PHAGE DISPLAY OF NAÏVE HUMAN T-CELL RECEPTORS IN SINGLE CHAIN FORMAT AGAINST ISOCITRATE LYASE (ICL) AND TUMOR NECROSIS FACTOR ALPHA (TNFα)

ANGELA CH'NG CHIEW WEN

UNIVERSITI SAINS MALAYSIA

2024

PHAGE DISPLAY OF NAÏVE HUMAN T-CELL RECEPTORS IN SINGLE CHAIN FORMAT AGAINST ISOCITRATE LYASE (ICL) AND TUMOR NECROSIS FACTOR ALPHA (TNFα)

by

ANGELA CH'NG CHIEW WEN

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

April 2024

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to those who have helped me throughout my phD study.

First and foremost, I am grateful to my supervisor Associate Professor Dr. Lim Theam Soon for his unwavering guidance, invaluable support and continuous encouragement throughout this research journey. His invaluable constructive criticism and insightful feedback have helped me in improving my research ideas and skills. His dedication in teaching regardless of knowledge given or personal growth shall be brought with me even after I have graduated.

Besides, I would also like to express my gratitude to Associate Professor Dr. Choong Yee Siew as my co-supervisor and collaborators in helping my research. Their help and guidance led to my success in completing this dissertation. I would also like to thank all the science officers and staff in INFORMM for their help in equipment using and official documents handling. Apart from that, I would like to thank my beloved colleagues, especially Jing Yi that always lend me their hands whenever I need help.

Last but not least, I would like to extend my heartfelt thanks to my family and friends for their unwavering love, encouragement, and understanding throughout this challenging journey. Their constant support, belief in my abilities, and sacrifices have been the driving force behind my achievements.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF ABBREVIATIONS

3D	3 dimensional
ABTS	2,2’-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid)
APC	Antigen presenting cell
APS	Ammonium persulfate
ATP	Adenosine triphosphate
βME	Beta-mercaptoethanol
bp	Base pair
BSA	Bovine serum albumin
Cα	Constant region of alpha
Cβ	Constant region of beta
CD3	Cluster of differentiation 3
cDNA	Complementary deoxyribonucleic acid
CDR	Complementarity determining region
CFU	Colony forming unit
C_H1	Constant of heavy chain 1
Cκ	Constant of kappa light chain
CV	Column volume
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
E. coli	*Escherichia coli*
ELISA	Enzyme-linked immune assay
EtBr	Ethidium bromide
Fab	Fragment antigen-binding
eGFP	Enhanced Green fluorescence protein
HC	Heavy chain
His-tag	Histidine tag
HLA	Human leukocyte antigen
HRP	Horseradish peroxidase
ICL	isocitrate lyase

Ig	Imuunoglobulin
IPTG	Isopropyl-beta-D-thiogalactopyranoside
LC	Light chain
MAIT	Mucosal-associated Invariant T
MHC	Major histocompatibility complex
MTB	*Mycobacterium tuberculosis*
OD	Optical density
PBS	Phosphate-buffered saline
PBST	Phosphate-buffered saline with Tween 20
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PFU	Phage forming unit
PMHC	Peptide major histocompatibility complex
PTM	Milk powder in PBST (blocking buffer)
RAG	Recombinant activating gene
RE	Restriction site
RF	Replicative form
rICL	Recombinant isocitrate lyase
RNA	Ribonucleic acid
RT	Room temperature
rTNFα	Recombinant tumour necrosis factor alpha
scFv	Single chain fragment variable
scTCR	Single chain T-cell receptor
SDS	Sodium dodecyl sulphate
SDS PAGE	SDS-Polyacrylamide gel electrophoresis
ssDNA	Single stranded DNA
TCR	T-cell receptor
TCRAV	T-cell receptor alpha variable
TCRBV	T-cell receptor beta variable
TCRV	T-cell receptor variable
TNFα	Tumour necrosis factor alpha
TRAC	T-cell receptor alpha constant

TRBCII	T-cell receptor beta constant II
TRAV	T-cell receptor alpha variable primer
TRBV	T-cell receptor alpha variable primer
Vα	Variable alpha
Vβ	Variable beta
V	Volt

LIST OF SYMBOLS

α	Alpha
β	Beta
γ	Gamma
δ	Delta
°C	Degree Celcius
dH_20	Distilled water
g	Gram
h	Hour
λ	Lambda
L	Liter
M	Molar
µg	Microgram
µL	Microliter
µM	Micromolar
mg	Milligram
mL	Milliliter
min	Minute
ng	nanogram
%	Percent
s	second
U	Unit of enzyme
v/v	Volume/volume
w/v	Weight/volume

LIST OF APPENDICES

PAPARAN FAJ RESEPTOR SEL-T MANUSIA YANG NAIF DALAM FORMAT RANTAI TUNGGAL TERHADAP ISOSITRAT LIASE (ICL) DAN FAKTOR NEKROSIS TUMOR ALFA (TNFα)

ABSTRAK

Reseptor sel T (TCR) mengikat kompleks peptida-MHC atau lipid-CD1 sementara antibodi mengikat secara langsung kepada antigen. Kawasan pengikatan TCR dan antibodi menunjukan struktur tiga dimensi (3D) yang serupa. Kedua-duanya terdiri daripada dua rantai dan setiap rantai mengandungi dua domain terlipat iaitu satu pembolehubah dan satu tetap yang bersekutu untuk membentuk tapak pengikatan antigen di antara mereka. Oleh kerana kedua-duanya serupa, terdapat potensi untuk TCR berfungsi sebagai landasan pengikatan kepada antigen protein sepenuhnya, sama seperti antibodi. Oleh itu, keseluruhan kajian eksperimen adalah bertujuan menyiasat sama ada TCR boleh mengikat kepada protein antigen sepenuhnya seperti cara antibodi mengikat kepada protein antigen. Tapak pengenalan TCR terdiri dari rantai alfa (α) dan beta (β). Oleh itu, kedua-dua rantai pembolehubah alfa dan beta dihubungkan melalui penghubung glisine-serin, menghasilkan pembentukan rantai Tunggal TCR. Selain itu, sebuah perpustakaan scTCR naif telah dibina dengan menggunakan teknologi paparan faj dengan saiz perpustakaan 10^{10}. Kemudian, perpustakaan scTCR digunakan untuk mengasingkan pengikat scTCR terhadap dua antigen terpilih, lyase isositrat rekombinan (rICL) dan faktor nekrosis tumor alfa rekombinat (rTNFα). Sebanyak tiga pengikat scTCR monoklonal kepada rICL dan dua kepada rTNFα masing-masing telah diasingkan. Semua protein gabungan scTCR

