

***IN SILICO* DESIGN AND CONSTANT-PH  
MOLECULAR DYNAMICS STUDY OF HUMAN  
IGG1 FC AT PH6.0 AND PH 7.5**

**LIM YEE YING**

**UNIVERSITI SAINS MALAYSIA**

**2020**

***IN SILICO* DESIGN AND CONSTANT-PH  
MOLECULAR DYNAMICS STUDY OF HUMAN  
IGG1 FC AT PH6.0 AND PH 7.5**

by

**LIM YEE YING**

**Thesis submitted in fulfilment of the requirements  
for the degree of  
Master of Science**

**June 2020**

## **ACKNOWLEDGEMENT**

I as the author would like to express my utmost gratitude towards my parents for being very supportive throughout the study. Besides, I cannot thank Assoc. Prof. Dr. Choong Yee Siew enough for her invaluable guidance and encouragement throughout the duration of the master's program. Appreciation is also given to Assoc. Prof. Dr. Lim Theam Soon for the additional guidance and assistance in antibody expression related studies and to my colleague in Structural Biology Lab, Ms. Soong Jia Xin for spending time and patience in guiding and supporting me in the completion of the project, together with all staffs and friends in INFORMM, Universiti Sains Malaysia to share knowledge, advices and guidance throughout my study. I am grateful to receive support from Universiti Sains Malaysia through USM fellowship. This work is supported by FRGS (203/CIPPM/6711680) from the Malaysia Ministry of Education.

## TABLE OF CONTENTS

<b>ACKNOWLEDGEMENT .....</b>	<b>ii</b>
<b>TABLE OF CONTENTS.....</b>	<b>iii</b>
<b>LIST OF TABLES .....</b>	<b>vi</b>
<b>LIST OF FIGURES .....</b>	<b>vii</b>
<b>LIST OF SYMBOLS .....</b>	<b>xi</b>
<b>LIST OF ABBREVIATIONS .....</b>	<b>xii</b>
<b>LIST OF APPENDICES .....</b>	<b>xiv</b>
<b>ABSTRAK .....</b>	<b>xv</b>
<b>ABSTRACT.....</b>	<b>xvii</b>
<b>CHAPTER 1 INTRODUCTION .....</b>	<b>1</b>
1.1 Problem statement.....	1
1.2 Study objectives and expected outcome .....	2
1.3 Thesis outline.....	2
<b>CHAPTER 2 LITERATURE REVIEW .....</b>	<b>3</b>
2.1 Antibody diversity .....	3
2.2 Antibody structure .....	4
2.2.1 Fragment of crystallizable (Fc) region.....	5
2.2.2 Fc-fusion protein.....	7
2.3 Therapeutic antibodies.....	7
2.4 Neonatal Fc receptor (FcRn).....	10
2.4.1 IgG-FcRn binding mechanism.....	10
2.4.2 Fc-engineering .....	12

2.5	Computational approach in molecular studies.....	15
2.5.1	Molecular dynamics (MD) simulation.....	16
2.5.2	Influence of environmental pH on titratable residues.....	17
2.5.3	Constant pH molecular dynamics (CpHMD) simulations.....	19
2.5.4	Assessing Fc-FcRn complex binding affinity via binding free energy calculation.....	21
<b>CHAPTER 3 METHODOLOGY .....</b>		<b>23</b>
3.1	The Fc-FcRn complex structure.....	23
3.2	Constant pH molecular dynamics (CpHMD) simulation.....	25
3.2.1	Minimization.....	27
3.2.2	Heating.....	27
3.2.3	Equilibration .....	27
3.2.4	Production.....	27
3.3	<i>In silico</i> design of Fc variants and constant pH Molecular Dynamics (CpHMD) simulation of Fc variants in complexed with FcRn.....	28
3.4	Analysis.....	29
3.4.1	MM-GBSA binding free energy calculation .....	31
<b>CHAPTER 4 RESULTS.....</b>		<b>34</b>
4.1	Fc-FcRn complex crystal structure analysis .....	34
4.2	Design of Mut <sub>M4</sub> .....	37
4.3	CpHMD simulation system stability.....	39
4.4	Protein backbone root mean square deviation (RMSD) .....	39
4.4.1	Fc-FcRn complex clustering.....	43
4.5	Fc-FcRn binding free energy analysis .....	48
4.6	The effects of residue protonation states on complex binding affinity.....	50

<b>CHAPTER 5</b>	<b>DISCUSSION.....</b>	<b>76</b>
<b>CHAPTER 6</b>	<b>SUMMARY .....</b>	<b>84</b>
6.1	Conclusion remarks .....	84
6.2	Limitations of the study .....	85
6.3	Suggestions for future work.....	85
<b>REFERENCES.....</b>		<b>87</b>
<b>APPENDICES</b>		
<b>LIST OF PUBLICATIONS</b>		

## LIST OF TABLES

	<b>Page</b>
Table 2.1	Multiple sequence alignment of Ab IgG Fc subclass isotypes with (Genebank ID -IgG1: AAA02914.1; IgG2: AXN93665.1; IgG3: AXN93659.2; IgG4: AAB59394.1).....6
Table 4.1	The pKa value and protonation fraction of crystal structure (PDB_ID: 4N0U) predicted using DelPhiPka web server. ....38
Table 4.2	Protein sequence of Homo sapien IgG1 Fc variants at CH <sub>2</sub> and CH <sub>3</sub> region in this work. ....40
Table 4.3	MM-GBSA binding free energy (kcal/mol) analysis for Fc-FcRn complexes simulated at pH 6.0 and 7.5. ....49
Table 4.4	Fc-FcRn MM-GBSA pair-wise decomposition binding free energy (kcal/mol). $\Delta G_{\text{Bind}}$ less than -1.5 kcal/mol are considered as the residue that shows consistent high binding affinity within complex in the average of 250 frames MMGBSA calculation and they were highlighted in <b>bold</b> . Contrary, positive value of $\Delta G_{\text{Bind}}$ are residue that shows repulsion in Fc-FcRn complex. ....52
Table 4.5	The pKa prediction of titratable residues in Fc-FcRn complex. Residues that do not incur in any protonation transition during simulation resulted in a (-∞) symbol. Residues labelled as (null) are mutated residues that are no longer titratable residues. The protonation states are labelled using numbers with Asp containing state (0 to 4), Glu containing state (0 to 4) and His containing state (0 to 2). The structural details for protonation states are stated in Appendix D. ....69
Table 4.6	Hydrogen bond formation between Fc and FcRn at pH 6.0 and 7.5. Only hydrogen bonds occurred > 50% are reported. Bold residues are titratable residues. ....72

## LIST OF FIGURES

		<b>Page</b>
Figure 2.1	The structure of an Ab Fc region. Fab regions are in green while the Fc region is in blue. The white sticks are glycans branched out from Asn297 (PDB-code: 1HZH (Saphire et al., 2001). Structure image was produced with PyMOL program (DeLano, 2002). .....	6
Figure 2.2	Schematic diagram of IgG scaffold being applied in Fc-fusion construction. The IgG domains were coloured in blue while proteinaceous compound in Fc-fusion protein were coloured in yellow. ....	9
Figure 2.3	Schematic diagram of murine, chimeric, humanized and fully human antibodies. The murine Abs were coloured in red while human Abs were coloured green. Figure adapted from (Chames et al., 2009). .....	9
Figure 2.4	The structure of FcRn. HC is in pink (Domain $\alpha$ 1-3) and LC ( $\beta$ 2m) is in blue. (PDB-code: 1FRT (Burmeister et al., 1994), Figure was produced using PyMOL program (DeLano, 2002). .....	11
Figure 2.5	Schematic diagram of FcRn-dependent IgG rescue pathway. ....	13
Figure 3.1	The methodology flow chart .....	26
Figure 4.1	Fc-FcRn complex structure presented with Fc (yellow), FcRn heavy chain (pink) and FcRn light chain ( $\beta$ 2-microglobulin, $\beta$ 2m). The Interface analysis for: (a) full Fc-FcRn complex, (b,c) Fc-FcRn light chain (Fc- $\beta$ 2m) complex and its interaction radar at binding interface, (d,e) Fc-FcRn heavy chain complex and its interaction radar at binding interface. The interaction parameter in (c,e) includes: Interface area, $\text{\AA}^2$ (IA); solvation energy, kcal/mol (DG); binding energy, kcal/mol (BE); hydrophobicity, kcal/mol (PV); number of hydrogen bonds (HB); number of salt bridges (SB); number of disulfide bonds (DS).	

