ISOLATION AND CHARACTERIZATION OF APTAMERS AGAINST *LEPTOSPIRA* LIPL32

YEOH TZI SHIEN

UNIVERSITI SAINS MALAYSIA

2023

ISOLATION AND CHARACTERIZATION OF APTAMERS AGAINST *LEPTOSPIRA* LIPL32

by

YEOH TZI SHIEN

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

June 2023

ACKNOWLEDGEMENT

This section is dedicated to those involved, without the key persons along the way, I never would have been able to finish this. First, I would like to express my gratitude to my supervisors, Dr Citartan Marimuthu, Prof Dr Tang Thean Hock and Dr Hazrina Yusof Hamdani for their guidance and advises in scientific research.

Next, to the members of RNA-Bio Research Group, past and present, thank you so much for your assistance during my time at the lab. Even though working on different projects, we do spend time helping each other through discussions and interactions and lift each other up at times of difficulties. People come and go, perhaps you guys know who you are and I cherish all those times we have spent together.

Lastly, to my parents, sibling and companion for always supporting my decision and giving all they can to help. I hope that have made you all proud for me reaching this stage of education.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS AND ABBREVATIONS

A	Adenine
A-PCR	Asymmetric PCR
AuNPs	Gold Nanoparticles
BSA	Bovine serum albumin
ddH_2O	double-distilled water
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
ELASA	Enzyme-linked Apta-sorbent Assay
ErBr	Ethidiium bromide
et. al.	And others
G	guanine
G4	G-quadruplex
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horseradish peroxidase
IPTG	Isopropyl-β-D-thiogalactopyranoside
Kd	Dissociation constant
kDa	Kilodalton
KOH	Potassium hydroxide
LB	Luria Bertani medium
LFA	Lateral Flow Assay
LINA	Lithium Chloride and Sodium Chloride
Min	Minute (s)

mL	Milliliter
mM	Millimolar
MRE	Molecular Recognition Element
Na^+	Sodium ion
NaCl	Sodium chloride
NaOAc.3H2O	Sodium acetate trihydrate
NaOH	Sodium hydroxide
ng	Nanogram
nM	Nanomolar
nt	Nucleotide (s)
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PBST	Phosphate buffered saline with Tween 20
PCR	Polymerase chain reaction
PLL	Poly-L-Lysine
pmol	Picomole
RNA	Ribonucleic acid
rpm	Rotations per minute
RT	Room temperature
RT-PCR	Reverse transcription-PCR
SDS	Sodium Dodecyl Sulfate
SELEX	Systematic Evolution of Ligands via Exponential Enrichment
ssDNA	Single-stranded DNA
T	Thymine
TAE	Tris–Acetic Acid–EDTA

TBE	Tris-Boric Acid-EDTA
TMB	3,3',5,5'-tetramethylbenzidine
Tris	Tris-(Hydroxymethyl)-Aminomethane
tRNA	Transfer RNA
U	Units of enzymatic activity
UV	Ultraviolet
V	Volt (s)
v/v	Volume per volume
VEGF	Vascular Endothelial Growth Factor
w/v	Weight per volume
xg	Relative Centrifugal Force
X–gal	5’-Bromo-4’-Chloro-3’-Indolyl-β-D galactoside
µg	Microgram
µL	Microliter
µM	Micromolar
%	Percentage
°C	Degrees Celsius

PENGASINGAN DAN PENCIRIAN APTAMER TERHADAP *LEPTOSPIRA* LIPL32

ABSTRAK

Digelar “äntibodi kimia”, aptamer ialah jujukan tunggal DNA/RNA yang mampu berinteraksi dengan pelbagai sasaran dengan pengikatan dan pengkhususan yang tinggi. Disebabkan kelebihan seperti ketiadaan variasi kelompok ke kelompok dan kos sintesis yang lebih murah berbanding antibodi, aptamer merupakan kelas ‘molecular recognition element’ (MRE) yang berpotensi, terutamanya dalam diagnosis awal leptospirosis, zoonosis yang sangat endemik di kawasan tropika dan subtropika di seluruh dunia termasuk Malaysia. Memandangkan Ujian Aglutinasi Mikroskopik (MAT) standard emas merangkumi beberapa kelemahan yang menghalangnya daripada diagnosis pantas, pengesanan langsung adalah wajar dengan menggunakan biomarker membran luar LipL32 yang dinyatakan secara eksklusif oleh *Leptospira* yang patogenik sebagai sasaran. Dalam kajian ini, kedua-dua aptamer DNA dan RNA terhadap LipL32 telah dibangunkan. LepRapt-11 telah muncul sebagai aptamer RNA yang paling baik. dengan nilai penceraian dan had pengesanan 350 ± 47.45 dan 100 nM seperti yang ditentukan oleh ELASA langsung, masing masing. Sementara itu, LepDapt-5a adalah calon aptamer DNA yang paling kuat seperti yang ditentukan oleh ELASA langsung, dengan nilai penceraian antara 33.97 ± 5.303 dan 46.35 ± 9.09 nM. Potensi diagnostik LepDapt-5a telah diuji selanjutnya pada platform ELASA langsung dan sandwic untuk pengesanan langsung Leptospira patogenik dalam 10% serum, dengan LOD masing-masing 10^5 dan 10^4 CFU/mL. Ujian dot blot yang dibangunkan mampu mencapai LOD 10^4 CFU/mL. Perbincangan umum antara kedua-dua aptamer

mendedahkan keunikan kedua-dua aptamer, yang terletak pada sifat pemilihan "*in vitro*", pembentukan "G-quadruplex", teknikal ELASA dan perbezaan dalam kos pengeluaran. Berbanding dengan LepRapt-11, LepDapt-5a adalah lebih baik daripada LepRapt-11 dalam pembentukan G-quadruplex, teknikal ELASA dan kos pengeluaran yang lebih murah tetapi bukan sifat pemilihan. Disimpulkan bahawa kedua-dua aptamer baru terhadap LipL32 mempunyai potensi dalam diagnostik, terutamanya dalam bentuk ujian ELASA dan boleh dilanjutkan kepada penilaian menggunakan sampel klinikal.

.