pemalar antibodi manusia (CH1 atau CK) dinyatakan terbaik pada 30 °C kecuali a-rICL-2G10 scTCR. Kedua-dua a-rICL-2D10-CK scTCR dan a-rICL-2G10-CK scTCR berprestasi terbaik dengan gabungan CK tanpa bantuan pendamping SKP manakala a-rICL-2C11-CH1 scTCR dan a-rTNFα-F4-CH1 scTCR dilakukan terbaik dengan bantuan pendamping SKP, kecuali a-rTNFα-2B3-CH1 scTCR. Kemudian, kedua-dua pengikat scFv dan scTCR yang telah dimurnikan telah dititrasi untuk menilai kecekapan pengikatan dan kereaktifan silang dengan antigen tidak spesifik. Akibatnya, pengikat protein gabungan scTCR terhadap rICL dapat mengikat antigen sasarannya rICL (> 2.4) dengan kereaktifan silang dengan protein tidak spesifik lain seperti eGFP dan α-kristal manakala scFv terhadap rICL mengikat pada nilai penyerapan yang lebih rendah (0.7) dengan kereaktifan silang yang lebih rendah (< 0.2). Kedua-dua pengikat scFv dan scTCR terhadap rTNFα dapat mengikat protein sasaran rTNFα (0.25-1.4) tanpa bertindak balas silang (< 0.1) dengan protein tidak spesifik. Kesimpulannya, scTCR monoklonal yang diasingkan boleh berpotensi untuk digunakan sebagai motif pengikat sasaran untuk aplikasi bioperubatan, sama seperti cara antibodi berfungsi walaupun kejuruteraan lanjut diperlukan untuk ciri-ciri dan peningkatan pengeluarannya.

PHAGE DISPLAY OF NAÏVE HUMAN T-CELL RECEPTORS IN SINGLE CHAIN FORMAT AGAINST ISOCITRATE LYASE (ICL) AND TUMOR NECROSIS FACTOR ALPHA (TNFα)

ABSTRACT

T-cell receptors (TCRs) bind with peptide-MHC or lipid-CD1 complexes while antibodies directly bind to antigen. The binding region of T-cell receptors (TCRs) and antibodies exhibit similar three-dimensional (3D) structures. Both composed of two chains and each chain contains two folded domains of one variable and one constant which associate together to create the antigen binding site between them. Since both antibodies and TCR have similar structures, there exists a potential for TCRs to serve as a binding scaffold to bind full protein antigens, much like antibodies. Consequently, the entire experimental study was designed with the aim of investigating whether TCRs can indeed bind to full antigen proteins just like how antibodies bind to antigen proteins. The recognition site of the TCR comprises of the alpha and beta variable regions. Consequently, both alpha and beta variable regions are interconnected through a glycine-serine linker, resulting in the formation of a single-chain TCR. Additionally, a naïve scTCR library was meticulously crafted using phage display technology with a library size of 10^{10}. Then, the scTCR library was used to isolate scTCR binders against two designated antigens recombinant isocitrate lyase (rICL) and recombinant tumor necrosis factor alpha (rTNFα). A total of three monoclonal scTCR binders to rICL and two to rTNFα respectively were isolated. All of the human antibody constant (C_H1 or C_K) scTCR fusion proteins expressed best at 30 °C except a-rICL-2G10 scTCR.

Both a-rICL-2D10-C_K scTCR and a-rICL-2G10-C_K scTCR performed best with C_K fusion without the help of the SKP chaperone while the a-rICL-2C11-C_H1 scTCR and a-rTNFα-F4-C_H1 scTCR performed best with the help of the SKP chaperone, except a-rTNFα-2B3-C_H1 scTCR. Then, both purified scFv and scTCR binders were titrated to assess their binding efficiency and cross-reactivity with nonspecific antigens. As a results, the scTCR fusion protein binders against rICL were able to bind to its target antigen rICL (> 2.4) but exhibited cross reactivity with other nonspecific proteins such as eGFP and α-crystalline. However, the scFv against rICL binds at a lower absorbance value (0.7) with lower cross reactivity (< 0.2). Both the scFv and scTCR binders against rTNFα were able to bind to the target protein rTNFα (0.25-1.4) without cross reacting (< 0.1) with the nonspecific proteins. In conclusion, the monoclonal scTCR isolated could have the potential to be used as a target binding motif for biomedical applications, similar to how antibodies work although further engineering is required for its characteristics and production improvement.

CHAPTER 1

INTRODUCTION

1.1 Research background

The primary role of the immune system is to safeguard our body from being harmed. Both B and T lymphocytes are the pivotal components within the immune system that works together to stop and eradicate intruders such as viruses, bacteria, fungi, and parasites from attacking our healthy cells (Janeway, 2001). Antibodies play a relatively more direct role in recognizing targets compared to T-cell receptors (TCRs). Antibodies directly bind to antigen proteins while TCRs bind with peptide-Major Histocompatibility Complex (pMHC) or lipid-CD1 complexes which identify antigenic peptides displayed on the major histocompatibility complex (MHC) (Rossjohn et al., 2015). However, both antibodies and TCRs have a target binding site known as the variable region that exhibits a similar three-dimensional (3D) structure (Knapp et al., 2017). This region in both TCRs and antibodies is composed of frameworks and complementarity determining regions (CDR), creating the binding site for antigens (Richman et al., 2009; Kunik et al., 2012). The high structural similarity between antibodies and TCRs suggests a potential for TCRs to serve as binding scaffolds for antigen proteins since the binding structure of TCRs is analogous to the binding structures of antibodies (Knapp et al., 2017).

The application of antibody-like structures as binding scaffolds is not a new concept with closely related structures being used successfully to bind to target antigens (Chidyausiku et al., 2022). The ability to generate useful scaffolds for use as

alternative affinity binding scaffolds is important as deeper understanding of molecular based immunology highlights the need for different biomolecules to be used for different biological and immunological functions. The introduction of recombinant DNA and protein engineering techniques has highlighted the potential of using alternative affinity scaffolds that are non-immunoglobulin like in structure for medical applications. This involved the use of different proteins from the body that are evolved to generate binding sites based on the CDR concept of antibodies. Although antibodies have been successful as a binding scaffold because of their high binding specificities and generation by animal immunization against various target. However, the limited application of TCRs to only bind a limited class of targets remains an elusive enigma. The ability to have an Ig-like scaffold that has a higher similarity to antibodies would be useful and closer to nature (Koenning & Schaefer, 2021).

