Kode Technology peptide-based kodecyte diagnostics using Spirochetes and SARS CoV-2 as models

Radhika Nagappan

A thesis submitted to
Auckland University of Technology
in fulfilment of the requirements for the degree of
Doctor of Philosophy (PhD)

2021

Centre for Kode Technology Innovation

School of Engineering

Abstract

Despite advanced developments in molecular testing of disease, antibodies, as biomarkers, play an important role in diagnosing disease, checking immune response and for serological surveillance. There is an increasingly urgent need for rapidly adaptable, sensitive, low-cost antibody diagnostics not only for existing diseases but also for the proper case management and control of emerging and re-emerging infectious diseases, particularly in resource constrained settings. Peptide based diagnostics are a potential alternative to tests that involve recombinant protein antigens, which, while generally effective, have constraints with respect to reproducibility and adaptability.

Kode technology is a highly adaptable platform that uses function-spacer-lipid (FSL) constructs to attach epitopes to cells (kodecytes) for use in diagnostic assays. The feasibility of designing more refined synthetic antigens (short peptide epitopes) provides potential for enhanced sensitivity and specificity. Kodecytes can be implemented easily into a simple, rapid, sensitive and relatively less expensive diagnostic using the Kode technology platform. One objective of this study was to develop an algorithm for designing FSL peptide constructs using bioinformatics.

This research initially selected two different complex pathogens (*T. pallidum* and *Leptospira*) to develop an algorithm and develop Kode technology antibody diagnostics compatible with existing routine serologic platforms. However, with the unprecedented appearance of the COVID-19 (SARS CoV-2) pandemic, this research was extended to create a potential antibody diagnostic assay for COVID-19.

Candidate peptides were made for *T. pallidum*, *Leptospira* and SARS CoV-2 using the peptide identification and FSL peptide selection algorithm. Validation and the functional prediction of candidate peptides were performed using blood samples and the kodecyte assay. Among the candidate peptides, *T. pallidum* and SARS CoV-2 had one potential candidate peptide suitable for diagnostics. Assessing the datasets over time will help in further refining the algorithm.

The syphilis kodecyte assay and COVID-19 kodecyte assay achieved specificity and sensitivity at least equivalent to an established EIA antibody diagnostic. The *Leptospira* kodecyte assay was more challenging to validate, as there was minimal access to samples, but the preliminary results were good, and while most of the candidate peptides worked, further validation is required.

This research built successful kodecyte assays for three diseases. Kodecyte assays described in this thesis are β versions and will ultimately need to undergo regulatory approval processes and product development trials to be able to be implemented in clinical use. However, despite not yet being optimized, the assays reported are functional and usable.

Table of Contents

List of Figures

List of Tables

Attestation of Authorship

I hereby declare that this submission is my own work and that to the best of my knowledge and belief it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements) nor material which to a substantial extent has been accepted for qualification of any other degree or diploma of a university or other institutions of higher learning except where the acknowledgement is made.

Radhika Nagappan

Radhika

Dec 2021

Intellectual Property Rights

Intellectual property rights comprising all aspects of the projects reported in this thesis belong to Kode Biotech Ltd and shall not be passed on to a third party without explicit approval in writing from Kode Biotech Ltd.

Ethical Approval

Ethical approval was obtained for this study from Auckland University of Technology Ethics Committee (AUTEC) to use de-identified samples of human origin (AUTEC#13/300).

Dedication

This thesis is dedicated to my parents.

You both have been my inspiration.

Thank you both for the endless love and support.

Acknowledgements

I would like to thank my primary supervisor, Professor Steve Henry, for providing me an opportunity to pursue my research under his guidance and support, for being patient and for giving me very flexible working hours throughout this journey. It was a great privilege to have a supervisor who challenges one's thoughts and abilities, provides excellent technical support and is an endless source of inspiration.

I would also like to thank my supervisors, Dr Eleanor Williams and Dr Fabrice Merien, for their valuable suggestions and guidance. I would like to thank Fabrice for organising the control samples for *Leptospira* from the Institute of Pasteur, New Caledonia.

I would like to express my thanks to Professor Nicolai Bovin for his advice and guidance, and for sharing his rich experience in academics and research.

I wish to thank LabPlus at Auckland District Health Board, New Zealand Blood Services, and Professor Cyrille Goarant from the Institute of Pasteur, New Caledonia, for providing the blood samples for this study. I also want to thank the Auckland University of Technology for the PhD scholarship.

Thank you to the Kode Biotech team and my fellow PhD students for their friendship and support, and for providing an enjoyable work environment.

Finally, and most importantly I would like to thank my family, my husband and children, for their love, support and sacrifices. Without them, I would not have achieved this personal endeavor.

List of Abbreviations

ACC	Accident Compensation Corporation
ACE2	Angiotensin Converting Enzyme 2
ASSURED	Affordable, Sensitive, Specific, User friendly, Rapid and Robust
Biokit	Bio Elisa Syphilis EIA
CBS	Centre for Biological Sequence Analysis
CDC	Centre for Disease Control and Prevention
CMG	Carboxymethylated Oligoglycine
CNS	Central Nervous System
CAT	Column Agglutination Test
COVID-19	Coronavirus Disease
CSF	Cerebro Spinal Fluid
DFA- ABS	Direct Fluorescent Antibody – Absorption Test
DFA-TP	Direct Fluorescent Antibody test for *T. pallidum*
DFC	Dark Field Microscopy
DOPE	Dioleoyl Phosphatidylethanolamine
ECDC	The European Centre for Disease Prevention and Control
EIA/ELISA	Enzyme Immunoassay/Enzyme Linked Immunosorbent Assay
EMJH	Ellinghausen-McCullough-Johnson-Harris
EQA	External Quality Samples
EUA	Emergency Use Authorisation
ESR	The Institute of Environmental Science and Research Ltd
FDA	Food and Drug Administrator
FITC	Fluorescent Isothiocyanate
FSL	Function Spacer Lipid
FTA -ABS	Fluorescent Treponemal Antibody Absorption
HIV	Human Immunodeficient Virus
HCoV	Human Coronavirus
ICE	ICE syphilis (Abbott -murex) EIA
IF	Immunofluorescence
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IHA	Indirect Hemagglutination Test
LOOP	Loop Mediated Isothermal Amplification
LPS	Lipopolysaccharides
LERG	Leptospirosis Burden Epidemiology Reference Group
MAT	Microscopic Agglutination Test
MSAT	Microscopic Slide Agglutination Test
MERS	Middle Eastern Respiratory Syndrome
NAT	Nucleic Acid Test
NMR	Nuclear Magnetic Resonance

ACC	Accident Compensation Corporation
PC2	Pathogen Containment Level 2
PCR	Polymerase Chain Reaction
PPA	Polyacrylamide conjugated peptides
ProSA	Protein Structure Analysis Program
RBD	Receptor Binding Domain
ROC	Receiver Operating Characteristic curve
RPR	Rapid Plasmin Reagin
SARS CoV	Severe Acute Respiratory Syndrome Virus
STD	Sexually Transmitted Disease
STDI	Sexually Transmitted Diseases Diagnostics Initiative
VDRL	Venereal Disease Research Laboratory
SARS CoV-1	Severe Acute Respiratory Syndrome Virus-1
SARS CoV-2	Severe Acute Respiratory Syndrome Virus-2
WHO	World Health Organisation
WHO-SDI	WHO Sexually Transmitted Disease Diagnostic Initiative

Key Definitions

Term	Definition
Specificity	Negative reaction rate against expected negative samples
Sensitivity	Positive reaction rate against expected positive samples

Chapter 1. Introduction

In this thesis, three infectious diseases were evaluated for their potential to be diagnostically screened using Kode Technology. Two of the assays were designed to screen the antibodies to the spirochetes causing the diseases of syphilis and leptospirosis. The third was designed in response to the COVID-19 pandemic and was designed to screen COVID-19 antibodies.

1.1. Syphilis

Syphilis is caused by *Treponema pallidum,* a spirochete. It is a sexually transmitted disease. The reappearance of syphilis is a public threat in developed nations and remains endemic in developing countries despite affordable and efficient antibiotic therapy. Syphilis is a global medical issue, and in fact, one of the most infectious diseases affecting fetuses and newborns all over the world is congenital syphilis. Twelve million new cases of syphilis occur annually, according to the World Health Organization (WHO), most of which in settings where there is limited or no access to laboratory services. Individuals who are human immunodeficiency virus (HIV) positive generally have concurrent infections with *Treponema pallidum.* Estimates of increased risk of acquiring and transmitting HIV range from 2 to 9 % in individuals with syphilis infection[1i].

1.1.1. Microbiology

T. pallidum is an anaerobic bacterium belonging to the order Spirochaetales, the family Treponemataceae, and the genus *Treponema*. Spirochaetaceae are fastidious microaerophilic organisms which require little oxygen. In the genus Treponema, there are three other treponemes which cause non-venereal disease. These three pathogens are morphologically and antigenically similar, but they can be distinguished from each other by their clinical manifestations, mode of transmission, epidemiological characteristics, and genetic markers[2].
T. pallidum is the most virulent subspecies in the family since it crosses the blood brain barrier and the maternal fetal placental barrier. It is the only pathogen in the family transmitted through sexual contact (Table1).

Table 1. Clinical features of human trypanematoses

Treponema Species	Disease	Mode of transmission	Primary lesion
T. Pallidum spp. pallidum	Syphilis	Sexual and congenital	Genitalia, oral mucosa
T. Pallidum spp. endemicum	Bejel	Mucous membrane	Oral mucosa
T. Pallidum spp. pertenue	Yaws	Skin to skin	Lower extremities
T. carateum	Pinta	Skin to skin	Extremities

T. pallidum endemicum is the etiology for bejel, which is a chronic skin and tissue disease affecting children and adults in Africa and the Middle East. *T. pallidum pertenue* is the etiology for yaws. It is a tropical childhood infection of skin, bones and joints seen in sub-Saharan Africa, South America and the Caribbean. *Treponema carateum* is the etiology for pinta. It is a skin infection seen both in children and adults and is prevalent in warm areas of Central and South America. *Treponema carateum* is a distinct species primarily because no isolates were available to study the bacteria.

T. pallidum is a spiral shaped bacterium and the spiral (Figure 1) is evenly distanced at a distance of 1µm. surrounded by a cytoplasmic membrane with an attached outer membrane. Structural stability is provided by the peptidoglycan layer found between the cytoplasmic and outer membrane. The spirochaete has characteristic corkscrew motility, and it moves with the help of three to six flagella which extend into periplasmic space from both ends toward the centre the spirochete. The composition of the cell envelope is very different from the gram-negative bacteria. The outer membrane of *T. pallidum* is exceptionally fragile, it has very few surface exposed proteins and does not contain lipopolysaccharide. The lack of outer membrane targets has led to *T. pallidum* being labelled as a stealth pathogen[3].

Figure 1: Electron photomicrograph of *T. pallidum.* Modified from the CDC website.

The spirochete *T. pallidum is* six to 20 µm long with tapering ends and an approximate diameter of 0.2 µm and typically consists of six to 14 waves. Metabolic studies revealed that *T. pallidum* is a delicate organism to study, due to its rudimentary biosynthetic capacity and complete dependency on the mammalian host for continuous viability. The scarcity of information about the virulence mechanism of *T. pallidum* is attributed to the complexity of culturing this pathogen. The genome of *T. pallidum* (Nichol's strain) is a circular chromosome, which is one of the smaller genomes, with 1.1 million base pairs (bp).

1.1.2. Epidemiology and global burden

Syphilis remains globally a continuing public health challenge and concern even though it can be diagnosed with inexpensive simple tests and treated with Benthazine Penicillin G. Congenital syphilis is a global disease, causing adverse issues in pregnancies and newborns. Every year six million new syphilis cases occur among people aged 15 to 49. Annually, the spirochete infects two million pregnant women, when they are not treated or inadequately treated[4]. Complications include spontaneous abortion or stillbirth (25%), prematurity (13%), infant mortality (11%) and neonatal complications (20%)[5]. Syphilis has re-emerged in the UK, the USA, Canada, Australia, New Zealand, Russia, and China[6]. There has been an alarming increase in syphilis rates in China in the last two decades, and this is attributed to the societal and economic changes as well as ease of travel in the country[7]. There has also been a sustained increase in syphilis in metropolitan Adelaide (South Australia). According to ProMED[8], syphilis outbreaks are reported in remote areas among indigenous Australians. In Japan, the number of syphilis cases hit 6000 in November 2018 for the first time in 50 years[9]. Epidemiological and biological aspects of syphilis and HIV are considered to be responsible for the re-emergence of syphilis as a global public health threat. Active syphilis infections, both primary and secondary stage, enhance the risk of HIV transmission by twofold to fivefold[10]. The macrophages are genetically altered by the spirochete, making the individual who has contracted syphilis more prone to HIV infection[11]. Modelling studies reveal that if syphilis is effectively controlled, this would have a major impact on HIV[12].

According to the literature, transfusion transmitted syphilis is extremely rare in developed countries, due to the use of refrigerated blood and the loss of viability of spirochetes in the refrigerated blood 13. Although the infection cannot be transmitted by the transfusion of refrigerated blood, syphilis positive serology can be induced into the recipient, which can confound medical diagnosis and has social consequences. Several clinical and animal studies have tested the stored refrigerated blood donor samples for the survival time, infectivity and number of treponemes. It has been confirmed there is a correlation between the number of treponemes and infectivity. The survival time of *T. pallidum* was between 72 and 102 hours[14]. In developing countries, the reason for transfusion transmitted syphilis is because, due to limited blood supplies, blood is collected from family donors and transfused within an hour, enabling the transfer of *T. pallidum*. Syphilis screening tests should be made mandatory to prevent the incidence of transfusion transmitted syphilis in such situations[15].

Transmission of syphilis via solid organ transplantation, though rare, has been reported[16,17]. Testing for syphilis is mandatory in the donor selection process, as it is a potential surrogate for high risk behaviour[13].

1.1.3. Clinical manifestations

Syphilis, with its highly variable symptoms and multiple clinical stages, has earned the name 'The Great Imitator' due to its propensity to mimic other diseases[18]. There is a poor understanding of the factors involved in syphilis pathogenesis. Rapid widespread dissemination seems to be a major pathogenic mechanism for *T. pallidum.* Syphilis, if left untreated, progresses in four overlapping stages based on clinical findings. The overlapping stages are referred to as primary syphilis, secondary syphilis, latent syphilis and tertiary syphilis (Table 2).

Table 2. Clinical manifestations of syphilis

Congenital		Acquired	
Mother to child in utero		Sexual contact	
Early syphilis	Late syphilis	Early syphilis	Tertiary syphilis
Birth to 2 years age	After 2 years of life	Primary	Neurological
		Secondary	Cardiovascular
		Early latent	Gummatous
			Late latent

Unique symptoms and clinical manifestations characterise these stages[2,19]. Congenital syphilis occurs when a syphilis infected mother passes the infection during the pregnancy to the fetus. Early congenital syphilis occurs from birth to age 2, and late congenital syphilis after two years of life.

A. Primary syphilis

In primary infection, the spirochete disseminates through the blood from the site of inoculation to potentially any organ. A chancre appears at the inoculation site between nine and ninety days of infection. The syphilitic chancre initiates first as a papule, then turns into a painless, non-purulent ulcer with a clean base and a clear border, ranging in size from 0.3 to 3.0 cm. When the treponeme inoculation site is other than the genital area, the chancre developed will have an atypical presentation. The majority of the extragenital chancre occurs in the mouth but can occur in multiple body parts, including the hands, as typically happens among health care workers[20]. Regional lymphadenopathy develops in 80% of the cases in genital lesions. The chancre heals spontaneously within three to eight weeks, and this is followed by an asymptomatic period lasting between six weeks to six months. One third of the infected individuals will go on to the secondary stage if not treated [21].

B. Secondary syphilis

Secondary syphilis occurs six to eight weeks after the primary infection and the appearance of primary chancre results in systemic manifestations and will stay for months. The typical clinical symptoms are erythematous skin rashes on the palms of the hands, the soles of the feet and the trunk and mucocutaneous lesions, in addition to fever and lymphadenopathy. When primary syphilis is not treated, it will progress to secondary syphilis and both the primary and secondary stages of syphilis can overlap.

C. Latent syphilis

Latent syphilis is a consequence of untreated secondary syphilis. Clinical symptoms of secondary stage syphilis will disappear, but the spirochete will remain dormant in the body. In this stage, the patients are seroreactive but asymptomatic. Latency is classified into two stages based upon the time of infection. Latent syphilis falls within the first year of infection. Recurrent secondary infection can occur in some individuals. This stage is not contagious. The late latent stage is asymptomatic and occurs more than one year from the primary infection. Serological testing during the late latent stage will be positive, but sexual transmission at this stage is very rare[2].

D. Tertiary syphilis

In the tertiary stage of syphilis, the organs will be affected, resulting in myriads of complications and, often, death. Roughly 30% of syphilis cases end up in tertiary stage ten to thirty years after the primary infection. Both the cardiovascular and central nervous system (CNS) will be affected.

*i) **Cardiovascular syphilis***

Incidence of cardiovascular syphilis, one of the common manifestations of untreated syphilis, is rare, but sporadic cases are still seen. Clinical manifestations include calcification of the aortic wall, incompetence of the aortic valve or ventricular failure. The disease itself starts after five to ten years of infection, but the cardiovascular clinical symptoms may not be evident for more than two decades.

*ii) **Neurosyphilis***

Neurosyphilis occurs when *T. pallidum* enters the central nervous system, and this can happen in all stages of syphilis. The symptoms seen in neurosyphilis are headaches, seizures, confusion, depression, paralysis and weakness in the extremities and other parts of the body[22].

E. Congenital syphilis

Congenital syphilis occurs when syphilis is transmitted to the foetus during pregnancy. According to the CDC morbidity and mortality report, congenital syphilis can be prevented if the maternal

infection is detected early, and a minimum of 30 days of treatment is given before delivery. In an untreated case, the *T. pallidum* passes from the mother to the foetus through the placenta and enters the foetal bloodstream. This can cause deformities in the bones, central nervous system impairment, hearing loss, and inflammation of the skin and mucous membranes[19]. Transmission of syphilis to the foetus through placenta resulting in congenital syphilis remains a global problem.

1.1.4. Laboratory diagnosis - syphilis

The clinical diagnosis of syphilis is very complex, since the clinical presentation can mimic other infectious and non infectious conditions. Hence, laboratory testing of syphilis plays a crucial role in the diagnosis of this disease.

1.1.5. Direct microscopy

Darkfield microscopy can be used for the direct detection of *T. pallidum* from fresh lesions and serous exudates. The direct fluorescent antibody test detects the antigen and fluorescent isothiocyanate (FITC) labelled *T. pallidum* antibody is used for the identification of the antigen using a fluorescent microscope. However, dark field microscopy and direct fluorescent antibody staining require specimens be obtained from moist lesions (lesion exudates or lymph node aspirate) and need to be performed by an experienced microbiologist[23].

A. Polymerase chain reaction (PCR)

Several PCR methods, both real time and conventional, are currently used for the laboratory diagnosis of *T. pallidum.* Although PCR methods are not standardized and no commercial kits are available, PCR is a sensitive and reliable approach for the diagnosis of primary disease[24,25] when lesions are present, but the sensitivity of the PCR is poor (<50%) in later stages of the disease, when blood or urine samples are tested. PCR testing requires special instruments and technically skilled personnel to perform the test and is therefore available only in reference laboratories.

B. Serology

The common approach for diagnosing syphilis is to test for the presence of antibodies. IgM antibodies develop two weeks after exposure to *T. pallidum*, followed by IgG antibodies two weeks after IgM production. In primary syphilis, within three to five days, both IgM and IgG antibodies may be detectable[23]. There are two categories of syphilis diagnostic tests, non-treponemal and treponemal.

*i) **Non treponemal tests***

The non-treponemal tests for syphilis diagnosis are the venereal disease research laboratory (VDRL) test and the rapid plasma reagin (RPR) test. These tests detect anti-phospholipid antibodies. The damaged host cell releases the lipoidal material early in the infection, and as a result, the anti-phospholipid antibody is produced [20]. Anti-phospholipid antibodies are also detected in several other clinical conditions, including autoimmune disease, viral infections, malaria and systemic lupus erythematosus. The limitations of non-treponemal tests include false negativity, false positivity and reduced sensitivity. Reduced sensitivity and false negativity are seen in primary, latent, and, especially, early primary syphilis. False positive results occur due to cross reactivity of the nature of the reagin antibodies produced in diseases and conditions unrelated to syphilis. The VDRL test has better sensitivity and specificity for neurosyphilis detection using the CSF sample and is the recommended test for the diagnosis of neurosyphilis. The median sensitivity for the RPR and VDRL tests during primary infections is 86% and 78%, respectively, and, during late-stage infections, they have median sensitivities of 73% and 71%, respectively [26,27].

*ii) **Treponemal tests***

Syphilitic infections also induce treponemal specific antibodies and, based on this reaction, treponemal assays were designed which are more sensitive and specific compared to non-treponemal tests. These are used as confirmatory tests. They are expensive, technically difficult and performed in reference laboratories.

Fluorescent treponemal antibody absorption (FTA-ABS)

FTA-ABS is an antibody detection test. The *T. pallidum* antigen (Nichol's strain) is coated onto slides. The patient serum is diluted in a suspension containing *T. phagedenis* (non-pathogenic treponeme) in order to remove antibodies shared by *T. pallidum* and the non-pathogenic treponemes. If the antibody is present in the serum, it will attach to the slide coated with the treponeme antigen (Nichol's strain). When fluorescein isothiocyanate (FITC) labelled antihuman immunoglobulin is added, this binds with the patient's IgG and IgM antibodies that are attached to the *T. pallidum* antigen, which, when viewed under the fluorescent microscope, will fluoresce. Due to variation in equipment, reagent and interpretations, the FTA-ABS test is known to present both false positive and false negative results[2]. False positive results have been reported among patients with other spirochaetal infections, and in immunological disorders like systemic lupus erythematosus, rheumatoid factors, anti-nuclear and anti-DNA antibodies or cross-reacting antibodies, such as those directed against *Borrelia burgdorferi*. For the above reasons, FTA-ABS should not be used as a primary screening test for syphilis[27,28].

Treponema pallidum particle agglutination (TPPA)

TPPA is a confirmatory assay used by reference laboratories. The antigen is coated onto the coloured gelatine particles, which serve as carriers for the *T. pallidum* antigen. Agglutination occurs when the antigen binds to the antibodies present in the serum sample, resulting in a smooth mat of agglutination. The TPPA test has a high specificity rate. It has a low sensitivity rate in early primary syphilis disease detection. False negative TPPA results were reported during a retrospective laboratory audit in the United Kingdom[29].

Enzyme immunoassays (EIAs)

Enzyme immunoassay (EIA) is generally used in the serodiagnosis of syphilis. EIA methods can incorporate a direct or indirect sandwich method or a competitive capture method. Two types of EIA methods are commercially available: EIA assay with sonicated antigen and EIA assay with recombinant antigens. With the arrival of recombinant technology, several treponemal recombinant antigens have come into use in EIA. The most common commercial EIAs used are the ICE Syphilis (Abbott Murex) EIA and the Bioelisa Syphilis (Biokit) EIA. The ICE syphilis assay is a two step sandwich assay in EIA format. It uses recombinant antigens TpN15, TpN17 and TpN47. (See chapter 2 for a full description of these antigens). The EIA Bioelisa Syphilis 3.0 kit is a direct sandwich assay which detects antibodies (IgG and IgM) to *T. pallidum.* In addition to TpN15, TpN17 and TpN47, an extra recombinant protein is coated onto the microtiter wells. The sensitivity and specificity of EIA is equivalent to the FTA-ABS and TPPA tests. EIAs can be automated for high-throughput analysis. False positive results in EIAs can occur. The disadvantages of the EIAs are the complexity of the test and the need for specialist equipment to conduct it.

Rapid point of care diagnostic tests

Several rapid agglutination tests are available and are used mainly in developing countries. Rapid agglutination tests use treponema specific antigen either in strips or in carrier particles. In immunochromatographic tests (ICT) treponemal antigens are impregnated in a strip that reacts with syphilis antibodies when present in the clinical sample. In agglutination tests, gelatine particles are coated with treponemal antigen and a positive reaction result in the form of agglutination. The data available on the performance characteristics of rapid tests is minimal. One report states that ICT and particle agglutination tests have sensitivity ranging from 62 to 100% and 72 to 100% and specificity ranging from 83 to 100%, respectively[30].

According to the literature, several evaluation studies of existing syphilis assays have been conducted. Table 3 shows the data published by the WHO in 2013 regarding the treponemal and non-treponemal syphilis diagnostic tests in use at that time. It shows the different parameters and the performance characteristics of each test. The various tests were assessed into three

categories, easy, moderate, and complex, based on the simplicity of the test and the need for specialised equipment and skilled personnel to perform it. Another parameter, the level of use, was analysed, and identified whether the tests can be performed only in the reference laboratories or performed as a point-of-care testing. This study reveals that every method has constraints, including the possibility for both false positive and false negative results[31].

1.1.6. Criteria for syphilis diagnosis

To date, the criteria for reporting syphilis diagnosis has been a topic of debate, and numerous test interpretations have perplexed clinicians. There are currently three algorithms used to confirm syphilis: a traditional algorithm, a reverse screening algorithm, and a third algorithm. The third algorithm is a version of the reverse screening method that the European Centre for Disease Control and Prevention (ECDCP) recommends[32-36].

1.1.7. Need for new diagnostic assays for syphilis

The complexity of syphilis diagnosis, which relies upon multiple tests and interpretations, is a challenge for clinicians. Several recent studies have shown that patients often prefer point-of-care testing (POCT) and are more likely to be treated if needed[37]. The number of patients who come for retesting after the primary RPR screening is very low, according to a retrospective study conducted from 1998 to 2008 in the USA. The use of the POCT test increased the percentage of patients treated for active primary syphilis from 44.8 to 68.3%.[38]. Individuals infected with STDs are often asymptomatic yet infectious, so the need for simple and specific assays is warranted for early detection and treatment, and to prevent transmission.

Before the 1950s, syphilis was one of the oldest and most commonly described transfusion transmitted diseases, but the incidence of transfusion transmitted syphilis is very extremely rare now, due to the shift from the usage of fresh blood to the use of refrigerated blood and blood components. The US Food and Drug Administration (FDA) and the World Health Organisation (WHO) state that syphilis testing should continue to be a mandatory requirement for blood donation. Another reason for compulsory syphilis testing is the fact that syphilis is a potential surrogate for high-risk behaviour for HIV infections. Screening for syphilis using the traditional algorithm is not ideal for blood donor screening or in specialized settings such as blood banks. Assays which detect all stages of *T. pallidum* infection are needed in such environments. The ideal assay should be simple and offer good specificity and sensitivity. It should also cover all the three stages of syphilis infection, unlike non-treponemal tests and current treponemal screen tests that use multiple recombinant antigens. Approximately 90% of the new syphilis reported cases every year occur in low-income countries, the primary reason being the lack of resources necessary to

purchase and perform the non-treponemal or treponemal tests[39]. The World Health Organization's Sexually Transmitted Diseases Diagnostics Initiative (WHO-SDI) has set a yardstick based on the ASSURED (affordable, sensitive, specific, user-friendly, rapid and robust, equipment free and deliverable) criteria to determine if the tests employed meet the disease control requirements[39,40].

Table 3. Comparison of syphilis diagnostic assays

TC	Non treponemal test		Treponemal test					Rapid test		
	RPR	VDRL	EIA	TPPA	TPHA	CLIA	FTA-ABS	TT	TT and NTT	DT
Specimen	SE/PL	SE/PL/CSF	SE/PL	SE/PL	SE/PL	SE	SE/PL/CSF	WB/PL/SE	WB/PL/SE	WB/PL/SE
Ease of use	Easy	Easy	MOD	COM	COM	COM	COM	Easy	Easy	Easy
Sensitivity %	73-100	71-100	82-100	85-100	85-100	82-100	70-100	86 WB-80	87.1-89.4	Syp-85 HIV-98
Specificity %	98-100	98-100	97-100	98-100	98-100	98-100	94-100	94-98	94-98	94-98
Level of use	On-site lab	On-site lab	RF	RF	RF	RF	RF	On-site lab	On-site lab	On-site lab
Equipment								None	None	None
Training	MIN	MIN	MOD	EXT	EXT	EXT	EXT	MIN	MIN	MIN
False negative	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
False positive	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

TC: Test characteristics, CLIA: Chemiluminescence, FTA-ABS: Fluorescent Treponemal Antibody Test, TT: Treponemal Test, NTT: Non Treponemal Test, DT: Dual test - syphilis and HIV, RPR: Rapid Plasma Reagin, Veneral Disease Research Laboratory, EIA: Enzyme Immuno Assay, TPPA: Treponema pallidum Particle Agglutination, SE: Serum, PL: Plasma, WB: Whole Blood, CSF: Cerebro spinal fluid. RT: Reference lab, MOD: Moderate, COM: Complex, MIN: Minimal, EXT: Extensive, Sensitivity and specificity data from WHO: Laboratory diagnosis of sexually transmitted infections 2013.

Table 3 shows the comparison of different syphilis serological assays. As seen in the table, the non-treponemal tests and rapid tests are easy to perform in an onsite lab and require minimal training, whereas the treponemal test are specific tests and are done only in reference laboratories. Both false positive and false negative results are seen in treponemal and non-treponemal tests. The sensitivity and specificity of treponemal tests (EIA, TPPA, and CLIA) are higher than the treponemal tests, with the exception of the FTS-ABS test, which has a sensitivity similar to the non-treponemal tests. Multiple tests and interpretations are involved in syphilis diagnosis.

1.2. Leptospirosis

Leptospirosis is a global zoonotic disease that is now recognised as a re-emerging infectious disease of public health concern in both developed and developing nations[41,42]. It is caused by the spirochete *Leptospira* (Greek – leptos "fine" and Latin – spiral "coil"). Humans acquire infection through direct and indirect contact. Direct contact can be made with animals, infected body fluids, tissues and urine of carrier animals, typically through occupation (butchers) and recreational sports. Indirectly, leptospirosis can be acquired through rain, flooding, and contaminated water. The health impacts of rain and flooding in South East Asia and other countries has led to increased awareness of leptospirosis and the need for improved diagnostic tests for this disease. A group of lead experts from the World Health Organisation (WHO) estimated the annual global burden of leptospirosis to be 873,000 cases and 49,000 deaths[43].

Leptospira is known for its variable signs and symptoms. Leptospirosis often mimics many other diseases, including influenza, dengue, malaria, Hantavirus infection, enteric diseases, aseptic meningitis, pneumonia and other illness with febrile syndromes. This feature makes it difficult to differentiate leptospirosis from other illness based on clinical signs alone. Hence, it is often not appropriately diagnosed [44,45]. Laboratory tests play a significant role in confirming leptospirosis[46].

1.2.1. Microbiology

With a length of between 6 and 20 µm in diameter, the spirochete *Leptospira* is very thin and coiled and has a high motility. It has two periplasmic flagella which extend from either end along the cell body without overlapping. The presence of lipopolysaccharide and the double membrane resembles the characteristics and features of Gram-negative bacteria. It also resembles Gram positive bacteria in terms of molecular architecture. Leptospires are obligate aerobes and slow-growing bacteria that require a temperature between 28 and 30°C for optimal growth [42,45,47,48].

There are two modes of classification for the genus *Leptospira,* phenotypic and genotypic. Phenotypic classification*s* involve *s*erotyping results of *Leptospira.* There are nearly 250 serovars, and phenotypic classification is based on the structural heterogeneity in the carbohydrate component of the lipopolysaccharide (LPS)[49]. There are twenty serogroups, and deoxyribonucleic acid (DNA) similarities in *Leptospira* are used for genetic classification[50]. More than fifty percent of the pathogenic *Leptospira* spp and serovars seen in infections are *L. interrogans* and *L. borgpetersenii* [51].

1.2.2. Epidemiology

Leptospires are widespread, and their abundance is due to their ability to survive in a moist environment outside the host and go on to infect a range of animal species and humans. Rodents,

marsupials, bats and other feral and domestic animals are the natural carriers of pathogenic leptospires. Leptospires persistently colonise the renal tubules of the infected animals, and are excreted by the infected animals intermittently for months or years. Humans are dead-end hosts for *Leptospira* species, which are transmitted to humans by direct contact with infected material or indirectly through water or wet soil that has been contaminated by the urine of infected animals [52]. Leptospirosis is occupationally related to farmers, veterinarians, abattoir workers and hunters[53]. In developed countries, the risk profile for infection is changing as international travel and recreational activities like swimming, canoeing or caving have become increasingly common sources of exposure to infection.

1.2.3. Clinical features of leptospirosis

The specific clinical presentation and the severity of the disease relies upon on the infecting *Leptospira* strain involved, the infectious dose received and the age and immunological status of the individual[42]. Infection in humans is most commonly attributed to contact of the leptospires with mucous membrane of eyes, mouth or damaged skin with urine from animals that are persistent shedders. The time of onset of the illness after the exposure varies in individuals, and the symptoms can persist for several months[44].

The clinical signs and symptoms in leptospirosis range from the very common, like fever and influenza-like symptoms, to severe systemic illness, which develops if left untreated. The reason that leptospirosis is often not diagnosed correctly, under-reported and overlooked by the patient, is the initial flu like illness that presents.

Leptospirosis has a broad clinical manifestation that varies from a mild self-limiting illness to a life-threating disease affecting multiple organs. There are two stages in leptospirosis: the acute stage and the immune stage. The acute stage of the disease will be present for up to a week and be followed by the immune stage. The clinical symptoms include fever, vomiting, headache, anorexia, myalgia, nausea and conjunctival suffusion. In the acute phase, leptospires can be detected in blood, body fluids, CSF, tissues, and in the eye. In the immune stage, antibodies are produced, and the shedding of the spirochete in the urine will take place. In the immune stage of the illness, the spirochete will disappear from the blood, but it will be present in the urine and aqueous humor. Leptospiruria will last for many months[44], and five to ten percent of people with leptospirosis will end up with Weil's disease, a serious and complicated stage of the disease that can involve multi organ dysfunction like kidney failure, hepatomegaly, myocarditis, meningitis, uveitis and pulmonary involvement. Leptospirosis has a case fatality of 5 to 15%.

i) Incidence and burden of leptospirosis

With reported increases in its incidence, leptospirosis has re-emerged as a disease of global public health concern. The ability of the spirochete to survive outside the host is a contributing factor to infection in reservoirs. As a result, humans and other accidental hosts acquire the infection[53]. In New Zealand, according to the Institute of Environmental Science and Research (ESR), most of the notified cases of leptospirosis are acquired through occupation, with dairy farmers and meat workers accounting for the majority of the notified cases.

In New Zealand, only six serovars from the two *Leptospira* species, *L. interrogans* and *L. Borgpetersenii,* have been isolated and are endemic. . The six serovars are *L. interrogans* Pomona, *L. interrogans* Hardjobovis, *L. interrogans* Copenhageni, *L. borgpetersenii* Balcanica, *L. borgpetersenii* Ballum, and *L. borgpetersenii* Tarassovi [54].

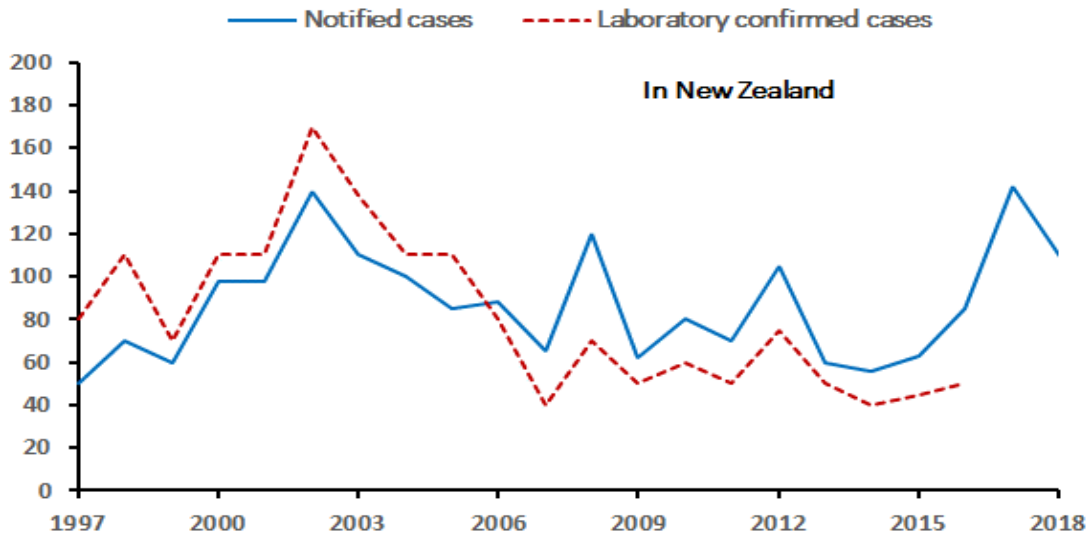

Figure 2: Incidence of leptospirosis in New Zealand

Notified and laboratory confirmed cases of leptospirosis each year recorded by ESR, New Zealand, since 1999.

The 2020 ESR annual public health surveillance notification report (Figure 2) for leptospirosis shows that there was a gradual increase in the number of notified cases of leptospirosis from the year 2015 to the year 2018. The highest notification of leptospirosis was reported from the West Coast and Tairawhiti District Health Boards (18.4 and 14.2 per 100,000 cases, respectively). According to the ESR reports, Hardjo and Ballum are the most notified serovars in humans. Rural New Zealanders are most affected by leptospirosis, and around 50 % of the ESR notified cases are hospitalised[55].

Globally, adverse weather conditions such as heavy rainfall, flooding, and high temperatures, as well as international travel, recreational activities, and overcrowding, have resulted in several outbreaks of leptospirosis[56]. In New South Wales, 50 people contracted leptospirosis from rat urine in July 2019. In 2008, approximately 35% of cases (international travel- and recreation-related) were reported to the enhanced surveillance system of the WHO Collaborating Centre for Reference and Research in Brisbane[57]. In the Philippines, 1030 cases of leptospirosis were reported with 93 deaths in 2018, a 41% increase in the number of reported cases compared to 2017. Several outbreaks of leptospirosis have been reported throughout the world in countries including Sri Lanka [58], Nicaragua [59], the USA[60], New Caledonia[61], and India[62]. The disease is generally under-reported in many countries due to the difficulty in diagnosis and the lack of laboratory facilities. Because of a lack of annual incidence reports for and data associated with leptospirosis outbreaks from various countries, there are no global health policies and guidelines. The WHO started the Leptospirosis Burden Epidemiology Reference Group (LERG) in 2010 to estimate the global burden of this disease and to establish relevant health policies and guidelines to reduce it. The LERG is an outside group which will work with disease-burdened groups in different countries. They will advise WHO regarding the worldwide incidence of leptospirosis, any scientific, technical, and research areas which need action[63].

1.2.4. Laboratory diagnosis of leptospirosis

The confirmation of leptospirosis relies heavily on laboratory diagnosis, since the illness is often either under-diagnosed due to the mild respiratory symptoms that are experienced or misdiagnosed due to the difficulty associated with differentiating it from other febrile illnesses. Several tests are available for laboratory diagnosis, but due to the biphasic (acute and convalescent) nature of the disease, a negative test result may not be genuinely negative and thus does not rule out the disease. Diagnosis depends on the selection of the test and the stage of the disease [45,64,65].

1.2.5. Direct examination

Darkfield microscopic examination (DFM) can be used to visualize the leptospires in blood, CSF, urine and other body fluids. The approximate detection threshold is 10^4 leptospires/mL per cell per field to be seen[66]. Though it is simple and easy to perform the test, it does not provide definitive results. Leptospire will not be present all the time in the blood. The presence of fibrin and other artifacts in the blood will give false positive results[67]. Hence, an experienced person is required to do the examination. The primary constraint is the cost of the dark field microscope, and not all laboratories can afford this tool.

A. Culture

The definitive diagnosis of leptospirosis is made through the isolation of the leptospires in a culture[41]. The Ellinghausen-McCullough-Johnson-Harris (EMJH) culture medium is the medium used for growing *Leptospira.* The collection of the samples should be done before giving the antibiotics to the patient. Blood, CSF and other body fluids can be used for culture during the first week, at the acute stage of the illness[65]. The urine sample should be cultured from the second week onwards, and should be processed immediately, since leptospires do not survive in the collected sample. Cultures are incubated at 28 to 30°C and examined weekly using a darkfield microscope. Leptospires are slow-growing organisms. Hence, the culture should be kept for 13 weeks before reporting the clinical sample as negative for leptospirosis[68]. Though culture is the confirmatory test, it is not suitable for immediate management of the disease for the following reasons. This method is tedious, time-consuming, technically demanding, less sensitive, needs a biosafety cabinet and can only be done in a reference laboratory[41].

B. Serological tests

Serology plays an important role in the diagnosis of leptospirosis, and it is a commonly used approach. Several serological methods have been developed to diagnose leptospirosis during the early stages of the disease for the treatment to be effective.

The Microscopic agglutination test (MAT)

The microscopic agglutination test (MAT), which uses live antigen suspension of the leptospiral serovars, is the standard test used for the serological diagnosis of leptospiral infections. MAT detects the total antibodies, both IgM and IgG. Paired sera (samples collected at the acute and convalescent stages of the illness to see a four-fold increase in antibody titre) are essential to confirm the disease. Each serum sample is tested with a panel of antigens. Live antigen suspension and the serum are mixed together and visualised under a dark field microscope for agglutination. A four-fold increase in titre between the two samples confirms the acute illness[69]. No increase in the titre between the samples suggests past infection. MAT will identify the serovar, but if the infecting serovar is not in the panel of antigen suspension, then a false negative result will be produced. This test is not suitable for the early or acute stage of the illness. MAT has a sensitivity of 41% in the first week, 82% from weeks two to four and 96% beyond week four[70]. MAT requires a large panel of live suspensions of *Leptospira* of different serogroups to provide adequate coverage of the antigenic diversity represented in a given area. It requires constant maintenance of the different serovars, and the handling of the live culture is hazardous for the performing employee, so this should be handled in a biosafety cabinet. MAT has low sensitivity [71], and it also requires paired sera for testing. Cross reactivity with other leptospiral

serovars has been documented. MAT also requires a dark field microscope to see the agglutination, equipment that is not available in routine diagnostic laboratories [68,70].

i) Enzyme linked immunosorbent assay (ELISA/EIA)

The EIA test is simple and easy to perform in a routine laboratory, unlike the reference test MAT, which is complicated. There are several EIA kits available in the market, and the majority of these kits use the saprophytic *Leptospira* biflexa Patoc 1 strain as the whole antigen. This strain shares many surface antigens with pathogenic *Leptospira.* Recombinant surface proteins or lipoproteins of *Leptospira* as antigens are used to increase the specificity of the EIA assay. Variation in the sensitivity and specificity of EIA is reported. Various factors like the study population, previous exposure and the antigen used may result in the test recognising the local circulating strains. The Panbio IgM ELISA kit (Abbott Diagnostics) recommended by the WHO has variable sensitivity and specificity in different countries. This suggests that the ELISA assays need to be validated for distinct epidemiological situations using specific parameters like the difference in the study population and the prevalence of various infecting serogroups. The advantage of EIA is that, if the antigen is selected appropriately, it can detect both IgG and IgM separately. EIA can detect the antibodies earlier than MAT. IgM antibodies appear by day two[72], and if the sample collected is from the acute stage of illness, positive IgM results can be used to help manage the disease. The test can be performed with minimal training.

ii) Immunofluorescence assay (IFA)

The immunofluorescence technique can detect both antigen and antibodies in body fluids and antigen in the tissue section. The sensitivity and specificity of IFA are similar to ELISA. The assay sensitivity ranges from 78% to 92% and specificity from 85% to 98%. IFA needs a fluorescent microscope, and it is available only in reference laboratories.

iii) Microscopic slide agglutination test (MSAT)

The microscopic slide agglutination test (MSAT) uses killed suspension of *Leptospira* serovars. Due to the known strain variation in the regions, the antigen should consist exclusively of locally prevalent strains or the saprophyte *Leptospira biflexa* Patoc I strain, which shares many surface antigens with pathogenic strains. MSAT is relatively insensitive and has a low reproducibility; it is not acceptable in a clinical setting to rely on MSAT alone[73].

iv) Indirect hemagglutination test (IHA)

The indirect hemagglutination (IHA) test uses human blood group O erythrocytes sensitised with genus-specific leptospiral antigens derived from the *Leptospira biflexa* Patoc I strain. When the antigen-coated cells are mixed with the serum sample containing the antibody, the antigen

antibody will combine resulting in agglutination of the cells. In areas where leptospirosis is endemic, the sensitivity of the IHA assay varies from satisfactory to unsatisfactory. A possible reason is due to the study design, including the inclusion of distinct epidemiological populations and case control studies[74].

*v) **Microcapsule agglutination test***

The microcapsule agglutination test uses a synthetic polymer in place of RBC coated with sonicated antigens from leptospires of different serogroups. This test has been evaluated in many reference centres and has shown sensitivity similar to MAT and IgM ELISA in the acute phase, but fails to detect infections caused by some serovars[75].

*vi) **Lepto dipstick***

The Lepto dipstick test is simple and easy to perform for detecting the presence of *Leptospira* specific IgM antibodies. The test strip contains two bands: the control and the antigen band. The serum sample is added to the strip and incubated at room temperature. If the serum contains the *Leptospira* IgM antibody, it will react with the strip's antigen and form a stained coloured band. The intensity of the antigen staining is compared with the intensity of the control band. The test shows cross reactivity in patients with other disease conditions. Given the cross reactivity and variability, it is recommended to test a second sample. A total of 2665 acute serum samples from 12 countries were tested as part of the evaluation. The assay's sensitivity range is from 35% to 86%, and the specificity range is from 88% to 99% [44].

*vii) **Lateral flow assay***

The lateral flow assay uses heat treated whole cell lysate antigen derived from *Leptospira biflexa* Patoc I strain. The undiluted serum is added to the sample well. Distinct staining of the antigen is observed in the form of a line if the serum sample has the antibody. A blood sample collected finger prick blood can also be used. According to the literature, the assay has a sensitivity range from 53% to 86% and a specificity from 89% to 94%. Cross reactivity was observed when the samples from meningitis and rheumatoid arthritis were tested. A repeat sample is required at a later stage to rule out both false positive and false negative results[76,77].

C. Molecular methods

*i) **Polymerase chain reaction***

The polymerase chain reaction (PCR) test is sensitive for the detection of the acute stage of leptospirosis, which aids in early diagnosis. Several conventional and real-time PCRs are in use, but very few PCRs have validation for use in humans and animals with various samples[78-81]. Serovar identification and quantification are not a requirement for patient management. For PCR

to be efficient and for early diagnosis of leptospirosis, the test has to be performed on an adequate sample from the acute stage of the illness. There are few methods reported in the literature to identify the leptospiral serovar using melt curve analysis and sequencing [82]. PCRs are available only in reference laboratories, they are expensive and they require specialised equipment and technical expertise.

ii) ***Matrix-assisted laser desorption ionization time-of-flight mass spectrometry***

The matrix-associated laser desorption ionisation time-of-flight mass spectrometry (MALDI TOF-MS)[83-85] is used for the species level identification of the leptospiral isolates. Protein peaks can differentiate the leptospiral strains. Species level identification with the MALDI-TOF MS is faster and more reliable than other molecular typing methods. The MALDI-TOF MS system cannot detect pathogens directly from clinical material other than urine. Various attempts have been made to analyse other clinical sample types directly, however, no working protocols are in place to date for routine diagnostics. The use of a direct clinical sample is not suitable for this method, which requires culture material with a high bacterial count, a specimen that contains little human protein and skilled personnel to perform the test. For analysis, MALDI-TOF MS requires a reference database that covers all the existing strains. The cost of the instrument can be substantial, and the annual maintenance costs are estimated to be 8 to 10 percent of the instrument cost. Therefore, MALDI-TOF MS is not an ideal test method for routine diagnostics.

iii) ***Isothermal methods***

The isothermal amplification method, particularly loop mediated isothermal amplification (LOOP), is growing popular in the field of molecular diagnostics. Several LOOP assays are in use for the identification of viral, bacterial and parasitic pathogens. Unlike conventional or real-time PCRs, this method does not require thermal cyclers. The LOOP method is an attractive, simple alternative to the PCR, as it only requires the sample to be tested, primers, an enzyme and a water bath at constant temperature (60 to 65°C). LOOP is a one-tube reaction with all the reagents heated in a reaction tube at a constant temperature for around 45 minutes. Fluorescent dyes or turbidity check are the methods of choice to detect the amplified product. For example, the addition of fluorescence SYBR green dye in the amplified product will change the colour of the product from orange to green in a positive reaction. The diagnostic sensitivity and specificity of LOOP for leptospiral diagnosis is moderate to low, particularly in endemic areas[86].

1.2.6. Need for better diagnostic assays for leptospirosis

Accurate diagnosis is essential for the initiation of effective antibiotic therapy and to prevent severe complications. Prediction of the causative pathogen based on clinical signs and symptoms alone is not possible[64]. Leptospirosis is known for its nonspecific clinical features during the initial

stage of the disease. The gold standard test for leptospirosis diagnosis is MAT, which relies on live cultures; this results in problems with respect to maintenance and safety. Rapid diagnostic tests like EIA/ELISA have been reported to have varying diagnostic performance in distinct countries or regions. Tests based on whole cell antigens lead to false positive results due to persistent or cross reacting antibodies, and recombinant antigens will have specificity with antigen homologous infections only[87]. Saprophytic *Leptospira* strain, *L. biflexa,* is not an ideal strain to be used, because the *L. biflexa* strain does not have the virulence factors identified in pathogenic *Leptospira* spp, like the immunoglobulin like Lig proteins. Ninety lipoproteins present in the pathogenic *Leptospira* have no orthologs in *L. biflexa.* Similarly pathogenic *Leptospira* is missing one-third of the *L. biflexa* genes[88]. Sequence variation between pathogenic *Leptospira* species and serovars should be taken into account while designing a test in order to have good sensitivity.

Both commercial and in-house diagnostic tests are in use without robust validation schemes, thus hampering the selection of the test of choice[89]. The lack of sufficient and simple-to-use lab diagnostics is an issue in leptospirosis. Public health authorities and clinicians have pointed out there is an urgent need to develop more effective technologies to support case detection and diagnosis[87].

Molecular assays like conventional and real-time PCR are sensitive during the acute stage of an illness, but they require expensive instruments and technically skilled personnel to perform the tests. They are not suitable for laboratories with limited resources, especially those in developing countries where leptospirosis is endemic.

Serology is the primary test used for diagnosing leptospirosis. In acute infections, a quick answer is needed for treatment and prognosis that will eliminate alternative diagnoses which may confuse strategies for management and treatment. Effective and accurate testing also helps to indicate the direction that can be taken for epidemiological control and preventive measures.

There is a demand for a simple, easy-to-perform diagnostic assay which can be used with minimal training as POCT in any resource-limited settings.

1.3. Severe Acute Respiratory Syndrome Coronavirus (SARS CoV-2)

Coronaviruses are a large and diverse viral family that infect a wide range of hosts, including bats, reptiles, birds, rodents, camels, civet cats, raccoon dogs and humans. Coronaviruses cause a range of diseases including enteritis in animals, upper respiratory infection in birds, and lethal respiratory infection in humans[90,91].

1.3.1. Classification of coronaviruses

Coronaviruses are classified based on genome structure and phylogenetic relationships. Coronaviruses belong to order Niroviridale, family Coronaviridae, and subfamily Coronavirinae. There are four genera in the subfamily: Acoronavirus (α - HCoV), β-coronavirus (β - HCoV), γ-coronavirus (γ - HCoV), and δ-coronavirus. (δ - HCoV) (Figure 3). The α - HCoV and β - HCoV coronaviruses infect animals and humans. The γ-coronavirus and δ-coronavirus mainly infect birds[92].

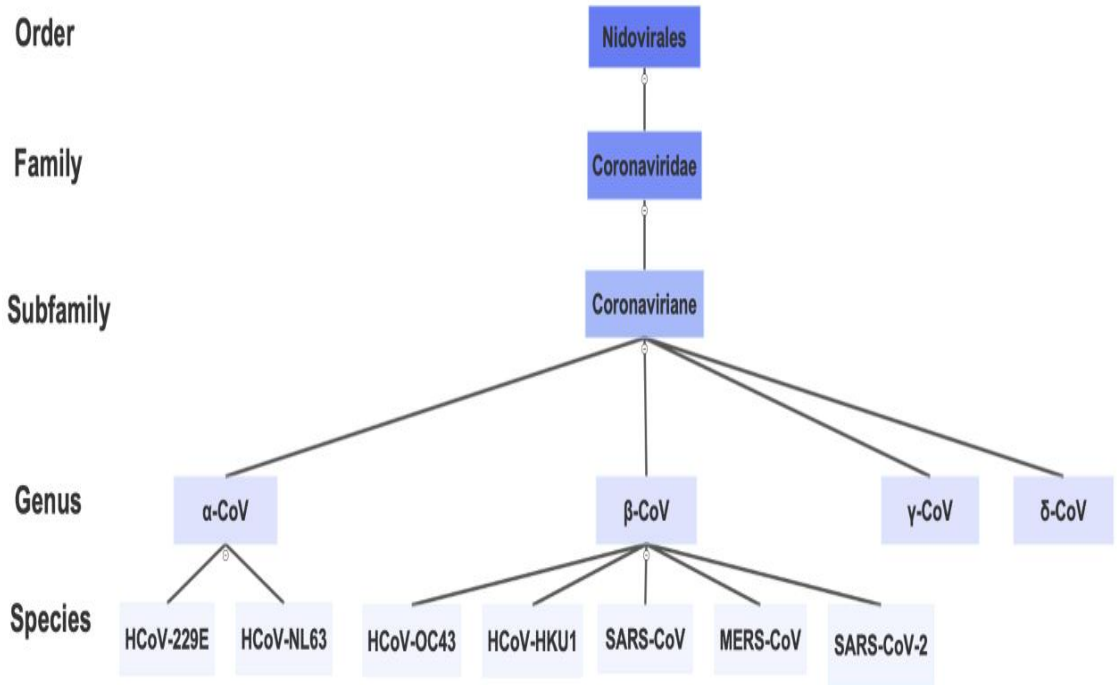

Figure 3: Classification of coronaviruses

HCoV: Human Coronavirus, CoV: Coronavirus, SARS CoV: Severe Acute Respiratory Syndrome Coronavirus, MERS CoV: Middle East Respiratory Syndrome Coronavirus, SARS CoV-2: Severe Acute Respiratory Syndrome Coronavirus-2, HCoV-NL63 HCoV-OC43 and HCoV-HKU1: Human coronavirus strain types.

The four pathogenic human coronaviruses (HCoV) found globally are the HC0V-NL63, HC0V-229E, HC0V-HKU1, and HC0V-OC43. They all cause mild upper respiratory illness similar to the common cold in immunocompetent individuals and account for 30% of upper respiratory tract infections annually.

In 2002, the human pathogenic coronavirus, severe acute respiratory syndrome coronavirus (SARS CoV), emerged and the index case was identified as a human on 16 November in Guangdong, China. SARS CoV caused a global outbreak infecting 8000 people and causing 774 deaths. The WHO announced that the SARS CoV was contained on 5 July 2003 and it disappeared in 2004. The 6th human pathogenic coronavirus, the Middle Eastern Respiratory Syndrome Coronavirus (MERS CoV) emerged in the Middle East in September 2012, causing severe respiratory infections with a high fatality rate of more than 50%[93]. The genomic analysis of this

virus shows that the virus may have originated from bats and then was transmitted to camels. Humans became infected from contact with the dromedary camels, which are the reservoirs for MERS CoV. The WHO MERS CoV case report issued in March 2016 stated that globally there were 1651 laboratory confirmed cases and 590 deaths. MERS CoV still causes localized outbreaks. In 2019 the novel coronavirus, COVID-19 (CO-Corona, VI-Virus, D-Disease, Year "19") disease outbreak occurred with respiratory illness emerging from Wuhan, China in December 2019. The virus was officially named the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS CoV-2) by the International Committee on Virus Taxonomy (ICTV). SARS CoV-2 belongs to the genus β - HCoV and the subgenus sarbecovirus. The virus spread rampantly worldwide and the WHO declared it to be a Public Health Emergency of International Concern on 30 January 2020 and proclaimed it to be a pandemic on 11 March 2020. The human coronaviruses types NL63, 229E, SARS CoV, MERS and SARS CoV-2 originated from bats and OC43 and HKU1 originated from rodents as per the available sequencing data.

