THEORETICALLY ENGINEERED DARP$_{ins}$ TARGETTING DENV-2 ENVELOPE PROTEIN

CHONG WEI LIM

FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2021

THEORETICALLY ENGINEERED DARP**ins** TARGETTING DENV-2 ENVELOPE PROTEIN

CHONG WEI LIM

THESIS SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF CHEMISTRY
FACULTY OF SCIENCE
UNIVERSITI MALAYA
KUALA LUMPUR

2021

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **CHONG WEI LIM** Matric No: **SHC140078**

Name of Degree: **Doctor of Philosophy**

Title Thesis :

THEORETICALLY ENGINEERED DARPins TARGETTING DENV-2 ENVELOPE PROTEIN

Field of Study: **COMPUTATIONAL AND THEORETICAL CHEMISTRY**

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date:

Subscribed and solemnly declared before,

Witness's Signature Date:

Name:

Designation:

ABSTRACT

THEORECTICALLY ENGINEERED DARPins TARGETTING DENV-2 EVELOPE PROTEIN

Infectious diseases caused by dengue virus (DENV) have been threatening human health worldwide particularly tropical and sub-tropical regions. Infected population, as reported by WHO, reach 390 million yearly and the global incidence of dengue has tremendously grown. Specific drugs for treating dengue are under development while several antibodies bound to domain III (DIII) envelope (E) protein of DENV were found to prevent the viral entry process. In this study, the potential of Designed Ankyrin Repeat Proteins (DARPins), one of the non-immunoglobulin protein scaffolds, mimicking the binding interactions of antibodies and DIII of DENV envelope (E) protein was explored. Selected DARPins retrieved from the Protein Data Bank were docked to the epitope of domain III as recognised by antibodies under the HADDOCK web server. Representative docked complexes then underwent molecular dynamics simulations (MDs) with AMBER forcefield ff14SB to study their molecular properties. Binding affinity of DARPins to DIIIs was elucidated by free energy calculation using Molecular Mechanics–Poisson-Boltzmann Surface Area/Generalized Born Surface Area (MM-PPSA/GBSA) protocols while the important residues for protein-protein interactions were identified by further decomposing the binding free energy per residue basis. DARPin residues found with unfavourable energy within the binding vicinity then underwent computational site direct mutagenesis. Improved and engineered DARPins in complexed with DIII were simulated under the same parameters for free energy calculations. Gaussian Network Model (GNM), an elastic network model, was then used to investigate the dynamics of the proteins in terms of global mode shape and their dynamic cross-correlations. Global mode shape was observed corresponding to binding free energy, in which a low binding free energy was accompanied by a lower mode shape, or lower mobility. The simulated

techniques provide valuable tools in understanding the structural dynamics and energy contribution in designing the DARPins in their binding to the E protein of DENV-2. The protein-protein complexes with different binding activities can clearly be identified by evaluating the binding free energy and global mode shape of the structures from the long timescale MDs.

Keywords: DARPins, Dengue, Molecular dynamics simulations, Gaussian network model, Global mode shape

REKABENTUK PROTEIN BERULANG ANKYRIN YANG MENYASARKAN PROTEIN SARUNG DENV-2

ABSTRAK

Penyakit berjangkit yang disebabkan oleh virus denggi mengancam kesihatan manusia seluruh dunia terutamanya kawasan tropikal dan sub-tropikal. Populasi yang dijangkiti mencapai 390 milion setiap tahun berdasarkan WHO dan insiden global denggi telah bertambah secara mendadak. Ubat khas merawat denggi masih di bawah pembangunan manakala beberapa antibodi yang terikat kepada domain III (DIII) protein envelope (E) denggi didapati boleh menghalang proses kemasukan virus. Dalam kajian ini, potensi *Designed Ankyrin Repeat Proteins* (DARPins) sebagai salah satu perancah protein bukan immunoglobulin yang mencontohi interaksi pengikatan antibodi dengan DIII protein E DENV diterokai. DARPin terpilih daripada *Protein Data Bank* dilabuhkan dengan epitop DIII yang dikenali oleh antibodi menggunakan perkhidmatan HADDOCK web. Kompleks labuhan kemudiannya menjalani simulasi molekul dinamik (MD) menggunakan daya medan ff14SB *AMBER* bagi mengkaji sifat molekul. Kecederungan interaksi antara DARPins dan DIII dapat dijangkakan dengan pengiraan tenaga bebas pengikatan menggunakan protokol *Molecular Mechanics–Poisson-Boltzmann Surface Area/Generalized Born Surface Area (MM-PPSA/GBSA)* manakala residu penting bagi interaksi antara protein dikenal pasti melalui penguraian tenaga bebas per residu. Residu DARPins yang didapati memiliki tenaga yang tidak memuaskan dalam kawasan interaksi ditukarganti melalui mutagenesis langsung lokasi secara pengkomputeran. DARPins berkompleks dengan DIII yang telah dibaiki menjalani simulasi menggunakan parameter yang sama melalui pengiraan tenaga bebas. *Gaussian Network Model* (GNM), sebuah model jaringan elastik, juga digunakan untuk mengkaji dinamika protein dari segi bentuk mod *global* dan analisis *korelasi bersilang dinamik.* Bentuk mod global diperhati berhubungkait dengan tenaga bebas pengikatan, di mana tenaga bebas pengikatan yang

rendah mempunyai bentuk mod yang rendah, ataupun mobiliti yang rendah. Teknik simulasi yang digunakan adalah sesuatu yang penting untuk memahami dinamika struktur dan sumbangan tenaga dalam mereka bentuk DARPins yang berinteraksi dengan protein E DENV-2. Kompleks protein-protein yang berbeza aktiviti berikat boleh dikenalpasti dengan mengira tenaga pengikatan bebas dan *bentuk mod global* struktur selepas melalui simulasi molekul dinamika yang berjangkang panjang.

Kata Kunci: DARPins, Denggi, Simulasi Molekul Dinamik, Gaussian network model, Global mode shape

ACKNOWLEDGEMENTS

First, I would like to thank my supervisors Assoc. Prof Dr. Vannajan Sanghiran Lee, Prof Dr. Sharifuddin and Assoc. Prof. Dr. Chee Hui Yee for motivation and guidance throughout the course of my study.

Also, I would like to extend my grateful thank to my lab-mates, Ms Vertika, Dr. Chin Sek Peng, Mr. Yam See Chuan and Mr. Ramtin Ranji for their kind support and useful suggestions. It was really great to learn from each other and sharing knowledge together.

Of course, a big thank my beloved family member too for their understanding and endless love. They have been supporting me and encouraged me along the way especially the moment when I am emotional.

Thanks also go to the Data Intensive Computing Centre (DICC) University Malaya for high performance computing facility.

Last but not least, I have to thank my friends too for their support and care. Many thanks!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

2J	:	DARPin 2J8S
3C	:	DARPin 3NOC
ADE	:	Antibody Dependent Enhancement
Fab	:	Antigen Binding Fragment
C	:	Capsid
CAPRI	:	Critical Assessment of PRedicted Interactions
CDR	:	Complementarity Determining Region
DAA	:	Direct-acting Antivirals
DARPin	:	Designed Ankyrin Repeat Proteins
DCC	:	Dynamics Cross-Correlations
DENV	:	Dengue virus
DHF	:	Dengue Hemorrhagic Fever
DF	:	Dengue Fever
DSS	:	Dengue Shock
DIII	:	Domain III
E	:	Envelope
ER	:	Endoplasmic reticulum
FFT	:	Fast Fourier transform
GNM	:	Gaussian Network Model
HADDOCK	:	High Ambiguity Driven Docking
HDA	:	Host-directed Antivirals
HGF	:	Hepatocyte Growth Factor
HIV	:	Human Immunodeficiency Virus
M	:	Membrane

mAbs	:	Monoclonal Antibodies
MDs	:	Molecular Dynamics simulations
MM-GBSA	:	Molecular Mechanics Generalized Boltzmann Surface Area
MM-PBSA	:	Molecular Mechanics Poisson Boltzmann Surface Area
NMR	:	Nuclear Magnetic Resonance
NS	:	Non-Structural protein
ns	:	Nano Second
NTP	:	Nucleoside Triphosphate
PDB	:	Protein Data Bank
prM	:	Pre-Membrane
RNA	:	Ribonucleic Acid
RdRP	:	RNA dependent RNA polymerase
SARS-Cov-2	:	Severe Acute Respiratory Syndrome Coronavirus 2
scFv	:	Single Chain Fragment Variable
RMSD	:	Root Mean Square Deviation
VEGF	:	Endothelial Growth Factor

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Background of study

Of the emerging and re-merging viral infections, vector borne infections have raised global concerns as their incidence has been increasing accompanied by geographical expansion (Oliver et al., 2019). Arboviruses or arthropod-borne viruses are responsible for millions of infections on humans and animals every year (Huang et al.,2019). These arboviruses fall under the genus *Flavivirus* with more than 70 of them transmitted by mosquito or tick vector. Mosquito borne flaviviruses like Zika, dengue (DENV), Japanese encephalitis, West Nile and yellow fever are very well-known for causing diseases to humans (Holbrook, 2017).

Among these flaviviruses, DENV has been classified as the most widely and rapidly spreading arboviral disease in the world and being a burden from the aspect of disease and socioeconomy for countries in tropical and sub-tropical regions (Bhatt et al., 2013). DENV is now reported in more than 128 countries and infecting approximately 400 million of humans annually although it was once endemic in less than 10 countries in the 1970s (Laureti et al., 2018). In Malaysia, dengue outbreak was endemic since 1902 and become epidemic in 1973 with almost 1500 cases and 3.6% death rate (Shekhar & Huat, 1992). Suitable climate and high population density of humans and vectors put Asia and Americas at high risk for the infection while Asia alone has contributed to almost 70% of global infections (Bhatt et al., 2013).

In this study, the potential of existing DARPins in replacing antibody to interact with DIII of DENV-2 had been explored through MDs. Under MD simulations, the 3-dimensional structure of the docked conformation of DARPin and DIII generated could be refined and hence the affinity in binding accounted would be more comparable to the experimental results. Free energy of binding and the decomposition energy calculated

under MMGBSA protocol using the trajectories collected could identify potent DARPins in binding with DIII. Dynamics and flexibility of proteins were also explored under GNM to understand the correlation between mobility of proteins and their binding free energy as calculated by MMGBSA approach.

1.2 Problem statement

Since all dengue serotypes could cause hemorrhage (Fried et al., 2010) therefore a deeper understanding of the influence of dengue serotypes on disease severity including clinical manifestations become indispensable (Fried et al., 2010; Suppiah et al., 2018; Vicente et al., 2016). Severity of disease is associated with the production rate of antibodies while production rate of antibodies is dependent on the replication rate of the DENV serotypes. Numerous findings suggested DENV-2 has caused more severities (Thomas et al., 2008) like dengue hemorrhagic fever (DHF) (Fried et al., 2010; Nisalak et al., 2003; Vaughn et al., 2000), dengue shock syndrome (DSS) (Huy et al., 2013) and it was responsible for the incidence of severe dengue to increase in several regions (Dussart et al., 2012).

There is no antiviral drug approved for treatment. Despite a tetravalent vaccine (Vannice et al., 2016) is now available, however low efficacy was observed for DENV-2, and also age and serostatus are lowering the efficacy (Hadinegoro et al., 2015; Halstead & Russell, 2016; Villar et al., 2015). Considering the facts that global burden of dengue and the efficacy of vaccine, effort in finding antiviral for treating DENV infection is needed (Hadinegoro et al., 2015; Hammond et al., 2005). Although neutralizing antibodies bind specifically to target with high affinity (Reichert, 2001; Schrama et al., 2006) but they are structurally complicated and having issue like low tissue penetration (Waltz, 2007). These limitations have therefore promoted the development and search of other protein scaffold for substituting antibodies. Designed ankyrin repeat proteins

(DARPins) are one of the alternatives for replacing antibodies (Mosavi & Cammett, 2004; Stumpp et al., 2008) to bind with DIII DENV.

1.3 Objectives

The main theme of this study is to design DARPin that could bind to DIII DENV-2 mimicking the interactions mediated by mAbs using computational approach. Three main objectives have been established:

1. To understand the molecular interaction of DARPin upon binding with DIII DENV-2 by generating the docked complex under protein-protein docking method
2. To identify crucial residues in the binding interface and DIII DENV-2-DARPin interactions contribution, as well as the dynamics by analyzing the MD trajectories
3. To improve the binding affinity of DARPin towards DIII DENV-2 through computational site-directed mutagenesis and hence evaluating the *in-silico* binding activity of the novel engineered DARPin

Energetics and dynamics of the proteins were studied following the workflow as illustrated in Figure 1.1 while details of the methods are described later in chapter 3.

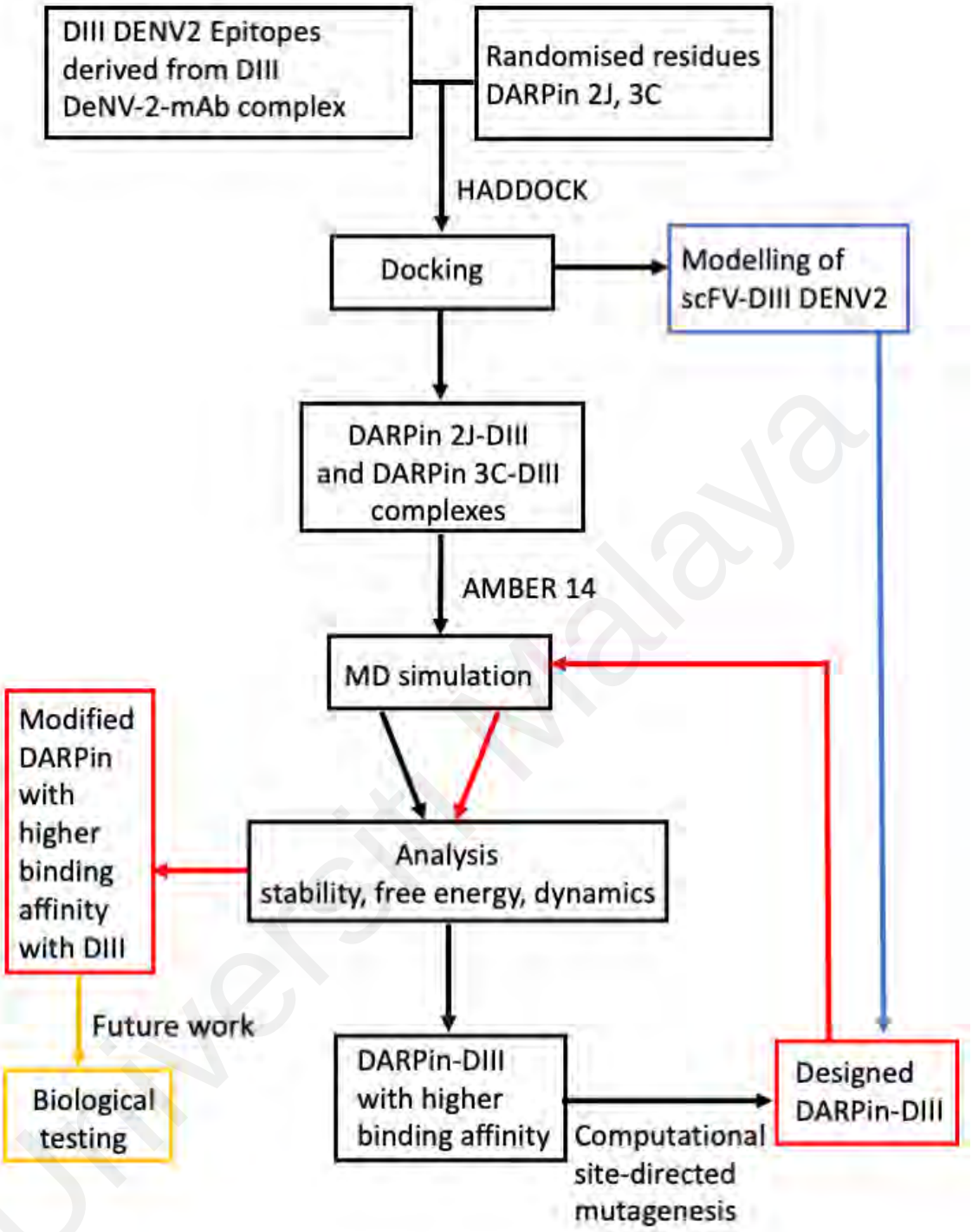


Figure 1.1 Workflow of study

1.4 Thesis outline and the connection of each chapter

There are 5 Chapters presented in this thesis. It begins with Chapter 1, an introduction that describes the problem statements and objectives apart from background study. Next, literature review in Chapter 2 summarizes previous research on dengue, including potential targets and drugs. Methods and approaches adopted in this study would be discussed in Chapter 3. Results and discussion in Chapter 4 explain the findings of the affinity of wild type DARPin and performance of engineered DARPins towards DIII DENV-2. In Chapter 5, conclusion and further works are discussed.

CHAPTER 2: LITERATURE REVIEW

2.1 Dengue virus (DENV)

There are four dengue serotypes and classification were made based on their antigenic properties. All the four dengue serotypes, DENV 1- 4 share a similar genome up to 65 %. The fatal complication that can cause by DENV includes DHF and DSS. However, dengue fever is the most common form of the disease. An infected person may develop symptoms such as fever, rash, headache, vomiting, arthralgia, myalgia and minor hemorrhagic manifestation (Hayes & Gubler, 1992) . Nonetheless, severity of infections varies according to individual immune response. People living in endemic area with more than one serotype present may be at a higher risk to develop DHF and DSS due to antibody dependent enhancement (ADE) phenomenon.

Apart from being enveloped, DENV has an icosahedral symmetry with diameter of 500 Å when it is matured and 600 Å when it is immature (Kuhn et al., 2002). The 11 kb length of positive stranded RNA contained in dengue viral genome has an important role in making structural and non-structural proteins. Structural proteins of DENV are capsid, membrane and envelope (E) while non-structural proteins included NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5.

Viruses are generally known as microscopic parasites and appeared to be much smaller than bacteria. They are one of the infectious agents with single type of nucleic acid contained in a protein coat called capsid and this capsid is surrounded by lipid bilayer in some viruses. To cause infection and diseases, pathogenic viruses would first bind to host cells, release viral content into the cells and start copying its viral components required for forming new viral particles (Lucas, 2010). Lifecycle of DENV was explain in Figure 2.1. DENV first enter the host cell through membrane fusion followed by replicating the

viral genome for protein translation and finally budding and leaving the cell after maturation.

Receptor-mediated endocytosis has been the way employed by the virus to enter the host cell. In acidic condition, E protein that anchored to the viral membrane underwent conformational changes by forming trimers and mediates the fusion between virus and membrane of host cell (Allison et al., 1995). Translation and replication of viral proteins occurs after viral genome is released to cytoplasm. Assembly of viral particles containing the viral genome and proteins take place in endoplasmic reticulum (ER) membrane. Host furin cleaved the pre-membrane (prM) protein to membrane (M) protein in Golgi vesicles for viral maturation. Matured DENV leaving the infected host cell via exocytosis (Tuiskunen Bäck & Lundkvist, 2013; Welsch et al., 2009).

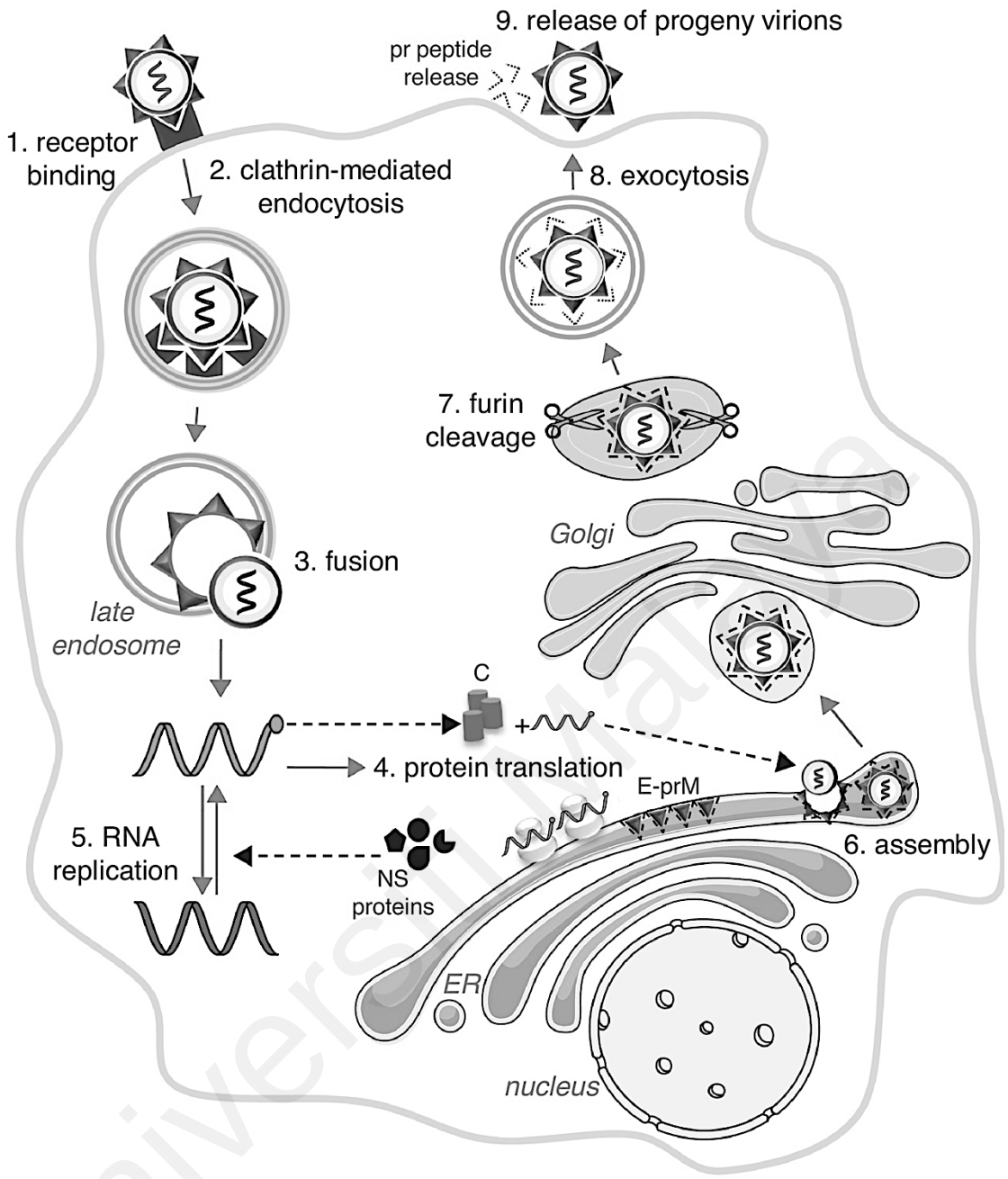


Figure 2.1 Life cycle of DENV. (Copyright and permission from Troost & Smit, 2020)

2.2 Targets for anti-dengue

2.2.1 Structural proteins

Structural proteins are the key protein for viral entry process. Capsid (C) protein is required for exocytosis process that releases matured virus from the infected cell. The C-terminal of C protein contains a hydrophobic region that attach itself to the ER membrane and cleaving prM protein to membrane (Nowak et al., 1989). Solved NMR structure of C protein suggested that highly basic region and a small part of the hydrophobic region of the C protein interacted with the genome of virus and lipid membrane, respectively (Ma et al., 2004). Among the four helices in C protein, α2 region was observed to be a potential target site (Nasar et al., 2020)

Besides being the precursor for membrane protein, glycosylated pre-domain coded by the hydrophilic N-terminal of prM protects E protein to have premature fusion under acidic environment of trans-golgi network (Guirakhoo et al., 1992). Hence, the cleavage process can be targeted for antiviral activity as cleavage of prM is vital for fusion of E protein to host cell membrane (Knobeloch et al., 2010).

E protein consists of three domains, DI, DII and DIII. DII has been the most targeted part owing to its role in receptor binding (Hung et al., 2004; Mazumder et al., 2007). An immunoglobulin like fold was found on DIII in which a feature of cell receptor. Furthermore, the most efficient inhibition on the viral attachment was observed when monoclonal antibody (mAb) bind with the epitopes on the DIII (Crill & Roehrig, 2001). Later, Roehrig (Roehrig, 2003) suggested that vital, virus type-specific, neutralization sites were located on DIII of flavivirus. Crystallographic study from Lok and colleagues (Lok et al., 2008) identified a total of 11 important residues on DIII of DENV-2 which served as a target site. Compiling the data from the mentioned literatures, it encourages

the exploration on DIII for inhibition purpose. Structure of virus particle and E protein of DENV was presented in Figure 2.2.

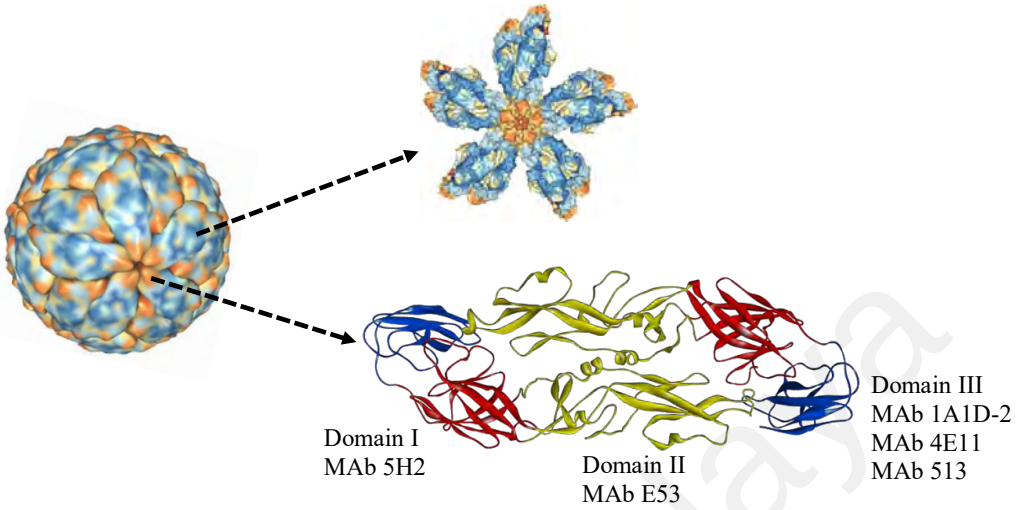


Figure 2.2 Structure of virus particle and E protein of DENV generated from X-ray crystal structure with PDB ID 3J27. Image of the dengue virion and the starfish like structure were produced using NGL viewer. Domains of E protein are labeled in the dimer structure.

2.2.2 Non-structural protein

Among the non-structural proteins, protease NS2B/NS3, helicase NS3 and methyltransferase and RNA dependent RNA polymerase (RdRP) of NS5 are more often to be targeted. Helicase NS3 unwind RNA (Chiang & Wu, 2016) for virus to replicate while nucleoside triphosphate (NTPase) is utilized in NTP hydrolysis to provide helicase activity energy (Wu et al., 2005). Interfering NS3 protease in cleaving precursor protein and host protease could stop viral protein synthesis and hence prevent virus from replicating. Later, NS2B was observed to have its role in facilitating NS3 protease activity (Clum et al., 1997; Pambudi et al., 2013). Methyltransferase NS5 plays its part in viral replication via methylation and capping of viral RNA.

2.3 Antivirals for dengue infection

Although approximately 200 viruses are known to-date, however, only nine of them are treatable with approved antiviral drugs (De Clercq & Li, 2016). There are peptides (Chew et al., 2017; Hrobowski et al., 2005; Lok et al., 2012; Prusis et al., 2013; Yin et al., 2006), small molecules (Marks et al., 2001; Modis et al., 2003) and also antibodies (Cockburn et al., 2012; Lok et al., 2008; Megret et al., 1992; Robinson et al., 2015) being explored and identified as potential anti-dengue by targeting the structural and non-structural proteins of DENV. Both structural and NS proteins of DENV serve as a target for antivirals of dengue since structural proteins involve in virus entry process while non-structural proteins play their role in viral replication. E protein is the most studied structural protein as it is associated with cell entry process (Lim, 2019; Lok et al., 2008) whereas NS5 (El Sahili & Lescar, 2017) and NS3 (Gu et al., 2000) are widely studied as they involve in viral replication process.

2.3.1 Small molecules and peptides

Combining several antivirals that interfere the replication cycle of virus (Yeo et al., 2015) improved the efficacy of anti-dengue in *in vitro* and *in vivo* studies (Chang et al., 2011; Yang et al., 2019; Yeo et al., 2015). There are two groups of antivirals for dengue, direct-acting antivirals (DAA) (Low et al, 2018) and host-directed antivirals (HDA) (Kaufmann et al., 2018). The classification was made according to function of the antiviral, whether they are interacting with the viral particles or host. Detailed list of antivirals for DENV was reviewed (Troost & Smit, 2020) and those anti-dengues tested *in vivo* or clinical trials were tabulated in Table 2.1 and Table 2.2.

Table 2.1 List of direct-acting antivirals

Drug	Target protein	Inhibition Mechanism	Reference
Compound-6	E	Fusion	(Wang et al., 2009)
PI-88	E	Virus entry / binding	(Lee et al., 2006)
ST-148	C	Capsid protein	(Byrd et al., 2013)
7DMA	NS5 RdRP	Viral replication	(Olsen, et al., 2004)
Balapiravir			(Chen et al., 2014)
NITD008			(Yin et al., 2009)
BG-323	NS5 MTase	Viral RNA capping	(Stahla-Beek et al., 2012)
Suramin	NS3 helicase	NS3 helicase	(Basavannacharya & Vasudevan, 2014)
ST-610		NS3 helicase	(Byrd et al., 2013)
NITD-618	NS4B	NS4B	(Xie et al., 2011)
Compound 14a	NS4B	IFN signaling	(Wang et al., 2015)

DAA targets specific viral protein with small toxicity but it could cause resistance to develop (Boldescu et al, 2017). Inhibitor for NS3 mainly targeting serine protease of NS2B/NS3. Meanwhile RdRP inhibitor, balapiravir, that targets NS5 protein is the only DAA proceeded to clinical studies despite a very large number of DAA were reported (Botta et al., 2018).

HDA functions as antiviral for DENV by disrupting the cellular pathway making it unfavorable for DENV to replicate (Kaufmann et al., 2018). The greatest advantages offered by HDA is that it is broadly applicable in targeting different arboviruses as these viruses require similar factors from the host for replication (Acosta & Bartenschlager, 2016). However, HDA could disrupt cell homeostasis (Boldescu et al., 2017; Kaufmann et al., 2018) process and hence it is not as effective as DAA despite low toxicity. Of the cellular pathway, α−glucosidase and inosine monophosphate dehydrogenase are targeted

as former is essential for protein folding (Low et al., 2018) while latter is important for biosynthesis of nucleotide (Botta et al., 2018).

Table 2.2 List of host-directed antivirals

Drug	Target protein	Inhibition Mechanism	References
Prochlorperazine	Dopamine receptor D2 antagonist Clathrin lattices formation	Clathrin-mediated inhibition	(Simanjuntak et al., 2015)
Bromocriptine	Dopamine receptor D2 and D3 antagonist	Replication	(Kato et al., 2016)
Chloroquine	Low-pH dependent entry steps and furin-dependent virus maturation	Fusion and maturation	(Randolph et al, 1990) (Tricou et al., 2010) (Farias et al., 2015)
Lovastatin	HMG-CoA reductase	Entry and assembly	(Rothwell et al., 2009) (Whitehorn et al., 2016)
AR-12	3-phosphoinositide-dependent kinase 1	Replication	(Chen et al., 2017)
Bortezomib	Proteasome pathway	Egress	(Choy et al., 2015)
NN-DNJ	α–glucosidase	Reduce E-protein and NS1	(Schul et al., 2007) (Wu et al., 2002)
UV-4 UV-4B UV-12 CM-9-78/ CM-10-18		Interfere viral glycoprotein folding	(Perry et al., 2013) (Warfield et al., 2016) (Warfield et al., 2015) (Chang, et al., 2011; Chang et al., 2013)
Prednisolone	Anti-inflammatory/ anti-hemorrhagic		(Nguyen et al., 2012)

Table 2.2, continued

Schisandrin A	STAT1/2 mediated antiviral interferon responses	Replication	(Yu et al., 2017)
Ketotifen Cromolyn Montelukast	mast cell modulator	Reduce vascular leakage	(St. John, 2013)

Although there were DAA (balapiravir) and HDA (UV-4B) made it to clinical trial, but undesirable outcomes have prevented them from being further tested. For a compound to be used as antiviral, it must be safe with low toxicity, stable and not likely to cause resistance to develop in viruses.

For peptides, they could be engineered using a short sequence of amino acids derived from viral protein so that binding of viral protein and host is prevented. Peptides designed from combining several amino acids from all domains of E protein (John et al., 2019) and stem domain of E protein (residue 412-444 and 419-447) (Hrobowski et al., 2005; Schmidt et al., 2010) have shown their ability to prevent host cells from being infected by DENV. Besides, peptide inhibitors have also been screened for their potential in binding with specific targets like NS1 protein combining both experimental and computational approaches (Songprakhon et al., 2020)

2.3.2 Monoclonal Antibodies

There are 82 approved monoclonal antibodies and the number continues to grow since the first therapeutic monoclonal antibody introduced in 1986 (Shilova & Deyev, 2019). Neutralizing antibodies have been found binding with E protein of flaviviruses (Sukupolvi-Petty et al., 2007). Antibodies that bind to the E protein was discussed in a review paper (Fibriansah & Lok, 2016) and presented in Table 2.3.

Table 2.3 mAbs that bind to monomer of E protein DENV

mAb	Source	Domain	DENV	PDB ID
1A1D-2	mouse	DIII	1, 2, 3	2R69 2R6P
4E11	mouse	DIII	1, 2, 3, 4	3UZQ 3UZV 3UZE 3UYP
2H12	mouse	DIII	1, 3, 4	4AL8 4ALA 4AM0
E111	mouse	DIII	1	4FFY 4FFZ
E104	mouse	DIII	2	3J8D
E53	mouse	DII	WNV and DENV 2	3I50 3IXX 3IXY
5H2	chimpanzee	DI	4	3UC0 3UAJ

1A1D-2 is complex-reactive antibody that manifests strong inhibition for DENV 1-3 (Sukupolvi-Petty, Austin, et al., 2007). Analyzing three-dimensional structure of complex containing Fab fragment of 1A1D-2 and recombinant DIII DENV 2 confirmed 11 epitope residues (K305-I312, Q325, P364, K388 and N390) and breathing structure of DIII

epitope was observed suggesting binding was influenced by temperature (Lok et al., 2008). Although amino acid sequence of both heavy and light chain of 4E11 and 1A1D-2 shares more than 80% similarity but 4E11 could neutralise all DENV serotypes (Megret et al., 1992). In addition, structural analysis of scFv 4E11-DIII DENV-2 revealed more than 11 epitopes residues which are important, 305-312 (A-strand), 323 (B-strand), 327 (BC-loop), 361-364 (DE-loop) and 385-391 (G-strand) (Cockburn et al., 2012). Later, one heavy chain residue and four light chain residues of 4E11 were re-designed using computational approach. Affinity for DIII DENV-2 and DENV-4 of the resulted mAb 4E5A had increased 15 folds and 450 folds, respectively.

Indeed, many of the mAbs bind to DIII are not fully cross-reactive and ability of neutralizing varied genotypes among all DENV serotypes was not shown. Using epitope-paratope connectivity network, 4E5A was further improved and to engineer mAb 513. Editing six amino acids in sequence mAb 4E11 and removal of residue 26 in the heavy chain makes mAb 513 bind with DIII DENV-3 and DENV-4 at higher affinity by 13- and 22-fold (Robinson et al., 2015) while affinity for DENV-1 and DENV-2 almost remained the same. This newly design mAb was also tested against 21 DIIIs of different serotypes and exhibited its highly neutralizing property with EC_{50} values less than 200 ng/ml.

2.3.2.1 Alternative scaffold to antibody

Despite antibodies bind specifically to target with high affinity to induce neutralizing effect (Reichert, 2001; Schrama et al., 2006), however ADE remains as the biggest challenge together with issues like structural complicated and low tissue penetration (Waltz, 2007) owing to their size ~150 kDa. Non-immunoglobulin protein scaffolds such as affibodies, affilins and DARPins (designed ankyrin repeat proteins) are alternatives of antibodies that have been very much explored for their therapeutics potentials as drug (Binz et al., 2005). Other than higher stability, these scaffolds distinguish themselves from monoclonal antibody (mAb) by their small size, range from 2-20 kDa in molecular weight (Vazquez-Lombardi et al., 2015). Ability of ankyrin repeat proteins in mediating protein-protein interactions (Andrade et al., 2001), together with their abundance in the nature (Bork, 1993) have catalyzed the construction of combinatorial libraries of DARPins under consensus design approach (Binz et al., 2003).

DARPin are genetically engineered antibody mimetic proteins typically exhibiting highly specific and high-affinity target protein binding. They are used as investigational tools, and diagnostic and therapeutic applications. A general structure of a DARPin with four ankyrin repeat motifs is shown in Figure 2.3. They are composed of stacked repeats containing 33 amino acids. Each repeats is formed by two antiparallel α-helices and a β-turn connecting to the next repeat (Binz et al., 2003; Stumpp et al., 2008). These repeats are flanked by constant capping regions, forming one contiguous polypeptide chain. Target protein binding involves the tips of the β-hairpins and the surface of the helical bundle facing the concave ankyrin groove (Krzywda et al., 2004). In addition, DARPins have framework and randomized residues (variable) like the constant domain and CDR region (variable) in antibodies, where editing amino acids in the variable regions would keep the protein structure remained unchanged. Six of the thirty-three amino acids of a repeat unit are randomized residues and susceptible for further design by replacing the

wildtype amino acids with other sixteen natural amino acids. Furthermore, complex containing DARPins and their target was found crystallized readily (Bukowska & Grütter, 2013; Gilbreth & Koide, 2012; Huber et al., 2007) and it facilitated DARPin engineering as detailed protein-protein interactions could be dissected at molecular level using the three-dimensional structure obtained from X-ray crystallography (Plückthun, 2015).