ISOLATION AND CHARACTERIZATION OF APTAMERS AGAINST *LEPTOSPIRA* LIPL32

ABSTRACT

Dubbed “chemical antibodies”, aptamers are single-stranded DNA/ RNA capable of binding to various targets with high affinity and specificity. Owing to the advantages such as absence of batch-to-batch variation and cheaper cost of synthesis as compared to antibodies, aptamers are a promising class of molecular recognition element (MRE) especially in early diagnosis of leptospirosis, a zoonosis highly endemic in tropical and subtropical regions worldwide including Malaysia. As the gold standard Microscopic Agglutination Test (MAT) is encompassed with several drawbacks that hinder it from rapid diagnosis, direct detection is desirable by using the outer membrane biomarker LipL32 exclusively expressed by pathogenic *Leptospira* as the target. In this study, both DNA and RNA aptamers were developed against LipL32. LepRapt-11 was shown to be the most potent RNA aptamer, with dissociation constant and limit of detection of 350 ± 47.45 and 100 nM as determined by direct Enzyme-linked Aptasorbent Assay (ELASA), respectively. Meanwhile, LepDapt-5a is the most potent DNA aptamer candidate as determined by direct ELASA, with a Kd value between 33.97 ± 5.303 and 46.35 ± 9.09 nM. The diagnostic potential of LepDapt-5a was further tested on direct and sandwich ELASA platform for direct detection of pathogenic *Leptospira* in 10 % serum, with LOD of 10^5 and 10^4 CFU/ mL, respectively. The dot blot assay developed was able to attain a LOD of 10^4 CFU/ mL. A general discussion between both aptamers unveils the uniqueness of both aptamers, which lies in the nature of *in vitro* selection, formation of G-quadruplex, technicality

of ELASA and differences in the cost of production. As compared to LepRapt-11, LepDapt-5a is better than LepRapt-11 in formation of G-quadruplex, technicality of ELASA and cheaper cost of production but not nature of selection. It is concluded that both novel aptamers against LipL32 have potentiality in diagnostics, especially in the form of ELASA assay and can be extended to evaluation using clinical samples. .

CHAPTER 1

INTRODUCTION

1.1 Introduction

Since their discovery in 1990 (Ellington and Szostak, 1990, Tuerk and Gold, 1990), aptamers have gradually gained popularity as molecular recognition elements (MRE), especially in diagnostic applications. They are able to form a myriad of three-dimensional (3D) structures that can bind with high affinity and specificity to various target molecules, aided by electrostatic charges, hydrogen bonding, and van der Waals forces (Hermann and Patel, 2000, Nomura *et al*., 2010, Piganeau and Schroeder, 2003). Also dubbed "chemical antibodies", aptamers have many advantages compared to antibodies. They are associated with a lower cost of synthesis, are able to undergo reversible denaturation and renaturation, have low-to-no immunogenicity, and are easily functionalized. To date, many aptamers have been isolated against a variety of targets such as small molecules, proteins, viruses, and whole cells (Elskens *et al*., 2020, McKeague and Derosa, 2012, Odeh *et al*., 2019)

To isolate highly specific and affine aptamers, an in vitro selection process known as "Systemic Evolution of Ligands via Exponential Enrichment" or SELEX is employed (Ellington and Szostak, 1990, Tuerk and Gold, 1990). It is an iterative selection strategy comprising four major steps, which are the incubation of the initial randomized single-stranded nucleic acid pool (ssDNA or RNA) with a target molecule, partitioning, and recovery of the target-bound nucleic acid. The eluted molecules are then amplified via Polymerase Chain Reaction (PCR) or Reverse Transcription-Polymerase Chain Reaction (RT-PCR) for DNA SELEX or RNA SELEX, respectively. The ssDNA/RNA pool is then regenerated for the subsequent rounds of SELEX. There

are approximately 1.03 million cases of leptospirosis globally, with nearly 60,000 deaths per annum (Costa *et al*., 2015, Torgerson *et al*., 2015). It is a potentially life-threatening zoonosis caused by pathogenic *Leptospira*. Diagnosis of this disease relies on the gold standard microscopic agglutination test (MAT) (World Health Organization, 2003). Although offering an unsurpassed specificity as compared to other diagnostic techniques such as indirect serological-based enzyme-linked Immunosorbent assay (ELISA) and direct detection such as culture method and PCR, the gold standard suffers from several shortcomings. These include the need for special culture facilities, the requirement to maintain panels of live leptospires, the technically demanding nature of the assay, the time-consuming process characteristics of the method, and the issue of undetectability of the corresponding antibodies when the causative strain is absent in the panel (World Health Organization, 2003).

Even though most human leptospirosis in endemic areas is mild or asymptomatic, a delay in diagnosis could potentially cause the transition of the disease into a more severe form known as Weil's syndrome. This stage has an estimated global fatality rate from <5% to 30% and is characterized by jaundice, renal failure, hemorrhage, and myocarditis with arrhythmias (World Health Organization, 2003). As such, direct detection is always favored for the rapid diagnosis of leptospirosis. However, direct detection techniques such as PCR and culture methods are time-consuming, technically demanding, and highly susceptible to false-negative results due to the presence of inhibitors in the clinical samples (World Health Organization, 2003). To alleviate the issues pertaining to PCR and culture method, direct detection of a *Leptospira*-specific biomarker by a MRE can be a suitable strategy.

Among the characterized outer membrane biomarkers of pathogenic *Leptospira*, LipL32 has many interesting features that render it usable as a primary

diagnostic target of pathogenic *Leptospira*. It is an outer membrane lipoprotein that is not entirely surface-exposed like LipL41 and LipL21 (Cullen *et al*., 2003, Pinne and Haake, 2013, Shang *et al*., 1996). Several studies have indicated the presence of surface-exposed epitopes of LipL32 that are accessible by antibodies (Kumar *et al*., 2016, Maneewatch *et al*., 2014, Pissawong *et al*., 2020). Moreover, RT-PCR analysis of *LipL32* gene evidenced that LipL32 is the protein with the highest copy number of 38000 per cell (Podgoršek *et al*., 2020). LipL32 is also an extracellular matrix-interacting protein that can mediate pathogen-host interactions (Vieira *et al*., 2014).

Hence, diagnostics based on the direct detection of LipL32 can be an efficient strategy as it can reduce the time of experimental execution for immediate diagnostics and treatment. This entails the generation of a molecular recognition element that can directly target LipL32 protein. Owing to the advantages of aptamers such as the absence of batch-to-batch variation, cheaper cost of synthesis, and ease-of functionalization, isolating an aptamer against LipL32 protein is a promising strategy.

