IN VITRO SELECTION OF RNA APTAMERS THAT SPECIFICALLY BIND TO 50kDa OUTER MEMBRANE PROTEIN OF *Salmonella enterica* serovar Typhi

By

SITI MARIAM MAWARNI RAMLI

Thesis submitted in fulfillment of the requirements for the degree of

Master of Sciences

February 2016

ACKNOWLEDGEMENT

With this thesis, at last I am finishing my master's degree at Institute for Research in Molecular Medicine (INFORMM), Universiti Sains Malaysia (USM). I would like to express my gratitude to all who had given me the possibility to complete this research study and thesis. Thank you to God, for His blessing, so that I am able to finish my study and present this thesis. I am deeply indebted to my supervisor Associate Prof. Dr. Hamid Fauzi whose gave me this opportunity to purse my study in master degree. Distinctive thanks also to my latest main supervisor Professor Dato' Dr. Asma Ismail and Dr. Khairul Mohd Fadzli Mustaffa**,** whose help, stimulating scientific questions and arguments, suggestions and encouragement throughout this research study and led me writing of this thesis. I am also deeply grateful to INFORMMs lecturers who gave constructive comments to my study. During this study I also work together with many colleagues for whom I have great regards, and I wish to extend my warmest thanks to all who had helped me in at INFORMM. I would like to thank the lab staff and fellow friends who helped me throughout my study. I wish to thank all the administration staff for helping me to get through the difficult times, and the care they provided. I cannot end without thanking my family especially my parents, Mr. Ramli Hussain and Mrs. Siti Zaharah Yaacob, who constantly encourage and love me throughout my life time. Last but not least, special thanks to my husband Zul Fahmi Che Mahmud, for his patience and love which enabled me to complete this thesis. I also would like to thank the E-science fund (305/CIPPM/6113210) from MOSTI and the USM Research University Grant for supporting this study.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

Abbreviation		Description
%	=	Percentage
°C	=	Degree (temperature)
µg	=	Microgram
µl	=	Microliter
µm	=	Micrometer
A	=	Adenosine
A_{260}	=	absorption at λ=260 nm
A_{280}	=	absorption at λ=280 nm
ATP	=	adenosine triphosphate
bp	=	Base pair
C	=	cytosine
CTP	=	cytidine triphosphate
dH_2O	=	Distilled water
DNA	=	Deoxyribonucleic acid
DNase	=	deoxyribonuclease
dNTP	=	deoxyribonucleotide triphosphate
ds	=	Double stranded
DTT	=	Dithiothreitol
EDTA	=	Ethylene diamine tetraacetic acid
EMSA	=	Electrophoretic Mobility Shift Assay
g (for acceleration)	=	Gravitational force (relative centrifugal force)
g	=	Gram
G	=	Guanosine
GTP	=	Guanosine triphosphate
HCl	=	Hydrochloric acid
HPLC	=	high performance liquid chromatography
IgE	=	Immunoglobulin E
IgG	=	Immunoglobulin G
IgM	=	Immunoglobulin M
K_d	=	Dissociation constant
kDa	=	Kilo Dalton
LB	=	Luria Bertani
M	=	Molar or Molarity

MBU	=	Molecular Biology Unit
mg	=	miligram
$MgCl_2$	=	Magnesium Chloride
ml	=	Mililliter
MW	=	Molecular Weight
NaCl	=	Sodium Chloride
NaOH	=	Sodium Hydroxide
mM	=	milimolar
nm	=	Nanomolar
nmole	=	Nanomole
OD	=	Optical density
OMP	=	Outer membrane protein
PAGE	=	Polyacrylamide gel electrophoresis
PCR	=	Polymerase Chain Reaction
pmol	=	Picomole
RNA	=	Ribonucleic acid
RNAse	=	Ribonuclease
RT-PCR	=	Reverse trancription PCR
S.	=	*Salmonella*
SDS	=	Sodium Dodecyl Sulfate
SELEX	=	Systematic Evolution of Ligands by Exponential Enrichment
ss	=	Single stranded
T	=	Thymine
TBE	=	Tris Base EDTA
TEMED	=	TEMED (N,N,N',N'-tetramethylethylenediamine)
TLC	=	Thin layer chromatography
Tris	=	Tris(hydroxymethyl)aminomethane
U	=	Uracil (RNA)
U	=	units of enzymatic activity
UTP	=	uridine triphosphate
UV	=	Ultra violet
V	=	Volt (s)
WHO	=	World Health Organization
w/v	=	weight per volume
X-gal	=	5-bromo-4-chloro-3-indolyl- beta-D-galactopyranoside

ABSTRAK

Demam kepialu merupakan penyakit bawaan makanan yang disebabkan oleh jangkitan bakteria *Salmonella enterica* serovar Typhi (*S.* Typhi). Demam kepialu masih lagi menjadi satu permasalahan kesihatan awam di kebanyakan negara mundur dan negara yang sedang membangun termasuk Malaysia. Dianggarkan kira-kira 22 juta kes dan 216,000 kematian akibat daripada jangkitan *S.* Typhi telah dilaporkan di seluruh dunia setiap tahun. Membran protin luar (OMP) spesifik dengan berat molekul pada 50 ribu Dalton (50 kDa), daripada *S.* Typhi telah dilaporkan berkemungkinan memainkan peranan dalam jangkitan demam kepialu. Oleh sebab itu OMP 50 kDa daripada *S.* Typhi telah digunakan sebagai molekul sasaran dalam mengesan *S.* Typhi. Terkini, antibodi monoklonal dan poliklonal telah banyak digunakan untuk pembangunan kit diagnostik bagi mengesan antigen daripada organisma penyebab penyakit. Namun demikian disebabkan ketidaktahanan antibodi tersebut ke atas suhu yang tinggi atau keperluan mengekalkan rantaian suhu telah menyebabkan kesukaran kit diagnostik tersebut digunakan dilapangan. Sebagai pendekatan baharu, aptamer RNA telah dipilih untuk mengesan antigen daripada bakteria *S.* Typhi untuk mendiagnosis demam kepialu.. Berbanding antibodi, beberapa ciri aptamer seperti ketahanan kepada suhu yang tinggi menyebabkan ia lebih bagus untuk digunakan dalam teknologi diagnostic dimana ia boleh mengantikan antibodi pada masa akan datang. Aptamer merupakan satu jalinan asid nukleik sintetik yang membentuk struktur dimensi yang unik, membolehkan ia berinteraksi khusus kepada molekul sasaran. Aptamer dihasilkan melalui teknik evolusi ligan bersistematik dengan cara penkayaan eksponen (SELEX). Satu proses menyaring ligan tertentu dari kumpulan besar “oligonucleotides” melalui ulangan proses pemilihan dan penggandaan, Melalui kajian ini, SELEX telah berjaya mengasingkan enam calon