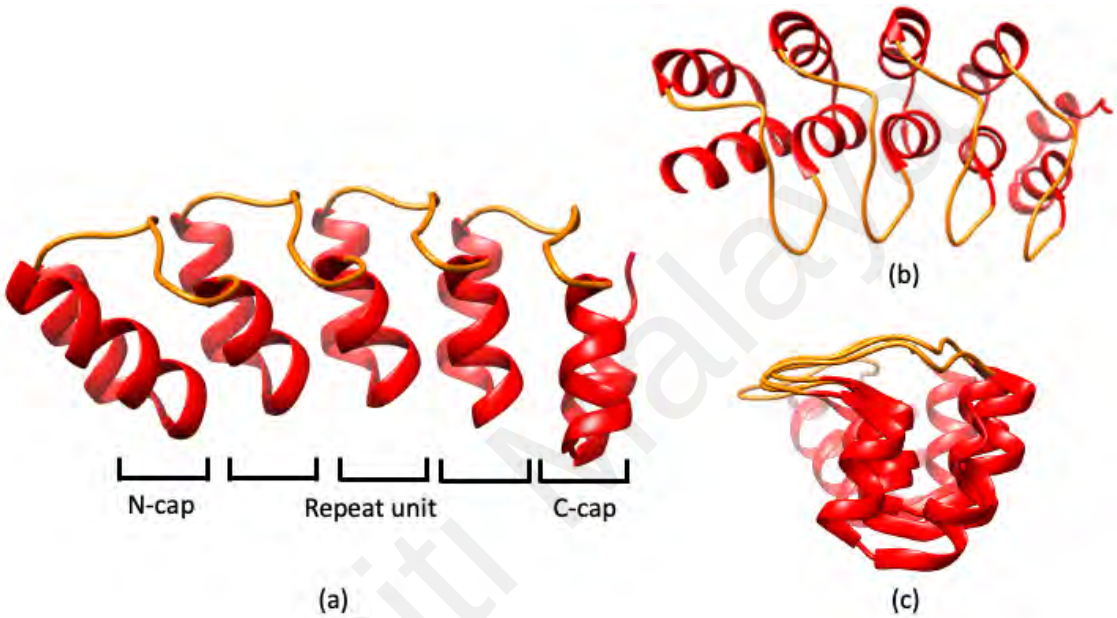


Figure 2.3 Framework of DARPin in (a) front (b) top and (c) side view using DARPin obtained from PDB 1SVX

Several sequencing projects have shown that relatively short, tandemly repeating motifs are common in many proteins (Heringa, 1998; Marcotte at al., 1999). Biochemical and structural characterization reveal that these repeats act as building blocks that stack side by side, forming the underlying architecture of a modular, specific protein-binding interface. Unlike globular proteins, the structure of repeat domains is dominated by local, short-range interactions, which represent a new paradigm for understanding the principles of protein stability and the mechanism of folding. In addition, a number of proteins with destabilizing mutations in their repeat domains have been implicated in several human

diseases. Advantages of using DARPin are that it displays tight specific binding at the nanomolar range to its target protein, and each repeat can contribute to target binding (Zahnd et al., 2007). This provides direct evidence that DARPins can bind rapidly and selectively to their intended target cells *in vivo* and is one promising avenue of research in which problems caused by protein mutation and alteration may be less severe. The repetitive modular architecture of DARPins allow the shuffling of repeats between repeat domains and adaptation of amino acid residues in its binding sites is simple to be designed (Binz et al., 2003; Kobe & Kajava, 2000; Marcotte et al., 1999).

DARPins could distinguish phosphorylated signal-regulated kinase (ERK) from non-phosphorylated upon binding and that make DARPins a potent biosensor as DARPins bind specifically for either form of ERK (Kummer et al., 2012). Besides, DARPins bound to many of the cancer-related proteins (Bery et al., 2019; Binz et al., 2017; Kramer et al., 2017; Steiner et al., 2008; Winkler et al., 2009) and binding affinity of DARPin for HER2 is very high and even overshadowed the approved antibody with higher specificity (Theurillat et al., 2010). Apart from being agent for cancer diagnosis, DARPins were engineered for cancer therapy too, whether interfere the cell signaling process or delivering toxic module (Shilova & Deyev, 2019). MP0274 that targets HER2 has started Phase I clinical trial in 2017 (Baird et al., 2018). MP0250 disrupts the signaling process involving endothelial growth factor (VEGF) and hepatocyte growth factor (HGF) (Binz, Bakker, et al., 2017) and demonstrated its strength in inhibiting VEGF in phase I clinical trial (Rodon et al., 2015) and proceeded to phase II since 2017. Abicipar that specifically targeting VEGF is now under phase III clinical study for chorio-retinal vascular diseases (Simeon & Chen, 2018).

There are specific DARPins which are anti-viral and have demonstrated inhibitory effect towards HIV (Nangola et al., 2012; Praditwongwan et al., 2014) and even SARS-

CoV-2 (Walser et al., 2020). MP0420 that recognizes three epitopes of the domain for receptor binding on spike has entered clinical trial phase I in November 2020.

Search for complexes containing DARPins that bound to membrane protein was done in Protein Data Bank since E protein dengue targeted by mAbs is a class II membrane fusion protein (Rey et al., 1995). Nine DARPins were studied for their interactions with membrane protein AcrB (Monroe et al., 2011) that is part of the multidrug efflux system in *Escherichia coli* (Tseng et al., 1999). Of the nine DARPins that were studied, x-ray crystal structure of three DARPins, PDB ID 2J8S (2J) and 3NOC (3C) and 3NOG were made available and deposited in Protein Data Bank. DARPin 3C has the greatest affinity for AcrB while DARPin 2J was the weakest among the three. DARPin 2J and 3C differ at 18 amino acid residues, involving capping units and repeat units as showed in Figure 2.4. Potential of these two DARPins were then evaluated for mimicking mAbs interacting with DIII DENV-2.

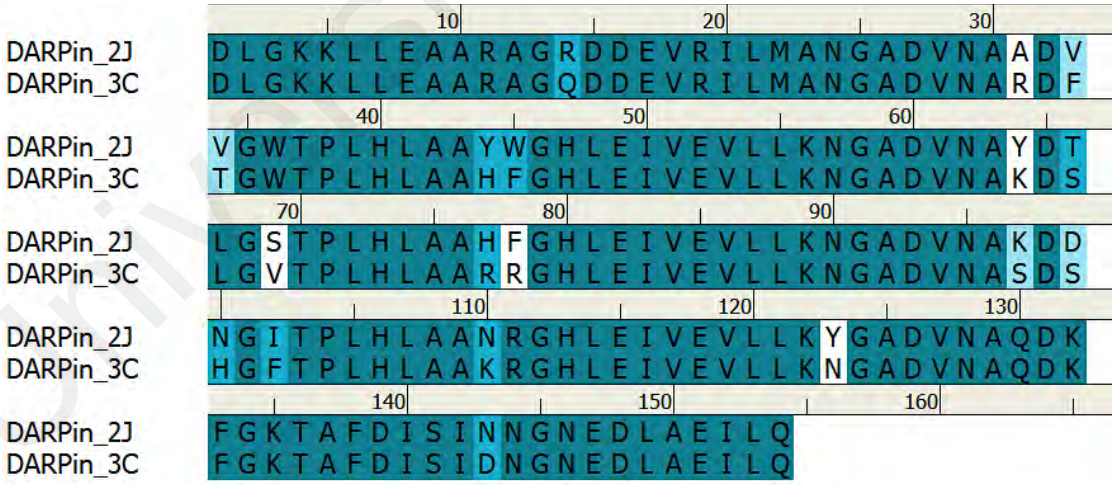


Figure 2.4 Sequence alignment of DARPin 2J and 3C

2.4 Computational aided drug design study

Availability of target protein and drug in their three-dimensional structure make structure-based design possible. In addition, major advances in high performance computers and software development would not only accelerate identification but also optimization in drug development process. Virtual screening, molecular docking and MDs are the most used computational approaches in structure-based drug design. Apart from structure based drug design, protein modelling and MDs have also successfully made potential drug available for its therapeutic application (Anderson, 2003; Clark, 2006; Rutenber & Stroud, 1996; Wlodawer & Vondrasek, 1998).

2.4.1 Protein-protein docking

For large, multidomain proteins in this study, computational oriented protein–protein docking would be rational to predict the binding complex from the two proteins (Rodrigues et al., 2013), mimicking the hierarchical nature of protein folding. The complexity of the problem is huge due to flexibility of proteins upon association and rotational-conformational space exploration when the two proteins interact. The field of protein–protein docking is highly computationally oriented, and shares approaches with small-molecule docking.

Docking algorithms available for protein-protein docking have been reviewed. Checking for surface complementarity using Fast Fourier transform (FFT) correlation-based method (Katchalski-Katzir et al., 1992) has been a fast and popular technique. Rigid-body docking algorithm of ZDOCK (Pierce et al., 2014) and HADDOCK (High Ambiguity Driven DOCKing) (Vangone et al., 2017; Vries et al., 2010) are FFT based while flexible docking algorithm offered by SwarmDock (Moal et al., 2018; Torchala et al., 2013) optimizes the atomic coordinates of proteins including orientation and position using population-based memetic method. Performance of available protein-protein

docking algorithms are often evaluated and benchmarked under Critical Assessment of Predicted Interactions (CARPI) community-wide experiment (Lensink et al., 2017; Vreven et al., 2015).

Prediction obtained from performing ab-initio docking is still questionable since experimental data has not been concerned (Rodrigues & Bonvin, 2014). Among the docking programmes reported, HADDOCK (Zundert et al., 2016) has included the interfacial water molecules during docking process (van Dijk & Bonvin, 2006). Unlike other docking programme, HADDOCK makes use of biochemical, biophysics or even bioinformatics data in docking the molecules instead of filtering the pre-generated structures to increase the efficiency of sampling and scoring (Rodrigues & Bonvin, 2014). Both the experimental and bioinformatics data can be integrated to improve the predicted outputs.

Ambiguous interaction restraints (AIRs) of the residues of different molecules are defined by computing the effective distance (d^{eff}). Of the residues, they are further defined by active of passive. Effective distances are restrained to 2Å, maximum. Any violations larger than 2Å, AIR forces would be constant. The docking process is completed in three stages, rigid docking by energy minimization driven by interaction restrains (it0), semi-flexible refinement in torsion angle space that permitting the atoms of the interface residues to move (it1), and lastly refinement in explicit solvent (water).

2.4.2 Molecular Dynamics Simulations (MDs)

Together with X-ray crystallography, nuclear magnetic resonance (NMR) and cryo-electron microscopy (EM) are the major techniques employed to solve the three-dimensional structure of protein-protein complex for protein-protein interactions study (Takemura, 2012). However, not all protein-protein interactions can be elucidated as solving structures for some protein-protein complexed can be challenging (Anishchenko

et al., 2014) and hence MDs have been an alternative to close the gap. MDs were not only being exploited in mapping the important binding residues on the binding surface (Hernandez-Rodriguez et al., 2016; Zhao & Caflisch, 2015), exploring the mechanism of drug upon association with it protein target but also enhancing the virtual screening steps (Durrant & McCammon, 2011).

Simulation time of MDs could possibly cause problem in sampling (Zuckerman, 2011). In order to improve the performance of sampling, multiple independent MDs begin with different parameters were attempted and proven to be useful (Grossfield et al., 2018; Grossfield & Zuckerman, 2009). Implementation of graphics processing units has accelerated MDs for biomolecular systems which are usually large and require longer computing time under CPU (Le et al., 2013).

MDs empowered by Newton's second law of motion numerically to study interactions of a system at atomic level. Changes of the atomic coordinates and velocities generated as a function of time resulted trajectory that describe the positions of atoms. Differential equations of the Newton's second law were solved for capturing the variation occurred with respect to time:

$$a_i = \frac{d^2 x_i}{dt^2} = \frac{F_i}{m_i} \qquad eq\ 2.1$$

At time t, Force (x_i) exerted on atom i with mass i has acceleration of ai. Potential energy as a function of position V (x_i) is obtained by calculating force (x_i) using the derivative of the expression under molecular mechanics. Properties of the trajectories collected from MDs could be compared with experimental findings. Overall energy of the system and forces of two atoms that come to interact are accounted by a force field like CHARMM (MacKerell et al., 1998) and AMBER (Cornell et al., 1995). Apart from

describing amino acids, parameters for nucleic acids, carbohydrates and lipids were also added to enrich the library.

Force field is computed by two terms, bonded and non-bonded interactions as found in expression below:

$$V = \sum_{i}^{bonds} \frac{k_{l,i}}{2}\left(l_i - l_{0,i}\right)^2 + \sum_{i}^{angles} \frac{k_{\alpha,i}}{2}\left(\alpha_i - \alpha_{0,i}\right)^2 + \sum_{i}^{torsions}\left\{\sum_{k}^{M} \frac{V_{ik}}{2}\left[1 + \cos\left(n_{ik} * \theta_{ik} - \theta_{0,ik}\right)\right]\right\} + \sum_{i,j}^{pairs} \varepsilon_{ij}\left[\left(\frac{r_{0,ij}}{r_{ij}}\right)^{12} - 2\left(\frac{r_{0,ij}}{r_{ij}}\right)^{6}\right] + \sum_{i,j}^{pairs} \frac{q_i q_j}{4\Pi\varepsilon_0\varepsilon_r r_{ij}} \qquad eq\ 2.2$$

Bonded interactions are described first three terms by computing how potential energy varies with respect to stretching, bending and torsions of bond between atoms directly involved in bonding relationships. l_0 and α_0 are reference value while k_l and k_α are force constants under harmonic potentials. Every dihedral angle of torsion term described by cosine series of M terms. Multiplicity of k_{th} term of the series defined by n_{ik}, with corresponding phase angle $\theta_{0,ik}$ and energy barriers V_{ik}.

Meanwhile, non-bonded interactions, are accounted by the fourth and fifth terms. Van der Waals interactions computed by the fourth term has 12-6 Lennard Jones potential. Of

the two interacting atoms, least energy distance that equivalent to the summation of van der Waal radii represented by $r_{0,ij}$ while ε_{ij} of the term describes the energy well. Fifth term computed electrostatic interactions using Coulomb potential. Partial charges of the atom pairs represented by q_i and q_j, permittivity free spaced denoted by ε_0 and dielectric constant represented by ε_r (ε_r=1 in vacuum).

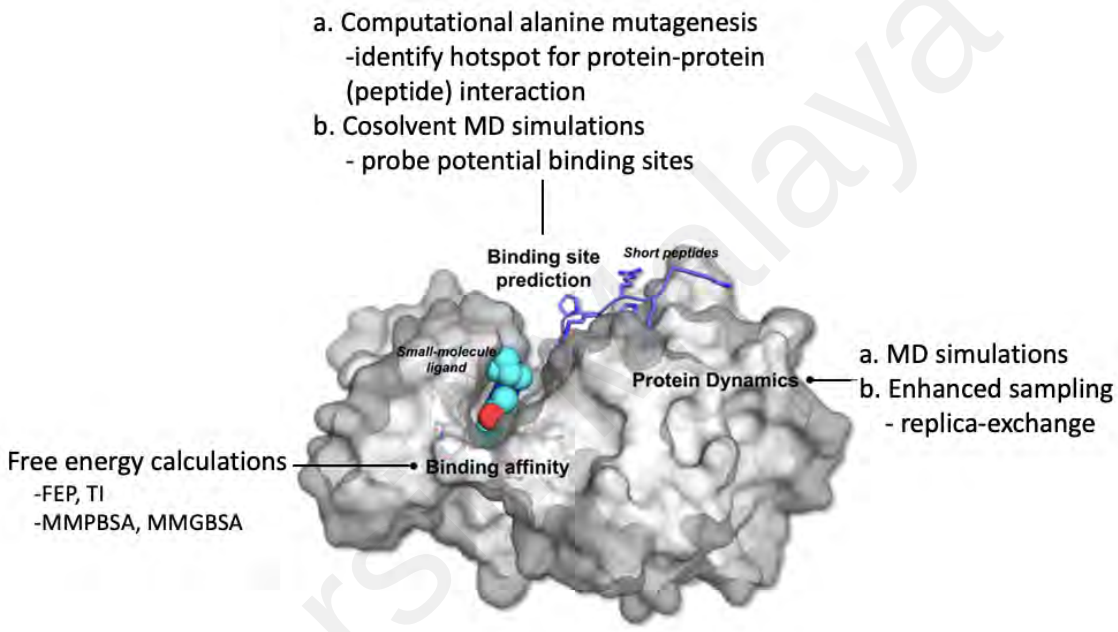


Figure 2.5 Utilization of MD simulations and enhanced sampling methods for protein-protein (peptide) and protein-ligand complexes.

Simulation techniques used to study protein upon binding with small molecule or other protein domain, including peptide was summarized in Figure 2.5. Under MDs and enhanced sampling methods, dynamics and structure of protein can be analysed in depth to understand the mechanism mediated interactions of protein with its targets (Lazim et al., 2020).

Among the algorithms in predicting binding affinity, molecular mechanics Poisson-Boltzmann (MM-PBSA) and molecular mechanics Generalised Born (MM-GBSA) are

preferred compared to free energy perturbation (FEP) and thermodynamic integration (TI) as they offer good balance between computing time and accuracy (Chen et al., 2016; Feng et al., 2017; Maffucci & Contini, 2016). Performance of MM-PBSA and MM-GBSA have been compared and evaluated for different parameters. Both methods were observed to be able to predict binding affinity of protein-peptide systems besides identifying the correct poses of binding (Weng et al, 2019). In an earlier study, MM-GBSA was observed to rank affinity of inhibitors better than MM-PBSA (Hou et al., 2011). In addition, MM-GBSA is more computational efficient considering computing time and power required for calculation.

2.4.3 Gaussian Network Model

Elastic network (EN) model has been introduced to probe the dynamics of protein-protein complexes. It addresses the shortcomings of MD simulations that consume times and relying on sophisticated architecture of computation. Among the EN models, Gaussian Network Model (GNM) is one of the most recommended models (Bahar et al., 1997) and it has been utilized to study the magnitude of motion of individual component of a structure and the connections between the motions and the component. *i*GNM (Li et al., 2016) is the online server available that compute the folded structure near native state and describe the structure under EN representation. Then, spectrum of motions could be decomposed into slow modes and fast modes. It is more efficient than normal mode analysis with additional simplicity.

Mode shape obtained from the slowest motions describes the global motions of the protein complex by plotting the residues square displacements of a particular mode against the residue index. The local minima observed from the global mode shapes correspond to hinge residues that coordinate the domain motions (Bahar et al., 1999; Bahar et al., 1997; Durrant & McCammon, 2011; Keskin et al., 2002; Wang et al., 2004).

Performance of GNM has been evaluated and benchmarked using experimental data like B-factor from X-ray crystallography (Kundu et al., 2002) and NMR data (Yang et al., 2007). Applications of GNM is not limited to protein-ligand system (Micheletti et al, 2002) but also a much larger system involve protein-protein complex (Haliloglu et al., 2008).

2.4.4 Computational mutagenesis

Protein-protein interactions are crucial for many biological processes and hence being targeted for drug discovery (Qiu et al., 2018). Crucial amino acid residues or hotspots at the interface of two interacting proteins in complex can be determined by performing computational alanine scanning (Ibarra et al., 2019).

Upon mutation of a protein residue in protein-protein complex, interactions mediated between the two proteins would change, such as folding free energy of interacting residues increase and missing of crucial non-covalent interactions (Gao et al., 2015; Sahni et al., 2015). The mCSM-PPI2 server uses machine learning method that utilizes graph-based signatures to describe the changes of environment of a protein when mutation occurs. The additional six features added for mCSM-PPI2 are (1) environment of wildtype residue including solvent accessibility, torsional variation, neighbouring amino acids of mutated residue, (2) nature of wild-type and mutant residue, identifying if glycine or proline are involved, (3) evolutionary conversation, (4) non-covalent interaction network analysis, (5) energy terms (prediction of folding free energy change and interaction energy change of the protein chains) and (6) atomic fluctuations. Even before the additional six features were added, quality of prediction of mCSM, the first generation of mCSM-PPI2, was already observed (Ibarra et al., 2019; Lazim et al., 2020). Method and workflow utilized by mCSM-PPI2 was illustrated in Figure 2.6.

On the other hand, BeAtMuSiC (Dehouck et al., 2013) is a webserver that allows rapid prediction on changes in binding affinity upon mutations for protein-protein complexes. It is a coarse-grained based algorithm with statistical potentials adapted. It remains as one of those top algorithms for prediction experiment CAPRI.

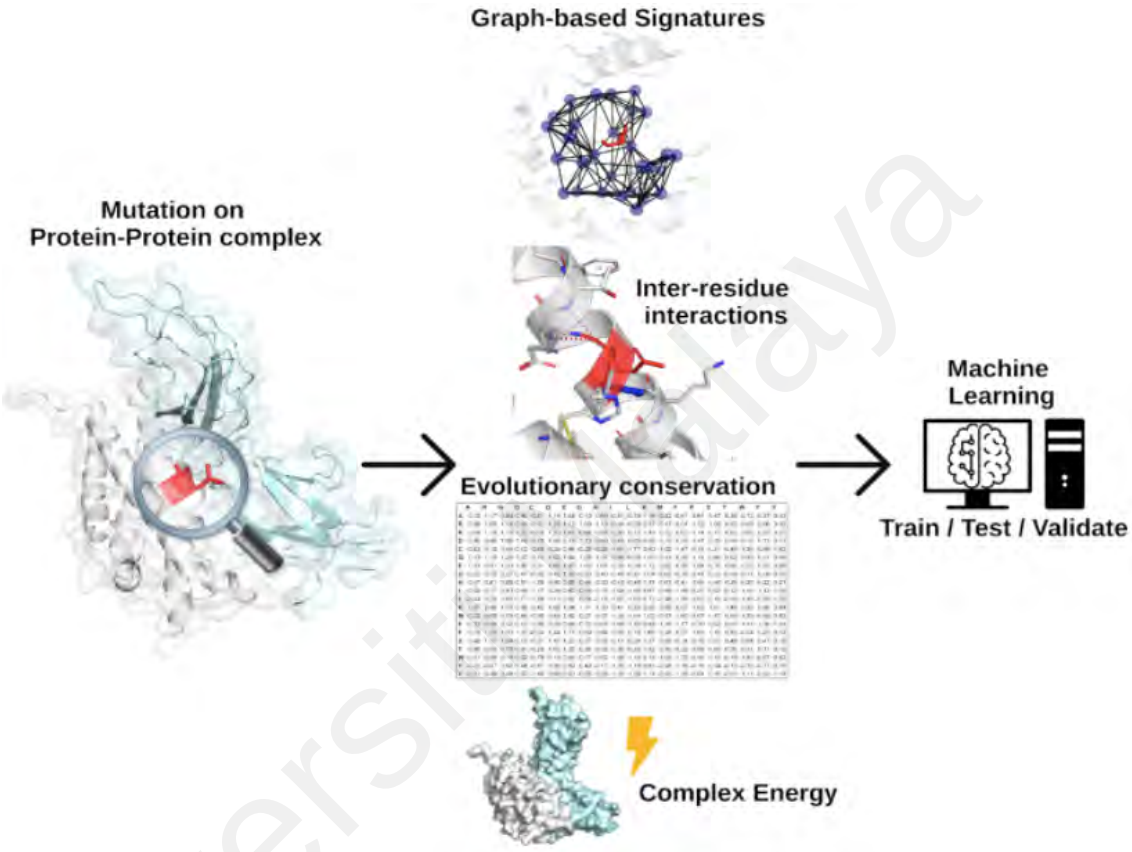


Figure 2.6 Workflow of methodology used by mCSM-PPI2. (Copyright and permission from Rodrigues et al., 2019)

CHAPTER 3: METHODOLOGY

3.1 Protein Preparation

3.1.1 Refinement of three-dimensional structure of DARPins and DIII

Two DARPins, termed as 2J and 3C were taken from the X-ray crystal structure deposited in Protein Data Bank with PDB ID 2J8S (2J) (Sennhauser et al., 2007) and 3NOC (3C) (Monroe et al., 2011) while DIII DENV-2 was taken from complex with PDB ID 2R29. The selected DARPins were found to be bound with membrane protein. Each of the protein structure was first cleaned by removing the water molecules, heteroatoms, alternate conformation of amino acids and followed by checking and correcting the bonds and bond orders under Discovery Studio (DS) programme version 3.0. Using Smart Minimizer protocol embedded in DS, all the three protein structures were minimized and reaching final RMS 0.01 kcal/mol. It is not unusual for atomic coordinates of the protein atoms misplaced and deviated from their original position during X-ray crystallization process. It may cause overlapping of the side chain of amino acids or distortion of bond lengths and angles. At the same time, steric clashes might present between the atoms. Hence, energy minimization comes into place to solve the problem.

3.1.2 Protein-protein docking

Minimized DIII and DARPins structures were submitted to HADDOCK web server for the protein-protein docking process (Vries et al., 2010; Zundert et al., 2016). Active residues of DIII were defined using the epitope residues recognized by the mAb, K305, F306, K307, V308, V309, K310, I312, Q325, P364, K388 and N390, while active residues on DARPin were defined using the randomized residues. A maximum of 200 docked conformations were used for clustering purpose after refinement in presence of water. The top conformation from the cluster with the largest cluster size was selected as starting structure for both DARPin 2J-DIII DENV-2 and DARPin 3C-DIII DENV-2 complexes.

3.1.3 Modelling scFv-DIII DENV-2 complexes

Altogether, three scFv-DIIIs and an antigen binding fragment (Fab) Fab-DIII DENV-2 complexes were prepared for MD simulations. Although these antibodies were capable of neutralizing DENV-2 with their binding activity reported, but their binding free energy under computational study remained unclear. Furthermore, some of the three-dimensional structures of mAb-DIII complexes are not available in the Protein Data Bank, hence they were modelled for MDs to obtain the binding free energy as reference value for comparison and assess affinity of DARPins for DIII.

X-ray crystal structure of Fab 1A1D-2-DIII and scFv 4E11-DIII were taken from protein data bank, with PDB ID 2R29 (Lok et al., 2008) and 3UZV (Cockburn et al., 2012). After cleaning the protein by removing all heteroatoms and alternate conformations of amino acids as described in section 3.1.1, gaps in each protein chain of the complexes were identified. Amino acids QN, GSE, KDS, HKTST were added to close the gap found in light chain of Fab 1A1D-2 between S126 -L129, D155 - R159, S172 – T176 and T201 - S203 accordingly. Linker (residue 118 – 131) connecting the two protein domains of all the three mAb and a small loop region (residue 343 - 345) of DIII DENV2 in PDB 3UZV were missing. Amino acids GGGGSGGGGSGGGG and ERK were added to fill the linker gap and loop gap, respectively. All the missing residues have been added to close the gaps using loop building function, a feature of UCSF Chimera programme, candidate version 1.14, (Pettersen et al., 2004) interface to Modeller version 9 (Fiser et al., 2000; Sali & Blundell, 1993). However, missing residues located at both end of the protein structures were not added. Altogether, 5 models were generated and the model with the lowest normalized discrete optimized protein energy (zDOPE) score (Shen & Sali, 2006) was selected. The zDOPE score is a statistical score calculated based on atomic distance.

ScFv 513-DIII DENV-2 complex was modelled using scFv 513 structure in PDB 5AAW (Robinson et al., 2015) and DIII structure in PDB 2R29. DIII DENV-2 taken from PDB 2R29 was superimposed to DIII DENV-4 in PDB 5AAW. Both scFv and DIII DENV-2 structures were kept in the complex while DIII DENV-4 was removed. Gap found between residue 116 and 130 was filled by amino acids GGGGSGGGGSGGGG. Whereas three-dimensional structure of scFv 4E5A-DIII (Tharakaraman et al., 2013) was built using X-ray crystal structure of scFv 4E11-DIII (PDB ID 3UZV). Amino acid sequence of the two scFvs varied at 5 positions, CDR-L1 R31K, CDR-L2 N57E, E59Q, S60W and CDR-H2 A55E. These five position of scFv 4E5A-DIII were modelled through computational site-directed mutagenesis using UCSF Chimera programme (Pettersen et al., 2004). Conformation of amino acids with the highest probability score was selected.

3.1.4 Protonation state prediction

Initial docked structure of the two DARPin-DIII DENV-2 and all the antibody-DIII complexes were submitted to PDB2PQR server for assigning protonation states of the protein residues (http://server.poissonboltzmann.org/pdb2pqr) at pH 7.0 under AMBER forcefield (Dolinsky et al., 2007; Dolinsky et al., 2004). The pqr outputs were then used for generating prmtop and inpcrd file under tleap module of AMBER14 programme (Case et al., 2014) with AMBER forcefield ff14SB (Maier et al., 2015) loaded to describe the molecular properties of the proteins. Using tleap, all the complexes were solvated in an octahedral box of TIP3P water and neutralized by adding counterions Na+ or Cl- ions.

3.2 MDs of protein-protein complexes in water

MDs were performed using a time step of 2 fs. Long range interactions were taken into account using periodic boundary condition based on the Particle Mesh Ewald (PME) method with a non-bonded cutoff of 12 Å were used. The SHAKE algorithm was applied to constrain the bond with the presence of hydrogen and Langevin dynamics was

employed to control the temperature. Large forces of an initial structure that is far from equilibrium could stop the MDs process from running. Hence, all protein-protein complexes underwent energy minimization to avoid steric clashes between atoms. Determining the number of steps required for minimization in the search of global minima could be challenging as problem arise if the steps of minimization is insufficient or excess (Zhang, 2015). A common protocol to ensure sufficient local optimization was discussed, protein is first solvated, and only the solvent is optimized while solute is hold and fixed using a large force. Heavy atoms are then hold, followed by CA atoms, and eventually the whole solvated complex is optimized with all restrains removed. All the complexes underwent two minimization cycles.

First minimization worked under constant volume, involved a total of 1000 steps, beginning with 500 steps of steepest descent and followed by another 500 steps of conjugate gradient, using cut off radius 12 Å. At this stage, all protein residues were fixed using a force constant of 100 kcal/mol $Å^2$. Second minimization worked under the same parameters except the force constant was removed, allowing proteins and water molecules move freely. The output structure after minimization steps were then proceeded with MD simulations for 100 ns. The starting temperature of each system was 0 and being raised to 310 K within 60 ps of NVT dynamics. During the heating step, the force constant that hold the protein fixed in position had been reduced to 10 kcal/mol. Prior to NPT-MD for 100 ns, the complexes underwent NPT equilibration at 310.15 K and 1 atm pressure for 1ns. All complexes were simulated using PMEMD.CUDA (Goetz et al., 2012; Le Grand et al., 2013; Salomon-Ferrer et al., 2013) from AMBER14 (Case et al., 2012) on GPUs using TITAN graphic card manufactured by NVIDIA that runs on a node with 32 CPUs and 54 GB memory. Computing time required to collect 100 ns long of MDs on DARPin-DIII system was less than a day with average speed > 100 ns/day.

3.2.1 Stability of protein complexes from MDs

Trajectories collected from the production run were analysed using CPPTRAJ (Roe & Cheatham III, 2013) module as embedded in AMBER14 programme. Potential energy of the system was plotted in order to examine if the changes along the simulation time arrived at equilibrated state to rule out the possibility of having inappropriate simulation protocol such as time step, parameters as well as initial structure (Simmerling et al., 2002). Aided by perl script, all the output files of production phase from 1 ns to 100 ns were taken to make the plot pictured in Figure 3.1.

The root mean square deviations (RMSDs) of all Cα-atoms of protein complexes were calculated with using the initial MD structure as reference over 100 ns to account for the structural stability and conformational change during dynamics process. A root-mean-square deviation (RMSD) plot pictures how a protein deviates from its initial conformation steadily along the simulations. When the atomic coordinates of two structures are superposed for RMSD calculation, the differences of the structures and dynamics is hence revealed. Analysing RMSD of MD trajectories collected allows clustering of similar conformations and more importantly reflecting the equilibration period as trajectories extracted from this period are more reasonable for free energy calculations (Zhong et al., 2013). RMSD was computed according to equation 3.1:

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{i=0}^{N}[m_i\ (X_i - Y_i)^2}{M}} \qquad eq.\ 3.1$$

N denotes the number of atoms, m_i denotes the atom mass, X_i denotes the coordinate vector for target atom i, Y_i is the coordinate vector for reference atom i while M account the total mass.

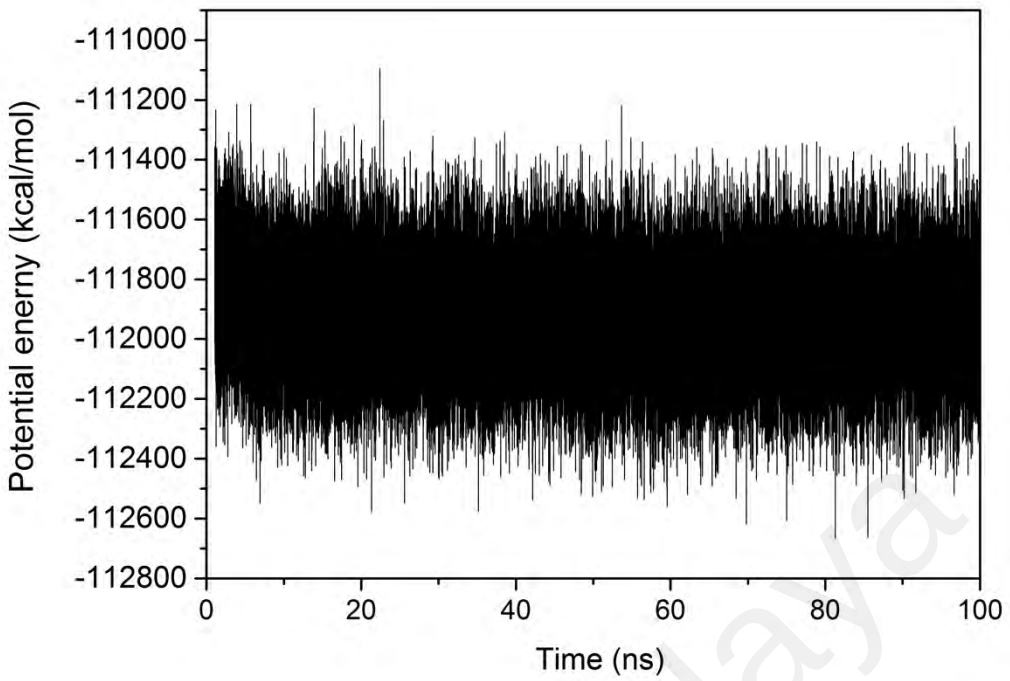


Figure 3.1 Potential energy of DARPin 2J-DIII DENV2 complex along 100 ns of simulation time

3.2.2 Free energy and hydrogen bond calculations

Molecular Mechanics-Generalised Born Solvent Area (MM-GBSA) (Onufriev et al., 2004) method has been used to compute the binding free energy and decomposed binding free energy under MMPBSA.py module (Miller et al., 2012). Binding free energy (ΔG_{bind}) was calculated by using the equations shown below (Çınaroğlu & Timuçin, 2019) :

$$\Delta G_{bind} = G_{complex} - (G_{receptor} + G_{ligand})$$

$$= \Delta H - T\Delta S$$

$$= \Delta E_{MM} + \Delta G_{sol} - T\Delta S$$

$$\Delta E_{MM} = \Delta E_{int} + \Delta E_{ele} + \Delta E_{vdw}$$

$$\Delta G_{sol} = \Delta G_{polar} + \Delta G_{non\text{-}polar}$$

From the first equation, difference in free energy between the complex and the receptor and ligand would give total binding free energy, ΔG_{bind}. At the same time, ΔG_{bind} can be obtained by subtracting entropy ($T\Delta S$) from enthalpy (ΔH) or summation of molecular

mechanical energy (ΔE_{MM}) and solvation free energy (ΔG_{sol}). Although enthalpy could be computed by MM-GBSA however calculating entropy is computationally expensive and hence entropy was not calculated. ΔE_{MM} can be further decomposed into intra-molecular energy (ΔE_{int}), electrostatic energy (ΔE_{ele}) and van der Waals interactions (ΔE_{vdw}) while solvation free energy (ΔG_{sol}) is the sum of polar contribution (ΔG_{polar}) and non-polar contribution ($\Delta G_{non\text{-}polar}$). Under MM-GBSA calculation, generalized Born model is used to account ΔG_{polar} term while SASA based LCPO algorithm is applied to calculate $\Delta G_{non\text{-}polar}$ term.

Altogether, 1000 snapshots were extracted from the last 20 ns of production trajectories for free energy calculation. In order to map the important residues that are contributing for protein binding, both DARPin and DIII residues found within 4Å of binding vicinity were included for pairwise residue interaction energy calculations. Interaction of one residue with all other residues in 4Å was calculated. Apart from energy calculations, hydrogen bond (H-bond) was also accounted to understand the intermolecular interactions of the proteins in depth. Hydrogen bonds with occupancy percentage lower than 30% were excluded for discussion. Energy calculations and trajectory analysis were running on a node that has 64 CPUS and 249 GB memory.

3.3 Protein analysis using elastic network model

Final structure of the complexes obtained from MDs subjected to GNM server *i*GNM (Li, et al., 2016) for computing the global mode shape. Structure of the complex was treated as an elastic network and denoted by the α-carbon (Bahar et al., 1997; Haliloglu et al., 1997). The α-carbon of each residue is treated as a node in GNM and connected to other nodes by springs within a cut-off distance at 7.3 Å (Bahar et al., 1997; Kundu et al., 2002). Kirchhoff matrix at dimension N x N is used to describe the network topology with -1 assigned for the off-diagonal elements if nodes *i* and *j* were joined, else 0 would

be assigned. Summation of each of the column would bring the number to zero. The cross-correlations of the residue fluctuations, ΔRi and ΔRj of *i* and *j* obtained from (Bahar et al., 1997) based on the statistical mechanics theory which applied for polymer gel network (Flory et al., 1976):

$$<\Delta Ri \,.\, \Delta Rj> = (3k_BT/\gamma)\,[\Gamma^{-1}]_{ij} \qquad \text{eq. 3.2}$$

where k_B refers to Boltzmann constant, T refers to absolute temperature and $[\Gamma^{-1}]_{ij}$ refers the ij^{th} element of Γ inverse. N-1 modes defined the equilibrium dynamics of structures (Haliloglu, Bahar, et al., 1997) whereas k^{th} eigenvalue (u_k) defining the residues displacements along the k^{th} mode and the eigenvalue, λ_k scaled with its frequency. k^{th} mode contributed to the mean-square fluctuation of residue *i* in such a way:

$$<<\Delta R_i>>_k = (3k_BT/\gamma)(1/\lambda_k)\,(u_k\, u_k^T)_{ii} \qquad \text{eq. 3.3}$$

where $(u_k\, u_k^T)_{ii}$ denotes the i^{th} diagonal element of u_k.