1.2 Objectives of the study

Intrigued by the prowess of aptamers, the major aim of this study is to isolate aptamers against LipL32 and to use them in diagnostics of pathogenic *Leptospira,* especially in the form of Enzyme-linked aptasorbent assay (ELASA). The objectives of this study are:

i) To express and purify recombinant LipL32 as the antigenic target of pathogenic *Leptospira.*

ii) To isolate and characterize RNA and DNA aptamers against recombinant LipL32.

iii) To develop enzyme-linked aptasorbent assay (ELASA) based on the isolated aptamers

iv) To generally discuss about the performance between the isolated RNA and DNA aptamers against LipL32

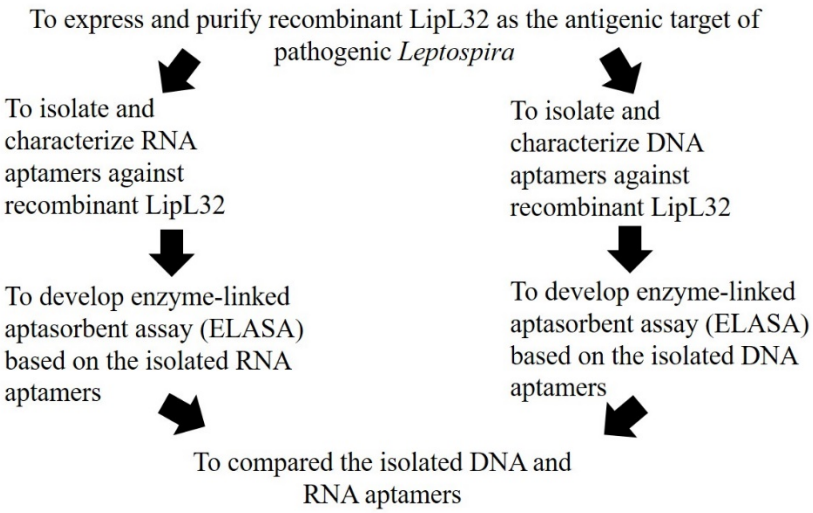

Figure 1.1 Objectives of the study.

CHAPTER 2

LITERATURE REVIEW

2.1 Potentiation of Aptamers as Molecular Recognition Element (MRE)

Derived from the Latin word “aptus” and “meros” which means “to fit” and “particles”, respectively (Ellington and Szostak, 1990), aptamers are an emerging class of molecular recognition element comprised of single-stranded nucleic acids such as ssDNA or RNA capable of forming special three-dimensional (3D) structures and hence enabling them to bind to a myriad of targets with high specificity and affinity similar to antibodies. As such, they are also dubbed as “chemical antibodies”.

The wide repertoire of 3D structures formed by aptamers is attributable to the usage of combinatorial nucleic acid sequences in the randomized region which could promote the formation of vast variety of aptamer structures such as G-quadruplexes, stem loop and pseudoknot (Ditzler *et al*., 2011), making it available for virtually all sorts of target as evidenced in a number of aptamers previously isolated against small molecules, peptides, proteins, whole cells as well as virus particles (Zhou and Rossi, 2017). Moreover, the interaction between aptamers and targets are in general forged by van der Waals forces, electrostatic interactions and hydrogen-bonding (Hermann and Patel, 2000, Nomura *et al*., 2010, Piganeau and Schroeder, 2003).

To qualify as a promising class of molecular recognition element, first, aptamers have demonstrated dissociation constant values ranging from high picomolar to low nanomolar (Kovacevic *et al*., 2018, Maier and Levy, 2016). Secondly, they also exhibited high selectivity by being able to distinguish cognate targets from their structural resemblances of different functional groups, a single amino acid mutation and enantiomeric (Zhou and Rossi, 2017). Next, as aptamers are essentially nucleic

acid molecules, they can be easily functionalized towards diagnostic and therapeutic applications (Odeh *et al*., 2019). Lastly, potentiation of aptamers is further corroborated when several aptamers targeting cancer, cardiovascular disease, macular degeneration, anaemia of chronic diseases and diabetes have been developed and clinically evaluated (Kovacevic *et al*., 2018, Maier and Levy, 2016). For instance, vascular endothelial growth factor (VEGF) is a biomarker responsible for the onset of age-related macular degeneration whereby Pegaptanib RNA aptamer was clinically approved by FDA to reduce vision loss via administration by intravitreal injection (0.3 mg) every 6 weeks (Kovacevic *et al*., 2018, Ng *et al*., 2006).

2.2 Advantages of Aptamers over Antibodies

Historically, aptamers were discovered at a much later period as compared to antibodies (Groff *et al*., 2015), however, it does not affect the impression aptamers have given to the researchers as evidenced in consistent publications of over 1,000 research articles annually since 2010 (Ku *et al*., 2015), owing to the advantages of aptamers possessed over antibodies. First, a much smaller in size of aptamers at about 6-30 kDa as compared to typical antibodies of 150-180 kDa significantly enhanced tissue penetration as observed in solid tumor and even intact human skin (Lenn *et al*., 2018, Xiang *et al*., 2015). Secondly, as another benefit of having a small size, cost of aptamer synthesis is comparatively cheaper than antibodies as well and is further supported by a recent study conducted by Sun and Zu back in 2015 that an aptamer-based flow cytometry is 1,000 times lesser than its antibody counterpart (Sun and Zu, 2015).

Next, an animal-free production of aptamers enables quality of aptamers to be free from batch-to-batch variation, a ‘reproducibility crisis’ encompassed during

production of antibodies (Baker, 2015). Moreover, applicability of aptamers is even profound when they possess a high target selectivity regardless of target immunogenicity while demonstrating a low-to-no immunogenicity (Zhou and Rossi, 2017). In addition, synthetic nature of nucleic acid aptamers renders them able to undergo reversible heat denaturation and renaturation.

2.3 DNA SELEX Versus RNA SELEX

Systemic Evolution of Ligands via Exponential Enrichment (SELEX) is an *in vitro* four-step selection strategy comprised of (i) incubation of nucleic acid pool (ssDNA/ RNA) with target molecules (ii) partitioning and recovery of target-bound nucleic acid molecules (iii) amplification of target-bound nucleic acid molecules using Polymerase Chain Reaction (PCR) for DNA SELEX or Reverse Transcription-Polymerase Chain Reaction (RT-PCR) for RNA SELEX and (iv) regeneration of RNA/ssDNA for next round of SELEX (Figure 1.1).

