THESIS APPROVAL

GRADUATE SCHOOL, KASETSART UNIVERSITY

Doctor of Philosophy (Biochemistry)
DEGREE

Biochemistry
FIELD

Biochemistry
DEPARTMENT

TITLE: Characterization of Single Domain – Antibodies and Small Molecular Inhibitors to ErbB2 Tyrosine Kinase

NAME: Mr. Kunan Bangphoomi

THIS THESIS HAS BEEN ACCEPTED BY

THESIS ADVISOR
(Assistant Professor Kiattawee Choowongkomon, Ph.D.)

THESIS CO-ADVISOR
(Professor Wanpen Chaicumpa, D.V.M. (Hons.), Ph.D.)

THESIS CO-ADVISOR
(Miss Sasimanas Unajak, Ph.D.)

DEPARTMENT HEAD
(Assistant Professor Kiattawee Choowongkomon, Ph.D.)

APPROVED BY THE GRADUATE SCHOOL ON _____

DEAN
(Associate Professor Gunjana Theeragool, D.Agr.)

THESIS

CHARACTERIZATION OF SINGLE DOMAIN-ANTIBODIES AND SMALL MOLECULAR INHIBITORS TO ERBB2 TYROSINE KINASE

KUNAN BANGPHOOMI

A Thesis Submitted in Partial Fulfillment of
the Requirements for the Degree of
Doctor of Philosophy (Biochemistry)
Graduate School, Kasetsart University
2013

Kunan Bangphoomi 2013: Characterization of Single Domain-Antibodies and Small Molecular Inhibitors to ErbB2 Tyrosine Kinase. Doctor of Philosophy (Biochemistry), Major Field: Biochemistry, Department of Biochemistry. Thesis Advisor: Assistant Professor Kiattawee Choowongkomon, Ph.D. 163 pages.

Erythroblastosis oncogene B 2, also known as ErbB2, is a cell membrane surface-bound receptor tyrosine kinase protein that important for cell signaling pathway. ErbB2 protein overexpresses in ErbB2 positive cancer especially mammary cell cancer and lung cancer. Currently, trend of anti-cancer molecule is the development of target specific molecule such as lead compound and monoclonal antibody. Although, there are many anti-cancer drugs that launched in the market, the new candidate of anti-cancer drug is still important. In this study, the anti-cancer molecules including monoclonal antibody and lead compound were developed by using bio-panning and virtual screening technique. The VH/V_HH antibody against ErbB2-TK clone number 4, 17, 22, and 25 provided the ELISA binding signal with ErbB2-TK and EGFR-TK. Moreover, the VH/V_HH antibody clone number 4 and 22 can inhibit the TK activity of EGFR. The tetrad amino acid hallmark identification revealed that, the amino acid sequence analysis revealed that clone number 4, 22, and 25 were VH antibodies, while clone number 17 was V_HH antibody. The molecular docking of VH/V_HH antibody clone number 4, 17, 22, and 25 revealed that, there are many binding sites of all antibodies on ErbB2-TK and EGFR-TK. The docking results of all antibodies correlated with ELISA and inhibiting TK activity assay. In addition, the structure of ErbB2-TK was used to screen lead compound from NCI database and ChemBridge Diverset™. There are 10 lead compounds that analyzed the binding residue from each library.

_____ _____ __ / __ / __

Student's signature Thesis Advisor's signature

ACKNOWLEDGEMENT

I owe a depth of gratitude to my advisor and co-advisor, Assist. Prof. Kiattawee Choowongkomon (KC), Prof. Wanpen Chaicumpa (WC), and Sasimanas Unajak (SU). Moreover, I would like to Prof. Pa-thai Yenchitsomanus (PY). They give me the vision and foresight which inspired to conceive this project. I would like to express my appreciation to member of KC laboratory, WC laboratory, and PY laboratory for hospitality and friendliness. They give me a chance to use the laboratory facility in their lab. They also give valuable advice to me all the time. I am particularly in debited to Miss Orathai Sawatdichaikul, Miss Ornnutchar Poungpair, Miss Siriluk Ratanabunyong, Miss Kanyarat Thueng-in, Miss Wanarat Yim-im, Mr. Napat Songtawee and, Mr. Jeeraphong Thanongsaksrikul for their frequent aid and best suggestion.

Finally, I would like to thank National Nanotechnology Center (NANOTEC), Thailand for providing the Discovery Studio 2.5 package.

Kunan Bangphoomi

January, 2013

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF ABBREVIATIONS

AREG	=	amphiregulin
ATP	=	adenosine-triphosphate
bp	=	base pairs
BTC	=	betacellulin
CD	=	circular dichroism
CDR	=	complementary-determining region
CR	=	cysteine-rich domain
DAG	=	diacylglycerol
DARPINs	=	designed ankyrin repeat protein
ECD	=	extracellular domain
EGF	=	epidermal growth factor
EGFR	=	epidermal growth factor receptor
EGFR-TK	=	epidermal growth factor receptor tyrosine kinase domain
EGFR-TKIs	=	EGFR-TK inhibitors
ELISA	=	enzyme linked immunosorbant assay
ER	=	endoplasmic reticulum
ERK	=	extracellular signal-regulated kinase
ErbB	=	erythroblastosis oncogene B
EREG	=	epiregulins
FGFR	=	fibroblast growth factor receptor
FP	=	fluorescence polarization
FR	=	framework
FRET	=	fluorescence resonance energy transfer
Grb2	=	growth factor receptor-bound protein-2
GST	=	glutathione S-transferase
HB-EGF	=	heparin binding EGFR-like growth factor
H-bond	=	hydrogen bond
HCAbs	=	heavy chain antibody
HER	=	human Epidermal Receptor

