GENERATION OF RNA APTAMER AGAINST PROGESTERONE RECEPTOR DNA BINDING DOMAIN AND ITS POTENTIAL IN DIAGNOSTIC AND THERAPEUTIC APPLICATIONS IN BREAST CANCER

PRESELA A/P RAVINDERAN

UNIVERSITI SAINS MALAYSIA

2021

GENERATION OF RNA APTAMER AGAINST PROGESTERONE RECEPTOR DNA BINDING DOMAIN AND ITS POTENTIAL IN DIAGNOSTIC AND THERAPEUTIC APPLICATIONS IN BREAST CANCER

by

PRESELA A/P RAVINDERAN

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

January 2021

ACKNOWLEDGEMENT

First and foremost, thank you God for giving me an opportunity to pursue my goal. You were there being my strong support system from the very beginning till now in completing my study. Next appreciation and gratitude to my supervisor, Dr. Citartan Marimuthu for being a tremendous mentor for me. All the guidance, advice and encouragement given by him throughout the research made me to grow as an independent and diligent researcher. I would like to also thank my co-supervisors, Prof. Dr. Tang Thean Hock and Dr. Ch'ng Ewe Seng for all the guidance and consultation given throughout my study. A special thanks to the members of RNA-Bio Group, Histopathological Unit of IPPT and administration staffs of IPPT who provided all the help and guidance throughout the project. To my parents, Mr. Ravinderan and Mrs. Maliga, family members and companion Dr. Prabu, thank you for the unconditional support, patience and love shown towards me while I complete the journey I embarked. My sincere thanks to my fellow friends; Lee Lee, Rogini, Shawn, Alwin Yeoh, Alvin, Nadia and Yik Wei for all the support, stimulating ideas and discussions we had. My sincere gratitude to Public Service Department (JPA) and Graduate Research Assistant Scheme, IPPT for the financial support received throughout the study. Lastly, I would like to thank everyone who supported me and involved directly and indirectly throughout the years in completion of this study.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS AND ABBREVIATIONS

A	Adenine
aa	Amino acid
AF	Activation function
ALISA	Aptamer-Linked Immunosorbent Assay
AMD	Age macular degeneration
AML	Acute myeloid leukaemia
APS	Ammonium persulfate
ASCO	American Society of Clinical Oncology
ATP	Adenosine 5’-triphosphate
Bis	N, N’-methylene bisacrylamide
bp	Base pair (s)
BSA	Bovine serum albumin
C	Cytosine
CAP	College of American Pathologists
cDNA	Complementary DNA
CTP	Cytidine 5’-triphosphate
DAB	3,3'-Diaminobenzidine Tetrahydrochloride
ddH_2O	Double-distilled water
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
E. coli	Escherichia coli
EDTA	Ethylenediaminetetraacetic Acid
ELISA	Enzyme-Linked Immunosorbent Assay
et al.	And others
EtBr	Ethidium bromide
ECL	Enhanced Chemiluminescence
ER	Estrogen Receptor

FDA	USA Food and Drug Administration
FFPE	Formalin Fixed Paraffin Embedded
g	Gram
G	Guanine
GTP	Guanosine 5’-triphosphate
HCl	Hydrochloric acid
HER-2	Human Epidermal Growth Factor 2
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HRP	Horseradish peroxidise
IPTG	Isopropyl-β-D-thiogalactopyranoside
KCl	Potassium chloride
K_d	Dissociation constant
kDa	Kilodalton
KOH	Potassium hydroxide
LB	Luria Bertani medium
LiCl	Lithium Chloride
LINA	Lithium Chloride and Sodium Chloride
LOD	Limit of Detection
M	Molar, [(Mole)/(Litre)]
Mg^{2+}	Magnesium ion
Min	Minute (s)
mL	Milliliter
mM	Millimolar
MWCO	Molecular Weight Cut-Off
Na^+	Sodium ion
NaCl	Sodium chloride
$NaOAc.3H_2O$	Sodium acetate trihydrate
NaOH	Sodium hydroxide
ng	Nanogram
nM	Nanomolar
nt	Nucleotide (s)
N-terminal	Amino-terminal
-OH	Hydroxyl

OTA	Ochratoxin A
PAGE	Polyacrylamide gel electrophoresis
PBST	Phosphate buffered saline with Tween 20
PCR	Polymerase chain reaction
PI3K	Phosphatidylinositol 3-kinase
pmol	Picomole
PR	Progesterone Receptor
PR DBD	Progesterone Receptor DNA binding domain
PRapt-3	Progesterone Receptor aptamer - 3
RNA	Ribonucleic acid
RNase	Ribonuclease
rpm	Rotations per minute
RT	Room temperature
RT-PCR	Reverse transcription-PCR
s	Second (s)
SEER	Surveillance, Epidemiology and End Results
SELEX	Systematic Evolution of Ligands via Exponential Enrichment
ssDNA	Single-stranded DNA
T	Thymine
TAE	Tris–Acetic Acid–EDTA
TBE	Tris-Boric Acid-EDTA
TEMED	N,N,N’,N’-Tetramethylethylenediamine
TMB	3,3',5,5'-tetramethylbenzidine
Tris	Tris-(Hydroxymethyl)-Aminomethane
tRNA	Transfer RNA
U	Units of enzymatic activity
UTP	Uridine 5’-triphosphate
UV	Ultraviolet
V	Volt (s)
VEGF	Vascular Endothelial Growth Factor
v/v	Volume per volume
w/v	Weight per volume
X–gal	5’-Bromo-4’-Chloro-3’-Indolyl-β-Dgalactoside

x g	Relative Centrifugal Force
WHO	World Health Organisation
µg	Microgram
µL	Microliter
µM	Micromolar
γ32P	Gamma Phosphorus
°C	Degrees Celsius
%	Percentage

LIST OF APPENDICES

PENJANAAN APTAMER RNA BAGI DOMAIN PENGIKAT DNA RESEPTOR PROGESTERON DAN POTENSINYA DALAM APLIKASI DIAGNOSTIK DAN TERAPI BAGI BARAH PAYUDARA

