COMPARISON OF FAB AND SCFV PHAGE DISPLAY EFFICIENCY: APPLICATION OF BEST FORMAT FOR THE DEVELOPMENT OF A NOVEL DNA G-QUADRUPLEX BASED IMMUNOASSAY

LOH QIUTING

Universiti Sains Malaysia

2015

COMPARISON OF FAB AND SCFV PHAGE DISPLAY EFFICIENCY: APPLICATION OF BEST FORMAT FOR THE DEVELOPMENT OF A NOVEL DNA G-QUADRUPLEX BASED IMMUNOASSAY

LOH QIUTING

Thesis submitted in fulfilment of the requirements for the Degree of Master of Science (Molecular Medicine)

SEPTEMBER 2015

ACKNOWLEDGEMENTS

After three years of ups and downs in accomplishing my study, I had discovered many things that we can't learn from a book. Determination and the fighting spirit with a little luck could lead you to success. And, yes. I am lucky enough to meet many kind people whereby I would like to take this opportunity to express my gratitude and appreciation to them.

First foremost,I would like to express my gratitude to my supervisor, Dr. Lim for giving me the opportunity to further my studies, for the knowledge that he shared and for the passion that he always have to educate his students. He had given me many chances to try something new to broaden my knowledge with guidance and also moral support throughout my studies. I would also like to thank Dr. Choong for her support and encouragement during my studies.

My heartfelt appreciation to my fellow colleagues in Antibody Engineering Lab for their help, suggestion and discussion that we had in the lab. I believed that this project was able to accomplish because of our teamwork. I hope that our spirit of togetherness will never fade. Thank you for my friends in Bioinformatics lab for the assistance. Special thanks to Siew Wen who always stand by me and walk through the difficult times with me.

I would also like to show my gratitude to my family members for them whom never stop giving me full support and encouragement. To all my friends who have been giving me moral supports, thank you so much for being there for me. Last but not least, I would like to express my gratitude to the Malaysian Ministry of Education for the MyMaster scholarship.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

°C	Degree Celsius
µg	Microgram
µL	Microlitre
µM	Micromolar
2×YT	2×Yeast and tryptone
4-MU	4-methylumbelliferone
4-MUP	4-methylumbelliferyl phosphate
ABTS	2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
AFP	α-fetoprotein
APS	Ammonium peroxodisulfate
bp	Base pairs
BSA	Bovine Serum Albumin
cDNA	Complementary DNA
CDR	Complementarily determining region
cfu	Colony forming unit
CH	Constant domain of heavy chain
CL	Constant domain of light chain
CN/DAB	4-chloro1-naphthol/3,3' diaminobenzidine
D segment	Diversity segment
DNA	Deoxyribonucleic acid
dNTPs	Deoxynucleoside-triphosphate
dsDNA	Double stranded DNA
DTT	Dithiothreitol
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid

ELISA	Enzyme-linked immunosorbent assay
Fab	Fragment of antigen binding
Fc	Constant fragment
Fc	fragment crystallisable
fM	Femtomolar
g	Gram
GNP	Gold nanoparticles
h	Hour
His-Tag	Histidine tag
HNE	human neutrophil elastase
HRP	Horseradish peroxidase
Ig	Immunoglobulin
IgG	Immunoglobulin gamma
IPTG	Isopropyl-ß-D-thiogalactoside
J segment	Joining segment
kb	Kilo base pairs
kDA	Kilo Dalton
L	Litre
M	Molar
MCS	Multiple cloning site
mg	Milligram
min	Minutes
mL	Millilitre
mM	Millimolar
MMP	Magnetic microparticles
ng	Nanogram
Ni-NTA	Nickel-nitrilotriacetic acid

nM	Nanomolar
NMR	Nuclear magnetic resonance
OD	Optical density
OD_{405nm}	OD at 405 nm wavelength
OD_{600nm}	OD at 600 nm wavelength
PBS	Phosphate buffer saline
PBST	PBS-Tween20
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
pM	Picomolar
PPIase	Peptidylprolyl isomerase
PTM	PBST-Milk
QPA	Quadurplex priming amplification
RCA	Rolling circle amplification
rf	Replicative form
RNA	Ribonucleic acid
rpm	Revolutions per minute
scFv	Single chain variable fragment
SDS	Sodium dodecyl sulfate
ssDNA	Single stranded DNA
Taq	DNA polymerase from *Thermus aquaticus*
TBE	Tris-Boric Acid-EDTA
TEMED	Tetramethylethylenediamine
TES	Trie-EDTA-Sucrose
TMB	3,3′,5,5′-Tetramethylbenzidine
Tween 20	Polyoxyethylenesorbitan monolaurate
U	Enzyme units

UHQ	Ultra-high quality
UV	Ultraviolet
UV-VIS	Ultraviolet-visible
V	Volt
v/v	Volume/volume
VH	Variable domain of heavy chain
VL	Variable domain of light chain
w/v	Weight/volume

LIST OF PUBLICATIONS

Loh, Q., Leong, S. W., Tye, G. J., Choong, Y. S., & Lim, T. S. (2015). Improved Fab presentation on phage surface with the use of molecular chaperone coplasmid system. *Analytical Biochemistry, 477*(0), 56-61.

Loh, Q., Omar, N., Glökler, J., & Lim, T. S. (2014). IQPA: Isothermal nucleic acid amplification-based immunoassay using DNAzyme as the reporter system. *Analytical Biochemistry, 463*(0), 67-69.

Loh, Q., Bahara, N. H. H., Choong, Y. S., & Lim, T. S. (2012). Assembly of highly diverse genes using degenerate oligonucleotides by temperature cascade. *Analytical Biochemistry, 431*(1), 54-56.

Loh, Q., Konthur, Z., Glökler, J., Choong, Y. S., & Lim, T. S. (2013). Terbium(III)-Binding Tag as Reporter System for Antibody-Antigen Interaction. 1st International Conference on Molecular Diagnostics and Biomarker Discovery, Penang Malaysia.

Loh, Q., Glökler, J., & Lim, T. S. (2013). Development of a DNAzyme reporter system using the isothermal quadruplex priming amplification for antibody-antigen interaction.International Congress on Chemical, Biological and Environmental Sciences (ICCBES), Bali Indonesia.