1.3.2. Structure of SARS CoV-2

SARS CoV-2, when analysed by scientists using electron microscope imaging techniques, was found to have crown like projections (Figure 4); it was later confirmed that the virus belongs to the coronavirus family and the subfamily β coronavirus. It is a single stranded RNA virus and the genome of SARS CoV-2 is around 25 to 30 kb [94].

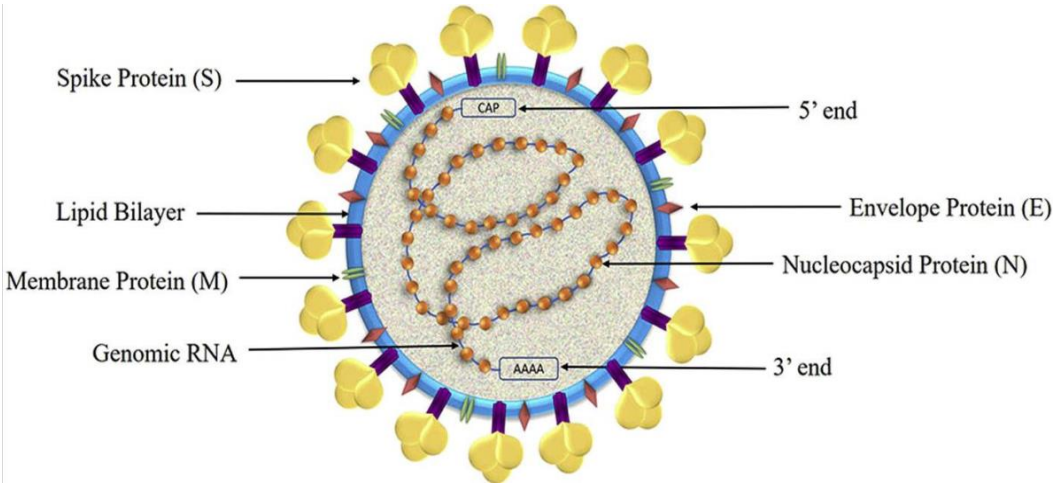

Figure 4: Schematic representation of SARS CoV-2

There are structural, non-structural and accessory proteins in the genome. The four structural proteins (Figure 4) are spike protein (S), nucleocapsid protein (N), envelope protein (E), and membrane protein (M)[95]. They all play a crucial role in the viral replication and infection process. The spike protein consists of two subunits, S1 and S2, and promotes receptor binding and membrane fusion. The receptor binding domain in the S1 subunit first recognises the human

angiotensin converting enzyme 2 (ACE2) receptor and latch, followed by viral and cellular membrane fusion. It also plays a role in spreading the virus to the uninfected cells from the infected cells by influencing the adhesion[96]. The S protein sequence of SARS CoV-2 has ~ 66 % similarity with the SARS CoV-1 S protein, particularly in the S1 subunit. The nucleocapsid protein is an abundant protein that helps in virus penetration and is involved in the cellular processes required for virus survival in the host. The N protein sequence of SARS CoV-2 has a 90 % similarity with the SARS CoV-1 N protein sequence[97]. The envelope protein is a small integral protein with 109 amino acids located at the C-terminal. It promotes virion formation, which is essential for viral assembly, morphogenesis and pathogenesis. The membrane protein produces the ribonucleoproteins and mediates the inflammatory response of the host. However, the precise function of many minor viral glycoprotein components has yet to be determined. To further understand the structural properties of SARS CoV-2 that underpin diverse pathogenic pathways, more research is needed[98].

1.3.3. Pathogenesis of SARS CoV-2

Since the start of the SARS CoV-2 outbreak in December 2019, a massive global effort has been underway to characterise the virus and determine its clinical course. Coronavirus disease (COVID-19) has a biphasic pattern, an early viral response phase and an inflammatory secondary phase[99]. The virus is transmitted primarily by infected respiratory droplets through the nasal, oral and conjunctival mucosa[100]. The clinical symptoms of SARS CoV-2 are similar to flu-like illness, with fever, cough, headache and myalgia. In some cases, symptoms can be severe, like those experienced with pneumonia, or the illness can lead to multi organ failure due to the spike protein attachment to the ACE2 receptor present in the epithelial cells of various organs and blood vessels, resulting in massive replication of the virus[101].

1.3.4. Laboratory diagnosis of SARS CoV-2

The US Center for Disease Control (CDC) constantly updates the guidelines for laboratory testing and information regarding new methods being approved by the WHO through the Emergency Use Authorisation (EUA) for the laboratory diagnosis of SARS CoV-2. The current laboratory diagnostic methods for SARS CoV-2 are nucleic acid tests (NATs) and antibody detection.

As soon as the virus was identified by PCR testing, sequencing of the SARS CoV-2 genome led to the development of nucleic acid tests with several viral targets. The viral RNA amplification by PCR is the gold standard and the sole method of choice for detecting the SARS CoV-2 infection and for contact tracing. The preferred specimen for PCR is a nasopharyngeal swab. There are several challenges to date globally with this workflow. The biggest bottleneck is the supply of

reagents and proper pieces of equipment, not to mention the time, intense labour and cost involved in the RNA extraction process. False negative results are seen in PCR[102], perhaps due to low viral genome count or inadequate sample collection. NAT will miss asymptomatic and mild infections.

Serological testing for antibodies plays a vital role in understanding the immune response to many infectious diseases. There are several reasons for the delay in developing serological tests for SARS CoV-2, mainly the lack of positive control sera and the requirement for thorough specificity and sensitivity testing in the setting of seasonal coronavirus immunity[103]. Several commercially available antibody assays using different platforms, like ELISAs, LFTs and the Western blot method, are widely used, and new antibody assays are being developed. SARS CoV-2 antibody assays are critical to developing an understanding of antibody responses following SARS-CoV-2 infection and vaccination. The limitations for such antibody assays are that they cannot detect early infection and that the chances of false positive results due to cross reacting antibodies are significant[104].

1.3.5. Need for SARS CoV-2 antibody assay

SARS CoV-2 antibody detection will not only complement molecular detection and contact tracing. More importantly, this serosurveillance will also give a complete estimation at a global level, which is very important from a public health perspective in the current pandemic situation and for intervention strategies. The WHO has approved and still is approving several commercial antibody assays using spike protein under Emergency use Authorisation (EUA), where sensitivity and specificity are not thoroughly tested in all the EUA assays.

The need for SARS CoV-2 antibody assay is to:

- check the immune status of the individual for clinical management
- monitor the efficiency of the vaccine to reach herd immunity in the community
- look at antibodies for therapeutic applications.

1.4. Peptides based diagnostics

Recombinant antigens have been used extensively for specific antibody detection in clinical diagnostic laboratories since the advent of recombinant DNA technology in the early 1990s. Certain recombinant antigens appear to be less immunoreactive than the corresponding purified human antigens due to the absence of post translational modifications of the expressed proteins or misfolded structures adopted by the recombinant antigens. The production of recombinant antigens of high quality and in large quantities is laborious and expensive. Problems with the

reproducibility of the assay methods are documented because of the inter assay reactivity variations among different protein batches used to make the antigen[105].

Synthetic peptides are a beneficial tool for the laboratory diagnosis of infectious diseases. Peptides are chemically stable, suitable for long-term storage and can be generated in huge quantities at a minimal cost[106]. Using synthetic peptide antigens to develop the new diagnostic system offers an excellent alternative to traditional assays. The use of synthetic peptide antigens will remove the challenges associated with reproducibility, variability and nonspecific reactions encountered in the traditional immunoassays. Another advantage is that the synthetic peptide structure can be modified to enable the production of chemically well-defined antigens. Synthetic peptides provide enhanced specificity and can be implemented easily into a simple, rapid, sensitive and relatively less costly diagnostic kit[105,107,108].

Peptide based diagnostic tests aim to identify specific antibodies produced by the full proteins using shorter protein fragments containing the strong antigenic determinants. A good peptide selection process is imperative to facilitate an excellent diagnostic assay.

1.4.1. Peptide antigens

An antigen is a molecule that stimulates the immune response and binds to the antibody produced by the immune response. Antigenicity is defined as an antigen's ability to be recognised by one or several antibody components. Antigens can be proteins, carbohydrates, lipids and nucleic acids, either on their own or in combination, as in lipopolysaccharide.

Protein antigens, from a biological perspective, can produce a potent immune response and hence are considered the most suitable antigens. Proteins are conformationally defined molecules made up of long stretched covalently linked amino acids called polypeptides, whereas peptide antigens are short sequences of ten to 30 amino acids. The functional aspect of the proteins is structure related. Proteins form particular folding patterns to create the secondary structural elements called helices, β sheets, turns and loops to perform the specific functions. These structural elements, when tightly packed, result in the tertiary structure. The folding pattern depends on the type, location and side chain interaction of the amino acids in the sequence[109,110]. The tertiary structure, in association with other proteins or subunits, will form the quaternary structure.

The key event in the immune system is the antigen and antibody interaction that serves to clear the invading pathogens, and they are mostly the protein molecules from the pathogens. The immune system recognises fragments of antigen rather than the entire antigen. The host body recognises the antigens in two distinct ways (T and B cells). Antigens are processed and then

presented to the T cells by the antigen presenting cells, whereas antibodies on the B cell surface can recognise the structure of the protein antigen. The fragments of an antigen that come into contact with the antigen binding site (paratope) of a B cell antibody are called epitopes. In the case of proteins, they are called peptide epitopes. Protein epitopes are of two kinds: linear (or continuous) and discontinuous (or conformational). In linear epitopes, the amino acids in the protein's peptide chain are found adjacent to each other or in a continuous sequence. In discontinuous epitopes, the amino acid residues are located in different areas of the peptide chain of the protein. The residues come in close proximity when the protein structure folds and forms an epitope with a specific structural conformation, hence they are also referred to as conformational epitopes. Most epitopes are discontinuous epitopes. What makes an epitope act as an antigen, why a particular region of the antigen becomes the B cell epitopes and the exact selection mechanism involved, are still unclear [111].

A. Identification methods of B cell epitopes

Identifying B cell epitopes is a critical step in immunodiagnostics, vaccine development and treatment. Since the protein function is structure related, both the functional and structural epitopes can be identified by experimental methods. Structural mapping methods define and decipher the protein structure and amino acid's interactions with the antibody. X-ray crystallography and nuclear magnetic resonance (NMR) methods are used to predict structural epitopes[112]. Functional epitope mapping methods (PIN ELISA, Teabag positional scanning, overlapping peptide library screening, site directed mutagenesis and phage display method) are comparatively more straightforward than structural methods. Both structural and functional prediction methods, though ideal in a number of ways, are labour intensive and require many resources.

B. Current approach B cell epitope predictions

Predicting B cell epitopes can now be done using a variety of methods. There are three types of approaches: sequence-based methods, structure-based methods and combinational methods.

1. Sequence based methods: The advantage of sequence-based methods is that they do not require the protein's 3D structure. They incorporate propensity scale methods with different physicochemical properties of amino acids and assign a score for prediction. Propensity scale methods date back to 1981, and have been developed with the addition of several new propensity scales for different physicochemical properties like hydrophilicity[113], flexibility[114] surface accessibility[115] and several others.

2. Structure based methods: Structure genomics predicts the protein structure. Advanced computer modelling techniques and programs in structure genomics help in experimental structure prediction and comparative modelling. There are two methods of 3D protein structure prediction: template-based structure prediction and the free folding method.

Template based structure prediction: One of the prerequisites for the right protein structure prediction is the availability of a similar structure in the protein database. This is the basis for the template-based structure prediction method: finding a similar protein structure for the sequence which needs structure predictions. This step is very crucial and challenging, and several computer algorithms can perform this step. The second step involves aligning the structure with the target sequence and the process termed threading. The last step comprises structural alignment and refinement.

Free folding modelling method: This method is for predicting the protein sequences when there is no similar structure available in the protein database, and as a result, the threading method is not possible. It involves building the protein structure from scratch. In other words, it is a de novo method. All the current computer-based experimental approaches incorporate the free modelling work pioneered by Bewsie and Eisenberg[116], who designed many new protein structures by cutting and assembling the small fragments from the existing protein structures in the database. Several commercial structural prediction programs are available, and though this approach is more accurate than prediction based on sequence alone, it is highly complex, the number of 3 D structures available in the protein database are minimal, and the computational cost is high.

3. Combinational prediction methods: Because prediction using a single propensity scale method is not reliable for epitope location, this method combines sequence based and structure-based methods. The correlation between propensity scale based profiles and the location of B cell epitopes for 50 proteins was investigated using exhaustive assessments of 484 amino acid propensity scales[117]. According to the study, even the finest combination of propensity scale approaches was just modestly better. The poor performance of the amino acid propensity scale predictions led to the development of machine learning based algorithms, which incorporate multiparametric algorithms, utilising the advantages and information from both the sequence and 3D structure models such as the Hidden Markov model[118].

Despite the development of a number of methods for predicting B cell epitopes described above, the number of approaches for predicting discontinuous epitopes is still limited[119].

C. Bioinformatic tools for epitope prediction

There is broad recognition that computer based analysis currently available is reliable and can be used to find peptide epitopes. It is very diverse, and each algorithm has its own set of criteria to define antigenicity. Several limitations exist in the available algorithms. For example, some of the software will only accept a single protein sequence of limited size, and in these cases whole proteome analysis cannot be performed. Most of the software is designed based on any one or two protein characters, and thus testing through multiple servers is required to test a single protein. As seen in the literature, no single method is perfect for B cell epitope predictions[120].

As per the literature review, the following bioinformatic tools are frequently used for B cell epitope prediction.

i) I-TASSER

I-TASSER is a protein structure prediction software which has fully automated inbuilt algorithms that allow the user to generate high-quality 3D structures and to predict the biological functions of the protein from the submitted protein sequence. The prediction process involves three steps. In the first step, the server matches its threads with the previously tested proteins and produces the 3D image of all threading patterns for the submitted sequence. In step 2, the server matches the functionality with enzyme classification, gene ontology and ligand binding site libraries for the sequence. In the third step, a confidence score for the prediction is provided based on accuracy estimates with the existing libraries. A confidence "C" score obtained for each image reveals the maximum and minimum relatedness with the matched structures. The C-score runs between [-5, 2], with higher C-scores suggesting more confident models and lower C-scores indicating less confident model[121-124]. Finally, the generated 3D image helps the viewer understand the outline structure of the peptide, the possibility of secondary structure models and the available surface area (Figure 6).

The picture shown in Figure 5 is an example of the secondary structure prediction for the Oropouche virus (OROV) polyprotein. The image shows both the C-terminal and the N-terminal of the protein molecule.

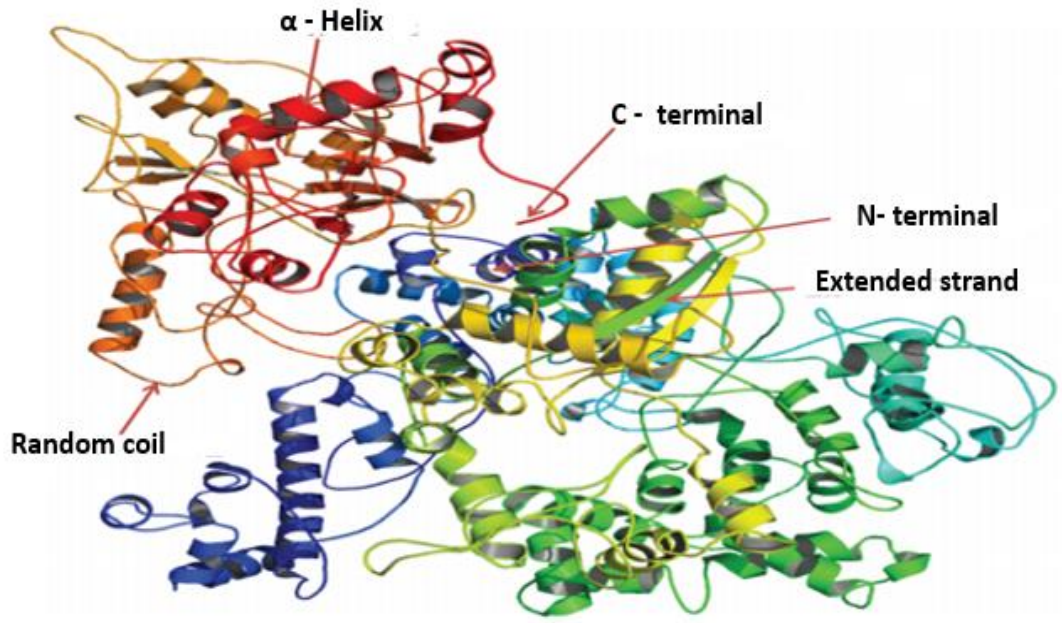

Figure 5: Example of Oropouche virus protein structure predicted by I-TASSER

The I-TASSER predicted secondary structures of the Oropouche virus protein shows mainly α-helix and random coils. The protein structure has few extended strands. More detailed information and the link to I-TASSER database is in Table 5.

ii) DNASTAR NovaFold and protean 3D

DNASTAR NovaFold protein structure prediction software is a modified version of I-TASSER, and predicts the 3D structure, protein function, ligand binding sites and enzyme activity for any given protein sequences. Protean 3D is a built-in application for examining macromolecular structure, motion, and function in DNASTAR. it can predict B cell epitopes and also construct molecular and solvent accessible surfaces to see the predicted epitopes (Figure 6)[125]. The protein threading method is better than the homology model because the homology model is based only on sequence identity. The threading techniques take into account the expected secondary structure, predicted internal interactions and approximate solvent accessibility, in addition to sequence similarity. By presenting the 3D molecular structure, annotated sequence, and analysing the used prediction algorithms, DNASTAR protean 3D tool enables for easy identification and analysis of secondary structure elements all at the same time. Another feature in this program is the colour coding of structural characteristics (binding sites, disulfide bonds, helices and β-sheets). Models from NovaFold can be linked to the protean 3D application, and protean 3D can do the structure and function analysis of the predicted structure. Protean 3D also has a B cell epitope prediction program and a bioinformatics approach, using experimental antibody recognition data. As

mentioned in the earlier section, the protean 3D program also has in-built propensity scale measurements for antigenicity, hydrophilicity and surface accessibility and other parameters to check the predicted epitope. An example of epitope prediction using DNASTAR protean 3D software is shown in Figure 6.

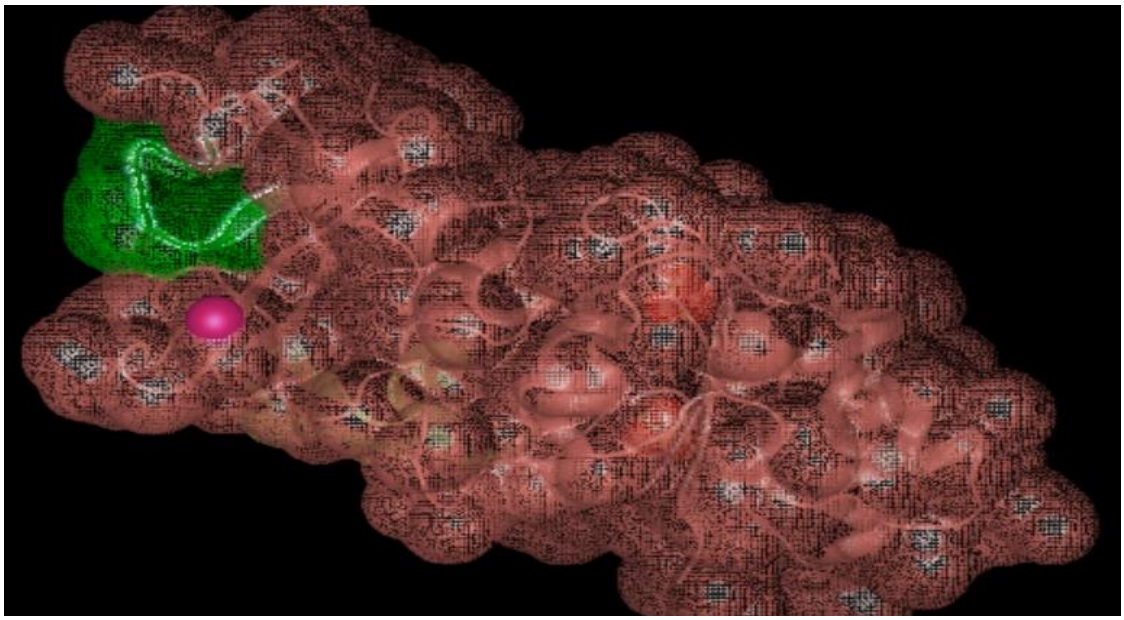

Figure 6: Structural view of a predicted B cell epitope using DNASTAR Protean 3D

The image shows the structural view of the protein, and the epitope is on the solvent accessible surface, very well exposed on the surface of the protein, a better candidate for B cell epitope. Information regarding the DNASTAR database is in Table 5.

iii) ***BepiPred-2.0***

BepiPred-2.0 is a sequence based, sequential B cell linear epitope prediction method. The BepiPred-2.0 server is constructed with a more significant number of structurally defined epitopes and non-epitopes. Each input sequence predicts the positions of the linear B cell epitopes located at the residue, annotated with an "E", predicted as being part of a linear B cell epitope. The default epitope threshold is 0.5. BepiPred-2.0 has an advanced output format which will predict the secondary structures and the surface accessibility. The secondary structures are colour coded.

iv) ***ABCpred***

This prediction method uses artificial neural network trained data sets, both positive and negative (B cell epitope and non-B cell epitope). Negative data sets used in the ABCpred method are random peptides. The ABCpred method predicts epitopes with 66% accuracy, 67% sensitivity and 65% specificity[126].

*v) **LBtope***

Unlike ABCpred, this prediction method uses a huge dataset of B cell epitopes and non B cell epitopes from the IEDB database (12,063 epitopes and 20,589 non epitopes). LBtope gives a percentage score for each amino acid, from 0% to 100%, with the default set at 60%. Any area or residue above 60 % qualifies as an epitope[127].

*vi) **BCPred***

This prediction method uses a subsequence kernel method and both B cell and non B cell epitopes from the in-built database and has similar built-in physiochemical properties to ABCpred. BCPred has a prediction accuracy of 75.8%[128].

*vii) **Netsurf predictor scale***

The netsurf predictor scale uses a neural network program. The prediction is made through a two-step process. The first step predicts the individual amino acid surface accessibility, indicating "B" for buried and "E" for exposed amino acid residues. The predicted values are then assigned a Z score. The greater the Z score, the more reliable the results will be. However, this score is more reliable for predicting buried residues than for predicting exposed ones[129].

*viii) **Propensity scale methods***

Propensity scale techniques assign each amino acid a propensity value which compares the likelihood of the amino acid being part of the B cell epitope to the background. To locate the linear B cell epitopes in the query protein sequence, they rely on observed connections between specific physicochemical features of amino acids and antigenic determinants in protein sequences. Several bioinformatics tools such as Parker's hydrophilicity[113], Emini surface accessibility[115], Karplus and Schultz flexibility[114] and Kolaskar and Tongaonkar antigenicity[130] are available to check features including, hydrophilicity, surface accessibility, flexibility and antigenicity are used to develop these approaches.[128].

1.5. Kode Technology

Kode Technology is a surface engineering technology that can modify any biological or non-biological surface with a small molecule[131-135]. Kode constructs or function spacer lipids (FSL) are water dispersible, self-assembling molecules that are used to modify the surface. Modification is achieved by a simple and easy process that involves keeping the amphipathic FSL constructs in contact with the desired surface for a short duration. Kode Technology can modify and emulate many surfaces of interest. In general, modified biological and non-biological surfaces have applications in diagnostics, oncology, drug targeting and release, biosensors, transplantation and much more. To establish quality control systems and diagnostic panels, FSL constructs can be

used to modify embryos, spermatozoa, zebrafish, epithelial cells, endometrial cells, red blood cells and virions without affecting the cell or its functionality in the way genetic and chemical modification of cells does. Non biological surfaces like metals, glass, plastics and fibers have been successfully modified with FSL constructs[136-140].

The name Kode Technology was coined to differentiate the surface modification of this FSL construct from other surface modifications. Any living (or dead) cell modified with FSL constructs is called a kodecyte. The process of applying Kode constructs to a surface is called "koding", while any biological or non-biological surface modified with FSL constructs is generically referred to as "koded"[131].

FSL constructs are made of three components, each of which play unique roles in surface modification and functionalisation. FSL construct components (function-spacer-lipid) are first individually synthesised and then assembled together.

Functional groups

The functional head group (F) of the FSL construct can be almost any hydrophilic bioactive or functional small molecules like carbohydrates, peptides, and fluorophores. FSL constructs can impart any functional head feature to the surface.

Spacers

The FSL spacer (S) has the ability to disperse in water while also spontaneously incorporating into a membrane. The spacer is unique and is the core component of this technology. The spacers exist in several variations: short, long, branched, clustered and functionalised. The structure of a particular spacer reflects the needs of the specific application, including the distance required from the membrane and the solubility demands[141]. The short adipate linker and carboxymethylglycine (CMG) are the spacers used in the FSL construct.

Lipids

The lipid tail (L) is used for anchoring the construct to a cell membrane or non-biological surface. Different membrane lipids will have different physicochemical characteristics and thus can affect the biological function of FSL. Dioleoyl phosphatidyl (DOPE), sterol, and ceramide are examples of lipids used for FSL constructs. The most common lipid chosen for the FSL construct, which is also most suitable for biological assays, is DOPE.

The hydrophobic lipid tail and the hydrophilic spacer work together to make the amphipathic Kode construct self-assemble on the surface. Multiple forces like van der Waals, electrostatic,

hydration and steric forces are involved in the attachment and retention of the FSL construct onto the surface.

The three distinct parts of the FSL constructs can be compared to the parts of a Lego toy. A schematic representation of the Kode FSL constructs compared with the Lego toy structure is shown in Figure 7[138].

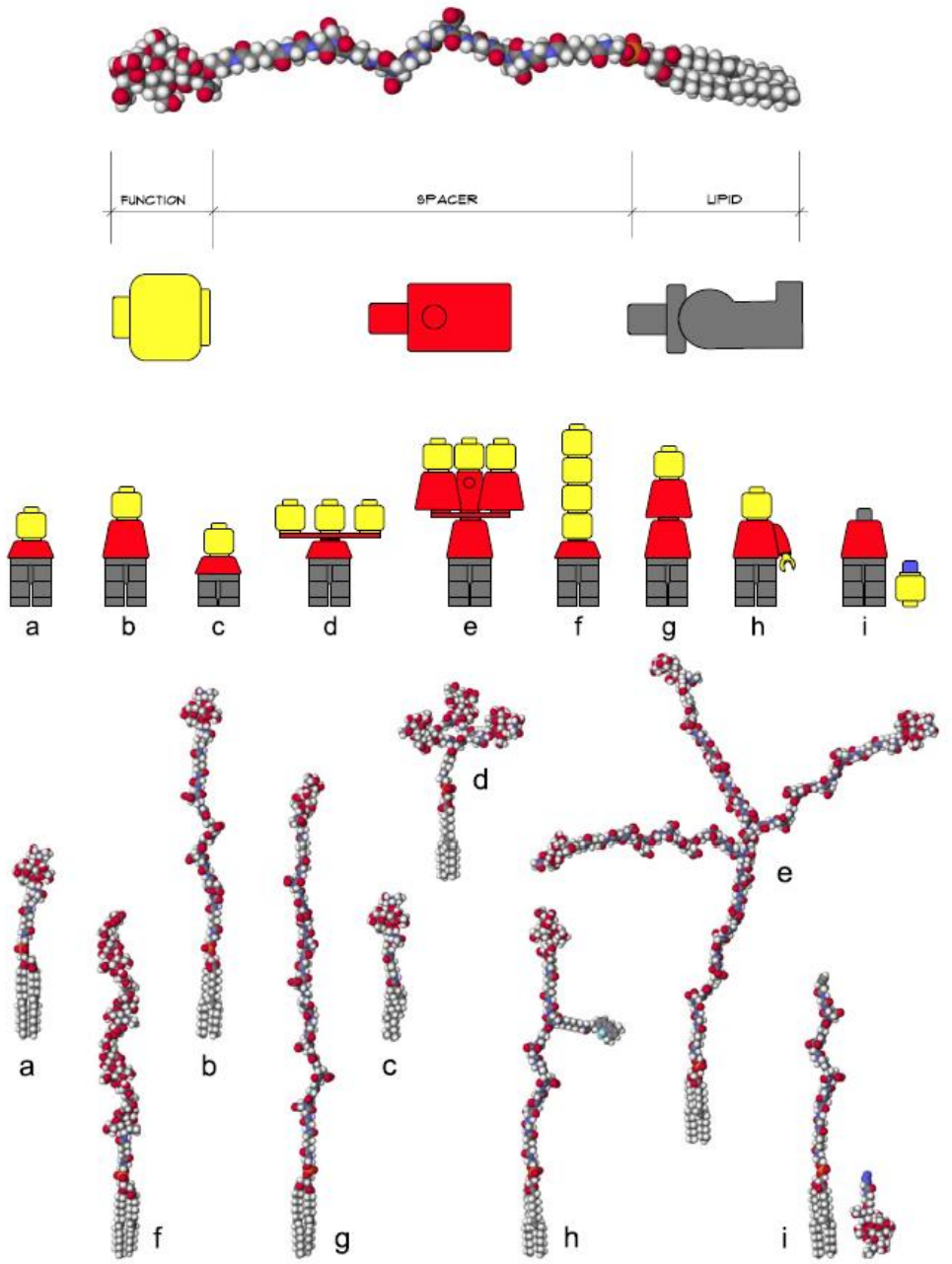

Figure 7: Schematic diagram of different Kode function-spacer-lipid (FSL) constructs

The upper image shows a generic Kode construct based on a carboxymethyl glycine spacer linked to a DOPE lipid tail. The 'building block toy figure' representations beneath show a yellow head, representing a single type of functional head, a red body, representing a spacer, and grey legs, representing a lipid tail. The nine structures shown at the bottom of the figure are space-filling molecular models of the building block toy figures, with each having the same tetra saccharide blood group A functional head, except model f, which has an (8-mer) hyaluronic acid functional head. Variation representations shown are: (a) short 1 nm adipate spacer, (b) CMG 7 nm spacer, (c) sterol lipid instead of DOPE, (d) clustered head, (e) trimeric CMG spacer, (f) linear repeating functional heads, (g) double CMG spacer and (h) functionalized CMG spacer where the spacer can undertake a secondary function. In example (h), the fluorophore BODIPY is attached, secondarily attached functional head, which in this case uses click coupling chemistry. Figure is copyright of Kode Biotech and reproduced with permission.

The functional group 'F' is comparable to the toy head that gives the functionality. The spacer 'S' is similar to the body that separates the head from the legs and provides some flexibility. The lipid "L" is comparable to the legs that help to firm the membrane attachment or self-assembly on

surfaces. There are several configurations available for functional heads, spacers and lipids. According to the needs associated with specific applications, the functional head, spacer and lipid can be interchanged, optimised and tailored.

Earlier in the laboratory diagnostic methods section, it was mentioned that IHA assays are suitable for leptospirosis diagnostics, and that the antigen coated red cells, covalently attached, are used in IHA[142-144]. In Kode technology, instead of covalently attaching the peptides to the red cells, it can non-covalently insert lipid linked peptide antigens on to the surface of the red cells. Moreover, the prominent feature of Kode is the ability to control and optimise the amount of peptide antigen coated onto the cell membrane. By varying the concentration of the FSLs during kodecyte preparation, Kode Technology can control the amount of antigen or the functional moiety on the surface[145,146]. The spacer, which is the core component of the FSL, is designed to be inert, devoid of any detectable affinity towards serum and cellular components. It is suitable for use with undiluted serum. Amino acid selection, peptide size and the nature and structure of the FSL spacer determine the kodecyte specificity and sensitivity.

Peptides are compatible biomolecules for FSL technology, as seen for the blood group Miltenberger antigen positive red cells preparation[147]. The use of cytomegalovirus FSL peptide constructs coated red cells for the Fc function assessment of immunoglobulin also proved to be a convenient method with good specificity and precision compared with the European pharmacopeia test, and it remains in current use[148].

Identifying the epitope region on an antigen which is diagnostic of disease biomarkers, and analysing the selected peptide for FSL construction compatibility, is critical for building FSL constructs. Some of the peptide identification issues that arise will be related to peptides, while others will be specific to FSL chemistry. FSL constructs are simple and straightforward to build once optimal peptide sequences have been identified and FSL compatibility has been established.[141].

The use of synthetic peptides as biosensors is an excellent alternative to minimize the non-specific reaction seen in traditional immunoassays and to remove the reproducibility and variability sometimes observed in these assays. The advantageous property of synthetic peptides is the possibility of modifying their chemical structure. Hence, the use of synthetic peptides in designing diagnostics has attracted interest.

Kode Technology is a potent tool which can optimise biomarker concentration with controlled attachment to the cell surface. The nature and structure of the spacer potentially enhances the presentation of biomarkers. The sensitivity and specificity of peptide biomarkers for infectious

disease diagnostics can potentially be achieved by the good selection of peptide epitopes combined with FSL compatibility.

One of the aims of this research project was to develop an informative algorithm for peptide diagnostic assays using FSL constructs and design/evaluate diagnostic assays.

1.6. Aims and Objectives

This research initially selected two different complex pathogens to evaluate Kode Technology as a peptide-based diagnostic. *T. pallidum* was chosen because it was considered relatively easy to develop as a potential diagnostic and samples were readily available. It was also useful to develop the initial algorithms using the diverse range of peptide antigen prediction tools for the peptide epitope selection, as significant published serological data existed.

The learnings of the *T. pallidum* assay were then used to guide the development of the more difficult *Leptospira* assay. This assay was expected to be extremely challenging because no current assay for the diagnosis of leptospirosis is satisfactory. It was also more challenging to validate, as there was minimal access to the samples, and those available would not be validated.

Additionally, with the unprecedented appearance of the COVID-19 pandemic, this research was extended to also create a potential antibody diagnostic assay.

The aims and objectives of this research project were:

- To develop a generic algorithm for diagnostic assays using FSL peptide constructs.
- To build and validate a sensitive and specific syphilis antibody diagnostic assay using FSL constructs.
- To build and validate a sensitive and specific leptospirosis antibody diagnostic assay using FSL constructs.
- To build and validate a sensitive and specific SARS CoV-2 diagnostic antibody assay using FSL constructs for external sensitivity and specificity validation.

Chapter 2. Method and algorithm - FSL peptide prediction

2.1. Peptide antigen design strategy

Theoretical prediction of epitopes is challenging, and considerable progress has been made in the identification of synthetic peptides which represent epitopes. The following are the significant issues and factors to be considered for peptide antigen design and synthesis, which could possibly affect the FSL peptide biology.

2.1.1. Protein structure

The peptide epitope needs to have a certain degree of resemblance to the native protein to be identified by the specific antibody[149]. The antibody will bind to certain structural patterns as the antigenic nature of the protein is structurally related. Conformation is structure/shape. A protein molecule is elaborately folded into three levels of organisation, secondary, tertiary and quaternary structures into unique conformation. The most common secondary structure motif of a protein is α-helix followed by β-sheet; it and usually contains a right hand helix confirmation in which every backbone N-H group hydrogen bonds to the backbone C=O group of the amino acid on the 4th amino acid away from it, resulting in 3.6 residues per turn[150]. In general, peptide antigens have α-helix conformation, which is the most common structure found in proteins. Conformation is necessary for a protein to be antigenic or immunogenic. However, in water, α-helices do not necessarily gain their native conformation and may adopt an extended helix, a favourable characteristic for peptides in water[151-153]. Peptides can adopt a stabilized helix structure if the peptide contains alanine and leucine or two more side chains with carbon. The presence of isoleucine and valine may destabilize the peptide chains with carbon[154]. The presence of charged residues in a peptide sequence relates to helix formation and stability. The composition of the peptide is vital while choosing peptides with possible α-helix conformation. In particular, the presence of alanine can form a stable helix with good water solubility. Based on the amino acid residues in a protein sequence, the secondary structure can be predicted to a certain extent.

The β-sheet is the next prominent secondary structure in a protein. It usually forms complementary hydrogen bonds to a parallel polypeptide chain. The adjacent chains can be aligned in the same direction or in opposite directions, which means hydrogen bonding can be in any pattern. That is not ideal for peptide epitope design. Also, β-folding can occur as barrels, helices, propellers, single sheets or multimers, and in any pattern, they are insoluble and could be problematic in immunoassays[110,155]. Turns have fewer amino acid residues, they do not have a defined structure and they may or may not have hydrogen bonding. The loops are longer, non-

repeating, and omega (Ω) shaped structures encompassing interloop disulfide bonds. The polar and charged residues usually form these surface-exposed structures. Hydrophilic areas tend to reside on the surface of proteins in most natural environments, thus, ideal antigenic epitopes are hydrophilic, surface oriented, and flexible. B cells identify hydrophilic (surface exposed) sections of proteins, whereas hydrophobic areas are found in the protein's core. Turns and loops are usually surface exposed structures, and the immune system makes antibodies to surface exposed structures in the first instance. Hence, they are promising structures to be used for epitope selection. However, epitopes may be found inside a protein structure and become exposed during degradation of the protein and be immunogenic. Also, turns will have high conformational flexibility[156]. Turns, loops and helices can be made water-soluble, usually, are surface exposed and probably contain B cell epitopes.

2.1.2. Amino acid composition

The composition of amino acid dictates the functionality of the peptide. The amino acid composition of a peptide will have an impact on the peptide synthesis, assembly, purification and solubility. The sequence of the peptide antigen should have both hydrophobic and hydrophilic residues. In particular, amino acid residues like cysteine, methionine and tryptophan are susceptible to oxidation and side reactions that will affect the purity of the peptide and may affect the conjugation chemistries and antibody binding. If possible, it is better to choose sequences that minimize these residues. The sequences of the designed peptide antigens should closely match the sequences found on the surface of the native protein.

2.1.3. Cysteines

There are several reasons to avoid cysteines in the peptide sequence. They can undergo rapid oxidation and multiple cysteines will form disulfide linkages. Furthermore, the probability of the cysteine not being recognized as part of the native protein is greater, the reason being that, generally, free cysteines are not seen in vivo. The alternative is to replace the cysteine with the conservative residue if possible. In the FSL chemistry, the cysteine residue is used to bind the peptide to the spacer.

2.1.4. Hydrophilicity

Protein sequences that are likely to be hydrophilic, surface oriented and flexible are used for the design and production of peptide antigens. When a protein folds, the hydrophobic residues are buried inside and the hydrophilic residues are exposed on the outer surface of the protein. Antibodies prefer to attach to epitopes on the protein's outer surface, which is important for identifying epitopes on the hydrophilic surface. It is ideal to have fewer hydrophobic residues. If

the peptide sequence has more than 50% hydrophobic residues, it will be challenging to synthesize and will have solubility issues.

2.1.5. N-terminal glutamate

N-terminal glutamine (Gln) is unstable, and it can undergo notoriously fast cyclization to form pyroglutamic acid. Hence peptide sequence with N-terminal glutamine (Gln) should be avoided[157]. This should be excluded or fixed when the expected epitope starts with N-terminal glutamine since this will affect immunogenicity. The formation of pyroglutamic acid can be prevented by acetylating the N-terminal glutamine or choosing a peptide with an additional residue preceding glutamine[158].

2.1.6. Glycosylation

Almost all proteins are glycosylated. Proteins can be glycosylated at specific amino acid chains. These modifications are termed as N and O glycosylation sites, and both are equally considered during the peptide selection process. Glycosylation impacts folding, solubility, antigenicity, conformation, interactions and the half-life of a peptide. It is a common post translational modification. The consensus sequence for N-glycosylation is AsnXxxSer/Thr/Cys, where X can be any amino acid except proline[159]. There is no standard O-glycosylation sequence, but such sequences usually occur with a high content of serine, threonine and proline residues [160]. Glycosylation is often associated with signal peptides, and known glycosylation motifs, when present, should be avoided, in particular when there is a signal peptide in the native protein[161]. Although glycosylated peptides are relevant antigenic components, their synthesis is too complex for routine use. As a result, glycosylated peptides are avoided.

2.1.7. Oxidation

Tryptophan and methionine residues are prone to oxidation. Amino acid oxidation is a post translational modification that may adversely affect peptide conformation, activity and stability[162]. These residues should be avoided in the peptide sequence.

2.1.8. Asparagine deamidation

The amino acid asparagine (Asn) is known for deamidation. The amide group of the side chain is converted to another functional group like aspartic acid or is aspartic acid, this results in altered antigenicity. The presence of a non-hydrophobic residue Asn-Xaa, followed by an asparagine (Asn) in a peptide sequence can lead to spontaneous deamidation. Deamidation is accomplished by forming a five-membered aspartimide intermediate, which results in a combination of related

α- and β-aspartyl peptides[163]. This will cause peptide instability, and although conservative replacement will fix the issue, Asn-Gly should be excluded[164].

2.2. FSL peptide selection consideration

The peptide antigen selection process was limited to linear epitopes in this study for the following reasons:

1. Linear epitopes (antigens) consisting of a continuous stretch of amino acids are easy to manufacture.
2. Discontinuous epitopes, as the name suggests, are not a continuous stretch of amino acids and, in the absence of mimetics, they are avoided. It is difficult and challenging to create the exact synthetic peptide mimic for discontinuous epitopes.

2.2.1. FSL peptide design strategy

The theoretically predicted peptide sequences must be analysed for suitability and refined for FSL compatibility and chemistry. By ligating the functional head (peptide) to a spacer (carboxymethylated oligoglycine-CMG) linked to a lipid tail (DOPE), suitable peptides will be synthesized into FSL constructs and assessed for functionality.

In the FSL constructs, the functional head is a peptide antigen. There are a couple of significant issues to be considered for FSL compatibility. First, the predicted peptides should comply with all the significant factors stated in the peptide antigen design strategy and considerations about proteins, peptides and amino acid chemistry described in section 2.1. Second, the predicted peptides should comply with FSL synthesis chemistry.

2.2.2. Solubility

Solubility is a crucial factor in FSL chemistry, particularly for the functional groups of the FSL constructs. Peptide solubility depends on the number of charged residues in the peptide sequence and avoiding long stretches (more than five amino acids) of uncharged residues during the sequence design. Amino acid composition is essential for peptide solubility. Leucine (Leu), valine (Val), isoleucine (Ile), methionine (Met), phenylalanine (Phe), and tryptophan (Trp), are hydrophobic residues, and a high content of these residues in peptides will render the peptide with limited solubility in aqueous solution, or leave it completely insoluble. Manipulation of the peptide sequence can be done to improve the solubility issue; if manipulation does not help, then cosolvents can be used before ligation. This will improve the insolubility issues, although doing so involves modifying the pH. Ideally, the FSL constructs should be dispersed in saline without manipulating the solution with cosolvent or changing the pH. Dispersing the FSL in saline is critical

for downstream applications like modifying the cells without affecting their vitality and functionality.

2.2.3. Internal cysteine

In FSL ligation chemistry, the cysteine residue is thiol based and is used to bind the peptide to the spacer (S). Depending on the orientation required, a cysteine residue is inserted at either the amino or carboxyl ends of the final peptide sequence[148]. In theory, this restriction necessitates the absence of cysteine(s) in the chosen peptide sequence. The solutions are either to use the native cysteine side chain for spacer lipid ligation or to use the isosteric substitution of cysteine for α-aminobutyric acid[165].

2.2.4. Peptide length

In theory, the minimum number of amino acids necessary for antibody recognition is four to six[166]. However, using this minimum range is not recommended in practice because the additional amino acid residues also affect the presentation of the epitope and binding affinity[167,168]. The ideal peptide length for FSL construction appears to be in the range of 12 to 20 residues; this size will minimize synthesis issues, be reasonably soluble in an aqueous solution and potentially enhance polyclonal antibody recognition.

2.3. FSL peptide prediction algorithm

The objective is to make epitope predictions using the combination method as discussed in 1.3.1 (c) and then develop an initial algorithm to select peptides suitable for constructing into FSL constructs. The first step is to select suitable proteins with known antigenicity from the literature cited proteomic and microarray studies. The second step is to retrieve protein sequences from the Entrez database and then analyse the sequences using the BLAST program to determine the degree of cross reactivity with other microbial sequence relatedness. The sequence retrieval database and the BLAST program can be accessed from the National Centre for Biotechnology Information (NCBI home page http://www.ncbi.nlm.nih.gov). The next step involves the prediction and analysis of the protein structure using the selected protein sequence, followed by the identification of B cell epitopes using the programs described earlier. The peptide sequence, predicted multiple times as epitope, is further analysed for all the peptide design parameters and FSL peptide selection. The FSL selection of peptides is based on the initial FSL selection algorithm shown in Table 4, which is divided into two sections. Table 4a shows the methods used for epitope prediction and Table 4b shows the methods used for checking the biophysical properties of the epitope.

Table 4. FSL peptide selection algorithm criteria

4a: **Epitope prediction**

Peptide details	Method	Target	Units
Sequence retrieval	NCBI database††		
Similarity check		No microbial relatedness	Consensus
Protein structure	I-TASSER	High C score	C score [range -5, 2] ††
	DNASTAR	Structure analysis	Loops / turns helix/
Epitope prediction	ABCpred	>0.51	Auto threshold value: 0.51
	BCPred	>80	% Specificity
	DNASTAR	Purple bar	Antigenic region
	BepiPred-2.0	>0.50	Auto threshold value:0.5

4b: **Epitope biophysical properties**

Peptide details	Method	Target	Units
Internal cysteine	Sequence analysis	0	C
Terminal Glutamine	Sequence analysis	0	Q
Binding site		1 or more	
O-glycosylation		0	
N-glycosylation		0	
Net charge		0 to +1	Z
Isoelectric point		6 to 7	pl
Peptide length		15 to 20	Amino acids
Flexibility		>1	Auto threshold value:0-2
Parker hydrophilicity		>2.5	Auto threshold
Surface probability		>1	Threshold value:1
Secondary structure		Turns/loops	Turns/loops/helix
Antigenicity	Kolaskar/Tongaonkar	≥1	Threshold value:1
	Antigen profiler	>3.6	2.7-3.5: moderate antigen 3.6-5.0: excellent antigen

Tables 4a and 4b show the parameters used in this study for the peptide prediction and FSL peptide construct selection developed and refined during this research. Table 4a: †The NCBI suite of database resources are available from the NCBI home page - link mentioned in Table 5. †† C score [range -5, 2], where a C score of higher value signifies a model with a high confidence and vice-versa.

The antigenicity, hydrophilicity, presence of glycosylated residues and other parameters of the predicted peptide is checked (Table 4b) for its suitability to be built as FSL construct.

The functional head (peptide) is subsequently ligated to a spacer (carboxymethylglycine-CMG) connected to a lipid (DOPE) tail, resulting in FSL constructs.

Table 5. Resources and databases used for FSL peptide selection

Prediction tool	Web address	Description	Reference
NCBI†	www.ncbi.nlm.nih.gov	†	[169]
MultAlin	http://multalin.toulouse.inra.fr/multalin/	Alignment of the GenBank sequences for consensus	[170]
BCPred	ailab-projects1.ist.psu.edu:8080/bcpred/predict	Linear B cell prediction	[128]
BepiPred-2.0	www.cbs.dtu.dk/services/BepiPred	B cell linear epitope prediction	[171]
ABCPred	www.imtech.res.in/raghava/abcpred/	Artificial neural network linear B cell epitope prediction	[126]
LBtope	webs.iiitd.edu.in/raghava/lbtope/protein.php	B cell linear epitope prediction	[127]
I-TASSER	zhanglab.ccmb.med.umich.edu/I-TASSER	3-Dimensional structure prediction	[124]
SOPMA	https://npsa-prabi.ibcp.fr/cgi-bin/secpred_sopma.pl	Structure prediction	[172]
NetNGlyc	www.cbs.dtu.dk/services/NetNGlyc/	NetNGlyc prediction	[173]
NetOGlyc	www.cbs.dtu.dk/services/NetOGlyc/	NetOGlyc prediction	[174]
Hydrophilicity	tools.iedb.org/bcell/result/	Hydrophilicity prediction	[113]
Antigenicity	tools.iedb.org/bcell/result/	Epitope antigenicity prediction	[130]
Surface Accessibility	tools.iedb.org/bcell/result/	Epitope surface accessibility prediction	[115]
DNASTAR	www.dnastar.com/software/structural-biology/	Protein structure analysis and B cell epitope prediction	[125]

† NCBI database provides data retrieval system and a number of other computational resources for the analysis of GenBank data, which can be accessed from the NCBI home page.

Table 5 shows the resources and links for the databases used in this study for the peptide predictions. Although a vast amount of data is available in the public domain, for this study relevant databases are selected based on availability and accessibility.

Chapter 3. Methods and Results: Syphilis diagnostic using Kode FSL peptide

3.1. Literature review of *Treponema pallidum* proteome

T. pallidum proteome studies date back to 1975[175]. A literature review of the *T. pallidum* proteome shows a number of *T. pallidum* proteins that are identified in immunoproteomic studies. These studies are reviewed in this section as part of the protein selection process for FSL peptide design.

3.1.1. Outer membrane proteins (OMPs)

There are several reasons why the quest for the *T. pallidum* OMPs continues even though the genome sequence of Nichols strain of *T. pallidum* was published twenty years ago[176]. Primarily, the OMPs are very fragile (they do not even withstand centrifugation) and any experimental manipulation to localize the proteins usually completely disrupts them. Structure prediction methods and bioinformatic tools are used to circumvent the technical challenges associated with *T. pallidum. T. pallidum* research shows that the spirochete has about 1% of the OMPs, present in the gram-negative bacteria *E. coli*. The hallmark of many chronic multistage diseases causing pathogens like *Borrelia burgdorferi* and *Trypanosoma brucei*, and *T. pallidum* is the antigenic variation of outer membrane antigens. In the case of *T. pallidum*, there is a low abundance of OMPs. There are only 17 proteins identified as OMPs in *T. pallidum*[177], out of which 12 belong to the TprK family of proteins. TprK proteins undergo antigenic variations in seven variable regions. Substantial differences in TprK sequences are seen between and within individual strains. As a result, a single patient may have several clones of *T. pallidum* expressing different variants[178]. These variants are selected by immune pressure. The TprK family of proteins is thought to play a key role in *T. pallidum* pathogenesis and immune evasion. TprK proteins are not identified as seroreactive antigens in immunoproteomic studies, probably for the above reasons, and they are of limited value as diagnostic antigens (Table 6).

The non TprK surface exposed proteins are known as treponema rare outer membrane proteins (TROMP). TROMP-2, the FlaA homolog, was reactive only in primary syphilis, and this protein was not reactive with human sera [179,180]. The Tp92 is another outer membrane protein expressed in low levels with β-barrel structure and variations in the antigen's β-barrel portion. It is not seroreactive in the latent stage of syphilis[181]. Tp0453 is another OMP that has shown moderate reactivity in primary syphilis cases[182], and Tp0453 is present on the inner surface of the outer membrane. Recombinant proteins with Tp92, Tp0453 and chimeric Tp0453-Tp92 as diagnostic antigens are reported[183] with a sensitivity of 86%, 98% and 98%, respectively. That study reported

three false negative results in primary syphilis cases, which were then identified as positive by the confirmatory treponemal assay the TPPA and a low level of RPR positive result.

3.1.2. Flagellar proteins

Flagellar proteins, represented by FlaB1, FlaB2 and FlaB3 proteins (Table 6), make up a major portion of the *T. pallidum* proteome and are found in the periplasmic region.

Extensive homology between the FlaB of *T. pallidum* and flagellin of other bacteria and broad cross reactivity with different sera have been documented[184,185]. These features restricted the diagnostic significance of flagellar protein. Recent work with flagellar protein has shown that the middle portion of the protein has better specificity and that it is more useful in primary syphilis[186]. Overall, it is less specific than immunodominant lipoprotein antigens and hence are not used in any current routine commercial syphilis diagnostic tests.

3.1.3. Adhesion proteins

Adhesion proteins were not seroreactive in proteome studies, and they showed only 9% positivity in syphilis positive sera and only in primary syphilis[182]. Adhesion protein Tp0136 is exposed to the outer membrane of *T. pallidum*. This protein reacted with syphilis human sera than with infected rabbit sera. Tp0136 showed 85.5 % positivity in primary syphilis patients[187].

Tp0751 is a laminin binding adhesion protein (Table 6) molecule of *T. pallidum*[188]. This protein attaches to the extracellular matrix component laminin, and plays a crucial role in the *T. pallidum* dissemination in the host. Tp0751 had a very low seroreactivity[182]; because of the immunoprotective property of the protein, it was unsuccessfully trialled as a vaccine[189].

3.1.4. Inner membrane proteins

Proteome analysis of *T. pallidum* predicted that a large number of lipoproteins (Tp15, Tp47, Tp17, TmpA, Tp32) are located in the inner membrane[190]. Proteins located in the inner membrane are highly immunogenic, although they are not exposed to the outer surface. Some of the lipoproteins, Tp15, Tp17, Tp47 and TmpA, typically reacted with all stages of syphilis infected sera (Table 6). The function and role of Tp15 are unknown. Protein ligand interaction, treponemal membrane architecture maintenance, and the stimulation of intercellular adhesion molecules expression are all possible functions of this protein. Tp47 is carboxypeptidase, and it is involved in host pathogen interaction. Human fibronectin was in high homology with the sequence of the Tp47 antigen, leading to false positive reactions[191]. Tp32, an L-methionine-binding lipoprotein located in the inner membrane of *T. pallidum*, has been found to have a sensitivity of 91% and a

specificity of 94% for IgM detection by some researchers[192] despite not being discovered in any immunoproteomic investigations.

Several novel antigens are reported by researchers to be promising target genes for syphilis diagnostics. Still, ongoing studies reveal the low efficacy of some of these antigens compared to the recombinant lipoproteins currently in use. Nearly 45 membrane proteins have been postulated so far, which accounts for 4.3% of the total genome of *T. pallidum*. However, treponemal lipoprotein functions have remained mostly obscure. The spirochete's intricate parasitic strategy is probably due to *the* membrane protein's functions of *T. pallidum*

The humoral immune response develops at different rates during *T. pallidum* infection for each *T. pallidum* protein. Expression of different *T. pallidum* proteins are seen at each stage of the disease. Protein array studies[180] showed that out of the 908 proteins tested, 34 were considered antigenic. Only 14 proteins showed positivity with all three stages of infection, and not all of them are expressed equally in all stages of the disease. Seroreactivity studies of syphilis proteins have determined this protein expression feature in *T. pallidum* (Table 6).

The novel periplasmic proteins Gpd and TpF1 are more useful in primary syphilis detection, but they are less sensitive. The flagellar proteins Flab1, Flab2, and Flab3, though detected in all stages of syphilis infection, have low specificity due to cross reactivity. Although *T. pallidum* proteomic research revealed novel antigens, ongoing studies showed that most of these novel recombinant antigens are less effective than the immunodominant inner membrane lipoproteins.

Initially, the inner membrane proteins Tp15, Tp17 and Tp47 were developed individually into syphilis diagnostics. Though Tp47 is an immunogenic and highly expressed protein, false positivity is seen in Tp47 based assays due to homology with human fibronectin[193,194]. To improve the sensitivity and specificity, a combination of these proteins is currently used in the majority of the commercial assays. In recent years, a number of commercial tests (for example, recombinant Western blot assay) have added TmpA and Tp0453 in addition to the other three proteins (Tp15, Tp17 and Tp47).

A summary of the *T. pallidum* proteins listed in the literature is shown in Table 6. The outer membrane protein is not seronegative in all three stages of syphilis infection and it is not used in commercial assays. The flagellar proteins Flab 1, Flab 2, and Flab 3 are non-reactive in certain immunoproteomic studies[179] and are known to cross react with similar flagellar proteins[184,185]. The adhesive proteins, as per the literature review, are not found in all stages of syphilis infection. The most of the inner membrane proteins are found to be reactive in all

Table 6. *T. pallidum* proteins identified in immunoproteomic studies

Protein location	Identified in immunoproteomic studies	Syphilis stages (P, S, L, N) Sero-reactivity	Commercial assay
Outer membrane	TprK	N	No
	TROMP-2	P	No
	Tp92	P/S	No
	Tp0453	P	No
Flagellar †	Flab1	P/S/L††	No
	Flab2	P/S/L††	No
	Flab3	P/S/L††	No
Adhesive	Tp0155	N	No
	Tp0483	N	No
	Tp0136	P	No
	Tp0750	P/L	No
	Tp0751	N	No
Inner membrane ‡	Tp15	P/S/L	Yes
	Tp17	P/S/L	Yes
	Tp47	P/S/L	Yes
	TmpA	P/S/L	Yes
	Tp32	N	No

P: Primary syphilis, S: Secondary syphilis, L: Latent syphilis, N: Non-reactive, †: Cross reactivity seen with similar flagellar proteins, ††: Flagellar protein is non-reactive in certain immunoproteomic studies. ‡: Tp15, Tp17, Tp47 not sensitive on their own so all three proteins are used together as a recombinant antigen in the current commercial assays.

three stages of syphilis. It is also found that the inner membrane proteins are not sensitive on their own, so the inner membrane proteins are used together in combination as recombinant antigens in the commercial assays. False positive results are reported with the inner membrane protein Tp47 antigen[193,194].