aptamer. Analisis jujukan menunjukkan urutan konsensus GUU, GUUU, dan GUUUU muncul di kebanyakan calon-calon aptamer yang berpotensi menjadi tapak lekatan pada molekul sasaran. Aktiviti pelekatan antara aptamer dan protin sasaran telah dinilai melalui ujian peralihan pergerakan secara elektroforetik (EMSA). Daripada enam calon aptamer yang diperolehi hanya aptamer ST01, ST03 dan ST04 berinteraksi khusus dengan protin OMP 50 kDa *S.* Typhi. Oleh kerana protin OMP 50 kDa merupakan satu gabungan protin komplek yang terdiri daripada 3 subunit, subunit 1 protin flagellin, subunit 2 "glycerol kinase" dan subunit 3 protin TolC, maka tiga aptamer terpilih telah diuji seterusnya ke atas subunit-subunit tersebut secara berasingan. Keputusan menunjukkan bahawa aptamer ST01 dan ST03 dapat berinteraksi khusus kepada kedua-dua subunit 2 dan 3, manakala ST04 tidak berinteraksi pada mana-mana subunit. Kesimpulannya, ST04 hanya melekat kepada gabungan komplek protin OMP 50 kDa dan aptamer ST01 dan ST03 boleh mengikat kepada gabungan komplek protin OMP 50 kDa dan dua subunitnya secara berasingan. Tiga aptamer ini ST01, ST03 ST04 mempunyai potensi untuk digunakan sebagai ligan pelekatan yang kuat untuk menangkap dan seterusnya mengesan *S.* Typhi dalam usaha untuk mengesan agen penyebab demam kepialu.

ABSTRACT

Typhoid fever is a food borne illness caused by the bacteria, *Salmonella enterica* serovar Typhi (*S.* Typhi). Typhoid fever remains a public health problem in many of the underdeveloped and developing countries including Malaysia. It has been estimated around 22 million cases and 216,000 related deaths occurred worldwide annually. The 50 kDa outer membrane protein (OMP) of *S.* Typhi has been considered as a possible candidate that plays role during the infection of typhoid. Therefore 50kDa OMP of *S.* Typhi was used as target molecule in detecting of *S.* Typhi. Currently, monoclonal and polyclonal antibodies were widely used in diagnostic kit development for antigen detection from the organisms that caused the disease. However, this diagnostic kits are not suitable to be used in the fields because of the antibody are not stable at high temperature and need to maintain the temperature chain. Thus, as a new approach, RNA aptamers were selected for *S.* Typhi antigen detection in the diagnosis of typhoid fever. Compare to antibody, a few properties of aptamer such as more stable at high temperature make this aptamer more suitable to be used in diagnostic technology, where it can substitute the antibody in the future. Aptamer is a synthetic, single stranded nucleic acid that folds up into a unique two or three-dimensional structure, allowing them to bind specifically to the target molecules. The aptamer was generated using Systematic Evolution of Ligands by Exponential Enrichment (SELEX) technique. SELEX is a process for screening specific ligands from large libraries of oligonucleotides by an iterative process of selection and amplification. SELEX demonstrated six potential aptamer candidates that dominated the final pool of oligonucleotides. The alignment analysis showed the consensus sequences GUU, GUUU, and GUUUU occurred in most of the aptamer candidates which could be the potential binding site of the aptamers towards target

molecule. The binding activity between the aptamers and the target protein was evaluated by electrophoretic mobility shift assays (EMSA). From the six aptamer candidates isolated, only aptamer ST01, ST03 and ST04 showed binding affinity towards 50 kDa OMP of the *S.* Typhi. Since 50 kDa OMP of *S.* Typhi is in a complex formation, comprising 3 subunits, subunit 1, marker flagellin protein, subunit 2, glycerol kinase and subunit 3 TolC protein, the three selected aptamers were further tested against the subunits individually. The results showed that aptamers ST01 and ST03 bound to both subunits 2 and 3, while ST04 could not bind to any subunits. As a conclusion, aptamers ST01 and ST03 can bind to the 50 kDa OMP complex and its subunits individually and ST04 would only bind to the 50 kDa OMP complex. These three aptamers, ST01, ST03 and ST04 have the potential to be used as high affinity ligands for the capture and subsequently detection of *S.* Typhi in the diagnosis of typhoid fever.

CHAPTER ONE:

INTRODUCTION

1.1 Introduction to typhoid fever and *Salmonella enterica* serovar Typhi

1.1.1. Introduction of typhoid fever

Typhoid fever is a food borne illness, which is an acute systemic infection disease. The disease is transmitted through the ingestion of the food or drink contaminated by the feces and urine infected with bacteria *Salmonella enteric* serovar Typhi (*S.* Typhi) (Fischer *et al*., 2007, Ismail *et al*., 1991). Thus, the typhoid fever usually occurs in countries or places that lack clean water and basic hygienic practices (Fischer *et al*., 2007). Children especially in endemic areas, travelers and microbiological laboratory technicians are particularly at risk of exposed to the disease (Hamid and Jain, 2008). This febrile illness can also be transmitted in crowded and impoverished populations with inadequate sanitation that led to unsafe water and food for consumption.

Typhoid fever is a global health problem. Variable estimates of typhoid fever have been published in the scientific literature (Crump and Mintz, 2010). The true incidence of typhoid fever is not known. A study published in 2004 estimated 22 million cases and 216,000 related deaths occurring worldwide each year particularly among children and adolescent (Newton and Mintz, 2014, WHO, 2013). It has been shown that the incidence of typhoid was high especially in south central and southeast Asia, also southern Africa with more than 100 cases per 100,000 population per year (Ismail *et. al.,* 2009). However, for the rest of Asia, Africa and Latin America the

incidence was moderate with 10 – 100 cases per 100,000 populations (Figure 1.1) (Crump *et al*., 2004, Crump and Mintz, 2010).