PERBANDINGAN KECEKAPAN PAMERAN FAB DAN SCFV : APLIKASI FORMAT TERBAIK UNTUK PEMBANGUNAN ESEI IMUNO NOVEL BERASASKAN DNA G-KUADRUPLEKS

ABSTRAK

Kebanyakan platform diagnostik konvensional lazimnya bergantung kepada antibodi monoklonal tikus yang tersedia ada dan khusus untuk kegunaan platform diagnostik seperti esei immuno dan aliran sisi. Namun, salah satu daripada isu yang terlibat dalam kaedah diagnostik konvensional adalah kepekaan dan ketersediaan biomolekul. Untuk mereka bentuk platform diagnostik, terdapat dua komponen utama yang perlu dipertimbangkan terutamanya pembangunan bio-pengikat dan sistem pelapor. Bagi pembangunan bio-pengikat, antibodi rekombinan telah dipilih dengan menggunakan teknologi pameran faj. Perbandingan antara serpihan rantai tunggal boleh ubah (scFv) dan serpihan pengikat antigen (Fab) telah dianalisisdari segi pembangunan.identifikasi dan pengeluaran. Cabaran utama dalam penggunaan serpihan Fab adalah kerumitan pembentukan ikatan disulfida yang sering menyebabkan kecekapan paparan serpihan Fab dalam paparan faj lebih rendah jika dibandingkan dengan scFv. Oleh itu, molekul penghantar seperti DsbA dan DsbC yang dikodkan oleh plasmid pembantu, pTUM4 telah digunakan dalam kajian ini dengan tujuanmenambahbaikkan pembentukan ikatan disulfidauntuk membantu kecekapan pameran Fab atas partikel faj. Sehubungan dengan itu, pengubahsuaian dalam proses penyaringan dengan menggunakan sistem plasmid penolong turut dinilai serta digunakan untuk pembangunan antigen pengikat Fab tertentu. Di samping itu, proses penyaringan dengan menggunakan perpustakaan scFv juga

dilaksanakan. Walaupun pameran Fab atas partikel faj telah bertambah baik, format scFv didapati lebih berfaedah dalam proses pembangunan bio-pengikat berbanding dengan proses pembangunan bio-pengikat Fab dan hal ini telah membuat scFv lebih sesuai untuk digunakan dalam pembentukan esei imuno. Selain itu, esei imuno konvensional sering menggunakan konjugasi antibodi-enzim sebagai sistem pelapor yang sering mengalami masalah sensitiviti, kestabilan, ketahanan dan kos esei. Justeru, kaedah pengesanan baru telah diperkenalkan bagi menangani masalah-masalah tersebut. Kebanyakan esei imuno menggunakan kaedah amplifikasi nukleik asid untuk meningkatkan kepekaan esei imuno. Pengikat scFv rekombinan yang terpilih telah berjaya diasimilasikan dalam esei imuno yang menggunakan G-kuadrupleksDNAzimsebagai sistem pelapor. Kaedah "immuno-quadruplex priming amplification" (IQPA) ini menggunakan struktur G-kuadrupleks yang boleh memisahkan dirinya daripada templat semasa amplifikasi sebagai DNAzim untuk menjana bacaan kolorimetrik. Pengoptimuman esei ini telahmenyediakan bukti konsep platform tersebut dalam pembentukan esei imuno yang baru ini. Bagi IQPA secara langsung, had pengesanan esei tersebut didapati hany 0.5 µM, tetapi, bagi IQPA secara kompetitif, esei tersebut boleh mengesan antigen serendah 1 fM. Kesimpulannya, gabungan teknologi antibodi rekombinan dan DNA nanoteknologi dapat menyediakan saluran alternatif untuk pembentukan platform diagnostik baru yang mempunyai sensitiviti yang lebih baik dan mudah digunakan.

COMPARISON OF FAB AND SCFV PHAGE DISPLAY EFFICIENCY: APPLICATION OF BEST FORMAT FOR THE DEVELOPMENT OF A NOVEL DNA G-QUADRUPLEX BASED IMMUNOASSAY

ABSTRACT

The conventional diagnostic platforms are largely dependent on the availability of murine monoclonal antibodies and are designed to be used on platforms such as immunoassays and lateral flow. One of the main issues associated with conventional diagnostic methods is the sensitivity and availability of biomolecules. In designing a diagnostic platform, there are two main components that have to be considered. It is mainly the development of bio-binders and a reporter system. In the case of a bio-binder development, recombinant antibodies were selected using phage display technology. A comparison between the scFv (single-chain fragment variable) and Fab (fragment of antigen binding) was carried out in terms of development, identification and production. The main challenge in using Fab fragments is the complexity of the disulphide bond formation that normally results in lower display efficiencies of Fab fragments during phage display. Therefore, in this study, molecular chaperones such as DsbA and DsbC that were encoded by the helper plasmid, pTUM4 were introduced in order to improve the disulphide bond formation to aid display efficiencies during phage display. A modified version of the panning using the helper plasmid system was evaluated and used for the development of antigen specific Fab binders. On the other hand, panning using scFv library was also carried out. Although the presentation of Fab was improved, the scFv format was

advantageous in the development process making it more ideal for immunoassay development. Conventional immunoassays often uses enzyme-conjugated secondary antibody as the reporter which sometimes suffers from sensitivity, stability, durability and the costs of the assay. Subsequently, new detection methods other than antibody-enzyme conjugates have been developed to counter these issues. Many have incorporated nucleic acid amplification methods into immunoassays with an amplification step for increased sensitivity. The recombinant scFv binder was successfully adapted to a novel immunoassay method which uses G-quadruplexes DNAzymes as the reporter system. The immuno-quadruplex priming amplification (IQPA) method uses self-dissociating G-quadruplex structures to function as DNAzymes to generate a colorimetric readout. Optimization of the assay provides a proof-of-concept of the platform for immunoassay development. In the direct IQPA, the limit of detection was found to be 0.5 μM, nevertheless, in competitive IQPA, it can detect as low as 1 fM of antigen. In conclusion, the fusion of recombinant antibody technology and DNA nanotechnology provides an alternative avenue for the development of novel diagnostic platforms with improved sensitivities and ease of use.

CHAPTER 1

Introduction

1.1 Immunoassay

In the year 1960, immunoassay was first introduced by Berson and Yalow. They utilized radioisotopically labelled analytes as the reporter system by measuring the radioactivity of the bound analyte (Kahn & Roth, 2004; Yalow & Berson, 1960). Like all immunoassays, it is based on the quantification of the bound analyte to the specific antibody in the system. The use of radioisotopic label in immunoassays has given rise to many disadvantages which includes health hazards, stability of the radioisotope and the need of special requirements for licensing, handling and disposal of the radioisotopic substrates (Booth et al., 1982).

