CHAPTER 4

DISCUSSION

Human Immunodeficiency Virus (HIV) is a retrovirus that caused the acquired immunodeficiency syndrome (AIDS). A patient will continuously progress a failure of immune system leading to the increasing susceptibility to opportunistic infections and cancers. The average survival time post HIV infection is estimated to be around 10 years without treatment, depending on the HIV subtype. According to World Health Organization (WHO) and UNAIDS (www.unaids.org), there were more than 35 million people living with HIV/AIDS worldwide in 2013. HIV is in the top rank of infectious disease which causes approximately 39 million people to expire since the first case reported in 1981. The introduction of highly active antiretroviral drugs therapy (HAART) has aided both patients and healthcare personnel to manage virus level, restore immunological response and enhance quality of life. With the goal to make ART globally accessible, more HIV infected patients are being treated every year. However, the excessive use of HAART has led to the more severe obstacle, multidrug resistance and HIV mutation. Resistance-associated mutations (RAMs) is the ART-treated patients who are incapable to suppress viral replication. Moreover, these RAMs are prone to transmit and raise the prevalence of multidrug resistance HIV strain to other untreated or uninfected individuals as seen in many parts of the world [49-53].

HIV entry is an attractive target for inhibitor development since it is the first key step to HIV life cycle. These inhibitors are divided into three major groups which are attachment and CD4 binding inhibitor, coreceptor (CXCR4 or CCR5) binding inhibitor and membrane fusion inhibitor. To date, only 2 entry inhibitors were approved by FDA. The first one is a

fusion inhibitor Enfuvirtide, approved in 2003 (Fuzeon; T-20). It is a mimetic peptide of HR-C domain which binds to the central coiled coils of HR-N. This binding leads to the inhibition of 6HB formation [54, 55]. Despite to its inhibition efficiency, the drug administration is only limited to subcutaneous injection. Once the lyophilized drug has been reconstituted with supplied sterile water, it must be injected immediately or within 24 hours when refrigerated. The indicated dose for adult patients are 90 mg twice daily. To overcome this limitation, the orally bioavailable form of Enfuvirtide should be developed and still keep its inhibition efficiency [21]. Another FDA approved HIV entry inhibitor is Maraviroc (Selzentry in the U.S., Celsentri in other countries; MVC). It is a CCR5 antagonist which have been approved since 2007. The drug binds to the hydrophobic pocket within CCR5 transmembrane domain which leads to prevention of HIV Env binding [56-58]. However, the use of Maraviroc is restricted to only HIV strain which use CCR5 as coreceptor. Other targets for HIV entry inhibitor development, such as CD4, CXCR4, gp120 and gp41 are under process in preclinical and various stages of clinical trials.

Despite these challenges, there have been successes and promising signs as more alternative HIV therapy have been studied and developed. One strategy is a protein-based therapy. Ankyrin repeat protein is an outstanding scaffold as it represents an alternative binding property to antibodies. The development of this protein structure has introduced a new chapter of this protein which is called Designed Ankyrin Repeat Protein (DARPin). This repeat protein is disulfide bond-free scaffold which can be produced in large quantity with high level expression and outstanding thermostability [4, 27].

A novel HIV entry inhibitor based on DARPin technology, CD4-specific DARPin had been raised from DARPin library using ribosome display technique. Selected DARPin candidates bound specifically to D1 of CD4 molecules with K_D in low nanomolar range and successfully competed with gp120 of many HIV-1 isolates. Among the selected DARPin from the study of Schweizer *et al.* [5], there were some DARPin candidates which had cross species reactivity by recognizing rhesus macaque's CD4 molecules and effectively inhibit SIV from infecting cells from donor rhesus macaque.