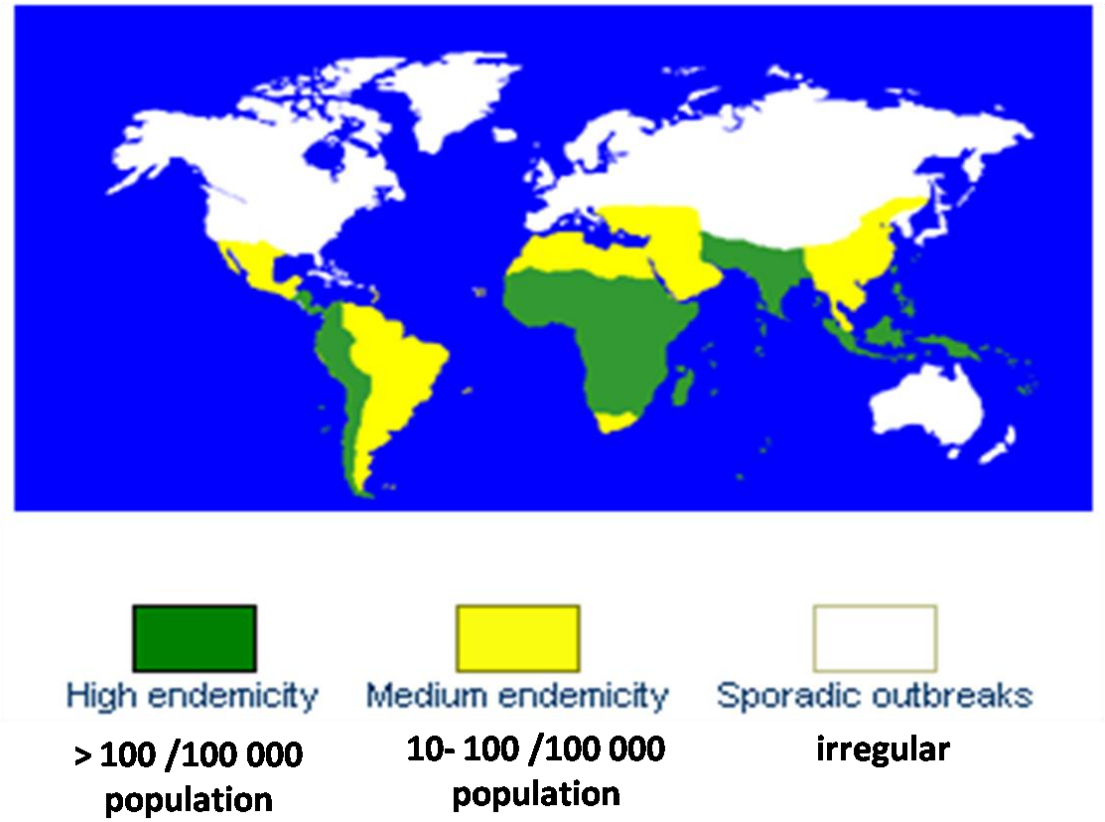

Figure 1.1 : Worldwide distribution of typhoid fever.
Source : Crump *et al.* (2004)

In Malaysia, typhoid fever is endemic and periodically gives rise to outbreak (Anita *et al.*, 2012). According to the reports from Ministry of Health from 2004 – 2012 there were between 200 – 1,072 cases (Table 1.1) and incidence rate was less than 5 per 100,000 population for over the country (Table 1.2) (MOH, 2011, Malik and Malik, 2001). . The last big outbreak was reported in 2005 in state of Kelantan involving 887 cases (Table 1.1) with incidence rate of 58.9 per 100,000 population (Table 1.2), but for overall in Malaysia, the incidence rate on that year was 4.1 per 100,000 population (Table 1.2) which is still not considered as highly endemic (MOH, 2011, Department of Statistic Malaysia, 2014).

Table 1.1 Typhoid incidence case for each state in Malaysia (2004-2013) which showed that Kelantan has a higher incidence compared to other states

State	Year									
	2004	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
Malaysia	**484**	**1,072**	**204**	**325**	**201**	**303**	**210**	**242**	**218**	**218**
Johor	14	12	13	44	8	14	13	7	NA	NA
Kedah	21	12	8	9	5	18	2	8	NA	NA
Kelantan	111	887	76	136	72	85	46	84	31	NA
Melaka	2	1	2	7	1	3	NA	NA	NA	NA
N.Sembilan	4	7	NA	1	2	5	2	1	NA	NA
Pahang	9	12	2	17	7	5	3	31	NA	NA
Penang	5	9	3	16	3	NA	2	1	NA	NA
Perak	24	18	13	21	9	11	5	8	NA	NA
Perlis	3	NA	1	NA	1	NA	2	NA	NA	NA
Selangor	50	26	17	22	22	88	31	38	NA	NA
Terengganu	32	8	9	3	2	6	4	NA	NA	NA
Kuala Lumpur	12	2	2	7	6	7	12	7	NA	NA
Sabah	163	62	41	36	59	54	82	185	NA	NA
Sarawak	34	16	17	6	4	7	6	50	NA	NA

NA = Not available

Source: Department of Statistic Malaysia, (2014)

Table 1.2 Typhoid incidence rate for each state in Malaysia (2004-2013) which showed that Kelantan has a higher incidence compared to other states

State	Year									
	2004	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**
Malaysia	**1.9**	**4.1**	**0.8**	**1.2**	**0.7**	**1.1**	**0.7**	**0.8**	**0.7**	**0.7**
Johor	0.5	0.4	0.4	1.4	0.2	0.4	0.4	0.2	0.2	0.4
Kedah	1.2	0.7	0.4	0.5	0.3	0.9	0.1	0.4	0.2	0.3
Kelantan	7.5	58.9	5.0	8.7	4.5	5.2	3	5.2	1.9	2.5
Melaka	0.3	0.1	0.3	1.0	0.1	0.4	NA	NA	NA	0
N. Sembilan	0.4	0.7	NA	0.1	0.2	0.5	0.2	0.1	0.2	0.6
Pahang	0.6	0.8	0.1	1.2	0.5	0.3	0.2	2.0	0.4	03
Perak	1.1	0.8	0.6	0.9	0.4	0.5	0.2	0.3	0.7	0.2
Perlis	1.4	NA	0.4	NA	0.4	NA	0.9	NA	NA	1.2
Penang	0.4	0.6	0.2	1.1	0.2	NA	0.1	0.0	0.1	0.6
Sabah	5.5	2.1	1.3	1.1	1.8	1.6	2.5	1.5	1.1	1.6
Sarawak	1.5	0.7	0.7	0.3	0.2	0.3	0.2	0.3	1.4	0.6
Selangor	1.1	0.6	0.4	0.4	0.4	1.7	0.6	0.7	0.6	0.8
Terengganu	3.2	0.8	0.9	0.3	0.2	0.5	0.4	NA	0.3	0.2
Kuala Lumpur	0.8	0.1	0.1	0.4	0.4	0.4	0.7	0.4	2.0	0.4

