

TABLE OF CONTENT

	Page
ACKNOWLEDGEMENT	iii
ABSTRACT (ENGLISH)	v
ABSTRACT (THAI)	vii
LIST OF TABLES	xiii
LIST OF FIGURES	xv
ABBREVIATIONS	xvii
CHAPTER 1 INTRODUCTION	1
Statement and Significance of the Problem	1
HIV infection	3
T cell Responding	4
HIV Life Cycle	5
CD4 Binding	7
HIV/AIDS Treatment	8
Ankyrin Repeat Proteins	10
Structure of Ankyrin	10

DARPin	11
CD4-specific DARPin	15
Protein-protein Docking	16
ZDOCK Protocol	18
RDOCK Protocol	21
Other Docking Software	21
Key Binding Site Residues	22
Physical Properties of Binding Residue	22
Evaluation of Web Servers	27
Histogram Analysis	32
Validation	33
CHAPTER 2 LITERATURE REVIEW	35
Literature review	35
Purpose of the study	39
Research scope	40
CHAPTER 3 RESEARCH METHODOLOGY	41
Computational Simulation	41

Searching and Preparing for 3D Structures	41
Predicting the Protein-Protein Complex	43
Finding the Intermolecular/Interface Neighbor	45
Designing Algorithm for finding key residue	45
Extracting Considered CD4's Residues	46
Identifying Key Binding Residues	52
Validating the Predicted Key Residues	55
CHAPTER 4 RESULTS AND DISCUSSIONS	57
Computational Simulation	57
Searching and Preparing for 3D Structures	57
Predicting the Protein-Protein Complex	60
Finding the Intermolecular/Interface Neighbor	67
Designing Algorithm for finding key residue	69
Extracting Considered CD4's Residues	69
Identifying Key Binding Residues	73
Validating the Predicted Key Residues	81

CHAPTER 5 CONCLUSIONS	87
BIBLIOGRAPHY	92
APPENDICES	102
VITA	149

LIST OF TABLES

Table	Page	
1	The properties of 20 standard amino acid	26
2	Performance values of various empirical prediction method of HotPoint	29
3	Weight of energy terms in the scoring functions	32
4	Characters of atom using in criterion 3 and the number of backbone and side chain atoms using in criterion 4	51
5	The potential energy and RMS gradient of CD4 and DARPin 23.2	59
6	The Ramachandran plot of minimizing structures of CD4 and DARPin 23.2	60
7	The interface residue of CD4 bound to gp120, MHCII, and DARPin 23.2	62
8	Docking scores of 11 complexes	66
9	Hydrogen bond in 11 poses with the distance cutoff and angle as 2.5 Å and 120-180°	68
10	The considered CD4 residues that were a combination of top 10 in criterion 1-4	70
11	The renormalized values of considered CD4 residues in six pattern of 11 complexes	73
12	The top 3 key CD4 residues bound to DARPin 23.2	80

13	The hot spot propensity analyzed by Bogan & Thorn was divided into 4 classes	81
14	The hot spots predicted by HotPOINT	85
15	The hot spots predicted by HSPred	85
16	The intersection of hot spots between HotPOINT and HSPred	86
17	PDBid of CD4 crystal structures reported in protein data bank (updated 6 June, 2013)	103
18	The frequency value in criterion 1	104
19	The frequency value in criterion 2	107
20	The frequency value in criterion 3	109
21	The frequency value in criterion 4	111
22	The raw frequency value of considered CD4 residues in 11 poses	113
23	The normalized value of considered CD4 residues in 5 criteria of 11 poses	119
24	The hot spots and non-hot spots with their scoring value of HotPOINT prediction	125
25	The hot spots and non-hot spots with their scoring value of HSpred prediction	136

LIST OF FIGURES

Figure		Page
1	The estimated number of people living with HIV by UNAIDS in 2001	4
2	Schematic representation of HIV life cycle within a CD4+ cell	6
3	The x-ray structures of CD4, CD4-MHCII complex, and CD4-gp120 complex	8
4	Structural features of the ankyrin repeat using AR consensus refinement of Binz et al.	12
5	Structure of ankyrin repeat	13
6	Locations of the interface in four different AR co-crystal structure	14
7	The AR consensus from Binz et al. (2003)	15
8	Overview of the protein-protein docking stage, which consist of initial and refinement stage	18
9	The schematic of intramolecular interactions	23
10	The schematic overview of protein regions defined the different energy contributions	30
11	The binding and non-binding residue of DARPin 23.2 relied on AR consensus by Binz et al.	44

12	The designed algorithm to identify the number of interaction pair in each CD4's amino acid (criterion 2)	48
13	The designed algorithm to identify the number of interaction pair in each CD4's amino acid (criterion 1)	49
14	The designed algorithm to identify the number of interaction pair in each CD4's amino acid (criterion 3)	50
15	Flowchart of algorithm for finding key residues of CD4 binding DARPin 23.2	54
16	Diagram data flow of selecting proposed CD4 showing in Venn diagram and tree diagram	58
17	The Ramachndran plot of minimized CD4 and minimized DARPin 23.2	60
18	The structures of CD4-gp120 complex, CD4-MHCII complex and CD4-DARPin 23.2 complexes	63
19	The binding area on the tip of domain 1 of CD4	64
20	The docked scoring of docked poses	66
21	The superimposition of 3 rational CD4 – DARPin 23.2 complexes	67
22	The characteristic of intermolecular neighbor	69
23	The intermolecular neighbors in all 11 poses are shown in histogram values in four criteria	71

ABBREVIATIONS

3D	Three dimensional
6D	Three dimensional
Å	Angstrom
°	Degree
A	Alanine
AIDS	Acquired immune deficiency syndrome
ACE	Atom Contact Energy
Ala	Alanine
ANK	Ankyrin repeat
AR	Ankyrin repeat
ARV	Antiretroviral
Arg	Arginine
ASA	Accessible surface area
Asn	Asparagine
Asp	Asparic acid

CONJ	Conjugate Gradient
CTLs	Cytotoxic T lymphocytes
D	Asparic acid
D1-D4	Domain 1 – domain 4
DARPin	Designed Ankyrin Repeat Protein
DS	Discovery studio
DE	Desolvation
E	Glutamic acid
ELEC	electrostatic
F	Phenylalanine
FFT	Fast Fourier Transform
G	Glycine
Gly	Glycine
Gln	Glutamine
Glu	Glutamic acid
H	Histidine
His	Histidine

HAART	Highly Active Antiretroviral Therapy
HIV	Human immunodeficiency virus
I	Isoleucine
IgC	Ig constant
IgV	Ig variable
Ile	Isoleucine
kDa	Kilo dalton
kcal/mol	Kilogram calorie per mol
L	Leucine
Leu	Leucine
M	Methionine
Met	Methionine
MHCII	Histocompatibility complex class II
N	Asparagine
Nabs	Neutralizing antibodies
Phe	Phenylalanine

PP	Pair potential
PPI	Protein-protein interface
PID ^{predict}	Percentage of identity prediction
PSC	Pairwise shape complementarity
Q	Glutamine
R	Arginine
RMSD	Root mean square deviation
SD	Steepest Descent
SPF	Spherical polar fourier
SVM	Support Vector Machine
Trp	Thryptophan
Tyr	Tyrosine
V	Valine
Val	Valine
W	Thryptophan
Y	Tyrosine