Furthermore Pugach *et al.*, had chosen CD4-specific DARPin 57.2 to be studied in SHIV-infected rhesus macaque. The intravenously injected DARPin 57.2 could be detected on CD4 molecules of T-cells in peripheral blood as well as in lymph node as early as 30 min after injection. Nonetheless, the signal from the binding of DARPin 57.2 to CD4 molecules were decreasing with time and almost undetectable at 24 h post injection. In plasma free DARPin was also detectable but were rapidly cleared from circulation [6].

In this study CD4-specific DARPin 23.2 was chosen due to its effective ability in binding to CD4 and inhibition of HIV infection. The purpose of this study is aiming to elucidate the key amino acids in the binding of CD4 molecule and CD4-specific DARPin 23.2 by combining the computational data and mutational analysis. Plasmid carrying along DARPin 23.2 had been constructed earlier. Based on the sequence that Schweizer *et al.* reported on the database (AM997265.1) [5], 6×His tag at N-terminus was utilized for purification and detection of this DARPin. Production of CD4-specific DARPin 23.2 was performed in bacterial culturing system. Purified DARPin 23.2 produced in this study bound specifically to CD4 molecules on both cell line (SupT1) and primary T cells from PBMC as shown in Figure 3.3 and 3.4 respectively. Although the binding affinity of CD4-specific DARPin 23.2 to CD4 molecules on primary T cells is comparable to the anti-CD4 mAb (MT4), low mean fluorescent intensity (MFI) can be observed when CD4 was detected by DARPin 23.2. As the suitable concentration of DARPin 23.2 was determined from SupT1 which display different CD4 expression level, DARPin 23.2 at 100 ng/ml might not yield the best result for primary T cells. To obtain an improved CD4 profile using DARPin 23.2 for primary T cells, the proper concentration of DARPin 23.2 should be re-analyzed. Moreover, DARPin 23.2 also exhibited negative binding profile in CD4 negative cells (HeLa) as shown in Figure 3.12.

Computational study of protein has become popular initially in drug discovery industry. The field can be applicable to study protein-protein interaction to aid in better understanding the underlying mechanism in the reaction. In order to study this interaction, three dimensional structure (3D) of the studied protein is required, either from X-ray crystallography, NMR, cryoelectron tomography or homology modeling. Nimmanpipug *et al*. built homology model of CD4-specific DARPin 23.2 from template DARPin in order to identify the crucial amino acid residues lying in the interaction interface of

DARPin 23.2 with its target CD4. The interaction between DARPin 23.2 and CD4 were evaluated by molecular docking using ZDOCK protocol. The selected complex was tested for stability using molecular dynamics simulations. Calculated free energy obtained from the interaction revealed three plausible key amino acids [33]. Phenylalanine at position 100 was further suggested to be mutated into arginine to improve the binding affinity between the two proteins. Unexpectedly, binding affinity of CD4-specific DARPin (F100R) was totally depleted comparing to wild type (Figure 3.8). This indicated that F100 had a major role in binding interaction between DARPin 23.2 and CD4 and required to be untouched since computational mutation of DARPin 23.2 at F100 exhibited no other amino acid candidate which improved this binding affinity (Table 1.2).

The internal repeat of ankyrin repeat (AR) protein is protected by hydrophobic interface of N-terminal and C-terminal domain. Terminal capping, especially C-terminal cap has a major role in maintaining structural stability and providing flexibility. From the study of Binz *et al.*, both caps were originally taken from natural AR [27]. Further study proposed the molecular dynamics (MD) calculations of full-consensus DARPin. Internal repeat-inspired modification of C-cap suggested from MD simulation improved the stability and packing of C-terminal cap interface to the last internal repeat [59, 60]. X-ray crystallography of selected redesigned C-terminal module mutations confirmed the MD calculations [61].