LIST OF ABBREVIATIONS (Continued)

HTS	=	high-throughput screening
ICD	=	intracellular domain
Ig	=	immunoglobulin(s)
IP3	=	inositol triphosphate 3
LB	=	Luria-Bertani medium
LB-AG	=	Luria-Bertani medium with ampicillin and glucose
mAbs	=	monoclonal antibodies
MAPK	=	mitogen-activated protein kinase
mdm2	=	mouse double minute 2 homolog
MMP	=	matrix metalloproteinase
mTOR	=	mammalian target of rapamycin
Nb	=	nanobody
NC	=	nitrocellulose
NCI	=	national Cancer Institute
NMR	=	nuclear magnetic resonance
NRG	=	neuregulins
OD	=	optical density
PAbs	=	polyclonal antibodies
PARAs	=	pro-apoptotic receptor agonist
PB	=	peripheral blood
PBS	=	phosphate buffered saline
PCR	=	polymerase chain reaction
PDB	=	protein Data Bank
PDK1	=	pyruvate dehydrogenase lipoamide kinase isoenzyme 1
PI3K	=	phosphoinositide 3-kinase
PIP3	=	phosphatidylinositol (3, 4, 5)-triphosphate
PKs	=	protein kinases
PKB	=	protein kinase B
PLC	=	phospholipase
PTKs	=	protein tyrosine kinases

LIST OF ABBREVIATIONS (Continued)

PTP	=	protein tyrosine phosphatases
Ras	=	rat sarcoma
RECK	=	reversion-inducing-cysteine-rich protein with Kazal motifs
RF	=	replicative form
RMSD	=	root mean square deviation
RTKs	=	receptor tyrosine kinases
RTP	=	rivavirin triphosphate
SH	=	Src homology
SOS	=	Son of sevenless homolog
SPA	=	Staphylococcal surface display of immunoglobulin A
SPA	=	Scintillation proximity assay
Src	=	sacroma
ssDNA	=	single-stranded DNA
STAT	=	signal transducers and activators of transcription factors
TEMED	=	N, N, N', N'-tetramethylenediamine
TGF	=	transforming growth factor
TK	=	tyrosine kinase
TRIAL	=	tumor necrotic factor apoptosis-inducing ligand
USP	=	ubiquitin specific protease
UV	=	ultraviolet
VDW	=	Van der Waals
VEGF	=	vascular endothelial growth factor
VH	=	Variable domain of heavy chain
VL	=	Variable domain of light chain
VS	=	virtual screening

CHARACTERIZATION OF SINGLE DOMAIN-ANTIBODIES AND SMALL MOLECULAR INHIBITORS TO ERBB2 TYROSINE KINASE

INTRODUCTION

Cancer is a generic term for a large group of disease that abnormal growth and divides without control. The cancer can spread and invade into the other tissue via blood vessel and lymphatic vessel. Subsequently, the cancer harms affected tissue. There are many damage of tissue resulting from cancer such as ischemia and apoptosis of the normal cells. These lesions result from invasion of cancer cell upon adjacent normal cells. Despite, the cancer cell should bear some structural and functional resemblance to the cells and tissues of origin, but they are not responds to usual physiologic regulation. The prognosis of cancer is depends on type of cancer and its complication. For example, low metastasis cancer is prognosis good while high metastasis cancer is prognosis poor. Nowadays, there are several treatments for cancer such as surgery, radiation, chemotherapy, and combination therapy. However, the treatment of cancer should be considered about type of cancer, stage, and patient's condition. In the past, there are many types of conventional chemotherapy; alkylating agent, anti-metabolite, plant alkaloid, vinca alkaloid, podophyllotoxin, taxane, topoisomerase inhibitor, and anti-neoplastic. Although the conventional chemotherapy provides many advantages, there are many adverse effects such as depression of the immune system, fatigue, gastrointestinal distress, and hair loss. Currently, targeted therapies such as monoclonal antibodies (mAbs) and kinase inhibitor are the new era of cancer therapy. Targeted therapies provide many advantages and less adverse effects because they focus at a molecular abnormality certain types of cancer such as overexpression of tyrosine kinase (TK) in lung and breast cancer.

Growth factor is one of important factor for cancer cell growth. The functions of growth factor are apoptosis control, cell proliferation, cell differentiation, gene

expression, immune response, and organization of cell cytoskeleton. Growth factor has a receptor at a cell surface; especially this receptor is overexpressed in a cancer cell. There are many types of growth factor receptors, including epidermal growth factor family (EGFR) and fibroblast growth factor receptor family (FGFR) (Haugsten *et al.*, 2010; Turner and Grose, 2010). Nowadays, there are many studies reveal the relationship between EGFR and cancer, especially human epidermal growth factor receptor 2 (ErbB 2 or HER 2). ErbB2 is a cell membrane surface-bound receptor tyrosine kinase (RTKs) and is normally involved in the signal transduction pathways leading to cell growth and differentiation. There are many evidences that reveal the role of ErbB2 and metastasis-associated properties of cancer cell (Yu and Hung, 2000), especially ErbB2 positive type breast cancer. There are twenty-five percent of breast cancer patients have ErbB2 positive breast cancer. The prognosis of ErbB2 positive breast cancer patient is poor. Moreover, the life span of patient with ErbB2 positive breast cancer is lower than other type of breast cancer patients.