NA = not available

Source: Department of Statistic Malaysia (2014)

1.1.2. Introduction to *Salmonella enterica* serovar Typhi

Salmonella is a member of family Enterobacteriaceae. The genus *Salmonella* is divided into two species: *Salmonella enterica* and *Salmonella bongori*. *S. enterica* was then divided into six distinct subspecies according to biochemical differences, which are *Salmonella enterica* subsp. *enterica* (subsp. I), *Salmonella enterica* subsp. *salamae* (subsp. II), *Salmonella enterica* subsp. *arizonae* (subsp. IIIa), *Salmonella enterica* subsp. *diarizonae* (subsp. IIIb), *Salmonella enterica* subsp. *houtenae* (subsp. IV) and *Salmonella enterica* subsp. *indica* (subsp. VI). Meanwhile *S. bongori* was originally identified as subsp. V, but now was recognized as separated species from *S. enterica* (McQuiston *et. al.*, 2008).

In addition, the salmonellae are further subdivided by using a subtyping method based on two surface structures, the somatic O antigen of lipopolysaccharide and the flagellar H antigen. There are 2 463 serotypes (serovars) of *Salmonella*. The majority of the *Salmonella* serotypes belong to *Salmonella enterica* subsp. *enterica* with 59% from all serotypes as inTable 1.3. *Salmonella* Typhi is one of the example of serotypes under *Salmonella enterica* subp. enterica. For named serotypes, the serotype name is not italicized and the first letter is capitalized. *Salmonella* serotype Typhi would be referred as *Salmonella enterica* subsp. *enterica* serotype Typhi. At the first citation of a serotype the genus name is given followed by the word “serotype” (serovar) or the abbreviation “ser.” and then the serotype name, for example,*Salmonella* serotype or ser. Typhi. Subsequently, the name may be written with the genus followed directly by the serotype name for example, *Salmonella* Typhi or *S.* Typhi (Brenner *et. al.*, 2000).

Table 1.3 : Number of serotypes within subspecies of *Salmonella* (Brenner *et. al.*, 2000)

***Salmonella* species and subspecies**	**No. of serotypes within subspecies**
S. enterica subsp. *enterica* (I)	1 454
S. enterica subsp. *salamae* (II)	489
S. enterica subsp. *arizonae* (IIIa)	94
S. enterica subsp. *diarizonae* (IIIb)	324
S. enterica subsp. *houtenae* (IV)	70
S. enterica subsp. *indica* (VI)	12
S. bongori (V)	20
Total	**2 463**

.

S. Typhi is an important pathogen exclusively for humans that causes typhoid or enteric fever (Pan et al., 2005). It is a gram-negative bacteria, rod-shaped, noncapsulated, nonsporulating, facultative anaerobic bacilli, with characteristic flagellars, somatic and outer coated antigens (Hamid and Jain, 2008). It is motile by means of peritrichous flagella (H-d antigen) which is also came across in 80 other bioserotypes of *Salmonella* (Ismail, 2009). It is also serogically positive for lipopolysaccharides antigens O9 and O12 (somatic antigen), protein flagellar (H-d antigen) and capsular polysaccharide (Vi antigen) (Ismail, 2009). The Vi antigen is both virulence factor and protective antigen of *S.* Typhi. The presence of the Vi antigen on the isolated samples can be detect by using antiserum agar method (Nolan *et al.*, 1980)

S. Typhi is distinguished from other salmonellae species by its metabolic characteristics such as lactose or sucrose fermentation, production of hydrogen sulfide (H_2S) which is usually observed as a crescent-shaped wedge of black precipitate forming at the interface of the slant and butt in triple sugar iron media (Ismail, 2009). Its capsular polysaccharide (Vi antigen) is a virulence factor and can act as protective antigen.. In addition, inability of *S.* Typhi to cause to similar disease in other animal species showed that it is a human adapted pathogen (Crump and Mintz, 2010) and ability to establish a chronic infection in the host gall bladder. According to World Health Organization (WHO), (2013) about 3 – 5% of individuals develop chronic carrier phases after a year of an acute infection (Ismail, 2009, Crawford *et al.*, 2008).

1.1.3. Diagnosis of Typhoid fever

The isolation of *S.* Typhi from feces, urine, bone marrow and blood via bacteria culture method remained as the gold standard for the diagnosis of the typhoid. The other methods available include the Widal test (Ismail *et al.*, 1991), dot Enzyme immunoassay (EIA) (Choo *et al.*, 1994) and molecular approaches such as Polymerase Chain Reaction (PCR) (Joshi *et al.*, 2009). However, all of these methods have several limitations that will be discussed later. Until 1991, Ismail *et. al.*, 1991, had shown that the 50 kDa specific outer membrane protein (OMP) of *S.* Typhi is a specific antigenic protein to *S* Typhi and has been used for the development of typhoid detection kit called TyphiDot TH (Ismail *et al.*, 1991).

TyphiDotTH is a dot enzyme immunosorbent assay kit, which has been developed for the detection of specific IgM and IgG antibodies to the 50 kDa protein in typhoid patients. The assay was developed to cater for the need of a rapid, simple and inexpensive test for diagnosis of acute typhoid fever, as compared to the gold standard culture method and the Widal test which is difficult to interpret. The techniques using the outer membrane protein was found to be sensitive and specific and could produce results within 3 hours compared to 3 to 7 days via the conventional culture method. (Choo *et al.*, 1994, Choo *et al.*, 1997).

When developing a diagnostic test, a diagnostic kit must be rapid, specific and sensitive. The traditional culture based assay is still considered as the gold standard method for diagnosis the typhoid fever, simply for the lack of a better alternative method. Unfortunately, the culture based method requires five to seven days to

produce final results and the technique is laborious (Joshi *et al.*, 2009). Therefore, there is an urgent need for a diagnostic method for typhoid that is rapid and accurate. For effective management, a rapid diagnosis results is needed to provide the information for initial clinical decision making (Ismail *et al.*, 1991).