Phenylalanine is an amino acid with hydrophilic aromatic side chain. Its position in C-terminal domain of DARPin is highly conserved among the synthetic libraries and also the mutated C-terminal module described above. The X-ray crystallographic study of DARPin and its target found that, Phe100 placed itself within the 5-Å radius of the interaction. This close proximity indicates the possible interaction to its partner target protein. The study of Veesler *et al.* built the DARPin libraries against the tip of receptor-binding protein (RBP, the BppL trimer) of the lactococcal phage TP901-1. Among them, DARPin 20 was selected as a candidate to be further evaluated in the interaction by performing crystallographic study (PDB ID: 3HG0). The binding interface indicated that the amino acid involving in this interaction were the randomized amino acid residues in internal repeats and also C-terminal cap. Phenylalanine at the same position as F100 in

CD4-specific DARPin 23.2 was within 5 Å radius, stated to be one of the key amino acid residues of DARPin 20 in the interaction to its target [62].

There is another study to support the importance of F100 in the interaction. The study of Kummer *et al.*, two DARPins specific to phosphorylated and unphosphorylated form of protein kinase ERK2 (extracellular signal-regulated kinase 2) were selected. DARPin against unphosphorylated form, E40 and its target ERK2 were examined at atomic level by forming crystal structure complex (PDB ID: 3ZU7). The same method was also performed with DARPin against phosphorylated ERK2, pE59 (PDB ID: 3ZUV). The atomic evaluation revealed that the amino acid residues involving in the interaction which included all variable residues in internal repeats and C-terminal cap. For both complex, C-terminal domain had a significant role in the interaction with kinase which included phenylalanine at the same position as F100 in CD4-specific DARPin 23.2 [63].

In order to stabilize protein structure, there are a number of interaction forces involving, which includes hydrogen bonds, electrostatic interactions, van der Waals interactions and hydrophobic interactions. The mutation of F100 to arginine or glycine might alter the hydrophobic interface or conformation of C-terminal cap of DARPin. In contrast to phenylalanine, arginine is categorized as a positively charged amino acid. The negative binding to its target, CD4-specific DARPin 23.2 (F100R) and also CD4-specific DARPin 23.2 (F100G) might be a result from the lost in hydrophobicity and favorable van der Waals contacts [64].

Since CD4 molecules has the true 3D structure obtained from X-ray crystallography and also the interaction between CD4 and its interaction pair, class II MHC [65] and HIV gp120 [14, 17] had been firmly studied, this would raise the predictive power from molecular docking. Wisitponchai *et al*., in contrast, studied the same binding interaction and identified key amino acids in this interaction on CD4 molecules instead. By performing a molecular docking using ZDOCK protocol combined with a set of physical criteria and *in vitro* information from Schweizer *et al.*, the researcher indicated three plausible key amino acids on CD4 molecules from group 1 which are K35, Q40 and Q25.

Using alanine scanning approach to evaluate the key amino acid residues on CD4 in the binding interface with DARPin 23.2, two versions of mutated CD4 had been generate. The first version is single mutation to alanine at lysine position 35 (K35A). Regarding to interaction model proposed by Wisitponchai (Figure 1.13), the second version of mutated DARPin 23.2 contained another mutation at glutamine position 40 (K35AQ40A). Adenovirus system was used as a tool for gene delivery and expression of CD4 in CD4 negative cells. CD4 mutants were cloned into transfer vector, pAdTrack-CMVGFP. Homologous recombination between pAdTrack-CMVGFP-CD4 mutant and pAdEasy-1 was held in *E. coli* BJ5183. The recombined adenoviral vectors were transfected into packaging 293A cells. Concentrated virus was used in this experiment.

HeLa cell line was used to be the target cells to express CD4 as it expressed Ad5 entry receptors (CAR and αv-integrins) but showed negative profile for CD4 surface expression when stained with both mAb against CD4 (MT4) and purified CD4-specific DARPin 23.2 (Figure 3.11 and 3.12 respectively). Low expression of CAR was observed in this experiment (44.61% positive). HeLa is known to express high level of CAR for both mRNA and protein [66, 67]. One factor that may affect the low expression of CAR is the unstability of protein expression in cancer and also cancer cell line. Moreover, mouse anti-CAR mAb used in this experiment is needed to be optimized.