ErbB2 is an interested protein for used as a target for targeted therapy. There are two types of anti-ErbB2 drug including monoclonal antibody and small molecule. Trastuzumab (Herceptin®) is monoclonal antibody that targeted an ECD of ErbB2. However, this drug contains important disadvantages including drug resistant and induced heart failure (Chen *et al.*, 2012; Gottesman, 2002). The mechanism of Trastuzumab (Herceptin®) resistance is the alteration of the drug target. The alteration of ErbB2 is the lack of ECD that used for binding with Trastuzumab (Herceptin®). Moreover, the ECD targeted drug cannot overcome breast cancer metastasis because monoclonal antibody unable penetrated to blood brain barrier that is a target organ for metastasis cancer cell. For small molecule anti–TK activity such as Sorafenib (Nexavar®) and Lapatinip (Tykerb®), these small molecules have TK domain of EGFR family as a target. Drug targeted TK domain can solve drug resistant problems that resulted from truncated ErbB2 and dimerization of EGFR family. From previous studies reveal that, there are many disadvantages of currently anti-ErbB2 drug, including drug resistant and adverse effects. As the disadvantages of current anti-cancer drug, the finding of new drug candidate and bio-molecule are still important for cancer therapy. For drug candidate discovery, there are two techniques

that used for this approach: high-throughput screening technique and virtual screening technique. Virtual Screening (VS also called *in silico* screening) is a computational approach which used for drug discovery. This technique is applied for finding new lead compound that can react with target of interest. The criterion for this selection is appropriate binding energy and residue. For bio-molecules discovery, the monoclonal antibodies are interest candidates because these antibodies provide many advantages when compared with known anti-ErbB2 drug such as the binding affinity and the anti-drug resistance properties. For the aspect of binding affinity, the monoclonal antibody provides high binding affinity to the antigen of interest resulting from the number of interacting residue. For the aspect of anti-drug resistance properties, the monoclonal antibodies have many binding residues when compared with known anti-ErbB2 drug. As such number of binding residues, the antibodies can inhibit the activity of ErbB2 when the mutation occurs in cancer. There are many format of monoclonal antibody including single chain variable fragment (ScFv) and heavy chain antibody fragment (VH/V_HH) from camel (*Camelus dromedarius*). The VH/V_HH monoclonal antibodies have many advantages when compared with ScFv including penetration capacity and structural stability. Moreover, the lack of constant region of antibody VH/V_HH monoclonal antibodies can reduce the adverse effect resulting from the remaining of constant region (Klement *et al.*, 2012). According to the antibody production technique, the antibody can be produced by using hybridoma and biopanning technique. The hybridoma technique is a technology of forming hybrid cell line (also called hybridoma) by fusing a specific antibody producing B cell with a myeloma cell (Milstein, 1999). Although, the hybridoma technique provides many advantages, it provides very large time and effort commitment (Osterhaus and UytdeHaag, 1985). For the biopanning technique, it provides small time and effort commitment when compared with conventional technique. Moreover, this technique also provides the various type of antibody formats (Ellis *et al.*, 2012).

The study of antibody-antigen interaction is a useful knowledge for the development of antibody. There are many techniques that used for study the binding of antibody and antigen. These techniques include mimotope mapping (Thueng-In *et al.*, 2012), yeast two-hybrid (Uetz *et al.*, 2000), mass spectrometry (Gavin *et al.*,

2002), X-ray crystallography, protein NMR, and computational techniques. For X-ray crystallography and protein NMR, both techniques are extremely time-consuming contrary to the computational techniques that are extremely fast technique. For computational techniques, there are including protein modeling and protein-protein docking. The accuracy of computational technique was approved by using mimotope mapping and inhibition assay (Chavanayarn *et al.*, 2012; Chulanetra *et al.*, 2012; Thueng-In *et al.*, 2012).

In this study, The VH/V_HH anti-TK domain of ErbB2 was screened from VH/V_HH library by using biopanning technique. The screened VH/V_HH anti-TK domain of ErbB2 was subjected to inhibition assay. Subsequently, the DNA of screened antibody was subjected to DNA sequencing and used for computational analysis. For computational analysis, the antibody-antigen interaction was studied by using Autodock4 (Scripps Research Institute), ZDOCK and RDOCK module that embedded in Discovery Studio 2.5 (Accelrys USA). The accuracy of docking calculation was determined by using data comparison with mimotope mapping and inhibition assay. The samples for antibody docking method verification are including ScFv against tetrodotoxin (Chulanetra *et al.*, 2012), and VH/V_HH against RNA dependent RNA polymerase of hepatitis C virus (Thueng-In *et al.*, 2012). The verified docking method was subjected to study the interaction of VH/V_HH anti-TK domain of ErbB2. Computational analysis provides much information including model of VH/V_HH, the interactions of amino acid residues that reside in the binding pocket of antibodies and ErbB2, binding energy of antibodies and ErbB2. In addition, this study was shown the screening of lead compound that bound TK domain of ErbB2 by using virtual screening technique. The lead compounds were screened from two libraries including National cancer institute database (NCI) and Chembridge Diverset™. Screened lead compounds were studied the binding properties including binding energy and residues by using computational approaches. The computational data will be a useful knowledge for further drug developments.