	Figure was prepared using PyMOL (DeLano, 2002) and with interface radar interpreted using jsPISAv2.0.5 (Krissinel & Henrick, 2007). ....	35
Figure 4.2	Fc-FcRn complex structure zoomed in interacting interface with Fc (yellow), Fc interacting interface (dark pink), FcRn heavy chain (light pink), FcRn interacting interface (blue). Figure was prepared using PyMOL (DeLano, 2002). ....	36
Figure 4.3	Fc-FcRn complex structure with Fc (yellow), FcRn heavy chain (pink) and Cys residues (blue). The disulfide bond between (a) heavy chain FcRn: Cys96 and Cys159, (b) Fc: Cys261 and Cys321, (c) Fc: Cys367 and Cys425. ....	36
Figure 4.4	Hotspot analysis on Mut <sub>YTE</sub> Fc variant Fc-Rn interacting residues were imposed with single amino acid mutation. The results were presented using heatmap in respect to change in binding free energy upon mutation ( $\Delta\Delta G$ ) originate from representative complex structure at pH 6.0 with negative and positive $\Delta\Delta G$ indicating favourable and unfavourable improvement in binding affinity, respectively. ....	38
Figure 4.5	The total energy of CpHMD simulation system stability evaluation at (a) pH 6.0 and (b) pH 7.5 for Fc-FcRn complexes for Fc WT, Mut <sub>AAA</sub> , Mut <sub>YTE</sub> and Mut <sub>M4</sub> , respectively, in complex with FcRn. ....	41
Figure 4.6	Root mean square deviation of the protein backbone atoms (RMSD <sub>C<math>\alpha</math>,N,C,O</sub> ) with the function of time at (a) pH 6.0 and (b) pH 7.5 for Fc WT, Mut <sub>AAA</sub> , Mut <sub>YTE</sub> and Mut <sub>M4</sub> , respectively, in complex with FcRn. ....	42
Figure 4.7	Clustering analysis on the last 5 ns for Fc-FcRn complex trajectories based on protein backbone atom root-mean-square deviation (RMSD <sub>C<math>\alpha</math>,N,C,O</sub> ) with the function of time at (a) pH 6.0 and (b) pH 7.5 for Fc WT, Mut <sub>AAA</sub> , Mut <sub>YTE</sub> and Mut <sub>M4</sub> , respectively, in complex with FcRn. Number of clusters and their respective trajectories allocation were stated around the clustered region. ....	45

Figure 4.8	The superimposition of the MD representative structures at pH 6.0 (A-D) and pH 7.5 (E-H) with the crystal structure (PDB ID 4NOU; Fc: black & FcRn: grey ribbon presentation). The IgG1-Fc (A, E) WT-FcRn, (B, F) Mut <sub>AAA</sub> -FcRn, (C, G) Mut <sub>YTE</sub> -FcRn, and (D, H) Mut <sub>M4</sub> -FcRn complexes were represented as Fc: cyan & FcRn: pink ribbon presentation at pH 6 and Fc: yellow & FcRn: orange ribbon presentation at pH 7.5. The red arrow represents the quaternary structure deviation between the superimposed complexes .....46
Figure 4.9	The superimposition of <b>(a)</b> WT-FcRn, <b>(b)</b> Mut <sub>AAA</sub> -FcRn, <b>(c)</b> Mut <sub>YTE</sub> -FcRn, and <b>(d)</b> Mut <sub>M4</sub> -FcRn complexes at pH 6.0 (Fc: cyan and FcRn: pink ribbon presentation) and pH 7.5 (Fc: yellow & FcRn: orange ribbon presentation). The red arrow represents the quaternary structure deviation between the superimposed complexes. ....47
Figure 4.10	Total binding energy ( $\Delta G_{\text{Bind}}$ ) bar chart of Fc-FcRn complex.....49
Figure 4.11	The energy contribution denoted as hydrophilic ( $\Delta E_{\text{elec+polar}}$ ) and hydrophobic ( $\Delta E_{\text{vdw+np}}$ ) interaction of Fc-FcRn complex.....51
Figure 4.12	The electrostatic interaction of <b>(a)</b> WT, <b>(b)</b> Mut <sub>AAA</sub> , <b>(c)</b> Mut <sub>YTE</sub> , and <b>(d)</b> Mut <sub>M4</sub> with FcRn at pH 6.0. IgG1-Fc is positioned at the right (cyan) and FcRn is positioned at the left (pink). Residues involved in electrostatics interaction are in red and blue stick presentation for Fc and FcRn, respectively. The Fc residues names are underlined while FcRn residues names are without underline. Yellow dots labelled with (H) are denote the hydrogen bonding between Fc and FcRn.....73
Figure 4.13	The hydrophobic interaction of <b>(a)</b> WT, <b>(b)</b> Mut <sub>AAA</sub> , <b>(c)</b> Mut <sub>YTE</sub> , and <b>(d)</b> Mut <sub>M4</sub> with FcRn at pH 6.0. IgG1-Fc is positioned at the right (cyan) and FcRn is positioned at the left (pink). Residues involved in hydrophobic interaction are in yellow and white ball presentation for Fc and FcRn, respectively. The Fc residues names are underlined while FcRn residues names are without underline.....74

Figure 4.14 The hydrophobic interaction of **(a)** WT, **(b)** Mut<sub>AAA</sub>, **(c)** Mut<sub>YTE</sub>, and **(d)** Mut<sub>M4</sub> with FcRn at pH 7.5. IgG1-Fc is positioned at the right (yellow) and FcRn is positioned at the left (orange). Residues involved in hydrophobic interaction are in turquoise and white ball presentation for Fc and FcRn, respectively. The Fc residues names are underlined while FcRn residues names are without underline.....75

## LIST OF SYMBOLS

$\Delta G$	Change in absolute free energy
$\Delta G_{bind}$	Change in binding free energy
AS4	Aspartatate (renamed for dynamic protonation procedures)
C	Carbon
$C_{\alpha}$	Carbon alpha
$\Delta G_{np}$	Contribution of non-polar solvation
$\Delta G_{pol}$	Contribution of polar solvation
$\Delta T_B$	Contribution of residues backbone
$\Delta T_S$	Contribution of residues side chain
$\Delta$	Difference
fs	Femtosecond ( $10^{-15}$ s)
GL4	Glutamate (renamed for dynamic protonation procedures)
HIP	Histidine (renamed for dynamic protonation procedures)
$E_{int}$	Internal energy
ns	Nanosecond ( $10^{-9}$ s)
O	Oxygen
pKa	Ionization constant
ps	Picosecond ( $10^{-12}$ s)
s	Seconds
$S$	Solvation free energy
$\gamma$	Surface tension proportionality
T	Temperature
3D	Three dimensional

## LIST OF ABBREVIATIONS

Ab	Antibody
Ala	Alanine
ADCP	Antibody-dependent cell-mediated phagocytosis
ADCC	Antibody-dependent cellular cytotoxicity
Ag	Antigen
Arg	Arginine
Asn	Asparagine
Asp	Aspartate
$\beta$ 2m	beta-2-microglobulin
CDC	Complementary-dependent cytotoxicity
CH	Constant heavy
CpHMD	Constant pH molecular dynamics
Cys	Cysteine
FDA	Food and Drug Administration
Fc	Fragment crystallizable
FcRn	Neonatal fragment crystallizable receptor
Fab	Fragment of antigen binding
Glu	Glutamate
Gln	Glutamine
Gly	Glycine
HC	Heavy chain
His	Histidine
HAMA	Human anti-mouse antibody
Ig	Immunoglobulin
Ile	Isoleucine
Leu	Leucine
LC	Light chain
LCPO	linear combination of pairwise overlap
Lys	Lysine
MHC	Major histocompatibility complex
Met	Methionine
MD	Molecular dynamics
MM-GBSA	Molecular Mechanics/Generalized Born Surface Area
MW	Molecular weight
mAb	Monoclonal antibody
MC	Monte Carlo

Mut <sub>AAA</sub>	Mutant AAA (Fc mutation: Ile252Ala/His310/His435Ala)
Mut <sub>M4</sub>	Mutant M4 (Fc mutation: Asp294Ala/Thr254Phe/Pro257Trp/Asp312Gly)
Mut <sub>YTE</sub>	Mutant YTE (Fc mutation: Met252Tyr/ Ser254Thr/Thr256Glu)
N	Nitrogen
NMR	Nuclear magnetic resonance
Phe	Phenylalanine
Pro	Proline
RMSD	Root mean square deviation
Ser	Serine
scFv	Single-chain variable fragment
sdAb	Single-domain antibody
SASA	Solvent accessible surface area
Thr	Threonine
Trp	Tryptophan
Tyr	Tyrosine
Val	Valine
vdW	van der Waals
WT	wildtype