1.2 Problem statement and rational of study

The binding and recognition of TCR to protein antigens is indirect as it requires either MHC to display out the single antigenic peptide sequence or CD protein to display out antigen lipid target. In addition, early studies showed that TCRs have relatively low binding affinity for their peptide MHC and lipid-CD1 protein (Stone et al., 2009). Meanwhile, antibodies bind to antigen without having it to be displayed on a carrier protein. Since TCR and antibody have similar structure and gene rearrangement for their 3D protein confirmation, it is likely TCR may be able to bind antigen directly without the display proteins albeit at lower affinities. Therefore, it is critical to establish the basic understanding of TCR-antigen binding to determine the possibility for TCR binders being used as binders to bind protein antigens as an alternative binding scaffold. As most scaffolds are derived from the basic binding site

design as antibody CDRs but with a non-Ig like backbone, the possibility of using an Ig-like backbone with naturally occurring CDRs in TCR as a scaffold proposes an attractive alternative especially for biomedical applications to be applied for situations that are not suitable for antibodies. Currently, therapies like chimeric antigen receptor T cell (CAR-T) utilize T cells with antibodies as recognition sites rather than TCR due to various limitations, such as stability, affinity and solubility (Hiltensperger & Krackhardt, 2023). However, if these issues can be resolved, TCR could be a viable candidate for CAR-T therapy given their close resemblance to the human body's natural immune system. A naïve scTCR phage display library was generated and tested against rICL and rTNFα to determine the possibility of TCR binding to antigens without being displayed on MHC or CD1 proteins. The two antigens were selected due to their negative effects to patients if not inhibited. Additionally, existing inhibitors against these two antigens have side effects. Thus, scTCR can be an alternative if it is able to bind to the antigen protein. The isolated scTCR binders against rICL and rTNFα were expressed as proteins for the antigen binding and cross-reactivity test via enzyme-linked immunosorbent assay (ELISA). Then, the scTCR proteins were compared with the in-house anti-rICL and anti-rTNFα scFv proteins via titration test for binding. This would allow for the initial determination if scTCR scaffolds, especially the complementarity determining regions (CDRs) are capable of reproducing a similar binding characteristic as antibodies against a set of two targets.

1.3 Objectives of study

The objectives of the study are as follows:

1. To determine if scTCR can function as binding scaffold against protein by isolating scTCR against rICL and rTNFα from a naïve scTCR phage display library.
2. To evaluate monoclonal scFv and scTCR against rICL and rTNFα.
3. To compare the binding of scTCR and scFv with rICL and rTNFα

CHAPTER 2

LITERATURE REVIEW

2.1 Phage display technology

Phage display (PhD) technology has played a crucial role in studying protein-protein interactions for the past 25 years. Additionally, phage display has evolved to become a platform for drug discovery and development and immunological studies. Thus, it is a commonly used platform for displaying monoclonal antibodies, peptides (Wu et al., 2016; Rami et al., 2017) and epitope mapping (Chin et al., 2017). This method physically links the phenotypic characteristics with a retrievable genotypic information, making it a valuable method for displaying peptides or proteins on phage surfaces for selection with the possibility of downstream identification. Genotypic diversification contributes to a diverse range of phenotypes which is critical to the success in phage display experiments. This technology was initially developed by George Smith for displaying short peptides on bacteriophage surfaces (Smith, 1985) and it later expanded to include larger molecules like antibodies (Winter et al., 1994) and proteins such as TCR (McCafferty et al., 1990). This technology utilizes filamentous bacteriophages like M13 to infect bacteria without killing them, enabling the bacteria to replicate and present recombinant targets (antibody or peptide) on their surface by fusing them to M13 coat proteins (McCafferty et al., 1990).

2.2.1 Filamentous phage

The main cornerstone of phage display technology is the application of various bacteriophage system, such as filamentous phage (M13, fd and f1), lytic phages (T4 and T7) and temperate phage lambda (λ) to present antibodies (Barbas et al., 1991; Winter et al., 1994; Burritt et al., 1996) and TCRs.

The first filamentous phage, f1, was identified in *E. coli* as part of the Ff group, including f1, M13, and fd phage in the early 1960s (Loeb, 1960; Hofschneider, 1963; Marvin & Hoffmann-Berling, 1963). Ff bacteriophages dominates display technology due to several key advantages. They have titres approximately 100 times higher than other known phages, ensuring large library sizes can be generated. These phages are highly stable at high temperatures (Pershad & Kay, 2013), in detergents, and under extreme pH conditions (Bicalho et al., 2012), enabling a wide range of screening conditions in phage display libraries . These phages are specific for *E. coli* with a conjugative F pilus. Over the past 50 years, various research groups have significantly advanced our understanding of filamentous phage structure, biology, and life cycle. M13 filamentous phages are used for the phage library packaging and biopanning. M13 virions, which include the receptor-binding protein pIII, retain their integrity and infectivity when exposed to a range of pH levels (pH 3-11) and elevated temperatures (below 80 °C) (Branston et al., 2013).

2.2.1(a) The structure of M13 filamentous phage

M13 filamentous phage virions are characterized by their elongated, supple and slender morphology, measuring roughly 6.6 nm in diameter and extending to a length of 880 nm. These virions are composed of thousands of pVIII major coat protein

subunits and organized in a helical structure that envelops a single-stranded DNA (ssDNA) core (Marvin et al., 2006) as shown in Figure 2.1.

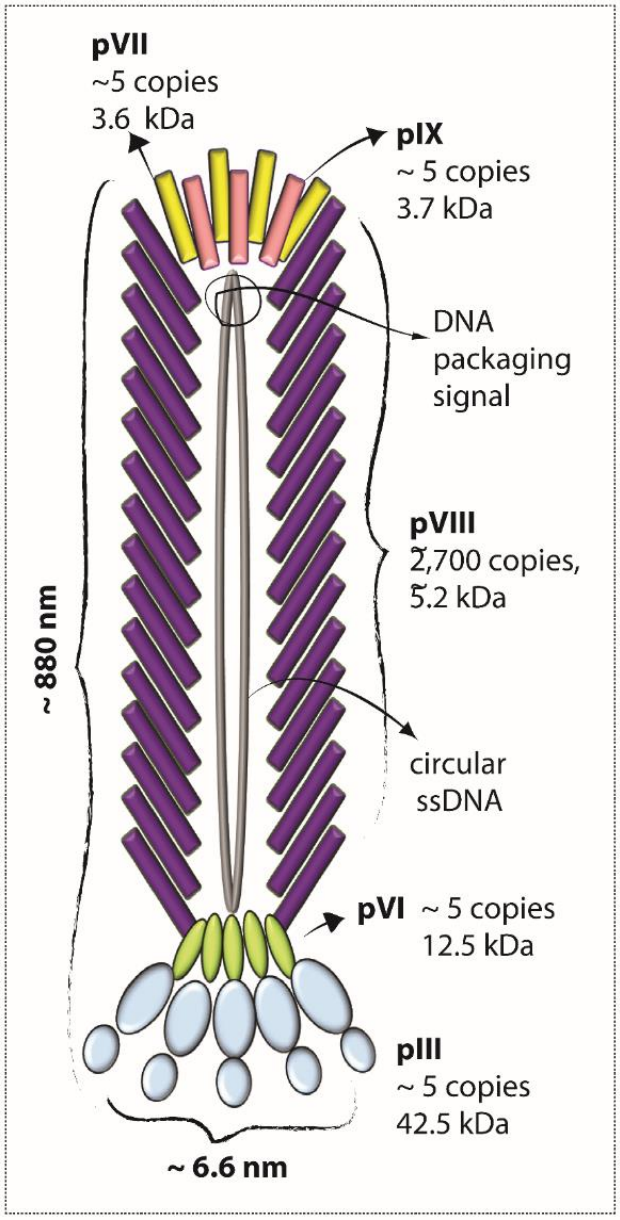

Figure 2.1 The structure of M13 filamentous phage. The M13 bacteriophage features a filament with a length of 880 nm and a diameter of 6.6 nm. Comprising five structural proteins, it envelops a 6.4 kb single-stranded DNA genome. The major coat protein, pVIII is present in approximately 2700 copies while each minor coat protein, including pVII and pIX at the proximal end and pIII and pVI at the distal end, is found in 3-5 copies.