3.2. Rationale for choosing TmpA protein for FSL peptide design

Lipoproteins Tp47, Tp17, Tp15 and TmpA are the primary mediators of the immune response to *T. pallidum.* Reliability of the *T. pallidum* serological tests is currently based mainly on the inner membrane proteins Tp17, Tp15, Tp47 and TmpA.

Several independent studies have shown TmpA is the most predominantly recognised and identified antigen (95%) followed by Tp47, Tp17 and Tp15 (92.5%, 89.5% and 67.5%, respectively)[193-195]. These and other studies offer other relevant key findings to support the use of TmpA in this study:

- A total of 34 reactive antigens have been identified so far in *T. pallidum* immunoproteomic studies. Only very few proteins matched with seroreactivity and clinical illness; TmpA is one among them. It is a highly expressed protein during syphilis infection.
- The serological performance of recombinant TmpA, its relevance to detect syphilis in all stages of infections and its potential use in assessing the efficacy of antibiotic therapy, feature in published studies [192,193,195,196].
- The high cost of syphilis EIAs is due to multiple recombinant antigen formulation required for such tests; the more antigens involved in the test, the more costly the kit. Therefore, if an assay can perform well with a single synthetic peptide, it would be ideal for screening purposes. Hence, TmpA protein was selected for this study.

3.3. Methods and results: Analysis of TmpA protein sequence

The TmpA protein sequence used in this study to design the peptides was derived from the NCBI database, GenBank no: AGN75941.1.

3.3.1. Structure prediction using I-TASSER/NovaFold

The TmpA protein structure prediction was done using the I-TASSER/NovaFold software. Based on simulation and alignment of the threading pattern, the calculated confidence score (C score), of the predicted structure was -4.35. The C-score runs between [-5, 2], with higher C-scores

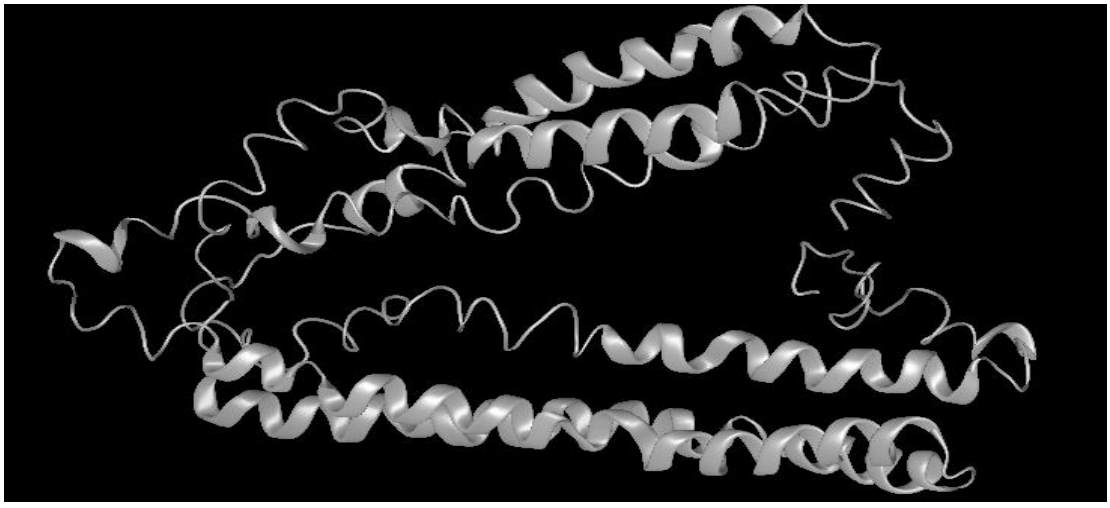

Figure 8: Predicted three-dimensional structure of the TmpA protein.

C-score of the model: -4.35. The model has no β sheets, it has mainly α helical structures, coils and turns, which fits in with the results from the SOPMA server.

suggesting more confident models and lower C-scores indicating less confidence. The score suggested that the predicted model is good. When tested in the I-TASSER program, TmpA protein showed many α-helical structures and coils (Figure 8). The residues 25, 29 and residues between

37 and 44, 161 and 179, 184 and 193, 199 and 220, and 272 and 274 are expected to be binding residues, according to the I-TASSER predictions.

3.3.2. Structure analysis using protein structure analysis program (ProSA)

The I-TASSER/NovaFold predicted structure was further analysed with the ProSA program. The ProSA program calculates the quality score (Z score) by comparing the query model (TmpA) with all the experimental structures available in the protein database. If this score is outside a range characteristic for native proteins, it indicates the structure probably contains errors. On the

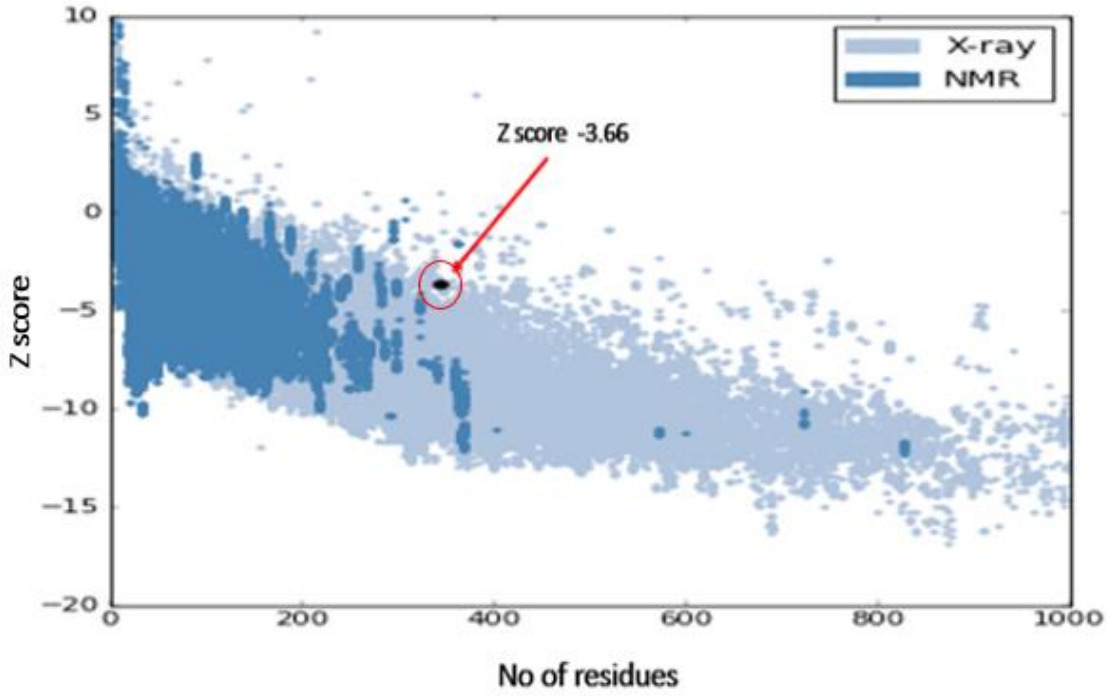

Figure 9: Z score plot for predicted model of TmpA protein

x axis, the number of residues is plotted, while the Z score is plotted on the y axis (Figure 9). The light blue dots in the score plot represent the scores of all the currently available protein structures recorded by X-ray crystallography, and the dark blue represent those captured by the NMR spectroscopy method. The user input structure score is circled in red. The Z score of the input structure should be within the range of scores obtained in native proteins of similar size. The Z score represents the model's overall structural prediction [197,198]. The calculated Z score for the TmpA protein is -3.66, and the predicted score is within the range of -10 and 5 of the native protein conformations, suggesting the structure does not contain any errors in general.

3.3.3. Structure analysis using self optimized structure prediction method (SOPMA)

The TmpA protein sequence was submitted to SOPMA software[172]. The secondary structure analysis of TmpA protein (Table 7) sequence, using the SOPMA server, revealed that the TmpA

Table 7. Secondary structures of TmpA protein (SOPMA)

Secondary structure	Percentage (~100 %)
α helix *	74
Extended strand	1.2
β turn	2.6
Random coil	22

*The secondary structure of the TmpA protein is dominated by α helix.

structure is dominated by α helix (74%), followed by the coil (22%). The TmpA protein has few β turns (2.6 %), and extended strands (1.2 %). Both the secondary structure predictions of SOPMA and I-TASSER/NovaFold are similar, finding that the TmpA protein structure has a greater percentage of α helix.

3.3.4. Prediction of hydrophilicity using the Parker hydrophilicity program

The TmpA protein sequence was analysed using the Parker hydrophilicity prediction method to get an idea of the hydrophilic regions present in the entire length of the protein.

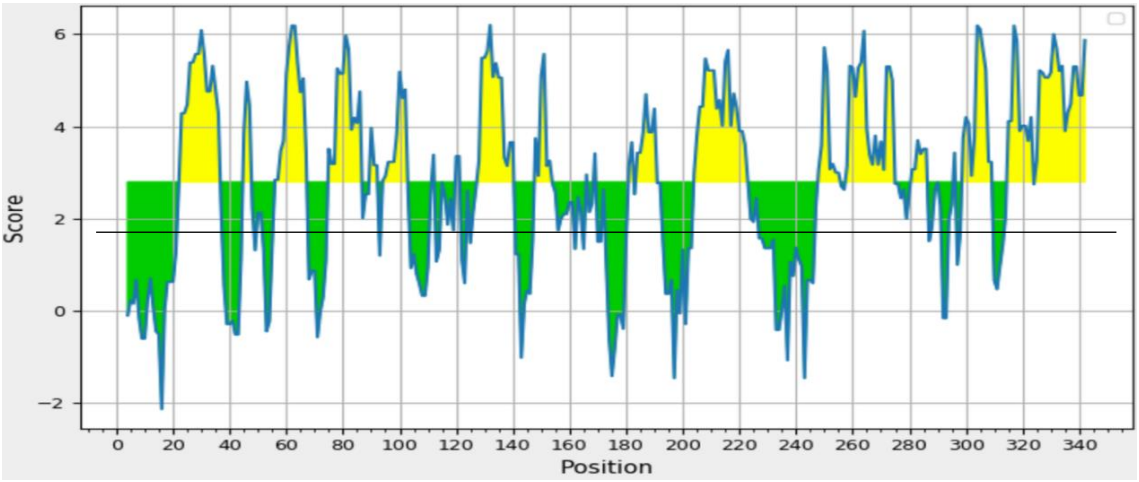

Figure 10: Parker hydrophilicity predictions for the TmpA protein

The graph, as can be seen above, has a line in the middle which is the threshold line. The yellow area above the threshold line in the graph represents residues with prediction values above the threshold, indicating the residues are hydrophilic. The green area below the line represents residues below threshold value. The image shows that TmpA protein has alternating stretches of hydrophilic residues.

In this method, the epitope region was analysed for a window of seven amino acid residues where the hydrophilic value was assigned to seven corresponding residues and the arithmetic mean of those were assigned to the fourth residue. The score is generated for every seven residues and

hence output of the scale completely depends on the N-terminal amino acid of the input sequence. TmpA protein, when tested with the Parker prediction algorithm (Figure 10), showed the presence of alternate stretches of hydrophilic and hydrophobic residues. The highest hydrophilic peaks on this scale, above the threshold value of 2.8 and reaching a maximum of 6.2, are seen between residues 20 and 40, 55 and 85, 125 and 140, 200 and 220 and 320 and 335, which means those residues are extremely hydrophilic.

3.4. Methods and Results: Identification of B cell epitopes

As discussed earlier, no single method is perfect for B cell epitope prediction, hence, the consensus method was selected for predicting B cell epitopes in this study.

3.4.1. Epitope prediction by ABCpred

The TmpA protein sequence with 345 amino acids was submitted to the ABCpred program. The epitope length was set to 16, 18 and 20 amino acids. The default threshold for the ABCpred program is 0.5. The scoring threshold was set to 0.8 to increase the accuracy and probability of prediction.

Table 8. Epitopes predicted by ABCpred method and ranking scores

Rank	Sequence†	Start position	End position	Score
1	RGVEDGGRSPKSSMNE	325	341	0.94
2	PSRVQSRGVEDGGRSP	319	334	0.93
3	KEPIEVEPLPNDRLNT	287	303	0.91
4	YQEAREAEEVNARAKA	211	226	0.88
5	EGFGRLLPDMKARAGA	171	287	0.88
6	DAKIEGDLKKAAGVAS	83	99	0.87
7	KYEILRNRVEVADLQS	103	119	0.87
8	PIPISDTSSPSRVQSR	310	326	0.86
9	GVASEAADKYEILRNR	95	111	0.85
10	AGAAKTDVGGLKVAVE	184	200	0.85
11	EEAEKKAAEQRALLVE	28	44	0.85
13	EKALQSAKTKQKASSDLARS	252	270	0.82
14	AAEQRALLVESAHADR	34	50	0.81

†: The stretch of amino acid in the underlined sequences (areas of interest) identified as epitopes by other software.

The software assigns a score from 0 to 1 to each epitope prediction. A score of one or closer to one indicates a higher probability of the epitope existing, and as the score goes down, the amino acid sequence is decreasingly likely to be an epitope. A total of 14 epitopes were predicted by this method (Table 8). All the 14 epitopes predicted had scores above the set threshold with ranking scores close to one, suggesting the probability of being an epitope.

3.4.2. Epitope prediction by BCPred

The TmpA protein sequence with 345 amino acids was submitted to the BCPred epitope prediction program. BCPred predicts antigens with varying residue lengths from the input sequence. The default setting of the BCPred is at 75% specificity; this was set at 80% to increase the accuracy and probability of prediction.

Table 9. Epitopes predicted by BCPred method and ranking scores

Rank	Sequence†	Start position	End position	Score
1	KEEAEKKAAEQRALL	27	41	1
2	SADKSAPLPENAQG	271	285	1
3	EPIEVEPLPNDRLN	288	302	1
4	VVRLKKTEAEKALQ	243	257	1
5	GGRSPKSSMNEEGASR	330	345	0.998
6	AMLGSCASGAKEEAEKKAAEQR	17	39	0.931
7	ESAPIPISDTSSPS	307	321	0.996
8	ADSQYQEAREAEEV	207	221	0.976
9	ALQSAKTKQKASSD	254	268	0.893
10	PIEVEPLPNDRLNT	289	303	0.889
11	RIGAQESGADTQHP	55	69	0.868
13	DKYEILRNRVEVAD	102	116	0.821

†: The stretch of amino acid in the underlined sequences (areas of interest) identified as epitopes by other software.

Similar to the settings in the ABCpred program, a score of one or close to one using the BCPred method indicates high probability of the epitope existing. Thirteen epitopes were predicted by this method for the submitted TmpA protein sequence (Table 9). As seen from the above table, all 13 epitopes predicted were above the set threshold and the ranking scores were closer to one, suggesting those amino acid sequences can be identified as epitopes. The BCPred method also predicted the sequences similar to the ABCpred method as epitopes.

3.4.3. Epitope prediction by BepiPred-2.0

The TmpA protein sequence with 345 amino acids was submitted to the BepiPred-2.0 program for the prediction of B cell linear epitopes. Table 10 shows the predicted potential B cell epitopes of varying lengths.

Table 10. Epitopes predicted by BepiPred-2.0 method

Peptide	Sequence†	Start position	End position
1	SGAKEEAEKKAAEQRALL	21	39
2	SAHADRRLMEA	44	54
3	IGAQESGADTQHPELFSQIQDVERQST DAKIEGDLKKAAGVASEAADKYEILRNR VEVADLQSKIQTHQLAQYDGDSANAAEES WKKALELYETDSAQCLQSTVEALE	55	164
4	KTDVGGLKVAVELRPQLEEADSQYQEAREA EEVNARAKAFSGYH	187	232
5	KKTEAEKALQSAKTKQKASSDLARSADKSAPLPENA QGFSKEPIEVEPLPNDRLNTTQADESAPIPISDTSS PSRVQSRGVEDGGRSPKSSMNEEGA	247	342

†The amino acid in underlined sequences (areas of interest) identified as epitopes by other software.

The BepiPred-2.0 default threshold setting of 0.5 was used for the prediction of the epitopes. A total of five stretches of amino acids were marked as 'E' for the epitope by the software. The amino acid sequences underlined were identified as epitopes by both ABCpred and BCpred methods.

3.4.4. Epitope prediction by LBtope method

The TmpA protein sequence with 345 amino acids was submitted to the LBtope program for the prediction of B cell linear epitopes. The LBtope gives scores between 0 and 100% to each predicted peptide. The default value of 80% was set in the program to increase the accuracy of prediction.

Table 11. Epitopes predicted by LBtope method and ranking scores

Rank	Sequence†	Start position	End position	Score
1	YRKVAHEGFGRLLPDMKARA	165	184	100.01
2	GVEDGGRSPKSSMNEEGASR	326	345	91.17
3	KEEAEKKAAEQRALLVESAH	27	46	88.28
4	APLPENAQGFSKEPIEVEPL	276	295	86.48
5	EKALQSAKTKQKASSDLARS	252	271	83.96
6	SDTSSPSRVQSRGVEDGGRS	314	333	82.93
7	EVNARAKAFSGYHRALEIYT	219	238	80.01

†: The stretch of amino acid in the underlined sequences (areas of interest) identified as epitopes by other software.

A total of seven amino acid sequences with scores between 80 and 100 %, indicating the possibility of being epitopes, were selected by the LBtope software (Table 11. Any predicted result

above 80% is a good score. Out of the seven predicted epitopes, three underlined sequences were also predicted by other prediction methods.

DNASTAR protean 3D epitope prediction method further showed that the underlined sequences are within the B cell epitope region (Figure 12).

The overlapping epitopes predicted by all the prediction programs used (ABCpred, BCPred, BepiPred-2.0 and LBtope) were considered as potential epitopes. Three candidate peptide epitope sequences consistently predicted by all the methods were chosen for further analysis. A peptide length of 15 amino acids to fit in the FSL selection of peptides, based on the solubility, was selected, and the chosen peptides were named TmpA1: KEEAEKKAAEQRALL, TmpA2: GGRSPKSSMNEEGASR and TmpA3: ALQSAKTKQKASSDLA.

3.5. Methods and Results: Analysis of the predicted TmpA peptides

The structure and the biophysical properties of the predicted peptides were analysed for their suitability to be built as FSL constructs as per the FSL selection algorithm table described earlier.

3.5.1. Structural analysis of the predicted TmpA peptides

A. Structural analysis using the protean 3D software

The predicted TmpA peptide sequences were analysed in the protean 3D program to locate and

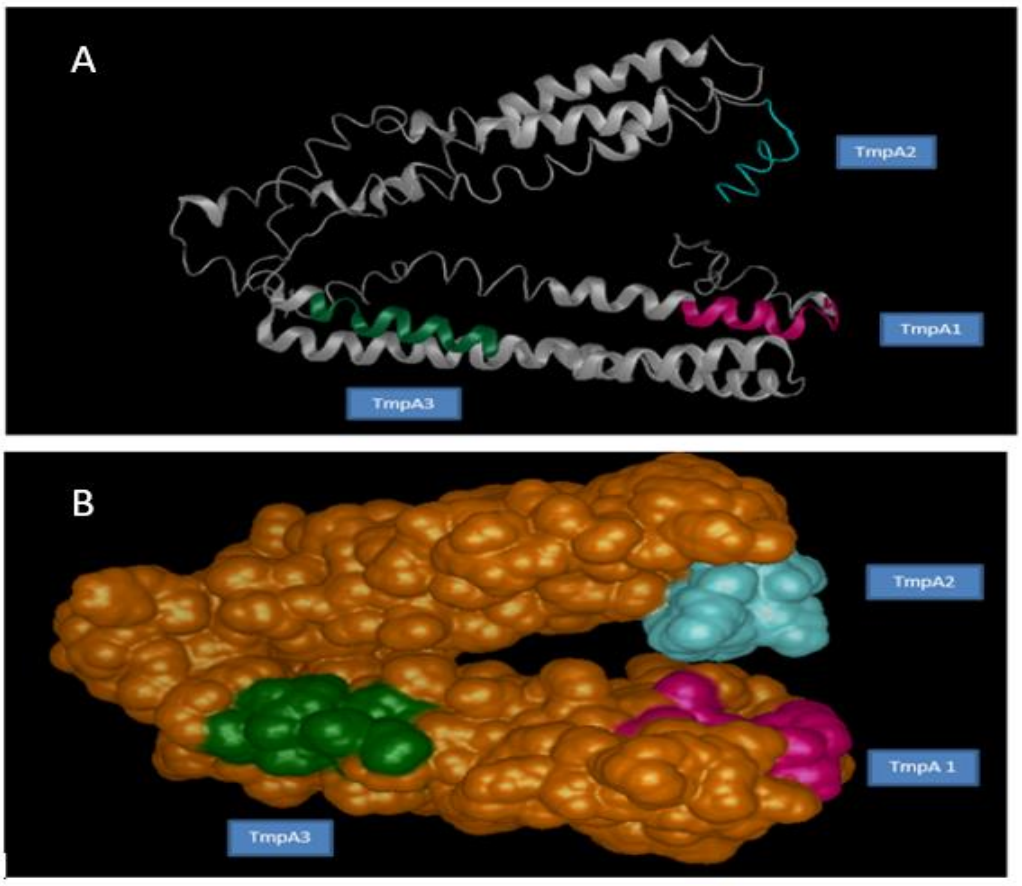

Figure 11: Location of syphilis epitopes, structural view (A) and solid view (B)

visualise the secondary structures of the predicted peptides. The location of the predicted TmpA peptide epitopes is shown in Figure 11. The structural view of the syphilis epitopes is shown in Plate A and the solid view of the syphilis epitopes are shown in Plate B. From the structure, it is evident that TmpA1 and TmpA3 are helices, TmpA2 is a coil and they are all exposed when seen in the solid view.

B. Structural prediction of the TmpA peptides using the I-TASSER server

The predicted TmpA peptide sequences were analysed in the I-TASSER server and the structure of the three predicted syphilis peptides is shown in Figure 12. TmpA1 is from amino acid 27 to amino acid 41, with 15 residues in total. Its predicted structure by I-TASSER is a helix/coil, and it is located in the N termini of the peptide. TmpA2 is from amino acid 330 to amino acid 345, with 16 residues in total. Its predicted structure by I-TASSER is a coil. TmpA3 is from amino acid 254 to amino acid 269, with 16 residues in total. Its predicted structure by I-TASSER is a helix/coil, and it

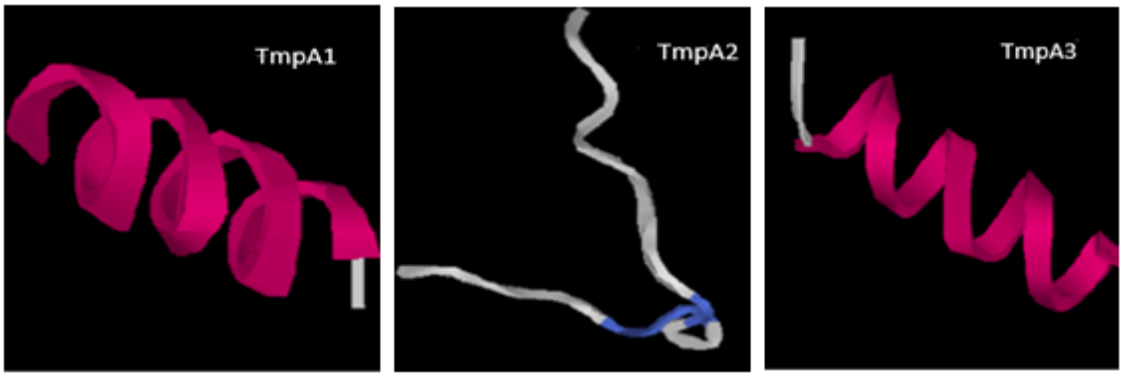

Figure 12: I-TASSER predicted models of TmpA candidate peptides

is located in the middle of the peptide. Locating peptides on the N- or C-terminal of the protein has traditionally been a popular approach for peptide design, since there seems to be a higher probability that these regions will be exposed. The results of the I-TASSER prediction are identical to the sequence-based structures predicted by the BepiPred-2.0 program. Results from both the sequence and structure-based prediction programs are in agreement for all three syphilis peptides.

3.5.2. Analysis of biophysical properties of the predicted TmpA peptides

A. Antigenicity of TmpA peptides

The antigenicity predictions of TmpA peptides from DNASTAR protean 3D are shown in Figure 13. The B cell epitope prediction bar is shown at the top of the figure, followed by results from Jameson-Wolf's and Welling's antigenicity prediction. The location of the peptides is within the boxed area. TmpA1 peptide sequence is from amino acid 27 to 41, TmpA2 is from 330 to 345, and TmpA3 is from 254 to 269. The blue line divides the B cell epitope predictions from the antigenicity predictions attained by the Jameson-Wolf and Welling methods. Prediction of antigenicity in the Welling method compares the amino acid present in known antigenic determinants to the percentage of the amino acid in the average protein sequence. The Jameson-Wolf technique combines known methods used for the prediction of protein structures to predict potential antigenic determinants.

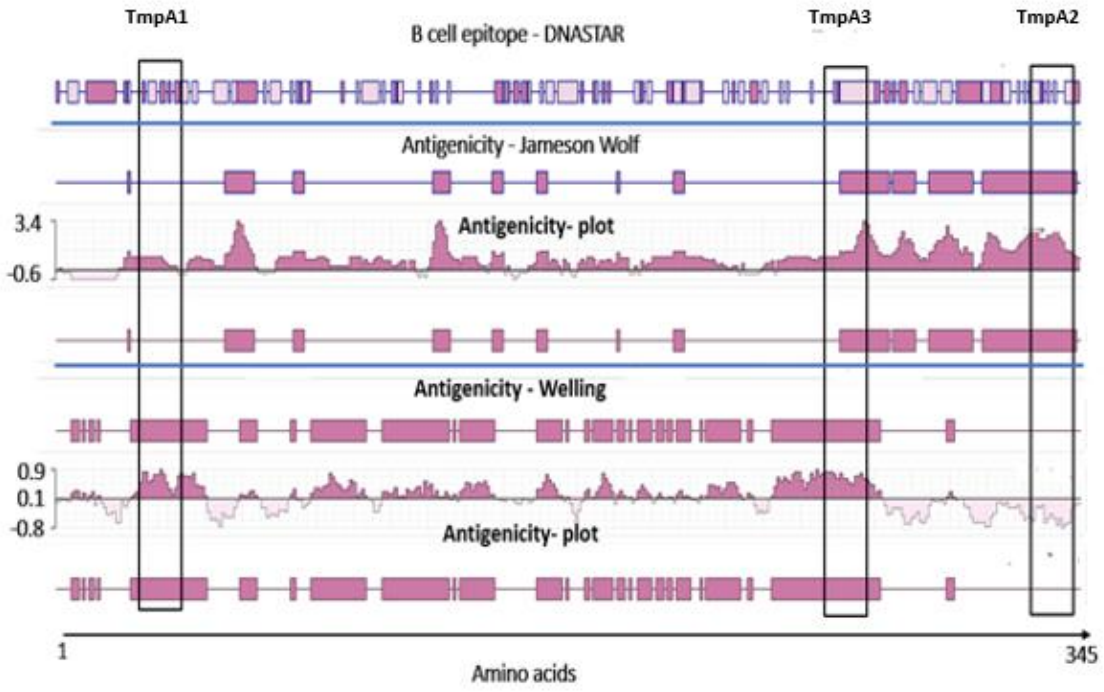

Figure 13: Antigenicity predictions of TmpA peptides using DNASTAR protein 3D

In the antigenicity plot, the result is shown as peaks, with each peak indicating a potential antigenic determinant. The antigenic region graph is plotted using a threshold value of 0.1 for the Welling method and 0.6 for the Jameson-Wolf method. Peptides TmpA1 and TmpA3 are predicted as antigenic by both the methods, with results well above the threshold; TmpA2 is predicted as highly antigenic only by the Jameson-Wolf method.

The peptides were also tested using the Kolaskar/Tongaonkar antigenicity program[130] and all three peptides were found to be highly antigenic. The Kolaskar/Tongaonkar gives antigenicity scores (Table 12) unlike threshold settings given in the Jameson-Wolf and Welling methods.

B. Surface accessibility prediction of TmpA peptides

The surface accessibility prediction by the Emini method from DNASTAR protean 3D is shown in Figure 14. The boxed area indicates the location of the three candidate syphilis peptides. TmpA1 peptide sequence is from amino acid 27 to 41, TmpA2 is from 330 to 345, and TmpA3 is from 254 to 269. The peaks seen in the middle of the figure are the surface probability plot. The line seen underneath the surface probability plot is the threshold line. The predicted peptide sequence has an increased probability of being found on the surface if the surface probability of that peptide is more than 1.0 (threshold).

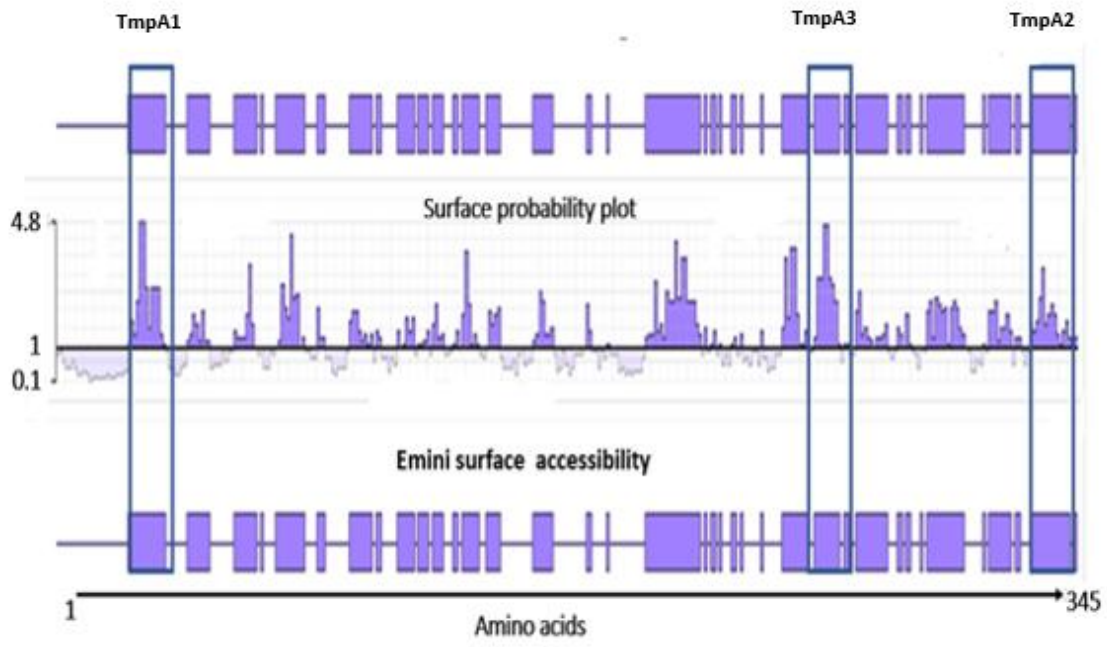

Figure 14: Surface probability predictions of TmpA peptides

All three peptides have peaks considerably above the threshold, suggesting the epitopes are surface exposed and therefore suitable candidates. Peptides TmpA1 and TmpA2 are located in the N-and C-terminal. It is well known that protein's C- and N-termini are frequently surface exposed and have a high degree of flexibility. Antibody access to target epitopes may be hindered by protein shape. If the epitope is concealed within the interior of the folded protein, it may become inaccessible. The above results show that all three peptides have surface accessibility.

C. Hydrophilicity prediction of TmpA peptides

Hydrophilicity is a vital feature to consider when predicting antigenic determinants. Parker hydrophilicity prediction is used extensively in peptide design and for selecting appropriate candidate antigens. The Parker hydrophilicity prediction results for syphilis peptides using DNASTAR protean 3D are shown in Figure 15. All three peptide sequences are in the boxed area. TmpA 1 peptide sequence is from amino acid 27 to 41, TmpA 2 is from 330 to 345, and TmpA3 is from 254 to 269. The blue peaks in the lower half of the figure are the hydrophilicity plot. The line seen underneath the hydrophilicity plot is the threshold line. The predicted peptide sequence has an increased probability of being hydrophilic if the hydrophilicity plot is greater than the set threshold of 1.8. The peaks seen in orange in the upper half of the figure are the hydrophobicity plot. The line seen underneath the hydrophobicity plot is the threshold line. The predicted peptide sequence has an increased probability of being hydrophobic if the hydrophobicity plot is more than the set threshold of 1.6. TmpA1 peptide has both hydrophilic and hydrophobic residues, whereas TmpA 2 and TmpA 3 peptides are not hydrophobic.

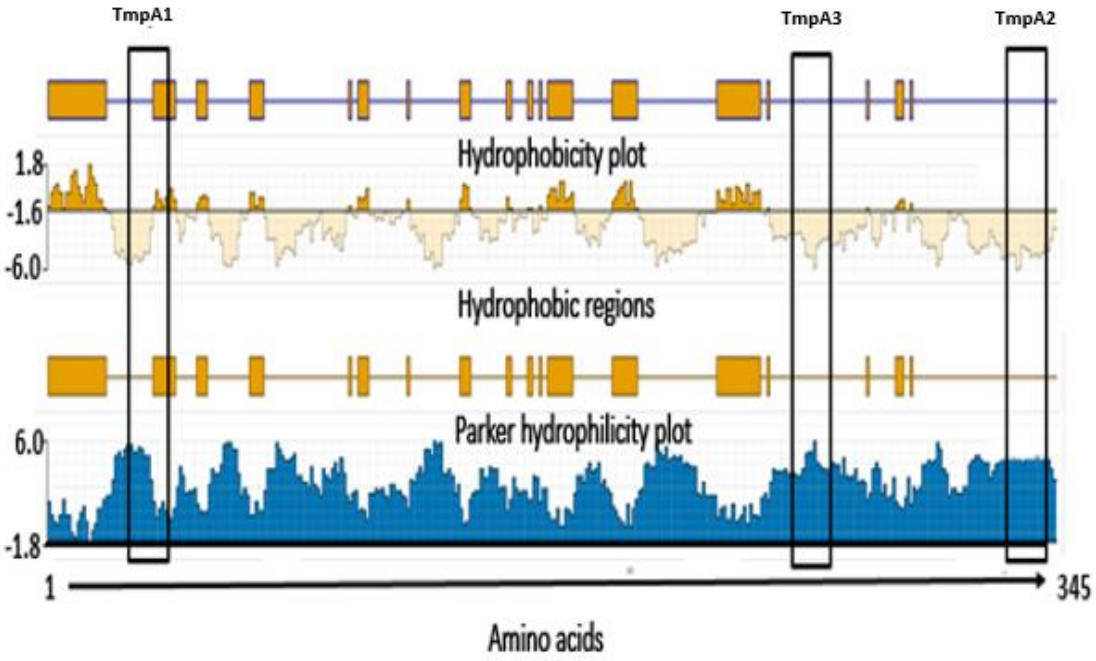

Figure 15: Hydrophilicity predictions of TmpA peptides by the Parker method

Hydrophilic peptides pose fewer problems during peptide synthesis; hydrophobic peptides can be more difficult to dissolve. The hydrophobic residues in TmpA1 are less than 25%. The peptide sequences with less than 25% hydrophobic residues are not expected to have solubility issues, and this suggests TmpA1 peptide will be suitable. However, the actual solubility of the peptides can be determined only by empirical testing. The peaks seen in the blue colour in the lower half of the picture are the hydrophilicity plot. All three peptides have peaks way above the threshold, suggesting that the epitopes are hydrophilic and suitable candidates.

The Kyte Doolittle hydrophilicity prediction results for syphilis peptides using DNASTAR protean 3D is shown in Figure 16. TmpA 1 peptide sequence is from amino acid 27 to 41, TmpA 2 is from 330 to 345, and TmpA3 is from 254 to 269, and they are shown in the boxed area. The peaks seen in the orange colour in the upper half of the picture are the hydrophobicity plot. The line seen underneath the hydrophobicity plot is the threshold line. The predicted peptide sequence has an increased probability of being hydrophobic if the hydrophobicity plot is greater than 0 (threshold). TmpA1 has hydrophilic and some hydrophobic residues, whereas TmpA2 and TmpA3 peptides are not hydrophobic. The peaks seen in blue in the lower half of the picture are the hydrophilicity plot. The line seen underneath the hydrophilicity plot is the threshold line. The predicted peptide sequence has an increased probability of being hydrophilic if the hydrophilicity plot is greater than 0 (threshold).

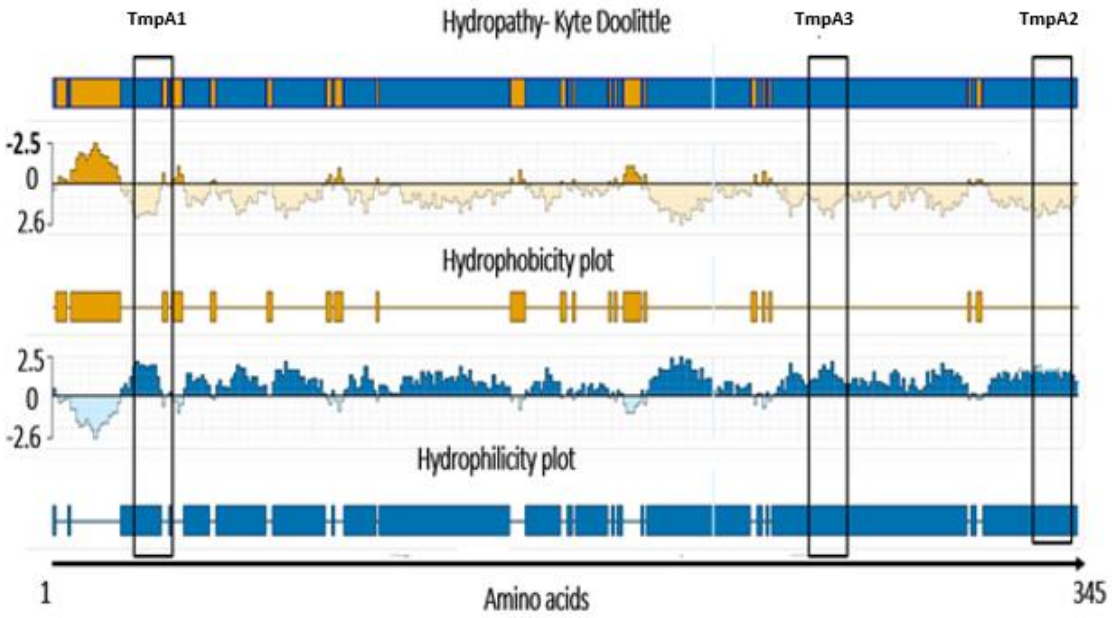

Figure 16: Hydrophilicity predictions of TmpA peptides by Kyte Doolittle method

All three peptides have peaks way above the threshold, suggesting that the epitopes are hydrophilic and suitable candidates. Kyte Doolittle's hydrophilicity predictions match the results from the Parker method of prediction.

The predicted peptides were also analysed for the glycosylation sites using the NetOGlyc and NetNGlyc program and no glycosylation sites were found in the predicted peptides. There was no internal cysteine or terminal glutamine in the predicted peptide epitopes.

3.5.3. Syphilis TmpA peptides: results of the theoretical prediction

The theoretical prediction of the syphilis peptides using various parameters and the results obtained are shown in Table 12a and Table 12b. Table 12a shows the epitope predictions, basic structure, antigenicity and specificity of the peptides. Table 12b shows the biophysical properties of the peptides. All the three peptides were predicted as B cell epitopes by the BepiPred-2.0, BCpred, ABCpred, DNASTAR and LBtope methods. All three syphilis peptides have good antigenicity scores from Kolaskar/Tongaonkar. A score close to 1 and above is considered as antigenic, and all three peptides have an antigenicity score close to 1. The peptides TmpA1 and TmpA2 showed good antigenicity scores of 4.2 and 3.7 by Antigen profiler method. Any score above 3.6 is considered as excellent antigen. The peptide TmpA 3 showed an antigenicity score of 2.1, which is moderately antigenic. The Jameson-Wolf method predicted all the three peptides

3.5.4. Results of the theoretical prediction of the TmpA FSL peptides

Table 12 Results of theoretical prediction of the leptospiral peptides

12a Epitope prediction			
Peptide details	TmpA1	TmpA2	TmpA3
Basic structure	Helix/Coil	Coil	Helix/Coil
Epitope prediction			
BepiPred-2.0	Yes	Yes	Yes
BCPred	Yes	Yes	Yes
ABCpred	Yes	Yes	Yes
DNASTAR	Yes	Yes	Yes
LBtope	Yes	Yes	Yes
Antigenicity			
Kolaskar/Tongaonkar	0.973	0.921	0.996
Antigen profiler	4.2	3.7	2.1
Similarity			
BLAST (%)	100	100	100

12b Epitope Biophysical properties			
Peptide details	TmpA1	TmpA2	TmpA3
Internal cysteine	0	0	0
Terminal Glutamine	0	0	0
NetOGlyc	0	0	0
NetNGlyc	0	0	0
Net charge	0	2	1
Surface Accessibility	<1	<1	<1
Isoelectric point	6.94	10	10
Parker scale	4.25	4.5	4.9
Peptide length	15	16	16

to be antigenic, as the scores were above the set threshold of 0.6 and they were closer to the value 3.0, indicating they have good antigenicity. The peptides showed no glycosylation site when tested in NetOGlyc and NetNGlyc programs. The predicted peptides did not have terminal glutamine or internal cysteine residues. The predicted peptides, when tested in the NCBI BLAST program for similarity check, did not have any relatedness to other microbial organisms.

3.5.5. Candidates of syphilis peptides for Kode FSL construction

A total of three syphilis peptides were selected for Kode FSL construction. Table 13 shows the predicted peptide sequences, peptide length and the location of the peptides in *T. pallidum*

Table 13 List of syphilis peptides for FSL construction

Peptide name	Location in the protein sequence	Length	FSL construction Peptide sequence[c]*
TmpA1	N-terminal	15	KEEAEKKAAEQRALL[C]
TmpA2	C-terminal	15	[C]GGRSPKSSMNEEGASR
TmpA3	Middle	16	[C]ALQSAKTKQKASSDLA

[C]*: C is the conjugation cysteine which has been added to the peptide, and also determines the orientation to the membrane as it is the linkage to the spacer.

TmpA protein sequence. The peptides TmpA1 is located in the N-terminal, TmpA2 is located in the C-terminal and TmpA3 Is located in the middle of the protein. A cysteine residue is added to the final peptide sequence at either the amino or carboxyl ends to conjugate the peptide to the spacer and this determines the orientation of the peptide to the membrane.

3.6. Syphilis FSL constructs

The Laboratory of Carbohydrates at the Shemyakin Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences in Moscow synthesised the candidate syphilis peptides from the FSL feasibility algorithm for this research. The purity, integrity and quality control of the syphilis Kode constructs was done by Nuclear Magnetic Resonance spectroscopy and thin layer chromatography and were >98%[132,133]. By ligating the functional head (peptide) with the cysteine residue to a spacer (CMG) linked to a lipid (DOPE) tail, the peptides were synthesized into FSL constructs.

The structures of the syphilis FSL constructs and the molecular weight are shown in Figures 17 to 19. The dotted line denotes the beginning and end of the CMG based spacer that separates the function and lipid segments[141].

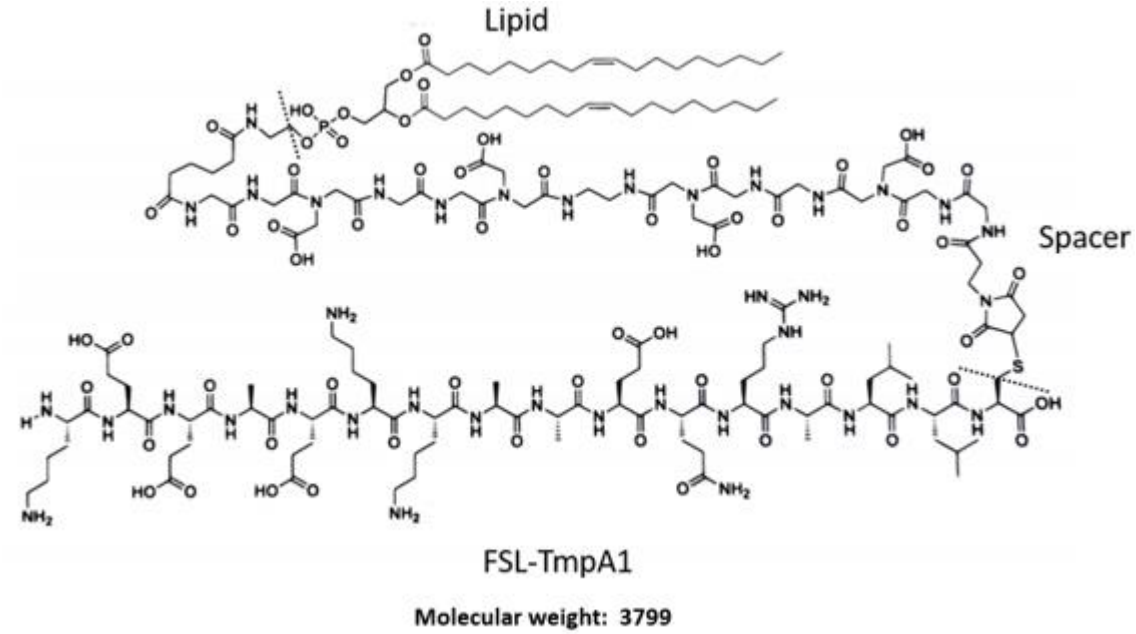

Figure 17: Schematic diagram, syphilis peptide TmpA1 FSL (KEEAEKKAAEQRALL) construct

Figure 17 shows the function-spacer-lipid (FSL) construct consisting of a lipid phosphate moiety conjugated to the spacer which is conjugated via a cysteine SH group to the peptide functional head (TmpA1-KEEAEKKAAEQRALL). The dotted line denotes the beginning and end of the carboxymethylated oligoglycine-based spacer that separates the function and lipid segments. Figures 18 and 19 feature the same structural logic and denotations.

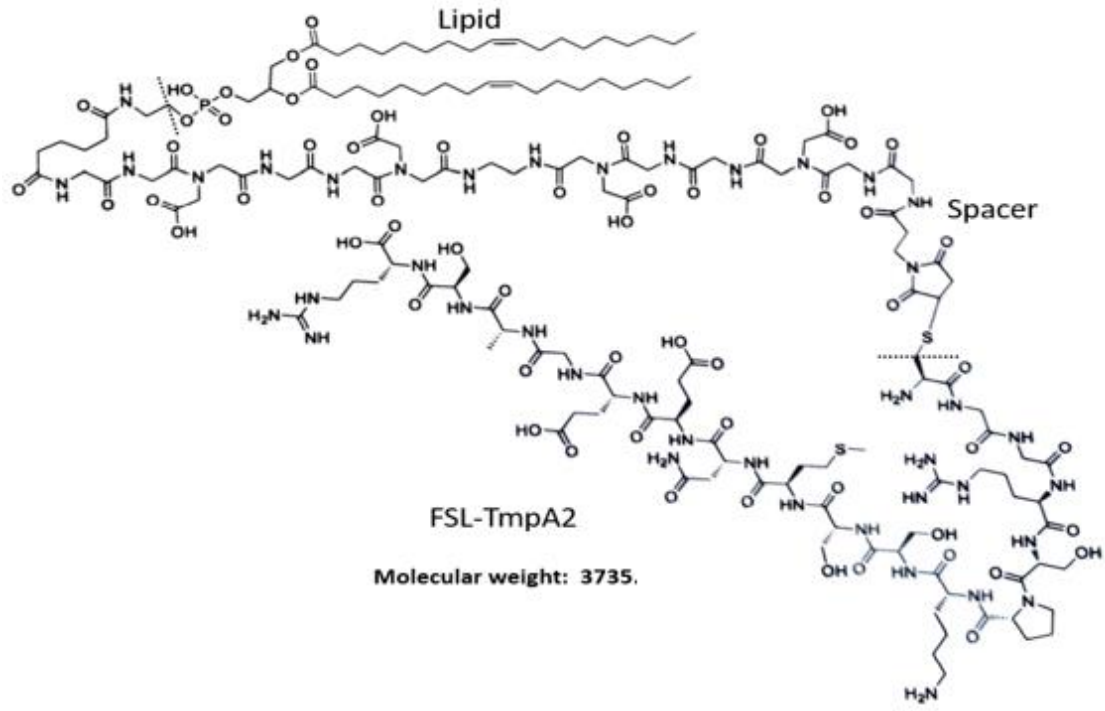

Figure 18: Schematic diagram, syphilis peptide TmpA2 (GGRSPKSSMNEEGASR) FSL construct

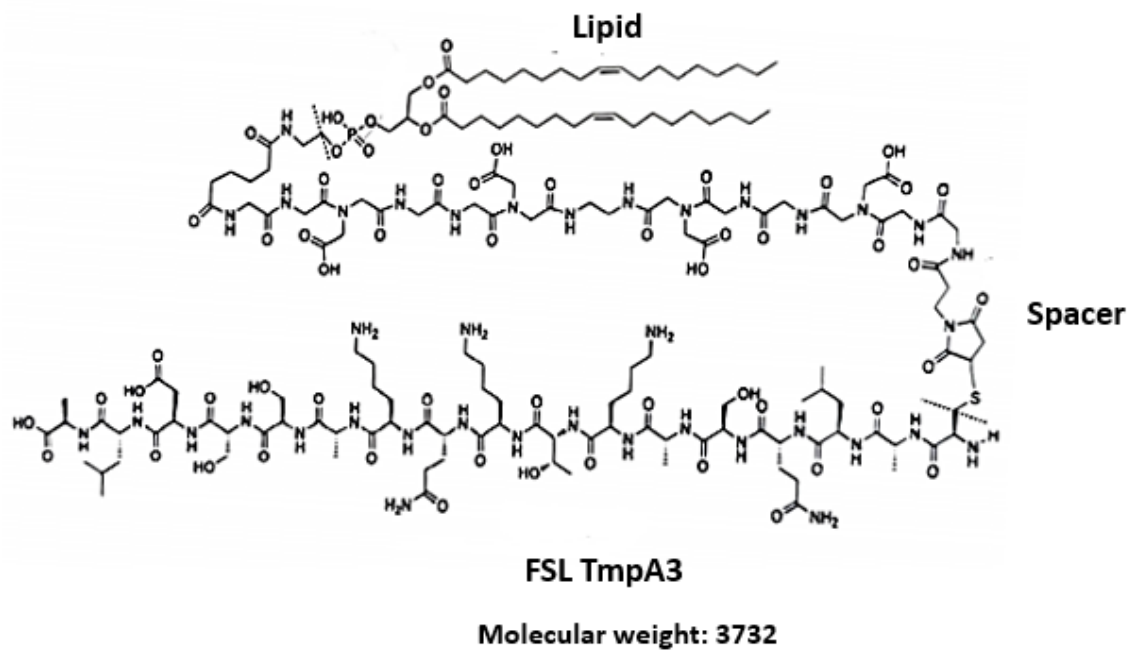

Figure 19: Schematic diagram, syphilis peptide TmpA3 (ALQSAKTKQKASSDLA) FSL construct

A generic look at the full protein, using the algorithm, selected out three top candidates.

3.7. Methods and Results: Functional prediction of syphilis FSL peptides

Syphilis peptides were tested against clinical samples as a proof of concept approach.

3.7.1. Materials

1. Syphilis positive and negative samples

De-identified true syphilis positive and true syphilis negative serum samples (routinely identified in the clinical laboratory), no longer required for laboratory analysis, were obtained from LabPlus at Auckland City Hospital to validate the peptides. Once received in the laboratory, each sample was assigned a number and stored at -80° C.

The ethics approval granted for this study is only for the use of de-identified samples for the initial validation of the peptides and does not include the access to either demographic or clinical data such as the age of the patient, the stage of the disease (primary, secondary or latent syphilis) and the days of onset of symptoms.

The following are the sera bank data received for the de-identified positive and negative samples used in this study.

True positive samples are the samples that were positive by the commercial syphilis EIA screen assay, positive or borderline positive by the syphilis confirmatory TPPA assay and positive or borderline positive by RPR assay.

False positive samples are the samples that were positive by the commercial syphilis EIA screen assay and negative by the syphilis confirmatory assay TPPA, regardless of the RPR assay result.

True negative samples are the samples that were negative by the commercial syphilis EIA screen assay and not confirmed by syphilis confirmatory TPPA assay as per the hospital/laboratory reporting guidelines.

2. Control samples

De-identified healthy blood donor samples were used as normal control samples for this study.

3. Red cell preservative solutions

- Celpresol, Cat no: CSL 06332301, Immulab, Melbourne, Australia.
- Cellstab, Cat no: 005650, Bio-Rad Laboratories Inc, Hercules, CA, USA.

4. Phosphate buffered solution (PBS)

PBS solution (137mM NaCl, 10mM Phosphate, 2.7mM KCl, pH 7.4)

5. Ortho ORTHO BioVue cards – column agglutination test (CAT)

ID (Infectious Diseases) Micro Typing system, Polyclonal rabbit anti human IgG cards, Cat no: MTSO84029, Ortho clinical diagnostics, Raritan, NJ, USA.

6. Natural O cells

Red blood cells from an O phenotype individual. These were obtained throughout this study from a single individual. The EDTA blood tube was spun and the plasma was removed. The packed cells were washed 3 times with 1% PBS solution and 1 time with celpresol solution, the hard packed cells are stored in celpresol for four weeks at 4° C[131,145,199].

7. Syphilis FSL constructs

The selected syphilis candidate peptides were constructed into FSL constructs (Kode Biotech materials, Auckland, New Zealand). The FSL constructs were obtained as lyophilised, based on the synthesised weight from the materials data sheet, the FSL constructs were reconstituted in celpresol to make 1mg/mL stock solutions and the aliquots were stored at -80° C.

3.7.2. Methods and Results

Preparation of syphilis kodecytes

The kodecytes are described by the identification (ID) of the functional head of the FSL and the micromolar concentration (µmol/L) of the FSL solution used to make the kodecytes. For example, syphilis TmpA1-30 kodecytes is made up of the syphilis FSL peptide TmpA1 at a concentration of 30 µm.

A total of 12 syphilis kodecytes, TmpA1-100, TmpA1-50, TmpA1-30, TmpA1-10, TmpA2-100, TmpA2-50, TmpA2-30, TmpA2-10, TmpA3-100, TmpA3-50, TmpA3-30 and TmpA3-10, were made for the syphilis kodecytes assay using the following protocol.

- Removed the stock vials of syphilis FSL constructs (1 mg/mL) from the -80° C freezer.
- Allowed the vials to come to room temperature.
- Reconstituted the vials by adding 1 mL of red cell preservative solution (celpresol).
- Made four different concentrations (100 µM, 50 µM, 30 µM, and 10 µM) of FSL constructs from each syphilis FSL construct (TmpA1, TmpA2 and TmpA3).
  Example: TmpA1 stock FSL vial is 1 mg/mL, which is 1g in 1L. Molecular weight of TmpA1 is 3799.22. Molarity at 1 mg/mL (µM) is 1/3799.22*1000000 = 263.2 µM. The concentration required is 100 µM and the volume of FSL construct required is 250 µl. Added 95.05 µl to 154.95 µl celpresol solution to get 250 µl of FSL TmpA1 constructs at 100 µl concentration.
- Added 250 µl of each concentration of the three syphilis FSL peptides to the labelled tubes.
- Added 250 µl of packed group O red blood cells to each tube.
- Incubated the tubes at 37°C for 2 hours.
- Washed 1 time with Phosphate buffered saline and 1 time with celpresol.
- The packed Kodecytes were resuspended in 5% celpresol and stored at 4° C.
- The control cells were made without adding the FSL peptides In other words, the control cells were unmodified group O red cells in celpresol solution.
- The kodecytes were rested overnight before use and were used within 21 days from the date of preparation.

Syphilis kodecytes assay using ORTHO BioVue gel cards (CAT)

The syphilis positive samples, syphilis negative samples, control samples from healthy blood donors, WHO external quality control samples and seroreactive samples from other infectious agents were all tested using the syphilis kodecytes assay.

