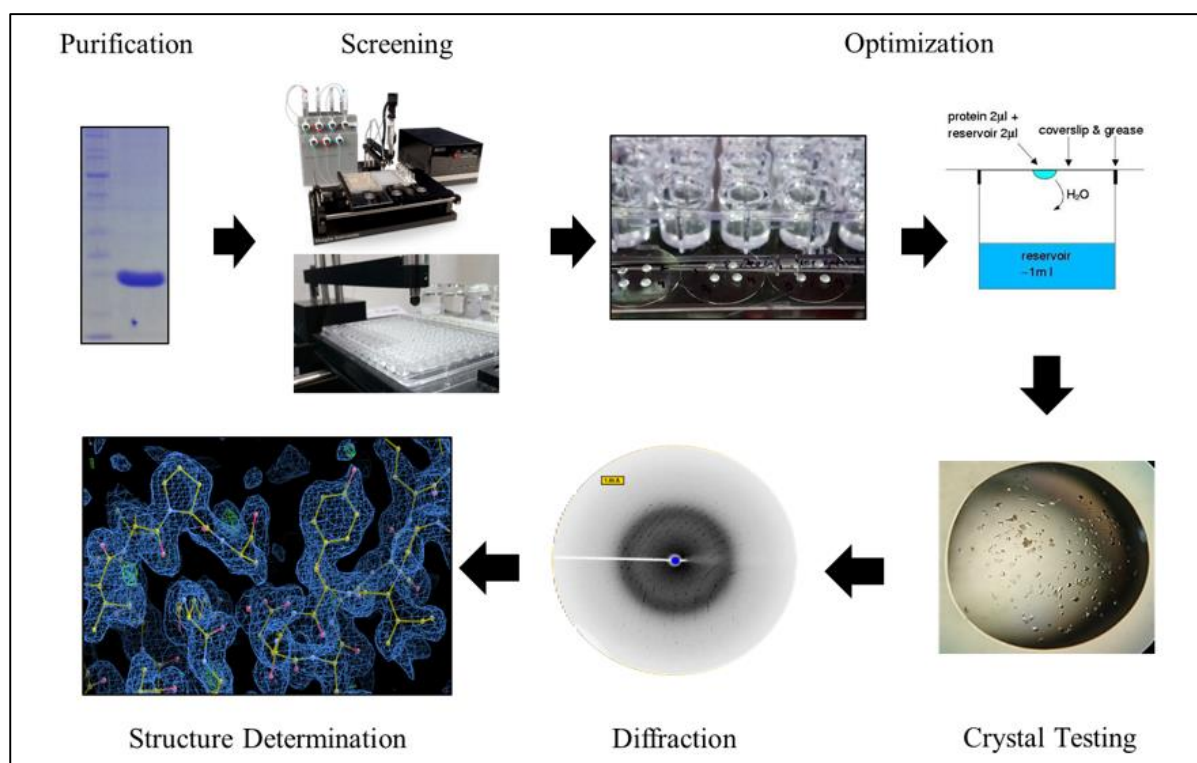


Figure 3.1: The experimental approach followed for protein crystallization and structure determination

3.3.3. Crystallization optimization

The successful crystallization screening hits were further optimized by the vapor diffusion hanging drop method in a 24-well plate. The concentration of all the components of the hits was varied one-by-one in 24-well plates. In addition, various crystallization seeding techniques such as macro seeding, micro seeding, and streak seeding were also used to improve the quality of crystals.

3.3.4. X-ray data collection and processing

The diffraction quality crystals were obtained for various LigA and LigB fragments in different crystallization conditions. The single crystal was tested *in-house* diffraction facility at the Centre for Cellular and Molecular Biology (CCMB), Hyderabad, and the Indian Institute of Science (IISc) Bangalore. Briefly, a single crystal was fished out from the coverslip with the help of a suitable loop. Crystals were aligned manually in an X-ray beam to test their diffraction. Only a single crystal of LigA8-9 diffracted at 1.8Å. A total of 360 images of LigA8-9 were collected using the Mar345dtb imaging plate detector. No Cryo-protectant was used to keep the crystal at a constant temperature (100K) under the liquid nitrogen jet during data collection.

3.3.5. Structure determination and model building

Initial phasing for structure solution was obtained using the molecular replacement routines of the program PHASER (McCoy, 2006). The atomic coordinates of the single terminal domain, LigB 12 (PDB code 2MOG), were used as a search model. The resultant model was refined using REFMAC5 and the model was adjusted using COOT (Emsley and Cowtan, 2004; Murshudov et al., 2011). Water molecules were added at positions where $F_o - F_c$ electron density peaks exceeded 3σ contour level, and potential H-bonds could be made. Based on

electron density interpretation, divalent ions were added, and further refinement was carried out. Figures were drawn with PyMOL (DeLano Scientific, San Carlos, CA, USA).

3.3.6. Validation and deposition of LigA8-9 structure

The final structure model of LigA8-9 is validated for the quality using PROCHECK, WHAT IF, and PDBRedo (Joosten et al., 2014; Laskowski et al., 1993). Protein structural properties such as bond angle, length, chirality, and omega angle are validated. Moreover, the quality of the structure was also analyzed by the Ramachandran plot. The suggestions on rotameric states, packing quality, and backbone conformation were fixed manually in Coot.

3.3.7. Molecular docking with the ECM

The structure of the GBD domain of fibronectin, an extracellular matrix (ECM), was retrieved from the protein database (PDB ID 1E8B). The structural coordinates of LigA8-9 and gelatin binding domains (GBD) of fibronectin were subjected to the protein-protein docking server ClusPro (<https://cluspro.bu.edu/home.php>). The best-docked structure was selected based on the binding energy of interaction, and interaction was analyzed among the two structures.

3.4. Results

3.4.1. Crystallization screening

The crystallization screening leads to the identification of many conditions in which crystals were observed for single and multiple domain fragments of LigA and LigB proteins. The observed crystallized droplets showed various morphologies including clear droplets, phase separations, precipitates, and crystallites. Multiple crystallization conditions such as A1, D4, E4, and H7 of JCSG++, and F11 of Wizard 1&2, yielded hits for the LigA8-13 protein fragment. Different shapes and morphologies of crystals were observed in different drops as shown in **Figure 3.2**.

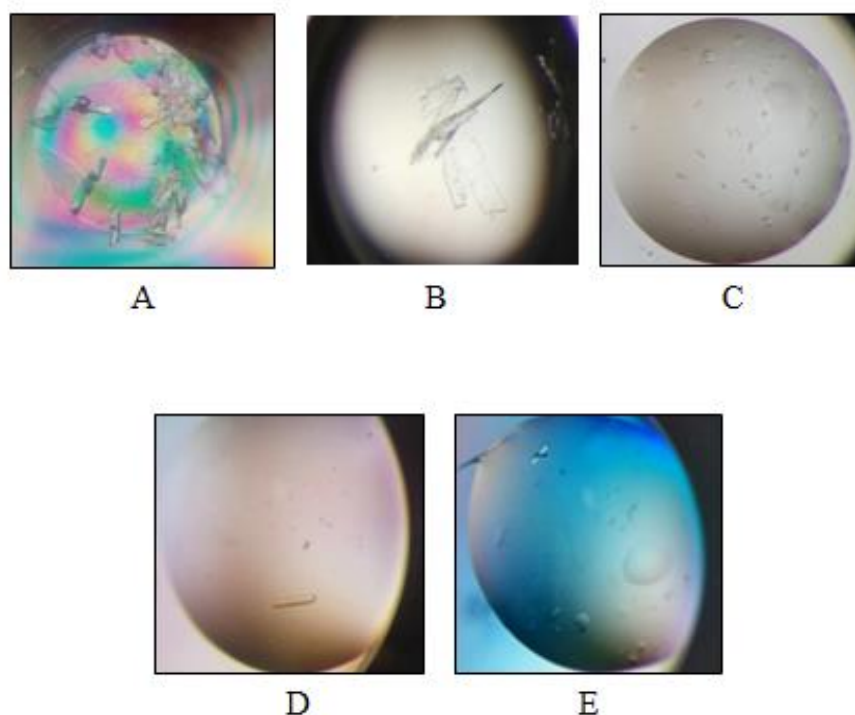


Figure 3.2: Crystallization condition obtained for LigA8-13 fragment (A) JCSG++, E4 (B) Wizard 1&2, F11 (C) JCSG++, A1 (D) JCSG++, D4 (E) JCSG++, H7. The name of the commercial screens and their specific condition number is indicated as the alphabet followed by numerals.

Similarly, various crystallization hits were obtained for the LigB7-12 fragment of LigB. The screening conditions such as A11, B8, and F11 of the JCSG++ and F11 of Wizard 1&2 and E1 of Wizard 3&4 yielded some needles and plate-like crystals in various crystallization drops (**Figure 3.3**).

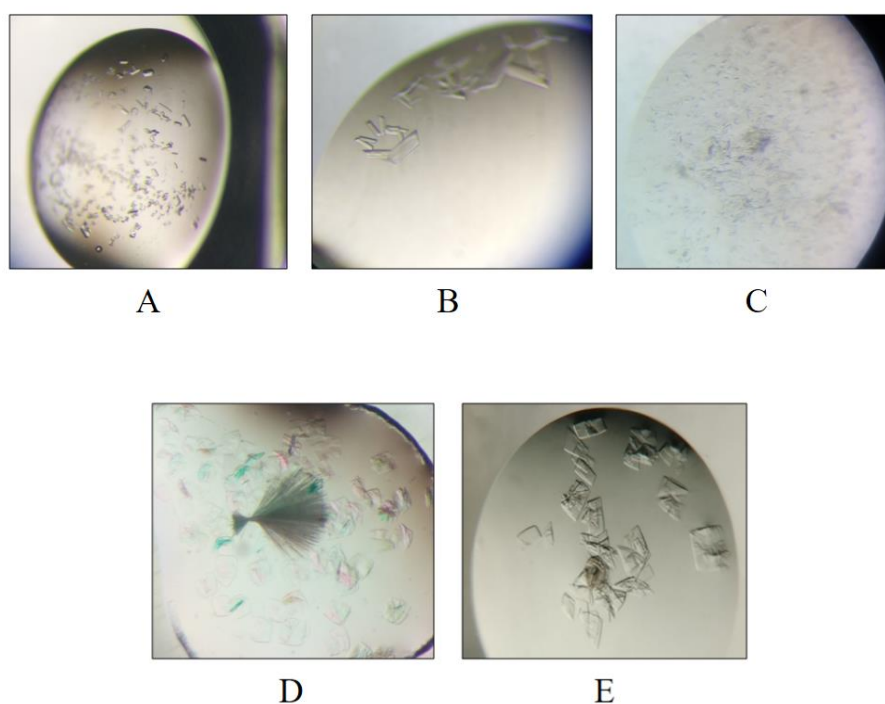


Figure 3.3: Crystallization condition obtained for LigB7-12 fragment of LigB (A) Wizard 1&2, F11 (B) JCSG++, A11 (2) (C) JCSG++, F11 (D) Wizard 3&4, E1 (E) JCSG++, B8. The name of the commercial screens and their specific condition number is indicated as the alphabet followed by numerals.

Few conditions such as D10 and E3 of Wizard 1&2 and E3 of Wizard 3&4 were obtained for the protein-containing six domains LigA8-13 fragment of LigA. The observed crystals hits are shown in **Figure 3.4**.

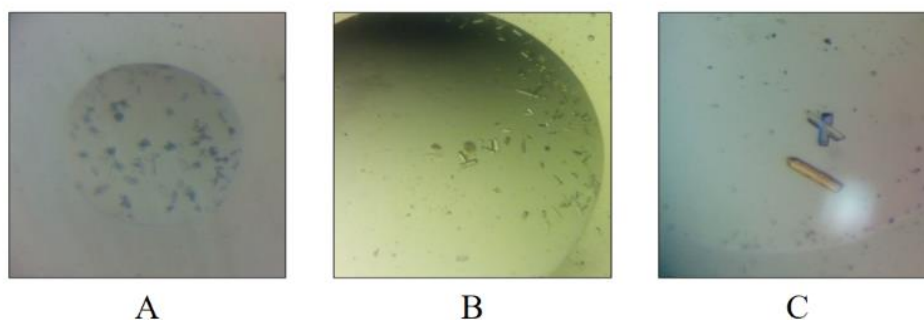


Figure 3.4: Crystallization condition obtained for Ig-like domains LigA8-13. (A) Wizard 1&2, D10 (B) Wizard 3&4, E3 (2) (C) Wizard 1&2, E3. The name of the commercial screens and their specific condition number is indicated as the alphabet followed by numerals.

Three and four Ig-like domains containing LigA such as, Lig11-13 and LigA10-13 fragments were, also observed to show needle-shaped crystals in conditions H11 of Wizard 1&2 and A8 of PACT Premier, respectively (**Figure 3.5**).

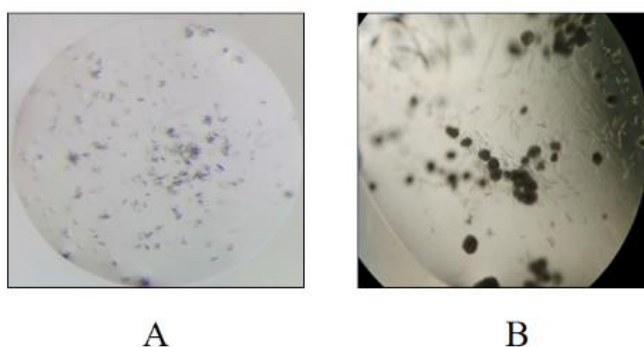


Figure 3.5: Crystallization condition obtained for Ig-like domains. (A) LigA10-13, PACT Premier, A8 (B) LigA11-13, Wizard 1&2, H11. The name of the commercial screens and their specific condition number is indicated as the alphabet followed by numerals.

The protein fragments containing two domains such as LigA8-9, LigA12-13, and LigB11-12 also yielded a few crystallization conditions in the screening conditions. The LigA8-9 gave crystallization hits in conditions such as C11, E4, E11, and F11 of Wizard 3&4 and G3 of PACT Premier (**Figure 3.6**).

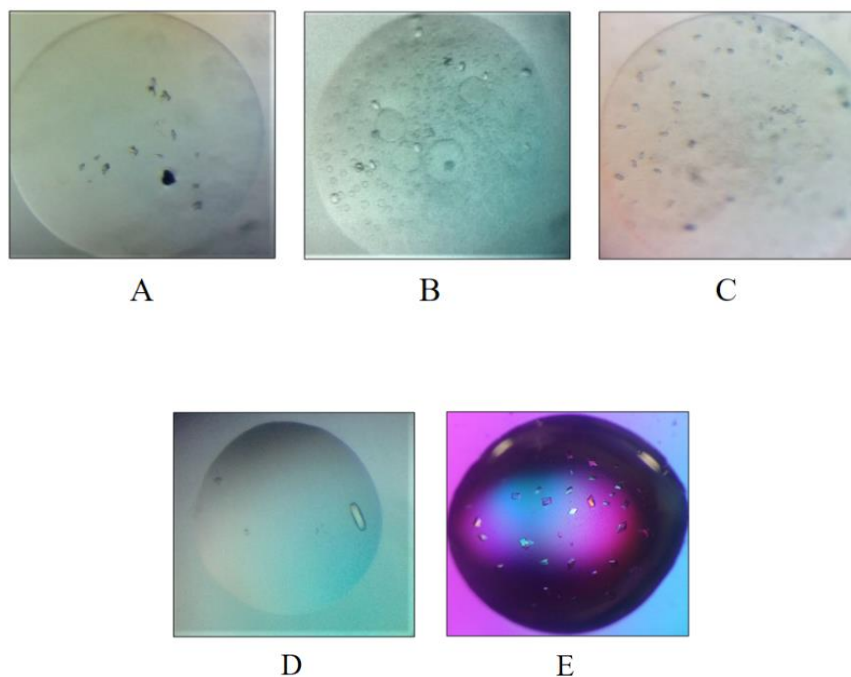


Figure 3.6: Crystallization condition obtained for LigA8-9 fragment (A) Wizard 3&4, C11 (B) Wizard 3&4, E4 (2) (C) Wizard 3&4, E11 (D) Wizard 3&4, F11 (E) PACT Premier, G3. The name of the commercial screens and their specific condition number is indicated as the alphabet followed by numerals.

Maximum numbers of crystallization hits were obtained for LigA12-13 fragment. Conditions such as D2 of JCSG++, A2, A9 of Morpheus, D10 of PACT Premier, B12, C9, E2 and E4 of Structure screens 1&2 yielded few needle-shaped crystals for LigA12-13 (**Figure 3.7**).

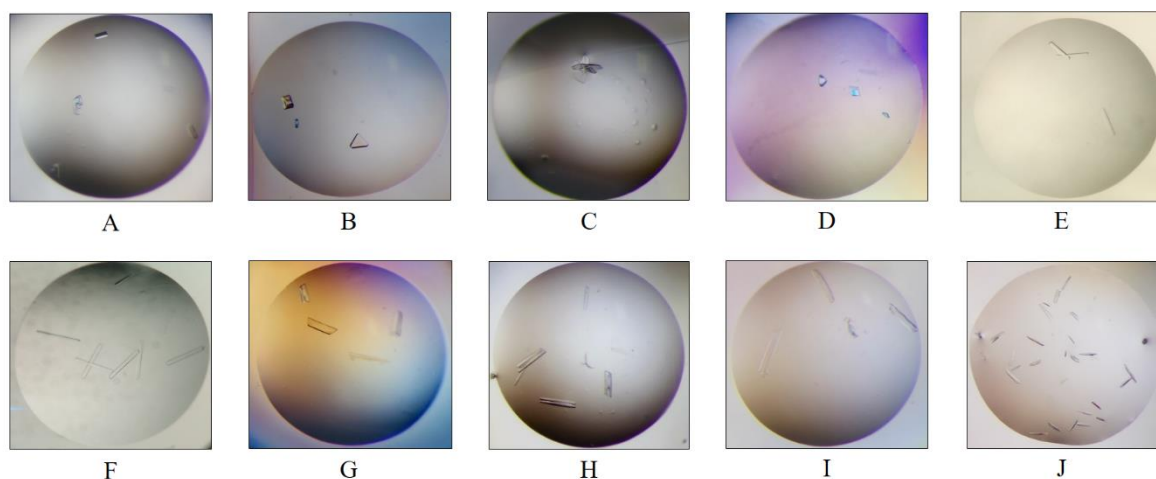


Figure 3.7: Crystallization condition obtained for LigA12-13 fragment (A) JCSG++, D2 (B) JCSG++, D2 (C) Morpheus, A2 (D) Morpheus, A9 (E) PACT Premier, D10 (F) PACT Premier, D10 (G) Structure screen, B12 (H) Structure screen, C9 (I) Structure screen, E2 (J) Structure screen E4. The name of the commercial screens and their specific condition number is indicated as the alphabet followed by numerals.



Figure 3.8: Crystallization condition obtained for LigB11-12 fragment in B11 of the PACT Premier screen.

Few conditions for single Ig-like domains, LigA9 and Lig13, were also obtained. The conditions D7, and D11 of JCSG++ gave crystallization hits for LigA9 (**Figure 3.9**).

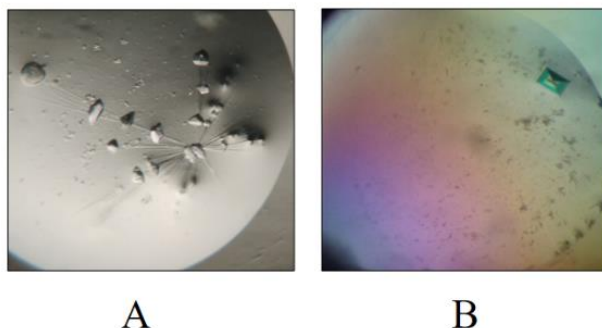


Figure 3.9: Crystallization condition obtained for LigA9 single Ig-like domain of LigA (A) JCSG++, D7 (B) JCSG++, H11. The name of the commercial screens and their specific condition number is indicated as the alphabet followed by numerals.

Similarly, single Ig-like domain LigA13 was also observed to give crystallization hits in the conditions such as F9 of JCSG++ and C9, E4, and F12 of Wizard 1&2 (**Figure 3.10**).

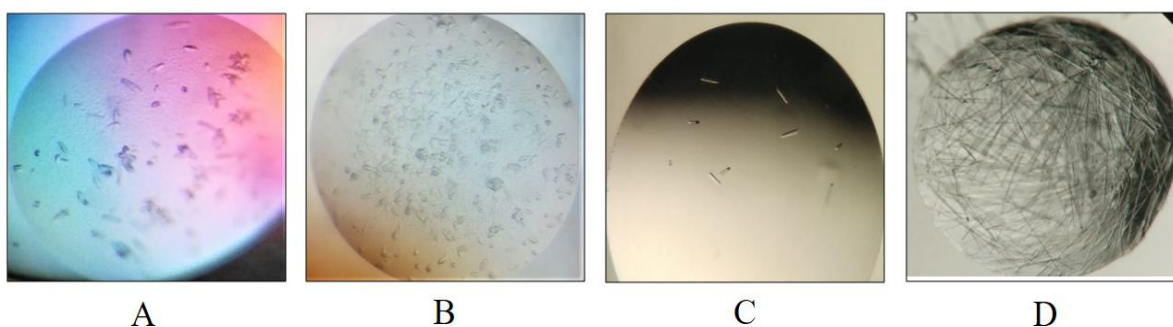


Figure 3.10: Crystallization condition obtained for Lig A13 fragment (A) JCSG++, F9 (B) Wizard 1&2, F12 (2) (C) Wizard 1&2, E4 (D) Wizard 1&2, C9. The name of the commercial screens and their specific condition number is indicated as the alphabet followed by numerals.

3.4.2. Crystallization optimization:

The crystallization conditions thus obtained for various truncated fragments LigA and LigB were further optimized manually in 24 well plates by vapour diffusion hanging drop method. The techniques such as micro and macro seeding were also employed to improve the quality of the crystals. After optimization, the diffraction quality of crystals were obtained for a few truncated fragments such as LigA12-13, LigB7-12, LigA8-13 and, LigA8-9 (**Figure 3.11**).

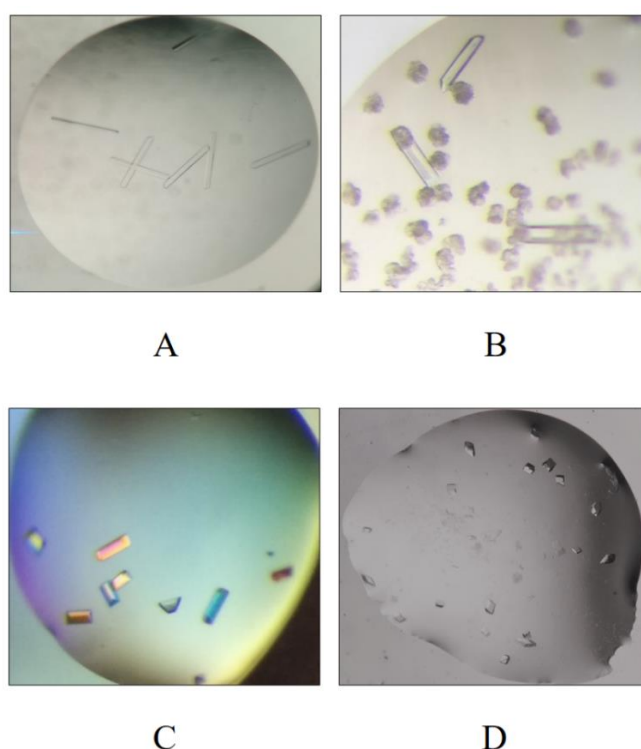


Figure 3.11: Crystals obtained after optimization (A) LigA12-13 (B) LigB7-12 (C) LigA8-13 and (D) LigA8-9

3.4.3. X-ray data collection

Single crystals of LigB7-12, LigA8-13, LigA8-9, and LigA12-13 obtained in various conditions were exposed to an *in-house* X-ray source for diffraction. Unfortunately, the crystals of LigB7-12, LigA8-13, and LigA12-13 either tuned out to be salt crystals or diffracted very poorly. Hence, data was not collected. A poor diffraction pattern was observed for LigB7-12

and LigA8-13 crystals (**Figure 3.12**). This could be related to crystal packing issues during the crystallization.

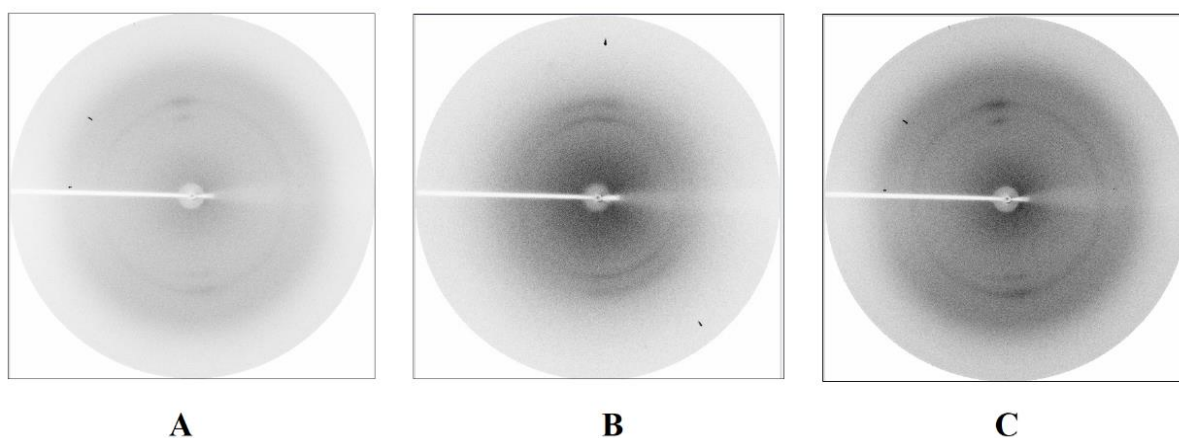


Figure 3.12: X-Ray diffraction pattern observed. (A&B) LigB7-12, and (C) LigA8-13

No improvement in the diffraction was observed even after many steps of crystallization condition optimization of LigA8-13 and LigB7-12 crystals.

The crystals obtained for two Ig-like domains LigA8-9 showed a better diffraction pattern. Two high-resolution data sets were collected from these crystals at resolutions of 2.3 Å and 2.1 Å. However, these two data sets were observed to be twinned and were difficult to deduce the initial phase. The L-test of the collected data also suggested the twinned in the data sets. The crystallization conditions were further optimized to grow better crystals of LigA8-9. Fortunately, the diffraction quality crystals were obtained in a crystallization condition containing 23% PEG3350, 260mM potassium iodide, and 100mM Bis-Tris Propane pH 7.5 and we were able to collect a higher resolution (1.8 Å) data set this time. The diffraction pattern of the crystal LigA8-9 observed at a resolution of 2.3 Å and 1.86 Å is shown in **Figure 3.13** and **Figure 3.14**, respectively.

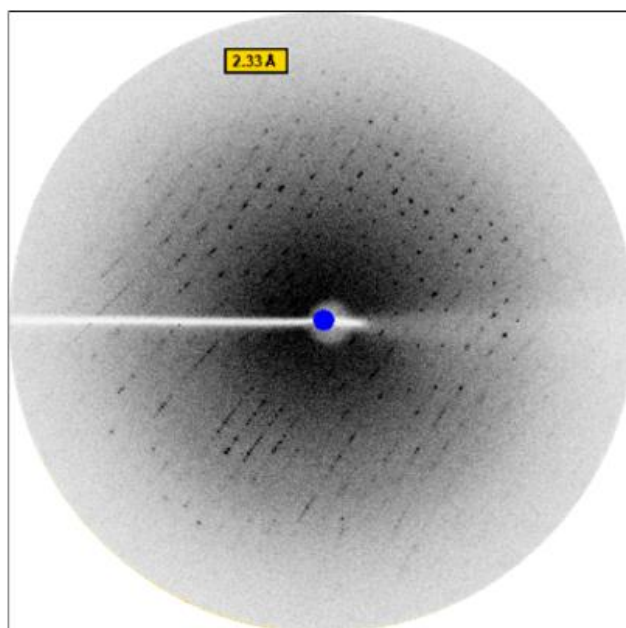


Figure 3.13: Diffraction pattern of one of the LigA8-9 crystals covering 1° oscillation range and diffracted up to 2.33 Å.

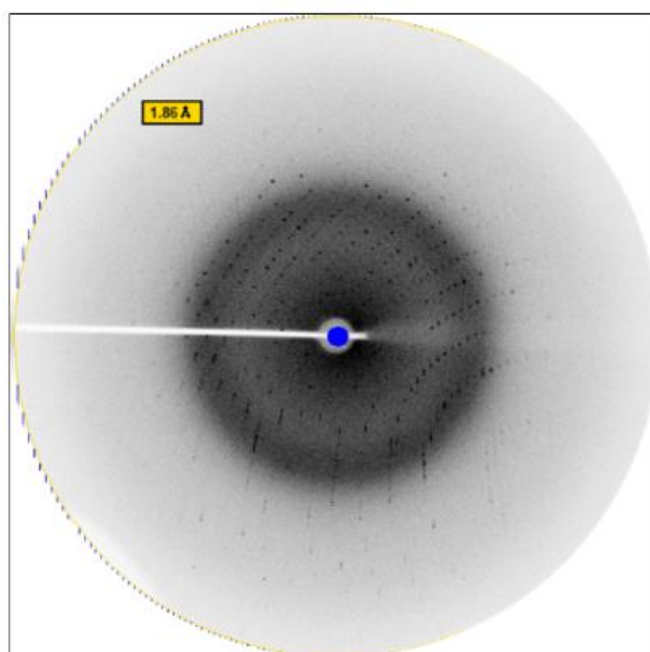


Figure 3.14: Diffraction pattern of one of the LigA8-9 crystals. Covering 1° oscillation and diffracted up to 1.86Å.

3.4.4. Data Processing of LigA8-9

The initial processing statistics are given in **Table 3.1**. The diffraction data were indexed into the space group $P2_12_12_1$. The unit parameters includes $a=34.219$, $b= 63.905$, $c= 171.839$ and $\alpha=\beta=\gamma=90^\circ$. The R_{merge} of the data was calculated to be 11.3 and Matthew's coefficient calculated two molecules in an asymmetric unit. The completeness of the data was calculated to be 96.5%. The total number of unique reflections collected was calculated to be 31943.

Table 3.1: Data collection and processing statistics

Data Collection & Processing Stats	
Diffraction source	CCMB Hyderabad
Detector	Mar345
Wavelength (Å)	1.5417
Space group	$P2_12_12_1$
Unit cell Parameters (Å)	$a=34.219$, $b= 63.905$, $c= 171.839$ $\alpha=\beta=\gamma=90^\circ$
Resolution range (Å)	85.9-1.87
No. of total reflections	213687 (8909)
Number unique reflections	31943(1705)
Mean (I)/ σ (I)	9.1 (2.9)
R_{merge}^I (%)	11.3 (34.9)
Completeness (%)	96.5
Probable Solvent content (%)	49.8
No. of molecules in asymmetric unit	2
Multiplicity	5.5

Values in parentheses refer to the last resolution shell (1.91-1.87Å)

I- $R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $I_i(hkl)$ are intensities of symmetry redundant reflections and $\langle I(hkl) \rangle$ is the average intensities over all observations

The possibility of twinning of data was ruled out by analyzing the L-test of the data set that was collected up to 1.86 Å. The test showed the cumulative distribution function for perfect twinned and untwinned data. The expected cumulative distribution function was observed towards untwinned data, suggesting that the collected data set may not be twinned (**Figure 3.15**).

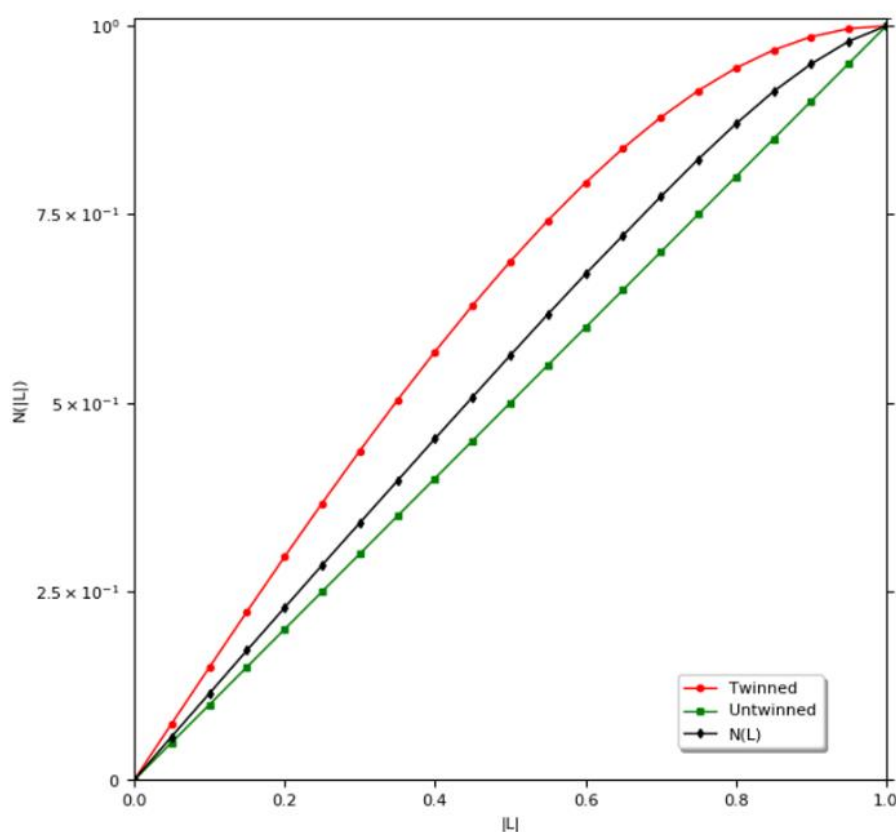


Figure 3.15: Cumulative distribution function of the centric and acentric reflections observed in the LigA8-9 data. Acentric reflection, perfect twin (—•—), centric, untwinned (—•—), and observed acentric, untwinned (—•—).

Wilson plot analysis showed a typical distribution of intensities of the observed data is the same as the reference distribution of the characteristic protein (**Figure 3.16**). Probable solvent content was estimated to be 49.8%. The crystals of LigA8-9 contain two molecules in an asymmetric unit.

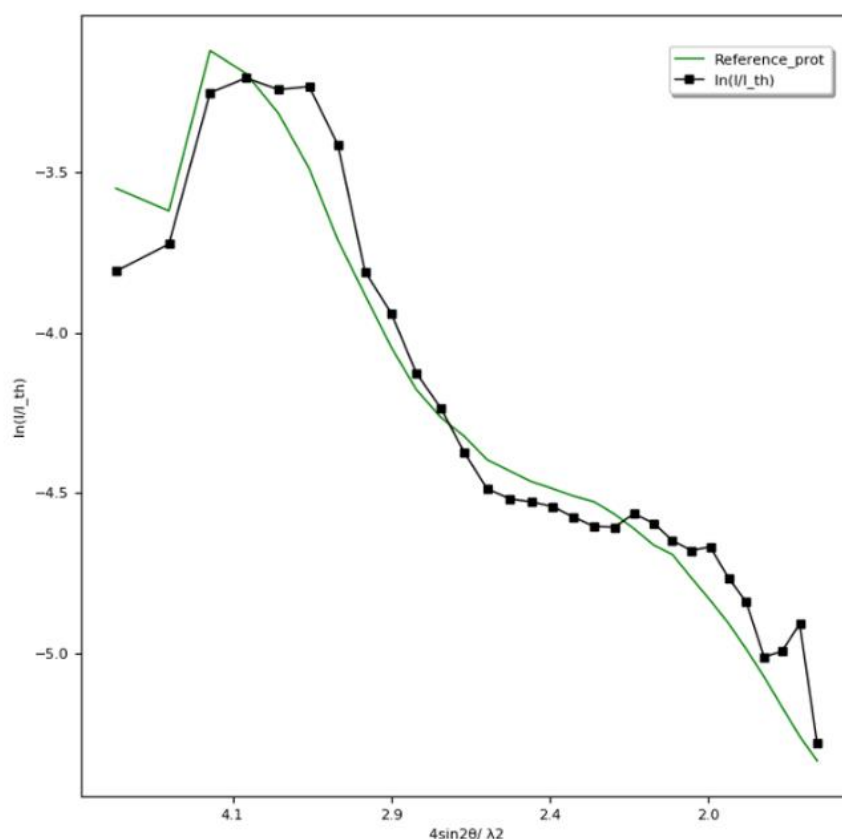


Figure 3.16: Wilson plot for LigA8-9 data extending to 1.86Å. The fall-off intensity is expected for protein crystals and is similar to the reference protein.

3.4.5. Structure solution and refinement of LigA8-9 data

A single crystal of LigA8-9 fragment (residues 672-851) consisting of two Ig-like domains of LigA (**Figure 3.17**), was diffracted and diffraction data were indexed into the space group $P2_12_12_1$. The initial phase was calculated by using the solution structure of LigB12 (2MOG) as a template. The final structure of LigA8-9 has been resolved up to 1.8 Å resolutions, with the R_{work} and R_{free} 21% and 24%, respectively (**Table 3.2**). Two molecules of LigA8-9 were observed in the asymmetric unit and named chains A and B. Almost all the residues were observed in the electron density. The final model consists of a total of 185 residues in each monomer.

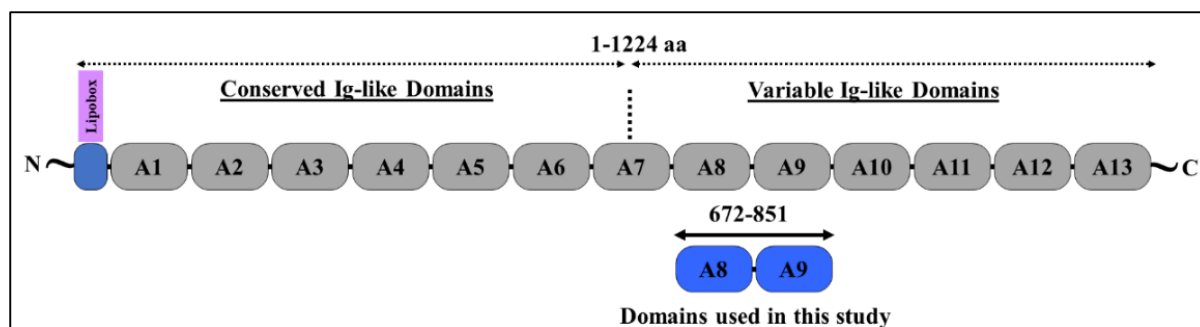


Figure 3.17: A schematic picture of LigA protein domains organization, indicating constant and variable domains. A total number of residues are indicated. Fragment of two domains (LigA8-9) of LigA is also represented.

Table 3.2: Model building and refinement statistics

Refinement Statistics	
Resolution range (Å)	85.9-1.85
Working R-factor (%)	21
Final $R_{\text{free}}^{\#}$ (%)	24
Rmsds	
Bond lengths (Å°)	0.019
Bond angles (°)	1.42
Average B values	
Main chain (Subunit A)	18.36
Side chain (Subunit A)	21.46
Main chain (Subunit B)	19.68
Side chain (Subunit B)	23.59
Ions	36.12
Ligands	25.35
Water	25.2
Number of Ions	33
Number of ligands	3
Number of water	247

$\#$ Five percent of reflections have been chosen as R_{free} set.

The refined model of LigA8-9 was subjected to structure validation to assess its quality. The final model showed an overall good geometry as assessed by Ramachandran plot analysis (**Figure 3.18**). The final structure comprised 99.4% of amino acids in the allowed regions and only 0.6% of amino acids in the disallowed regions. Overall statistics of the final structure from the PROCHECK are given in **Table 3.3**.

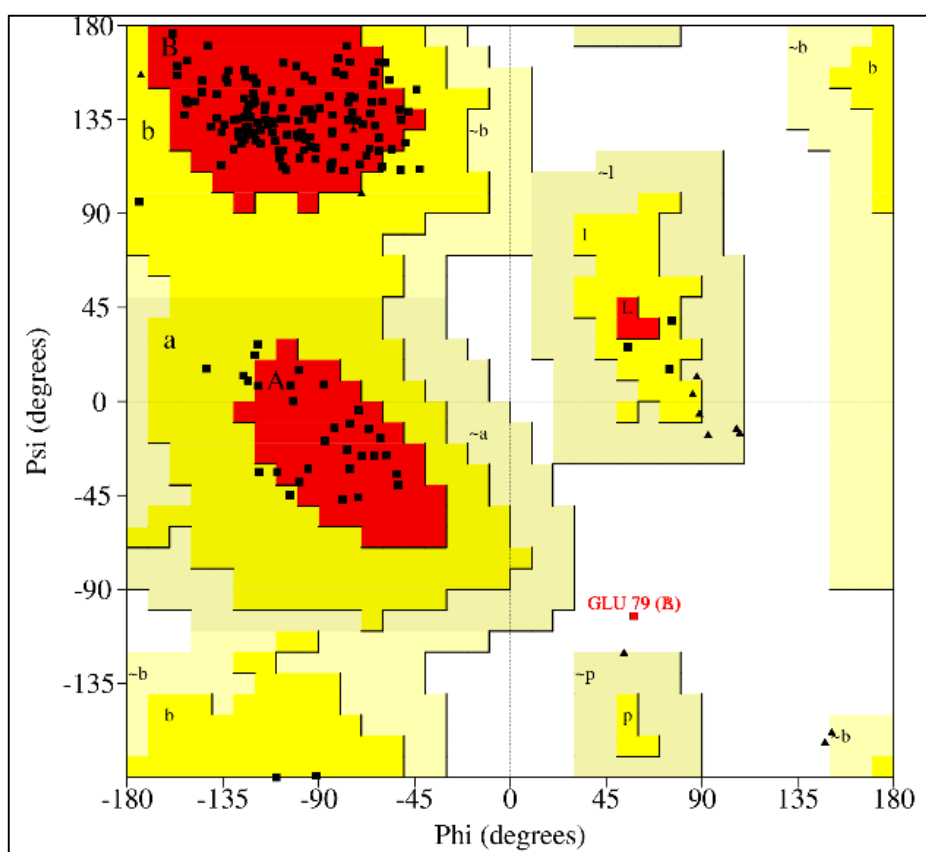


Figure 3.18: Ramachandran plot for the final structure of LigA8-9 fragment. All the residues are within allowed regions. The residue Glu79 of the B-chain is the outlier and lies in the high-temperature factor region of the structure.

Table 3.3: Summary of PROCHECK results of the final refined structure of the LigA8-9 fragment

Property	Statistics
<u>Ramachandran Plot</u>	
Core region	88.4% (290)
Additional allowed	11% (36)
Generously allowed	0%
Disallowed	0.6% (2)
<u>Residue Properties (deviation of bond length and bond angle)</u>	
Main chain bond length	99% within limits
Main chain bond angle	99.2% within limits
Planer group	97.4% within limits
<u>G- factors</u>	
Dihedrals	-0.47
Covalent	-0.13
Overall	0.32

3.4.6. Overall structure

The overall structure of the LigA8-9 fragment (residues **672-851**) consists of two Ig-like domains of LigA, and is composed of β -sheets similar to other Ig-like domains. The structural architecture of the single subunit of LigA8-9 consists of antiparallel β -strands (A-G) that form a classical Ig-like fold. Ig-like fold includes a β -sandwich-like structure based on a Greek key folding arrangement. The topology diagram reveals that each domain possesses almost ten strands arranged into two sheets each, in which six antiparallel β -strands are distributed in two sub-sheets, Ia and Ib, of sheet I, and the 2nd sheet is formed by four strands (**Figure 3.19**).

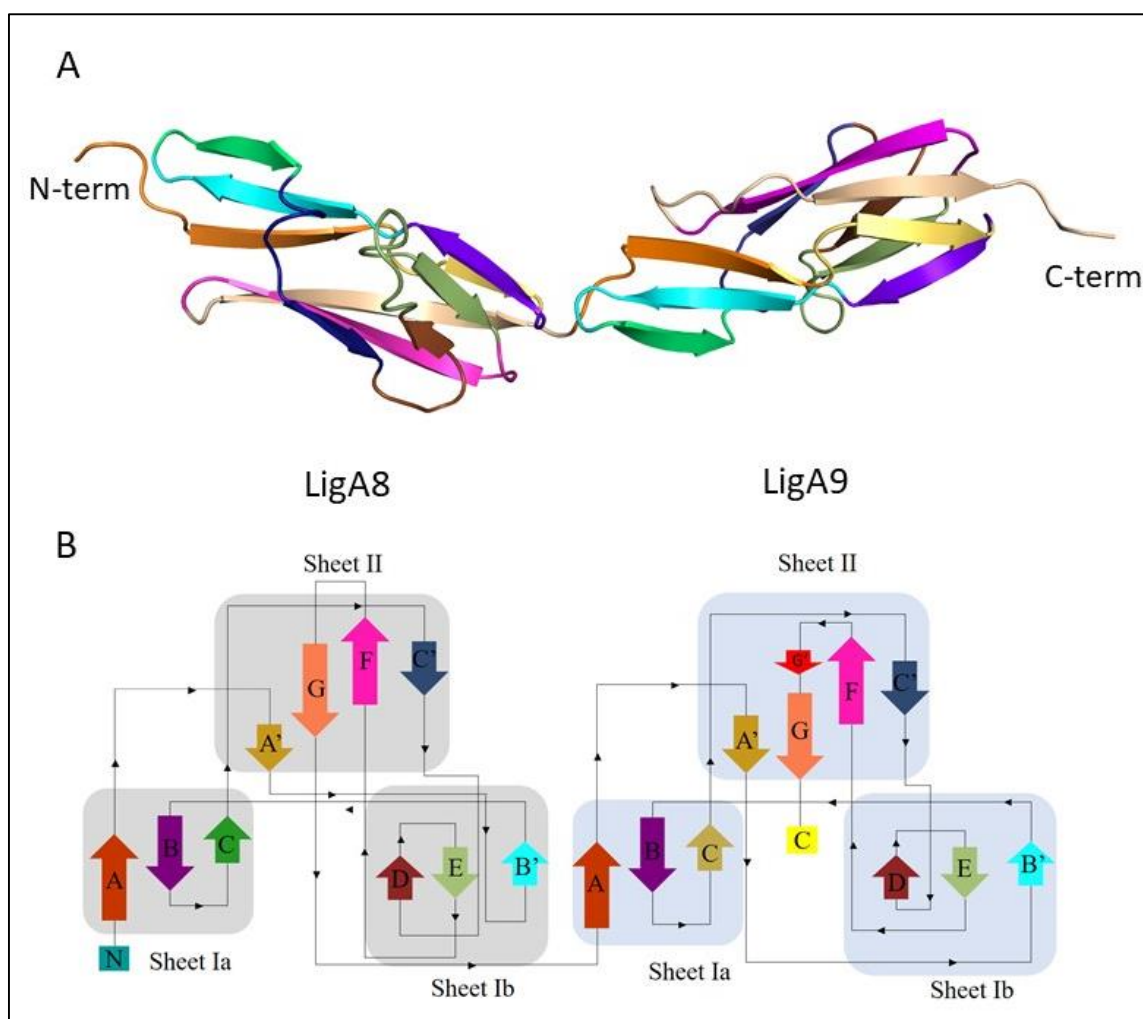


Figure 3.19: The crystal structure LigA8-9 subunit. (A) The overall structure of Ig-like domains of LigA8-9. N-and C-termini are indicated. (B) A schematic secondary structure topology of LigA8-9. The total β -strands of each Ig-like domain are arranged in two sheets. The β -strands in topology and structure are depicted with the same color.

Sheet Ia of LigA8 is composed of β -strands, A, B, and C, and sheet Ib is composed of B', D, and E β -strands. A similar number of β -strands were also observed among sheets 1a and Ib of LigA9. The sheet II of LigA8 was comprised of four β -strands, A', C', F, and G, while the sheet II of LigA9 was composed of five β -strands A', C', F, G, and G' (**Figure 3.19 B**). The number of β -strands was mainly conserved among the two. However, a slight variation of β -strand's size was observed. These β -strands among the sheets are stabilized by hydrogen bond

interactions. Unlike typical Ig fold, the Ig domains of LigA8-9 did not possess any disulphide bond. The amino acid sequence of LigA8 and LigA9 contains only one cysteine residue.

3.4.7. Conserved Hydrophobic core

Two layers of β -sheets close around a hydrophobic side-chain core. The core region is conserved among the Ig-like folds. The core region is generally composed of alternating buried hydrophobic and surface-exposed hydrophilic residues. The core of the LigA8 structure is lined up with F693, T708, V711, W713, L720, I740, I742, and T753. The hydrophobic residues F782, T797, V800, W802, I812, I832, and L846 are organized in the core of the LigA9 structure (Figure 3.20 A&B).