OBJECTIVES

1. To screen and produce antibody that against TK activity of ErbB2 by using bio-panning technique.

2. To verify the antibody-antigen docking by using Autodock4 (Scripps Research Institute), ZDOCK and RDOCK module that embedded on Discovery Studio 2.5 (Accelrys USA).

3. To screen lead compounds from various substance libraries against TK activity of ErbB2 by using virtual screening technique

LITERATURE REVIEWS

1. ErbB2

Human epidermal growth factor receptor 2 (ErbB2, HER2, or *neu*) is a receptor protein which important for cell development and cell proliferation. Although, this receptor is expressed in almost every cell but it is overexpressed in breast and lung cancer cell (Coussens *et al.*, 1985; Quirke *et al.*, 1989). ErbB2 is a member of epidermal growth factor receptor family that composed of ErbB1 (EGFR), ErbB2 (HER2), ErbB3 (HER3), and ErbB4 (HER4) respectively (Holbro *et al.*, 2003). These receptors are activated by ligand-activated oligomers which regulated intracellular processes through an oligomeric tyrosine kinase scaffold. The ErbB2 structure is composed of extracellular domain (ECD), transmembrane domain, juxtamembrane domain, tyrosine kinase domain (TK), and carboxyl-terminal tail (Figure 1.). The region from juxtamembrane domain till carboxy-terminal tail is called intracellular domain (ICD).

ErbB2 consists of 1255 amino acids residues with molecular weight of 185 kDa.The cDNA of erbB2 (c-erbB2) protooncogene is located on human chromosome 17. This gene is encoded into ErbB2 protein which has founded in several human tissues in different organs such as normal epithelium, endometrium epithelium, ovarian epithelium, neuromuscular tissue, prostate, pancreas, lung, kidney, liver, heart, and hematopoietic cells. Low level of ErbB2 expression is detected in mononuclear cells from bone marrow, peripheral blood (PB) and mobilized PB. The higher expression has been found in cord blood-derived cells. Although, low expression of ErbB2 involved with the quiescence of $CD34^{+}$ progenitor cells from all blood sources and resting lymphocytes, but the high expression of this receptor is up-regulated during cell-cycle recruitment of progenitor cells. Similarly, it increases in mature, hematopoietic proliferating cells, underlying the correlation between ErbB2 and the proliferating status of hematopoietic cells.

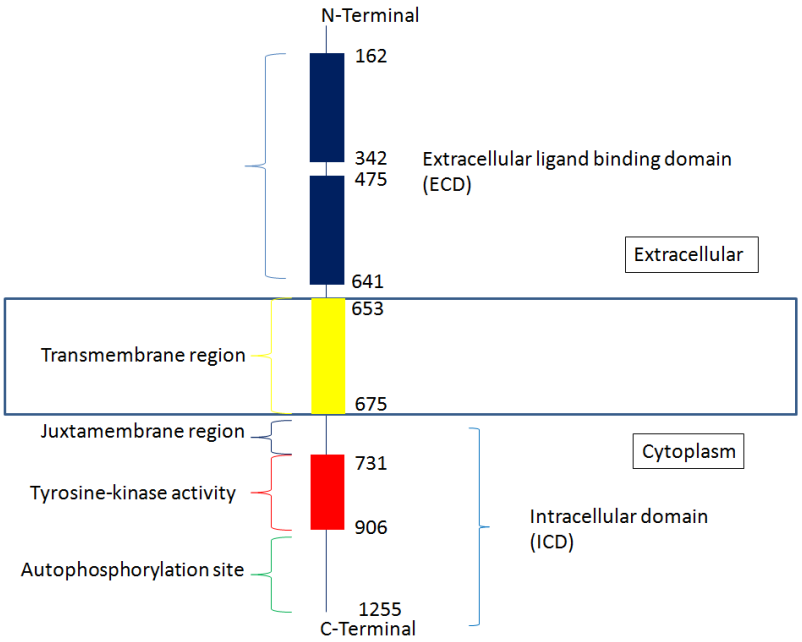


Figure 1 Schematic representation of ErbB2. Receptor tyrosine-kinases (RTKs) are cell surface allosteric enzymes consisting of an extracellular ligand-binding domain (ECD) (blue), a single transmembrane (TM) domain has an extensive homology to the epidermal growth factor receptor (EGFR) (brown), and a cytoplasmic domain with catalytic activity (green).