Alternatively, immunological and molecular assays have been developed and may even be more sensitive compared to the culture method. Immunological assay involves the antibody or antigen based detection. However, sometimes, this assay could produce negative results to detect for antibody response in immune-compromised patients. This is due to the light chain and heavy chain of the antibody that could produce cross reactive results (Yoshida *et al.*, 2008). Even molecular technique such as PCR could shorten the time, it is still cannot reduce the time for enrichment steps to increase the concentration of targets. In addition, the residual matrix-associated inhibitors also need to be removed because they oftentimes compromised molecular detection which will have an impact on the sensitivity and specificity of the assays (Joshi *et al.*, 2009). Thus, a rapid as well as sensitive and specific method is still needed for detection of *S.* Typhi.

Nevertheless, the detection of *S.* Typhi in complex sample matrices such as feces, foods and environmental samples are challenging for a number of reasons. First, time consuming culture-based enrichment steps are usually necessary to increase the target count numbers prior to the application of detection methods. Second, although molecular technique, such as PCR can shorten the detection time, this technique has not allowed for the elimination of enrichment step, largely because high levels of target is still necessary. This is due to the small amplification volumes (1-10 µl) in

contrast to much larger sample volumes (1 – 25 mg or more g or ml) (Joshi *et al.*, 2009).

Although new rapid test methods appear frequently, these almost always focused on the detection aspects and neglect the need for pre analytical sample processing prior to detection. The fact remains that detection of pathogens in these complex matrices would be more sensitive if the agent was concentrated and purified from the matrix prior to detection.

1.1.4. The 50 kDa outer membrane protein (OMP) of *S*. Typhi

The OMP of *S.* Typhi has been considered as a possible candidate for conferring protection against typhoid (Hamid and Jain, 2008). Due to its location on the surface of the bacteria, OMPs have been considered as valuable immunogens (Ismail *et al.*, 1991). Over the past years, several *Salmonella* OMPs have been investigated as potential vaccine candidates, virulence factors and diagnostic antigens. The molecular structure and function of OMPs and their respective genes also have been studied (Hamid and Jain, 2008). In the previous study the 50 kDa OMP from *S.* Typhi was revealed as a specific as well as antigenic protein for *S.* Typhi (Ismail *et al.,* 1991). the 50 kDa OMP is a protein by nature and not a Vi (capsular), dH (flagellar) or O9 (somatic) antigen of *S.* Typhi (Ismail *et al.*, 1991). The gene that encodes the 50 kDa OMP of *S.* Typhi was submitted to Genbank as ST050 gene (1476 bp) with accession number BD079162 (Ismail *et al.,* 2009).and was patented under Universiti Sains Malaysia with US pattern number 20, 020, 012, 668 (Ismail *et al.,* 2009).

In addition, the protein was also used in the development of detection kit for typhoid fever, TyphidotTH. However the kit was designed to detect the antibodies produced in the patient during infection (Gopalakrishnan *et al.*, 2002) instead of directly detecting the organism itself. Thus, in this study, we attempted to produce candidates in the form of RNA aptamers that can bind specifically to 50 kDa OMP of *S.* Typhi and to be use in the development of diagnostic assays. In addition, it also can be applied in detection of *S.* Typhi in food or environmental samples.

1.2. Introduction to aptamers

1.2.1. Theory and history of aptamers

Nucleic acids were considered as template or linear carriers of information. Meanwhile, protein molecules carried out most of the cell function, which possess complex three dimensional structures. With the discovery of the catalytic activity of RNAs by Guerrier and Altman, (1984), they found that nucleic acids were possessed a selectable both genotype and phenotype (Guerrier and Altman 1984, Wang 2008).. The formation of specific complexes with protein also has been known since the 1960s. However, it was not shown that the formation of ribonucleoprotein complexes can be used as a directly selectable phenotype (Burke and Berzal-Herranz, 1993).

This discovery supports the "RNA World" theory by Gilbert, (1986), where RNA carried out most of the biochemical functions in the cell (Gilbert, 1986, Klaussman, 2006). Because of the ability of RNA to catalyse all biological reactions similar to enzymes, such RNA are called ribozymes (Pyle, 1993, Wang, 2008). For example the ribozymes with RNA polymerase activity replicated all the functional RNA structure

in the cell. Thus, the probability of finding RNA in a pool of random molecules with any binding or catalytic activity could be like to the polymerase function (Klussmann, 2006).

The aptamers were discovered in early 1990s by Ellington and Szostak (Ellington and Szostak, 1990, Klussmann, 2006). A group of Joyce, Gold and Szostak had carried out a preliminary work of *in vitro* selection of RNA to identify unique RNA structures that displayed functions such as binding specifically to target molecules and also the enzymatic activity of the reaction (Klussmann, 2006). Ellington and Szostak demonstrated that large libraries of RNAs could be screened *in vitro* for RNA ligands that bind to a variety of organic dye and they named these RNAs as aptamers (Ellington and Szostak, 1990). The term aptamer was derived from the Latin word "aptus", which means to fit and to emphasize the "lock and key" relationship between aptamers and the target molecules (Becker and Becker, 2006) and the Greek word "meros" which means particles (Ellington and Szostak, 1990, Wang, 2008). In addition to Ellington and Szostak's discovery, Tuerk and Gold discovered the RNA ligand that bind to Bacteriophage T4 DNA Polymerase (Tuerk and Gold, 1990, Wang, 2008). Tuerk and Gold termed the *in vitro* selection process of the functional RNA as "SELEX" which stand for Systemic Evolution of Ligands by EXponential enrichment (Tuerk and Gold, 1990, Klussmann, 2006).