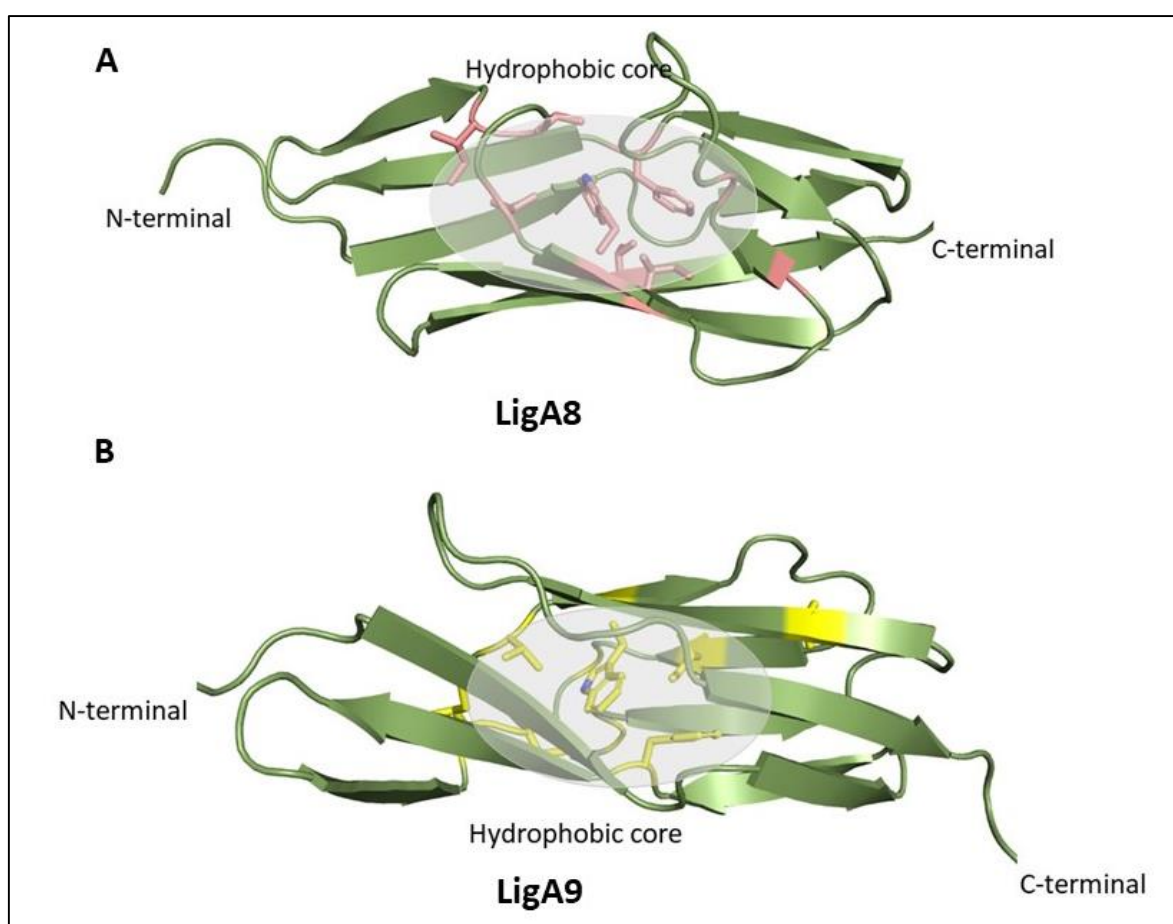


Figure 3.20: Hydrophobic core in Ig-like domain. (A) LigA8 structure showing the core residues in the tick model. (B) LigA9 structure showing the core residues in the stick model.

3.4.8. Dimerization of LigA8-9

The crystal structure of the LigA8-9 fragment consisted of two sub-units named as chain A and chain B. The two chains of LigA8-9 were in contact through the LigA8 domain. (**Figure 3.21**). Moreover, the dimer interface of the two chains was performed and it was found that the interface covered an area of 1236Å². The physiological role of dimerization of this fragment is not known. The LigA8-9 fragment existed as a monomer in solution in our experimental condition. The dimerized form of the LigA8-9 structure may be a crystallization artifact, which might have been required for crystal packing. The residues present on the interface of the dimer and are involved in the formation of Hydrogen bond between the two monomers are shown in **Figure 3.21**.

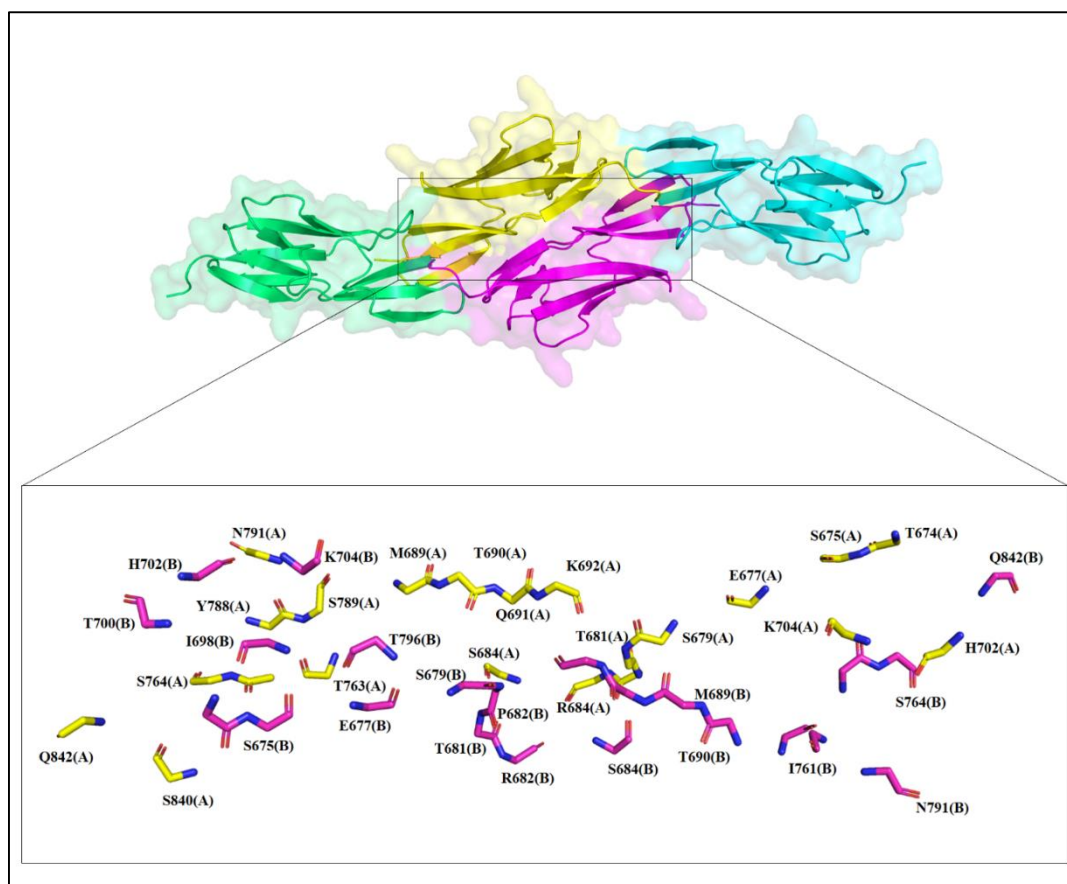


Figure 3.21: Structure of LigA8-9 dimer. Dimerization is mediated by the LigA8 domain. The two chains of LigA8-9 are shown as yellow (LigA8_chainA), Cyan (LigA9_ChainA) and magenta (LigA8_chainB), lime green (LigA9_ChainB) cartoon, respectively. The residues present at the dimer interphase is represented in stick model as yellow and magenta from Chain A and Chain B, respectively.

The structure analyzed by PDBePISA showed and compared the contribution of each monomer to dimer formation (**Table 3.4**). The residues of each monomer were approximately the same at the dimerization interface (42 and 40). About 10% of all atoms were involved in dimer formation. Dimer formation tends to increase the solvation energy of each monomer by about 2%, classifying the dimer as stable.

Table 3.4: Interface Summary analyzed by PDBePISA server

	Structure 1		Structure 2	
Selection range	B		A	
class	Protein		Protein	
symmetry operation	x, y, z		x, y, z	
symmetry ID	1_555		0_555	
Number of atoms				
interface	129	9.6%	128	9.5%
surface	829	61.7%	828	61.7%
total	1343	100.0%	1343	100.0%
Number of residues				
interface	42	22.7%	40	21.6%
surface	176	95.1%	174	94.1%
total	185	100.0%	185	100.0%
Solvent-accessible area, Å				
interface	1236.1	11.6%	1224.5	11.5%
total	10664.6	100.0%	10675.3	100.0%
Solvation energy, kcal/mol				
isolated structure	-158.8	100.0%	-159.1	100.0%
gain on complex formation	-3.3	2.1%	-3.1	2.0%
average gain	-3.5	2.2%	-3.4	2.1%

3.4.9. Binding of metals with the LigA8-9 structure

Several *fo*-*fc* electron density blobs (with an σ contour level of more than 6) were observed on the dimer interface of LigA8-9. These electron densities indicated the presence of metals. Since the crystallization condition had potassium iodide, potassium and Iodide were modeled in the electron density. The coordination geometry of the potassium ions was manually checked. Modeled potassium ions were observed forming bonds with the side chains of amino acids and water molecules (**Figure 3.22**). The presence of potassium ions mostly at the dimer interface suggests that potassium ions might be involved in LigA8-9 dimerization. However,

no differences were observed in analytical size-exclusion chromatography. The protein did not elute as a dimer in the presence of potassium iodide (Data not shown).

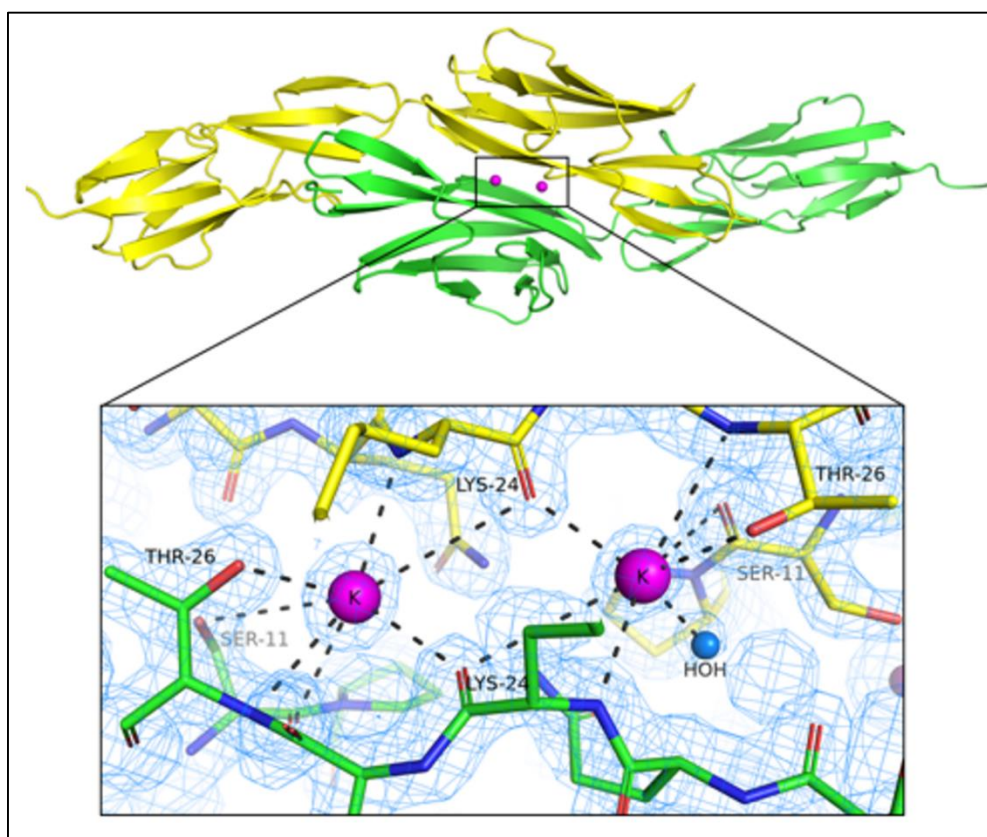


Figure 3.22: Representation of dimer structure of LigA8-9 with potassium ions. $F_o - F_c$ electron density (blue mesh) around the potassium ions. Residues participate in potassium-binding and water molecules are also shown. Chains A, and B of LigA8-9 are represented in the yellow and green cartoon, respectively. Potassium ions and water molecules are shown in magenta and blue spheres, respectively.

In addition, a few bigger *fo-fc* electron densities blobs were observed which were too big to be water or potassium. The CheckMyBlob tool (Brzezinski et al., 2021) (<https://checkmyblob.bioreproducibility.org/server/>), suggested calcium ions be preferentially modeled in these blobs. Hence, calcium ions were modeled, and their coordinated geometry was analyzed. The Ca^{2+} ions were found to coordinate perfectly with the side as well as main

chains of valine, lysine, and two threonine residues and water molecules (**Figure 3.23**). All four model Ca^{2+} ions were present in the LigA8. Surprisingly, no similar densities were observed in LigA9. Several previous reports suggest that the Lig proteins are calcium-binding proteins. Many in-vitro studies have demonstrated Ca^{2+} ions binding with Lig domains. Recently, the NMR structure of the LigA4 domain also showed Ca^{2+} ion interaction and reported amino residues involved in Ca^{2+} ions binding. Many of these residues were found conserved in LigA8-9 and observed to interact with Ca^{2+} ions. However, the physiological role of Ca^{2+} ions is not fully established.

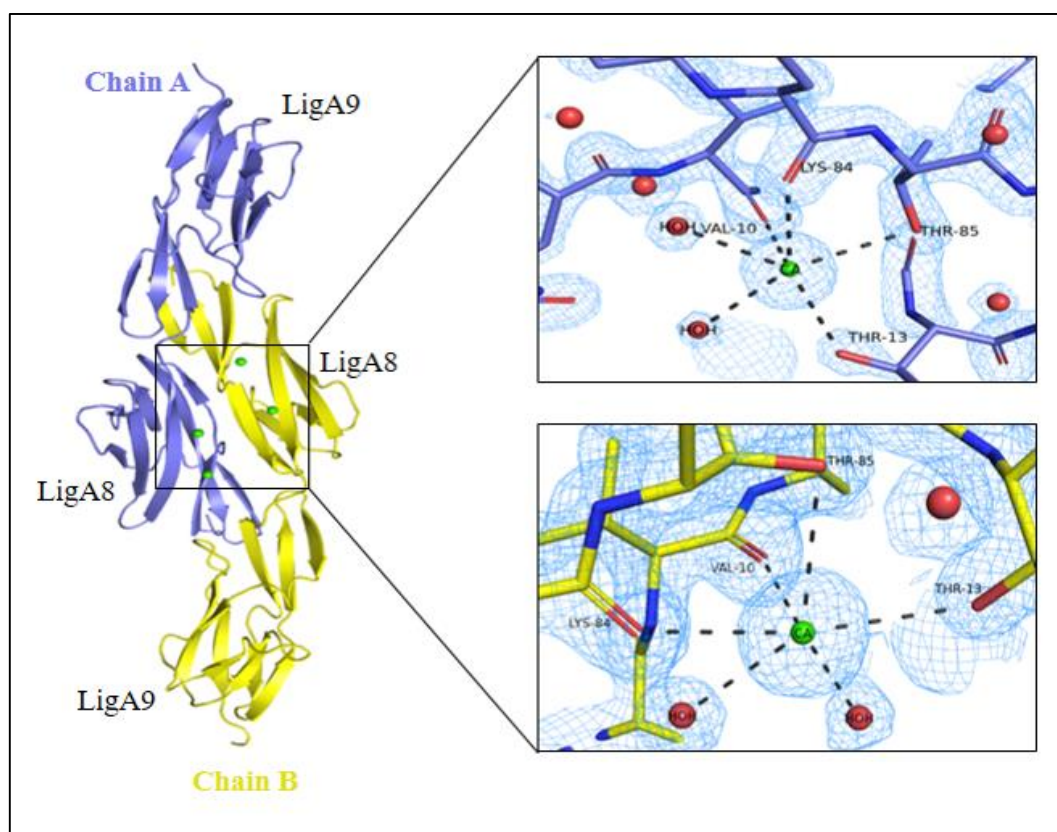


Figure 3.23: Representation of dimer structure of LigA8-9 with calcium ions. $F_o - F_c$ electron density (blue mesh) around the calcium ions. Residues participate in calcium-binding and two water molecules are also shown. Chains A, and B of LigA8-9 are represented in the blue and yellow cartoon, respectively. Calcium ions and water molecules are shown in green and red spheres, respectively.

3.4.10. Structural comparison among two domains

The two domains LigA8 and LigA9 showed an RMSD value of 0.507 Å when their structures were aligned. Structurally these two domains are well aligned with each other. Their core regions are well conserved. However, the lengths of a few strands – F and G were different. These strands were longer in LigA8. Moreover, the loop connecting strands A to A' in both structures was also different (**Figure 3.24A**). In addition, LigA8 and LigA9 displayed different electrostatic surface charge, suggesting a difference in surface residues among the two domains (**Figure 3.24B**).

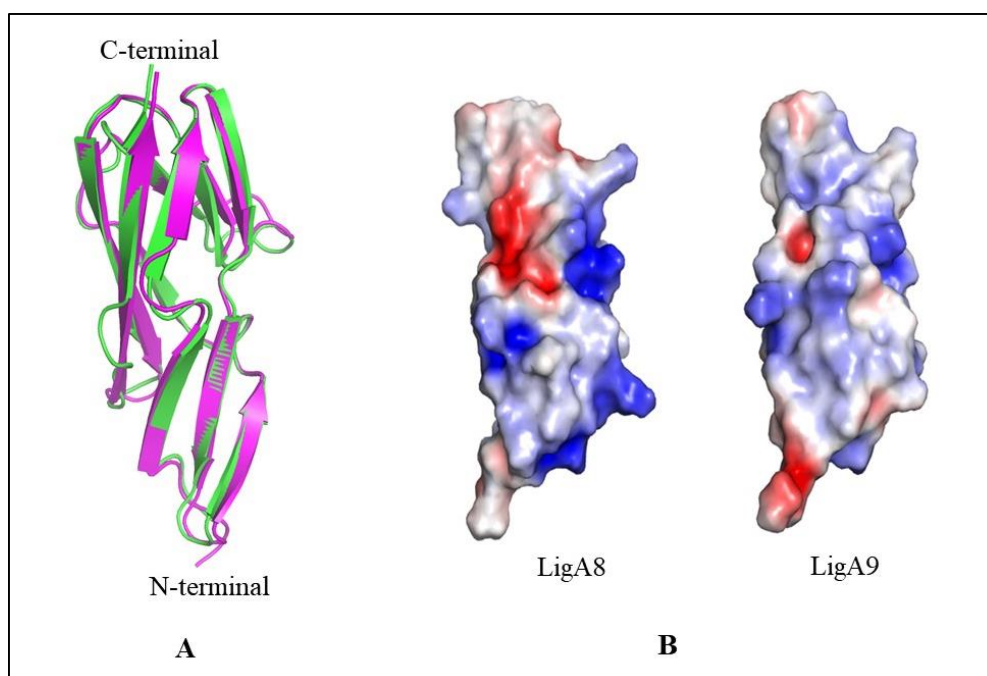


Figure 3.24: Comparison among LigA8 and LigA9 domains. (A) Superimposition of LigA8 with LigA9 structures represented as green and magenta cartoons, respectively. (B) Electrostatic surfaces of LigA8 and LigA9. Positive and negative charges distributions are indicated with blue and red color surfaces, respectively.

3.4.11. Structural comparison of LigA8-9 with other Big domains

The model structure of the protein, containing both individuals and the two-domain, was used to search for the closest homologs using the DALI server. The server showed that the protein

belongs to the B1g_2 family of proteins, as expected. The single-domain LigA8 showed the closest homology to LigBCen2R (PDB id-2MH4), with RMSD and z-score of 3.8 Å and 6.5, respectively. The structures of LigA9 and LigA8-9 showed the closest homology to LigA4 (PDB id- 2N7S) with RMSD of 3.1 Å (z-score-8.6) and 3.0 Å (z-score-8.6). The multiple sequence alignment of LigA8 and LigA9 and its closest homolog is given in **Figure 3.25** & **Figure 3.26**. An additional β -strand was observed in the case of LigA9, which was not present in any other reported domains. Furthermore, the closest homologous structures other than *Leptospira* are the Bacteriophage T5 tail protein Pb6 (PDB ID 4UID) for LigA8 and the crystal structure of the *Geobacillus stearothermophilus* sugar-binding S-layer protein SbsC (PDB ID 5NGJ) for LigA9 and LigA8-9.

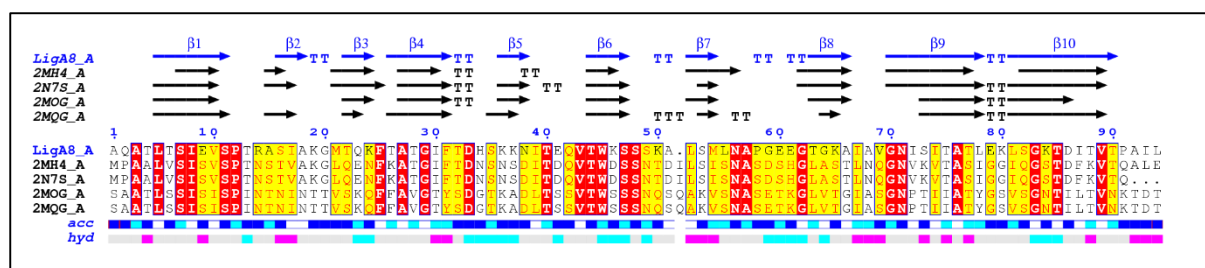


Figure 3.25: Sequence comparison between LigA8 and its closest structural homolog

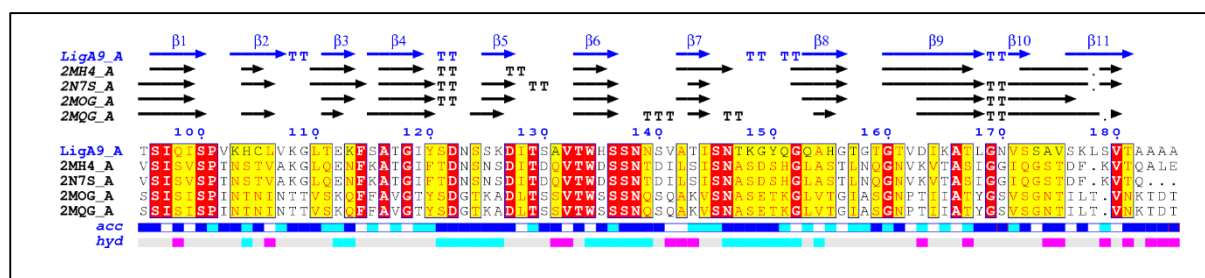


Figure 3.26: Sequence comparison between LigA9 and its closest structural homolog

The structures of the individual Ig-like domain of LigA8-9 were superimposed well with the NMR structures of LigA4 (PDBID- 2N7S), LigB12 (PDBID- 2MOG). LigA8 showed an RMSD value of 2.363 Å and 1.728 Å with the structures LigA4, and LigB12, respectively (**Figure 3.27 A&B**). On the other hand, the LigA9 showed an RMSD value of 2.426 Å and

1.604 Å with LigA4 and LigB12 structures, respectively (**Figure 3.27 E&F**). Interestingly, both LigA8 and LigA9 had fewer structural differences with the LigB12 compared to LigA4. Suggesting that LigA8 and LigA9 are structurally closer to the terminal domain of LigB. Furthermore, the structural comparison with intimin showed an RMSD of 2.217 Å with LigA8 and 0.408 Å with LigA9, suggesting LigA9 is structurally similar to intimin (**Figure 3.27 C&G**). However, LigA8 and LigA9 showed significant deviation with the Invasin (**Figure 3.27 D&H**).

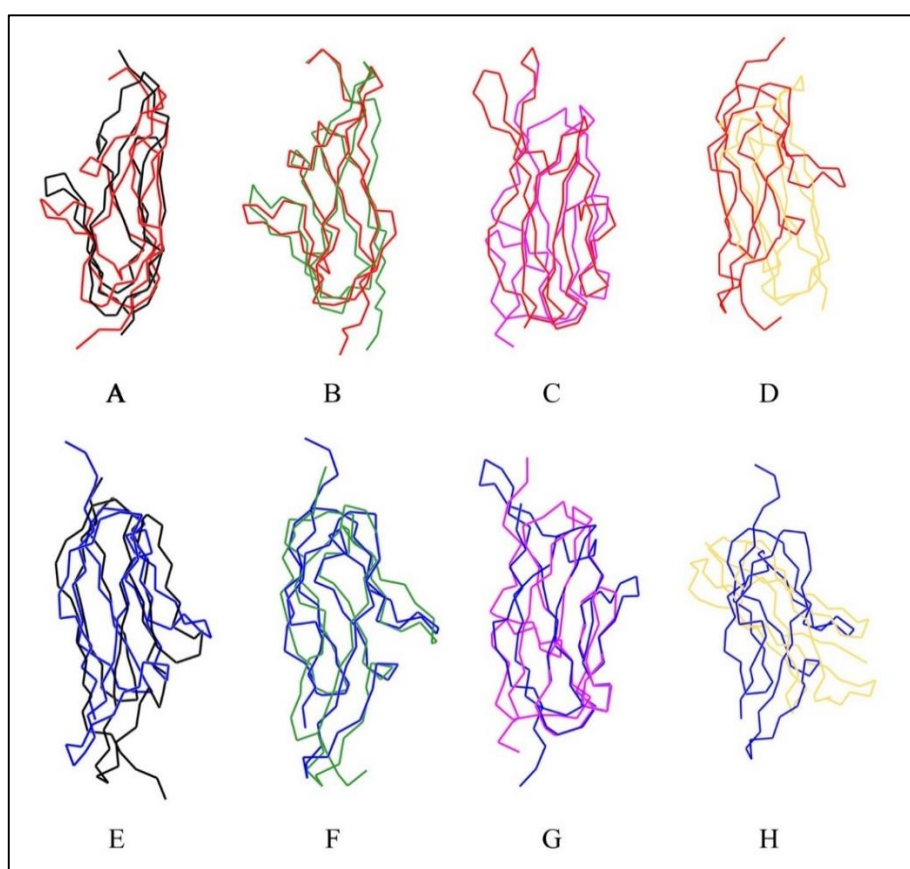


Figure 3.27: Structural comparison with Ig-like domain-containing proteins. Structures are represented as ribbon. (A) LigA8 (red) superimposed on LigA4 (black), (B) LigA8 (red) superimposed on LigB12 (green), (C) LigA8 (red) superimposed on Intimin (magenta) (D) LigA8 superimposed on Invasin (yellow-orange) (E) LigA9 superimposed on LigA4 (black) (F) LigA9 superimposed on LigB12 (green), (G) LigA9 superimposed on Intimin (magenta), (H) LigA9 superimposed on Invasin (yellow-orange).

3.4.12. Domain orientation in LigA8-9

The two domains of LigA8-9 were oriented at an angle of 124.8° to each other. The total shape of the LigA8-9 was observed as an elongated structure with a dimension of 914 Å length, 32.4 Å height and 28.4 Å breadth (**Figure 3.28 A&B**).

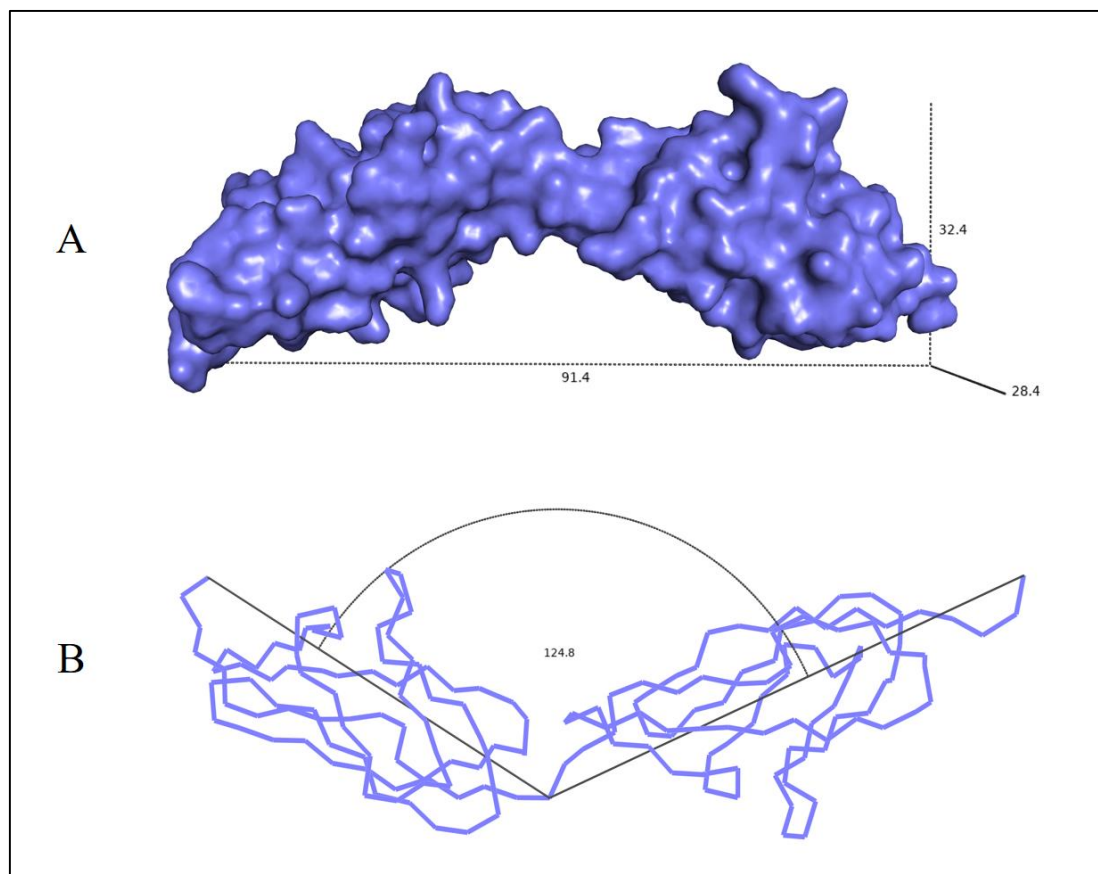


Figure 3.28: Size and orientation of LigA8-9. (A) Representation of the total shape and size of the LigA8-9, (B) Orientation angle between the two domains.

Orientation of these two domains in LigA7-8 is controlled by an inter-domain salt bridge near the loop region between LigA8 and Lig9. The lysine of LigA8 forms a salt bridge with aspartic acid of LigA9 (**Figure 3.29 A**). These residues were conserved in all domains (**Figure 3.29 B**). Electrostatic interactions between similar positions are also present in between the three BIG (bacterial Ig-like) domains found in the invasin (*Yersinia pseudotuberculosis*) structure. The involvement of salt bridges near the loop region promotes stability between multi-domain

LigA. Based on the angle formed between the two domains and the conservancy of the salt bridge forming residues, a probable domain orientation of the variable LigA7-13 is proposed (**Figure 3.29 C**). The domains may be oriented in an elongated zigzag fashion such a way that the conserved lysine residue from one Ig-like domain is readily positioned toward a conserved aspartic acid on the loop between β -strands of the subsequent Ig-like domain and results in the formation of salt-bridges.

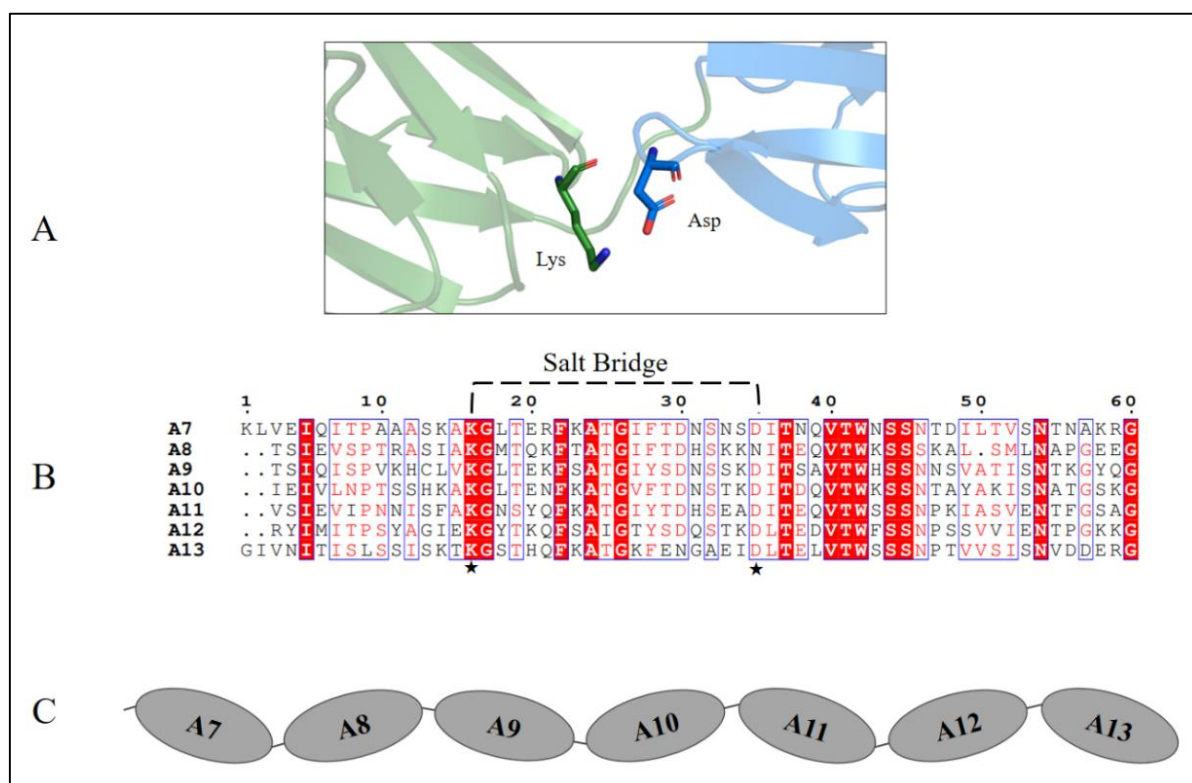


Figure 3.29: Proposed orientation of LigA variable domain (A) Lys and Asp interactions between LigA8 and LigA9 domains (B) Conservation of Lys and Asp residues among all Ig-like domains of LigA (C) Probable domains orientation in the variable domain of LigA.

3.4.13. Interaction with fibronectin

Lig proteins bind to the number of host ECMs proteins to establish colonization such as fibrinogen, fibronectin, integrin, and collagen. Several reports suggest that LigA and LigB proteins interact with the gelatin binding domain of fibronectin. The probable interacting residues of LigA8-9 with GBD were identified through molecular docking studies (**Figure 3.30**). GBD interacted with the loop region of LigA8 and LigA9 domains via 11 hydrogen bonds and 1 salt bridge interaction, involving 6 residues of GBD and 7 residues of LigA8-9 (**Figure 3.30 B&C**). The residues Lys126 (LigA9) and Asp99 (GBD) are involved in the formation of both H-bond and salt bridges.

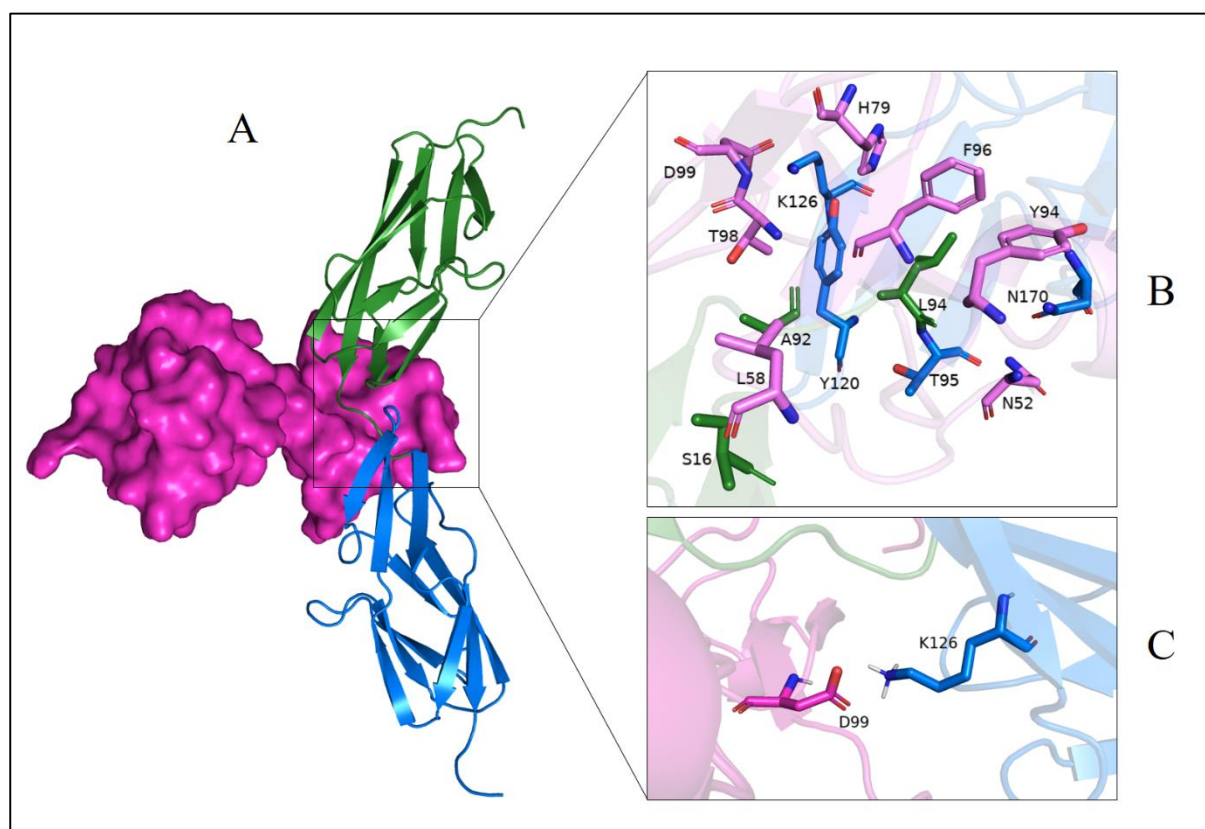


Figure 3.30: Docking GBD domain of fibronectin with the structure of LigA8-9 (A) Docked structure of GBD (Magenta, surface representation) and LigA8-9 (Forest green for LigA8 and marine blue for LigA9, cartoon representation) (B) Residues contributing in the hydrogen bond formation (C) Residues forming salt bridges.

3.5. Conclusions

Lig proteins play an important role in *Leptospira* infection and knowledge of its structure at atomic level can help in designing therapeutic against it. The NMR structure of only single domain is available. In this chapter we have tried to determine the structure of Lig proteins using x-ray crystallography method. The purified proteins (mentioned in Chapter 2) were subjected for crystallization screening using various commercially available screens. Many crystallization hits were obtained for the truncated proteins containing single and multiple domains. On optimization, we obtained diffraction quality crystals for truncated proteins such as LigA8-9, LigA12-13, LigA8-13 and LigB7-12. The high-resolution data for LigA8-9 were collected and solved at a resolution of 1.8Å. The structure determination using molecular replacement, reveals that the protein is majorly composed of β -sheets and followed Greek-key like folding pattern. The protein showed close homology with other Ig-like domains containing proteins such as intimin, LigA4 and LigB12. The angle of orientation between the two domains was calculated to be 124.8°. Using the structural details conserved salt bridge forming residues were identified in the loop region between the two domains. The structural details were further utilized to perform molecular docking with human fibronectin GBD. On analyzing the docked complex, we identified the probable residues involved in forming polar contacts with the human GBD. This chapter deals with the structure determination of LigA8-9.

Chapter - 4

Identification of antigenic epitopes within Lig Proteins

Work of this chapter is published as:

Kumar P., Lata S., Shankar UN, Akif M. Immunoinformatics-Based Designing of a Multi-Epitope Chimeric Vaccine From Multi-Domain Outer Surface Antigens of *Leptospira*. **Front Immunol.** 2021 Nov 30; 12:735373. doi: 10.3389/fimmu.2021.735373.

4.1. Introduction

Leptospirosis is considered one of the most neglected zoonotic diseases. The symptoms of the disease were identified over 100 years ago, but there is still a shortage of appropriate treatments and effective vaccines for the pathogenic *Leptospira*. Hence, an advancement in the development of new therapeutics and effective subunit vaccines is warranted. The current vaccines are based on inactivated whole-cell or membrane preparations from pathogenic *Leptospira* species. These are associated with severe side effects and do not provide cross-protection among the pathogenic *Leptospira* species (Adler and de la Peña Moctezuma, 2010b; Rodrigues de Oliveira et al., 2021b). In order to control the disease, several virulent factors involved in leptospiral pathogenesis are being investigated as both therapeutic targets and vaccine candidates. Outer membrane/surface proteins (OMPs) from pathogenic *Leptospira* play a significant role in establishing infection. They are being explored as attractive vaccine targets as most of them are conserved across the serovars. Moreover, they can be recognized by the host immune systems in the early phase of infection (Levett, 2015). One of the outer surface proteins, the Leptospiral immunoglobulin-like (Lig) family, is reported to be the most antigenic and attractive-vaccine candidate (Odir et al., 2011).

The family of Lig proteins is present exclusively in pathogenic species. Lig domains possess a similar structural fold as the adhesin domains from enterobacterial pathogens (Raghavan U M Palaniappan et al., 2007). The family of Lig proteins binds to host extracellular matrix components (ECM) and helps pathogens to invade and help in host tissue colonization (Christopher P. Ptak et al., 2014). In addition, Lig proteins are known to bind to the complement factors to evade innate immunity and establish the infection (Castiblanco-Valencia et al., 2012a). It has been reported that Lig proteins could induce significant protection against lethality in a hamster challenge model (Conrad et al., 2017; Evangelista et al., 2017; Silva et al., 2007b).

Moreover, in a mouse model, ~90% protection was induced by Lig proteins (Koizumi and Watanabe, 2004b). The ability of Lig proteins to bind varieties of host factors and contribute toward immune induction made it possible to consider Lig proteins as putative virulent factors and the most important vaccine candidate identified to date (Haake and Matsunaga, 2020). Recently, antigenic motifs within a single domain have been used to generate a chimeric immunoglobulin-like fold that showed enhanced leptospiral protection compared to whole Lig protein (Hsieh et al., 2017).

The Lig family of proteins is a big multi-domain protein, and it is thought that its immunogenic region may not be well accessible. Hence, knowledge of the immunogenic region/epitope will provide opportunities for further improvement in the Lig family vaccine efficacy through rational design. Many methods can be used to instigate the immunogenic regions or epitopes of the Lig family of proteins. Most of these methods are expensive and time-consuming for vaccine development. However, with the availability of many computational algorithms, identifying immunogenic regions or epitopes has been an easy task and has expedited vaccine development research (Dellagostin et al., 2017). There are various successful efforts toward epitope-based vaccines against diverse pathogens such as HIV (Xu et al., 2018), Influenza virus (Gottlieb and Ben-Yedidia, 2014), and hepatitis B and C viruses. Similar approaches have been used to identify potential epitopes against various antigens of Covid19 (Dong et al., 2020; Uttamrao et al., 2021). The two crucial arms that counter and eliminate pathogens are both humoral and cell-mediated immune responses. Hence, identifying BCL and CTL epitopes is essential for designing epitope-based engineered vaccines (De Gregorio and Rappuoli, 2012).

In the present study, the immunoinformatics approaches were applied for the comprehensive evaluation of antigenic epitopes present in the Lig family of proteins from *Leptospira interrogans*. Various T-cell epitopes were generated, and their binding interactions with major histocompatibility complexes (MHC) were analyzed. Subsequently, structural and continuous

B-cell epitopes were also predicted. In order to increase the confidentiality of prediction, three different tools were used to select common and overlapping epitopes. These epitopes were fused together with suitable linkers and incorporated with an adjuvant to generate a multi-epitope chimeric vaccine construct.

4.2. Methodology

4.2.1. Protein sequence retrieval

The amino acid sequences of full-length LigA (**Accession no. ACH89909.1**) and LigB (**Accession no. AAP74956.1**) were downloaded from NCBI in FASTA format. Using the domain information from the UniProt database, we separated the sequence of individual Ig-like domains from the full-length Lig sequence and saved them separately. All the prepared sequences of Ig-like domains were subsequently subjected to antigenic epitopes prediction. The methodology followed for the identification of most antigenic regions in LigA and LigB is given as a flow chart in **Figure 4.1**. Tools/servers used is listed in **Table 4.1**.

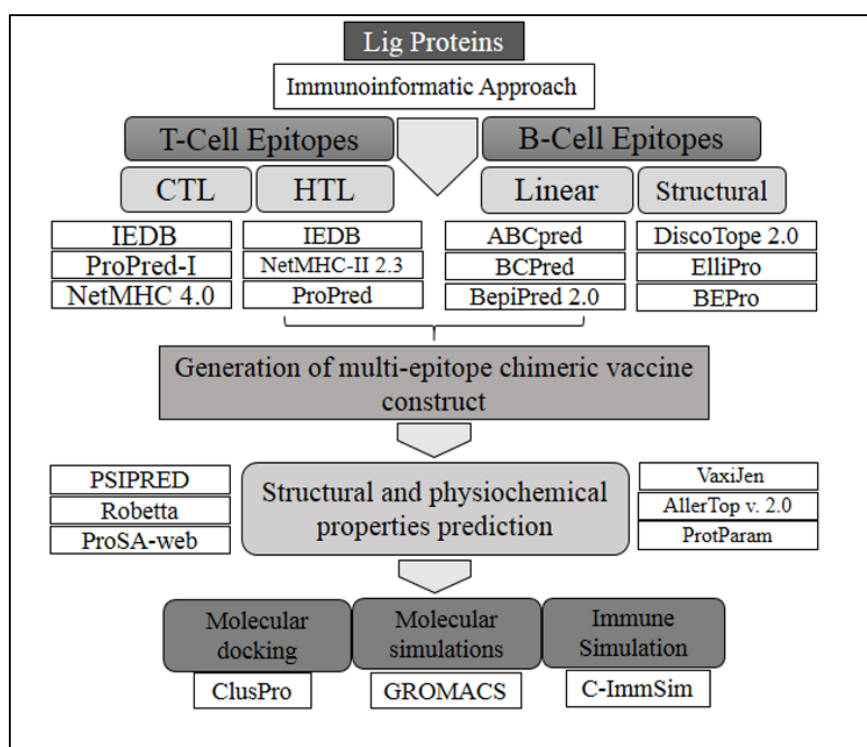


Figure 4.1: Flowchart of the methodology used in this study.

Table 4.1: List of tools and servers used in this study

Server used	Server link
NCBI	https://www.ncbi.nlm.nih.gov/protein
I-TASSER	https://zhanglab.dcmf.med.umich.edu/I-TASSER/
GalaxyRefine	http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=REFINE
SPDBV	https://spdbv.vital-it.ch/
ProSA Web	https://prosa.services.came.sbg.ac.at/prosa.php
Procheck	https://servicesn.mbi.ucla.edu/PROCHECK/
DiscoTope	http://www.cbs.dtu.dk/services/DiscoTope/
ElliPro	http://tools.iedb.org/ellipro/
BEPro	http://pepito.proteomics.ics.uci.edu/
ABCpred	https://webs.iiitd.edu.in/raghava/abcpred/
BCPred	http://ailab-projects1.ist.psu.edu:8080/bcpred/
BepiPred	http://www.cbs.dtu.dk/services/BepiPred/
IEDB (MHC-I)	http://tools.iedb.org/mhci/
ProPred-1	https://webs.iiitd.edu.in/raghava/propred1/
NetMHC 4.0	http://www.cbs.dtu.dk/services/NetMHC/
ProPred	https://webs.iiitd.edu.in/raghava/propred/page2.html
IEDB (MHC-II)	http://tools.iedb.org/mhcii/
NetMHC 2.3	http://www.cbs.dtu.dk/services/NetMHCII/
ProteoMapper	http://www.peptideatlas.org/map/
IFNepitope	https://webs.iiitd.edu.in/raghava/ifnepitope/predict.php
AllerTop	https://www.ddg-pharmfac.net/AllerTOP/
VaxiJen	http://www.ddg-pharmfac.net/vaxijen/VaxiJen/
ProtParam	https://www.expasy.org/resources/protparam
PSIPred	http://bioinf.cs.ucl.ac.uk/psipred/
Robetta	https://rosetta.bakerlab.org/
ClusPro	https://cluspro.bu.edu/login.php
RCSB	https://www.rcsb.org/
C-ImmSim	https://150.146.2.1/C-IMMSIM/index.php?page=1
JCAT	http://www.jcat.de/
UniProt	https://www.uniprot.org/

Cell-PLoc	http://www.csbio.sjtu.edu.cn/bioinf/Cell-PLoc-2/
TMHMM server	http://www.cbs.dtu.dk/services/TMHMM/
NCBI-BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi
Population Coverage	http://tools.iedb.org/population/
ToxinPred	https://webs.iiitd.edu.in/raghava/toxinpred/design.php
SOLPro	http://scratch.proteomics.ics.uci.edu/
PDBSum Generate	http://www.ebi.ac.uk/thornton-srv/databases/pdbsum/Generate

4.2.2. Prediction of BCL epitopes

Three different tools were used to predict the conformational and linear B-cell lymphocytes (BCL) epitopes. Since, for the conformational epitope prediction, structural information was required; hence all the individual Ig-like domains (7 variable domains from LigA and all 12 domains from LigB) of Lig proteins were modeled using an online server I-TASSER (Roy et al., 2010). The i-TASSER took the terminal Ig-like domain (LigB12) NMR structure as a template to model other individual Ig-like domains (Christopher P. Ptak et al., 2014). Using the online tool GalaxyRefine (Heo et al., 2013), the models were further refined, and energy was minimized using the SPDBV tool (Guex and Peitsch, 1997). The quality of the models was checked by ProSA-web (Wiederstein and Sippl, 2007) and validated by Ramachandran plot analysis (Laskowski et al., 1993). Three different online tools such as DiscoTope2.0 (Kringelum et al., 2012), ElliPro (Ponomarenko et al., 2008), and BEPPro (Jensen et al., 2018), with the default parameters, were employed for conformational epitope prediction.