They are typically found in *Escherichia coli* (*E. coli*) which known as Gram-negative bacterias. A distinctive feature of these phages is their capability to reproduce without causing harm to the host cell. They can exist as integrated elements in the host genome or replicate episomally (Young et al., 2000).

The M13 phage is mainly composed of five structural proteins, namely pVIII, pIII, pVI, pVII and pIX which cover the ssDNA molecule. The filament structure of the M13 virion is formed by 2700 copies of the pVIII major coat protein arranged in a helical fashion. In contrast, pIX and pVII minor proteins are compact and contain 33 and 32 amino acids, respectively which cover the tip end of the M13 phage particles. Both of these proteins are hydrophobic (Lubkowski et al., 1998). Additionally, there are 5 copies each of pIII and pVI protein which play pivotal roles in capping another end of the M13 filament. The primary function of the largest pIII protein within the virion is to bind to receptor and facilitate entry of the phage into the host cell (Boeke & Model, 1982).

2.2.1(b) The biology of M13 filamentous phage

The filamentous phage, Ff (comprising f1, M13, and fd phage), is extensively studied. Its core genome encompasses up to 11 genes categorized into three groups: replication, assembly, and structural genes. These genes play a vital role in the entire cycle of replication of Gram-negative bacteria, including infection, replication, and assembly/secretion processes (Mai-Prochnow et al., 2015). The nomenclature for proteins and genes associated with Ff phage uses Roman numerical symbols.

Among these genes, *gII*, *gX* and *gV* are involved in replicative form (RF) replication and in the preparation of newly synthesized ssDNA for assembly. Gene gII

encodes the nickase which allows rolling-circle replication to occur (P. & M., 1988). Additionally, genes *gIII*, *gVI*, *gVII*, *gVIII* and *gIX* encode structural proteins constituting the phage particle. Among these, *gVIII* encodes the major coat protein responsible for constructing the filament's shaft (Endemann & Model, 1995). Genes *gVII* and *gIX* code for compact coat proteins found at single end of the phage particle (Endemann & Model, 1995), whereas *gIII* and *gVI* encodes minor proteins situated at the opposite end of the virion filament (Gailus & Rasched, 1994). The genes *gI*, *gIV* and *gIX* encode proteins that create a transenvelope complex necessary for the filamentous phage particle assembly and secretion (Feng et al., 1999). The pI and pXI proteins form an inner membrane complex where phage assembly occurs with pI containing an ATP-binding Walker motif crucial for its function (Russel, 1991) whereas pIV is an outer membrane protein responsible for forming a sizable gated channel through which the growing phage particle can pass (Spagnuolo et al., 2010). Loss of pI, pIV or pXI prevents assembly in Ff phage (Russel, 1995).

2.2.1(c) The life cycle of M13 filamentous phage

Filamentous phages such as M13 exhibit a unique parasitic relationship with their host compared to the lytic bacteriophages like T4 or T7 (Loh et al., 2019). Generally, bacteriophages are classified based on their replication within the host cell, adhering to either the lytic or lysogenic life cycle (Elois et al., 2023). However, some phages like P1 and λ phages (temperate phages) follow both lytic and lysogenic pathways (Kuo et al., 1987). Lysogenic phages infect bacteria without lysing them but will reduce their growth (Orlova, 2012). Conversely, lytic phages utilize the host DNA replication machinery and ultimately lyse the host cell (Clokie et al., 2011). Notably, M13 (f1 or fd) phages are frequently employed as phage host for phage display. The

M13 phage is a non-lytic filamentous phage that only infects *E. coli* expressing the F pilus, allowing phage adsorption through binding to the phage coat protein (Tzagoloff & Pratt, 1964). Remarkably, M13 employs an energy-consuming process known as non-lytic replication cycle as shown in Figure 2.2 to assemble and exit the host envelope without causing host cell death (Model & Russel, 1988).