Global mode shape of the complexes was compared, particularly the binding regions of the two interacting proteins and hence the dynamics of the residues were studied.

3.4 Further design on DARPin through computational mutagenesis

Approach in constructing DARPin library including modification of the randomized residues within the internal repeats of DARPins had been described (Binz et al., 2003). Each repeating unit consists 33 residues, with 27 framework residues and 6 randomised residues (Seeger et al., 2013; Tamaskovic et al., 2012) that allow further modification or replacement with other amino acids, as shown in Figure 3.2. 2J has 3 repeating units, altogether 18 residues that were possible for further design to improve the interactions with DIII DENV-2.

3.4.1 Single point mutation on DARPin

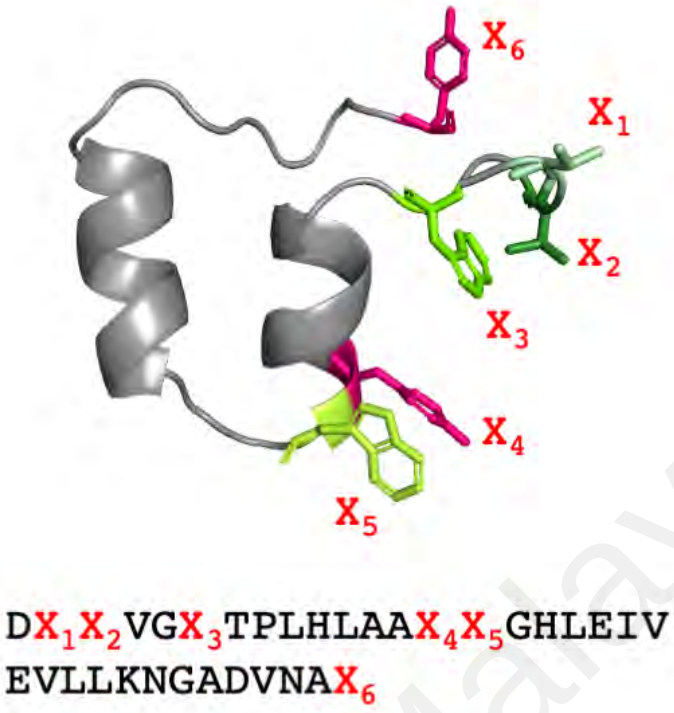


Figure 3.2 Repeating unit of DARPin showcases the location of all the six randomised residues for further design

Single point mutation was first performed to evaluate the potential of the 18 randomised residues for mutation. Table 3.1 summarises the number and amino acid of the 18 residues that were subjected to mutation study using two different algorithms while Figure 3.2 showing the location of the six randomized residues on a repeat unit. The final structure of 2J-DIII obtained from MD simulation was submitted to mCSM-PPI2 (http://biosig.unimelb.edu.au/mcsmpp2/) (Rodrigues et al., 2019) and BeAtMuSiC server (http://babylone.ulb.ac.be/beatmusic) (Dehouck, Kwasigroch, et al., 2013) for predicting the effect of single point mutation by replacing the wild type amino acid with other natural amino acids except glycine, proline and cysteine. Output generated from mCSM-PPI2 server describes the affinity of binding of mutants with details such as distance from closet partner, predicted $\Delta\Delta G^{Affinity}$, interactions change upon mutation apart from predicted affinity as shown in Figure 3.3 (a) while output obtained from BeAtMuSiC

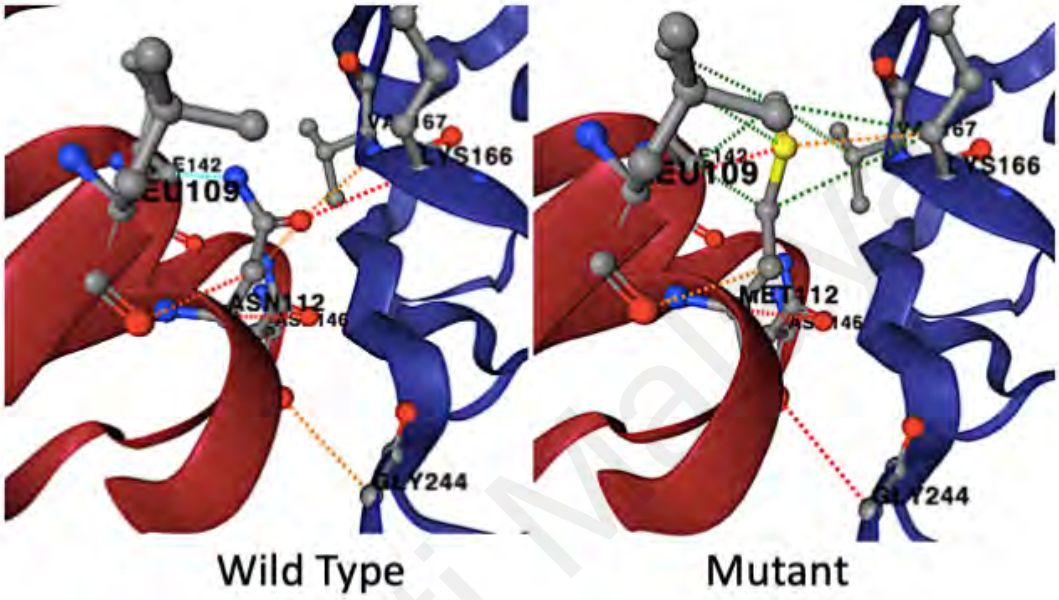
server is much simpler as shown in Figure 3.3 (b). Residues predicted with increasing affinity were first identified and cross-examined with results obtained from both servers. Mutation points predicted with increasing affinity were first identified.

Table 3.1 Randomised residues of each repeating unit of DARPin 2J

Repeating unit	Residue Number	Randomised residue
1	34-66	V35, V36, W38, Y46, W47, Y66
2	67-99	T68, L69, S71, H79, F80, K99
3	100-132	D101, N102, I104, N112, R113, Q132

a

#	Chain	Wild Type	Position	Mutant	Distance to interface	Predicted $\Delta\Delta G^{Affinity}$ (kcal/mol)	Affinity
31	A	ASN	112	ALA	2.981	-0.867	Decreasing
32	A	ASN	112	MET	2.981	0.181	Increasing



b

Chain(s)	Position	Mutation	$\Delta\Delta G_{Bind}$ (kcal/mol)	Solvent accessibility (in partner(s))	Solvent accessibility (in complex)	Interface
A	112	N → A	1.59	40.48 %	2.42 %	✓
A	112	N → C	-0.20	40.48 %	2.42 %	✓

Figure 3.3 Results obtained from (a) MCSM-PPI2 and (b) BeAtMuSiC server upon single point mutation

3.4.2 Verification of the predicted affinity under MD simulations

Further designed DARPin-2J-DIII DENV-2 complexes with single point mutation were modelled using rotamer function (Shapovalov & Dunbrack, 2011) embedded in UCSF Chimera programme (Pettersen et al., 2004). Rotamer with highest probability for each mutation was chosen as the initial three-dimensional structure. Each mutant complex

was then minimized prior to MD simulations following the same criteria as described in earlier section. RSMD was computed to examine the stability of the complexes throughout the simulation time followed by binding free energy calculations.

3.4.3 Two-point-mutation on DARPin

Complexes with binding free energy lower than the wild type were identified. Subsequently, two-point-mutation were attempted by mutating two randomized residues of DARPin 2J-DIII complex, in which these single point mutations were reported with better binding affinity from MDs. Initial structures of the DARPin-DIII DENV2 with double point mutation were modelled under UCSF Chimera programme and later simulated under the same protocol and being analysed using same post analyses.

3.4.4 Root mean square fluctuations (RMSF)

Flexibility of proteins after point mutation was described by calculating the Cα of all the protein residues of wild type and mutants. Residues with flexibility increased after mutation could affect the properties of DARPin especially stability of the protein as protein stability is associated with protein function and activity (Khan et al., 2019)

CHAPTER 4: RESULTS AND DISCUSSION

4.1 DARPins in complexed with DIII DENV-2 through molecular docking

Predicted three-dimensional structures of DARPin-DIII were generated using molecular docking protocol under HADDOCK. From 1000 poses, HADDOCK proceeded with the top 200 models for clustering and scoring processes. All the 200 models underwent refinement step using explicit water so that the interface has its energetics improved. It is crucial for the resulting docked models in term of proper scoring (Dominguez et al., 2003). Docked conformations with deviation less than 2 Å from each other were placed in the same cluster.

Altogether 185 structures and 189 structures of 2J-DIII and 3C-DIII complexes being clustered under HADDOCK, giving 11 clusters and 9 clusters each. HADDOCK arranged the clusters of a docking system according to their haddock score, from lowest to the highest. HADDOCK score was calculated by combining several energy terms like vdW intermolecular energy (E_{vdw}), electrostatic intermolecular energy (E_{elec}), desolvation energy (E_{desolv}) and distance restraint energy (E_{air}) and hence suggesting stability upon association of the two proteins. The four energy terms carried different weightage according to the equation 4.1.

$$HADDOCK\ score = 1.0 * Evdw + 0.2 * Eelec + 1.0 * Edesol + 0.1 * Eair \quad eq.4.1$$

Negative HADDOCK score is preferred as it accounts the magnitude of free energy decreased for docking the two proteins, and hence a more negative value indicating the generated docked structure is more energetically favorable. Cluster size of the top cluster of 2J-DIII and 3C-DIII was 51 and 64, while their second cluster containing 6 and 55 structures. Apart from lowest HADDOCK score, top ranked cluster based on the HADDOCK score has the largest cluster size and the most negative Z-score. Since Z-score accounts how many standard deviations from the average this cluster is in term of

score, hence the more negative the Z-score the better the results. Summary of the docking scores for both systems, 2J-DIII and 3C-DIII were tabulated in Table 4.1 and 4.2.

Cluster 1, the top cluster of 2J-DIII, has the largest number of docked models and also the lowest HADDOCK score, -152.9 and Z-score, -1.5. Apart from the lowest scores, intermolecular energy of this cluster also lowest among other clusters, -377.1 kcal/mol, with vdW energy -69.3 kcal/mol and electrostatic energy -307.8 kcal/mol. HADDOCK score of the second cluster, cluster 8 was very close to cluster 1, -151.8, as well as Z-score, -1.4. However, intermolecular energy of the models in cluster 8, -306.7 kcal/mol, was way higher than cluster 1. Cluster size of cluster 8 was only 6, much smaller compared to 51 models of cluster 1. Plus, restraints violation energy of cluster 8 was higher than that of cluster 1, hence the cluster 1 would be further considered for being selected as the representative structure for 2J-DIII. Other clusters of 2J-DIII suffered from high violation energy, high intermolecular energy and also small cluster size.

Like 2J-DIII, cluster 1 for 3C-DIII was ranked as the top cluster due to its lowest HADDOCK score, 163.2 and Z-score -1.4. Again, second cluster, cluster 2 has a very close HADDOCK score to cluster 1, -160.8. Cluster 1 has the largest cluster size with 64 models while cluster 2 has 9 model lesser than cluster 1. Nevertheless, cluster 1 has a higher intermolecular energy and restrains violation energy to that of cluster 2 in this case.

Table 4.1 Docking scores of 2J-DIII complexes

Cluster	Parameters								
	Haddock Score	Cluster Size	RMSD	VdW	ELE	Desolvation Energy	Restraints Violation energy	Buried Surface Area	Z-score
1	-152.9 ± 6.3	51	13.0 ± 0.1	-69.3 ± 1.8	-307.8 ± 33.8	-28.1 ± 3.9	60.8 ± 21.58	2238.7 ± 37.5	-1.5
8	-151.8 ± 8.0	6	0.6 ± 0.4	-74.7 ± 4.1	-232.0 ± 48.5	-39.7 ± 3.6	89.5 ± 39.69	2247.7 ± 63.7	-1.4
7	-142.9 ± 4.9	6	6.6 ± 0.2	-57.5 ± 6.0	-290.0 ± 10.1	-31.6 ± 2.1	42.0 ± 15.56	2026.3 ± 76.7	-0.6
5	-140.7 ± 6.3	17	5.9 ± 0.1	-59.2 ± 6.8	-295.2 ± 30.0	-30.3 ± 8.0	77.7 ± 16.57	2186.0 ± 82.3	-0.5
3	-139.8 ± 2.8	25	13.5 ± 0.1	-66.4 ± 5.3	196.4 ± 31.3	-46.7 ± 7.5	125.5 ± 42.23	2117.3 ± 43.2	-0.4
2	-135.0 ± 1.0	35	12.7 ± 0.3	-55.1 ± 7.0	-278.4 ± 27.3	-34.7 ± 4.0	104.5 ± 34.20	1958.6 ± 677.7	0
11	-128.5 ± 4.1	4	8.1 ± 0.3	-62.3 ± 3.4	-225.5 ± 48.6	-34.8 ± 6.1	136.0 ± 12.92	1966.5 ± 23.2	0.6
6	-121.4 ± 8.7	9	11.2 ± 0.3	-47.9 ± 9.3	-216.5 ± 26.5	-35.1 ± 5.0	48.7 ± 4.24	1936.9 ± 136.2	1.2
9	-121.3 ± 17.6	5	10.0 ± 0.3	-58.2 ± 4.9	-222.4 ± 52.5	-30.5 ± 1.5	119.4 ± 43.23	1810.0 ± 88.9	1.2
4	-119.1 ± 11.0	22	13.4 ± 0.3	-57.9 ± 4.6	-196.1 ± 35.9	-30.1 ± 9.6	81.5 ± 22.55	1948.7 ± 132.1	1.4

Table 4.2 Docking scores of 3C-DIII complexes

Cluster	Parameters								
	Haddock Score	Cluster Size	RMSD	VdW	ELE	Desolvation Energy	Restraints Violation energy	Buried Surface Area	Z-score
1	-163.2 ± 9.5	64	0.8 ± 0.4	-84.7 ± 6.0	-174.6 ± 50.5	-52.8 ± 9.7	91.9 ± 47.81	2364 ± 70.3	-1.4
2	-160.8 ± 4.1	55	6.3 ± 0.6	-68.3 ± 4.1	-297.8 ± 60.8	-37.7 ± 8.2	47.8 ± 22.88	2255.3 ± 76.5	-1.2
4	-153.1 ± 4.6	22	14.9 ± 0.1	-57.9 ± 4.1	-280.5 ± 48.9	-46.0 ± 6.8	68.7 ± 27.82	1990.2 ± 128.5	-0.7
5	-152.4 ± 8.6	7	12.6 ± 0.2	-72.0 ± 4.5	-207.3 ± 12.7	-47.7 ± 4.2	87.2 ± 29.17	2209.1 ± 43.3	-0.7
6	-140.5 ± 7.2	5	14.1 ± 0.4	-56.5 ± 6.5	-231.6 ± 35.1	-50.2 ± 7.2	124.7 ± 40.26	1925.9 ± 103.1	0.1
9	-134.2 ± 7.7	4	5.6 ± 0.5	-62.9 ± 5.9	-203.6 ± 40.1	-35.2 ± 5.8	45.6 ± 25.93	2063.5 ± 100.5	0.5
8	-128.4 ± 3.5	4	4.1 ± 0.2	-49.0 ± 6.3	-232.6 ± 17.9	-41.2 ± 5.5	83.3 ± 48.18	1773.6 ± 130.6	0.9
3	-123.3 ± 7.7	24	14.1 ± 0.1	-48.4 ± 4.6	-201.8 ± 40.0	-48.1 ± 6.5	134.8 ± 23.23	1776.9 ± 99.3	1.2
7	-120.0 ± 16.9	4	13.5 ± 0.2	-58.9 ± 1.8	-176.2 ± 43	-36.0 ± 9.8	101.9 ± 19.7	1920.5 ± 122.7	1.4

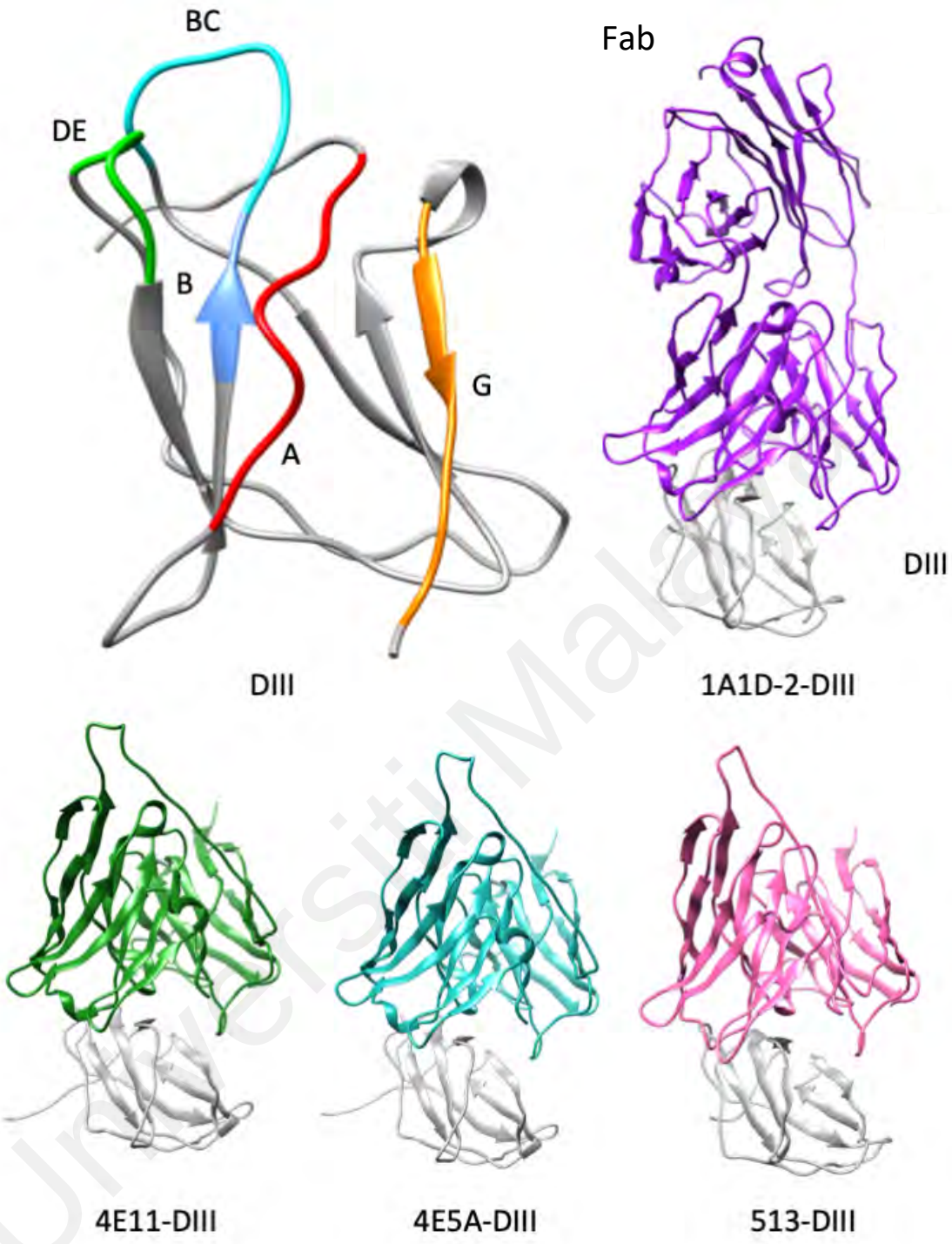


Figure 4.1 Minimized X-ray crystal structures of DIII, Fab-DIII and scFv-DIII complexes. Epitopes of DIII were highlighted and labeled.

Structural analysis of the three-dimensional structure of Fab-DIII and scFv-DIII complexes revealed the paratope of antibodies and epitope of DIII for binding affinity as illustrated in Figure 4.1. X-ray crystal structure of 4E11-DIII was analysed (Cockburn et al., 2012) and it showed that CDR of heavy-chain domain positioning above A-strand while CDR of the light-chain domain lies above G-strand of DIII. Similar observation obtained from minimized X-ray crystal structure of 1A1D-2-DIII, 4E5A-DIII and 513-DIII as depicted in Figure 4.1. Overall, variable domain of the heavy chain was not making contact with A-strand alone but also other regions, B-strand, BC-loop and DE-loop. Hence, interactions mediated between Fab 1A1D-2 and DIII as found in X-ray crystal structure has been used as a reference in judging the docking conformations of DARPin and DIII.

Docked conformation of the top model from each cluster of 2J-DIII system were shown in Figure 4.2 to showcase if DARPin could lie above A-, B- and G- strand, BC-loop and DE-loop. DARPin in cluster 1 has all its loop structures positioned above A- and B-strand, BC-loop and DE-loop while all its domains lie above G-strand. Similar orientation observed in cluster 3, 4, 6 and 9. In contrast, DARPin in cluster 8, 7, 5 and 11 was in the opposite orientation, where loop structures found closer to G-strand and the domains positioned near A- and B-strand, BC-loop and DE-loop. DARPin cluster 2 docked to DIII with a very different orientation with both loops and domains lies across DIII.

Similarly, docked pose of the top model from each cluster of 3C-DIII system were shown in Figure 4.3 to describe the orientation of DARPin in complex. DARPin 3C in cluster 1 was also having all its loop structures orientated closer to A- and B-strand and BC-loop while all its domains ride atop of G-strand and F-strand. Similar orientation observed in cluster 8, but DARPin experienced a light shift as its loops were closer to

DE-loop while its domains was no longer making contact with F-strand. Cluster 2 has a comparable cluster size to cluster 1 but DARPins from this cluster docked to DIII with different orientation. DARPin 3C in cluster 3, 4, 5, 6 and 7 was however having loops facing G-strand while domains docked at the position closer to A- and B-strand, BC-loop and DE-loop.

Although HADDOCK score was used to rank the clusters however, it was not correlated with binding affinity between the two proteins. Furthermore, top ranked cluster based on HADDOCK is commonly reliable in generating representative structure of two proteins in complex. Taken together all the observations from the docking output, cluster size and orientation of the two interacting proteins, DARPin and DIII, top model from cluster 1 of both systems was selected for undergoing MD simulations.

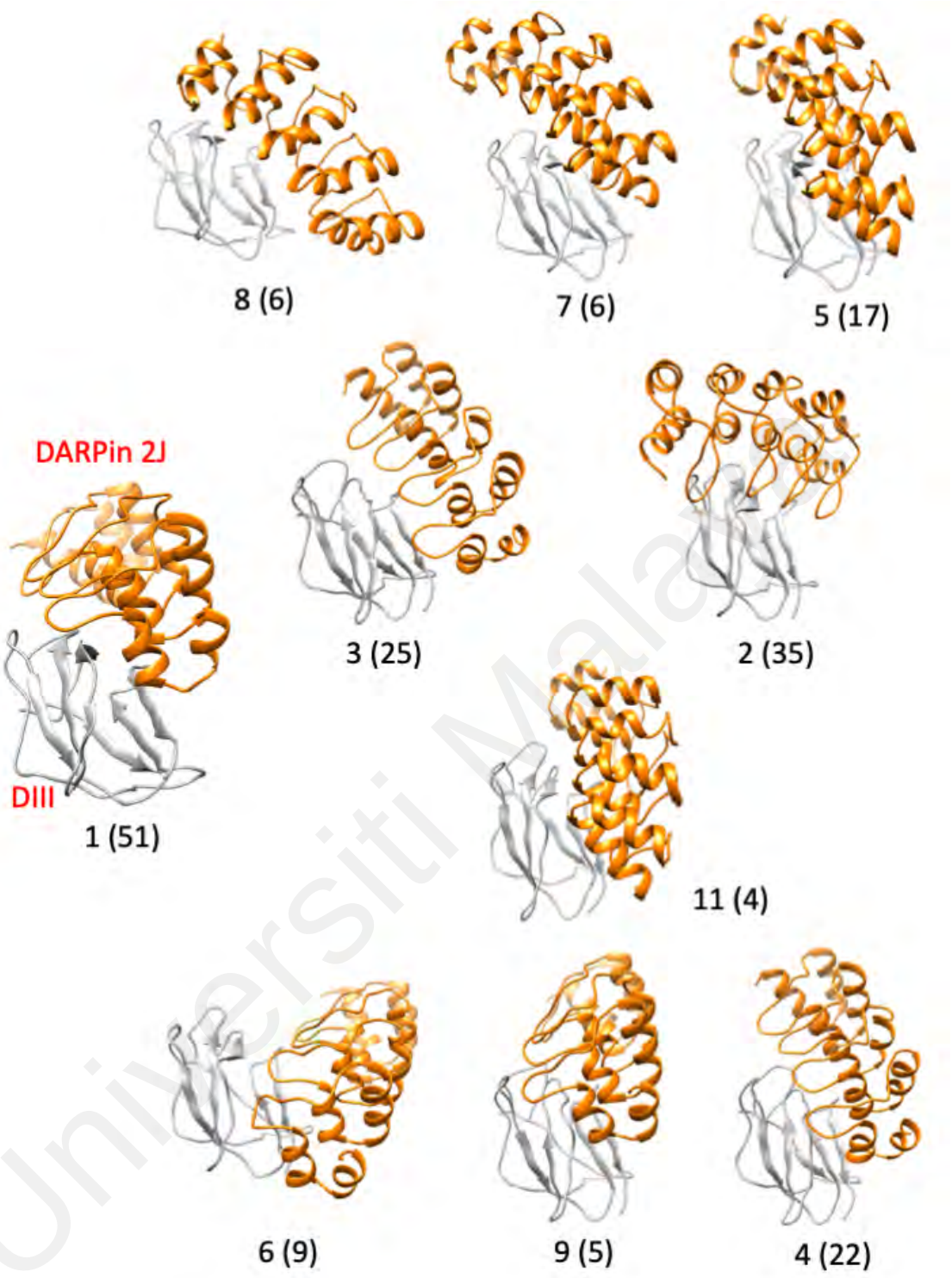


Figure 4.2 Top docked conformation of DARPin2J-DIII taken from the top ten clusters. Cluster number and number of strcutures in the cluster were shown.

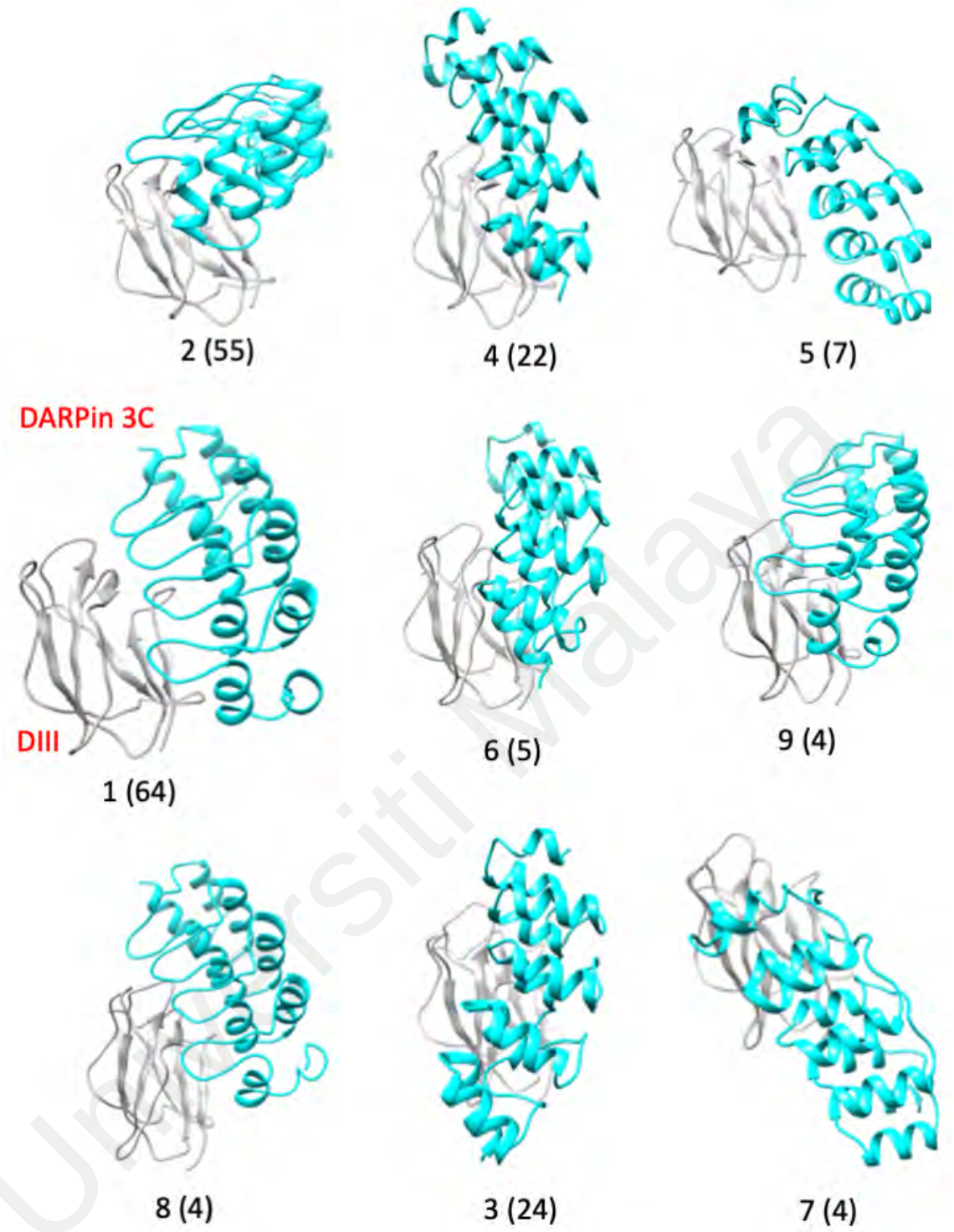


Figure 4.3 Top docked conformation of DARPin 3C-DIII of the nine clusters. Cluster number and number of structures of each cluster were shown.

4.2 Molecular dynamics analysis of trajectories of simulated protein complexes

Selected three-dimensional structure of 2J-DIII and 3C-DIII underwent MDs and trajectory of 100 ns long was collected from production run for detailed analysis. Similarly, 1A1D-2-DIII, 4E11-DIII, 4E5A-DIII and 513-DIII complexes were also simulated under same parameters and analysed for comparison purpose.

4.2.1 Stability of the scFv-DIII and DARPin-DIII complexes

Series of analyses began with computing the RMSD to assess the stability of the proteins in complex after having their potential energy checked. RMSD of all Cα-atoms in the protein complexes were calculated using all their backbone atoms with respect to the initial minimized structure. RMSD plot helps in determining the equilibrated state where snapshots are sampled for further analyses. Stability and conformational changes of the protein complexes would be reflected under RMSD plot, answering the questions like whether the proteins experienced large conformational changes and staying in the associated form during MDs.

Under different velocities, RMSD of the three independent MDs of DARPin-DIII complexes were not superposable throughout the simulation as found in Figure 4.4. For 2J-DIII, fluctuations of the three MDs occur within 1- 2 Å while 3C-DIII fluctuated in a greater range, 1- 4 Å. Conformational changes in 2J-DIII was much lesser as compared to 3C- DIII. After the first 10 ns, RMSD of 2J-DIII fluctuated steadily without much significant changes in its conformation. It indicated the structural differences of 2J-DIII before and after MDs were small. Among the three runs, 2J-2 appeared to be the most stable system owing to its steady fluctuations. For 3C-DIII systems, 3C-3 underwent the largest conformational changes as it began from 3 Å, increased steadily for the first 60 ns and decreased for following 20 ns before risen to 7 Å and ended around 6 Å. Fluctuating pattern of RMSD of 3C-3 suggested that simulated structures of 3C-DIII were different

from their initial structure in a higher order compared to another two runs, 3C-1 and 3C-2. Despite amplitude of fluctuation of the 3C1 and 3C-3 observed to be larger, however, these two systems came to a plateau region or known as equilibrated state during the last 20 ns of production run like 3C-1.

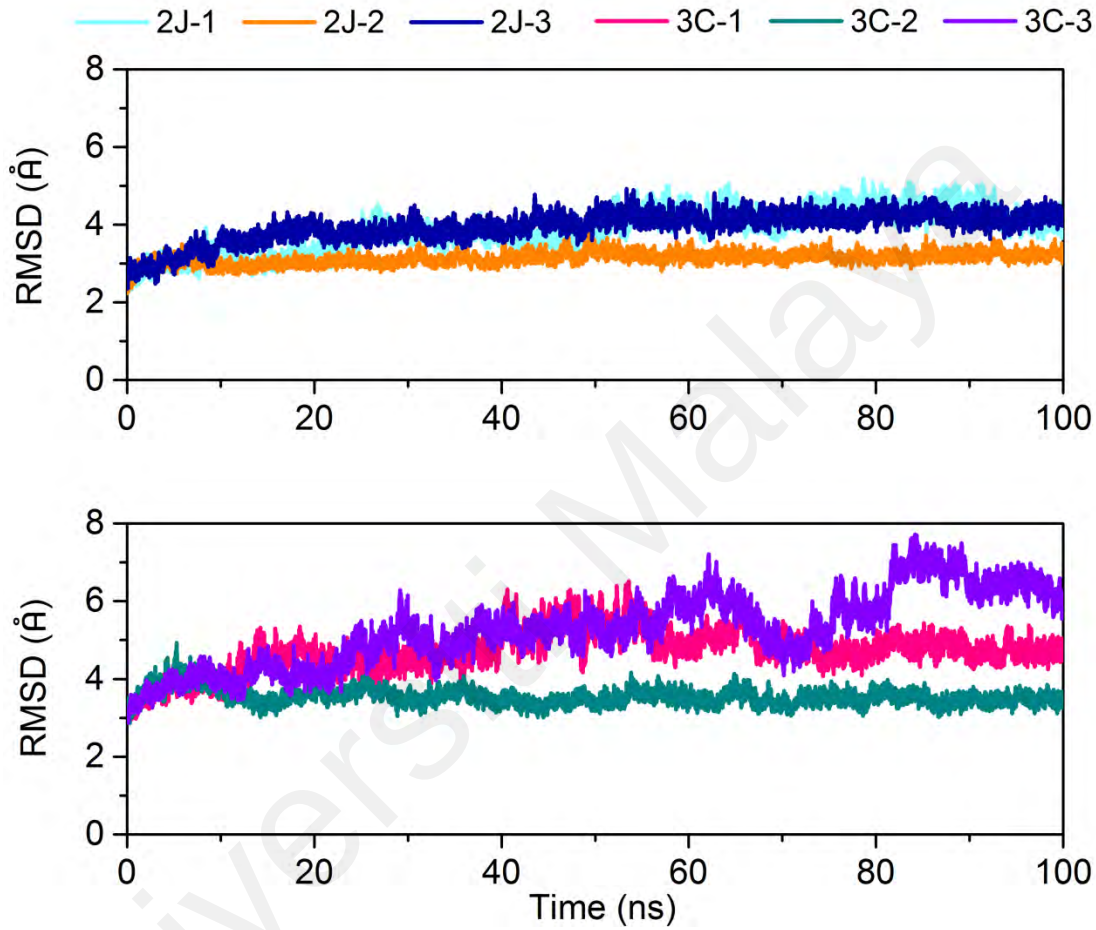


Figure 4.4 RMSDs of DARPin-DIII complexes throughout the multiple independent MDs.

RMSD of 1A1D-2-DIII was very different from the other three scFv-DIII complexes and it could be due to its larger size as presented in Figure 4.5. It has 530 amino acid residues while the other three scFv-DIII complexes having about 343 amino acid residues. Fluctuation within 3 to 6 Å occurred in 1A1D-2-DIII along the simulations and reaching a steadier state after 80 ns. Meanwhile, other scFv-DIII complexes fluctuate much lesser,

within 1 Å along the simulations. Despite 4E5A-DIII was modelled from X-ray crystal structure of 4E11-DIII (PDB ID 3UZV), stability of the complex was observed without significant conformational change. Both 4E11-DIII and 4E5A-DIII underwent a larger conformational change during the first 10 ns as RMSD increased from 2 to 3 Å and above but it has stabilized and continue to fluctuate steadily below 3 Å since then. Overlapping of the RMSD of 4E11-DIII and 4E5A-DIII occurred after the first 10 ns towards the end of the MDs. 513-DIII experienced the least fluctuation implying stability of the two proteins in complex. After 20 ns during production run, RMSD of the three scFv-DIII complexes were almost superposable.

Since a plateau region or equilibrated period was observable after 80 ns in all DARPin-DIII and scFv-DIII systems, hence snapshots were extracted from the last 20 ns (81 ns to 100 ns) for further analyses.

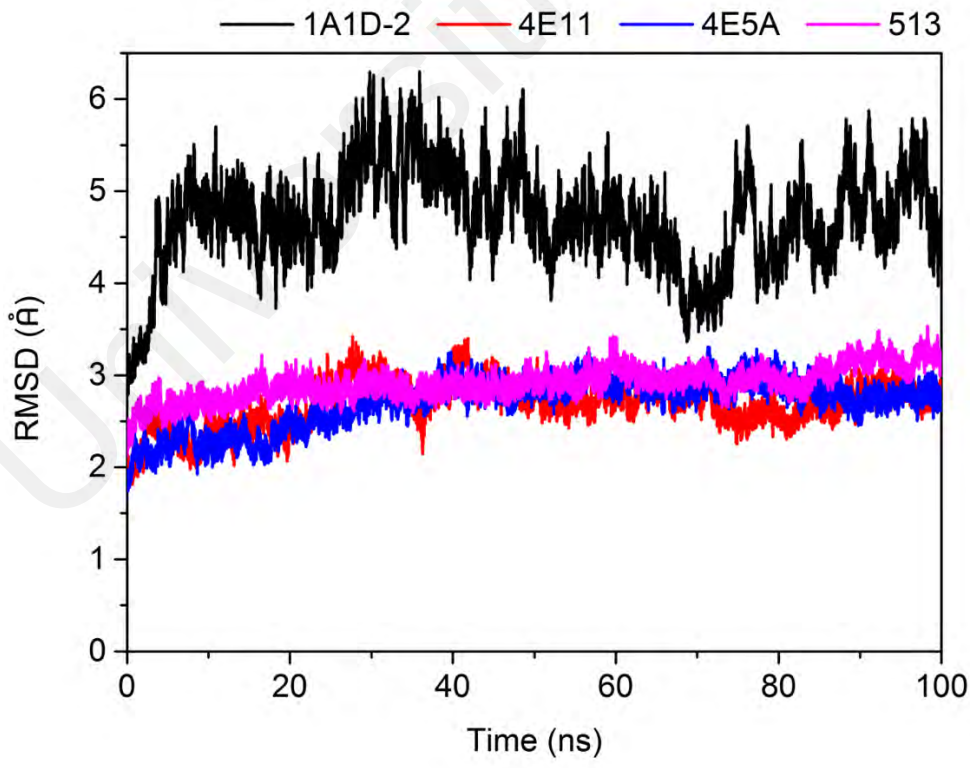


Figure 4.5 RMSD of scFv-DIII DENV-2 complexes.