ABSTRAK

Aptamer merupakan satu kelas elemen pengecaman molekul baru yang mempunyai spesifisiti dan keafinan yang tinggi pada sasaran. Di dalam kajian ini, satu aptamer RNA terhadap "Progesterone Receptor DNA binding domain (PR DBD)" yang merupakan penanda diagnostik dan mempunyai potensi terapeutik untuk kanser payudara telah dijana. Pada mulanya, gen *PR DBD* telah dihasilkan daripada jumlah RNA sel MCF-7 dan diklonkan ke vektor pET15b. Pengekspresan dan penulenan protein telah dilaksanakan menggunakan strain *E. coli* Rosetta 2(DE3)pLysS. Identiti protein telah disahkan melalui peblotan Western menggunakan antibodi poliklonal anti-PR dan penjujukan protein menggunakan spektrometri jisim MALDI-TOF/TOF. Protein yang dihasilkan telah tertakluk kepada SELEX bertujuan untuk mengasingkan aptamer RNA. Sejumlah lapan kitaran SELEX telah dilaksanakan dan himpunan asid nukleik daripada kitaran 8 menunjukkan pemerkayaan pada PR DBD. Penjujukan himpunan asid nukleik kitaran 8 telah dilaksanakan melalui penjujukan terus dan penjujukan 'crush and soak-based elution'. Kedua-dua penjujukan memdedahkan kehadiran tiga kelas urutan yang berbeza, di mana satu kelas bernama PRapt-3 menunjukkan pengikatan yang terkuat terhadap PR DBD. PRapt-3 mempunyai pemalar permisahan yang dianggarkan pada 380 nM ± 35 nM. PRapt-3 dengan jayanya telah digunakan untuk menghasikan pengujian diagnostik berasaskan aptamer seperti ALISA, blot titik

berasaskan aptamer, peblotan Western berasaskan aptamer, ‘aptacytostaining’ dan ‘aptahistostaining’. PRapt-3 telah mengesan PR DBD dengan had pengesanan 69.44 nM di dalam ALISA langsung. Di dalam asai blot titik berasaskan aptamer, PRapt-3 telah mengesan sehingga 6.25 pmol protein PR DBD. PRapt-3 juga dengan jayanya mengesan protein PR DBD yang dihasilkan dan PR DBD di dalam sel MCF-7 dan HeLa melalui asai peblotan Western berasaskan aptamer. PRapt-3 memnujukkan pewarnaan nuklear di ‘aptacytostaining’ berserta penembusan yang baik di ‘aptahistostaining’ dalam blok sel dan tisu kanser payudara yang dibenamkan parafin dan difiksatif formalin. Selain pengujian diagnostik, fungsi PRapt-3 aptamer RNA telah dikaji di dalam aplikasi terapeutik. Pada mulanya, ciri antagonis PRapt-3 telah dikaji di dalam analisis kitaran sel. Analisis aliran sitometri menunjukkan PRapt-3 mengurangkan sel MCF-7 di fasa S dan pada masa yang sama meningkatkan jumlah peratus sel di fasa G_0/G_1. PRapt-3 juga berpotensi menjadi agen induksi apoptotik. Di dalam kajian ini, PRapt-3 didapati telah mengurangkan gen anti-apoptotik, *BCL-2* dan meningkatkan gen pro-apoptotik, *BAX*. PRapt-3 mengurangkan ekspresi bagi gen berkaitan percambahan. Tahap ekspresi bagi gen berkaitan percambahan, *PI3K* dan *AKT* telah dikurangkan oleh RNA PRapt-3. Pengurangan gen ini menunjukkan PRapt-3 dengan jayanya menunjukkan kesan antagonis di antara PR dan progesteron. Secara kesuluruhannya, PRapt-3 mempunyai potensi sebagai agen diagnostik dan terapeutik.

GENERATION OF RNA APTAMER AGAINST PROGESTERONE RECEPTOR DNA BINDING DOMAIN AND ITS POTENTIAL IN DIAGNOSTIC AND THERAPEUTIC APPLICATIONS IN BREAST CANCER

ABSTRACT

Aptamers are a new class of molecular recognition element that exhibit high binding affinity and specificity against the target. In this study, an RNA aptamer was generated against Progesterone Receptor DNA binding domain (PR DBD) that acts as a diagnostic biomarker and potential therapeutic target for breast cancer. Firstly, *PR DBD* gene was amplified and isolated from the total RNA of MCF-7 cells and cloned into pET15b plasmid. Protein expression and native purification was performed using *E. coli* Rosetta 2(DE3)pLysS bacterial strain. The identity of the protein was confirmed by western blot using polyclonal anti-PR antibody and by protein sequencing using MALDI-TOF/TOF mass spectrometry. The purified protein was subjected to SELEX in order to isolate RNA aptamer. A total of eight SELEX cycles were executed and the resulting nucleic acid pool from cycle 8 showed enrichment against PR DBD. The cycle 8 nucleic acid pool was sequenced using direct sequencing and crush and soak elution-based sequencing methods. Both sequencing methods revealed the presence of three different classes of sequences, with one class termed, PRapt-3 showed the strongest binding against PR DBD. The dissociation constant of PRapt-3 RNA aptamer was estimated at 380 nM ± 35 nM. PRapt-3 was successfully used to develop aptamer-based diagnostic assays such as ALISA, aptamer-based dot blot, aptamer-based western blot, aptacytostaining and

aptahistostaining. PRapt-3 detected PR DBD in direct ALISA with a LOD of 69.44 nM. In the aptamer-based dot blot assay, PRapt-3 detected up to 6.25 pmol of PR DBD. PRapt-3 also successfully detected recombinant PR DBD and endogenous PR DBD in MCF-7 and HeLa cells on aptamer-based western blot assay. PRapt-3 demonstrated nuclear staining in aptacytostaining and with better penetration in aptahistostaining using the formalin fixed paraffin embedded breast cancer cell and tissue blocks. Apart from diagnostic application, the functionality of the PRapt-3 RNA aptamer was also investigated in the therapeutic applications. First, the antagonistic property of the PRapt-3 was investigated on the cell cycle analysis. Flow cytometry analysis showed that PRapt-3 RNA reduced the number of cells in S phase of MCF-7 cell cycle while reciprocally increased the percentage of cells in G_0/G_1. PRapt-3 can also act as a promising apoptotic inducing agent. In this study, it was found that PRapt-3 downregulated the anti-apoptotic gene, *BCL-2* and upregulated the pro-apoptotic gene, *BAX*. PRapt-3 mitigated expression of the proliferation-related genes. The expression level of proliferation-related genes, *PI3K* and *AKT* were reduced by PRapt-3 RNA. The downregulation of these genes showed that PRapt-3 was successfully employed to antagonize the interaction between PR and progesterone. In this entirety, PRapt-3 is a promising diagnostic and therapeutic agent.

CHAPTER 1

INTRODUCTION

1.1 Introduction to Aptamers

Nucleic acids were recognised to solely store genetic materials for a long time but were later discovered to own enzymatic catalysis properties in early 1980s (Lakhin et al., 2013; Burnett et al., 2012; Gilbert, 1986). In the late 90s, the ability of nucleic acids to bind to various targets was achieved through a new class of molecular recognition element termed ‘aptamer’ (Song et al., 2012).