## LIST OF APPENDICES

- Appendix A Topology file preparation
- Appendix B CpHMD simulation input files
- Appendix C *cpout* output file.
- Appendix D Structure of Asp, Glu and His protonation state

## REKA BENTUK *IN SILICO* DAN KAJIAN SIMULASI DINAMIK MOLEKUL

### PH TETAP UNTUK IGG1 FC MANUSIA PADA PH 6.0 DAN PH 7.5

#### ABSTRAK

Penggunaan antibodi monoklonal terhadap pelbagai penyakit seperti kanser, penyakit berjangkit atau gangguan autoimun, telah menjadi laluan utama dalam bidang terapeutik perubatan. Khususnya, IgG1 yang mempunyai separuh hayat serum yang panjang telah menyebabkan penggunaannya semakin popular sebagai perancah dadah terapeutik di mana ciri-ciri ini berkaitan dengan mekanisme pengikatan IgG1 Fc dengan FcRn yang bergantung terhadap pH. Kajian ini bermula dengan penubuhan sistem simulasi dinamik molekul pH tetap (CpHMD) menggunakan kompleks Fc-FcRn yang telah disah daripada eksperimen-eksperimen lain yang merangkumi Fc (jenis liar, Mut<sub>AAA</sub> dan Mut<sub>YTE</sub>) dengan setiap kompleks berikat pada pH 6.0 dan berpisah pada pH 7.5. Kiraan tenaga bebas pengikatan ( $\Delta G_{\text{Bind}}$ ) melalui pendekatan MMGBSA telah digunakan untuk menjelaskan pertalian pengikat bagi kompleks kawalan (jenis liar-, Mut<sub>AAA</sub>- dan Mut<sub>YTE</sub>-FcRn) dan keputusan menunjukkan varian Fc Mut<sub>YTE</sub> IgG1-Fc mempunyai pengikat yang paling kuat terhadap FcRn pada pH 6.0. Untuk mendapatkan reka bentuk alternatif Fc yang mempunyai pengikatan lebih baik terhadap FcRn, satu trajectori pH 6.0 kompleks Mut<sub>YTE</sub>-FcRn telah diekstrak menggunakan program MMTSB dan diteruskan dengan mutasi titik tunggal melalui aplikasi web SAAMBE pada wakil kompleks tersebut. Hasil mutasi mencadangkan empat lokasi Fc yang boleh meningkatkan tenaga pengikat ( $\Delta G$ ) Fc terhadap FcRn. Kesemua cadangan mutasi telah digunakan untuk memaksimumkan kesan pengikatan terhadap FcRn dan menghasilkan reka bentuk varian IgG1 Fc baru, Mut<sub>M4</sub> (Asp294Ala/Thr254Phe/Pro257Trp/Asp312Gly). Ciri-ciri pengikatan kompleks Mut<sub>M4</sub>-

FcRn telah dibandingkan dengan kompleks kawalan dan keputusan menyimpulkan bahawa pengikatan Mut<sub>M4</sub> ( $\Delta G_{\text{Bind}} = -23.19$  kcal/mol) lebih baik berbanding dengan varian Fc kawalan ( $\Delta G_{\text{Bind}} (\text{Mut}_{\text{YTE}}) = -20.67$  kcal/mol (jenis liar) =  $-13.94$  kcal/mol) pada pH 6.0. Kesemua kompleks Fc-FcRn menunjukkan affiniti pengikatan yang rendah pada pH 7.5 ( $\Delta G_{\text{Bind}} (\text{Mut}_{\text{AAA}}) = +37.02$  kcal/mol; (Wildtype) =  $-3.86$  kcal/mol; ( $\text{Mut}_{\text{YTE}}) = -6.77$  kcal/mol; ( $\text{Mut}_{\text{M4}}) = -8.51$  kcal/mol). Ini menunjukkan pemisahan kesemua Fc-FcRn complex berkekalan pada pH yang tinggi. Analisis menggunakan trajektori 5 ns terakhir berdasarkan program *cphstats* menunjukkan bahawa pemprotonan residu untuk kesemua kompleks berubah daripada pH 6.0 hingga pH 7.5. Pengikatan kuat kompleks pada pH 6.0 berkaitan dengan protonasi Fc:His310 dan menyebabkan ikatan hidrogen berkekalan dengan FcRn:Glu115. Sebarang peningkatan pH pelarut akan menyebabkan deprotonasi Fc:His310, menyebabkan kehilangan pembentukan ikatan hidrogen dengan FcRn:Glu115, menyebabkan peningkatan penolakan interaksi antara kompleks dan menyebabkan pemisahan kompleks Fc-FcRn. Pada peringkat teori, kajian ini menunjukkan keupayaan pendekatan *in silico* dalam mereka bentuk varian Fc baharu yang mungkin berguna untuk aplikasi masa depan.

***IN SILICO* DESIGN AND CONSTANT-PH MOLECULAR DYNAMICS STUDY  
OF HUMAN IGG1 FC AT PH6.0 AND PH 7.5**

**ABSTRACT**

The use of monoclonal antibody against various illnesses such as cancer, infectious diseases or autoimmune disorders has become a mainstream in the field of medical therapeutics. Specifically, the IgG1 with long serum half-life has led to its popular usage as a therapeutic drug scaffold where this property is related to the pH-dependent binding mechanism of IgG1 with FcRn. In this work, constant pH molecular dynamics (CpHMD) simulation setup was first established by testing against experimentally validated Fc-FcRn complex (Wildtype, Mut<sub>AAA</sub> and Mut<sub>YTE</sub>), associating and dissociating at pH 6.0 and pH 7.5, respectively. Binding free energy ( $\Delta G_{\text{Bind}}$ ) calculation via the MMGBSA approach was used to describe the binding affinity for the control complexes (Wildtype-, Mut<sub>AAA</sub>- and Mut<sub>YTE</sub>-FcRn) and have showed Fc variant Mut<sub>YTE</sub> as the strongest FcRn binder at pH 6.0. To obtain alternative Fc design with better binding properties for FcRn, pH 6.0 single trajectory Mut<sub>YTE</sub>-FcRn representing complex was extracted by MMTSB toolset clustering followed by SAAMBE server single point mutations imposed on the Mut<sub>YTE</sub> Fc of the complex. Mutagenesis result suggested four Fc positions with improved binding energy ( $\Delta G$ ) for FcRn. They were all applied to maximize the binding effect for FcRn, resulting the new IgG1 Fc variant design, Mut<sub>M4</sub> (Asp294Ala/Thr254Phe/Pro257Trp/Asp312Gly). The binding properties of Mut<sub>M4</sub>-FcRn complex were compared with the control complexes and deduced a better binding affinity of the designed Mut<sub>M4</sub> ( $\Delta G_{\text{Bind}} = -23.19$  kcal/mol) than control Fc ( $\Delta G_{\text{Bind}}(\text{Mut}_{\text{YTE}}) = -20.67$  kcal/mol; (Wildtype) = -13.94

kcal/mol) with FcRn at pH 6.0. All Fc-FcRn complexes showed low binding affinity at pH 7.5 ( $\Delta G_{\text{Bind}}(\text{Mut}_{\text{AAA}}) = +37.02$  kcal/mol; (Wildtype) = -3.86 kcal/mol; ( $\text{Mut}_{\text{YTE}} = -6.77$  kcal/mol; ( $\text{Mut}_{\text{M4}} = -8.51$  kcal/mol). This indicates the dissociation effect for all Fc-FcRn complexes retained at an increased pH. Last 5 ns *cphstats* program protonation analysis showed that the protonation state of titratable residues changed from pH 6.0 to pH 7.5 for all complexes. The complex strong binding effect at pH 6.0 has been correlated to the protonation of Fc:His310, leading to persistent hydrogen bonding with the FcRn:Glu115. Any increase in solvent pH had led to the deprotonation of Fc:His310, causing a loss in hydrogen bond formation with FcRn:Glu115 and led to increase in complex interface electrostatic repulsion, causing the Fc-FcRn complex dissociation. At theoretical level, this study demonstrates the ability of *in silico* approach in the design of a new Fc variant that could be useful for the future application.