4.2.2 Binding free energy of the protein complexes

Strength of the protein-protein interactions mediated was accounted using the stable simulated complex. Binding affinity of the two DARPin-DIII complexes was accounted using MM-GBSA method as tabulated in Table 4.3. MM-GBSA is computationally less expensive compared to MM-PBSA and also it has demonstrated its strength in predicting the binding affinities of protein-protein complexes with high accuracy (Chen et al., 2016). Also, contribution of entropy term was not accounted as calculating entropy is computational expensive and accuracy of estimating entropy under normal mode analysis remained low (Sun et al., 2018). Meanwhile, contribution from vdW and non-polar salvation energy was counterbalanced by sum of the electrostatic and polar solvation energy (EEL+EGB) that was very large and positive.

Table 4.3 Binding free energy of DARPin-DIII complexes and scFv-DIII complexes under MM-GBSA

Protein complex	Energy Term (kcal/mol)				
	VDW	EEL	EGB	ESURF	Δ_{GB}
2J-1	-99.15 ± 0.19	-85.40 ± 1.17	144.18 ± 1.11	-10.15 ± 0.013	-50.52 ± 0.21
2J-2	-83.29 ± 0.21	-78.66 ± 1.07	141.98 ± 0.97	-9.29 ± 0.017	-29.26 ± 0.22
2J-3	-102.60 ± 0.17	-61.14 ± 1.03	131.24 ± 0.99	-10.27 ± 0.015	-42.77 ± 0.21
3C-1	-73.57 ± 0.19	-4.35 ± 1.63	62.68 ± 1.51	-7.22 ± 0.019	-22.46 ± 0.23
3C-2	-68.47 ± 0.28	-21.64 ± 1.28	69.92 ± 1.17	-6.87 ± 0.028	-27.06 ± 0.25
3C-3	-57.98 ± 0.24	-36.33 ± 1.89	80.71 ± 1.77	-5.69 ± 0.030	-19.29 ± 0.33
1A1D-2	-97.34 ± 0.18	-372.62 ± 1.19	414.08 ± 0.92	-14.28 ± 0.01	-70.17 ± 0.45
4E11	-81.77 ± 0.17	-359.38 ± 1.20	411.75 ± 1.15	-13.62 ± 0.015	-43.02 ± 0.27
4E5A	-97.78 ± 0.25	-593.38 ± 1.54	618.31 ± 1.19	-16.17 ± 0.014	-89.02 ± 0.57
513	-106.63 ± 0.22	-543.06 ± 1.05	576.45 ± 0.92	-11.82 ± 0.012	-85.06 ± 0.24

Notes: VDW and EEL denote van der Waals and electrostatic contribution from molecular mechanics, respectively. EGB and ESURF denote polar and nonpolar contributions under GB model. Δ_{GB} is the binding free energy without taking entropy into calculation and is accompanied with the standard error mean of the calculation.

In the three independent runs of 2J-DIII complex, the lowest energy was -50.52 kcal/mol whereas the highest energy was -29.26 kcal/mol. For 3C-DIII systems, the lowest and the highest energy was -27.06 kcal/mol and -19.29 kcal/mol, respectively. For comparison purpose, systems with the binding free energy lies between the highest and the lowest energies were taken for further discussion. Hence, binding free energy of 2J and 3C in complex with DIII would be reported to be -42.77 kcal/mol and -22.46 kcal/mol, respectively.

Also, binding free energies of the Fab-DIII and scFv-DIII complexes were calculated using same parameters. Free energies obtained would serve as reference value to assess the potential of DARPin in substituting antibody to bind with DIII. All the four mAbs were previously reported to be effective in binding with DIII DENV-2 for neutralizing effect (Lok et al., 2008; Robinson et al., 2015; Tharakaraman et al., 2013) have binding free energy range from -89.02 kcal/mol to -43.02 kcal/mol (Table 4.3). According to experimental findings, 50% neutralization titers of mAb 1A1D-2 and mAb 4E11 for DENV-2 were found to be approximately 0.3 μg/ml (Crill & Roehrig, 2001; Lok et al., 2008; Sukupolvi-Petty et al., 2007) and 2.25 μg/ml (Cockburn et al., 2012; Thullier et al., 1999), respectively. Meanwhile, binding affinity of mAb 4E5A towards DIII DENV-2 had increased 15-fold as compared to its parent, mAb 4E11 (Tharakaraman et al., 2013). However, affinity for DENV-2 increased in a small extent when mAb 4E5A was redesigned to mAb 513 (Robinson et al., 2015).

Binding free energies of the Fab and scFvs correlated well with experimental finding and have successfully ranked the binding affinity of the mAbs towards DENV-2. Binding free energy of 4E5A and 513 were low and comparable, -89.02 kcal/mol and -85.06 kcal/mol, respectively. A small deviation between the binding free energy of 4E5A and 513 has further supported experimental finding where both 4E5A and 513 demonstrated

almost equivalent strength for neutralizing effect upon binding with DIII DENV-2. It was then followed by 1A1D-2 at -70.17 kcal/mol and lastly 4E11 at -43.02 kcal/mol. MAb with higher affinity in binding with DIII DENV-2 was observed to have a lower binding free energy. Noted that all the mAbs are tetravalent that bind with DIII of all dengue serotypes except 1A1D-2.

Although direct comparison of the binding free energies could be meaningless as the calculated energies were not the absolute binding free energy, however the range of the binding free energies obtained from scFv-DIII complexes could be used to rank or suggest if a protein has potential interacting with DIII DENV-2 like mAbs. 2J has a binding free energy (-42.77 kcal/mol) that was very close to the binding free energy range of the Fab/ scFvs (-43.02 to -89.02 kcal/mol) and hence its ability to manifest strong interaction with DIII was twice higher than 3C according to the binding free energy calculated from MM-GBSA method. Considering the comparisons made between complexes containing Fab-DIII, scFv-DIII and DARPin DIII, 2J was more likely and potent to interact with DIII under stable and strong protein-protein interaction like mAbs.

4.2.3 Residue interactions for binding affinity

Of the interfacial residues of two interacting proteins, not all of them are critical in mediating interactions for binding affinity. Residues that are responsible for key interactions within the binding interface of two proteins are recognized as hot spots and contribute for the binding free energy largely (Andrew et al., 1998; Clackson & Wells, 1995; Moreira et al., 2007a). Simulated structures obtained from MDs have been useful to provide an insight into key protein-protein interactions at atomic level besides predicting the key residues (Gonzalez-Ruiz & Gohlke, 2006; Huo et al., 2002; Lee et al., 2010; Moreira et al., 2007a; Rajamani et al., 2004; Yogurtcu et al., 2008). Therefore, binding free energy has been decomposed in a residue pairwise basis to identify the

interacting residues of the two interfaced proteins within the binding interface in 4 Å. In DARPin-DIII complexes, DIII residues that have been interacting with DARPin within 4 Å were determined and vice versa using CPPTRAJ analysis embedded in AMBER14 program. Similar calculation had been applied to Fab-DIII and scFv-DIII complexes as well, acted as a reference, since epitope of DIII for these complexes has been determined.

4.2.3.1 Important residues identification

Binding free energies of the protein-protein systems have been further decomposed into vdW interactions, total polar contribution obtained from summation of electrostatic energy and polar solvation energy (Ele+P) and polar solvation energy per residue pairwise basis. It involves interfacial residues found within 4 Å in the binding vicinity. The decomposed binding free energies allowed identification of important interfacial residues of DARPins and DIII that came to interact for binding affinity.

Altogether, 26 2J residues and 23 3C residues were categorized as interfacial residues while 24 residues and 27 residues of DIII were found in the binding interface upon complexation with 2J and 3C. Figure 4.6 and 4.7 described decomposition energy of the interfacial residues according to van der Waals, electrostatic + polar and non-polar energy terms. Very often, electrostatic and polar solvation terms offset each other and hence reflecting the screening effect of the solvent (Genheden & Ryde, 2015).

Residues with decomposition energy less than -5 kcal/mol were identified in the DARPin-DIII complexes, hence residue R13, V36, W38, Y46, D67, T68, H79, D100, N112, K137, N145 and N146 of 2J and H102, F135, T36, R113, N146, K112, L69, R79, H46, F35, W38, S68 of 3C were identified as important residues.

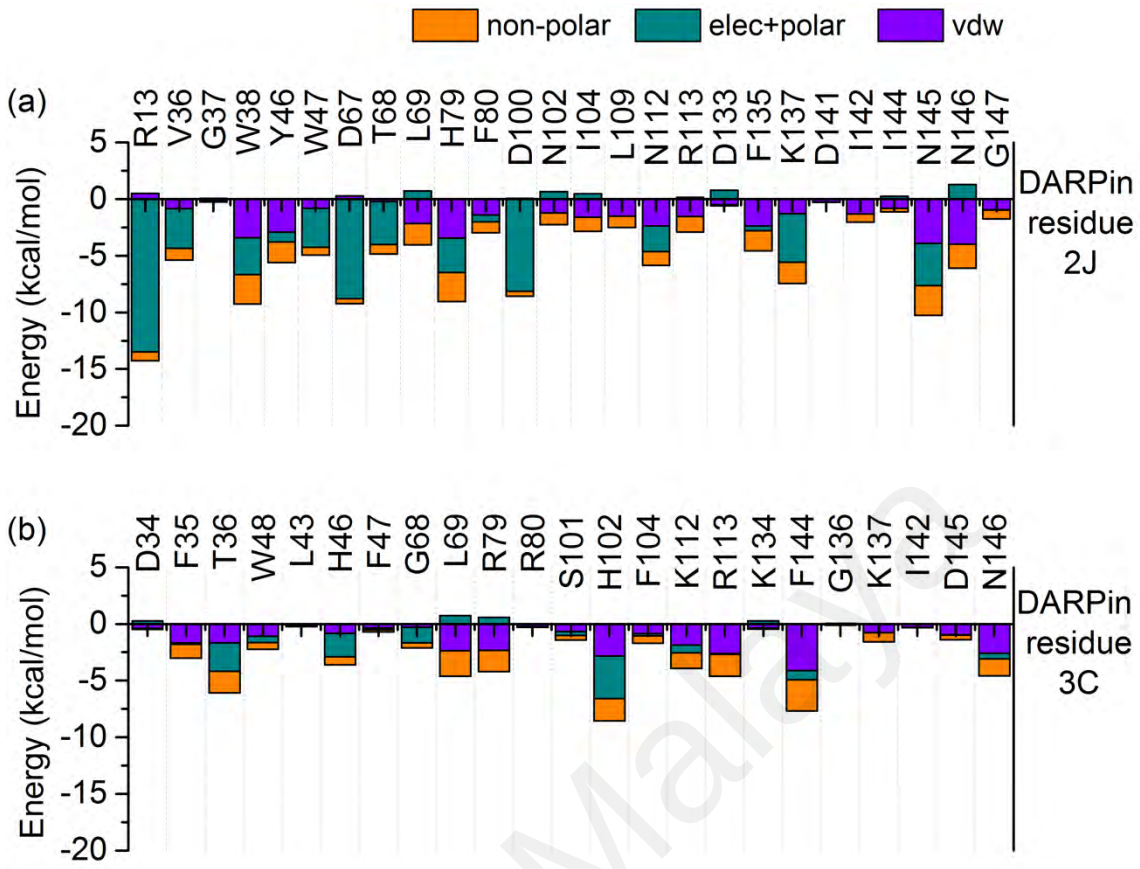


Figure 4.6 Decomposition Energy of DARPin residues of (a) 2J and (b) 3C upon interaction with DIII in 4 Å.

Sum of electrostatic energy and polar solvation energy or total polar contribution was the important component contributing for the binding free energy in 2J-DIII. However, contributions from non-polar and vdW terms were more significant for the binding free energy than the total polar contribution in 3C-DIII. On the other hand, DIII having more residues with low decomposition energy when it interacted with 2J. Under the same criteria, residue K305, K307, V308, V309, D329, K361 and K388 of DIII were identified as important residues when it associated with 2J while K305, K307, V308, and K388 were considered as important DIII residues in 3C-DIII system.

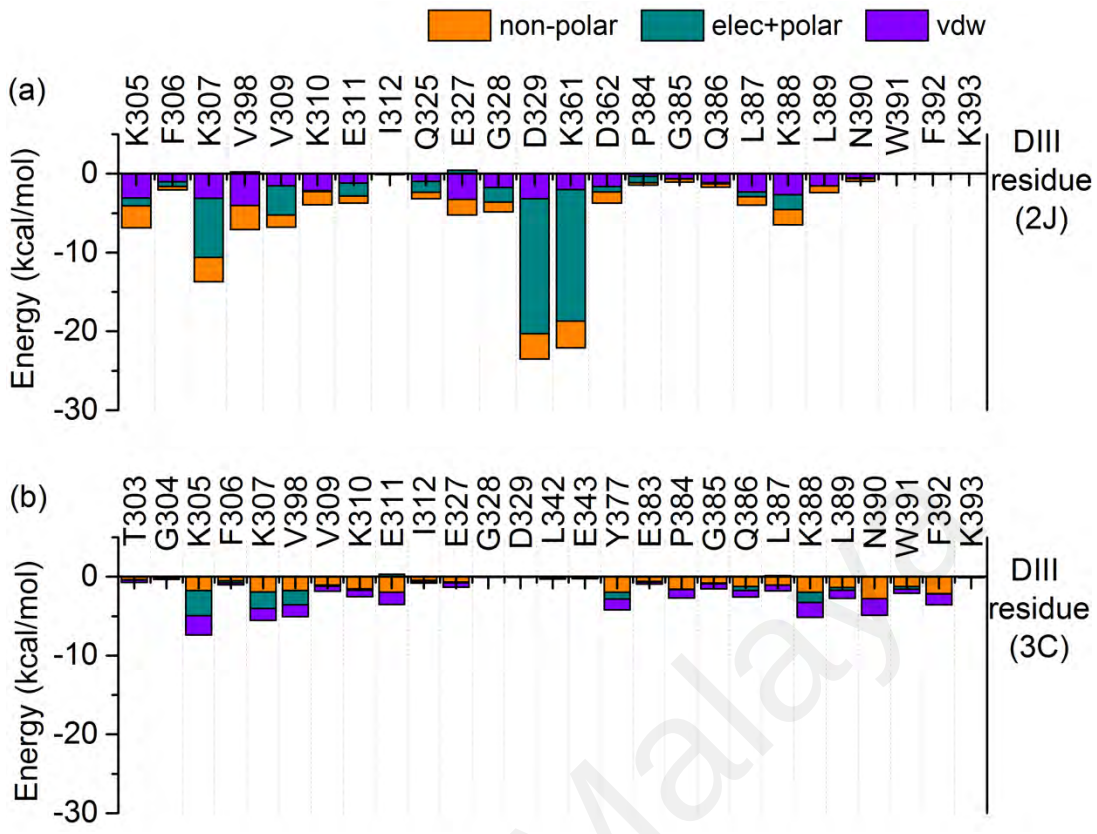


Figure 4.7 Decomposed binding free energy of DIII residues upon interacting with (a) DARPin 2J and (b) DARPin 3C within 4 Å in binding interface on a per residue pairwise basis.

Ample studies have also supported that 1A1D-2, 4E11, 4E5A and 513 interacted with DIII through A-strand, B-strand, DE loop and G-strand (Cockburn et al., 2012; Lok et al., 2008; Robinsonet al., 2015; Tharakaraman et al., 2013). Although epitope of DIII as mapped by mAbs has been determined experimentally, however from the simulated complex, more residues have been found interacting with Fab and scFvs within 4 Å as illustrated in Figure 4.8. It involved residues from A-strand (K305, F306, K307, V308, V309, K310, E311 and I312), B-strand (R323 and Q325), BC-loop (E327 and G328), DE-loop (K361, D362, S363, P364) and G-strand (G385, L387, K388, L389, N390, W391, F392 and K393).

In Fab-DIII and scFv-DIII complexes, distinctive low decomposition energy of residue K307 and K310 of DIII suggested the importance of these two residues for binding affinity as binding avidity reduced more than 140-fold when side chain at residue 307 and 310 was truncated by altering lysine to alanine, glycine or glutamic acid (Gromowski et al., 2010). Important residues of DIII would be discussed using 513-DIII system since mAb 513 has the greatest strength in combating all dengue serotypes. With decomposition energy less than -5 kcal/mol, DIII residues, K305, F306, V308, V309, E311, R323, E327 and D362 in 513-DIII complex were classified as important residues together with K307 and K310.

The decomposition energy profile of interfacial residues of DIII when it complexed with scFv and DARPins was compared in Figure 4.8. DIII in 2J-DIII has an energy profile that was closer to that of DIII in Fab- and scFv-DIII systems. Decomposition energy profile of DIII revealed several remarkable findings, (1) DIII was not interacting with the two DARPins through R323 from B-strand like how it interacted with the scFvs, (2) residue Q325 DIII, part of the epitope recognized by scFvs made contact with 2J but not 3C, (3) residues G328 and D329 of BC-loop DIII interacted with 2J but not scFvs and 3C, (4) negatively charged residues K361 and D362 of DIII interacted with 2J like the scFvs with a noticeable low decomposition energy, < -20 kcal/mol on K361.

Summation of the decomposition energy of the interfacial residues of respective strand or loop showed that A-strand has the lowest total energy (-101.82 kcal/mol), followed by B-strand (-27.17 kcal/mol), DE-loop (-20.15 kcal/mol), G-strand (-18.64 kcal/mol) and BC-loop (-8.49 kcal/mol) in 513-DIII system. For 2J-DIII, lowest total energy -44 kcal/mol found at A-strand, followed by BC-loop (-33.1 kcal/mol), DE-loop (-25.8 kcal/mol), G-strand (-15.53 kcal/mol) and B-strand (-3.14 kcal/mol). A-strand remained

as the lowest total energy region in 3C-DIII system, -27.34 kcal/mol, followed by G-strand (-22.67 kcal/mol) and BC-loop (-1.46 kcal/mol).

Both N- and C-capping units of 2J participated in DARPin-DIII interactions but not 3C. There was no N-cap residue of 3C found in the 4 Å binding interface and total decomposition energy contribution of C-cap (-15.78 kcal/mol) was almost two times higher to that of 2J, -31.82 kcal/mol. In 2J-DIII, lowest total decomposition energy contributed by repeat unit 2 (-29.13 kcal/mol), followed by repeat unit 1 (-25.37 kcal/mol) and repeat unit 3 (-23.55 kcal/mol). However, in 3C-DIII, repeat unit 3 has the lowest total decomposition energy -20.22 kcal/mol, followed by repeat unit 1 and 3, at -16.11 kcal/mol and -9.87 kcal/mol, respectively. Since important DIII residues have been identified, hence DARPin residues that interacted with them in pair were studied and analysed for their interaction type.

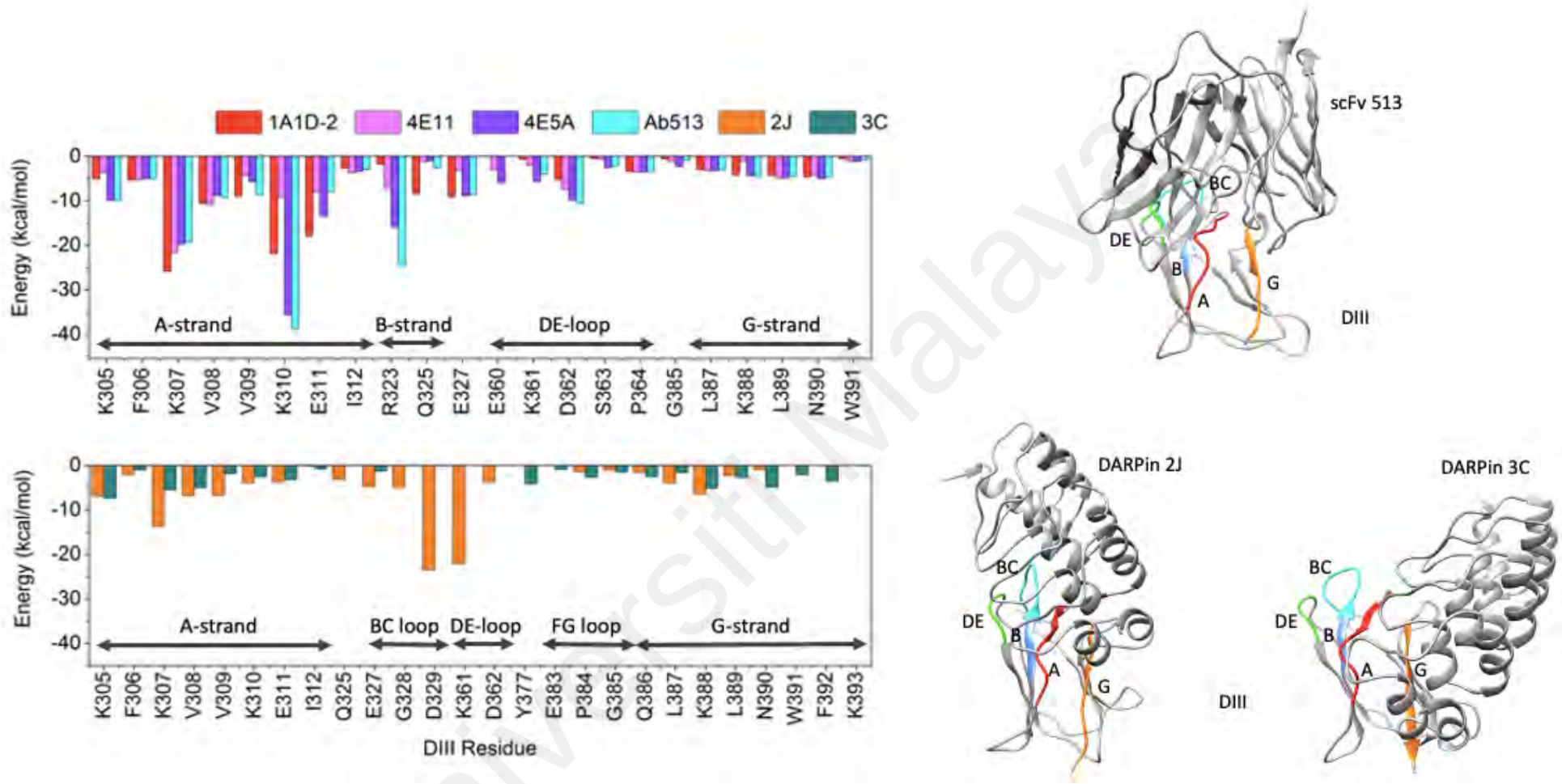


Figure 4.8 DIII residues participated in protein-protein interactions upon binding with scFvs and DARPins.

4.2.3.2 Residue pair interactions for binding affinity

Being one of the important non-covalent interactions in molecular systems, hydrogen bonds are electrostatic and formed when hydrogen atom is shared between two heavy electronegative atoms whereby one is covalently bonded to hydrogen bond donor and one bonded to hydrogen bond acceptor. Energy comes from a single hydrogen bond is too small to be considered but the contribution to protein energetics would be significant when large number of hydrogen bonds are involved and hence, hydrogen bonds are indispensable for protein binding (Erijman et al., 2014). Although the strength of a hydrogen bond is defined by distance and angle but under Amber forcefield, it is treated as a non-bonded terms and angle has no role in determining the strength. In order to determine if hydrogen bonds are well preserved under employed forcefield, their angle must be greater than 120.

Structural analysis on 513-DIII complex confirmed a total of 20 hydrogen bonds across the interface of the (Robinson et al., 2015). However, number of hydrogen bonds obtained from MD trajectories differ from the reported literature with additional four hydrogen bonds. Considering hydrogen bonds with occupancy greater than 21%, there were fifteen hydrogen bonds formed in 2J-DIII while 3C-DIII was observed to have two hydrogen bonds as tabulated in Table 4.4. The greater the occupancy percentage of a hydrogen bond, the greater the stability of that hydrogen bond (Garcia et al., 2012). Besides hydrogen bonds, hydrophobic and vdW have also made contribution for binding interactions. Detailed interaction of residue pairs for binding affinity in the protein complexes were analyzed and tabulated.

Table 4.4 Hydrogen bond formed in 2J-DIII, 3C-DIII and 513-DIII

Complex	Acceptor	Donor	%	Distance (Å)	Angle	DIII	Repeat module
2J-DIII	D329@OD2	R13@HH21	90.60	2.79	163.19	BC	N
	D329@OD1	R13@HE	89.00	2.82	163.95	BC	N
	N145@O	K388@H	71.20	2.85	159.50	G	C
	D329@OD1	W47@HE1	68.80	2.85	157.21	BC	1
	D67@OD2	K361@HZ3	34.90	2.75	154.32	DE	2
	D362@OD2	T68@HG1	33.50	2.73	158.21	DE	2
	K305@O	H79@HE2	33.10	2.86	154.10	A	2
	Q325@OE1	N102@HD22	33.00	2.83	157.46	B	3
	V309@O	K137@HZ2	31.30	2.80	156.98	A	C
	D67@OD2	K361@HZ1	24.50	2.74	153.28	DE	2
	V36@O	K361@HZ2	24.10	2.83	153.71	DE	1
	D100@OD1	K307@HZ1	22.40	2.76	154.64	A	3
	D100@OD1	K307@HZ2	22.20	2.77	154.82	A	3
	D67@OD2	K361@HZ2	22.00	2.78	152.36	DE	2
	E327@OE2	S71@HG	21.60	2.75	165.82	BC	2
3C-DIII	D375@OD1	R113@HH21	63.80	2.82	160.43	F	3
	D375@OD2	R113@HH11	58.80	2.80	160.85	F	3
513-DIII	F306@O	R184@HH22	94.90	2.79	160.66	A	
	E53@OE2	R323@HH22	80.10	2.81	162.82	B	
	E53@OE1	R323@HH12	78.00	2.78	154.97	B	
	K388@O	N164@HD22	71.80	2.85	159.49	G	
	E100@OE1	V308@H	67.60	2.87	161.48	A	
	E327@OE2	Q189@HE21	63.10	2.85	162.33	BC	
	E311@OE2	Y162@HH	60.60	2.64	161.39	A	
	D362@OD2	Y104@HH	43.60	2.68	163.69	DC	
	Y162@O	N390@H	43.30	2.89	163.37	G	
	D30@OD1	R323@HH21	42.70	2.80	154.46	B	
	D362@OD1	Y104@HH	42.40	2.68	163.74	DC	
	Q189@OE1	K307@HZ3	33.50	2.78	149.81	A	
	E53@OE1	K310@HZ1	30.90	2.78	157.18	A	
	E53@OE1	K310@HZ2	30.70	2.79	157.70	A	
	D329@OD1	W190@HE1	26.80	2.84	159.33	BC	
	E53@OE1	K310@HZ3	26.50	2.77	157.34	A	
	D362@OD2	Q3@HE21	26.50	2.83	158.19	DC	
	D51@OD1	K310@HZ3	26.00	2.78	154.91	A	
	E311@OE1	Y162@HH	25.80	2.63	161.62	A	
	D362@OD2	Q3@HE21	25.30	2.84	158.14	DC	
	D51@OD1	K310@HZ1	23.60	2.78	157.54	A	
	K29@O	K310@HZ2	21.50	2.82	155.97	A	
	E187@OE2	K305@HZ1	21.30	2.75	158.44	A	
	K29@O	K310@HZ2	21.00	2.81	156.36	A	

Table 4.5 Residue pair interactions in 513-DIII complex

DIII strand/ loop	Residue-pair		Energy (kcal/mol)				Interaction
	DIII	mAb	vdW	Ele+P	NP	Total	
A	K305	Y183	-0.63	-0.76	-0.62	-2.01	H-bond
		E187	0.067	-6.11	-1.00	-7.06	Salt bridge
	F306	R184	0.48	-4.61	-0.31	-4.44	H-bond
	K307	E100	-1.50	-5.75	-1.19	-8.44	H-bond
		Y183	-1.57	-2.36	-1.38	-5.31	Pi-cation
		Q189	0.35	-4.60	-0.44	-4.69	H-bond
	V308	W99	-1.21	-0.37	-0.78	-2.37	Pi-Alkyl
		E100	-0.89	-2.97	-0.97	-4.83	H-bond
	V309	D30	-0.44	-0.53	-0.21	-1.19	Alkyl
		V31	-1.14	-0.68	-0.95	-2.78	Alkyl
		Y32	-0.84	-0.66	-0.32	-1.81	H-bond
		W99	-1.22	-0.02	-0.66	-1.90	Pi-Alkyl
	K310	K29	0.14	-4.25	-0.32	-4.43	H-bond
		D30	-0.81	-0.49	-0.44	-1.74	Alkyl
		V31	-0.94	0.09	-0.61	-1.45	Alkyl
		Y32	-1.94	-1.87	-1.39	-5.20	Pi-Alkyl
		D51	0.17	-14.39	-0.79	-15.01	Salt-bridge
		E53	0.36	-9.08	-0.65	-9.37	Salt-bridge
		W99	-0.93	-0.80	-0.56	-2.29	Pi-Alkyl
	E311	Y32	-0.63	-0.57	-0.70	-1.91	H-bond
		W99	-1.08	0.20	-0.70	-1.58	Anion-Pi
		Y162	0.40	-4.24	-0.54	-4.39	H-bond
	I312	Y162	-1.12	-0.51	-1.04	-2.67	Pi-Alkyl
B	R323	D30	-0.85	-7.56	-1.22	-9.63	Salt-bridge
		E53	0.93	-13.85	-0.73	-13.65	Salt-bridge
	Q325	Y104	-0.52	-0.76	-0.61	-1.89	H-bond
BC	E327	Q189	-0.47	-3.71	-0.54	-4.72	H-bond
		W190	-1.21	-0.08	-0.76	-2.06	Anion-Pi
DC	D362	Q3	-0.25	-2.33	-0.44	-3.02	H-bond
		F26	-1.10	-0.20	-0.93	-2.23	Pi-Anion
		R97	-0.42	-0.37	-0.30	-1.09	Electrostatic
		Y104	-0.04	-3.29	-0.75	-4.07	H-bond
G	K388	N164	-0.24	-2.47	-0.56	-3.27	H-bond
	L389	Y162	-1.47	-0.7	-1.13	-3.30	Pi-Alkyl
	N390	Y162	-1.31	-1.15	-0.94	-3.40	H-bond
		G163	-0.5	-0.2	-0.43	-1.14	H-bond

In 513-DIII, scFv 513 interacted with A-strand, B-strand, G-strand and BC-loop of DIII through strong and stable hydrogen bonds with occupancy percentage greater than 60%. Strongest hydrogen bond was found between F306 and R184 (94.9%), followed by two hydrogen bonds between R323 and E53 (80.1% and 78%) and K388 and N164 (71.8%). All the seven hydrogen bonds found between K310 and 513 residues was low in occupancy percentage, in the range of 20 - 30%, however this extensive hydrogen bond network could explain the very low decomposition energy of K310 as observed in Figure 4.8. Detailed interaction of residue pairs for binding affinity in 513-DIII complex found in Table 4.5 revealed that residue pairs have hydrogen bond or salt bridge formed between them were low in decomposition energy, K310-D51 (-15.01 kcal/mol), R323-E53 (-13.65 kcal/mol), K305-E187 (-7.06 kcal/mol), K307-E100 (-8.44 kcal/mol), V308-E100 (-4.83 kcal/mol), E327-Q189 (-4.72 kcal/mol), F306-R184 (-4.44 kcal/mol), E311-Y162 (-4.39 kcal/mol) and D362-Y104 (-4.07 kcal/mol). Other than the favorable hydrogen bonds and salt bridges, hydrophobic interaction (alkyl) observed in V309-V31 (-2.78 kcal/mol).

Of the fifteen hydrogen bonds formed in 2J-DIII, four involved A-strand (K305, K307, V309), four involved BC-loop (E327 and D329), five involved DE-loop (K361 and D362) while one each for B-strand (Q325) and G-strand (K388). The two hydrogen bonds between D329 and R13 have the highest occupancy percentage, 90.6% and 89%, respectively. It was followed by hydrogen bond formed between K388 and N145 (71.2%) and D329 and W47 (68.8%). Residues of both loop and α-helices of 2J involved in protein-protein interactions with DIII and detailed interactions were tabulated in Table 4.6. Like 513-DIII complex, residues pairs in 2J-DIII with low decomposition energy involved hydrogen bonds or salt bridge. K305-HIE 79 (-3.5 kcal/mol), K307-D100 (-9.32 kcal/mol), V308-N112 (-2.23 kcal/mol), V309-K137 (-4.41 kcal/mol), D329-R13(-13.82 kcal/mol), K361-D67 (-9.56 kcal/mol) and K388-N145 (-3.32 kcal/mol) in 2J-DIII

complex were accounted as crucial residue pairs owing to their lowest decomposition energy.

Table 4.6 Important residue pair for 2J-DIII interactions

Residue pair		Energy			Total	Interaction	Module	Strand
DIII	DARPin	vdW	Ele + P	NP				
K305	HIE 79	-0.92	-1.65	-0.97	-3.54	H-bond	2	A
	F80	-1.03	-0.54	-0.95	-2.52	Pi-Alkyl	2	
K307	HIE 79	-0.53	-0.31	-0.31	-1.15	Pi-hydrophobic	2	
	D100	0.41	-9.35	-0.38	-9.32	Salt-bridge	3	
	I104	-0.80	0.56	-0.84	-1.07	Alkyl	3	
	N112	-0.72	-1.07	-0.51	-2.30	H-bond	3	
V308	N112	-0.64	-1.12	-0.47	-2.23	H-bond	3	
	I142	-0.91	0.02	-0.70	-1.59	Alkyl	C	
	N145	-0.83	0.21	-0.72	-1.34	vdW	C	
V309	F135	-0.87	-0.13	-0.71	-1.71	Pi-alkyl	C	
	K137	0.37	-4.42	-0.35	-4.41	H-bond	C	
D329	R13	0.55	-13.68	-0.69	-13.82	Salt bridge	N	BC
	W38	-1.12	-0.08	-0.66	-1.86	vdW	1	
	Y46	-1.62	-0.05	-1.08	-2.74	Pi-donor H-bond	1	
	W47	-0.72	-3.27	-0.70	-4.70	H-bond	1	
K361	V36	-0.35	-3.38	-0.85	-4.58	H-bond	1	DE
	W38	-1.08	-2.19	-1.08	-4.35	Pi-Cation	1	
	D67	0.54	-9.63	-0.47	-9.56	Salt bridge	2	
	T68	-0.14	-2.36	-0.38	-2.88	vdW	2	
K388	N145	-0.46	-2.43	-0.43	-3.32	H-bond	C	G
	N146	-1.07	0.39	-0.56	-1.24	vdW	C	
	G147	-0.71	0.02	-0.79	-1.48	vdW	C	

Remarkable low decomposition energy of D329 and K361 as shown in Table 4.6 could be explained by the formation of multiple hydrogen bonds involving D329-R13, D329-W47, K361-D67 and K361-V36 residue pairs. 2J manifested strong and more stable hydrogen bonds with BC-loop and G-strand but a weaker hydrogen bond with A-strand, B strand and DE-loop due to their lower occupancy percentage. Besides, hydrogen bonds were found between 2J and the negatively charged region of DIII, K361 and D362. Unlike 513-DIII, DARPin 2J was interacting with K310 through Pi-Alkyl interaction (-1.56 kcal/mol) instead of hydrogen bond.

There were only two hydrogen bonds observed in 3C-DIII, D375-R113 (63.8%) and D375-R113 (58.8%). Although the two hydrogen bonds were of stable and strong, but it involved F-strand, which was not one of the important epitope regions. A-strand DIII was observed to have interactions with the whole DARPin 3C except N-cap like 2J-DIII. In contrast, (1) N-cap did not interact with the important residues; (2) repeat unit 1 and repeat unit 2 was interacting with K388 from G-strand and (3) salt bridge was not found in any of the crucial residue pair. K305-T34 (-3.68 kcal/mol; K307-HID 100 (-3.39 kcal/mol); V308-HID 100 (-3.10 kcal/mol) and K388-HID 44 (-2.93 kcal/mol) were identified as the important residue pairs with their lowest decomposition energy as depicted in Table 4.7. Apart from K310, hydrogen bond was not formed between K307 and 3C as well. Side chain interaction of the important residue-pairs in 2J-DIII and 3C-DIII were described in Figure 4.9.

Most of the important residue-pairs identified in 2J-DIII and 513-DIII involved hydrogen bond and accompanied with low decomposition energy. All the five important regions of DIII made contact with scFv 513 through hydrogen bonds indicated stability of scFv-DIII interactions and crucial role of hydrogen bonds for binding affinity. Similarly, hydrogen bond network was observed between 2J and the five important

regions of DIII. Taken binding free energy, decomposition energy, important residue pairs and hydrogen bond network together, potential of 2J in mimicking mAb was predicted to be higher to that of 3C. Hence, 2J would be further theoretically designed and engineered for higher affinity in binding with DIII.