Based on this strategy, diversity of the nucleic acid pool will be gradually deconvoluted until the convergence of sequences which can only be identified following cloning and sequence analysis of selected nucleic acid pool. Generally, number of SELEX cycles between 8 to 15 cycles in a conventional SELEX is sufficient to successfully identify potent aptamers (Wang *et al*., 2012, Blind and Blank, 2015). Typically, a SELEX library comprised of 20-60 randomized nucleotides of adenine, guanine, cytosine, and thymine for DNA SELEX, whereby uracil is substituted in RNA SELEX and hence giving rise to approximately 4^n sequences (Tsao *et al*., 2017, Komarova and Kuznetsov, 2019). Furthermore, two constant primer binding regions are also incorporated to flank the randomized region to facilitate recovery of target-bound sequences throughout entire SELEX process.

The major difference between DNA SELEX and RNA SELEX lies in the way single-stranded DNA or RNA is generated which can be achieved by tweaking primer design. For RNA SELEX, a T7 promoter region is included during the design of forward primer recognizable by an efficient and highly selective T7 RNA polymerase (Tabor, 2001). On the other hand, as for DNA SELEX, generation of ssDNA is mainly reliant on those major ssDNA-generating techniques such as conventional biotin-streptavidin separation, lambda exonuclease enzymatic digestion and asymmetric PCR (A-PCR) that focus on modification, labeling or primer ratios of reverse primer (Marimuthu *et al*., 2012, Hao *et al*., 2020). A combination of these techniques even enhanced production of ssDNA (Svobodová *et al*., 2012, Yeoh *et al*., 2022).

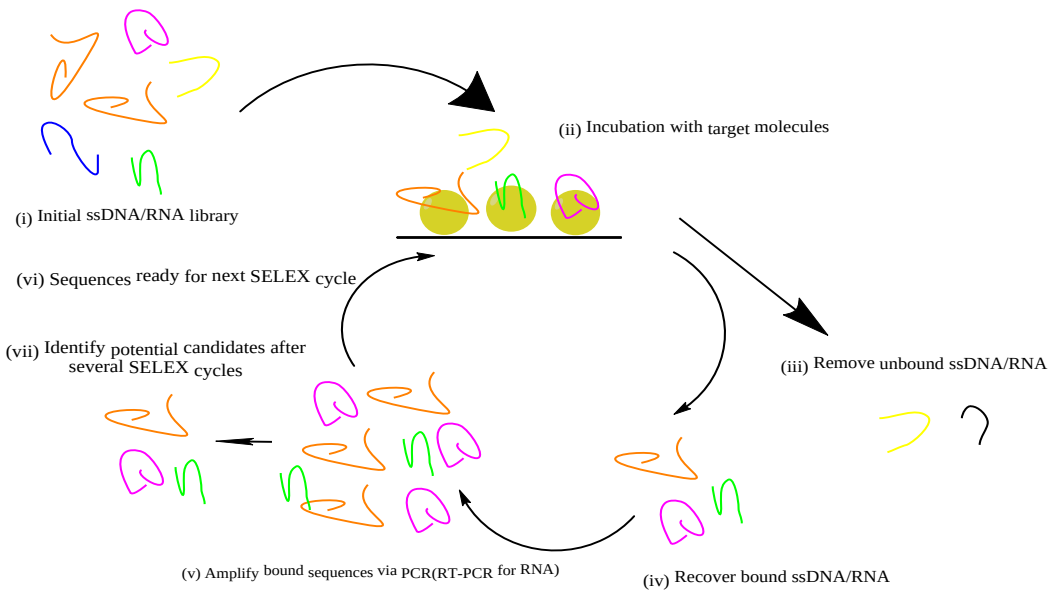

Figure 2.1: A general flow of SELEX. (i) Initial single stranded DNA/ RNA library (ii) is incubated with target molecules (protein) (iii) and partitioning and washing steps are performed to separate and remove the unbound nucleic acids. (iv) Bound sequences are recovered (v) and amplified via PCR (ssDNA) or RT-PCR (RNA) before starting a new SELEX cycle. (vi) The pool will be converted back to either ssDNA or RNA before starting a new SELEX cycle. (vii) After 8-15 cycles, candidate aptamers can be identified for further analyses.

2.4 Aptamers Isolated Against Bacterial Pathogens

Aptamers hold a promising outlook especially in diagnostic application. Over the years, this statement is further corroborated when a plethora of aptamers have been successfully generated against bacterial pathogens such as *Salmonella enterica*, *Staphylococcus aureus*, *Escherichia coli*, *Mycobacterium tuberculosis*, *Campylobacter jejuni* and others (Davydova *et al*., 2016). For instance, S-$PS_{8.4}$ RNA aptamer isolated against a structural protein type IVB pili of *S. enterica* Serovar *Typhi* was able to detect a single colony-forming unit (CFU) of target *S. enterica* in solution with a comparable detection limit close to real-time PCR assay using a potentiometric aptamer-based biosensor (aptasensor) (Zelada-Guillén *et al*., 2009). Moreover, in another report, the DNA aptamer 37 against exclusively expressed fimbriae protein of enterotoxigenic *E. coli* strain K88 (ETEC K88) successfully demonstrated high specificity against *E. coli* strain K88 from other bacteria such as ETEC K99, *S. aureus*, *E. coli* TOP 10 . Lastly, the 2'-Fluoro I-2 RNA aptamer of high affinity against *OmpC* of *S. Typhimurium* was able to recognize intact bacterial cells as compared to other aptamers while exhibiting a high target selectivity against Gram-positive *S. aureus* or Gram-negative *E. coli* O157:H7 (Han and Lee, 2013).

2.5 Brief History of Leptospirosis

Leptospirosis is a potentially life-threatening zoonosis caused by pathogenic *Leptospira* whereby it was first reported back in 1886 by Adolf Weil, a German physician who described a specific type of jaundice associated with renal dysfunction, splenomogaly, skin rashes and conjunctivitis (Adler, 2015). Owing to his contribution in this field, leptospirosis is also known as Weil's syndrome.

Despite being discovered in 1886, it was believed that leptospirosis began to exist for millennia prior to the advent of modern medical and scientific literature which was described in previous outbreaks such as Japan syndromes were termed as “seven-day fever” or “autumn fever” and “rice field jaundice” as described in ancient Chinese texts (Adler, 2015).