In the case of sequence-based linear B-cell lymphocytes (BCL) epitopes prediction, ABCpred (Sudipto Saha and G. P. S. Raghava, 2007), BCPred (El-Manzalawy et al., 2008), and BepiPred 2.0 (Jespersen et al., 2017) were used with the default parameters.

4.2.3. CTL and HTL Epitope prediction

Potential cytotoxic T-lymphocyte (CTL) antigenic epitopes were predicted from each Ig-like domain, using three different servers 1. IEDB (Andreatta and Nielsen, 2016), 2. ProPred-I (Singh and Raghava, 2003) and 3. NetMHC 4.0 (Andreatta and Nielsen, 2016) with the default parameters. The IEDB tool works on an artificial neural network (ANN) to cover 80 different MHC-I alleles, which include 36 HLA-A alleles, 34 HLA-B alleles, and 10 HLA-C alleles. The peptides that showed IC₅₀ value ≤ 250 nM were selected. The ProPred-I uses a matrix-based method to predict potential epitopes by covering 47 MHC Class-I alleles to predict potential epitopes and also predicts proteasomal cleavage sites present on antigens. At the default threshold of 4%, the prediction showed equal sensitivity and specificity towards all the alleles; thus default threshold was selected for our analysis. The NetMHC 4.0 uses an artificial neural network algorithm to predict 9mer peptides by covering 81 HLA-alleles belonging to HLA-A, HLA -B, HLA-C, and HLA-E classes. Only strong binding peptides that showed a % rank lower than 0.5 were selected. In all cases, peptides binding to MHC-I have a strong preference for nonamer peptides; thus, nonamer peptides with good antigenic scores were selected from each server.

For predicting helper T-lymphocytes (HTL) epitopes, three different tools, namely ProPred (Singh and Raghava, 2002), IEDB (Wang et al., 2010), and NetMHC-II 2.3 (Jensen et al., 2018), were used. The ProPred uses quantitative matrices to predict 9-18 amino acid length peptides by covering 51 HLA-DR alleles of MHC-II. The peptides with a threshold of $\leq 3\%$ were selected. The IEDB server uses a consensus approach that predicts 15 mer peptides using different methods available. Peptides showed adjusted rank ≤ 1 were selected. The NetMHC-II 2.3 uses an artificial neural networking algorithm to predict 15 mer peptides. Peptides showed a % rank ≤ 2 were selected. NetMHC-II 2.3 covers a total of 61 different alleles consisting of HLA-DR alleles, HLA-DQ, HLA-DP, and 7 mouse H2 class II alleles. Peptides

that showed binding to ≥ 3 different alleles of MHC-I and MHC-II alleles were considered promiscuous.

4.2.4. Removal of self-peptides and selection of IFN- γ -inducing epitopes

The presence of self-peptides can induce autoimmune responses. To overcome this problem, the peptides that showed overlapping with human peptides were removed using the ProteoMapper tool (Desiere et al., 2006) at the peptide atlas server. The IFN- γ positive epitopes present among the HTL epitopes were selected using the IFNepitope server (Dhanda et al., 2013). The IFN- γ epitopes play an important role in innate as well as adaptive immune responses by inducing T-helper cells. SVM hybrid method was used to categorize the epitopes into IFN- γ versus Non-IFN- γ epitopes.

4.2.5. Epitope selection and construction of a multi-epitopes vaccine

Common and overlapping BCL, CTL, and HTL epitopes predicted from all three servers were selected. Among them, the promiscuous, IFN- γ positive, and non-self-epitopes with the highest antigenic score were shortlisted. Further, the epitopes common and overlapping among all three lymphocytes were considered for a vaccine to construct. Selected epitopes were linked with suitable linkers such as EAAAK, GPGPG, and AAY by giving sufficient space and flexibility among the peptides to fold correctly (Arai et al., 2001; Livingston et al., 2002a; Shiraz et al., 2021). Additionally, a mycobacterial heparin-binding hemagglutinin (HBHA) adjuvant is attached at the N-terminal to generate a multi-epitope chimeric vaccine construct (Nezafat et al., 2014a).

4.2.6. Immunoinformatic and physicochemical analysis of the vaccine construct

The constructed vaccine was analyzed for Allergenicity using the AllerTop v. 2.0 tool (Dimitrov et al., 2014). This tool works on auto cross-covariance (ACC) transformation of protein. Allergenicity of the vaccine was predicted using the amino acid sequence of the final

multi-epitope vaccine. The antigenicity of the vaccine was analyzed by the VaxiJen 2.0 server (Doytchinova and Flower, 2007). Physicochemical parameters of vaccine construct such as molecular weight, molar extinction coefficient, in-vivo and in-vitro half-life time, grand average hydropathicity (GRAVY) index, and instability index were evaluated by using the ProtParam tool (Wilkins et al., 1999).

4.2.7. Prediction of secondary and tertiary structure

The 2° structure contents in the multi-epitope vaccine construct were predicted by the PSIPRED server (Jones, 1999). The PSIPRED uses two steps artificial neural networking algorithm based on position-specific scoring matrices generated by PSI-BLAST. To model the 3-D structure of the multi-epitope vaccine construct, the Robetta server (Kim et al., 2004) was used. ProSA-web server was used for further validating the quality of the modeled structure.

4.2.8. Molecular docking of multi-epitope vaccine with immune receptor

For evaluating the interaction of the multi-epitope vaccine construct with the immune receptor, TLR4, the ClusPro server (Kozakov et al., 2017) was used with the default parameters. The coordinate of the TLR4 (PDB ID: 3FXI) (Park et al., 2009) was downloaded from the RCSB, followed by removing non-TLR molecules. The monomeric form of TLR4 was used to dock with the vaccine construct. The complex between the two was generated using three subsequent steps rigid-body docking, clustering of the lowest energy structure, and structural refinement. The best-docked complex structure was selected based on maximum cluster size and lowest energy score. The interaction in the complex was analyzed on LIGPLOT v2.2 and visualized on PyMOL.

4.2.9. Molecular Dynamics Simulation studies

To study the stability of the interaction between the multi-epitope vaccine construct and the TLR4, molecular dynamics simulation was carried out using OPLS all-atom force

(Oostenbrink et al., 2004) field implemented in GROMACS 5.1.2. (Hess et al., 2008; Pronk et al., 2013). The SPC (Simple point charge) was selected as a water model (H J C Berendsen et al., 1987). The complex system was solvated in a cubic box with water molecules of 1.5 nm to the box wall from the protein's surface. The system was further neutralized by adding Na⁺ counter ions. The steepest descent method for 25,000 steps was used to cut down the internal steric clashes until the largest force acting in the system was smaller than 1,000 kJ/mol/nm. The complex was further equilibrated first in an NVT ensemble at 300 K for 50 ps using a modified V-rescale Berendsen thermostat with a time constant of 0.1 ps, followed by an NPT ensemble to 1 atm using a Parrinello–Rahman coupling method with a time constant of 2 ps for 50 ps (Bussi et al., 2007; Parrinello and Rahman, 1981). The complex was submitted for the final MD run for 100 ns at 300 K. The equations of motion were integrated with time steps of 2 fs and the coordinates were saved for every 2500-time step (5ps), which resulted in a total of 4000 frames for a 50 ns simulation. The Particle Mesh Ewald (PME) method with a real space cut-off of 10 Å was used for long-range electrostatic interactions (H. J.C. Berendsen et al., 1987). Constraints were applied to bonds involving hydrogens by implementing the P-LINCS algorithm (Hess, 2008). A total of 20000 frames were used for graph analysis from the production run. RMSD, RMSF, and gyrate were calculated using `gmx rms`, `gmx rmsf`, and `gmx gyrate` commands implemented in the GROMACS. The graphs were plotted using XMGRACE.

4.2.10. Immune simulation

The immunogenic response of the multi-epitope chimeric vaccine construct was predicted using the online server C-ImmSim (Rapin et al., 2010). Using the sequence information of vaccine construct, both arms of immune responses were predicted. An immune simulation was performed with 100 simulation steps and one injection without LPS. The response recorded was analyzed further.

4.2.11. *In-silico* cloning

Finally, in-silico cloning was done using a suitable plasmid containing a multi-epitope construct nucleotide sequence. The JCat server was used for codon optimization of DNA sequence based on *E.coli* K12 strain (No et al., 2005). The percentage of GC ranging between 30 and 70 is treated as an optimal value. The cloning was done in the pET28a (+) expression vector using SnapGene software at BamHI and HindIII restriction sites (Adames et al., 2015).

4.3. Results

4.3.1. Collection and primary sequence analysis

The sequences corresponding to individual Ig-like domains were separated from LigA and LigB. LigA consists of 1224 residues which form 13 Ig-like domains. On the other hand, LigB possesses 1890 residues that are divided into 12 Ig-like domains and 766 residues as non-random C-terminal extensions. Each Ig-like domain consists of approximately 90 residues and is connected to the adjacent Ig-like domains by three amino acid linkers. Since the first six and a half domains are conserved among the LigA and LigB proteins, the sequences corresponding to all 12 Ig-like domains from constant and variable regions of LigB (without non-repetitive C-terminal domain) were considered for the analysis. While sequences corresponding to only variable Ig-like domains of LigA were used for antigenic epitope identification.

4.3.2. Identification of linear and conformational BCL epitopes

The family of proteins (LigA and LigB) has been reported to be the most potent vaccine candidate against leptospirosis (Choy et al., 2011; Henry A Choy et al., 2007b; Croda et al., 2007). Humoral, as well as cell-mediated immunities are triggered in *Leptospira* infection. Humoral immunity is majorly involved in the clearance of *Leptospira* by the production of antibodies like IgG and IgM against the leptospiral lipopolysaccharides (LPS) and other antigenic proteins (Adler and Faine, 1977; Evangelista and Coburn, 2010). In fact, many other

reports demonstrated the implication of the cell-mediated immune responses against leptospirosis. Ying-Jun Guo *et al.* have identified a set of peptides among the Lig that were associated with the human CD8⁺ T-lymphocytes (Y.-J. Guo *et al.*, 2010). Similar studies also illustrated the induction of CD4⁺ and γ/δ T-cells to stimulate type 1 cytokine production against the heat-killed *Leptospira* (De Fost *et al.*, 2003; B M Naiman *et al.*, 2001). The potential immune protection of Lig proteins and their truncated version in the context of vaccine candidates and their role in T-cells and B-cell mediated immune response against *Leptospira* have been highlighted (Haake and Matsunaga, 2021). In order to identify B-cell epitopes, we generated a 3-dimensional model of all individual Ig-like domains from both proteins (**Figure 4.2**). All the individual domains showed a similar Ig-like fold with the Greek-Key motif. The good Z-score values and fairly good Ramachandran statistics verify the quality of 3-D models of individual domains (**Table 4.2**). We used DiscoTope, ElliPro, and BEPro, three different servers, to identify conformational (Non-continuous) B-cell epitopes. All the predicted epitopes, for each Ig-like domain, from three servers were matched, and only common and overlapping epitopes for each case were chosen. The three servers yielded a different number of discontinuous epitopes, and common epitope residues were selected. As in the case of LigA7, DiscoTope and BEPro predicted 13 residues that are involved in the formation of a conformational B-cell epitope, while the ElliPro for the same yielded three epitopes. One among the three was observed to be common and overlapping with the former.

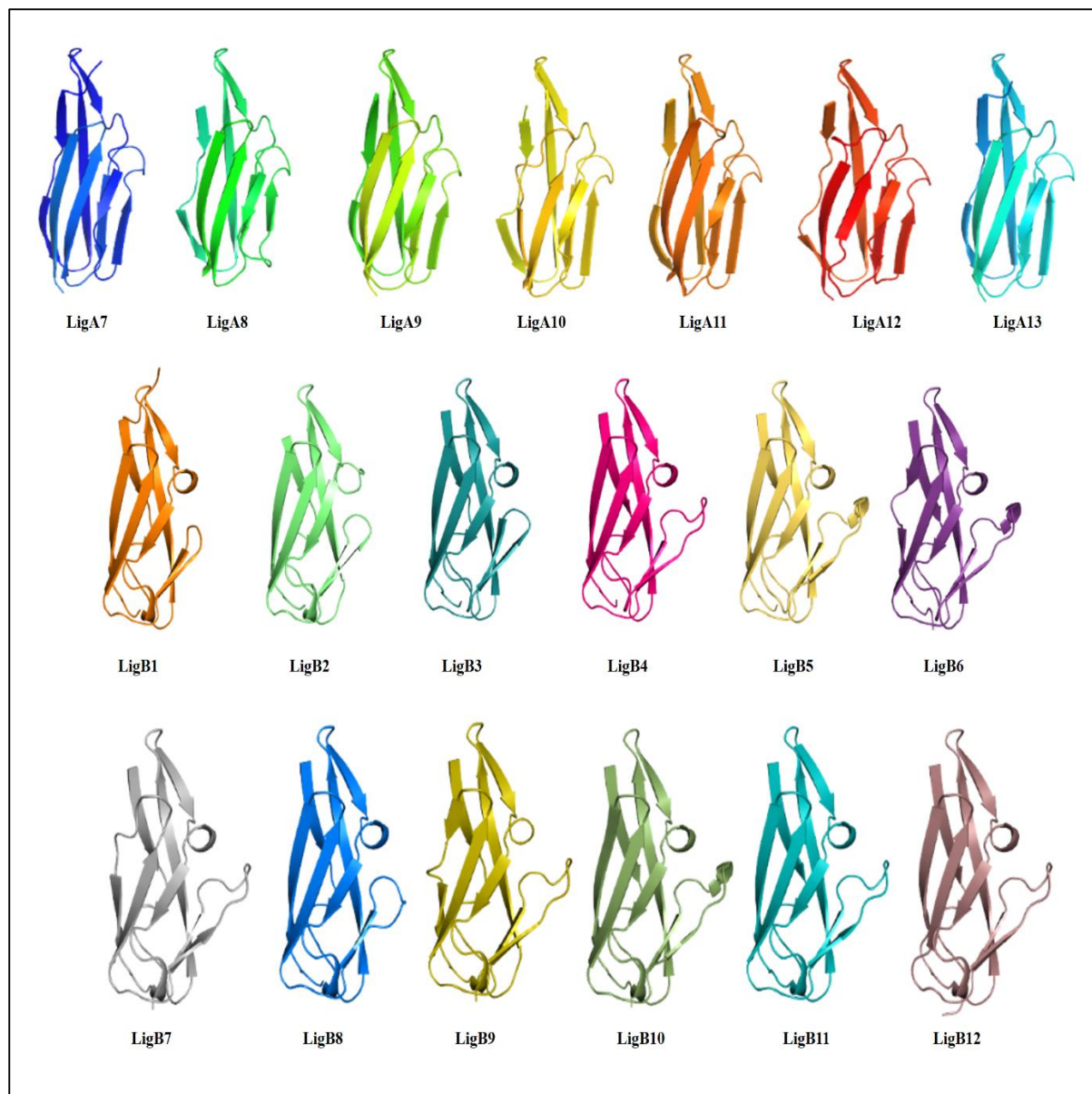


Figure 4.2: Homologous models of single variable LigA Ig-like domains and all individual LigB Ig-like domains

Table 4.2: Ramachandran statistics and ProSA Z score

Protein	Favoured Region (%)	Allowed Region (%)	Outlier Region (%)	ProSA Z-Score
LigA7	94.3	5.7	0	-3.64
LigA8	92.7	6.1	1.2	-4.47
LigA9	93	7	0	-4.47
LigA10	95.3	3.5	1.2	-4.18
LigA11	94.1	5.9	0	-4.66
LigA12	92	6.9	1.1	-5.17
LigA13	92	6.9	1.1	-4.35
LigB1	95.3	4.7	0	-4.96
LigB2	96.4	2.4	1.2	-3.88
LigB3	96.3	3.7	0	-4.01
LigB4	96.5	2.3	1.2	-4.10
LigB5	97.6	1.2	1.2	-3.50
LigB6	97.7	1.1	1.1	-3.82
LigB7	94.3	4.5	1.1	-2.98
LigB8	94	6	0	-3.10
LigB9	94.3	4.5	1.1	-3.73
LigB10	94.3	4.5	1.1	-3.17
LigB11	95.3	3.5	1.2	-2.77
LigB12	95.5	3.4	1.1	-4.70

The commonly selected epitope also showed a good antigenic score (**Appendix Table 1**). A similar approach was followed to select the overlapping or partial overlapping epitopes from other Ig-like domains of Lig proteins. The number of epitopes and residues involved and their antigenic scores predicted for each Ig-like domain are listed in **Appendix Table 1**. The bold residues in the list depict common and overlapping epitopes. Prediction with different servers increases the confidence in the selection of epitopes. Similarly, three servers such as ABCpred, BCPred, and BepiPred 2.0, were used for the linear BCL epitope identification. A total of 97

epitopes were predicted from the ABCpred. LigA8 shared maximum linear BCL epitopes. The BCPred predicted a total of 68 linear epitopes, and LigA8, LigB4, and LigB6 Ig-like domains possess a significantly high number of epitopes. The BepiPred 2.0 server generated a total of 71 epitopes, and LigA11 was observed to have maximum epitopes. For each Ig-like domain, the epitopes predicted from all three servers were matched, and common and overlapping epitopes were sorted out. Linear epitopes for each Ig-like domain generated from each server and their positions are listed in **Appendix Table 2**. The final BCL epitopes were selected based on their presence as completely or partially in both linear and conformational predicted epitopes. This selection criterion yielded five such epitopes from each Lig protein. The presence of these epitopes on the Ig-like domains and their length and location are listed in **Table 4.3**.

4.3.3. Identification of CTL and HTL Epitopes

Cytotoxic T Lymphocytes epitopes were identified among the Lig protein using three different servers such as IEDB, ProPred-1, and NetMHC. The IEDB server generated 400 strong MHC-I binding epitopes covering 35 HLA alleles for LigA and 36 HLA alleles for LigB, and the maximum number of epitopes are observed to bind with HLA-A*68:02 allele from both proteins. LigA contributed 27 peptides, and LigB shared 52 peptides towards binding to the HLA-A*68:02 allele. The ProPred-I generated 1032 epitopes covering all 47 HLA alleles and showed maximum binding with the HLA-B*58:01 allele. LigA and LigB contributed 176 and 112 peptides towards binding this allele, respectively. The NetMHC 4.0 generated 327 epitopes that cover 72 alleles for LigA and 76 alleles for LigB. The NetMHC 4.0 predicted the maximum number of peptides for HLA-A*11:01 allele in the case of LigA (18 peptides) and HLA-C*15:02 in the case of LigB (26 peptides) (**Appendix Table 3**). Epitopes that are overlapping or partially overlapping with at least two out of three servers and promiscuous nature, i.e., binding to three or more HLA alleles, are selected further for vaccine design.

In the case of HTL epitopes, the IEDB server predicted a total of 185 strong MHC-II binding epitopes (maximum binds to HLA-DRB1*07:01 for LigB and HLA-DRB1*04:23, HLA-DRB1*04:26, HLA-DRB1*04:10 equally for LigA), ProPred server predicted 177 epitopes (DRB1_0402 allele produced maximum epitopes for LigB and DRB1_0423 for LigA) and NetMHCII 2.3 server predicted 665 epitopes (HLA-DQA10201-DQB10303 generated maximum for both LigA and LigB) (**Appendix Table 4**). After considering promiscuous HTL epitopes, the number of epitopes for each case was sorted as 24, 129, and 215 from IEDB, ProPred, and NetMHCII2.3 servers, respectively. The best candidates from each server were selected based on IFN- γ positive epitopes. Finally, the HLA epitopes showing their presence either completely or partially overlapping with at least two out of three servers are considered for the vaccine design. The final selected epitopes were able to map only on four Ig-like domains of each Lig protein (**Table 4.5**).

4.3.4. Generation of multi-epitope chimeric vaccine

The best five CTL, HTL, and BCL epitopes from LigA and LigB were used to construct multi-epitope vaccines. The selected epitopes were predominantly present on Ig-like domains, LigA7, LigA9, LigA11, and LigA13 of LigA protein, and Ig-like domains, LigB1, LigB8, LigB9, and LigB12 from LigB protein. The antigenic score of these epitopes varies from 0.4 to 1.77 (**Table 4.3**). These epitopes were linked by short amino acid linkers such as AAY and GPGPG. To enhance the possibility of induction good immune response, the chimeric vaccine was additionally fused with a well-known TLR4 agonist mycobacterial heparin-binding hemagglutinin (HBHA) by EAAAK linker. The best combination of these epitopes was selected based on its stability and solubility. The final chimeric vaccine construct was 360 amino acids long (**Figure 4.3**) and was devoid of any sequence homology with the human protein sequences.

Table 4.3: List of selected common epitopes for vaccine construction

Domain	MHC-I	MHC-II	B-Cell Epitopes	Selected region	Score
LigA9	ATISNTKGY (48-56)	SNNSVATISNTKG (43-55)	NTKGYQGQAHGTGT (52-65)	ATISNTKGYQGQAHGTGT (48-65)	1.779
LigA7	VEIQITPAA (3-11)	VEIQITPAA (3-11)	IQITPAAASKAKGLT (5-19)	VEIQITPAAASKAKGLT (3-19)	0.950
LigA7	GTVKVTASM (68-76)	TVKVTASMGG (69-78)	LGSTLKQGTVKVTA (61-74)	LGSTLKQGTVKVTASMGG (61-78)	0.879
LigA13	TISLSSISK (6-14)	IVNITISLS (2-9)	SSISKTKGSTHQFK (10-24)	TISLSSISKTKGSTHQFK (6-24)	0.844
LigA11	EVIPNNISF (4-12)	VIPNNISFA (5-13)	VIPNNISFAKGNSYQFKATG (5-24)	VIPNNISFAKGNSYQFKATG (5-24)	0.648
LigB12	TVSKQFFAV (15-23)	ISPINTNINTTV (5-17)	ISPINTNINTTVS (5-17)	ISPINTNINTTVSKQFFAV (5-23)	0.740
LigB12	KQFFAVGTY (18-26)	FFAVGTYSA (20-28)	FFAVGTYSAGTKAD (20-34)	KQFFAVGTYSAGTKAD (18-34)	0.694
LigB8	MVNNVTGSV (49-57)	MVNNVTGSV (49-57)	VNNVTGSVTTVA (50-61)	MVNNVTGSVTTVA (49-61)	0.648
LigB9	TSIEITPTI (1-9)	SIEITPTINS (2-10)	PTINSITHGLTKQF (7-20)	TSIEITPTINSITHGLTKQF (1-20)	0.563
LigB1	IKAEYNGLY (68-76)	IQGNRVRGI (53-61)	RVRGIASGSSIIKAEYNGLYSEQKIT V (57-83)	IQGNRVRGIASGSSIIKAEYNGLYSEQKITV (53-83)	0.433

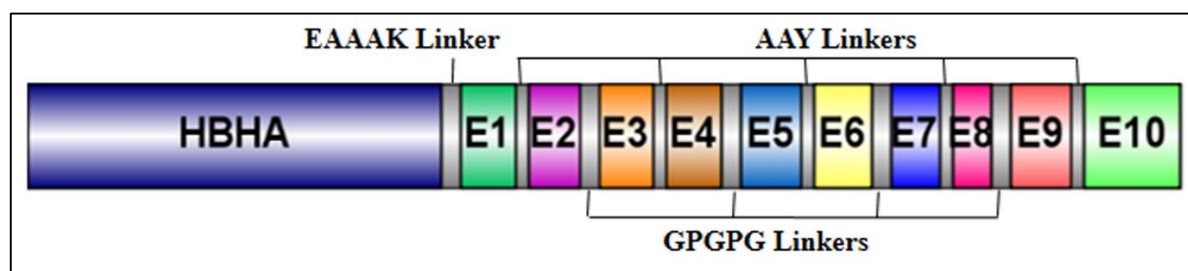


Figure 4.3: Schematic representation of the multi-epitope vaccine construct. HBHA is an adjuvant (TLR4 agonist mycobacterial heparin-binding hemagglutinin). Epitopes are represented as E and numbered from 1 to 10.

4.3.5. Physicochemical features of chimeric vaccine construct

The predicted solubility score of the vaccine construct was observed as -0.007, which suggests that the vaccine construct is soluble upon overexpression in a heterologous expression system. The stability index was observed to be 13.99, which categorized it as a stable protein molecule. The overall molecular weight was calculated to be 36.6 kDa. The estimated half-life in mammalian reticulocytes (in-vitro) was observed 30 hours, >20 hours in yeast (in-vivo), and >10 hours in *E. coli* (in-vivo). Finally, the antigenicity and allergenicity score of the vaccine classified it as non-allergen for humans and a probable antigen and can induce an immune response when used as a vaccine. All predicted features of the vaccine construct are summarized in **Table 4.4**.

Table 4.4: Physicochemical features of chimeric vaccine construct

Physicochemical properties	
Molecular weight (kDa)	36.6
The instability index	13.99 (Stable)
GRAVY score	-0.007 (soluble)
The estimated half-life	30 hours (mammalian reticulocytes, in vitro) >20 hours (yeast, <i>in-vivo</i>) >10 hours (<i>E. coli</i> , <i>in-vivo</i>)
Ext. coefficient	14900
Allergenicity (AllerTop 2.0)	Non-Allergen
Antigenicity Score (VaxiJen 2.0)	0.6848 (Antigen)

4.3.6. Secondary and Tertiary structure prediction

The secondary structure content of the vaccine construct includes 16.96% α -helix, 41.38% extended strand, and 41.66% random coil. The residues and region of the vaccine construct involved in the formation of secondary structure are highlighted in **Figure 4.4**. The 3-D model of the chimeric vaccine construct showed fairly good Ramachandran plot statistics (**Figure 4.5**). Approximately 99.7% of the residues were observed in the favored and additionally allowed region, and only 0.3% of residues were present in generously allowed regions of the Ramachandran plot (**Figure 4.5 B**). These statistics suggested that the 3-D model possessed fairly ideal bond lengths and bond angles. The calculated Z-score of 5.61 from ProSA server also validates the 3-D model of the vaccine construct. The Ramachandran and Z-Score plots of the same are shown in **Figure 4.5 B**. The 3D model of the final vaccine construct consisting of epitope regions from different Ig-like domains is shown in **Figure 4.5 A**.



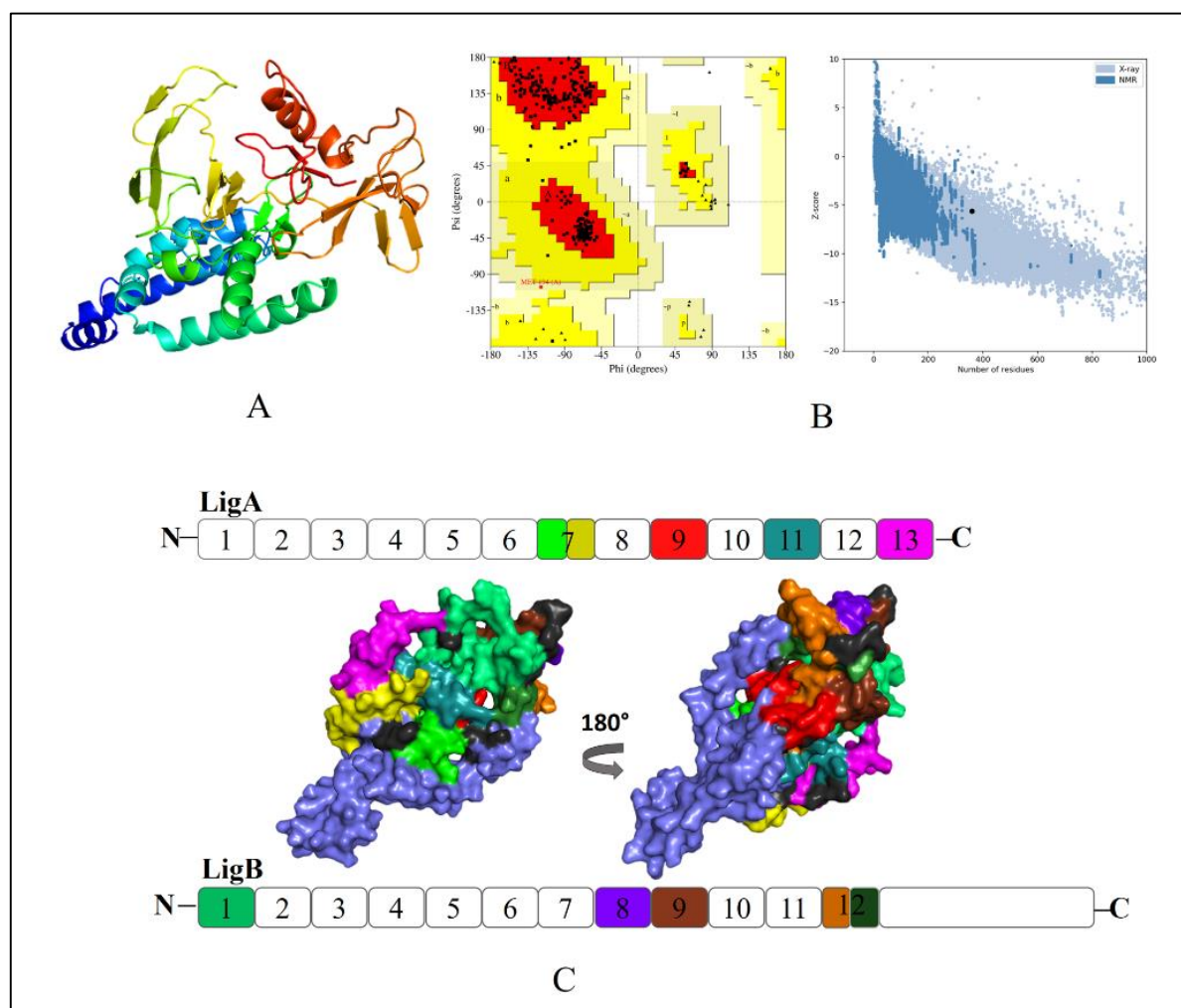


Figure 4.5: (A) 3-D model of vaccine construct. (B) Plot showing validation of the 3-D model by Ramachandran and ProSA server. (C) Promiscuous epitopes used in the vaccine construct are mapped with the same color code on the Ig-like domains of Lig proteins.

4.3.7. Docking complex of final vaccine construct with immune receptor

To order to get an insight into the binding of the final vaccine construct with the relative immune receptors, a molecular docking study was executed with the TLR4. The molecular docking analysis generated more than 30 different clusters. Cluster one showed the highest binding score and was thus considered for further analysis. The interacting residues between the two were analyzed using LIGPLOT v2.2, showing a strong interaction involving 32 hydrogen bonds and five salt bridges (**Table 4.6**). In the 3-D model of the docked complex, the

vaccine construct was observed to interact at the concave side of TLR4 (**Figure 4.6 A**). The residues from both that are involved in making interactions such as polar as well as salt bridges are represented in the 2-D plot (**Figure 4.6 B**).

4.3.8. Molecular Dynamics Simulation

To evaluate the stable interactions and dynamic behavior of the multi-epitope vaccine construct with the TLR4, Molecular Dynamics Simulation was performed for the docked complex at 100ns. The stability of the complex was investigated by some important parameters. The complex Root Mean Square Deviations (RMSD) trajectory represents the structural variations associated within the overall structure of the complex formation. The RMSD trajectory of the complex indicates that the system has attained adequate stability after 50ns of the MDS (**Figure 4.7 A**). The RMSD trajectories of each protein in the complex were analyzed (**Figure 4.8 A**). The RMSD of the complex started with 0.2nm and reached 0.65nm after stabilization; it was maintained till the end of the simulations, which suggests that the structural arrangement for the formation of the complex was stable. However, the RMSD plot reflects that there are significant structural deviations at 40ns, 45ns & 70ns time points. So, to validate the convergence after 50ns, we compared the change in interface interactions at each deviated time point along with the stable time frames. Although there is a slight deviation in the backbone, overall interactions were maintained till the end of the simulations (**Figure 4.7 B**).

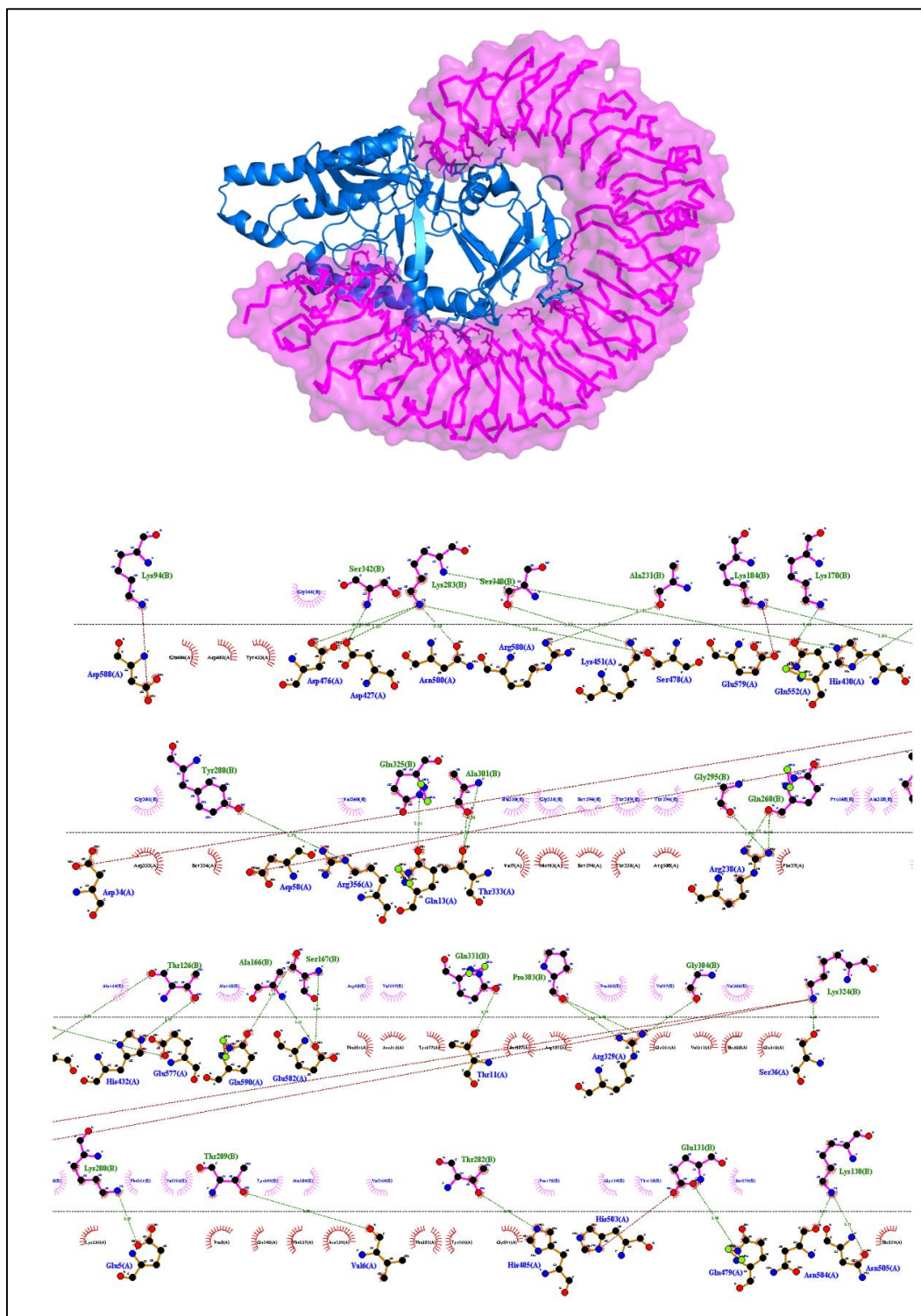


Figure 4.6: (A) Represent a docked complex of vaccine construct and TLR4 and shown as marine blue cartoon and magenta surface, respectively. (B) 2-D representation of interactions between vaccine construct and TLR4. Red arcs represent salt bridges, and green dotted lines represent hydrogen bonds.

Table 4.5: List of residues involved in forming polar interaction and salt bridges

Interactions	Receptor Residues	Ligand Residues
Hydrogen bonds	GLU-5, VAL-6, THR-11, GLN-13, SER-36, ARG-238, ARG-329, THR-333, ARG-356, HIS-405, ASP-427, HIS-430, HIS-432, LYS-451, ASP-476, SER-478, GLN-479, ASN-500, ASN-504, GLN-505, GLN-552, GLU-577, ARG-580, GLU-582, GLN-590	LYS-208, THR-209, GLN-331, GLN-325, LYS-324, GLN-260, GLY-304, PRO-303, GLY-304, TYR-288, THR-282, THR-126, LYS-283, GLU-131, LYS-130, LYS-184, ALA-231, ALA-166, SER-167
Salt-bridges	GLU-5, ASP-34, ASP-58, ASP-476, HIS-503, GLU-577, GLU-579, ASP-588	LYS-208, LYS-324, LYS-324, LYS-283, GLU-131, LYS-184, LYS-184, LYS-94

The Root Mean Square Fluctuations (RMSF) of all residues were computed for the docked complex. The RMSF trajectory highlights the flexibility of the residue in the docked complex. From our results, the residues 0-600 showed relatively low RMSF, which indicates that these residues may be involved in certain interactions to stabilize that region (**Figure 4.7 C**). In contrast, the construct residues 800-820 were found to have elevated RMSF values indicating slightly higher flexibility in these regions. Additionally, we performed separate 50ns MDS for only the construct part to check its flexibility. It was observed that the residual RMSF is also following the same trend as in the complex and not exceedingly more than 1.1nm (**Figure 4.8 B**).

The compactness of the complex was evaluated with the change in Radius of Gyration (Rg). It was observed that during the start of simulations, the complex had higher Rg values, but after 50ns, it attained a steady Rg trend (**Figure 4.7 D**). This fluctuation in Rg might be due to the folding and unfolding nature of the protein. For precision, we have analyzed the separate Rg

trajectory for TLR4 and construct, which clearly indicates that both proteins have attained a flat trajectory suggesting its compactness (**Figure 4.8 C**)

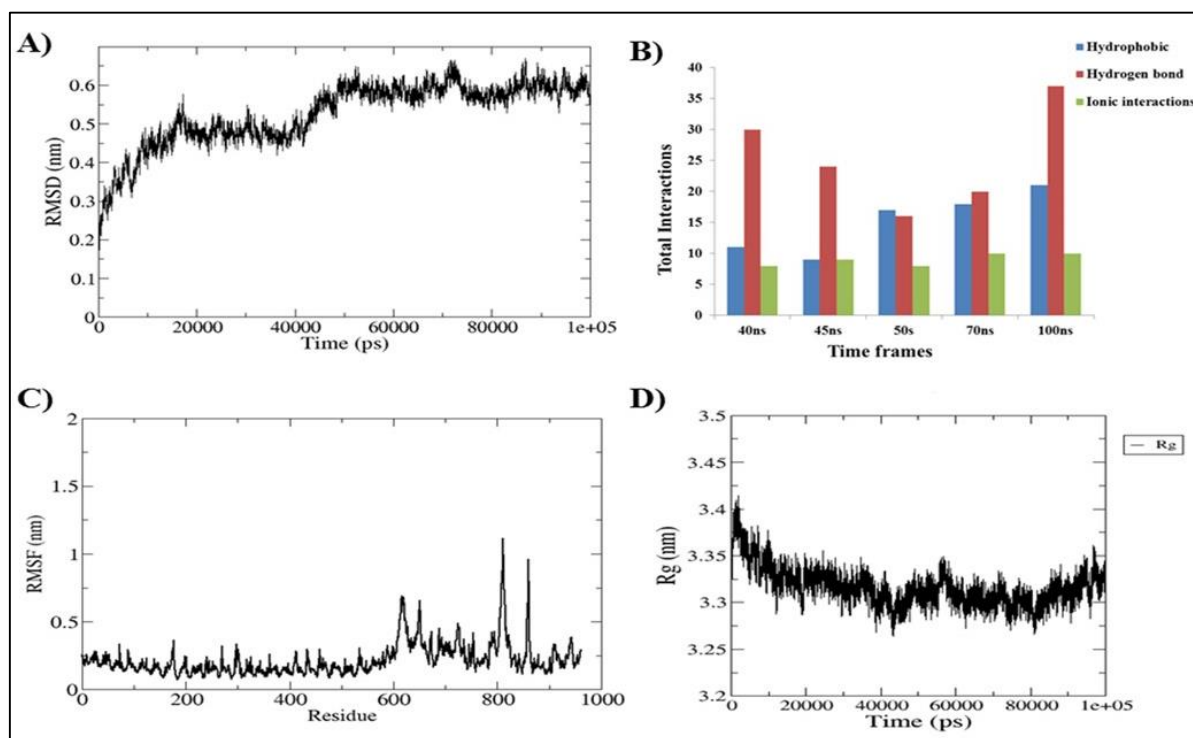


Figure 4.7: MDS plots of vaccine construct and TLR4 complex. (A) Backbone RMSD. (B) RMSF of all atoms. (C) The radius of gyration.

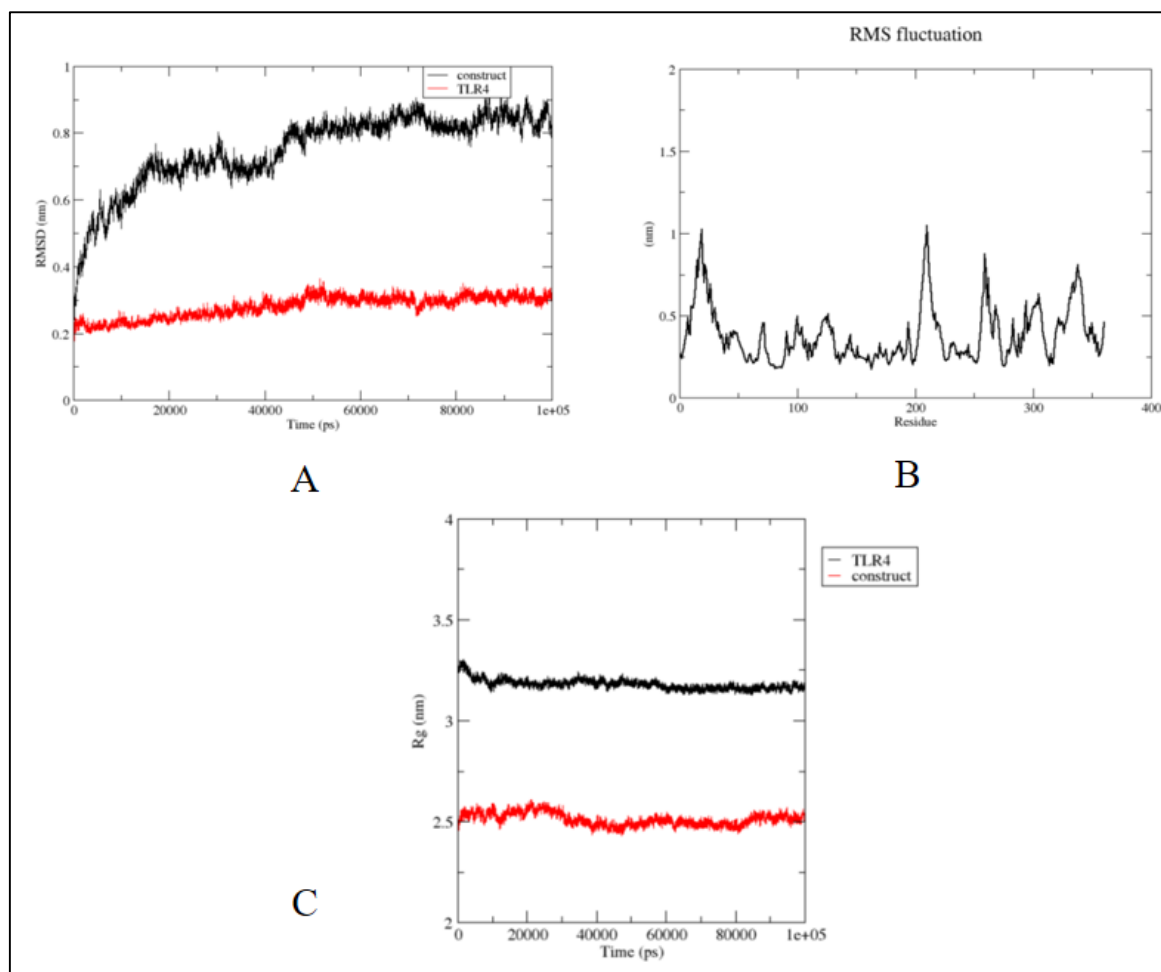


Figure 4.8: (A) Separate RMSD plot for the vaccine construct and TLR4 at 100ns (B) RMSF plot for vaccine construct only at 50ns of MDS (C) Separate Rg plot for TLR4 and vaccine construct at 100ns of MDS.

4.3.9. Immune response profile

Immune profiles of the designed vaccine construct were analyzed using an in-silico immune simulation approach. Immune simulation response results revealed the consistency in the actual immune response of the vaccine construct. A high level of B-cell population was observed upon vaccine construct administration (**Figure 4.9 A**). At the same time, a higher titration value of IgM+IgG and IgM was recorded after 15 days (**Figure 4.9 B**). This indicates induction of secondary and tertiary immune responses. The rise in IFN- γ level was observed after five days of incubation (**Figure 4.9 C**). The constructed vaccine showed an increase in the T-helper cell

population, thus indicating the activation of cell-mediated immune response (**Figure 4.9 D**).

This analysis ensures the clearance of antigens by induction of immune responses.

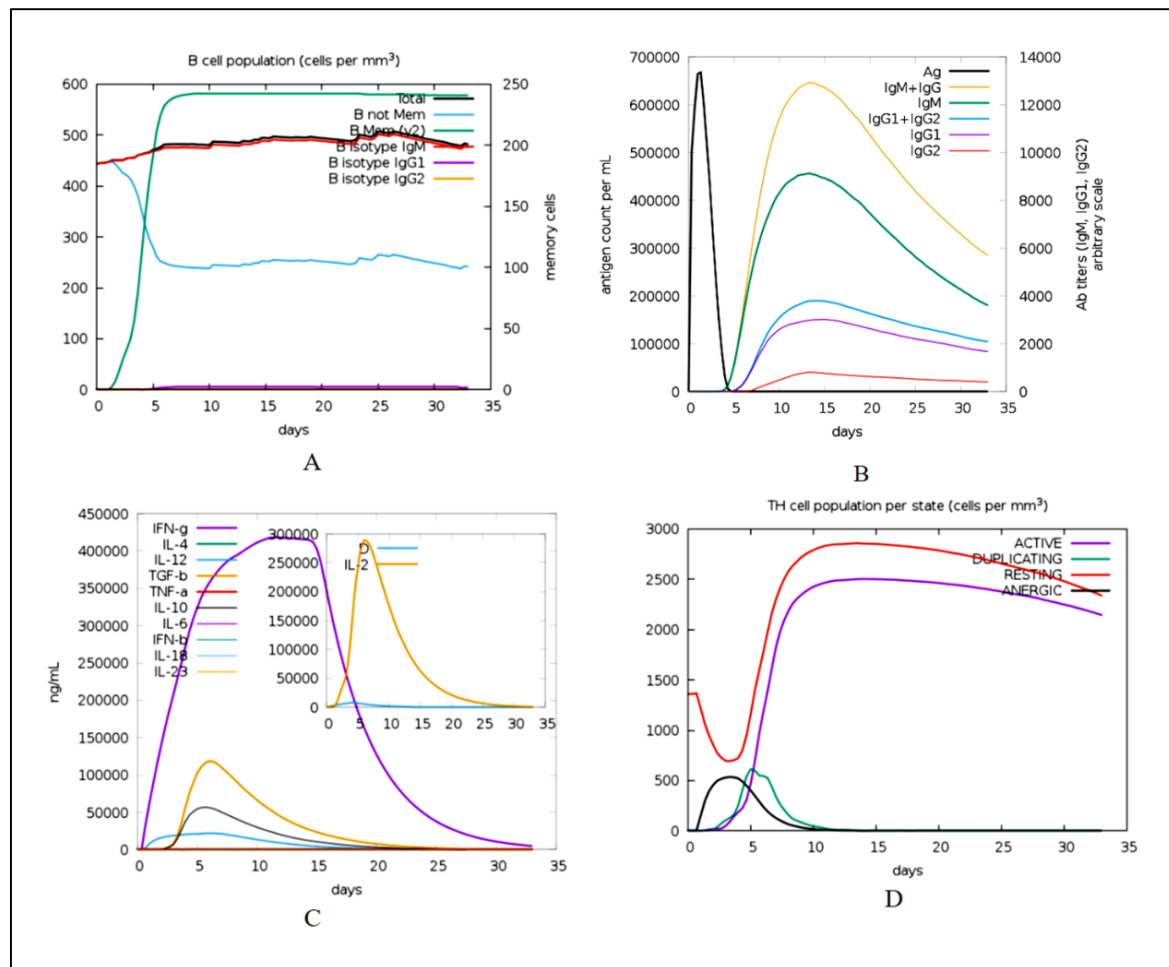


Figure 4.9: Prediction of immune response against the generated chimeric construct (A)

Total B-Cell responses (B) Total antibody responses (C) Interleukins responses (D) Total T-Cell responses.

4.3.10. *In-silico* cloning

The CAI value of the optimized DNA sequence was calculated to be 1, which is an ideal value and increases the chance of expression in *E.coli*. The GC content was also in the optimal range (49.62 %). Expression vector pET28a (+) carrying multi-epitope vaccine is shown in **Figure 4.10**. The final clone consists of a total of 6430bp of DNA, including a vaccine sequence (1080bp).