Table 4.7 Important residue pair for 3C-DIII interactions

Residue pair		Energy			Total	Interaction	Repeat module	Strand
DIII	DARPin	vdW	Ele + P	NP				
K305	T34	-0.30	-2.57	-0.81	-3.68	H-bond	1	A
	S66	-0.17	-1.36	-0.47	-2.00	H-bond	2	
	L67	-0.84	0.73	-0.99	-1.09	Alkyl	2	
K307	S99	-0.42	-0.31	-0.41	-1.15	vdW	3	
	HID 100	-0.81	-2.01	-0.56	-3.39	Pi-Alkyl	3	
V308	HID 100	-0.79	-1.47	-0.84	-3.10	H-bond	3	
	F133	-0.58	-0.34	-0.49	-1.40	vdW	C	
K388	HID 44	-0.35	-2.02	-0.56	-2.93	H-bond	1	G
	R77	-1.16	0.79	-1.24	-1.61	vdW	2	

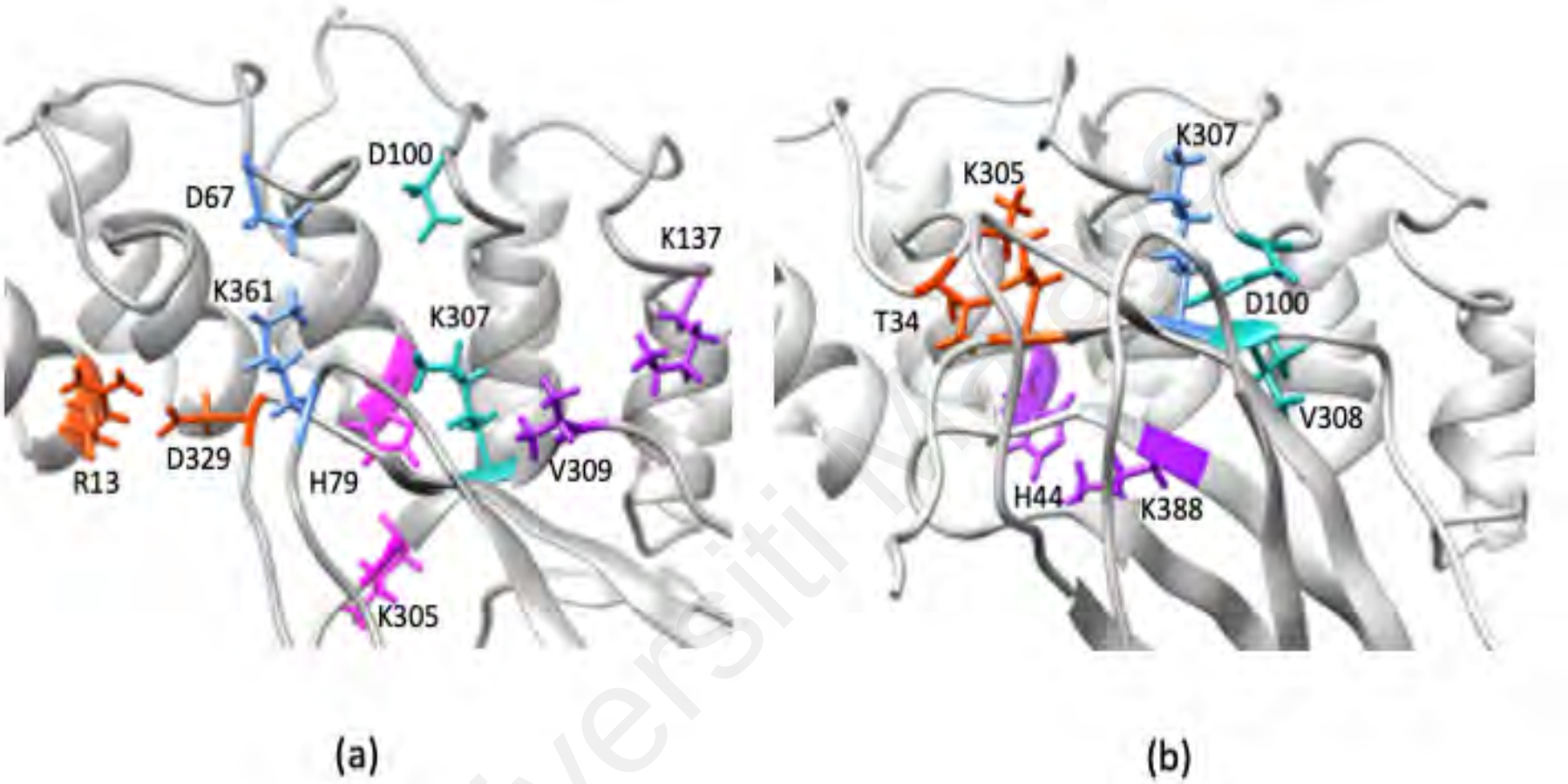


Figure 4.9 Interactions of crucial residue-pairs for binding affinity of (a) 2J-DIII and (b) 3C-DIII. Orange residue pair has the lowest decomposition energy, followed by blue, green, purple and magenta.

4.3 Further design of DARPin under site directed computational mutagenesis

Amino acids of the randomized residues located in all repeat units of DARPin could be replaced with other natural amino acids except glycine, proline and cysteine. Hence, 2J was subjected to mutation study and have its DARPin-DIII binding affinity predicted upon changes. Each repeat unit of DARPin consist of six randomized residues and three repeat units of 2J would have eighteen randomized residues in total. Altogether, 288 single point mutations were performed. Each time, only one residue was targeted and has its wild type amino acids replaced with other sixteen amino acids.

4.3.1 Effects of single point mutation and predicted affinity

Results obtained from single point mutation study under mSCM-PPI2 and BeAtMuSiC server were shown in Table 4.8. Both algorithms suggested amino acid replacement at 7 of the 18 randomised residues able to improve the binding affinity between 2J and DIII. A more positive predicted $\Delta\Delta G^{Affinity}$ obtained from mSCM-PPI2 implying a higher affinity upon mutation while a more negative $\Delta\Delta G_{Bind}$ obtained from BeAtMuSiC implying affinity improved between the proteins. V36, T68, L69, N102, I104, N112 and R113 were the seven residues predicted under mSCM-PPI2 while Y66, L69, S71, D101, N102, N112 and R113 were the seven residues predicted by BeAtMuSiC. Both algorithms have suggested L69, N102, N112 and R113 to be the possible point mutation to increase the binding affinity of 2J and DIII. Besides, both protocols proposed more possible residues located from third repeat unit and that could be explained by its highest total decomposition energy (-23.55 kcal/mol) compared to another two repeat units (-25.37 kcal/mol and -29.13 kcal/mol). Residues recommended by mSCM-PPI2 algorithm were those interfacial residues found within 4 Å binding interface while BeAtMuSiC had suggested three residues, Y66, S71 and D101 which were not part of interfacial residues.

Table 4.4 Effect of single point mutation in binding affinity calculated by mSCM-PPI2 and BeAtMuSiC server

mSCM-PPI2				mSCM-PPI2				BeAtMuSiC				BeAtMuSiC			
Res. #	Amino acid		$\Delta\Delta G^{Affinity}$ (kcal/mol)	Res. #	Amino acid		$\Delta\Delta G^{Affinity}$ (kcal/mol)	Res. #	Amino acid		$\Delta\Delta G_{Bind}$ (kcal/mol)	Res. #	Amino acid		$\Delta\Delta G_{Bind}$ (kcal/mol)
	WT	Mut			WT	Mut			WT	Mut			WT	Mut	
36	V	Y	0.581			D	0.293	66	Y	K	-0.16	113	R	W	-0.28
		W	0.391			R	0.192			R	-0.13				
		Q	0.367			E	0.186			Q	-0.08				
		R	0.366			H	0.166	69	L	R	-0.25				
		K	0.346	104	I	W	0.697			K	-0.13				
68	T	W	0.401			Y	0.536	71	S	L	-0.08				
		E	0.325			F	0.472	101	D	S	-0.07				
		K	0.27			R	0.160			T	-0.04				
		D	0.262			K	0.152	102	N	L	-0.08				
		H	0.187	112	N	Y	0.186	112	N	W	-0.98				
69	L	W	0.751			M	0.181			F	-0.68				
		F	0.538			W	0.177			Y	-0.59				
		Y	0.478	113	R	H	0.361			I	-0.42				
		D	0.300			W	0.157			L	-0.38				
		H	0.256			Y	0.001			V	-0.13				
102	N	Y	0389							M	-0.08				

$\Delta\Delta G^{Affinity}$ and $\Delta\Delta G_{Bind}$ denote predicted binding affinity, a more positive $\Delta\Delta G^{Affinity}$ and a more negative $\Delta\Delta G_{Bind} < 0$ indicated affinity increase

Under mSCM-PPI2 protocol, mutation L69W reported the highest $\Delta\Delta G^{Affinity}$ at 0.751 kcal/mol, followed by I104W at 0.697 kcal/mol and V36Y at 0.581 kcal/mol. To understand how amino acids replacement at these residues altered the interactions between 2J and DIII for increased affinity, residue with the highest $\Delta\Delta G^{Affinity}$ were examined using the interactive viewer feature offered by mCSM-PPI2 server. It snapshots the interatomic interactions between the wild type residue and the surrounding residues. In Figure 4.10, the interactions mediated before and after mutation were compared for the number and type of the interactions formed or disappeared.

Residue L69 uses its side chain interacting with OD2 atom and CB atom residue D221 DIII via polar interaction and hydrophobic, respectively. It also interacted with its own DARPin residue through CB atom D101 for hydrophobic interaction. Apart from side chain-side chain interaction, backbone of L69 formed hydrogen bond with OD2 residue D101 as well. Replacing Leucine with Tryptophan have increased the strength of interaction between residue 69 and residue D221 DIII as additional hydrogen bond and vdW interaction formed between the two residues. The ring structure of tryptophan was observed to have hydrophobic interaction with DARPin residue D100 and D101 while its backbone interacted with DARPin residue D67 through polar interaction.

Meanwhile, having phenylalanine replaced leucine was predicted to be lower in affinity compared to tryptophan. The ring structure of the phenylalanine side chain of making contact with OD2 atom and CB atom with 2 polar and 3 hydrophobic interactions, respectively. However, interaction between phenylalanine and nearby residues was mainly polar with absence of hydrogen bond.

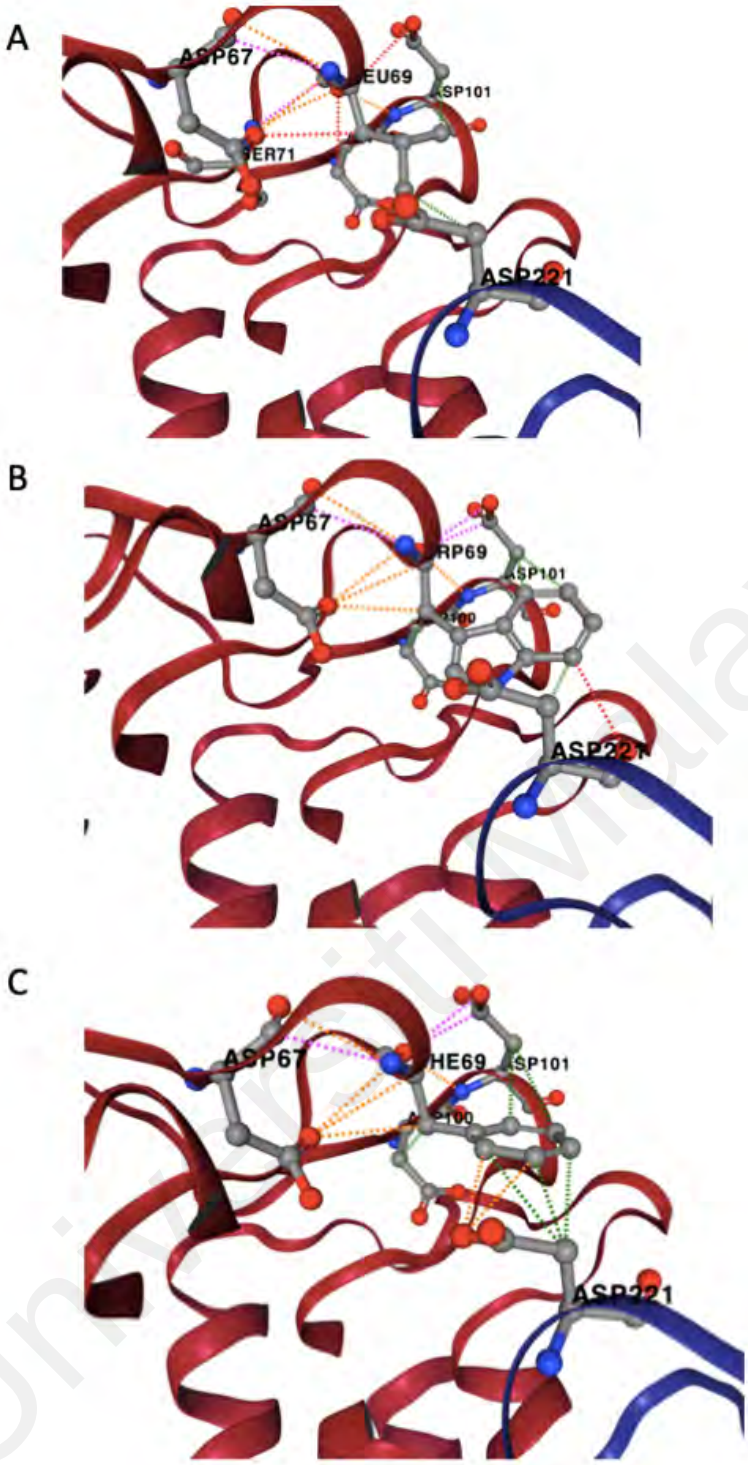


Figure 4.10 Interatomic interaction before and after single point mutation at residue L69. Green dotted line denotes hydrophobic; light blue denotes vdW; red denotes hydrogen; orange denotes polar and pink denotes clash interaction.

4.3.2 Free energy computed from MD simulations

Despite interactions formed in 2J-DIII after point mutation supported and explained the predicted affinity, however it was only considering a single structure. Hence, 2J in 2J-DIII underwent computational site-directed mutagenesis according to the top two choices of amino acids suggested and subjected to MD simulation using same parameters for binding free energy calculation. Altogether, 23 systems of single point mutations were simulated, their binding free energies were presented in Table 4.9.

Effect of mutating just one amino acid of DARPin could be huge and significant. DARPin Ank1D4 that bound to capsid protein HIV-1 has its binding affinity increased upon single point mutation proven experimentally and computationally (Karim et al., 2019). In another case, binding free energy of DARPin bound to ERK2 kinase protein was lowered for a stronger binding interaction by mutating a randomized residue under computational approach (Gautam, 2021).

Most of the 2J derivatives under single point mutation were calculated to have a higher binding free energy and hence decreasing the affinity to bind with DIII. Nevertheless, binding free energy of five mutants observed to be lower. N102Y has the lowest binding free energy (-57.17 kcal/mol) followed by Y66K (-56.91 kcal/mol), N112Y (-47.82 kcal/mol), N112M (-45.82 kcal/mol) and N102W (-45.14 kcal/mol). It was, however, mutants of 2J predicted with high affinity for DIII initially under mSCM-PPI2 and BeAtMuSiC were not observed to have lower binding free energy than the wild type. The top choice recommended by mSCM-PPI2 and BeAtMuSiC, L69W and N112W has binding free energy -21.35 kcal/mol and -26.69 kcal/mol, which was almost two times higher than the wild type (-42.77 kcal/mol).

Table 4.5 Binding free energy of 2J and mutans in complex with DIII

System	Energy Term				
	VDW	EEL	EGB	ESURF	Δ_{GB}
Wild type	-102.60 ± 0.17	-61.14 ± 1.03	131.24 ± 0.99	-10.27 ± 0.015	-42.77 ± 0.21
V36W	-93.59 ± 0.23	-83.10 ± 1.31	150.22 ± 1.26	-10.23 ± 0.019	-36.70 ± 0.27
V36Y	-82.89 ± 0.20	-36.98 ± 1.36	110.83 ± 1.38	-7.73 ± 0.026	-16.77 ± 0.20
Y66K	-104.01± 0.19	-94.04 ± 1.18	151.71 ± 1.13	-10.56 ± 0.012	-56.91 ± 0.20
Y66R	-96.01 ± 0.29	-71.53 ± 1.58	141.34 ± 1.43	-9.40 ± 0.025	-35.60 ± 0.45
T68W	-87.23 ± 0.233	-50.97 ± 1.49	112.66 ± 1.31	-9.34 ± 0.019	-34.89 ± 0.29
T68K	-81.68 ± 0.24	-151.08 ± 1.43	201.28 ± 1.28	-8.74 ± 0.018	-40.21 ± 0.29
T68E	-91.23 ± 0.18	-59.89 ± 1.05	120.67 ± 0.99	-9.14 ± 0.014	-39.60 ± 0.20
L69W	-82.77 ± 0.19	57.05 ± 1.75	12.38 ± 1.62	-8.00 ± 0.019	-21.35 ± 0.29
L69F	-81.26 ± 0.18	81.06 ± 2.40	-1.18 ± 2.23	-7.95 ± 0.019	-9.33 ± 0.23
S71L	-103.61 ± 0.17	-50.33 ± 1.14	129.49 ± 1.14	-8.90 ± 0.015	-33.35 ± 0.18
D101S	-52.19 ± 0.25	102.98 ± 1.06	-49.47 ± 1.03	-5.01 ± 0.021	-3.69 ± 0.19

Table 4.9. Continued

D101T	-89.05 ± 0.21	-31.68 ± 1.15	98.05 ± 0.99	-9.64 ± 0.015	-32.31 ± 0.23
N102Y	-96.50 ± 0.30	-75.00 ± 0.98	123.95 ± 0.96	-9.62 ± 0.026	-57.17 ± 0.29
N102D	-82.14 ± 0.21	-33.56 ± 1.57	98.84 ± 1.50	-8.12 ± 0.021	-25.00 ± 0.20
N102L	-88.14 ± 0.22	0.63 ± 1.50	64.50 ± 1.42	-8.74 ± 0.022	-31.74 ± 0.24
N102W	-99.54 ± 0.19	-57.88 ± 1.51	121.34 ± 1.48	-9.07 ± 0.013	-45.14 ± 0.21
I104W	-84.00 ± 0.18	-21.64 ± 1.53	89.97 ± 1.37	-8.90 ± 0.019	-24.57 ± 0.31
I104Y	-17.70 ± 0.15	-35.05 ± 1.31	57.34 ± 1.22	-1.77 ± 0.015	2.82 ± 0.19
N112W	-79.89 ± 0.19	-57.02 ± 0.98	118.29 ± 0.93	-8.07 ± 0.018	-26.69 ± 0.21
N112Y	-92.32 ± .23	-60.30 ± 0.96	114.76 ± 0.91	-9.97 ± 0.018	-47.82 ± 0.22
N112M	-93.30 ± 0.21	-114.36 ± 1.56	171.42 ± 1.47	-9.58 ± 0.015	-45.82 ± 0.21
R113H	-87.74 ± 0.19	-5.19 ± 1.18	67.81 ± 1.13	-8.52 ± 0.018	-33.64 ± 0.22
R113W	-73.04 ± 0.27	-119.16 ± 1.31	160.44 ± 1.27	-7.77 ± 0.026	-39.54 ± 0.25

VDW and EEL denote van der Waals and electrostatic contribution from molecular mechanics, respectively. EGB and ESURF denote polar and nonpolar contributions under GB model. Δ_{GB} is the binding free energy without taking entropy into calculation and is accompanied with the standard error mean of the calculation.

Of the five single point mutations that decreased the binding free energy, three were interfacial residues, one located in the first repeat (Y66) and two located in the third repeat unit (N102 and N112). It was a surprise to observe non-interfacial residue has its role in improving the binding affinity. Apart from performing point mutation, double-point mutation was also attempted since multiple point mutation was successfully to increase the strength of mAb binding with DIII DENV-2 and other three dengue serotypes. Therefore, double-point mutation was attempted on 2J to investigate if it enhances binding interactions with DIII by combining those successful single point mutations. Three two-mutant combinations were constructed by combining the point mutation occurred in the first and the third repeat unit, Y66K+N102Y, Y66K+N112Y and Y66K+N112M.

On the other hand, residue T68 was also targeted to combine with N102 and N112 for two-mutant combinations. L69 from the second repeat was predicted by mSCM-PPI2 algorithm to have high affinity when tryptophan replacing leucine at this position, but the calculated binding free energy increased almost 50%. Hence, T68 was investigated if this neighboring residue of L69 could have its role in lowering the binding free energy further.

Altogether, three two-mutant combinations were having a lower binding free energy than the wild type as described in Table 4.10. Y66K+N112M (YKNM) has the lowest binding free energy (-63.9 kcal/mol), followed by T68W+N112M (-51.04 kcal/mol) and T68W+N112Y (-50.28 kcal/mol). Binding free energy of YKNM was 50% lower than the wild type. In addition, combination of single point mutation from the first and third repeat unit improve the binding interactions better than mutation that involved second and third repeat unit. Hence, binding free energies calculated under MMGBSA protocol, have ranked affinity of 2J and mutants binding with DIII accordingly, YKNM> N102Y> Y66K > 2J (wild type).

Table 4.6 Binding free energy of two-mutant combinations

System	Energy Term				
	VDW	EEL	EGB	ESURF	Δ_{GB}
Wild type	-102.60 ± 0.17	-61.14 ± 1.03	131.24 ±0.99	-10.27 ± 0.015	-42.77 ± 0.21
Y66K+N102Y	-82.74 ± 0.27	-78.97 ± 1.18	143.01 ± 1.19	-8.89 ± 0.029	-27.59 ± 0.22
Y66K+N112M (YKNM)	-106.63 ± 0.27	-90.40 ± 1.13	144.03 ± 1.12	-10.90 ± 0.025	-63.90 ± 0.27
Y66K+N112Y	-82.50 ± 0.17	1.64 ± 1.07	68.70 ± 1.00	-8.22 ± 0.013	-20.38 ± 0.17
T68K+N112M	-83.50 ± 0.18	-63.87 ± 1.23	122.76 ± 1.17	-9.09 ± 0.014	-33.70 ± 0.20
T68K+ N112Y	-73.37 ± 0.30	-154.44 ± 1.44	194.37 ±1.44	-8.34 ± 0.026	-41.78 ± 0.29
T68K +N102Y	-94.28 ± 0.19	-59.61 ± 0.89	123.33 ± 0.85	-9.60 ± 0.016	-40.16 ± 0.21
T68W+N112M	-97.83 ± 0.19	-117.07 ± 0.92	175.27 ±0.88	-10.50 ± 0.015	-51.04 ± 0.18
T68W+N112Y	-79.61 ± 0.19	-43.36 ± 1.17	81.54 ± 1.05	-8.84 ± 0.015	-50.28 ± 0.22
T68W+N102Y	-79.19 ± 0.28	-97.50 ± 1.77	160.91 ± 1.79	-8.01 ± 0.03	-23.79 ± 0.25

VDW and EEL denote van der Waals and electrostatic contribution from molecular mechanics, respectively. EGB and ESURF denote polar and nonpolar contributions under GB model. Δ_{GB} is the binding free energy without taking entropy into calculation and is accompanied with the standard error mean of the calculation.

4.3.3 Root mean square fluctuation (RMSF)

RMSF analysis has been used to study the dynamics of protein residues after mutations (Aier et al., 2016; Doss et al., 2012; Khan et al, 2019; Kumar et al., 2014). Also, flexibility study like RMSF could be useful to assess whether a biomolecule remains active after mutations (Zhu et al., 2015). In general, residues with high RMSF value indicated their flexibility is higher compared to those with lower RSMF value. Residues DARPin 2J has RMSF values in the range of 1.19 – 12.09 Å, while 1.18 – 12.62 Å for Y66K, 1.11 – 12.29 Å for N102Y and 1.09 -14.73 Å for YKNM. Terminal residues of DARPins and DIII were flexible and hence contributed to the highest RMSF value. That also explained a larger range of RMSF value observed in mutant complexes as illustrated in Figure 4.11.

Upon point mutation, residue 91-100 (repeat unit 2) Y66K and N102Y appeared to be less flexible but their C-capping unit was more flexible as compared to 2J. When two amino acids of 2J were mutated, all residues of YKNM were lower in RMSF value. For DIII, the four regions that have recorded distinctive low flexibility were the binding regions corresponded to A-strand, BC-loop, DE-loop and G-strand. DIII residues in YKNM-DIII complex has the lowest RMSF value compared to DIII complexed with Y66K and N102Y. RMSF plots of DARPin after single point and two-point mutation were highly similar and overlapped in most of the regions. There were no protein residues experienced large fluctuation suggesting both DARPin and DIII remained stable after mutations.

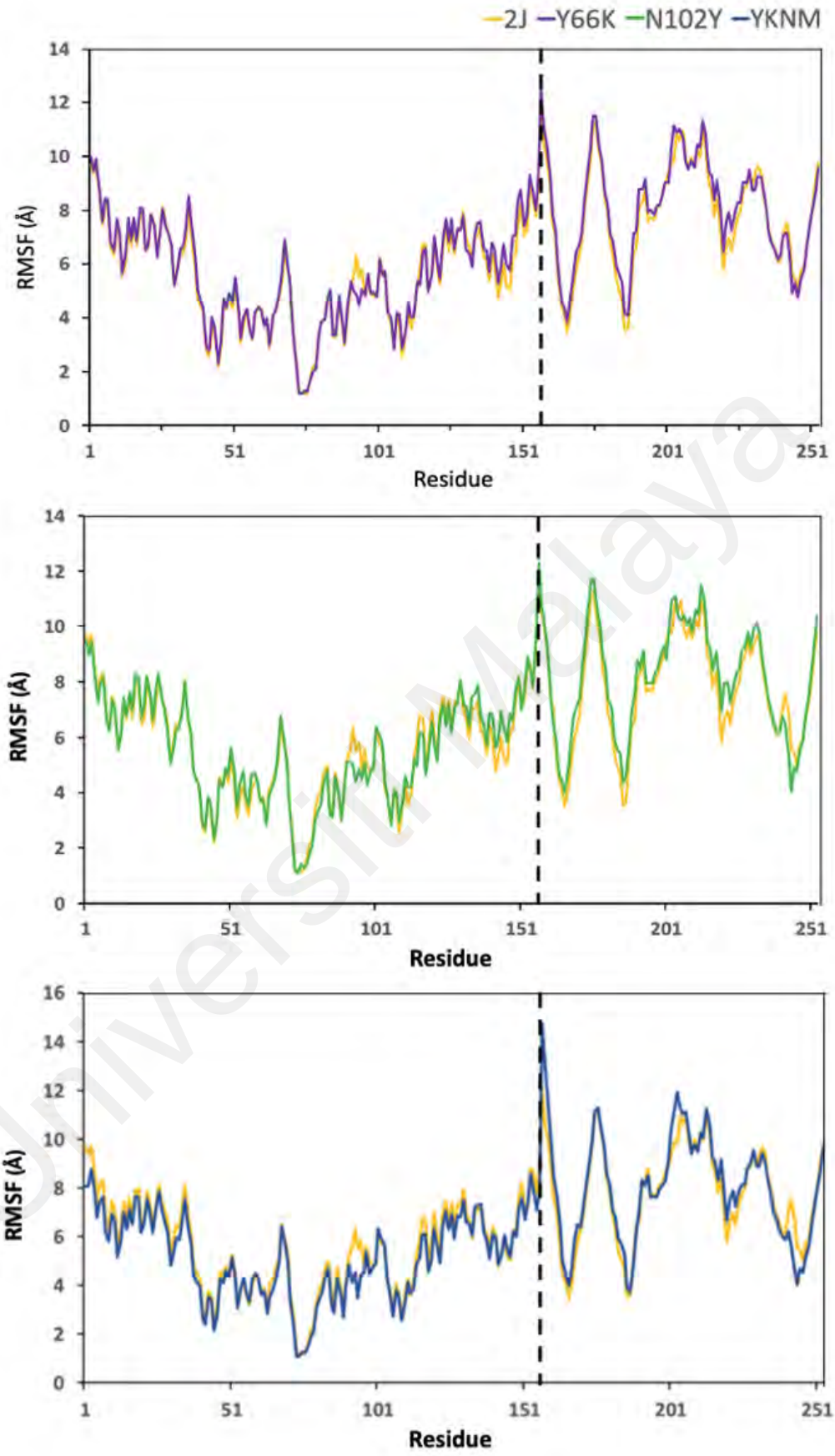


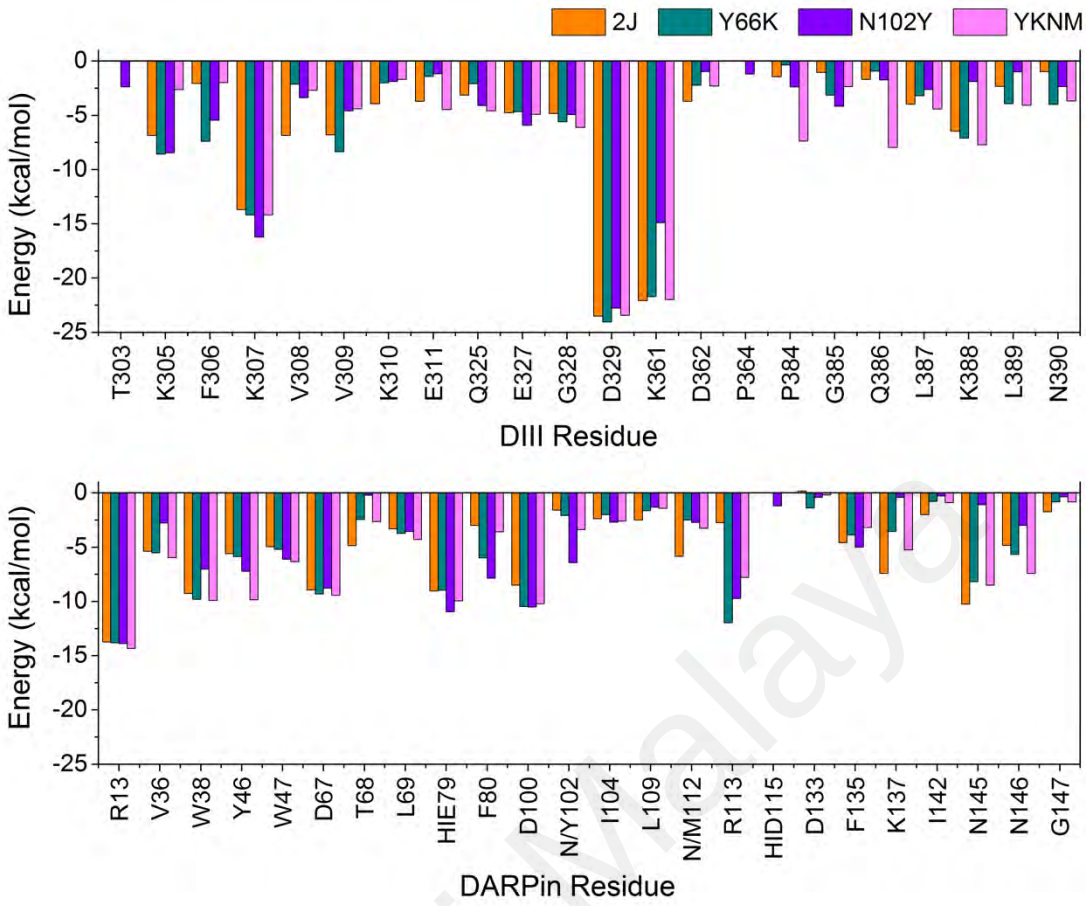
Figure 4.11 RMSF of DARPin 2J and mutants. Residue 1-156 belong to DARPin while residue 157-253 belong to DIII.

4.3.4 Insight of residue interactions between DIII and mutants

Decomposition energy of interfacial residues of wild type and mutants in complex with DIII was depicted in Figure 4.12. Also, the total decomposition energy was summed according to the five important DIII regions and repeat modules of DARPin.

K305, K307, V308, V309, D329, K361 and K388 of DIII were identified as the important residues upon association with 2J. Like wild type, K307, D329 and K361 of the top three mutants have distinctive low decomposition energy. Interaction between the mutants and V308 experienced a reduction more than 50%. Meanwhile interaction between Y66K and V309 improved despite energy of V309 remained higher for N102Y and YKNM mutants. Other than YKNM, both Y66K and N102Y interacted with K305 with lower decomposition energy. For K388, decomposition energy was lowered in Y66K-DIII and YKNM-DIII systems but not in N102Y-DIII. Energy profile of mutants was different from wild type implying binding interactions between mutants and DIII have changed slightly.

Despite energy of A-strand has reduced approximately 26% or 11.7 kcal/mol, YKNM has improved its interaction with G-strand as the energy was 80% or 12.53 kcal/mol lower than wild type. In addition, B-strand and BC-loop have strengthened their interactions with YKNM as the total decomposition energy of the two regions went lower. On the other hand, Y66K improved its interactions with DIII via BC-loop and G-strand while N102Y has strengthened its interaction with DIII via B-strand and FG-loop. DARPin YKNM and Y66K have lower total decomposition energy in N-cap and all the three repeat units but DARPin N102Y has a different energy profile as lower energy observed in repeat unit 2 and 3.



DIII/ DARPin	Region	Total decomposition energy kcal/mol					
		WT	Y66K	N102Y	YKNM	513	4E5A
DIII	A	-44.00	**-44.23**	**-43.63**	-32.27	-101.82	-101.49
	B	-3.14	-2.08	**-4.10**	**-4.60**	-27.17	-16.95
	BC	-33.10	**-34.30**	-33.60	**-34.45**	-8.49	-8.84
	DE	**-25.80**	-23.93	-17.07	-24.47	-20.15	27.49
	G	-15.53	**-19.41**	-9.70	**-28.06**	-18.64	-18.47
DARPin	N	-13.75	-13.80	-13.88	**-14.32**		
	1	-25.37	**-26.69**	-23.24	**-32.33**		
	2	-29.13	**-30.48**	**-31.38**	**-29.92**		
	3	-23.55	**-30.67**	**-34.70**	**-28.68**		
	C	**-31.82**	-24.98	-10.67	-27.34		

Figure 4.12 Decomposition energy of interfacial residues of DIII and 2J mutants in complex. Regions with lower energy than wild type were highlighted.

DARPin residues that were interacting with the important DIII residues with lowest decomposition energy were first identified. Type of interaction mediated in the residue pair was analyzed under Discovery Studio program and results were tabulated in Table 4.11.

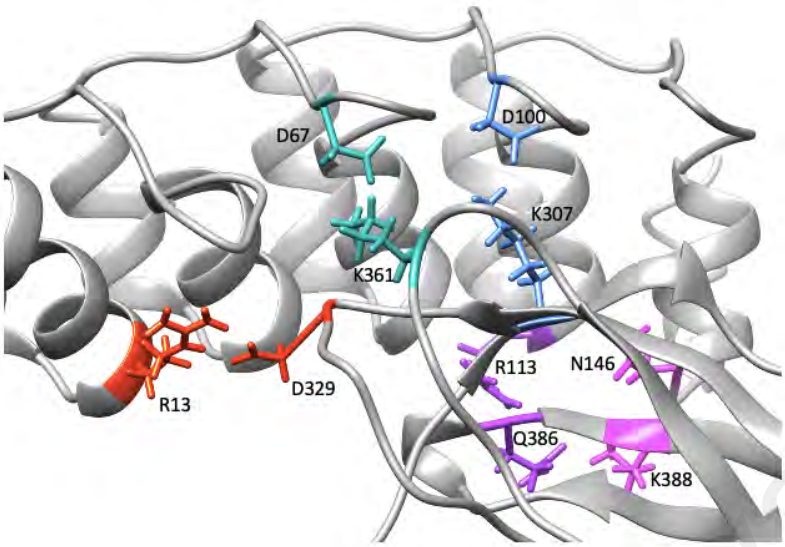
Effect of combining single point mutation from repeat unit 1 and 3 was greater and offered better enhancement of binding interactions as binding free energy went lower to -63.9 kcal/mol. Although number of residue pair interactions involved YKNM and A-strand has reduced, however participation of residue pair G328-W38 (-2.62 kcal/mol, hydrogen bond) and Q386-R113 (-7.29 kcal/mol, hydrogen bond) have increased interaction between BC-loop, G-strand and YKNM. Furthermore, interaction of K307-D100 and D329-R13 appeared to be stronger as suggested by their lower energy. Interestingly, hydrogen bond formed between K388 and N145 in wild type had switched to K388 and N146 in YKNM-DIII. Important residue pair interactions of the two single-point mutation systems were similar to wild type except involvement of an additional residue pair G328-W38 (-2.49 kcal/mol, hydrogen bond) in Y66K-DIII and absence of K388-N145 in N102Y-DIIII. Besides, hydrogen bond interaction K305-HIE79 in wild type become Pi-cation in Y66K-DIII and N102Y-DIII owing to the fact that K305 was not making contact with HIE 79 but F80. Instead of interacting with K305, HIE79 in N102Y-DIII made contact with E327 through hydrogen bond (-3.77 kcal/mol). Detailed hydrogen bond network of all the three mutant-DIII complexes were presented in Table 4.12.

Table 4.7 Important residue pairs interaction in mutants

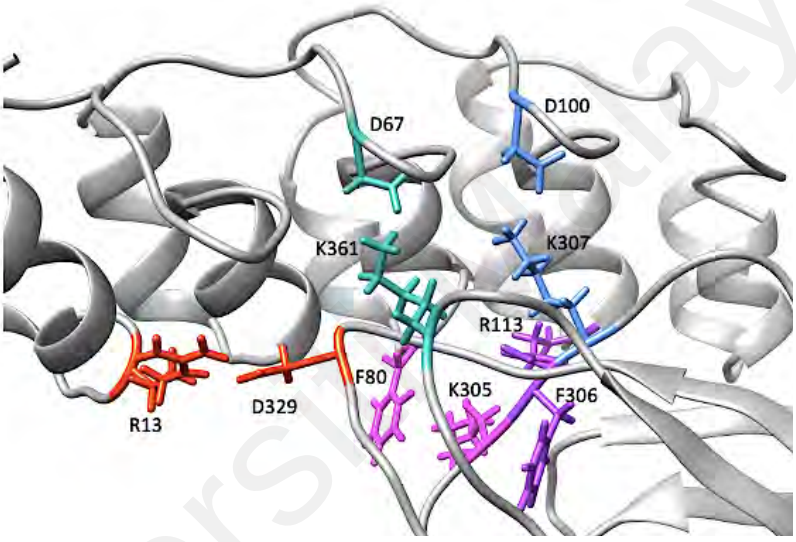
System	Residue pair		Energy	Interaction	Repeat	Strand
	DIII	DARPin	(kcal/mol)		module	
YKNM	K307	D100	-10.79	Salt bridge	3	A
	G328	W38	-2.62	H-bond	1	BC
	D329	R13	-14.42	Salt bridge	N	BC
	K361	D67	-9.12	Salt bridge	2	DE
	P384	F80	-3.02	Pi-Alkyl	2	FG
	Q386	R113	-7.29	H-bond	3	G
	K388	N146	-4.77	H-bond	C	G
Y66	K305	F80	-5.25	Pi-cation	2	A
	F306	R113	-6.51	H-bond	3	A
	K307	D100	-10.61	Salt bridge	3	A
	V309	K137	-2.50	H-bond	C	A
	G328	W38	-2.49	H-bond	1	BC
	D329	R13	-13.88	Salt bridge	N	BC
	K361	D67	-9.99	Salt bridge	1	DE
	K388	N146	-4.06	H-bond	C	G
N102	K305	F80	-4.53	Pi-cation	2	A
	F306	R113	-4.53	H-bond	3	A
	K307	D100	-10.24	Salt bridge	3	A
	E327	HIE 79	-3.77	H-bond	2	BC
	D329	R13	-13.91	Salt bridge	N	BC
	K361	D67	-9.90	Salt bridge	2	DE

Among the top five residue pairs of the three mutant-DIII complexes shown in Figure 4.12, three were three strong opposite charged interactions (D329-R13, K361-D67 and K307-D100). Remaining two residue pairs in YKNM-DIII were electrostatic interactions (Q386-R113 and K388-N146) while the other two pairs in Y66K-DIII and N102Y-DIII were interactions between charged amino acids (R and K) and non-polar amino acid (F).