2.6 Global Epidemiology of Leptospirosis

Globally, leptospirosis has an estimated epidemiology of 1.03 million cases with nearly 60, 000 cases per annum (Costa *et al*., 2015, Torgerson *et al*., 2015). However, this figure is only a gross estimate as majority of the cases are either undiagnosed or misdiagnosed due to the lack of awareness, difficulties in performing laboratory-based confirmatory tests and the overlapping clinical presentations with that of dengue and other haemorrhagic fevers. The highest median annual incidence of Leptospirosis occurs in African regions (95.5 per 100, 000) followed by Western Pacific (66.4), the Americas (12.5), South-East Asia (4.8) and Europe (0.5) (World Health Organization, 2011). Despite being endemic in tropical and subtropical regions, a previous study has indicated that this disease could readily turn epidemic after heavy rainfall and flooding (Haake and Levett, 2015). Even though prospective surveillance studies proposed that most human leptospirosis in endemic areas are mild or asymptomatic (Haake and Levett, 2015), the initial stage of *Leptospira*l infection, upon neglection, could potentially transition into to the stage characterized by Weil’s syndrome with an estimated global fatality rate ranging from <5 to 30 % due to jaundice, renal failure, haemorrhage, and myocarditis with arrhythmias (World Health Organization, 2003).

2.7 Epidemiology of Leptospirosis in Malaysia

Due to the location of Malaysia, which is in tropical region of South-East Asia, leptospirosis is endemic in Malaysia which was evidenced in the first international leptospirosis outbreak in the Echo-challenge in Borneo island which involved 304 athletes from 26 different countries, with 29 athletes hospitalized with no death (Garba *et al.*, 2017).

Due to increasing number of cases of leptospirosis, from 263 cases in 2004 to 1976 cases in 2010, this disease has become a notifiable disease in 2010 whereby probable or confirmed cases must be notified to relevant health district, with 3,665 and 4,457 probable and laboratory confirmed cases reported in 2012 and 2013, respectively and an overall case fatality rate of 1.47 % over 2-year period (Garba *et al*., 2017, Tan *et al*., 2016). Next, as compared to the number of reported cases of leptospirosis over the past 10 years, a spike in the number of cases in recent years can be attributable to better awareness on the mode of transmission of leptospirosis among Malaysians and diagnostic techniques (Garba *et al*., 2017). Moreover, evidence of 16 confirmed outbreaks whereby most are associated with residential areas indicates that leptospirosis is an endemic disease in Malaysia (Mohd Hanapi *et al*., 2021).

2.8 Classification and Pathogenesis of *Leptospira*

Leptospira is a Gram-negative, spiral-shaped and flexible spirochete with internal flagella which can be examined microscopically using dark-field microscope. It was classified under pathogenic (*Leptospira interrogans*), intermediate and non-pathogenic or saprophytic strains (*Leptospira biflexa*) (Wilkinson *et al*., 2021). The differences between these strains are (1) saprophytic strains prefer to live between 1 and 35 °C and do not cause an infection, (2) intermediate strains prefer to live between

1 and 37 °C and live as pathogens or saprophytes (3) pathogenic strains prefer to live between 20 and 37 °C and do cause infection in humans and rodents (Samrot *et al.*, 2021).

On the other hand, pathogenesis of *Leptospira* begins when the pathogenic *Leptospira* gains entry to the body via open wounds based on direct or indirect contact with infected animals or environmental water and soil contaminated with urine of infected rodents which is followed by penetration and multiplication in the host organs such as kidney, liver or central nervous system (Samrot et al., 2021). Upon clearance of spirochetes from blood or host tissues, inevitably, the pathogenic *Leptospira* persist and multiply for a certain period in the kidney tubules and subsequently shed into the urine.

2.9 Diagnosis of Leptospirosis

In general, diagnostics of pathogenic spirochetes are performed via gold standard microscopic agglutination test (MAT), indirect detection by serological-based assay such as enzyme-linked Immunosorbent assay (ELISA) and direct detection such as culture method and Polymerase Chain Reaction (PCR) (World Health Organization, 2003). Although offering an unsurpassed specificity, the gold standard suffers from several downsides such as the need for special facilities to culture, the requirement to maintain panels of live leptospires, technically demanding nature of the assay, time-consuming process characteristics of the method and the issue of undetectabilities of the corresponding antibodies when the culprit strain is absent in the panel (World Health Organization, 2003).

Direct detection is preferred over indirect serological detection as all serological-based tests are heavily reliant on the time required for a sufficient amount

of the anti-*Leptospira*l antibodies (seroconversion) to be reached and as such rapid diagnosis of leptospirosis is impossible. On the other hand, direct detection techniques such as PCR and culture method are time-consuming, technically demanding and highly susceptible to false-negative results due to the presence of inhibitors in the clinical samples (World Health Organization, 2003).

2.10 LipL32 Protein is the Most Immunodominant Outer Membrane Protein Effective as the Diagnostic Target of Pathogenic *Leptospira*

To date, more than 250 Serovars of *Leptospira* have been identified whereby the antigenic diversity between Serovars is attributable to the variation at the carbohydrate moiety of lipopolysaccharides (LPS) located at the surface of *Leptospira*l Serovars (Adler and de la Peña Moctezuma, 2010), rendering the development of a diagnostic assay which is highly specific against pathogenic *Leptospira* difficult. On the other hand, a strategy based on the usage of an outer membrane biomarker exclusively expressed by the culprit pathogen can be an alternative as compared to LPS.

Outer membrane proteins that are related to bacterial pathogenesis of Gram-negative bacteria can be particularly useful in differentiating pathogenic strains from non-pathogenic ones (Keenan *et al*., 2000, Cullen *et al*., 2004, Ellis and Kuehn, 2010). Derived from '32 kDa Lipoprotein from *Leptospira'*, LipL32 is an outer membrane lipoprotein associated with pathogenesis owing to its high-level expression during both cultivation and infection while being highly conserved among pathogenic *Leptospira* (Haake *et al*., 2000). With an apparent molecular weight of 32 kDa on SDS-PAGE analysis (Haake *et al*., 2000), an intact mass profile of outer membrane vesicles from a clinical isolate of *Leptospira interrogans* Serovar *Copenhageni* following a LC-MS+ analysis unveiled its actual molecular weight to be in between 28, 468 to 28, 580

dalton and is lipid-incorporated at cysteine 20 (Haake *et al*., 2000, Nally *et al*., 2005). Although multiple isoforms of LipL32 have been detected due to the iron-mediated truncation by cysteine protease from its carboxyl-terminus, the major isoform remained to be the intact LipL32 (Zuerner *et al*., 1991, Cullen *et al*., 2002).

Several interesting features inherent in LipL32 have enabled this antigen a prime diagnostic target of pathogenic *Leptospira* are as follows: First, as an outer membrane lipoprotein that is not entirely surface exposed like LipL41 and LipL21 (Shang *et al*., 1996, Cullen *et al*., 2003, Pinne and Haake, 2013), several studies have in fact indicated the presence of surface exposed epitopes accessible by antibodies (Maneewatch *et al*., 2014, Kumar *et al*., 2016, Pissawong *et al*., 2020).