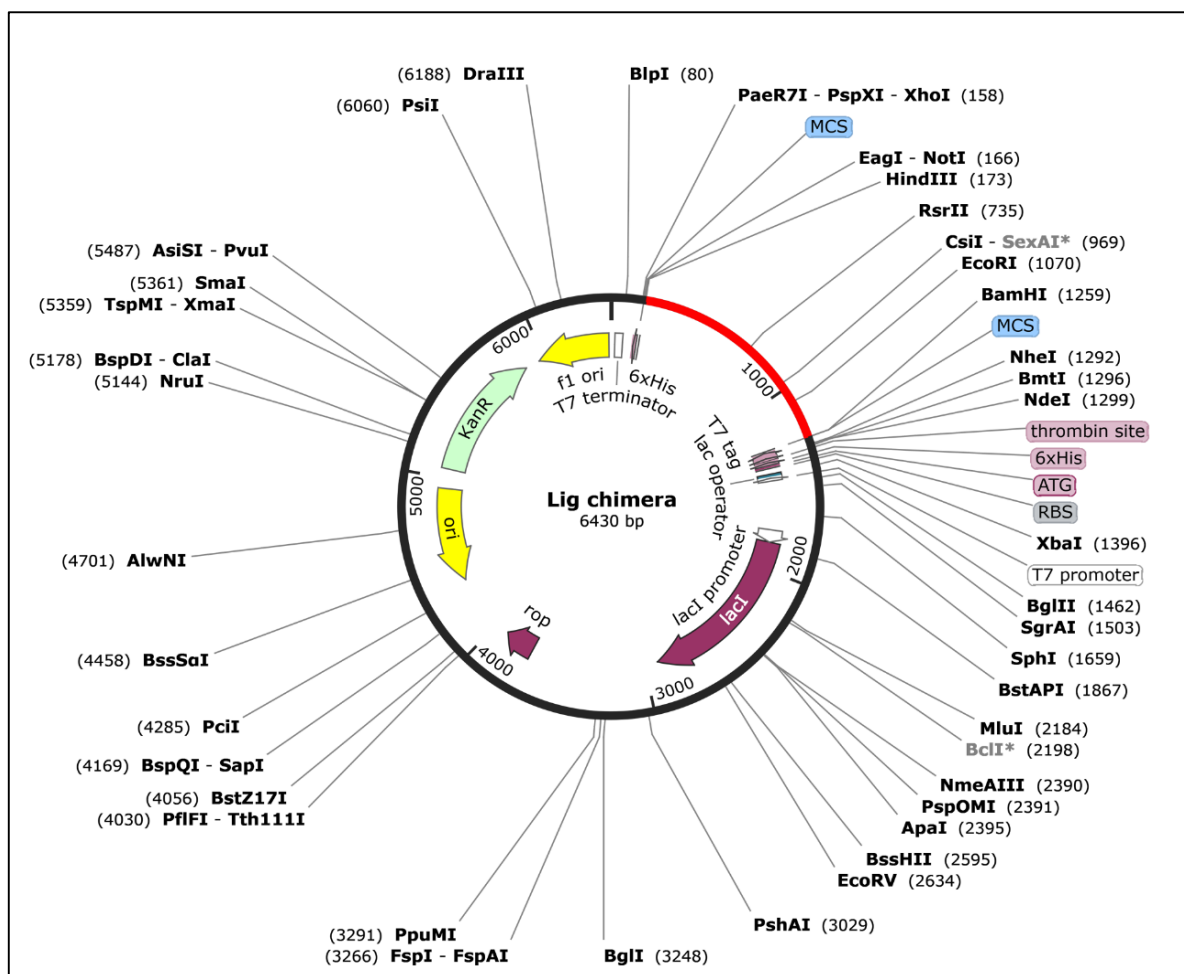


Figure 4.10: Representation of *In-silico* cloning of multi-epitope vaccine construct. Codon optimized gene sequence of the vaccine candidate (represented in red color) was cloned between *Bam*HI and *Hind*III restriction sites of pET-28a (+) expression vector (shown as a black circle).

4.4. Conclusions

Multi-domain outer surface proteins, such as LigA and LigB, have been reported to be vaccine candidates. These two consist of several (12–13) Ig-like domains with an overall size of proteins ranging from approximately 110 to 220 kDa. It has always been challenging to express and purify these proteins for immunological studies. Our immunoinformatic studies have exploited multiple tools to identify the best antigenic epitopes between the two proteins. The common and overlapping epitopes generated from these multiple tools were considered for the design of a multi-epitope chimeric vaccine construct. The use of multi-epitopes has provided confidence in epitope selection. We believe that our multi-epitope chimeric vaccine construct might provide better results both for *in-vitro* and *in-vivo* assays.

Chapter 5

Characterization of novel nuclease and protease activities among Lig proteins

Work of this chapter is published as:

Kumar P., Yung-Fu Chang, Akif M. (2022) Characterization of Novel nuclease and protease activities among Leptospiral immunoglobulin-like proteins. *Arch Biochem Biophys.* 2022 July 09; 109349. <https://doi.org/10.1016/j.abb.2022.109349>

5.1. Introduction

Immunoglobulin (Ig) like domains are widely distributed domains among the proteins and ubiquitously present across the different phyla. Structural features of the Ig-like domain among the various proteins are quite conserved. The domain consists mainly of a β -barrel composed of seven to nine antiparallel β -strands with the typical Greek-key fold and β sandwich topology that displays a structurally conserved core (Halaby et al., 1999). The length and regularity of the β -strands, as well as the length and structure of the connecting loops, are extremely variable. The Ig- like domains are usually grouped into Ig superfamily with having a sequence homology. In contrast, there are many Ig-like domains in proteins that bear no sequence homology also. The presence of Ig-like domains has been reported in a large number of proteins with diverse biological functions. They are found in proteins involved in cell-cell recognition, cell-surface receptors, muscle structure protein, and the immune system (Chothia, 1953). Ig-like proteins are also reported in bacterial species and are known as Bacterial Ig-(Big) domain frequently found on cell surface proteins. The Big domain plays a vital role in the adhesion of bacteria to the host cell and helps in an invasion of pathogenic strains. Intimin and invasin family of outer membrane (OM) adhesins from *E. coli* and *Yersinia*, respectively, containing Ig-like domains, have been studied in detail (Fairman et al., 2012; Seo et al., 2012). Apart from the structural component of adhesion, Ig-like domains in PapD protein play a chaperone function in the bacterial periplasm to help in pilus assembly. Moreover, the Ig-like domains are found in a few oxidoreductases family of proteins, sugar-binding proteins, and a few transcription factors, as well as in hydrolyzing enzymes (Bork et al., 1994; Holmgren and Bränden, 1989).

Bacterial Ig-like domains/proteins are also found in pathogenic *Leptospiral* species named Immunoglobulin-like (Lig) proteins. The Lig proteins (LigA and LigB) are primarily located on the surface of the pathogenic *Leptospira* and are reported as a potential virulent factor as

demonstrated by various mutational studies (Christopher J. Pappas, 2015; Fernandes et al., 2022). The role of many of these virulent factors has been investigated as vaccine candidates, as well as their roles in interaction with the host proteins (Castiblanco-Valencia et al., 2012b; Y. P. Lin et al., 2009). Moreover, the solution structure of one of the terminal domains, LigB12, suggests that the Ig-like domain consists of mainly β -strands that form Ig-like fold (Christopher P. Ptak et al., 2014). Other Ig-like domains from LigA and LigB proteins have relatively similar folds but are found to have electrostatically different (Christopher P. Ptak et al., 2014).

Additionally, individual domains from each Lig family protein display different thermostability (Ptak et al., 2019). The Ig-like domains of Lig proteins are well characterized in interaction with host extracellular matrix (ECM) components such as fibronectin, fibrinogen, tropoelastin, and laminin (Henry A. Choy et al., 2007; Y.-P. Lin et al., 2009; Lin et al., 2010). The role of LigA and LigB has also been demonstrated in serum resistance (Fernandes et al., 2021). These roles and functions attribute Lig proteins as multifaceted proteins. However, the hydrolyzing function of Lig proteins has not been reported yet.

In this chapter, we report a novel *in-vitro* nuclease activity in the Ig-like domain of the Lig family of proteins. All Ig-like domains were able to cleave the DNA in the presence of a divalent ion, but not RNA. Site-directed mutagenesis revealed Mg^{+2} binding residues in the Ig-like domain of LigA7. The basis of nuclease activity in the protein may be associated with adopting different conformations in the presence of divalent ions and substrate, as investigated by our fluorescence spectroscopy data. The electrostatic surface on each domain indicated the positive charge region contributed by either lysine or arginine residues may provide a DNA interaction site on the Ig-like domain. Docking of a stretch of double-strand DNA, shown to bind on the positive surface of the protein. In addition, the said protein is also found to have the ability to cleave available protease substrate, β -casein, in our experimental condition. This is the first report of *in-vitro* nuclease and protease activities in Ig-like domains and Lig proteins.

Our study contributes toward the characterization of novel functions among the Lig family of proteins.

5.2. Materials

Necessary chemicals and reagents have been described in chapter 2 of this thesis. The plasmid DNA for nicking studies was isolated using either plasmid isolation kits from Qiagen or following conventional mini prep plasmid isolation protocol. Magnesium chloride was purchased from Sigma-Aldrich. Single-stranded Φ X174 virion DNA was procured from NEB. Un-methylated DNA was prepared by polymerase chain reaction (PCR) using Pfu polymerase from NEB. DNase (ML068) and RNase (MB087) were purchased from Himedia. Tris used for preparing Tris-HCL buffer was purchased from BioRad. The universal protease substrate β -casein was purchased from Sigma-Aldrich (C6905). The EDTA was purchased from Amresco, and 0.5M stock was prepared in autoclaved double distilled water. DpnI enzyme to cleave the methylated DNA was purchased from NEB. DNA marker was purchased from Thermo scientific and protein ladder BioRad. The polymerase for generating the mutants was purchased from Takara.

5.3. Experimental Procedures

5.3.1. Nuclease activity assay

To monitor nuclease/DNase activity in the Lig protein containing single as well as multiple Ig-like domains, a supercoiled plasmid DNA (pET28a) in a reaction mixture containing 50mM Tris-HCl (pH 8) and 10mM $MgCl_2$ was incubated with different concentrations (1 μ M- 10 μ M) of purified recombinant proteins (mentioned in chapter 2). The reaction sample was incubated at 37° C for 10 min, and the reaction was stopped by adding a sample buffer containing 10mM EDTA. Plasmid cleavage, as well as nicking, were analyzed on 1% agarose gels. Single-stranded RNA isolated from human macrophage (Provided by Prof. G. Ravi Gutti, Department

of Biochemistry, UoH) was used to check the protein's ability to cleave RNA by following the same protocol. RNA cleavage activity was monitored on 2% agarose gel. Nuclease/DNase activity was also checked under varying pH conditions ranging from 4 to 8.5. Cleavage activity with the single-stranded DNA (ssDNA) was checked using Φ X174 virion DNA as a substrate. RNase and DNase were used as positive control.

5.3.2. Effect of divalent ions on nuclease activity

Most of the known nuclease required magnesium as a co-factor. The divalent metal-ion dependence for nuclease activity was determined by substituting Mg^{2+} ions with other metal ions like Mn^{2+} , Ca^{2+} , Fe^{2+} , Co^{2+} , Ni^{2+} , Cu^{2+} , and Zn^{2+} in the reaction mixture. The reaction sample was incubated at 37° C for 10 min, and the reaction was stopped by adding a sample buffer containing 10mM EDTA. Cleavage of supercoiled plasmid DNA was monitored on the agarose gel.

5.3.3. Intrinsic fluorescence measurement

The binding of Mg^{2+} with the LigA7 protein was monitored using JASCO (Model No. FP-8500) fluorescence spectrophotometer. The purified LigA7 protein (5 μ M) in PBS buffer pH 7.4 was excited at 280 nm, and the emission spectra were recorded between 300-400nm at 25 °C. An increasing concentration of $MgCl_2$ (2.5-7.5 μ M) was titrated against the same concentration of LigA7 protein. The data collected was plotted using excel.

5.3.4. Site-directed mutagenesis

The potential Mg^{2+} ion-binding residues among individual LigA domains were selected and their conservation was analyzed in all variable Ig-like domains of LigA. In order to check the involvement of these residues in Mg^{2+} dependent nuclease activity, mutation of these was performed in the LigA7 gene by site-directed mutagenesis. The codon of conserved Mg^{2+} ion-binding residues was replaced with the alanine codon using mutagenesis primers (**Table 5.1**)

and generated various *ligA* mutant clones in the pET28a-SUMO vector (**Table 5.2**). Variants were generated by performing PCR with primers containing the desired substitution mutations.

Table 5.1: List of primers used to generate substitution mutants in *ligA7*.

Construct	Sequence
pET28a/ <i>ligA7</i> T22A	F: 5' GTAGCCTGAATCTTTCTGCGAGTCCTTTTGCTTTGGAA 3' R: 5' TTCAAAGCAAAAGGACTCGCAGAAAGATTCAAGGCTAC 3'
pET28a/ <i>ligA7</i> T28A	F: 5' TTGTCCGTAAAGATACCAGCAGCCTTGAATCTTTCTGTG 3' R: 5' CACAGAAAGATTCAAGGCTGCTGGTATCTTTACGGACAA 3'
pET28a/ <i>ligA7</i> D33A	F: 5' GGCTACTGGTATCTTTACGGCCAACTCAAATTCCGATATTA 3' R: 5' TAATATCGGAATTTGAGTTGGCCGTAAAGATACCAGTAGCC 3'
pET28a/ <i>ligA7</i> D38A	F: 5' TCCAAGTAACTTGATTGTAAATAGCGGAATTTGAGTTGTCCGTAAAG 3' R: 5' CTTTACGGACAACCTCAAATTCCGCTATTACAAATCAAGTTACTTGGA 3'
pET28a/ <i>ligA7</i> T40A	F: 5' GAATTCCAAGTAACTTGATTTGCAATATCGGAATTTGAGTTGTCCG R: 5' CGGACAACCTCAAATTCCGATATTGCAAATCAAGTTACTTGGAATTC 3'
pET28a/ <i>ligA7</i> T41A	F: 5' AGGAATTCCAAGTAACTTGAGCTGTAATATCGGAATTTGAGTTGTCCGTAA 3' R: 5' TTACGGACAACCTCAAATTCCGATATTACAGCTCAAGTTACTTGGAATTCCT 3'
pET28a/ <i>ligA7</i> T44A	F: 5' ATCCGTATTAGAGGAATTCCAAGCAACTTGATTTGTAATATCGGAAT 3' R: 5' ATTCCGATATTACAAATCAAGTTGCTTGGAATTCCTCTAATACGGAT 3'
pET28a/ <i>ligA7</i> N49A	F: 5' GTTGGAACGCGTAAGAATATCCGTAGCAGAGGAATTCCAAGTAACTTGATTT 3' R: 5' TCCGCGTTTGGCGTTTGTGGCGGAAACGGTAAGAATATCC 3'
pET28a/ <i>ligA7</i> N57A	F: 5' TTCTTACCGTTTCCGCCACAAACGCCAAAC 3' R: 5' GTTTGGCGTTTGTGGCGGAAACGGTAAGAA 3'
pET28a/ <i>ligA7</i> E23A	F: 5' GCAAAAGGACTCACAGCAAGATTCAAGGCTACT 3' R: 5' AGTAGCCTTGAATCTTGCTGTGAGTCCTTTTGC 3'
pET28a/ <i>ligA7</i> S47A	F: 5' CAAGTTACTTGGAATGCCTCTAATACGGATATT 3' R: 5' AATATCCGTATTAGAGGCATTCCAAGTAACTTG 3'

The PCR product was treated with DpnI and incubated at 37 ° C for 1 hour. DpnI treatment removes mutation-free parent DNA. Samples treated with DpnI were then transformed into XL-Blue competent cells and plated on LB agar plates containing kanamycin antibiotics. Single colonies from the transforming plate were used to keep primary cultures at 37 ° C, overnight in a shaking incubator. The plasmid was isolated from the primary culture grown overnight using the plasmid isolation kit. The isolated plasmid was further sent for sequencing. The mutations were confirmed by sequencing, and recombinant mutant LigA7 proteins were purified with the same protocol as used for the wild-type LigA7.

Table 5.2: List of substitution mutants generated in *ligA7*.

Mutant plasmids/clones	
pET28a/ligA7T22A	pET28a-SUMO bearing <i>ligA7</i> T22A mutation
pET28a/ligA7T28A	pET28a-SUMO bearing <i>ligA7</i> T28A mutation
pET28a/ligA7D33A	pET28a-SUMO bearing <i>ligA7</i> D33A mutation
pET28a/ligA7D38A	pET28a-SUMO bearing <i>ligA7</i> D38A mutation
pET28a/ligA7T40A	pET28a-SUMO bearing <i>ligA7</i> T40A mutation
pET28a/ligA7T41A	pET28a-SUMO bearing <i>ligA7</i> T41A mutation
pET28a/ligA7T44A	pET28a-SUMO bearing <i>ligA7</i> T44A mutation
pET28a/ligA7N49A	pET28a-SUMO bearing <i>ligA7</i> N49A mutation
pET28a/ligA7N57A	pET28a-SUMO bearing <i>ligA7</i> N57A mutation
pET28a/ligA7E23A	pET28a-SUMO bearing <i>ligA7</i> E23A mutation
pET28a/ligA7S47A	pET28a-SUMO bearing <i>ligA7</i> S47A mutation

5.3.5. Purification of the mutant protein

All the generated mutants of LigA were expressed and purified using the protocol mentioned on page numbers 33-50 of Chapter 2.

5.3.6. Homology modeling and molecular docking

The homologous model of all single domains of variable LigA was generated using the online server I-TASSER (<https://zhanggroup.org/I-TASSER/>) and following the strategy published before (P. Kumar et al., 2021; Yang et al., 2015). The quality of the predicted model was verified by testing the Ramachandran stats and Z-score at the PROCHECK server (<https://saves.mbi.ucla.edu/>) (Laskowski et al., 1993) and the ProSA-web server (<https://prosa.services.came.sbg.ac.at/prosa.php>) (Wiederstein and Sippl, 2007), respectively. In order to check the interaction of LigA7 with DNA, molecular docking was performed using the HADDOCK server (<https://wenmr.science.uu.nl/haddock2.4/>) (De Vries et al., 2010). The model of a 19bp double-stranded DNA stretch (5'-CCGGAGGACAGTCCTCCGG-3') was taken from a DNA-protein complex reported in a protein data bank (<https://www.rcsb.org/>). The calculation of the electrostatic potential surface of the LigA7 modeled structure was done by the PyMOL tool.

5.3.7. Protease assay

To monitor the protease activity, a universal protease substrate, β -casein (Sigma cat no - C6905) was used. Purified proteins (10 and 15 μ g) were incubated with casein (5 μ g) at 37° C for 12 and 24hrs in a reaction condition containing 20mM Tris 100mM NaCl, pH 8.0. The β -casein cleavage was checked on 15% SDS-PAGE.

5.3.8. Mass spectrometry analysis of β -casein cleavage product

The proteolytic cleavage products of β -casein and the LigA7 protein bands were excised-out from the SDS-PAGE and sent for the MALDI-MS/MS analysis to the Sandor Specialty Diagnostics Pvt. Ltd, Hyderabad.

5.4. Results

5.4.1. Ig-like domains of LigA possess novel metal-dependent nuclease activity

The LigA containing one and more than one variable Ig-like domain were proficiently cleaving double-stranded supercoiled plasmid DNA in the presence of Mg^{2+} ion. The extent of DNA cleavage varies from nick formation to complete DNA degradation. High concentrations of LigA7 displayed complete degradation of the plasmid DNA in a fixed concentration of Mg^{2+} ion. The cleavage effect was minimal as the concentration of LigA was decreased in the reaction (Figure 5.1). The cleavage product was not observed when only purified proteins or only Mg^{2+} ions were incubated with the same plasmid DNA.

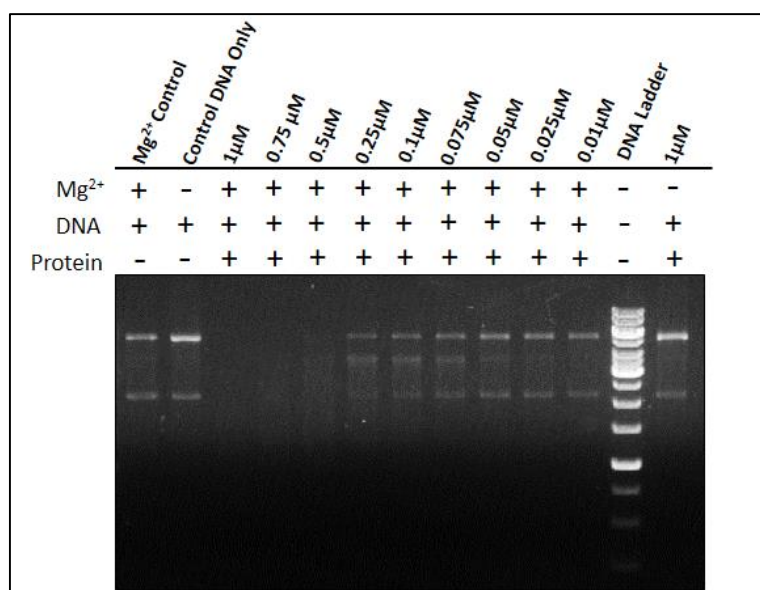


Figure 5.1: DNA cleavage activity in LigA7 on supercoiled double-standard DNA (pET28a). DNA cleavage activity in the presence of Mg^{2+} ions with different concentrations of protein was monitored on 1% agarose gel. Only DNA and protein were used as control.

Further, a decreased DNA cleavage activity was observed when EDTA was added to the reaction mixture (Figure 5.2).

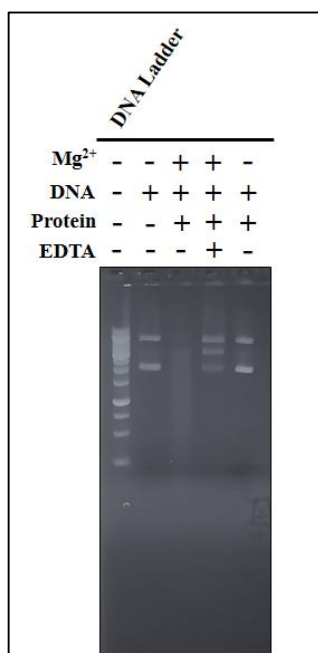


Figure 5.2: Agarose gel showing metal-dependent DNA cleavage activity of LigA7 in the presence of both Mg²⁺ ions and EDTA

Moreover, DNA cleavage activity was reduced with the boiled protein (**Figure 5.3**). The complete activity was abolished when protein was boiled with SDS (**Figure 5.4**).

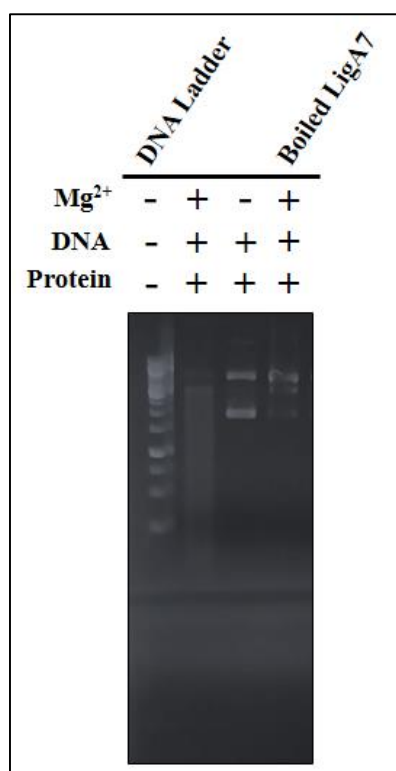


Figure 5.3: Agarose gel showing metal-dependent DNA cleavage activity of boiled LigA7

In addition, similar activity as LigA7 was displayed by the other purified LigA variable single domains such as LigA8, LigA9, LigA10, LigA11, LigA12, and LigA13. Since Lig proteins are reported to be calcium binding proteins, the activity was performed in the presence of Ca²⁺ as well (Figure 5.5).

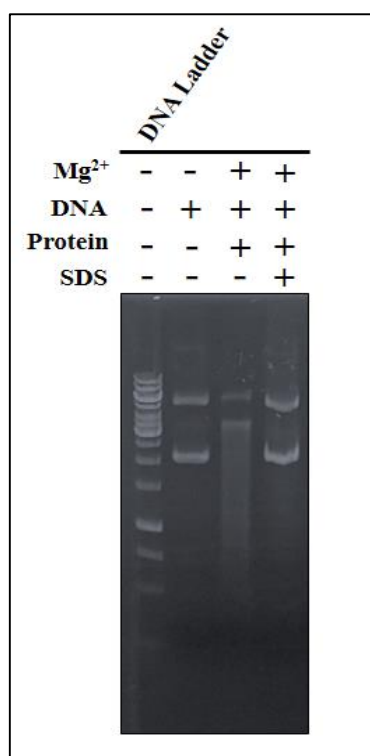


Figure 5.4: Agarose gel showing metal-dependent DNA cleavage activity of boiled LigA7 in the presence of SDS.

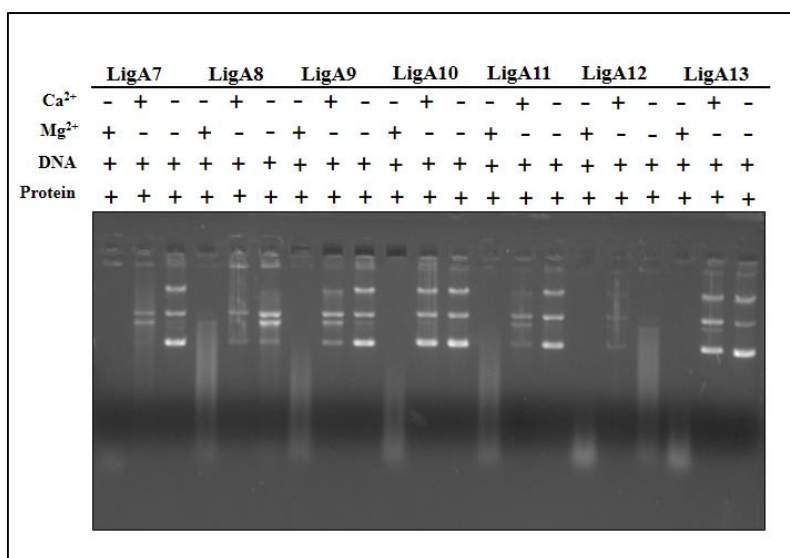


Figure 5.5: Agarose gel showing metal-dependent DNA cleavage activity of different variable single domains of LigA.

In addition, the nuclease activity with the protein containing multiple domains such as LigA12-13 (**Figure 5.6**), LigA11-13 (**Figure 5.7**), LigA10-13 (**Figure 5.8**), LigA9-13 (**Figure 5.9**) and LigA8-13 (**Figure 5.10**) was performed. This suggests that LigA7 and other variable domains of LigA protein possess Mg^{2+} ion-dependent nuclease activity. To our surprise, the Ig-like domain from the LigA protein did not display any homology with any known nuclease domains. This is the first report that demonstrated this novel nuclease activity within the LigA domains to the best of our knowledge.

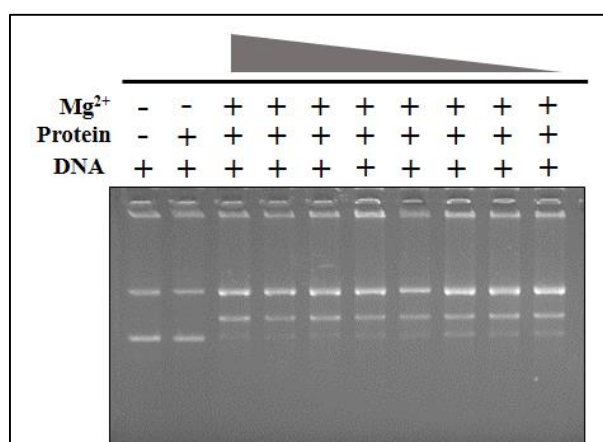


Figure 5.6: Agarose gel showing metal-dependent DNA cleavage activity of LigA12-13

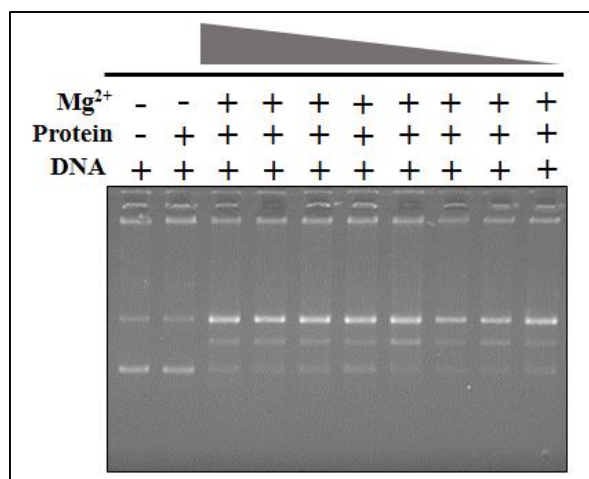


Figure 5.7: Agarose gel showing metal-dependent DNA cleavage activity LigA11-13

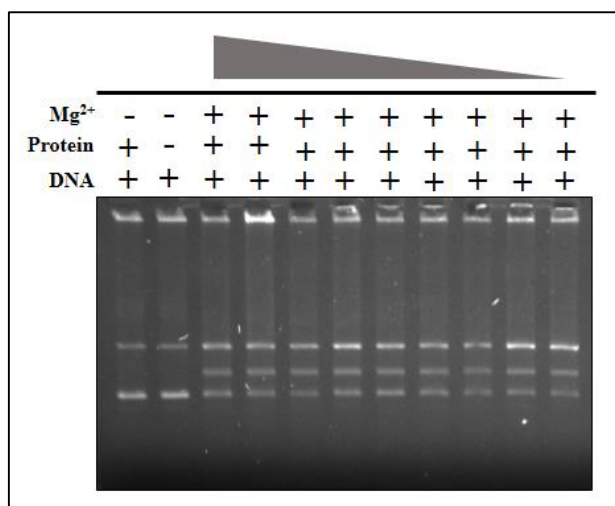


Figure 5.8: Agarose gel showing metal-dependent DNA cleavage activity LigA10-13.

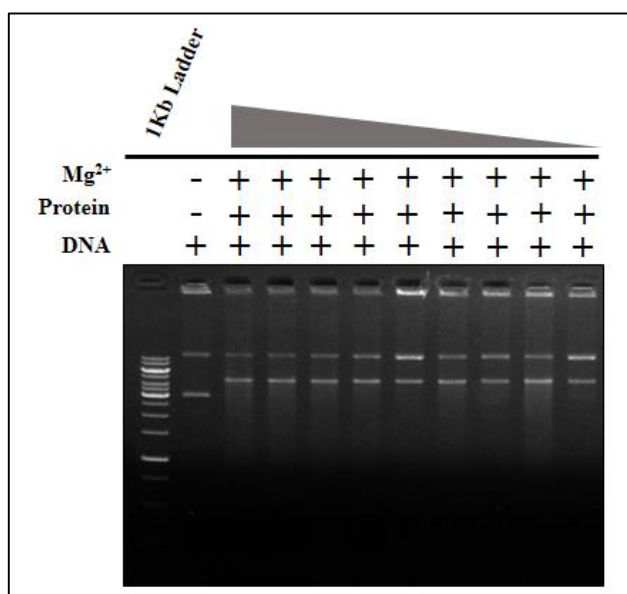


Figure 5.9: Agarose gel showing metal-dependent DNA cleavage activity LigA9-13.

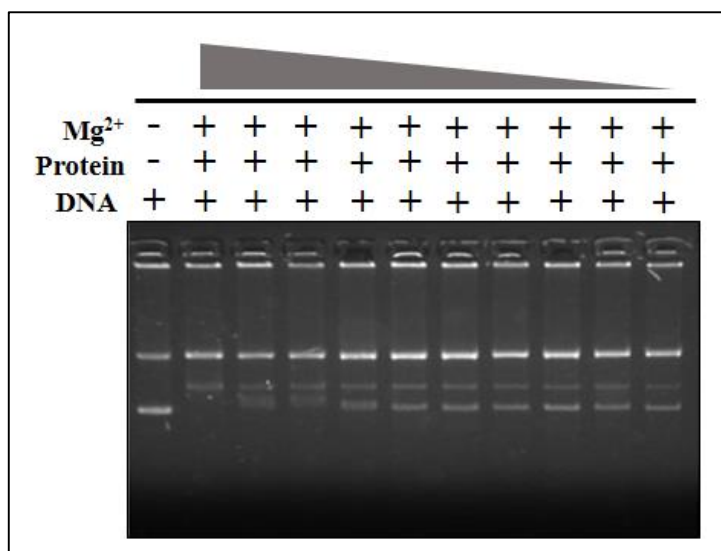


Figure 5.10: Agarose gel showing metal-dependent DNA cleavage activity LigA8-13.

5.4.2. Preference of divalent ions on the LigA nuclease activity

Since the LigA7 showed nuclease activity in the presence of Mg²⁺ ions, the preference of LigA7 for other divalent ions on the nuclease activity was also investigated. Similar to Mg²⁺ ions concentration, MnCl₂, CaCl₂, FeCl₂, CoCl₂, NiCl₂, CuCl₂, and ZnCl₂ were added with the purified proteins in the reaction sample. Interestingly, the LigA7 showed complete degradation of plasmid DNA in the presence of MgCl₂, MnCl₂, and CoCl₂ to an almost similar extent, whereas in the presence of other divalent ions, the LigA7 displayed less or no activity (**Figure 5.11**).

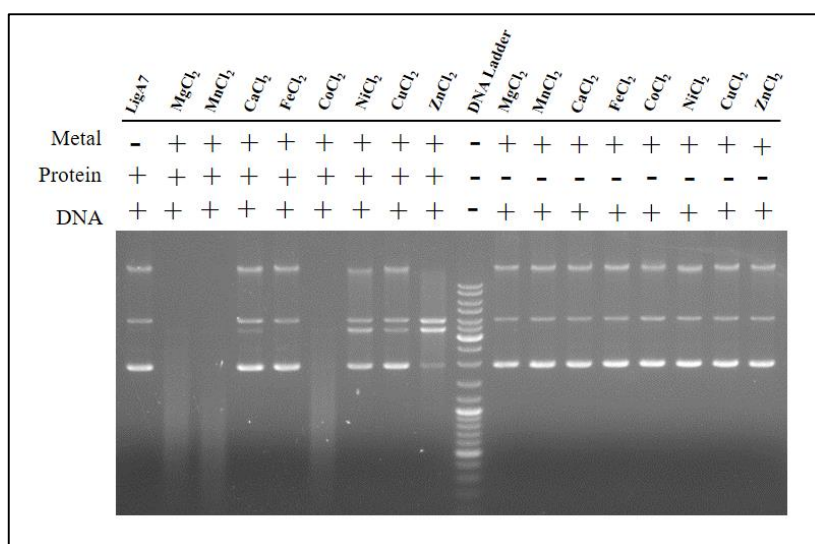


Figure 5.11: 1% agarose gel showing DNA cleavage activity of LigA7 on circular double-strand DNA in the presence of different divalent ions.

Moreover, the quantification of DNA on the agarose gel in the presence and absence of divalent ions revealed the maximum activity of LigA7 in the presence of MgCl₂, suggesting LigA7 has mainly preference towards the Mg²⁺ ion for this activity. Henceforth, all the other reactions were carried out further in the presence of Mg²⁺ ions.

The effect of pH on the nuclease activity was also investigated. The LigA7 is observed to show a partial degradation of the plasmid DNA in the presence of a similar concentration of Mg²⁺ ions at pH 4.0. However, the activity was high and displayed a complete degradation at pH 8.5 of the reaction mix (**Figure 5.12**). Similar to the LigA7, multiple domains from LigA exhibited the same observation at different pH. This observation suggested that all the individual domains as well as a combination of multi-domains possess better nuclease activity at pH 8.5 (**Figure 5.12**).

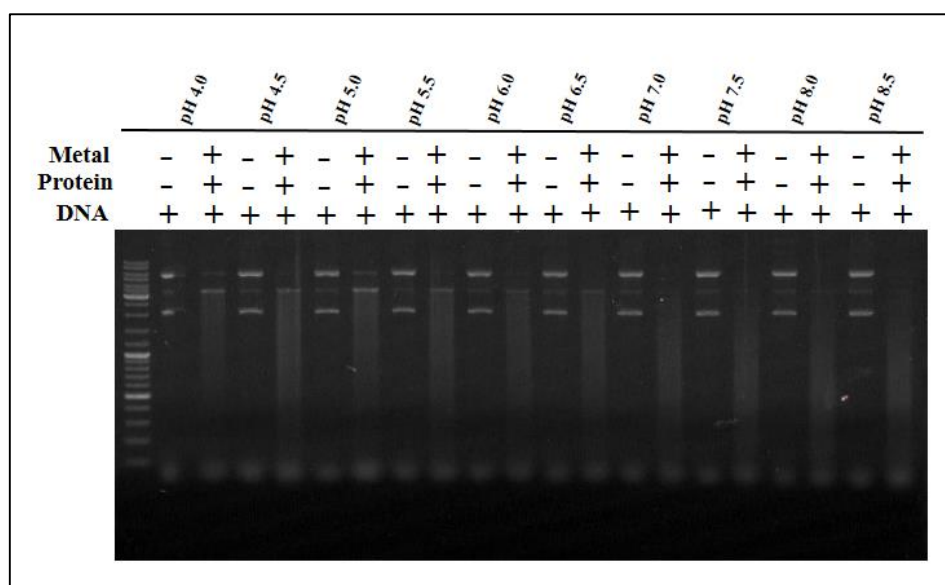


Figure 5.12: Representation of DNase activity in LigA7 with pH variation.

5.4.3. LigA7 nuclease activity on single-stranded DNA and RNA

Whether the metal-dependent nuclease activity exhibited by LigA7 has any substrate preference towards single or double-stranded DNA, we observed a complete degradation of single-stranded Φ X174 virion DNA by LigA7 in the presence of metal ions (**Figure 5.13**). This result suggested that the LigA7 can degrade both types of DNA, and no specific preference was displayed in this context.

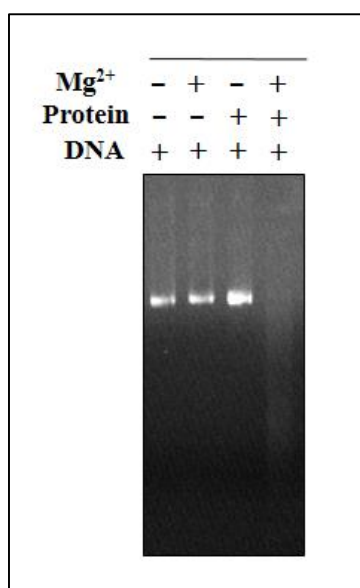


Figure 5.13: Agarose gel showing metal ion depended on DNA degradation of ssDNA by LigA7.

Moreover, LigA7 was also observed to cleave the PCR amplified product, which was considered an unmethylated DNA (**Figure 5.14**).



Figure 5.14: Agarose gel showing metal ion-dependent DNA degradation of Un-methylated DNA by LigA7.

However, the LigA7 could not display any activity to degrade RNA in the presence of divalent ions. The same activity was checked at both lower and higher concentrations of LigA7 (**Figure 4**). In both conditions, no activity was noticed. This suggests that the LigA7 prefers only DNA and not RNA for the cleavage activity. The reason for this preference is not known at this moment but opens up an avenue to investigate in this context. This observation could be concluded that LigA7 acts as DNase instead of nucleases.

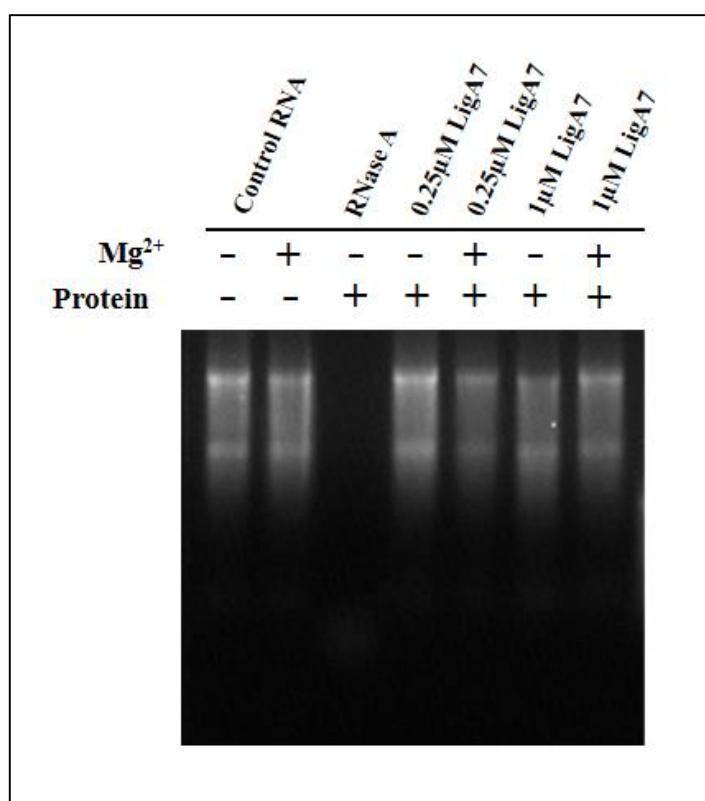


Figure 5.15: Metal depended DNA cleavage activity of LigA7 on RNA monitored on 2% agarose gel. RNase A was used as a positive control. Two different concentrations of protein were used with the fixed concentration of Mg²⁺ and RNA

5.4.4. Binding of Mg^{2+} ion with LigA7

Given our observation of the involvement of the Mg^{2+} ion in the DNase activity of LigA7, we asked whether the Mg^{2+} ion binds to the LigA7. To address this question, we performed a fluorescence titrations experiment wherein quenching of LigA7 intrinsic tryptophan emissions was monitored with an increasing concentration of Mg^{2+} ion ($MgCl_2$). **Figure 5.16** presents representative intrinsic tryptophan emissions spectra. Emissions were observed at a maximum wavelength of ~ 317 nm in the absence of $MgCl_2$ and underwent significant quenching as the $MgCl_2$ concentration was increased. Mg^{2+} ion-dependent quenching of LigA7 emissions spectra results in a concurrent decrease in the intrinsic fluorescence intensity, suggesting the binding of Mg^{2+} ions with the protein residues.

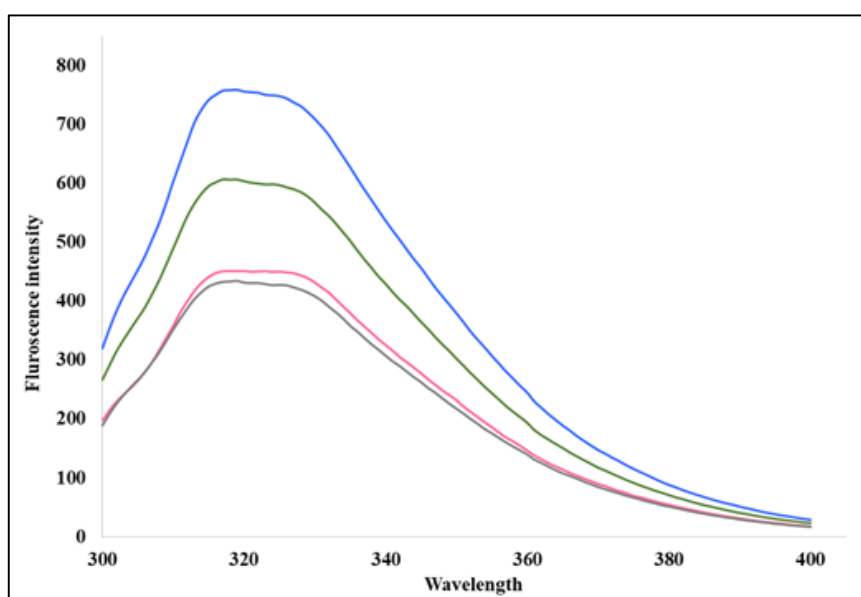


Figure 5.16: Fluorescence measurements of LigA7. Intrinsic fluorescence profile shows emission maxima at ~ 317 nm using $5 \mu M$ protein with varying concentrations of Mg^{2+} ion $0 \mu M$, $2.5 \mu M$, $5 \mu M$, and $7.5 \mu M$ are displayed as blue, green, magenta, and grey curves, respectively.

5.4.5. Effect of LigA7 point mutations on the nuclease activity

In order to investigate the residues which bind the Mg^{2+} ions, we have predicted the possible Mg^{2+} binding residues in LigA7. Many potential residues, such as T22, D33, D38, T41, T44, N49, and N57, were conserved among other variable Ig-like domains of LigA and depicted as stars in the sequence alignment (**Figure 5.17**).

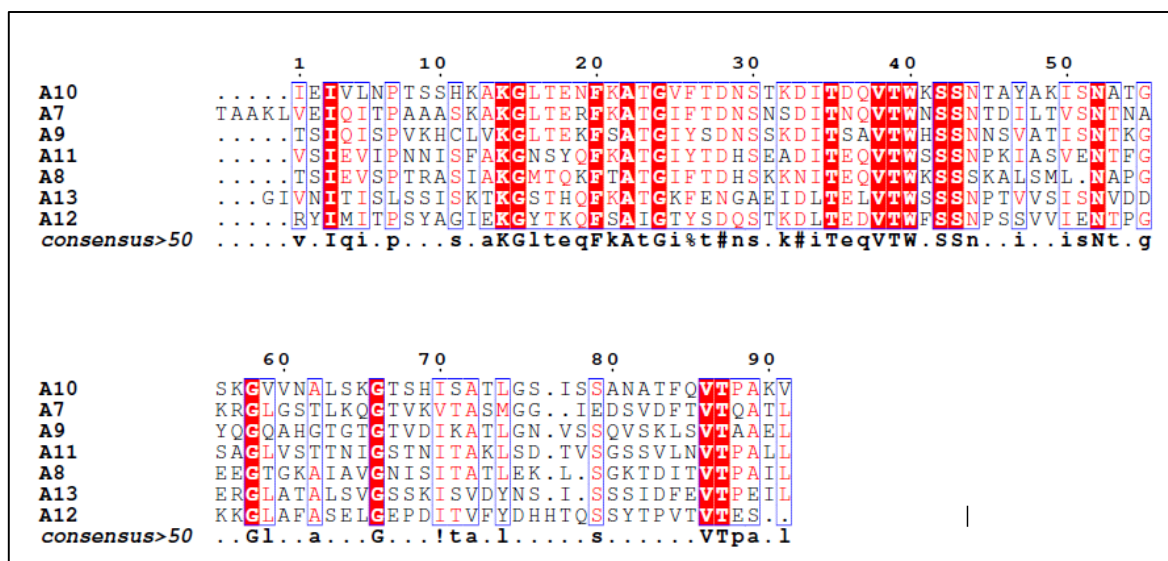


Figure 5.17: Multiple sequence alignment of all single Ig-like domains from the variable region of LigA. Star signs below the sequence denote conserved probable Mg^{2+} ion-binding residues.

The point mutations of LigA7 were generated by replacing these residues with Ala to examine individual residue's contribution toward DNA cleavage activity. All the plasmids bearing point mutations were transformed into expression cells. The protein was overexpressed and purified using the protocol mentioned in Chapter 2. All the mutants showed good expression and purified in a good amount. The purification profile of all the mutants is shown in **Figure 5.18**.

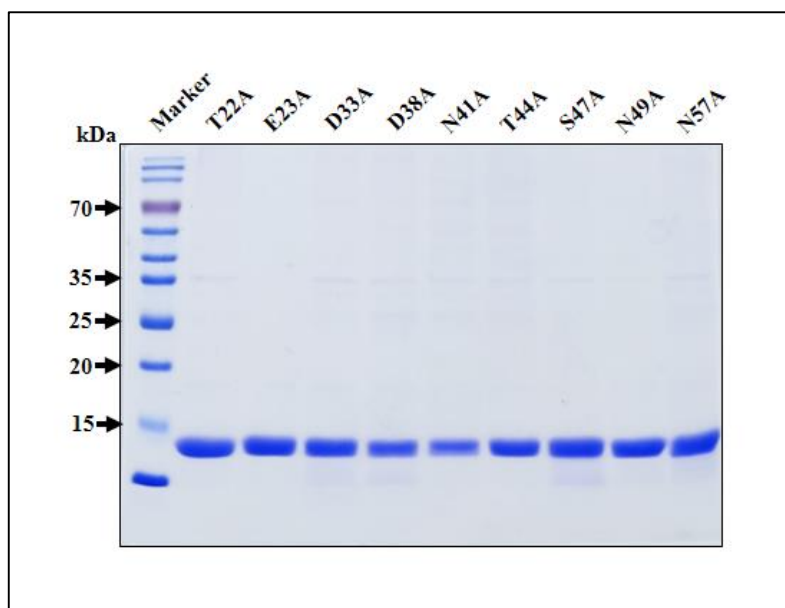


Figure 5.18: 12% SDS-PAGE showing purification of all the generated mutants.

Mutants D38A and N41A of LigA7 displayed Mg^{2+} ion-dependent DNA cleavage activity. Other mutants are observed to have no or less activity in our condition (**Figure 5.19**). The mutational analysis suggests that the residue T22, D33, T44, N49, and N57 from LigA7 binds with the Mg^{2+} ion and contributes to the activity.