(a)



(b)



(c)

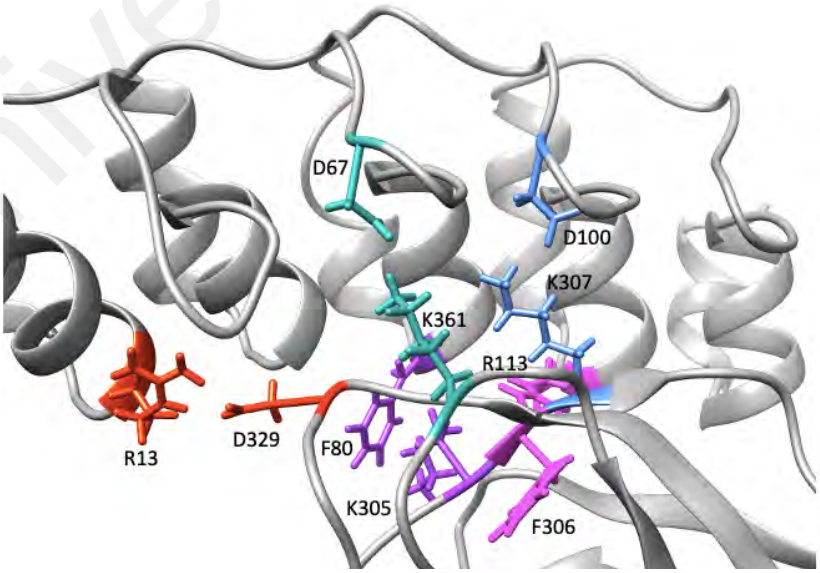


Figure 4.13 Top five important residue pairs in (a) YKNM-DIII, (b) Y66K-DIII and (c) N102Y-DIII. Orange residue pair has the lowest decomposition energy, followed by blue, green, purple and magenta.

Number of hydrogen bonds in YKNM-DIII was five more than the wild type. Eleven out of twenty hydrogen bonds were having occupancy percentage greater than 50% and hence these interactions were stable. Also, there were altogether eight salt bridges formed between YKNM and DIII. Extensive hydrogen bonds and salt bridges formed in K361-V36 and K361-D67 with occupancy percentage 20-30% while two strongest salt bridges formed in D329-R13 with occupancy percentage 96.4% and 86.2%. Other than salt bridges, D329 also form hydrogen bond with W47 with occupancy percentage 68.8%. Hence, noticeably low decomposition energy of K361-D67 and D329-R13 could be explained by formation of salt bridges and extensive hydrogen bonds. To be part of the top five crucial residue pairs, salt bridge was found in K307-D100 (54%) while two hydrogen bonds formed between Q386 and R113 (71.4% and 67.8%) and lastly, hydrogen bond found in K388-N146 was 45.4%.

Increase in occupancy percentage and number of hydrogen bond or salt bridge formed accompanied with a reduction of decomposition energy. D329-R13 decreased to -14.42 kcal/mol from -13.82 kcal/mol; K361-V36 decreased to -5.24 kcal/mol from -4.58 kcal/mol and K307-D100 reduced to -10.79 kcal/mol from -9.32 kcal/mol. Strongest hydrogen bond between K388 and 2J (K388-N145) was -3.32 kcal/mol while K388 and YKNM in YKNM-DIII (K388-N146) reduced to -4.77 kcal/mol.

Table 4.8 Hydrogen bond network of mutant-DIII

Complex	Acceptor	Donor	%	Distance (Å)	Angle	DIII	Repeat module
YKNM	D329@OD1	R13@HH21	96.4	2.78	163.22	BC	N
	D329@OD2	R13@HE	86.2	2.84	163.03	BC	N
	P364@O	Y46@HH	81.8	2.77	162.12	DE	1
	Q325@OE1	N102@HD22	73.1	2.84	162.80	B	3
	Q386@O	R113@HH12	71.4	2.83	151.10	FG	3
	D329@OD2	W47@HE1	68.8	2.85	159.51	DE	1
	Q386@O	R113@HH12	67.8	2.82	150.58	G	3
	E327@O	HIE79@HE2	63.3	2.84	152.48	BC	2
	D100@OD2	K307@HZ1	57.0	2.75	160.99	A	3
	E327@OE1	S71@HG	54.2	2.70	163.35	BC	2
	G328@O	W38@HE1	50.3	2.86	146.85	BC	1
	N146@OD1	K388@H	45.4	2.88	154.98	G	C
	E327@OE2	S71@HG	39.3	2.69	163.40	BC	2
	D100@OD2	K307@HZ2	31.4	2.76	160.22	A	3
	D67@OD1	K361@HZ1	30.1	2.76	153.20	DE	2
	D67@OD1	K361@HZ3	30.0	2.77	153.79	DE	2
	V36@O	K361@HZ1	27.0	2.84	159.50	DE	1
	V36@O	K361@HZ2	26.1	2.84	159.04	DE	1
	D67@OD1	K361@HZ2	24.6	2.77	152.90	DE	2
	V36@O	K361@HZ3	22.0	2.84	159.18	DE	1
Y66K	F306@O	R113@HH22	91.7	2.80	155.79	A	3
	D329@OD2	R13@HE	90.8	2.82	163.56	BC	N
	D329@OD1	R13@HH21	90.1	2.90	163.05	BC	N
	E327@OE1	S71@HG	80.8	2.68	164.59	BC	2
	D329@OD2	W47@HE1	64.8	2.85	155.97	BC	1
	F306@O	R113@HH12	41.3	2.88	147.08	A	3
	G328@O	W38@HE1	36.5	2.86	149.14	BC	1
	D67@OD2	K361@HZ3	32.4	2.74	155.45	DE	2
	N145@OD1	N390@H	27.7	2.88	158.33	G	C
	D100@0D1	K307@HZ2	26.7	2.76	155.75	A	3
	D100@0D1	K307@HZ1	26.0	2.76	156.18	A	3
	D67@OD2	K361@HZ2	25.7	2.74	154.84	DE	2
	D100@OD1	K307@HZ3	23.9	2.77	154.17	A	3
	E327@O	HIE79@HE2	21.7	2.86	146.53	BC	2
	V36@O	K361@HZ2	21.1	2.84	152.44	DE	1

Table 4.12, continued

N102Y	D329@OD1	R13@HH21	94.6	2.78	160.52	BC	N
	D329@OD2	R13@HE	82.9	2.84	163.17	BC	N
	D329@OD2	W47@HE1	70.5	2.85	159.35	BC	1
	E327@O	HIE79@HE2	65.5	2.82	154.50	BC	2
	E327@OE1	S71@HG	36.2	2.66	162.31	BC	2
	F306@O	R113@HH12	33.1	2.85	155.25	A	3
	D100@OD2	K307@HZ2	32.1	2.74	157.32	A	3
	F306@O	R113@HH22	31.2	2.85	154.95	A	3
	Q325@OE1	Y102@HH	30.3	2.76	159.34	B	3
	D100@OD2	K307@HZ1	28.5	2.74	157.93	A	3
	D100@OD2	K307@HZ3	27.8	2.74	156.30	A	3
	D67@OD2	K361@HZ3	24.1	2.78	158.59	DE	2

Figure 4.14 Electrostatic potential surface map of 2J and mutants in complex with DIII

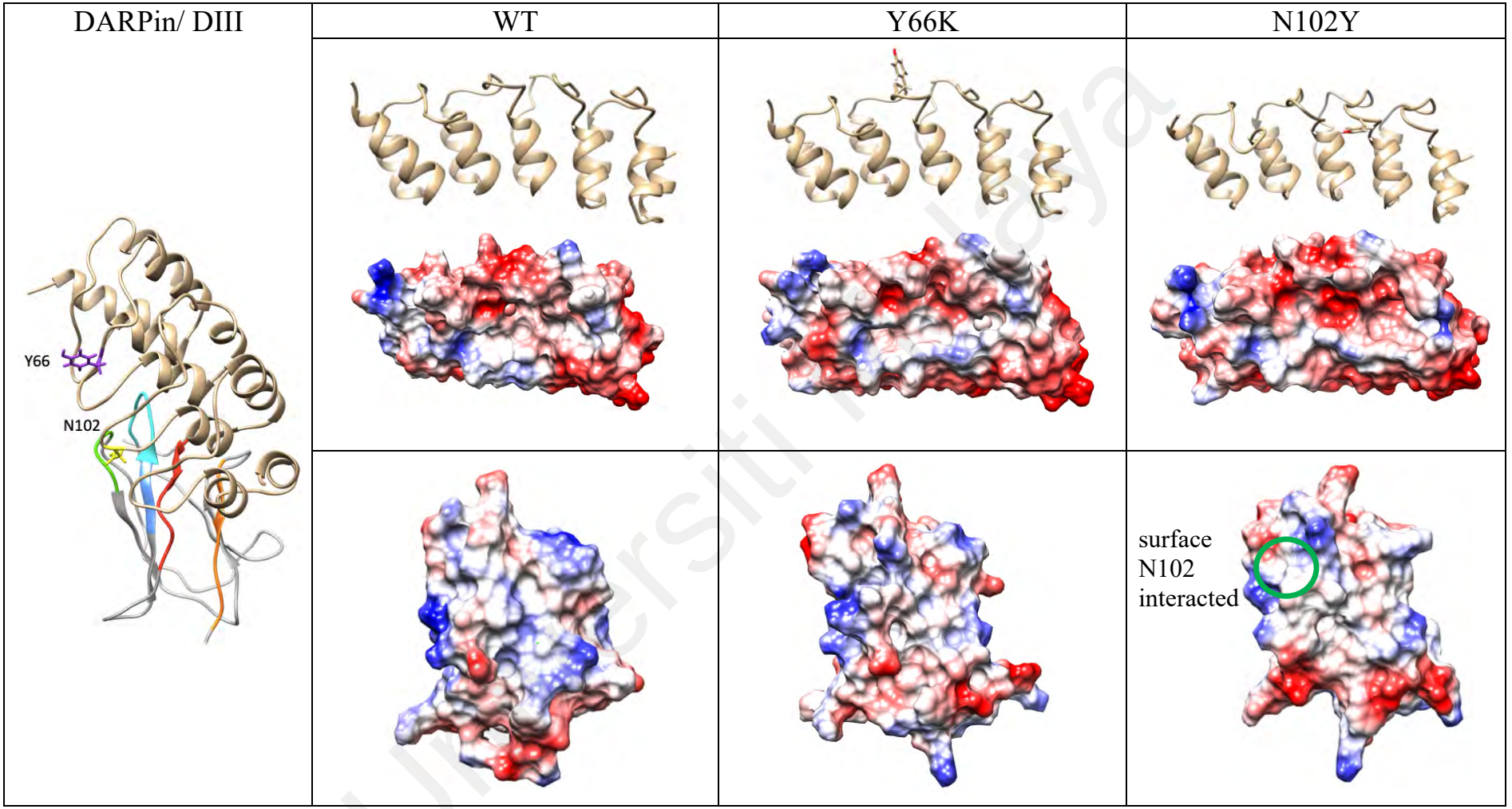
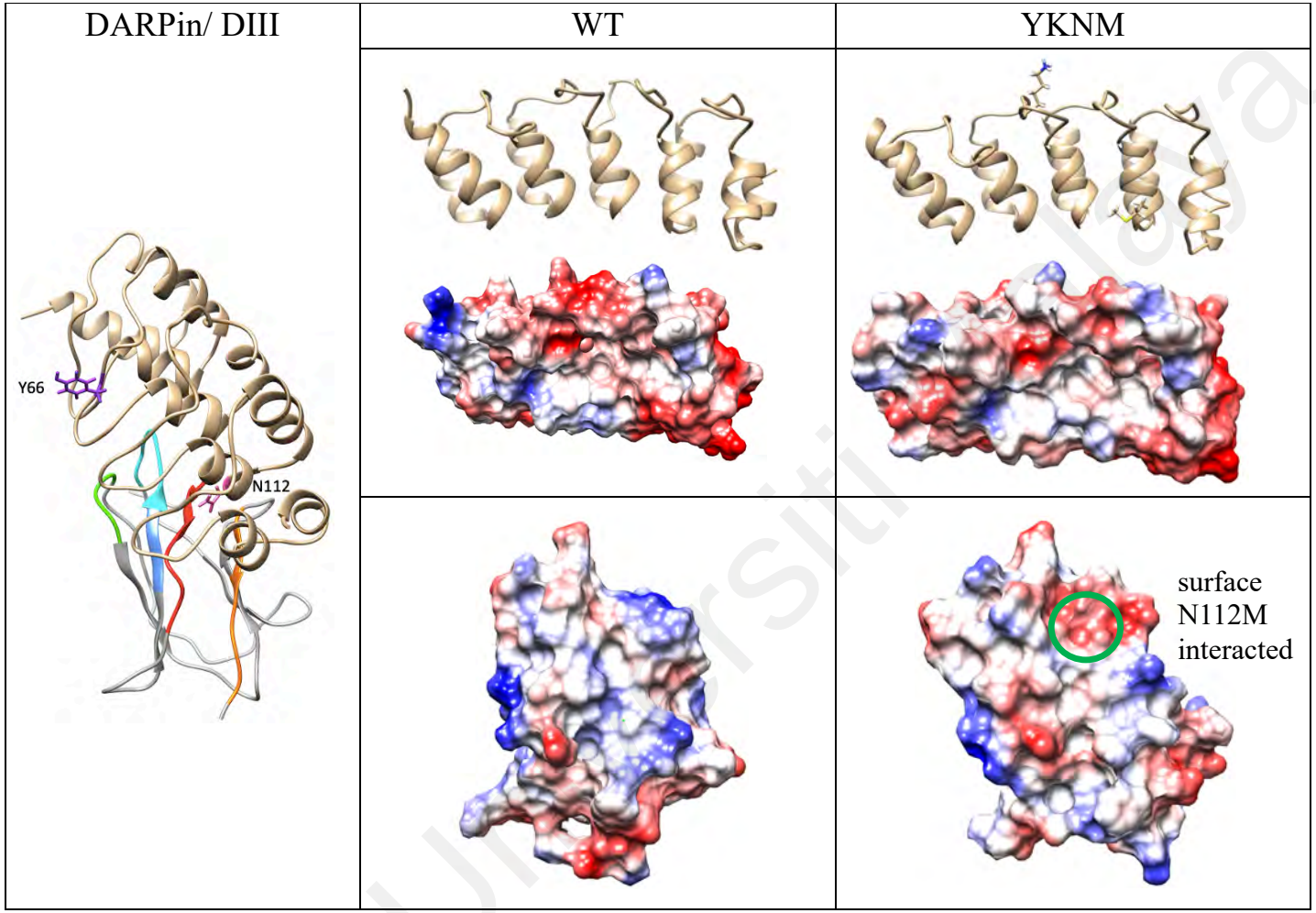


Figure 4.14, Continued



Electrostatic potential of protein surface was coloured using Coulombic Surface Coloring functions embedded in UCSF Chimera programme. Electrostatic potential was calculated using Coulomb's law and surface was coloured from red to blue for negative to positive potential while near neutral appear to white. Energy from negative potential to positive potential ranged from -10 kcal/mol to 10 kcal/mol. Electrostatic interactions are long-range and essential for protein-protein interactions (Zhang et al., 2011) and hence playing important role for binding and recognition in protein-protein systems (Campbell et al., 2014; Chakavorty et al., 2016). All DARPins, 2J and mutants were separated from DIII to observe charge distribution change on the protein surface upon a point mutation as illustrated in Figure 4.14.

To gain an insight on how mutating an amino acid at particular position would change the charge distribution and hence the electrostatic potential, YKNM-DIII was examined and compared with 2J-DIII by focusing on the top five residue pairs in YKNM-DIII and 2J-DIII. In Figure 4.15, significant changes occurred in region where R113 of both systems located. In 2J-DIII, positively charged R113 was found interacting with P364 which was non-polar while R113 in YKNM-DIII interacting with Q386 which is polar. DIII region that was interacting with HIE79 and R113 residues in YKNM-DIII was colored red indicated a negatively charged region. Together with its adjacent HIE 79 which is also positively charged, R113 in YKNM-DIII have interacted electrostatically with DIII and hence decreasing the decomposition energy.

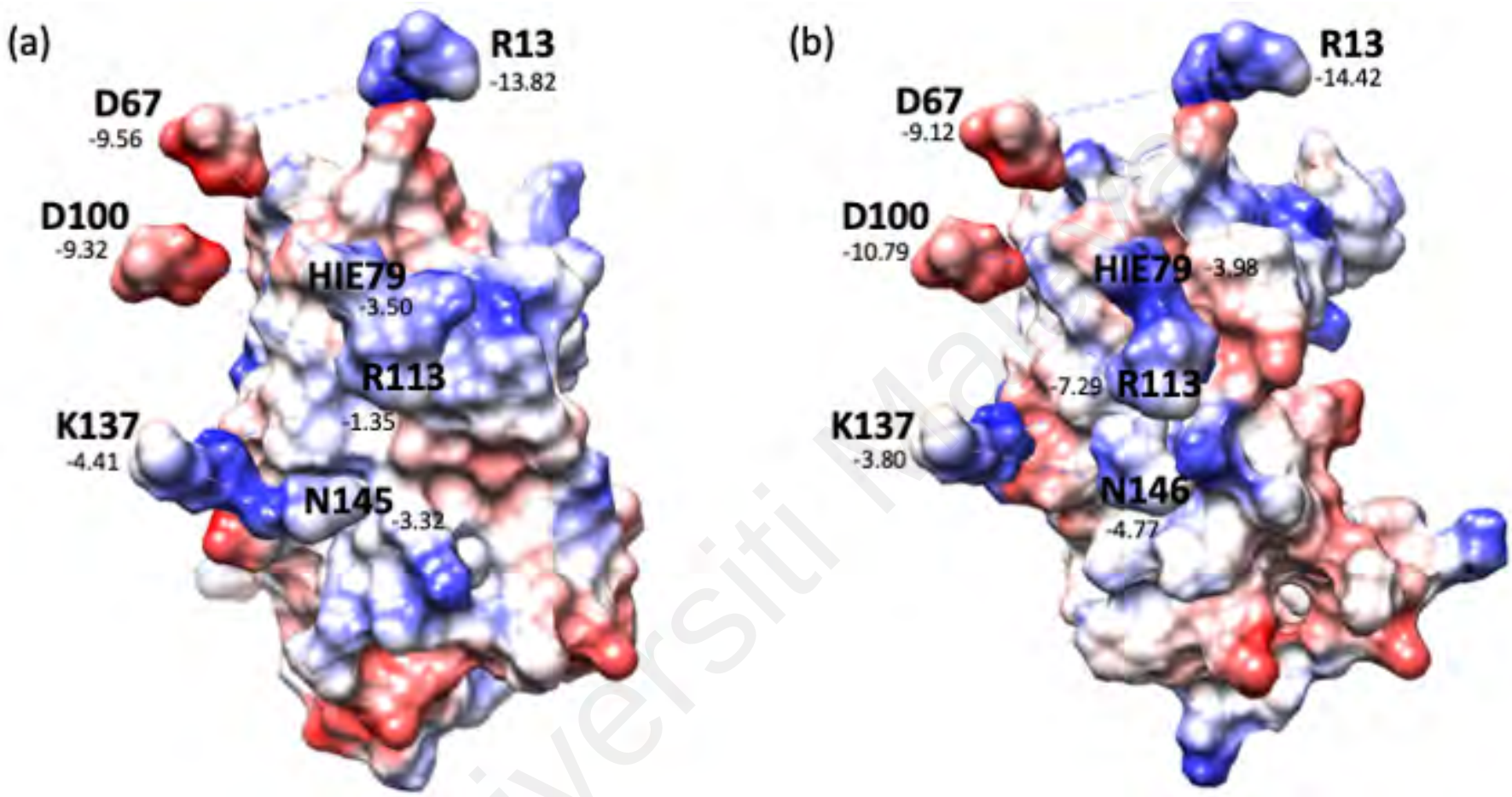


Figure 4.15. Electrostatic potential surface of important residue pair in (a) 2J-DIII and (b) YKNM-DIII. Decomposition energy of the respective residue pair was displayed.

4.4 GNM analysis of DIII in complex with scFv and DARPin

Other than free energy calculations, dynamics and flexibility of proteins have been extensively explored and studied for understanding the fluctuations that correlated to the protein functions. In X-ray crystallography study, Debye-Waller temperature factors or B-factors explains the atomic vibrations under thermal influence or in a simpler term, flexibility of the structure. B-factors, or the distribution of the displacement of the atoms as found in the protein crystal structure would reveal the flexibility together with the dynamics corresponding to the structure or even its sequence alone (Yuan et al., 2005 & Teasdale, 2005) . Flexibility of the residues in active sites of protein is very often observed to be lower than other residues. Therefore, the binding sites of proteins are distinguishable from other residues.

Conformational dynamics of a large and flexible molecule like protein is essential to understand and correlate protein structure and functions (Karabencheva-Christova et al., 2013; Karplus, 2002; Singh et al., 2016). In protein-ligand binding systems, mutations occur affected the flexibility of the residues particularly those located at catalytic site or active site for binding mutant more flexible (Singh et al., 2016; Zhao et al., 2015). It highlighted the mobility of specific region of the proteins.

4.4.1 Mode shape of scFv-DIII complexes

Under GNM, one of elastic network models, mobility of protein residues or mode shapes of the protein complexes was first analyzed and explored to find the correlation between protein dynamics and affinity in protein binding. Mobility of the protein residues or mode shape of DIII complexed with 4E11 and 4E5A were shown in Figure 4.16. Significant changes in mobility were observed on DIII residues especially on A-strand using the final MD structure of the two complexes, as highlighted in Figure 4.16 (a) and (b). Mode shapes in Figure 4.16 (c) captured the detailed differences in mobility of every

individual residue in the scFv-DIII complexes. Overall, mode shape of 4E5A was lower to that of 4E11 in complex. Amino acids sequence of the two mAbs differ at 5 positions (heavy chain, VH-A55; light chain, VL-R31, VL-N57, VL-E59, VL-S60) and these regions were highlighted. A lower mode shape was seen in these regions indicating mobility of the five protein residues and their neighboring residues have been reduced. Reduced mobility was found correlated with the trend of decreasing total decomposition energy of the region containing the mutation point. Total decomposition energy of region containing R31K reduced to -30.38 kcal/mol from -21.53 kcal/mol while region containing the three mutation points fall to -25.74 kcal/mol from -15.76 kcal/mol. The greatest reduction of total decomposition energy occurred in the region containing A55E mutation point, from -4 kcal/mol to -31.35 kcal/mol.

Besides, mobility of the protein residues lies within the epitope DIII have also decreased as portrayed by the four common distinct sharp peaks of DIII in Figure 4.16 (c). The four distinct minima corresponded to V309, Q325, P364 and Q385 while two additional minima occurred in 4E5A-DIII complex were due to E327 and N390. The four common minima of the scFv-DIII complexes corresponded to residue in A-strand (305-312), B-strand (320-325), DE-loop (361-364) and FG-loop (381-385), while additional minima around E327 and N390 belong to BC-loop and G-strand, respectively. Mobility of the minima residues and their neighboring residues were restricted and hence the residues in that small region (3-5 residues) has a lower mode shape compared to other residues.

Overall, mode shape of DIII was lower when it complexed with 4E5A. Largest difference of the total decomposition energy happened on A-strand, from -66.55 kcal/mol decreased to -101.48 kcal/mol. Meanwhile, total decomposition energy of B-strand decreased to -16.95 kcal/mol from -8.45 kcal/mol, -3.18 kcal/mol to -8.84 kcal/mol in

BC-loop, -13.69 kcal/mol to -21.52 kcal/mol in DE-loop and lastly, -14.92 kcal/mol to -18.47 kcal/mol in G-strand.

Although those region in 4E5A were lower in mobility and accompanied with a lower total decomposition energy, however, difference in total decomposition energy was not proportional to the difference in mobility or mode shape. Mode shape was useful to snapshot the dynamics of the scFv-DIII complexes that were contributing for the binding free energy. It serves as a simple yet powerful tool to examine the dynamics of parent mAb and re-designed mAb towards their target protein.

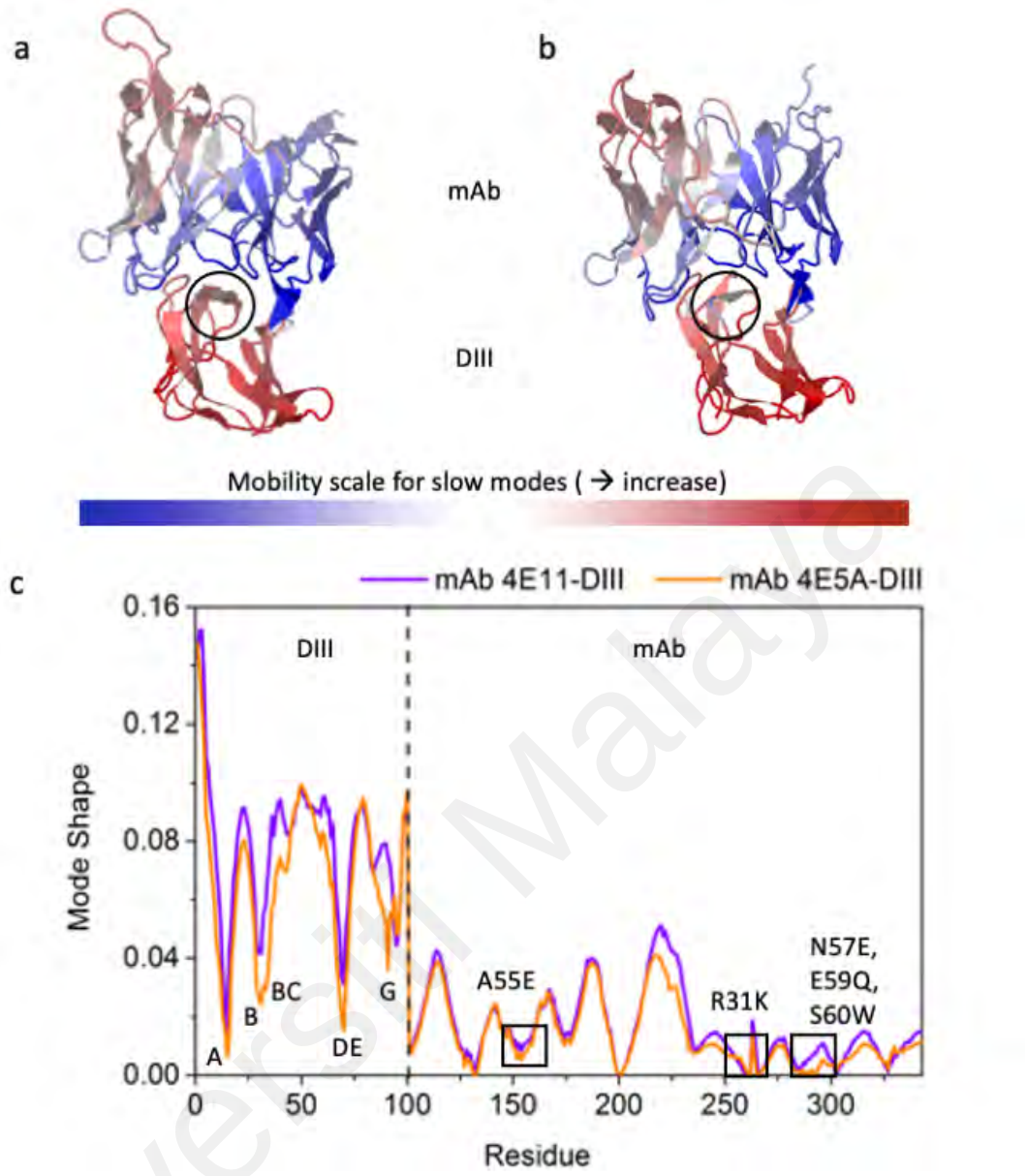


Figure 4.16 Mobility of the protein residues of (a) 4E11-DIII and (b) 4E5A-DIII complex while their mode shape is compared and showed in (c) using final MD structure. DIII regions containing epitope residues and CDR residues underwent mutations were highlighted.

4.4.2 Mode shape of DARPin-DIII complexes

To analyze the dynamics of the DARPin-DIII complexes, mode shape of the residues in complexes has been derived from the protein motions as described in Figure 4.17 using final MD structure extracted from (a) 20 ns, (b) 50 ns, (c) 80 ns and (d) 100 ns of simulation time interval. In 2J-DIII complex, four remarkable sharp peaks or minima were due to very low mobility of residue K308, D329, D361 and G385. These residues were part of the A-strand, BC-loop, DE-loop and FG-loop. These four regions constantly remain a low mobility and behaved as minima along the simulation. Like scFv-DIII complex, these residues and their neighboring residues experienced restricted mobility and hence can be termed as small regions. Mode shape of DIII complexed with 2J was very similar to mode shape of DIII complexed with the two scFvs, 4E11 and 4E5A as the same four small regions involving residues of A-strand, BC-loop, DE-loop and FG-loop were observed to have their mobility restricted.

However, mode shape of DIII in 3C-DIII complex behaved very differently to that of scFvs throughout simulation time. DIII became highly mobile during 50 ns and mobility decreased and fall back to previous magnitude at 80ns and maintained a similar dynamic profile towards the end of simulation. In Figure 4.17 (d), mode shape of DIII revealed regions with minima at residue V308, E311, R323, D341, D375, G385 and F392. These residues were part of A-strand, B-strand, C-strand, F-strand, FG-loop and G-strand. Nevertheless, minima at BC-loop and DE-loop regions were not observed in 3C-DIII complex.

2J and 3C were differ in their amino acids sequence at 18 residues and that explained the mode shape of these two DARPins were not superposable like the two scFvs. Total decomposition energy of A-strand, BC-loop, DE-loop, FG loop of DIII was lower in 2J-DIII complex except G-strand. Despite a lower total decomposition energy found in G-

strand of 3C-DIII but mode shape of that region was not observed to be lower than that of 2J-DIII. Overall, mobility of DIII and DARPin regions that were important for protein-protein interactions normally low and accompanied with low total decomposition energy.

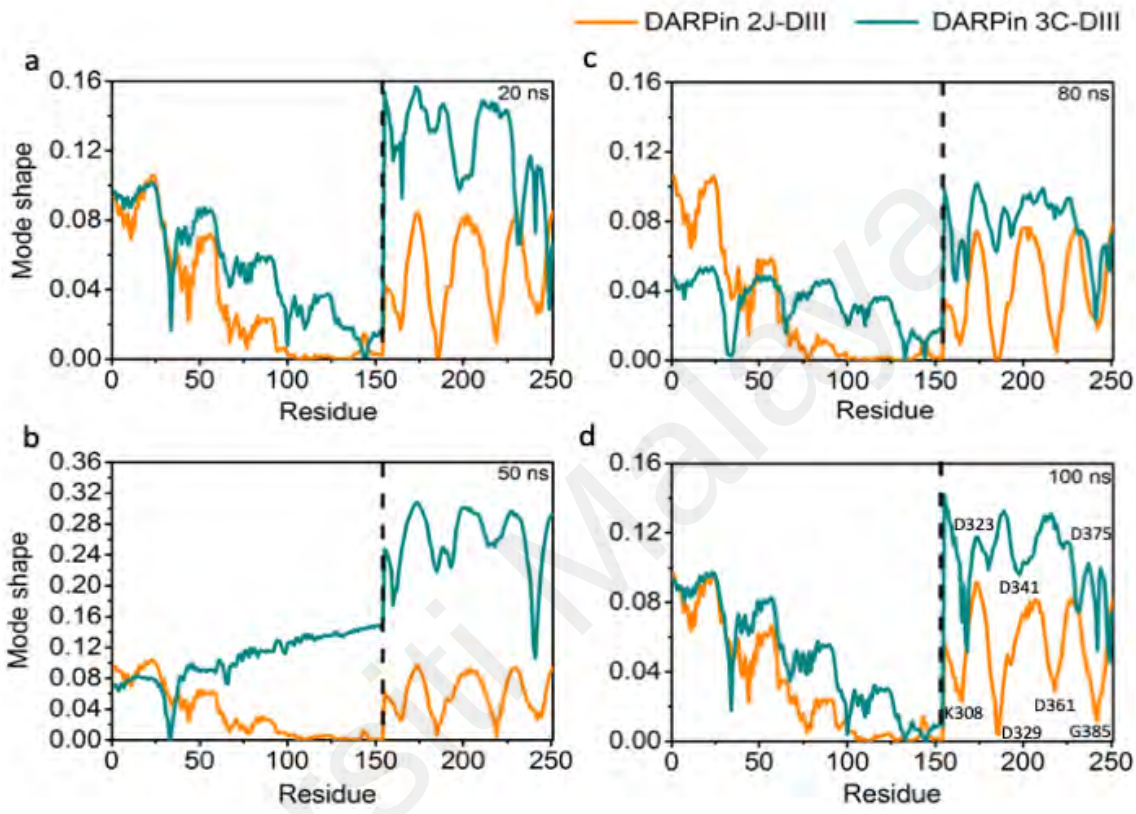


Figure 4.17 Mode shape of DARPins-DIII complexes after (a) 20 ns, (b) 50 ns, (c) 80ns and (d) 100ns of MD simulations. Residue 1 – 154 belong to DARPin whereas residue 155-250 belong to DIII DENV-2. Minima of DIII were labelled according to the numbering of PDB structure.

In an earlier report, protein kinases exhibited higher global mode fluctuations in its inactive conformation to that of active form (Kalaivani & Srinivasan, 2015). In addition, higher number of residues possessed larger fluctuations in the inactive conformation compared to those in active conformation. From an enzyme dynamics study, mobility of residues positioned in the catalytic site are highly restricted to that of average residues (Yang & Bahar, 2005). Association occur between two proteins could rely on several key amino acids and or groups of amino acids that are distributed across the protein interfaces and contributing for the binding free energy. Therefore, it postulated a fact that proteins

with greater affinity in binding would be indicated by a lower global mode shape or possessing more residues with restricted mobility compared to average residues, particularly residues responsible for the binding interactions. Ample studies have also suggested key amino acid and or group of amino acids contributed to the protein-protein interactions were very often found on the protein interfaces and distinguishable from other amino acids with significant binding free energy (Bogan & Thorn, 1998; Clackson & Wells, 1995; Keskin et al., 2005; Ma et al., 2001).

Hence, binding affinity of the two DARPins towards DIII was not only distinguishable by binding free energy but also their respective mode shape. 2J has a lower binding free energy and more residues or regions with restricted mobility implying its affinity in binding with DIII was higher to that of 3C. Mode shape was seen to be useful to locate the important regions of protein for protein-protein interactions and also differentiate the affinity of binding between two proteins. Next, mode shape of the best ranked designed DARPins was studied and compared to 2J.

As expected, mode shape of 2J and its mutants, Y66K, N102Y and YKNM were more superposable owing to their high similarity of amino acids sequence as illustrated in Figure 4.18.

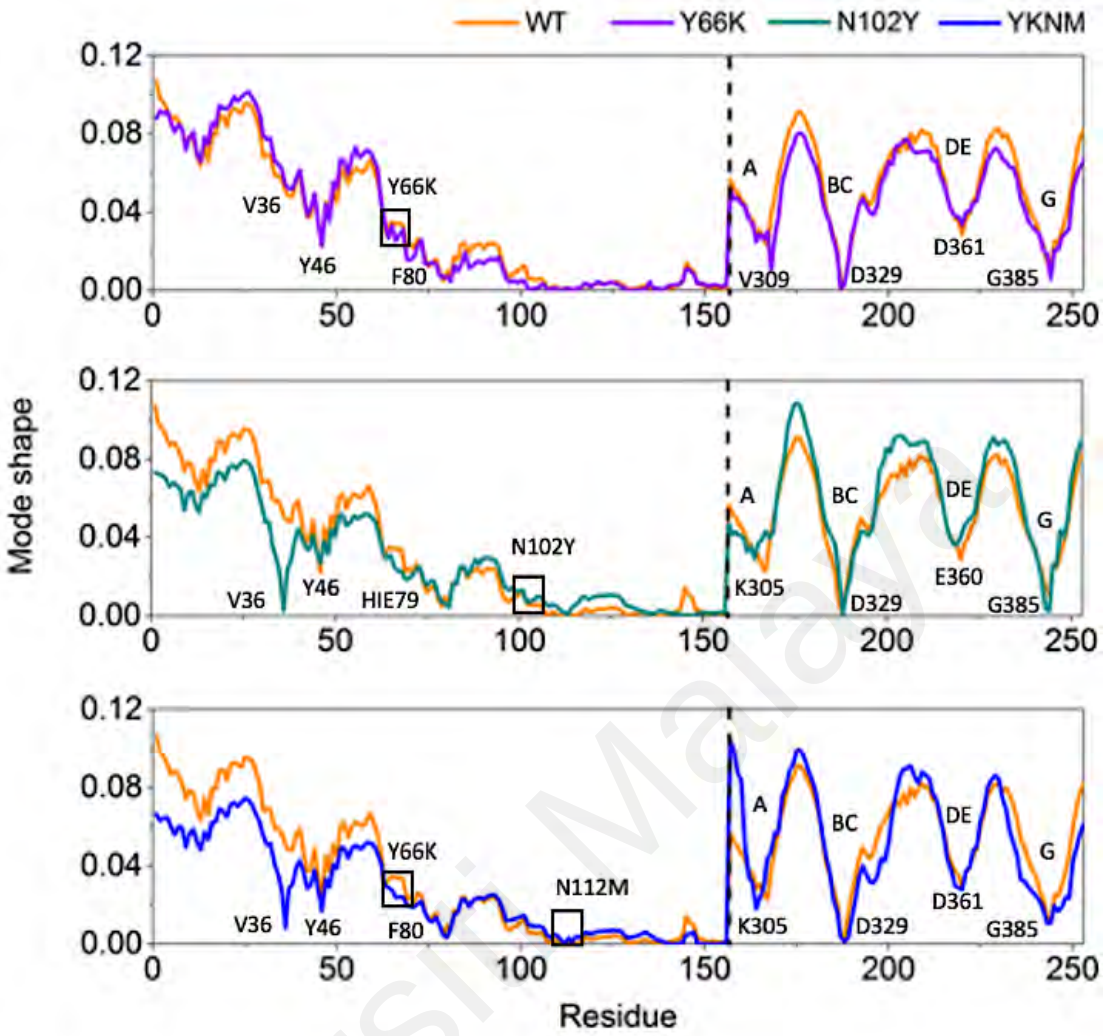


Figure 4.18 Mode shapes of designed DARPin in complex with DIII DENV-2. Mutation points on DARPin and DIII regions that are important for binding affinity were highlighted. Residue 1-156 belong to DAPRin while residue 157-253 belong to DIII.