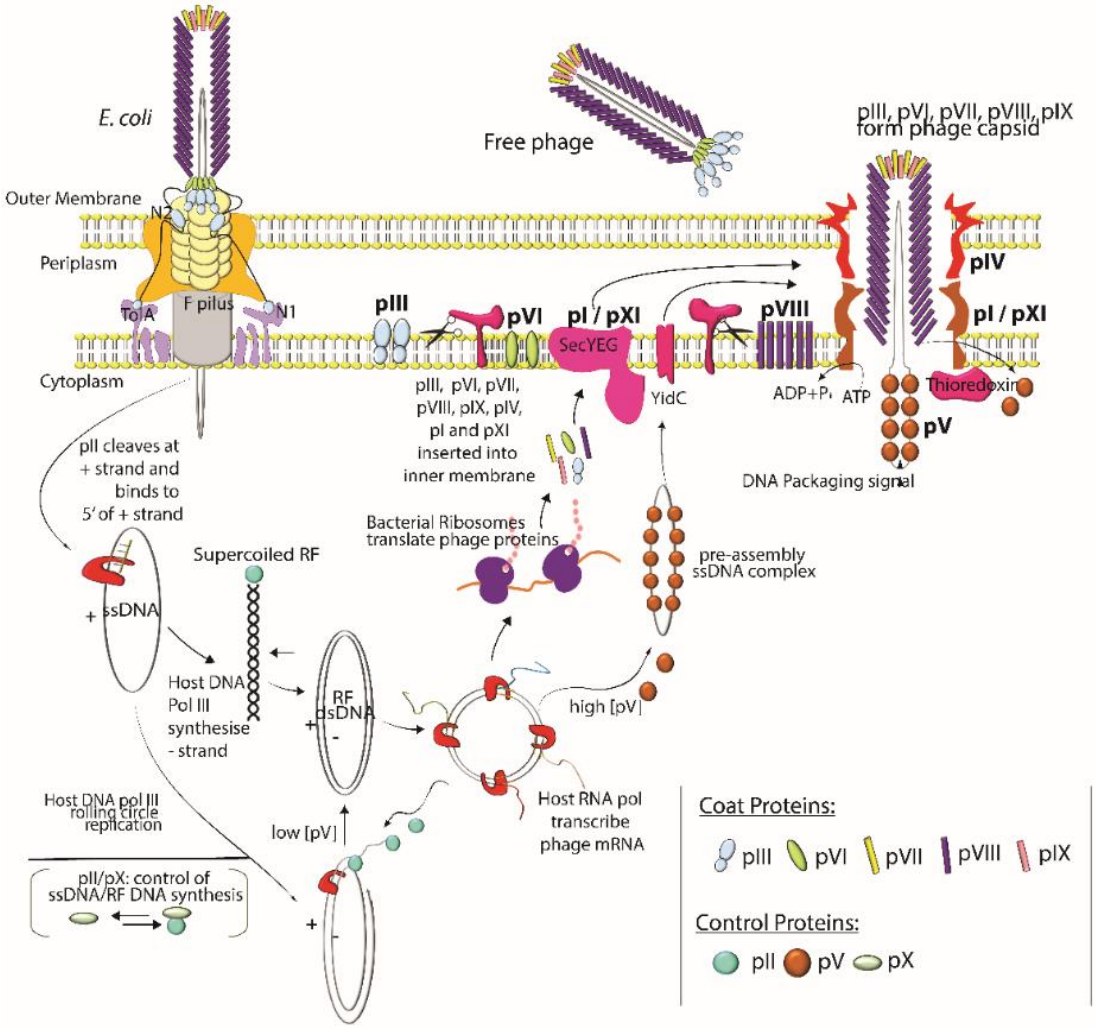

Figure 2.2 Life cycle of M13 filamentous phage. The M13 life cycle begin with the phage genome entering a host cell facilitated by coat protein pIII. Typical stages of a viral life cycle include infection, replication of the viral genome, the assembly of new viral particles, and the subsequent release of the progeny.

The life cycle of M13 filamentous phage starts with the phage binding to the host by attaching to the F-pilus of *E. coli*, facilitated by the coat protein pIII (Marvin & Hohn, 1969; Jazwinski et al., 1973). The pIII pre-protein comprises a signal sequence and three domains, designated as N1, N2, and C with glycine-rich linkers separating them (Armstrong et al., 1981; Stengele et al., 1990). During penetration, the N2 domain of the minor coat protein pIII recognizes and binds to the receptor located at the tip of the F pilus of *E. coli* while the N1 domain interacts with the inner membrane protein TolA which functions as a co-receptor (Riechmann & Holliger, 1997). The C domain of pIII serves the dual role of anchoring the protein within the capsid and facilitating the phage particle's release from the host cell (Riechmann & Holliger, 1997). Following adsorption, the phage releases its single-stranded genomic DNA into the host cytoplasm and sheds its major coat protein, pVIII (Skalka, 1977; Dotto et al., 1984). The host's machinery synthesizes the complementary DNA strand of the genome which can either integrate into the host's genome or exist as plasmid-like vectors. In the latter case, host-encoded enzymes directly convert the phage ssDNA into double-stranded Replicative Form (RF) (Rakonjac et al., 2017) in order to initiate the transcription of mRNAs for phage protein synthesis. The RF DNA can be propagated with the phage protein pV binds to and protects the phage ssDNA from nuclease, resulting in its transformation into a flexible rod to form the precursor to the extracellular virion. Approximately 750 dimers of pV bind to the DNA, forming a protective structure except for a short section known as the packaging sequence (PS) which is essential for assembly and extrusion (Russel, 1991). The pV-bound ssDNA is recognized by a membrane-spanning phage assembly complex which consists of proteins pI, pIV, and pXI. This complex then attaches minor coat proteins, pVII and pIX and facilitates the passage of the ssDNA through the cell membranes. As this

process unfolds, pVIII proteins are assembled around the ssDNA in order to encapsulate it within the virion. Additional minor coat proteins of pIII and pVI, are added before the phage particle is released from the cell. Two phage-encoded protein, namely pIV and pI together with pXI are essential for assembly and extrusion of the bacteriophage. The pIV forms a pore in the outer membrane while pI and pXI are inserted into the inner membrane through the translocase SecYEG (Russel, 1993; Haigh & Webster, 1999). Forming a trans-envelope membrane complex, these proteins contribute to the stabilization of phage assembly. Specifically, pI is involved in an active phage extrusion process and has the potential to interact with a host factor known as thioredoxin. This complex process is energy-consuming and requiring ATP for phage assembly, extrusion, and coat protein synthesis (Stopar et al., 1996). The assembly of this complex hinders cell growth, potentially attributed to the creation of a pore that allows the passage of protons or ions, disrupting the proton-motive force. Furthermore, the precursor protein pVIII is synthesized and then inserted into the membrane with the assistance of the insertase YidC (Feng et al., 1997; Samuelson et al., 2001). Following insertion, signal peptidase I removes the signal peptide, and pVIII oligomers may bind to the pI–pXI complex (Kuhn & Wickner, 1985).

The assembly process concludes by capping the phage with proteins pIII and pVI. The mature virion is then released into the environment. This termination of phage assembly occurs in two steps. Initially, the process begins with the recruitment and incorporation of pIII and pVI, which is then followed by the release of the phage from the host membrane by pIII. The entire process is a coordinated and energy-intensive cycle (Rakonjac et al., 1999).

2.2.2 Phage display selection

Antibody display technologies allow the successful isolation of antigen-specific antibodies (Hoogenboom, 2005). In this context, display technologies such as phage display (McCafferty et al., 1990), bacterial display (Francisco et al., 1993), yeast display (Boder & Wittrup, 1997), ribosome display (Hanes & Plückthun, 1997; He & Taussig, 1997), mRNA display (Xu et al., 2002) and mammalian display (Higuchi et al., 1997) have been widely used as an alternative to hybridoma technology (Köhler & Milstein, 1975) over the past three decades (Hoogenboom, 2005). The phage display of antibodies is the most widely used although bacteria or yeast display have been implemented successfully (Löfblom et al., 2005). In addition to antibodies, various molecules including peptides (Smith & Petrenko, 1997) and T-cell receptors (Onda et al., 1995) are also displayed using phage display technologies for the purpose of molecule isolation against specific targets.

Phage display is a platform being commonly applied for the generation of antibodies (Barbas et al., 1991; Hoogenboom et al., 1991). It exploits the natural replication cycle of bacteriophage to display a particular peptide or protein to the surface of either pIII or pVIII coat protein for selection purpose (Smith, 1985; Oh et al., 2007; Cho et al., 2012). The phage display technology is preferred as a rapid and cost-effective method for isolating molecules (Kay et al., 2001) such as antibody and TCR. In addition, phage display method is a more ethical approach as no animal is required along the process as compared to hybridoma technology (Köhler & Milstein, 1975; Gray et al., 2016b, 2016a; Gray et al., 2020). Furthermore, phage display offers a versatile and highly efficient platform for the display of customized antibody libraries encompassing a wide array of antibody variant (Sblattero & Bradbury, 2000). This capability facilitates rapid and high-throughput screening of these libraries

against specific targets, including peptides (Griffiths et al., 1993), haptens (Charlton et al., 2001; Moghaddam et al., 2001; Kramer, 2002; Strachan et al., 2002) and proteins (Sheets et al., 1998). Typically, the production of recombinant antibodies follows a sequence of steps, starting with the creation of an antibody library. The next phase involves the selection of phage-displayed antibodies that specifically bind to particular antigens through a procedure known as panning, ultimately leading to the isolation of monospecific antibodies. These isolated antibodies are then produced using recombinant expression systems. Phage-derived recombinant antibodies may initially have lower affinities, but additional affinity maturation can be introduced to enhance their binding strength (Larimer et al., 2014).