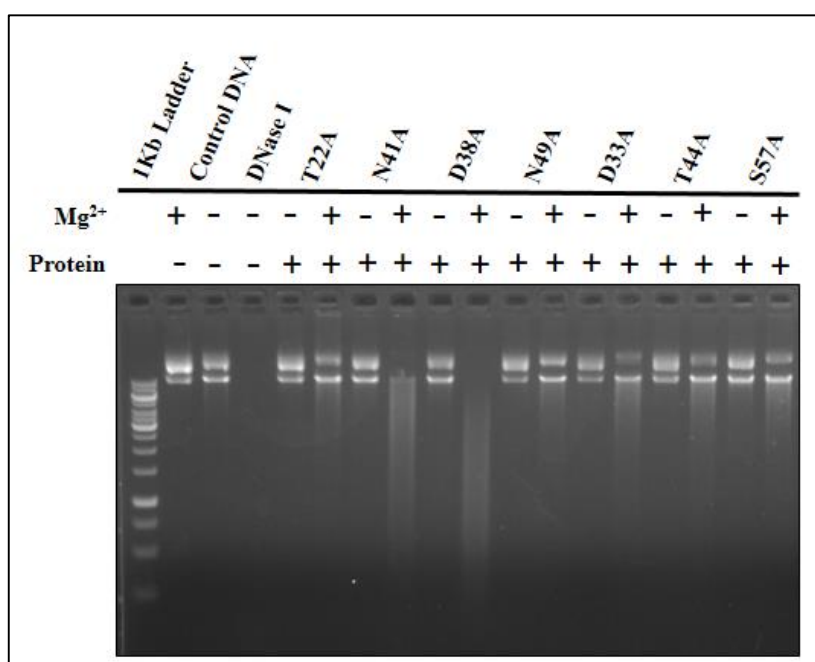


Figure 5.19: Effect of Mg^{2+} ions on LigA7 mutant proteins for the DNA cleavage activity.

5.4.6. Molecular docking study reveals potential DNA binding surface

In order to map potential DNA binding surfaces of the LigA7, we calculated the electrostatic surface charge potential for LigA7. Since the LigA7 structure was unavailable, a 3-D model of LigA7 was generated using the LigB12 NMR structure as a template. The homologous model of LigA7 was validated by checking the Ramachandran plot and overall Z-score of the model (**Figure 5.20**). The energy-minimized structure of LigA7 was used for molecular docking with a 19 bp double-stranded DNA stretched with the HADDOCK server. The best-docked complex having the highest energy of -315 kJ/mol was considered to examine the interactions with the DNA stretch. The calculated electrostatic potential surface charge showed a positively charged patch on LigA7 (**Figure 5.21**).

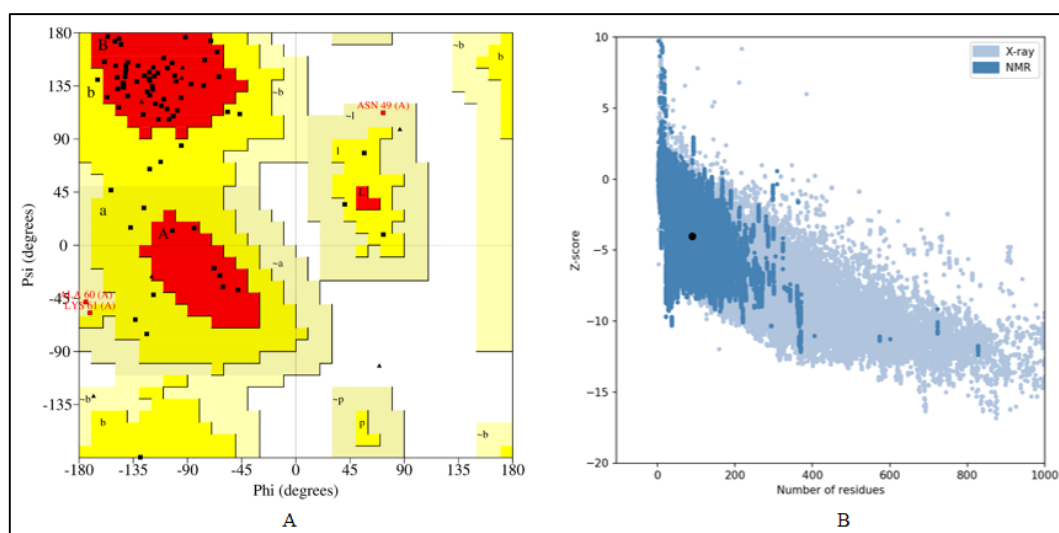


Figure 5.20: Validation of homologous model of LigA7 (A) Ramachandran plot (B) overall Z-score.

The positive residues involved in the formation of polar interaction with the negatively charged DNA strand include mostly basic residues such as Gln7, Lys15, Lys17, Thr20, Arg22, Phe23, Asn34, Lys59, Arg60, and Ser64. This region is the most plausible binding site for the negatively charged DNA, facilitating the DNase activity of the protein.

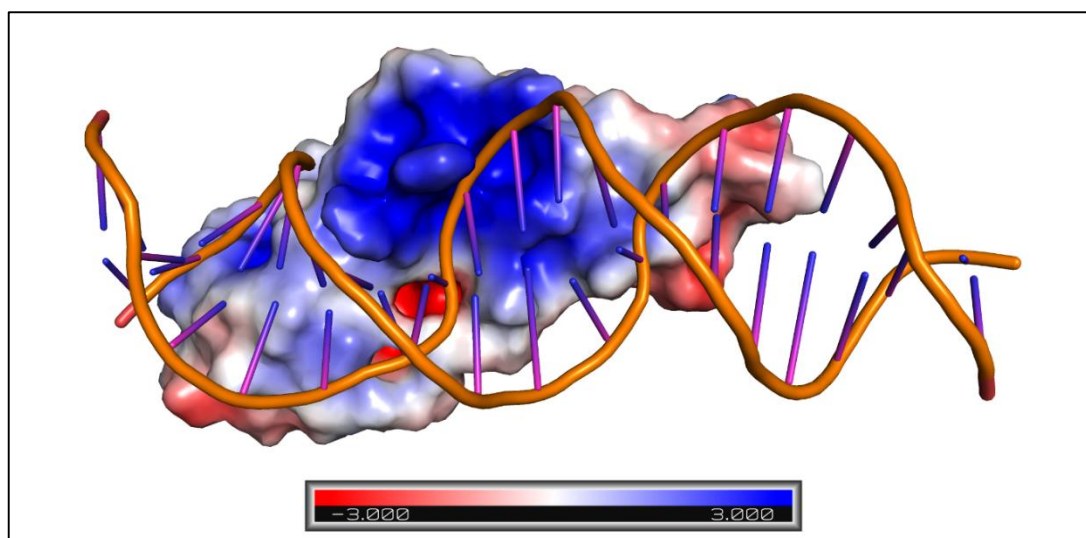


Figure 5.21: Representation of LigA7 and DNA complex. Eight nucleotides length of double-strand was docked on LigA7 modeled structure. LigA7 model is shown as surface electrostatic charge potential, and the DNA double-strand is represented as a cartoon.

5.4.7. LigA7 also possesses novel protease activity

The Lig family of proteins is known as multifaceted proteins. So far, our study has established its novel DNase activity. In order to probe other functions of this protein, we have subjected the sequence of LigA7 to the MoonProt 2.0 (<http://moonlightingproteins.org>) (Mani et al., 2015a) and ProtIdent server (<http://www.csbio.sjtu.edu.cn/bioinf/Protease/>) (Chou and Shen, 2008). The MoonProt server predicts moonlighting proteins that contain more than one distinct function in a polypeptide. The MoonProt predicted peptidase function in the LigA7. In order to confirm this predicted function, we have examined the degradation of β -casein by the LigA7 in *in-vitro* conditions. The degradation of β -casein was observed on SDS-PAGE when incubated with the LigA7. A decrease in the intensity of the β -casein and a smaller product of the same was observed compared to the control (**Figure 5.22**).

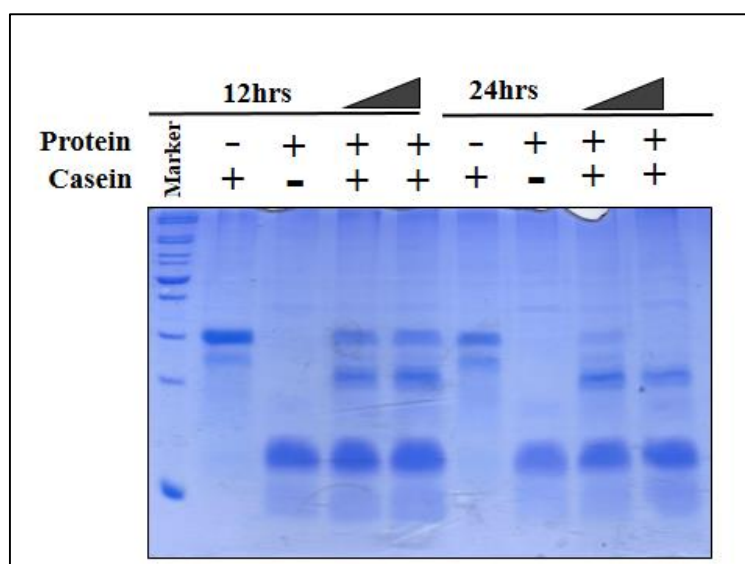


Figure 5.22: Protease activity. Degradation of β -casein in the presence of LigA7 incubated at the different time points and monitored on 12% SDS-PAGE. Only LigA7 and casein were also kept in control.

It was observed that LigA7 degraded the native form of β -casein completely at a higher concentration of LigA7. A 20 kDa and other small fragments of β -casein were observed when incubated for 24hrs. These bands were cut from the SDS-PAGE and sent for MALDI analysis (**Figure 5.23** and **Figure 5.24**).

The MALDI-MS spectra of the degraded β -casein band confirmed that, indeed, the degraded band is a cleaved product of β -casein (**Figure 5.23 A&B** and **Figure 5.24 A&B**).

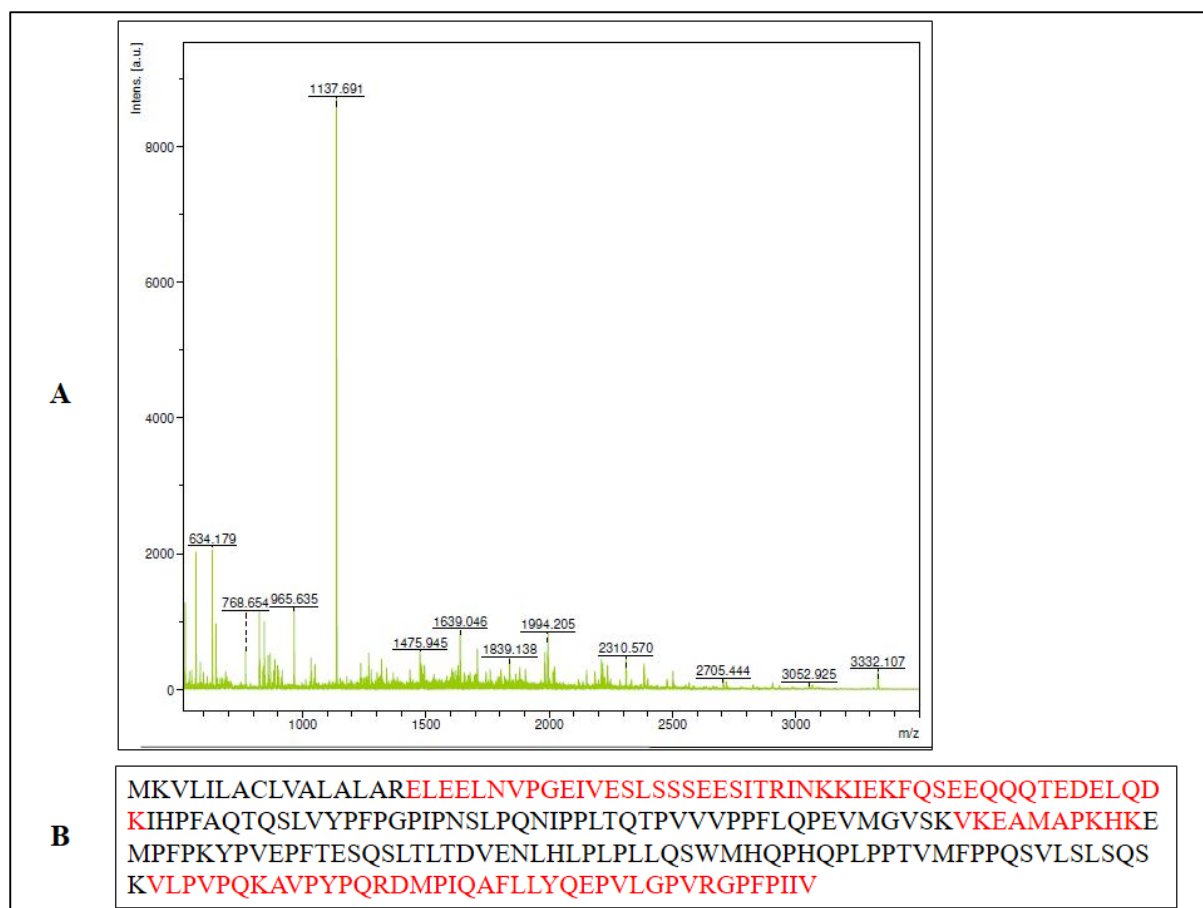


Figure 5.23: (A) MALDI-MS spectra of the degraded casein band. (B). A sequence of the degraded product.

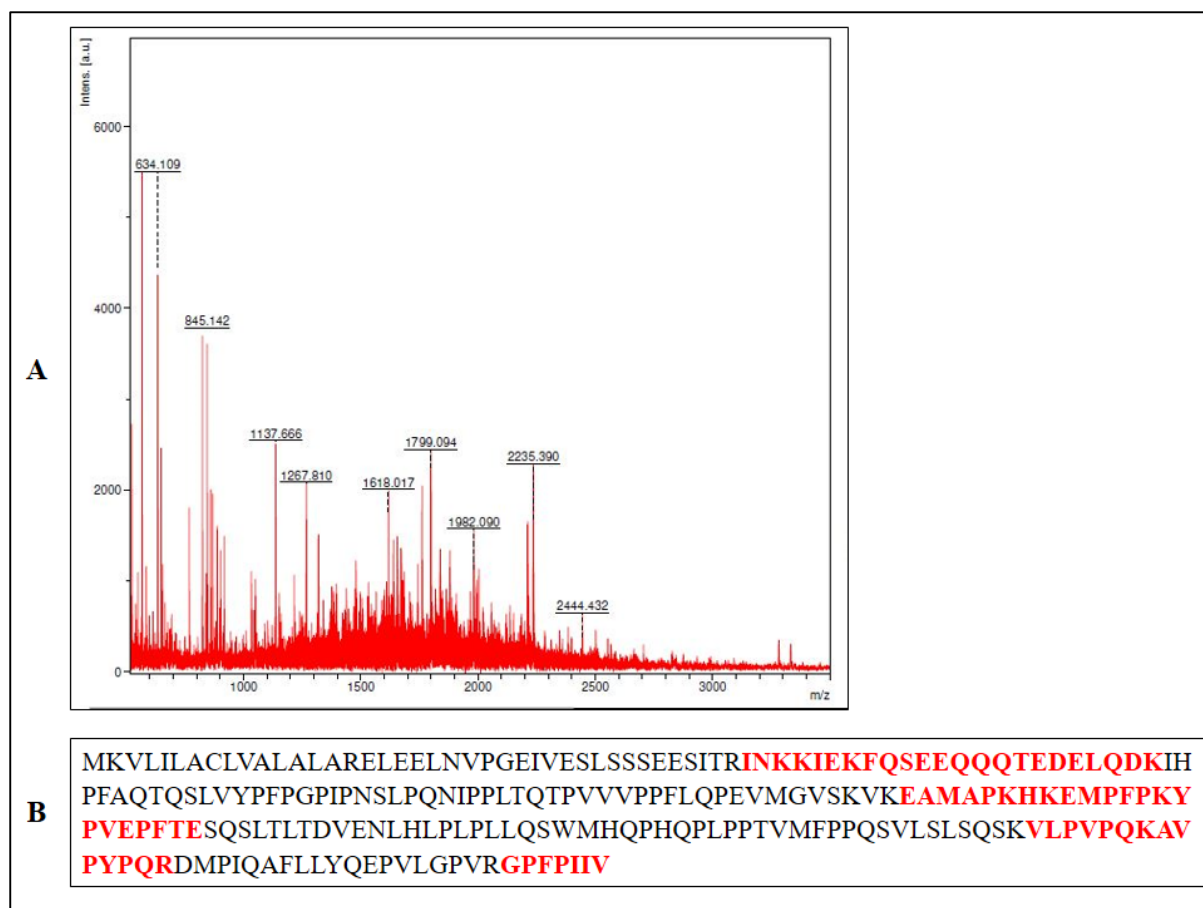


Figure 5.24: (A) MALDI-MS spectra of the degraded casein band. (B). A sequence of the degraded product.

Protease activity LigA7 showed similar activity in the absence and presence of the Mg^{2+} ions. Moreover, the optimum pH required for maximum cleavage activity was observed at pH 8.0.

In order to check the effect of protease inhibitors on the activity of LigA7, a β -casein degradation assay was performed in the presence of an increasing concentration of protease inhibitor cocktail. No cleavage in the β -casein was observed after incubation for 12hr (**Figure 5.25**) and 24hr (**Figure 5.26**).

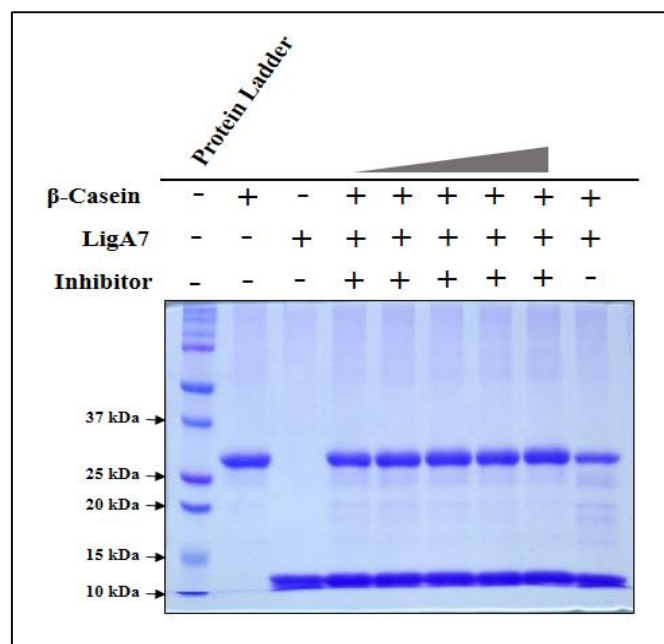


Figure 5.25: SDS-PAGE showing degradation pattern of β -Casein. The reaction was incubated for 12hr in the presence of different concentrations of protease inhibitor cocktail.

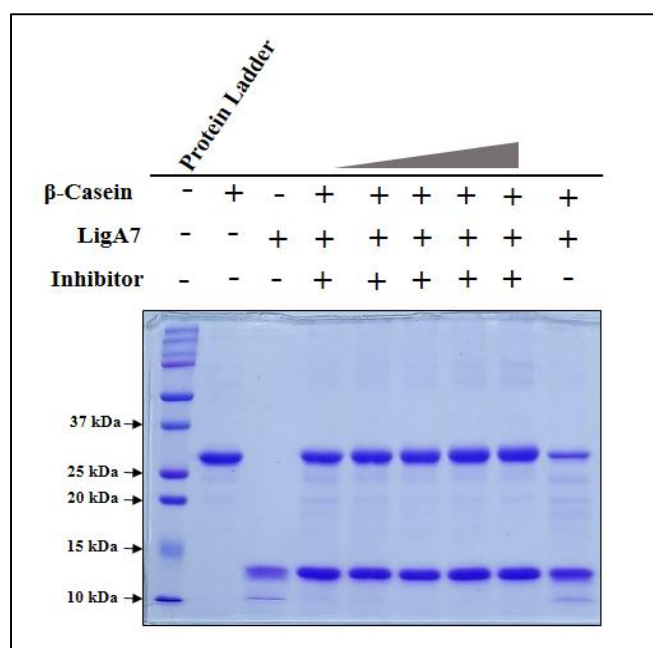


Figure 5.26: SDS-PAGE showing degradation pattern of β -Casein. Incubated for 24hr in the presence of different concentrations of protease inhibitor cocktail.

Furthermore, the protease activity of LigA7 was compared with boiled and native protein. The reaction mixture without the substrate was boiled at 95° C for 20mins. The mixture was cooled down on the ice for 5 minutes, followed by the addition of the substrate β -Casein. The mixture was incubated at 37° C for 12hr and 24 hrs. The sample was then boiled with SDS sample buffer and loaded on SDS-PAGE to monitor the cleavage.

As anticipated, no cleavage was observed in the boiled sample incubated at both 12 and 24hr. On the other hand, degradation was observed in the sample containing native protein. (**Figure 5.27 & Figure 5.28**). These observations suggest that LigA7 protein also possesses *in-vitro* non-specific protease activity.

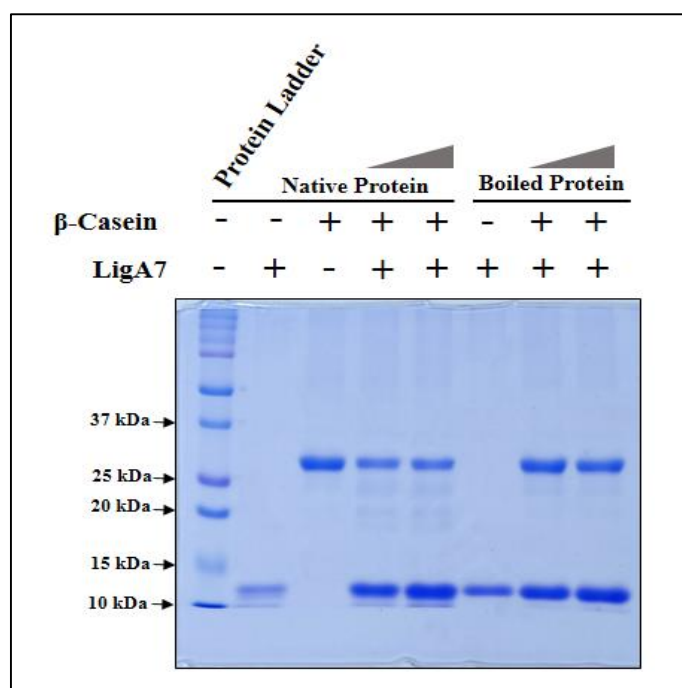


Figure 5.27: SDS-PAGE showing degradation pattern of β -Casein. Incubated for 12hr in the presence of boiled and native LigA7.

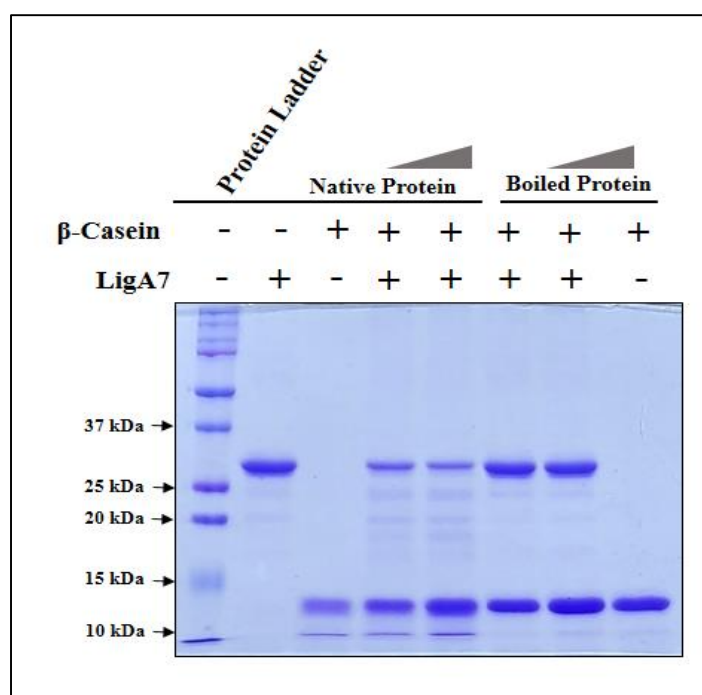


Figure 5.28: SDS-PAGE showing degradation pattern of β -Casein. Incubated for 24hr in the presence of boiled and native LigA7.

5.5. Proposed model of the novel activity

Lig proteins are surface-expressed and are associated with multiple functions. Reports suggest that when humans encounter *Leptospira* infection, neutrophil extracellular traps (NETs) formation takes place to neutralize the pathogens. Many pathogens express surface nuclease to cleave the NETs and escape the immune-mediated killing. However, the pathogenic *Leptospira* escape the NET-mediated killings, but during the genome mining, no known nuclease has been reported. The Lig proteins are highly conserved across the pathogenic *Leptospira spp.* and through our study, we were able to establish the novel nuclease and protease activity in Lig. Based on these observations, we hypothesize a model stating that Lig proteins might be involved in *Leptospira* NET evasion (**Figure 5.29**). However, these are preliminary results and need further investigation.

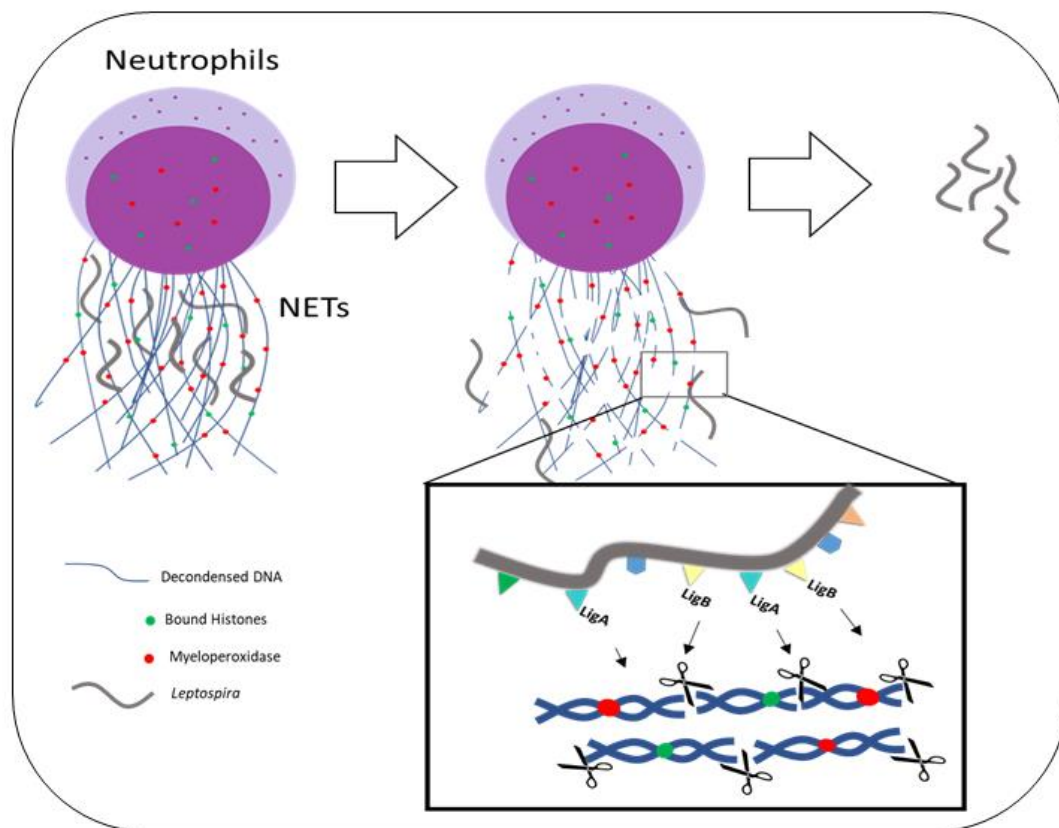


Figure 5.29: Proposed model of this novel activity in the NET escape.

5.6. Conclusions

In this chapter, characterization of novel nuclease and protease activities among Lig proteins has been done. This work is the first report on the comprehensive investigation of novel nuclease and protease activities among the Ig-like variable domains of the LigA protein from pathogenic *Leptospira*. Though Lig domains possess no homology with any known nucleases, they still can cleave DNA non-specifically in the presence of Mg^{2+} ions. Moreover, the Lig domains are specific to nucleic acid sugar, as LigA7 does not cleave RNA. The presence of basic patches on the LigA7 indicates DNA binding surface and may facilitate DNA catalysis in the presence of Mg^{2+} ions. Through mutational studies, we were able to identify the probable residues involved with Magnesium interactions. Our results indicate moonlighting properties among LigA protein. Moreover, our data provide a better understanding of the involvement of non-nucleases in NET evasion.

Chapter 6

Discussion

Leptospirosis is a global tropical zoonotic disease that accounts for approx. 60000 human lives annually (Costa et al., 2015). Outer membrane proteins (OMPs) from *Leptospira* played a crucial role in establishing infection and in immune evasion strategies (Fraga et al., 2014). Many of these have been investigated for their potency of vaccine candidates. To date, the Lig family of proteins (LigA and LigB) has been reported to be the most potent vaccine against leptospirosis (Henry A. Choy et al., 2007; Coutinho et al., 2011; Silva et al., 2007a). Lig proteins consist of multiple repetitive 12-13 homologous Ig-like domains divided within constant and variable regions. The family of Lig proteins is a multifaceted protein and is associated with Leptospiral infection (Matsunaga et al., 2003a). The LigA and LigB interact with several host extracellular matrix components. The role of these proteins in serum resistance has been demonstrated by knockdown mutants in *L. interrogans* (Fernandes et al., 2021). They also bind with the complement regulator proteins such as FH and C4BP, allowing pathogens to escape the immune-mediated clearance (Barbosa and Isaac, 2020; Fraga et al., 2014). The Ig-like domain is present in a diverse class of proteins and has been involved in various functions. Their conservancy across the pathogenic serovars and involvement in multiple functions makes them a potential vaccine target and one of the critical biomarkers. There are several reports which highlight the importance of Lig proteins in terms of pathogenesis and vaccine development.

The Lig protein is one of the most studied proteins but lacks a 3-dimension structure and antigenic regions. As mentioned in chapter 2, various recombinant LigA and LigB fragments were cloned overexpressed and purified to homogeneity with a good yield (**Table 2.10, Chapter 2**). Gel filtration chromatography profiles show that most LigA and LigB fragments exist as monomers in the solution. All purified fragments were subjected for crystallization trials. Many yielded crystallization hits. Only LigA8-9 crystals could diffract and data set was collected. The three-dimensional structure of LigA8-9 was determined using X-ray

crystallography. The structure reveals a classical immunoglobulin-like fold. The fragment containing the two domains showed a Greek key-like arrangement. Each domain consists of 10 anti-parallel β -strands arranged in two sheets. This arrangement is typical of a conventional Ig-like fold (**Figure 3.19**). However, the length of β -strands and loop regions vary among Ig-like folds. Overlaying the structure of individual Lig domains of LigA8-9 showed similarities to Intimin and other Lig domains such as LigA4 and LigB12 (Mei et al., 2015a; Ptak, 2014). The core of the Ig fold of LigA8 and LigA9 consists of many hydrophobic residues. Conserved Lys and Asp residues involved in stabilizing domain flexibility by salt bridge formation have been identified. The involvement of potassium ions between the dimer interfaces appears to be a crystallization artifact, as no dimer was detected during gel filtration chromatography. The identified calcium-binding residues were consistent with those previously determined for LigA4 (Mei et al., 2015a). (Chapter 3).

Getting an accurate knowledge of antigenic epitopes from many big or multi-domain proteins has been a laborious task due to constraints associated with the lack of high expression and purification as well as subsequent structural studies. Moreover, immunological research in vaccine design is experimentally costly and very intensive. However, the development of many powerful computational tools in the last decade has made tremendous progress towards handling a larger junk of genomics and proteomics data for valuable purposes. One such example of reverse vaccinology and structural bioinformatics has been exploited successfully to reduce the effort of screening many antigens and immuno-dominant epitopes to design effective vaccines. Similarly, immunoinformatics approaches have been used to unveil crucial immune-dominant determinants from available databases to develop vaccines.

This study uses various immunoinformatics approaches to construct a multi-epitope chimeric vaccine from multi-domain outer surface proteins, *Leptospira* immunoglobulins like (Lig) proteins (Chapter 4). Humoral, as well as cell-mediated immunities are triggered in *Leptospira*

infection. Humoral immunity is majorly involved in the clearance of *Leptospira* by producing antibodies like IgG and IgM against the leptospiral lipopolysaccharides (LPS) and other antigenic proteins (Adler and Faine, 1977; Karen V Evangelista & Jenifer Coburn, 2010). Many other reports demonstrated the implication of the cell-mediated immune responses against leptospirosis. One such is Ying-Jun Guo et. al. 2010, who identified Lig peptides associated with the human CD8⁺ T-lymphocytes (Y. J. Guo et al., 2010). Similar studies also illustrated the induction of CD4⁺ and γ/δ T-cells to stimulate type 1 cytokine production against the heat-killed *Leptospira* (De Fost et al., 2003; B. M. Naiman et al., 2001). The potential immunoprotection of Lig proteins and their truncated version in the context of vaccine candidates and their role in T-cells and B-cell mediated immune response against *Leptospira* have been highlighted (Haake and Matsunaga, 2021).

The study predicted most antigenic BCL and CTL epitopes present on both LigA and LigB proteins. The advantage of using numerous tools was greater confidence in predicting potential epitopes. The selected CTL epitopes had strong binding with the respective HLA alleles covering the wide distribution among the global population. Moreover, promiscuous peptides were selected, followed by separating non-self and IFN- γ inducing epitopes. Choosing the non-self-peptides may help to avoid any harmful auto-immune responses. The selection of IFN- γ epitopes is believed to induce innate and adaptive immunity (McLoughlin et al., 2008). B-cell epitopes in a vaccine construct help in the induction of cell-mediated immunity and contribute to antibody-based immune therapeutics (Getzoff et al., 1988; Van Regenmortel, 2006). Our study's generation of sequence-based linear and structure-based conformational BCL epitopes using multiple tools helped us select the best overlapping potential epitopes. Interestingly, comparing BCL and TCL epitopes further helps to sort out common and overlapping among them and yield final peptides for vaccine design. These peptides were mapped on the Ig-like domains of Lig proteins and observed to be present mainly on 7, 9, 11,

and 13 Ig-like domains from LigA and 1, 8, 9 and 12 Ig-like domains from LigB protein (**Table 4.3**). Interestingly, it was observed that primarily Ig-like domains from the C-terminal portion of Lig proteins possess high antigenic epitopes. In the case of LigA, a previous study reported that the carboxyl-terminal region of LigA protein provides protective immunity against *Leptospira* infection (Silva et al., 2007a). Moreover, Coutinho et al. 2011 have screened within the carboxyl-terminal of LigA and reported that the immunoprotective effect is mainly present in 10-13 Ig-like domains of LigA (Coutinho et al., 2011). No protective effect was reported from the N-terminal region of Lig proteins. In another study, the LigA variable region incorporated with PLGA confers a protective immune response (Faisal et al., 2009).

Similarly, the involvement of LigB and its variant in inducing a protective immune response against leptospiral infection have already been highlighted (Luo et al., 2009; Souza et al., 2017). Our prediction also showed the presence of the highest antigenic epitopes from the variable region of Lig proteins, which is in line with the previously reported studies, except for the existence of antigenic epitopes on the 1st Ig-like domain from the constant region of LigB. The chimeric vaccine constructs generated from the selected epitopes may provide excellent immune response. Our final construct consists of the highest antigenic epitopes from these two proteins linked together by AAY and GPGPG linkers. The linkers AAY and GPGPG used for the vaccine construct were reported to have better epitope presentation ability and better sites for proteasomal system cleavage (Farhadi et al., 2015; Livingston et al., 2002b; Nezafat et al., 2014b). Adjuvants used in the vaccine construct induce robust immune response via the immune receptor. Our vaccine construct is also augmented with an adjuvant HBHA at the N-terminal. This novel adjuvant is a TLR4 agonist expressed by *Mycobacterium* known to induce a Th1-type response that helps in effective immunotherapeutic strategies (Grange et al., 2008; Jung et al., 2011). It is believed that the final construct reported in our study consists of all the

required components for a vaccine construct that may also provide a better immune response (**Figure 4.3**).

The success of any vaccine construct relies on the fact that it should confer any deleterious effect on the host. The vaccine construct passes the test and has been categorized as non-allergic and antigenic, which may provide the required response. Moreover, it is predicted to be stable on overexpression and thus qualifies under a good vaccine category. Many promising vaccine constructs follow a well-defined immunological mechanism to trigger the expected response in the host. These responses mainly lie in the successful interaction of the vaccine constructs with toll-like receptors. In context to *Leptospira*, it is reported that mice defects in TLR4, have increased susceptibility to *Leptospira* infection (Koizumi et al., 2005). One of the surface adhesins, Lsa21 from *Leptospira*, induces TLR4 mediated inflammatory response (Faisal et al., 2016). These studies signify the importance of host TLR4 in immunological mechanisms against leptospiral antigens. Here, we observed that our multi-epitope vaccine construct also establishes a protein-protein docking interaction with the TLR4, through several hydrogen bonds and salt bridges (**Figure 4.6, Table 4.5**). Molecular dynamic simulation generated RMSDs emphasized stable interaction between our vaccine construct and TLR4 receptor. Moreover, the interaction and compactness of the complex between the two are also indicated in RMSF and RG plots (**Figure 4.7**). This suggests that our vaccine construct may follow TLR4 mediated immune response.

The ultimate aim of any vaccine is to produce a good immune response in terms of antibody and cytokines productions. Hence, immune profiling of the vaccine construct is essential. Our vaccine construct shows an increase in IgM antibodies after five days of vaccination, a rise in the B-cell population, and an elevated level of IFN- γ . These predicted immune profiles support our vaccine construct's strong cellular and humoral responses. Finally, producing a recombinant vaccine construct requires a suitable heterologous system for expression. Since

the half-life of our vaccine construct is predicted to be reasonably stable (>10 hr.) in the *E. coli* expression system, *E. coli* strain is used as a source for heterologous expression. Based on *E. coli* strain K12, codons optimization of the vaccine construct gene was performed for the in-silico cloning of the same in a standard expression vector, pET28a (+).

The host's innate immune system plays a crucial role in recognizing and eliminating *Leptospira* in the initial phase of infection (Fraga et al., 2011). However, the pathogen has adopted various ways to combat the host immune response by using the related proteins.

This study establishes novel activities within Ig-like domains of LigA and LigB (Chapter 5). Interestingly, single and multiple Ig-like domains from the variable region of LigA protein exhibit Mg^{2+} ion-dependent nuclease activity. However, a previous study reported that proteins with Ig-like folds bearing the His-Me finger domain act as an endonuclease (Jablonska et al., 2017b). Unfortunately, the His-Me finger domain is not found in any of the Ig-like domains of the Lig proteins. There was a total absence of His residues among the Ig-domains of LigA. Surprisingly, a concentration-dependent DNA degradation was observed, and complete degradation of supercoiled DNA was also observed at higher enzyme concentrations. The observed DNA degradation by various Ig-like fragments of LigA was very similar to the previously reported Vvn nuclease from *Vibrio vulnificus* (Li et al., 2003). To our surprise, no structural homology with any of the known nucleases was displayed by Lig domains. Most of the reported nucleases require metals for cleavage activity (Dupureur, 2008; Yang, 2011a). Our study also demonstrated a metal-dependent nuclease activity. Since the Lig protein has been classified as a calcium-binding protein, the presence of calcium is known to enhance its interaction with the host fibronectin (Lin et al., 2008b; Raman et al., 2010). However, in our study, the Lig protein favored the Mg^{2+} ion over the Ca^{2+} ion for its nuclease activity. The metal interaction with the nucleases plays a vital role in DNA catalysis in substrate affinity,

specificity, and structural integrity (Bowen and Dupureur, 2003; Conlan and Dupureur, 2002; Sigel, 2005; Sigel and Pyle, 2007).

Fluorescence spectroscopy is an important method for studying protein stability, hydrodynamics, kinetics, and ligand interactions (Sen et al., 2009). A change of fluorescence emission and quenching experiments have been applied to study protein metal interaction (Witkowska and Rowińska-Żyrek, 2019). A change of intrinsic fluorescence emission of LigA7 upon Mg^{2+} ion in a concentration-dependent manner indicates Mg^{2+} ion-protein interaction (**Figure 5.16**). Metal ions usually coordinate to the specific residues in the protein. Mg^{2+} ion mainly coordinates with Cys, His, Asp, Glu, Asn, Thr, etc., residues in the protein (Khrustalev et al., 2016). Except for Cys and His, many of these residues were conserved in the Ig-domain of the LigA protein (**Figure 5.17**). The mutation of a few of these residues in the LigA7 demonstrated their inability to cleave DNA in the presence of Mg^{2+} ion. A reduction in the nuclease activity with the protein bearing N49A mutation clearly indicates the coordination of Asn49 with the Mg^{2+} ions. In addition, other mutants such as T22A, D38A, T44A, and N57A also showed a reduction in the activity, highlighting their role in metal interaction and nuclease activity (**Figure 5.19**). The results demonstrate that, indeed, Mg^{2+} ions bind to LigA7. However, the binding was only investigated with LigA7, as many metal-binding residues are conserved among other Ig-domains of LigA proteins; the same conclusion may be extended for the different domains of LigA protein also. Nucleases exhibit cleavage activity with various nucleotide substrates such as single-stranded DNA, double-stranded DNA, supercoiled DNA, unmethylated DNA, and RNA. Based on their substrate preference, nucleases have been classified as DNases or RNases (Hsia et al., 2005; Yang, 2011b). The result demonstrates an inability of LigA7 to cleave RNA, which behaves more like DNases (**Figure 5.15**).

For any enzyme activity, a substrate has to interact with an enzyme. Similarly, the DNase/nuclease interacts with DNA substrate and catalyzes DNA cleavage. Few DNases, such as DNase-I, have specific binding to the minor groove of DNA for its cleavage. Other endonucleases such as Periplasmic endonuclease, Vvn from *Vibrio vulnificus* bind to the phosphate backbone of a non-specific DNA and catalyze its cleavage. The electrostatic surface of LigA7 showed a patch of basic amino acid residues, which indicates that this surface may provide a site for the negatively charged phosphate groups of the non-specific DNA to bind (**Figure 5.21**). This surface offers contact with the DNA for hydrolysis in the presence of Mg^{2+} ions. Our result also identified the probable DNA binding residues in the LigA7 by molecular docking analysis study. The DNA hydrolyzing activity of LigA7 was very similar to the Vvn endonuclease (Li et al., 2003). But, the DALI search did not yield any homology either with Vvn or other known nuclease/DNases. At this point, the mechanism of DNA hydrolysis by LigA7 or other LigA domains is not yet precise, but it left an avenue to investigate further. However, this is the first report of DNA hydrolysis within the Lig proteins. Lig protein plays a multifaceted function leading to speculation that Lig proteins may have moonlighting functions. Surprisingly, the MoonProt server predicted that the Lig protein would exhibit peptidase-like activity as a secondary function (Mani et al., 2015b). This activity was confirmed by the LigA7 protein's ability to cleave β -casein (**Figure 5.22**). Many proteins do different functions under different conditions/environments, such as pH, salt, or oxidizing and reducing conditions. Such types of proteins are known as moonlighting proteins (Jeffery, 2016). Similarly, we believe that Lig proteins may also perform different functions under different conditions. The novel nuclease/DNase process in the Ig-domain of LigA protein without having a nuclease/DNase motif may be associated with a transient structural change in the presence of substrate, DNA, or probably Mg^{2+} ion.

The present study demonstrates a unique and novel aspect of Lig harboring protease and nuclease activities. However, the physiological significance of this activity is unknown. As reported earlier, *Leptospira* infection-induced neutrophil extracellular trap (NET) formation is reported to evade the NET-mediated killings with unknown factors (Scharrig et al., 2015). Neutrophils are important players involved in the manifestation of host innate response, which aid in the clearance or neutralization of the pathogen burden in numerous ways, such as antimicrobial peptide production, reactive oxygen formation, and pathogen phagocytosis, and neutrophil extracellular trap formation (Amulic and Hayes, 2004). Interestingly, many pathogens, such as *Streptococcus pneumoniae*, *Vibrio cholerae*, *Paracoccidioides brasiliensis*, *Prevotella intermedia*, and *Trypanosoma brucei* has been reported to express surface-associated/secretory nucleases, which aid them to degrade DNA which is a major component of NETs (Doke et al., 2017; Jhelum et al., 2018; Seper et al., 2013; Zhang et al., 2021; Zonta et al., 2021). Although the pathogenic *Leptospira* exhibits the ability to degrade the DNA of NETs, in the course of genome mining, no annotated nuclease was found in the *Leptospira* (Nascimento et al., 2004). In a recent report, the single domain of LigA11 was shown to have nuclease activity that is reported to aid in NET escape (Kumar et al., 2021). This study establishes nuclease activity in all single Ig-like and multiple Ig-like domains of LigA. Based on this observation, It is hypothesized that upon *Leptospira* infection, neutrophils get activated, leading to NET induction. The NET traps the pathogen and neutralizes it using various antimicrobial components. In the absence of nucleases, the pathogenic *Leptospira* may use the surface-expressed Lig proteins to cleave the DNA of the NET and escape from NET-mediated killing. It is believes that this study enhances the basic knowledge of non-nuclease proteins involved in the NET evasion phenomenon in *Leptospira* and makes the foundation to explore other pathogens. This information may be utilized to develop preventive strategies to prevent *Leptospira* immune evasion. Overall finding of the study is summarized in **Figure 6.1**.