Despite residue 66 was not part of the interfacial residues, however, replace tyrosine with lysine at this point has reduced flexibility of the loop region containing residue 66 and the whole domain of repeat 2 (67-99). Total decomposition energy of all the three repeat units were lower in Y66K. Total decomposition energy of C-capping unit of Y66K was not as low as wildtype, -25 kcal/mol compared to -31.8 kcal/mol while N-capping unit maintained at -13.80 kcal/mol. Although the magnitude of difference in flexibility was small in C-capping unit (133-156), lower mode shape of 2J could be seen. Of the four regions in DIII, total decomposition energy of DE-loop in 2J was lower and hence gave

rise to a lower mode shape. Having tyrosine replaced asparagine at residue 102, the most significant change in total decomposition energy occurred in repeat unit 3 (100-132), about -11 kcal/mol lower, however higher mode shape was observed. In addition, mode shape of N102Y was lower in N-capping (1-33) and repeat unit 1 (34-66) region despite their total decomposition energy was higher than 2J. Similar incident also happened in G-strand of N102Y-DIII where a much higher total decomposition energy accompanied with lower global mode shape. In YKNM, double point mutations have brought the total decomposition energy and mode shape lower for DAPRin. On DIII, minimum near A-strand was different from that of 2J, it has shifted to K305 from V308 since YKNM was not interacting with V308 and it was consistent with decomposition energy of the two residues. Much lower total decomposition energy of FG-loop to G-strand in YKNM resulted a lower mode shape while total decomposition energy near BC-loop and DE-loop of YKNM and 2J was comparable and hence overlapping of the two mode shapes was observed.

Based on previous experience on global mode shape of DARPin E40/ERK2 kinase protein, binding region of ERK2 in the complex would have less fluctuation whereas mobility of active DARPin, DARPin E40 was less compared to another inactive DARPin upon binding with ERK2 (Gautam et al., 2019). It has further supported the findings obtained from global mode shape of scFv-DIII complexes and DARPin-DIII complexes as the epitope residues of DIII as well as scFv region that came to interact were underlined. However, magnitude of mode shape of a residue or the small region was not observed to be proportional to the total decomposition energy.

GNM offered a very fast, easy-to-use dynamics technique. Observing the mode shapes of scFV-DIII and DAPRin-DIII complexes, binding affinity and flexibility of protein were found to be correlated. It simplified the process to distinguish the binding affinity

protein and its derivatives upon point mutation especially their affinity in binding with target protein. Different regions of protein would experience different magnitude of flexibility and respective energy. Region of the protein with lower flexibility could have a higher energy and similarly energy of a flexible region have a lower mode shape, like N102Y-DIII system. Performing site-directed mutagenesis on protein would alter the dynamics of that region as well as neighboring regions. Assuming that a strong and tight binding interaction is less likely to have much fluctuation, hence, region of DARPin that underwent site-directed mutagenesis was expected to result less fluctuation and low energy in the complex.

4.4.3 Dynamic cross-correlations

A closer view on dynamics of proteins would be useful to understand if movement of the amino acid residues found in different regions of the protein is correlated and how they move in relation to each other. Correlations between residue pairs of two interacting proteins could be revealed by residue fluctuations under elastic network model like GNM without calculating the decomposition energy using the interfacial residues within the binding interface. Protein residues having correlated motions are normally part of the structural domains of the protein and their respective motions and protein function are related (Mishra & Jernigan, 2018). DCC maps was then plotted to capture the motions of the protein residues that are correlated after and before computational site-directed mutagenesis performed on scFv-DIII and DARPin-DIII complexes.

DCC maps of 4E11-DIII and 4E5A-DIII in Figure 4.19 (a) and (b) were similar except the three highlighted regions. It was in line with the findings obtained from mode shapes where dynamical differences mostly occurred near the mutation residues in 4E5A. Notable correlated motion included:

(1) 4E5A residues 288-292 (corresponded to mutation N57E, E59Q and S60W) and A-strand residues (K305-F306), B-strand and BC-loop (323-329), DE-loop (360-364) and G-strand (386-390).

(2) 4E5A residues 152-155 (154: A55E) and A-strand (13: K307 and 14: I308), DE-loop (68: D362, 69: S363)

(3) 4E5A 130-133 and A-strand (11-16, 305-310), B-strand (320-323) and DE-loop (66-70, 361-363). Although this region was not in the same CDR region as 152-155, but very close to residue 152-155 and located near the adjacent CDR region.

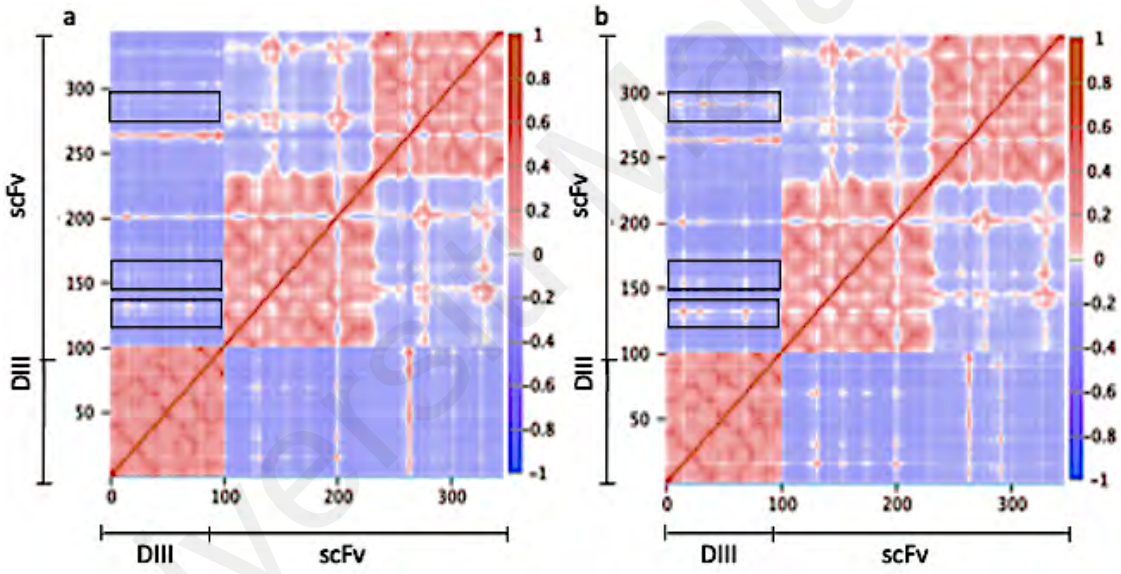


Figure 4.19 Dynamics cross-correlations for DIII DENV-2 complexed with (a) scFv 4E11 and (b) scFv 4E5A. Residue 1-100 belong to DIII DENV-2 while residue 101-343 belong to the scFvs. Red color denotes highly correlated residues while blue color denotes uncorrelated residues.

Changes in fluctuations of those regions supported that re-designed 4E5A could interact with DIII better than its parent 4E11. Interestingly, motion of mutation residue R31K (residue 263) was highly correlated to DIII after and before mutation. It suggested amino acid with basic side chain at this position was essential for correlated motion with DIII.

Cross-correlations between residue fluctuations of 2J-DIII and YKNM-DIII presented in Figure 4.20 described four major changes after double point mutation:

(1) Lower total decomposition energy of C-capping unit 2J was not only resulting a lower mode shape (Figure 4.17) but also a higher correlation motion with DIII.

(2) Motion of residue pair P384-R113 and their neighbouring residues were correlated in 2J-DIII but not in YKNM-DIII as YKNM did not make contact with V308. Correlated motion related to V308-N112 in 2J-DIII and K30Q386-R113 in YKNM-DIII was also observed.

(3) P384 in YKNM-DIII has its motion correlated with F80 unlike P384 in 2J-DIII that interacted with N112.

(4) Cross-correlations between residue fluctuations for residue pair G328-W38 and G328-Y46 in YKNM-DIII were more visible as G328 appeared to be important residue with decomposition energy < -5 kcal/mol

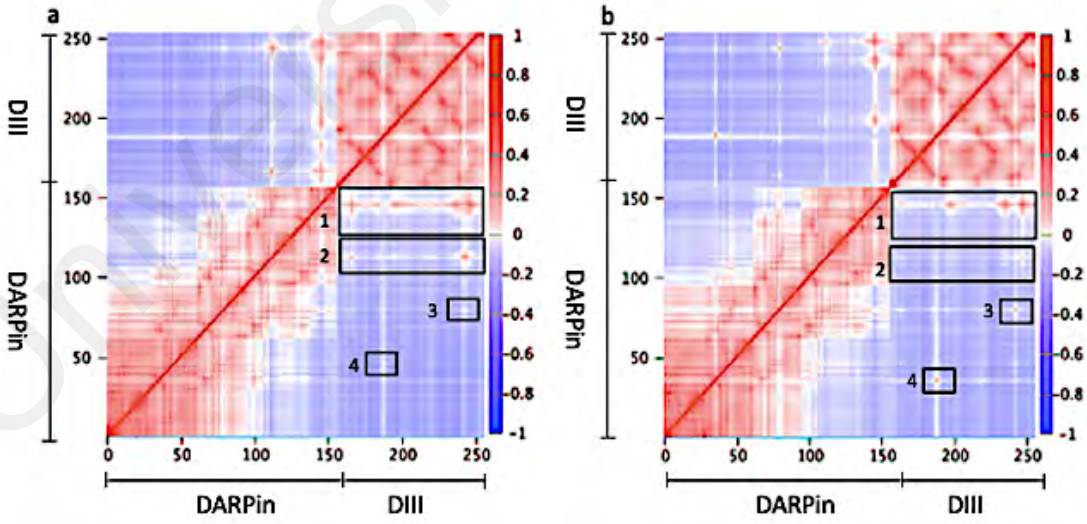


Figure 4.20 Dynamic cross-correlations map of (a) 2J-DIII, and (b) YKNM-DIII. Residue 1-156 belong to DARPin while residue 157-253 belong to DIII. Numbering of DIII follows numbering in protein crystal structure, 298-393.

CHAPTER 5: CONCLUSION AND FUTURE PROPECTS

Emerging or re-emerging of infections caused by virus remained as the deadly threat to public health globally. When specific treatments for dengue are not available, developing vaccine becomes an urgent effort to prevent dengue infection. However, shallow understanding of immune response and pathogenesis of DHF and DSS become the drag force in dengue vaccine development. DARPins are non-immunoglobin scaffold that could be taken as alternative to mAb as therapeutics for DENV due to several key advantages.

It is realistic to engineer DARPins that recognize and bind with DIII DENV-2 under computer-aided drug design method as we have seen how computational techniques were employed in re-designing 4E11 to 4E5A and later to Ab 513 that bind with DIII of all DENV serotypes with higher affinity. Two DARPins (2J and 3C) have been docked to DIII DENV-2 epitopes recognized by mAbs and their potential in mimicking mAbs to bind with DIII of DENV-2 was explored under long time scale MDs of 100 ns. 2J has a binding free energy of -42.77 kcal/mol that was closer to the binding free energy range of the Fab-DIII and scFv-DIIIs (-43.02 to -89.02 kcal/mol) and hence its affinity in binding with the antibody epitope was ~90% higher than 3C (-22.44 kcal/mol) as determined from free energy calculations using MMGBSA approach.

Through MDs and computational site-directed mutagenesis, DARPin 2J was further designed and engineered to bind with DIII DENV-2 with lower binding free energy. 23 systems of single point mutations and 9 systems of double-point mutation were simulated and their binding free energies were calculated under the same parameters. Affinity of 2J and the mutants for DIII was ranked by binding free energy obtained from MMGBSA method, YKNM (-63.90 kcal/mol) > N102Y (57.70 kcal/mol) > Y66K (56.91 kcal/mol) >N112Y (-47.82 kcal/mol) > N112M (-45.82 kcal/mol) > 2J (-42.77 kcal/mol). The best

two-mutant combination, YKNM formed strong and extensive hydrogen bonds and salt bridges with DIII like 513-DIII and improved affinity ~50%. Improved binding affinity of YKNM for DIII could be explained by the polar contributions of the top five residue pairs.

Apart from MDs, dynamics and flexibility of the proteins had been probed using GNM. Global motions of DARPin-DIII and scFv-DIII were observed correlating binding free energy calculated from MD simulations. Minima from global mode shape distinguished important residues from other residues. In addition, lower binding free energy of 2J-DIII was accompanied with a lower mode shape and larger number of residues with lower mobility. Residue or protein region with lower mobility were likely to characterize a tight binding interaction. Nevertheless, energy of a flexible region could be low while energy of a less mobile region could be high. Therefore, coupling MDs and GNM was useful in DARPin design as flexible regions that accounted a high energy could be targeted for further design so that both flexibility and energy would be lowered. However, to extract the details on dynamics and flexibility of the complex, a representative structure is essential and vital. Complex generated from molecular docking should undergo longer time scale MDs to get the reliable result of binding affinity prediction for energy and global mode shape analysis.

It is undeniable that biological study is indispensable to verify affinity of engineered DARPins for DIII and fine tune the workflow used in designing DARPin. Hence, as for the future works, plaque reduction neutralization test would be first needed to distinguish the DARPins if they have affinity for DIII as predicted and later, determination of the concentration required for inhibition has to be ascertained. Meanwhile, computational approach with machine learning and automation would continue to play a vital role to engineer DARPin to be tetravalent like mAbs that bind with DIII of all DENV serotypes.

REFERENCES

Acosta, E. G., & Bartenschlager, R. (2016). The quest for host targets to combat dengue virus infections. *Current Opinion in Virology*, *20*, 47–54.

Aier, I., Varadwaj, P. K., & Raj, U. (2016). Structural insights into conformational stability of both wild-type and mutant EZH2 receptor. *Scientific Reports*, *6*(1), 34984.

Allison, S. L., Schalich, J., Stiasny, K., Mandl, C. W., Kunz, C., & Heinz, F. X. (1995). Oligomeric rearrangement of tick-borne encephalitis virus envelope proteins induced by an acidic pH. *Journal of Virology*, *69*(2), 695–700.

Anderson, A. C. (2003). The process of structure-based drug design. *Chemistry & Biology*, *10*(9), 787–797.

Andrade, M. A., Perez-Iratxeta, C., & Ponting, C. P. (2001). Protein repeats: Structures, functions, and evolution. *Journal of Structural Biology*, *134*(2–3), 117–131.

Anishchenko, I., Kundrotas, P. J., Tuzikov, A. V, & Vakser, I. A. (2014). Protein models: The grand challenge of protein docking. *Proteins*, *82*(2), 278–287.

Bahar, I, Erman, B., Jernigan, R. L., Atilgan, A. R., & Covell, D. G. (1999). Collective motions in HIV-1 reverse transcriptase: Examination of flexibility and enzyme function. *Journal of Molecular Biology*, *285*(3), 1023–1037.

Bahar, Ivet, Atilgan, A. R., & Erman, B. (1997). Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential. In *Folding and Design* (Vol. 2, Issue 3, pp. 173–181).

Baird, R., Omlin, A., Kiemle-Kallee, J., Fiedler, U., Zitt, C., Feurstein, D., Herbst, J., Dawson, K., vom Baur, E., Stumpp, M., Hermann, F., Harstrick, A., & Schneeweiss, A. (2018). Abstract OT1-03-02: MP0274-CP101: A phase 1, first-in-human, single-arm, multi-center, open-label, dose escalation study to assess safety, tolerability, and pharmacokinetics of MP0274 in patients with advanced HER2-positive solid tumors. *Cancer Research*, *78*(4 Supplement), OT1-03–02--OT1-03–02.

Basavannacharya, C., & Vasudevan, S. G. (2014). Suramin inhibits helicase activity of NS3 protein of dengue virus in a fluorescence-based high throughput assay format. *Biochemical and Biophysical Research Communications*, *453*(3), 539–544.

Bery, N., Legg, S., Debreczeni, J., Breed, J., Embrey, K., Stubbs, C., … Rabbitts, T. H. (2019). KRAS-specific inhibition using a DARPin binding to a site in the allosteric lobe. *Nature Communications*, *10*(1), 2607.

Bhatt, S., Gething, P. W., Brady, O. J., Messina, J. P., Farlow, A. W., Moyes, C. L., … Wint, G. R. W. (2013). The global distribution and burden of dengue. In *Nature* (Vol. 496).

Binz, H. K., Amstutz, P., & Pluckthun, A. (2005). Engineering novel binding proteins from nonimmunoglobulin domains. *Nature Biotechnology*, *23*(10), 1257–1268.

Binz, H. K., Bakker, T. R., Phillips, D. J., Cornelius, A., Zitt, C., Göttler, T., … Stumpp, M. T. (2017). Design and characterization of MP0250, a tri-specific anti-HGF/anti-VEGF DARPin® drug candidate. *MAbs*, *9*(8), 1262–1269.

Binz, H. K., Stumpp, M. T., Forrer, P., Amstutz, P., & Plückthun, A. (2003). Designing repeat proteins: Well-expressed, soluble and stable proteins from combinatorial libraries of consensus ankyrin repeat proteins. *Journal of Molecular Biology*, *332*(2), 489–503.

Bogan, A A, & Thorn, K. S. (1998). Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology*, *280*(1), 1–9.

Boldescu, V., Behnam, M. A. M., Vasilakis, N., & Klein, C. D. (2017). Broad-spectrum agents for flaviviral infections: Dengue, Zika and beyond. *Nature Reviews. Drug Discovery*, *16*(8), 565–586.

Bork, P. (1993). Hundreds of ankyrin-like repeats in functionally diverse proteins: Mobile modules that cross phyla horizontally? *Proteins: Structure, Function, and Bioinformatics*, *17*(4), 363–374.

Botta, L., Rivara, M., Zuliani, V., & Radi, M. (2018). Drug repurposing approaches to fight Dengue virus infection and related diseases. *Frontiers in Bioscience (Landmark Edition)*, *23*, 997–1019.

Bukowska, M. A., & Grütter, M. G. (2013). New concepts and aids to facilitate crystallization. *Current Opinion in Structural Biology*, *23*(3), 409–416.

Byrd, C. M., Dai, D., Grosenbach, D. W., Berhanu, A., Jones, K. F., Cardwell, K. B., … Jordan, R. (2013). A novel inhibitor of dengue virus replication that targets the capsid protein. *Antimicrobial Agents and Chemotherapy*, *57*(1), 15–25.

Byrd, C. M., Grosenbach, D. W., Berhanu, A., Dai, D., Jones, K. F., Cardwell, K. B., … Jordan, R. (2013). Novel benzoxazole inhibitor of dengue virus replication that targets the NS3 helicase. *Antimicrobial Agents and Chemotherapy*, *57*(4), 1902–1912.

Campbell, B., Petukh, M., Alexov, E., & Li, C. (2014). On the electrostatic properties of homodimeric proteins. *Journal of Theoretical and Computational Chemistry*, *13*(03), 1440007.

Case, D. A., Babin, V., Berryman, J., Betz, R. M., Cai, Q., Cerutti, D. S., … others. (n.d.). AMBER 14, 2014. *University of California, San Francisco*.

Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., … Kollman, P. A. (2012). *AMBER 12 OR - University of California, San Francisco*. University of California, San Francisco. citeulike-article-id:10779586

Chakavorty, A., Li, L., & Alexov, E. (2016). Electrostatic component of binding energy: Interpreting predictions from Poisson–Boltzmann equation and modeling protocols. *Journal of Computational Chemistry*, *37*(28), 2495–2507.

Chang, J., Schul, W., Butters, T. D., Yip, A., Liu, B., Goh, A., … Block, T. M. (2011). Combination of α-glucosidase inhibitor and ribavirin for the treatment of dengue virus infection *in vitro* and *in vivo. Antiviral Research*, *89*(1), 26–34.

Chang, J., Warren, T. K., Zhao, X., Gill, T., Guo, F., Wang, L., … Block, T. M. (2013). Small molecule inhibitors of ER α-glucosidases are active against multiple hemorrhagic fever viruses. *Antiviral Research*, *98*(3), 432–440.

Chen, F., Liu, H., Sun, H., Pan, P., Li, Y., Li, D., & Hou, T. (2016). Assessing the performance of the MM/PBSA and MM/GBSA methods. 6. Capability to predict protein-protein binding free energies and re-rank binding poses generated by protein-protein docking. *Physical Chemistry Chemical Physics : PCCP*, *18*(32), 22129–22139.

Chen, H.-H., Chen, C.-C., Lin, Y.-S., Chang, P.-C., Lu, Z.-Y., Lin, C.-F., Chen, C.-L., & Chang, C.-P. (2017). AR-12 suppresses dengue virus replication by down-regulation of PI3K/AKT and GRP78. *Antiviral Research*, *142*, 158–168.

Chen, Y.-L., Abdul Ghafar, N., Karuna, R., Fu, Y., Lim, S. P., Schul, W., … Shi, P.-Y. (2014). Activation of peripheral blood mononuclear cells by dengue virus infection depotentiates balapiravir. *Journal of Virology*, *88*(3), 1740–1747.

Chew, M.-F., Poh, K.-S., & Poh, C.-L. (2017). Peptides as therapeutic agents for dengue virus. *International Journal of Medical Sciences*, *14*(13), 1342–1359.

Chiang, P.-Y., & Wu, H.-N. (2016). The role of surface basic amino acids of dengue virus NS3 helicase in viral RNA replication and enzyme activities. *FEBS Letters*, *590*(14), 2307–2320. h

Choy, M. M., Zhang, S. L., Costa, V. V, Tan, H. C., Horrevorts, S., & Ooi, E. E. (2015). Proteasome Inhibition Suppresses Dengue Virus Egress in Antibody Dependent Infection. *PLOS Neglected Tropical Diseases*, *9*(11), 1–16.

Çınaroğlu, S. S., & Timuçin, E. (2019). Comprehensive evaluation of the MM-GBSA method on bromodomain-inhibitor sets. *Briefings in Bioinformatics*, *21*(6), 2112–2125.

Clackson, T., & Wells, J. A. (1995). A hot spot of binding energy in a hormone-receptor interface. *Science (New York, N.Y.)*, *267*(5196), 383–386.

Clark, D. E. (2006). What has computer-aided molecular design ever done for drug discovery? *Expert Opinion on Drug Discovery*, *1*(2), 103–110.

Clum, S., Ebner, K., & Padmanabhan, R. (1997). Cotranslational Membrane Insertion of the Serine Proteinase Precursor NS2B-NS3 (Pro) of Dengue Virus. *Journal of Biological Chemistry*, *272*(49), 30715–30723.

Cockburn, J. J. B., Navarro Sanchez, M. E., Fretes, N., Urvoas, A., Staropoli, I., Kikuti, C. M., … Rey, F. A. (2012). Mechanism of dengue virus broad cross-neutralization by a monoclonal antibody. *Structure*, *20*(2), 303–314.

Cockburn, J. J. B., Navarro Sanchez, M. E., Goncalvez, A. P., Zaitseva, E., Stura, E. A.,

Kikuti, C. M., Duquerroy, S., … Rey, F. A. (2012). Structural insights into the neutralization mechanism of a higher primate antibody against dengue virus. *The EMBO Journal*, *31*(3), 767–779.

Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., … Kollman, P. A. (1995). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, *117*(19), 5179–5197.

Crill, W. D., & Roehrig, J. T. (2001). Monoclonal antibodies that bind to domain III of dengue virus E glycoprotein are the most efficient blockers of virus adsorption to Vero cells. *Journal of Virology*, *75*(16), 7769–7773.

De Clercq, E., & Li, G. (2016). Approved antiviral drugs over the past 50 years. *Clinical Microbiology Reviews*, *29*(3), 695 – 747.

Dehouck, Y., Kwasigroch, J. M., Rooman, M., & Gilis, D. (2013). *BeAtMuSiC : prediction of changes in protein – protein binding affinity on mutations*. *41*(Web Server issue), 333–339.

Dolinsky, T. J., Czodrowski, P., Li, H., Nielsen, J. E., Jensen, J. H., Klebe, G., & Baker, N. A. (2007). PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Research*, *35*(Web Server issue), W522–W525.

Dolinsky, T. J., Nielsen, J. E., McCammon, J. A., & Baker, N. A. (2004). PDB2PQR: An automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Research*, *32*(Web Server issue), W665-7.

Dominguez, C., Boelens, R., & Bonvin, A. M. J. J. (2003). HADDOCK: A protein−protein docking approach based on biochemical or biophysical information. *Journal of the American Chemical Society*, *125*(7), 1731–1737.

Doss, C. G. P., Rajith, B., Garwasis, N., Mathew, P. R., Raju, A. S., Apoorva, K., … Dike, I. P. (2012). Screening of mutations affecting protein stability and dynamics of FGFR1—A simulation analysis. *Applied & Translational Genomics*, *1*, 37–43.

Durrant, J. D., & McCammon, J. A. (2011). Molecular dynamics simulations and drug discovery. *BMC Biology*, *9*(1), 71.

Dussart, P., Baril, L., Petit, L., Beniguel, L., Quang, L. C., Ly, S., …Vasconcelos, P. F. da C. (2012). Clinical and virological study of dengue cases and the members of their households: The multinational DENFRAME Project. *PLoS Neglected Tropical Diseases*, *6*(1), e1482.

El Sahili, A., & Lescar, J. (2017). Dengue virus non-structural protein 5. *Viruses*, *9*(4), 91.

Erijman, A., Rosenthal, E., & Shifman, J. M. (2014). How structure defines affinity in protein-protein interactions. *PLOS ONE*, *9*(10), e110085.

Farias, K. J. S., Machado, P. R. L., Muniz, J. A. P. C., Imbeloni, A. A., & da Fonseca, B. A. L. (2015). Antiviral activity of chloroquine against dengue virus type 2 replication in Aotus monkeys. *Viral Immunology*, *28*(3), 161–169.

Feng, T., Chen, F., Kang, Y., Sun, H., Liu, H., Li, D., Zhu, F., & Hou, T. (2017). HawkRank: A new scoring function for protein-protein docking based on weighted energy terms. *Journal of Cheminformatics*, *9*(1), 66.

Fibriansah, G., & Lok, S.-M. (2016). The development of therapeutic antibodies against dengue virus. *Antiviral Research*, *128*, 7–19.

Fiser, A., Do, R. K., & Sali, A. (2000). Modeling of loops in protein structures. *Protein Science : A Publication of the Protein Society*, *9*(9), 1753–1773.

Flory, P. J., Gordon, M., & McCrum, N. G. (1976). Statistical thermodynamics of random networks [and discussion]. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *351*, 351–380.

Fried, J. R., Gibbons, R. V, Kalayanarooj, S., Thomas, S. J., Srikiatkhachorn, A., Yoon, I.-K., … Cummings, D. A. T. (2010). Serotype-specific differences in the risk of dengue hemorrhagic fever: an analysis of data collected in Bangkok, Thailand from 1994 to 2006. *PLoS Neglected Tropical Diseases*, *4*(3), e617.

Gao, M., Zhou, H., & Skolnick, J. (2015). Insights into disease-associated mutations in the human proteome through protein structural analysis. *Structure (London, England : 1993)*, *23*(7), 1362–1369.

Garcia, M. V. R., Marciela, W. N., Gomes, R. da S., & do Amaral, M. S. (2012). A tutorial for molecular dynamics simulations using Amber package. *Orbital: The Electronic Journal of Chemistry*, *4*(3), 222–234.

Gautam, V., Chong, W. L., Chin, S. P., Zain, S. M., Rahman, N. A., Vao-soongnern, V., & Lee, V. S. (2019). Loop dynamics behind the affinity of DARPins towards ERK2: Molecular dynamics simulations (MDs) and elastic network model (ENM). *Journal of Molecular Liquids*, *274*, 612–620.

Gautam, V., Nimmanpipug, P., Zain, S., & Rahman, N. A. (2021). *Molecular Dynamics Simulations in Designing DARPins as*. 1–18.

Genheden, S., & Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opinion on Drug Discovery*, *10*(5), 449–461.

Gilbreth, R. N., & Koide, S. (2012). Structural insights for engineering binding proteins based on non-antibody scaffolds. *Current Opinion in Structural Biology*, *22*(4), 413–420.

Goetz, A. W., Williamson, M. J., Xu, D., Poole, D., Grand, S. Le, & Walker, R. C. (2012). Routine microsecond molecular dynamics simulations with amber - part i: Generalized born. *Journal of Chemical Theory and Computation.*, *8*, 1542–1555.

Gonzalez-Ruiz, D., & Gohlke, H. (2006). Targeting protein-protein interactions with small molecules: challenges and perspectives for computational binding epitope detection and ligand finding. *Current Medicinal Chemistry*, *13*(22), 2607–2625.

Gromowski, G. D., Roehrig, J. T., Diamond, M. S., Lee, J. C., Pitcher, T. J., & Barrett, A. D. T. (2010). Mutations of an antibody binding energy hot spot on domain III of the dengue 2 envelope glycoprotein exploited for neutralization escape. *Virology*, *407*(2), 237–246.

Grossfield, A., Patrone, P. N., Roe, D. R., Schultz, A. J., Siderius, D. W., & Zuckerman, D. M. (2018). Best practices for quantification of uncertainty and sampling quality in molecularsimulations [Article v1.0]. *Living Journal of Computational Molecular Science*, *1*(1), 5067.

Grossfield, A., & Zuckerman, D. M. (2009). Quantifying uncertainty and sampling quality in biomolecular simulations. *Annual Reports in Computational Chemistry*, *5*, 23–48.

Gu, B., Liu, C., Lin-Goerke, J., Maley, D. R., Gutshall, L. L., Feltenberger, C. A., & Del Vecchio, A. M. (2000). The RNA helicase and nucleotide nriphosphatase Activities of thebovine viral diarrhea virus NS3 protein are essential for viral replication. *Journal of Virology*, *74*(4), 1794 LP – 1800.

Guirakhoo, F., Bolin, R. A., & Roehrig, J. T. (1992). The Murray Valley encephalitis virus prM protein confers acid resistance to virus particles and alters the expression of epitopes within the R2 domain of E glycoprotein. *Virology*, *191*(2), 921–931.

Hadinegoro, S. R., Arredondo-García, J. L., Capeding, M. R., Deseda, C., Chotpitayasunondh, T., Dietze, R., … Saville, M. (2015). Efficacy and long-term safety of a dengue vaccine inregions of endemic disease. *The New England Journal of Medicine*, *373*(13), 1195–1206.

Haliloglu, T., Bahar, I., & Erman, B. (1997). Gaussian dynamics of folded proteins. *Physical Review Letters*, *79*(16), 3090–3093.

Haliloglu, T., Seyrek, E., & Erman, B. (2008). Prediction of binding sites in receptor-ligand complexes with the Gaussian Network Model. *Phys. Rev. Lett.*, *100*(22), 228102.

Halstead, S. B., & Russell, P. K. (2016). Protective and immunological behavior of chimeric yellow fever dengue vaccine. *Vaccine*, *34*(14), 1643–1647.

Hammond, S. N., Balmaseda, A., Pérez, L., Tellez, Y., Saborío, S. I., Mercado, J. C., … Harris, E. (2005). Differences in dengue severity in infants, children, and adults in a 3-year hospital-based study in Nicaragua. *The American Journal of Tropical Medicine and Hygiene*, *73*(6), 1063–1070.

Hayes, E. B., & Gubler, D. J. (1992). Dengue and dengue hemorrhagic fever. *The Pediatric Infectious Disease Journal*, *11*(4), 311–317.

Heringa, J. (1998). Detection of internal repeats: How common are they? *Current Opinion in Structural Biology*, *8*(3), 338–345.

Hernandez-Rodriguez, M., Rosales-Hernandez, M. C., Mendieta-Wejebe, J. E., Martinez-Archundia, M., & Basurto, J. C. (2016). Current tools and methods in Molecular Dynamics (MD) simulations for drug design. *Current Medicinal Chemistry*, *23*(34), 3909-3924.

Holbrook, M. R. (2017). Historical perspectives on flavivirus research. *Viruses*, *9*(5), 97.

Hou, T., Wang, J., Li, Y., & Wang, W. (2011). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of Chemical Information and Modeling*, *51*(1), 69–82.

Hrobowski, Y. M., Garry, R. F., & Michael, S. F. (2005). Peptide inhibitors of dengue virus and West Nile virus infectivity. *Virology Journal*, *2*(1), 49.

Huang, Y.-J. S., Higgs, S., & Vanlandingham, D. L. (2019). Arbovirus-Mosquito Vector-Host interactions and the impact on transmission and diseasepathogenesis of arboviruses. *Frontiers in Microbiology*, *10*, 22.

Huber, T., Steiner, D., Röthlisberger, D., & Plückthun, A. (2007). In vitro selection and characterization of DARPins and Fab fragments for the co-crystallization of membrane proteins: The Na(+)-citrate symporter CitS as an example. *Journal of Structural Biology*, *159*(2), 206–221.

Hung, J.-J., Hsieh, M.-T., Young, M.-J., Kao, C.-L., King, C.-C., & Chang, W. (2004). An external loop region of domain III of dengue virus type 2 envelope protein is involved in serotype-specific binding to mosquito but not mammalian cells. *Journal of Virology*, *78*(1), 378–388.

Huo, S., Massova, I., & Kollman, P. A. (2002). Computational alanine scanning of the 1:1 human growth hormone-receptor complex. *Journal of Computational Chemistry*, *23*(1), 15–27.

Huy, N. T., Van Giang, T., Thuy, D. H. D., Kikuchi, M., Hien, T. T., Zamora, J., & Hirayama, K. (2013). Factors associated with dengue shock syndrome: a systematic review and meta-analysis. *PLoS Neglected Tropical Diseases*, *7*(9), e2412.

Ibarra, A. A., Bartlett, G. J., Hegedüs, Z., Dutt, S., Hobor, F., Horner, K. A., Hetherington, K., … Wilson, A. J. (2019). Predicting and experimentally validating hot-spot residues at protein–protein Interfaces. *ACS Chemical Biology*, *14*(10), 2252–2263.

John, A. M., Jittmittraphap, A., Chattanadee, S., Alwin Prem Anand, A., Shenbagarathai, R., & Leaungwutiwong, P. (2019). *In vitro* analysis of synthetic peptides in blocking the entry of dengue virus. *Virus Research*, *260*, 142–150.

Kalaivani, R., & Srinivasan, N. (2015). A Gaussian network model study suggests that structural fluctuations are higher for inactive states than active states of protein kinases. *Molecular BioSystems*, *11*(4), 1079–1095.

Karabencheva-Christova, T. G., Carlsson, U., Balali-Mood, K., Black, G. W., & Christov, C. Z. (2013). Conformational effects on the circular dichroism of Human Carbonic Anhydrase II: A multilevel computational study. *PLoS ONE*, *8*(2), e56874.

Karim, H. A. A., Rungrotmongkol, T., Zain, S. M., Rahman, N. A., Tayapiwattana, C., & Lee, V. (2019). Designed antiviral ankyrin – A computational approach to combat HIV-1 via intracellular pathway by targeting the viral capsid of HIV-1. *Journal of Molecular Liquids*, *277*, 63–69.

Karplus, M. (2002). Molecular dynamics simulations of biomolecules. *Accounts of Chemical Research*, *35*(6), 321–323. h

Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A. A., Aflalo, C., & Vakser, I. A. (1992). Molecular surface recognition: Determination of geometric fit between proteins and their ligands by correlation techniques. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(6), 2195–2199.

Kato, F., Ishida, Y., Oishi, S., Fujii, N., Watanabe, S., Vasudevan, S. G., … Hishiki, T. (2016). Novel antiviral activity of bromocriptine against dengue virus replication. *Antiviral Research*, *131*, 141–147.

Kaufmann, S. H. E., Dorhoi, A., Hotchkiss, R. S., & Bartenschlager, R. (2018). Host-directed therapies for bacterial and viral infections. *Nature Reviews Drug Discovery*, *17*(1), 35–56.

Keskin, O., Bahar, I., Flatow, D., Covell, D. G., & Jernigan, R. L. (2002). Molecular mechanisms of chaperonin GroEL-GroES function. *Biochemistry*, *41*(2), 491–501.

Keskin, Ozlem, Ma, B., & Nussinov, R. (2005). Hot regions in protein-protein interactions: the organization and contribution of structurally conserved hot spot residues. *Journal of Molecular Biology*, *345*(5), 1281–1294.

Khan, M. T., Khan, A., Rehman, A. U., Wang, Y., Akhtar, K., Malik, S. I., & Wei, D.-Q. (2019). Structural and free energy landscape of novel mutations in ribosomal protein S1 (rpsA) associated with pyrazinamide resistance. *Scientific Reports*, *9*(1), 7482.

Knobeloch, D., Schmidt, A., Scheerer, P., Krauss, N., Wessner, H., Scholz, C., … Höhne, W. (2010). A coleopteran triosephosphate isomerase: X-ray structure and phylogenetic impact of insect sequences. *Insect Molecular Biology*, *19*(1), 35–48.

Kobe, B., & Kajava, A. V. (2000). When protein folding is simplified to protein coiling: The continuum of solenoid protein structures. In *Trends in Biochemical Sciences* (Vol. 25, Issue 10, pp. 509–515).

Kramer, L., Renko, M., Završnik, J., Turk, D., Seeger, M. A., Vasiljeva, O., … Turk, B. (2017). Non-invasive *in vivo* imaging of tumour-associated cathepsin B by a highly selective inhibitory DARPin. *Theranostics*, *7*, 2806–2821.

Krzywda, S., Brzozowski, A. M., Higashitsuji, H., Fujita, J., Welchman, R., Dawson, S., … Wilkinson, A. J. (2004). The crystal structure of gankyrin, an oncoprotein found in complexes with cyclin-dependent kinase 4, a 19 S proteasomal ATPase regulator, and the tumor suppressors Rb and p53. *The Journal of Biological Chemistry*, *279*(2), 1541–1545.

Kuhn, R. J., Zhang, W., Rossmann, M. G., Pletnev, S. V, Corver, J., Lenches, E., … Strauss, J. H. (2002). Structure of dengue virus: Implications for flavivirus organization, maturation, and fusion. *Cell*, *108*(5), 717–725.