1.2.2. Aptamer properties

Aptamers are synthetic, single stranded nucleic acids of either DNA or RNA that can fold into different two or three dimensional structures and bind to target molecules with high affinity and specificity (Tuerk and Gold, 1990, Joshi *et. al.*, 2009, Ye *et. al.*, 2012). With the specific and complex three dimensional structure characteristic such as stems, loops, hairpin, pseudoknots, bulges, triplexes, and/or quadruplexes, they can bind to a wide variety of targets (Wang, 2008). These distinct three dimensional structures, allow them to form stable and specific complexes with wide target of complementary shape (Ye *et. al.*, 2012). The binding of the aptamers to the target molecules are based on structure compatibility such as aromatic rings, electrostatic and van der Waals interactions, hydrogen binding or a combination of these effects (Wang, 2008). These interactions form stable complexes, since the aptamers exhibit high affinity binding to their targets with kinetic dissociation constants (K_d) value in the low nanomolar (nM) to picomolar (pM) range (Jin and Bowser, 2011, Polisky, 1998). Hence, aptamers could bind with high affinity (Tombelli et al., 2005) and can be specific against a wide variety of targets (Joshi *et. al.*, 2009) ranging from small molecules such as metal ions, organic dye, amino acids or short peptides (Table 1.4), to large proteins including cell membrane proteins (Table 1.4) (Proske *et. al*., 2005, Tombelli *et. al*., 2005) or complex target such as whole cells, viruses, virus-infected cells or bacteria (Table 1.5).

Table 1.4 : The target molecules that have been used for aptamers selection through SELEX process (Hamula *et. al.*, 2011a)

Pathogen/cell	**Target molecules**	**Reference**
HIV-1	Integrase	Allen *et. al.,* 1996
	Reverse transcriptase	Tuerk and Gold, 1990
	Nucleocapsid protein	Kim *et. al.,* 2002
	Tat protein	Yamamoto *et. al,* 2000
	R5 SV glycoprotein	Khati *et. al.,* 2003
	Drug-resistant reverse transcriptase	Li *et. al.,* 2008
Hepatitis C virus	RdRp	Jones *et. al.,* 2006
	NS3	Kumar *et. al.,* 1997
	NS3 helicase	Hwang *et. al.* 2004
	30'X tail	Fukuda *et. al.,* 2008
	NS3 protease	Fukuda *et. al.,* 2000
	NS5B RNA polymerase	Biroccio *et. al.,* 2002
	IRES (internal ribosome entry site)	Fukuda *et. al.,* 2008
Hepatitis B virus	HBsAg surface antigen	Liu *et. al.,* 2010
Influenza virus	H5N1 HA protein	Cheng *et. al.,* 2008
SARS coronavirus	NTPase, Helicase	Jang *et. al.,* 2008
Apple stem pitting virus	Coat proteins	Lautner *et. al.,* 2010
Foot and mouth disease virus	VP1 protein	Bruno *et. al.,* 2008a
Prion proteins	PrPsc	Proske *et. al.,* 2002
	PrPsc fibrils	Rhie *et. al.,* 2003
	rPrPsc	Weiss *et. al.,* 1997
	rPrPc	Takemura *et. al.,* 2006
	Mammalian prion proteins	Bibby *et. al.,* 20008
Escherichia coli	Release factor 1	Sando *et. al.,* 2007
	Core RNA Polymerase	Kulbachinskiy *et. al.,* 2004
	Lipopolysaccharide O111: B4	Bruno *et. al.,* 2008b

Table 1.4 (Continued)

Pathogen/cell	Target molecules	Reference
Mycobacterium	*M. avium sub. paratuberculosis* MAP0105c gene product	Bannantine *et. al.*, 2007
	M. tuberculosis MPT64 protein	Qin *et. al.*, 2009
	M. tuberculosis polyphosphate kinase 2	Shum *et. al.*, 20011
Francisella tularensis	Protein lysate	Vivekananda and Kiel, 2006
Campylobacter jejuni	Surface extract	Bruno *et. al.*, 2009
	Protein lysate	McMasters and Stratis-Cullum, 2006
Salmonella enteric	serovar Typhi Type IVB pilus	Pan *et. al.*, 2005
	Typhimurium outer membrane proteins	Joshi *et. al.*, 2009
Listeria monocytogenes	Internalin A	Ohk *et. al.*, 2010
Leishmania infantum	H2 Antigen	Ramos *et. al.*,2003
Burkholderia pseudomallei	BipD/BopE/BPSL2748	Gnanam *et. al.*,2008
Ustilago maydis (corn pathogen)	RNA-binding protein Rrm4	Konig *et. al.*, 2007
Venezuelan equine encephalitis virus	Capsid protein	Kang *et. al.*, 2007
Bacterial toxins	Staphylococcal enterotoxin B,	Bruno and Kiel 2002
	Cholera toxin	Bruno and Kiel 2002
	Botulinum neurotoxin	Tok and Fischer, 2008
	Shiga toxin	Fan *et. al.*, 2008

Table 1.5 Whole bacterial cells and viral particles used in selection of aptamers by SELEX (Hamula *et. al.*, 2011a, Hamula *et. al.*, 2006)

Target molecule	**Pathogen**	**Reference**
Viral particles	Rous Sarcoma Virus particles	Pan *et. al.*, 1995
	Bacillus anthracis spores	Fan *e. al.*, 2008
	Live African Trypanosomes	Homann *et. al.*, 2006
	Trypanosoma cruzi	Ulrich *et. al.*, 2002
	Bacillus thurigensis spores	Ikanovic *et. al.*, 2007
	Human Influenza A virus particles	Gopinanth *et. al.*, 2006
	Vaccinia virus particles	Nitsche *et. al*, 2007
Whole bacterial cell	*Mycobacterium tuberculosis*	Chen *et. al.*, 2007
	Lactobacillus acidophilus	Hamula *et. al.*, 2008
	Escherichia coli DH5α	So *et. al.*, 2008
	MS-2 Bacteriophage particles	Fan *et. al.*, 2008
	Mammalian cells expressing Hepatitis C E2 envelope glycoprotein	Chen *et. al*, 2009
	Vaccinia-infected mammalian cells	Tang *et. al.* , 2009
	Staphylococcus aureus	Cao *et. al.*, 2009
	Streptococcus pyogenes	Hamula *et. a.*, 2011b
	Pseudomonas aeruginosa	Wang *et. al.*, 2011
	Campylobacter jejuni	Dwivedi *et. al.*, 2010
Whole cell	Human red blood-cell ghosts	Morris *et. al.*, 1998
	Heat-killed anthrax spores	Bruno and Kiel, 1999
	Trypsanoma cruzi	Ulrich *et. al.*, 2004
	Transformed YPEN-1 endothelial cells	Blank *et. al.*, 2001
	Glioblastoma-derived U251 cells	Daniels *et. al.*, 2003
	Differentiated P12 cells	Wang *et. al.*, 2003
	Receptor tyrosine kinase-expressing P12 cells	Cerchia *et. al.*, 2005
	Human osteoblasts	Guo *et. al.*, 2005
	Francisella tularensis antigens	Vivekananda and Kiel, 2006
	Recombinant growth factor-βType III receptor-expressing Chinese hamster ovary (CHO) cells	Ohuchi *et. al.*, 2006