Phage display of TCR is normally applied to improve the affinity of TCR towards the specific peptide major histocompatibility complex (pMHC) target (Li et al., 2005; Campillo-Davo et al., 2020). In addition to affinity maturation, phage display was also used to improve the stability of TCR through phage display (Gunnarsen et al., 2013) and to isolate recombinant TCR against a specific target. In order to isolate recombinant TCR, a TCR phage display library has to be first constructed for the biopanning selection.

2.2.2(a) T-cell receptor phage display library

A phage library that contains a diverse population of TCRs is a crucial component in TCR phage display experiments (Li et al., 2005). Either phage or phagemid vector system can be used for the cloning of TCR libraries (Løset et al., 2007). The phage system involves the integration of the TCR sequences into the phage genome, most commonly with the *gIII* or *gVIII* coat protein gene whereas the

phagemid system introduces the *gIII* or *gVIII* fused TCR gene with the assistance from a helper phage (Løset et al., 2015).

A naïve TCR library refers to a collection of native TCRs repertoire isolated from healthy individual T cells that have not been exposed to specific antigens (Macedo et al., 2009; Tian et al., 2022). The naïve TCR libraries are developed by cloning diverse sequence repertoires into suitable vectors. TCR phage display libraries, including scTCR, aim to enhance the affinity and stability of existing TCRs against specific targets. Similar to immunoglobulins, TCR α and β chains comprise variable (V) and constant C regions. To obtain TCR sequences for cloning, T-cell sequences are amplified from isolated T-cells. Peripheral blood mononuclear cells (PBMCs) are first obtained from human blood and their RNA is extracted for TCR usage analysis. An increasing number of *TCRV* gene sequences in public databases have led to the design of new *TCRV* gene primers (Ch'ng et al., 2019) for more comprehensive cloning.

Alternatively, TCRs can be chemically synthesized based on reported sequences in databases. Mutations can be introduced into scTCR sequences to create a "mini library" of scTCR mutants for screening using site-directed mutagenesis (Gunnarsen et al., 2013; Løset et al., 2015). These mutations are often focused on the CDR3 region of the Vα and Vβ domains (Holler et al., 2000). Directed evolution of TCR sequences using these methods allows for the creation of mutant libraries with a high degree of diversity, suitable for screening to improve TCR properties such as solubility, stability, and affinity maturation.

Generally, the type of TCR library is either naïve (de Greef et al., 2020; Li et al., 2020) or semi-synthetic TCR library (Løset et al., 2015) which are mainly known as combinatorial TCR libraries. These libraries enable the shuffling of alpha and beta

chain pairs to create a diverse pool of the TCR repertoire. The size limitation of phage display meant that full TCR molecules are not efficiently displayed. This meant a smaller scTCR format is preferred for the phage library construction (Løset et al., 2015).

2.2.2(b) Phage display system

An understanding of the proteins in the M13 genome has significantly influenced the design of phage display vectors. These vectors predominantly depend on the use of pIII or pVIII as anchor proteins to display diverse polypeptides on the surface of M13 phage (Smith, 1985; McCafferty et al., 1990). The choice of coat protein significantly impacts the valency of the polypeptide presentation in which pIII with a lower copy number favours monovalent presentation while pVIII with higher copy numbers enables multivalent presentation to enhance the avidity effect (Molek et al., 2011).

There are two main methods of either phage vector system or phagemid system to achieve this display mechanism on phage surfaces (Soltes et al., 2003). A phage vector system enables the expression of scTCR as a fusion protein with the pIII coat protein, governed by the natural phage promoter, and it does not necessitate the use of a helper phage for phage packaging. In contrast, the phagemid system necessitates the use of an artificial promoter, such as the *lac* promoter for scTCR-pIII fusion protein expression and a helper phage is required for phage packaging (O'Connell et al., 2002). However, because phagemid vectors lack a phage genome, achieving complete phage packaging is only possible with the assistance of helper phages, which introduce competition between wild-type pIII and mutant pIII (Rondot et al., 2001; O'Connell et

al., 2002). This difference in vector design also affects display efficiency; phage systems offer polyvalent antibody display on pIII while phagemid systems allow only monovalent presentation (Jaroszewicz et al., 2021). The isolation of scTCR-presenting phages after they bind to a target antigen simplifies the process of identifying clones through standard sequencing methods (Li et al., 2020). This approach has been used to introduce various scTCR fragment sequences into the phagemid vector, creating a collection of distinct clones known as a scTCR library in this project.

The *in-vitro* selection system of phage display libraries is classified based on the chosen coat protein for library design, with pIII and pVIII being common options known as Type 3 and Type 8 vectors, respectively. In this design, the scTCR gene insert is displayed either on all copies of pIII or pVIII (Petrenko, 2018). However, modified phage vectors known as 33 and 88 systems contain both wild type and recombinant pIII or pVIII genes (Ebrahimizadeh & Rajabibazl, 2014). In this system, only the fusion pIII or pVIII phages display scTCR on the bacteriophage surface (Petrenko, 2018; Barderas & Benito-Peña, 2019) and preserve the phage re-infectivity while displaying the proteins in an immunologically accessible form. The 3+3 system and 8+8 system involve concurrent use of a phagemid and a helper phage where the foreign DNA fragments are fused to the *gIII* gene in the phagemid and the helper phage supplies additional *gIII* for phage packaging (Wang & Yu, 2004; Petrenko, 2018). The 3+3 system is a relatively common display setup used in many instances involving antibodies and TCR display.

Phagemid vectors function with the assistance of helper phages in order to replicate like bacterial plasmids before being packaged as single-stranded DNA in viral particles. These vectors are designed to include both an origin of replication and f1 origin, facilitating both double-stranded replication and single-stranded DNA

synthesis before being packaged into phage particles. The scTCR gene fragments and *gIII* coat protein gene are linked together after scTCR gene fragments are cloned into the multiple cloning sites of the phagemid (Primrose & Twyman, 2006). Additionally, a signal leader peptide such as PelB is introduced for transmembrane transport of the scTCR fusion protein during virus particle assembly (Barbas et al., 1991). However, as phagemid vectors lack the necessary genes for virus replication and packaging, a helper phage like M13KO7 or Hyperphage is required (Qi et al., 2012).