Kumar, C. V., Swetha, R. G., Anbarasu, A., & Ramaiah, S. (2014). Computational Analysis Reveals the Association of Threonine 118 Methionine Mutation in PMP22 Resulting in CMT-1A. *Advances in Bioinformatics*, *2014*, 502618. https://doi.org/10.1155/2014/502618

Kummer, L., Parizek, P., Rube, P., Millgramm, B., Prinz, A., Mittl, P. R. E., Kaufholz, M., Zimmermann, B., Herberg, F. W., & Plückthun, A. (2012). Structural and functional analysis of phosphorylation-specific binders of the kinase ERK from designed ankyrin repeat protein libraries. *Proceedings of the National Academy of Sciences*, *109*(34), E2248--E2257.

Kundu, S., Melton, J. S., Sorensen, D. C., & Phillips, G. N. (2002). Dynamics of proteins in crystals: Comparison of experiment with simple models. *Biophysical Journal*, *83*(2), 723–732.

Laureti, M., Narayanan, D., Rodriguez-Andres, J., Fazakerley, J. K., & Kedzierski, L. (2018). Flavivirus receptors: Diversity, identity, and cell entry. *Frontiers in Immunology*, *9*, 2180.

Lazim, R., Suh, D., & Choi, S. (2020). Advances in molecular dynamics dimulations and enhanced sampling methods for the study of protein systems. In *International Journal of Molecular Sciences* (Vol. 21, Issue 17).

Le Grand, S., Götz, A. W., & Walker, R. C. (2013). SPFP: Speed without compromise - A mixed precision model for GPU accelerated molecular dynamics simulations. *Computer Physics Communications*, *184*(2), 374–380.

Lee, E., Pavy, M., Young, N., Freeman, C., & Lobigs, M. (2006). Antiviral effect of the heparan sulfate mimetic, PI-88, against dengue and encephalitic flaviviruses. *Antiviral Research*, *69*(1), 31–38.

Lee, V. S., Tue-ngeun, P., Nangola, S., Kitidee, K., Jitonnom, J., Nimmanpipug, P., … Tayapiwatana, C. (2010). Pairwise decomposition of residue interaction energies of single chain Fv with HIV-1 p17 epitope variants. *Molecular Immunology*, *47*(5), 982–990.

Lensink, M. F., Velankar, S., & Wodak, S. J. (2017). Modeling protein-protein and protein-peptide complexes: CAPRI 6th edition. *Proteins*, *85*(3), 359–377.

Li, H., Chang, Y.-Y., Yang, L.-W., & Bahar, I. (2016). iGNM 2.0: the Gaussian network model database for biomolecular structural dynamics. *Nucleic Acids Research*, *44*(D1), D415–D422.

Lim, S. P. (2019). Dengue drug discovery: Progress, challenges and outlook. *Antiviral Research*, *163*, 156–178.

Lok, S.-M., Kostyuchenko, V., Nybakken, G. E., Holdaway, H. A., Battisti, A. J., Sukupolvi-Petty, S., … Rossmann, M. G. (2008). Binding of a neutralizing antibody to dengue virus alters the arrangement of surface glycoproteins. *Nat Struct Mol Biol*, *15*(3), 312–317.

Lok, S. M., Costin, J. M., Hrobowski, Y. M., Hoffmann, A. R., Rowe, D. K., Kukkaro, P., … Michael, S. F. (2012). Release of dengue virus genome induced by a peptide inhibitor. *PLoS ONE*, *7*(11).

Low, J. G., Gatsinga, R., Vasudevan, S. G., & Sampath, A. (2018). *Dengue antiviral development: A continuing journey BT - Dengue and Zika: Control and antiviral treatment strategies* (R. Hilgenfeld & S. G. Vasudevan (eds.); pp. 319–332). Springer Singapore.

Lucas, W. (2010). Viral capsids and envelopes: Structure and function. In *eLS*.

Ma, B., Wolfson, H. J., & Nussinov, R. (2001). Protein functional epitopes: Hot spots, dynamics and combinatorial libraries. *Current Opinion in Structural Biology*, *11*(3), 364–369.

Ma, L., Jones, C. T., Groesch, T. D., Kuhn, R. J., & Post, C. B. (2004). Solution structure of dengue virus capsid protein reveals another fold. *Proceedings of the National Academy of Sciences*, *101*(10), 3414–3419.

MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., … Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The Journal of Physical Chemistry B*, *102*(18), 3586–3616.

Maffucci, I., & Contini, A. (2016). Improved computation of protein–protein relative binding energies with the Nwat-MMGBSA method. *Journal of Chemical Information and Modeling*, *56*(9), 1692–1704.

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., & Simmerling, C. (2015). ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *Journal of Chemical Theory and Computation*, *11*(8), 3696–3713.

Marcotte, E. M., Pellegrini, M., Yeates, T. O., & Eisenberg, D. (1999). A census of protein repeats. *Journal of Molecular Biology*, *293*(1), 151–160. h

Marks, R. M., Lu, H., Sundaresan, R., Toida, T., Suzuki, A., Imanari, T., … Linhardt, R. J. (2001). Probing the interaction of dengue virus envelope protein with heparin: Assessment of glycosaminoglycan-derived inhibitors. *Journal of Medicinal Chemistry*, *44*(13), 2178–2187.

Mazumder, R., Hu, Z. Z., Vinayaka, C. R., Sagripanti, J. L., Frost, S. D. W., Kosakovsky Pond, S. L., & Wu, C. H. (2007). Computational analysis and identification of amino acid sites in dengue e proteins relevant to development of diagnostics and vaccines. *Virus Genes*, *35*(2), 175–186.

Megret, F., Hugnot, J. P., Falconar, A., Gentry, M. K., Morens, D. M., Murray, J. M., … Deubel, V. (1992). Use of recombinant fusion proteins and monoclonal antibodies to define linear and discontinuous antigenic sites on the dengue virus envelope glycoprotein. *Virology*, *187*(2), 480–491.

Micheletti, C., Cecconi, F., Flammini, A., & Maritan, A. (2002). Crucial stages of protein folding through a solvable model: predicting target sites for enzyme-inhibiting drugs. *Protein Science : A Publication of the Protein Society*, *11*(8), 1878–1887.

Miller, B. R., McGee, T. D., Swails, J. M., Homeyer, N., Gohlke, H., & Roitberg, A. E. (2012). MMPBSA.py: An efficient program for end-state free energy calculations. *Journal of Chemical Theory and Computation*, *8*(9), 3314–3321.

Mishra, S. K., & Jernigan, R. L. (2018). Protein dynamic communities from elastic network models align closely to the communities defined by molecular dynamics. *PloS ONE*, *13*(6), e0199225–e0199225.

Moal, I. H., Chaleil, R. A. G., & Bates, P. A. (2018). Flexible protein-protein docking with SwarmDock. *Methods in Molecular Biology (Clifton, N.J.)*, *1764*, 413–428.

Modis, Y., Ogata, S., Clements, D., & Harrison, S. C. (2003). A ligand-binding pocket in the dengue virus envelope glycoprotein. *Proc Natl Acad Sci U S A*, *100*(12), 6986–6991.

Monroe, N., Sennhauser, G., Seeger, M. A., Briand, C., & Grütter, M. G. (2011). Designed ankyrin repeat protein binders for the crystallization of AcrB : Plasticity of the dominant interface. *Journal of Structural Biology*, *174*(2), 269–281.

Moreira, I. S., Fernandes, P. A., & Ramos, M. J. (2007a). Hot spots—A review of the protein–protein interface determinant amino-acid residues. *Proteins: Structure, Function, and Bioinformatics*, *68*(4), 803–812.

Moreira, I. S., Fernandes, P. A., & Ramos, M. J. (2007b). Computational alanine scanning mutagenesis--an improved methodological approach. *Journal of Computational Chemistry*, *28*(3), 644–654.

Mosavi, L., & Cammett, T. (2004). The ankyrin repeat as molecular architecture for protein recognition. *Protein Science*, *13*, 1435–1448.

N, N., Zhu, H., Liu, J., V, K., C, G. P. D., Chakraborty, C., & Chen, L. (2015). Analysing the effect ofmutation on protein function and discovering potential inhibitors of CDK4: Molecular modelling and dynamics studies. *PLos ONE*, *10*(8), 1–21.

Nangola, S., Urvoas, A., Valerio-Lepiniec, M., Khamaikawin, W., Sakkhachornphop, S., Hong, S.-S., … Tayapiwatana, C. (2012). Antiviral activity of recombinant ankyrin targeted to the capsid domain of HIV-1 Gag polyprotein. *Retrovirology*, *9*(1), 17.

Nasar, S., Rashid, N., & Iftikhar, S. (2020). Dengue proteins with their role in pathogenesis, and strategies for developing an effective anti-dengue treatment: A review. *Journal of Medical Virology*, *92*(8), 941–955.

Nguyen, T. H. T., Nguyen, T. H. Q., Vu, T. T., Farrar, J., Hoang, T. L., Dong, T. H. T., … Simmons, C. P. (2013). Corticosteroids for dengue – Why don't they work? *PLOS Neglected Tropical Diseases*, *7*(12), 1–6.

Nisalak, A., Endy, T. P., Nimmannitya, S., Kalayanarooj, S., Thisayakorn, U., Scott, R. M., … Vaughn, D. W. (2003). Serotype-specific dengue virus circulation and dengue disease in Bangkok, Thailand from 1973 to 1999. *The American Journal of Tropical Medicine and Hygiene*, *68*(2), 191–202.

Nowak, T., Färber, P. M., Wengler, G., & Wengler, G. (1989). Analyses of the terminal sequences of West Nile virus structural proteins and of the in vitro translation of these proteins allow the proposal of a complete scheme of the proteolytic cleavages involved in their synthesis. *Virology*, *169*(2), 365–376.

Oliver, G. F., Carr, J. M., & Smith, J. R. (2019). Emerging infectious uveitis: Chikungunya, dengue, Zika and Ebola: A review. *Clinical & Experimental Ophthalmology*, *47*(3), 372–380.

Olsen, D. B., Eldrup, A. B., Bartholomew, L., Bhat, B., Bosserman, M. R., Ceccacci, A., … Carroll, S. S. (2004). A 7-deaza-adenosine analog is a potent and selective inhibitor of hepatitis C virus replication with excellent pharmacokinetic properties. *Antimicrobial Agents and Chemotherapy*, *48*(10), 3944–3953.

Onufriev, A., Bashford, D., & Case, D. A. (2004). Exploring protein native states and large-scale conformational changes with a modified Generalized Born model. *Proteins: Structure, Function and Genetics*, *55*(2), 383–394.

Pambudi, S., Kawashita, N., Phanthanawiboon, S., Omokoko, M. D., Masrinoul, P., Yamashita, A., … Kurosu, T. (2013). A small compound targeting the interaction between nonstructural proteins 2B and 3 inhibits dengue virus replication. *Biochemical and Biophysical Research Communications*, *440*(3), 393–398.

Perry, S. T., Buck, M. D., Plummer, E. M., Penmasta, R. A., Batra, H., Stavale, E. J., Shresta, S. (2013). An iminosugar with potent inhibition of dengue virus infection in vivo. *Antiviral Research*, *98*(1), 35–43.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera-A visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612.

Pierce, B. G., Wiehe, K., Hwang, H., Kim, B.-H., Vreven, T., & Weng, Z. (2014). ZDOCK server: Interactive docking prediction of protein-protein complexes and symmetric multimers. *Bioinformatics (Oxford, England)*, *30*(12), 1771–1773.

Plückthun, A. (2015). Designed Ankyrin Repeat Proteins (DARPins): Binding proteins for research, diagnostics, and therapy. *Annual Review of Pharmacology and Toxicology*, *55*(1), 489–511.

Praditwongwan, W., Chuankhayan, P., Saoin, S., Wisitponchai, T., Lee, V. S., Nangola, S., … Tayapiwatana, C. (2014). Crystal structure of an antiviral ankyrin targeting the HIV-1 capsid and molecular modeling of the ankyrin-capsid complex. *Journal of Computer-Aided Molecular Design*, *28*(8), 869–884.

Prusis, P., Junaid, M., Petrovska, R., Yahorava, S., Yahorau, A., Katzenmeier, G., … Wikberg, J. E. S. (2013). Design and evaluation of substrate-based octapeptide and non substrate-based tetrapeptide inhibitors of dengue virus NS2B-NS3 proteases. *Biochemical and Biophysical Research Communications*, *434*(4), 767–772.

Qiu, L., Yan, Y., Sun, Z., Song, J., & Zhang, J. Z. H. (2018). Interaction entropy for computational alanine scanning in protein–protein binding. *WIREs Computational Molecular Science*, *8*(2), e1342.

Rajamani, D., Thiel, S., Vajda, S., & Camacho, C. J. (2004). Anchor residues in protein-protein interactions. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(31), 11287–11292.

Randolph, V. B., Winkler, G., & Stollar, V. (1990). Acidotropic amines inhibit proteolytic processing of flavivirus prM protein. *Virology*, *174*(2), 450–458.

Reichert, J. M. (2001). Monoclonal antibodies in the clinic. *Nature Biotechnology*, *19*(9), 819–822.

Rey, F. A., Heinz, F. X., Mandl, C., Kunz, C., & Harrison, S. C. (1995). The envelope glycoprotein from tick-borne encephalitis virus at 2 A resolution. In *Nature* (Vol. 375, Issue 6529, pp. 291–298).

Robinson, L. N., Tharakaraman, K., Rowley, K. J., Costa, V. V., Chan, K. R., Wong, Y. H., … Sasisekharan, R. (2015). Structure guided design of an anti-dengue antibody directed to a non-immunodominant epitope. *Cell*, *162*(3), 493–504.

Rodon, J., Omlin, A., Herbschleb, K. H., Garcia-Corbacho, J., Steiner, J., Dolado, I., … Baird, R. D. (2015). Abstract B25: First-in-human Phase I study to evaluate MP0250, a DARPin blocking HGF and VEGF, in patients with advanced solid tumors. *Molecular Cancer Therapeutics*, *14*(12 Supplement 2), B25--B25.

Rodrigues, C. H. M., Myung, Y., Pires, D. E. V, & Ascher, D. B. (2019). mCSM-PPI2: predicting the effects of mutations on protein–protein interactions. *Nucleic Acids Research*, *47*(W1), W338–W344.

Rodrigues, J P G L M., Melquiond, A. S. J., Karaca, E., Trellet, M., van Dijk, M., van Zundert, G. C. P., … Bonvin, A. M. J. J. (2013). Defining the limits of homology modeling in information-driven protein docking. *Proteins*, *81*(12), 2119–2128.

Rodrigues, João P G L M, & Bonvin, A. M. J. J. (2014). Integrative computational modeling of protein interactions. *The FEBS Journal*, *281*(8), 1988–2003.

Roe, D. R., & Cheatham III, T. E. (2013). PTRAJ and CPPTRAJ: Software for processing and analysis of molecular synamics trajectory data. *Journal Chemical Theory and Computation*, *9*(7), 3084–3095.

Roehrig, J. T. (2003). Antigenic structure of flavivirus proteins. *Advances in Virus Research*, *59*, 141–175.

Rothwell, C., Lebreton, A., Young Ng, C., Lim, J. Y. H., Liu, W., Vasudevan, S., … Gaither, L. A. (2009). Cholesterol biosynthesis modulation regulates dengue viral replication. *Virology*, *389*(1–2), 8–19.

Rutenber, E. E., & Stroud, R. M. (1996). Binding of the anticancer drug ZD1694 to E. coli thymidylate synthase: Assessing specificity and affinity. *Structure (London, England : 1993)*, *4*(11), 1317–1324.

Sahni, N., Yi, S., Taipale, M., Fuxman Bass, J. I., Coulombe-Huntington, J., Yang, F., … Vidal, M. (2015). Widespread macromolecular interaction perturbations in human genetic disorders. *Cell*, *161*(3), 647–660.

Sali, A., & Blundell, T. L. (1993). Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology*, *234*(3), 779–815.

Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S., & Walker, R. C. (2013). Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh ewald. *Journal of Chemical Theory and Computation*, *9*(9), 3878–3888.

Schmidt, A. G., Yang, P. L., & Harrison, S. C. (2010). Peptide inhibitors of flavivirus entry derived from the E protein stem. *Journal of Virology*, *84*(24), 12549–12554.

Schrama, D., Reisfeld, R. a, & Becker, J. C. (2006). Antibody targeted drugs as cancer therapeutics. *Nature Reviews. Drug Discovery*, *5*(2), 147–159.

Schul, W., Liu, W., Xu, H.-Y., Flamand, M., & Vasudevan, S. G. (2007). A dengue fever viremia model in mice shows reduction in viral replication and suppression of the inflammatory response after treatment with antiviral drugs. *The Journal of Infectious Diseases*, *195*(5), 665–674.

Seeger, M. A., Zbinden, R., Flütsch, A., Gutte, P. G. M., Engeler, S., Roschitzki-Voser, H., & Grütter, M. G. (2013). Design, construction, and characterization of a second-generation DARP in library with reduced hydrophobicity. *Protein Science : A Publication of the Protein Society*, *22*(9), 1239–1257.

Sennhauser, G., Amstutz, P., Briand, C., Storchenegger, O., Gru, M. G., & Screening, I. (2007). *Drug Export Pathway of Multidrug Exporter AcrB Revealed by DARPin Inhibitors*. *5*(1), 1–8.

Shapovalov, M. V, & Dunbrack, R. L. J. (2011). A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure (London, England : 1993)*, *19*(6), 844–858.

Shekhar, K. C., & Huat, O. L. (1992). Epidemiology of dengue/dengue hemorrhagic fever in Malaysia-A retrospective epidemiological study. 1973-1987. Part II: Dengue fever (DF). *Asia Pacific Journal of Public Health*, *6*(3), 126–133.

Shen, M.-Y., & Sali, A. (2006). Statistical potential for assessment and prediction of protein structures. *Protein Science : A Publication of the Protein Society*, *15*(11), 2507–2524.

Shilova, O. N., & Deyev, S. M. (2019). DARPins: Promising scaffolds for theranostics. *Acta Naturae*, *11*(4), 42–53.

Simanjuntak, Y., Liang, J.-J., Lee, Y.-L., & Lin, Y.-L. (2015). Repurposing of prochlorperazine for use against dengue virus infection. *The Journal of Infectious Diseases*, *211*(3), 394–404.

Simeon, R., & Chen, Z. (2018). *In vitro*-engineered non-antibody protein therapeutics. *Protein & Cell*, *9*(1), 3–14.

Simmerling, C., Strockbine, B., & Roitberg, A. E. (2002). All-atom structure prediction and folding simulations of a stable protein. *Journal of the American Chemical Society*, *124*(38), 11258–11259.

Singh, W., Karabencheva-Christova, T. G., Black, G. W., Sparagano, O., & Christov, C. Z. (2016). Conformational flexibility influences structure–function relationships in tyrosyl protein sulfotransferase-2. *RSC Adv.*, *6*(14), 11344–11352.

Songprakhon, P., Thaingtamtanha, T., Limjindaporn, T., Puttikhunt, C., Srisawat, C., Luangaram, P., … Noisakran, S. (2020). Peptides targeting dengue viral nonstructural protein 1 inhibit dengue virus production. *Scientific Reports*, *10*(1), 12933.

St. John, A. L. (2013). Influence of mast cells on dengue protective immunity and immune pathology. *PLOS Pathogens*, *9*(12), 1–4.

Stahla-Beek, H. J., April, D. G., Saeedi, B. J., Hannah, A. M., Keenan, S. M., & Geiss, B. J. (2012). Identification of a novel antiviral inhibitor of the flavivirus guanylyltransferase enzyme. *Journal of Virology*, *86*(16), 8730–8739.

Steiner, D., Forrer, P., & Plückthun, A. (2008). Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display. *Journal of Molecular Biology*, *382*(5), 1211–1227.

Stumpp, M. T., Binz, H. K., & Amstutz, P. (2008). DARPins: A new generation of protein therapeutics. In *Drug Discovery Today* (Vol. 13, Issues 15–16, pp. 695–701).

Sukupolvi-Petty, S., Austin, S. K., Purtha, W. E., Oliphant, T., Nybakken, G. E., Schlesinger, J. J., … Diamond, M. S. (2007). Type- and subcomplex-specific neutralizing antibodies against Domain III of dengue virus type 2 envelope protein recognize adjacent epitopes. *Journal of Virology*, *81*(23), 12816–12826.

Sun, H., Duan, L., Chen, F., Liu, H., Wang, Z., Pan, P., Zhu, F., Zhang, J. Z. H., & Hou, T. (2018). Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches. *Phys. Chem. Chem. Phys.*, *20*(21), 14450–14460.

Suppiah, J., Ching, S.-M., Amin-Nordin, S., Mat-Nor, L.-A., Ahmad-Najimudin, N.-A., Low, G. K.-K., Abdul-Wahid, M.-Z., Thayan, R., & Chee, H.-Y. (2018). Clinical manifestations of dengue in relation to dengue serotype and genotype in Malaysia: A retrospective observational study. *PLOS Neglected Tropical Diseases*, *12*(9), 1–20.

Takemura, K., Guo, H., Sakuraba, S., Matubayasi, N., & Kitao, A. (2012). Evaluation of protein-protein docking model structures using all-atom molecular dynamics simulations combined with the solution theory in the energy representation. *The Journal of Chemical Physics*, *137*(21), 215105.

Tam, D. T. H., Ngoc, T. V., Tien, N. T. H., Kieu, N. T. T., Thuy, T. T. T., Thanh, L. T. C., … Wills, B. A. (2012). Effects of short-course oral corticosteroid therapy in early dengue infection in Vietnamese patients: a randomized, placebo-controlled trial. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *55*(9), 1216–1224. h

Tamaskovic, R., Simon, M., Stefan, N., Schwill, M., & Plückthun, A. (2012). Chapter five - Designed Ankyrin Repeat Proteins (DARPins): From research to therapy. In K. D. Wittrup & G. L. B. T.-M. in E. Verdine (Eds.), *Protein Engineering for Therapeutics, Part B* (Vol. 503, pp. 101–134). Academic Press.

Tharakaraman, K., Robinson, L. N., Hatas, A., Chen, Y.-L., Siyue, L., Raguram, S., … Sasisekharan, R. (2013). Redesign of a cross-reactive antibody to dengue virus with broad-spectrum activity and increased in vivo potency. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(17), E1555-64.

Theurillat, J.-P., Dreier, B., Nagy-Davidescu, G., Seifert, B., Behnke, S., Zurrer-Hardi, U., Ingold, F., Pluckthun, A., & Moch, H. (2010). Designed ankyrin repeat proteins: a novel tool for testing epidermal growth factor receptor 2 expression in breast cancer. *Modern Pathology : An Official Journal of the United States and Canadian Academy of Pathology, Inc*, *23*(9), 1289–1297.

Thomas, L., Verlaeten, O., Cabié, A., Kaidomar, S., Moravie, V., Martial, J., … Césaire, R. (2008). Influence of the dengue serotype, previous dengue infection, and plasma viral load on clinical presentation and outcome during a dengue-2 and dengue-4 co-epidemic. *The American Journal of Tropical Medicine and Hygiene*, *78*(6), 990–998.

Thullier, P., Lafaye, P., Megret, F., Deubel, V., Jouan, A., & Mazie, J. C. (1999). A recombinant Fab neutralizes dengue virus in vitro. *Journal of Biotechnology*, *69*(2–3), 183–190.

Torchala, M., Moal, I. H., Chaleil, R. A. G., Fernandez-Recio, J., & Bates, P. A. (2013). SwarmDock: a server for flexible protein-protein docking. *Bioinformatics (Oxford, England)*, *29*(6), 807–809.

Tricou, V., Minh, N. N., Van, T. P., Lee, S. J., Farrar, J., Wills, B., Tran, H. T., & Simmons, C. P. (2010). A randomized controlled trial of chloroquine for the treatment of dengue in Vietnamese adults. *PLOS Neglected Tropical Diseases*, *4*(8), 1–12.

Troost, B., & Smit, J. M. (2020). Recent advances in antiviral drug development towards dengue virus. *Current Opinion in Virology*, *43*, 9–21.

Tseng, T. T., Gratwick, K. S., Kollman, J., Park, D., Nies, D. H., Goffeau, A., & Saier, M. H. J. (1999). The RND permease superfamily: An ancient, ubiquitous and diverse family that includes human disease and development proteins. *Journal of Molecular Microbiology and Biotechnology*, *1*(1), 107–125.

Tuiskunen Bäck, A., & Lundkvist, Å. (2013). Dengue viruses – An overview. *Infection Ecology & Epidemiology*, *3*(1), 19839.

van Dijk, A. D. J., & Bonvin, A. M. J. J. (2006). Solvated docking: introducing water into the modelling of biomolecular complexes. *Bioinformatics*, *22*(19), 2340–2347.

Vangone, A., Rodrigues, J. P. G. L. M., Xue, L. C., van Zundert, G. C. P., Geng, C., Kurkcuoglu, Z., … Bonvin, A. M. J. J. (2017). Sense and simplicity in HADDOCK scoring: Lessons from CASP-CAPRI round 1. *Proteins: Structure, Function, and Bioinformatics*, *85*(3), 417–423.

Vannice, K. S., Durbin, A., & Hombach, J. (2016). Status of vaccine research and development of vaccines for dengue. *Vaccine*, *34*(26), 2934–2938.

Vaughn, D. W., Green, S., Kalayanarooj, S., Innis, B. L., Nimmannitya, S., Suntayakorn, S., … Nisalak, A. (2000). Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity. *The Journal of Infectious Diseases*, *181*(1), 2–9.

Vazquez-Lombardi, R., Phan, T. G., Zimmermann, C., Lowe, D., Jermutus, L., & Christ, D. (2015). Challenges and opportunities for non-antibody scaffold drugs. *Drug Discovery Today*, *20*(10), 1271–1283.

Vicente, C. R., Herbinger, K.-H., Fröschl, G., Malta Romano, C., de Souza Areias Cabidelle, A., & Cerutti Junior, C. (2016). Serotype influences on dengue severity: a cross-sectional study on 485 confirmed dengue cases in Vitória, Brazil. *BMC Infectious Diseases*, *16*, 320.

Villar, L., Dayan, G. H., Arredondo-Garcia, J. L., Rivera, D. M., Cunha, R., Deseda, C., … Noriega, F. (2015). Efficacy of a tetravalent dengue vaccine in children in Latin America. *The New England Journal of Medicine*, *372*(2), 113–123.

Vreven, T., Moal, I. H., Vangone, A., Pierce, B. G., Kastritis, P. L., Torchala, M., … Weng, Z. (2015). Updates to the integrated protein-protein interaction benchmarks: Docking benchmark version 5 and affinity benchmark version 2. *Journal of Molecular Biology*, *427*(19), 3031–3041.

Vries, S. J. De, Dijk, M. Van, & Bonvin, A. M. J. J. (2010). The HADDOCK web server for data-driven biomolecular docking. *Nature Protocols*, *5*(5), 883–897.

Walser, M., Rothenberger, S., Hurdiss, D. L., Schlegel, A., Calabro, V., Fontaine, S., … Amstutz, P. (2020). Highly potent anti-SARS-CoV-2 multi-DARPin therapeutic candidates. *BioRxiv*. Retrieved from https://doi.org/10.1101/2020.08.25.256339

Waltz, E. (2007). Supreme Court boosts licensees in biotech patent battles. *Nature Biotechnology*, *25*(3), 264–265.

Wang, Q.-Y., Dong, H., Zou, B., Karuna, R., Wan, K. F., Zou, J., … Shi, P.-Y. (2015). Discovery of dengue virus NS4B inhibitors. *Journal of Virology*, *89*(16), 8233–8244.

Wang, Q.-Y., Patel, S. J., Vangrevelinghe, E., Xu, H. Y., Rao, R., Jaber, D., … Vasudevan, S. G. (2009). A small-molecule dengue virus entry inhibitor. *Antimicrobial Agents and Chemotherapy*, *53*(5), 1823–1831.

Wang, Y., Rader, A. J., Bahar, I., & Jernigan, R. L. (2004). Global ribosome motions revealed with elastic network model. *Journal of Structural Biology*, *147*(3), 302–314.

Warfield, K. L., Plummer, E., Alonzi, D. S., Wolfe, G. W., Sampath, A., Nguyen, T., … Ramstedt, U. (2015). A novel iminosugar UV-12 with activity against the diverse viruses influenza and dengue. *Viruses*, *7*(5), 2404–2427.

Warfield, K. L., Plummer, E. M., Sayce, A. C., Alonzi, D. S., Tang, W., Tyrrell, B. E., … Shresta, S. (2016). Inhibition of endoplasmic reticulum glucosidases is required for in vitro and in vivo dengue antiviral activity by the iminosugar UV-4. *Antiviral Research*, *129*, 93–98.

Welsch, S., Miller, S., Romero-Brey, I., Merz, A., Bleck, C. K. E., Walther, P., … Bartenschlager, R. (2009). Composition and three-dimensional architecture of the dengue virus replication and assembly sites. *Cell Host & Microbe*, *5*(4), 365–375.

Weng, G., Wang, E., Chen, F., Sun, H., Wang, Z., & Hou, T. (2019). Assessing the performance of MM/PBSA and MM/GBSA methods. 9. Prediction reliability of binding affinities and binding poses for protein-peptide complexes. *Physical Chemistry Chemical Physics : PCCP*, *21*(19), 10135–10145. h

Whitehorn, J., Nguyen, C. V. V., Khanh, L. P., Kien, D. T. H., Quyen, N. T. H., Tran, N. T. T., … Wills, B. (2016). Lovastatin for the treatment of adult patients with dengue: A randomized, double-blind, placebo-controlled trial. *Clinical Infectious Diseases : An Official Publication of the Infectious Diseases Society of America*, *62*(4), 468–476.

Winkler, J., Martin-Killias, P., Plückthun, A., & Zangemeister-Wittke, U. (2009). EpCAM-targeted delivery of nanocomplexed siRNA to tumor cells with designed ankyrin repeat proteins. *Molecular Cancer Therapeutics*, *8*(9), 2674–2683.

Wlodawer, A., & Vondrasek, J. (1998). Inhibitors of HIV-1 protease: A major success of structure-assisted drug design. *Annual Review of Biophysics and Biomolecular Structure*, *27*, 249–284.

Wu, J., Bera, A. K., Kuhn, R. J., & Smith, J. L. (2005). Structure of the flavivirus helicase: Implications for catalytic activity, protein interactions, and proteolytic processing. *Journal of Virology*, *79*(16), 10268–10277.

Wu, S.-F., Lee, C.-J., Liao, C.-L., Dwek, R. A., Zitzmann, N., & Lin, Y.-L. (2002). Antiviral effects of an iminosugar derivative on flavivirus infections. *Journal of Virology*, *76*(8), 3596–3604.

Xie, X., Wang, Q.-Y., Xu, H. Y., Qing, M., Kramer, L., Yuan, Z., & Shi, P.-Y. (2011). Inhibition of dengue virus by targeting viral NS4B protein. *Journal of Virology*, *85*(21), 11183–11195.

Yang, C.-C., Hu, H.-S., Lin, H.-M., Wu, P.-S., Wu, R.-H., Tian, J.-N., … Yueh, A. (2019). A novel flavivirus entry inhibitor, BP34610, discovered through high-throughput screening with dengue reporter viruses. *Antiviral Research*, *172*, 104636.

Yang, L.-W., Eyal, E., Chennubhotla, C., Jee, J., Gronenborn, A. M., & Bahar, I. (2007). Insights into equilibrium dynamics of proteins from comparison of NMR and X-Ray data with computational predictions. *Structure*, *15*(6), 741–749.

Yang, L. W., & Bahar, I. (2005). Coupling between catalytic site and collective dynamics: A requirement for mechanochemical activity of enzymes. *Structure*, *13*(6), 893–904.

Yeo, K. L., Chen, Y.-L., Xu, H. Y., Dong, H., Wang, Q.-Y., Yokokawa, F., & Shi, P.-Y. (2015). Synergistic suppression of dengue virus replication using a combination of nucleoside analogs and nucleoside synthesis inhibitors. *Antimicrobial Agents and Chemotherapy*, *59*(4), 2086 – 2093.

Yin, Z., Chen, Y.-L., Schul, W., Wang, Q.-Y., Gu, F., Duraiswamy, J., … Shi, P.-Y. (2009). An adenosine nucleoside inhibitor of dengue virus. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(48), 20435–20439.

Yin, Z., Patel, S. J., Wang, W.-L., Wang, G., Chan, W.-L., Rao, K. R. R., … Keller, T. H. (2006). Peptide inhibitors of dengue virus NS3 protease. Part 1: Warhead. *Bioorganic & Medicinal Chemistry Letters*, *16*(1), 36–39.

Yogurtcu, O. N., Erdemli, S. B., Nussinov, R., Turkay, M., & Keskin, O. (2008). Restricted mobility of conserved residues in protein-protein interfaces in molecular simulations. *Biophysical Journal*, *94*(9), 3475–3485.

Yu, J.-S., Wu, Y.-H., Tseng, C.-K., Lin, C.-K., Hsu, Y.-C., Chen, Y.-H., & Lee, J.-C. (2017). Schisandrin A inhibits dengue viral replication via upregulating antiviral interferon responses through STAT signaling pathway. *Scientific Reports*, *7*, 45171.

Yuan, Z., Bailey, T. L., & Teasdale, R. D. (2005). Prediction of protein B-factor profiles. *Proteins: Structure, Function and Genetics*, *58*(4), 905–912.

Zahnd, C., Wyler, E., Schwenk, J. M., Steiner, D., Lawrence, M. C., McKern, N. M., P… Plückthun, A. (2007). A Designed Ankyrin Repeat Protein evolved to picomolar affinity to Her2. *Journal of Molecular Biology*, *369*(4), 1015–1028.

Zhang, J. (2015). The hybrid idea of (energy minimization) optimization methods applied to study prion protein structures focusing on the beta2-alpha2 loop. *Biochem Pharmacol (Los Angel)*, *4*(175), 501–2167.

Zhang, Z., Witham, S., & Alexov, E. (2011). On the role of electrostatics in protein-protein interactions. *Physical Biology*, *8*(3), 35001.

Zhao, H., & Caflisch, A. (2015). Molecular dynamics in drug design. *European Journal of Medicinal Chemistry*, *91*, 4–14.

Zhao, Y., Zeng, C., & Massiah, M. A. (2015). Molecular dynamics simulation reveals insights into the mechanism of unfolding by the A130T/V mutations within the MID1 zinc-binding Bbox1 Domain. *PLOS ONE*, *10*(4), e0124377.

Zhong, H., Huang, W., He, G., Peng, C., Wu, F., & Ouyang, L. (2013). Molecular dynamics simulation of tryptophan hydroxylase-1: binding modes and free energy analysis to phenylalanine derivative inhibitors. *International Journal of Molecular Sciences*, *14*(5), 9947–9962.

Zuckerman, D. M. (2011). Equilibrium sampling in biomolecular simulations. *Annual Review of Biophysics*, *40*(1), 41–62.

Zundert, G. C. P. Van, Rodrigues, J. P. G. L. M., Trellet, M., & Schmitz, C. (2016). The HADDOCK 2.2 web server : User-friendly integrative modeling of biomolecular complexes. *Journal of Molecular Biology*, *428*(4), 720–725.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

List of publications

1. Lee, V. S., **Chong, W. L**., Sukumaran, S. D., Nimmanpipug, P., Letchumanan, V., Goh, B. H., ... & Abd Rahman, N. (2020). Computational screening and identifying binding interaction of anti-viral and anti-malarial drugs: Toward the potential cure for SARS-CoV-2. *Progress in Drug Discovery & Biomedical Science, 3(*1), a0000065.

2. Isa, D. M., Chin, S. P., **Chong, W. L**., Zain, S. M., Abd Rahman, N., & Lee, V. S. (2019). Dynamics and binding interactions of peptide inhibitors of dengue virus entry. *Journal of Biological Physics*, *45*(1), 63-76.

3. Gautam, V., **Chong, W.** L., Chin, S. P., Zain, S. M., Rahman, N. A., Vao-soongnern, V., & Lee, V. S. (2019). Loop dynamics behind the affinity of DARPins towards ERK2: Molecular dynamics simulations (MDs) and elastic network model (ENM). *Journal of Molecular Liquids, 274*, 612-620.

4. Tan, C. W., Sam, I. C., **Chong, W. L.,** Lee, V. S., & Chan, Y. F. (2017). Polysulfonate suramin inhibits Zika virus infection. *Antiviral Research*, *143*, 186-194.

5. Hematpoor, A., Liew, S. Y., **Chong, W. L**., Azirun, M. S., Lee, V. S., & Awang, K. (2016). Inhibition and larvicidal activity of phenylpropanoids from Piper sarmentosum on acetylcholinesterase against mosquito vectors and their binding mode of interaction. *PLoS ONE, 11*(5), e0155265.

6. Tee, J. T., Chee, C. F., Buckle, M. J., Lee, V. S., **Chong, W. L**., Khaledi, H., & Abd Rahman, N. (2015). Model studies on construction of the oxabicyclic [3.3. 1] core of the mulberry Diels–Alder adducts morusalbanol A and 441772-64-1. *Tetrahedron letters, 56*(36), 5082-5085.

Papers Presented

1. **Chong, W. L**., Zain, S. M., & Lee, V. S. (2017). Binding of designed ankyrins to dengue envelope protein from molecular dynamics simulations and Gaussian Network Model. 9th Conference of the Asian Consortium on Computational Materials Science (ACCMS-9), 8-11 August, Kuala Lumpur, Malaysia.

2. **Chong, W. L.**, Chin, S. P., Gautam, V., Zain, S. M., Vao-soongnern, V., & Lee, V. S. (2016). Gaussian Network Model (GNM) for binding affinity of DARPins towards membrane protein. Paper presented at 20th Annual Symposium on Computational Science and Engineering (ANSCSE20), 27-29 July, Kasetsart University, Bangkok, Thailand.

3. **Chong, W. L**., Zain, S. M., & Lee, V. S. (2015). Binding of a designed ankyrin to E protein of dengue virus type 2 from molecular dynamics simulations. Paper presented at 8th Conference of the Asian Consortium on Computational Materials Science (ACCMS-8), 16-18 June, Taipei, Taiwan.

4. **Chong, W. L**., Zain, S. M., Rahman, N. A., Othman, R., Othman, S. B., Nimmanpipug, P., & Lee, V. S. (2015). Exploration of residue binding energy of potential ankyrin for Dengue virus II from MD simulations. In *Proceedings of the 3rd International Conference on Computation for Science and Technology* (pp. 100-103). Paris: Atlantis Press.