The exceptional specificity displayed by aptamers can even discriminate between closely related molecules or different conformational states of the same target molecules (Conrad and D.Ellington, 1996). As an example Zueva *et. al.,* 2011 developed aptamers for the specific recognition of highly metastatic cells. Two malignant isogenic hamster cell lines, HET-SR-1 (HM) and HET-SR (LM) were used as samples. These two cells were similar in many aspects, including tumorigenicity and growth properties, but the former shows greater metastatic potential *in vivo* compared to the latter. Aptamer E10 and E37 were successfully identified as specifically bound to the highly metastatic cell line HM and not to LM (Zueva *et. al.,* 2011). The other example is from Cerchia *et. al.,* 2009, where, they generated aptamers specifically for malignant human glioma cell line U87MG. The aptamer can discriminate between target cells with the highly related phenotype but poorly tumorgenic human glioma cell line, T98G within the same tumor (Cerchia *et. al.,* 2009).

1.2.2.1. Aptamer versus antibody

Until recently, monoclonal and polyclonal antibodies were the most commonly used to capture ligands or antigen from pathogen in a wide range of medical, biology and diagnostic applications. However, there are certain limitations associated with antibodies. Several properties of aptamers make them more attractive diagnostic agent and in some cases even surpass antibodies (Table 1.6). Thus it could be the rival for the antibodies for the future.

Table 1.6 : Properties of aptamers and antibodies (Jayasena, 1999, Nimjee *et al.*, 2006)

	Antibodies	**Aptamers**
Selection procedure	requires a biological system, therefore difficult to raise antibodies to toxins (not tolerated by animal) or non-immunogenic targets	chemical process carried out in vitro, therefore can target any molecules
Working condition	Limited to physiologic conditions for optimizing antibodies for diagnostics	Can select for ligands under a variety of conditions for in vitro diagnostics
Time and expense	Screening monoclonal antibodies time consuming and expensive	Iterative rounds against known target limits screening processes
activity	Activity of antibodies vary from batch to batch	Uniform activity regardless of batch synthesis
Target site	Immune system determines target site of protein	Investigator determines target site of protein
modification	Limited modifications of molecule	Wide variety of chemical modifications to molecule for diverse functions
immunogenicity	Significant	No evidence
Cross-reactive compound	No method for isolating cross-reactive compound	Cross-reactive compounds can be isolated utilizing toggle strategy to facilitate preclinical studies
Reverse activity	No rational method to reverse molecules	Aptamer-specific antidote can be developed to reverse the inhibitory activity of the drug
Kinetic parameters	Kinetic parameters of Ab-Ag interactions are difficult to be changed on demand	Kinetic parameters such as on/off rates can be changed on demand
Shelf-life and temperature stability	limited shelf life and sensitive to temperature and may undergo denaturation	unlimited shelf life and denatured aptamers can be regenerated, stable to long term storage and can be transported at ambient temperature
labelling	Labeling of antibodies can cause loss in affinity	Reporter molecules can be attached to aptamers at precise locations not involved in binding
Size	Large, MW : 150 kDa	Small, MW : 8 – 12 kDa

The preparation and selection process of aptamers are carried out *in vitro* via chemical process compared to antibody production which required a biological system such as mouse. These processes need the iterative rounds against the target, which is faster and cheaper compared to the production of antibodies. The production of antibody is laborious and could become very expensive especially when searching for rare antibodies that require screening of large number of colonies (Jayasena, 1999, Wang 2008). The performances of the same antibody could vary from batch to batch requiring immunoassays to be reoptimize with each new batch of antibodies. But for the aptamer, the activity is uniform for every batch of synthesized (Jayasena, 1999). The size of the aptamer is also small ranging from 25 – 100 nucleotides long with molecular weigh 8 to 12 kDa. This small size enables them to travel in circulation faster, penetrate tissue better and distribute in organs more efficient (Wang, 2008).

Aptamer acts similarly as antibodies by folding into two or three dimensional structures and binding to the target molecules. Aptamers bind to their targets with affinities comparable to the monoclonal antibodies with a K_d is in low nM to pM, and even sometimes higher than antibodies (Nimjee *et. al.*, 2006). For example, in a study where aptamer had been used against the IgG substitute the secondary antibody, the K_d showed lower than 15pM to the IgG (Yoshida *et. al.*, 2008). In addition, the binding specificity of aptamers is also very high and they could distinguish between related protein or cells that share common sets of structural domain (Wang, 2008).

Therefore, aptamers could target a wide range of target under a variety of condition for *in vivo* diagnostic. In comparison it would be difficult to raise antibodies to toxins or non-immunogenic targets and limited to physiologic condition (Nimjee *et. al.*, 2006). The pharmacokinetic parameters of binding reaction could be changed according to the demand of the reaction, but it is difficult to modify antibody reactions. On the other hand, the chemical process in the making of aptamers could be modified for diverse functions (Jayasena, 1999). Commonly the aptamers detect the target protein by three dimensional structure compared to antibodies which detect the amino acids itself.

After being exposed to a certain temperature, the aptamer still could return to the original conformation structure but antibodies are well known to be sensitive to temperature and undergo irreversible denaturation (Jayasena, 1999). Therefore the shelf-life of the aptamers were longer compared to antibodies (Nimjee *et. al.*, 2006). When using aptamers as ligands, the immunogenicity reaction that arises is miniscule when compared to antibodies where the immunogenicity reaction might be significant. Cross reactive compounds can be isolated utilizing a strategy to facilitate preclinical studies unlike antibodies where there is no method for isolating cross reactive compound. Aptamer-specific antidote could be developed to reverse the inhibitory activity of the drug but there are no rational method to reverse molecules in antibody reaction (Nimjee *et. al.*, 2006).

Therefore, based on the many advantages described above, aptamers are considered to be an alternative to antibodies in many biological applications. In fact aptamers could be a rival to commercialised monoclonal and polyclonal antibodies due to the fact that they are inexpensive, stable, and can be synthetically manufactured and chemically manipulated with relative ease (Joshi *et. al.*, 2009).