Reverse Transcription-Polymerase Chain Reaction (RT-PCR) analysis of *LipL32* gene by Podgoršek *et.al*. clearly highlighted LipL32 as the most copious protein of pathogenic *Leptospira* with a copy number of 38, 000 per cell (Podgoršek *et al*., 2020). Moreover, several pieces of evidence that LipL32 is an extracellular matrix-interacting protein further shed a light on the presence of surface exposed epitopes in mediating pathogen-host interactions (Vieira *et al*., 2014).

In a study conducted by Guerreiro *et. al.*, quantitative and qualitative immunoblot analysis based on the humoral immune response exhibited by human patient sera from 105 patients from Brazil and Barbados unveiled anti-LipL32 reactivity as the highest sensitivity of 37 and 84 % in acute and convalescent phase, respectively while demonstrating the highest specificity with only 5 % of positive reactions in healthy community control (Guerreiro *et al*., 2001). Apart from immunoblot analysis, serodiagnosis have also been carried out using an Enzyme-linked Immunosorbent Assay (ELISA) for the detection of anti-LipL32 IgM and IgG antibodies in human sera by using recombinant LipL32 as antigen (Saengjaruk *et al*.,

2002, Bomfim *et al*., 2005, Boonyod *et al*., 2005, Aviat *et al*., 2010, Chalayon *et al*., 2011, Vedhagiri *et al*., 2013, Ye *et al*., 2014, Pissawong *et al*., 2020), corroborating the high diagnostic value of this protein effective in rapid diagnosis of leptospirosis.

2.11 Enzyme-linked Apta-sorbent Assay (ELASA) as the Ideal Platform for the Rapid Detection of Antigens

Enzyme-linked Immunosorbent Assay (ELISA) is regarded as one of the most established assay expedient for a rapid detection of analytes in a given sample based on the usage of antibodies and was previously shown to recognize bacterial pathogens such as *Escherichia coli, Salmonella* spp., *Campylobacter* spp. and *Bacillus cereus* (Verma *et al*., 2013, Zhao *et al*., 2014).

Among several configurations in ELISA, sandwich ELISA is the most effective form by relying on the usage of two antibodies (Zhao *et al*., 2014). By replacing antibodies with aptamers as the MRE, this assay is dubbed as sandwich Enzyme-linked Apta-sorbent Assay (Sandwich ELASA) (Toh *et al*., 2015). For example, the TK1_apta37 and TK1_apta69 DNA aptamer sandwich pair isolated against Thymidine kinase 1 (TK1) serum biomarker that is elevated during early stages of malignancies had a dynamic concentration range of 54-3,500 pg/ mL, covering the clinically relevant serum levels (Nazari *et al*., 2019).

CHAPTER 3

EXPRESSION AND PURIFICATION OF RECOMBINANT LIPL32 AND ISOLATION OF RNA APTAMER BY SELEX

3.1 Introduction

RNA aptamers are better than DNA aptamers due to its vast structural complexity far more than DNA aptamers which is facilitated by the presence of 2'-OH group, which could enhance the chance of isolating RNA aptamers of high affinity and specificity. Meanwhile, LipL32 is an outer membrane biomarker that is exclusively expressed by pathogenic *Leptospira*, the culprit organism responsible for leptospirosis. Owing to its high degree of conservation among pathogenic *Leptospira* and highest expression level as compared to other biomarkers (Haake and Levett, 2015, Podgoršek *et al*., 2020), development of RNA aptamer against purified LipL32 protein as the diagnostic agent is therefore highly desirable.

As recombinant LipL32 protein is no longer available commercially, this protein was obtained via *in-house* purification by using pAE-LipL32 plasmid which was given by Prof Dellagostin. In this study, *LipL32* gene which was previously subcloned into pAE-LipL32 plasmid was expressed using a *E. coli*-based bacterial expression system and purified under native condition (Seixas *et al*., 2007). To enhance the purity of expressed LipL32 protein, purification was performed by using Talon Resin (Clontech, CA, USA) under high salt condition. Upon purification, western blot was then carried out to confirm the identity of purified LipL32 protein. Following this, a total of 12 cycles of tripartite-hybrid SELEX were carried out, the selected RNA pool was reverse-transcribed, PCR-amplified and subjected to cloning and sequence analysis. After cloning and sequence analysis of 29 plasmids, the

cumulative effect of several parameters such as frequency of appearance (%), percentage of guanine (%) and the lowest predicted Gibbs free energy (kcal/ mol) were examined via an *in-house* python-aided unbiased data sorting for identification of potent aptamers whereby the top 5 candidates that fulfilled the criteria were chosen for subsequent validation by direct ELASA. Lastly, the most potent aptamer was then subjected to several characterizations such as "rational truncation" approach guided by Mfold analysis for structural optimization (Rockey *et al*., 2011), web-based QGRS Mapper for prediction of G-quadruplex (Kikin *et al*., 2006), dissociation constant, Kd estimation and limit of detection, LOD determination. Details of this chapter is depicted as shown in Figure 3.1.

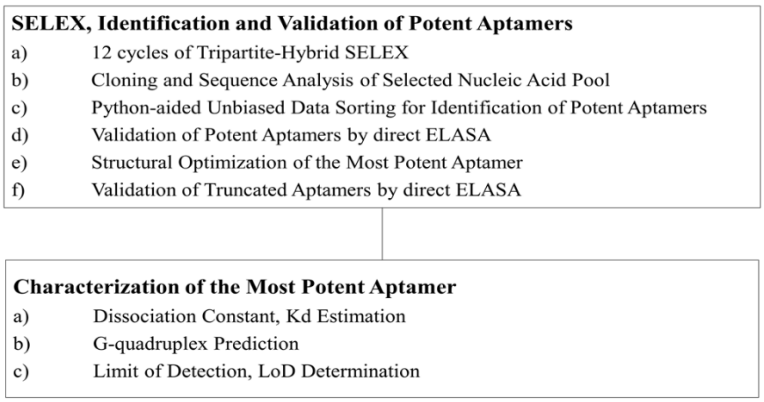

Figure 3.1 General flow of this chapter from isolation to characterization of RNA aptamer.