1. Made 0.8% suspension (in celpresol) of each syphilis kodecyte (TmpA1-100, TmpA1-50, TmpA1-30, TmpA1-10, TmpA2-100, TmpA2-50, TmpA2-30, TmpA2-10, TmpA3-100, TmpA3-50, TmpA3-30 and TmpA3-10).
2. Labelled the ORTHO BioVue gel cards and removed the foil.
3. Added 50 µl of 0.8% suspension of each syphilis kodecyte to the corresponding labelled wells in the ORTHO BioVue gel cards.
4. Added 40 µl of serum samples to the corresponding wells.
5. Incubated the gel cards at 37° C for 15 minutes.
6. Centrifuged for 5 minutes at preset rpm as per protocol.
7. Graded results based on the agglutination pattern.
8. Samples reactive to syphilis constructs were also tested with unmodified group O red cells as part of the quality control, in addition to the known syphilis negative and syphilis positive control.

Interpretation of column agglutination test results using ORTHO BioVue gel cards:

1. 4+: Agglutinated cells form a band at the top of the bead column.
2. 3+: Most agglutinated cells remain at the top of the column.
3. 2+: Agglutinated cells observed throughout and a small button at the bottom of the column is seen.
4. 1+: Most agglutinated cells remain in the lower half and a button at the bottom of the column is seen.
5. Weak positive (w): weak reaction on top of the button.
6. Negative (Neg): compact clear button at the bottom of the column is seen.

The methodology and the scoring system were as recommended by the manufacturer, including the grade “w” for the weak positive reaction. An example of the serological grades of the kodecytes TmpA1-10 are shown in Figure 20. The syphilis positive samples are in Columns 1 to 5, and the syphilis negative control sample is in the “Neg” column. The syphilis positive sample in column 1 shows the serological grade of 4+ (interpretations, 4+: is a strong positive result and the agglutinated cells form a band at the top of the bead column). The samples in columns 3 to 5

show a serological grade of 3+ (interpretations, 3+: is a positive result and most agglutinated cells remain at the top of the column).

Figure 20: Example of Ortho BioVue card syphilis kodecyte results
Serological grades indicated against five positive samples (columns 1 to 5).

The sample in column 2 shows a serological grade of 2+ (interpretations, 2+: is a moderately positive result, and agglutinated cells observed throughout the column, with some cells forming a button at the bottom of the column). The negative result is shown in column 6, marked as "Neg" (interpretations, Neg: no agglutination and compact clear button at the bottom).

Initial evaluation of syphilis FSL constructs for specificity as kodecytes

Method overview

As part of the initial evaluation of the syphilis FSL constructs as kodecytes, a range of concentrations from 100 μM to 10 μM were tested with expected negative samples from healthy blood donors. A total of 94 healthy blood donor samples (expected negative) were tested with syphilis kodecytes (TmpA1-100, TmpA1-50, TmpA1-30, TmpA1-10, TmpA2-100, TmpA2-50, TmpA2-30, TmpA2-10, TmpA3-100, TmpA3-50, TmpA3-30, TmpA3-10) using the syphilis kodecyte assay (CAT).

Results

The results of the syphilis kodecytes assay for the healthy blood donor samples (expected negatives) is shown in Figure 21. Kodecytes TmpA1-100 and TmpA1-50 showed weak positive reactions with a percentage specificity of 94% and 95%. Kodecytes TmpA1-30 and TmpA-10 did

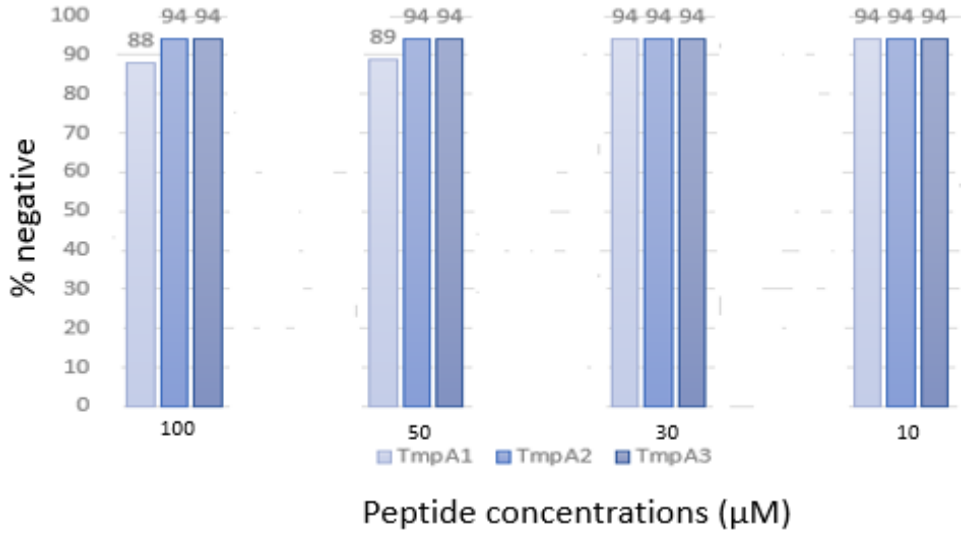

Figure 21: Syphilis kodecytes results for healthy blood donors (expected negatives)

not show any positivity (100 % specificity). Kodecytes TmpA2-100, TmpA2-50, TmpA2-30, TmpA2-10, TmpA3-100, TmpA3-50, TmpA3-30, TmpA3-10, were negative for all the 94 samples, and showed 100 % specificity. The results show that the syphilis kodecytes have overall good specificity, except TmpA1-100 and TmpA1-50, which showed weak reactions, suggesting lower concentrations for syphilis FSL TmpA1 in terms of specificity.

Performance of syphilis TmpA1 kodecytes with EIA positive and negative samples

Method overview

A total of 211 EIA positive and 197 EIA negative syphilis samples were used to verify the functional prediction of the syphilis kodecyte TmpA1. The kodecytes TmpA1-100, TmpA1-50, TmpA1-30 and TmpA1-10 were tested against both EIA syphilis screen positive and EIA syphilis screen negative clinical samples using the syphilis kodecyte assay. As part of the quality control, unmodified group O cells were also used as controls, in addition to known positive and negative controls.

Results

The syphilis kodecytes TmpA1-100, TmpA1-50 and TmpA1-30 performed well, identifying all the positive samples correctly with 100 % sensitivity (Table 14). The kodecytes TmpA1-10 identified 208 EIA screen positive samples correctly and identified three false negative results out of the

Table 14. Results of TmpA1 kodecytes for EIA positive and negative samples

FSL-TmpA1 (μM)		EIA Positive (n=211)	EIA Negative (n=197)
		Positive	Negative
TmpA1-100	Positive	211 (100%) †	46 (70%) ††
	Negative	0	151
TmpA1-50	Positive	211 (100%) †	34 (79%) ††
	Negative	0	163
TmpA1-30	Positive	211 (100%) †	20 (89%) ††
	Negative	0	177
TmpA1-10	Positive	208 (98.6%) †	4 (98%) ††
	Negative	3	193

†: Percentage sensitivity of the TmpA1 kodecytes for EIA screen positive samples. ††: Percentage specificity of the TmpA1 kodecytes for EIA screen negative samples.

total 211 samples tested, with a sensitivity of 98.6%. Among the 197 EIA negative samples tested, false positive results were seen in all four kodecytes tested. Kodecytes TmpA1- 100 gave 46 false positive results, TmpA1-50 gave 34 false positive results, TmpA1-30 gave 20 false positive results and TmpA1-10 gave four false positive results, with a percentage specificity of 70%, 79%, 89% and 98%, respectively. The control sample results were as expected. The above results show that the TmpA1 kodecytes perform best at 10 μM concentrations with good sensitivity and specificity (98.6% and 98%), which is similar to the findings in the initial specificity evaluation experiment with the healthy blood donor samples.

Performance of syphilis TmpA2 kodecytes with EIA positive and negative samples

Method overview

The performance evaluation of syphilis kodecytes TmpA2 was done following the protocol used for the TmpA1 kodecytes. The kodecytes TmpA2-100, TmpA2-50, TmpA2-30 and TmpA2-10 were tested using the kodecytes assay against 211 EIA syphilis screen positive and 197 EIA syphilis screen negative clinical samples.

Results

The TmpA2 kodecytes showed false negative results at higher concentrations when tested against the 211 EIA syphilis screen positive samples table (Table 15). The kodecytes TmpA2-100, TmpA2-50, TmpA2-30 and TmpA2-10 showed a percentage sensitivity of 96.2%, 92.4%, 89% and 75%, respectively.

Table 15. Results of TmpA2 kodecytes for EIA positive and negative samples

FSL-TmpA2 (μM)		EIA Positive (n=211)	EIA Negative (n=197)
		Positive	Negative
TmpA2-100	Positive	203 (96.2%) †	13 (93.4%) ††
	Negative	8	184
TmpA2-50	Positive	195 (92.4%) †	6 (97%) ††
	Negative	16	191
TmpA2-30	Positive	188 (89%) †	1 (99.4%) ††
	Negative	23	196
TmpA2-10	Positive	158 (75%) †	1 (99.4%) ††
	Negative	53	196

†: Percentage sensitivity of the TmpA2 kodecytes for EIA screen positive samples. ††: Percentage specificity of the TmpA2 kodecytes for EIA screen negative samples.

The TmpA2 kodecytes gave false positive results when tested against 197 EIA syphilis screen negative samples. The kodecytes TmpA2-100 gave 13 false positive results, TmpA2-50 gave six false positive results, TmpA2-30 and TmpA2-10 gave one false positive each, showing the percentage specificity of 93.4%, 97%, 99% and 99.4%, respectively. TmpA2 kodecytes have good specificity at lower concentrations, but the sensitivity decreases at lower concentrations. Based on the results obtained, the performance of TmpA2 kodecytes was better in sensitivity (92.4%) and specificity (97%) at 50 μM concentration.

Performance of syphilis TmpA3 kodecytes with EIA positive and negative samples

Method overview

The functional predictions of syphilis FSL peptide TmpA3 were done using the kodecytes TmpA3-100, TmpA3-50, TmpA3-30 and TmpA3-10. The kodecytes were tested against 211 EIA syphilis screen positive and 197 EIA syphilis screen negative clinical samples. The unmodified group O red cells were used as controls as part of the testing protocol.

Results

TmpA3 kodecytes showed false negative results when tested against 211 EIA syphilis screen positive (Table 16). The kodecytes TmpA3-100 gave one false negative result, TmpA3-50 gave five false negative results, TmpA3-30 gave 11 false negative results and TmpA3-10 gave 48 false negative results, with a percentage sensitivity of 99.5%, 98%, 95% and 77.2%, respectively. This clearly suggests that the syphilis TmpA2 FSL peptide does not have good sensitivity at lower

Table 16. Results of TmpA3 kodecytes for EIA screen positive and negative samples

FSL-TmpA3 (μM)		EIA Positive (n=211)	EIA Negative (n=197)
		Positive	Negative
TmpA3-100	Positive	210 (99.5%) †	36 (84.7%) ††
	Negative	1	161
TmpA3-50	Positive	206 (97.6%) †	17 (91.8%) ††
	Negative	5	180
TmpA3-30	Positive	200 (94.7%) †	4 (97.9%) ††
	Negative	11	193
TmpA3-10	Positive	163 (77.2%) †	0 (100%) ††
	Negative	48	197

†: Percentage sensitivity of the TmpA3 kodecytes for EIA screen positive samples. ††: Percentage specificity of the TmpA3 kodecytes for EIA screen negative samples.

concentrations, similar to the TmpA2 syphilis FSL peptides. Out of the 197 EIA syphilis screen negative samples tested, the kodecytes TmpA3-100 showed 36 false positive results, TmpA3-50 showed 17 false positive results, TmpA3-30 showed four false positive results and TmpA3-10 showed zero false positive results, with a percentage specificity of 85%, 92%, 98% and 100%, respectively. TmpA3 kodecytes perform well at 30 μM concentration with sensitivity of 95% and specificity of 98%.

Comparison of TmpA1 kodecytes results with TPPA results

The results obtained from the syphilis TmpA1 kodecytes assay for the 211 EIA syphilis screen positive samples were then used to compare the performance of the TmpA1 kodecyte versus

Table 17. Comparison of TmpA1 kodecytes results with TPPA results

FSL TmpA1(μM)		TPPA results (n=211)				Total
		Positive (P)	Borderline (B)	P+B	Negative	
TmpA1-100	Positive	188 (100%) †	11	199	12	211
	Negative	0	0		0	0
TmpA1-50	Positive	188 (100%) †	11	199	12	211
	Negative	0	0		0	0
TmpA1-30	Positive	188 (100%) †	11	199	12	211
	Negative	0	0		0	0
TmpA1-10	Positive	187 (99.5%) †	11	198	10	208
	Negative	1	0		2	3

The results show the TmpA1 kodecytes performance versus TPPA test results with a total of 211 EIA screen positive syphilis samples. †: Percentage sensitivity of the TmpA1 kodecytes for TPPA confirmed positive samples.

the confirmatory test TPPA results (Table 17). The TPPA results were obtained along with the de-identified samples for validation purposes. Analysis of TmpA1 kodecytes results with the TPPA results showed that 199 samples which were TPPA positive were also positive for TmpA1-100, TmpA1-50, and TmpA1-30 kodecytes. The kodecyte TmpA1-30 gave one negative result, but the TPPA test was positive. All four kodecytes, TmpA1-100, TmpA1-50, TmpA1-30 and TmpA1-10, picked up the 11 TPPA borderline positive samples which were also EIA syphilis screen test positive. Out of the 211 syphilis EIA screen positive samples received for validation purposes, 12 samples were syphilis EIA screen positive and TPPA negative. The kodecytes TmpA1-100, TmpA1-50 and TmpA1-30 showed false positive results for the 12 TPPA negative samples. The kodecytes TmpA1-10 showed 10 false positive results for the 12 TPPA negative samples. Out of the 12 TPPA negative samples tested, three were RPR positive, so there is a possibility of kodecyte assay results being true positive, but no evidence, from either the clinical details or the follow up test results, is available to prove this.

Overall, TmpA1 kodecytes results compared well with the EIA Screen test results and the TPPA positive results. Three of the 11 TPPA borderline positive samples were also RPR test positive, suggesting that the syphilis kodecytes assay is more sensitive than the TPPA test.

Comparison of TmpA2 kodecytes results with TPPA results

The results obtained from the syphilis TmpA2 kodecytes assay for the 211 EIA syphilis screen positive were used to evaluate the performance of the syphilis kodecyte TmpA2 against TPPA test results (Table 18). Analysis of TmpA2 kodecytes results with the TPPA results showed that the

TmpA2-100 showed correlation for 181 TPPA positive samples, and missed seven samples. The kodecytes TmpA2-50 showed correlation for 176 samples, and missed 12 samples. The

Table 18 Comparison of TmpA2 kodecytes results with TPPA results

FSL TmpA2 (µM)		TPPA results (n=211)				Total
		Positive (P)	Borderline (B)	P+B	Negative	
TmpA2-100	Positive	181 (96 %) †	10	191	12	203
	Negative	7	1		0	8
TmpA2-50	Positive	176 (94%) †	8	184	12	196
	Negative	12	3		0	15
TmpA2-30	Positive	169 (90%) †	6	175	9	184
	Negative	19	5		3	27
TmpA2-10	Positive	145 (77%) †	6	151	7	158
	Negative	43	5		5	53

†: Percentage sensitivity of the TmpA2 kodecytes for TPPA confirmed positive samples.

TmpA2-30 showed correlation for 169 samples and TmpA2-10 showed correlation for 145 samples. Out of the 11 TPPA borderline positive samples, which were also EIA syphilis screen positive, the kodecytes TmpA2-100 gave positive results for ten samples. The TmpA2 kodecytes missed the TPPA borderline positive samples at lower concentrations, suggesting the sensitivity of the TmpA2 kodecytes is better, at higher concentrations, and TmpA2 performs well at 100 µM concentrations compared to lower concentrations. This finding is similar to what is indicated by the EIA screen test results, but the specificity is lower at higher concentrations. Out of the twelve TPPA negative samples and EIA screen positive samples, the kodecytes TmpA2-100 and TmpA2-50 showed 12 false positive results, TmpA2-30 showed nine false positive results and TmpA2-10 showed 7 false positive results. Out of the twelve TPPA negative samples tested, three were RPR positive. Because the TmpA2 kodecytes gave positive results, there is a possibility of kodecyte assay results being true positive, but no evidence either from the clinical details or the follow up test results are available to prove this. Overall, the TmpA3 kodecytes have less specificity at higher concentrations and have less sensitivity at lower concentrations; the optimum concentration is 30 µM, which is similar to what was seen in the previous TmpA3 peptide results (Table 16).

Comparison of TmpA3 kodecytes with TPPA results

The 211 EIA syphilis screen positive results and the syphilis confirmatory test TPPA results were compared to access the performance of the syphilis TmpA3 kodecytes assay. The results of the comparision of the TmpA3 kodecytes with the confirmatory test (TPPA) are shown in Table 19.

Table 19 Comparison of TmpA3 peptide results with TPPA results

FSL TmpA3 (µM)		TPPA results (n=211)				
		Positive (P)	Borderline (B)	P+B	Negative	Total
TmpA3-100	Positive	187 (99.5%) †	11	198	12	210
	Negative	1	0		0	1
TmpA3-50	Positive	184 (98%) †	10	194	12	206
	Negative	4	1		0	5
TmpA3-30	Positive	178 (95%) †	10	188	12	200
	Negative	10	1		0	11
TmpA3-10	Positive	149 (79%) †	5	154	10	164
	Negative	39	6		2	47

†: Percentage sensitivity of the TmpA3 kodecytes for TPPA confirmed positive samples.

The kodecytes TmpA3-100 showed positive results for 187 samples, TmpA3-50 showed positive results for 184 samples, TmpA3-30 showed positive results for 178 and TmpA3-10 showed positive results for 149. The percentage correlation of the TmpA3 kodecytes versus TPPA results were 99.5%, 98%, 95% and 79%, respectivel. As seen in the above table, a total of 11 samples showed borderline positive results with the TPPA test and the TmpA3 kodecytes picked up all the 11 TPPA borderline positive samples. At 50 and 30 µM concentrations, the TmpA3 kodecytes picked up ten TPPA borderline samples and, at 10 µM concentrations, five TPPA borderline positive samples were identified. TmpA3 peptide showed 12 false positive results at 100 µM, 50 µM and 30 µM concentrations and 10 false positive results at 10 µM concentration, a similar number of positive results as seen in TmpA2 and TmpA1 kodecytes. It is seen from the resukts that theTmpA3 kodecytes perform well at higher concentrations.

Receiver operating characteristic (ROC) curve analysis of TmpA1 kodecytes

Receiver operating characteristic curve (ROC) illustrates the sensitivity, specificity and accuracy of the syphilis peptides based on the validation results obtained when tested against the known

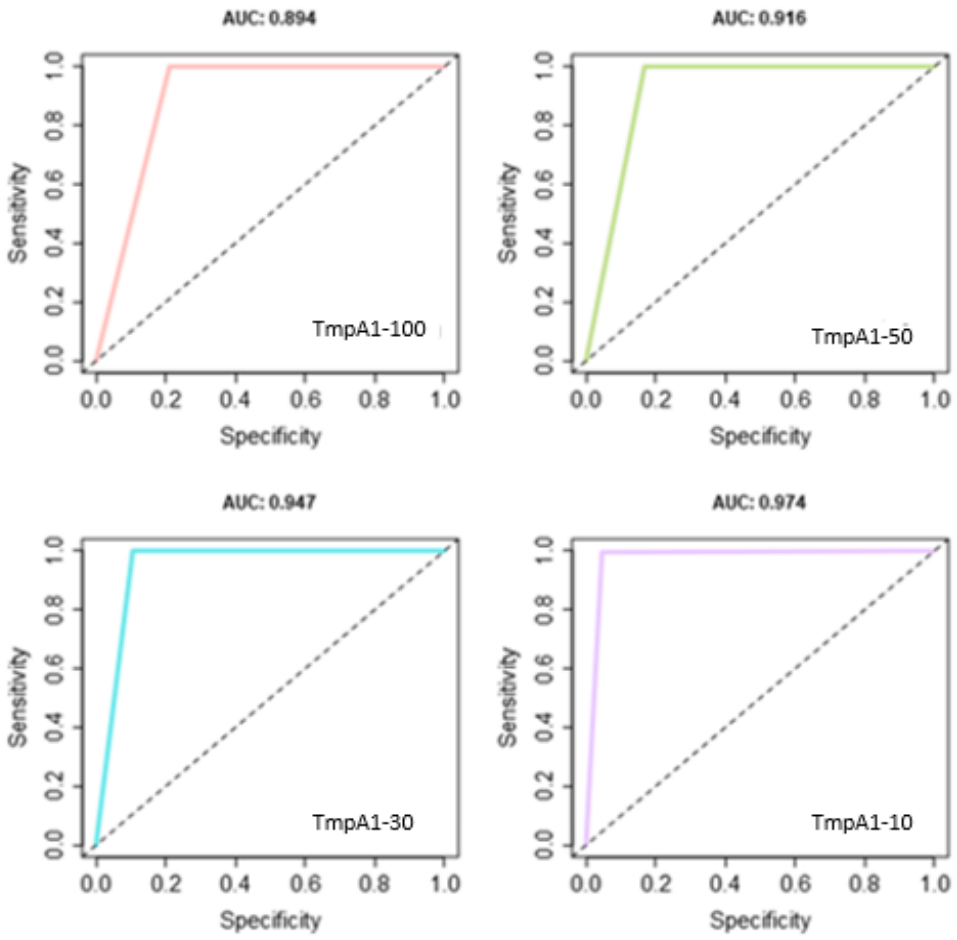

Figure 22: ROC analysis of TmpA1 kodecytes at various µM concentrations

the known positive and negative samples. Figure 22 shows the ROC curve for the TmpA1 kodecytes. Sensitivity (true positive) is shown in the y axis. Specificity (true negative) is shown in the x axis. Sensitivity is best at 1.0 and the specificity is best at 0.

The area under the ROC curve (AUC) encapsulates the accuracy of the diagnostic test. It has the value 0 to 1, where 0 means the test is inaccurate and 1 indicates the test is highly accurate. For use as a diagnostic test, ideally, the result of the ROC analysis should be very close to 1. If the ROC curve is more closely aligned to the upper left corner, the overall accuracy of the test will be high[200].

As shown in Figure 22, the AUC values for the TmpA1 kodecytes TmpA1-100, TmpA1-50, TmpA1-30 and TmpA1-10 are 0.894, 0.916, 0.947 and 0.974, respectively. Of the kodecytes, TmpA1-10 has the highest AUC value of 0.974, and the top of its curve is closer to the upper left corner than those of the other three peptide concentrations. TmpA1 kodecytes have good sensitivity at higher concentrations, but the specificity decreases at high concentrations. At 10 µM, it has both good sensitivity and specificity. The AUC value of the TmpA1 kodecytes at 10 µM is a good candidate to be used in syphilis diagnostic assay in terms of both sensitivity and specificity.

ROC curve analysis of TmpA2 kodecytes

The ROC curve analysis and the AUC values of the TmpA2 kodecytes at various concentrations are shown in Figure 23. Sensitivity (true positive) is shown in the y axis. Specificity (true negative) is shown in the x axis. The AUC values for the TmpA2 kodecytes TmpA2-100, TmpA2-50, TmpA2-30 and TmpA2-10 are 0.94, 0.93, 0.92 and 0.87, respectively. TmpA2 kodecytes TmpA2-100 has an AUC of 0.94, and the top of the curve is closer to the upper left corner than those of the other three concentrations. TmpA2 kodecytes perform well at higher concentrations. The sensitivity

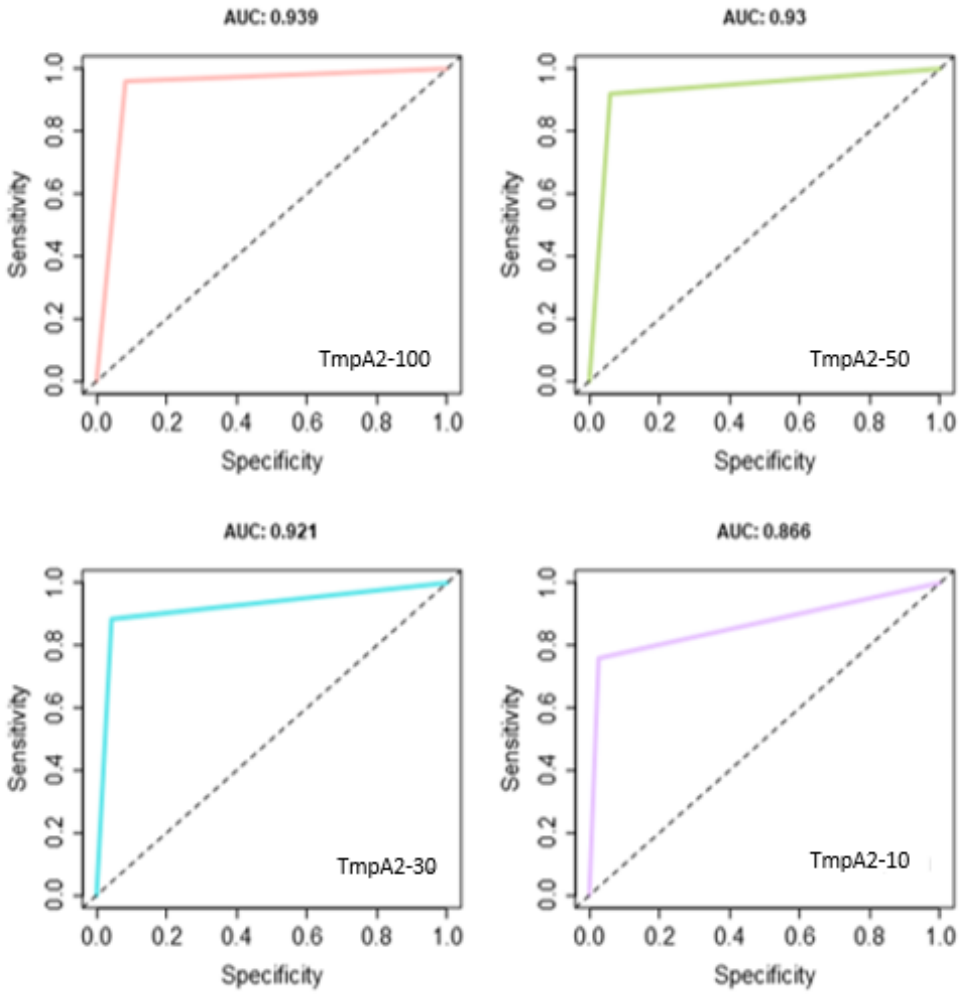

Figure 23: ROC curve analysis of TmpA2 kodecytes at various µM concentrations

and specificity of the TmpA2 kodecytes are better at 50 µM concentrations compared to the sensitivity and specificity at the other three concentrations, has an AUC value of 0.93.

ROC curve analysis of TmpA3 kodecytes

The ROC curve analysis for TmpA3 kodecytes at various concentrations and the AUC values for the TmpA3 kodecytes are shown in Figure 24. Sensitivity (true positive) is shown in the y axis. Specificity (true negative) is shown in the x axis. The value of AUC should be 1 or close to 1 for the test to be highly accurate. In other words, the top of the curve should be close to the upper left corner. The AUC values for the TmpA3 kodecytes TmpA3-100, TmpA3-50, TmpA3-30, TmpA3-10 are 0.92, 0.94, 0.95 and 0.87, respectively.

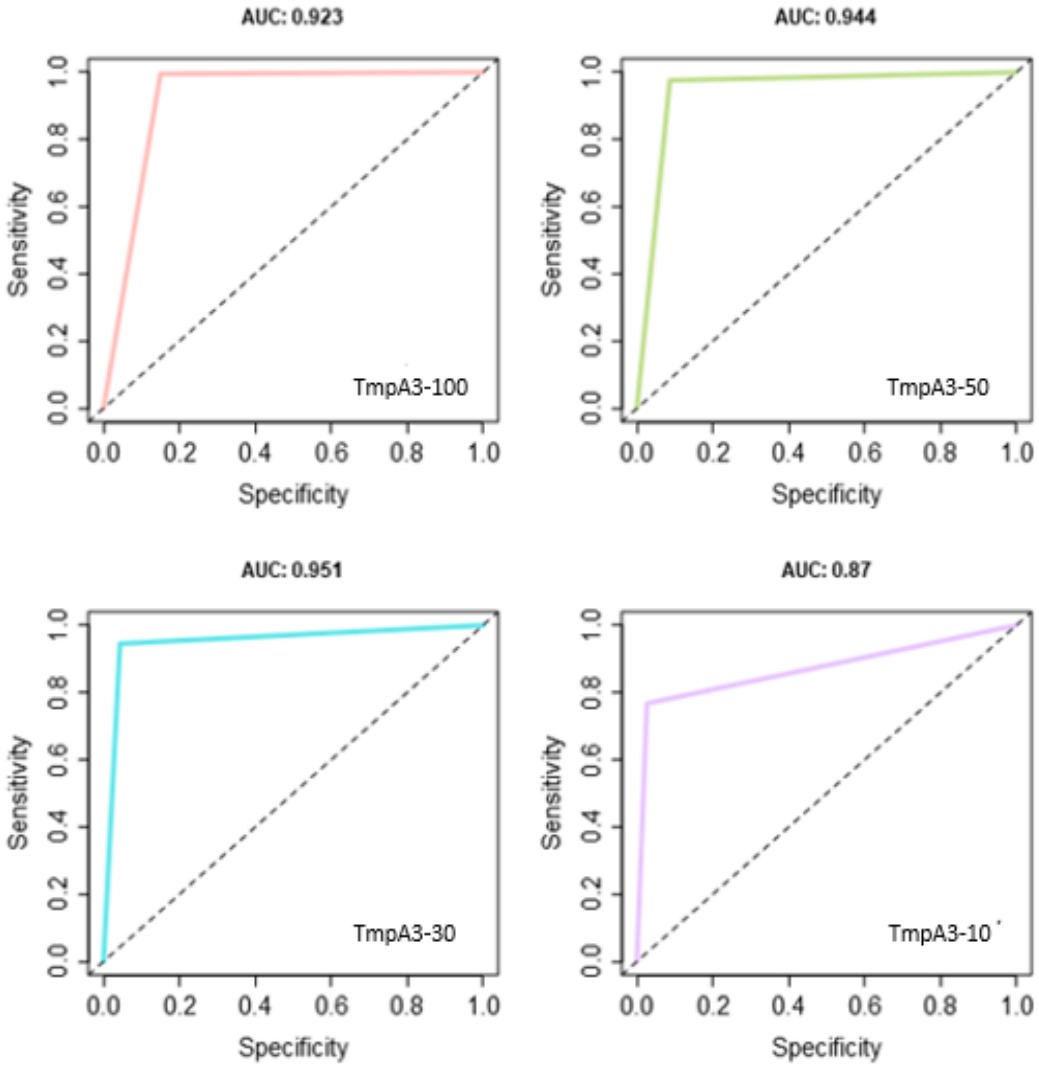

Figure 24: ROC analysis of TmpA3 kodecytes at various µM concentrations

0.94, 0.95 and 0.87, respectively. The sensitivity and specificity of the TmpA3 kodecytes are better at 30 µM concentrations compared to the sensitivity and specificity of the other three concentrations, and has an AUC value of 0.951. Although the kodecytes have good specificity at 10 µM concentration, the peptide's sensitivity is low at 10 µM concentration.

Performance of syphilis kodecytes with WHO/CDC EQA samples

Syphilis kodecytes were tested with WHO/CDC syphilis EQA samples. EQA samples are sent once a year to participating laboratories throughout the world as part of the quality control program. For this study, the EQA syphilis samples received from WHO/CDC as part of the annual testing program and then tested and stored at -80° C in the clinical laboratory were used for validation purposes along with the de-identified syphilis positive and negative samples. An aliquot of the EQA samples was used to test the performance of the syphilis kodecytes. The ORTHO BioVue column agglutination method described earlier was used to test undiluted EQA syphilis quality control samples to all three syphilis kodecytes. Table 20 shows the performance of syphilis kodecytes at four different concentrations when tested against WHO/CDC Syphilis EQA (External Quality Assessment) samples QC1, QC4, QC5 and QC7.

Table 20. Syphilis kodecytes results for WHO/CDC Syphilis EQA samples

EQA No		Kodecyte concentration (µM)				WHO/CDC syphilis EQA results††			
		100	50	30	10				
	TmpA1	Kodecyte grading				EIA result	EIA score	RPR result	TPPA result
QC1		4+	4+	3+	**3+**	4.9	P	P2	R
QC4		2+	2+	2+	**2+**	4.9	P	N	R
QC5		2+	2+	2+	**2+**	5.7	P	P1	R
QC7		4+	4+	4+	**3+**	89.5	P	P8	R
	TmpA2								
QC1		2+	**2+**	2+	1+	4.9	P	P2	R
QC4		1+	**1+**	1+	1+	4.9	P	N	R
QC5		1+	**1+**	1+	1+	5.7	P	P1	R
QC7		2+	**2+**	2+	1+	89.6	P	P8	R
	TmpA3								
QC1		3+	2+	**2+**	1+	4.9	P	P2	R
QC4		1+	1+	**1+**	1+	4.9	P	N	R
QC5		1+	1+	**1+**	1+	5.7	P	P1	R
QC7		2+	2+	**2+**	1+	89	P	P8	R

††: EIA, RPR and TPPA results in the above table are the expected results for the EQA samples as provided by CDC in the EQA result analysis report. The best ROC concentrations of the syphilis peptides are shown in boxes.

The syphilis kodecytes results were compared with the copy of the WHO/CDC EQA test results received in the clinical laboratory. TmpA1 performed well at a lower concentration of 10 µM with good sensitivity and specificity. TmpA2 and TmpA3 are less sensitive than TmpA1 at 10 µM

concentration, as seen from earlier results. Still, they perform well at a higher concentration. All three syphilis FSL peptides passed the external QC program for syphilis testing, and the EQA panel results show that the syphilis FSL peptides are suitable for diagnostic use. The best concentrations for the syphilis FSL peptides TmpA1, TmpA2 and TmpA3 to be used for further clinical validation studies are 10 µM, 50 µM and 30 µM, respectively.

Verification of analytical specificity of syphilis kodecytes

To verify the analytical specificity of the syphilis peptides, syphilis kodecytes TmpA1-100, TmpA1-50, TmpA1-30, TmpA1-10, TmpA2-100, TmpA2-50, TmpA2-30, TmpA2-10, TmpA3-100, TmpA3-50, TmpA3-30 and TmpA3-10 were tested against 157 samples seroreactive for other pathogens and samples seroreactive for autoantibodies (Table 21).

Table 21. Verification of analytical specificity of syphilis kodecytes

Category of samples	No of samples	Syphilis kodecytes No reactive
Measles	4	0
Mumps	6	0
Rubella	2	0
Hepatitis B	16	1 †
Hepatitis C	10	0
Hepatitis D	1	0
Hepatitis E	1	0
Varicella Zoster virus	9	0
Epstein Barr virus	13	0
Cytomegalovirus	31	0
Bordetella	5	0
Parvovirus	7	0
Toxoplasma	2	0
Autoantibodies	50	0

†: All three kodecytes showed one positive reaction out of the 16 hepatitis B samples.

A total of 157 samples, which were seronegative for other pathogens, were tested with syphilis kodecytes and 156 samples gave negative results except one sample. A total of 16 hepatitis B seropositive samples were included in the analytical sensitivity experiment. Out of the 16 hepatitis B seropositive samples tested, only one sample was seropositive with syphilis kodecytes. The results suggests that syphilis kodecytes have excellent specificity.

3.7.3. Conclusion and summary

This study successfully designed and built three syphilis kodecytes (TmpA1, TmpA2 and TmpA3) using the FSL peptide selection algorithm. The proof of concept for the syphilis kodecytes design, sensitivity and specificity of the designed kodecytes was validated using known syphilis positive

and syphilis negative samples pre tested in a clinical laboratory. The analysis of the syphilis kodecytes results when tested against known pre tested clinical samples, as a proof of concept, show that the syphilis kodecytes design worked. Syphilis kodecytes, when tested against syphilis EIA, screened positive and negative samples (Table 22) and showed that the TmpA1 kodecytes perform well at 10 µM concentration with a sensitivity of 98.6% and a specificity of 98%. The syphilis TmpA2 kodecytes perform well at 50 µM concentration with a sensitivity of 92.4% and a

Table 22 Sensitivity and specificity of syphilis kodecyte assay with syphilis EIA samples

Peptide	Syphilis EIA positive samples				Syphilis EIA negative samples			
	Peptide concentration (µM)				Peptide concentration (µM)			
	100	50	30	10	100	50	30	10
	Sensitivity (%)				Specificity (%)			
TmpA1	100	100	100	98.6	70	79	89	98
TmpA2	96.2	92.4	89	75	93.4	97	99.4	99.4
TmpA3	99.5	97.6	94.7	77.2	84.7	91.8	97.9	100

specificity of 97%. The syphilis TmpA3 kodecytes perform well at 30 µM concentration with a sensitivity of 94.7% and a specificity of 97.9%. It is evident from the results shown in Table 22 that the syphilis TmpA1, TmpA2 and TmpA3 kodecytes have good sensitivity and specificity at 10 µM, 50 µM and 30 µM concentrations, respectively.

The ROC and AUC analysis shown in Figures 22 to 24 shows that the optimum concentration for good sensitivity and specificity for syphilis kodecytes are TmpA1 - 10 µM, TmpA2 - 50 µM and TmpA3 -- 30 µM concentrations. The results from the ROC and AUC analysis show that the syphilis TmpA1 kodecytes perform well with good sensitivity and specificity and at a lower concentration (10 µM).

Syphilis TmpA1 kodecytes picked up the 11 TPPA borderline positive samples with good serological grades, showing better sensitivity than TPPA, the current standard confirmatory test, when compared to results for TmpA2 and TmpA3 kodecytes (Tables 17 to 19). Three of the 11 TPPA borderline positive samples were also RPR test positive, suggesting that the syphilis TmpA1 kodecytes are more sensitive than the TPPA test.

Syphilis TmpA1 kodecytes gave positive results for 12 TPPA negative samples which were EIA syphilis screen test positive. Three of the 12 TPPA negative samples were both EIA syphilis screen and RPR test positive. The TPPA assay is known for producing false negative results[29]. The caveat

of this study is that there should be no access to the clinical information and follow up details of the positive samples used to test the syphilis kodecytes, to see whether syphilis TmpA1 kodecytes results were true positive or false positive results.

The external syphilis quality control samples received from WHO/CDC and tested and stored at 80° C in a clinical laboratory were used in this study to evaluate the performance of the syphilis kodecytes. The results show (Table 20) that all three peptides identified the external quality control samples from WHO/CDC correctly, validating the use of syphilis peptides in diagnostics

A total of 157 samples seronegative for other pathogens and samples seronegative for autoantibodies were used to verify the analytical specificity of the syphilis peptides (Table 21). All the three syphilis kodecytes showed very good analytical specificity (99.3%).

To summarise, the proof of concept results for the syphilis kodecytes (TmpA1, TmpA2, and TmpA3) show that the FSL design algorithm is working. TmpA1 peptide has good sensitivity at higher concentrations but the specificity decreases at high concentrations. At 10 µM, TmpA1 has both good sensitivity and specificity, as well as good accuracy. TmpA2 and TmpA3 peptides have good specificity at lower concentrations, but the sensitivity for both is very low. TmpA3 peptides perform well at 30 µM concentration. TmpA2 peptides have better results at 50 µM concentration and TmpA2 peptides have lower sensitivity and specificity when compared to TmpA1 and TmpA3. The reason that TmpA1 kodecytes perform consistently well at 10 µM concentrations with good sensitivity and specificity at above 98% may be that they have high antigenicity scores compared to TmpA2 and TmpA3. This might explain the lower sensitivity seen in TmpA2 and TmpA3 when compared to TmpA1. The fact that all three syphilis peptides passed the external quality samples at different peptide concentrations shows that the FSL peptide selection algorithm needs more fine-tuning.

Future work

This study has successfully designed syphilis FSL peptides as per the study design and protocol for the intended purpose of developing syphilis diagnostics. This study has so far completed stage1 of any assay development pathway (analytical characteristics), that is, comparing the syphilis kodecyte assay with existing standard methods, checking the analytical sensitivity and specificity. This study has also completed stage 2 of any assay development pathway, that is, determining diagnostic characteristics by using reference samples such as WHO EQA samples, known positive and negative samples and seropositive samples from other pathogens for checking the cross reactivity and diagnostic specificity.

The caveat of this study is the lack of samples representing different stages of the disease.

Initial validation of the syphilis FSL peptides suggests that it is feasible to use the TmpA1 peptide at 10 µM concentration for high throughput automation of screening assays as well as a confirmatory tests. To fully validate the syphilis kodecytes assay for the intended purpose (stage 3 of any assay development), a further large scale diagnostic trial must be done, with samples representing different stages of the disease, with defined evaluating panel, clinical follow up, regulatory approval and assay optimisation.

Chapter 4. Methods and Results: Leptospiral diagnostic using Kode FSL peptides

4.1. Literature review of *Leptospira* proteome

The pathogenesis of *Leptospira* at the molecular level is still unclear, despite significant progress made to determine its origin and development over the years. Researchers have identified several virulence factors like Lipopolysaccharide (LPS), outer membrane proteins, adhesion molecules and surface proteins that might play a role in the pathogenesis of this spirochete.

Several OMPs of *Leptospira* have been identified and this focus has been on the outer membrane proteins for the following reasons. LPS is serovar specific and more than 300 serovars are reported in the literature, and the immunity elicited will be against homologous strains. In contrast, the protein antigens are relatively conserved in the OMP, in addition to the broad cross reactivity, will also be serovar specific. Hence the interest in the OMPs in the field of diagnostics and vaccine research. Another factor is that more than half the bacterial outer membrane mass contains the leptospiral outer membrane proteins. These proteins play a crucial role in the host pathogen interaction due to their location in the leptospiral architecture. They also help in differentiating the pathogenic leptospires from non-pathogenic leptospires.

There are three groups of leptospiral OMPs (Table23): lipoproteins, transmembrane proteins and surface proteins[201,202].

Table 23. Leptospiral outer membrane proteins

Leptospiral outer membrane proteins	
Lipoproteins	LipL21, LipL32, LipL36, LipL40, LipL41, LipL46, LipL48, LipL53, LigA, LigB, LigC, Loa22, LruC, Lsa21, Lsa27, LepA, LenA-F, Lp29, Lp49, MPL36, Lp49, MPL36
Transmembrane	OmpL1, OmpL36, OmpL37, OmpL47, OmpL54 Tly7, HbpA, FecA, Mce
Surface proteins	P31, LipL48

Several major outer membrane proteins have been identified through microarray studies and various complementary approaches like surface immunofluorescence, proteolysis, biotinylation and surface ELISA[203-205].

Although several immunodominant lipoproteins have been identified through various studies and experiments, little information is available about the biological role they play. The LipL32 protein accounts for 75% of the outer membrane proteome, with about 40,000 copies per cell. It interacts with several components (collagen, fibronectin and laminin) and has host extracellular binding

properties. LipL32 is found only in pathogenic leptospires, and does not have orthologs among saprophytic leptospires[88]. Even though several studies show evidence that LipL32 plays a role in the pathogenesis of *Leptospira*, unexpected results found by one of the research groups denied the role of LipL32 in pathogenesis[206].

The next abundant protein is Loa22, with about 30,000 copies per cell. It is highly immunogenic and is the only leptospiral protein that fulfils Koch's postulate. Twenty five percent of the Loa22 protein is surface exposed. There is more evidence to suggest that LipL41, the surface exposed protein, plays a role in virulence and pathogenesis. LipL41 is highly conserved antigenically among pathogenic leptospires[207] and does not have ortholog with saprophytes. It has vaccine efficacy when coadministered with OmpL and it has hemin binding properties[208]. LipL21 was initially thought to be the second most abundant protein and to be conserved among pathogenic leptospires[209]. Later study shows it has 50% similarity with *L. biflexa*[88]. Surface biotinylation study shows that LipL21 protein is surface exposed and immunogenic. There is not much evidence to indicate that proteins LipL36 and LipL48 have diagnostic or vaccine efficacy. LipL53 is another surface exposed adhesion protein that binds to the extracellular matrix and other host proteins. This protein is upregulated when there is a change in the temperature, and is a known feature of *Leptospira*, regulating the protein synthesis in the case of in vitro changes in temperature.

The Lig proteins are immunoglobulin like proteins and are expressed only during infection. All three Lig genes, LigA, LigB and LigC, encode for virulence factors and bind to a variety of host proteins. The protein LigC was found to be a pseudogene, LigB is present in all the *Leptospira* isolates and LigA is present in *L. interrogans* and *L. kirschneri*[210].

The lipoprotein LruC was identified from the eye fluid of the horses with uveitis, and is present in the inner part of the OMPs[211]. The other OMPs found on the spirochete surface are Lsa21, Lsa24, Lsa27, Lsa30, Lsa63 and Lsa66. Their role as adhesion molecules is still not very clear. The protein Lsa24 has been renamed as Len protein. There are six types of Len protein, Len A through F. These proteins are present in the pathogenic leptospires only, and they bind to the coagulation cascade, suggesting they may play a role in pathogenesis. Lipoproteins like Lp29, Lp30, Lp49, MPL17, MPL21 and MPL36 have also been identified as having a role in pathogenesis[212].

Several leptospiral transmembrane proteins have been identified and their functions, such as cell structure maintenance, substrate attachments and nutrient transport, have been elucidated. OmpL1 was the first identified transmembrane protein. It is highly conserved in pathogenic leptospires and offers very high immunoprotection. There are three types of OmpL1 protein: OmpL/1, OmpL/2 and OmpL/3[202 213]. Later silico studies revealed several other transmembrane proteins, including OmpL36, OmpL37, OmpL47 and OmpL54[204]. OmpL37 protein binds to the

human skin and facilitates the entry of the spirochete[203]. TlyC protein is thought to be found only in pathogenic strains and has 58% homology with the saprophytic leptospires.

The periplasmic flagellar proteins FlaA and FlaB have been identified through genome sequencing. The study conducted shows that, compared with the MAT, they have a positive and negative predictive value as a biomarker of 81.7% and 83.3%, respectively.[214]. The FlaB proteins have similarities not only to other spirochetes but also to other Gram positive and Gram negative bacteria[215].

Leptospira microarray studies show around 18 antigens, but only a few of these antigens have high sensitivity and specificity. LigA and LigB proteins ranked first in acute phase patients, followed by LipL32. In convalescent patients, LipL32 came first in terms of sensitivity and specificity, followed by Lig proteins. The novel protein LIC10215 ranked next to LigA, LigB and LipL32. The heat shock protein GroEL had high sensitivity, but its specificity was very low both in acute and convalescent patients[216].

4.2. Rationale for choosing *Leptospira* proteins and FSL peptide design

Although several leptospiral proteins, and in particular, their outer membrane proteins, have been studied extensively, their role in diagnostics is still problematic. Challenges include protein selection and the existence of more than two hundred and fifty circulating leptospiral serovars, given that and no single assay will pick up all serovars.

There is a need for a diagnostic assay which will detect the acute phase of illness, for an assay to precisely identify the convalescent phase of illness and for sero epidemiological studies with good sensitivity and specificity.

The aim is to first select the proteins that can be used to diagnose both the acute and convalescent stages of illness. The second step is to use Kode technology to design and build the best leptospiral FSL peptide constructs and test them for sensitivity and specificity. Once the functionality of the leptospiral FSL peptides is established, the next step is to use the multiple functional head concept of Kode to make leptospiral FSL constructs with multiple epitopes designed from the locally circulating strains for any given area. This approach will not only improve the sensitivity and specificity of the assay but also be a good solution to current problems facing leptospiral serological assays due to the large number of circulating serovars.

Rationale used for the selection of proteins for FSL peptide design:

- Not present in non-pathogenic *Leptospira*
- Conserved among other leptospires

- Present in the acute or convalescent phases of illness
- Cited in the literature

Protein selection:

- The immunoglobulin like proteins LigB and LigA were chosen because they are found only in pathogenic leptospires. Microarray studies show that antibody to Lig proteins is detected in samples collected during the acute phase of illness, which indicates it will be a good candidate for IgM detection.
- The LipL32 protein is heavily expressed, and it is conserved among pathogenic leptospires. LipL32 is highly immunogenic and expressed both in the acute and convalescent phases of leptospirosis.
- The LipL41, another outer membrane protein, plays a vital role as a determinant of virulence, and is expressed during leptospiral infection. Lipl41 is highly conserved among pathogenic leptospires. It is expressed in both the acute and convalescent phases of leptospirosis, according to microarray studies[217].
- The hypothetical protein LIC10251 can detect and differentiate acute and convalescent cases from healthy individuals. It is ranked next to Lig proteins[216].
- LipL21 is a highly expressed leptospiral protein found in pathogenic leptospires only. It is detected in the patient's immune sera[209] and also used in PCR to differentiate pathogenic from non-pathogenic (saprophytic) leptospiral serovars.
- The importance of the protein Loa22 in pathogenesis is well documented, and it also plays a role in virulence. Loa22 is a pathogen specific protein that helps the spirochete adhere to and penetrate the host cell.
- The periplasmic protein FlaB reacts with sera from leptospirosis patients detecting both IgM and IgG antibodies. It is a highly conserved protein[218]. It has a significant role in diagnostics and vaccine development[214].

As seen in the literature, even though 14 to 18 leptospiral antigens are detected in microarray studies using patient sera, only a handful of these antigens have been widely tested. Also, the sensitivity of the assays with individual proteins is low, and results have shown that more than one recombinant protein improves the sensitivity of the leptospiral serodiagnostic assay.

The aim of the study is to design an FSL peptide construct for leptospiral diagnostics from the selected proteins.

4.3. Method and Results: analysis of leptospiral protein sequences

The leptospiral protein sequences used in this study to design the peptides were derived from the NCBI database (http://ncbi.nlm.nih.gov). Multiple amino acid sequences from different serovars and strains of each leptospiral protein were obtained.

The following reference sequence for each leptospiral protein was selected. LIC10215: WP_004459793, LipL41: WP_001228945, LipL21: WP_000610520, Loa22: WP_002634744, Flab: WP_002618149, LipL32: WP_000736494, LigA: WP_000000000, LigB: WP_017862321.

4.3.1. Structure prediction using I-TASSER/NovaFold

The leptospiral protein structure prediction was done using the I-TASSER/NovaFold software. The predicted 3D structures of the leptospiral proteins are shown in Figure 26. Based on simulation

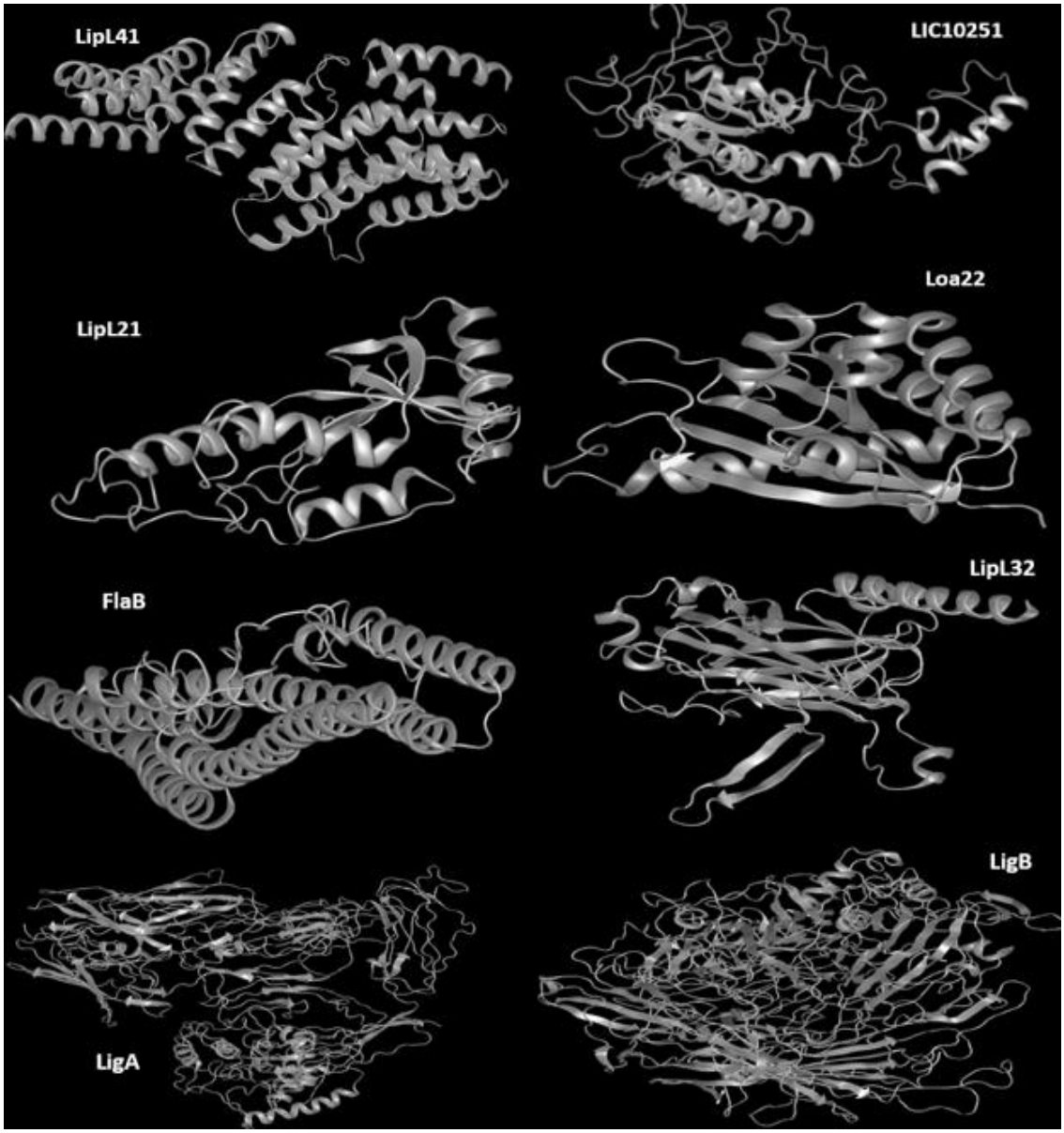

Figure 25: I-TASSER/NovaFold predicted structures of leptospiral proteins

and alignment of the threading pattern, the program calculates the C-score. The confidence of each model is quantitatively measured by the C-score, which is typically in the range of (-5, 2); a C-score of a higher value signifies a model with a higher confidence. The calculated confidence score (C-score), of the predicted leptospiral proteins are as follows: Lip41 (-1.15), LIC10215 (-4.93), Lipl21 (-3.71), Loa22 (-2.56), Flab (0.44), LipL32 (0.46), LigA (-0.27) and LigB (-0.79). The secondary structure predictions of the leptospiral proteins by I-TASSER/NovaFold/ protean 3D program showed fewer strands and more helices and coils in LipL41, Loa22 and FlaB, whereas LigA and LigB proteins showed few helices and more strands and coils. The leptospiral protein Lipl32 has more beta sheets followed by coil, with very few helices.

4.3.2. Structure analysis using the protein structure analysis program (ProSA)

The I-TASSER/NovaFold predicted structure was further analysed with the ProSA program. The ProSA program calculates the quality score (Z score) by comparing the query model with all the experimental structures available in the protein database. If this score is outside a range characteristic for native proteins, it indicates that the structure probably contains errors. The number of residues is plotted on the x axis, while the Z score is plotted on the y axis (Figure 26). The user input structure score is circled in red. The Z score of the input structure should be within the range of scores obtained in native proteins of similar size.