Bibliography

- Adames, N.R., Wilson, M.L., Fang, G., Lux, M.W., Glick, B.S., Peccoud, J., 2015. GenoLIB: a database of biological parts derived from a library of common plasmid features. *Nucleic Acids Res.* 43, 4823–4832. <https://doi.org/10.1093/nar/gkv272>
- Adler, B., de la Peña Moctezuma, A., 2010a. *Leptospira* and leptospirosis. *Vet. Microbiol.* 140, 287–296. <https://doi.org/10.1016/j.vetmic.2009.03.012>
- Adler, B., de la Peña Moctezuma, A., 2010b. *Leptospira* and leptospirosis. *Vet. Microbiol.* 140, 287–296. <https://doi.org/10.1016/j.vetmic.2009.03.012>
- Adler, B., Faine, S., 1977. Host immunological mechanisms in the resistance of mice to leptospiral infections. *Infect. Immun.* 17, 67–72. <https://doi.org/10.1128/iai.17.1.67-72.1977>
- Agampodi, S.B., Peacock, S.J., Thevanesam, V., Nugegoda, D.B., Smythe, L., Thaipadungpanit, J., Craig, S.B., Burns, M.A., Dohnt, M., Boonsilp, S., Senaratne, T., Kumara, A., Palihawadana, P., Perera, S., Vinetz, J.M., 2011. Leptospirosis Outbreak in Sri Lanka in 2008 : Lessons for Assessing the Global Burden of Disease 85, 471–478. <https://doi.org/10.4269/ajtmh.2011.11-0276>
- Ajay R Bharti, 2003. Leptospirosis: a zoonotic disease of global importance. *Lancet Infect Dis* 3, 176–180.
- Amulic, B., Hayes, G., 2004. Quick guide Neutrophil extracellular traps. *CURBIO* 21, R297–R298. <https://doi.org/10.1016/j.cub.2011.03.021>
- Andre, G.O., Converso, T.R., Politano, W.R., Ferraz, L.F.C., Ribeiro, M.L., Leite, L.C.C., Darrieux, M., 2017. Role of *Streptococcus Pneumoniae* proteins in evasion of complement-mediated immunity. *Front. Microbiol.* 8, 1–20. <https://doi.org/10.3389/fmicb.2017.00224>
- Andreatta, M., Nielsen, M., 2016. Gapped sequence alignment using artificial neural networks: Application to the MHC class I system. *Bioinformatics* 32, 511–517. <https://doi.org/10.1093/bioinformatics/btv639>
- Arai, R., Ueda, H., Kitayama, A., Kamiya, N., Nagamune, T., 2001. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng.* 14, 529–532. <https://doi.org/10.1093/protein/14.8.529>
- Atzingen, M. V., Barbosa, A.S., De Brito, T., Vasconcellos, S.A., De Moraes, Z.M., Lima,

- D.M.C., Abreu, P.A.E., Nascimento, A.L.T.O., 2008. Lsa21, a novel leptospiral protein binding adhesive matrix molecules and present during human infection. *BMC Microbiol.* 8, 1–16. <https://doi.org/10.1186/1471-2180-8-70>
- Atzingen, M. V., Gómez, R.M., Schattner, M., Pretre, G., Gonçalves, A.P., de Moraes, Z.M., Vasconcellos, S.A., Nascimento, A.L.T.O., 2009. Lp95, a novel leptospiral protein that binds extracellular matrix components and activates e-selectin on endothelial cells. *J. Infect.* 59, 264–276. <https://doi.org/10.1016/j.jinf.2009.07.010>
- Atzingen, M. V, Gonc, A.P., Moraes, Z.M. De, Arau, E.R., Brito, T. De, Vasconcellos, S.A., Nascimento, A.L.T.O., 2010. Characterization of leptospiral proteins that afford partial protection in hamsters against lethal challenge with *Leptospira interrogans* Printed in Great Britain 1005–1015. <https://doi.org/10.1099/jmm.0.021485-0>
- Atzingen, M. V, Vieira, M.L., Oliveira, R., Domingos, R.F., Renata, S., Barros, A.T., Gonçalves, A.P., Moraes, Z.M. De, Vasconcellos, S.A., Nascimento, A.L.T.O., 2012. Evaluation of Immunoprotective Activity of Six Leptospiral Proteins in the Hamster Model of Leptospirosis 79–87.
- Barbosa, Angela S., Abreu, P.A.E., Neves, F.O., Atzingen, M. V., Watanabe, M.M., Vieira, M.L., Moraes, Z.M., Vasconcellos, S.A., Nascimento, A.L.T.O., 2006. A newly identified leptospiral adhesin mediates attachment to laminin. *Infect. Immun.* 74, 6356–6364. <https://doi.org/10.1128/IAI.00460-06>
- Barbosa, Angela S, Abreu, P.A.E., Neves, F.O., Atzingen, M. V, Watanabe, M., Vieira, L., Moraes, Z.M., Vasconcellos, A., Nascimento, A.L.T.O., Sa, U., 2006. A Newly Identified Leptospiral Adhesin Mediates Attachment to Laminin □ 74, 6356–6364. <https://doi.org/10.1128/IAI.00460-06>
- Barbosa, A.S., Abreu, P.A.E., Vasconcellos, S.A., Moraes, Z.M., Gonc, A.P., Silva, A.S., Daha, M.R., Isaac, L., 2009. Immune Evasion of *Leptospira* Species by Acquisition of Human Complement Regulator C4BP □ 77, 1137–1143. <https://doi.org/10.1128/IAI.01310-08>
- Barbosa, A.S., Isaac, L., 2020. Strategies used by *Leptospira* spirochetes to evade the host complement system. *FEBS Lett.* 594, 2633–2644. <https://doi.org/10.1002/1873-3468.13768>
- Barthel, D., Schindler, S., Zipfel, P.F., 2012. Plasminogen Is a Complement Inhibitor *. *J. Biol.*

- Chem. 287, 18831–18842. <https://doi.org/10.1074/jbc.M111.323287>
- Beck, M., Schmidt, A., Lange, V., Deutsch, E.W., 2009. Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans* 460, 762–766. <https://doi.org/10.1038/nature08184>
- Berendsen, H J C, Grigera, J.R., Straatsma, T.P., 1987. The Missing Term in Effective Pair Potentialst 6269–6271.
- Berendsen, H. J.C., Grigera, J.R., Straatsma, T.P., 1987. The missing term in effective pair potentials. *J. Phys. Chem.* 91, 6269–6271. <https://doi.org/10.1021/j100308a038>
- Bolin, C.A., Alt, D.P., 2001. Use of a monovalent leptospiral vaccine to prevent renal colonization and urinary shedding in cattle exposed to *Leptospira borgpetersenii* serovar hardjo 995–1000.
- Bork, P., Holm, L., Sander, C., 1994. The Immunoglobulin Fold. *J. Mol. Biol.* <https://doi.org/10.1006/jmbi.1994.1582>
- Bowen, L.M., Dupureur, C.M., 2003. Investigation of restriction enzyme cofactor requirements: a relationship between metal ion properties and sequence specificity. *Biochemistry* 42, 12643–12653. <https://doi.org/10.1021/bi035240g>
- Branger, C., Chatrenet, B., Gauvrit, A., Aviat, F., Aubert, A., Bach, J.M., Andre, G., 2005. Protection against *Leptospira interrogans* Senu Lato Challenge by DNA Immunization with the Gene Encoding Protein 1 73, 4062–4069. <https://doi.org/10.1128/IAI.73.7.4062>
- Branger, C., Sonrier, C., Chatrenet, B., Klonjowski, B., Aubert, A., Andre, G., 2001. Identification of the Hemolysis-Associated Protein 1 as a Cross- Protective Immunogen of *Leptospira interrogans* by Adenovirus-Mediated Vaccination 69, 6831–6838. <https://doi.org/10.1128/IAI.69.11.6831>
- Breda, L.C.D., Hsieh, C.L., Castiblanco Valencia, M.M., da Silva, L.B., Barbosa, A.S., Blom, A.M., Yung-Fu, C., Isaac, L., 2015. Fine Mapping of the Interaction between C4b-Binding Protein and Outer Membrane Proteins LigA and LigB of Pathogenic *Leptospira interrogans*. *PLoS Negl. Trop. Dis.* 9, 1–18. <https://doi.org/10.1371/journal.pntd.0004192>
- Brown, R.A., Blumberman, S., Gay, C., Bolin, C., Duby, R., Baldwin, C.L., 2003. Comparison of three different leptospiral vaccines for induction of a type 1 immune response to *Leptospira borgpetersenii* serovar Hardjo 21, 4448–4458. <https://doi.org/10.1016/S0264->

410X(03)00439-0

- Brzezinski, D., Porebski, P.J., Kowiel, M., Macnar, J.M., Minor, W., 2021. Recognizing and validating ligands with CheckMyBlob. *Nucleic Acids Res.* 49, W86–W92. <https://doi.org/10.1093/nar/gkab296>
- Budihal, S.V., Perwez, K., 2014. Leptospirosis diagnosis: competency of various laboratory tests. *J. Clin. Diagn. Res.* 8, 199–202. <https://doi.org/10.7860/JCDR/2014/6593.3950>
- Bulach, D.M., Zuerner, R.L., Wilson, P., Seemann, T., Mcgrath, A., Cullen, P.A., Davis, J., Johnson, M., Kuczek, E., Alt, D.P., Peterson-burch, B., Coppel, R.L., Rood, J.I., Davies, J.K., Adler, B., 2006. Genome reduction in *Leptospira borgpetersenii* reflects limited transmission potential.
- Bussi, G., Donadio, D., Parrinello, M., 2007. Canonical sampling through velocity rescaling. *J. Chem. Phys.* 126. <https://doi.org/10.1063/1.2408420>
- Cabanes, D., Dehoux, P., Dussurget, O., Frangeul, L., Cossart, P., 2002. Surface proteins and the pathogenic potential of *Listeria monocytogenes*. *Trends Microbiol.* 10, 238–245. [https://doi.org/10.1016/s0966-842x\(02\)02342-9](https://doi.org/10.1016/s0966-842x(02)02342-9)
- Carlisle, J.T., Greer, D.L., Hyswp, N.E., 1988. Infectious Disease Section , Department of Medicine , Tulane Medical School ' and the Mycology and Mycobacteriology Laboratory , Department of Pathology and Dermatology , Charity Hospital , Louisiana State University Medical Center , Human-to-Human Transmi 246–247.
- Castiblanco-Valencia, M.M., Fraga, T.R., Silva, L.B. da, Monaris, D., Abreu, P.A.E., Strobel, S., Józsi, M., Isaac, L., Barbosa, A.S., 2012a. Leptospiral immunoglobulin-like proteins interact with human complement regulators factor H, FHL-1, FHR-1, and C4BP. *J. Infect. Dis.* 205, 995–1004. <https://doi.org/10.1093/infdis/jir875>
- Castiblanco-Valencia, M.M., Fraga, T.R., Silva, L.B. Da, Monaris, D., Abreu, P.A.E., Strobel, S., Józsi, M., Isaac, L., Barbosa, A.S., 2012b. Leptospiral immunoglobulin-like proteins interact with human complement regulators factor H, FHL-1, FHR-1, and C4BP. *J. Infect. Dis.* 205, 995–1004. <https://doi.org/10.1093/infdis/jir875>
- Chapman, A.J., Everard, C.R., Faine, S., Adler, B., 1991. Antigens recognized by the human immune response to severe leptospirosis in Barbados 143–155.
- Chothia, C., 1953. THE MOLECULAR STRUCTURE OF CELL ADHESION

- MOLECULES. Annu. Rev. Biochem. 75, 4116–4117.
<https://doi.org/10.1021/ja01112a543>
- Chou, K.C., Shen, H. Bin, 2008. ProtIdent: A web server for identifying proteases and their types by fusing functional domain and sequential evolution information. *Biochem. Biophys. Res. Commun.* 376, 321–325. <https://doi.org/10.1016/j.bbrc.2008.08.125>
- Choy, H.A., 2012. Multiple Activities of LigB Potentiate Virulence of *Leptospira interrogans* : Inhibition of Alternative and Classical Pathways of Complement 7. <https://doi.org/10.1371/journal.pone.0041566>
- Choy, Henry A., Kelley, M.M., Chen, T.L., Møller, A.K., Matsunaga, J., Haake, D.A., 2007. Physiological osmotic induction of *Leptospira interrogans* adhesion: LigA and LigB bind extracellular matrix proteins and fibrinogen. *Infect. Immun.* 75, 2441–2450. <https://doi.org/10.1128/IAI.01635-06>
- Choy, Henry A, Kelley, M.M., Chen, T.L., Møller, A.K., Matsunaga, J., Haake, D.A., 2007a. Physiological Osmotic Induction of *Leptospira interrogans* Adhesion : LigA and LigB Bind Extracellular Matrix Proteins and Fibrinogen □ 75, 2441–2450. <https://doi.org/10.1128/IAI.01635-06>
- Choy, Henry A, Kelley, M.M., Chen, T.L., Møller, A.K., Matsunaga, J., Haake, D.A., 2007b. Physiological osmotic induction of *Leptospira interrogans* adhesion: LigA and LigB bind extracellular matrix proteins and fibrinogen. *Infect. Immun.* 75, 2441–2450. <https://doi.org/10.1128/IAI.01635-06>
- Choy, H.A., Kelley, M.M., Croda, J., Matsunaga, J., Babbitt, J.T., Ko, A.I., Picardeau, M., Haake, D.A., 2011. The multifunctional LigB adhesin binds homeostatic proteins with potential roles in cutaneous infection by pathogenic *Leptospira interrogans*. *PLoS One* 6, e16879. <https://doi.org/10.1371/journal.pone.0016879>
- Christopher J. Pappas, M.P., 2015. Control of Gene Expression in *Leptospira* spp. by Transcription Activator-Like Effectors Demonstrates a Potential Role for LigA and LigB in *Leptospira interrogans* Virulence Christopher. *Appl. Environ. Microbiol.* 81, 7888–92. <https://doi.org/10.1128/AEM.02202-15>.Editor
- Claus, A., Maele, I. Van De, Pasmans, F., Gommeren, K., 2008. Leptospirosis in dogs : A retrospective study of seven clinical cases in Belgium Leptospirosis in dogs : a

- retrospective study of seven clinical cases in Belgium Leptospirose bij honden : een retrospectieve studie van zeven klinische gevallen.
- Clerke, A.M., Leuva, A.C., Joshi, C., Trivedi, S. V, 2002. Clinical profile of leptospirosis in South gujarat. *J. Postgrad. Med.* 48, 117–118.
- Coggins, W.J., 1962. Leptospirosis Due to *Leptospira pomona*: An Outbreak of Nine Cases. *JAMA J. Am. Med. Assoc.* 181, 1077–1078. <https://doi.org/10.1001/jama.1962.03050380055017b>
- Conlan, L.H., Dupureur, C.M., 2002. Dissecting the metal ion dependence of DNA binding by PvuII endonuclease. *Biochemistry* 41, 1335–1342. <https://doi.org/10.1021/bi015843x>
- Conrad, N.L., Cruz McBride, F.W., Souza, J.D., Silveira, M.M., Félix, S., Mendonça, K.S., Santos, C.S., Athanazio, D.A., Medeiros, M.A., Reis, M.G., Dellagostin, O.A., McBride, A.J.A., 2017. LigB subunit vaccine confers sterile immunity against challenge in the hamster model of leptospirosis. *PLoS Negl. Trop. Dis.* 11, 1–20. <https://doi.org/10.1371/journal.pntd.0005441>
- Costa, F., Hagan, J.E., Calcagno, J., Kane, M., Torgerson, P., Martinez-Silveira, M.S., Stein, C., Abela-Ridder, B., Ko, A.I., 2015. Global Morbidity and Mortality of Leptospirosis: A Systematic Review. *PLoS Negl. Trop. Dis.* 9, 0–1. <https://doi.org/10.1371/journal.pntd.0003898>
- Coutinho, M.L., Choy, H.A., Kelley, M.M., Matsunaga, J., Babbitt, J.T., Lewis, M.S., Aleixo, J.A.G., Haake, D.A., 2011. A ligA three-domain region protects hamsters from lethal infection by *leptospira interrogans*. *PLoS Negl. Trop. Dis.* 5, 1–10. <https://doi.org/10.1371/journal.pntd.0001422>
- Cox, T.E., Smythe, L.D., Leung, L.K.-P., 2005. Flying foxes as carriers of pathogenic *Leptospira* species. *J. Wildl. Dis.* 41, 753–757. <https://doi.org/10.7589/0090-3558-41.4.753>
- Croda, J., Ramos, J.G.R., Matsunaga, J., Queiroz, A., Homma, A., Riley, L.W., Haake, D.A., Reis, M.G., Ko, A.I., 2007. *Leptospira* immunoglobulin-like proteins as a serodiagnostic marker for acute leptospirosis. *J. Clin. Microbiol.* 45, 1528–1534. <https://doi.org/10.1128/JCM.02344-06>
- David A. Haake, 2015. *Leptospira* and leptospirosis, *Journal of Biological Education*.

- <https://doi.org/10.1080/00219266.1991.9655201>
- David A Haake, 2015. The Leptospiral Outer Membrane 187–221.
<https://doi.org/10.1007/978-3-662-45059-8>
- De Brito, T., Silva, A.M.G. da, Abreu, P.A.E., 2018. Pathology and pathogenesis of human leptospirosis: a commented review. *Rev. Inst. Med. Trop. Sao Paulo* 60, e23.
<https://doi.org/10.1590/s1678-9946201860023>
- De Fost, M., Hartskeerl, R.A., Groenendijk, M.R., Van der Poll, T., 2003. Interleukin 12 in part regulates gamma interferon release in human whole blood stimulated with *Leptospira interrogans*. *Clin. Diagn. Lab. Immunol.* 10, 332–335.
<https://doi.org/10.1128/CDLI.10.2.332-335.2003>
- De Gregorio, E., Rappuoli, R., 2012. Vaccines for the future: learning from human immunology. *Microb. Biotechnol.* 5, 149–155. <https://doi.org/10.1111/j.1751-7915.2011.00276.x>
- De Vries, S.J., Van Dijk, M., Bonvin, A.M.J.J., 2010. The HADDOCK web server for data-driven biomolecular docking. *Nat. Protoc.* 5, 883–897.
<https://doi.org/10.1038/nprot.2010.32>
- Dellagostin, O.A., Grassmann, A.A., Rizzi, C., Schuch, R.A., Jorge, S., Oliveira, T.L., McBride, A.J.A., Hartwig, D.D., 2017. Reverse Vaccinology: An Approach for Identifying Leptospiral Vaccine Candidates. *Int. J. Mol. Sci.* 18.
<https://doi.org/10.3390/ijms18010158>
- Deneke, Y., Sabarinath, T., Gogia, N., Lalsiamthara, J., Viswas, K.N., Chaudhuri, P., 2014. Evaluation of recombinant LigB antigen-based indirect ELISA and latex agglutination test for the serodiagnosis of bovine leptospirosis in India. *Mol. Cell. Probes* 28, 141–146.
<https://doi.org/10.1016/j.mcp.2014.01.001>
- Desiere, F., Deutsch, E.W., King, N.L., Nesvizhskii, A.I., Mallick, P., Eng, J., Chen, S., Eddes, J., Loevenich, S.N., Aebersold, R., 2006. The PeptideAtlas project. *Nucleic Acids Res.* 34, 655–658. <https://doi.org/10.1093/nar/gkj040>
- Dhanda, S.K., Vir, P., Raghava, G.P.S., 2013. Designing of interferon-gamma inducing MHC class-II binders. *Biol. Direct* 8, 1–15. <https://doi.org/10.1186/1745-6150-8-30>
- Dias, J.P., Teixeira, M.G., Costa, M.C.N., Mendes, C.M.C., Guimarães, P., Reis, M.G., Ko,

- A., Barreto, M.L., 2007. Factors associated with *Leptospira* sp infection in a large urban center in northeastern Brazil. *Rev. Soc. Bras. Med. Trop.* 40, 499–504. <https://doi.org/10.1590/s0037-86822007000500002>
- Dimitrov, I., Bangov, I., Flower, D.R., Doytchinova, I., 2014. AllerTOP v.2 - A server for in silico prediction of allergens. *J. Mol. Model.* 20. <https://doi.org/10.1007/s00894-014-2278-5>
- Doke, M., Fukamachi, H., Morisaki, H., Arimoto, T., Kataoka, H., Kuwata, H., 2017. molecular oral Nucleases from *Prevotella intermedia* can degrade neutrophil extracellular traps 4, 288–300. <https://doi.org/10.1111/omi.12171>
- Dong, R., Chu, Z., Yu, F., Zha, Y., 2020. Contriving Multi-Epitope Subunit of Vaccine for COVID-19: Immunoinformatics Approaches. *Front. Immunol.* 11, 1784. <https://doi.org/10.3389/fimmu.2020.01784>
- Doytchinova, I.A., Flower, D.R., 2007. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics* 8, 1–7. <https://doi.org/10.1186/1471-2105-8-4>
- Dupureur, C.M., 2008. Roles of metal ions in nucleases. *Curr. Opin. Chem. Biol.* 12, 250–255. <https://doi.org/10.1016/j.cbpa.2008.01.012>
- Eduardo, C., Bettin, E.B., Farid, A., Aziz, A., Bakry, Y., Clair, A., Seixas, P., Amaral, M.G., Dellagostin, O.A., 2019. Evaluation of different strategies to promote a protective immune response against leptospirosis using a recombinant LigA and LigB chimera. *Vaccine* 37, 1844–1852. <https://doi.org/10.1016/j.vaccine.2019.02.010>
- El-Manzalawy, Y., Dobbs, D., Honavar, V., 2008. Predicting linear B-cell epitopes using string kernels. *J. Mol. Recognit.* 21, 243–255. <https://doi.org/10.1002/jmr.893>
- Ellinghausen, H.C.J., McCullough, W.G., 1965. NUTRITION OF *LEPTOSPIRA* POMONA AND GROWTH OF 13 OTHER SEROTYPES: FRACTIONATION OF OLEIC ALBUMIN COMPLEX AND A MEDIUM OF BOVINE ALBUMIN AND POLYSORBATE 80. *Am. J. Vet. Res.* 26, 45–51.
- Ellis, W.A., 2015. Animal Leptospirosis. <https://doi.org/10.1007/978-3-662-45059-8>
- Ellis, W.A., Bryson, D.G., Neill, S.D., McParland, P.J., Malone, F.E., 1983. Possible involvement of leptospire in abortion, stillbirths and neonatal deaths in sheep. *Vet. Rec.*

- 112, 291–293. <https://doi.org/10.1136/vr.112.13.291>
- Ellis, W.A., Michno, S.W., 1976. Bovine leptospirosis: a serological and clinical study. *Vet. Rec.* 99, 387–391. <https://doi.org/10.1136/vr.99.20.387>
- Ellinghausen H C & McCullough W G, 1965. Nutrition of *Leptospira pomona* and growth of 13 other serotypes: fractionation of oleic albumin complex and a medium of bovine albumin and polysorbate 80. *Amer. J. Vet. Res.* 1989.
- Emsley, P., Cowtan, K., 2004. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 60, 2126–2132. <https://doi.org/10.1107/S0907444904019158>
- Enterotoxin, H., Fagundes, M.Q., Seixas, F.K., Silva, F., Conceição, F.R., Dellagostin, O.A., 2012. Protection against Lethal Leptospirosis after Vaccination with LipL32 Coupled or Coadministered with the B Subunit of *Escherichia coli* 740–745. <https://doi.org/10.1128/CVI.05720-11>
- Evangelista, K. V., Coburn, J., 2010. *Leptospira* as an emerging pathogen: A review of its biology, pathogenesis and host immune responses. *Future Microbiol.* 5, 1413–1425. <https://doi.org/10.2217/fmb.10.102>
- Evangelista, K. V, Lourdault, K., Matsunaga, J., Haake, D.A., 2017. Immunoprotective properties of recombinant LigA and LigB in a hamster model of acute leptospirosis. *PLoS One* 12, e0180004. <https://doi.org/10.1371/journal.pone.0180004>
- Factor, H., Castiblanco-valencia, M., Fraga, T.R., Bezerra, L., Monaris, D., Abreu, E., Strobel, S., Isaac, L., Barbosa, A.S., 2012. Leptospiral Immunoglobulin-like Proteins Interact With Human Complement Regulators 205, 995–1004. <https://doi.org/10.1093/infdis/jir875>
- Faine, S., 1999. *Leptospira* and Leptospirosis 14, 296–326. <https://doi.org/10.1128/CMR.14.2.296>
- Faine, S., Adler, B., Palit, A., 1974. CHEMICAL , SEROLOGICAL AND BIOLOGICAL PROPERTIES OF A SEROTYPE-SPECIFIC POLYSACCHARIDE ANTICEN IN *LEPTOSPIRA* 52, 311–319.
- Faine, S., Stallman, N.D., 1982. Amended descriptions of the genus *Leptospira* Noguchi 1917 and the species *L. interrogans* (Stimson 1907) Wenyon 1926 and *L. biflexa* (Wolbach and Binger 1914) Noguchi 1918. *Int. J. Syst. Bacteriol.* 32, 461–463.

<https://doi.org/10.1099/00207713-32-4-461>

- Fairman, J.W., Dautin, N., Wojtowicz, D., Liu, W., Noinaj, N., Barnard, T.J., Udho, E., Przytycka, T.M., Cherezov, V., Buchanan, S.K., 2012. Crystal structures of the outer membrane domain of intimin and invasin from enterohemorrhagic *E. coli* and enteropathogenic *Y. pseudotuberculosis*. *Structure* 20, 1233–1243. <https://doi.org/10.1016/j.str.2012.04.011>
- Faisal, S.M., Varma, V.P., Subathra, M., Azam, S., Sunkara, A.K., Akif, M., Baig, M.S., Chang, Y.F., 2016. *Leptospira* surface adhesin (Lsa21) induces Toll like receptor 2 and 4 mediated inflammatory responses in macrophages. *Sci. Rep.* 6, 1–14. <https://doi.org/10.1038/srep39530>
- Faisal, S.M., Yan, W., McDonough, S.P., Chang, C., Pan, M., Chang, Y., 2009. Leptosome-entrapped leptospiral antigens conferred significant higher levels of protection than those entrapped with PC-liposomes in a hamster model 27, 6537–6545. <https://doi.org/10.1016/j.vaccine.2009.08.051>
- Faisal, S.M., Yan, W.W., Chen, C.S., Palaniappan, R.U.M., McDonough, S.P., Chang, Y.F., 2008. Evaluation of protective immunity of *Leptospira* immunoglobulin like protein A (LigA) DNA vaccine against challenge in hamsters. *Vaccine* 26, 277–287. <https://doi.org/10.1016/j.vaccine.2007.10.029>
- Farhadi, T., Nezafat, N., Ghasemi, Y., Karimi, Z., 2015. Designing of Complex Multi-epitope Peptide Vaccine Based on Omps of *Klebsiella pneumoniae* : An In Silico Approach. *Int. J. Pept. Res. Ther.* 325–341. <https://doi.org/10.1007/s10989-015-9461-0>
- Fernandes, C.H., Hartwig, D.D., 2007. Evaluation of different ways of presenting LipL32 to the immune system with the aim of developing a recombinant vaccine against leptospirosis 479, 472–479. <https://doi.org/10.1139/W06-138>
- Fernandes, L.G.V., Hornsby, R.L., Nascimento, A.L.T.O., Nally, J.E., 2021. Genetic manipulation of pathogenic *Leptospira*: CRISPR interference (CRISPRi)-mediated gene silencing and rapid mutant recovery at 37 °C. *Sci. Rep.* 11, 1–12. <https://doi.org/10.1038/s41598-021-81400-7>
- Fernandes, L.G.V., Putz, E.J., Stasko, J., Lippolis, J.D., Nascimento, A.L.T.O., Nally, J.E., 2022. Evaluation of LipL32 and LigA/LigB Knockdown Mutants in *Leptospira*

- interrogans Serovar Copenhageni: Impacts to Proteome and Virulence. *Front. Microbiol.* 12, 1–14. <https://doi.org/10.3389/fmicb.2021.799012>
- FERRIS, D.H., HANSON, L.E., HOERLEIN, A.B., BEAMER, P.D., 1960. Experimental infection of white-tailed deer with *Leptospira pomona*. *Cornell Vet.* 50, 236–250.
- Figueira, C.P., Croda, J., Choy, H.A., Haake, D.A., Reis, M.G., Ko, A.I., 2011. Heterologous expression of pathogen-specific genes *ligA* and *ligB* in the saprophyte *Leptospira biflexa* confers enhanced adhesion to cultured cells and fibronectin.
- Forster, K.M., Hartwig, D.D., Seixas, F.K., Bacelo, K.L., Amaral, M., Hartleben, C.P., Dellagostin, O.A., 2013. A Conserved Region of Leptospiral Immunoglobulin-Like A and B Proteins as a DNA Vaccine Elicits a Prophylactic Immune Response against Leptospirosis. <https://doi.org/10.1128/CVI.00601-12>
- Fraga, T.R., Barbosa, A.S., Isaac, L., 2011. Leptospirosis: Aspects of innate immunity, immunopathogenesis and immune evasion from the complement system. *Scand. J. Immunol.* 73, 408–419. <https://doi.org/10.1111/j.1365-3083.2010.02505.x>
- Fraga, T.R., Courrol, S., Castiblanco-valencia, M.M., Hirata, I.Y., Vasconcellos, S.A., Juliano, L., Barbosa, A.S., Isaac, L., 2014. Immune Evasion by Pathogenic *Leptospira* Strains : The Secretion of Proteases that Directly Cleave Complement Proteins 209. <https://doi.org/10.1093/infdis/jit569>
- Gendron, K., Christe, A., Walter, S., Schweighauser, A., Francey, T., Doherr, M.G., Lang, J., 2014. Paper Serial CT features of pulmonary leptospirosis in 10 dogs. <https://doi.org/10.1136/vr.102046>
- Getzoff, E.D., Tainer, J.A., Lerner, R.A., Geysen, H.M., 1988. The Chemistry and Mechanism of Antibody Binding to Protein Antigens. *Adv. Immunol.* 43, 1–98. [https://doi.org/10.1016/S0065-2776\(08\)60363-6](https://doi.org/10.1016/S0065-2776(08)60363-6)
- Gottlieb, T., Ben-Yedidia, T., 2014. Epitope-based approaches to a universal influenza vaccine. *J. Autoimmun.* 54, 15–20. <https://doi.org/10.1016/j.jaut.2014.07.005>
- Grange, J.M., Bottasso, O., Stanford, C.A., Stanford, J.L., 2008. The use of mycobacterial adjuvant-based agents for immunotherapy of cancer. *Vaccine* 26, 4984–4990. <https://doi.org/10.1016/j.vaccine.2008.06.092>
- Guex, N., Peitsch, M.C., 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment

- for comparative protein modeling. *Electrophoresis* 18, 2714–2723. <https://doi.org/10.1002/elps.1150181505>
- Guo, Y.-J., Wang, K.-Y., Sun, S.-H., 2010. Identification of an HLA-A*0201-restricted CD8(+) T-cell epitope encoded within Leptospiral immunoglobulin-like protein A. *Microbes Infect.* 12, 364–373. <https://doi.org/10.1016/j.micinf.2010.01.010>
- Guo, Y.J., Wang, K.Y., Sun, S.H., 2010. Identification of an HLA-A*0201-restricted CD8+ T-cell epitope encoded within Leptospiral immunoglobulin-like protein A. *Microbes Infect.* 12, 364–373. <https://doi.org/10.1016/j.micinf.2010.01.010>
- Haake, D.A., Chao, G., Zuerner, R.L., Barnett, J.K., Barnett, D., Mazel, M., Matsunaga, J., Levett, P.N., Bolin, C.A., 2000. The Leptospiral Major Outer Membrane Protein LipL32 Is a Lipoprotein Expressed during Mammalian Infection 68, 2276–2285.
- Haake, D.A., Levett, P.N., 2015. Leptospirosis in humans. *Curr. Top. Microbiol. Immunol.* 387, 65–97. https://doi.org/10.1007/978-3-662-45059-8_5
- Haake, D.A., Martinich, C., Summers, T.A., Shang, E.S., Pruetz, J.A.Y.D., Coy, A.M.M.C., Mazel, M.K., Bolin, C.A., Mmun, I.N.I., 1998. Characterization of Leptospiral Outer Membrane Lipoprotein LipL36: Downregulation Associated with Late-Log-Phase Growth and Mammalian Infection 66, 1579–1587.
- Haake, D.A., Matsunaga, J., 2021. Leptospiral Immunoglobulin-Like Domain Proteins: Roles in Virulence and Immunity. *Front. Immunol.* 11, 1–21. <https://doi.org/10.3389/fimmu.2020.579907>
- Haake, D.A., Matsunaga, J., 2020. Leptospiral Immunoglobulin-Like Domain Proteins: Roles in Virulence and Immunity. *Front. Immunol.* 11, 579907. <https://doi.org/10.3389/fimmu.2020.579907>
- Haake, D.A., Matsunaga, J., Angeles, L., Angeles, L., Angeles, L., Genetics, M., Angeles, L., 2010. *Leptospira*: A Spirochete with a Hybrid Outer Membrane 77, 805–814. <https://doi.org/10.1111/j.1365-2958.2010.07262.x>.*Leptospira*
- Haake, D.A., Mazel, M.K., Coy, A.M.M.C., Milward, F., Chao, G., Matsunaga, J., Wagar, E.A., Biodivision, M., Ce, L., 1999. Leptospiral Outer Membrane Proteins OmpL1 and LipL41 Exhibit Synergistic Immunoprotection 67, 6572–6582.
- Halaby, D.M., Poupon, A., Mornon, J.P., 1999. The immunoglobulin fold family: Sequence

- analysis and 3D structure comparisons. *Protein Eng.* 12, 563–571. <https://doi.org/10.1093/protein/12.7.563>
- Harrison, N.A., Fitzgerald, W.R., 1988. Leptospirosis - can it be a sexually transmitted disease ? 163–164.
- Hartskeerl, R.A., Collares-Pereira, M., Ellis, W.A., 2011. Emergence, control and re-emerging leptospirosis: Dynamics of infection in the changing world. *Clin. Microbiol. Infect.* 17, 494–501. <https://doi.org/10.1111/j.1469-0691.2011.03474.x>
- Hartwig, D.D., Forster, K.M., Oliveira, T.L., Amaral, M., McBride, A.J.A., Dellagostin, A., 2013. A Prime-Boost Strategy Using the Novel Vaccine Candidate , LemA , Protects Hamsters against Leptospirosis 20, 747–752. <https://doi.org/10.1128/CVI.00034-13>
- He, H., Wang, W., Wu, Z., Lv, Z., Li, J., Tan, L., 2008. Protection of Guinea Pigs against *Leptospira interrogans* Serovar Lai by LipL21 DNA Vaccine 5, 385–391.
- Heo, L., Park, H., Seok, C., 2013. GalaxyRefine: Protein structure refinement driven by side-chain repacking. *Nucleic Acids Res.* 41, 384–388. <https://doi.org/10.1093/nar/gkt458>
- Hess, B., 2008. P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* 4, 116–122. <https://doi.org/10.1021/ct700200b>
- Hess, B., Kutzner, C., Van Der Spoel, D., Lindahl, E., 2008. GRGMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theory Comput.* 4, 435–447. <https://doi.org/10.1021/ct700301q>
- Heterologous, A., Maneewatch, S., Sakolvaree, Y., Saengjaruk, P., Srimanote, P., Tapchaisri, P., Tongtawe, P., Klaysing, B., Wongratanacheewin, S., 2008. Monoclonal Antibodies to LipL32 Protect 27.
- Hoke, D.E., Egan, S., Cullen, P.A., Adler, B., 2008. LipL32 is an extracellular matrix-interacting protein of *Leptospira* spp. and *Pseudoalteromonas tunicata*. *Infect. Immun.* 76, 2063–2069. <https://doi.org/10.1128/IAI.01643-07>
- Holmgren, A., Bränden, C.L., 1989. Crystal structure of chaperone protein PapD reveals an immunoglobulin fold. *Nature* 342, 248–251. <https://doi.org/10.1038/342248a0>
- Hovind-Hougen, K., 1979. Leptospiraceae, a new family to include *Leptospira* Noguchi 1917 and *Leptonema* gen. nov. *Int. J. Syst. Bacteriol.* 29, 245–251.

- <https://doi.org/10.1099/00207713-29-3-245>
- Hsia, K.-C., Li, C.-L., Yuan, H.S., 2005. Structural and functional insight into sugar-nonspecific nucleases in host defense. *Curr. Opin. Struct. Biol.* 15, 126–134. <https://doi.org/10.1016/j.sbi.2005.01.015>
- Hsieh, C.L., Ptak, C.P., Tseng, A., de Souza Suguiura, I.M., McDonough, S.P., Sritrakul, T., Li, T., Lin, Y.P., Gillilan, R.E., Oswald, R.E., Chang, Y.F., 2017. Extended low-resolution structure of a leptospira antigen offers high bactericidal antibody accessibility amenable to vaccine design. *Elife* 6, 1–25. <https://doi.org/10.7554/eLife.30051>
- Hughes, M.J.G., Moore, J.C., Lane, J.D., Wilson, R., Pribul, P.K., Younes, Z.N., Dobson, R.J., Everest, P., Reason, A.J., Redfern, J.M., Greer, F.M., Paxton, T., Panico, M., Morris, H.R., Feldman, R.G., Santangelo, J.D., 2002. Identification of major outer surface proteins of *Streptococcus agalactiae*. *Infect. Immun.* 70, 1254–1259. <https://doi.org/10.1128/IAI.70.3.1254-1259.2002>
- Hüttener, M., Hergueta, J., Bernabeu, M., Prieto, A., Aznar, S., Merino, S., Tomás, J., Juárez, A., 2022. Roles of Proteins Containing Immunoglobulin-Like Domains in the Conjugation of Bacterial Plasmids. *mSphere* 7, e0097821. <https://doi.org/10.1128/msphere.00978-21>
- Inada, R.I., Y., 1915. A report on the discovery of the causative organism (a new species of *Spirochaeta*) of Weil's disease. *Tokyo Med J* 7, 10–15. <https://doi.org/10.1007/s101560170028>
- Inada, R., Ido, Y., Hoki, R., Kaneko, R., Ito, H., 1916. THE ETIOLOGY, MODE OF INFECTION, AND SPECIFIC THERAPY OF WEIL'S DISEASE (SPIROCHAETOSIS ICTEROHAEMORRHAGICA). *J. Exp. Med.* 23, 377–402. <https://doi.org/10.1084/jem.23.3.377>
- Isogai, E., Isogai, H., Fujii, N., 1990. Macrophage Activation by Leptospiral Lipopolysaccharide 208, 200–208. [https://doi.org/10.1016/S0934-8840\(11\)80250-1](https://doi.org/10.1016/S0934-8840(11)80250-1)
- Ito, T.O.S.H.I.H.I.R.O., Yanagawa, R.Y.O., 1987. Leptospiral Attachment to Extracellular Matrix of Mouse Fibroblast (L929) Cells 15, 89–96.
- Izurieta, R., Galwankar, S., Clem, A., 2008. Leptospirosis : The “ mysterious ” mimic 21–33.
- J.D Allen, C.L.M., 1982. LEPTOSPIRURIA IN CATTLE EXPOSED TO A NATURAL

CHALLENGE 58, 93–96.

- Jablonska, J., Matelska, D., Steczkiewicz, K., Ginalski, K., 2017a. Systematic classification of the His-Me finger superfamily 45, 11479–11494. <https://doi.org/10.1093/nar/gkx924>
- Jablonska, J., Matelska, D., Steczkiewicz, K., Ginalski, K., 2017b. Systematic classification of the His-Me finger superfamily. *Nucleic Acids Res.* 45, 11479–11494. <https://doi.org/10.1093/nar/gkx924>
- James Matsunaga, 2003. Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily 49, 929–945.
- Jan-Roblero, J., García-Gómez, E., Rodríguez-Martínez, S., Cancino-Díaz, M.E., Cancino-Díaz, J.C., 2017. Surface Proteins of *Staphylococcus aureus*. *Rise Virulence Antibiot. Resist. Staphylococcus aureus*. <https://doi.org/10.5772/65976>
- Jeffery, C.J., 2016. Protein species and moonlighting proteins: Very small changes in a protein's covalent structure can change its biochemical function. *J. Proteomics* 134, 19–24. <https://doi.org/10.1016/j.jprot.2015.10.003>
- Jensen, K.K., Andreatta, M., Marcatili, P., Buus, S., Greenbaum, J.A., Yan, Z., Sette, A., Peters, B., Nielsen, M., 2018. Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* 154, 394–406. <https://doi.org/10.1111/imm.12889>
- Jeong, H., Barbe, V., Lee, C.H., Vallenet, D., Yu, D.S., Choi, S.-H., Couloux, A., Lee, S.-W., Yoon, S.H., Cattolico, L., Hur, C.-G., Park, H.-S., Séguens, B., Kim, S.C., Oh, T.K., Lenski, R.E., Studier, F.W., Daegelen, P., Kim, J.F., 2009. Genome sequences of *Escherichia coli* B strains REL606 and BL21(DE3). *J. Mol. Biol.* 394, 644–652. <https://doi.org/10.1016/j.jmb.2009.09.052>
- Jespersen, M.C., Peters, B., Nielsen, M., Marcatili, P., 2017. BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res.* 45, W24–W29. <https://doi.org/10.1093/nar/gkx346>
- Jhelum, H., Sori, H., Sehgal, D., 2018. A novel extracellular vesicle-associated endodeoxyribonuclease helps *Streptococcus pneumoniae* evade neutrophil extracellular traps and is required for full virulence. *Sci. Rep.* 1–17. <https://doi.org/10.1038/s41598-018-25865-z>

- Jones, D.T., 1999. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* 292, 195–202. <https://doi.org/10.1006/jmbi.1999.3091>
- Joosten, R.P., Long, F., Murshudov, G.N., Perrakis, A., 2014. The PDB_REDO server for macromolecular structure model optimization. *IUCrJ* 1, 213–220. <https://doi.org/10.1107/S2052252514009324>
- Jost, B.H., Adler, B., Faine, S., 1989. Experimental immunisation of hamsters with *Leptospira interrogans* 29, 115–120.
- Jung, D., Jeong, S.K., Lee, C.M., Noh, K.T., Heo, D.R., Shin, Y.K., Yun, C.H., Koh, W.J., Akira, S., Whang, J., Kim, H.J., Park, W.S., Shin, S.J., Park, Y.M., 2011. Enhanced efficacy of therapeutic cancer vaccines produced by Co-treatment with mycobacterium tuberculosis heparin-binding hemagglutinin, a novel TLR4 agonist. *Cancer Res.* 71, 2858–2870. <https://doi.org/10.1158/0008-5472.CAN-10-3487>
- Karen V Evangelista & Jenifer Coburn, 2010. *Leptospira* as an emerging pathogen: a review of its biology, pathogenesis and host immune responses. *Future Microbiol.* 1413–1425. <https://doi.org/10.1201/9781420052275-7>
- Khrustalev, V.V., Barkovsky, E.V., Khrustaleva, T.A., 2016. Magnesium and manganese binding sites on proteins have the same predominant motif of secondary structure. *J. Theor. Biol.* 395, 174–185. <https://doi.org/10.1016/j.jtbi.2016.02.006>
- Kim, D.E., Chivian, D., Baker, D., 2004. Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.* 32, 526–531. <https://doi.org/10.1093/nar/gkh468>
- Ko, A.I., Goarant, C., Picardeau, M., 2009. *Leptospira* : the dawn of the molecular genetics era for an emerging zoonotic pathogen. *Nat. Rev. Microbiol.* 7. <https://doi.org/10.1038/nrmicro2208>
- Koebnik, R., Locher, K.P., Van Gelder, P., 2000. Structure and function of bacterial outer membrane proteins: Barrels in a nutshell. *Mol. Microbiol.* 37, 239–253. <https://doi.org/10.1046/j.1365-2958.2000.01983.x>
- Koizumi, N., Watanabe, H., 2004a. Leptospiral immunoglobulin-like proteins elicit protective immunity. *Vaccine* 22, 1545–1552. <https://doi.org/10.1016/j.vaccine.2003.10.007>
- Koizumi, N., Watanabe, H., 2004b. Leptospiral immunoglobulin-like proteins elicit protective immunity. *Vaccine* 22, 1545–1552. <https://doi.org/10.1016/j.vaccine.2003.10.007>

- Koizumi, N., Watanabe, H., 2003. Molecular cloning and characterization of a novel leptospiral lipoprotein with OmpA domain 226, 215–219. [https://doi.org/10.1016/S0378-1097\(03\)00619-0](https://doi.org/10.1016/S0378-1097(03)00619-0)
- Koizumi, N., Watanabe, H., Watanabe, H., 2005. Leptospirosis vaccines: Past, present, and future.
- Kozakov, D., Hall, D.R., Xia, B., Porter, K.A., Padhorny, D., Yueh, C., Beglov, D., Vajda, S., Biology, Q., 2017. The ClusPro web server for protein-protein docking. *Nat. Protoc.* 12, 255–278. <https://doi.org/10.1038/nprot.2016.169>.The
- Kreikemeyer, B., Klenk, M., Podbielski, A., 2004. The intracellular status of *Streptococcus pyogenes*: Role of extracellular matrix-binding proteins and their regulation. *Int. J. Med. Microbiol.* 294, 177–188. <https://doi.org/10.1016/j.ijmm.2004.06.017>
- Kringelum, J.V., Lundegaard, C., Lund, O., Nielsen, M., 2012. Reliable B Cell Epitope Predictions: Impacts of Method Development and Improved Benchmarking. *PLoS Comput. Biol.* 8. <https://doi.org/10.1371/journal.pcbi.1002829>
- Kumar, A., Varma, V.P., Sridhar, K., Abdullah, M., Vyas, P., 2021. Deciphering the Role of *Leptospira* Surface Protein LigA in Modulating the Host Innate Immune Response 12, 1–18. <https://doi.org/10.3389/fimmu.2021.807775>
- Kumar, P., Lata, S., Shankar, U.N., Akif, M., 2021. Immunoinformatics-Based Designing of a Multi-Epitope Chimeric Vaccine From Multi-Domain Outer Surface Antigens of *Leptospira* 12, 1–14. <https://doi.org/10.3389/fimmu.2021.735373>
- Kumari, A., Premlatha, M.M., Raja, V., Solomon, C., Mercy, A., 2017. Original Article Protective immunity of recombinant LipL21 and I-LipL21 against *Leptospira interrogans* serovar Autumnalis N2 infection. <https://doi.org/10.3855/jidc.9545>
- Kuriakose, M., Eapen, C.K., Paul, R., 1997. Leptospirosis in Kolenchery , Kerala , India : Epidemiology , prevalent local serogroups and servoars and a new serovar 691–697.
- Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* 26, 283–291. <https://doi.org/10.1107/s0021889892009944>
- Lau, C.L., Watson, C.H., Lowry, J.H., David, M.C., Craig, S.B., Wynwood, S.J., Kama, M., Nilles, E.J., 2016. Human Leptospirosis Infection in Fiji : An Eco- epidemiological

- Approach to Identifying Risk Factors and Environmental Drivers for Transmission 1–25.
<https://doi.org/10.1371/journal.pntd.0004405>
- Lehmann, J.S., Matthias, M.A., Vinetz, J.M., Fouts, D.E., 2014. Leptospiral Pathogenomics 280–308. <https://doi.org/10.3390/pathogens3020280>
- Levett, 2001. Leptospirosis. *Eur. Soc. Clin. Infect. Dis.* 17, 494–501.
<https://doi.org/10.1128/CMR.14.2.296>
- Levett, P.N., 2015. Systematics of leptospiraceae. *Curr. Top. Microbiol. Immunol.* 387, 11–20.
https://doi.org/10.1007/978-3-662-45059-8_2
- Li, C.-L., Hor, L.-I., Chang, Z.-F., Tsai, L.-C., Yang, W.-Z., Yuan, H.S., 2003. DNA binding and cleavage by the periplasmic nuclease Vvn: a novel structure with a known active site. *EMBO J.* 22, 4014–4025. <https://doi.org/10.1093/emboj/cdg377>
- Libonati, H.A., Santos, G.B., Souza, G.N., Brandão, F.Z., Lilenbaum, W., 2018. Leptospirosis is strongly associated to estrus repetition on cattle.
- Lilenbaum, W., Martins, G., 2014. Leptospirosis in Cattle : A Challenging Scenario for the Understanding of the Epidemiology 61, 63–68. <https://doi.org/10.1111/tbed.12233>
- Lin, Y.-P., Lee, D.-W., McDonough, S.P., Nicholson, L.K., Sharma, Y., Chang, Y.-F., 2009. Repeated domains of leptospira immunoglobulin-like proteins interact with elastin and tropoelastin. *J. Biol. Chem.* 284, 19380–19391. <https://doi.org/10.1074/jbc.M109.004531>
- Lin, Y., Greenwood, A., Nicholson, L.K., Sharma, Y., Mcdonough, S.P., Chang, Y., 2009a. Fibronectin Binds to and Induces Conformational Change in a Disordered Region of Leptospiral Immunoglobulin-like Protein B * □ 284, 23547–23557. <https://doi.org/10.1074/jbc.M109.031369>
- Lin, Y., Lee, D., Mcdonough, S.P., Nicholson, L.K., Sharma, Y., Chang, Y., 2009b. Repeated Domains of Leptospira Immunoglobulin-like Proteins Interact with Elastin and Tropoelastin * □ 284, 19380–19391. <https://doi.org/10.1074/jbc.M109.004531>
- Lin, Y.P., Greenwood, A., Yan, W.W., Nicholson, L.K., Sharma, Y., McDonough, S.P., Chang, Y.F., 2009. A novel fibronectin type III module binding motif identified on C-terminus of *Leptospira* immunoglobulin-like protein, LigB. *Biochem. Biophys. Res. Commun.* 389, 57–62. <https://doi.org/10.1016/j.bbrc.2009.08.089>

- Lin, Y.P., Mcdonough, S.P., Sharma, Y., Chang, Y.F., 2010. The terminal immunoglobulin-like repeats of LigA and LigB of *Leptospira* enhance their binding to gelatin binding domain of fibronectin and host cells. *PLoS One* 5. <https://doi.org/10.1371/journal.pone.0011301>
- Lin, Y.P., Raman, R., Sharma, Y., Chang, Y.F., 2008a. Calcium binds to leptospiral immunoglobulin-like protein, LigB, and modulates fibronectin binding. *J. Biol. Chem.* 283, 25140–25149. <https://doi.org/10.1074/jbc.M801350200>
- Lin, Y.P., Raman, R., Sharma, Y., Chang, Y.F., 2008b. Calcium binds to leptospiral immunoglobulin-like protein, LigB, and modulates fibronectin binding. *J. Biol. Chem.* 283, 25140–25149. <https://doi.org/10.1074/jbc.M801350200>
- Lindahl, G., Stålhammar-Carlemalm, M., Areschoug, T., 2005. Surface proteins of *Streptococcus agalactiae* and related proteins in other bacterial pathogens. *Clin. Microbiol. Rev.* 18, 102–127. <https://doi.org/10.1128/CMR.18.1.102-127.2005>
- Liu, G.Y., 2009. Molecular pathogenesis of *Staphylococcus aureus* infection. *Pediatr. Res.* 65, 71R–77R. <https://doi.org/10.1203/PDR.0b013e31819dc44d>
- Liu, Y., Zheng, W., Li, L., Mao, Y., Yan, J., 2007. Pathogenesis of leptospirosis: interaction of *Leptospira interrogans* with in vitro cultured mammalian cells. *Med. Microbiol. Immunol.* 196, 233–239. <https://doi.org/10.1007/s00430-007-0047-0>
- Livingston, B., Crimi, C., Newman, M., Higashimoto, Y., Appella, E., Sidney, J., Sette, A., 2002a. A rational strategy to design multiepitope immunogens based on multiple Th lymphocyte epitopes. *J. Immunol.* 168, 5499–5506. <https://doi.org/10.4049/jimmunol.168.11.5499>
- Livingston, B., Crimi, C., Newman, M., Higashimoto, Y., Appella, E., Sidney, J., Sette, A., 2002b. A Rational Strategy to Design Multiepitope Immunogens Based on Multiple Th Lymphocyte Epitopes. *J. Immunol.* 168, 5499–5506. <https://doi.org/10.4049/jimmunol.168.11.5499>
- Luo, D., Xue, F., Ojcius, D.M., Zhao, J., Mao, Y., Li, L., Lin, X., Yan, J., 2009. Protein typing of major outer membrane lipoproteins from Chinese pathogenic *Leptospira* spp. and characterization of their immunogenicity. *Vaccine* 28, 243–255. <https://doi.org/10.1016/j.vaccine.2009.09.089>

- Mani, M., Chen, C., Amblee, V., Liu, H., Mathur, T., Zwicke, G., Zabad, S., Patel, B., Thakkar, J., Jeffery, C.J., 2015a. MoonProt: A database for proteins that are known to moonlight. *Nucleic Acids Res.* 43, D277–D282. <https://doi.org/10.1093/nar/gku954>
- Mani, M., Chen, C., Amblee, V., Liu, H., Mathur, T., Zwicke, G., Zabad, S., Patel, B., Thakkar, J., Jeffery, C.J., 2015b. MoonProt: a database for proteins that are known to moonlight. *Nucleic Acids Res.* 43, D277–82. <https://doi.org/10.1093/nar/gku954>
- Matsunaga, J., Barocchi, M.A., Croda, J., Young, T.A., Sanchez, Y., Siqueira, I., Bolin, C.A., Reis, M.G., Riley, L.W., Haake, D.A., Ko, A.I., 2003a. Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily 49, 929–945. <https://doi.org/10.1046/j.1365-2958.2003.03619.x>
- Matsunaga, J., Barocchi, M.A., Croda, J., Young, T.A., Sanchez, Y., Siqueira, I., Bolin, C.A., Reis, M.G., Riley, L.W., Haake, D.A., Ko, A.I., 2003b. Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily. *Mol. Microbiol.* 49, 929–946. <https://doi.org/10.1046/j.1365-2958.2003.03619.x>
- Matsunaga, J., Barocchi, M.A., Croda, J., Young, T.A., Sanchez, Y., Siqueira, I., Bolin, C.A., Reis, M.G., Riley, L.W., Haake, D.A., Ko, A.I., 2003c. Pathogenic *Leptospira* species express surface-exposed proteins belonging to the bacterial immunoglobulin superfamily. *Mol. Microbiol.* 49, 929–945. <https://doi.org/10.1046/j.1365-2958.2003.03619.x>
- Matsunaga, J., Lo, M., Bulach, D.M., Zuerner, R.L., Adler, B., Haake, D.A., 2007. Response of *Leptospira interrogans* to Physiologic Osmolarity: Relevance in Signaling the Environment-to-Host Transition □ † 75, 2864–2874. <https://doi.org/10.1128/IAI.01619-06>
- Matsunaga, J., Sanchez, Y., Xu, X., Haake, D.A., 2005a. Osmolarity , a Key Environmental Signal Controlling Expression of Leptospiral Proteins LigA and LigB and the Extracellular Release of LigA 73, 70–78. <https://doi.org/10.1128/IAI.73.1.70>
- Matsunaga, J., Sanchez, Y., Xu, X., Haake, D.A., 2005b. Osmolarity, a key environmental signal controlling expression of leptospiral proteins LigA and LigB and the extracellular release of LigA. *Infect. Immun.* 73, 70–78. <https://doi.org/10.1128/IAI.73.1.70-78.2005>
- Matsunaga, J., Schlax, P.J., Haake, A., 2013. Role for cis -Acting RNA Sequences in the Temperature-Dependent Expression of the Multiadhesive Lig Proteins in *Leptospira*

- interrogans. <https://doi.org/10.1128/JB.00663-13>
- Matsuo, K., 2000. Control of Immunologically Crossreactive Leptospiral Infection by Administration Lipopolysaccharides from a Nonpathogenic Strain of *Leptospira biflexa* 44, 887–890.
- McCoy, A.J., 2006. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. Sect. D Biol. Crystallogr.* 63, 32–41. <https://doi.org/10.1107/S09074444906045975>
- McLoughlin, R.M., Lee, J.C., Kasper, D.L., Tzianabos, A.O., 2008. IFN- γ Regulated Chemokine Production Determines the Outcome of *Staphylococcus aureus* Infection . *J. Immunol.* 181, 1323–1332. <https://doi.org/10.4049/jimmunol.181.2.1323>
- Medicina, F. De, Informa, D. De, Medicina, F. De, Sa, U. De, Insuficie, G. De, Nefrologia, D. De, Medicina, F. De, Sa, U. De, 1999. RISK FACTORS FOR DEATH AND CHANGING PATTERNS IN LEPTOSPIROSIS ACUTE 61, 630–634.
- Mei, S., Zhang, J., Zhang, X., Tu, X., 2015a. Biochemical and Biophysical Research Communications Solution structure of leptospiral LigA4 Big domain. *Biochem. Biophys. Res. Commun.* 467, 288–292. <https://doi.org/10.1016/j.bbrc.2015.09.170>
- Mei, S., Zhang, J., Zhang, X., Tu, X., 2015b. Solution structure of leptospiral LigA4 Big domain. *Biochem. Biophys. Res. Commun.* 467, 288–292. <https://doi.org/10.1016/j.bbrc.2015.09.170>
- Mei, S., Zhang, J., Zhang, X., Tu, X., 2014. STRUCTURE NOTE Solution structure of a bacterial immunoglobulin-like domain of the outer membrane protein (LigB) from *Leptospira* 195–200. <https://doi.org/10.1002/prot.24723>
- Meri, T., Murgia, R., Stefanel, P., Meri, S., Cinco, M., 2005. Regulation of complement activation at the C3-level by serum resistant leptospires 39, 139–147. <https://doi.org/10.1016/j.micpath.2005.07.003>
- Michna, S.W., Campbell, R.S., 1969. The isolation of *Leptospira sejroe* from the kidneys of aborting cattle. *Vet. Rec.* 84, 83–86. <https://doi.org/10.1136/vr.84.4.83>
- Microhrology, G., Britain, G., 1986. Ultrastructure and Chemical Composition of Lipopolysaccharide Extracted from *Leptospira interrogans* serovar *copenhageni*.