3.2 Materials and Methods

3.2.1 LipL32 Expression and Purification

Prior to protein expression, the open reading frame of LipL32 gene as per Seixas *et al*. (2007) was retrieved from NCBI database (GenBank: AE016823.1) and scrutinized manually on SnapGene Viewer 4.3.10 for the identification of rare codons. The

recombinant pAE-LipL32 plasmid was chemically transformed into BL21(DE3)pLysS (Merck, KGaA, Darmstadt, Germany) and Rosetta 2(DE3)pLysS (Merck, KGaA, Darmstadt, Germany) strains. LipL32 expression was first carried out in BL21(DE3)pLysS, aided by the induction with IPTG. The optimal induction time was determined by SDS-PAGE analysis, which is guided by ImageJ analysis of the band intensity (Schneider *et al*., 2012). The cells were pelleted by brief centrifugation and lysed by sonication on ice (4 times for 30 sec each time with 10 sec interval). The resulting lysate was resuspended either in low-salt Binding buffer (20 mM Tris-HCl (pH 8.0), 10 % glycerol, 500 mM NaCl) or high-salt Binding buffer (20 mM Tris-HCl (pH 8.0), 10 % glycerol, 1 M NaCl) containing pre-equilibrated TALON Resin (Clontech, CA, USA) and rotated end-over-end on a rotator for 30 minutes at 4 °C. The columns were washed 5 column volumes with low-salt washing buffer (20 mM Tris-HCl (pH 8.0), 10 % glycerol, 500 mM NaCl, 100 mM Imidazole) or high-salt washing buffer (20 mM Tris-HCl (pH 8.0), 10 % glycerol, 1 M NaCl, 100 mM Imidazole). After elution using ice-cold Elution buffer (20 mM Tris-HCl (pH 8.0), 10 % glycerol, 500 mM NaCl, 300 mM Imidazole), the protein was then dialyzed against 2 L of Dialysis buffer (20 mM Tris-HCl (pH 8.0), 10 % glycerol, 150 mM NaCl) by using a dialysis tubing (TOR-3K, 3.5k MWCO; Nippon Genetics Co. Ltd., Tokyo, Japan). Dialysis was carried out at 4 °C using a magnetic stirrer for a total period of 24 hours at 8-hour intervals with constant stirring.

3.2.2 Western Blot

The purified protein was resolved on 10 % (w/v) SDS-PAGE and electro-blotted onto a mini-Trans-Blot Turbo Pack PVDF membrane (Bio-Rad, California, USA) by using a Trans-Blot Turbo Semi-Dry Blotting System (Bio-Rad, California, USA) for 7 min

at 20 mA. The membrane was blocked for 30 min with 5 % (w/v) BSA in TBS-T (0.05 % Tween-20 in Tris buffered Saline). The membrane was probed using HisDetector™ Nickel-HRP (Sera Care, MA, USA) with the dilution of 1: 5000. Washing step was performed thrice using TBS-T at every step for 15 min each time. SuperSignal West Pico Chemiluminescent substrate solution (Thermo Scientific, Massachusetts, USA) was used to detect formation of purified protein-Nickel-HRP complexes by incubating it for 1 min followed by imaging by using VersaDoc 4000 MP (Bio-Rad, California, USA).

3.2.3 SELEX

SELEX library with a randomized region of 40-mer and primers were bought from Integrated DNA Technologies. The sequence of the combinatorial library is 5'-GGG GGA ATT TC*T AAT ACG ACT CAC TAT AG*G GAG GAC GAT GCG G-N40-GGC ACC ACG GTC GGA TCC AC-3'. The forward and reverse primers are 5'-GGG GGA ATT TCT AAT ACG ACT CAC TAT AG-3' and 5'-TCT CGG ATC CTC AGC GAG TCG TC-3', respectively, whereby the T7 promoter sequence is underlined and italicized. The initial RNA pool was derived from the PCR-amplified ssDNA pool using Ampliscribe™ T7 Flash™ Transcription Kit (Epicenter, Wisconsin, USA), following the manufacturer's instruction. Following transcription, RNA was selectively purified using "crush and soak" method (Citartan *et al*., 2012). In the first cycle of SELEX, the reaction mixture was prepared under the following condition: 6.02 µM of RNA pool was dissolved in 1X SELEX binding buffer (10 mM HEPES-KOH [pH 7.4], 150 mM NaCl), heat denatured at 95 °C for 2 min before cooling to room temperature (RT) for 10 min to allow the proper folding of RNA molecules and added with 12.5 µM of yeast tRNA (Invitrogen Corporation, Carlsbad, USA) prior to

the addition of 1.88 µM of LipL32 protein. The reaction mixture was then incubated at RT for 15 min followed by partitioning of LipL32-bound molecules from the unbound molecules. In this study, SELEX has been intermittently carried out using nitrocellulose filter membrane, microtiter plate-based and Native PAGE-based partitioning. To increase the stringency of SELEX condition, concentrations of RNA pool, LipL32 protein and yeast tRNA were progressively manipulated. Following the partitioning step, the LipL32-bound RNA molecules were heat-denatured using urea at 95 °C for 2 min, ethanol precipitation assisted by Dr. GenTLE^{TM} precipitation carrier (Takara Bio, Shiga prefecture, Japan) followed by reverse transcription by using AMV Reverse transcriptase (Promega, Wisconsin, USA), following manufacturer's instruction. Upon completion of reverse transcription, the resulting cDNA was subjected to PCR amplification, ethanol precipitation and *in vitro* transcription. The *in vitro* selection has been carried out for a total of 12 cycles.

3.2.4 Native PAGE-based partitioning method at SELEX cycle 12

After 11 cycles of SELEX, one additional SELEX cycle was carried out following 8% Native PAGE-based partitioning. The Native PAGE-based gel mobility shift assay was performed under the following conditions: A 10 µL reaction mixture was prepared, which contains the folded RNA and LipL32 protein before incubation at RT for 15 min. Following the incubation, the reaction mixture was added with 2 µL of 90 % glycerol to a final concentration of 15 %, loaded onto a 8% Native Polyacrylamide gel and run at 140 V for 45 min. Next, the polyacrylamide gel was carefully removed, stained with 0.5X TBE (40 mM Tris-HCl, pH 8.3, 45 mM boric acid, 1 mM EDTA) supplemented with ethidium bromide at a concentration of 0.5 µg/ mL (Sigma, St Louis, USA) for 10 min and rinsed with ddH2O before visualization on Bio-rad Gel

Doc XR+ System (Bio-rad Laboratories, Hercules, USA). ImageJ analysis (Schneider *et al*., 2012) was used to estimate the intensity of the band that constitutes the RNA-protein complex. Next, the band corresponding to the LipL32-bound RNA was selectively excised and purified as previously described by Citartan *et. al.* (Citartan *et al*., 2012), with the aid of Dr GenTLE™ Precipitation Carrier. Upon recovery of the LipL32-bound RNA molecules, the molecules were reverse-transcribed, PCR-amplified and subjected to cloning using pCR™ 2.1 TOPO™ TA Cloning Kit, following the manufacturer's instructions. Following blue-white screening, 29 white colonies were selected and subjected to plasmid extraction using Roche High Pure Plasmid Isolation Kit (Roche Diagnostics GmbH, Mannheim, Germany). The extracted plasmids were sent for sequencing performed by First BASE Laboratories Sdn. Bhd., Selangor, Malaysia.