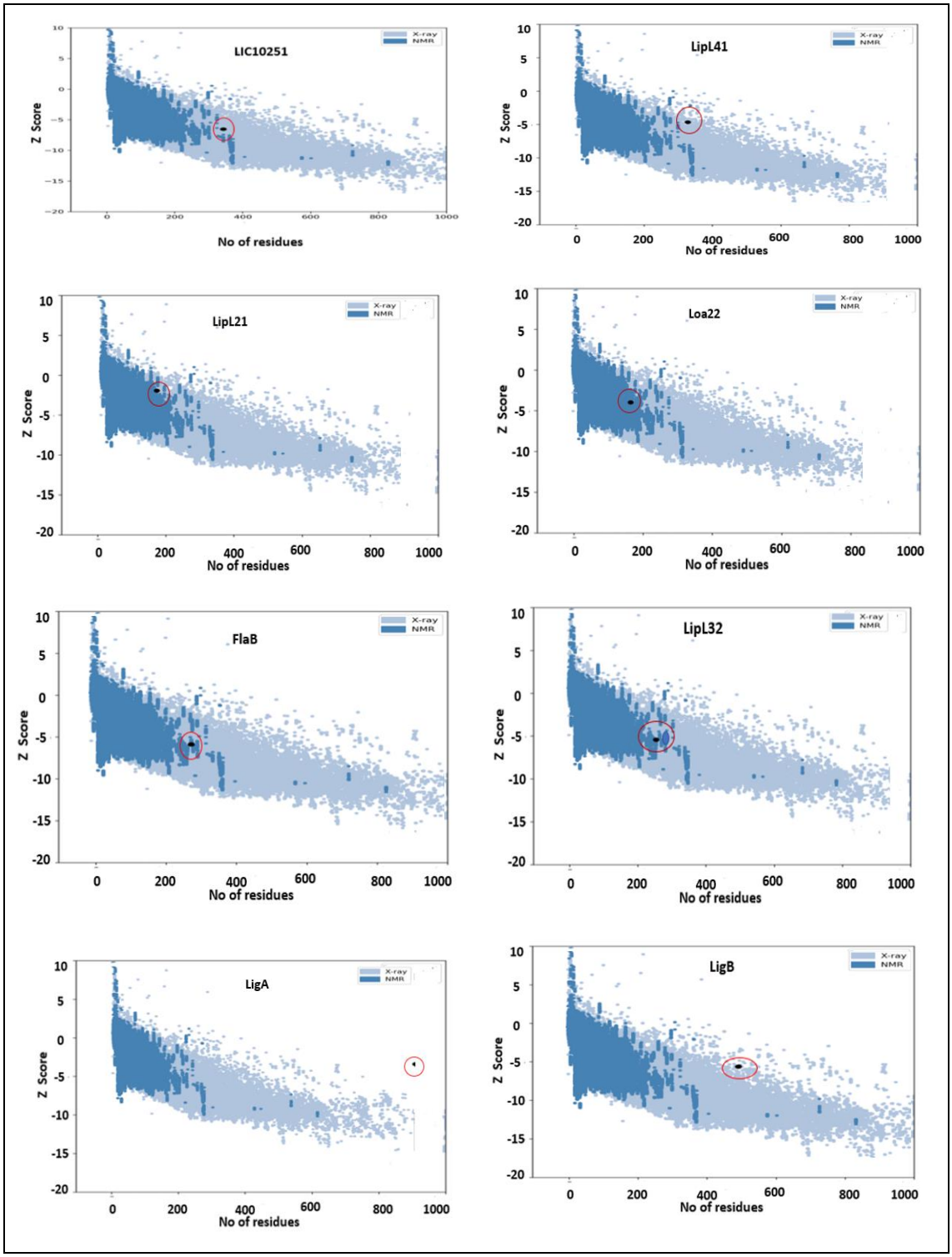

Figure 26: Z score plot for predicted model of the leptospiral proteins

The ProSA structure analysis is shown in Figure 26. The number of residues is plotted on the x axis, while the Z score is plotted on the y axis. The plot shows dots which are the scores of all the currently available protein structures in the protein database. The light blue dots are the protein structures determined by X-ray crystallography, and the dark blue are the structures determined by the NMR spectroscopy method. The user input structure score position is circled in red.

The Z scores of the leptospiral proteins are as follows: LipL32 (-6.59), LipL41 (-4.1), LipL21 (-2.65), Loa22 (-4.83), LIC10251 (-6.52), LigB (-4.13), Flab (-6.44) and LigA (-4.67). All the leptospiral protein structures fall within the Z score range (circled in red) characteristic for native protein structures determined by both the X-ray crystallography and the NMR spectroscopy methods, suggesting no error in the predicted structures.

4.3.3. Structure analysis using the self-optimized structure prediction method (SOPMA)

The protein's secondary structural features were analysed using the SOPMA[172] software with default settings. The secondary structures of the leptospiral proteins (Table 24) identified using the SOPMA

Table 24. Secondary structures of leptospiral proteins (SOPMA)

Secondary structure	Leptospiral proteins (%)							
	LIC10215	LipL41	LipL21	Loa22	FlaB	LipL32	LigA	LigB
α helix	39	52	37	30	69	33	6	6
Extended strand	23	15	19	17	8	18	41	36
β turn	9.4	4	7.5	6.1	1.8	7	5.1	6
Random coil	29	30	37	48	20	41	48	52

software revealed that they are dominated by α-helix and the coil followed by extended strand and β-turns, except for immunoglobulin-like proteins LigA and LigB, which have higher percentages of extended strands than the other proteins. LigA and LigB have a greater percentage of random coil structures. Overall, the selected proteins have both a good percentage of the helices and random coils.

4.4. Methods and Results: Identification of B cell epitopes

For this study, due to cost limitations, it was decided to design one peptide from each protein to build as FSL leptospiral peptide constructs.

4.4.1. Epitope prediction

Method overview

As discussed earlier, the B cell epitope prediction programs ABCPred, BCpred, BepiPred-2.0, LBtope and DNASTAR Protean 3D were used for the predictions. The detailed methodology is explained in the syphilis chapter. The leptospiral protein sequences were submitted to the B cell prediction programs.

Results

The candidate peptide epitope sequences consistently predicted by all the methods (ABCpred, BCPred, BepiPred-2, LBtope and DNASTAR protean 3D - B cell epitope prediction) were chosen and were considered as potential epitopes (Table 25).

Table 25 List of predicted leptospiral peptides

Peptide name	Peptide sequence†	Peptide length
LIC10251	VIKEGKDENQGIGYL	16
LipL41	ATGKDVNTGNEPVSKPTG	18
LipL21	QRNDGKTPRDTNPK	14
Loa22	TDAIGPEQAEGAKK	14
FlaB	VKADGRPIAISSPGEA	16
LipL32	IPNPPKSFDDLKNIDTK	17
LigA	DSKVVSISNSNDDRGL	16
LigB	SFDQENLQSSPKDRIN	16

†Peptide sequences identified as epitopes by ABCpred, BCPred, BepiPred—2.0, and LBtope B cell epitope prediction software.

A total of eight leptospiral peptides, one from each protein, were selected and they were further analysed for their suitability as FSL constructs.

4.5. Methods and Results: Analysis of the predicted *Leptospira* peptides

The structures and the biophysical properties of the predicted peptides were analysed for their suitability to be built as FSL constructs as per the FSL selection algorithm table described earlier in chapter 2.

4.5.1. Structural analysis of the predicted leptospiral peptides

A. Structural analysis of protein using DNASTAR Protean 3D program

The predicted leptospiral peptide sequences were analysed in the protean 3D program to locate and visualise the secondary structures of the predicted peptides. The location of the predicted leptospiral peptides on the full parent proteins are marked and shown in green and yellow in the solid view of the proteins (Figure 27) and they are exposed as seen in the solid view of the protean 3D structure analysis.

B. Structure prediction of the leptospiral peptides using I-TASSER/NovaFold

The predicted leptospiral peptide sequences were analysed using I-TASSER/ NovaFold software and the probable solvent confirmations of the peptide fragments and their structures are shown

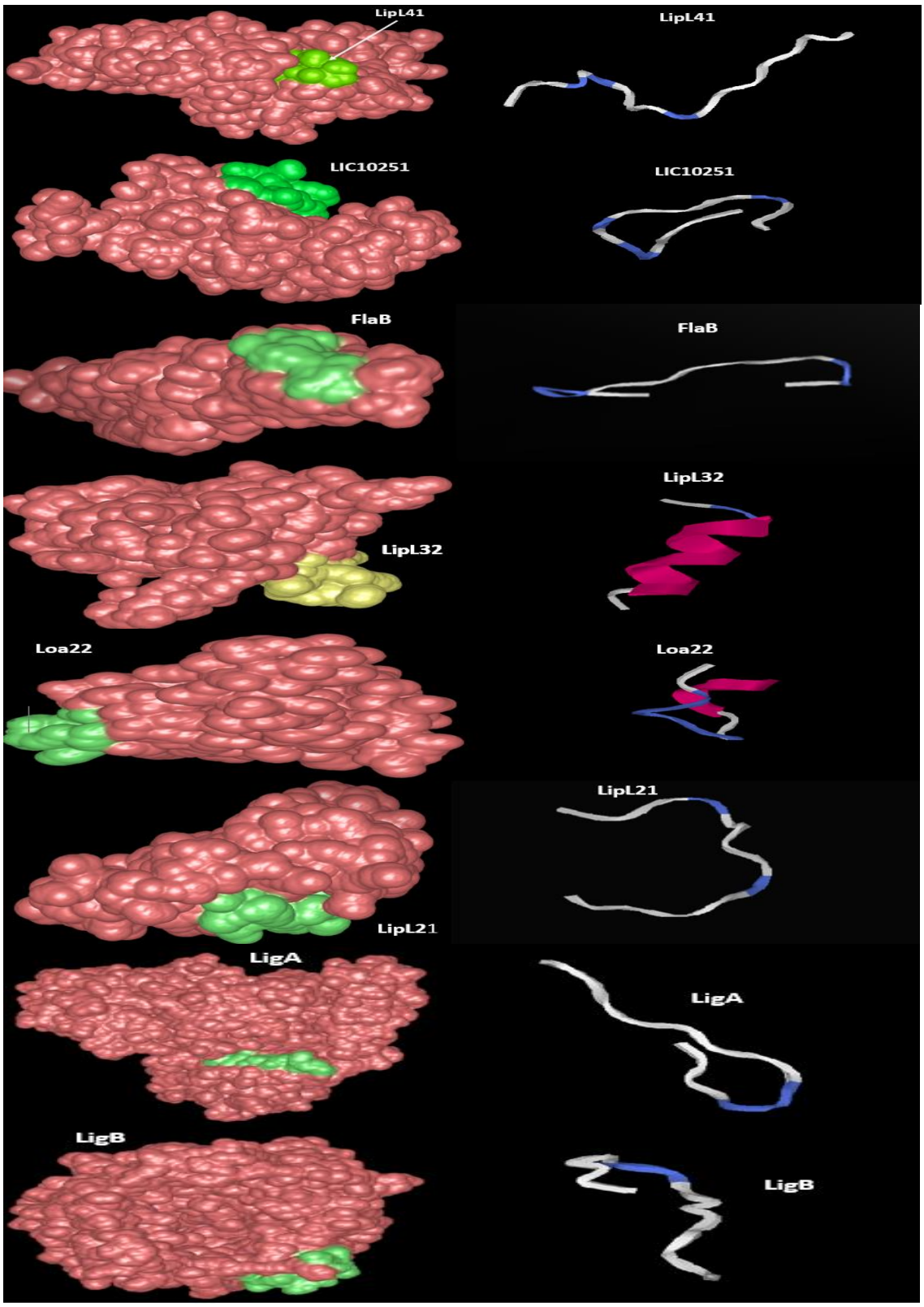

Figure 27: Location of the selected peptide sequences in leptospiral proteins

on the right side of Figure 27. The I-TASSER/NovaFold predicted structures of the peptides LipL41, LIC10215, Flab, LipL21, LigA and LigB are coils, and the predicted structures of LipL32 and Loa2 are coil/helix.

4.5.2. Results: Theoretical prediction of the leptospiral peptides

Theoretical predictions of the leptospiral peptides, using various parameters and the results obtained, are shown in Table 26a and Table 26b. Table 26a shows the epitope predictions, basic

Table 26. Results of the theoretical predictions of the leptospiral FSL peptides

26a Epitope prediction								
Peptide details	LIC10251	LipL41	LipL21	Loa22	FlaB	LipL32	LigA	LigB
Basic structure	Coil	Coil	Coil	Coil/ helix	Coil	Coil/ helix	Coil	Coil
Epitope prediction								
BepiPred-2.0	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BCPred	No	Yes	No	Yes	Yes	Yes	Yes	Yes
ABCpred	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
DNASTAR	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Antigenicity								
Kolaskar/Tongaonkar	0.93	0.97	0.90	0.97	1.02	0.98	0.99	0.98
Antigen profiler	4.4	3.4	4.2	4.0	3.6	3.9	2.3	3.5
Similarity								
BLAST (%)	100	100	100	100	100	100	100	100

26b Epitope Biophysical properties								
Peptide details	LIC10251	LipL41	LipL21	Loa22	FlaB	LipL32	LigA	LigB
Internal cysteine	0	0	0	0	0	0	0	0
Terminal Glutamine	0	0	0	0	0	0	0	0
NetOGlyc	0	0	0	0	0	0	0	0
NetNGlyc	0	0	0	0	0	0	0	0
Net charge	-1	0	2	-1	0	0	-1	-1
Surface Accessibility	1	1	1	1	1	1	1	1
Isoelectric point	4.43	6.95	10.6	4.32	6.9	6.84	4.01	4.17
Parker scale	4.8	4.5	5.64	4.2	2.0	2.73	3.95	3.94
Peptide length	16	18	14	14	16	17	16	16

of the peptides. All the leptospiral peptides are surface exposed and Parker hydrophilicity scores are good (Table 26b). This is in agreement with the structural analysis of the parent protein and

the location of the predicted peptides (Figure 27). The leptopsiral peptides were analysed for antigenicity in a process similar to that used for the syphilis peptides. When tested using the Kolaskar/Tongaonkar method, all eight leptospiral peptides received good antigenicity scores. The Antigen profiler method also predicted the peptides with excellent antigen scores, except for LigA peptide, which received a moderate antigenicity score. The peptides were analysed for glycosylation sites using the NetOGlyc and NetNGlyc and found to be non-glycosylated. The peptides were also analysed using the BLAST program, and found to have good similarity: the peptides matched with the correct *Leptospira* proteins and did not match with any other microorganisms.

4.5.3. Candidates of leptospiral peptides for Kode FSL construction

The list of leptospiral peptide sequences suitable for FSL construction and their predicted peptide

Table 27. List of leptospiral peptides for FSL construction

Peptide name	FSL construction - Peptide sequence[C]†	Peptide length
LIC10215	[C]VIKEGKDENQGIGYL	16
LipL41	[C]ATGKDVNTGNEPVSKPTG	18
LipL21	QRNDGKTPRDTNPK[C]	14
Loa22	[C]TDAIGPEQAEGAKK	14
FlaB	[C]VKADGRPIAISSPGEA	16
LipL32	[C]IPNPPKSFDDLKNIDTK	17
LigA	[C]DSKVVSISNSNDDRGL	16
LigB	[C]SFDQENLQSSPKDRIN	16

[C]† is the conjugation cysteine, which has been added to the peptide and which determines the orientation to the membrane.

sequences, peptide length and the location of the peptides is shown in Table 27. A cysteine residue is added to the final peptide sequence at either the amino or carboxyl ends to conjugate the peptide to the spacer. This determines the orientation to the membrane.

4.6. *Leptospira* FSL constructs

The Leptospiral peptides, LIC10215, LipL41, LipL21 Loa22, LipL21, FlaB, LipL32 LigA and LigB, were selected using FSL peptide selection algorithm and because they were deemed suitable for construction as FSL Constructs, were built as leptospiral FSL peptide constructs.

The peptides were synthesized into FSL constructs by conjugating the functional head (peptide) with the cysteine SH group to a spacer (carboxymethylglycine), which is conjugated to the lipid phosphate moiety (DOPE). The structures of the leptospiral FSL constructs and their molecular weights are shown in Figures 29 to 36.

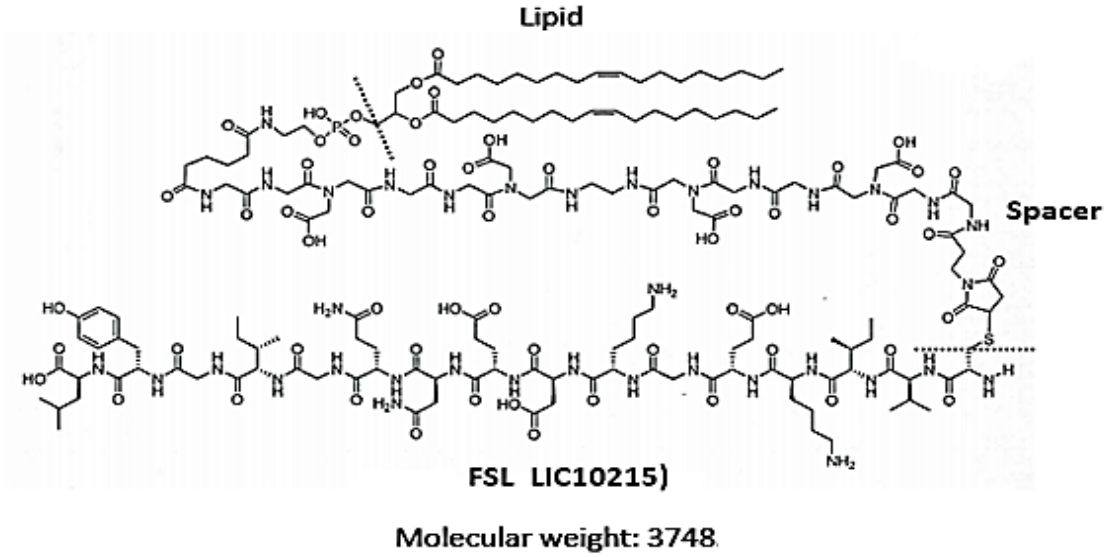

Figure 28: Schematic diagram of FSL LIC10251 construct

Figure 28 shows the function-spacer-lipid (FSL) construct consisting of a lipid phosphate moiety conjugated to the spacer, which is conjugated via a cysteine SH group to the peptide functional head (LIC10215-VIKEGKDENQGIGYL). The dotted line denotes the beginning and end of the carboxymethylated oligoglycine based spacer that separates the Function and Lipid segments. The subsequent Figures (29 to 36) feature the same structural logic and denotations.

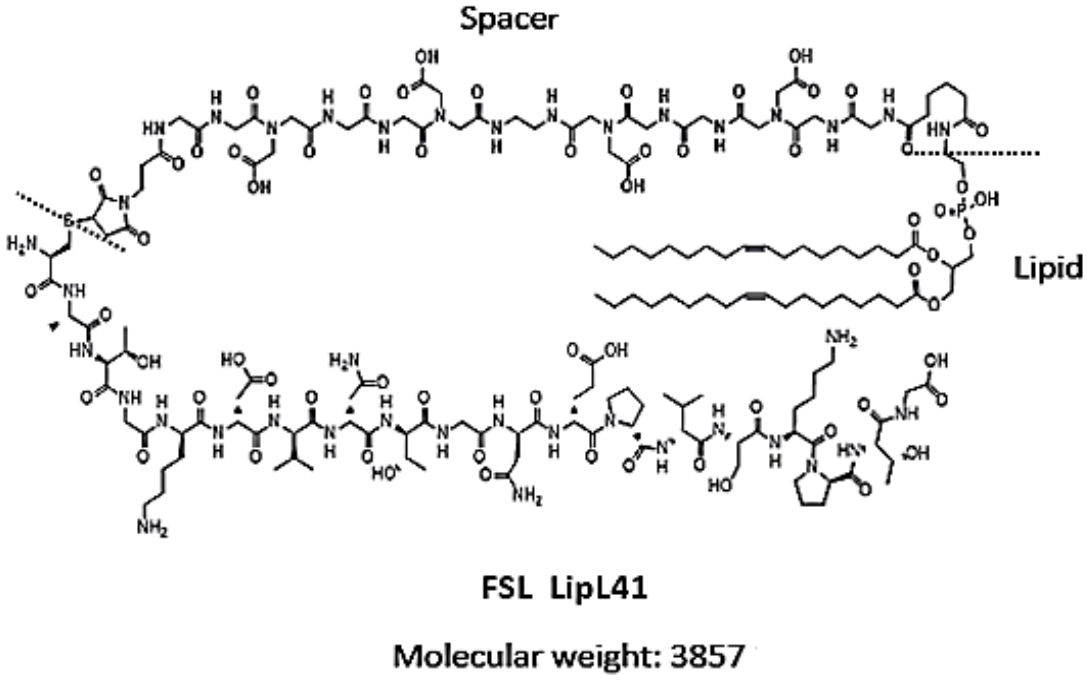

Figure 29: Schematic diagram of FSL LipL41 (ATGKDVNTGNEPVSKPTG) construct

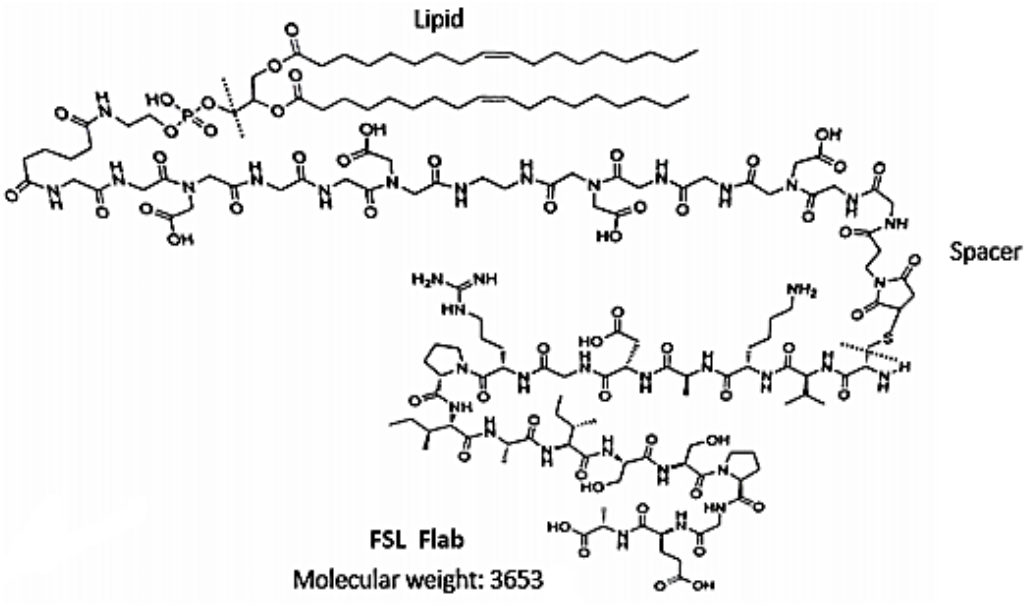

Figure 30: Schematic diagram of FSL Flab (ATGKDVNTGNEPVSKPTG) construct

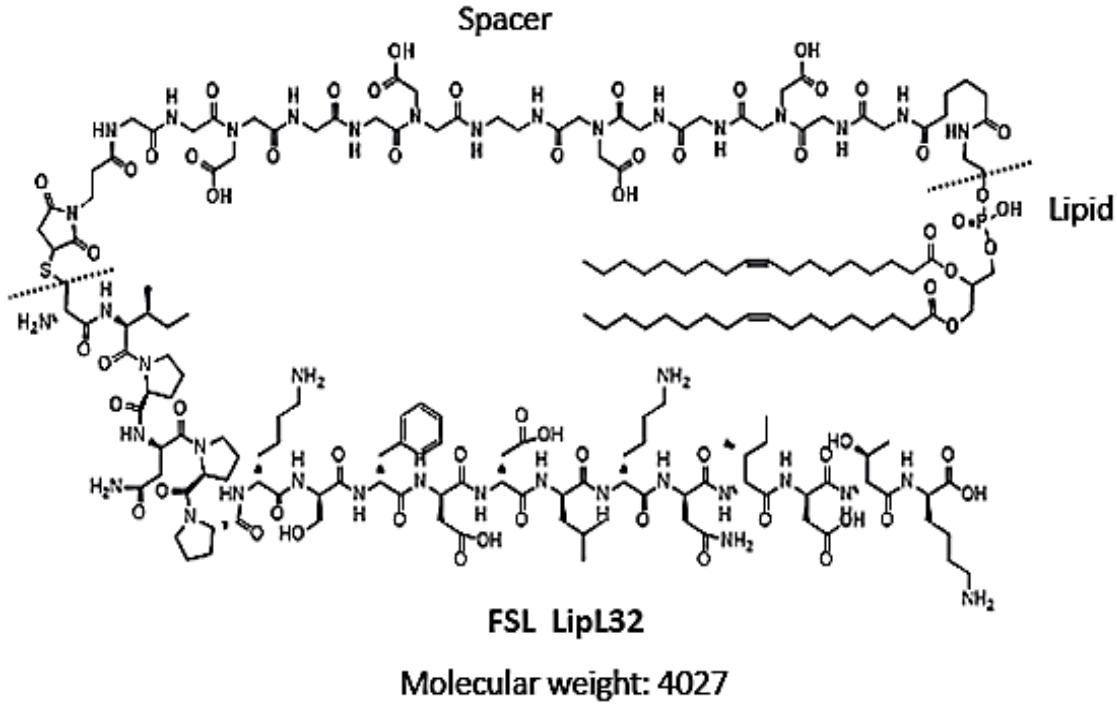

Figure 31: Schematic diagram of FSL LipL32 (IPNPPKSFDDLKNIDTK) construct

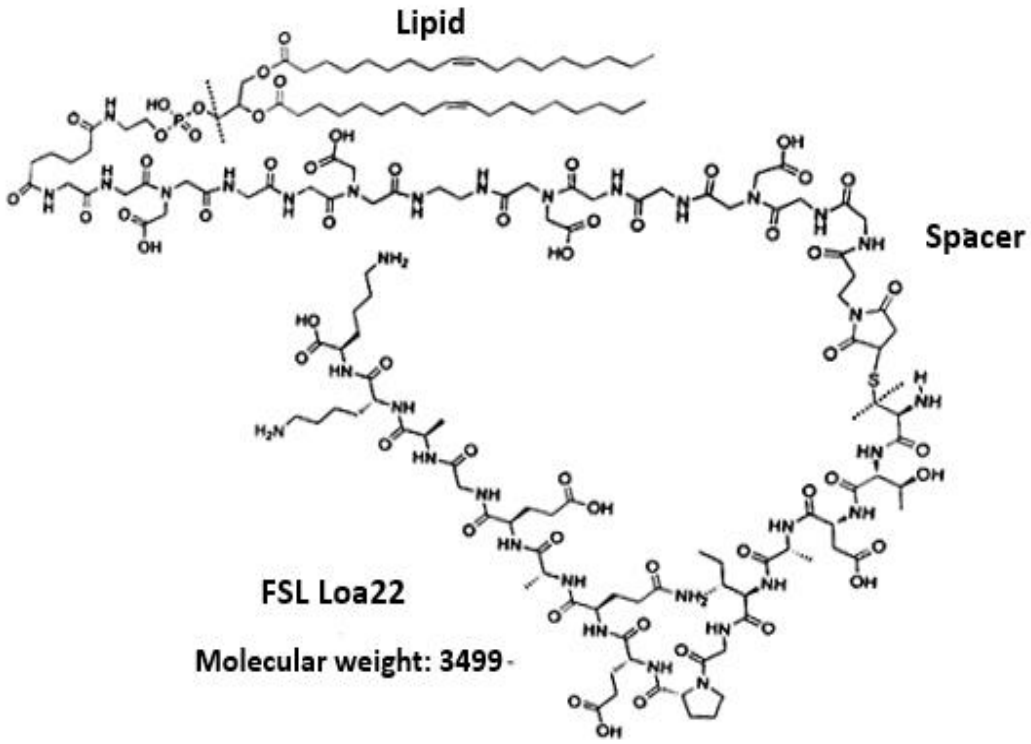

Figure 32: Schematic diagram of FSL Loa22 (TDAIGPEQAEGAKK) construct

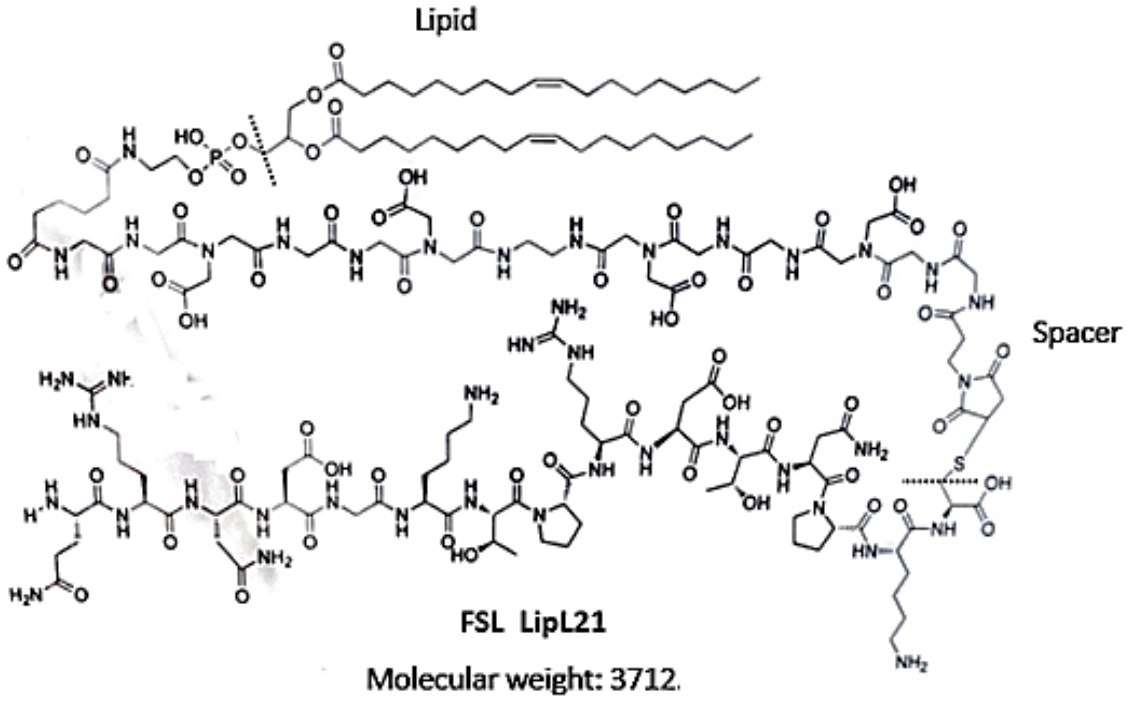

Figure 33: Schematic diagram of FSL LipL21 (QRNDGKTPRDTNPK) construct

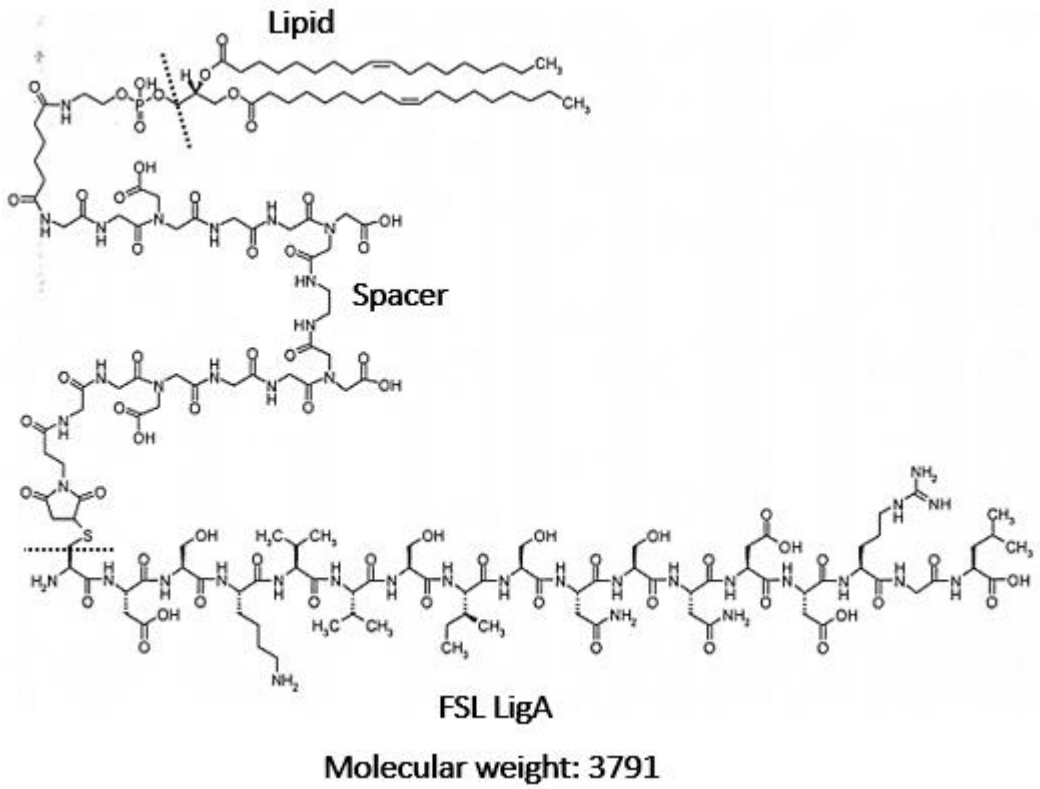

Figure 34: Schematic diagram of FSL LigA (DSKVVSISNSNDDRGL) construct

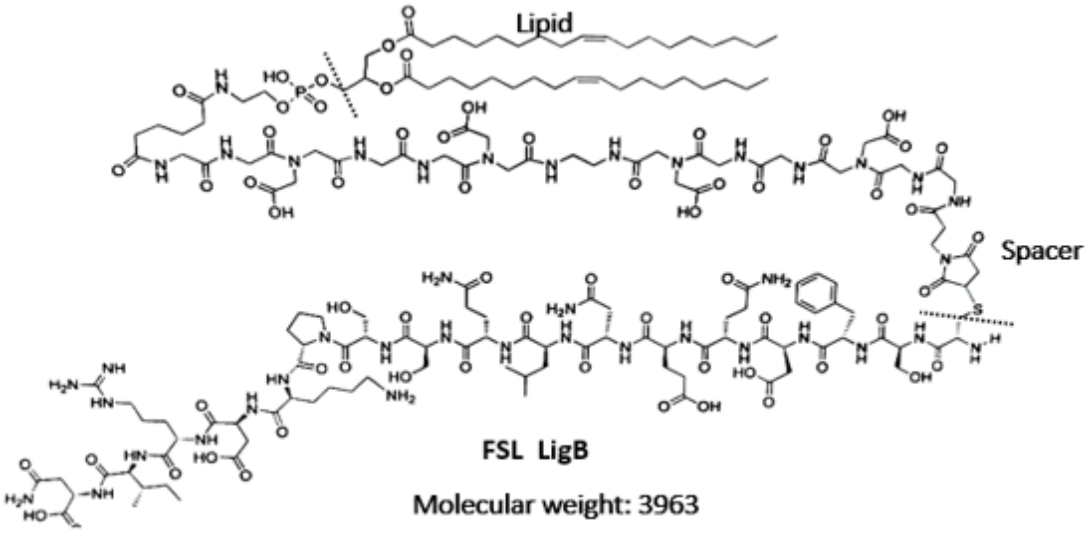

Figure 35: Schematic diagram of FSL LigB (SFDQENLQSSPKDRIN) construct

A generic look at the full protein of the spirochete *Leptospira* using the FSL algorithm resulted in the selection of the eight best candidates.

4.7. Methods and Results: Functional prediction of leptospiral FSL peptides

Functional prediction of leptospiral peptides was expected to be challenging because there was minimal access to validated positive samples. Leptospiral FSL peptides were tested against positive samples as a proof of concept approach for the peptide design and as a preliminary test to check the performance and the suitability of the peptides in diagnostics.

4.7.1. Materials

1. *Leptospira* positive and Negative samples

A small number of de-identified *Leptospira* positive samples, MAT confirmed with a rise in titre, and negative samples were kindly provided for this study by Prof. Cyrille Goarant, Institute Pasteur, New Caledonia.

The ethics approval granted for this study only covers the use of de-identified samples for the initial validation of the peptides and does not allow access to either demographic or clinical data such as the age of the patient, and the days of onset of symptoms.

The following are the sera bank data received for the de-identified positive and negative samples used in this study.

True positive samples are the samples that were lepotospiral antibody positive for different serovars (paired samples tested positive for leptospiral antibody using the gold standard MAT test). *Leptospira* serovars Panama, Pyrogens, Australis, Tarassovi and Icetrohemorrhagiae were received.

True negative samples are the samples that were lepotospiral antibody negative

2. Control samples

De-identified healthy blood donor samples were used as normal control samples for this study.

De-identified samples which were positive for other antibodies and no longer required for further testing were obtained from LabPlus, Auckland City Hospital to check the specificity of the *Leptospira* peptides.

3. Phosphate buffered solution (PBS)

PBS solution (137mM NaCl, 10mM Phosphate, 2.7mM KCl, pH 7.4)

4. Ortho OrthoBioVue cards – column agglutination test (CAT)

ID (Infectious Diseases) - Micro Typing system, Polyclonal rabbit anti human IgG cards, Cat no: MTSO84029, Ortho clinical diagnostics, Raritan, NJ, USA.

5. Natural O cells

Red blood cells from an O phenotype individual were obtained throughout this study from a single individual. The EDTA blood tube was spun and the plasma was removed. The packed cells were washed three times with 1 % PBS solution and one time with celpresol solution. The hard packed cells were stored in celpresol for four weeks at 4° C.

6. *Leptospira* FSL constructs

The selected *Leptospira* candidate peptides were constructed into FSL constructs (Kode Biotech materials, Auckland, New Zealand). The FSL constructs were obtained as lyophilised 5 mg/vial. The FSL constructs were reconstituted in celpresol to make 1mg/mL stock solutions, and the aliquots were stored at -80° C.

4.7.2. Methods and Results

Preparation of *Leptospira* kodecytes

The kodecytes are described by the identification (ID) of the functional head of the FSL and the micromolar concentration (µmol/L) of the FSL solution used to make the kodecytes. For example, *Leptospira* LIC10215-30 kodecytes are made up of the *Leptospira* FSL peptide LIC10215 at a concentration of 30 µM. Based on the learnings from the syphilis diagnostics, and also given the inability to test four concentrations of the eight leptospiral peptides (32 kodecytes) due to the insufficient volume of the *Leptospira* positive samples, it was decided to use a single concentration of 30 µM FSL constructs for the kodecyte assay. A total of eight *Leptospira* kodecytes, LIC10215-30, LipL41-30, LipL21-30, Loa22-30, Flab-30, LipL32-30, LigA-30 and LigB-30 were made using the following protocol.

- Removed the stock vials of leptospiral FSL constructs (1 mg/mL) from the -80° C freezer.
- Allowed the vials to come to room temperature.
- Reconstituted the vials by adding 1 mL of red cell preservative solution (celpresol).
- Made 250 µl of 30 µM concentrations of each leptospiral FSL construct as explained in chapter 3.
- Added 250 µl of each of the eight leptospiral FSL peptides to the labelled tubes.
- Added 250 µl of packed group O red blood cells to each tube.

- Incubated the tubes at 37° C for two hours.
- Washed one time with Phosphate buffered saline and one time with celpresol.
- The packed kodecytes were resuspended in 5% celpresol and stored at 4° C.
- The control cells were made without adding the FSL peptides. In other words, the control cells were group O red cells in celpresol solution.
- The kodecytes were rested overnight before use and were used within 21 days from the date of preparation.

Leptospira kodecytes assay using OrthoBioVue gel cards (CAT)

Protocol for kodecyte assay using column agglutination test (Ortho BioVue gel cards):

1. Made 0.8% suspension (in celpresol) of each *Leptospira* kodecytes (LIC10215-30, LipL41-30, LipL21-30, Loa22-30, Flab-30, LipL32-30, LigA-30 and LigB-30).
2. Labelled the Ortho BioVue gel cards and removed the foil.
3. Added 50ul of 0.8% suspension of each *Leptospira* kodecytes to the corresponding labelled wells in the Ortho BioVue gel cards.
4. Added 40ul of serum samples to the corresponding wells.
5. Incubated the gel cards at 37°C for 15 minutes.
6. Centrifuged for 5 minutes at preset rpm as per protocol.
7. Graded results based on the agglutination pattern.
8. Samples reactive to leptospiral constructs were also tested with unmodified group O red cells as part of the quality control, to ensure positive reactions were not due to anti RBC antibodies.

Interpretation of column agglutination test results:

1. 4+: Agglutinated cells form a band at the top of the bead column.
2. 3+: Most agglutinated cells remain at the top of the column.
3. 2+: Agglutinated cells observed throughout and a small button at the bottom of the column is seen.
4. 1+: Most agglutinated cells remain in the lower half and a button at the bottom of the column is seen.
5. Weak positive (wk+): weak reaction on top of the button.
6. Negative (Neg): compact clear button at the bottom of the column is seen.

An example of the agglutination pattern and the serological grades of the CAT is shown in syphilis chapter 3.

Performance of leptospiral kodecytes with MAT confirmed positive samples

Method overview

A total of 11 true lepotospiral antibody positive samples (paired samples tested positive for leptospiral antibody using the gold standard MAT test) with five different serovars were used to check the performance of the leptospiral kodecytes. All eight leptospiral kodecytes (LIC10215-30, LipL41-30, LipL21-30, Loa22-30, Flab-30, LipL32-30, LigA-30 and LigB-30) were tested against the *Leptospira* positive and negative samples using the kodecyte assay mentioned earlier. As part of the quality control, unmodified group O cells were also used as controls, to ensure positive reactions were not due to anti RBC antibodies.

Results

The *Leptospira* kodecytes performed well. The results for the *Leptospira* kodecytes against MAT positive samples are shown in Table 28. The *Leptospira* kodecytes LigA-30, LipL32-30, and LipL41-30 identified all the results correctly. The LigA kodecytes reacted with all the serovars except Panama. LigA is not present in the serovar Panama, which suggest that LigA-30 is performing well. The kodecytes LipL32-30 identified all the samples correctly; this protein, as per the literature, reacts with both acute and convalescent MAT positive samples. Sample 11 is a convalescent sera sample and LipL32-30 gave a strong signal, unlike other peptides, and those proteins are expressed in mainly acute stage of the illness. The kodecytes LigB-30 and LIC10215-30 identified ten samples correctly out of the 11 samples tested and they picked up all the serovars correctly. The LigB and the LIC10215 proteins are expressed mainly in the acute stage of leptospirosis, and LIC10215 can differentiate between the acute and convalescent stages of the disease. Pending validation with a large-scale number of samples representing the different stages of the disease, these peptides have good potential to be used in *Leptospira* diagnostic testing. The kodecytes Flab-30 identified ten samples correctly and did not identify one sample; Flab-30 also correctly identified all the serovars. The kodecyte LipL41-30 also identified all the 11 samples tested, with results similar to those found for the LipL32-30 kodecytes. The LipL41 protein plays a role in pathogenesis and virulence and is highly conserved among pathogenic leptospires, making it another candidate that, with further validation, the potential for diagnostic use. Loa22 kodecytes and LipL21 kodecytes both did not identify three positive samples. As anticipated, it was challenging to get the clinical samples for testing.

Table 28. Results for *Leptospira* kodecytes against MAT positive samples

Leptospira kodecytes 30 µM	MAT titre Sample ID	*Leptospira* serovars										
		Panama 800 2175	Pyrogens 800 1916	Australis 800 1634	Australis 3200 222-3	Tarassovi 400 112-8	Tarassovi 400 112-7	Ictero 400 112-4	Ictero 800 2186	Ictero 800 1604	Ictero 1600 112-2	Ictero 3200
LigA		-	1+	3+	2+	2+	2+	2+	1+	2+	2+	1+
Flab		wk+	1+	2+	2+	2+	-	1+	2+	2+	1+	wk+
LigB		2+	2+	1+	1+	2+	2+	2+	1+	2+	1+	-
LIC10215		2+	2+	3+	1+	wk+	wk+	wk+	3+	2+	2+	-
Loa22		2+`	-	2+`	2+`	2+`	1+	-	-	2+	2+	1+
LipL21		2+`	-	1+	1+	2+	2+	-	1+	2+	-	wk+
LipL32		1+	1+	2+	2+	3+	2+	wk+	2+	2+	2+	3+
LipL41		2+	3+	1+	1+	1+	2+	1+	1+	2+	2+	1+

The above table shows the results for the sensitivity of the *Leptospira* kodecytes. Eight leptospiral kodecytes were tested with the known positive samples using gel cards with anti-human globulin (infectious diseases micro typing system). The scoring system was followed as per the manufacturer's instructions: 3+ (strong positive), 2+ (moderately positive), 1+ (positive), wk+ (assigned to weak positive samples).

The caveat of this experiment is the limited number of samples used for the evaluation and the nature of the samples (not fresh and having undergone freeze/thaw conditions). To prove the true sensitivity of the *Leptospira* FSL peptides constructs, a large number of fresh clinical samples need to be tested and different concentrations of the need to be used. From the above table the preliminary results show that the *Leptospira* kodecytes are working but need further large-scale testing with well-defined samples and clinical follow up to validate the kodecytes for clinical use.

Validation of analytical specificity *of Leptospira* kodecytes

The specificity of the *Leptospira* kodecytes was evaluated using two methods. The first method was testing the kodecytes with healthy blood donor samples. The second method was testing the kodecytes for cross reactivity with samples that had been tested and had antibodies/reactive serology for other pathogens. A number of disease agents and serum samples positive for autoimmune diseases have been reported by other investigators to cross react in leptospirosis serologic assays[219-222].

1. *Leptospira* kodecytes against healthy blood donor samples

Method overview

A total of 100 de-identified healthy blood donor samples from the New Zealand blood service were tested with eight *Leptospira* kodecytes (LIC10215-30, LipL41-30, LipL21-30, Loa22-30, Flab-30, LipL32-30, LigA-30 and LigB-30) using the kodecytes assay (CAT). The results were graded according to the test protocol. The further confirmation of the positive results was done using the commercial *Leptospira* ELISA test to see whether the positive results obtained from the kodecyte assay were true positive or false positive.

Results

The *Leptospira* kodecytes showed positive results initially when tested against healthy blood donor samples (Table 29). The kodecytes LIC10215-30, LigB-30 and LipL21-30 gave eight positive results each. The kodecytes LipL41-30 gave seven positive results, LipL32-30 gave one positive result, and the remaining kodecytes gave one to four positive results out of the 100 samples tested (Table 29). These positive samples were further tested with the *Leptospira* ELISA test for confirmation. The kodecytes LipL21-30 and LigB-30 gave eight positive results initially, and the ELISA method confirmed five positive results (two IgM positive results and three IgG positive results) out of eight, showing that the two kodecytes have 97 % specificity. The revised results of kodecytes LigA-30 and Loa22-30 showed 98% specificity, with two false positive results. The kodecytes Flab-30 and LigB-30 showed 97% specificity and the kodecytes LipL32-30 showed 99%

Table 29. Results for *Leptospira* peptides with healthy blood donor samples

Leptospira Peptides 30 µM	BD IR n=100		Positive sample ELISA confirmation			Confirmed	False results	Revised results	Specificity (%)
	-	+	IgM+IgG	IgM	IgG				
LigA	97	3	0	1	0	1	2	98	98
FlaB	96	4	0	1	0	1	3	97	97
LigB	92	8	3	2	0	5	3	97	97
LIC10215	92	8	0	0	1	1	7	92	93
Loa22	97	3	0	1	0	1	2	98	98
LipL21	92	8	1	2	2	5	3	96	97
LipL32	99	1	0	0	0	0	1	99	99
LipL41	93	7	0	0	1	1	6	94	94

BD: Blood donor sample, IR: Initial results.

specificity. Among the *Leptospira* kodecytes tested with blood donor samples, LIC10215-30 and LipL41-30 have a lower specificity of 93% and 94%. The revised results showed that the *Leptospira* kodecytes have reasonable specificity, except for LIC10215 and LipL41.

2. *Leptospira* kodecytes against hepatitis B seroreactive samples

Method overview

A total of 11 de-identified samples with reactive serology for hepatitis B virus antibody were tested with eight *Leptospira* kodecytes (LIC10215-30, LipL41-30, LipL21-30, Loa22-30, Flab-30, LipL32-30, LigA-30 and LigB-30) using the *Leptospira* kodecytes assay (CAT) to see whether they are cross reacting with *Leptospira* kodecytes. The results were graded according to the test protocol. The initial positive results obtained were further tested using the commercial *Leptospira* ELISA test to see whether the positive results obtained from the kodecyte assay were true positive or false positive.

Results

The *Leptospira* kodecytes assay results (Table 30) showed positive results initially when tested against the hepatitis B antibody positive samples. The kodecytes LigA-30 gave six positive results and LigB gave eight positive results. The Kodecytes, Flab-30, Loa22-30 and LipL41-30 gave negative results for all the 11 samples tested. The Kodecytes, LIC10215-30, LipL21 and LipL32 -30 gave one to two positive results each. The positive results from kodecytes LigA-30 and LigB-30 were confirmed by the commercial *Leptospira* ELISA test; three of these were IgM positive and the rest IgG positive for *Leptospira* antibody, suggesting possible recent infection of leptospirosis. Similarly, initial positive results were seen for the kodecytes LIC10215-30 and LipL21-30 were confirmed positive by *Leptospira* ELISA test.

Table 30. Cross reactivity results for *Leptospira* kodecytes against hepatitis B samples

Leptospira Peptides 30 µM	LKIR IR n=11		Positive sample ELISA confirmation			Confirmed	False results	Revised results	Specificity (%)
	-	+	IgM+IgG	IgM	IgG				
LigA	5	6	1	3	2	6	0	11	100
FlaB	11	0	0	0	0	0	0	11	100
LigB	3	8	1	3	2	6	2	9	98
LIC10215	10	1	0	1	0	1	0	11	100
Loa22	11	0	0	0	0	0	0	11	100
LipL21	10	1	0	1	0	1	0	11	100
LipL32	9	2	1	0	1	2	0	11	100
LipL41	11	0	0	0	0	0	0	11	100

LKIR: *Leptospira* kodecytes initial results in hepatitis B antibody positive samples.

The revised *Leptospira* kodecytes results showed that the *Leptospira* kodecytes have 100% specificity, and that most did not cross react with hepatitis B antibody positive samples, except for the LigB-30 kodecytes, which gave two false positive results with a specificity of 98%.

3. *Leptospira* kodecytes against dengue seroreactive samples

Method overview

A total of ten de-identified samples with reactive serology for dengue virus antibody were tested with eight *Leptospira* kodecytes (LIC10215-30, LipL41-30, LipL21-30, Loa22-30, Flab-30, LipL32-30, LigA-30 and LigB-30) using the *Leptospira* kodecytes assay (CAT) to see whether they cross react with *Leptospira* kodecytes. The results were graded according to the test protocol. The initial positive results obtained were further tested using the commercial *Leptospira* ELISA test to see whether the positive results obtained from the kodecyte assay were true positive or false positive.

Results

The *Leptospira* kodecytes, when tested against the dengue virus antibody positive samples, initially showed positive results for the kodecytes LigB-30, Flab-30 and LipL32-30. The kodecytes LigB-30 initially gave eight positive results, out of which six were confirmed by ELISA test. Flab-30 and LipL32-30 gave seven positive results, out of which four were confirmed by ELISA test. The revised results showed two false positive results for LigB-30 and three false positive results for

Table 31. Cross reactivity results for *Leptospira* kodecytes against dengue samples

Leptospira Peptides 30 µM	LKIR results n=10		Positive sample ELISA confirmation			Confirmed	False results	Revised results	Specificity (%)
	-	+	IgM+IgG	IgM	IgG				
LigA	6	4	0	2	1	3	1	9	90
FlaB	3	7	0	2	2	4	3	7	70
LigB	2	8	1	3	2	6	2	8	80
LIC10215	9	1	0	0	0	0	1	9	90
Loa22	9	1	0	0	0	0	1	9	90
LipL21	9	1	0	0	0	0	1	9	90
LipL32	3	7	1	3	2	6	1	9	90
LipL41	9	1	0	0	0	0	1	9	90

LKIR: *Leptospira* kodecytes initial results in dengue antibody positive samples.

FlaB-30. The remaining kodecytes, LIC10215-30, Loa22-30, LipL21-30 and LipL41-30 gave one false positive result each out of the ten dengue positive samples tested. *Leptospira* ELISA antibody test confirmed some of the positive reactions and collates leptospirosis and dengue coinfection as stated in the literature[223].

4. *Leptospira* kodecytes against autoantibody positive samples

Method overview

A total of eight de-identified samples serologically reactive for autoantibodies were tested with eight *Leptospira* kodecytes (LIC10215-30, LipL41-30, LipL21-30, Loa22-30, Flab-30, LipL32-30, LigA-30 and LigB-30) using the *Leptospira* kodecytes assay (CAT) to see whether they are cross reacting with *Leptospira* kodecytes. The results were graded according to the test protocol. The initial positive results obtained were further tested using the commercial *Leptospira* ELISA test.

Results

The *Leptospira* kodecytes assay results showed good specificity (100%) for the kodecytes LigA-30, Flab-30, LipL21-30, Loa22-30 and LipL41-30. Initially, LigB kodecytes gave one positive result, LIC10215 gave two positive results and LipL32 gave two positive results (Table 32). The initial positive results of the kodecytes LigB-30 and LIC10215-30 were confirmed by the *Leptospira* ELISA

Table 32. Cross reactivity results for *Leptospira* kodecytes against autoantibody samples

Leptospira Peptides 30 µM	LKIR n=8		Positive sample ELISA confirmation			Confirmed	False results	Revised results	Specificity (%)
	-	+	IgM+IgG	IgM	IgG				
LigA	8	0	0	0	0	0	0	8	100
FlaB	8	0	0	0	0	0	0	8	100
LigB	7	1	0	0	1	1	0	8	100
LIC10215	6	2	0	0	2	2	0	8	100
Loa22	8	0	0	0	0	0	0	8	100
LipL21	8	0	0	0	0	0	0	8	100
LipL32	6	2	1	0	1	2	0	8	100
LipL41	8	0	0	0	0	0	0	8	100

LKIR: *Leptospira* kodecytes initial results in autoantibody positive samples.

antibody test as IgG positive. The two initial positive results from the kodecytes LipL32-30 were confirmed by the ELISA test as IgG positive and IgM+IgG positive. The results show that leptospiral kodecytes do not cross-react with autoantibodies, as reported in the literature[219].

A. Conclusion and summary

The preliminary results of leptospiral FSL peptides show that the leptospiral FSL peptide constructs are working, as seen from the proof of concept results. The caveat of this study is the small number of positive samples used for the functional analysis of the *Leptospira* kodecytes. The number of positive samples tested was very limited and the samples were not fresh but frozen samples which had undergone freeze/thaw conditions.

Leptospira kodecytes gave reasonably good results when tested using known true MAT test confirmed positive samples. The kodecytes LipL41-30 and LipL32-30 identified all the positive samples correctly. The kodecytes LigA-30 and LigB-30 identified all the positive samples except for one sample. The kodecytes LigA-30 did not identify the serovar Panama. LigA gene is not present in the serovar Panama, which was rightly missed by the LigA-30 kodecytes. The kodecytes LigB-30 missed one sample out of the five positive serovar Iceterohaemorrhagiae samples. The kodecytes LIC10215-30, like the LigB-30, identified all the positive samples except one. The kodecytes FlaB-30 also identified all the positive samples except one sample, serovar Tarassovi. Out of the eight *Leptospira* kodecytes tested, the LipL21-30 and Loa22-30 kodecytes failed to identify the largest number of samples (3). Overall, the leptospiral peptide constructs are working, as they picked up all the serovars tested. As mentioned earlier, further testing of the peptides with a large number of defined clinical samples with clinical follow up is required to validate these peptides for their suitability in diagnostics

The *Leptospira* kodecytes showed good analytical specificity (range 93% to 99%) as seen from the results obtained from 100 healthy blood donor samples. The kodecytes LipL32-30 showed a maximum specificity of 99% followed by LigA-30 and Loa22-30, both with a specificity of 98%. The kodecytes FlaB-30, LigB-30 and LipL21-30 showed 97% specificity. The kodecytes LIC10215-30 and LipL41-30 showed the lowest percentage specificity of 93% and 94%. These results show that the designed *Leptospira* FSL peptides have good specificity.

There was no cross reactivity seen when the peptides were tested with samples serologically reactive for autoantibodies. The kodecytes were also tested using samples positive for hepatitis B antibody. Though the initial results showed more positive results, further testing of the positive samples with *Leptospira* ELISA method confirmed they were true positives and not false positives. This shows that the kodecyte assay is comparable to the commercial ELISA test; again, the caveat is the small number of samples tested.

The *Leptospira* kodecytes, when tested against ten dengue virus antibody positive samples, showed positive results. Further testing of the positive samples with commercial *Leptospira* ELISA methods confirmed only some of the positives. Most of the *Leptospira* kodecytes showed one false positive result with dengue seroreactive samples. This has to be further validated with a greater number of samples.

Future work

The results obtained from the functional analysis of the leptospiral FSL peptides shows that the peptides are working with reasonably good specificity. The *Leptospira* FSL peptides need further validation studies with a greater number of *Leptospira* positive samples with different serovars, negative samples, and limits for detection, either using a quantitative commercial positive control or in house positive control, to validate the suitability in *Leptospira* diagnostic.