M13KO7 is a commonly used helper phage that provides the viral components needed for single-stranded DNA replication and packaging of the phagemid DNA into phage particles (Vieira & Messing, 1989). Typically, helper phages are engineered to package less efficiently than phagemids, reducing the packaging of wild-type helper phages and allowing the phagemid to predominate (Paschke, 2006; Qi et al., 2012). Filamentous phage infection relies on the presence of a pilus, which allows F-plasmid bacterial hosts to generate phage particles (Qi et al., 2012). M13KO7, due to its weaker packaging signal, still allows for the presentation of wild-type pIII during phage assembly, resulting in a monovalent display (Løset et al., 2007). In contrast, some helper phage systems delete the pIII encoding gene from the phage genome, enabling five copies of scTCR fused pIII to be displayed on each phage particle, resulting in a polyvalent display configuration (Kramer et al., 2003).

2.2.2(c) Phage display biopanning

The process of isolating specific and valuable binders from a collection of clones is often referred to as panning which draws parallels with the ancient Roman gold panning technique. Both involve a continuous sieving process, aiming to separate and enrich the targeted antibodies or scTCR from a diverse pool of antibody or scTCR

sequences based on their affinity for a target (McCafferty et al., 1994; Onda et al., 1995) as shown in Figure 2.3.

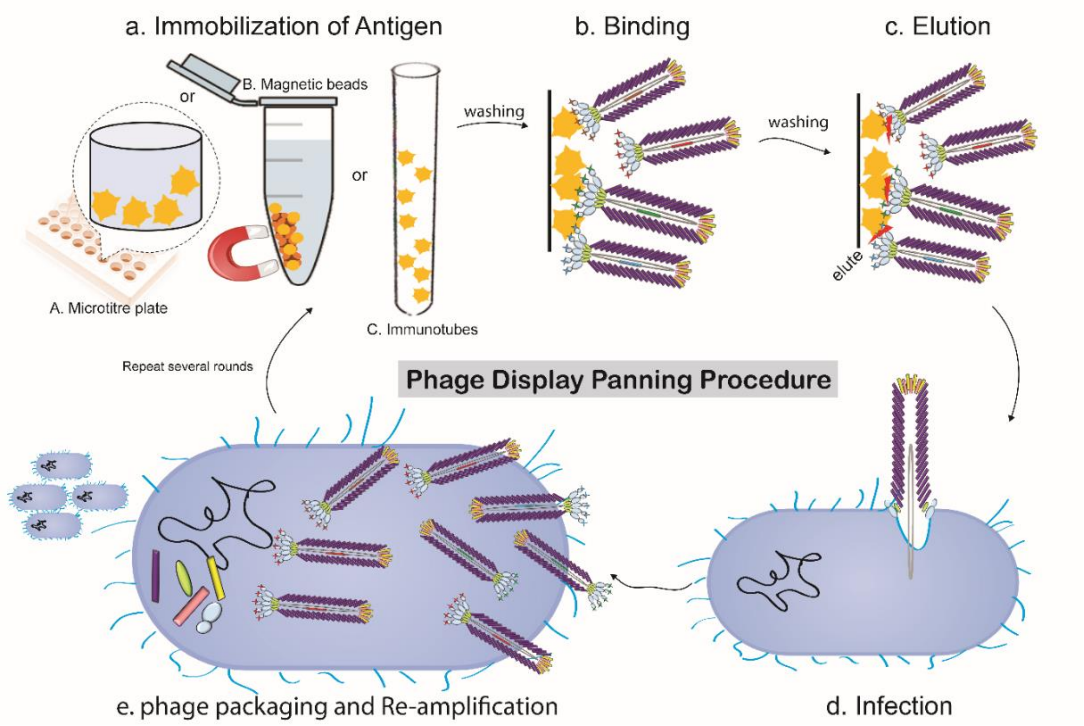

Figure 2.3 Phage display panning schematic diagram. The process of phage display panning begins with immobilizing the protein on a solid support. Subsequently, the phage display library is applied to the immobilized protein to facilitate the binding of specific variants. Multiple rounds of washing are conducted to eliminate non-binders before eluting and re-amplifying the remaining bound phages.

The aim of such panning campaigns is to isolate high affinity scTCR against recombinant antigen proteins immobilized on solid surfaces such as magnetic beads (Nilsen, 2014), microtiter plates (Nilsen, 2014), immunotubes (Gunnarsen et al., 2013)or immunoassay tips (Chin et al., 2016). The biopanning procedure typically begins with washing the solid surface, followed by blocking with a blocking agent. Simultaneously, a pre-blocked immunotube is used to pre-incubate the phage library

to deplete background binding before applying it to the blocked antigen-coated surface. Unbound scTCR phages are then removed through a series of washing cycles (Onda et al., 1995). The remaining scTCR phages that are bound to the target are then eluted from the solid surface and recovered through reinfection using *E. coli* containing F-pilus. To propagate more target-specific phages, the infected *E. coli* is cultured until reaching an OD_{600nm} of 0.5 before being packaged into phages with the assistance of the M13KO7 helper phage. This iterative process can be repeated for several rounds until a enriched pool of binding scTCR is identified at the polyclonal level (Onda et al., 1995).

To enhance the isolation of high affinity binders, optimization can be employed on the washing condition and reagent. This include increasing the number of washing cycles in subsequent panning rounds or using different washing reagent such as phosphate-buffered saline (PBS), PBS with 0.1 % Tween 20 (PBST), PBS with milk (MPBS) or combination of these buffers. Tris buffered saline is also commonly employed as a washing reagent (Pande et al., 2010; Thie, 2010; Rahbarnia et al., 2016). Additionally, the elution condition of target-bound phage particles can be fine-tuned with options such as 10 μg/mL of trypsin, 100 mM Triethylamine pH 11 or 0.1 M Tris-glycine pH 2.2. Elution can also be achieved using the target antigen itself (Lee et al., 2007).

In addition to the conventional plate panning method, a semi-automated magnetic beads panning technique has also been introduced. In this approach, automation is used to replace human involvement in the processes such as blocking, affinity binding and most importantly washing before eluting the target specific scTCR phages. These eluted phages are then transferred into *E. coli* harbouring F-pili for infection. However, it is worth noting that the phage packaging and amplification

processes are still performed manually after removing the infected *E. coli* from the KingFisher™ Flex processor (Ch'ng et al., 2023). After confirming the polyclonal results of the biopanning selection, the enriched phages are used for monoclonal selection with the aim of isolating a single clone. Monoclonal phage ELISA is then conducted to verify the binding efficiency of the scTCR clones before confirming their sequences using Sanger sequencing service (Camacho et al., 2013).

2.3 Antigen binding receptors of adaptive immunity

2.3.1 Immune system

The human immune system is a complex defence system comprising various organs, cells, and proteins. Its primary role is to protect the body from harmful invaders like bacteria, viruses, parasites, and fungi. It also detects and neutralizes dangerous substances from the environment and combats disease-related challenges, including cancerous cells. This defence system is triggered by foreign substances called antigens, which bind to T-cell receptors (TCRs) or antibodies, initiating a series of responses.