Source: Adopt from Casalini *et al.* (2004)

The understanding of ErbB2-TK structure is important to design potent inhibitors. Before 2011, structure of EGFR-TK was used as a model for study of ErbB2-TK structure. There are two reasons for application of EGFR-TK structure as a template for understanding of ErbB2-TK structure. First, EGFR-TK and ErbB2-TK are closely related member of EGFR family. Second, the various conformation of ErbB2-TK structure is unavailable. For today, more than thirty crystal structures of EGFR kinase domain were deposited in PDB database (Blair *et al.*, 2007; Brewer *et al.*, 2009; Stamos *et al.*, 2002; Wood *et al.*, 2004; Xu *et al.*, 2008a; Xu *et al.*, 2008b; Yun *et al.*, 2007; Yun *et al.*, 2008; Zhang *et al.*, 2007; Zhang *et al.*, 2004). The summary of ligand for EGFR was shown in Table 1. According to the deposited EGFR-TK structure, there are two types of conformation such as active conformation

and inactive conformation. The activation of EGFR from inactive to active conformation is resulted from the binding of ligand. The EGFR-TK is composes of N-lobe and C-lobe. The differences of active conformation and inactive conformation are arrangement for each motif including helix-alpha-C, leucine pack residues (Leu834 and Leu837), salt-bridge residues (Lys721 and Glu738) and activation loop. 3D of EGFR-TK (PDB entry 1M17) (Stamos *et al.*, 2002) and 1XKK (Wood *et al.*, 2004) were used as example for study of active conformation and inactive conformation of EGFR-TK. Data that summary from all three dimensional structure of EGFR-TK reveal that the most of deposited structures of EGFR-TK are active conformation and water molecules are not necessary for binding of ligand.

Nowadays, there are two crystal structures of ErbB2 which deposited in protein data bank (PDB). Both structures were introduced in 2011. The first structure is HER2 kinase domain complex with TAK-285 (PDB entry 3RCD) (Ishikawa *et al.*, 2011). The latest structure of ErbB2-TK that deposited in PDB database is crystal structure of the kinase domain of human HER2 (ErbB2). The PDB entry of this structure is 3PP0. The resolution of 3RCD and 3PP0 are 3.21 angstrom (A°) and 2.25 A° respectively. Moreover, both structures are active-like conformation.

EGFR is activated after ligand binding. The ligand induces homodimerization or heterodimerization (Alroy and Yarden, 1997). Dimerization of EGFR is essential for TKs activation and subsequent recruitment of target proteins. For initiating cascade, a complex signaling cascade leads into distinct transcriptional programs (Garrett *et al.*, 2003). There are the several EGFR binding ligands (Table 1). The ligand binds to ECD of EGFR family, then downstream signal is transferred to ICD and it activates the cell function. The schematic of dimerization of EGFR family was shown in Figure 2. Although, ErbB2 has no specific ligand, it can play a major coordination role in EGFR signaling network via heterodimerization mechanism (Tzahar *et al.*, 1998). The important role of coordination of ErbB2 family is enhance and stabilizes the dimerization of EGFR. Other EGFR with specific ligand prefers to cooperate with ErbB2 as its heteromeric partner. ErbB2 in heterodimeric form is characterized by extremely high signaling potency due to dramatically reduction of

ligand dissociation, allowing strong and prolonged activation of downstream signaling pathways (Tzahar *et al.*, 1997).

Table 1 Ligands of EGFR family.

Group	Ligand	Receptor
1	Epidermal growth factor (EGF)	ErbB1
1	Transforming growth factor a (TGF- α)	ErbB1
1	Amohiregulin	ErbB1
2	Heparin-binding EGF (HB-EGF)	ErbB1, ErbB4
2	Betacellulin	ErbB1, ErbB4
2	Epiregulin	ErbB1, ErbB4
3	Heregulin (neuregulin, NRG family)	ErbB3, ErbB4

Source: Adopt from Rowinsky (2004)

All proteins in EGFR family are a member of receptor tyrosine kinase (RTK) activity, except ErbB3 that lack of TK activity. Then TK activities of ErbB3 proteins are supported by other EGFR. For example, ErbB2 plays an important role for signaling pathway of ErbB3 protein. For downstream signaling activation of ErbB3 via ErbB2, ErbB2 activated the signal of ErbB3 by using heterodimerization (Arteaga, 2003). RTK of EGFR is also plays an important role for cell differentiation, cell transformation, angiogenesis, and cell migration (Yeon and Pegram, 2005).