Chapter 5. Methods and Results: SARS CoV-2 diagnostic using Kode FSL peptides

5.1. Literature review of SARS CoV-2 proteome

SARS CoV-2 genome consists of ten distinctive proteins (four structural proteins, five non-structural proteins and one non-structural polyprotein)[224]. The spike protein and the nucleocapsid protein are reported to be immunogenic among the four structural proteins involved in SARS CoV-2 viral replication (Spike glycoprotein-S, Membrane-M, Envelope-E, and Nucleocapsid-N)[225,226]. The N protein, though highly immunogenic, shares 90% sequence similarity with SARS CoV-1 and around 30% similarity with human coronavirues[227,228]. Cross reactivity of SARS CoV-2 N protein with other coronaviruses, in particular SARS CoV-1, is reported in two of the proteome microarray studies [226,229] and hence N protein was not selected for this study. The spike protein is highly immunogenic[226], and has 76% similarity to SARS CoV-1 compared to N protein. The S1 region of the spike protein has only 66% similarity compared to the full length of S protein. The S protein is widely used in the development of antibody assays, targeting the receptor binding domain (RBD) of the S1 region, since the ACE2 receptor of the spike protein binds to the host cell [230-232] and the notion that the human immune system and the antibody will recognize.

5.2. Rationale for choosing spike protein for FSL peptide design

Spike protein was selected for the SARS CoV-2 peptide design for FSL constructs, based on its immunogenicity and sequence differences from other coronaviruses. Most of the current SARS CoV-2 antibody assays are targeting the S1 region of the spike protein (RBD). The aim of this study is to look for B cell epitopes throughout the S protein, covering both S1 and S2 domains of the protein, and see whether the human immune system recognises other regions of the S protein, rather than the receptor binding domain only. An additional reason for looking at the S2 region is that it is more conserved than S1 and not many mutations are seen in this region.

5.3. Method and Results: Analysis of SARS CoV-2 protein sequences

The same approach used for syphilis and *Leptospira* was used for SARS CoV-2.

Protein sequence retrieval

The SARS CoV-2 published S protein sequence used in this study to design the peptides was derived from the NCBI database, GenBank no: QHD43416.1. To verify for sequence similarity of the SARS CoV-2 protein against other coronaviruses, the S protein sequences of normal HCoVs (229E, HKU1, NL63 and OC43), SARS CoV-1 and MERS CoV reference sequences were also used.

The following are the GenBank accession numbers used for other coronaviruses. HCoV-229E-ALK28781, HCoV-HKU1- AGT17777, HCoV-NL63- AGT51331, HCoV-OC43-QBP84731.1, SARS CoV-1- ABD72968.1, MERS CoV- AHI48572.

5.3.1. Structure prediction using I-TASSER/NovaFold

The SARS CoV-2 spike protein structure prediction was done using the I-TASSER/NovaFold software. Based on simulation and alignment of the threading pattern, this softwarre calculates a confidence score (C-score) of the predicted I-TASSER model. The C-score runs between [-5, 2], with higher C-scores suggesting more confident models and lower C-scores indicating less confident models. The predicted C-score of the SARS CoV-2 structure of the submitted protein sequences is -1.90, this is a good score and model.

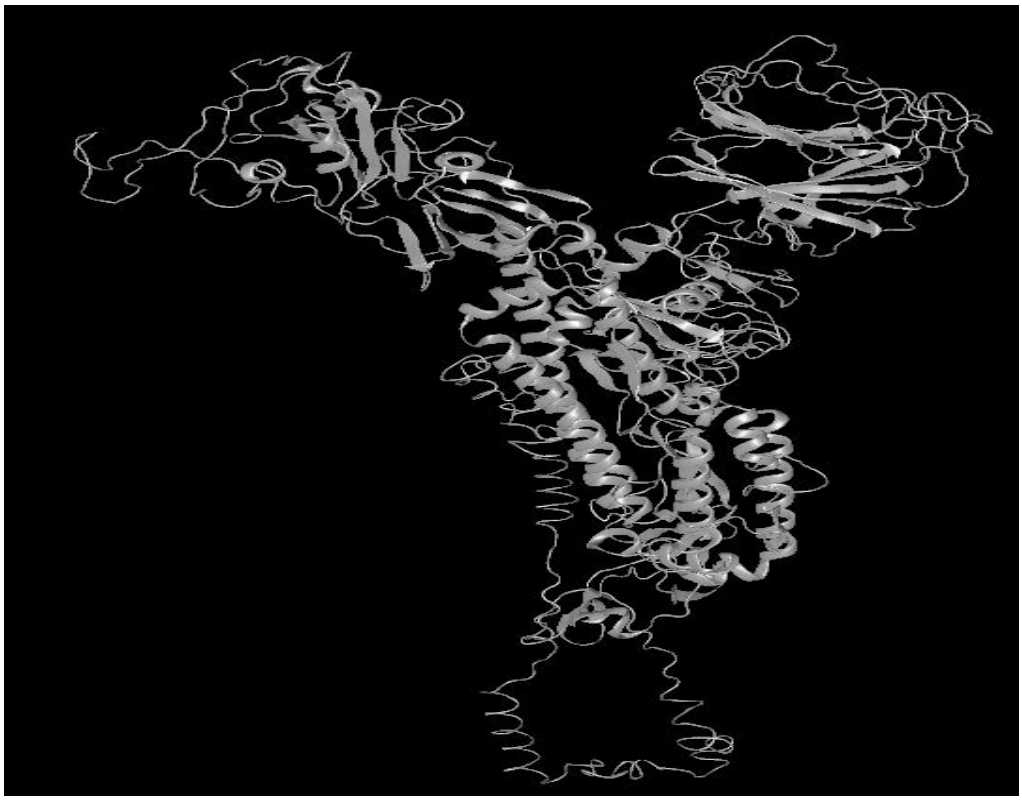

Figure 36 Structure of SARS CoV-2 spike protein

The predicted 3D structure of the spike protein is shown in Figure 36. The SARS CoV-2 spike protein structure showed more random coils than β strands and helices. The I-TASSER predicted binding site residues are 295 to 298, 1032, 1043, 1048, 1001 to 1004, 1212, 1213 and 874 to 878.

5.3.2. Structure analysis using protein structure analysis program (ProSA)

The I-TASSER/NovaFold predicted structure was further analysed with the ProSA program. The number of residues is plotted on the x axis, while the Z score is plotted on the y axis (Figure 37).

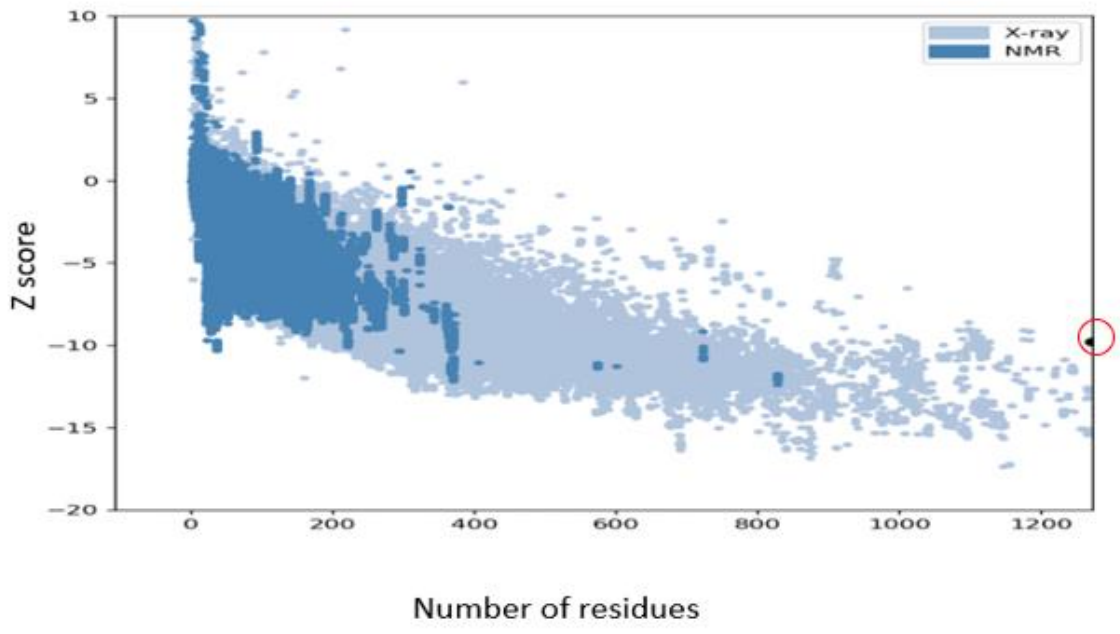

Figure 37 Seven score plot for the predicted model of the SARS CoV-2 protein

The user input structure score is circled in red. The Z score of the input structure should be within the range of scores obtained in native proteins of similar size. The predicted Z score is -9.78, which is within the range of the estimated Z score (4 to -12, as seen in the figure) of all the available protein structures in the protein database at the time of prediction, using both NMR and X-ray crystallography methods.

5.3.3. Structure analysis using the self-optimized structure prediction method (SOPMA)

The secondary structural features of the protein were analysed using SOPMA software. The secondary structures of the SARS CoV-2 spike protein using the SOPMA software (Table 33) revealed that

Table 33. Secondary structures predictions of SARS CoV-2 spike protein - SOPMA

Secondary structure	Percentage
α helix	28.6
β sheet	23.3
β turn	3.4
Bend region	0
Random coil	44.8

the protein structure has an equal percentage of helices and β sheets, but the percentage of random coil was closer to 50. Even though the secondary structure showed large percentages of helices and coil regions, because the spike protein was heavily glycosylated, it was difficult to find surface exposed epitopes in non-glycosylated regions.

5.4. Methods and Results: Identification of B cell epitopes

For this study, it was decided to design eight to ten SARS CoV-2 peptides covering the full length of the spike protein and build them as FSL SARS CoV-2 peptide constructs.

5.4.1. Epitope prediction

Method overview

As discussed earlier, the B cell epitope prediction programs ABCpred, BCPred, BepiPred-2.0, LBtope and DNASTAR Protean 3D were used for the predictions. The detailed methodology is explained in the syphilis chapter. The SARS CoV-2 protein sequences were submitted to the B cell prediction programs.

Results

The candidate peptide epitope sequences consistently predicted by all the methods were chosen and were considered as potential epitopes. The selected peptides are shown in Table 34. A total

Table 34. List of SARS CoV-2 peptide sequences

ID*	CoV†	Peptide sequence
178	2	DLEGKQGNFKNLREF
406	2	EVRQIAPGQTGKIAD
458	2	KSNLKPFERDISTEI
808	2	DPSKPSKRSFIEDLL
1147	1, 2	SFKEELDKYFKNHTS
1255	1, 2	KFDEDDSEPVLKGVK

*ID is based on the initial amino acid in the SARS CoV-2 consensus sequence and includes an H if the sequence has an additional histidine tail sequence appended. † CoV indicates if specific to SARS CoV-2 (2) or common to both SARS-CoV-1 and SARS CoV-2 (1, 2).

of six unique peptide sequences were predicted for FSL construction. Out of the six predicted peptides, four were specific to SARS CoV-2 and 2 of them were common to both SARS CoV-1 and SARS CoV-2 (1, 2).

5.4.2. Methods and Results: analysis of the predicted SARS CoV-2 peptides

The structure and the biophysical properties of the predicted peptides were analysed for the suitability to be built as FSL constructs, as per the FSL selection algorithm table described earlier.

A. Structural analysis of protein using I-TASSER and DNASTAR Protean 3D program

The predicted leptospiral peptide sequences were analysed in the protean 3D program to locate and visualise the secondary structures of the predicted peptides (Figure 38). The location of the

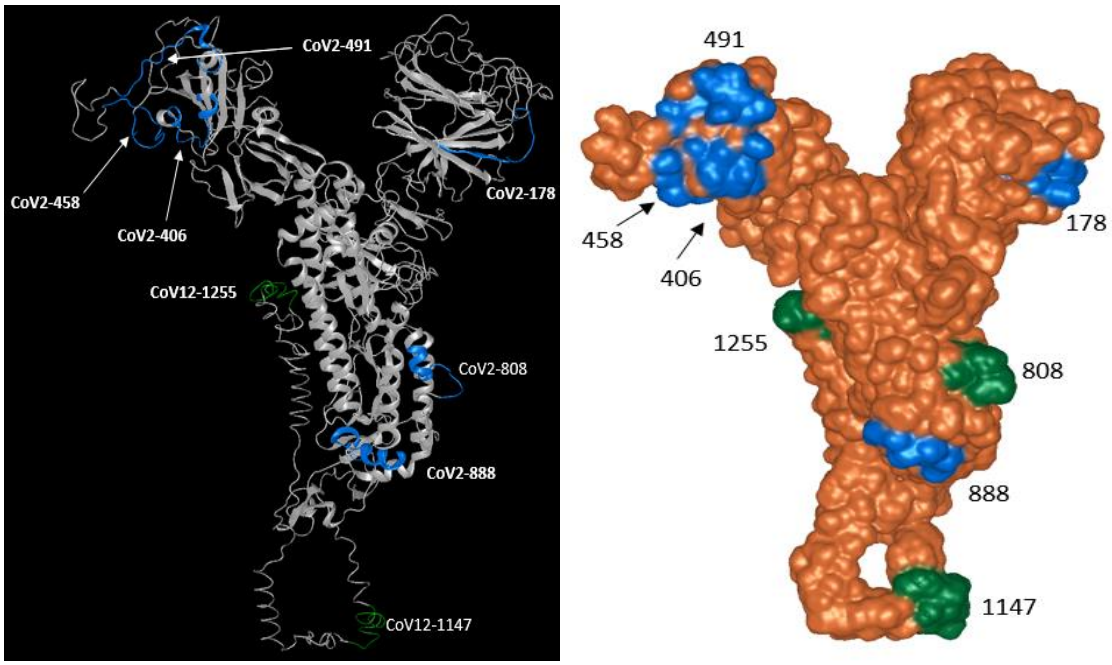

Figure 38: Location of the selected peptide sequence in SARS CoV-2 spike protein

SARS CoV-2 peptides on the full parent protein are marked and shown in blue and green on the left side of Figure 38. The location of the peptides is marked (blue for low and green for high sensitivity and specificity as FSL constructs) in the solid view of the protein shown on the right side of the figure. The peptides are exposed, as seen in the solid view of the protean 3D structure analysis.

B. Structure prediction of the SARS CoV-2 peptides using I-TASSER/NovaFold

The predicted SARS CoV-2 peptide sequences were analysed in the I-TASSER/ NovaFold program and the probable solvent confirmation of the peptide fragments and their structures are shown in Figure 39.

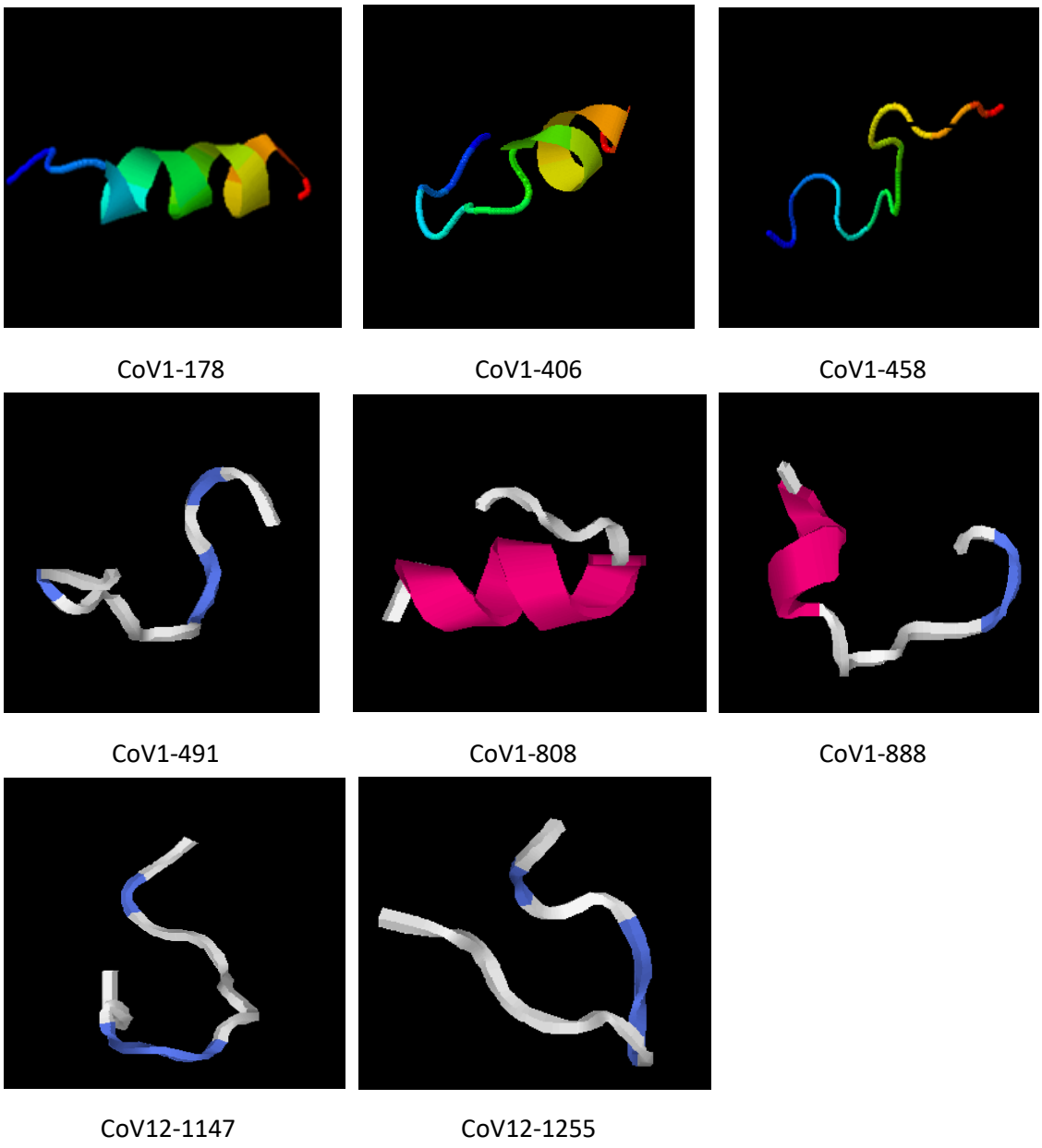

Figure 39 : I-TASSER predicted model of the selected SARS CoV-2 peptides

The peptides CoV1-178, CoV1-406, CoV1-808 and CoV1-888 were predicted as coil/helix by the I-TASSER program, whereas the peptides CoV1-458, CoV1-491, CoV12-1147 and CoV12-1255 were predicted as coils.

5.4.3. SARS CoV-2 peptides: Results of the theoretical prediction

The theoretical predictions of the SARS CoV-2 peptides using various parameters and the results obtained are shown in Table 35a and Table 35b. Table 35a shows the epitope predictions, the basic structure and antigenicity of the peptides and the sequence similarity to other coronaviruses.

Table 35. Results of the theoretical prediction of the SARS CoV-2 FSL peptides

35a **Epitope prediction**										
Peptide details	178	406	458	491	491H	808	888	888H	1147	1255
Basic structure	Coil/ helix	Coil/ helix	Coil	Coil	Coil	Coil/ helix	Coil/ helix	Coil/ helix	Coil	Coil
Epitope prediction										
BepiPred-2	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes
BCPred	No	Yes	No	Yes	Yes	Yes	No	No	Yes	Yes
ABCpred	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
DNASTAR	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes
Antigenicity										
Kolaskar/Tongaonkar	0.95	1.0	1.0	1.1	1.1	0.98	0.99	0.99	1.0	1.1
Antigen profiler	3.9	3.2	3.8	2.3	1.6	4.1	2.1	2.0	3.9	4.5
Similarity (SARS-CoV-n)										
BLAST	2	2	2	2	2	2	2	2	1, 2	1, 2

BLAST: "2" denotes sequence similarity to SARS CoV-2 and "1, 2" denotes sequence similarity to both SARS CoV-1 and SARS Cov-2.

Table 35b **Epitope biophysical properties**										
Peptide details	178	406	458	491	491H	808	888	888H	1147	1255
Internal cysteine	0	0	0	0	0	0	0	0	0	0
Terminal Glutamine	0	0	0	0	0	0	0	0	0	0
NetOGlyc	0	0	0	0	0	0	0	0	0	0
NetNGlyc	0	0	0	0	0	0	0	0	0	0
Net charge	0	0	0	0	0	0	0	0	0.09	-2
Isoelectric point	6.18	6.17	6.18	5.95	5.95	6.12	5.52	5.52	6.48	4.44
Parker scale	4.5	4.6	4.8	4.0	4.0	5.5	2.0	2.0	3.7	7.7
Surface Accessibility	1.0	2.0	2.0	1.3	1.4	2.1	1.6	2.4	1.7	2.3
Peptide length	16	16	16	16	21	16	16	20	16	16

Table 35b shows the biophysical properties of the peptides. The Kolaskar/Tongaonkar antigenicity scores are good for all the peptides, but the antigen profiler method gave low antigenicity scores for four peptides (published peptides 491, 888, 491H and 888H. The peptides 491H and 888H are variations of 491 and 888, with added histidine residues for solubility purposes).

5.4.4. Candidates of SARS CoV-2 peptides for Kode FSL construction

The list of SARS CoV-2 peptide sequences deemed suitable for FSL construction are shown in Table 36.

Table 36. SARS CoV-2 peptide sequences selected for construction into FSL constructs.

ID*	CoV†	Peptide sequence‡
178	2	DLEGKQGNFKNLREF[C]
406	2	EVRQIAPGQTGKIAD[C]
458	2	KSNLKPFERDISTEI[C]
491	2	PLQSYGFQPTNGVGY[C]
491H	2	PLQSYGFQPTNGVGY[HHHH][C]
808	2	DPSKPSKRSFIEDLL[C]
888	2	FGAGAALQIPFAMQM[C]
888H	2	FGAGAALQIPFAMQM[HHH][C]
1147	1, 2	[C]SFKEELDKYFKNHTS
1255	1, 2	[C]KFDEDDSEPVLKGVK

*ID is based on the initial amino acid in the SARS CoV-2 consensus sequence and includes an H if the sequence has an additional histidine tail sequence appended. † CoV indicates if specific to SARS CoV-2 (2) or common to both SARS CoV-1 and SARS CoV-2 (1, 2). ‡ SARS CoV-2 peptide sequence used from the first amino acid (relating to the ID number) with additional residues not part of the natural peptide sequence, including the conjugation cysteine [C] and solubilization histidine [H] residues. The [C] cysteine residue (which is employed to bind the peptide to the spacer) also shows the peptide's closest proximity to the cell membrane.

The ID numbers shown in the table is based on the initial amino acid in the spike protein sequence and the numbers in the CoV† column indicates whether the peptide sequence is specific to SARS CoV-2 alone or to both SARS CoV-1 and SARS CoV-2. Out of the ten SARS CoV-2 peptides listed in the table, six peptides were predicted in this study (178,406,458,808,1147 and 1255) and two from the published data [233] (491 and 888), plus two additional peptides (491H and 888H) with histidine tail sequences designed and selected using FSL peptide selection algorithm, were found to be suitable for construction of FSL constructs, and were selected for FSL construction.

5.5. SARS CoV-2 FSL constructs

The peptides were synthesized into FSL constructs by conjugating the functional head (peptide) with the cysteine SH group to a spacer (carboxymethylglycine), which is conjugated to the lipid phosphate moiety (DOPE). The schematic diagrams of the ten SARS CoV-2 function-spacer-lipid (FSL) constructs and their molecular weights are shown in Figures 40 to 49.

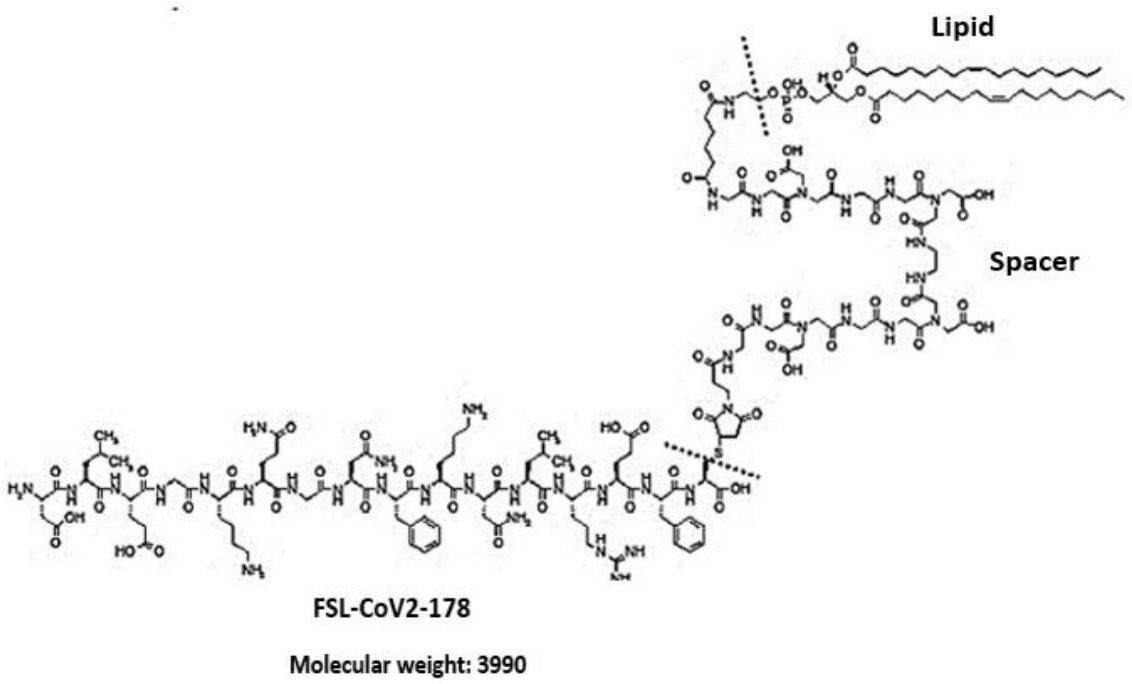

Figure 40: Schematic diagram of FSL- CoV 2-178 construct

Figure 40 shows the function-spacer-lipid (FSL) construct consisting of a lipid phosphate moiety conjugated to the spacer, which is conjugated via a cysteine SH group to the peptide functional head (CoV2-178-DLEGKQGNFKNLREF). The dotted line denotes the beginning and end of the carboxymethylated oligoglycine based spacer that separates the Function and Lipid segments. The subsequent Figures (41 to 49) feature the same structural logic and denotations.

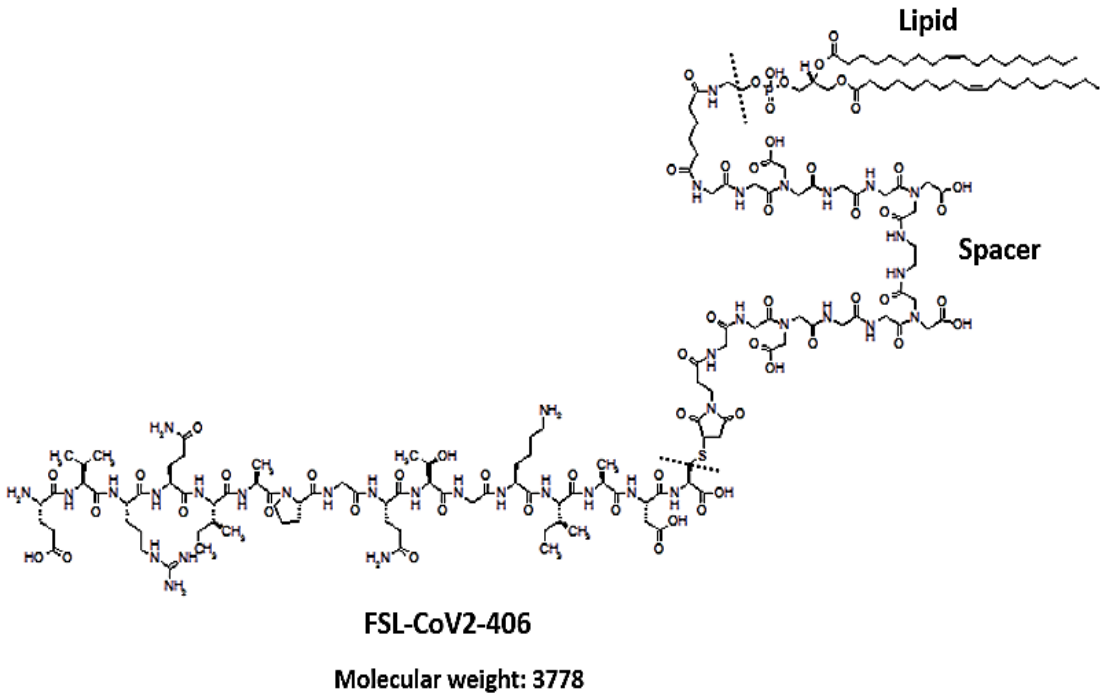

Figure 41: Schematic diagram of FSL- CoV 2- 406 (EVRQIAPGQTGKIAD) construct

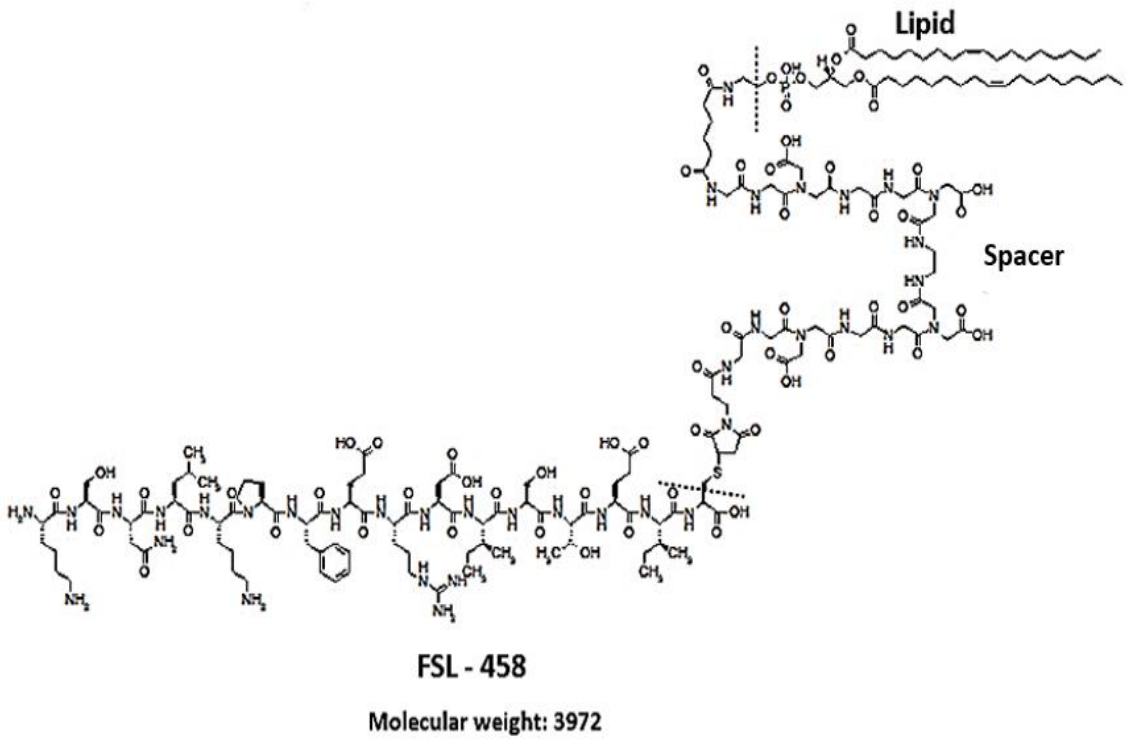

Figure 42: Schematic diagram of FSL-CoV 2-458 (KSNLKPFERDISTEI) construct

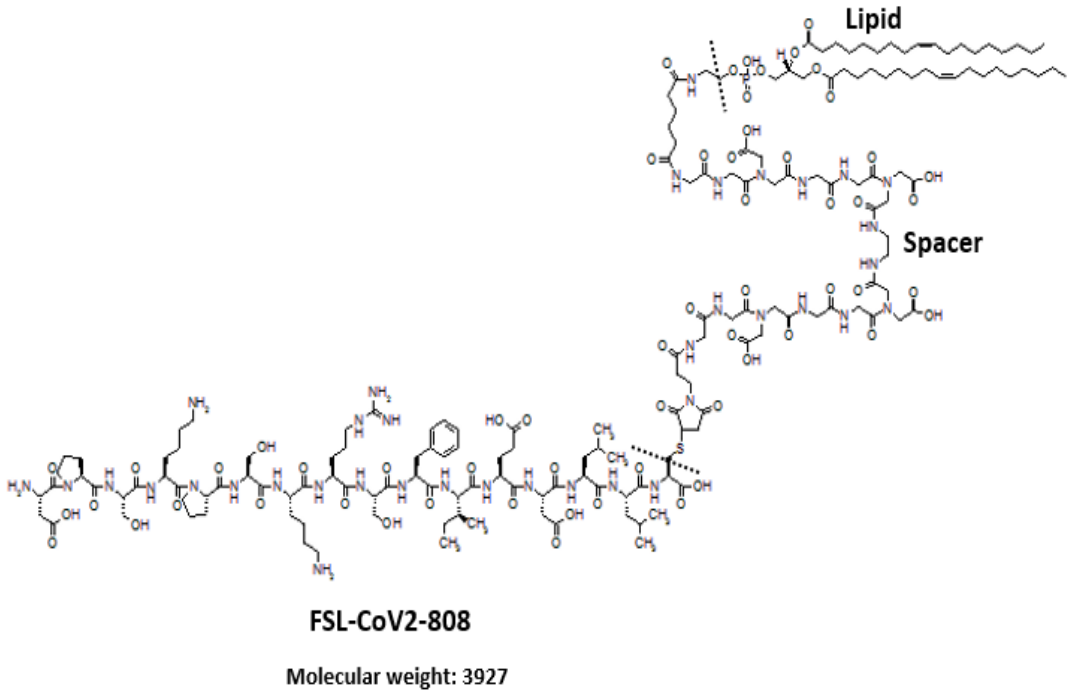

Figure 43: Schematic diagram of FSL-CoV-2-808 (DPSKPSKRSFIEDLL) construct

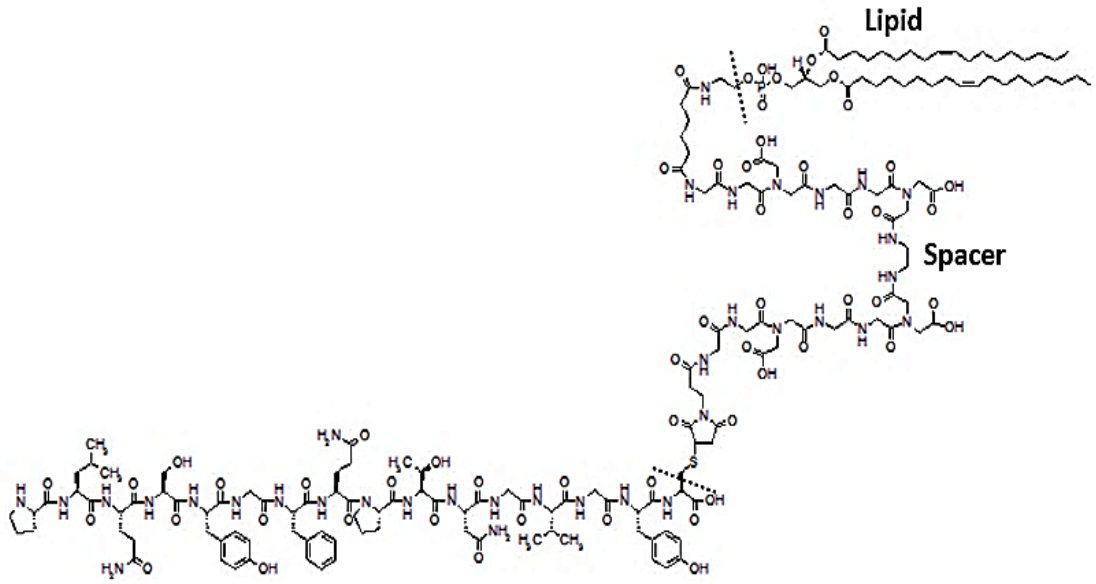

Figure 44: Schematic diagram of FSL-CoV-2-491 (PLQSYGFQPTNGVGY) construct

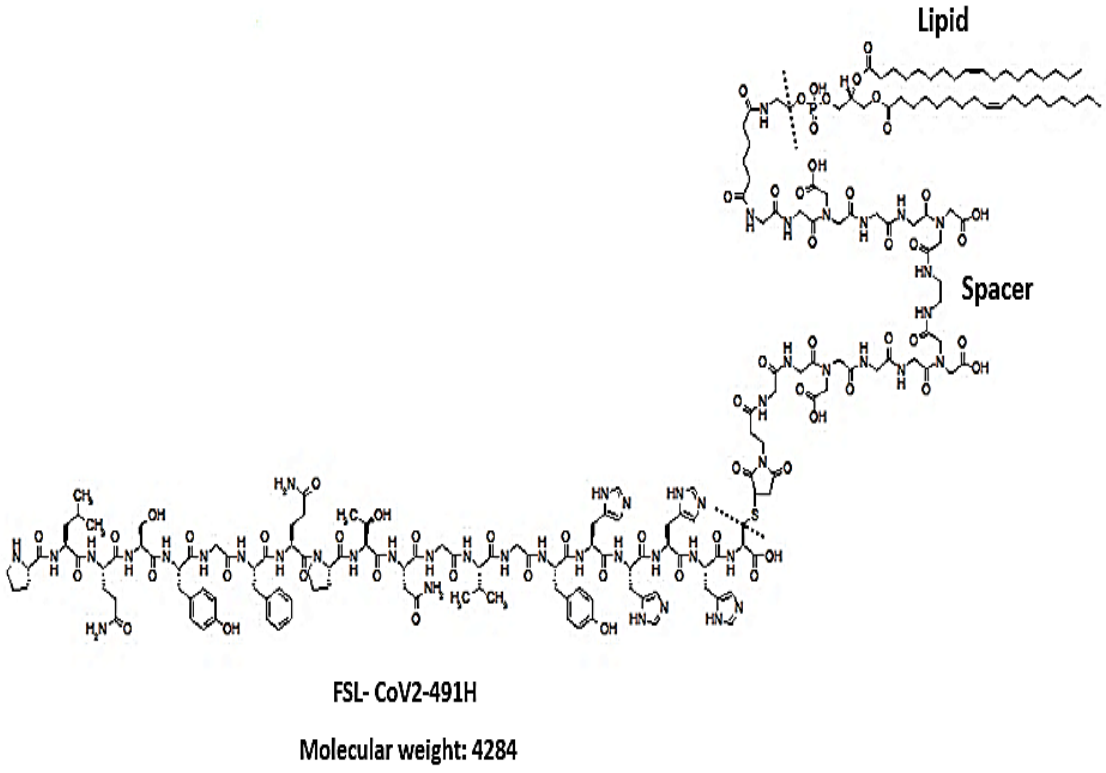

Figure 45: Schematic diagram of FSL-CoV-2-491H (PLQSYGFQPTNGVGY-HHHH) construct

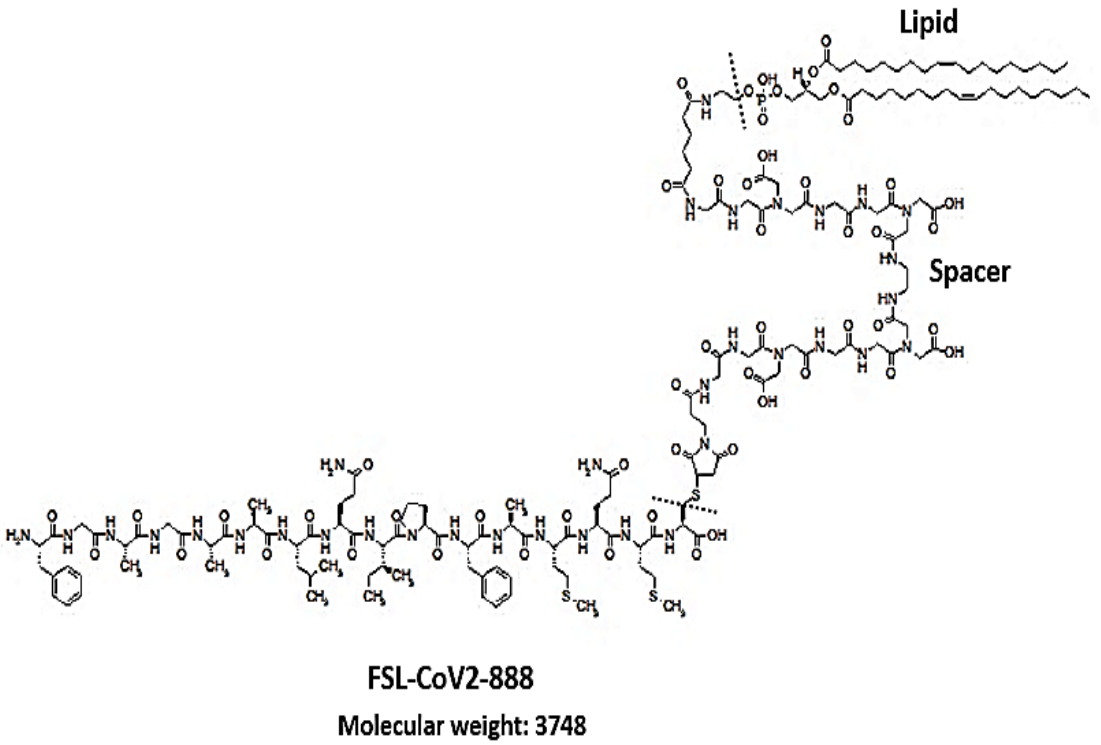

Figure 46: Schematic diagram of FSL-CoV-2-888 (FGAGAALQIPFAMQM) construct

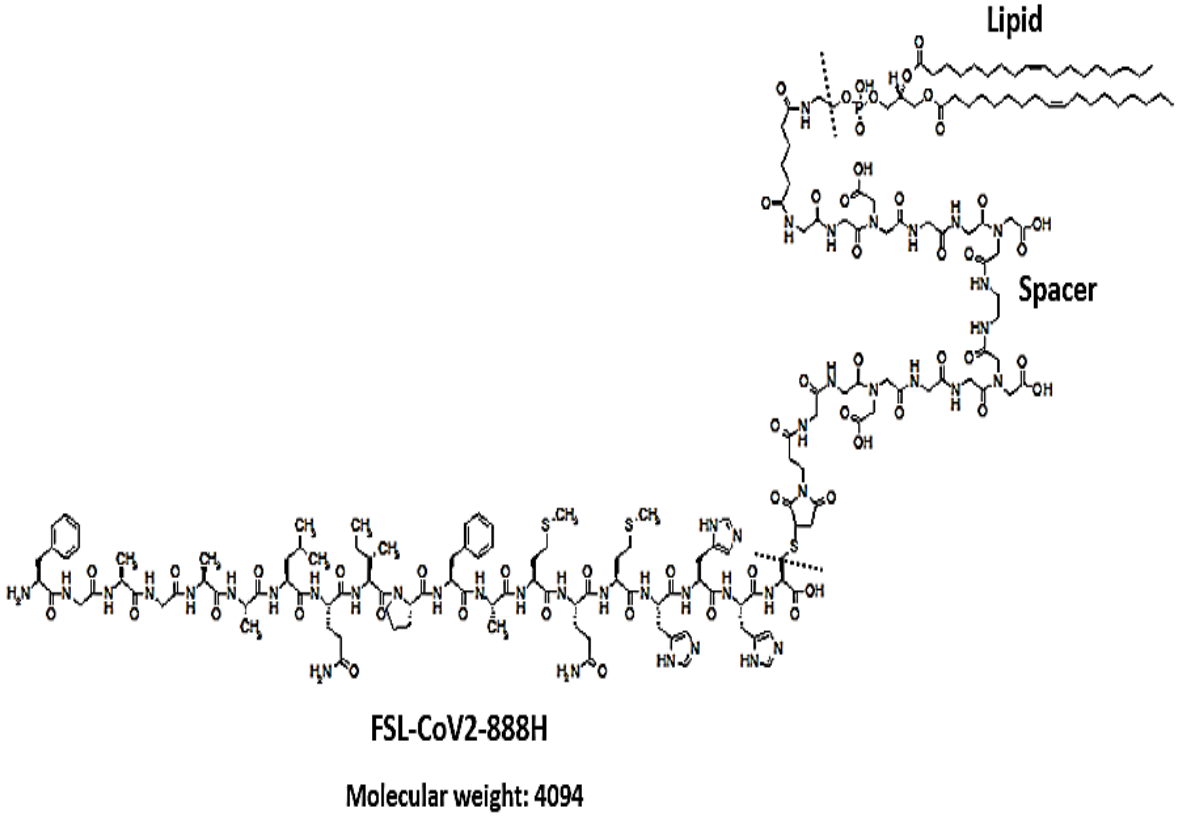

Figure 47: Schematic diagram of FSL-CoV-2-888H (FGAGAALQIPFAMQM-HHH) construct

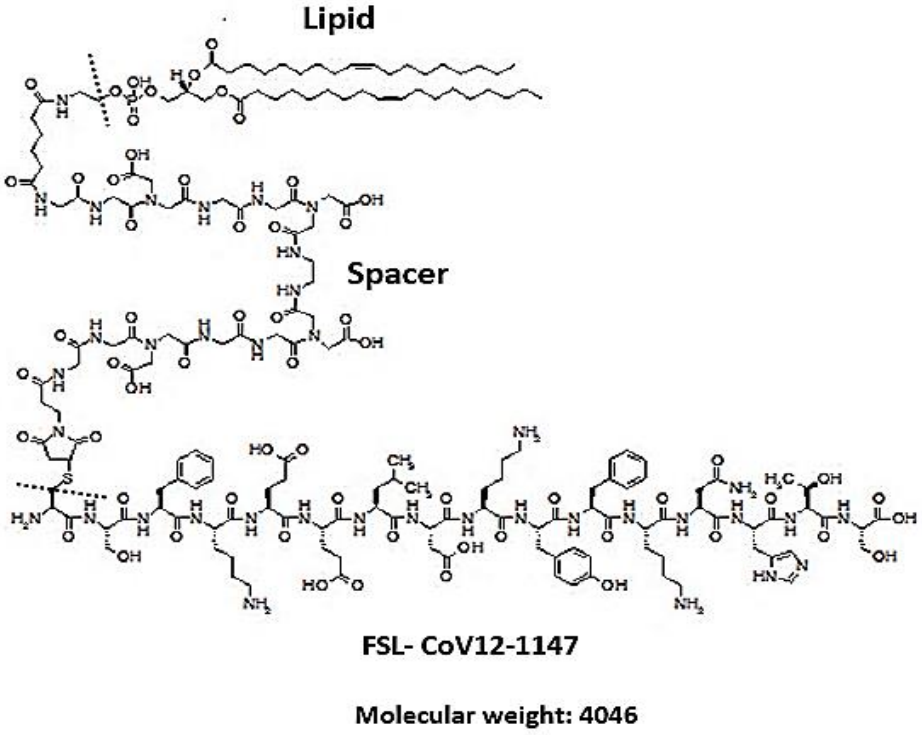

Figure 48: Schematic diagram of FSL-CoV- (1,2)-1147 (SFKEELDKYFKNHTS) construct

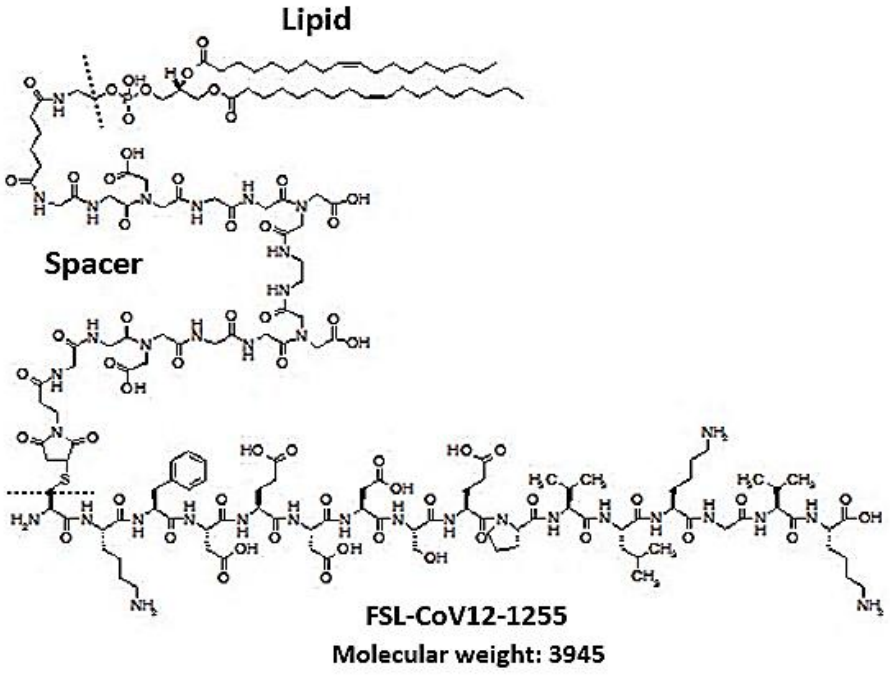

Figure 49: Schematic diagram of FSL-CoV- (1,2)-1255 (KFDEDDSEPVLKGVK) construct

A generic look at the full spike protein of SARS CoV2, using the FSL algorithm, selected out the eight best candidates, and identified two additional peptides with histidine tail sequences.

5.6. Methods and Results: Functional prediction of SARS CoV-2 FSL peptides

This study designed ten SARS CoV-2 peptides for FSL construction. The functional prediction of the SARS CoV-2 FSL constructs was done externally. The National Institute of Health (NIH), Bethesda, Maryland, USA, Centre for Kode Technology Innovation (c4KTI) and the Institute of Bioorganic Chemistry, Russian Academy of Science, Moscow, Russian Federation, undertook the serological analysis using SARS CoV-2 kodecytes (C19 kodecytes) for the functional prediction of SARS CoV-2 FSL peptide constructs (Appendix).

5.6.1. COVID-19 antibody screening with SARS CoV-2 red cell kodecytes

Detection of COVID-19 antibody using SARS CoV-2 kodecytes in routine serologic diagnostic platforms was done by NIH and c4KTI and subsequently published[234].

Materials

Donor and convalescent samples used for the functional prediction of the SARS CoV-2 FSL peptide constructs.

1. Plasma samples from New Zealand blood donors (control samples)
2. Convalescent samples (77) from Southern Community Laboratories (sera bank), New Zealand. These samples were confirmed SARS CoV-2 PCR positive and 59/77 samples were recorded as IgG antibody positive (Nucleocapsid) by Abbott SARS CoV-2 IgG antibody assay.
3. Convalescent samples (62) from donors recovered from PCR confirmed COVID-19 cases and 20 healthy blood donor samples from the Department of Transfusion Medicine, NIH, Bethesda, USA.

Preliminary screening of the kodecytes

Method overview

SARS CoV-2 kodecytes were made by mixing a solution of FSL construct with washed red cells, incubating at 37°C for two hours. The washed kodecytes were stored at 4° C overnight before use (see detailed protocol for kodecyte assay in Chapter 3). As part of the initial screening of the kodecytes for sensitivity and specificity, all the ten kodecytes were tested over a range of 3 to 20 µm/L concentration against expected negative samples. The kodecytes concentrations were then tested against convalescent samples to select the kodecytes with appropriate sensitivity and specificity. Kodecyte positive samples were also tested against non kodecytes and unrelated FSL control cells.

Results

The results of the SARS CoV-2 kodecyte assay, performed using Column Agglutination Test (CAT) to detect the COVID-19 antibody is shown in Table 37. The three kodecytes made with FSLs 808, 1147 and 1255 showed adequate sensitivity out of the ten kodecytes tested.

Table 37. Initial specificity and sensitivity analysis of all SARS CoV-2 kodecytes

			Specificity (expected negative)					Sensitivity (expected positive)		
			Manual serology (NZ)			CAT‡		CAT		
Kodecyte	µM*	N (NZ+US) †	IgM	IgM	AHG	nPos	% Pos	n§ (US)	N Pos	% Pos
ID			RT	37°		NZ+				
178	10	68+20	1	0	0	1+0	1	55	1	2
406	10	72+20	1	0	2			6	0	0
458	10	60+20	3	0	0	2+1	4	54	4	7
495	10	72+20	3	0	0			6	0	0
495H	10	72+20	4	0	2			6	0	0
808	10	100+20	0	0	0	3+3	5	50	23	46
888	10	72+20	2	0	0			6	0	0
888H	10	72+20	1	0	0			6	0	0
1147	3	100+20	0	0	2	5+2	6	52	44	85
1255	5	100+20	3	0	1	1+0	1	48	25	52
Untreated	0	§	0	0	0	0			0	

* µM refers to the concentration of FSL used to make the kodecyte. Results are only shown for the single concentration considered most appropriate for diagnostic use. Gaps in the table indicate analysis not done. † The number of samples is divided into samples from NZ and the US (separated by the + symbol) ‡ Column agglutination technology (CAT) results are split into results from NZ and the US, representing BioRad and MTS platforms, respectively. NZ+US nPos indicates the number of positive reactions by tester while % Positive uses combined CAT data to calculate percentage positive rate. § All samples reactive as positives also tested negative against unmodified cells used to make kodecytes. RT: Room Temperature.

FSL 808 showed only half the strength of FSL-1255 reaction, so was dropped from further validation. The two constructs which had good sensitivity, FSL-1147 and FSL-1255, were then formulated as single and dual epitope bearing kodecytes with two different micromolar concentrations for further analysis (3+5 µm/L and 1.2+2.5 µm/L). Better sensitivity and specificity were achieved with dual kodecytes (1147+1255-1.5+2.5), which were renamed as C19-kodecytes.

Evaluation of C19 kodecytes

Extended sensitivity analysis of the C19 kodecytes was done using the SCL-SB samples and the NIH convalescent donor samples recovered from PCR confirmed COVID-19 infection (Table 38).

Table 38. C19 kodecytes results with positive samples

COVID status	Antibody results	CAT FSL 1147+FSL1255 MTS (US)		BioRad (NZ)		Grifols (NZ)		EIA results	
PCR positive	Positive	49/54	91%	63/77	82%	75/77	97%	59/77	77%
Expected* Negatives	Positive	0/19	100%	4/100	4%	9/100	9%		

* Expected negative samples were from blood donors.

The Grifols DG system (NZ), reacted with 97%, the ortho MTS sytem with 91% and the Bio-Rad ID with 82% of these confirmed COVID-19 positive samples, compared to the 77% EIA result. Overall, the kodecytes assay was able to achieve sensitivity at least equivalent to the EIA method.

5.6.2. Summary

This study, as a first step in developing a SARS CoV-2 diagnostic Kode FSL peptide, used the FSL selection algorithm and various online tools to identify non-glycosylated epitopes on the S protein. The selected peptides were further refined to be compatible with FSL chemistry and eight candidate peptides were selected for FSL construction. SARS CoV-2 kodecytes were generated independently in two centres and the kodecyte assay was performed using three column agglutination test platforms (Grifols DG system, Ortho MTS system, and Bio-Rad ID). Evaluation samples included >120 expected negative samples and >140 expected positive samples from COVID-19 positive convalescent donor samples, with independent serological analysis from two centres. On initial evaluation of all the FSL constructs, three FSL constructs (808, 1147, and 1255) showed the most promising sensitivity (Table 37). Further testing of the peptides eliminated FSL-808 due to its lower sensitivity when compared to FSL-1147 and FSL-1255. The two remaining constructs, FSL-1147, and FSL-1255 were then formulated as single and dual epitope bearing kodecytes for further analysis[234]. The C19 kodecytes (1147+1255-1.5+2.5) which strongly bound to the plasma antibodies of COVID-19 convalescent patients when analyzed using kodecytes were used for further analysis. The specificity in three CAT platforms against the C19 kodecytes is >91 %, which correlates with the published literature. The sensitivity ranged from 82% to 97% compared to 77% with Abbott Architect SARS CoV-2 IgG assay. The C19 kodecytes assay gave 9% reactivity with the expected negative samples in the Grifols system and 4% in the Bio-Rad system. In addition to the kodecyte assay, external validation was also done by designing two solid phase EIA methods with the same peptides used to make the FSL constructs in order to observe the performance of these peptides in solid phase. In one of the solid phase EIA, the peptide was conjugated to PAA and was immobilized onto a microplate (PAA-EIA). In the second solid phase

EIA, the same FSL construct was immobilized in a solid phase lipid layer onto a microplate (FSL-EIA). The FSL-EIA outperformed (polyacrylamide conjugated peptides (peptide-PAA) PAA-EIA, suggesting the presentation of a small peptide epitope on FSL, as described in the earlier section, is a favorable form of presentation.