- Murgia, R., Garcia, R., Cinco, M., 2002. Leptospire Are Killed In Vitro by Both Oxygen-Dependent and -Independent Reactions 70, 7172–7175. <https://doi.org/10.1128/IAI.70.12.7172>
- Murray, G.L., Srikram, A., Henry, R., Puapairoj, A., Sermswan, R.W., Adler, B., 2009. *Leptospira interrogans* requires heme oxygenase for disease pathogenesis. *Microbes Infect.* 11, 311–314. <https://doi.org/10.1016/j.micinf.2008.11.014>
- Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., Vagin, A.A., 2011. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D. Biol. Crystallogr.* 67, 355–367. <https://doi.org/10.1107/S0907444911001314>
- Mwachui, M.A., Crump, L., Hartskeerl, R., Zinsstag, J., 2015. Environmental and Behavioural Determinants of Leptospirosis Transmission: A Systematic Review 1–15. <https://doi.org/10.1371/journal.pntd.0003843>
- Naiman, B. M., Alt, D., Bolin, C.A., Zuerner, R., Baldwin, C.L., 2001. Protective killed *Leptospira borgpetersenii* vaccine induces potent Th1 immunity comprising responses by CD4 and $\gamma\delta$ T lymphocytes. *Infect. Immun.* 69, 7550–7558. <https://doi.org/10.1128/IAI.69.12.7550-7558.2001>
- Naiman, B M, Alt, D., Bolin, C.A., Zuerner, R., Baldwin, C.L., 2001. Protective killed *Leptospira borgpetersenii* vaccine induces potent Th1 immunity comprising responses by CD4 and gammadelta T lymphocytes. *Infect. Immun.* 69, 7550–7558. <https://doi.org/10.1128/IAI.69.12.7550-7558.2001>
- Nally, J.E., Whitelegge, J.P., Bassilian, S., Blanco, D.R., Lovett, M.A., 2007. Characterization of the Outer Membrane Proteome of *Leptospira interrogans* Expressed during Acute Lethal Infection □ 75, 766–773. <https://doi.org/10.1128/IAI.00741-06>
- Nascimento, A.L.T.O., Ko, A.I., Martins, E.A.L., Ho, P.L., Haake, D.A., Hartskeerl, R.A., Marques, M. V, Oliveira, M.C., Menck, C.F.M., Leite, L.C.C., Carrer, H., Coutinho, L.L., Degraeve, W.M., Dellagostin, O.A., Ferro, E.S., Ferro, M.I.T., Furlan, L.R., Gamberini, M., Giglioti, E.A., Lemos, M.V.F., Marino, C.L., Nunes, L.R., Oliveira, R.C. De, Pereira, G.G., Reis, M.S., Gene, E., 2004. Comparative Genomics of Two *Leptospira interrogans* Serovars Reveals Novel Insights into Physiology and Pathogenesis † 186, 2164–2172. <https://doi.org/10.1128/JB.186.7.2164>

- Nezafat, N., Ghasemi, Y., Javadi, G., Khoshnoud, M.J., Omidinia, E., 2014a. A novel multi-epitope peptide vaccine against cancer: an in silico approach. *J. Theor. Biol.* 349, 121–134. <https://doi.org/10.1016/j.jtbi.2014.01.018>
- Nezafat, N., Ghasemi, Y., Javadi, G., Khoshnoud, M.J., Omidinia, E., 2014b. A novel multi-epitope peptide vaccine against cancer: An in silico approach. *J. Theor. Biol.* 349, 121–134. <https://doi.org/10.1016/j.jtbi.2014.01.018>
- Niemann, H.H., Schubert, W.D., Heinz, D.W., 2004. Adhesins and invasins of pathogenic bacteria: A structural view. *Microbes Infect.* 6, 101–112. <https://doi.org/10.1016/j.micinf.2003.11.001>
- No, B., Grote, A., Hiller, K., Scheer, M., Mu, R., Hempel, D.C., Jahn, D., 2005. JCat : a novel tool to adapt codon usage of a target gene to its potential expression host 33, 526–531. <https://doi.org/10.1093/nar/gki376>
- Odir, A.D., Grassmann, A.A., Hartwig, D.D., Félix, S.R., Da Silva, É.F., McBride, A.J.A., 2011. Recombinant vaccines against leptospirosis. *Hum. Vaccin.* 7, 1215–1224. <https://doi.org/10.4161/hv.7.11.17944>
- Oliveira, T.R., Longhi, M.T., Gonc, A.P., 2010. LipL53 , a temperature regulated protein from *Leptospira interrogans* that binds to extracellular matrix molecules 12. <https://doi.org/10.1016/j.micinf.2009.12.004>
- Oostenbrink, C., Villa, A., Mark, A.E., Gunsteren, W.F.V.A.N., 2004. A Biomolecular Force Field Based on the Free Enthalpy of Hydration and Solvation : The GROMOS Force-Field Parameter Sets 53A5 and 53A6. <https://doi.org/10.1002/jcc.20090>
- Palaniappan, R.U.M., Chang, Y.-F., Jusuf, S.S.D., Artiushin, S., Timoney, J.F., McDonough, S.P., Barr, S.C., Divers, T.J., Simpson, K.W., McDonough, P.L., Mohammed, H.O., 2002a. Cloning and molecular characterization of an immunogenic LigA protein of *Leptospira interrogans*. *Infect. Immun.* 70, 5924–5930. <https://doi.org/10.1128/IAI.70.11.5924-5930.2002>
- Palaniappan, R.U.M., Chang, Y., Jusuf, S.S.D., Artiushin, S., Timoney, J.F., McDonough, S.P., Barr, S.C., Divers, T.J., Simpson, K.W., McDonough, P.L., Mohammed, H.O., 2002b. Cloning and Molecular Characterization of an Immunogenic LigA Protein of *Leptospira interrogans* 70, 5924–5930. <https://doi.org/10.1128/IAI.70.11.5924>

- Palaniappan, Raghavan U M, Ramanujam, S., Chang, Y.-F., 2007. Leptospirosis: pathogenesis, immunity, and diagnosis. *Curr. Opin. Infect. Dis.* 20, 284–292. <https://doi.org/10.1097/QCO.0b013e32814a5729>
- Palaniappan, Raghavan U.M., Ramanujam, S., Chang, Y.F., 2007. Leptospirosis: Pathogenesis, immunity, and diagnosis. *Curr. Opin. Infect. Dis.* 20, 284–292. <https://doi.org/10.1097/QCO.0b013e32814a5729>
- Pappachan, M.J., Mathew, S., Aravindan, K.P., Khader, A., Bharghavan, P. V, Kareem, M.M.A., Tuteja, U., Shukla, J., Batra, H. V, 2004. Risk factors for mortality in patients with leptospirosis during an epidemic in northern Kerala. *Natl. Med. J. India* 17, 240–242.
- Park, B.S., Song, D.H., Kim, H.M., Choi, B., Lee, H., Lee, J., 2009. The structural basis of lipopolysaccharide recognition by the TLR4 – MD-2 complex. *Nature* 458, 1191–1196. <https://doi.org/10.1038/nature07830>
- Parrinello, M., Rahman, A., 1981. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* 52, 7182–7190. <https://doi.org/10.1063/1.328693>
- Petrakovsky, J., Bianchi, A., Fisun, H., Nájera-aguilar, P., Pereira, M.M., 2014. Animal Leptospirosis in Latin America and the Caribbean Countries : Reported Outbreaks and Literature Review 10770–10789. <https://doi.org/10.3390/ijerph111010770>
- Philip, N., Affendy, N.B., Nur, S., Ramli, A., Arif, M., Raja, P., Nagandran, E., Renganathan, P., Taib, N.M., Masri, N., Yuhana, M.Y., Thian, L., Than, L., Seganathirajah, M., Id, C.G., Goris, M.G.A., Sekawi, Z., Kumari, V., Id, N., 2020. *Leptospira interrogans* and *Leptospira kirschneri* are the dominant *Leptospira* species causing human leptospirosis in Central Malaysia 1–14. <https://doi.org/10.1371/journal.pntd.0008197>
- Picardeau, M., 2017. Virulence of the zoonotic agent of leptospirosis: still terra incognita? *Mathieu. Nat. Publ. Gr.* <https://doi.org/10.1038/nrmicro.2017.5>
- Pinne, M., Matsunaga, J., Haake, D.A., 2012. Leptospiral outer membrane protein microarray, a novel approach to identification of host ligand-binding proteins. *J. Bacteriol.* 194, 6074–6087. <https://doi.org/10.1128/JB.01119-12>
- Pizarro-Cerdá, J., Cossart, P., 2006. Subversion of cellular functions by *Listeria monocytogenes*. *J. Pathol.* 208, 215–223. <https://doi.org/10.1002/path.1888>

- Ponomarenko, J., Bui, H.H., Li, W., Fusseder, N., Bourne, P.E., Sette, A., Peters, B., 2008. ElliPro: A new structure-based tool for the prediction of antibody epitopes. *BMC Bioinformatics* 9, 1–8. <https://doi.org/10.1186/1471-2105-9-514>
- Prasadarao, N. V., Wass, C.A., Weiser, J.N., Stins, M.F., Huang, S.H.E., Kim, K.S., 1996. Outer membrane protein A of *Escherichia coli* contributes to invasion of brain microvascular endothelial cells. *Infect. Immun.* 64, 146–153. <https://doi.org/10.1128/iai.64.1.146-153.1996>
- Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., Shirts, M.R., Smith, J.C., Kasson, P.M., Van Der Spoel, D., Hess, B., Lindahl, E., 2013. GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* 29, 845–854. <https://doi.org/10.1093/bioinformatics/btt055>
- Ptak, C.P., 2014. NMR Solution Structure of the Terminal Immunoglobulin-like Domain from the *Leptospira* Host-Interacting Outer Membrane Protein, LigB.
- Ptak, C.P., Akif, M., Hsieh, C.L., Devarajan, A., He, P., Xu, Y., Oswald, R.E., Chang, Y.F., 2019. Comparative screening of recombinant antigen thermostability for improved leptospirosis vaccine design. *Biotechnol. Bioeng.* 116, 260–271. <https://doi.org/10.1002/bit.26864>
- Ptak, Christopher P, Hsieh, C., Lin, Y., Maltsev, A.S., Raman, R., Sharma, Y., Oswald, R.E., Chang, Y., 2014. NMR Solution Structure of the Terminal Immunoglobulin-like Domain from the *Leptospira* Host-Interacting Outer Membrane Protein, LigB.
- Ptak, Christopher P., Hsieh, C.L., Lin, Y.P., Maltsev, A.S., Raman, R., Sharma, Y., Oswald, R.E., Chang, Y.F., 2014. NMR solution structure of the terminal immunoglobulin-like domain from the *leptospira* host-interacting outer membrane protein, LigB. *Biochemistry* 53, 5249–5260. <https://doi.org/10.1021/bi500669u>
- Putz, E.J., Nally, J.E., 2020. Investigating the Immunological and Biological Equilibrium of Reservoir Hosts and Pathogenic *Leptospira*: Balancing the Solution to an Acute Problem? *Front. Microbiol.* 11. <https://doi.org/10.3389/fmicb.2020.02005>
- Raman, R., Rajanikanth, V., Palaniappan, R.U.M., Lin, Y.P., He, H., McDonough, S.P., Sharma, Y., Chang, Y.F., 2010. Big domains are novel Ca²⁺-binding modules: Evidences from big domains of *Leptospira* immunoglobulin-like (Lig) proteins. *PLoS One* 5.

- <https://doi.org/10.1371/journal.pone.0014377>
- Rapin, N., Lund, O., Bernaschi, M., Castiglione, F., 2010. Computational Immunology Meets Bioinformatics : The Use of Prediction Tools for Molecular Binding in the Simulation of the Immune System 5. <https://doi.org/10.1371/journal.pone.0009862>
- Ren, S.X., Fu, G., Jiang, X.G., Zeng, R., Miao, Y.G., Xu, H., Zhang, Y.X., Xiong, H., Lu, G., Lu, L.F., Jiang, H.Q., Jia, J., Tu, Y.F., Jiang, J.X., Gu, W.Y., Zhang, Y.Q., Cai, Z., Sheng, H.H., Yin, H.F., Zhang, Y., Zhu, G.F., Wan, M., Huang, H.L., Qian, Z., Wang, S.Y., Ma, W., Yao, Z.J., Shen, Y., Qiang, B.Q., Xia, Q.C., Guo, X.K., Danchin, A., Saint Girons, I., Somerville, R.L., Wen, Y.M., Shi, M.H., Chen, Z., Xu, J.G., Zhao, G.P., 2003. Unique physiological and pathogenic features of *Leptospira interrogans* revealed by whole-genome sequencing. *Nature* 422, 888–893. <https://doi.org/10.1038/nature01597>
- Rino Rappuoli, 2000. Reverse vaccinology 445–450.
- Rodrigues de Oliveira, N., Jorge, S., Andrade Colares Maia, M., Thurow Bunde, T., Kurz Pedra, A.C., Pinto Seixas Neto, A.C., Larré Oliveira, T., Dellagostin, O.A., 2021a. Protective efficacy of whole-cell inactivated *Leptospira* vaccines made using virulent or avirulent strains in a hamster model. *Vaccine* 39, 5626–5634. <https://doi.org/10.1016/j.vaccine.2021.08.014>
- Rodrigues de Oliveira, N., Jorge, S., Andrade Colares Maia, M., Thurow Bunde, T., Kurz Pedra, A.C., Pinto Seixas Neto, A.C., Larré Oliveira, T., Dellagostin, O.A., 2021b. Protective efficacy of whole-cell inactivated *Leptospira* vaccines made using virulent or avirulent strains in a hamster model. *Vaccine* 39, 5626–5634. <https://doi.org/10.1016/j.vaccine.2021.08.014>
- Roy, A., Kucukural, A., Zhang, Y., 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* 5, 725–738. <https://doi.org/10.1038/nprot.2010.5>
- Sambasiva, R.R., Naveen, G., 2003. Leptospirosis in India and the Rest of the World 7, 178–193.
- Samrot, A. V., Sean, T.C., Bhavya, K.S., Sahithya, C.S., Chandrasekaran, S., Palanisamy, R., Robinson, E.R., Subbiah, S.K., Mok, P.L., 2021. Leptospiral infection, pathogenesis and its diagnosis—a review. *Pathogens* 10, 1–30. <https://doi.org/10.3390/pathogens10020145>

- Saul, F., 1973. Three-Dimensional Structure of the Fab' Fragment of a Human Immunoglobulin at 2.8-A Resolution 0 70, 3305–3310. <https://doi.org/10.1073/pnas.70.12.3305>
- Scharrig, E., Carestia, A., Ferrer, M.F., Cédola, M., Pretre, G., 2015. Neutrophil Extracellular Traps are Involved in the Innate Immune Response to Infection with *Leptospira* 1–18. <https://doi.org/10.1371/journal.pntd.0003927>
- Sen, P., Fatima, S., Ahmad, B., Khan, R.H., 2009. Spectrochimica Acta Part A : Molecular and Biomolecular Spectroscopy Interactions of thioflavin T with serum albumins : Spectroscopic analyses 74, 94–99. <https://doi.org/10.1016/j.saa.2009.05.010>
- Seo, K.S., Kim, J.W., Park, J.Y., Viall, A.K., Minnich, S.S., Rohde, H.N., Schnider, D.R., Lim, S.Y., Hong, J.B., Hinnebusch, B.J., O'Loughlin, J.L., Deobald, C.F., Bohach, G.A., Hovde, C.J., Minnich, S.A., 2012. Role of a new intimin/invasin-like protein in *Yersinia pestis* virulence. *Infect. Immun.* 80, 3559–3569. <https://doi.org/10.1128/IAI.00294-12>
- Seper, A., Hosseinzadeh, A., Gorkiewicz, G., Lichtenegger, S., Roier, S., Grutsch, A., Reidl, J., Urban, C.F., Leitner, D.R., Ro, M., Schild, S., 2013. *Vibrio cholerae* Evades Neutrophil Extracellular Traps by the Activity of Two Extracellular Nucleases 9. <https://doi.org/10.1371/journal.ppat.1003614>
- Serruto, D., Bottomley, M.J., Ram, S., Giuliani, M.M., Rappuoli, R., 2012. The new multicomponent vaccine against meningococcal serogroup B , 4CMenB : Immunological , functional and structural characterization of the antigens. *Vaccine* 30, B87–B97. <https://doi.org/10.1016/j.vaccine.2012.01.033>
- Sethi, S., Sharma, N., Kakkar, N., Taneja, J., Chatterjee, S.S., Singh, S., Sharma, M., 2010. Increasing Trends of Leptospirosis in Northern India : A Clinico-Epidemiological Study 4, 1–7. <https://doi.org/10.1371/journal.pntd.0000579>
- Shang, E.S., Summers, T.A., 1996. Molecular Cloning and Sequence Analysis of the Gene Encoding LipL41 , a Surface-Exposed Lipoprotein of Pathogenic *Leptospira* Species 64, 2322–2330.
- Shiraz, M., Lata, S., Kumar, P., Shankar, U.N., Akif, M., 2021. Immunoinformatics analysis of antigenic epitopes and designing of a multi-epitope peptide vaccine from putative nitro-reductases of *Mycobacterium tuberculosis* DosR. *Infect. Genet. Evol. J. Mol.*

- Epidemiol. Evol. Genet. Infect. Dis. 94, 105017.
<https://doi.org/10.1016/j.meegid.2021.105017>
- Shivakumar, S., 2008. Leptospirosis - Current Scenario in India.
- Sigel, R.K.O., 2005. Group II Intron Ribozymes and Metal Ions – A Delicate Relationship 2281–2292. <https://doi.org/10.1002/ejic.200401007>
- Sigel, R.K.O., Pyle, A.M., 2007. Alternative roles for metal ions in enzyme catalysis and the implications for ribozyme chemistry. Chem. Rev. 107, 97–113.
<https://doi.org/10.1021/cr0502605>
- Silva, É.F., Medeiros, M.A., McBride, A.J.A., Matsunaga, J., Esteves, G.S., Ramos, J.G.R., Santos, C.S., Croda, J., Homma, A., Dellagostin, O.A., Haake, D.A., Reis, M.G., Ko, A.I., 2007a. The terminal portion of leptospiral immunoglobulin-like protein LigA confers protective immunity against lethal infection in the hamster model of leptospirosis. Vaccine 25, 6277–6286. <https://doi.org/10.1016/j.vaccine.2007.05.053>
- Silva, É.F., Medeiros, M.A., McBride, A.J.A., Matsunaga, J., Esteves, G.S., Ramos, J.G.R., Santos, C.S., Croda, J., Homma, A., Dellagostin, O.A., Haake, D.A., Reis, M.G., Ko, A.I., 2007b. The terminal portion of leptospiral immunoglobulin-like protein LigA confers protective immunity against lethal infection in the hamster model of leptospirosis. Vaccine 25, 6277–6286. <https://doi.org/10.1016/j.vaccine.2007.05.053>
- Singh, H., Raghava, G.P.S., 2003. ProPred1: Prediction of promiscuous MHC class-I binding sites. Bioinformatics 19, 1009–1014. <https://doi.org/10.1093/bioinformatics/btg108>
- Singh, H., Raghava, G.P.S., 2002. ProPred: Prediction of HLA-DR binding sites. Bioinformatics 17, 1236–1237. <https://doi.org/10.1093/bioinformatics/17.12.1236>
- Souza, D., Silveira, M.M., Conrad, N.L., Mendonça, K.S., Santos, C.S., Athanazio, D.A., Medeiros, A., Reis, M.G., Dellagostin, O.A., McBride, A.J.A., 2017. LigB subunit vaccine confers sterile immunity against challenge in the hamster model of leptospirosis 1–20.
- Stein, C., 2007. The Global Burden of Disease Assessments — WHO Is Responsible ? PLoS Negl. Trop. Dis. 1. <https://doi.org/10.1371/journal.pntd.0000161>
- Stevenson, B., Choy, H.A., Pinne, M., Rotondi, M.L., Miller, M.C., Demoll, E., Kraiczy, P., Anne, E., 2007. Leptospira interrogans Endostatin-Like Outer Membrane Proteins Bind Host Fibronectin , Laminin and Regulators of Complement.

- <https://doi.org/10.1371/journal.pone.0001188>
- Stimson, A.M., 1907. Note on an Organism Found in Yellow-Fever Tissue. *Public Heal. Reports* 22, 541. <https://doi.org/10.2307/4559008>
- Storisteanu, D.M.L., Pocock, J.M., Cowburn, A.S., Juss, J.K., Nadesalingam, A., Nizet, V., Chilvers, E.R., 2017. Evasion of Neutrophil Extracellular Traps by Respiratory Pathogens. *Am. J. Respir. Cell Mol. Biol.* 56, 423–431. <https://doi.org/10.1165/rcmb.2016-0193PS>
- Sudipto Saha and G. P. S. Raghava, 2007. Prediction of Continuous B-Cell Epitopes in an Antigen Using Recurrent Neural Network 48, 659–664. <https://doi.org/10.1002/prot>
- Sulzer, K.R., Rogers, F., 2005. Deoxyribonucleic Acid Relatedness between Serogroups and Serovars in the Family 37.
- Taylor, P., Schollum, L.M., Blackmore, D.K., 2011. The serological and cultural prevalence of leptospirosis in a sample of feral goats 1–4. <https://doi.org/10.1080/00480169.1981.34813>
- Taylor, R.G., Walker, D.C., McInnes, R.R., 1993. E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing. *Nucleic Acids Res.* 21, 1677–1678. <https://doi.org/10.1093/nar/21.7.1677>
- Terpstra, W.J., 1992. Typing leptospira from the perspective of a reference laboratory. *Acta Leiden.* 60, 79–87.
- Thomas, D.D., Higbie, L.M., 1990. In vitro association of leptospires with host cells. *Infect. Immun.* 58, 581–585. <https://doi.org/10.1128/iai.58.3.581-585.1990>
- Torgerson, P.R., Hagan, J.E., Costa, F., Calcagno, J., Kane, M., Martinez-Silveira, M.S., Goris, M.G.A., Stein, C., Ko, A.I., Abela-Ridder, B., 2015. Global Burden of Leptospirosis: Estimated in Terms of Disability Adjusted Life Years. *PLoS Negl. Trop. Dis.* 9, 1–14. <https://doi.org/10.1371/journal.pntd.0004122>
- Toshihiro ITO, 1987. Leptospiral attachment to four structural components of Extracellular Matrix.
- Uttamrao, P.P., Sathyaseelan, C., Patro, L.P.P., Rathinavelan, T., 2021. Revelation of Potent Epitopes Present in Unannotated ORF Antigens of SARS-CoV-2 for Epitope-Based Polyvalent Vaccine Design Using Immunoinformatics Approach. *Front. Immunol.* 12,

692937. <https://doi.org/10.3389/fimmu.2021.692937>
- Van Regenmortel, M.H.V., 2006. Immunoinformatics may lead to a reappraisal of the nature of B cell epitopes and of the feasibility of synthetic peptide vaccines. *J. Mol. Recognit.* 19, 183–187. <https://doi.org/10.1002/jmr.768>
- Vedhagiri, K., Velineni, S., Timoney, J.F., Shanmughapriya, S., Vijayachari, P., Narayanan, R., Natarajaseenivasan, K., 2013. Detection of LipL32-specific IgM by ELISA in sera of patients with a clinical diagnosis of leptospirosis. *Pathog. Glob. Health* 107, 130–135. <https://doi.org/10.1179/2047773213Y.0000000088>
- Verma, A., Brissette, C.A., Bowman, A.A., Shah, S.T., Zipfel, P.F., Stevenson, B., 2010. Leptospiral Endostatin-Like Protein A Is a Bacterial Cell Surface Receptor for Human Plasminogen 78, 2053–2059. <https://doi.org/10.1128/IAI.01282-09>
- Verma, A., Hellwage, J., Artiushin, S., Zipfel, P.F., Kraiczy, P., Timoney, J.F., Stevenson, B., 2006. LfhA , a Novel Factor H-Binding Protein of *Leptospira interrogans* 74, 2659–2666. <https://doi.org/10.1128/IAI.74.5.2659>
- Verma, A., Stevenson, B., Adler, B., 2013. Leptospirosis in horses. *Vet. Microbiol.* 167, 61–66. <https://doi.org/10.1016/j.vetmic.2013.04.012>
- Victoriano, A.F.B., Smythe, L.D., Gloriani-Barzaga, N., Cavinta, L.L., Kasai, T., Limpakarnjanarat, K., Ong, B.L., Gongal, G., Hall, J., Coulombe, C.A., Yanagihara, Y., Yoshida, S.-I., Adler, B., 2009. Leptospirosis in the Asia Pacific region. *BMC Infect. Dis.* 9, 147. <https://doi.org/10.1186/1471-2334-9-147>
- Vieira, M.L., Vasconcellos, S.A., Gonc, A.P., Morais, Z.M. De, Nascimento, A.L.T.O., 2009. Plasminogen Acquisition and Activation at the Surface of *Leptospira* Species Lead to Fibronectin Degradation □ 77, 4092–4101. <https://doi.org/10.1128/IAI.00353-09>
- Wang, H., Wu, Y., Ojcius, D.M., Yang, X.F., Zhang, C., Ding, S., Lin, X., Yan, J., 2012. Leptospiral Hemolysins Induce Proinflammatory Cytokines through Toll-Like Receptor 2-and 4-Mediated JNK and NF- κ B Signaling Pathways 7. <https://doi.org/10.1371/journal.pone.0042266>
- Wang, P., Sidney, J., Kim, Y., Sette, A., Lund, O., Nielsen, M., Peters, B., 2010. Peptide binding predictions for HLA DR, DP and DQ molecules. *BMC Bioinformatics* 11. <https://doi.org/10.1186/1471-2105-11-568>

- Wang, Z., Jin, L., Alicja, W., 2007. Leptospirosis vaccines 10, 1–10. <https://doi.org/10.1186/1475-2859-6-39>
- Weil, I.V.A., 1886. Ueber eine eigenthümliche, mit Milztumor, Icterus und Nephritis einhergehende, acute Infectionskrankheit / von A. Weil. | Wellcome Collection 44.
- Werts, C., Tapping, R.I., Mathison, J.C., Chuang, T.H., Kravchenko, V., Saint Girons, I., Haake, D.A., Godowski, P.J., Hayashi, F., Ozinsky, A., Underhill, D.M., Kirschning, C.J., Wagner, H., Aderem, A., Tobias, P.S., Ulevitch, R.J., 2001. Leptospiral lipopolysaccharide activates cells through a TLR2-dependent mechanism. *Nat. Immunol.* 2, 346–352. <https://doi.org/10.1038/86354>
- Widiyanti, D., Koizumi, N., Fukui, T., Muslich, L.T., Segawa, T., Villanueva, S.Y.A.M., Saito, M., 2013. Development of Immunochromatography-Based Methods for Detection of Leptospiral Lipopolysaccharide Antigen in Urine. <https://doi.org/10.1128/CVI.00756-12>
- Wiederstein, M., Sippl, M.J., 2007. ProSA-web: Interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* 35, 407–410. <https://doi.org/10.1093/nar/gkm290>
- Wilkins, M.R., Gasteiger, E., Bairoch, A., Sanchez, J.C., Williams, K.L., Appel, R.D., Hochstrasser, D.F., 1999. Protein identification and analysis tools in the ExPASy server. *Methods Mol. Biol.* 112, 531–552. <https://doi.org/10.1385/1-59259-584-7:531>
- Witkowska, D., Rowińska-Żyrek, M., 2019. Biophysical approaches for the study of metal-protein interactions. *J. Inorg. Biochem.* 199, 110783. <https://doi.org/10.1016/j.jinorgbio.2019.110783>
- Xu, K., Acharya, P., Kong, R., Cheng, C., Chuang, G.-Y., Liu, K., Louder, M.K., O'Dell, S., Rawi, R., Sastry, M., Shen, C.-H., Zhang, B., Zhou, T., Asokan, M., Bailer, R.T., Chambers, M., Chen, X., Choi, C.W., Dandey, V.P., Doria-Rose, N.A., Druz, A., Eng, E.T., Farney, S.K., Foulds, K.E., Geng, H., Georgiev, I.S., Gorman, J., Hill, K.R., Jafari, A.J., Kwon, Y.D., Lai, Y.-T., Lemmin, T., McKee, K., Ohr, T.Y., Ou, L., Peng, D., Rowshan, A.P., Sheng, Z., Todd, J.-P., Tsybovsky, Y., Viox, E.G., Wang, Y., Wei, H., Yang, Y., Zhou, A.F., Chen, R., Yang, L., Scorpio, D.G., McDermott, A.B., Shapiro, L., Carragher, B., Potter, C.S., Mascola, J.R., Kwong, P.D., 2018. Epitope-based vaccine design yields fusion peptide-directed antibodies that neutralize diverse strains of HIV-1. *Nat. Med.* 24, 857–867. <https://doi.org/10.1038/s41591-018-0042-6>

- Yan, W., 2009. Immunogenicity and protective efficacy of recombinant *Leptospira* immunoglobulin-like protein B (rLigB) in a hamster challenge model. *Microw. Opt. Technol. Lett.* 11, 230–237. <https://doi.org/10.1016/j.micinf.2008.11.008>
- Yang, C., Wu, M., Pan, M., Hsieh, W., 2002. The *Leptospira* Outer Membrane Protein LipL32 Induces Tubulointerstitial Nephritis-Mediated Gene Expression in Mouse Proximal Tubule Cells 2037–2045. <https://doi.org/10.1097/01.ASN.0000022007.91733.62>
- Yang, H., Zhu, Y., Qin, J., He, P., Jiang, C., Zhao, G., Guo, X., 2006. In silico and microarray-based genomic approaches to identifying potential vaccine candidates against *Leptospira interrogans* 12, 1–12. <https://doi.org/10.1186/1471-2164-7-293>
- Yang, J., Zhang, Y., I-tasser, T., 2015. I-TASSER server : new development for protein structure and function predictions 43, 174–181. <https://doi.org/10.1093/nar/gkv342>
- Yang, W., 2011a. Nucleases: Diversity of structure, function and mechanism, *Quarterly Reviews of Biophysics*. <https://doi.org/10.1017/S0033583510000181>
- Yang, W., 2011b. Nucleases: diversity of structure, function and mechanism. *Q. Rev. Biophys.* 44, 1–93. <https://doi.org/10.1017/S0033583510000181>
- Zenebe, T., Abdi, R.D., 2013. Global Epidemiological Overview of Leptospirosis Global Epidemiological Overview of Leptospirosis. <https://doi.org/10.5829/idosi.ijmr.2013.4.1.7134>
- Zhang, K., Jiang, N., Chen, H., Zhang, N., Sang, X., Feng, Y., Chen, R., Chen, Q., 2021. TatD DNases of African trypanosomes confer resistance to host neutrophil extracellular traps 64, 621–632.
- Zonta, Y.R., Laura, A., Dezen, O., Manoel, A., Coletta, D., Shu, K., Yu, T., Carvalho, L., Alves, L., Depra, I.D.C., Kubes, P., Ximenes, V.F., Alarcão, L., 2021. *Paracoccidioides brasiliensis* Releases a DNase-Like Protein That Degrades NETs and Allows for Fungal Escape 10, 1–13. <https://doi.org/10.3389/fcimb.2020.592022>
- Zuerner, R.L., 2015. Host response to leptospira infection. *Curr. Top. Microbiol. Immunol.* 387, 223–250. https://doi.org/10.1007/978-3-662-45059-8_9

Appendix

Appendix Table 1: List of predicted conformation BCL epitopes

Domain	BEPro	DiscoTope 2.0		ElliPro		
	Residue No.	Residue No	Total	Residue	Total	Score
LigA7	1-2, <u>30-33,36</u> , 54, 56-59	2, 5, <u>29-36</u> , 56, 58, 59	13	A:A1, A:K2, A:L3, A:V4, A:E5, <u>A:F29, A:T30, A:D31, A:N32, A:S33, A:N34</u> , A:G78, A:G79	13	0.791
				A:N55, A:T56, A:N57, A:A58, A:K59	5	0.771
				A:A11, A:A12, A:A13, A:S14, A:K15, A:A16, A:K17, A:G18, A:S46, A:N47, A:T48, A:D49, A:I50, A:L66, A:K67, A:Q68, A:G69, A:T70, A:T87, A:T89, A:Q90	21	0.67
LigA8	1, 12, <u>24-32</u> , 51-54, 74	1, 22, <u>24-32</u> , 36, 52-54,	15	A:S1, <u>A:I24, A:F25, A:T26, A:D27, A:H28, A:S29, A:K30</u>	8	0.833
LigA8	1, 12, <u>24-32</u> , 51-54, 74	1, 22, <u>24-32</u> , 36, 52-54,	15	A:I33, A:A51, A:P52, A:G53, A:E54, A:E55	6	0.72
				A:T34, A:E35, A:Q36, A:E73, A:K74	5	0.652
				A:T7, A:A9, A:S10, A:A12, A:K13, A:G14, A:M15, A:S42, A:S43, A:K44, A:A45, A:L46, A:I61, A:A62, A:V63, A:G64, A:N65, A:T82, A:T84	19	0.647
LigA9	1, <u>26-31</u> , 51, 53-56	2, 4, <u>23, 25, 27-33</u> , 54, 55,	13	A:T1, A:S2, A:I25, <u>A:Y26, A:S27, A:D28, A:N29, A:S30, A:S31</u> , A:K32, A:I34, A:T35, A:S36, A:S51, A:N52, A:T53, A:K54, A:G55, A:Y56, A:G75, A:N76	21	0.725
LigA9	1, <u>26-31</u> , 51, 53-56	2, 4, <u>23, 25, 27-33</u> , 54, 55,	13	A:H10, A:C11, A:L12, A:V13, A:K14, A:G15, A:S43, A:N44, A:N45, A:S46, A:V47, A:G64, A:T65, A:G66, A:T67, A:S85, A:T87, A:A88	18	0.705
				A:V8, A:K9, A:K83	3	0.561
LigA10	13, <u>27-30</u> , 54-56	13, <u>28, 29</u> , 87	4	A:A53, A:T54, A:G55, A:S56, A:K57	5	0.823

LigA10	13, <u>27-30</u> , 54-56	13, <u>28, 29</u> , 87	4	A:I1, A:E2, A:V25, <u>A:F26, A:T27, A:D28, A:N29, A:S30</u> , A:T31, A:K32, A:D33, A:I34, A:T35	13	0.804
				A:T8, A:S9, A:S10, A:H11, A:K12, A:A13, A:K14, A:G15, A:S43, A:N44, A:T45, A:A46, A:Y47, A:L63, A:S64, A:K65, A:G66, A:T67, A:Q85, A:T87	20	0.699
LigA11	<u>13-15</u> , 27-30, <u>47</u> , 55, 65,	<u>13-15</u> , 28, 29, <u>47</u> , 65, 67, 87	9	A:V1, A:S2, A:I25, A:Y26, A:T27, A:D28, A:H29, A:S30, A:E31	9	0.815
LigA11	<u>13-15</u> , 27-30, <u>47</u> , 55, 65,	<u>13-15</u> , 28, 29, <u>47</u> , 65, 67, 87	9	<u>A:A13, A:K14, A:G15</u> , A:T87	4	0.784
				A:I34, A:E51, A:T53, A:F54, A:G55, A:S56, A:A57	7	0.686
				A:N8, A:N9, A:I10, A:S11, <u>A:S43, A:N44, A:P45, A:K46, A:I47</u> , A:N64, A:I65, A:G66, A:S67, A:N85	14	0.677
				A:T35, A:E36, A:Q37, A:V38, A:S75, A:D76, A:T77	7	0.575
LigA12	1-2, 4, 6, 13, <u>25-33</u> , 53-55, <u>76-78</u> , 89	1, 2, 4, <u>23-33</u> , 35-37, 55, 73, <u>75-79</u>	24	A:R1, A:Y2, <u>A:T25, A:Y26, A:S27, A:D28, A:Q29, A:S30, A:T31, A:K32</u> , A:L34, A:T35, A:E36, A:N52, A:T53, A:P54, A:G55, A:K56, A:K57, A:Y74, <u>A:D75, A:H76, A:H77, A:Q79</u>	24	0.682
	1-2, 4, 6, 13, <u>25-33</u> , 53-55, <u>76-78</u> , 89	1, 2, 4, <u>23-33</u> , 35-37, 55, 73, <u>75-79</u>	24	A:S8, A:Y9, A:A10, A:G11, A:E13, A:K14, A:G15, A:Y16, A:S43, A:N44, A:P45, A:S46, A:S47, A:S63, A:E64, A:L65, A:G66, A:E67, A:T86, A:T88, A:E89	21	0.660
LigA13	1, 15, <u>29-32</u> , 56-57, 89	1, <u>29-31</u> ,	4	A:S53, A:N54, A:V55, A:D56, A:D57, A:E58	6	0.732
	1, 15, <u>29-32</u> , 56-57, 89	1, <u>29-31</u> ,	4	A:G1, A:I2, A:V3, A:N4, <u>A:K27, A:F28, A:E29, A:N30, A:G31, A:A32, A:E33, A:I34, A:L36, A:T37, A:E38</u> , A:N77, A:S78	17	0.697
				A:S10, A:S11, A:I12, A:S13, A:K14, A:T15, A:K16, A:G17, A:S45, A:N46, A:P47, A:T48,	23	0.646

				A:V49, A:L65, A:S66, A:V67, A:G68, A:S69, A:K71, A:D84, A:E86, A:T88, A:P89		
LigB1	<u>1-6</u> , 18-19, <u>29-37</u> , 40, 49-50, 55-56, 74-76, 86-87	<u>1-6</u> , 27, <u>29-37</u> , 40, 41, 56, 74-77	23	<u>A:N1, A:P2, A:T3, A:T5, A:R6, A:I29, A:F30, A:D31, A:N32, A:G33, A:T34, A:N35, A:Q36</u>	13	0.8
LigB1	<u>1-6</u> , 18-19, <u>29-37</u> , 40, 49-50, 55-56, 74-76, 86-87	<u>1-6</u> , 27, <u>29-37</u> , 40, 41, 56, 74-77	23	A:S14, A:S15, A:I16, A:A17, A:N18, A:G19, A:S48, A:Q49, A:S50, A:V51, A:I62, A:A63, A:S64, A:G65, A:S66, A:T83, A:V84, A:T85, A:P86, A:A87	20	0.647
LigB2	<u>1</u> , 14-15, <u>25-32</u> , 38, 48, 73-75, 85-86	<u>1</u> , 14,15, <u>25-32</u> , 38, 48, 73, 74	15	<u>A:S1, A:Q3, A:I25, A:F26, A:S27, A:D28, A:G29, A:S30, A:H31</u>	9	0.832
LigB2	<u>1</u> , 14-15, <u>25-32</u> , 38, 48, 73-75, 85-86	<u>1</u> , 14,15, <u>25-32</u> , 38, 48, 73, 74	15	A:S9, A:G10, A:I11, A:L12, A:P13, A:K14, A:G15, A:T16, A:N17, A:S45, A:N46, A:P47, A:D48, A:L49, A:I61, A:N62, A:L63, A:G64, A:T65, A:E80, A:T82, A:V83, A:G84, A:D85, A:A86	25	0.623
				A:I34, A:S35, A:N36, A:D37, A:P38, A:L39, A:D53, A:D54, A:S55, A:G56, A:Q73, A:S74, A:K75	13	0.573
LigB3	1-2, 6-7, 9, 14-15, 17, 19, 23, <u>25-33</u> , 36, 45-46, <u>51-53</u> , 55, 71-72, 83-84	1, 2, 7, 9, 14, 15, 17, <u>23-33</u> , 36, 46, <u>51-53</u> , 59, 71, 72, 83, 84	28	A:S1, A:Q2, <u>A:I25, A:Y26, A:S27, A:D28, A:N29, A:S30, A:N31, A:R32</u> , A:I34	11	0.767
LigB3	1-2, 6-7, 9, 14-15, 17, 19, 23, <u>25-33</u> , 36, 45-46, <u>51-53</u> , 55, 71-72, 83-84	1, 2, 7, 9, 14, 15, 17, <u>23-33</u> , 36, 46, <u>51-53</u> , 59, 71, 72, 83, 84	28	A:E71, A:N72, A:I73, A:I74	4	0.726
				A:N8, A:N9, A:P10, A:N11, A:I12, A:P13, A:L14, A:G15, A:K16, A:K17, A:S43, A:N44, A:S45, A:T46, A:I47, A:A59, A:D60, A:T61, A:G62, A:I63, A:I80, A:T82, A:P83, A:A84	24	0.634
				A:S35, A:S36, A:S37, <u>A:Q51, A:N52, A:N53</u>	6	0.552
LigB4	1, <u>26-31</u> , 46, <u>53-55</u> , 66, 75-76, 88	<u>25, 27-31, 54, 55</u> , 76	9	A:V1, A:S2, <u>A:I25, A:F26, A:T27, A:D28, A:N29, A:S30, A:N31, A:S32, A:I34</u>	11	0.758

LigB4	1, <u>26-31</u> , 46, <u>53-55</u> , 66, 75-76, 88	<u>25, 27-31</u> , <u>54, 55</u> , 76	9	A:G75, A:G76, A:I77, A:Q78	4	0.716
				A:S51, A:N52, <u>A:A53, A:S54</u> , A:D55, A:S56, A:H57	7	0.695
				A:T8, A:N9, A:S10, A:T11, A:V12, A:A13, A:K14, A:G15, A:L16, A:S43, A:N44, A:T45, A:D46, A:I47, A:L63, A:N64, A:Q65, A:G66, A:N67, A:T84, A:V85, A:T86, A:Q87, A:A88	24	0.649
LigB5	1, <u>25-32</u> , 36, 52-55, 74-75	<u>25-32</u> , 52-54,	11	A:T1, A:S2, <u>A:I25, A:F26, A:T27, A:D28, A:N29, A:S30, A:K31</u>	9	0.854
LigB5	1, <u>25-32</u> , 36, 52-55, 74-75	<u>25-32</u> , 52-54,	11	A:G74, A:K75, A:V76, A:S77	4	0.713
				A:V8, A:L9, A:P10, A:S11, A:I12, A:A13, A:K14, A:G15, A:L16, A:S43, A:S44, A:A45, A:I46, A:V47, A:H62, A:A63, A:V64, A:G65, A:D66, A:T83, A:V84, A:V85, A:P86, A:A8	24	0.610
				A:T35, A:D36, A:Q37, A:S48, A:V49, A:S50, A:N51, <u>A:L52, A:D53, A:D54</u> , A:N55, A:K56	12	0.573
LigB6	1-2, 7, <u>25-33</u> , 36, <u>53-56</u> , 76, 90	1, 2, <u>23-33</u> , 36, <u>53-55</u> , 76	18	A:T1, A:S2, <u>A:I25, A:Y26, A:S27, A:D28, A:N29, A:S30, A:N31, A:K32</u>	10	0.831
LigB6	1-2, 7, <u>25-33</u> , 36, <u>53-56</u> , 76, 90	1, 2, <u>23-33</u> , 36, <u>53-55</u> , 76	18	A:G75, A:K76, A:V77, A:S78	4	0.711
				A:S36, <u>A:S51, A:N52, A:A53, A:Q54, A:K55</u> , A:N56, A:Q57	8	0.697
				A:S11, A:L12, A:A13, A:K14, A:G15, A:S43, A:D44, A:S45, A:S46, A:I47, A:A63, A:A64, A:T65, A:G66, A:A67, A:V86, A:T87, A:A88, A:A89, A:K90	20	0.670
LigB7	1, <u>25-32</u> , 36, 45-46, <u>53-56</u> , 76, 89-90	<u>26-31</u> , 46, <u>53-55</u> , 90	11	A:V1, A:E2, <u>A:I25, A:F26, A:T27, A:D28, A:N29, A:S30, A:N31, A:S32, A:I34</u>	11	0.768
LigB7	1, <u>25-32</u> , 36, 45-46, <u>53-56</u> , 76, 89-90	<u>26-31</u> , 46, <u>53-55</u> , 90	11	A:A10, A:S11, A:K12, A:A13, A:K14, A:G15, A:L16, A:S43, A:N44, A:T45, A:D46, A:I47, A:E49, A:L63, A:T64, A:P65, A:G66, A:S67, A:K85, A:V86, A:T87, A:P88, A:A89, A:Q90	24	0.653

				A:T35, A:N36, A:Q37, <u>A:T51, A:N52, A:T53, A:S54, A:G55, A:S56, A:K57</u>	10	0.607
				A:L74, A:G75, A:S76, A:I77, A:K78, A:S79, A:S80	7	0.579
LigB8	<u>27-30</u> , 45-46, 52, 73, 85	<u>28, 29</u> , 46	3	A:D72, A:S73, A:I74	3	0.776
LigB8	<u>27-30</u> , 45-46, 52, 73, 85	<u>28, 29</u> , 46	3	A:I1, A:S2, <u>A:T25, A:Y26, A:T27, A:D28, A:H29, A:S30, A:V31, A:Q32, A:V34, A:T35, A:A36, A:L37, A:A38</u> , A:N51, A:N52, A:V53, A:T54, A:G55	20	0.640
				A:P7, A:I8, A:N9, A:P10, A:S11, A:V12, A:A13, A:K14, A:G15, A:L16, A:S43, A:N44, A:P45, A:R46, A:K47, A:V60, A:A61, A:T62, A:G63, A:N64, A:N81, A:V82, A:T83, A:P84, A:A85	25	0.625
LigB9	1-2, 7, 9, <u>14-15, 26-31</u> , 45-46, <u>51, 53-56</u> , 66, 75, 76, 81, <u>89-90</u>	2, 7, 9, 13-15, <u>26-31</u> , 33, 44-46, <u>51, 53-56</u> , 63, 64, 66, 67, <u>88-90</u>	28	A:T1, A:S2, <u>A:I25, A:F26, A:S27, A:D28, A:K29, A:S30, A:T31, A:Q32</u>	10	0.832
LigB9	1-2, 7, 9, <u>14-15, 26-31</u> , 45-46, <u>51, 53-56</u> , 66, 75, 76, 81, <u>89-90</u>	2, 7, 9, 13-15, <u>26-31</u> , 33, 44-46, <u>51, 53-56</u> , 63, 64, 66, 67, <u>88-90</u>	28	A:F76, A:I77, A:Q78	3	0.719
				A:T35, A:Q36, A:L37, A:E51, A:N52, <u>A:T53, A:S54, A:G55, A:K56, A:K57</u>	10	0.627
				A:T8, A:I9, A:N10, A:S11, A:I12, A:T13, A:H14, A:G15, A:L16, A:S43, A:D44, A:P45, A:S46, A:K47, A:S63, A:K64, A:L65, A:G66, A:S67, A:V86, A:T87, <u>A:D88, A:L89</u>	23	0.611
LigB10	1-2, 7, <u>25-31, 53-55</u> , 75-77, 90	1, 2, <u>25-31, 54, 55</u> , 76	12	A:K1, A:S2, <u>A:T25, A:F26, A:I27, A:D28, A:G29, A:S30, A:E31, A:Q32</u>	10	0.844
LigB10	1-2, 7, <u>25-31, 53-55</u> , 75-77, 90	1, 2, <u>25-31, 54, 55</u> , 76	12	A:N75, A:S76, A:I77	3	0.770
				A:S10, A:S11, A:I12, A:A13, A:K14, A:G15, A:L16, A:S43, A:K44, A:S45, A:D46, A:V47, A:A48, A:A62, A:L63, A:S64, A:I65, A:G66,	26	0.607