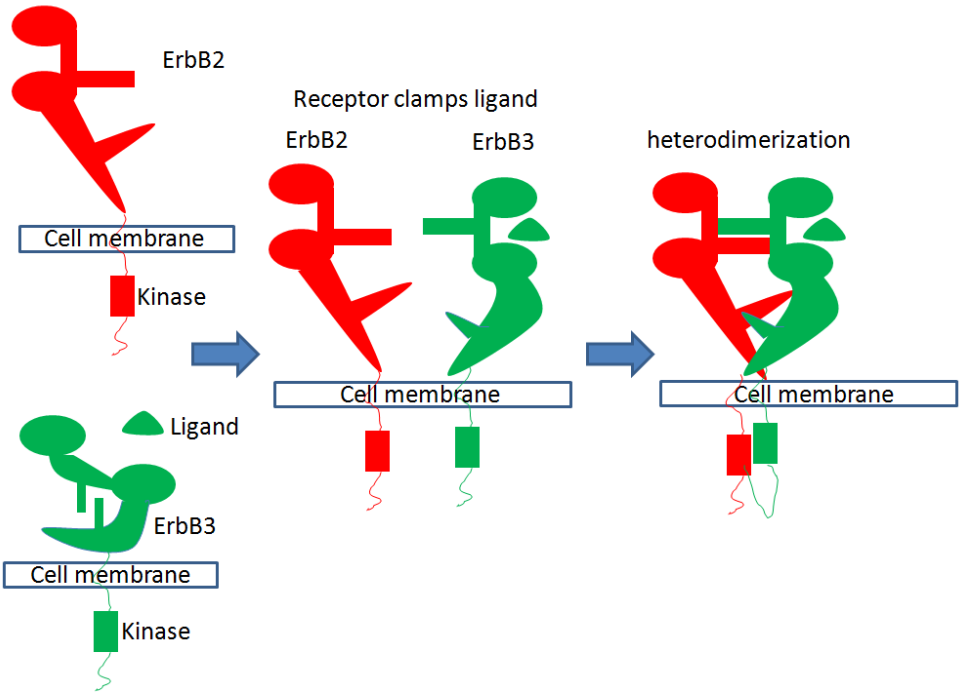


Figure 2 Schematic represented ErbB2 heterodimer dimerization. ErbB2 is originally in an active conformation, poised which ready to interact with other EGFR. By contrast, the EGF receptor and another family member, ErbB3 are held in an inactive form involving an interaction between the cysteine-rich 1 (CR1) loop and a loop in the CR2 domain. On ligand binding, the tethering between CR1 loop and CR2 domain is released, rendering the CR1 loop available for dimerization. Heterodimerization occurs resulting in signaling. When TK domain of ErbB2 is inhibited, the downstream signaling is blocked.

Source: Casalini *et al.* (2004)

The gene encoded EGFR family is a proto-oncogene. The normal gene that switches to oncogene is caused by mutations or increasing of its expression level. There are many mechanisms that induce carcinogenesis including overexpression of proto-oncogene, self-activation of this receptor, over synthesis of growth hormone,

and RTK activation by using other receptor. However, these mechanisms are resulting from gene mutation and environment.

2. Cell signaling in cancer

Signal transductions of cells are involved with survival, growth, and dead of cancer cells. There are six important pathways that control these mechanisms including cellular apoptosis pathway, Akt signal pathway, EGF pathway, p53 pathway, matrix metalloproteinase pathway, and vascular endothelial growth factor (VEGF) pathway. These pathways have response to a variety of ligands such as growth factor, toxin, cytokine, and hormone. Cancer cells also used these mechanisms like normal cells, but different in number of involved proteins or activate by involved protein mutated. The abnormality of involved protein can induce the normal cells to cancer cell.

2.1 Cellular apoptosis pathway

Apoptosis or program cell death is an important process for cancer development and progression. Cancer cell has property to avoid this program. Apoptosis is triggered via extrinsic pathway and intrinsic pathway. The extrinsic pathway is including pro-apoptotic signal such as Apo2L/ Tumor necrotic factor apoptosis-inducing ligand (TRAIL) and other pro-apoptotic receptor agonists (PARAs). The intrinsic pathway is initiated by cellular development or as a result of severe cellular stress such as DNA damage and toxin.

2.2 Akt signal pathway

Akt also known as protein kinase B (PKB) is a serine/threonine-specific protein kinase that play an important role in multiple cellular processes such as glucose metabolism, cell proliferation, apoptosis, transcription and cell migration. Moreover, Akt/PKB pathways involved with cellular survival pathway by inhibiting apoptosis process and inducing protein synthesis process. Akt is used the

PI3K/Akt/mTOR pathway and other signaling pathways. There are many actions of Akt in signaling pathways including phospholipids binding, phosphorylation, ubiquitination, lipid phosphatases, and protein phosphatases.

2.3 EGF pathway

The epidermal growth factor pathway (EGF) consists of four related receptors: EGFR (EGFR1, ErbB1); ErbB2 (Neu, HER2); ErbB3 (HER3); and ErbB4 (HER4) that activated by EGF, TGFα, amphiregulin (AREG), betacellulin (BTC), epiregulin (EREG), heparin-binding EGF-like growth factor (HB-EGF), and the neuregulins (NRG1, NRG2, NRG3 and NRG4)(Table1.). The EGF family is a member of receptor tyrosine kinase (RTK). This family is activated by homo- and hetero-dimerization. Individual and specific EGF/neuregulin receptor monomeric proteins contribute to the ligand and response specificity of the EGFR dimer and initiate downstream signaling pathways. EGFR family promotes cell survival, growth and differentiation via the activation of several integrated signaling pathways. These frequently involve the early activation of the key upstream signaling elements: phospholipase C; phosphoinositide 3-kinase (PI3K), Ras-GTPase, and Src kinase. EGFR binds phospholipase C-γ1 through its Src homology 2 (SH2) domain and activates this lipase through phosphorylation. Activation of PLC-γ1 leads to the formation of inositol triphosphate 3 (IP3) and diacylglycerol (DAG). Phosphatidylinositol (3, 4, 5)-triphosphate (PIP3) induces the releasing of the second messenger, calcium ion (Ca^{2+}) from the endoplasmic reticulum (ER), which together with DAG activates various conventional and novel protein kinase Cs. Activation of PI3K results in the activation of the pyruvate dehydrogenase lipoamide kinase isoenzyme 1 (PDK1); Akt/PKB pathway. Phosphorylated EGFR binds the growth factor receptor-bound protein 2 (Grb2): son of sevenless homolog (SOS) complex which activates the p21Ras-GTPase pathway leading to extracellular signal-regulated kinase (ERK) activation. Src kinase can be linked to EGFR through adaptor proteins p66Shc and p52Shc. STAT (signal transducers and activators of transcription factors) proteins (STAT1, STAT3 and STAT5) can be differentially activated directly by the

EGFR tyrosine kinase and/or associated Src kinase. The integration and mitigation of these signaling pathways contribute to the final cell response to EGFR agonists.