With the advent of new technologies, the development of immunoassay has been greatly improved from the utilization of radioisotope labels to nonisotopic alternatives such as chromogenic, fluorogenic and chemiluminogenic substrates (Deshpande, 1996). In developing an immunoassay method, there are four main components that are taken into account: (a) the antibody used for detection, (b) the antigen to be detected, (c) the detection method and (d) the method to eliminate the unbound reactants. The most essential components of these four for a good immunoassay are the antibody and also the detection method (Andreotti et al., 2003; Koivunen & Krogsrud, 2006). Regardless of the immunoassay format, the specificity of the assay can be achieved with higher affinity antibodies. With higher affinity antibodies, antigen can be easily detected and the efficiency of the antibody-antigen complex formation can be increased tremendously (Liddell, 2013; Saerens et al., 2008). Besides, the detection method used in an assay can affect the sensitivity level of the assay. The reactants can be detected visually, electronically, chemically or

physically depending on the nature of the signal produced. There is a wide range of instruments available to measure the signal produced with a high degree of sensitivity (Weeks et al., 2013).

1.2 Bio-binders/Antibody

Invasion of foreign molecules into the human body will trigger the immune system of the body to generate a response to shield it from being harmed by these invaders. The B lymphocytes will start to produce antibodies or also known as immunoglobulins to attack, kill or neutralize these foreign entities (Wood, 2006). Antibodies have the ability to recognize foreign molecules presented on the surface of the pathogens in the body. They are highly sought after because of their high specificity and affinity against target molecules. It is this feature that makes antibodies a useful tool for diagnostics and therapeutics (Filpula, 2007).

An antibody is also known as immunoglobulin, in short, Ig which is produced by the B-lymphocytes. A basic antibody molecule is often depicted as a Y-shaped structure which is made up of four protein subunits (Wood, 2006). As shown in Figure 1.1, the two longer subunits are called the heavy chains while the shorter ones are called the light chains. The heavy chains and light chains are linked together by interchain disulphide bonds. It is also made up of repeating substructures called domains which are flanked by intrachain disulphide bonds. The domains at the C-terminal are the constant domain as they do not differ much from antibody to antibody and some of the antibodies have the same amino acid sequences. Heavy chains have three to four constant domains while light chains have only one constant domain. On the other hand, the domains at the N-terminal ends of heavy chain and light chain are the variable domains because of the variability in amino acid

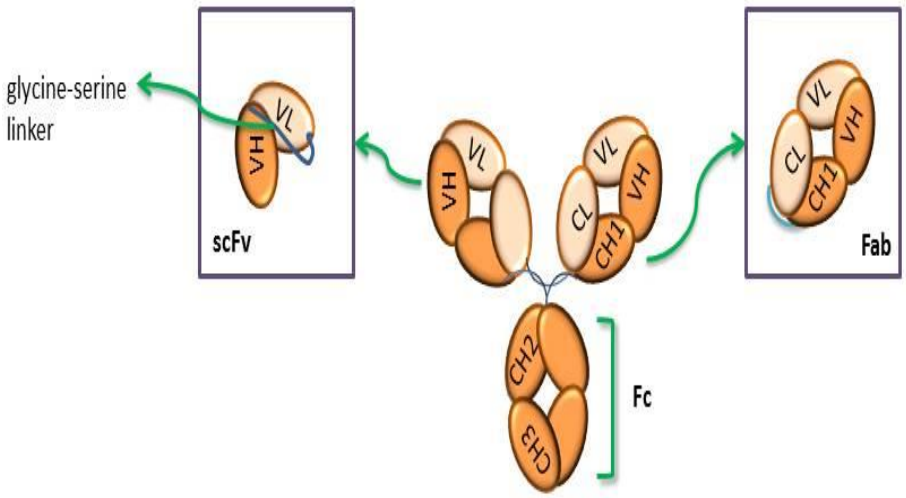


Figure 1.1 : Illustration of Generic IgG Antibody Structure

sequence of these domains among the antibodies. These domains are made up of strands of the polypeptide chain that forms a pair of anti-parallel β-sheets, held together by a disulphide bond, surrounding a hydrophobic core in between these strands. These sheets then form a β-barrel structure known as immunoglobulin fold (Janeway et al., 2001).

The variability is confined to three regions within each heavy chain and light chain. These regions are called the hyper variable regions or more commonly known as complementary determining regions (CDRs) (Liddell, 2013). The parts between the hyper-variable regions that do not differ much between antibodies are called the framework regions. These frameworks contribute to the overall structure of the antibody. It provides a framework on which the hyper-variable regions sit. Thus, when the protein folds, the three hyper-variable regions of each heavy and light chain are brought together at the end of the molecule and form the antigen binding sites or paratope (Wood, 2006). This site binds to a specific part of the antigen which is called the epitope as it is the complement to the shape of the epitope. As this antigen-binding site is located at the variable domains of the heavy and light chains, different antibodies will have different antigen-binding sites consisting of different amino acid sequences against different antigens. Nevertheless, such variability ensures the total repertoire of the antibodies is sufficient to recognize any foreign entities or antigens that attack the body (Mian et al., 1991).

1.3 Antibody Engineering

The advent of recombinant antibody engineering has successfully allowed the introduction of engineered monoclonal antibody of smaller antibody fragments. The well-established antibody fragment is Fab (Fragment of antigen binding). It has a

complete set of light chains (VL and CL) paired with the variable domain and one of the constant domain of the heavy chain (VH and CH_1), linked by the interchain disulphide bond (Janeway et al., 2001). The size of a Fab construct is at approximately 65 kDa which is double the the size of a scFv (Doshi et al., 2014). Thus, this can cause the display efficiency on phage to be less efficient compared to scFv display. Fab is lacking of a fragment consisting two constant domain of the heavy chain that has no antigen-binding activity, namely the Fc (fragment crystallisable). The Fab fragment is very much dependent on the formation of the interchain disulphide bond in between the CL and CH1. As the cytoplasm is in the reduced state that prevents the formation of the disulphide bridge, hence, the heavy chain (VH and CH1) and light chain (VL and CL) are expressed as two independent entities and is translocated to the the oxidizing periplasmic region where they can form disulphide bonds. Therefore, in the periplasmic region, the light chain will pair with the heavy chain and form the disulphide bond in between them, holding them together as one Fab construct (Barbas et al., 1991).The disulphide bond between the light and heavy chains plays an essential role for the stability of the Fab (Lilie et al., 1995; Rothlisberger et al., 2005). For example, for IgG1, the disulphide bond is formed between the C-terminal cysteines of the CH1 domain at position H233 and light chain constant domain at position L214. For other Ig families or even other subtypes, the positions of cysteines of the CH1 differ, resulting in different location of the interchain disulphide bond. Even so, the spatial location of the chains is still able to command a similar backbone to the Ig fold (Loh et al., 2015; Rothlisberger et al., 2005).