# CHAPTER 1

## INTRODUCTION

### 1.1 Problem statement

Monoclonal antibody (mAb) has been used to treat against various conditions such as cancer, infectious diseases or autoimmune disorders (Lobner et al., 2016). They remain as the mainstream in the field of medical therapeutics with ~80 drugs being Food and Drug Administration (FDA) approved till date (Kumar, 2019). Fc-fusion proteins are any proteinaceous compound incorporate with the fragment of crystallizable (Fc) of an immunoglobulin G (IgG). The resulted Fc-fusion protein hence inherits various antibody (Ab) effector functions in therapeutics including the antibody-dependent cellular cytotoxicity (ADCC), antibody-dependent cell-mediated phagocytosis (ADCP) and complementary-dependent cytotoxicity (CDC) (Hulett & Hogarth, 1994; Nimmerjahn & Ravetch, 2008). The long IgG circulating half-life (~21 days for IgG1) was found induced by interaction between the IgG fragment of crystallizable region (Fc) and neonatal Fc receptor (FcRn), also known as the Fc-FcRn complex. The Fc-FcRn complex binds strongly strong at acidic pH ( $\leq 6.5$ ) but weaker around neutral pH ( $\geq 7.4$ ) (Raghavan et al., 1995; Oganessian et al., 2014), improves the drug pharmacokinetic by lowering requirement of drug administration frequencies and doses (Lobner et al., 2016). However, there are only a few Fc variants showed improved therapeutic serum half-life compared to the wildtype. Therefore, a more cost-effective method applying *in silico* approach for the design of IgG1 Fc with increase affinity for FcRn at acidic pH (pH ~6.0) yet retaining low affinity at pH 7.5 was considered.

## 1.2 Study objectives and expected outcome

The objectives of this work include:

- i. To study the IgG1 Fc-FcRn dynamics at pH 6.0 and pH 7.5.
- ii. To design IgG1 Fc variants with improved binding affinity for FcRn receptor via *in silico* approach

## 1.3 Thesis outline

Chapter 1 presents chapter outlined the problem statement and general objective of this work.

Chapter 2 included an overview of IgG structure, FcRn and the computational approach mechanism for the study of IgG Fc-FcRn binding.

Chapter 3 described the methodology for CpHMD simulation. The design of Fc variant was also described in this chapter.

Chapter 4 presented the result analysis including the Fc-FcRn interacting interface analysis, observation of quaternary structural shift and binding affinity of Fc-FcRn complexes.

Chapter 5 discussed and explained the results obtained from the analysis of the work.

Finally, Chapter 6 summarized overall results of the study. It also includes the study limitations and potential directions which can be considered to expand the performed work.

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1 Antibody diversity

The host immune system is responsible in protecting an organism against pathogenic diseases and it involves many biological processes that has been categorized as: (1) innate and; (2) adaptive immune responses. A deeper view, the adaptive immune response consists of two major types, namely the (1) cell-mediated immunity, mediated by T lymphocytes (T cells) and; (2) humoral immunity, mediated mainly by the antibodies (Abs) produced by B lymphocytes (B cells). Antigens (Ags) are characterized as host foreign molecules such as pathogens and toxins where they need to be eliminated to prevent pathogenic diseases (Sormanni et al., 2018). Contrary, Abs are glycoproteins produced by the host B lymphocytes that serve as Ag receptors exist to elicit intracellular signalling pathways upon Ag recognition. The Abs can recognize, neutralize and opsonize pathogenic antigens (Ag) through their paratope also known as Complementary Determining Regions (CDRs) (Irani et al., 2015; Sormanni et al., 2018). Following the recognition of Ab for Ag, both adaptive immune systems may interlink via Abs effector mechanism. The downstream effector mechanisms upon Ag recognition, is dependent on the type of Ab that recognizes the pathogenic Ag specifically, reported Abs show specific binding for a wide range of Ags via non-covalent interactions at great affinities ( $K_d = 10^{-7}$  to  $10^{-11}$  M) with the (Abbas et al., 2018).

The Ab type plays a role in different downstream effector mechanism upon Ag recognition while the diversity of Ab is based on the type heavy chain incorporated in the protein itself. The diversity of Ab hence evokes different effector immune responses, with each carrying different lifespan in human circulation. The main factor that leads to Ab

diversity is the B lymphocyte class-switching mechanism that in general, has created different Ab heavy chain including immunoglobulin G (IgG) , immunoglobulin M (IgM), immunoglobulin A (IgA), immunoglobulin D (IgD) and immunoglobulin E (IgE) (Roopenian & Akilesh, 2007). Among, IgG was found highly abundant in serum, at the concentration of 10-20 mg/ml and accounts around 10-20% of total plasma proteins. IgG subclass numbering were done based on its subclass abundance in human serum (Shakib & Stanworth, 1976). In short, IgG1 is the most prevalent IgG subclass in serum. The IgG subclass was found to share 90% identical amino acid sequences, each of them not only present in various amount and responded differently upon antigen recognition (Bruhns et al., 2009; Vidarsson et al., 2014).

## **2.2 Antibody structure**

Ab molecules shared the same basic tetrameric Y-shaped structural characteristics, consisting a symmetric core composed of two identical light chains (LCs) and two identical heavy chains (HCs) (Padlan, 1994), with each composed of two and four immunoglobulin domains, respectively. The Ig domains are secondary architecture-antiparallel  $\beta$ -sheets compactly folded into a globular structure and stabilized by interchain Cys-Cys disulfide bonds (Liu & May, 2012). The LCs are linked by disulfide bonds formed between Cys residues of each monomeric HCs.

In general, Ab structure accommodates two regions based on their functions: (1) N-terminal antigen-binding fragment (Fab) region which is involved in Ag binding, composed of the first two domains of LC : variable light (VL) and constant light (CL) domains, together with HC: variable heavy (VH) and constant heavy 1 (CH1) domains; and (2) C-terminal fragment of crystallizable (Fc) region that is composed of two monomeric HCs (N-glycosylated-constant heavy 2 (CH2) and constant heavy 3 (CH3) domains) (Figure

2.1). The non-Ag binding Fc region exhibits lesser variation compared to the paratope of Fab region and is responsible in regulating the communication with other parts of the immune system including the aforementioned cell-mediated immunity. Since the downstream reactions initiates upon Ab Fc interaction with a variety of related accessory molecules, leading to direct or indirect elicitation of downstream effector mechanisms such as antibody-dependent cytotoxicity (ADCC), antibody-dependent cellular phagocytosis (ADCP) and complement-dependent cytotoxicity (CDC) (Peipp et al., 2007). The Ab Fc region hence was described to carry the physicochemical properties in mediating pathogen clearance via cellular and complement-mediated pathways (Irani et al., 2015). Besides, by binding to neonatal Fc receptor (FcRn), Ab such as the IgG1 shows prolonged serum half-life that has been rank as second long lasting serum circulating protein (sustained ~21 days) (Choe et al., 2016).

### **2.2.1 Fragment of crystallizable (Fc) region**

The two heavy chains (HCs) of Fc are covalently linked by disulfide bridges at the hinge region. Interactions between the monomeric Fc involve intermolecular CH2 glycans interaction also the non-covalent intermolecular interactions at the CH3 domains (Figure 2.1). Due to presence of N-linked carbohydrate covalently branched out from Asn297 in CH2 domain, an “opened” interior surface was exhibited where both domains are apart. The Fc region plays important biological roles including: (1) increase stability of overall Ab structure; (2) enhance Fc-mediated downstream effector functions and; (3) extend IgG serum half-life. IgG subclass hinge structure differentiates prominently (Table 2.1). The difference between subclasses has conferred to many unique properties of an Ab such as stability, flexibility and the ability of Fab arms to reach distantly sparse Ag.



### **2.2.2 Fc-fusion protein**

Numerous endogenous or exogenous proteins have the potential to treat various pathological conditions. However, the therapeutic efficacies may be confined by limitations such as protein instability or rapid clearance. Application of utilizing the Fc and FcRn (Fc-FcRn) complex binding mechanism for Fc-fused constructions can be introduced with the Fab region being replaced with any proteinaceous compound used targeted pathogenic Ags. Hence, the Fc-fused construction is expected to retain its downstream effects of the immune mechanisms. Theoretically, any cell death can be enhanced using Ab as a vehicle to deliver cytotoxic drugs to act directly on targeted cells since the Fc-fusion protein inherits the biological advantage (including extended serum half-life) of Ab Fc region (Yang et al., 2018) (Figure 2.2). Some marketed Fc-based therapeutic applications has been applied clinically such as Etanercept (Fc region fused with Tumour Necrosis Factor  $\alpha$  (TNF- $\alpha$ ) inhibitor) and Abatacept (Cytotoxic T-lymphocyte-associated protein 4 (CTLA-4)) have shown success in prolonging small recombinant proteins half-life and overall therapeutic effects (Higel et al., 2016).