As seen in independent study, good sensitivity and specificity were achieved with dual kodecytes (C19-kodecytes) assay at least equivalent to the commercial Abbott Architect SARS CoV-2 EIA assay. Kode Technology is highly adaptable and enables the epitopes to be easily redesigned and made within a few weeks to allow rapid response to new strains arising due to mutations.

Chapter 6. Discussion

Diagnostics plays a major role in the management of infectious diseases. Despite advancements in molecular testing, antibodies cannot be detected by molecular techniques. As a result, rapid diagnostic tests that can detect antibodies serve as more promising tools for case management and disease control. Antibodies have long been used as diagnostic, prognostic and predictive biomarkers, especially in illnesses with a strong autoimmune or infectious etiology. As blood biomarkers, they are very helpful in disease diagnostics, used to check the adaptive immune response, facilitate contact tracing, conduct serological surveillance at regional, national and global levels, and to evaluate herd immunity to vaccination.

There is a need for rapidly adaptable, sensitive, low-cost, antibody-based diagnostics for the mass screening of existing diseases and there will continue to be an ongoing need for such tools to respond to emerging and re-emerging infectious diseases, particularly in resource limited settings.

In general, EIA is used as the most common rapid mass screening antibody based diagnostic, employed before more sophisticated tests can be used.

The EIA assay is ubiquitous in the diagnostic world and is considered the gold standard, although it is not without significant limitations. Most EIA antibody-based assays are simply based on the detection of antibodies binding to solid phase bound recombinant antigens and their subsequent visualisation with anti-Ig reagents labelled with enzymes which will produce a colorimetric reaction in proportion to the amount of antibody bound. The methodology is relatively simple, but, like all diagnostic assays, requires finding a balance between sensitivity and specificity by selecting and managing the concentrations of the components of the assay.

Several key differences exist between EIA assays and kodecyte assays:

(1) The antigen in an EIA is usually a large and complex recombinant protein(s) while the antigen in the kodecyte assay is always a small peptide epitope(s).

(2) The antigen of the EIA assay is large and generally cannot be easily refined (or fine-tuned), unlike the kodecyte antigen, which can be easily modified by simply selecting a slightly different peptide sequence.

(3) The kodecyte assay is much more limited in the number of epitopes involved, compared to EIA, and therefore can only bind a portion of the antibodies present in a polyclonal antibody response.

(4) Unlike EIA, kodecytes do not appear to be as affected by low affinity and nonspecific antibody binding of undiluted serum[199,234].

(5) Kodecyte assays are able to work with undiluted serum[235], while EIA assays must use diluted serum.

(6) The antigen in EIA is solid phased, adhered to an absorbent plastic membrane, while the kodecyte antigen is flexible and mobile and presented in a spacer format on a cell membrane.

(7) EIA assays require nonspecific antibody blocking while kodecyte assays require no blocking.

(8) EIA assays are generally very good at detecting IgM antibodies while kodecytes are poor at this.

(9) EIA assay is able to be stored for longer periods of time than kodecyte assays, which are limited to the life span (usually 2 months) of reagent red cells.

(10) The kodecytes assays reported in this thesis are β version (i.e., generally unrefined), compared with the results of commercial assays which have undergone extensive product refinement.

As can be seen by the non-exhaustive list above, the EIA and kodecyte assays have significant differences, yet they are both trying to determine the same outcome in an antibody-based diagnostic assay. Although EIA is considered to be the gold standard assay, care should still be exercised in interpretation of results derived from this paradigm (especially as concordance between EIA assays is not perfect even when they use the same antigens). Ultimately, the effectiveness the diagnostic assay should be measured by its correlation with clinical diagnosis and prognosis. It is not yet known whether EIA or kodecytes are better diagnostic platforms, and only further large-scale diagnostic trials correlating results with clinical diagnosis and prognosis will resolve this. Therefore, assessing the performance of the kodecyte assay against EIA (or other platform results) is inevitably skewed to the existing paradigm that the EIA result is the gold standard.

There are specific advantages the kodecyte assay has over EIA assays. These are: its low cost, the way it fits directly into the existing blood centre infrastructure, and its rapid adaptability and flexibility, as a kodecyte assay can be easily adjusted in the future to new targets.

The creation of a good kodecyte assay is dependent on quality selection of the best epitopes (unlike an EIA assay, which has almost all epitopes present on the recombinant protein). One approach to this selection of the epitopes could be to make everything possible, however this

would be very inefficient and expensive. Therefore, one objective of this thesis was to develop algorithms to find peptides potentially suitable for the kodecyte assays.

6.1. Identification of peptide epitopes and FSL selection algorithm

There are two main challenges faced in designing peptide based diagnostics: firstly, the identification of the peptide epitopes, and secondly, the presentation of the peptide with the correct orientation in a technological platform to capture the desired antibody efficiently.

One objective of this thesis is to develop an algorithm for identifying candidate peptides that are compatible with, and successful as, FSL constructs for kodecyte assays. FSL constructs are synthetically made molecules, so their structures can be controlled and engineered for better peptide orientation, resulting in good sensitivity.

An algorithm was developed for FSL peptide identification which incorporates several steps, starting from the target protein (diagnostic potential) sequence retrieval from GenBank database, epitope prediction, and the exclusion of microbial sequences using BLAST, peptide antigenicity, and various biophysical properties pertinent to peptide synthesis chemistry and FSL chemistry.

The target proteins were first tested for hydrophilicity, surface accessibility and glycosylation, the rationale being that the production of amino acid residues (epitopes) involved in antibody recognition and binding is facilitated by hydrophilic motifs and surface accessibility, and if such residues are glycosylated in the native protein, they probably cannot be used as synthetic peptides.

The peptide epitope (antigen) selection process was limited to linear epitopes in this study, as linear epitopes are easy to synthesize, unlike discontinuous epitopes (it is challenging to have the exact synthetic peptide mimic for discontinuous epitopes). The selection process was also limited to the availability of free online prediction tools.

Despite the fact that a range of computer algorithms for identifying B cell epitopes are now available, the accuracy of existing methods is still far from optimal[236]. Both sequences based and structure-based epitope identification have their own advantages and limitations, as mentioned earlier in section 1.4.1. For this study, it was decided to use the consensus method. In this approach, several prediction methods with various parameters and strengths are combined to achieve better results. The sequences with high scores and which were predicted as epitopes several times by different methods were taken for further analysis (peptide antigenicity, synthesis chemistry and FSL chemistry).

Peptide antigenicity is determined by several key factors such as secondary structures, surface accessibility, hydrophilicity and chemical composition. In the first instance, the immune system produces antibodies against surface exposed structures, and hence antibodies are considered as promising structures for epitope selection. Nonetheless, epitopes can be found within a protein structure and become exposed during protein degradation, making them immunogenic.

The peptides were tested for FSL peptide selection chemistry using bioinformatic programs for hydrophilicity, surface accessibility, glycosylation, presence of terminal glutamine and solubility. Ideally, the peptide sequence is without internal cysteine in the sequence for FSL construct due to conjugation chemistry, but this limitation can be resolved (section 2.3).

Syphilis peptides did not have solubility issues and all three syphilis peptides performed well at different concentrations. TmpA1 had more antigenicity scores than TmpA2 and TmpA3 and performed well at lower peptide concentration compared to TmpA2 and TmpA3. This might explain the lower sensitivity seen in TmpA2 and TmpA3 compared to TmpA1. The fact that all three syphilis peptides passed the external quality samples, but at different peptide concentrations, shows that the FSL peptide selection algorithm needs more fine-tuning.

The *Leptospira* peptides designed using the FSL identification algorithm, worked well overall, in spite of the caveats explained earlier. The *Leptospira* peptides need further validation using a large number of defined clinical samples for the verification of the FSL peptide design algorithm.

A total of ten SARS CoV-2 peptides were designed using FSL selection algorithm. CoV-1147 and CoV-1255 worked well, followed by CoV-808. The analysis of the CoV-1147, CoV-1255 and CoV-808 peptides shows antigenicity, hydrophilicity, surface accessibility and solubility are key determinants and reveals that the FSL selection algorithm needs further refinement.

The final algorithm decided on for FSL selection used a set series of parameter ranges, which were then used to decide on the most likely candidates to construct as FSL. Clearly, the approach was at least partially successful in that good diagnostically valid peptides resulted for each assay. However, not all predicted candidates showed appropriate sensitivity and specificity, despite often being the highest ranked candidates. This means there is still significant opportunity to revise and refine the algorithm and/or the immune system response to peptides cannot be accurately predicted, and a range of candidates will always need to be evaluated in a clinical setting.

6.2. Syphilis

Syphilis is curable by a simple antibiotic treatment, yet it has re-emerged globally in developing and developed countries, creating a substantial public health concern, particularly regarding a significant increase in congenital syphilis[237]. This type of syphilis is on the rise in New Zealand, where eight cases of congenital syphilis were documented between 2017 and 2018, with two stillbirths and two live births each year[238].

Diagnosis of syphilis is very challenging and remains elusive, in spite of a multitude of available tests and interpretations. This thesis designed three FSL peptide constructs from *T. pallidum* TmpA protein (FSL-TmpA1, FSL-TmpA2, FSL-TmpA3) to make syphilis kodecytes able to detect diagnostics antibodies to *T. pallidum*.

The verification of the syphilis FSL peptide constructs were done by testing these constructs against clinical samples for sensitivity, against healthy blood donor samples and against samples which were seroreactive for other pathogens for the specificity.

Syphilis kodecytes with varying concentrations were tested for specificity and sensitivity against expected positive convalescent, PCR confirmed samples and established syphilis kodecyte assay. Increasing the FSL concentration increases the sensitivity (ability to detect weak reactions), but decreases the specificity by detecting unwanted reactions. It is important to appreciate that these undesired weak reactions are real antibody antigen reactions and are probably cross reactive low affinity antibodies.

ROC curve analysis of the syphilis kodecytes shows that the optimal concentration to be used for the best performance of syphilis kodecytes was 10 µM for TmpA1, 50 µM for TmpA2, and 30 µM for TmpA3.

The analytical specificity of all three syphilis kodecytes (kodecyte TmpA1, kodecyte TmpA2 and kodecyte TmpA3) was 99.3% when tested against samples seroreactive for other pathogens and for autoantibodies, indicating syphilis kodecytes do not cross react with other pathogens.

The percentage specificity and the best peptide concentration of the TmpA1 kodecytes when tested against 197 EIA syphilis screen negative samples were 98%. TmpA2 kodecytes showed 97% specificity, and TmpA3 kodecytes 98 %.

The analytical sensitivity of syphilis kodecytes, when tested against EIA screen positive samples, showed good percentage sensitivity, with kodecytes TmpA1, TmpA2, and TmpA3 having 98.6%, 92.4% and 94.7% sensitivity, respectively. This is similar to the kodecytes concentration seen in the specificity analysis.

Syphilis kodecytes, when compared against 211 TPPA positive samples, showed 99.5% sensitivity for TmpA1 kodecytes, 94% sensitivity for TmpA2 kodecytes and 95% for sensitivity for TmpA3 kodecytes. Syphilis kodecytes results are comparable with the syphilis confirmatory test (TPPA) results.

TmpA1 kodecytes identified all 11 TPPA borderline positive samples, and TmpA2 and TmpA3 kodecytes identified the borderline TPPA positive samples at higher concentration. Some of the samples gave very strong serological grades, particularly with TmpA1 kodecytes, suggesting syphilis kodecyte assay has better sensitivity than TPPA.

Among the 211 EIA syphilis screen positive samples, 12 samples were TPPA negative. Syphilis kodecytes TmpA1 showed positive results for the TPPA negative samples. TmpA2 kodecytes and TmpA3 kodecytes showed positive results at higher concentrations. The fact that the 12 TPPA negative samples were EIA screen positives indicated that they could be primary or latent infections. Among the 12 TPPA negative samples, four samples were RPR positive. TPPA is known to give false negative results for syphilis screen (EIA) reactive samples. A laboratory audit in the UK identified a significant number of retrospective TPPA negative sera that came out positive when tested by other confirmatory assays[29]. Also, four of the TPPA negative samples were RPR positive, which could indicate a very early infection with low level antibodies, another reason for the TPPA result being negative. Syphilis kodecytes assay, particularly TmpA1, showed positive results for TPPA negative samples, indicating that syphilis kodecyte assays have higher sensitivity which result in false positives. The caveat of this study is the lack of clinical data for correlation.

Syphilis kodecytes identified all the 211 syphilis EIA screen positive samples. Out of the 211 EIA screen positive samples, 78 samples were EIA, TPPA, and the syphilis kodecytes assay were positive but the RPR test was negative. The traditional algorithm of using a non-treponemal test such as RPR for screening followed by a treponemal test for RPR screen positive patients, recommended by the CDC, is reliable in predicting disease activity[239-241]. It is cost effective and ideal for small laboratories and for POC testing. Two disadvantages of the RPR method are that it is manual and subjective. Moreover, RPR as a screening assay will miss the very early infections and the late-stage infections. The most convincing finding from a study (the Toronto study) showed that out of the 9137 patients who had initial EIA positive and RPR negative results, around 1 % seroconverted to RPR reactive, a diagnosis of early syphilis, in follow up studies held after eight weeks [242].

Several laboratories have started using the reverse algorithm, wherein treponemal specific EIA screen assays are used, and the EIA screen positive samples are tested with a non-treponemal test (RPR) to access the disease status and treatment[35]. The reverse algorithm testing is a

response to the high demand in syphilis serology test requests. Such testing can be automated for high throughput, above all the cost, but there are reports of discordant results that find many EIA screen positive samples are RPR negative. The CDC did a significant study on this and came up with the recommendation that EIA screen positives and RPR negatives should be tested with a second test (TPPA). If the TPPA result is positive, the infection is current or latent syphilis. If the TPPA result is negative, it is a false positive result, but this is not always true. In one of the studies, 560 samples were collected from patients with syphilitic lesions. These samples were tested for syphilis with a direct fluorescent antibody (DFA) test and alongside serology tests (EIA and RPR). All the 560 samples were DFA positive, and among the 560 samples, 18 samples were EIA positive and RPR negative[242].

To summarise, both traditional and reverse algorithms have advantages and disadvantages, and syphilis kodecytes results confirm kodecyte assay offer the best of both algorithms.

- Syphilis kodecytes identified 211 EIA screen-positive samples, showing syphilis kodecytes can be used as a high throughput screen assay.
- Among the 211 EIA screen positive samples, 136 samples were RPR negative. All these 136 samples were positive for syphilis with the kodecyte assay and TPPA. The results indicate that the traditional algorithm of screening with non-treponemal assay (RPR) and testing RPR positive samples only with confirmatory assay, will miss true positive samples (early/latent syphilis).
- Syphilis kodecytes identified the TPPA positive and TPPA borderline positive samples, indicating syphilis kodocyte assay performance is equivalent to syphilis confirmatory assay (TPPA).
- EIA screen positive and RPR positive samples should be tested with TPPA (reverse algorithm) to rule out false positive results or to identify early/latent syphilis. The EIA screen positive and RPR positive results (78) were confirmed positive with TPPA and syphilis kodecyte assay. These results suggest that syphilis kodecyte assay is specific and might be helpful, even if the number of early infections detected by EIA positive and RPR positive is low in number the kodecyte assay will help identify the syphilis cases and reduce the disease at the population level to a certain extent.
- This study also proves that the syphilis FSL peptide constructs do not cross react with other antibodies. Only one sample had a positive reaction out of 157 samples tested, and 156 samples were negative, suggesting that syphilis kodecytes have excellent specificity.
- Syphilis FSL peptide constructs performed well in the WHO/CDC EQA samples. All three kodecytes identified the four EQA samples correctly with neat serum samples, validating the peptides for diagnostic use.

- Analysis suggests the feasibility of using just TmpA1 at 10 µM concentration (i.e., TmpA1-kodecytes), for screening assay as well as for a confirmatory test, would perform at least as well as any existing syphilis assay. To implement this, large scale diagnostic trials and clinical follow ups must be performed.

6.3. Leptospirosis

Leptospirosis is a re-emerging zoonotic disease of global public health concern. Leptospirosis mimics other febrile illnesses like malaria, rickettsia and arbovirus, hence the difficulty in differential diagnosis. Early identification is essential for clinical diagnosis and treatment of leptospirosis. The lack of simple, easy to use diagnostic tests is the main reason why leptospirosis is difficult to manage. It is a major endemic environmental illness, typically associated with natural disasters such as floods and hurricanes, with a high risk of the incidence of catastrophic epidemics. Rapid identification of leptospirosis is essential for effective disease management and outbreak control in both human and animal populations. Public health officials and clinicians have raised the urgent need to develop more effective technologies for case detection, diagnosis and outbreak management[43].

The aim of this study is to design *Leptospira* FSL peptides and validate a sensitive and specific leptospirosis antibody diagnostic assay using FSL constructs. This thesis designed eight leptospiral FSL peptide constructs (FSL-LipL32, FSL-LipL21, FSL-Loa22, FSL-LipL41, FSL-LigA, FSL-LigB and FSL-LIC10215) from different proteins of *Leptospira* for leptospiral diagnostics.

Leptospira kodecyte assay was expected to be extremely challenging to develop, unlike an assay for syphilis, because no current assay for the diagnosis of leptospirosis is satisfactory. It was also more challenging to validate, as there was minimal access to samples, and those samples available could not be validated.

The analytical specificity of the eight leptospiral FSL constructs was tested with 100 healthy blood donor samples. Based on the learnings from syphilis, it was decided to use 30 µM concentration of the FSL constructs for the leptospiral kodecyte assay. The initial kodecytes results were further confirmed for true positivity with the leptospiral ELISA test. Kodecytes Loa22 and LigA showed 98% specificity. Kodecytes LigB, LipL21 and FlaB showed 97% specificity. Lipl32 has 99% specificity. Kodecytes LIC10215 and LipL41 have a lower specificity of 93% and 94%, respectively. The results show that most leptospiral kodecytes have good specificity, except LIC10215 and LipL41.

Leptopsira kodecytes were then tested against a small number of clinical samples with antibodies for hepatitis B virus, dengue virus and samples positive for autoantibodies, in order to verify the

analytical specificity of the leptospiral kodecytes. The reason for the selection of hepatitis B virus and dengue virus is the availability of these samples and their known cross reactivity with *Leptospira* serology.

All eight leptospiral kodecytes were tested against 11 samples with hepatitis B virus antibody positive, to check for cross reactivity. Seven leptospiral kodecytes showed 100% specificity; kodecytes LigB gave two positive results, demonstrating a specificity of 98%.

The results show that the leptospiral kodecytes have good specificity and do not cross react with hepatitis B virus antibody positive samples.

Leptospiral kodecytes were tested against eight samples seroreactive for autoantibodies to test the cross reactivity/analytical specificity of the leptospiral kodecytes. The results show that leptospiral kodecytes have 100% specificity and do not cross react with autoantibodies[243].

Leptospiral kodecytes were tested against ten dengue virus antibody positive samples. *Leptospira* ELISA antibody test confirmed some of the positive reactions and this finding corelates with the leptospirosis and dengue coinfection as stated in the literature[223]. The kodecytes LigB gave two false positive results, kodecytes FlaB gave three false positive results and the remaining gave one false positive result each out of the ten dengue positive samples tested. The overall specificity was 90%.

The analytical sensitivity of the eight leptospiral FSL constructs was tested with a small number (11) of different serovars of *Leptospira* positive samples (confirmed cases of leptospirosis diagnosed by MAT seroconversion on paired sera – acute and convalescent). The kodecytes LigA, LipL32 and LipL41 identified all the results correctly. As seen in the literature, LipL32 reacts with both acute and convalescent MAT positive samples. Kodecytes LigA reacted with all the serovars except Panama. Kodecytes LigA is not present in the serovar Panama. Kodecytes LIC10215 and LigB did not identify one positive sample, and kodecytes Loa22 and LipL21 did not identify some of the positive samples. Leptospiral FSL constructs have reasonable specificity.

It was challenging to get the clinical samples for the functional prediction of leptospiral FSL constructs. The caveats of this validation study are that the samples were not fresh, and the positive samples tested were few. To prove the actual sensitivity of the *Leptospira* FSL peptides constructs, a large number of fresh clinical samples need to be tested. The preliminary results show that the *Leptospira* kodecytes are working but further testing is needed to validate the kodecytes for clinical use.

The challenge in *Leptospira* diagnostics is that more than two hundred and fifty circulating leptospiral serovars are present and no single assay will pick up all serovars. Given the fact that preliminary results are looking good, the next step is to validate the peptide with a greater number of defined positive samples and negative samples with clinical follow up, followed by a large-scale study with external collaborators to validate the leptospiral peptides for routine diagnostics.

6.4. SARS CoV-2 and COVID-19

Due to the unexpected appearance of the COVID-19 pandemic while this research was being conducted, the study was extended to create a potential SARS Cov-2 antibody diagnostic assay.

Laboratory diagnosis of COVID-19 is crucial for understanding epidemiology, contact tracing, patient management and preventing SARS CoV-2 transmission. Although molecular techniques are used in the tests of choice for detecting the virus and for contact tracing, serology also plays a vital role in patient management, detecting the immune response of the individuals after infection and vaccination, antibody therapeutics and serosurveillance studies. From a public health perspective, these are all important for implementing intervention strategies. SARS CoV-2 Spike (S) and receptor binding domain protein are now being used in a rising number of commercial serological assays as practical methods to determine the extent of COVID-19 immunity in a population[234]. Despite this, there is still a need for further assays as viral mutations resulting in new lineages can rapidly arise, and this may affect the robustness of sensitivity and specificity of any specific assay.

The spike protein (S) was chosen as the candidate protein, and SARS CoV-2 kodecytes were generated using the SARS CoV-2 FSL peptide constructs (FSL-178, FSL-406, FSL-458, FSL-491, FSL-808, FSL-888, FSL-1147, FSL-1255, FSL-491H and FSL-888H). SARS CoV-2 kodecyte assay was performed using the routine blood serology Column Agglutination Test (CAT) test platforms to detect COVID-19 antibodies. From the initial eight candidates, three FSL constructs (FSL-808, FSL-1147, and FSL-1255) showed the best sensitivity and specificity. Further analysis showed FSL-808 is less sensitive compared to FSL-1147 and FSL-1255, and hence FSL-808 was excluded. The two remaining constructs, FSL-1147 and FSL-1255 were then formulated as single and dual epitopes for further analysis. Better sensitivity and specificity were achieved with dual kodecytes in the SARS CoV-2 kodecyte assay; these were renamed as C19-kodecytes.

External validation[235,244] of the COVID-19 antibody assay showed that the C19 kodecyte assay performed well, and that the sensitivity is equivalent to an established EIA assay. The specificity of the kodecyte assay was >91%, and the sensitivity of the C19 kodecyte assay was between 82%

and 97%, compared to 77% with the Abbott Architect SARS CoV-2 IgG assay. The results show that the C19 kodecytes have good specificity that differentiates convalescent patients from healthy blood donors, as confirmed by the CAT test results.

Independent research at the Institute of Bioorganic Chemistry, Russian Academy of Science, Moscow, Russian Federation also found the sensitivity and specificity of the dual epitopes (FSL-1147 and FSL-1255) kodecytes to be promising[234]. Both predicted antigenic FSL peptides (1147 and 1255) strongly bound to the plasma antibodies of COVID-19 convalescent patients, when analysed using kodecytes.

Two further C-19 studies done by the National Institutes of Health in the US[244] and by the German Red Cross[245], also show that good sensitivity and specificity and were achieved with the dual kodecytes assay, and this was suitable for the determination of vaccination status.

Overall, the C19-kodecyte assay, despite still being a β version requiring further tuning, achieved specificity and sensitivity at least equivalent to an established EIA antibody diagnostic. Further refinement of the C19-kdoecyte assay is currently in progress.

6.5. Conclusion

It should be noted that all the kodecyte assays described in this thesis are β versions and all assays will ultimately need to undergo regulatory approval processes and product development trials to be able to be implemented in the clinical field. However, the assays reported are functional and usable despite not yet being optimized. Improved sensitivity and specificity are expected to be achieved by shifting the existing peptide frame either two steps to the right or two steps to the left, or by potentially looking for additional peptides elsewhere in the protein. Similarly, optimizing the concentrations and the ratios of the epitopes in the kodecyte still has to be done. As seen in the SARS CoV-2 antibody assay, FSL-1147 and FSL-1255 in low concentration were not sensitive enough to identify some samples with low levels of antibody, but when FSL-1147 and FSL-1255 were put together (dual epitopes), they reacted with two antibodies in the sample and consequently were able to produce a positive result. Thus, a combination of select epitopes will give higher sensitivity.

One component of the thesis was the use of an algorithm to increase the likelihood of developing functional FSL peptide constructs with limited resources (i.e., not every possible candidate could be made as an FSL).

In all three diseases investigated, a small range of less than 10 kodecytes was made, and within this range, at least one kodecyte was always found which had the potential to be suitable for a

diagnostic assay. Potentially, over time, as more data is accumulated, the opportunity exists to refine this algorithm and increase the success rate, particularly as some of the constructs predicted most likely to be successful (e.g., SARS CoV-2 peptides 178, 406 and 458) were non-functional when tested. This clearly means the value of the algorithm is still limited and although it may help in the process of selection, it is not (yet) the ultimate predictive tool of FSL success. Consequently, a range of constructs will still need to be made for testing with clinical samples to find those of highest diagnostic value.

There are several advantages of kodecyte assay. Undiluted serum can be used for testing and it can be easily fitted into the existing blood group infrastructure. The kodecyte assay is highly flexible and adaptable and can be rapidly modified by simply changing the peptide/antigenic epitope on the Kode FSL construct. Fine tuning can be done to avoid the use of undesired epitopes that cannot be avoided in an EIA platform. Unlike recombinant peptides which require time and multiple considerations to change, any change in the epitope can be potentially achieved within a few weeks, allowing for rapid response to new strains arising with novel antigenic mutations. Kodecyte assays can be rapidly implemented technically in response to any new epidemic and pandemic.

Future work

The performance of the syphilis and SARS CoV-2 kodecyte assays was found to be equivalent to the syphilis and SARS CoV-2 commercial EIA. The future work is to implement this, through external collaborations, by doing large scale diagnostic trials, clinical follow ups, assay optimisation, regulatory approval and product development. The preliminary results of *Leptospira* FSL peptide constructs were promising, the peptides need further validation with well-defined clinical samples and assay optimisation.

This research built kodecyte diagnostics for three diseases and had success each time. The opportunity exists to further develop these successes into clinical products, but there is also an opportunity now to expand this approach into other infectious diseases which are detected by antibodies, such as malaria and dengue.

References

1. Fleming DT, Wasserheit JN. From epidemiological synergy to public health policy and practice: the contribution of other sexually transmitted diseases to sexual transmission of HIV infection. Sexually transmitted infections 1999;75:3-17.
2. Lafond RE, Lukehart SA. Biological basis for syphilis. Clin Microbiol Rev 2006;19:29-49.
3. Radolf JD, Deka RK, Anand A, Smajs D, Norgard MV, Yang XF. Treponema pallidum, the syphilis spirochete: making a living as a stealth pathogen. Nat Rev Microbiol 2016;14:744-59.
4. Schmid GP, Stoner BP, Hawkes S, Broutet N. The need and plan for global elimination of congenital syphilis. Sex Transm Dis 2007;34:S5-S10.
5. Kahn JG, Jiwani A, Gomez GB, et al. The cost and cost-effectiveness of scaling up screening and treatment of syphilis in pregnancy: a model. PLoS One 2014;9:e87510.
6. Cameron CE, Lukehart SA. Current status of syphilis vaccine development: need, challenges, prospects. Vaccine 2014;32:1602-9.
7. Tucker JD, Cohen MS. China's syphilis epidemic: epidemiology, proximate determinants of spread, and control responses. Curr Opin Infect Dis 2011;24:50-5.
8. ProMED- Syphilis - Australia (02): increasing cases of syphilis in indigenous population, 2018-10-30 , Archive Number: 20181030.6119580
9. ProMED-Syphilis - Japan (02): rising incidence, heterosexual women & men, Published Date: 2018-12-02, Archive Number: 20181202.6175741
10. Douglas JM. Penicillin treatment of syphilis. Jama 2009;301:769-71.
11. Karp G, Schlaeffer F, Jotkowitz A, Riesenberg K. Syphilis and HIV co-infection. European journal of internal medicine 2009;20:9-13.
12. White RG, Orroth KK, Korenromp EL, et al. Can population differences explain the contrasting results of the Mwanza, Rakai, and Masaka HIV/sexually transmitted disease intervention trials?: A modeling study. JAIDS Journal of Acquired Immune Deficiency Syndromes 2004;37:1500-13.
13. Perkins HA, Busch MP. Transfusion-associated infections: 50 years of relentless challenges and remarkable progress. Transfusion 2010;50:2080-99.
14. Adegoke AO, Akanni OE. Survival of Treponema pallidum in banked blood for prevention of Syphilis transmission. North American journal of medical sciences 2011;3:329.
15. Kaur G, Kaur P. Syphilis testing in blood donors: an update. Blood Transfusion 2015;13:197.

16. Tariciotti L, Das I, Dori L, Perera MT, Bramhall SR. Asymptomatic transmission of Treponema pallidum (syphilis) through deceased donor liver transplantation. Transpl Infect Dis 2012;14:321-5.
17. Cortes NJ, Afzali B, MacLean D, et al. Transmission of syphilis by solid organ transplantation. Am J Transplant 2006;6:2497-9.
18. Peeling RW, Hook EW, 3rd. The pathogenesis of syphilis: the Great Mimicker, revisited. J Pathol 2006;208:224-32.
19. Follett T, Clarke DF. Resurgence of congenital syphilis: diagnosis and treatment. Neonatal Network 2011;30:320-8.
20. Ficarra G, Carlos R. Syphilis: the renaissance of an old disease with oral implications. Head and neck pathology 2009;3:195-206.
21. Hall CS, Klausner JD, Bolan GA. Managing syphilis in the HIV-infected patient. Current infectious disease reports 2004;6:72-82.
22. Marra CM. Neurosyphilis. Curr Neurol Neurosci Rep 2004;4:435-40.
23. Ratnam S. The laboratory diagnosis of syphilis. Canadian Journal of Infectious Diseases and Medical Microbiology 2005;16:45-51.
24. Heymans R, van der Helm JJ, de Vries HJ, Fennema HS, Coutinho RA, Bruisten SM. Clinical value of Treponema pallidum real-time PCR for diagnosis of syphilis. J Clin Microbiol 2010;48:497-502.
25. Gayet-Ageron A, Ninet B, Toutous-Trellu L, et al. Assessment of a real-time PCR test to diagnose syphilis from diverse biological samples. Sex Transm Infect 2009;85:264-9.
26. Sena AC, White BL, Sparling PF. Novel Treponema pallidum serologic tests: a paradigm shift in syphilis screening for the 21st century. Clin Infect Dis 2010;51:700-8.
27. Larsen SA, Steiner BM, Rudolph AH. Laboratory diagnosis and interpretation of tests for syphilis. Clinical microbiology reviews 1995;8:1-21.
28. Singh AE, Romanowski B. Syphilis: review with emphasis on clinical, epidemiologic, and some biologic features. Clin Microbiol Rev 1999;12:187-209.
29. Maple PA, Ratcliffe D, Smit E. Characterization of Treponema pallidum particle agglutination assay-negative sera following screening by treponemal total antibody enzyme immunoassays. Clin Vaccine Immunol 2010;17:1718-22.
30. Herring AJ, Ballard RC, Pope V, et al. A multi-centre evaluation of nine rapid, point-of-care syphilis tests using archived sera. Sexually transmitted infections 2006;82:v7-v12.
31. Binnicker MJ, Jespersen DJ, Rollins LO. Treponema-specific tests for serodiagnosis of syphilis: comparative evaluation of seven assays. J Clin Microbiol 2011;49:1313-7.
32. Tong ML, Lin LR, Liu LL, et al. Analysis of 3 algorithms for syphilis serodiagnosis and implications for clinical management. Clin Infect Dis 2014;58:1116-24.

33. Radolf J, Bolan G, Park I, et al. Discordant Results From Reverse Sequence Syphilis Screening-Five Laboratories, United States, 2006-2010 (Reprinted from MMWR, vol 60, pg 133-137, 2011). JAMA-JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION 2011;305:1189-91.
34. Lee K, Park H, Roh EY, et al. Characterization of sera with discordant results from reverse sequence screening for syphilis. Biomed Res Int 2013;2013:269347.
35. Binnicker MJ. Which algorithm should be used to screen for syphilis? Curr Opin Infect Dis 2012;25:79-85.
36. Pastuszczak M, Wojas-Pelc A. Current standards for diagnosis and treatment of syphilis: selection of some practical issues, based on the European (IUSTI) and U.S. (CDC) guidelines. Postepy Dermatol Alergol 2013;30:203-10.
37. Lee D, Fairley C, Cummings R, Bush M, Read T, Chen M. Men who have sex with men prefer rapid testing for syphilis and may test more frequently using it. Sexually transmitted diseases 2010;37:557-8.
38. Mishra S, Naik B, Venugopal B, et al. Syphilis screening among female sex workers in Bangalore, India: comparison of point-of-care testing and traditional serological approaches. Sex Transm Infect 2010;86:193-8.
39. Peeling RW, Holmes KK, Mabey D, Ronald A. Rapid tests for sexually transmitted infections (STIs): the way forward. Sex Transm Infect 2006;82 Suppl 5:v1-6.
40. Organization WH. Global prevalence and incidence of selected curable sexually transmitted infections: overview and estimates. 2001.
41. Adler B, de la Peña Moctezuma A. Leptospira and leptospirosis. Veterinary microbiology 2010;140:287-96.
42. Evangelista KV, Coburn J. Leptospira as an emerging pathogen: a review of its biology, pathogenesis and host immune responses. Future microbiology 2010;5:1413-25.
43. Picardeau M, Bertherat E, Jancloes M, Skouloudis AN, Durski K, Hartskeerl RA. Rapid tests for diagnosis of leptospirosis: current tools and emerging technologies. Diagn Microbiol Infect Dis 2014;78:1-8.
44. Plank R, Dean D. Overview of the epidemiology, microbiology, and pathogenesis of Leptospira spp. in humans. Microbes and infection 2000;2:1265-76.
45. Levett PN. Leptospirosis. Clin Microbiol Rev 2001;14:296-326.
46. Goris MG, Leeflang MM, Loden M, et al. Prospective evaluation of three rapid diagnostic tests for diagnosis of human leptospirosis. PLoS Negl Trop Dis 2013;7:e2290.
47. Vijayachari P, Sugunan A, Shriram A. Leptospirosis: an emerging global public health problem. Journal of biosciences 2008;33:557-69.

48. Picardeau M. Virulence of the zoonotic agent of leptospirosis: still terra incognita? Nature Reviews Microbiology 2017;15:297.
49. de la Peña-Moctezuma A, Bulach DM, Adler B. Genetic differences among the LPS biosynthetic loci of serovars of Leptospira interrogans and Leptospira borgpetersenii. FEMS Immunology & Medical Microbiology 2001;31:73-81.
50. Nalam K, Ahmed A, Devi SM, et al. Genetic affinities within a large global collection of pathogenic Leptospira: implications for strain identification and molecular epidemiology. PLoS One 2010;5:e12637.
51. Victoriano AFB, Smythe LD, Gloriani-Barzaga N, et al. Leptospirosis in the Asia Pacific region. BMC infectious diseases 2009;9:147.
52. Haake DA, Levett PN. Leptospirosis in humans. Leptospira and leptospirosis: Springer; 2015:65-97.
53. Hartskeerl R, Collares-Pereira M, Ellis W. Emergence, control and re-emerging leptospirosis: dynamics of infection in the changing world. Clinical microbiology and infection 2011;17:494-501.
54. Marshall R, Manktelow B. Fifty years of leptospirosis research in New Zealand: a perspective. New Zealand Veterinary Journal 2002;50:61-3.
55. Science IoE, Ltd. R. Notifiable diseases in New Zealand: Annual report 2016. Institute of Environmental Science and Research Ltd Porirua, New Zealand; 2017.
56. Lau CL, Smythe LD, Craig SB, Weinstein P. Climate change, flooding, urbanisation and leptospirosis: fuelling the fire? Transactions of the Royal Society of Tropical Medicine and Hygiene 2010;104:631-8.
57. Lau C, Smythe L, Weinstein P. Leptospirosis: an emerging disease in travellers. Travel medicine and infectious disease 2010;8:33-9.
58. Agampodi SB, Peacock SJ, Thevanesam V, et al. Leptospirosis outbreak in Sri Lanka in 2008: lessons for assessing the global burden of disease. The American journal of tropical medicine and hygiene 2011;85:471-8.
59. Schneider MC, Nájera P, Aldighieri S, et al. Leptospirosis outbreaks in Nicaragua: identifying critical areas and exploring drivers for evidence-based planning. International journal of environmental research and public health 2012;9:3883-910.
60. Gaynor K, Katz AR, Park SY, Nakata M, Clark TA, Effler PV. Leptospirosis on Oahu: an outbreak associated with flooding of a university campus. The American journal of tropical medicine and hygiene 2007;76:882-6.
61. Goarant C, Laumond-Barny S, Perez J, Vernel-Pauillac F, Chanteau S, Guigon A. Outbreak of leptospirosis in New Caledonia: diagnosis issues and burden of disease. Tropical Medicine & International Health 2009;14:926-9.

62. Maskey M, Shastri J, Saraswathi K, Surpam R, Vaidya N. Leptospirosis in Mumbai: post-deluge outbreak 2005. Indian journal of medical microbiology 2006;24:337.
63. Abela-Ridder B, Sikkema R, Hartskeerl RA. Estimating the burden of human leptospirosis. International journal of antimicrobial agents 2010;36:S5-S7.
64. Leptospirosis WH. Guidance for diagnosis, surveillance and control. World Health Organization 2003;109.
65. Ahmad S, Shah S, Ahmad FH. Laboratory diagnosis of leptospirosis. Journal of postgraduate medicine 2005;51:195.
66. Budihal SV, Perwez K. Leptospirosis diagnosis: competancy of various laboratory tests. Journal of clinical and diagnostic research: JCDR 2014;8:199.
67. Vijayachari P, Sugunan A, Umapathi T, Sehgal S. Evaluation of darkground microscopy as a rapid diagnosis procedure in leptospirosis. Indian Journal of Medical Research 2001;114:54.
68. Bharti AR, Nally JE, Ricaldi JN, et al. Leptospirosis: a zoonotic disease of global importance. The Lancet Infectious Diseases 2003;3:757-71.
69. Picardeau M. Diagnosis and epidemiology of leptospirosis. Med Mal Infect 2013;43:1-9.
70. Musso D, La Scola B. Laboratory diagnosis of leptospirosis: a challenge. Journal of Microbiology, Immunology and Infection 2013;46:245-52.
71. Limmathurotsakul D, Turner EL, Wuthiekanun V, et al. Fool's gold: Why imperfect reference tests are undermining the evaluation of novel diagnostics: a reevaluation of 5 diagnostic tests for leptospirosis. Clinical infectious diseases 2012;55:322-31.
72. Silva M, Camargo E, Batista L, et al. Behaviour of specific IgM, IgG and IgA class antibodies in human leptospirosis during the acute phase of the disease and during convalescence. The Journal of tropical medicine and hygiene 1995;98:268-72.
73. Marín-León I, Pérez-Lozano MJ, de Villar-Conde E, Dastis-Bendala C, Vargas-Romero J, Pumarola-Suñé T. Prospective evaluation of the macroagglutination slide test for Leptospira. Serodiagnosis and immunotherapy in infectious disease 1997;8:191-3.
74. Effler PV, Domen HY, Bragg SL, Aye T, Sasaki DM. Evaluation of the indirect hemagglutination assay for diagnosis of acute leptospirosis in Hawaii. Journal of clinical microbiology 2000;38:1081-4.
75. Arimitsu Y, Kmety E, Ananyina Y, et al. Evaluation of the one-point microcapsule agglutination test for diagnosis of leptospirosis. Bulletin of the World Health Organization 1994;72:395.
76. Smits HL, Eapen CK, Sugathan S, et al. Lateral-flow assay for rapid serodiagnosis of human leptospirosis. Clin Diagn Lab Immunol 2001;8:166-9.

77. Eapen C, Sugathan S, Kuriakose M, Abdoel T, Smits HL. Evaluation of the clinical utility of a rapid blood test for human leptospirosis. Diagnostic microbiology and infectious disease 2002;42:221-5.
78. Merien F, Portnoi D, Bourhy P, Charavay F, Berlioz-Arthaud A, Baranton G. A rapid and quantitative method for the detection of Leptospira species in human leptospirosis. FEMS Microbiol Lett 2005;249:139-47.
79. Villumsen S, Pedersen R, Borre MB, Ahrens P, Jensen JS, Krogfelt KA. Novel TaqMan® PCR for detection of Leptospira species in urine and blood: pit-falls of in silico validation. Journal of microbiological methods 2012;91:184-90.
80. Boonsilp S, Thaipadungpanit J, Amornchai P, et al. Molecular detection and speciation of pathogenic Leptospira spp. in blood from patients with culture-negative leptospirosis. BMC infectious diseases 2011;11:338.
81. Ahmed SA, Sandai DA, Musa S, et al. Rapid diagnosis of leptospirosis by multiplex PCR. The Malaysian journal of medical sciences: MJMS 2012;19:9.
82. Perez J, Goarant C. Rapid Leptospira identification by direct sequencing of the diagnostic PCR products in New Caledonia. BMC microbiology 2010;10:325.
83. Djelouadji Z, Roux V, Raoult D, Kodjo A, Drancourt M. Rapid MALDI-TOF mass spectrometry identification of Leptospira organisms. Vet Microbiol 2012;158:142-6.
84. Rettinger A, Krupka I, Grünwald K, et al. Leptospira spp. strain identification by MALDI TOF MS is an equivalent tool to 16S rRNA gene sequencing and multi locus sequence typing (MLST). BMC microbiology 2012;12:185.
85. Calderaro A, Arcangeletti MC, Rodighiero I, et al. Matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass spectrometry applied to virus identification. Sci Rep 2014;4:6803.
86. Suwancharoen D, Sittiwicheanwong B, Wiratsudakul A. Evaluation of loop-mediated isothermal amplification method (LAMP) for pathogenic Leptospira spp. detection with leptospires isolation and real-time PCR. Journal of Veterinary Medical Science 2016:15-0702.
87. Picardeau M, Bertherat E, Jancloes M, Skouloudis AN, Durski K, Hartskeerl RA. Rapid tests for diagnosis of leptospirosis: current tools and emerging technologies. Diagnostic microbiology and infectious disease 2014;78:1-8.
88. Picardeau M, Bulach DM, Bouchier C, et al. Genome sequence of the saprophyte Leptospira biflexa provides insights into the evolution of Leptospira and the pathogenesis of leptospirosis. PloS one 2008;3.

89. Goris MG, Boer KR, Bouman-Strijker M, Hartskeerl R, Lucas C, Leeflang MM. Serological laboratory tests for diagnosis of human leptospirosis in patients presenting with clinical symptoms. Cochrane Database of Systematic Reviews 2011.
90. Sharun K, Dhama K, Pawde AM, et al. SARS-CoV-2 in animals: potential for unknown reservoir hosts and public health implications. Veterinary Quarterly 2021;41:181-201.
91. Maurin M, Fenollar F, Mediannikov O, Davoust B, Devaux C, Raoult D. Current status of putative animal sources of SARS-COV-2 infection in humans: Wildlife, domestic animals and pets. Microorganisms 2021;9:868.
92. Tang Q, Song Y, Shi M, Cheng Y, Zhang W, Xia X-Q. Inferring the hosts of coronavirus using dual statistical models based on nucleotide composition. Scientific reports 2015;5:1-8.
93. Alfaraj SH, Al-Tawfiq JA, Memish ZA. Middle East Respiratory Syndrome Coronavirus (MERS-CoV) infection during pregnancy: Report of two cases & review of the literature. 2019.
94. Satarker S, Nampoothiri M. Structural proteins in severe acute respiratory syndrome coronavirus-2. Archives of medical research 2020.
95. Troyano-Hernáez P, Reinosa R, Holguín Á. Evolution of SARS-CoV-2 envelope, membrane, nucleocapsid, and spike structural proteins from the beginning of the pandemic to September 2020: a global and regional approach by epidemiological week. Viruses 2021;13:243.
96. Huang Q, Yu L, Petros AM, et al. Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. Biochemistry 2004;43:6059-63.
97. Gralinski LE, Menachery VD. Return of the Coronavirus: 2019-nCoV. Viruses 2020;12:135.
98. Kumar M, Al Khodor S. Pathophysiology and treatment strategies for COVID-19. Journal of translational medicine 2020;18:1-9.
99. Cevik M, Kuppalli K, Kindrachuk J, Peiris M. Virology, transmission, and pathogenesis of SARS-CoV-2. bmj 2020;371.
100. Hui KP, Cheung M-C, Perera RA, et al. Tropism, replication competence, and innate immune responses of the coronavirus SARS-CoV-2 in human respiratory tract and conjunctiva: an analysis in ex-vivo and in-vitro cultures. The Lancet Respiratory Medicine 2020;8:687-95.
101. Vincent J-L, Taccone FS. Understanding pathways to death in patients with COVID-19. The Lancet Respiratory Medicine 2020;8:430-2.
102. Pang J, Wang MX, Ang IYH, et al. Potential rapid diagnostics, vaccine and therapeutics for 2019 novel coronavirus (2019-nCoV): a systematic review. Journal of clinical medicine 2020;9:623.
103. Krammer F, Simon V. Serology assays to manage COVID-19. Science 2020;368:1060-1.

104. Asrani P, Afzal Hussain KN, AlAjmi MF, Amir S, Sohal SS, Hassan MI. Guidelines and safety considerations in the laboratory diagnosis of SARS-CoV-2 infection: A prerequisite study for health professionals. Risk Management and Healthcare Policy 2021;14:379.
105. Gomara M, Ercilla G, Alsina M, Haro I. Assessment of synthetic peptides for hepatitis A diagnosis using biosensor technology. Journal of immunological methods 2000;246:13-24.
106. Mucci J, Carmona SJ, Volcovich R, et al. Next-generation ELISA diagnostic assay for Chagas Disease based on the combination of short peptidic epitopes. PLoS neglected tropical diseases 2017;11:e0005972.
107. Gomara M, Haro I. Synthetic peptides for the immunodiagnosis of human diseases. Current medicinal chemistry 2007;14:531-46.
108. Mishra AR, Hutke VR, Satav AR, et al. Synthetic Peptides are Better Than Native Antigens for Development of ELISA Assay for Diagnosis of Tuberculosis. International Journal of Peptide Research and Therapeutics 2017;23:247-57.
109. Williams RW, Chang A, Juretić D, Loughran S. Secondary structure predictions and medium range interactions. Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology 1987;916:200-4.
110. Chou PY, Fasman GD. Empirical predictions of protein conformation. Annual review of biochemistry 1978;47:251-76.
111. Sela-Culang I, Kunik V, Ofran Y. The structural basis of antibody-antigen recognition. Frontiers in immunology 2013;4:302.
112. Abbott WM, Damschroder MM, Lowe DC. Current approaches to fine mapping of antigen–antibody interactions. Immunology 2014;142:526-35.
113. Parker J, Guo D, Hodges R. New hydrophilicity scale derived from high-performance liquid chromatography peptide retention data: correlation of predicted surface residues with antigenicity and X-ray-derived accessible sites. Biochemistry 1986;25:5425-32.
114. Karplus P, Schulz G. Prediction of chain flexibility in proteins. NW 1985;72:212-3.
115. Emini EA, Hughes JV, Perlow D, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. Journal of virology 1985;55:836-9.
116. Bowie JU, Eisenberg D. An evolutionary approach to folding small alpha-helical proteins that uses sequence information and an empirical guiding fitness function. Proceedings of the National Academy of Sciences 1994;91:4436-40.
117. Blythe MJ, Flower DR. Benchmarking B cell epitope prediction: underperformance of existing methods. Protein Sci 2005;14:246-8.
118. Larsen JE, Lund O, Nielsen M. Improved method for predicting linear B-cell epitopes. Immunome Res 2006;2:2.

119. Anderson PH, Nielsen M, Lund O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structure. Protein Science 2006;15:2558-67.
120. Yang X, Yu X. An introduction to epitope prediction methods and software. Rev Med Virol 2009;19:77-96.
121. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 2008;9:40.
122. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER Suite: protein structure and function prediction. Nature methods 2015;12:7-8.
123. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 2010;5:725-38.
124. Zhang Y. I-TASSER: fully automated protein structure prediction in CASP8. Proteins 2009;77 Suppl 9:100-13.
125. Darnell S, Riese M. Precise Predictions of Linear B Cell Epitopes in Protean 3D, 2012. www.dnastar.com
126. Saha S, Raghava GP. Prediction methods for B-cell epitopes. Immunoinformatics: Springer; 2007:387-94.
127. Singh H, Ansari HR, Raghava GP. Improved method for linear B-cell epitope prediction using antigen's primary sequence. PLoS One 2013;8:e62216.
128. El-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using string kernels. J Mol Recognit 2008;21:243-55.
129. Petersen B, Petersen TN, Andersen P, Nielsen M, Lundegaard C. A generic method for assignment of reliability scores applied to solvent accessibility predictions. BMC Struct Biol 2009;9:51.
130. Kolaskar A, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS letters 1990;276:172-4.
131. Henry SM, Bovin NV. Kode Technology–a universal cell surface glycan modification technology. Journal of the Royal Society of New Zealand 2019;49:100-13.
132. Henry S, Williams E, Barr K, et al. Rapid one-step biotinylation of biological and non-biological surfaces. Scientific reports 2018;8:1-6.
133. Korchagina E, Henry S. Synthetic glycolipid-like constructs as tools for glycobiology research, diagnostics, and as potential therapeutics. Biochemistry (Moscow) 2015;80:857-71.
134. Katki HA, Kovalchik SA, Berg CD, Cheung LC, Chaturvedi AK. Development and Validation of Risk Models to Select Ever-Smokers for CT Lung Cancer Screening. JAMA 2016;315:2300-11.

135. Ryzhov IM, Tuzikov AB, Perry H, Korchagina EY, Bovin NV. Blood Group O→ A transformation by chemical ligation of erythrocytes. ChemBioChem 2019;20:131-3.
136. Blake DA, Bovin NV, Bess D, Henry SM. FSL constructs: a simple method for modifying cell/virion surfaces with a range of biological markers without affecting their viability. J Vis Exp 2011.
137. Williams E, Barr K, Korchagina E, Tuzikov A, Henry S, Bovin N. Ultra-fast glyco-coating of non-biological surfaces. International journal of molecular sciences 2016;17:118.
138. Henry S, Perry H, Bovin N. Applications for kodecytes in immunohaematology. ISBT Science Series 2018;13:229-37.
139. Henry S, Barr K, Oliver C. Modeling transfusion reactions with kodecytes and enabling ABO-incompatible transfusion with function-spacer-lipid constructs. ISBT Science Series 2012;7:106-11.
140. Stephan MT, Irvine DJ. Enhancing cell therapies from the outside in: cell surface engineering using synthetic nanomaterials. Nano today 2011;6:309-25.
141. Henry S, Komarraju S, Heathcote D, Rodionov I. Designing peptide-based FSL constructs to create Miltenberger kodecytes. ISBT Science Series 2011;6:306-12.
142. Sulzer C, Glosser J, Rogers F, Jones W, Frix M. Evaluation of an indirect hemagglutination test for the diagnosis of human leptospirosis. Journal of Clinical Microbiology 1975;2:218-21.
143. Levett PN, Whittington CU. Evaluation of the indirect hemagglutination assay for diagnosis of acute leptospirosis. Journal of clinical microbiology 1998;36:11-4.
144. Zochowski W, Palmer M, Coleman T. An evaluation of three commercial kits for use as screening methods for the detection of leptospiral antibodies in the UK. Journal of clinical pathology 2001;54:25-30.
145. Henry SM. Modification of red blood cells for laboratory quality control use. Curr Opin Hematol 2009;16:467-72.
146. Hult AK, Frame T, Chesla S, Henry S, Olsson ML. Flow cytometry evaluation of red blood cells mimicking naturally occurring ABO subgroups after modification with variable amounts of function-spacer-lipid A and B constructs. Transfusion 2012;52:247-51.
147. Heathcote D, Carroll T, Wang JJ, et al. Novel antibody screening cells, MUT+Mur kodecytes, created by attaching peptides onto red blood cells. Transfusion 2010;50:635-41.
148. Georgakopoulos T, Komarraju S, Henry S, Bertolini J. An improved Fc function assay utilizing CMV antigen-coated red blood cells generated with synthetic function–spacer–lipid constructs. Vox Sanguinis 2012;102:72-8.

149. Van Regenmortel M. Molecular dissection of protein antigens and the prediction of epitopes. Laboratory techniques in biochemistry and molecular biology 1988;19:1-39.
150. Baldwin RL. Structure and mechanism in protein science. A guide to enzyme catalysis and protein folding, by A. Fersht. 1999. New York: Freeman. 631 pp. $67.95 (hardcover). Protein Science 2000;9:207-.
151. Henchey LK, Jochim AL, Arora PS. Contemporary strategies for the stabilization of peptides in the alpha-helical conformation. Curr Opin Chem Biol 2008;12:692-7.
152. Shoemaker KR, Kim PS, Brems DN, et al. Nature of the charged-group effect on the stability of the C-peptide helix. Proceedings of the National Academy of Sciences 1985;82:2349-53.
153. Guharoy M, Chakrabarti P. Secondary structure based analysis and classification of biological interfaces: identification of binding motifs in protein-protein interactions. Bioinformatics 2007;23:1909-18.
154. Allison JR, Müller M, van Gunsteren WF. A comparison of the different helices adopted by α-and β-peptides suggests different reasons for their stability. Protein science 2010;19:2186-95.
155. Richardson JS, Richardson DC. Natural β-sheet proteins use negative design to avoid edge-to-edge aggregation. Proceedings of the National Academy of Sciences 2002;99:2754-9.
156. Huang F, Nau WM. A conformational flexibility scale for amino acids in peptides. Angew Chem Int Ed Engl 2003;42:2269-72.
157. Abraham GN, Podell DN. Pyroglutamic acid. Molecular and Cellular Biochemistry 1981;38:181-90.
158. Dick LW, Jr., Kim C, Qiu D, Cheng KC. Determination of the origin of the N-terminal pyro-glutamate variation in monoclonal antibodies using model peptides. Biotechnol Bioeng 2007;97:544-53.
159. Christlet THT, Veluraja K. Database analysis of O-glycosylation sites in proteins. Biophysical journal 2001;80:952-60.
160. Medzihradszky KF. Characterization of site-specific N-glycosylation. Post-translational Modifications of Proteins: Springer; 2008:293-316.
161. Spiro RG. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology 2002;12:43R-56R.
162. Manning MC, Patel K, Borchardt RT. Stability of protein pharmaceuticals. Pharmaceutical research 1989;6:903-18.

163. Geiger T, Clarke S. Deamidation, isomerization, and racemization at asparaginyl and aspartyl residues in peptides. Succinimide-linked reactions that contribute to protein degradation. Journal of Biological Chemistry 1987;262:785-94.
164. Robinson N, Robinson A. Prediction of primary structure deamidation rates of asparaginyl and glutaminyl peptides through steric and catalytic effects. The Journal of peptide research 2004;63:437-48.
165. Wlodawer A, Miller M, Jaskolski M, et al. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. Science 1989;245:616-21.
166. Tang XL, Tregear GW, White D, Jackson D. Minimum requirements for immunogenic and antigenic activities of homologs of a synthetic peptide of influenza virus hemagglutinin. Journal of virology 1988;62:4745-51.
167. Van Regenmortel MH. What is a B-cell epitope? Epitope Mapping Protocols: Springer; 2009:3-20.
168. Duranti MA, Franzoni L, Sartor G, et al. Trypanosoma cruzi: conformational preferences of antigenic peptides bearing the immunodominant epitope of the B13 antigen. Experimental parasitology 1999;93:38-44.
169. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. Nucleic acids research 2008;36:W5-W9.
170. Corpet F. Multiple sequence alignment with hierarchical clustering. Nucleic acids research 1988;16:10881-90.
171. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic acids research 2017;45:W24-W9.
172. Geourjon C, Deleage G. SOPM: a self-optimized method for protein secondary structure prediction. Protein Engineering, Design and Selection 1994;7:157-64.
173. Gupta R, Jung E, Brunak S. Prediction of N-glycosylation sites in human proteins, 2002 Pacific Symposium on Biocomputing 7:310-322 (2002)
174. Steentoft C, Vakhrushev SY, Joshi HJ, et al. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. The EMBO journal 2013;32:1478-88.
175. Norris S. Polypeptides of Treponema pallidum: progress toward understanding their structural, functional, and immunologic roles. Treponema Pallidum Polypeptide Research Group. Microbiology and Molecular Biology Reviews 1993;57:750-79.
176. Fraser CM, Norris SJ, Weinstock GM, et al. Complete genome sequence of Treponema pallidum, the syphilis spirochete. Science 1998;281:375-88.