Smaller fragments are engineered from Fab fragments such as scFv (single-chain fragment variable). scFv is a truncated Fab that consists of the variable domain of a heavy chain and the light chain connected via a flexible polypeptide linker (Ewert et al., 2003). The polypeptide linker plays an essential role in the stability of the VH-VL interface of the scFv fragment. It may cause transient opening of the interface of the scFv, exposing the hydrophobic regions. This can lead to aggregation where the VH and VL domain will dissociate or even form dimers (Worn & Pluckthun, 2001). Therefore, the choice of a linker is very important in order to construct a stable scFv construct. There must be a certain degree of flexibility by the linkers that associate the VH and VL domains and form a functional VH-VL interface. For instance, the most commonly used flexible linker in constructing scFv is the glycine-serine (GGGS)n linker. This linker allows the mobility of the connecting VH-VL domain and at the same time, maintains its own stability in the aqueous solutions (Brichta et al., 2005; Chen et al., 2013). Besides, the stability of the scFv construct is also depending on the intrinsically stable VH and VL domains which involves the intradomain disulphide bonds. The positions of the cysteines of the two variable domains are highly conserved which are at position H22/H92 and L23/L88 respectively. Though, the correct folding of the VH and VL domain respectively can also be achieved even without the good formation of disulphide bond. However, some reports have claimed that many scFv are not stable in the absence of the disulphide bond (Worn & Pluckthun, 2001).

In order to acquire the stable and soluble antibody fragment, selection of such antibody fragments from large antibody libraries will have to be carried out. Generally, there are four types of libraries which are naïve, immune, semi-synthetic and synthetic antibody libraries. These libraries vary in their origins or source, their

library sizes, and diversity (Bahara et al., 2013; Mondon et al., 2008). Naïve libraries are usually made from a naïve collection of antibody genes where it does not shown any preference in terms of diversity towards a particular antigen. The common source for naïve libraries are B-lymphocytes from healthy donors. Immunized libraries on the other hand are generated by a collection of skewed antibody genes generated as a response normally during infections. The source from which the genes can be derived from donors recently infected with a disease (Bazan et al., 2012; Kim et al., 2005). Synthetic antibody libraries are made up of a naïve repertoire but the genes are derived from a synthetic platform and not from lymphocytes. Generation of synthetic libraries involves random mutation of both variable heavy (VH) and variable light (VL) chain complementarity determining regions (CDRs) by using oligonucleotides resulting in a diverse repertoire of human VH and VL genes. The difference between the synthetic antibody and semi-synthetic antibody library is that semi-synthetic antibody combines natural CDRs and artificial ones instead of being fully synthetized (Ponsel et al., 2011).

Nevertheless, the availability of these combinatorial libraries allows for the generation of engineered antibody with customized affinity and stability. There are various *in vitro* display technologies for selection of antibodies such as yeast display, phage display, ribosome display, mRNA display and bacterial display. Despite these technologies, antibody phage display, currently, is still the gold standard for recombinant antibody production due to its many success stories (Rakonjac et al., 2011).

1.4 Phage Display Technology

The developments in molecular biology have greatly improved the production of recombinant antibodies without the drawbacks of hybridoma technology using phage display. Phage display technology was first introduced by George Smith in the year 1985 (Smith & Petrenko, 1997). Phage display allows for the presentation of antibodies on phage surfaces that can be used to screen antigens using a process called biopanning (Barbas et al., 2004; Schirrmann et al., 2011). The biopanning process is a repetitive process that allows the enrichment and separation of specific antibodies by virtue of affinity. Therefore the availability of a combinatorial library on phage will allow the presentation of a diverse collection of antibody molecules in a small volume. This will allow easy isolation of monoclonal antibodies as phage display allows a physical interaction between the phenotype and genotype. Such technology has also lead to various techniques for screening and selection of the peptides and proteins, including antibodies against any target, specifically and with high affinity (Hust & Dübel, 2004; Schirrmann et al., 2011).

The bacteriophage that is mostly used in phage display technology is filamentous phage that can infect most Gram negative bacterias. M13 bacteriophage is the most widely used in this field. M13 phage is usually thin and cylindrical in shape with a single-stranded DNA genome as its genetic material that encodes for 5 coat proteins, pIII, pVI, pVII, pVIII and pXIX (Gao et al., 1999). pIII and pVIII are the most commonly used to present the peptide or protein at the surface of the phage. pIII, which is encoded by gene 3 protein, can only display 3-5 copies of each peptide while pVIII, encoded by gene 8 protein, can display up to 2700 copies of peptide (Figure 1.2). The bacteria that the phage infects are bacteria only such as *E.coli* bacterial cells that have the F plasmid that encodes for the F-pilus. This allows

infection to occur, mediated by pIII coat protein of the phage which comprises of three domains such as N1, N2 and C. These phages will not lyse the cell after infection, unlike lytic phages. Instead, they keep replicating in the host cell and are released from the cell membrane, at the same time, the host cell continues to grow and divide (Rakonjac et al., 2011; Russel & Model, 2006).

During infection, the N2 domain of pIII of the phage will bind to the tip of the F-pilus of the bacterial cell. The pilus then, contracts to bring the phage into contact with a co-receptor, TolA, followed by the binding of the N1 domain to TolA in the periplasmic region of the cell. C domain of the pIII forms an α-helical structure which allows the entry of phage ssDNA into the cytoplasm. Upon infection, the phage ssDNA enters the cytoplasm and it is replicated by host enzymes while host cell growth continues. It commences with the synthesis of the negative (-) strand by RNA polymerase, generating a double-stranded replicative form(rf). PII which is the replication protein of the filamentous phage binds to the positive strand of the rf. It acts as a nickase as it makes a nick at the positive strand, binds to the 5' end, leaving a free 3' hydroxyl end to serve as a primer for the host DNA polymerase III to synthesize the new positive strand. The rf then undergoes rolling-circle replication of the positive strand. The positive strand will be displaced as the replication starts. As each genome length is completed, it is cleaved by pII and then, the ends will be re-ligated and circularized. As the copy number of dsDNA increase, the synthesis of phage encoded proteins will also increase. The positive strands are then coated by the pV protein with the ssDNA-binding protein serving as a packaging signal. Besides, pV also inhibits the translation of pII in order to stop the ssDNA replication process.