The word aptamer is derived from the Latin word ‘aptus’ which means to fit and Greek word ‘meros’, which means particle (Ellington et al., 1990). Aptamers are single-stranded nucleic acids of either DNA or RNA that has high binding affinity and specificity towards their targets which ranges from ions, small organic molecules, proteins, viruses, bacteria and even live cells (Keefe et al., 2010; Navani et al., 2009; Hermann et al., 2000). Aptamers can fold into various secondary structures such as stem, loop, pseudoknot and G-quadruplex (Mayer, 2009). The coalescence of the variety of the secondary structures of the aptamer eventually result in a unique three-dimensional structure that is capable of recognising and binding to the respective target. The interaction of the aptamer-target complex is established via non-covalent bonds including hydrogen bonds, van der Waals forces, base stacking as well as hydrophobic and electrostatic interactions (Gelinas et al., 2016; Mayer, 2009; Gold, 1995). The high complexity of the aptamer-target interactions causes the binding

affinity of the aptamer can go as low as nanomolar to picomolar range (Hanif et al., 2019; Jenison et al., 1994).

Furthermore, aptamers can discriminate between the closely related molecules. These including theophylline and caffeine which are small molecules that differs only by a methyl group. The presence of hydrogen atom instead of a methyl group in theophylline increases the binding affinity of the aptamer by 10000 times. As the aptamer interacts through hydrogen bond and base stacking to theophylline (interaction between the cytosine and purine-like theophylline) as compared to caffeine, the presence of methyl group in the latter hinders the binding of the aptamer and results in lower binding affinity (Hermann et al., 2000; Jenison et al., 1994).

1.2 Systematic Evolution of Ligands by Exponential Enrichment (SELEX)

Aptamers are generated through a process called Systematic Evolution of Ligands by Exponential Enrichment (SELEX). SELEX is the gold-standard technique in isolating DNA or RNA aptamers with high binding affinity since 30 years ago (Ellington et al., 1990; Tuerk et al., 1990). The *in vitro* selection begins with an initial oligonucleotide pool or also known as combinatorial library comprising of 10^{12} - 10^{15} different sequences. Each of the oligonucleotide sequences is designed to have a randomised region of 25- to 60-nucleotides in which each of the nucleotide position is occupiable by one of the four bases. The randomised region is flanked by a constant region at the 5’ and 3’ -ends commonly used for amplification and transcription purposes (Nakamura et al., 2012; Hermann et al., 2000).

In general, SELEX consists of 3 main steps (Figure 1.1); (i) incubation of the randomized oligonucleotide library with the target (ii) separation to remove the unbound sequences from the target-bound sequences (iii) amplification of the bound

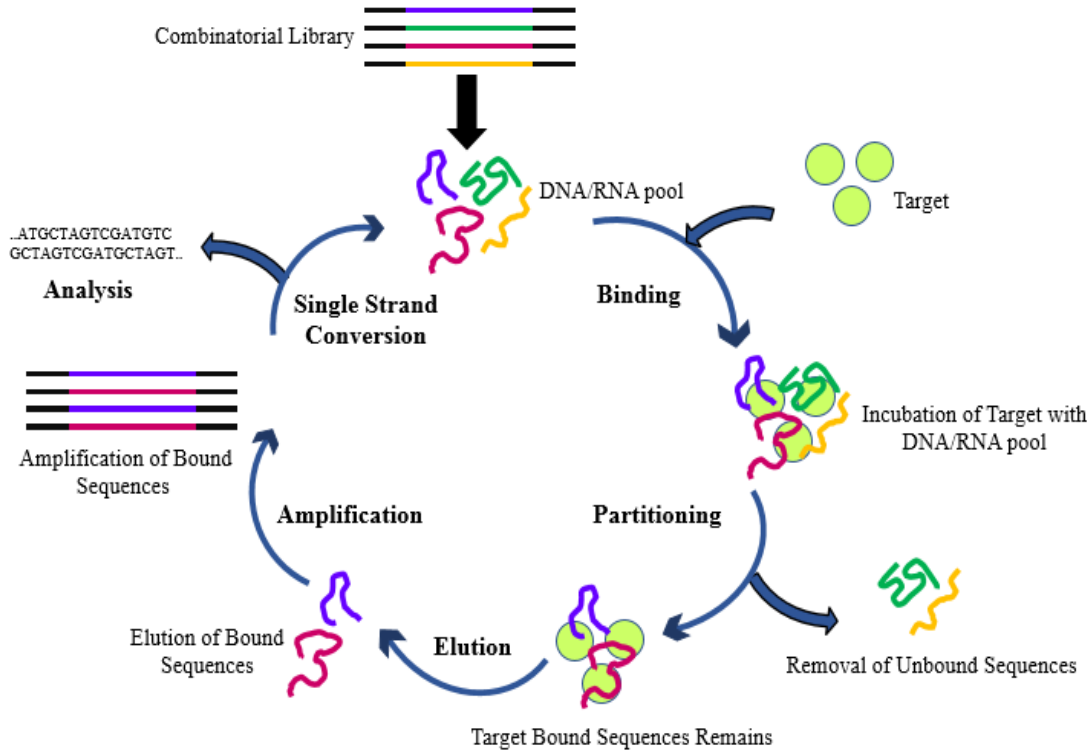


Figure 1.1 The conventional SELEX cycle

The initial single stranded DNA or RNA pool is incubated with the target to allow the binding to happen. The target bound sequences are removed from the unbound sequences through partitioning. The bound sequences are eluted from the target and amplified by PCR for DNA aptamer and RT-PCR for RNA aptamer generation. The pool is then converted back into ssDNA by double stranded DNA strand separation or to RNA by *in vitro* transcription for the next round of SELEX cycle. This process is repeated for at least 8 to 15 cycles before sequence analysis to identify the putative aptamers.

sequences. The SELEX cycles are repeated for at least 5 - 15 rounds with a selective pressure at each cycle with varying ratio of library to target, salt concentration and addition of competitors such as yeast tRNA and salmon sperm DNA in order to isolate an aptamer with high binding affinity amongst the weak binders (Zhou et al., 2017; Song et al., 2012). The stringency of the selection can also be increased with negative selection, a method to eliminate sequences that bind to non-specific target or separation medium resulting in aptamers with high specificity (Xiang et al., 2017; Thiel et al., 2009).