177. Cox DL, Luthra A, Dunham-Ems S, et al. Surface immunolabeling and consensus computational framework to identify candidate rare outer membrane proteins of Treponema pallidum. Infection and immunity 2010;78:5178-94.
178. Heymans R, Kolader M-E, Van Der Helm J, Coutinho R, Bruisten S. TprK gene regions are not suitable for epidemiological syphilis typing. European journal of clinical microbiology & infectious diseases 2009;28:875-8.
179. McGill MA, Edmondson DG, Carroll JA, Cook RG, Orkiszewski RS, Norris SJ. Characterization and serologic analysis of the Treponema pallidum proteome. Infection and immunity 2010;78:2631-43.
180. Brinkman MB, McKevitt M, McLoughlin M, et al. Reactivity of antibodies from syphilis patients to a protein array representing the Treponema pallidum proteome. Journal of clinical microbiology 2006;44:888-91.
181. Luthra A, Zhu G, Desrosiers DC, et al. The transition from closed to open conformation of Treponema pallidum outer membrane-associated lipoprotein TP0453 involves membrane sensing and integration by two amphipathic helices. Journal of Biological Chemistry 2011;286:41656-68.
182. Van Voorhis WC, Barrett LK, Lukehart SA, Schmidt B, Schriefer M, Cameron CE. Serodiagnosis of Syphilis: Antibodies to Recombinant Tp0453, Tp92, and Gpd Proteins Are Sensitive and Specific Indicators of Infection by Treponema pallidum. Journal of Clinical Microbiology 2003;41:3668-74.
183. Smith BC, Simpson Y, Morshed MG, et al. New proteins for a new perspective on syphilis diagnosis. J Clin Microbiol 2013;51:105-11.
184. Titz B, Rajagopala SV, Goll J, et al. The binary protein interactome of Treponema pallidum–the syphilis spirochete. PloS one 2008;3:e2292.
185. Rajagopala SV, Titz B, Goll J, et al. The protein network of bacterial motility. Molecular systems biology 2007;3:128.
186. Tan M, Xu M, Xiao Y, et al. Screening and identification of immunoactive FlaB protein fragments of Treponema pallidum for the serodiagnosis of syphilis. Pathogens and disease 2018;76:ftx122.
187. Yang J, Shen L, Zhang X, Sun Q. Soluble expression, purification and characterization of recombinant Tp0136 selective fragment from Treponema pallidum. Chinese Journal of Microbiology and Immunology 2011;31:119-23.
188. Cameron CE, Kuroiwa JM, Yamada M, Francescutti T, Chi B, Kuramitsu HK. Heterologous expression of the Treponema pallidum laminin-binding adhesin Tp0751 in the culturable spirochete Treponema phagedenis. Journal of bacteriology 2008;190:2565-71.

189. Lithgow KV, Hof R, Wetherell C, Phillips D, Houston S, Cameron CE. A defined syphilis vaccine candidate inhibits dissemination of Treponema pallidum subspecies pallidum. Nature communications 2017;8:1-10.
190. Setubal JC, Reis M, Matsunaga J, Haake DA. Lipoprotein computational prediction in spirochaetal genomes. Microbiology (Reading, England) 2006;152:113.
191. Baughn RE, Jiang A, Abraham R, Ottmers V, Musher DM. Molecular mimicry between an immunodominant amino acid motif on the 47-kDa lipoprotein of Treponema pallidum (Tpp47) and multiple repeats of analogous sequences in fibronectin. The Journal of Immunology 1996;157:720-31.
192. Xie Y, Xu M, Wang C, et al. Diagnostic value of recombinant Tp0821 protein in serodiagnosis for syphilis. Letters in applied microbiology 2016;62:336-43.
193. Sambri V, Marangoni A, Simone M, D'antuono A, Negosanti M, Cevenini R. Evaluation of recomWell Treponema, a novel recombinant antigen-based enzyme-linked immunosorbent assay for the diagnosis of syphilis. Clinical microbiology and infection 2001;7:200-5.
194. Backhouse JL, Nesteroff SI. Treponema pallidum western blot: comparison with the FTA-ABS test as a confirmatory test for syphilis. Diagnostic microbiology and infectious disease 2001;39:9-14.
195. Ijsselmuiden O, Schouls L, Stolz E, et al. Sensitivity and specificity of an enzyme-linked immunosorbent assay using the recombinant DNA-derived Treponema pallidum protein TmpA for serodiagnosis of syphilis and the potential use of TmpA for assessing the effect of antibiotic therapy. Journal of Clinical Microbiology 1989;27:152-7.
196. Yelton D, Limberger R, Curci K, et al. Treponema phagedenis encodes and expresses homologs of the Treponema pallidum TmpA and TmpB proteins. Infection and immunity 1991;59:3685-93.
197. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic acids research 2007;35:W407-W10.
198. Sippl MJ. Recognition of errors in three-dimensional structures of proteins. Proteins: Structure, Function, and Bioinformatics 1993;17:355-62.
199. Perry H, Bovin N, Henry S. A standardized kodecyte method to quantify ABO antibodies in undiluted plasma of patients before ABO-incompatible kidney transplantation. Transfusion 2019;59:2131-40.
200. Zweig MH, Campbell G. Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. Clinical Chemistry 1993;39:561-77.

201. Cullen PA, Cordwell SJ, Bulach DM, Haake DA, Adler B. Global analysis of outer membrane proteins from Leptospira interrogans serovar Lai. Infection and immunity 2002;70:2311-8.
202. Haake DA, Mazel MK, McCoy AM, et al. Leptospiral outer membrane proteins OmpL1 and LipL41 exhibit synergistic immunoprotection. Infection and immunity 1999;67:6572-82.
203. Pinne M, Haake DA. A comprehensive approach to identification of surface-exposed, outer membrane-spanning proteins of Leptospira interrogans. PLoS One 2009;4:e6071.
204. Haake DA, Matsunaga J. Leptospira: a spirochaete with a hybrid outer membrane. Molecular microbiology 2010;77:805-14.
205. Cullen PA, Xu X, Matsunaga J, et al. Surfaceome of Leptospira spp. Infection and immunity 2005;73:4853-63.
206. Murray GL, Srikram A, Hoke DE, et al. Major surface protein LipL32 is not required for either acute or chronic infection with Leptospira interrogans. Infection and immunity 2009;77:952-8.
207. Shang ES, Summers TA, Haake DA. Molecular cloning and sequence analysis of the gene encoding LipL41, a surface-exposed lipoprotein of pathogenic Leptospira species. Infection and immunity 1996;64:2322-30.
208. Asuthkar S, Velineni S, Stadlmann J, Altmann F, Sritharan M. Expression and characterization of an iron-regulated hemin-binding protein, HbpA, from Leptospira interrogans serovar Lai. Infection and immunity 2007;75:4582-91.
209. Cullen PA, Haake DA, Bulach DM, Zuerner RL, Adler B. LipL21 is a novel surface-exposed lipoprotein of pathogenic Leptospira species. Infection and immunity 2003;71:2414-21.
210. Cerqueira GM, McBride AJ, Picardeau M, et al. Distribution of the leptospiral immunoglobulin-like (lig) genes in pathogenic Leptospira species and application of ligB to typing leptospiral isolates. Journal of medical microbiology 2009;58:1173.
211. Verma A, Brissette CA, Bowman AA, Shah ST, Zipfel PF, Stevenson B. Leptospiral endostatin-like protein A is a bacterial cell surface receptor for human plasminogen. Infection and immunity 2010;78:2053-9.
212. Oliveira TR, Longhi MT, de Morais ZM, et al. Evaluation of leptospiral recombinant antigens MPL17 and MPL21 for serological diagnosis of leptospirosis by enzyme-linked immunosorbent assays. Clinical and Vaccine Immunology 2008;15:1715-22.
213. Haake D, Champion C, Martinich C, et al. Molecular cloning and sequence analysis of the gene encoding OmpL1, a transmembrane outer membrane protein of pathogenic Leptospira spp. Journal of bacteriology 1993;175:4225-34.

214. Kositanont U, Saetun P, Krittanai C, Doungchawee G, Tribuddharat C, Thongboonkerd V. Application of immunoproteomics to leptospirosis: towards clinical diagnostics and vaccine discovery. Proteomics – Clinical Applications 2007;1:400-9.
215. Picardeau M, Brenot A, Saint Girons I. First evidence for gene replacement in Leptospira spp. Inactivation of L. biflexa flaB results in non-motile mutants deficient in endoflagella. Molecular microbiology 2001;40:189-99.
216. Lessa-Aquino C, Rodrigues CB, Pablo J, et al. Identification of seroreactive proteins of Leptospira interrogans serovar copenhageni using a high-density protein microarray approach. PLoS Negl Trop Dis 2013;7:e2499.
217. Lessa-Aquino C, Wunder Jr EA, Lindow JC, et al. Proteomic features predict seroreactivity against leptospiral antigens in leptospirosis patients. Journal of proteome research 2015;14:549-56.
218. Lin M, Surujballi O, Nielsen K, Nadin-Davis S, Randall G. Identification of a 35-kilodalton serovar-cross-reactive flagellar protein, FlaB, from Leptospira interrogans by N-terminal sequencing, gene cloning, and sequence analysis. Infection and immunity 1997;65:4355-9.
219. Bajani MD, Ashford DA, Bragg SL, et al. Evaluation of four commercially available rapid serologic tests for diagnosis of leptospirosis. Journal of clinical microbiology 2003;41:803-9.
220. Gussenhoven GC, Van der Hoorn M, Goris M, et al. LEPTO dipstick, a dipstick assay for detection of Leptospira-specific immunoglobulin M antibodies in human sera. Journal of Clinical Microbiology 1997;35:92-7.
221. Smits HL, Eapen C, Sugathan S, et al. Lateral-flow assay for rapid serodiagnosis of human leptospirosis. Clin Diagn Lab Immunol 2001;8:166-9.
222. shaw s. Cross Reactivity Observed in Patients Whose Blood Samples Tested Positive for Leptospira IgM and Dengue IgM Antibodies.
223. Sachu A, Madhavan A, Vasudevan A, Vasudevapanicker J. Prevalence of dengue and leptospirosis co-infection in a tertiary care hospital in south India. Iranian journal of microbiology 2018;10:227.
224. Crooke SN, Ovsyannikova IG, Kennedy RB, Poland GA. Immunoinformatic identification of B cell and T cell epitopes in the SARS-CoV-2 proteome. Scientific reports 2020;10:1-15.
225. Musicò A, Frigerio R, Mussida A, et al. SARS-CoV-2 epitope mapping on microarrays highlights strong immune-response to N protein region. Vaccines 2021;9:35.
226. Jiang H-w, Li Y, Zhang H-n, et al. Global profiling of SARS-CoV-2 specific IgG/IgM responses of convalescents using a proteome microarray. MedRxiv 2020.

227. Dutta NK, Mazumdar K, Gordy JT. The nucleocapsid protein of SARS–CoV-2: a target for vaccine development. Journal of virology 2020;94.
228. Infantino M, Damiani A, Gobbi FL, et al. Serological assays for SARS-CoV-2 infectious disease: benefits, limitations and perspectives. Isr Med Assoc J 2020;22:203-10.
229. De Assis RR, Jain A, Nakajima R, et al. Analysis of SARS-CoV-2 antibodies in COVID-19 convalescent blood using a coronavirus antigen microarray. Nature communications 2021;12:1-9.
230. Iyer AS, Jones FK, Nodoushani A, et al. Dynamics and significance of the antibody response to SARS-CoV-2 infection. MedRxiv 2020.
231. Goh YS, Chavatte J-M, Jieling AL, et al. Sensitive detection of total anti-Spike antibodies and isotype switching in asymptomatic and symptomatic individuals with COVID-19. Cell Reports Medicine 2021;2:100193.
232. Perkmann T, Perkmann-Nagele N, Koller T, et al. Anti-Spike protein assays to determine post-vaccination antibody levels: a head-to-head comparison of five quantitative assays. MedRxiv 2021.
233. Alves D, Curvello R, Henderson E, et al. Rapid gel card agglutination assays for serological analysis following SARS-CoV-2 infection in humans. ACS sensors 2020;5:2596-603.
234. Nagappan R, Flegel WA, Srivastava K, et al. COVID-19 antibody screening with SARS-CoV-2 red cell kodecytes using routine serologic diagnostic platforms. Transfusion 2021;61:1171-80.
235. Ryzhov IM, Tuzikov AB, Nizovtsev AV, et al. SARS-CoV-2 Peptide Bioconjugates Designed for Antibody Diagnostics. Bioconjugate Chemistry 2021;32:1606-16.
236. El-Manzalawy Y, Honavar V. Recent advances in B-cell epitope prediction methods. Immunome Res 2010;6 Suppl 2:S2.
237. Stamm LV. Global challenge of antibiotic-resistant Treponema pallidum. Antimicrob Agents Chemother 2010;54:583-9.
238. Gilmour LS, Best EJ, Duncanson MJ, et al. High Incidence of Congenital Syphilis in New Zealand: A New Zealand Pediatric Surveillance Unit Study. The Pediatric Infectious Disease Journal 2021.
239. Control CfD, Prevention. Discordant results from reverse sequence syphilis screening--five laboratories, United States, 2006-2010. MMWR Morbidity and mortality weekly report 2011;60:133-7.
240. Owusu-Edusei Jr K, Peterman TA, Ballard RC. Serologic testing for syphilis in the United States: a cost-effectiveness analysis of two screening algorithms. Sexually transmitted diseases 2011;38:1-7.

241. Control CfD, Prevention. Syphilis testing algorithms using treponemal tests for initial screening--four laboratories, New York City, 2005-2006. MMWR Morbidity and mortality weekly report 2008;57:872.
242. Mishra S, Boily MC, Ng V, et al. The laboratory impact of changing syphilis screening from the rapid-plasma reagin to a treponemal enzyme immunoassay: a case-study from the Greater Toronto Area. Sex Transm Dis 2011;38:190-6.
243. Bajani MD, Ashford DA, Bragg SL, et al. Evaluation of Four Commercially Available Rapid Serologic Tests for Diagnosis of Leptospirosis. Journal of Clinical Microbiology 2003;41:803-9.
244. Srivastava K, West KA, De Giorgi V, et al. COVID-19 antibody detection and assay performance using red cell agglutination. medRxiv 2021.
245. Weinstock C, Flegel WA, Srivastava K, et al. Erytra Blood Group Analyser and Kode Technology testing of SARS-CoV-2 antibodies among convalescent patients and vaccinated individuals. medRxiv 2021.

Appendix: COVID-19 antibody screening with SARS CoV-2 kodecytes

Received: 15 December 2020 | Revised: 31 January 2021 | Accepted: 1 February 2021
DOI: 10.1111/trf.16327

DONOR INFECTIOUS DISEASE TESTING

TRANSFUSION

COVID-19 antibody screening with SARS-CoV-2 red cell kodecytes using routine serologic diagnostic platforms

Radhika Nagappan[1] | Willy A. Flegel[2] | Kshitij Srivastava[2] | Eleanor C. Williams[1] | Ivan Ryzhov[3] | Alexander Tuzikov[3] | Oxana Galanina[3] | Nadezhda Shilova[3] | Gennady Sukhikh[4] | Holly Perry[1,5] | Nicolai V. Bovin[1,3] | Stephen M. Henry[1]

[1]Centre for Kode Technology Innovation, School of Engineering, Computer and Mathematical Sciences, Faculty of Design and Creative Technologies, Auckland University of Technology, Auckland, New Zealand

[2]Department of Transfusion Medicine, NIH Clinical Center, National Institutes of Health, Bethesda, Maryland, USA

[3]Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Moscow, Russia

[4]Kulakov National Medical Research Center for Obstetrics, Gynecology and Perinatology, Moscow, Russia

[5]School of Science, Faculty of Health and Environmental Sciences, Auckland University of Technology, Auckland, New Zealand

Correspondence
Stephen M. Henry, Centre for Kode Technology Innovation, School of Engineering, Faculty of Design and Creative Technologies, Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand.
Email: kiwi@aut.ac.nz

Abstract

Background: The Coronavirus disease 2019 (COVID-19) pandemic is having a major global impact, and the resultant response in the development of new diagnostics is unprecedented. The detection of antibodies against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has a role in managing the pandemic. We evaluated the feasibility of using SARS-CoV-2 peptide Kode Technology-modified red cells (C19-kodecytes) to develop an assay compatible with existing routine serologic platforms.

Study Design and Methods: A panel of eight unique red cells modified using Kode Technology function-spacer-lipid constructs and bearing short SARS-CoV-2 peptides was developed (C19-kodecyte assay). Kodecytes were tested against undiluted expected antibody-negative and -positive plasma samples in manual tube and three column agglutination technology (CAT) platforms. Parallel analysis with the same peptides in solid phase by enzyme immunoassays was performed. Evaluation samples included >120 expected negative blood donor samples and >140 COVID-19 convalescent plasma samples, with independent serologic analysis from two centers.

Results: Specificity (negative reaction rate against expected negative samples) in three different CAT platforms against novel C19-kodecytes was >91%, which

Abbreviations: CAT, column agglutination technology; C19-kodecytes, 1147 + 1255–1.5 + 2.5-kodecytes prepared from FSL-1147 (1.5 μmol/L) & FSL-1255 (2.5 μmol/L); EIA, enzyme immunoassay; FSL, function-spacer-lipid Kode Technology construct also known as a Kode construct; N.Z., New Zealand; RUS, Russian Federation; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; SCL-SB, Southern Community Laboratories Southern Region SARS-CoV-2 serum-bank

correlated with published literature. Sensitivity (positive reaction rate against expected positive convalescent, PCR-confirmed samples) ranged from 82% to 97% compared to 77% with the Abbott Architect SARS-CoV-2 IgG assay. Manual tube serology was less sensitive than CAT. Enzyme immunoassay results with some Kode Technology constructs also had high sensitivity.

Conclusions: C19-kodecytes are viable for use as serologic reagent red cells for the detection of SARS-CoV-2 antibody with routine blood antibody screening equipment.

KEYWORDS

Infectious disease testing, intravenous immunoglobulin, kodecyte

The diagnostic detection of antibodies to the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus is critically important to address the Coronavirus disease 2019 (COVID-19) pandemic as such assays contribute to the understanding of the susceptibility of a given population[1,2] and to the preparation of antibody-enriched therapeutic plasma products.[3,4] Most importantly, SARS-CoV-2 antibody assays may help support public health efforts in distinguishing between natural infection and vaccination rates in populations.

There is a growing number of commercial serological assays that use SARS-CoV-2 Spike (S) and receptor-binding domain protein as practical methods to determine the extent of COVID-19 immunity in a population.[5–11] Despite this, there is still a need for further assays as viral mutations resulting in new lineages can rapidly arise, and this may affect the robustness of sensitivity and specificity of any specific assay. Similarly, different and evolving vaccination strategies may produce different serologic profiles, with some better suited than others. By having multiple validated assay options available, these risks will be somewhat mitigated, especially if the assay has an intrinsic ability to rapidly adjust its antigenic profile. Likewise, having assays that do not require advanced laboratory instrumentation will be of particular value for developing countries.

Laboratories in most countries are already equipped to undertake routine blood group antibody serology, ranging from simple manual tube serology to fully automated screening. The opportunity, therefore, exists to modify the red cell membrane with SARS-CoV-2 peptides and then use these modified red cells in existing serology platforms.

One such approach for SARS-CoV-2 antibody detection with red cells has been investigated already, where antibodies modified with SARS-CoV-2 peptides have been attached to red cells and used in a serology platform.[12] An alternative is to use the highly adaptable Kode Technology platform, which utilizes function-spacer-lipid (FSL) constructs to attach epitopes to cells (kodecytes). Kode Technology has successfully been applied to produce a range of carbohydrate blood group antigens, including ABO, Lewis, P, and FORS,[13,14] and peptide epitopes, such as Miltenberger[15] and cytomegalovirus.[16] Red cell kodecytes have been used for qualitative and quantitative antibody diagnostic purposes.[17]

With the peptide sequence of the SARS-CoV-2 virus well established, we investigated the opportunity to create SARS-CoV-2 kodecytes for use on existing routine blood antibody screening diagnostic platforms.

1 | MATERIALS AND METHODS

1.1 | Donor and convalescent plasma samples

Plasma samples from blood donors (July 2020) were obtained from the New Zealand Blood Service (Auckland, New Zealand)—ethics approval AUTEC 20/183. In addition, 77 convalescent plasma samples, predominantly from international repatriations to New Zealand (N.Z.), were collected over the months of June to August 2020 and supplied by the Southern Community Laboratories Southern Region SARS-CoV-2 serum-bank (SCL-SB)—ethics approval HDEC 20/NTB/101. Diagnostics for these SCL-SB samples were performed independently, and all were confirmed as SARS-CoV-2 positive by PCR on at least one occasion, and 59 of 77 were recorded as IgG (nucleocapsid) antibody positive using the Abbott Architect SARS-CoV-2 IgG assay (as reported by SCL-SB). U.S. convalescent plasma samples (July to August 2020) were obtained from 62 donors who had recovered from PCR-confirmed COVID-19 (ClinicalTrials.gov Identifier NCT04360278), and expected negative samples were obtained from 20 blood donors at the Department of Transfusion Medicine, NIH Clinical

Center. Russian convalescent plasma samples were obtained from 14 patients recovered from PCR-confirmed COVID-19 and expected negative samples from eight blood donors.

1.2 | SARS-CoV-2 FSL constructs

Published S protein peptide sequences for SARS-CoV-2 (Genebank QHD43416.1) were used to determine candidate peptide epitopes suitable for construction as FSL constructs.[13,14,18] Nonglycosylated peptides sequences were selected according to algorithms (Tables S1, S2 and S3), including the use of space-filling models,[19–21] such as DNASTAR[22] of the glycan naked peptide[23] (Figure 1). Eight unique peptide sequences, plus two variations (491H & 888H, with additional histidine tail sequences) were selected (Table 1) and constructed into FSL constructs (Figure 2). Peptides 491 and 888, compatible with FSL construction, were selected from published data.[12]

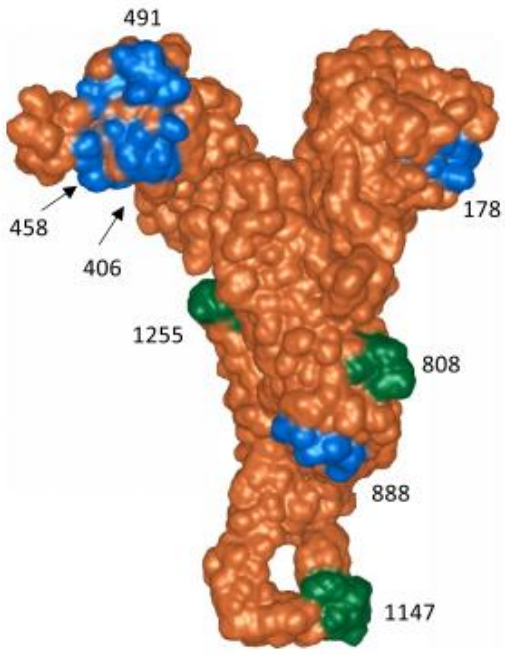

FIGURE 1 Location of the selected peptides sequences in the SARS-CoV-2 spike protein. This space-filling model is shown naked of its significant glycan coating.[23] Numbers refer to the peptide ID in Table 1, and the full length of the peptide epitope is highlighted with color (blue for low and green for high sensitivity and specificity as FSL constructs). Image produced using I-TASSER program for protein three-dimensional structure prediction[19-21] and DNASTAR[22]

1.3 | SARS-CoV-2 kodecytes

Terminology and methodology for describing FSL constructs and the resultant kodecytes are described in detail elsewhere.[24,25] Essentially, the kodecyte is described by the identification (ID) of the FSL's functional head (Table 1) and the micromolar (μmol/L) concentration of the FSL solution used to make it, for example, an 1147-3-kodecyte is a kodecyte made with peptide 1147 at an FSL concentration of 3 μmol/L, while an 1147 + 1255–3 + 5-kodecyte is a dual epitope-bearing kodecyte made with a blend of FSL-1147 and FSL-1255 at respective concentrations of 3 and 5 μmol/L. The abbreviated term C19-kodecytes was used for the final 1147 + 1255–1.5 + 2.5-kodecyte preparation.

Kodecytes were independently prepared in both Auckland (N.Z.) and Bethesda (U.S.) laboratories using parallel methodology.[25] Kodecytes and FSLs were diluted in red cell stabilizer solutions (ID-CellStab 005650, Bio-Rad Laboratories, Inc, Hercules, CA, USA) or Alsever's solution (no. A3551, Sigma-Aldrich, St. Louis, MO, USA).

In brief, the manufacture of kodecytes involves mixing a solution of FSL construct(s) with washed red cells, incubation at 37°C for 2 h, and storage in red cell stabilizer solution at 4°C with no washing required. Kodecytes were rested overnight before use and were used within 21 days.

TABLE 1 Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) peptide sequences selected for construction into function-spacer-lipid (FSL) constructs

FSL SARS-CoV-2 peptides		
ID[a]	SARS-CoV-n[b]	Peptide sequence[c]
178	2	DLEGKQGNFKNLREF[C]
406	2	EVRQIAPGQTGKIAD[C]
458	2	KSNLKPFERDISTEI[C]
491	2	PLQSYGFQPTNGVGY[C]
491H	2	PLQSYGFQPTNGVGY[HHHH][C]
808	2	DPSKPSKRSFIEDLL[C]
888	2	FGAGAALQIPFAMQM[C]
888H	2	FGAGAALQIPFAMQM[HHH][C]
1147	1,2	[C]SFKEELDKYFKNHTS
1255	1,2	[C]KFDEDDSEPVLKGVK

[a]ID is based on the initial amino acid in the SARS-CoV-2 consensus sequence and includes an H if the sequence has an additional histidine tail sequence appended.
[b]SARS-CoV-n indicates if specific to SARS-CoV-2 (2) or common to both SARS-CoV-1 and SARS-CoV-2 (1,2).
[c]SARS-CoV-2 peptide sequence (relating to the ID number) together with additional residues not part of the natural peptide sequence, including the conjugation cysteine [C] and solubilization histidine [H] residues. The location of the [C] cysteine residue (used to conjugate the peptide to the spacer) also indicates the region of peptide closest to the cell membrane.

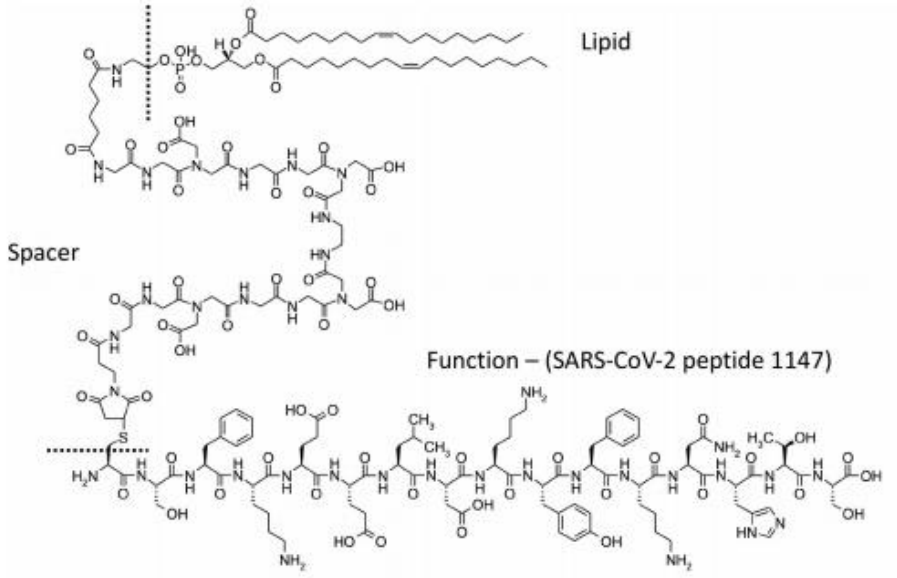

FIGURE 2 Representative schematic diagram a function-spacer-lipid (FSL) construct (FSL-1147). The FSL construct consists of a lipid phosphate moiety (1,2-dioleoyl-*sn*-glycero-3-phosphoethanolamine) conjugated to the spacer (carboxymethylglycine), which is conjugated via a cysteine SH group to the variable peptide functional head (in this example, the peptide is ID 1147, Table 1)

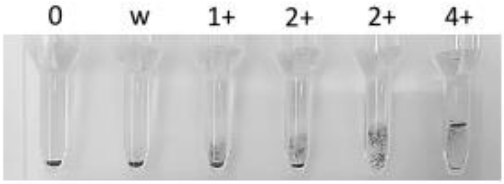

FIGURE 3 Typical agglutination reactions observed with C19-kodecytes in column agglutination technology. Shown are results of healthy donor sample (negative control) and 5 convalescent COVID-19 recovered plasma donor samples in the ID-micro typing system, anti-IgG card. Serologic grades assigned to the reactions are indicted above the microwells. Positive samples were selected and arranged as examples of observed increasing reactivity. Related images for Bio-Rad ID-system and the Grifols DG Gel system are shown in Figure S1

1.4 | SARS-CoV-2 kodecyte assay

For tube serology, 50 µl of plasma was mixed with 40 µl of a 5% suspension of kodecytes immediately centrifuged to grade IgM room temperature (IgM-RT) reactions, then incubated at 37°C for 60 min, and graded directly (IgM-37). After washing and addition of anti-human globin (Epiclone AHG Poly Anti-IgG-C3d, Seqirus, Australia), antiglobulin reactions were graded (AHG). Three CAT platforms were used, and methodologies and scoring systems were as recommended by the manufacturer, including the use of the grade "w" to indicate weak positive reaction (Figure 3). The Bio-Rad ID system used ID Cards LISS/Coombs (no. 50531, Bio-Rad Laboratories, Inc, Hercules, CA, USA); the Grifols DG Gel system used DG Gel Coombs cards and neutral cards (no. 210342 and 210343, Grifols S.A., Barcelona, Spain); and the ID-Micro Typing System used polyclonal rabbit anti-human IgG cards (no. MTS084024; Ortho Clinical Diagnostics, Raritan, NJ, USA). All samples reactive to SARS-CoV-2 kodecytes were also tested against unmodified and unrelated FSL control cells.

1.5 | Solid-phase enzyme immunoassays with peptides

Eight peptides (178, 406, 458, 808, 888, 888H, 1147, 1255) used to make FSL constructs for the preparation of kodecytes (Table 1) were also tested by in-house enzyme immunoassay (EIA). Two different approaches were used to immobilize the peptide epitopes in 96-well microtiter plates. The first approach (polyacrylamide [PAA]-EIA) used peptides conjugated to PAA, which were used to coat the microplates (described in detail elsewhere).[26] The second proof-of-concept approach (FSL-EIA) used FSL constructs 1147 & 1255 and attached these to 96-well microplates (Nunc Maxisorp, Sigma-Aldrich, St Louis, MO, USA) precoated with lipid. In brief, 100 µl of ethanol-

TABLE 2 Initial specificity and sensitivity analysis of all severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) kodecytes

			Specificity (expected negative)					Sensitivity (expected positive)		
			Tube serology (New Zealand [N.Z.])			Column agglutination technology (CAT)[a]		CAT		
Kodecyte ID	μM[b]	n (N.Z. + United States [U.S.])[c]	IgM-RT	IgM-37	AHG	nPos (N.Z. + U.S.)[c]	% Pos	n (US)	n Pos	% Pos
178	10	68 + 20	1	0	0	1 + 0	1%	62	1	2%
406	10	72 + 20	1	0	2			6	0	0%
458	10	60 + 20	3	0	0	2 + 1	4%	62	5	8%
491	10	72 + 20	3	0	0			6	0	0%
491H	10	72 + 20	4	0	2			6	0	0%
808	10	100 + 20	0	0	0	3 + 3	5%	62	28	45%
888	10	72 + 20	2	0	0			6	0	0%
888H	10	72 + 20	1	0	0			6	0	0%
1147	3	100 + 20	0	0	2	5 + 2	6%	62	48	77%
1255	5	100 + 20	3	0	1	1 + 0	1%	62	32	52%
Untreated	0	[d]	0	0	0	0		62	0	0%

[a]CAT results are split into N.Z. and U.S. results representing Bio-Rad and MTS platforms, respectively. nPos indicates the number of positive reactions by N.Z. and U.S. laboratory, while % Pos uses combined CAT data to calculate percentage positive rate. Gaps in the table indicate analysis not done.
[b]μM refers to the μmol/L concentration of function-spacer-lipid used to make the kodecyte. Results are only shown for the single concentration considered most appropriate for diagnostic use.
[c]The number of samples is divided into N.Z. and U.S. samples (separated by the + symbol).
[d]All samples reactive as positives were also tested and found negative against unmodified cells used to make kodecytes.

containing lecithin (25 μg/ml) and cholesterol (50 μg/ml) were added to the wells of the plate and dried at 37°C for 60 min and then left at RT overnight until completely dry. In the next stage, FSL peptides (10 μg/ml) in 100 μl of buffer were then added to the plate wells and incubated for 60 min at 47°C. After peptide coating the microplates, both PAA-EIA and FSL-EIA methods were standard procedures with secondary labeled antibodies and color development read using a microplate spectrophotometer.[26]

2 | RESULTS

Two terms are used to describe diagnostic accuracy in this article according to FDA criteria.[27] Specificity is the estimated proportion of subjects without the target condition in whom the test is negative. Sensitivity is the estimated proportion of subjects with the target condition in whom the test is positive.

2.1 | Preliminary screening kodecyte specificity and sensitivity

The initial method development in manual and CAT platforms involved evaluating each FSL construct for specificity as kodecytes over the range of 3–20 μmol/L against expected negative samples. Concentrations of kodecytes were then evaluated for sensitivity against convalescent samples to select those kodecytes that showed appropriate specificity and sensitivity. The sensitivity and specificity for the single concentration considered the most appropriate for diagnostic use for all kodecytes evaluated are summarized in Table 2. Only kodecytes made with FSLs 808, 1147, and 1255 showed adequate sensitivity. However, the average reaction strength of 808-kodecytes was half that of 1255-kodecytes and gave no additional information over 1147-kodecytes and so were not tested further. From these initial experiments, the two most promising FSL candidates (1147 and 1255) were selected for extended analysis along with a dual-epitope kodecyte made with a blend of FSLs 1147 and 1255.

2.2 | Tuning kodecyte specificity and sensitivity: 1147 and 1255

Kodecytes were prepared with FSLs 1147 and 1255 at two different micromolar concentrations (3 + 5 μmol/L and "half strength" 1.5 + 2.5 μmol/L), both singly and as dual 1147 + 1255-kodecytes. A comparison of the results obtained with the 1147 + 1255–3 + 5 kodecytes and half-strength 1147 + 1255–1.5 + 2.5 kodecytes against the 56

TABLE 3 C19-kodecytes in three different column agglutination technology (CAT) platforms and the Abbott Architect enzyme immunoassay (EIA) IgG antibody results against PCR SARS-CoV-2-positive samples

	Antibody positive results in Coronavirus Disease 2019 (COVID-19) PCR-positive samples							
	CAT (anti-IgG) C19-kodecyte[a] assay							
COVID-19 Status	MTS (U.S.)		Bio-Rad (N.Z.)		Grifols (N.Z.)		EIA[b]	
PCR Positive	49/54	91%	63/77[c]	82%	75/77[c]	97%	59/77[c]	77%
Expected Negative[d]	0/19	100%	4/100	4%	9/100	9%		

[a]1147 + 1255–3 + 5-kodecytes were also tested but did not differ (not shown).
[b]EIA – Abbott Architect SARS-CoV-2 IgG enzyme immunoassay results as reported by the Southern Community Laboratories Southern Region SARS-CoV-2 serum bank (SCL-SB).
[c]77 PCR-confirmed SARS-CoV-2-positive samples supplied by SCL-SB; 7% of these samples also had C19-kodecyte IgM saline CAT activity.
[d]Expected negative samples are blood donor samples.

TABLE 4 C19-kodecyte tube serology results against PCR severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)-positive samples

Coronavirus Disease 2019 (COVID-19) status	Tube serology C19-kodecyte assay					
	IgM-RT		IgM-37		AHG (anti-IgG)	
PCR Positive	0/77	0%	0/77	0%	51/77	66%
Expected Negative	1/100	1%	1/100	1%	2/100	2%

U.S. convalescent samples showed a slight reduction in score for five samples and no change in the number of positive samples. However, in the 20 negative samples, specificity improved with epitope dilution to 1.5 + 2.5 μmol/L. In addition, plasma samples diluted to 1:3 in PBS were evaluated; however, dilution did not improve specificity and instead substantially reduced sensitivity (not shown).

On balance, it was considered that the performance of the 1147 + 1255–1.5 + 2.5 kodecytes (hereinafter called C19-kodecytes) against undiluted plasma (Figure 3) was the best generic formulation for further evaluation.

2.3 | Evaluation of C19-kodecytes

Extended sensitivity analysis of the C19-kodecytes was undertaken with the SCL-SB samples (Abbott Architect SARS-CoV-2 IgG known status) and NIH convalescent donors who had recovered from PCR-confirmed COVID-19. The Grifols DG system reacted with 97%, Ortho MTS system with 91%, and the Bio-Rad ID system with 82% of these PCR-confirmed SARS-CoV-2 positive samples, compared with 77% by EIA (Table 3). It should also be noted that the specificity rate was also different, with the Grifols DG system reacting with 9% and Bio-Rad platform with 4% of expected negative samples. The specificity rate for the EIA assay was not reported by SCL-SB.

Further analysis of the differences between the Grifols DG, Bio-Rad-ID, and Abbott Architect systems (Figure S2 and Table S4) revealed that C19-kodecytes in the Grifols DG platform obtained 2+ or greater grades in 69 of 77 (90%) of samples, while the Bio-Rad-ID platform scored 2+ or greater grades in only 24% of samples (and 18% of these expected positive samples were negative).

To establish the contribution of IgM to the results observed in antiglobulin CAT cards, the SCL-SB samples were also tested in neutral CAT cards. Of the 73 available anti-IgG CAT-positive samples, 5 (7%) were also positive in the neutral cards, indicating the detection of IgM activity. In all examples, the anti-IgG card grades were at least one grade stronger than in neutral cards (results not shown), indicating the copresence of IgG. None of these IgM-reactive samples reacted with unmodified cells, nor were there any positive samples at IgM-RT or IgM-37 by manual serology with C19-kodecytes (Table 4).

Although manual tube serology showed 66% sensitivity and 98% specificity for IgG (Table 4), reactions were much weaker than those observed in CAT. Of the

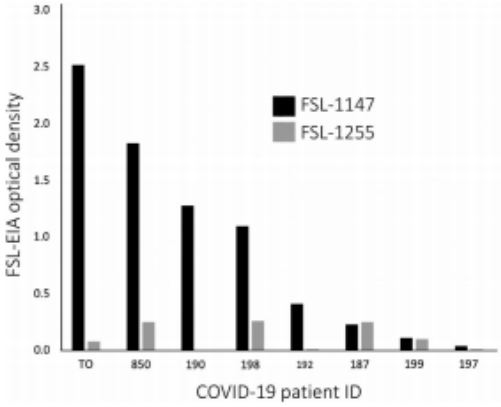

FIGURE 4 Function-spacer-lipid (FSL)-enzyme immunoassay (EIA) of convalescent plasma samples. Eight Coronavirus Disease 2019 (COVID-19) convalescent plasma Russian Federation samples diluted 1:3 and tested against FSL-1147 and FSL-1255 by EIA. Patients 850,190, and 187 had severe COVID-19, while the remainder were mild or asymptomatic. There was insufficient sample 190 to test against FSL-1255

manual tube serology-positive results recorded for AHG, 25 of 51 (49%) were w reactions, and 44 of 51 (86%) were either w or 1+. Only 7 of 51(14%) reactions were grade 2 +, and none were grade 3+ or 4+. Doubling the strength of the kodecyte formulation to 1147 + 1255–3 + 5-kodecytes did not improve the rate of sensitivity (not shown).

2.4 | Specificity and Sensitivity: Solid-phase EIA assays

The in-house PAA-EIA assay IgG/IgM results against PAA-conjugated peptides for Russian Federation (RUS) blood donors and convalescent plasma samples did not show clear specificity or sensitivity, although the results supported the use of peptides 1147 and 1255 (not shown).

The in-house proof-of-concept FSL-EIA assay with solid-phase FSL-1147 and 1:3 diluted plasma was clearly able to identify SARS-CoV-2 antibody in five of eight COVID-19 convalescent samples (Figure 4). In contrast, FSL-1255 did not show sensitivity.

3 | DISCUSSION

Plasma as a source of therapeutic antibodies is still considered a useful therapeutic approach,[28,29] and in early clinical trials with SARS-CoV-2 antibody-enriched intravenous immunoglobulin, improved clinical outcomes were observed.[3,4] Demand for these products will require mass screening of blood donors.

Kode Technology is a simple technique that attaches small molecules onto the outside of the red cell and then uses these modified red cells (kodecytes) in existing serologic diagnostic platforms.[13–16] Other than kodecyte reagent red cells, the kodecyte assay methodology that we have developed in this study is identical to that used for routine antiglobulin screening for red cell antibodies. Because red cell antibody detection by manual tube or CAT is present in almost every transfusion laboratory, the kodecyte assay can be easily used by almost any serologic laboratory, including those in most developing economies. We therefore considered the possibility that SARS-CoV-2 kodecytes could also be created using established kodecyte methods[14,15] with a view that these kodecytes would also be of value to blood services that could undertake mass population sample screening using surplus laboratory capacity.

The design approach taken to develop this SARS-CoV-2 kodecyte antibody diagnostic was to first identify potential nonglycosylated epitopes on the S protein using a range of predictive algorithms and online tools (Tables S1 and S2). Although Kode Technology is able to attach glycans onto cells,[13,14,16,17] the SARS-CoV-2 kodecyte assay was restricted to nonglycosylated peptide epitopes, both because of ease of manufacture and no diagnostic glycan antibody signatures for COVID-19 are as yet recognized. The selected peptides were then further refined to be compatible with FSL conjugation chemistry,[18] and eight candidates were chosen for construction as Kode FSL constructs (Table 1). These FSL constructs were then used to create kodecytes at different concentrations, and they were tested against expected negative (specificity) and positive (sensitivity) samples in manual and CAT platforms to find useable concentration ranges that gave acceptable specificity and sensitivity. This range was found to be 5 μmol/L or less. From the initial eight candidates, three FSL constructs (808, 1147, and 1255) showed the most promising sensitivity, which was further reduced to two after FSL-808 was sidelined. The two constructs FSL-1147 and 1255 were then formulated as single and dual epitope-bearing kodecytes for further analyses. We found that better sensitivity and specificity was achieved with 1147 + 1255–1.5 + 2.5-kodecytes, which were renamed C19-kodecytes. It is important to note this was a proof-of-principle trial and final optimizations for the different kodecyte formulations on different platforms have not yet been performed, and as expected, differences between platforms were observed (Table 3). We expected to achieve better specificity on the Grifols platform and improved sensitivity on the Bio-Rad and manual tube platforms than was observed with the current base formulation of the

C19-kodecytes described in this study. We are currently undertaking these extensive optimizations in preparation for clinical studies with a prototype assay, including determining extended stability and storage parameters.

Evaluating the performance of C19-kodecytes against convalescent plasma found that the three different CAT platforms were all able to detect antibodies in the majority of samples from COVID-19 convalescent patients (Table 3). The results observed were almost certainly due to IgG antibodies, and although IgM will contribute to the reactions in CAT anti-IgG cards, saline reactions indicated a low-level contribution (about 7%) of IgM. This was expected as, in previous studies with Miltenbeger kodecytes, it was found that IgM is poorly reactive with kodecytes made with peptide-FSLs.[14,15] Overall, the C19-kodecyte assay detected more antibody-positive samples than were detected with the Abbott Architect SARS-CoV-2 IgG enzyme immunoassay. However, it should be appreciated that alternative and next-generation EIA platforms would be expected to have improved sensitivity and specificity. Extensive comparative analysis of the CAT and EIA platforms, including a scatterplot analysis (Figure S2), did not reveal any clear reason for differences in performance. However, it must be appreciated that, together with the significant fundamental differences in assay methodology, the EIA assay uses diluted plasma, while the kodecyte assay uses undiluted plasma. It is reasonable to expect that lower levels of antibody will be easier to detect in undiluted plasma (although the level of epitope on the kodecyte may offset this somewhat).[14] When 1:3 dilutions of plasma were evaluated against kodecytes, there was a substantial loss in positive reactivity, supporting the loss of sensitivity caused by dilution.

Parallel to the sensitivity for diagnostic assays is specificity, where an undesired number of false-positive reactions will invalidate the usefulness of an assay. It is reported that antibodies to the coronavirus S protein (from endemic common corona viruses) are expected in up to about 10% of the general population.[30–32] Consequently, a nonspecificity rate of 3%–5% reactivity with samples that have not had contact with SARS-CoV-2 was considered acceptable. The C19-kodecytes were found to react with up to 9% of expected negative samples in the Grifols DG system (which showed the highest degree of sensitivity at 97%); however, the nonspecificity rate was a more acceptable 4% with the Bio-Rad platform, but the sensitivity rate was lower at 81%. The MTS platform appeared to have a performance in-between these two platforms, with a sensitivity of 91%, and further testing is required to more accurately define specificity. Manual tube serology was less sensitive than CAT and only reacted with two-thirds of the convalescent plasma samples and with generally w reaction grades. Further optimization of the assay for manual serological use is required for the detection of antibody; however, the assay as is would be suitable for the detection of samples with higher levels of antibody (especially as the kodecyte assay is semiquantitative).[17]

In addition to the kodecyte assay, two solid-phase assays using the same peptides as those used to make the FSL constructs were designed to evaluate the usefulness of these peptides by solid phase. These two solid-phase EIA included one where the peptide was conjugated to PAA and immobilized on a microplate (PAA-EIA), and the other used the same FSL constructs that were used to make kodecytes but instead immobilized them in a solid-phase lipid layer onto microplates. The PAA-EIA did not show sensitivity or specificity in contrast to the FSL-EIA, which demonstrated a viable proof of concept and could also be used with undiluted plasma (results not shown), but further analysis of a larger dataset is still needed.

Intriguingly, the FSL-EIA outperformed the PAA-EIA assay, suggesting that the presentation of small peptide epitopes on FSL constructs is a favorable form of presentation. Furthermore, the presentation of the FSL in the red cell membrane may be more favorable (than in EIA) as it allows for the use of undiluted plasma. It can be speculated that the glycocalyx of the red cell may be able to buffer nonspecificity from low-affinity cross-reactive antibodies (an effect managed by plasma dilution in EIA assays).

The manufacture of kodecytes is a very simple process,[13,25] requiring only the contacting of a solution of FSL construct(s) dispersed in buffered saline with washed red cells, incubation at 37°C for 2 h, and then dispensing them "ready for use" (washing is not required). The prepared kodecytes are then stored and used as normal reagent red cells. One mg of a 15-amino acid peptide FSL construct will make about 70 ml of 5 μmol/L kodecytes (packed cells). If these kodecytes are then diluted to 1% for CAT technology, this will result in 7 L of reagent red cells, and if 50 μl is used per assay, then 1 mg will enable more than 100,000 CAT assays. Although RBCs were used to prepare kodecytes in this study, FSL constructs can also be attached to noncellular surfaces,[24,33] including microspheres and lateral flow membranes or immobilized cell membranes, opening up further possibilities to develop novel diagnostics.

Overall, the kodecyte assay was able to achieve specificity and sensitivity at least equivalent to an established EIA antibody diagnostic. Due to the cassette design of Kode Technology, it is highly adaptable,[13-18,24–26] and changing the antigenic epitope on a Kode FSL construct can be achieved within a few weeks, allowing for rapid response to new strains arising with novel antigenic

mutations. Other than determining optimal concentrations for sensitivity and specificity, no other modifications to the methodology for use are required.

This article describes an adaptable platform technology able to be easily accommodated into almost all existing transfusion diagnostic laboratories, including those with limited infrastructure, and will allow for this sector to actively participate in the screening for SARS-CoV-2 antibodies, both for population needs and therapeutic uses.

ACKNOWLEDGMENTS

We acknowledge support in providing samples from Kamille A. West and Valeria de Giorgi of the NIH Clinical Center (NCT04360278); Arlo Upton and Alyson Craigie of the Southern Community Laboratory (Southern Region SARS-CoV-2 serum-bank); and Marina Ziganshina, and Nataliya Dolgushina of the Kulakov National Medical Research Center for Obstetrics, Gynecology and Perinatology, Moscow, Russian Federation. This work was partially funded by a research grant from the NZ Ministry of Business, Innovation & Employment COVID-19 Innovation Acceleration Fund, contract CIAF 0490, and the NIH Intramural Research Program at the NIH Clinical Center, projects ZIA CL002128 and RASCL727301. N.B. and N.S. were supported in part by the Russian Foundation for Basic Research grant #20-04-60335.

STATEMENT OF DISCLAIMER

The views expressed do not necessarily represent the view of the National Institutes of Health, the Department of Health and Human Services, or the U.S. Federal Government.

CONFLICT OF INTEREST

SMH, NVB, and ECW are employees and stockholders of Kode Biotech, the patent owner of Kode biosurface engineering technology. All other authors report no conflict of interest.

ORCID

Willy A. Flegel https://orcid.org/0000-0002-1631-7198
Stephen M. Henry https://orcid.org/0000-0002-9946-8441

REFERENCES

1. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, Oteo J, Hernán MA, Pérez-Olmeda M, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): A nationwide, population-based seroepidemiological study. Lancet. 2020;396(10250):535–44.
2. GeurtsvanKessel CH, Okba NM, Igloi Z, Bogers S, Embregts CW, Laksono BM, et al. An evaluation of COVID-19 serological assays informs future diagnostics and exposure assessment. Nat Commun. 2020;11:3436.
3. Gharebaghi N, Nejadrahim R, Mousavi SJ, Sadat-Ebrahimi SR, Hajizadeh R. The use of intravenous immunoglobulin gamma for the treatment of severe coronavirus disease 2019: A randomized placebo-controlled double-blind clinical trial. BMC Infect Dis. 2020;20:786.
4. Sakoulas G, Geriak M, Kullar R, Greenwood K, Habib M, Vyas A, et al. Intravenous immunoglobulin (IVIG) significantly reduces respiratory morbidity in COVID-19 pneumonia: A prospective randomized trial. medRxiv. 2020. https://doi.org/10.1101/2020.07.20.20157891.
5. Vashist SK. In vitro diagnostic assays for COVID-19: Recent advances and emerging trends. Diagnostics (Basel). 2020;10(4):202.
6. Burbelo PD, Riedo FX, Morishima C, Rawlings S, Smith D, Das S, et al. Sensitivity in detection of antibodies to nucleocapsid and spike proteins of severe acute respiratory syndrome coronavirus 2 in patients with coronavirus disease 2019. J Infect Dis. 2020;222(2):206–13.
7. Guo L, Ren L, Yang S, Xiao M, Chang D, Yang F, et al. Profiling early humoral response to diagnose novel coronavirus disease (COVID-19). Clin Infect Dis. 2020;71(15):778–85.
8. Li Z, Yi Y, Luo X, Xiong N, Liu Y, Li S, et al. Development and clinical application of a rapid IgM-IgG combined antibody test for SARS-CoV-2 infection diagnosis. J Med Virol. 2020;92:1518–24.
9. Amanat F, Stadlbauer D, Strohmeier S, Nguyen THO, Chromikova V, McMahon M, et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. Nat Med. 2020;26:1033–6.
10. Jin Y, Wang M, Zuo Z, Fan C, Ye F, Cai Z, et al. Diagnostic value and dynamic variance of serum antibody in coronavirus disease 2019. Int J Infect Dis. 2020;94:49–52.
11. Okba NM, Müller MA, Li W, Wang C, CH GK, Corman VM, et al. Severe acute respiratory syndrome coronavirus 2– specific antibody responses in coronavirus disease patients. Emerg Infect Dis. 2020;26(7):1478–88.
12. Alves D, Curvello R, Henderson E, Kesarwani V, Walker JA, Leguizamon SC, et al. Rapid gel card agglutination assays for serological analysis following SARS-CoV-2 infection in humans. ACS Sens. 2020;5(8):2596–603.
13. Henry S, Perry H, Bovin N. Applications for kodecytes in immunohaematology. ISBT Sci Ser. 2018;13(3):229–37.
14. Henry S. Kodecytes: Modifying the surface of red blood cells. ISBT Sci Ser. 2020;15(3):303–9.
15. Heathcote D, Carroll T, Wang JJ, Flower R, Rodionov I, Tuzikov A, et al. Immunohematology: Novel antibody screening cells, MUT+ Mur kodecytes, created by attaching peptides onto red blood cells. Transfusion. 2010;50(3):635–41.
16. Georgakopoulos T, Komarraju S, Henry S, Bertolini J. An improved Fc function assay utilizing CMV antigen-coated red blood cells generated with synthetic function–spacer–lipid constructs. Vox Sang. 2012;102(1):72–8.
17. Perry H, Bovin N, Henry S. A standardized kodecyte method to quantify ABO antibodies in undiluted plasma of patients before ABO-incompatible kidney transplantation. Transfusion. 2019;59(6):2131–40.
18. Henry S, Komarraju S, Heathcote D, Rodionov IL. Designing peptide-based FSL constructs to create Miltenberger kodecytes. ISBT Sci Ser. 2011;6(2):306–12.
19. Roy A, Kucukural A, Zhang Y. I-TASSER: A unified platform for automated protein structure and function prediction. Nat Protoc. 2010;5:725–38.

20. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y. The I-TASSER suite: Protein structure and function prediction. Nat Methods. 2015;12:7–8.
21. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008;9:40.
22. Darnell S, Riese M. Precise predictions of linear B cell epitopes in protean 3D. 2012.
23. Casalino L, Gaieb Z, Goldsmith JA, Hjorth CK, Dommer AC, Harbison AM, et al. Beyond shielding: The roles of glycans in the SARS-CoV-2 spike protein. ACS Cent Sci. 2020;6(10):1722–34.
24. Henry S, Williams E, Barr K, Korchagina E, Tuzikov A, Ilyushina N, et al. Rapid one-step biotinylation of biological and non-biological surfaces. Sci Rep. 2018;8:2845.
25. Williams E, Trent C, Bovin N, Nagappan R Technical Bulletin: Kode Technology SARS-CoV-2 kodecytes and Function Spacer Lipid constructs. AUT Centre for Kode Technology Innovation. Available from: https://openrepository.aut.ac.nz/handle/10292/12731.
26. Tuzikov A, Chinarev A, Shilova N, Gordeeva E, Galanina O, Ovchinnikova T, et al. 40years of glyco-polyacrylamide in glycobiology. Glycoconjugate J. In press. https://doi.org/10.1007/s10719-020-09965-5.
27. Food and Drug Administration. Guidance for industry and FDA staff: Statistical guidance on reporting results from studies evaluating diagnostic tests. Silver Spring, MD: US Food and Drug Administration; 2007. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/statistical-guidance-reporting-results-studies-evaluating-diagnostic-tests-guidance-industry-and-fda.
28. Ouyang J, Isnard S, Lin J, Fombuena B, Peng X, Routy JP, et al. Convalescent plasma: The relay baton in the race for coronavirus disease 2019 treatment. Front Immunol. 2020;570063:11.
29. Sun M, Xu Y, He H, Zhang L, Wang X, Qiu Q, et al. A potentially effective treatment for COVID-19: A systematic review and meta-analysis of convalescent plasma therapy in treating severe infectious disease. Int J Infect Dis. 2020;98:334–46.
30. Ng K, Faulkner N, Cornish G, Rosa A, Harvey R, Hussain S, et al. Pre-existing and de novo humoral immunity to SARS-CoV-2 in humans. Science. 2020;370:1339–43.
31. van der Heide V. SARS-CoV-2 cross-reactivity in healthy donors. Nat Rev Immunol. 2020;20(7):408–8.
32. Nguyen-Contant P, Embong AK, Kanagaiah P, Chaves FA, Yang H, Branche AR, et al. S protein-reactive IgG and memory B cell production after human SARS-CoV-2 infection includes broad reactivity to the S2 subunit. MBio. 2020;11(5):e01991–20.
33. Williams E, Barr K, Korchagina E, Tuzikov A, Henry S, Bovin N. Ultra-fast glyco-coating of non-biological surfaces. Int J Mol Sci. 2016;17(1):118.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Nagappan R, Flegel WA, Srivastava K, et al. COVID-19 antibody screening with SARS-CoV-2 red cell kodecytes using routine serologic diagnostic platforms. *Transfusion*. 2021; 61:1171–1180. https://doi.org/10.1111/trf.16327