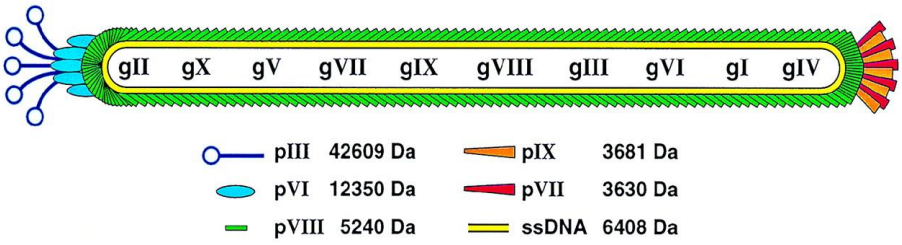


Figure 1.2 : M13 phage and its coat proteins (Gao et al., 1999).

The ssDNA-pV complex will give signal to the pI, pXI and pIV to form a phage export pore complex. PIV makes a channel in order to export the phages. At the same time, pVII and pIX will identify the phage ssDNA genomes to be packaged into the virions while pIII, pVI and pVIII are transported into the cytoplasmic membrane before the phage assembly. The ssDNA genome that is coated by pV will be replaced by pVIII as it passes through the membrane. After the completion of the protein assembly, the phage will pass through the pIV pore to emerge as mature phages and then, the pore is replugged by the pI and pXI without lysing the host cell (Russel et al., 2004; Sidhu, 2001).

In phage display technology, there are few types of pIII display systems such as Type 3, Type 33 or Type 3+3 (Paschke, 2006). The most commonly used system will be Type 3+3 and it requires a phagemid. Phagemid acts as a vector that consists of the origin of replication (ori) from the filamentous phage, the plasmid origin of replication of the *E. coli*, gene III or gene VIII protein, multiple cloning site (MCS), a promoter and also an antibiotic resistance gene (Mullen et al., 2006). The main two purpose of the phagemid is for cloning and also expression of desired protein on the phage's surface. It can be grown as a plasmid in *E. coli* and a recombinant filamentous phage can be packaged with the aid of helper phages (Bradbury & Marks, 2004; Paschke, 2006). Due to the insufficient amount of the filamentous phage genes that encodes for the structural and non-structural proteins in the genome, the helper phage is required to provide all of the necessary proteins for phage assembly. A phage display library that contains clones that carry different foreign DNA insert can be formed (Smith & Petrenko, 1997).

Proteins, in this case, antibodies will be presented at the coat protein of the phage. However, screening and selection of desired antibodies will have to be carried out by affinity selection on the immobilised target. This process is known as phage display panning (Figure 1.3). Targets will be immobilised and it is incubated with the phage library. The complementary protein presented on the phage will be bound and the unbound ones will be washed off. Then, the bound phages will be eluted and reinfected into *E. coli* for amplification. Basically, the panning process is a cycle where several rounds of panning and phage amplification are carried out as shown in Figure 1.3. Then, the protein or antibody presented will be analysed (Bazan et al., 2012; Bradbury & Marks, 2004; Gao et al., 1999).

1.5 Molecular Chaperones

The protein aggregation and the inefficient folding of protein secreted by *E. coli* often occurs during protein expression especially when the nascent proteins that needed to be translocated to the periplasmic layer for proper folding of the protein (Missiakas & Raina, 1997). The antibody fragments are no exception to this especially the Fab fragment that requires the formation of disulphide bridge for proper folding (Plückthun, 1990). This is often the problems associated with Fab expression. In order to overcome these problems, molecular chaperones have been introduced during protein expression. Molecular chaperones are the protein components that participate in protein folding by assisting the nascent proteins to fold into their native fold, prevent aggregation of the protein and some protect the proteins from heat shock (Goemans et al., 2014; Hendrick & Hartl, 1995).

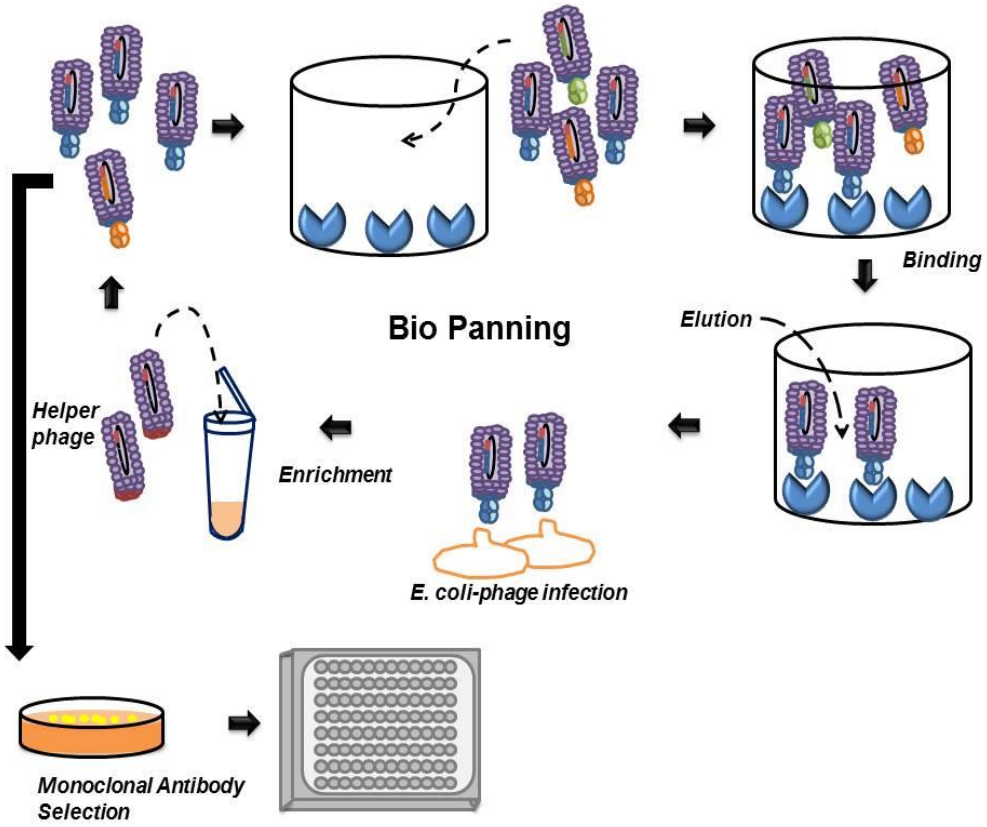


Figure 1.3 : Bio-panning or phage display panning cycle.