2.4 p53 pathway

The p53 pathway, known as protein 53 or tumor protein 53, involve with cell cycle of multi-cellular organisms. This protein has 53 kDa on SDS-PAGE. In a normal cell p53 is inactivated by its negative regulator such mouse double minute 2 homolog (mdm2). Upon DNA damage or other stresses, various pathways will lead to the dissociation of the p53 and mdm2 complex. Once activated, p53 will induce a cell cycle arrest to allow repair and survival of the cell or apoptosis to discard the damaged cell. There are two proteins kinase as targets of p53 protein. First target is mitogen-activated protein kinase (MAPK) family (JNK1-3, ERK1-2, p38 MAPK), which is known to respond to several types of stress, such as membrane damage, oxidative stress, osmotic shock, heat shock, etc. Second target is ATR, ATM, CHK1 and CHK2, DNA-PK, CAK which is implicated in the genome integrity checkpoint, a molecular cascade that detects and responds to several forms of DNA damage caused by genotoxic stress. Oncogenes also stimulate p53 activation, mediated by the protein p14ARF. Unstressed condition, the p53 level is decrease. A protein called mdm2 (a product of p53), binds to p53, preventing its action and transports it from the nucleus to the cytosol. Also mdm2 acts as ubiquitin ligase and covalently attaches ubiquitin to p53 and thus marks p53 for degradation by the proteasome. However, ubiquitination of p53 is reversible. A ubiquitin specific protease, USP7 (or HAUSP), can cleave ubiquitin off p53, thereby protecting it from proteasome-dependent degradation. This is one means by which p53 is stabilized in response to oncogenic insults.

2.5 Matrix metalloproteinase pathway

Matrix metalloproteinase (MMPs) is a protease that secreted by tumor cells. This protein induces the degradation of extracellular matrix and allowing the metastasis of tumor cell. The MMPs is inhibited by reversion-inducing, cysteine-rich

protein with reversion-inducing-cysteine-rich protein with Kazal motifs (RECK), the membrane anchored protein. The RECK is inhibited by RAS.

2.6 VEGF pathway

Vascular endothelial growth factor (VEGF) is the family of secreted polypeptides with a highly conserved in receptor-binding cystine-knot structure similar to that of the platelet-derived growth factors (PDGF). The VEGF involved with angiogenesis and vasculogenesis. Especially, the cancer with overexpression of PDGF has angiogenesis and vasculogenesis . The activation mechanism of VEGF receptor is similar to receptor tyrosine kinase. Dimerization of VEGF receptor induces trans-phosphorylation and activate downstream signal. The most important member of VEGF family is VEGF-A. The receptors of VEGF-A are VEGFR-1 (Flt-1) and VEGFR-2 (KDR/Flk-1). VEGFR-2 appears to mediate almost all of the known cellular responses to VEGF. The function of VEGFR-1 is thought to modulate VEGFR-2 signaling. Another function of VEGFR-1 may be to act as a dummy/decoy receptor, sequestering VEGF from VEGFR-2 binding (this appears to be particularly important during vasculogenesis in the embryo). VEGF-C and VEGF-D, but not VEGF-A, are ligands for a third receptor (VEGFR-3), which mediates lymphangiogenesis.

3. Development of inhibitors of EGFR family

According to the signaling pathway of EGFR family, the signaling pathway of EGFR family can be inhibited by using any molecules that can occupy at the active site or allosteric binding site of involved molecules in EGFR signaling pathway. The inhibiting molecules such as small molecule and antibody are currently used as a binding molecule for design of anti-cancer drug. The targets domain of EGFR molecules are included extracellular domain (ECD) and intracellular domain (ICD) (Figure 1.). The ECD is used as a targeting protein because it involved with initiation of signal transduction. Example of ECD targeting drug is Trastuzumab (Herceptin®),

monoclonal antibody that against ECD of ErbB2 (Carter *et al.*, 1992). Although, Trastuzumab has a great binding affinity to ECD of ErbB2, it provides an adverse effect because it cause of cardiac dysfunction in 2-7% of cases (Miles *et al.*, 2002). The limitations of Trastuzumab are penetration capacity and drug resistance (Fiszman and Jasnis, 2011). As the penetration capacity, this drug cannot used for treatment cancer with brain metastasis because of it cannot penetrate into blood brain barrier (Nielsen *et al.*, 2009).

The second domain of EGFR on anti-cancer drug development is ICD. There are three domain in ICD including juxtamembrane domain, tyrosine kinase domain (TK domain), and C-terminal domain. TK domain is usually used as a target for anti-ErbB2 drug development. This domain plays an important role for EGFR heterodimerization, especially ErbB3 and ErbB4. Moreover, this domain is highly conserved among different receptors and reduce drug resistant due to gene mutation of ECD (Zhang *et al.*, 2006).

Table 2 Small molecules that inhibit TK activity of EGFR family.