### **2.3 Therapeutic antibodies**

Therapeutic Abs have emerged as a potential field for diagnostic and therapeutics after the rise of efficient procedures for monoclonal antibody (mAb) production discovered by Kohler and Milstein (1975) . As a multifunctional glycoprotein, Ab facilitates numerous cellular and humoral responses against variety of host or foreign Ag. They are used in the field of medicine including the treatment of autoimmune diseases, cancers or infection conditions. The usage with several grams of mAb in therapeutics per patient is common, implying the significant impact of mAb on clinical benefits against diseases. With the

application of hybridoma technology, mAb can be produced in mass quantity from a single B-cell clone. Research and development on Abs started to bloom in 1980's since the launch of first commercial murine Ab, muromonab-CD3 to prevent rejection of kidney transplantation (Kuhn & Weiner, 2016; Yang et al., 2018). However, murine Ab administration may elicit of undesired human anti-mouse Ab (HAMA) responses. Development on mAbs modification to reduce immunogenicity hence, involves the production of murine, chimeric (rituximab), humanized (transtuzumab) and fully human mAbs (Figure 2.3) (Chames et al., 2009; Yang et al., 2018). These mAbs can be further optimized with desired specificity before they are ready to be produced in bulk for clinical use.

To date, approximately 40 mAbs have been clinically approved with the majority being full-sized Abs (Yang et al., 2018). Although these mAbs are very similar to naturally produced Abs but full-sized mAbs present poorer penetration and binding accessibility against target Ags which is permissible by smaller molecules. The understanding of Abs size and structure further allows possible alteration and design based on Abs size with a variety of derivatives drugs marketed for the past few decades including: (1) Fragment of antigen binding (Fab) (Ranibizumab, Abciximab); (2) single-chain variable fragment (scFv) (under clinical development: gancotamab, pexelizumab, brolocizumab) (Bates & Power, 2019); (3) single-domain antibody (sdAb) (Caplacizumab) (Duggan, 2018). These small Abs are lack of fragment of crystallizable (Fc) region. Therefore small-sized Abs are rapidly cleared from circulation via renal filtration (MW below 60kDa) or protein catabolism. The half-life of these drugs ranged from few minutes to several hours. These treatment hence require tight administration schedule in order to maintain therapeutically effective dose (Kontermann, 2009). In terms of therapeutic efficacy concerns, there is a rise

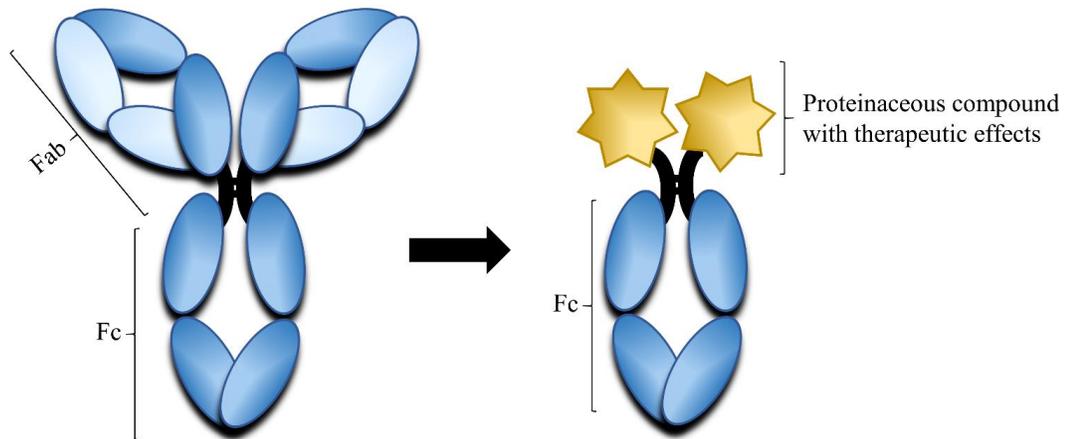


Figure 2.2 Schematic diagram of IgG scaffold being applied in Fc-fusion construction. The IgG domains were coloured in blue while proteinaceous compound in Fc-fusion protein were coloured in yellow.

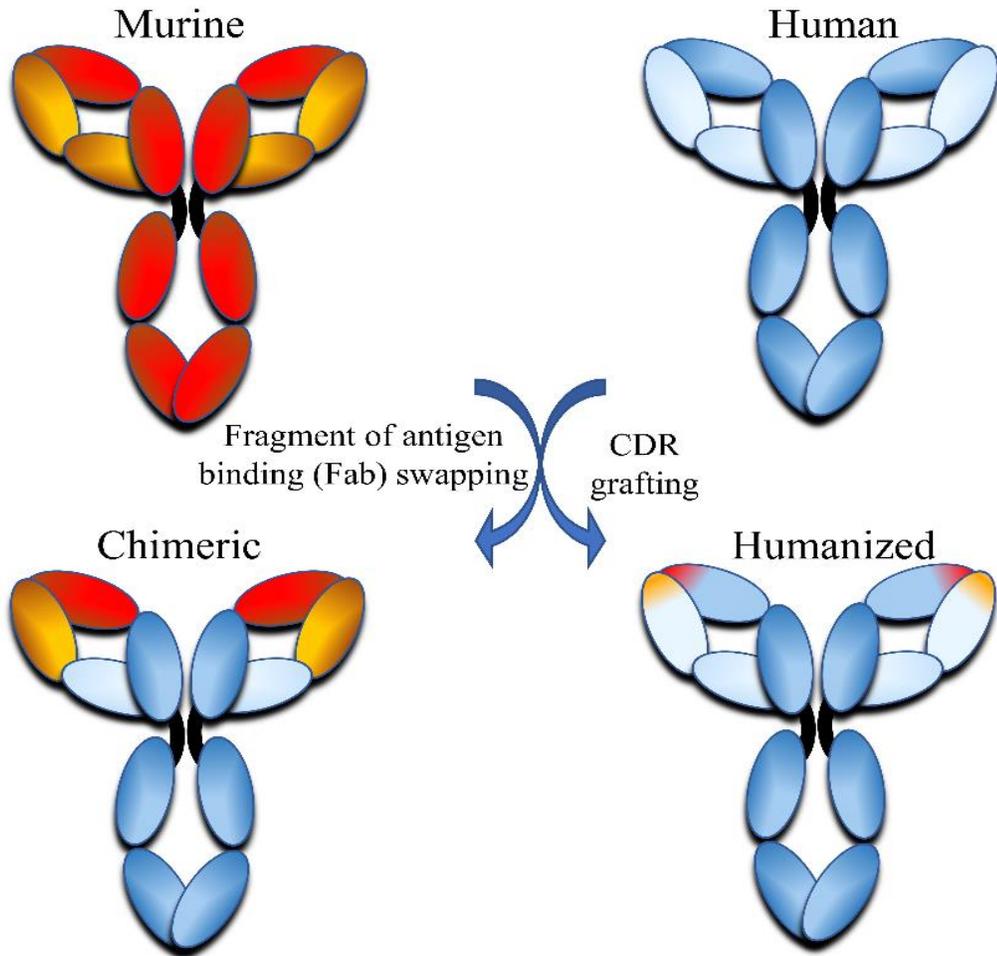


Figure 2.3 Schematic diagram of murine, chimeric, humanized and fully human antibodies. The murine Abs were coloured in red while human Abs were coloured green. Figure adapted from (Chames et al., 2009).

with the use of small sized antibodies for chronic diseases such as cancer whilst long-lasting activity of drugs is in need (Kontermann, 2009). All these concerns resulted in the need in pharmacokinetics-optimized designs for small size Abs to minimize premature clearance of these treatment.