				A:S67, A:N83, A:N85, A:V86, A:S87, A:A88, A:A89, A:T90		
				A:T35, A:N36, A:L37, A:I50, A:N51, A:N52, <u>A:A53, A:A54, A:N55,</u> A:E56, A:K57	11	0.600
LigB11	<u>1-2, 4,</u> 7, 9, 15, <u>25-32,</u> 45-46, 53-56, 75-77, 80, 87, 88	<u>1-4,</u> 6, 7, 9, <u>23-33,</u> 36, 37, 46, 53-56, 67, 73, 75-78, 80, 82	33	<u>A:D1, A:S2, A:K4, A:V25, A:Y26, A:S27, A:D28, A:S29, A:T30, A:I31, A:Q32</u>	11	0.806
LigB11	<u>1-2, 4,</u> 7, 9, 15, <u>25-32,</u> 45-46, 53-56, 75-77, 80, 87, 88	<u>1-4,</u> 6, 7, 9, <u>23-33,</u> 36, 37, 46, 53-56, 67, 73, 75-78, 80, 82	33	A:N75, A:S76, A:I77	3	0.754
				A:V8, A:N9, A:N10, A:N11, A:I12, A:A13, A:K14, A:G15, A:L16, A:S43, A:N44, A:S45, A:S46, A:S47, A:L63, A:Q64, A:I65, A:G66, A:N67, A:T84, A:V85, A:S86, A:A87, A:A88	24	0.642
				A:S35, A:D36, A:S37, A:S51, A:N52, A:S53, A:T54, A:E55, A:T56, A:K57, A:G58	11	0.544
LigB12	<u>14-15,</u> 27-30, 45-47, 53-55, 64, 66-67, 76, <u>86-90</u>	<u>14, 15,</u> 46, 63, 64, 66, 67, <u>86-90</u>	12	A:S1, A:S2, A:G24, A:T25, A:Y26, A:S27, A:A28, A:G29, A:T30, A:K31, A:A32, A:L34	12	0.749
LigB12	<u>14-15,</u> 27-30, 45-47, 53-55, 64, 66-67, 76, <u>86-90</u>	<u>14, 15,</u> 46, 63, 64, 66, 67, <u>86-90</u>	12	A:G75, A:S76, A:V77, A:S78	4	0.708
				A:P7, A:I8, A:N9, A:T10, A:N11, A:I12, A:N13, <u>A:T14, A:T15,</u> A:S43, A:N44, A:Q45, A:S46, A:Q47, A:K49, A:I63, A:A64, A:S65, A:G66, A:N67, A:T84, A:V85, <u>A:N86, A:K87, A:T88, A:D89, A:T90</u>	27	0.622
				A:T35, A:S36, A:S37, A:S51, A:N52, A:A53, A:S54, A:E55, A:T56, A:K57	10	0.614

Appendix Table 2: List of predicted linear BCL epitopes

Domain	BepiPred 2.0		BCPred			ABCPred,		
	Peptide	Position	Peptide	Position	Score	Peptides	Position	Score
LigA7	ASKAKGLT	12-19	FTDNSNSDITNQ	28-39	0.992	IQITPAAASKAKGL	5-18	0.87
	FTDNSNSDI	28-36	TWNSSNTDILTV	41-52	0.916	AKGLTERFKATGIF	15-28	0.87
	VSNTNAKRGL G	52-62	GLGSTLKQGTVK	60-71	0.695	QGTVKVTASMGGIE	67-80	0.80
	EDS	79-81	DSVDFTVTQATL	81-92	0.583	LGSTLKQGTVKVTA	61-74	0.80
						SNSDITNQVTWNSS	32-45	0.80
						FKATGIFTDNSNSD	22-35	0.78
						AAASKAKGLTERFK	10-23	0.75
LigA8	VSPTRASIAKG	5-15	TSIEVSPTRASI	1-12	0.898	NISITATLEKLSGK	66-79	0.86
	TDHKKKNIT	27-35	KKNITEQVTWKS	31-42	0.741	SSSKALSMLNAPGE	42-55	0.81
	KSSSKALS	41-48	LSGKTDITVTPA	76-87	0.715	VSPTRASIAKGMTQ	5-18	0.80
	APGEEGTGKAI	52-62	APGEEGTGKAIA	52-63	0.675	SMLNAPGEEGTGKA	48-61	0.79
			KGMTQKFTATGI	14-25	0.307	HSKKNITEQVTWKS	29-42	0.79
						GTGKAIAVGNISIT	57-70	0.75
						GIFTDHSKKNITEQ	24-37	0.75
						TQKFTATGIFTDHS	17-30	0.75
LigA9	ISPVKHCLVKG LT	5-17	TLGNVSSQVSKL	73-84	0.999	NTKGYQQGAHGTG T	52-65	0.94
	YSDNSSKDI	26-34	AHGTGTGTVDIK	60-71	0.946	GIYSDNSSKDITSA	24-37	0.78
	SSNN	42-45	VATISNTKGYQG	47-58	0.616	SKDITSAVTWHSSN	31-44	0.75
	NTKGYQQGQ	52-59	TGIYSDNSSKDI	23-34	0.373			
LigA10	NPTSSHKAKG	6-18	FTDNSTKDITDQ	26-37	0.948	TDNSTKDITDQVT	27-40	0.88
	LTE					W		
	TDNSTKDITD	27-36	ISNATGSKGVVN	50-61	0.720	LNPTSSHKAKGLTE	5-18	0.85
	SS	42-43	SSANATFQVTPA	78-89	0.411	ITDQVTWKSSNTAY	34-47	0.84
	NATGSK	52-57	HKAKGLTENFK	11-22	0.338	SHKAKGLTENFKAT	10-23	0.80
			A			LTENFKATGVFTDN	16-29	0.77

						KGTSHISATLGSIS	65-78	0.75
LigA11	<u>IPNNISFAKGN</u> <u>S</u>	<u>6-17</u>	VSTTNIGSTNIT	60-71	0.855	EADITEQVTWSSSN	31-44	0.85
	YTDHSEADI	26-34	<u>AKGNSYQFKAT</u> <u>G</u>	<u>13-24</u>	<u>0.639</u>	FKATGIYTDHSEAD	20-33	0.85
	S	43				GSTNITAKLSDTV	66-79	0.83
	P	45				<u>VIPNNISFAKGNSY</u>	<u>5-18</u>	<u>0.81</u>
	A	48				ASVENTFGSAGLVS	48-61	0.75
	VENTFGSA	50-57						
LigA12	YAGIEKGYT	9-17	<u>GTYSQSTKDLT</u>	<u>24-35</u>	<u>0.920</u>	<u>SDQSTKDLTEDVT</u> <u>W</u>	<u>27-40</u>	<u>0.91</u>
	<u>YSDQSTKDLT</u> <u>E</u>	<u>28-36</u>	GEPDITVFYDHH	66-77	0.589	TKQFSAIGTYSDQS	17-30	0.82
	S	43	QSSYTPVTVTES	79-90	0.447	SNPSSVVIENTPGK	43-56	0.75
	TPGKKG	53-58				GIEKGYTKQFSAIG	11-24	0.75
LigA13	<u>SSISKTKGST</u>	<u>10-19</u>	<u>ISKTKGSTHQFK</u>	<u>12-23</u>	<u>0.842</u>	SNPTVVSISNVDDE	45-58	0.93
	ENGAEIDL	29-37	TGKFENGAEIDL	25-36	0.585	<u>SSISKTKGSTHQFK</u>	<u>10-24</u>	<u>0.92</u>
	S	45	VVSISNVDDER	49-60	0.585	AEIDLTELVTWSSS	32-45	0.81
	VDDERGL	55-61	SSSIDFEVTPEI	80-91	0.368	STHQFKATGKFENG	18-31	0.79
LigB1	<u>ELSYQDSSIAN</u> <u>GTST</u>	<u>7-21</u>	DSTSIVPDSQSV	39-50	0.997	LEVTAIFDNGTNQN	23-36	0.89
	FDNGTNQNITD STSIVPDSQS	29-49	<u>SIANGTSTTLEV</u>	<u>14-25</u>	<u>0.996</u>	ITDSTSIVPDSQSV	37-50	0.88
	<u>RVRGIASGSSII</u> <u>KAEYNGLYSE</u> <u>QKITV</u>	<u>57-83</u>	YSEQKITVTPAT	76-87	0.613	<u>ELSYQDSSIANGTS</u>	<u>7-20</u>	<u>0.85</u>
			<u>TIQGNRVRGIAS</u>	<u>52-63</u>	<u>0.053</u>	SIKAEYNGLYSEQ	66-79	0.80
						IANGTSTTLEVTAI	15-28	0.78
						NPTIRIELSYQDSS	1-14	0.78
						<u>IQGNRVRGIASGSS</u>	<u>53-66</u>	<u>0.76</u>

LigB2	<u>TSLGSGILPKG</u> <u>TN</u>	<u>5-17</u>	SSNPDLVQVDDS	45-56	0.739	<u>IQVTSLGSGILPKG</u>	<u>3-16</u>	<u>0.88</u>
	<u>FSDGSHQDISN</u> <u>DPLIVWSSNP</u>	<u>27-48</u>	<u>SDGSHQDISNDP</u>	<u>28-39</u>	<u>0.635</u>	<u>SHQDISNDPLIVWS</u>	<u>31-44</u>	<u>0.81</u>
			HIRASFQSKQGA	68-79	0.556	VQVDDSGLASGINL	51-64	0.78
			<u>GSGILPKG</u> <u>TNRQ</u>	<u>9-20</u>	<u>0.553</u>	GINLGTAHIRASFQ	61-74	0.77
LigB3	<u>VTSNNPNIPLG</u> <u>KKQ</u>	<u>5-18</u>	SSNPDLVQVDDS	45-56	0.739	ILETADTGIVTISA	55-68	0.83
	<u>YSDNSNRDIS</u>	<u>26-35</u>	<u>SDGSHQDISNDP</u>	<u>28-39</u>	<u>0.635</u>	IANIQNNGILETAD	47-60	0.80
			HIRASFQSKQGA	68-79	0.556	<u>QIQVTSNNPNIPLG</u>	<u>2-15</u>	<u>0.78</u>
			<u>GSGILPKG</u> <u>TNRQ</u>	<u>9-20</u>	<u>0.553</u>	<u>LIATGIYSDNSNRD</u>	<u>20-33</u>	<u>0.77</u>
						PLGKKQKLIATGIY	13-26	0.75
LigB4	<u>SPTNSTVAKG</u> <u>LQ</u>	<u>6-17</u>	<u>FTDNSNSDITDQ</u>	<u>26-37</u>	<u>0.972</u>	<u>TDNSNSDITDQVTW</u>	<u>27-39</u>	<u>0.87</u>
	<u>FTDNSNSDITD</u>	<u>26-36</u>	TWDSSNTDILSI	39-50	0.851	NQGNVKVTASIGGI	64-77	0.81
	T	45	<u>ISVSPTNSTVAK</u>	<u>3-14</u>	<u>0.83</u>	FKATGIFTDNSNSD	20-33	0.79
	LSISNASDSHG	48-58	ASTLNQGNVKVT	60-71	0.352	<u>VSISVSPTNSTVAK</u>	<u>1-14</u>	<u>0.77</u>
			GGIQGSTDFTVT	75-86	0.316			
LigB5	SPVLPSIAKGLT	6-17	<u>FTDNSKKDITDQ</u>	<u>26-37</u>	<u>0.954</u>	SVSNLDDNKGLGKA	48-61	0.84
	<u>FTDNSKKDIT</u>	<u>26-35</u>	GDTTITATLGKV	65-76	0.944	KGLTQKFTAIGIFT	14-27	0.83
	<u>SSAIVSVSNLD</u> <u>DNKGLG</u>	<u>43-59</u>	<u>LDDNKGLGKAH</u> <u>A</u>	<u>52-63</u>	<u>0.764</u>	<u>TDNSKKDITDQVT</u> <u>W</u>	<u>27-40</u>	<u>0.77</u>
						<u>NKGLGKAHAVGDT</u> <u>T</u>	<u>55-68</u>	<u>0.76</u>
LigB6	NPVNPSLAKGL T	6-17	<u>AQKNQGNAYGA</u> <u>A</u>	<u>53-64</u>	<u>0.973</u>	<u>ISNAQKNQGNAYG</u> <u>A</u>	<u>50-63</u>	<u>0.87</u>
	<u>YSDNSNKDIT</u>	<u>26-35</u>	TFGKVSSPVSTL	83-94	0.847	<u>GIYSDNSNKDITSA</u>	<u>24-37</u>	<u>0.82</u>
	SSDS	42-45	<u>FSATGIYSDNSN</u>	<u>20-31</u>	<u>0.503</u>	QKFSATGIYSDNSN	18-31	0.78
	<u>NAQKNQG</u>	<u>52-58</u>	VTWFSSDSSIAT	38-49	0.469	KDITSAVTWFSSDS	32-45	0.77
			TSIQINPVNPSL	1-12	0.357	QGNAYGAATGATDI	57-70	0.75

LigB7	TPAAASKAKG LT	6-17	ITNTSGSKGITN	50-61	0.999	ITPAAASKAKGLTE	5-18	0.89
	<u>FTDNSNSDIT</u>	<u>26-35</u>	<u>FTDNSNSDITNQ</u>	<u>26-37</u>	<u>0.992</u>	<u>TDNSNSDITNQVTW</u>	<u>27-40</u>	<u>0.83</u>
	NTSGSKGI	52-59				ASKAKGLTERFKAT	10-23	0.82
						SNTDIAEITNTSGS	43-56	0.79
						NQVTWNSSNTDIAE	36-49	0.76
LigB8	<u>TPINPSVAKGL</u> <u>I</u>	<u>6-17</u>	VNNVTGSVTTVA	50-61	0.989	VQDVTALATWSSSN	31-44	0.89
	<u>YTDHSVQDVT</u>	<u>26-35</u>	<u>IAVTPINPSVAK</u>	<u>3-14</u>	<u>0.979</u>	<u>FKATGTYTDHSVQ</u> <u>D</u>	<u>20-33</u>	<u>0.84</u>
	SSSNPRKAMV	41-50	<u>ATGTYTDHSVQD</u>	<u>22-33</u>	<u>0.911</u>	KGLIRQFKATGTYT	14-27	0.83
						GNTNIKATIDSISG	63-76	0.82
						SVTTVATGNTNIKA	56-69	0.80
						NPRKAMVNNVTGS V	44-57	0.78
						<u>PINPSVAKGLIRQF</u>	<u>7-20</u>	<u>0.77</u>
LigB9	<u>PTINSITHGLT</u>	<u>7-17</u>	<u>FSDKSTQNLTLQ</u>	<u>26-37</u>	<u>0.986</u>	TQNLTLQLVTWISSD	31-44	0.88
	<u>FSDKSTQNLTL</u>	<u>26-35</u>	IENTSGKKGIAT	50-61	0.899	SDPSKIEIENTSGK	43-56	0.85
	ISSDPSKIEIENT SGKKGIA	41-60	IQSSPIPIVTD	77-89	0.595	<u>PTINSITHGLTKQF</u>	<u>7-20</u>	<u>0.82</u>
	L	65	<u>EITPTINSITHG</u>	<u>4-15</u>	<u>0.497</u>	<u>KATGIFSDKSTQNL</u>	<u>21-34</u>	<u>0.79</u>
	FI	76-77				THGLTKQFKATGIF	13-26	0.77
						NTSGKKGIATASKL	52-65	0.76
LigB10	PSSSSIAKGLTQ	7-18	VAPINNAANEKG	47-58	1	QQFKAIGTFIDGSE	18-31	0.90
	<u>FIDGSEQEI</u>	<u>26-34</u>	<u>GSEQEITNLVTW</u>	<u>29-40</u>	<u>0.741</u>	<u>ITNLVTWYSSKSDV</u>	<u>34-47</u>	<u>0.75</u>
	YSSKS	41-45						
	AANEKGLA	53-60						
LigB11	INPVNNNIAKG LT	5-17	<u>ISNSTETKGKAT</u>	<u>50-61</u>	<u>0.97</u>	QQYTALGVYSDSTI	18-31	0.90
	<u>YSDSTIQDIS</u>	<u>26-35</u>	<u>YSDSTIQDISDS</u>	<u>26-37</u>	<u>0.865</u>	ALQIGNSKITATYN	62-75	0.85
	SNS	43-45	IKINPVNNNIAK	3-14	0.853	SSNSSSISISNSTE	42-55	0.84

LigB12	<u>STETKGK</u>	<u>53-59</u>				<u>SDSTIQDISDSVTW</u>	<u>27-40</u>	<u>0.82</u>
						<u>NSTETKGKATALQI</u>	<u>52-65</u>	<u>0.79</u>
						AKGLTQQYTALGVY	13-26	0.77
	<u>ISPINTNINTTV</u> <u>S</u>	<u>5-17</u>	<u>SPINTNINTTVS</u>	<u>6-17</u>	<u>0.936</u>	KADLTSSVTWSSSN	31-44	0.91
	<u>SAGTKADL</u>	<u>27-34</u>	VTGIASGNPTII	60-71	0.916	VTGIASGNPTIIAT	60-73	0.90
	SNQ	43-45	QAKVSNASETKG	47-58	0.339	<u>FFAVGTYSAGTKA</u> <u>D</u>	<u>20-34</u>	<u>0.85</u>
	ETK	55-57	<u>GTYSAGTKADLT</u>	<u>24-35</u>	<u>0.327</u>	KVSNASETKGLVTG	49-62	0.84
						<u>PINTNINTTVSKQF</u>	<u>7-20</u>	<u>0.76</u>

Appendix Table 3: List of predicted CTL epitopes

Domain	IEDB	NetMHC 4.0	ProPred
LigA7	FTDNSNSDI	FTDNSNSDI	FTDNSNSDI
	GTVKVTASM	GTVKVTASM	GTVKVTASM
	LTVSNTNAK	SMGGIEDSV	SMGGIEDSV
	STLKQGTVK	VEIQITPAA	VEIQITPAA
	LTVSNTNAK	LTVSNTNAK	LTVSNTNAK
			STLKQGTVK
LigA8		ATLEKLSGK	ATLEKLSGK
		TWKSSSKAL	TWKSSSKAL
		EVSPTRASI	EVSPTRASI
		KSSSKALSM	KSSSKALSM
		KNITEQVTW	KNITEQVTW
	QVTWKSSSK		QVTWKSSSK
	SIAKGMTQK		SIAKGMTQK
	GIFTDHKK		GIFTDHKK
LigA9	ATISNTKGY	ATISNTKGY	ATISNTKGY
	SVATISNTK	SVATISNTK	SVATISNTK
		TLGNVSSQV	TLGNVSSQV
		KDITSAVTW	KDITSAVTW
		WHSSNNSVA	WHSSNNSVA
		SSKDITSAV	SSKDITSAV
		GTVDIKATL	GTVDIKATL
		YSDNSSKDI	YSDNSSKDI
LigA10	SSANATFQV	SSANATFQV	SSANATFQV
	GVFTDNSTK	GVFTDNSTK	GVFTDNSTK
	KISNATGSK	KISNATGSK	KISNATGSK
	ATFQVTPAK	ATFQVTPAK	ATFQVTPAK

		ALSKGTSHI	ALSKGTSHI
		GTSHISATL	GTSHISATL
		SISSANATF	SISSANATF
		TWKSSNTAY	TWKSSNTAY
	KSSNTAYAK		KSSNTAYAK
LigA11	FAKGNSYQF	FAKGNSYQF	FAKGNSYQF
	SVLNVTPAL	SVLNVTPAL	SVLNVTPAL
	KIASVENTF	KIASVENTF	KIASVENTF
	ISFAKGNSY	ISFAKGNSY	ISFAKGNSY
	VTWSSSNPK	VTWSSSNPK	VTWSSSNPK
	YQFKATGIY	YQFKATGIY	YQFKATGIY
	EVIPNNISF	EVIPNNISF	EVIPNNISF
		YTDHSEADI	YTDHSEADI
		ITAKLSDTV	ITAKLSDTV
		DTVSGSSVL	DTVSGSSVL
		NTFGSAGLV	NTFGSAGLV
LigA12	YSDQSTKDL	YSDQSTKDL	YSDQSTKDL
	TPGKKGLAF	TPGKKGLAF	TPGKKGLAF
	FSSNPSSVV	FSSNPSSVV	FSSNPSSVV
	HTQSSYTPV	HTQSSYTPV	HTQSSYTPV
	KQFSAIGTY	KQFSAIGTY	KQFSAIGTY
	YDHHTQSSY	YDHHTQSSY	YDHHTQSSY
	RYIMITPSY	RYIMITPSY	RYIMITPSY
		YIMITPSYA	YIMITPSYA
		ELGEPDITV	ELGEPDITV
		MITPSYAGI	MITPSYAGI
	VVIENTPGK		VVIENTPGK
	GTYSDQSTK		GTYSDQSTK
LigA13	KTKGSTHQF	KTKGSTHQF	KTKGSTHQF
	AEIDLTELV	AEIDLTELV	AEIDLTELV
		HQFKATGKF	HQFKATGKF

		WSSSNPTVV	WSSSNPTVV
	FENGAEIDL		FENGAEIDL
	SSNPTVVSI		SSNPTVVSI
	ISSSIDFEV		ISSSIDFEV
	TISLSSISK		TISLSSISK
LigB1	LANGTSTTL	LANGTSTTL	LANGTSTTL
	TTLEVTAIF	TTLEVTAIF	TTLEVTAIF
		IKAEYNGLY	IKAEYNGLY
LigB2	HIRASFQSK	HIRASFQSK	HIRASFQSK
	RQFSAIGIF	RQFSAIGIF	RQFSAIGIF
	EEMTVGDAV	EEMTVGDAV	EEMTVGDAV
		GTAHIRASF	GTAHIRASF
		LPKGTNRQF	LPKGTNRQF
		HQDISNDPL	HQDISNDPL
		ISNDPLIVW	ISNDPLIVW
	FSDGSHQDI		FSDGSHQDI
LigB3		KQKLIATGI	KQKLIATGI
		IPLGKKQKL	IPLGKKQKL
		YSDNSNRDI	YSDNSNRDI
		TADTGIVTI	TADTGIVTI
LigB4		FTDNSNSDI	FTDNSNSDI
		IQGSTDFTV	IQGSTDFTV
LigB5	TSIEVSPVL		TSIEVSPVL
	TTITATLGK		TTITATLGK
LigB6	AQKNQGNAY	AQKNQGNAY	AQKNQGNAY
	STLSVTAAK	STLSVTAAK	STLSVTAAK
		QKFSATGIY	QKFSATGIY
		KVSSPVSTL	KVSSPVSTL
		KDITSAVTW	KDITSAVTW
		GATDIKATF	GATDIKATF
LigB7	FTDNSNSDI	FTDNSNSDI	FTDNSNSDI

		EISAALGSI	EISAALGSI
		ERFKATGIF	ERFKATGIF
		VEIQITPAA	VEIQITPAA
LigB8	RQFKATGTY	RQFKATGTY	RQFKATGTY
	MVNNVTGSV	MVNNVTGSV	MVNNVTGSV
	SVLNVTPAL	SVLNVTPAL	SVLNVTPAL
	ATWSSSNPR	ATWSSSNPR	ATWSSSNPR
	YTDHSVQDV	YTDHSVQDV	YTDHSVQDV
	HSVQDVTAL	HSVQDVTAL	HSVQDVTAL
LigB9	FSDKSTQNL	FSDKSTQNL	FSDKSTQNL
	TSIEITPTI	TSIEITPTI	TSIEITPTI
	QSSPIPIV	QSSPIPIV	QSSPIPIV
	FIQSSPIPI	FIQSSPIPI	FIQSSPIPI
	YKFIQSSPI	YKFIQSSPI	YKFIQSSPI
		TINSITHGL	TINSITHGL
		KQFKATGIF	KQFKATGIF
		ITHGLTKQF	ITHGLTKQF
		SNIKAVYKF	SNIKAVYKF
		EITPTINSI	EITPTINSI
		KSTQNLTQL	KSTQNLTQL
	SSNIKAVYK		SSNIKAVYK
LigB10	GSSDIYAIY	GSSDIYAIY	GSSDIYAIY
	FIDGSEQEI	FIDGSEQEI	FIDGSEQEI
	LSIGSSDIY	LSIGSSDIY	LSIGSSDIY
	NEKGLATAL	NEKGLATAL	NEKGLATAL
	YAIYNSISS	YAIYNSISS	YAIYNSISS
	QEITNLVTW	QEITNLVTW	QEITNLVTW
		EITNLVTWY	EITNLVTWY
		DIYAIYNSI	DIYAIYNSI
		SSNKINFNV	SSNKINFNV
		QQFKAIGTF	QQFKAIGTF

LigB11	YSDSTIQDI	YSDSTIQDI	YSDSTIQDI
	QQYTALGVY	QQYTALGVY	QQYTALGVY
	WSSSNSSSI	WSSSNSSSI	WSSSNSSSI
		TIQDISDSV	TIQDISDSV
		ETKGKATAL	ETKGKATAL
	KITATYNSI		KITATYNSI
LigB12	KQFFAVGTY	KQFFAVGTY	KQFFAVGTY
		THIATYGSV	THIATYGSV
		TVSKQFFAV	TVSKQFFAV
		ETKGLVTGI	ETKGLVTGI
		YSAGTKADL	YSAGTKADL

Appendix Table 4: List of predicted HTL epitopes

Domain	ProPred-I	NetMHC 2.3	IEDB
LigA7	VEIQITPAA	VEIQITPAA	
	ILTVSNTNA	ILTVSNTNA	
	VKVTASMGG	VKVTASMGG	
	LVEIQITPAA	LVEIQITPAA	
	ILTVSNTNA	ILTVSNTNA	
	IQITPAAAS	IQITPAAAS	
LigA8	VTWKSSSK	VTWKSSSK	
	ISITATLEK	ISITATLEK	
	IEVSPTRAS	IEVSPTRAS	
LigA9	ITSAVTWHS	ITSAVTWHS	
	ITSAVTWHSSNNS	ITSAVTWHSSNNS	
	VDIKATLGNVS	VDIKATLGNVS	
	IKATLGNVSSQVS	IKATLGNVSSQVS	
	SNNSVATISNTKG	SNNSVATISNTKG	
LigA10	IVLNPTSSH	IVLNPTSSH	IVLNPTSSH
	WKSSNTAYA	WKSSNTAYA	
	IEIVLNPTS	IEIVLNPTS	
	IVLNPTSSHKA	IVLNPTSSHKA	
	LGSISSANA	LGSISSANA	

LigA11	SVLNVTPALL	SVLNVTPALL	
	VIPNNISFA	VIPNNISFA	
	IEVIPNNIS	IEVIPNNIS	
	YQFKATGIY	YQFKATGIY	
	LVSTTNIGS	LVSTTNIGS	
LigA12	YIMITPSYA	YIMITPSYA	YIMITPSYA
	IMITPSYAG	IMITPSYAG	IMITPSYAG
	WFSSNPSSV	WFSSNPSSV	
	QSSYTPVTVTES	QSSYTPVTVTES	
	TVVFYDHHTQS	TVVFYDHHTQS	
	VTWFSSNPSSV	VTWFSSNPSSV	
LigA13	TVVSISNVDD	TVVSISNVDD	
	LVTWSSSNPT	LVTWSSSNPT	
	IVNITISLS	IVNITISLS	
	VNITISLSS	VNITISLSS	
	ITISLSSIS	ITISLSSIS	
LigB1	VVTIQGNRV	VVTIQGNRV	
	VTIQGNRVR	VTIQGNRVR	
	IQGNRVRGI	IQGNRVRGI	
LigB3	VKLIVTPAA	VKLIVTPAA	VKLIVTPAA
	IVTISASSE	IVTISASSE	
	IWNSSNSTIA	IWNSSNSTIA	
	IQVTSNNPNI	IQVTSNNPNI	
	VKLIVTPAAL	VKLIVTPAAL	

	GIVTISASSE	GIVTISASSE	
	IVTISASSEN	IVTISASSEN	
LigB5	WLTVVPAVL	WLTVVPAVL	
	WNSSSAIVS	WNSSSAIVS	
	VLPSIAKGL	VLPSIAKGL	
LigB6	IQINPVNPS	IQINPVNPS	IQINPVNPS
	ITSAVTWFS	ITSAVTWFS	
	WFSSDSSIA	WFSSDSSIA	
	VNPSLAKGL		VNPSLAKGL
LigB7	ILKVTPAQL	ILKVTPAQL	ILKVTPAQL
	VEIQITPAA	VEIQITPAA	
	VILKVTPAQ	VILKVTPAQ	
	IKSSKVILK	IKSSKVILK	
	IQITPAAAS	IQITPAAAS	
LigB8	VLNVTPALL	VLNVTPALL	
	LIRQFKATGTY	LIRQFKATGTY	
	MVNNVTGSV	MVNNVTGSV	
	VTPINPSVATWSS	VTPINPSVATWSS	
LigB9	AVYKFIQSSPIPITV	AVYKFIQSSPIPITV	AVYKFIQSSPIPITV
	FIQSSPIPI	FIQSSPIPI	
	SIEITPTINS	SIEITPTINS	
	WISSDPSKI	WISSDPSKI	
LigB10	SDIYAIYNSISSNKI	SDIYAIYNSISSNKI	SDIYAIYNSISSNKI
	SSDIYAIYNSISSNK	SSDIYAIYNSISSNK	SSDIYAIYNSISSNK

	LVTWYSSKS	LVTWYSSKS	
	IYAIYNSIS	IYAIYNSIS	
	YAIYNSISS	YAIYNSISS	
	FKAIGTFID	FKAIGTFID	
	ITISPSSSS	ITISPSSSS	
LigB11	VTWSSSNSS	VTWSSSNSS	
	IKINPVNNN	IKINPVNNN	
LigB12	FFAVGTYSA	FFAVGTYSA	
	VTWSSSNQS	VTWSSSNQS	
	ISPINTNINTTV	ISPINTNINTTV	

Publications



Characterization of novel nuclease and protease activities among Leptospiral immunoglobulin-like proteins

Pankaj Kumar^a, Yung-Fu Chang^b, Mohd. Akif^{a,*}

^a Laboratory of Structural Biology, Department of Biochemistry, School of Life Sciences, University of Hyderabad, Gachibowli, Hyderabad, India

^b Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA

ARTICLE INFO

Keywords:

Leptospira
Immunoglobulin-like protein
LigA
Nuclease
Protease

ABSTRACT

Bacterial immunoglobulin-like (Big) domain containing proteins play a variety of biological functions. Leptospiral Immunoglobulin-like (Lig) proteins are well-known virulence factors located on the surface of the pathogenic *Leptospira* that act during adhesion, invasion, and immune evasion. The Lig proteins have many roles and have been designated as multifaceted proteins. However, the hydrolyzing function of Lig proteins is not yet investigated in detail. Here, we report novel *in-vitro* nuclease and protease activities in the Ig-like domain of LigA protein. All Ig-like domains were able to cleave DNA in the presence of a divalent ion, but not RNA. Site-directed mutagenesis revealed Mg^{+2} binding residues in the Ig-like domain of LigA7. The basis of novel nuclease activity may be associated with protein adopting different conformation in the presence of divalent ions and substrate as investigated by change of intrinsic fluorescence. The docking of a stretch of double-strand DNA shows the binding on the positive surface of the protein. In addition, the protein is also observed to cleave a general protease substrate, β -casein, in our experimental condition. Our results proposed that the novel functions may be associated with neutrophil extracellular Trap (NET) evasion. Overall this study enhances the basic knowledge of non-nuclease proteins involved in the DNA cleavage activity and makes the foundation to explore its *in-vivo* activity in pathogenic *Leptospira* and other pathogens as well. Moreover, this information may be utilized to develop preventive strategies to interfere with *Leptospira* immune evasion.

1. Introduction

Immunoglobulin (Ig) like domains are widely distributed domains among the proteins and ubiquitously present across the different phyla. Structural features of the Ig-like domain among the various proteins are quite conserved. The domain consists mainly of a β -barrel composed of seven to nine antiparallel β -strands with the typical Greek-key fold and β sandwich topology that displays a structurally conserved core [1]. The length and regularity of the β -strands, as well as the length and structure of the connecting loops, are extremely variable. The Ig-like domains are usually grouped into Ig superfamily with having a sequence homology. In contrast, there are many Ig-like domains in proteins that bear no sequence homology also. The presence of Ig-like domains has been reported in a large number of proteins with diverse biological functions. They are found in proteins involved in cell-cell recognition, cell-surface

receptors, muscle structure protein, and the immune system [2]. Ig-like proteins are also reported in bacterial species and are known as Bacterial Ig-(Big) domain frequently found on cell surface proteins. The Big domain plays a vital role in the adhesion of bacteria to the host cell and helps in an invasion of pathogenic strains. Intimin and invasin family of outer membrane (OM) adhesins from *E. coli* and *Yersinia*, respectively, containing Ig-like domains, have been studied in detail [3,4]. Apart from the structural component of adhesion, Ig-like domains in PapD protein play a chaperone function in the bacterial periplasm to help in pilus assembly. They are also found in a few oxidoreductases family of proteins, sugar-binding proteins, and a few transcription factors as well as in hydrolyzing enzymes [5,6].

Bacterial Ig-like domains/proteins are also found in pathogenic Leptospiral species named Immunoglobulin-like (Lig) proteins. The Lig proteins (LigA and LigB) are primarily located on the surface of the

Abbreviations: Big, Bacterial Immunoglobulin; Lig, Leptospiral Immunoglobulin-like; SUMO, Small Ubiquitin-like Modifier; EDTA, Ethylenediamine tetraacetic acid; ECM, extracellular matrix; C4BP, C4b Binding Protein; NET, neutrophil extracellular traps.

* Corresponding author. Laboratory of Structural Biology, Department of Biochemistry, School of Life Sciences, University of Hyderabad, Gachibowli, 500056, India.

E-mail address: akif@uohyd.ac.in (Mohd. Akif).

<https://doi.org/10.1016/j.abbi.2022.109349>

Received 8 March 2022; Received in revised form 1 June 2022; Accepted 3 July 2022

Available online 9 July 2022

0003-9861/© 2022 Elsevier Inc. All rights reserved.



Immunoinformatics-Based Designing of a Multi-Epitope Chimeric Vaccine From Multi-Domain Outer Surface Antigens of *Leptospira*

Pankaj Kumar[†], Surabhi Lata[†], Umate Nachiket Shankar and Mohd. Akif^{*}

Laboratory of Structural Biology, Department of Biochemistry, School of Life Sciences, University of Hyderabad, Hyderabad, India

OPEN ACCESS

Edited by:

Jochen Mattner,
University of Erlangen—Nuremberg,
Germany

Reviewed by:

Abdelrahman Hamza Abdelmoneim,
Al-Neelain University, Sudan
Alok Jain,
Birla Institute of Technology—Mesra,
India

*Correspondence:

Mohd. Akif
akif@uohyd.ac.in

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Vaccines and Molecular Therapeutics,
a section of the journal
Frontiers in Immunology

Received: 02 July 2021

Accepted: 08 November 2021

Published: 30 November 2021

Citation:

Kumar P, Lata S, Shankar UN and
Akif M (2021) Immunoinformatics-
Based Designing of a Multi-Epitope
Chimeric Vaccine From Multi-Domain
Outer Surface Antigens of *Leptospira*.
Front. Immunol. 12:735373.
doi: 10.3389/fimmu.2021.735373

Accurate information on antigenic epitopes within a multi-domain antigen would provide insights into vaccine design and immunotherapy. The multi-domain outer surface *Leptospira* immunoglobulin-like (Lig) proteins LigA and LigB, consisting of 12–13 homologous bacterial Ig (Big)-like domains, are potential antigens of *Leptospira interrogans*. Currently, no effective vaccine is available against pathogenic *Leptospira*. Both the humoral immunity and cell-mediated immunity of the host play critical roles in defending against *Leptospira* infection. Here, we used immunoinformatics approaches to evaluate antigenic B-cell lymphocyte (BCL) and cytotoxic T-lymphocyte (CTL) epitopes from Lig proteins. Based on certain crucial parameters, potential epitopes that can stimulate both types of adaptive immune responses were selected to design a chimeric vaccine construct. Additionally, an adjuvant, the mycobacterial heparin-binding hemagglutinin adhesin (HBHA), was incorporated into the final multi-epitope vaccine construct with a suitable linker. The final construct was further scored for its antigenicity, allergenicity, and physicochemical parameters. A three-dimensional (3D) modeled construct of the vaccine was implied to interact with Toll-like receptor 4 (TLR4) using molecular docking. The stability of the vaccine construct with TLR4 was predicted with molecular dynamics simulation. Our results demonstrate the application of immunoinformatics and structure biology strategies to develop an epitope-specific chimeric vaccine from multi-domain proteins. The current findings will be useful for future experimental validation to ratify the immunogenicity of the chimera.

Keywords: *Leptospira interrogans*, antigenic epitope, outer surface antigen, vaccine, *Leptospira* immunoglobulin-like protein, subunit vaccine, immunoinformatics

INTRODUCTION

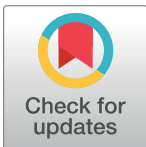
Leptospirosis is categorized as an emerging and neglected tropical zoonotic disease worldwide. It is considered a public health problem globally, with an estimated 1 million leptospirosis cases reported each year, causing deaths of around 60,000 (1–3). The infection usually shows symptoms such as headache, chills, illness, and muscle aches, and a more severe form of disease is associated with

RESEARCH ARTICLE

Phosphorylation of eukaryotic initiation factor eIFiso4E enhances the binding rates to VPg of turnip mosaic virus

Mateen A. Khan^{1*}, Pankaj Kumar², Mohd. Akif², Hiroshi Miyoshi³

1 Department of Life Science, College of Science and General Studies, Alfaisal University, Riyadh, Saudi Arabia, **2** Department of Biochemistry, School of Life Science, University of Hyderabad, Hyderabad, India, **3** Department of Microbiology, St. Marianna University School of Medicine, Kawasaki, Japan

* matkhan@alfaisal.edu

Abstract

Binding of phosphorylated eIFiso4E with viral genome-linked protein (VPg) of turnip mosaic virus was examined by stopped-flow, fluorescence, circular dichroism (CD) spectroscopy, and molecular docking analysis. Phosphorylation of eIFiso4E increased (4-fold) the binding rates as compared to unphosphorylated eIFiso4E with VPg. Stopped-flow kinetic studies of phosphorylated eIFiso4E with VPg showed a concentration-independent conformational change. The dissociation rate was about 3-fold slower for eIFiso4E·VPg complex upon phosphorylation. Phosphorylation enhanced the association rates and lowered the dissociation rates for the eIFiso4E·VPg binding, with having higher preferential binding to eIFiso4Ep. Binding rates for the interaction of eIFiso4Ep with VPg increased (6-fold) with an increase in temperature, 278 K to 298 K. The activation energies for binding of eIFiso4Ep and eIFiso4E with VPg were 37.2 ± 2.8 and 52.6 ± 3.6 kJ/mol, respectively. Phosphorylation decreased the activation energy for the binding of eIFiso4E to VPg. The reduced energy barrier suggests more stable platform for eIFiso4Ep·VPg initiation complex formation, which was further supported by molecular docking analysis. Moreover, far-UV CD studies revealed that VPg formed complex with eIFiso4Ep with substantial change in the secondary structure. These results suggested that phosphorylation, not only reduced the energy barrier and dissociation rate but also enhanced binding rate, and an overall conformational change, which provides a more stable platform for efficient viral translation.

OPEN ACCESS

Citation: Khan MA, Kumar P, Akif M., Miyoshi H (2021) Phosphorylation of eukaryotic initiation factor eIFiso4E enhances the binding rates to VPg of turnip mosaic virus. PLoS ONE 16(11): e0259688. <https://doi.org/10.1371/journal.pone.0259688>

Editor: Michael Massiah, George Washington University, UNITED STATES

Received: June 3, 2021

Accepted: September 29, 2021

Published: November 4, 2021

Copyright: © 2021 Khan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

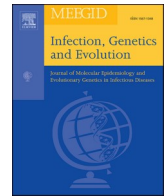
Data Availability Statement: All relevant data are within the manuscript and its [supporting information](#) files.

Funding: This work was supported in part by faculty enhancement research support from Alfaisal University, Riyadh, Saudi Arabia to M.A.K. Grant No. IRG20413. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors declare that they have no potential conflict of interest.

Introduction

Viruses depend on the translational apparatus of the host cells and have developed sophisticated mechanisms to suppress translation of cellular mRNAs whereas ensuring its own translation. Many viruses use a 5'-cap-dependent mechanism, while other viruses use a cap-independent mechanism for initiation. Turnip mosaic virus (TuMV) is a positive strand RNA virus belonging to the *potyviridae* family [1]. While *potyviridae* mRNAs possess poly(A) tail at their 3'-terminus, they lack m⁷G cap structures and instead carry viral genome-linked proteins (VPgs) that are covalently attached to the 5'-terminus [2]. VPg of *Potyvirus* mRNA plays an



Research paper

Immunoinformatics analysis of antigenic epitopes and designing of a multi-epitope peptide vaccine from putative nitro-reductases of *Mycobacterium tuberculosis* DosR

Mohd. Shiraz, Surabhi Lata, Pankaj Kumar, Umate Nachiket Shankar, Mohd. Akif^{*}

Department of Biochemistry, School of Life Sciences, University of Hyderabad, Hyderabad 500046, India

ARTICLE INFO

Keywords:

Mycobacterium tuberculosis
Nitro-reductases
Antigenic T-cell epitopes
Peptide vaccine
Dormancy
DosR

ABSTRACT

Mycobacterium tuberculosis (*Mtb*) resides in alveolar macrophages as a non-dividing and dormant state causing latent tuberculosis. Currently, no vaccine is available against the latent tuberculosis. Latent *Mtb* expresses ~48 genes under the control of DosR regulon. Among these, putative nitroreductases have significantly high expression levels, help *Mtb* to cope up with nitrogen stresses and possess antigenic properties. In the current study, immunoinformatics methodologies are applied to predict promiscuous antigenic T-cell epitopes from putative nitro-reductases of the DosR regulon. The promiscuous antigenic T-cell epitopes prediction was performed on the basis of their potential to induce an immune response and forming a stable interaction with the HLA alleles. The highest antigenic promiscuous epitopes were assembled for designing an *in-silico* vaccine construct. A TLR-2 agonist Phenol-soluble modulin alpha 4 was exploited as an adjuvant. Molecular docking and Molecular Dynamics Simulations were used to predict the stability of vaccine construct with the immune receptor. The predicted promiscuous epitopes may be helpful in the construction of a subunit vaccine against latent tuberculosis, which can also be administered along with the BCG to increase its efficacy. Experimental validation is a prerequisite for the *in-silico* designed vaccine construct against TB infection.

1. Introduction

Mycobacterium tuberculosis (*Mtb*) is a multifaceted pathogen causing Tuberculosis (TB), which remains one of the prominent reasons for mortality worldwide. According to the WHO global TB 2020 report, millions of people are currently infected with *Mtb*. Among that less than 10% of the total infected individuals are affected with active TB; and while more than 90% of individuals are afflicted with latent tuberculosis where *Mtb* resides inside infected macrophages for a longer period of time in an inactive metabolic state and non-transmissible form but reversible state, known as "dormancy". The population of latent tuberculosis infection (LTBI) individuals indeed responsible for a major obstruction to TB control strategies. Although antibiotic treatment offers standard care for active TB, its effectiveness against latent TB is doubtful. The current protective Bacillus-Calmette-Guerin (BCG) TB vaccine – a live-attenuated *Mycobacterium bovis* vaccine (Eddine and Kaufmann, 2005) – is known to protect against dreadful forms of TB in young children (Corbett et al., 2003). However, it does not show efficient and consistent protection in adults against pulmonary TB and also

it doesn't protect from reactivation of the latent TB infection. Hence, significant efforts for the development of new vaccines/therapies that can prevent latent *Mtb* infection are desperately needed. Dormancy in *Mtb* is characterized as a state of low pH, nutrient deprivation, hypoxia, and nitric oxide, which triggers an upregulation of a set of ~48 genes (~1.2% of the *Mtb* genome), known as Dormancy Survival Regulon or DosR (Voskuil et al., 2003). Functions of most DosR regulon gene products are unknown, but many of them are found to be immunodominant that elicit a strong T-cell response. It has been reported that DosR regulon triggers a T-cell response and IFN- γ inducing capability. Rv1733c, Rv2029c, Rv2627c, and Rv2628 were found to elicit a strong immune response as compared to CFP-10, a well-recognized antigen for *Mtb* infection (Leyten et al., 2006). Rv1813c, Rv2628, Rv2029c, and Rv2659c were also reported as latency antigens with a good humoral immune response and produce a higher number of CD4⁺ cells (Liang et al., 2019). Therapeutic effects of these antigens were studied in the endogenous resurgence mouse TB model. In fact, in the same study, Rv2626c and Rv2032 latency antigens were reported to induce a significant effect on CD4⁺ and CD8⁺ cells (Liang et al., 2019). These two

^{*} Corresponding author.

E-mail address: akif@uohyd.ac.in (Mohd. Akif).

<https://doi.org/10.1016/j.meegid.2021.105017>

Received 10 April 2021; Received in revised form 13 June 2021; Accepted 22 July 2021

Available online 28 July 2021

1567-1348/© 2021 Elsevier B.V. All rights reserved.

Anti-Plagiarism report

Characterization of structure, epitope and novel activities in multi-domain Bacterial Immunoglobulin-like (Big) proteins from Leptospira

by Pankaj Kumar

Submission date: 15-Jul-2022 02:16PM (UTC+0530)

Submission ID: 1870808502

File name: n_Bacterial_Immunoglobulin-like_Big_proteins_from_Leptospira.pdf (6.83M)

Word count: 30084

Character count: 167001

Characterization of structure, epitope and novel activities in multi-domain Bacterial Immunoglobulin-like (Big) proteins from Leptospira

ORIGINALITY REPORT

42%

SIMILARITY INDEX

25%

INTERNET SOURCES

39%

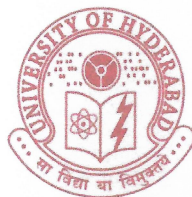
PUBLICATIONS

29%

STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of Hyderabad, Hyderabad Student Paper	14%
2	Pankaj Kumar, Yung-Fu Chang, Mohd Akif. "Characterization of novel nuclease and protease activities among Leptospiral immunoglobulin-like proteins", Archives of Biochemistry and Biophysics, 2022 Publication	14%
3	www.ncbi.nlm.nih.gov Internet Source	8%
4	www.frontiersin.org Internet Source	2%
5	Pankaj Kumar, Yung-Fu Chang, Mohd. Akif. "Characterization of novel nuclease and protease activities among Leptospiral immunoglobulin-like proteins", Archives of Biochemistry and Biophysics, 2022 Publication	1%



University of Hyderabad
Hyderabad- 500 046, India

PLAGIARISM FREE CERTIFICATE

This is to certify that the similarity index of this thesis as checked by the Library of the University of Hyderabad is 42%. Out of this, 37% similarity has been found from the candidate's (**Pankaj Kumar**) own publications which form the adequate part of the thesis. The details the of student's publication are as follows:

Paper	% Similarity
1. Kumar P., Chang YF., and Akif M. (2022) Characterization of novel nuclease and protease activities among Leptospiral immunoglobulin-like proteins. <i>Archives of Biochemistry and Biophysics</i> 727: 109349	15% (Source No. 2 & 5)
2. Kumar P., Lata S., Shankar UN., and Akif M. (2021) Immunoinformatics- Based Designing of a Multi-Epitope Chimeric Vaccine From Multi-Domain Outer Surface Antigens of <i>Leptospira</i> . <i>Front. Immunol.</i> 12:735373.	22% (Source No. 1, 3 & 4 as 13%, 7% & 2%)

About 5% similarity was identified from external sources in the present thesis which is in accordance with the regulations of the University. All the publications related to this thesis have been appended at the end of the thesis. Hence the present thesis may be considered to be plagiarism free.

Dr. Mohd. Akif

Supervisor

MOHD. AKIF Ph.D.
Assistant Professor
DEPARTMENT OF BIOCHEMISTRY
SCHOOL OF LIFE SCIENCES
UNIVERSITY OF HYDERABAD
HYDERABAD-500 046. (INDIA)