1.3. *In vitro* selection process

As mentioned earlier, SELEX or *in vitro* selection is a technique used to isolate aptamers with high affinity. Since the 1990's, *in vitro* selection was used to identify unique nucleic acid-based structure from random sequence of oligonucleotide libraries that could bind arbitrarily to the interest target molecules (Ellington and Szostak, 1990, Tuerk and Gold, 1990) as well as carry out a specific function (Brody and Gold, 2000, Jarosch *et. al.*, 2006, Keefe and Cload, 2008). It was a process for screening from large libraries of oligonucleotides by an iterative process of selection and amplification (Polisky, 1998). The basic scheme for *in vitro* selection is outlined in Figure 1.2.

The SELEX process begins by the generation of a large library of oligonucleotide sequences. The sequences are constructed with known flanking sequences at the 5' and 3' ends and a random region in the middle. The randomized region usually consisted between 40 -60 nucleotides. The library also contains between 10^{14} – 10^{15} different DNA/RNA species that fold into numerous sequence-dependent structures (Klussman, 2006). Then, the sequences from an input library were subjected to a selected target and went through a selective procedure. The active molecules that have the capability to bind

to the target were separated from the inactive molecules. The sequences that were bound to the target were eluted and subsequently amplified by PCR for DNA or reverse transcription PCR (RT-PCR) for RNA. Once amplified, the product was transcribed into RNA (for RNA aptamer) or the DNA was subjected to the target again. This repeating process usually was cycled between 7 – 15 rounds until functional molecules with higher affinity for the target protein dominated the population (Dausse *et al.*, 2005). The pool is then cloned and sequenced for further analysis (Klussmann, 2006, Proske *et al.*, 2005).

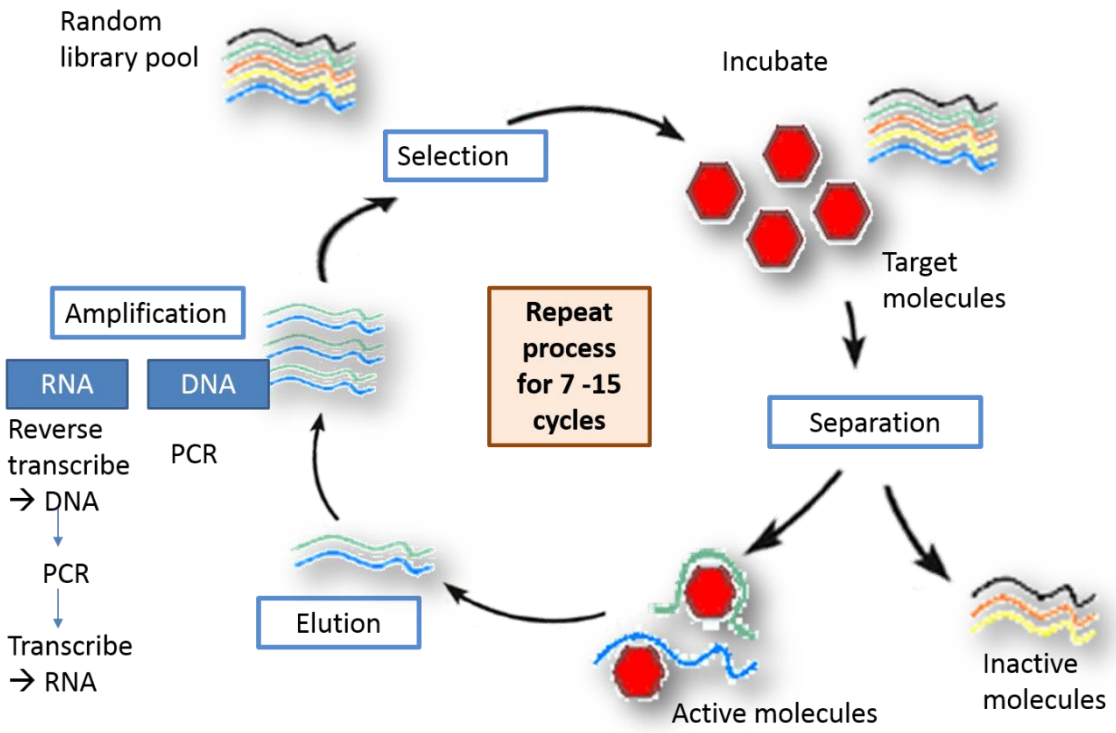

Figure 1.2 : General schemes of SELEX process starting with library pool of oligonucleotides, then the selection process and separation between the bound and unbound molecules. Subsequently, the bound molecules were subjected to amplification process and the whole processes were repeated for several cycles, resulting in the functional molecules, aptamer. (adapted from Klussman, 2006)

1.4. Applications of aptamer

Since its discovery in early 1990s (Ellington and Szostak, 1990, Tuerk and Gold, 1990), aptamer technology has progressed intensively (Proske *et. al.,* 2005). Its attributes, as previously mentioned, make aptamers to be very competitive with protein affinity reagents. Aptamers are successfully used in different areas of biotechnology such as therapeutics, purification processes, target validation, drug discovery, diagnostics and therapy. It also has been applied to various targets including proteins, whole cells and tissues (Table 1.5 and 1.6) (Hamula *et. al.*, 2011a). Their high degree of target specificity and sensitivity are well documented (Joshi *et. al.,* 2009). Aptamers have also been used in analytical methodologies such as chromatography and biosensor (aptasensor). These aptamer-based methods have been mainly employed in the clinical setting for the development of diagnostic assays (WHO, 2013).

Most of the applications of the aptamerS are in the therapeutics fields. A viable aptamer-based drug, macugen or pegaptanib sodium has been clinically developed to help patient manage the devastating disease of age-related macular degeneration (AMD). The use of the drug has recently received the approval by the US Food and Drug Administration (FDA). There also numerous aptamers that are in preclinical development and few of them are scheduled to begin clinical evaluation soon (Klussmann, 2006). For example; a novel sandwich ELISA based on an anti-MPT64 antibody aptamer had been developed by Zhu *et. al.*, (2012). Clinical validation showed that the ELISA method was a reliable test for the serological diagnostic of pulmonary tuberculosis. Thus the method could be used for diagnostic of the tuberculosis (Zhu *et. al.*, 2012).