The ability to form disulphide bridges during protein folding is usually the rate-limiting factor that results in inefficient protein folding and also decreases the yield of successfully folded proteins. Thus, recently, some molecular chaperones that catalyze the disulphide bridge formation have been recognized. As reported by Schlapschy et al., pTUM4 that encodes for few molecular chaperones such as DsbA, DsbC, FkpA and SurA was able to improve the formation of disulphide brigdes (Appendix G). DsbA and DsbC were able to catalyse the formation of disulphide bridge while FkpA and SurA help in increasing the yield of the protein produced (Schlapschy et al., 2006; Schlapschy & Skerra, 2011).

DsbA is a monomeric periplasmic protein that has two redox-active site cysteines residues present in a CXXC motif. This motif, more precisely, the Cys30, is in oxidised state at physiological pH, hence, the motif will be stabilized and this initiates the redox reaction of the DsbA and the substrate protein. As the unfolded substrate protein enters the periplasmic layer of bacteria, the oxidized DsbA will react with the substrate protein by transfering its disulphide to the adjacent cis-Pro motif of the substrate protein. After transferring its disulphide bond to the substrate protein, the reduced DsbA will be released and reoxidized by the membrane-boun protein, DsbB (Collet & Bardwell, 2002; Heras et al., 2007; Messens et al., 2007).

After introducing the non-native disulphide bridge, DsbC will take part in the isomerization of the disulphide bridge. DsbC, a V-shaped form protein, comprises of a catalytic thioredoxin domain and a dimerization domain linked by a α-helix motif (Nakamoto & Bardwell, 2004). It has a hydrophobic surface which is believed to be the binding site of substrate proteins. DsbC is able to shuffle the disulphide bridge to the correct position in the substrate protein and isomerize it. This significant event has allowed the substrate protein to fold correctly. Nevertheless, the process time of

the oxidation-isomerization of the disulphide bridge is greatly reduced by both of these chaperones (Collet & Bardwell, 2002; Goemans et al., 2014; Heras et al., 2007).

FkpA and SurA, on the other hand, are also involved in the periplasmic protein folding pathway (Duguay & Silhavy, 2004). FkpA is described as a V-shaped homodimer, comprising of two domains such as the N-terminal domain that is involved in dimerization and the C-terminal domain is responsible for the peptidyl-prolyl isomerase (PPIase) activity. Such PPIase activity catalyzes the cis/trans isomerization of the peptide bonds involving proline residues (Bothmann & Plückthun, 2000; Justice et al., 2005). Apart from that, FkpA acts as a folding enhancer during the protein folding activity. FkpA suppresses the formation of insoluble proteins in the periplasmic layer. In fact, FkpA does improve the yield of the protein expressed by *E. coli* (Schlapschy et al., 2006).

SurA, which is originally known as a protein needed for survival, has being recognized as protein that exhibits the PPIase activity too. SurA comprises of a N-terminal domain, two PPIase domain and a C-terminal helix domain. The N-terminal and C-terminal domain with one of the PPIase domain which has no PPIase activity forms a core structure that possesses the chaperone activity while the other PPIase domain is active in PPIase activity (Hennecke et al., 2005). SurA is involved in the assembly of the endogenous outer membrane porins which is essential for the cell. It can also promote the folding of the unstable proteins and prevents accumulation of misfolded proteins (Schlapschy et al., 2006). Although both FkpA and SurA can exhibit the PPIase activity, reports found that PPIase does not play an important role in increasing the rate of the periplasmic protein folding. However, SurA and FkpA can improve the efficiency of protein folding in the periplasmic area (Goemans et al., 2014).

1.6 Reporter Systems

The sensitivity of an immunoassay is not solely dependent on the antibody quality but also on the reporter system that is used. The detection of the reactants can be either visualized or measured by different instruments depending on the nature of the reporters. Many researchers or lab users utilize enzyme-labelled antibodies as the reporter to detect antigen. Such immunoassay is called the enzyme-linked immunosorbent assay (ELISA). This assay which was first introduced in the year of 1971 by Engvall and Perlmann has now become a well-known diagnostic tool in clinic, plant pathology and food industry (Andreotti et al., 2003). This immunoassay may give rise to different signal sources such as colorimetry, fluorometry and chemiluminescence, depending on the substrate used in the assay. The enzyme-labelled immunoassay with colormetric readout is one of the simplest assays to be carried out (Gan & Patel, 2013). The common enzymes used are the horseradish peroxidase and alkaline phosphatase. These enzymes act as catalyst in which they will convert the colourless substrate such as ABTS and TMB in the presence of hydrogen peroxide to a coloured end-product. These enzymes will keep on converting the substrate until a stopping reagent, for instance, acid, is added to stop the enzyme activity (Gan & Patel, 2013). In order to measure the intensity of the colour and estimate the concentration of the analyte, a spectrophotometer is required. However, the sensitivity of this immunoassay is still a limitation which led to the emergence of more sensitive indicators.

Fluorometric enzyme-labelled immunoassays utilizes enzyme to convert its substrate to a fluorescent end-product instead of a coloured end-product. Fluorophores in the substrate absorbs energy, in this case, light, at a specific wavelength and emit the light again at a longer wavelength. 4-methylumbelliferyl

phosphate (4-MUP) is the most widely used in fluorometric enzyme immunoassay while the enzyme label used is alkaline phosphatase. This enzyme will dephosphorylate 4-MUP to form 4-methylumbelliferone (4-MU) which acts a fluorophore in the assay. Then, it will excite the light at 365nm and emit at 448nm, measured using fluorescence spectrophotometer (Wild, 2005). Instead of using enzymes, a fluorophore can be directly labelled to the antibody such as fluorescent dyes by coupling chemistries or fluorescent protein by cloning. Such an assay is called the direct fluorescence immunoassay and it can also be designed into different formats too. The fluorescence signal given out is comparatively higher than the signal given out by colorimetric immunoassay. However, interferences may occur due to light scattering, background fluorescence and quenching of the sample used in the immunoassay. Light scattering and background fluorescence that is naturally present in biological samples can cause an increase in background reading. The signal can also be quenched by the sample and give rise to false negative result (Andreotti et al., 2003).