One of the most performed SELEX is the SELEX with purified protein. The easily tuneable selection conditions that eliminates the need to use highly advanced technology to carry out the selection process remains as one of the reason to use purified protein as a common target in SELEX (Cerchia et al., 2010; Pestourie et al., 2006). The advancement in the SELEX technology allows the researchers to develop various types of SELEX considering the final use of the aptamer. For instance, cell-SELEX utilises the whole cell as a target to generate an aptamer. The generated aptamer is not only capable of binding to the target proteins in its native conformation but also to the unknown receptors present on the membrane leading towards a powerful biomarker discovery (Ohuchi, 2012; Cerchia et al., 2010; Guo et al., 2008). Meanwhile in live animal-based SELEX, the nucleic acid library is intravenously injected into the animal model followed by harvesting the tissue and extracting the bound sequences. These methods are capable of isolating tissue-penetrating aptamers which eventually can be used to treat certain diseases or conditions in a living animal model (Keefe et al., 2010; Healy et al., 2004).

1.3 Advantages of Aptamers Over Antibodies

The unique properties of aptamer make it an elegant and attractive molecular recognition element that rival and in some cases, surpass antibodies. Hence, aptamers are widely known as antibody substitutes thus, termed as 'chemical antibodies' (Jayasena, 1999).

Aptamers are usually smaller than the counterpart antibodies in which a 50 nucleotides aptamer molecular weight is only 15 kDa while a typical antibody has a molecular weight of 150 kDa (Zhou et al., 2017). The smaller size of aptamer allows a more efficient entry into biological compartments for example an RNA aptamer generated by *in vivo* SELEX that could penetrate the blood-brain barrier (Cheng et al., 2013) as well as exhibiting better penetration into tissue samples that allows for enhanced staining during immunohistostaining applications (Zeng et al., 2010).

Aptamers can be easily synthesized *in vitro* via chemical reactions unlike antibodies which require a biological system for production. Hence, a great quantity of aptamers can be produced at a lower cost. Since, aptamers are produced synthetically there is no need for an *in vivo* immunization. As such, production of aptamers eliminates any batch to batch variations (Zhou et al., 2017). Property of an aptamer as a nucleic acid and its ability to be produced chemically eases the labelling or modifications processes. Aptamers can be conjugated to any signal molecules like biotin, fluorophores or quenchers for biosensor detection or modified to increase their stability and resistance against nucleases (Nimjee et al., 2017; Mascini, 2008).

The high thermal stability of aptamers is another added advantage to be utilised in a wide range of assay conditions. Aptamers being nucleic acids has a higher thermal stability and maintains their structures over repeated cycles of denaturation or

renaturation process. Aptamers are capable of folding back into their native conformation and bind to their target after the re-annealing process (Parashar, 2016). Besides, aptamers also exhibit low to no immunogenicity when used upon humans (Ng et al., 2006; Jayasena, 1999).

1.4 Diagnostic Potentials of Aptamers

Aptamers make a great contribution to the diagnostics field so much so that some of the aptamers have been successfully commercialized. Aptamers can be isolated to recognise and bind to various targets including toxins (Ruscito et al., 2016). Dr. Gregory Penner of Neoventures Biotechnology Inc. has designed OTA-Sense, an aptamer based diagnostic kit to detect Ochratoxin A (OTA). OTA is a toxin that produced by fungi commonly *Aspergillus* and *Penicillium* species in the agricultural products and the presence of the toxin is highly carcinogenic even at low doses. OTA-Sense kit utilises the OTA-specific aptamer whereby when the OTA is being detected, the intensity of the aptamer-conjugated fluorophores increases greatly. Alike OTA-Sense, AflaSense, which is an aptamer-based diagnostic kit that recognizes aflatoxin is also commercialized. These system employs biotin-streptavidin interactions that is used for the detection of aflatoxins in corn and peanuts (Soh et al., 2015; Penner, 2012; Clarke et al., 1993).

Pronucleotein, CibusDx (USA) selected a wide range of aptamers against foodborne and waterborne pathogens that undergo structural changes upon binding to the target. The electrochemical sensor company employed these pathogen binding specific aptamers to develop a rapid and highly sensitive biosensor platform to detect the pathogens at a shorter timer (Bruno et al., 2015).

Apart from toxin and pathogen detection, aptamers have also been developed for the diagnosis of human disease. OLIGOBIND®Thrombin activity assay, a product developed by Sekisui Diagnostics (Germany) measures the true *in vivo* level of active thrombin in plasma in order to identify patients who are susceptible to bleeding or thromboembolic conditions. The plasma samples containing thrombin is added into pre-coated thrombin aptamer microtitre plate, followed by washing and addition of fluorogenic substrate that results in signal emission at 460 nm. The aptamer-based enzyme-capture fluorescent assay is highly specific as it can detect thrombin 100-fold higher than the detection of prothrombin (Königsbrügge et al., 2017; Muller et al., 2011; Merlini et al., 1995).

Aptamers are also vastly used to isolate positive cells in flow cytometry applications. The Aptamer Science Inc. (Korea) (AptSciSci Inc) commercialized their product AptoCyto (http://www.aptsci.com/product/product_1.html.), which detects positive cells using a wide range of aptamers that were specifically selected against various biomarkers such as CD-31, EGFR, HGFR, HER-2, ICAM-2 and VEGFR-2. The aptamers used in these diagnostic kits are dual-labelled in which one end of the aptamer is conjugated with biotin to be captured by the streptavidin-coated magnetic beads and the other end is FITC-labelled to be measured during flow cytometry applications. These aptamers specifically recognise and bind to the biomarker-positive cells despite the presence of negative cells in a cell population (Kaur et al., 2018).

Aptamers are used as a probe to detect various biomarkers in the pathological assessment of tissues. A DNA aptamer successfully detected estrogen receptor-α in MCF-7 breast cancer cell line and human breast tissue sections using aptacytochemistry and aptahistochemistry, respectively (Ahirwar et al., 2016). Meanwhile, Zeng et al. (2010) proved that CD30 RNA aptamer requires a lower

temperature and shorter antigen retrieval time at 37 °C and 20 minutes compared to CD30 antibody. The antibody detects CD30 antigen at 96 °C and 90 minutes in formalin-fixed and paraffin-embedded lymphoma tissues (Zeng et al., 2010).

1.5 Therapeutic Potentials of Aptamers

The first therapeutic aptamer, Macugen was approved by FDA in December 2004 to treat neovascular age-related macular degeneration (AMD) (Gragoudas et al., 2004). The aptamer specifically binds to the 165 isoform of vascular endothelial growth factor ($VEGF_{165}$) that plays a vital role in angiogenesis which could lead to vision loss (Ruckman et al., 1998). The pegylated RNA aptamer is truncated to 27 nucleotides to reduce the cost of synthesis and chemically modified with 2'-fluoropyrimidines and 2'-O-methyl purines to prevent nucleases degradation. It is given intravitreally once every six weeks at 0.3 mg dose per eye (Chakravarthy et al., 2006; Ferrara, 2004).