There are three major immune responses, the external barriers (including physical and chemical), innate and adaptive immune responses (Moser & Leo, 2010). The innate immune system acts as the first line of defence. It provides immediate, nonspecific protection against pathogens but lacks immunologic memory. It employs immune cells like natural killer cells and phagocytes and is crucial in countering harmful substances that enter the body through various routes (Roos et al., 2005).

If the innate immune system cannot eliminate pathogens, the adaptive immune system takes over. It produces antibodies specifically tailored to combat previously encountered pathogens. This "learned" immunity adapts over time to fight evolving

bacteria or viruses. Adaptive immunity is antigen-specific and dependent, resulting in a delay between antigen exposure and a full response. Importantly, it forms a memory, enabling a quicker and more effective immune response upon future encounters with the same antigen (Sebina & Pepper, 2018).

2.3.2 The lymphocytes and their roles in the immune system

T-cells and B-cells are fundamental components of the human immune system. They work together to safeguard the body against invading pathogens. T-cells, including cytotoxic T-cells (CD8+ T cells), helper T-cells (CD4+ T cells), and regulatory T-cells (Tregs) form a diverse army with specialized functions. These T-cells express an antigen-specific receptors known as T-cell receptor (TCR) on their cell surface to detect the antigens displayed on carriers such as major histocompatibility complex (MHC) and CD1. Specifically, CD8+ T-cells recognize peptides displayed on the MHC I and eliminate infected or cancerous cells directly while CD4+ T cells act as conductors by regulating the adaptive immune responses. They achieve this by releasing cytokines to activate other immune cells when they detect peptide antigens displayed on MHC II. Tregs play a vital role in maintaining immune balance by suppressing excessive responses to prevent autoimmune reactions (Janeway et al., 2008).

Another crucial component of the immune system is B cells, responsible for adaptive immunity. When B cells encounter antigens, they transform into plasma cells which produce B-cell receptors or soluble form antibodies. These antibodies are highly specific and bind to antigens to neutralize pathogens. Additionally, memory B cells are

formed during infections to enable faster and more effective responses upon reinfection (Punt & Owen, 2013).

The synergy between T cells and B cells is pivotal in immune defence. The CD4+ T cells can enhance B cell antibody production and CD8+ T cell production while CD8+ T cells promptly eliminate infected cells. Memory B cells and the support from memory T cells help to establish immunological memory, ensuring a swifter and more potent response to previously encountered pathogens (Janeway, 2001; Punt & Owen, 2013).

2.3.3 Antigen recognition of receptors on T and B lymphocytes

T-cell receptor and B-cell receptor are the antigen recognition sites found on the surface of T cell and B cells, respectively as shown in Figure 2.4.

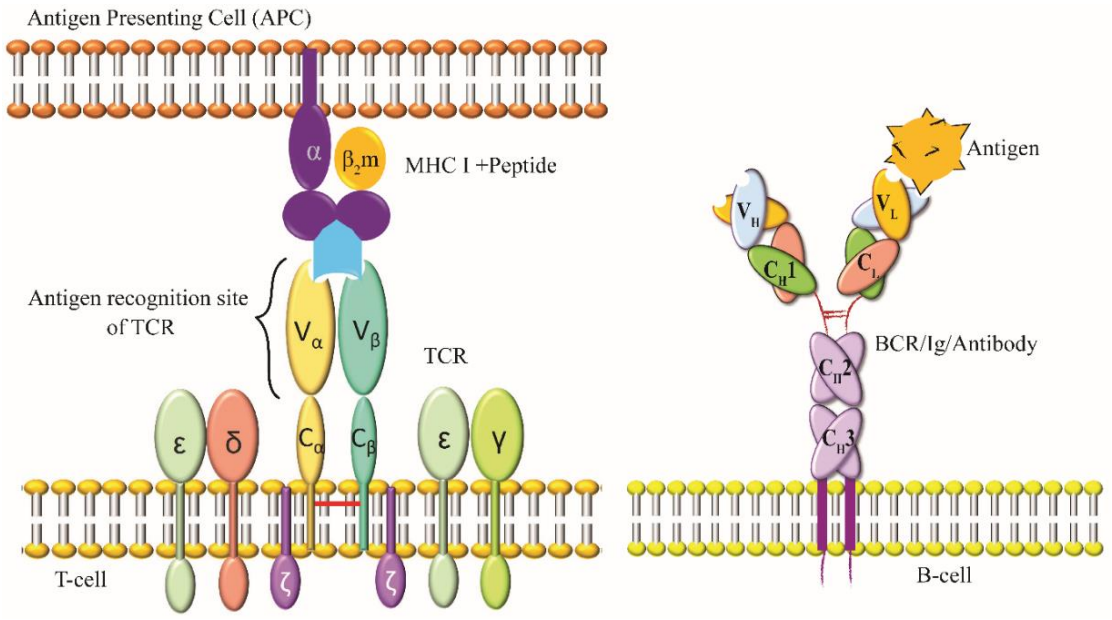

Figure 2.4 Antigen recognition of TCR and BCR on lymphocytes. Both TCR and BCR are presented on the surface of T cell and B cells, respectively. TCR bind to antigen complex whereas BCR binds to antigen.

Both receptors share similarities in their structure that are produced by the gene rearrangement. Furthermore, each receptor is capable of expressing a single species of receptor against an antigen.

TCRs on T-cells play a vital role in recognizing antigens presented by a variety of antigen presenting cells (APCs), including dendritic cells, macrophages, B cells, fibroblasts and epithelial cells (Neefjes et al., 2011). Specifically, TCRs bind to short peptide fragments displayed on the surface of MHC molecules located on APCs. MHCs are divided into two main classes: class I molecules found on the surface of all nucleated cells, and class II molecules (also referred as HLA DP, DQ and DR) primarily presented on macrophages, dendritic cells and B cells (Rock et al., 2016). MHCs showcase antigen fragments in peptide form of either from intracellular pathogens like viruses or from the phagocytosis of foreign proteins or organisms (Neefjes et al., 2011). Class I MHCs display endogenous peptides while class II MHC class II present exogenous peptides to T cells (Vyas et al., 2008).

The interaction between the antigen complex on APCs and TCRs triggers a series of events, prompting the respective cell to secrete interleukins or antibodies to counteract the antigen before it can infiltrate host cells. To increase the likelihood of T cells encountering APCs with suitable peptide-MHC complexes, T cells circulate throughout the body via the lymphatic system and bloodstream. Additionally, the accumulation of both T cells and APCs in lymph nodes facilitates these interactions (Knapp et al., 2019).

It is worth noting that TCRs can interact not only with peptide-MHC complexes but also non-peptide antigens like lipids mediated by some of the five isoforms of CD1 (a-e) (Mori & De Libero, 2012). Furthermore, research has shown