HeLa cells was transduced by adenovirus which can visualized by expression of reporter gene, GFP. Moreover, transduced HeLa cells also expressed CD4 (WT), CD4 (K35A) and CD4 (K35AQ40A) when stained with anti-CD4 mAb (MT4) (Figure 3.22). Binding affinity of DARPin 23.2 to both CD4 mutants remained stable when compared to the interaction with CD4 (WT) (Figure 3.23). This positive interaction indicated that K35 and Q40 had no role in the binding interface of CD4 and CD4-specific DARPin 23.2.

When critically evaluated the protocol Wisitponchai [34] used to identify the earlier stated hot spots, some bias and difference in critical criteria to identify the key amino acid residues were found. Discovery Studio (DS) is a software developed for study of protein, ligand and pharmacophore modeling. To predict protein-protein interaction, ZDOCK algorithm for rigid body protein can be used. According to Wisitponchai, the interaction complex obtained from the ZDOCK protocol were top 2,000 poses ranked with top ZDOCK score out of initial 54,000 poses. The top 20 poses of highest ZDOCK

score were chosen for further evaluation. ZDOCK score alone was not sufficient to recruit the true pose mimetic to native interaction. In the study of Pierce and Weng compared the scoring option for protein using the combination of ZDOCK and ZRANK. They found that the hit success rate (interface RMSD $\leq$ 2.5 Å) was around 10% within the ten predictions. On the other hand, combining ZDOCK and ZRANK increased the hit success rate ranging from 10 – 25%. The success rate were greater when expand the larger prediction numbers. Within 100 prediction, hit success rate from ZDOCK and ZRANK went up to 55%. The same pattern was observed with hit success rate and near-hit success rate (4 > RMSD $\geq$ 2.5 Å) [68].

Wisitponchai *et al.* had propose 3 clusters of CD4-specific DARPin 23.2 and CD4 complexes as seen in Figure 1.12. The cluster was identified after ranking the top 20 ZDOCK score. The complexes were evaluated for binding domain of CD4. Only 11 poses was chosen as they bound to D1 of CD4 according to Schweizer *et al.* [5]. As there were the highest member in group 1, proposed key amino acids were from this group as bias. The area of binding site on CD4 can be coarsely identified beforehand by choosing Receptor Binding Site Residue according to the study on gp120 binding. By doing this, DARPin will be docked close to or hit the native binding site. Moreover, combining this parameter with ZDOCK and ZRANK score will improve the predictive power and increase the member of clusters to be more diverse.

Nonetheless, the sampling size of 20 as used by Wisitponchai is enough. According to Epa *et al*. the molecular docking of HER2 and G3 was performed by ZDOCK algorithm. The top 2,000 docked solution was ranked by ZRANK. The top 20 ZRANK poses were further analyzed for total interaction energy (IE) and shape complementarity (SC). The outstanding candidate for this complexed with the lowest IE of -19.37 kcal/mol which was in the fourth ZRANK [69]. Interaction energy is one of the chemical property criteria to determine the protein-protein interaction. Wisitponchai's criteria counted the interaction pair rather than calculating energy. However, in molecular docking between $Ank^{GAG}1D4$ [70] and its interaction pair HIV-1 capsid protein (CA), top ranked score of ZDOCK and ZRANK were chosen from primary ZDOCK. The interaction atom within 5 Å of the interactive amino acid residues after running RDOCK

were counted. The key amino acid residue with highest frequency interaction pair were indicated for CA protein [71].

For further identification of the key amino acid residues in this interaction using molecular docking, especially on CD4 molecule are recommended, since solid information on 3D structure and the interaction with their interaction partner, either with HIV-1 gp120 or class II MHC are available. In combination with the selective criteria and parameters mention above, ZDOCK algorithm with defined Receptor Binding Site Residue along with ZRANK ranking will raise the predictive power and lead to more precise consideration for the binding hot spot.