Aptamers are also utilised for the treatment of human cancer. AS1411, the first aptamer to enter clinical trial for the treatment of acute myeloid leukaemia (AML), solid tumours and metastatic renal cell carcinoma is a 26 nucleotides DNA aptamer that has a high binding affinity towards nucleolin, a highly expressed protein in cancer cells that is involved in tumorigenesis. The guanine-rich DNA aptamer forms a stable G-quadruplex structure making it highly resistance against nucleases (Bates et al., 2009; Laber et al., 2006; Bates et al., 1999). It was proposed that the binding of the aptamer to the surface nucleolin protein outcompetes the binding of bcl-2 mRNA. As a result, bcl-2 mRNA is degraded and subsequently induces cell apoptosis. The aptamer is currently being tested in phase II clinical trials against renal cell carcinoma which showed a promising result (Platella et al., 2017; Choi et al., 2010; Soundararajan et al., 2008).

The potentiality of the aptamers is also investigated in targeted drug delivery where an aptamer can be conjugated with a drug to form aptamer-drug conjugate (ApDC) to ensure the drug is being delivered to the exact site mainly cancer cells while sparring the normal cells (Marimuthu et al., 2019). Yoon et al. (2017) designed an ApDC by linking the P19 RNA aptamer with anti-mitotic drugs named monomethyl auristatin E (MMAE) and maytansinoid DM1 in order to treat pancreatic ductal adenocarcinoma. The ApDC exhibited highly specificity by specifically internalizing the PANC-1 pancreatic cancer cells. The *in vitro* study further showed G2/M phase cell cycle arrest suggesting the drug acts as a strong antimitotic agent by halting the formation of microtubules and simultaneously inhibiting the cell proliferation in a dose-dependent manner (Yoon et al., 2017).

1.6 Human Progesterone Receptor (hPR)

Progesterone Receptor (*PR*) gene, a member of the nuclear receptor superfamily is located on chromosome 11q22.1. The complete amino acid sequence of human PR (hPR) was identified in 1987 in order to study the clinical and pharmacological functions of PR after the discovery of rabbit and chicken PR cDNA in 1986 (Misrahi et al., 1987; Conneely et al., 1986; Loosfelt et al., 1986). Human *PR* gene consists of eight coding exons with an open reading frame of 2799 bp that encodes for 933 amino acids yielding a 116 kDa full-length protein (Misrahi et al., 1993).

PR consists of three main domains; the N-terminal domain (NTD), the highly conserved and centrally located DNA binding domain (DBD) and the C-terminal ligand binding domain (LBD) (Figure 1.2). Human PR is expressed as two main isoforms from a single gene under two different promoters; PR B, the full-length of PR (116 kDa) and PR A, lacking the first 164 amino acid at the N-terminus producing

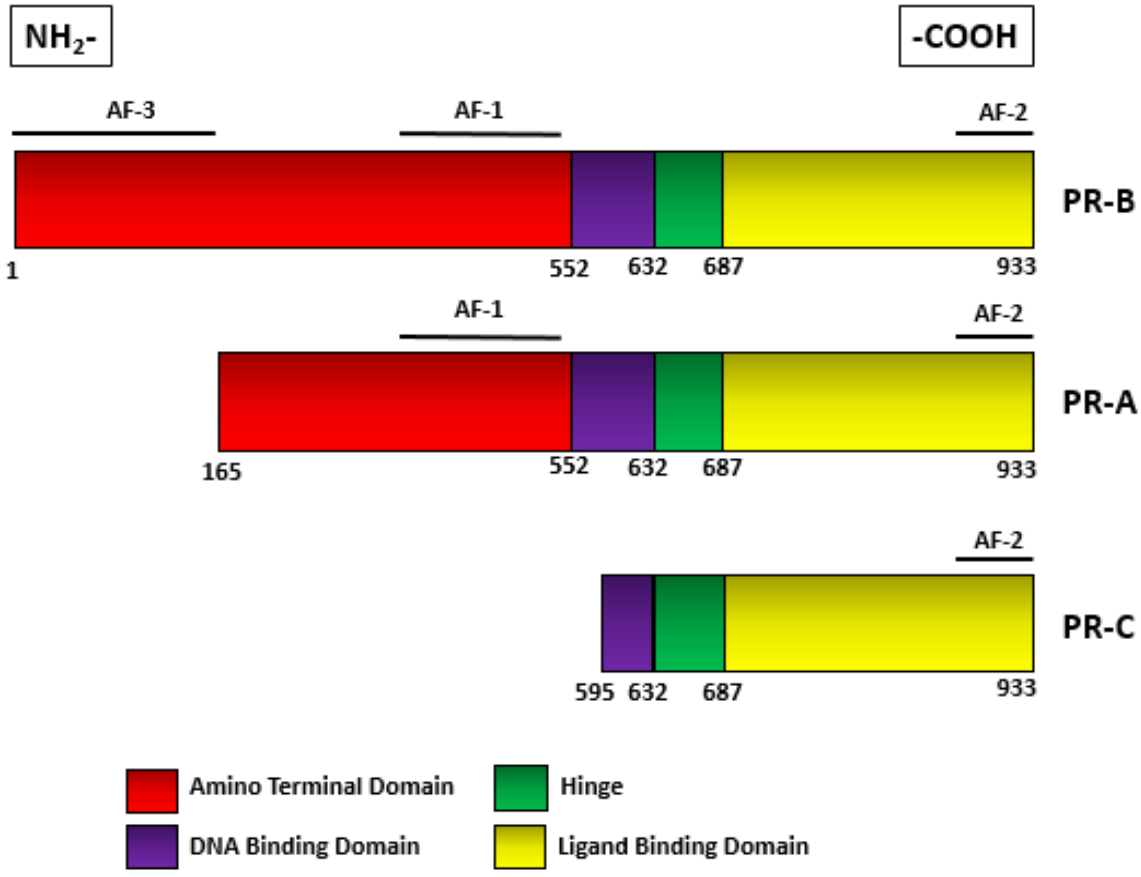


Figure 1.2 Progesterone Receptor (PR) protein structure

PR B is the full-length PR isoform. PR A lacks the first 164 amino acids at the amino terminal domain while PR C is the shortest isoform of PR. A DNA binding domain is present from amino acid residues 552 to 632 while ligand binding domain is located at amino acid residues 687 to 933. There are three activation function (AF) present in PR in which AF-1, AF-2 and AF-3 located at amino acid residue 904 to 919, 456 to 546 and 1 to 165 respectively.