## **2.4 Neonatal Fc receptor (FcRn)**

FcRn (also known as Brambell receptor) has a role in IgG transport and protection in human body. It protects IgG from protein catabolism in a pH-dependent manner, which explains the prolonged half-life of IgG compared to other endogenous proteins (Kuo & Aveson, 2011; Olafsen, 2012). FcRn that is encoded by *fcgrt* gene in chromosome 19 (Kuo et al., 2010) can be found expressed on various organs and cells such as endothelial cells (Shah et al., 2003) and syncytiotrophoblast of human placental (Simister et al., 1996). It is a major histocompatibility (MHC) class I-like molecule where it shares 22-29% sequence identity with the MHC class I molecule (Simister & Mostov, 1989). Both FcRn and MHC class I molecules are heterodimer composed of a soluble light chain (LC) known as beta-2-microglobulin ( $\beta$ 2m) and a membrane bound heavy chain (HC) encompassed of three extracellular domains ( $\alpha$ 1,  $\alpha$ 2,  $\alpha$ 3) as demonstrated in Figure 2.4 (Burmeister et al., 1994). FcRn differs from the MHC class I by means of its general structure, having a narrowed groove that leads to the inability to bind to peptides as antigen presenting cells.

### **2.4.1 IgG-FcRn binding mechanism**

There are certain IgG subclasses that can bind with FcRn including IgG1, IgG2 and IgG4. On the other hand, IgG3 cannot bind to the FcRn hence presented with lower serum half-life (Spiegelberg & Grey, 1968; Morell et al., 1970; Kim et al., 1999). The capability of these three subclasses to bind FcRn leads to observable prolonged serum half-life for

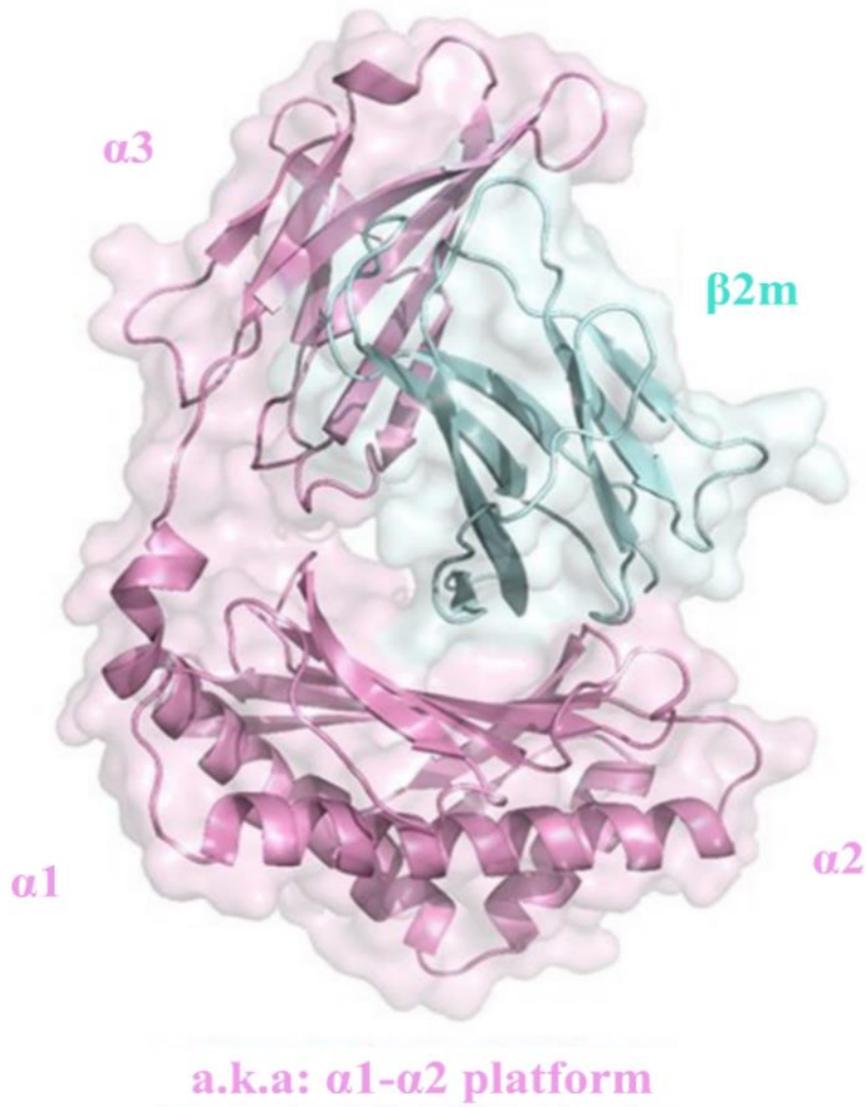


Figure 2.4 The structure of FcRn. HC is in pink (Domain  $\alpha 1$ -3) and LC ( $\beta 2m$ ) is in blue. (PDB-code: 1FRT (Burmeister et al., 1994), Figure was produced using PyMOL program (DeLano, 2002).

~21 days (Ghetie et al., 1996; Kim et al., 1999; Ghetie & Ward, 2000; Martin et al., 2001; Roopenian et al., 2003; Roopenian & Akilesh, 2007). The binding of IgG Fc region to neonatal Fc receptor (FcRn) has been denoted to enhance prolonging the Ab serum half-life. Fc-FcRn binding mechanism (Figure 2.5) starts from the intracellular mechanisms in both transport and protection that begins with: (1) pinocytosis of IgG from the membrane surface to form intracellular endosomes; (2) binding of FcRn and IgG in the acidified (~pH 6.0) endosome; (3) circulation and merging of endosome with the plasma membrane; and (4) detachment of FcRn and IgG in alkalinized endosomes (~pH 7.5) upon exposure to the blood circulatory (release IgG back into the circulation) (Kuo & Aveson, 2011). During the process, any proteins that do not bind to FcRn will be catabolized by lysosomes.

#### **2.4.2 Fc-engineering**

The use of small molecule drugs as therapeutics leads to low bioavailability that requires higher dosing frequency. Despite applying the optimal administration route (balanced by the avoidance of intravenous access for subcutaneous route), limitation still lies in the bioavailability of the treatment. Attesting to antibodies' importance in the role of protective immunity, long serum half-life for a few immunoglobulin subclasses (IgG1, IgG2 and IgG4) have been of the interest in several studies (Mankarious et al., 1988; Meredith et al., 1992; Kontermann, 2011). Although different effector functions elicitation by differing Ig subclasses, IgG1 which occupies highest serum count is highly applied for therapeutics purposes. The long serum half-life of IgG1 mediated by FcRn salvation mechanism promotes the use of IgG1 Fc as scaffold aim to improve pharmacokinetics of therapeutics.

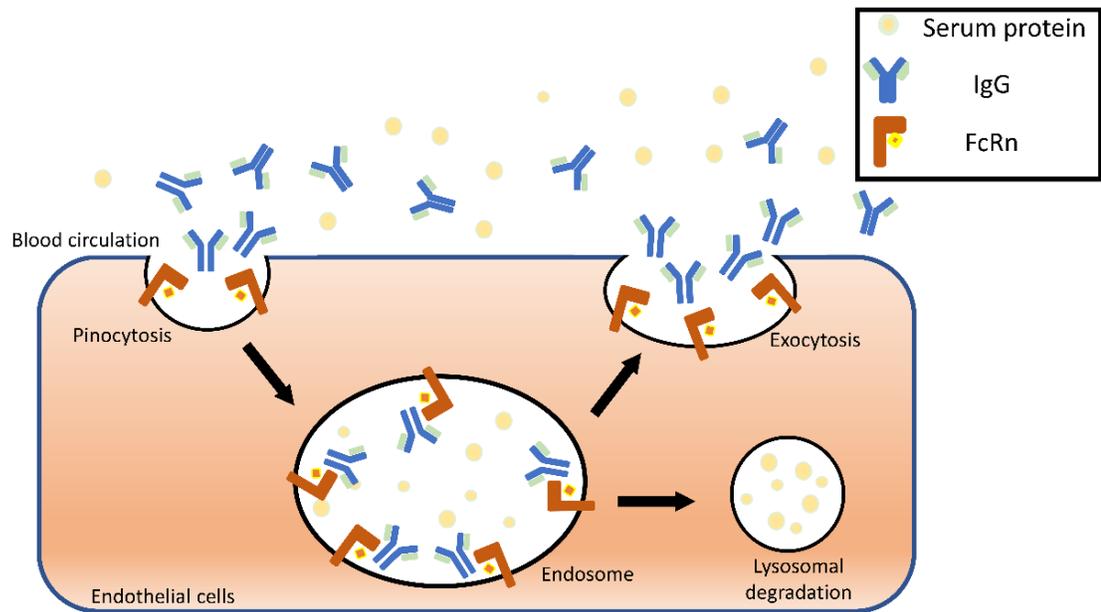


Figure 2.5 Schematic diagram of FcRn-dependent IgG rescue pathway.