Compared to the colorimetric and fluorometric immunoassays, chemiluminescent immunoassays have achieved higher levels of sensitivity. It utilizes chemiluminescent compounds such as luminol, isoluminol and acridinium ester that can emit light as a result of a chemical reaction. Like the fluorometric immunoassays, chemiluminescent compounds can either be directly labelled to the analyte or used as a substrate in an enzyme-labelled immunoassay (Andreotti et al., 2003). Chemiluminescence enzyme-labelled immunoassay utilizes an enzyme such as alkaline phosphatase to cleave the phosphate group of the chemiluminescent substrate, for example, adamantyl 1,2-dioxetane arylphosphate, and produce an unstable anion which then decays and causes the emission of light. Although the

enzymatic reactions may produce weak light emission which rapidly decays, the addition of another chemical such as hydrogen peroxide enables to enhance the light signal. The enhanced luminescence can persist for hours though the signal readout usually taken within minutes of the reaction (Weeks et al., 2013).

Apart from these immunoassays, many researchers used many new detection methods in order to amplify the signals and reduce the background reading (Weeks et al., 2013). Interference in fluorometric immunoassays that causes high background readout due to the protein or colloids can be circumvented using time-resolved fluorescence. The basic principle behind time-resolved fluorescence is having a time gap between excitation and measurement of the emitted light from the fluorophore. Therefore, the fluorophore used in this assay must have a longer decay time compared to the background fluorescence (Kwok & Chen, 2003; Liu et al., 2009). Lanthanides are the ideal fluorophores for this assay as they have an intense long-lived fluorescence signal and a large Stokes shift (Tully & O'Kennedy, 2014). They can form fluorescent chelates with certain organic ligands. Signal can be strongly emitted when the lanthanides are excited at a specific wavelength of light. Due to its long-lived fluorescence, the lanthanides can be pulse-excited and generate accumulated signal which can enhance the signal strength of the assay (Andreotti et al., 2003).

Some exploit the strong interaction of biotin and streptavidin in order to amplify the signal in an immunoassay (Koivunen & Krogsrud, 2006). Each streptavidin molecule possesses four binding sites for biotin and form a complex lattice. In the assay, the analyte is usually biotinylated whereas the streptavidin is conjugated with a signal-generating enzyme such as HRP and alkaline phosphatase.

The enzyme conjugated streptavidin will bind to the biotinylated analyte very quickly and form a complex lattice which can then amplify the signal produce in the assay (Andreotti et al., 2003). The interaction of streptavidin and biotin can be easily applied to any range of the immunoassay as long as the analyte is biotinylated. Utilization of the enzyme-conjugated streptavidin as a generic signal generation reagent has solved the need to develop different conjugation methods of the reagents that are used in different assay.

In recent years, polymerase chain reaction has been successfully integrated into immunoassays that utilize DNA as a label bringing about Immuno-PCR (Niemeyer et al., 2015). With PCR, a million-fold increase in the number of copies of DNA can be obtained if the DNA is replicated in approximately 25 cycles. Thus, the amplification of the DNA can lower the detection limit of the assay. As such, further improvements of immunoassay, from using polymerase chain reaction to isothermal amplication, have been achieved too. A single temperature is needed to amplify the DNA using an incubator instead of a thermocycler can simplify the setting of the immunoassay and also lower the cost of it (Schweitzer et al., 2000). Another alternative DNA-label detection method without the need of PCR is the fluorophore DNA barcodes. The single-stranded fluorescent-labelled DNA and the analyte are coated to the microbeads. These microbeads are then allowed to bind to the ligand and the bound fluorescent-labelled DNA will be released and quantified. The signal can be amplified as there are many fluorophore-DNA molecules that are released from one microbead (Weeks et al., 2013). As this signal can only be quantified using a fluorescent spectrophotometer, thus, the discovery of DNA G-quadruplex that mimics the peroxidase activity can be used in immunoassay to eliminate the need of fluorescent spectrophotometer (Ruttkay-Nedecky et al., 2013; Stefan et al., 2012).

Instead, it uses the UV spectrophotometer for quantitative analysis or using naked eyes for qualitative analysis. Nevertheless, the utilization of DNA as the reporter in immunoassays is the key for immunoassay improvements.

1.7 Overview of DNA G-Quadruplex as Reporter System

Deoxyribonucleic acid, DNA, plays a crucial role in storing genetic information of every living organism. DNA replication and transcription mechanisms have been well studied with many findings that show DNA to be a very dynamic molecule. It has the capability to form a number of spatial arrangements such as single-stranded hairpins, homoduplexes, triplexes and quadruplexes (Bochman et al., 2012; Davis, 2004). Later findings had also show that these structures are involved in DNA recombination, regulation of gene expression and potentially proliferation of tumour cells. Therefore, researchers have grown interest in studying the structures of nucleic acids and their application in medical science.

One of the DNA structures that is well studied is the G-quadruplex structure (Doluca et al., 2013). The structure comprises of a nucleic acid sequence rich in guanines (G). The typical base pairing in nucleic acid is called the Watson–Crick base-pairing which is the pairing of guanine and cytosine or adenine and thymine with hydrogen bond. However, nucleic acids can also be involved in another type of pairing known as Hoogsteen base pairing, allowing the formation of higher-order structures such as the quadruplex structures (Haider et al., 2011; Huppert, 2008; Keniry, 2000). This structure comprises of planar stacks of two or more of the G-quartets that are joined by the phosphodiester backbone and is stabilized in the presence of specific cations. G-quartet is a square planar array that is formed by the linkage of four guanine bases. Many NMR and crystallographic studies have

reported that G-quadruaplexes are highly polymorphic (Adrian et al., 2012). They can form many different structural arrangements depending on the length of the DNA, orientation of the chains, positions of the loops and nature of the cations. Figure 1.4A shows some of the most common G-quadruplex structures such as parallel (tetramolecular), hairpin (bimolecular), basket and chair (unimolecular) (Burge et al., 2006). In recent years, the properties of the G-quadruplex have been well studied and some of the properties are exploited in many biological applications such as biosensor and therapeutic application (Liu et al., 2009).

G-quadruplex is found to have the capability in generating enzymatic reaction and it is called the DNAzyme (Schlosser & Li, 2009). G-quadruplex DNAzymes, initially known as catalytic enzymes or DNA enzymes is able to exhibit catalytic capabilities as they can catalyze many reactions such as DNA modification, ligation, cleavage of DNA or RNA and also methylation of porphyrin rings (Breaker & Joyce, 1994; Haider et al., 2011; Li et al., 2000; Li & Sen, 1996). Therefore, DNAzymes can be exploited into many applications especially in the field of medicine, biology and material sciences. The utilization of DNAzymes has many advantages as DNAzymes are stable in a broad range, even at high temperatures unlike the classical protein enzyme that is usually active in a narrow temperature range. Besides, they can be easily prepared by chemical synthesis and by PCR whereas protein enzymes require tedious preparation and purification.