a 94 kDa truncated protein (Hill et al., 2012; Li et al., 2003). There is also another truncated isoform, PR C that is 60 kDa in size (Cork et al., 2008; Wei et al., 1996). Both the isoforms, PR A and PR B have an identical DBD (552 - 632 aa) and LBD (687 - 933 aa) which is parted by a hinge while PR C only has LBD at the C-terminus. The PR DBD recognises a unique consensus sequence on its target genes generally referred to as the progesterone response element (PRE) (Li et al., 2003; Cheung et al., 2000). The PRE comprises an inverted repeat hexanucleotide sequence 5'-AGAACA-3' separated by a 3 nucleotide (3N) spacer which is distinctively specific for PR (Hill et al., 2012; Umesono et al., 1989).

PR is found to be mainly localised in the nucleus (Scarpin et al., 2009; Arnett-Mansfield et al., 2007; Arnett-Mansfield et al., 2004). Both the isoforms are co-expressed at equimolar levels in normal mammary cells but showed dysregulated levels of expression in cancer cells (Mote et al., 2007; Sartorius et al., 2003). PR demonstrates distinct transactivation properties depending on the cell type and target gene. For most gene targets, PR B was found to be a stronger transcriptional activator compared to PR A (Vegeto et al., 1993). This was further shown in an *in vivo* study by Fernandez-Valdivia et al. (2005) that revealed mice lacking PR A have normal mammary gland development while mice lacking PR B showed a reduced mammary gland development (Fernandez-Valdivia et al., 2005). Meanwhile, the expression and role of PR C remains unclear but, it was proposed to interact with PR B resulting in increased transcriptional activity as well as involvement in the induction of labour (Condon et al., 2006; Wei et al., 1997; Wei et al., 1996).

The *PR* gene was found to be highly expressed in the female genitalia tract, breasts and brain which suggests a role leading towards the development, differentiation and maintenance of female reproductive tissues (Li et al., 2003). PR is

an inducible transcription factor which carries out the transcription activity through genomic action via interaction with the genome. The genomic action of PR begins upon the binding of progesterone on the LBD of PR, which confers conformational changes and subsequent dimerization of the receptor. This is followed by the direct binding of the DBD of PR to the PRE on the promoter of the target genes, eventually leading towards the transcription of the genes (Scarpin et al., 2009; Tata, 2002). The activated genes are those involve in cellular activities like transcription – S*TAT5A*, cell growth and apoptosis – *Bcl-X*$_L$, steroid and lipid metabolism – *NPC1* and as well as nucleic acid and protein processing – *ADARB1* (Bertucci et al., 2013; Obr et al., 2013; Cerliani et al., 2011; Xie et al., 2006; Graham et al., 2005; Richer et al., 2002). The binding of PR DBD to the PRE is due to the high binding affinity of PR DBD towards the nucleic acids (Hill et al., 2012; Roemer et al., 2006). On the other hand, PR activity can also be mediated through a non-genomic manner in which the progesterone-liganded PR does not bind to the PRE but instead interacts with cytokine and growth factors through protein-protein interaction thus activating the signalling cascade for example towards the proliferation and differentiation of the mammary cells (Leonhardt et al., 2003).

1.7 PR as an Important Diagnostic Biomarker for The Detection of Breast Cancer

Apart from the physiological function of PR in growth and development of female reproductive system, studies have linked the overexpression of PR to malignancies such as breast cancer, ovarian cancer, cervical cancer and endometrial cancer (Hong et al., 2017; Daniel et al., 2011; Lee et al., 2005; Kleine et al., 1990). PR is most commonly associated with the occurrence of breast cancer, the leading cause of death in women worldwide regardless in developed nor undeveloped countries (Torre et al., 2015; Daniel et al., 2011).

Initially, biochemical ligand-binding assays such as dextran coated charcoal assay (DCCA) was the 'gold standard' in pathological assessment of PR levels but the advent of immunohistochemistry (IHC) became a better technique as it is relatively cheaper, more efficient and allows for direct detection of the antigen on the tumour cells via the monoclonal antibodies raised against the protein (Allred, 1993). This was confirmed by a study conducted by Mohsin et al. (2004) in which PR measured using IHC provides better clinical information compared to ligand binding assay thus allowing a better decision-making for pathologists towards diagnosing malignancies (Mohsin et al., 2004). PR was found to be expressed in various breast cancer cell lines such as T47D, MCF-7, BT-474, ZR-75-1 as well as cervical cancer cell line, HeLa (Liu et al., 2017; Hevir et al., 2011; Subik et al., 2010).

In a pathological setting, breast cancer is typically diagnosed by determination of the PR expression levels in combination with other biomarkers such as estrogen receptor (ER), human epidermal growth factor 2 (HER2) and proliferation marker, Ki-67. Both the hormonal receptors PR and ER can be classified into different molecular

subtypes based on their expression level; ER^+/PR^+, ER^+/PR^-, ER^-/PR^+ and ER^-/PR^- (Colomer et al., 2018; Yip et al., 2014; Patani et al., 2013) (Table 1.1). The data analysis of breast cancer patients from 1990 to 2001 using Surveillance, Epidemiology and End Results (SEER) database from the National Cancer Institute of United States found that ER^+/PR^+ subtypes comprises of 63.5%, ER^+/PR^- were 12.8%, ER^-/PR^+ were 3.1% and ER^-/PR^- were 20.6%. Initially, ER^-/PR^+ entity was thought to be a rare occurrence and was suggested to be a technical artefact (Nadji et al., 2005). However, this entity was later recognised as a distinct biological phenotype after ER^-/PR^+ subtype constituted 3.2% out of 4053 cases reported (Rhodes et al., 2009). On the other hand, ER^-/PR^- phenotype is commonly associated with poor prognosis, later and advanced stages, larger tumour with higher grade and lymph node metastasis (Dunnwald et al., 2007). Among all the four subtypes, ER^+/PR^+ subtype has the lower risk of mortality compared to the other subtypes (Dunnwald et al., 2007).