3.2.5 Python-aided unbiased data sorting

To select potential aptamers for the binding assessment with LipL32, multiple parameters such as frequency of appearance (%), percentage of guanine (%) and lowest predicted Gibbs free energy (kcal/ mol) were analyzed. First, sequences were analyzed for sequence homology and similar sequences were clustered together. Next, the percentage of guanine (%) and the lowest Gibbs free energy (kcal/ mol) were estimated with the aid of Mfold program (Zuker algorithm) under the default setting (Zuker, 2003). The secondary structure with the lowest Gibbs free energy was selected as the most potent conformation for each sequence. Frequency of appearance (%) and percentage of guanine (%) were sorted in descending orders while Gibbs free energy (kcal/ mol) was ranked in ascending order using Python's Panda package. Each ranked sequence was given a weightage value and was subjected to all sorting conditions (all

possibilities) of the multiple parameters. This in-house developed program is to prevent any biased data sorting. The top 5 sequences that fulfilled all the parameters were considered eligible for the subsequent validation using ELASA.

3.2.6 Direct Enzyme-linked Apta-sorbent Assay (ELASA)

Prior to the development of direct ELASA, the selected sequences from the python-based unbiased data sorting were functionalized via extension of selected sequences with Poly-A tail by PCR followed by duplex formation using 5'-biotin functionalized dT-20 (Prabu *et al*., 2020). LipL32 protein was serially diluted from 0, 200 and 400 nM in 1X PBS (137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4, and 2 mM KH_2PO_4) and coated overnight at 4 °C onto the wells of a Nunc Maxisorp™ microtiter plate (Nunc, New York, USA). Next, the wells were washed once with 300 µL of PBST (1X PBS, 0.05 % Tween-20) and blocked at 37 °C for 2 hours with 300 µL of Superblock Blocking Solution (Thermo Scientific, Massachusetts, USA) in 1X PBS solution. The wells were then washed thrice with PBST. Parallelly, equimolar amount of the functionalized RNA sequences extended with poly (A) tail at the 3' end was mixed with 5'-biotinylated $(dT)_{20}$ in LINA buffer, heated at 95 °C for 2 min followed by cooling to RT for 10 min. Fifty picomoles of biotin-functionalized sequences were incubated in each well at RT for 1 hour. The wells were washed four times with LINA-T (150 mM NaCl, 150 mM LiCl, 0.05 % Tween-20). Signal production was carried out using Streptavidin-HRP (Thermo Scientific, Massachusetts, USA) and Poly HRP-Streptavidin (Thermo Scientific, Massachusetts, USA), both in LINA buffer containing 3 % BSA at the dilution of 1:1000 at RT for 1 hour. The unbound streptavidin-HRP conjugates were removed by washing four times with LINA-T followed by the addition of 100 µL of TMB Chromogen Solution (Thermo Scientific,

Massachusetts, USA) and allowed to react at RT for 30 min. Finally, the reaction was stopped by 100 µL of 1M HCl and the absorbance @ 450 nm was taken with an ELISA Plate Reader (Tecan, Männedorf, Switzerland).

3.2.7 Structural Optimization of the Aptamer based on Rational Truncation Approach

The most potent aptamer was then subjected to structural optimization based on rational truncation approach (Rockey *et al*., 2011). Two mini versions of the RNA aptamer namely LepRapt-11a and LepRapt-11b were constructed based on the most potent conformation predicted by Mfold (Zuker, 2003). The variants of the aptamers were validated by direct ELASA.

3.2.8 Dissociation Constant, Kd Estimation of the Optimized RNA Aptamer on ELASA

In this study, two different reactions were performed for the determination of the dissociation constant of the RNA aptamer. In the first reaction, the amount of LipL32 protein was varied while the amount of aptamer was fixed at a constant value. In the second reaction, the amount of aptamer was varied while the amount of LipL32 protein was kept constant. For a better signal amplification, Poly HRP-Streptavidin was used. In the first reaction, LipL32 was serially diluted from 0 to 2400 nM and coated onto the wells while a constant amount of 500 nM of biotin-functionalized RNA aptamer was used in each well. Parallelly in the second reaction, the RNA aptamer was serially diluted from 0 to 3200 nM and titrated against a constant amount of LipL32 of 400 nM coated onto each well. The $OD_{450\ nm}$ readings were recorded and used to plot a hyperbolic curve. Subsequently, the dissociation constant Kd was estimated using a

non-linear regression curve via GraphPad Prism software (version 6.05) (GraphPad Software Inc., California, USA).

3.2.9 Determination of Limit of Detection, LOD against LipL32

LipL32 was serially diluted from 0 to 100 nM and titrated against 50 pmol of the biotin-functionalized RNA Aptamer to determine its limit of detection, LOD. As the control, 200 nM of BSA (Promega, Wisconsin, USA) was also coated onto one of the well. In this study, the limit of detection is defined as the lowest concentration of LipL32 that produces a $OD_{450\ nm}$ reading that is statistically significant than that of the blank (0 nM).

3.2.10 Statistical Analysis

To test the nature of the data to assume Gaussian distribution, all $OD_{450\ nm}$ readings were subjected to Shapiro-Wilk test. Following this, the data was then subjected to analysis of variance (ANOVA) followed by Tukey's post hoc. All experiments were carried out in triplicates and the quantitative results were expressed as mean ± standard deviation. T-test and ANOVA were carried out via GraphPad Prism 6.05, with $P < 0.05$ considered to be statistically significant.

3.2.11 QGRS Mapper-assisted Prediction of G-Quadruplex

When a primary nucleic acid sequence possesses four runs of at least three guanines separated by short stretches of other bases, it could potentially fold into an intramolecular G-quadruplex structure (Kwok and Merrick, 2017), which can be predicted with the aid of QGRS Mapper (Kikin *et al*., 2006). The G-quadruplex formation of the aptamer was analyzed by using QGRS Mapper. The sequence of the