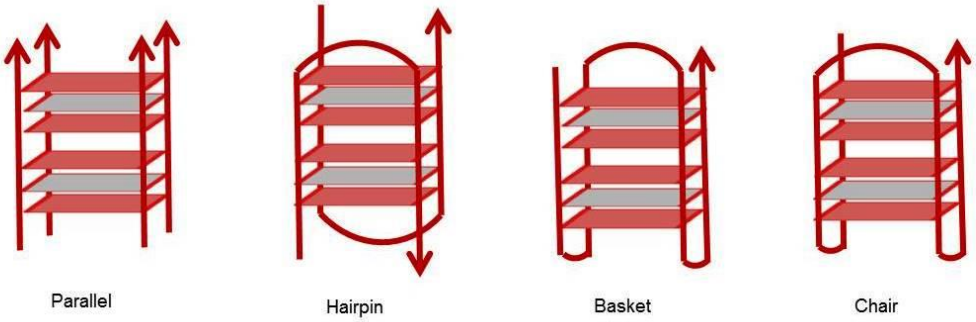
One of the most important features of DNAzyme activity of quadruplex-forming oligonucleotides is that it can mimic the peroxidase activity (Stefan et al., 2011). More precisely, this DNAzyme exhibits peroxidase activity when hemin is bound to the quadruplex (Kosman & Juskowiak, 2011; Saito et al., 2012). This complex

catalyzes the peroxide-mediated oxidation of the 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) diammonium salt, in short ABTS (Figure1.4B).

Thus, in the presence of H_2O_2, oxidation of the colorless ABTS will generate a colored product. This characteristic has enabled the use of DNAzyme as a reporter system in various assays as it also works well with TMB substrate. For instance, DNAzyme-functionalized gold nanoparticles (GNPs) used as nanoprobes in immunoassay to amplify detection of protein cancer biomarkers. A "sandwich-type" immunoassay with two types of probes was employed. One of the probes used is the magnetic microparticles (MMPs) functionalized with the protein cancer biomarker, α-fetoprotein (AFP) monoclonal antibodies for specific detection. Another probe is GNPs that are functionalized with double-stranded DNA and AFP polyclonal antibodies. Both GNPs and MMPs are mixed together in excess in order to capture the AFP effectively. These MMP-AFP-GNP complexes were collected magnetically and washed extensively. Then, the double stranded DNA on GNPs was denatured to form active DNAzymes in the presence of hemin and then reacted with ABTS as substrate (Zhou et al., 2009).

Apart from the peroxidase activity, G-quadruplex can also enhance the fluorescence or luminescence signal to almost two- to ten-fold higher quantum yields. For example, a switch-on detection platform for human neutrophil elastase (HNE) has exploited the duplex-to-quaduplex conversion strategy with iridium (III) complex by using the G-quadruplex as the probe. First, the HNE aptamer was hybrized with the complementary DNA strand. The iridium (III) complex binds only weakly to the duplex, therefore, it emits a low lumiscence signal. The additional of the HNE protein causes the duplex structure to dissociate and form the HNE-aptamer-quadruplex. Then, the iridium (III) complex strongly interacts with the newly formed

A



B

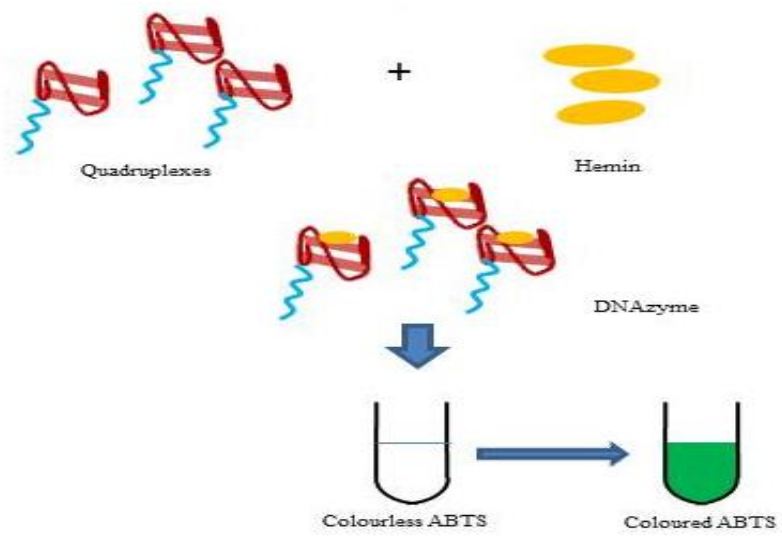


Figure 1.4 : A. Typical G-quadrplex structures B. Hemin is bound to the quadruplexes to acts as DNAzyme that can exhibit peroxidase activity. This complex catalyzes the peroxide-mediated oxidation of the 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) diammonium salt (ABTS). In the presence of H_2O_2, oxidation of the colorless ABTS will generate a colored product.

quadruplex, resulting in an enhancement in luminescence signal (Ma et al., 2013). Nevertheless, the development of immunoassays using G-quadruplex structures as the reporter system have now gain a lot of attention as they can achieve catalysis with high accuracy, stability and reusability.

1.8 Quadruplex Priming Amplification (QPA)

QPA is a new method that involves isothermal amplification of a linear GGGT sequence that allows the formation of DNA quadruplexes in solution. QPA was designed under the fundamentals that free energy of DNA quadruplexes can be used to initiate endergonic reactions at constant temperatures. Different G-rich sequences are capable of forming quadruplexes with different preferential thermodynamics. QPA makes use of special primers that are designed with sequences that allow spontaneous dissociation from the hybridization site and fold into intramolecular quadruplexes in the presence of cations upon polymerase elongation. The primer, a complementary sequence of the template is designed with one or two guanine residues missing which allows self-dissociation from the template after elongation. The amplification begins with the binding of the primer to the template, forming a DNA duplex and the polymerase will start to elongate the primer. During the elongation, the 5' end of the product will fold into an intramolecular quadruplex and self-dissociates, leaving its complementary template completely accessible to the incoming primer for the next priming cycle (Figure 1.5). This allows the continuous formation of quadruplex structures until all the primers are spent or the reaction is terminated (Johnson et al., 2013; Kankia, 2011; Taylor et al., 2013). To date, Kankia et al. had developed a fluoresecence based DNA detection method using QPA. However, this method has not been used for colorimetric based protein detection