PR is a powerful prognostic biomarker for breast cancer in which it's presence or absence in combination with that of ER reveals varying degree on the tumour grade, response to hormonal therapy, recurrence and survival rate of the patient (Purdie et al., 2014; Braun et al., 2013; Arpino et al., 2005). One of the largest population-based cohort studies using 1074 patients revealed that PR expression is a strong and independent prognostic biomarker in breast cancer (Purdie et al., 2014). The significance of PR expression in breast cancer was further consolidated by the panel of experts from the American Society of Clinical Oncology (ASCO) and the College of American Pathologists (CAP) who recommended that PR must be measured on every invasive breast cancer and breast cancer recurrences (Hammond et al., 2010).

Table 1.1 Molecular subtypes of breast cancer

Molecular Subtype	ER	PR	HER2	PI (Ki-67)
Luminal A	+ and/or +		-	Low
Luminal B	+ and/or + + and/or +		- +	High Any
HER2 Overexpression	-	-	+	Any
Basal-like	-	-	-	Any

The breast cancer molecular subtypes can be divided into four main groups based on the expression level of the biomarkers. (ER: estrogen receptor, PR: progesterone receptor, HER2: human epidermal growth factor 2, PI: proliferation index)

The expression of PR was also useful in identifying the good prognostic group of patients that could benefit from various therapies including the treatments targeting the growth factor receptor pathways (Purdie et al., 2014). It was demonstrated that ER^+/PR^+ tumour cells benefited more than ER^+/PR^- tumour cells in an adjuvant tamoxifen hormonal therapy suggesting the higher survival rate of ER^+/PR^+ subtype (Cui et al., 2005). Furthermore, it was proposed that the lack of PR expression on ER^+ tumour cells could act as a marker for abnormal growth factor actions that results in the resistance to selective ER modulator (SERM) causing the failure of hormonal therapy (Cui et al., 2005). Hence, ER^+/PR^- metastatic tumours are more aggressive, has a poorer survival rate and poor response to hormonal therapy (Balleine et al., 1999; Gross et al., 1984). Collectively, it is undeniable that PR is a valuable prognostic and predictive biomarker for the pathological diagnosis of breast cancer.

1.8 PR as a Target for Breast Cancer Therapeutics

PR as a ligand-dependent transcription factor was found to trigger the tumorigenesis of breast cancer by regulating a series of genes (Poole et al., 2006; Robertson et al., 1999; Perrault et al., 1996). A microarray analysis of T47D breast cancer cells revealed a subset of PR regulated genes that involved in breast cancer development. The *Bcl-X_L*, an anti-apoptosis gene was upregulated by 3.2-fold suggesting the resistance of the cancer cells to apoptosis. Apart from that, a gene coding for calcium binding protein known as *S100P* was also upregulated by PR (Richer et al., 2002). A study conducted by Guerreiro Da Silva et al. (2000) showed that the S100P protein was overexpressed in the MCF-10F cell line indicating the involvement of the protein in the immortalization of human breast epithelial cells *in vitro* and also in the progression of breast cancer *in vivo* (Guerreiro Da Silva et al., 2000).

Apart from being a transcriptional regulator, PR also provides proliferative or survival advantages to the breast cancer tumours by recruiting and activating the c-Src tyrosine kinase protein (Faivre et al., 2007). The PR N-terminal domain has a unique polyproline motif at amino acid residues 421 to 428 which binds directly to the SH3 domain of c-Src. The binding eventually leads to the rapid activation of signal transduction pathways such as proliferative Ras/Raf/MEK/MAK kinase, JAK-1/-2/Stat3 signalling pathway and also PI3K/AKT survival pathway (Boonyaratanakornkit et al., 2007; Proietti et al., 2005; Saitoh et al., 2005; Ballare et al., 2003). PR regulated genes such as *Wnt-1* is also involved in the proliferation signalling during the malignant transformation of cancer (Faivre et al., 2007; Brisken et al., 2000).

The activity of PR can be inhibited through PR antagonists. The first PR antagonist, mifepristone (RU 486) was discovered in 1981 (Philibert et al., 1981). Mifepristone is one of the widely used PR antagonist in breast cancer treatment that binds at the ligand binding domain (LBD) of PR with a binding affinity of five times more than the of hormone progesterone (Yin et al., 2012; Schreiber et al., 1983). Mifepristone recognises unique amino acid residues on PR that results in a conformational change that is relatively different compared to when the ligand (progesterone) binds to PR. The binding of mifepristone eventually leads to the recruitment of co-repressor proteins like nuclear receptor corepressor (NCoR) and silencing mediator for retinoid and thyroid hormone receptor (SMRT) that results in the downregulation of PR targeted genes (Wagner et al., 1998; Onate et al., 1995; Benhamou et al., 1992). In an *in vitro* study using MCF-7 breast cancer cell line, mifepristone showed a very potent antiproliferative and cytotoxic effects by inducing apoptosis and cell cycle arrest (Fjelldal et al., 2010). Similarly, the effect of mifepristone on the proliferation of breast tissue was studied using premenopausal fertile women. It was found that mifepristone significantly reduced the proliferation of the tissue when proliferation marker, Ki-67 was assessed after three months of treatment (Engman et al., 2008). Hence, PR is a suitable therapeutic target towards combating breast cancer.

1.10 Objectives of the Study

Generation of aptamer against PR DBD which can facilitate the development of aptamer-based diagnostic and therapeutic applications of breast cancer. Therefore, the objectives of this study are as follows:

i) To express and purify recombinant PR DBD

a) Construction of pET-15b/PR DBD plasmid

b) Expression and purification of recombinant PR DBD

c) Protein identity confirmation using western blot assay and MALDI-TOF/TOF mass spectrometry analysis

ii) To generate an RNA aptamer that has high binding affinity and specificity against recombinant PR DBD

a) Execution of dual-partitioning SELEX to isolate RNA aptamer against recombinant PR DBD

b) Cloning and sequence analysis of aptamer pool by 'direct sequencing' and 'crush and soak elution-based' method after 8 cycles of SELEX

c) Binding analysis of the sequence classes using nitrocellulose filter binding assay to search for RNA aptamer candidate

d) Dissociation constant and secondary structure determination of the candidate RNA aptamer against PR DBD

iii) To study the potential functionality of the RNA aptamer in diagnostic applications

a) Development of aptamer-linked immunosorbent assay (ALISA)
b) Development of aptamer-based dot blot and western blot assay
c) Development of aptacytostaining assay with MCF-7 and HeLa cell lines
d) Development of aptahistostaining assay with MCF-7, HeLa cell lines and breast tissues samples from patients

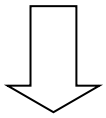
iv) To analyse the therapeutic potential of the RNA aptamer using flow cytometry and real time PCR

a) To investigate the cell cycle analysis of MCF-7 cells using flow cytometry
b) To analyse the PR regulated apoptosis and proliferation-related genes expression level via Real-Time PCR