Since IgG1 poses long serum half-life (approximately 21 days), the potential in Fc-fusion was observed. The long serum half-life for IgG1 is not solely due to the large molecular weight (MW) that retained them from kidney filtration but also its interaction with the FcRn receptor. Naturally, IgG1 Fc region binds to FcRn with high affinity at acidic pH (pH 6.0-6.5) but remain its dissociating effect at physiological pH (~pH 7.5) (Jones & Waldmann, 1972; Rodewald, 1976; Simister & Rees, 1985). Implementation of FcRn-mediated recycling processes hence enables the rescue of Fc-conjugated drugs from rapid clearance, improving the drug pharmacokinetic properties in many ways such as their distribution, metabolism and excretion that were strongly influenced by the size, shape, physiochemical properties and sensitivity to proteolytic degradations (Kontermann, 2009). However, Fc design with improved binding for FcRn at both acidic and physiological pH showed enhanced antibody degradation and thereby reduces their circulatory half-life (Roopenian & Akilesh, 2007).

The engineering on IgG1 Fc to improvise Fc-FcRn interaction in a pH dependent way was experimentally determined. The complex crystal structure revealed FcRn binds to the Fc at the CH<sub>2</sub>-CH<sub>3</sub> domain interface with mutation Fc variant (Mut<sub>YTE</sub>: Met252Tyr/Ser254Thr/Thr256Glu) being identified with improved affinity for FcRn at acidic but not neutral pH (Dall'Acqua et al., 2002; Dall'Acqua et al., 2006; Yeung et al., 2009; Oganessian et al., 2014). The binding improvement at acidic pH without compromising the dissociation from FcRn at basic pH environment is important as *in vivo* testing found that monoclonal antibodies Fc incorporated with Mut<sub>YTE</sub> such as cetuximab and bevacizumab outperformed the wildtype (WT) Fc with observable increased plasma half-life without losing their therapeutic effects (Zalevsky et al., 2010). Hence, regardless monoclonal antibody or Fc-fusion protein, optimization and modification of Fc for FcRn binding at acidic is expected

to improve the respective therapeutic half-life and efficacy (Sidhu & Fellouse, 2006; Czajkowsky et al., 2012).

## **2.5 Computational approach in molecular studies**

In the area of structure-based protein engineering, the use of experimental methods such as X-ray crystallographic and NMR spectroscopy has allowed protein structural information to be determined. With the availability of these biological data, biostructure modelling has taken an advantage of computational technology to predict the biology system structure, mechanism and functions. When macromolecule folding or binding involving complex inter- and intramolecular interactions which is distinctively depending on the macromolecules' three-dimensional (3D) shapes, methods such as X-ray crystallography and *in silico* analysis not only permits the generation and visualization of these tertiary structure conformations. These methods also allow selection of macromolecule position for amino acid mutation to be eased as the detailed protein three-dimension (3D) structures can be aided with computational binding energy predictor to determine the effect of predicted mutagenesis outcomes.

Protein-protein interactions were widely described to lead to sequential biological activities such as signal transduction and immune responses (Tuncbag et al., 2011; Z. P. Liu & Chen, 2012). Therefore, the structure-based knowledge could provide directional design of potential therapeutic drugs at macromolecular level. For example, comparative modelling and docking has been found capable in improving protein-protein binding affinity by two single-point mutations (Cannon et al., 2019). This strategy reveals the potential in visualising and estimating the mechanistic details ranging across various macromolecules including proteins, carbohydrates, lipids and nucleic acids. By means

understanding protein-protein interaction at the molecular level allows biomolecular utilization in the field of medical diagnostics and therapeutics.

Mutagenesis on protein-protein interacting surface can introduce diverse changes, ranging from protein-protein interacting affinity to biological interaction network (Panchenko & Babu, 2015). Besides structural-based approach, several methods either empirical-energy, statistical potential or machine-learning-based prediction can be used to study the behaviour of macromolecules (Fu et al., 2019). For example, complex binding affinity can be measured by sophisticated physical energies calculations instead of relying solely only on the complex structural arrangement.

### **2.5.1 Molecular dynamics (MD) simulation**

The MD simulation allows prediction of biomolecules dynamic properties such as the structure, dynamics, and thermodynamics of single or complex macromolecules. The capacity in macromolecule simulation events showed considerable accuracy to depict incomplete 3D structural information. Hence, MD simulation serves as a complementary analysis to conventional experiments that enables estimation of macromolecules assemblies and interactions at atomic level with the function of time. Computational molecular dynamics (MD) simulation are available in different programs (e.g., AMBER (DeLano, 2002), CHARMM (Peipp et al., 2007), GROMACS (Lindahl et al., 2001), GROMOS (Saphire et al., 2001), NAMD (Bruhns et al., 2009), TINKER (Shakib & Stanworth, 1976), and XPLOR (Vidarsson et al., 2014)). During simulation, molecular atoms are assigned with velocities under a given temperature which impose atomic physical movements during simulation and applying forcefield approximation to study the macromolecule physical properties. Based on the programs, different forcefield parameters

could be applied. For example, AMBER software utilizes the ff14SB (an improved version for protein conformation from ff99) forcefield while CHARMM software utilizes all atom CHARMM 36 forcefield (an improved version from CHARMM22). These forcefield information used to describe the protein conformation were obtained from experimental data incorporate with mathematical theories. Dynamics are applied on atoms at a time step of (e.g.  $2 \times 10^{-15}$  s) governed by Newtons's Law of motion along the simulation. The whole process is related to three independent qualities such as: (1) time; (2) molecule conformation; and (3) potential energy. MD simulation can also be applied to Ag and Ab interaction studies with studies found applying the technique to understand the Ag-Ab interaction and folding at a molecular level (Yanaka et al., 2017; Nimrod et al., 2018).

### **2.5.2 Influence of environmental pH on titratable residues**

The presence of titratable residues at the active site can influence macromolecules interactions pattern at different environmental pH (Harris & Turner, 2002). Environment pH key regulator in biological activities as it can directly impose effects on macromolecule in term of structure, stability, solubility and functions (Pace et al., 2009; Socher & Sticht, 2016). For example, some bacteria is able to survive the acidic conditions of their host (e.g. the stomach) by using acid-activated chaperones which protect substrate proteins upon binding (Hong et al., 2012). In virus, some fusion proteins mediate cell entry pH-dependently (Roche et al., 2006). Other proteins in vertebrates undergo pH changes during their maturation on the way through the endoplasmic reticulum and the Golgi apparatus (Thomson & Ananthanarayanan, 2000).

In protein, the chemistry theory regarding the weak acid-base equilibrium has been extended with six titratable residues (Asp, Glu, Cys, His, Lys, Tyr). The titratable residue

side chains will change in protonation state at different environmental pH. Based on the chemical properties, Asp, Glu, Tyr, and Cys remain neutral in acidic pH environments but negatively charged at basic pH environment. On the other hand, Arg, His, and Lys remain neutral in basic pH environment but are positively charged at acidic pH environments. From a molecular point of view, changes in environmental pH that further affects the protonation states and charge distribution of these amino acids will lead to conformational changes and protein interacting properties such as the formation or lose of non-covalent interactions.

Understanding the protein-protein binding mechanisms requires the knowledge of protein nature and changes in protein physical states. The protein shape complementary play a role in binding process where different structural arrangement may occur upon in protein binding or non-binding form. For real condition mimicry, solution pH remain as an unavoidable physicochemical property for all biochemical reactions. In differing pH environments, the protein net charge will in turn affect protein-protein interactions. Together with the high specificity in biological structure and function integration, a delicate balance between macromolecule attractive and repulsive forces remained as a key factor in complex formation. The changes in titratable residues protonation equilibrium displacement under the influence of pH change may lead to alteration in the residue net change (Gunner et al., 2000). Altered charge proteins may lead to apoprotein conformational changes or triggers changes in protein binding activities. In addition, complex formation involves surface water removal and reduces protein side chain degree of freedom. This will result in the whole binding process to be compensated by energy gained during non-covalent formation and other complex stabilizing factors such as the electrostatic effects which again, largely influenced by the protonation states of the protein

(Warshel & Russell, 1984; Honig & Nicholls, 1995; Bashford, 2004; García-Moreno & Fitch, 2004).

The pH-dependent binding mechanism of Fc with FcRn occurs at Fc CH<sub>2</sub>-CH<sub>3</sub> domain interface toggled by His residues between neutral and positively charged states at different pH (Raghavan et al., 1995; Medesan et al., 1997; Martin et al., 2001; Shields et al., 2001; Dall'Acqua et al., 2002). Similar pH-dependent mechanism related to His residue charges was also described in a study where protonated His promotes adsorption mechanism attributed to its net positive charge at acidic pH (Kurut et al., 2014). Later, crystallographic analysis supported that Fc-FcRn interaction was enhanced by positively charged Fc residues which interact with FcRn negatively charged residues. When the environment pH becomes basic, His residues tend to lose their protons and lead to neutralization of the side chain. This reduces the electrostatic attraction with its neighboring residues. Diminished electrostatic interaction between the Fc-FcRn complex was then enhanced the release of Ab from endosomes and back to the blood circulation (Schmidt et al., 2013; Oganessian et al., 2014).

### **2.5.3 Constant pH molecular dynamics (CpHMD) simulations**

The conventional MD simulation employs fixed, predetermined protonation states for titratable residues along the simulation (Mongan et al., 2004). However, the technique faces limitation in hotspot residues identification corresponding to their protonation states. The interplay between change in protonation state and the effects in protein conformation alternation will not be taken into account for conventional MD simulations. This constant protonation approach results in some drawbacks for studying pH-dependent effects because: (1) assigning the right protonation states for the titratable groups in the protein requires

knowledge of their pKa values; (2) if any of these pKa values are near the solvent pH, there may be no single protonation state that adequately represents the ensemble of protonation states appropriate at the stated pH and; (3) the invariable protonation states decouple the dynamic dependence of pKa and protonation state on conformation (Mongan et al., 2004; Socher & Sticht, 2016).

The atomic level of protein's pH-dependent conformational changes using experimental approaches such as X-ray crystallography or NMR spectroscopy is technically demanding. These approaches have limitations such as most proteins do not crystallize at very different pH values or the technique is limited for the identification of small proteins (Socher & Sticht, 2016). The static structures obtained from X-ray crystallography or NMR spectroscopy do not take into account of protein conformation changes induced by protonation states. Hence, MD simulation which allow the modelling of protein ionization states for titratable residues based on CpHMD simulation and allows the calculation of pKa of titratable residues can be a good alternative (Baptista et al., 2002; Bürgi et al., 2002; Dlugosz & Antosiewicz, 2004; Lee et al., 2004; Mongan et al., 2004; Khandogin & Brooks, 2005; Mongan & Case, 2005; Williams et al., 2011; Vila-Viçosa et al., 2013; Socher & Sticht, 2016). CpHMD simulation allows the dynamics simulation taking consideration in the protonation state changes for the six aforementioned titratable residues (Glu, Asp, His, Cys, Lys and Tyr). It applies computational simulation for complex conformation coupled with periodic Monte Carlo (MC) sampling of titratable residues protonation states interspersed throughout the simulation. The simulation is periodically halted while MC sampling comes into role for protonation states sampling on titratable residues while solution pH is set as an external thermodynamic parameter determining the

titratable residues protonation states to investigate the effect of solution pH on complex's atomic changes (Baptista et al., 1997; Börjesson & Hünenberger, 2001).

In short, dynamics simulation coupling pH dynamic adjustment for titratable residue protonation state changes implemented in CpHMD approach carries an advantage in terms of free energy calculations as the technique promotes better pH-dependent study for structural fluctuations or conformation changes (Chu et al., 2013). This allows the view for pH effects on biological structure alteration and shed light on existing ambiguities in pH-dependent activities. The dominant approach is to probe and analyse the static structure at desired pH environments and to demonstrate the importance of identifying the transient conformational state, which accounts a minor population of total conformations under physiological conditions. The pH-dependent molecule interaction applying CpHMD has been demonstrated by a few studies. For example, pH-dependent ligand releasing mechanisms from AegOBP1 protein or CquiOBP1 where both studies can highlight the protein structure deviation in relative to pH changes which leads to variation in protein-ligand binding affinity (Chu et al., 2012; Chu et al., 2013).

#### **2.5.4 Assessing Fc-FcRn complex binding affinity via binding free energy calculation**

Protein-protein complex formation is essential for various biological processes. Due to the importance of pH-dependent Fc-FcRn binding mechanism, CpHMD simulation incorporate Molecular Mechanics/Generalized Born Surface Area (MM-GBSA) binding free energy can be applied to determine the complex binding affinity at different pH. MM-GBSA method calculates complex binding affinity by including potential energies from molecular mechanics and solvation energies from both polar and non-polar components

(Petukh et al., 2006). The end-point binding free energy calculation MM-GBSA was claimed to be strongly appealing for the protein-protein interface interaction studies. However, the accuracy still strongly depends on the several components including the applied forcefields, dielectric constant and the trajectory sampling methods (Kollman et al., 2000; Miller III et al., 2012; Weng et al., 2019).

## CHAPTER 3

### METHODOLOGY

#### 3.1 The Fc-FcRn complex structure

This study has chosen IgG1 construction for the study due to its high population in serum and its potential to bind with FcRn. The Fc-FcRn initial complex structure for simulations was retrieved from Protein Data Bank (<https://www.rcsb.org/>), with accession code 4N0U at 3.8 Å resolution (Oganessian et al., 2014). Protein-protein interface of the retrieved crystal structure was analyzed using web server Protein Interface, Surfaces and Assemblies (PISA) version 2.0.4 (Krissinel & Henrick, 2007) to predict the intermolecular interacting interface. The software PISA enables user to input complex 3D structures and correlates the assembly of complex structure to biological macromolecule physiological functions. The software incorporated calculation for several components including complex macromolecular weight, interface area (IA), solvation energy (DG), complex binding energy (BE), interface hydrophobicity (PV), numbers of intermolecular hydrogen bonds (HB) and numbers of intermolecular disulfide bond (SB). The results were presented in percentage based on macromolecules interface scoring system for the aforementioned calculated criteria that classifies the molecular interaction as “biological relevant” (interaction radar result of 30% or more) or “insignificant” (interaction radar result lesser than 30%). Scoring parameters used were based on the Macromolecular Structure Database (Henrick & Thornton, 1998). The crystal structure (PDB\_ID: 4N0U) was uploaded to the PISA server to highlight the Fc-FcRn interface titratable residues and the domain molecules consist of titratable residues expected to be involved in pH interactions between both molecules were retained (Ying et al., 2014).

The corresponding X-ray crystallographic Fc-FcRn complex structure is originally comprised of IgG1-Fc region with YTE mutations (Met252Tyr/ Ser254Thr/Thr256Glu), denoted as Mut<sub>YTE</sub> throughout the project has been used as the starting structure in this study. Other Fc variants (wildtype and Mut<sub>AAA</sub>) were modelled using MODELLERv9.19 (Šali & Blundell, 1993), taking Mut<sub>YTE</sub> Fc-FcRn complex as a template. The Mut<sub>AAA</sub> with Ile252Ala/His310/His435Ala Fc mutation is experimentally confirmed to show less binding with FcRn at pH 6.0 (Popov et al., 1996; Medesan et al., 1997; Kim et al., 1999) was hence used as a negative control in CpHMD simulation setups (expected binding free energy to be less favourable than WT and Mut<sub>YTE</sub>). The wildtype (WT) is a revert Fc variant from the available crystal structure Mut<sub>YTE</sub> mutation (Met252Tyr/ Ser254Thr/Thr256Glu). Both WT and Mut<sub>YTE</sub> were included as positive controls to validate the CpHMD setups (expected binding free energy of Mut<sub>YTE</sub> to be higher than WT). All parameter and topology files were prepared using LEaP program available in AMBER 18 software suite (Case et al., 2018). The CpHMD method implemented in AMBER software issues the Monte Carlo sampling of protonation states for titratable residues where the protonation states of the residue will be changed by changing the partial charges of the titratable residue atoms along the MD simulation. To allow changes occur on titratable residues during the simulation, the titratable residues (Asp, Glu, His, Lys, Tyr, and Cys) naming have to be changed as ASP to AS4, GLU to GL4, HIS to HIP while Lys, Tyr, and Cys remained as the same name. For this study, only the complex interface titratable residues were redefined which includes: Fc (Asp249, Glu256, His310, Asp312, His428, Glu429, His433 and His435) and FcRn (Glu115, Glu116, Asp130 and Glu133) residues. All titratable residues were fully protonated during topology file preparation. The parameter files that define the protein was leaprc.constpH command file which loads the ff12SB force field for the definition of protein