Flow Chart

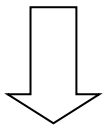
Expression and purification of recombinant PR DBD

- Construction of pET-15b/PR DBD plasmid
- Expression and purification of recombinant PR DBD
- Protein identity confirmation using western blot assay and MALDI-TOF/TOF mass spectrometry analysis



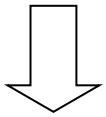
Execution of SELEX cycles with PR DBD

- Execution of dual-partitioning SELEX to isolate RNA aptamer against recombinant PR DBD
- Cloning and sequence analysis of aptamer pool by 'direct sequencing' and 'crush and soak elution-based' method
- Binding analysis of the sequence classes
- Dissociation constant and secondary structure determination of the putative aptamer



Diagnostic applications of PRapt-3 RNA aptamer

- Development of aptamer-linked immunosorbent assay (ALISA), aptamer-based dot blot and western blot assay, aptacytostaining and aptahistostaining with MCF-7 and HeLa cell lines and breast tissues samples from patients



Therapeutic applications of PRapt-3 RNA aptamer

- Cell cycle analysis of MCF-7 cells using flow cytometry
- Gene expression analysis via real time PCR on PR regulated apoptosis and proliferation-related genes

CHAPTER 2

EXPRESSION AND PURIFICATION OF RECOMBINANT PROGESTERONE RECEPTOR DNA BINDING DOMAIN (PR DBD) FOR RNA APTAMER SELECTION

2.1 Introduction

PR, a member of the nuclear receptor family is an important diagnostic biomarker for the detection of breast cancer (Lim et al., 2016; Purdie et al., 2014). The transcriptional regulation of the PR DBD in breast cancer carcinogenesis render it a suitable target for therapeutic intervention. As a promising diagnostic and therapeutic target, PR is an immensely suitable target for the isolation of aptamer. Successful isolation of aptamers against PR DBD requires the protein to be present in high yield and purity. Hence, prior to the aptamer generation, expression and purification of the recombinant PR DBD will be actualized. The identity of the protein was confirmed with Western blot assay using anti-PR antibody and MALDI-TOF/TOF mass spectrometry analysis.

2.2 Materials and Methods

2.2.1 MCF-7 Cell Line Maintenance

The cell culture maintenance was conducted inside the Biosafety Cabinet Class II type A2 (Esco Corporation, Plymouth, USA). MCF-7 cells were cultured in complete DMEM medium (Gibco, California, USA) supplemented with 10 % Fetal Bovine

Serum (Gibco, California, USA) and 1 % Penicillin-Streptomycin (Gibco, California, USA) in a T75 tissue culture flask and incubated in a 37 °C incubator (NuAire, Plymouth, USA) supplemented with 5 % CO_2. The cells were washed with 1X PBS to remove cell debris and degradation products. The culture medium and the 1X PBS buffer were pre-warmed at 37 °C prior the cell culture work.

2.2.2 Total RNA Extraction from MCF-7 Cell Line

MCF-7 cells were cultured to confluency in DMEM supplemented with 10 % Fetal Bovine Serum. The total RNA was extracted by using the TRIzol method (Rio et al., 2010a). Briefly, the cells were trypsinized using TrypLE™ Express (Gibco, California, USA) and pelleted by centrifugation at 1000 x *g* for 5 min (Thermo Scientific, Massachusetts, USA). The supernatant was aspirated, and the pellet was resuspended in 1 ml of TRIzol Reagent (Thermo Fisher Scientific, Massachusetts, USA). After 5 min of incubation at room temperature, 200 µl of chloroform (Merck KGaA, Darmstadt, Germany) was added to the mixture and vortexed briefly. The mixture was further incubated for another 5 min at RT and centrifuged at 12000 x *g* for 15 min at 4 °C. The aqueous phase was transferred into a new microcentrifuge tube and added with 500 µl of Isopropanol (Merck KGaA, Darmstadt, Germany). Incubation was carried out for 5 min at RT. The mixture was then centrifuged at 12000 x *g* for 15 min at 4 °C and the pellet was washed with 75 % Ethanol (Merck KGaA, Darmstadt, Germany). The pellet was air-dried in a vacuum concentrator (Eppendorf, Hamburg, Germany) for 5 mins and dissolved in water at an appropriate volume. The RNA concentration was quantified using a spectrophotometer (Implen Inc, Westlake Village, USA).

2.2.3 Agarose Electrophoresis using 'Bleach Gel'

The quality of the total RNA extracted from the MCF-7 cell line was analysed via agarose gel electrophoresis using a 'bleach gel' based on a protocol adapted from Aranda et al. (2012). The 'bleach gel', which utilises a bleaching agent is a safer and inexpensive way to evaluate the RNA integrity. 'Bleach gel' was prepared by adding 1 g agarose (Promega Corporation, Madison, USA) into 99mL of 1X TAE buffer (40 mM Tris-acetate, 1 mM EDTA) followed by 1 mL % commercial bleach which contain 6 % sodium hypochlorite (Clorox®). The solution was incubated at RT for 5 min with occasional swirling. The suspension was heated for 1 min and allowed to cool down before adding ethidium bromide (Sigma-Aldrich, St. Louis, USA) at a final concentration of 0.5 μg/mL. Two microgram of the total RNA was added with 1 μL of 6X Blue/Orange DNA loading dye (10 mM Tris-HCl pH 7.5, 50 mM EDTA pH 8.0, 0.4 % orange G, 0.03 % bromophenol blue, 0.03 % xylene cyanol FF, 15 % Ficoll® 400) (Promega Corporation, Madison, USA) which contained 0.5 μg/mL of ethidium bromide (Sigma, St. Louis, USA). The RNA was subjected to agarose electrophoresis using a mini gel apparatus (Bio-Rad Laboratories, California, USA) with 1X TAE buffer (40 mM Tris-acetate,1 mM EDTA) at 100 V for 35 min. The gel was visualized using Gel Doc™ XR+ Gel documentation (Bio-Rad Laboratories, California, USA).

2.2.4 Isolation of *PR DBD* Gene via Reverse Transcription-PCR

Two micrograms of the total RNA were resuspended in 10X AMV Reverse Transcriptase Buffer (50 mM Tris-HCl pH 8.3, 5 mM $MgCl_2$, 50 mM KCl, 0.5 mM spermidine and 1 mM DTT) (Promega Corporation, Madison, USA), 0.1 μM of Reverse primer (5'-CCG ctc gag TTA AAA TTT TCG ACC TCC AAG G -3) which contained the *Xho1* site (lowercase and underlined) and 1.6 mM dNTPs was subjected