Agents	Target kinase (s)	Phase of development
ZD1839 (Iressa ^{R})	EGFR	FDA approved
OSI774 (Tarceva ^{R})	EGFR	FDA approved
CI10333 (Canertinib ^{R})	EGFR/ErbB2/ErbB3/ErbB4	Phase II
PKI166	EGFR	Phase I
EKB569 (Pelitinib ^{R})	EGFR/ErbB2	Phase II
GW572016 (Lapatinib ^{R})	EGFR/ErbB2	Phase III

Despite the EGFR-TK inhibitors can increase the survival rate of patient with metastasis cancer (Efferth, 2012), the cancer become resistant to the EGFR-TK

inhibitors. The resistance mechanisms of cancer to EGFR-TK inhibitors are including secondary mutation that interferes with drug binding, molecular cross-talk that create the alternative pathway for signal transduction, and polymorphism in *EGFR* and EGFR pathway related genes that interrupt the pharmacokinetic of drug (Lin and Bivona, 2012; Murray *et al.*, 2012; Oxnard, 2012).

4. Antibody and its diversity

Antibodies are synthesized by B lymphocyte. B lymphocyte produced the various types of antibodies. The antibodies compose of two heavy and two light chains. There are two types of light chain including kappa and lambda. Kappa light chain is found predominantly in the mouse; both are equally used in the human, only lambda chains are found in the chicken. There are five types of heavy chain: μ, δ, γ (4 subtypes), ε and α. The type of heavy chain defines the class of immunoglobulin: IgM, IgD, IgG, IgE and IgA, respectively. Each heavy or light chain is composed of a relatively invariable constant part and a variable part. However, there are also hypervariable regions interspersed throughout each chain. All of the highly variable sequences are combined form the antigen binding site in the three dimensional structures.

4.1 Assembly of kappa (κ) light chain

A mature kappa light chain is assembled from approximately 300 variable (V) region genes, 5 junction (J) region genes, and 1 constant (C) region gene. The mature chain has the VJC structure - i.e. one variable region gene is joined to one junction region gene and to the single constant region gene. The joining of one of the V region genes to one of the J region genes proceeds by a non-homologous recombination strand exchange reaction which is directed by conserved repeated sequences that flank each V and J region. Each V gene is followed by a 7 bp pair conserved sequence and a 9 bp conserved sequence; these two sequences are 23 bp apart - i.e. two turns of the DNA helix. Each J gene is preceded by the same two conserved sequences except that they are separated by 12 bp - i.e. one turn of the

DNA helix. The conserved sequence elements next to the V genes can either be in inverted orientation or in direct orientation with respect to the elements next to the J genes. Either way, they can form a stem-loop structure by complementary base-pairing. The formation of such a stem loop brings a V gene and a J gene next to one another. The gene recombination of V and J genes and deletion of intervening DNA involved with recombination activation gene 1 (RAG1) and recombination activation gene 2 (RAG2). There are 5 different places at which breakage and rejoining can occur. When this happens, however, a single continuous reading frame must be maintained through the joined V and J genes. This breakage and rejoining adds diversity to the joined molecules since the place of breakage and joining will affect the identity of the two codons and ultimately the two amino acids that fall in this region. This recombination reaction involves inverted repeats but it does not result in any inversion of the intervening DNA. This recombination does not involve homologous strand pairing between these regions. In this reaction, the inverted repeats serve only as a signal for binding of the appropriate enzymes coded by the RAG1 and RAG2 genes. The resulting rearranged genome is now expressed. The final step in expression of a mature immunoglobulin chain requires that the RNA be spliced so that the final mRNA codes for a contiguous VJC polypeptide.

4.2 Assembly of lambda (λ) light chain

The lambda light chains are coded by fewer genes: 2 variable region genes, 4 junction region genes and 4 constant region genes.

4.3 Assembly of heavy chain

The assemblies of heavy chains are similar to that of light chains assembly. Although, they contain an additional region called a diversity (D) region, which is located between the variable region and the junction region. Heavy chain genes are organized similarly to the kappa light chain genes. There are approximately 200 variable chain genes, 12 diversity region genes, 4 junction region genes, and 8 constant region genes. Heavy chains are assembled in manner similar to that of the

kappa light chains. Initially a D region gene is joined to a J region gene. This mechanism of this joining also involves the recognition of 7 bp and 9 bp conserved sequences. The intervening DNA is discarded. In addition, a few nucleotides are randomly inserted between the D and J region by terminal deoxynucleotidyl transferase. Subsequent to DJ joining, V(D)J joining occurs - by the same mechanism already outlined for kappa chains. Following this, transcription occurs and mRNA splicing generates the final mRNA which will be translated into an immunoglobulin chain.

Finally, it has been found that B cells carry out the random somatic mutation of the hypervariable region. This increases the possibility of finding an antibody to any antigen.

4.4 Class switching

Initially, the constant chain that is expressed will be a class mu (μ) chain since that is the first constant region gene to be encountered. B lymphocytes make IgM class antibodies at first, switching to a different class depending upon the antigen of the purpose for which the antibody is needed. This change requires another site-specific recombination event later in the development of a B-lymphocyte. This will bring the appropriate constant region near to the V(D)J joined genes by removing the intervening constant region genes. Expression of the final heavy chain, however, will still require mRNA splicing to create an mRNA that has a joined V(D)JC coding sequence.

5. Heavy chain antibody fragment (VH/V_HH)

The antibody or immunoglobulin is the bio-molecule that used by immune system to identify and neutralize the foreign objects such bacteria and viruses. The antibody is secreted by the plasma cell. The antibody is known as immunoglobulin (Ig), because it contains structures that commonly found in many proteins. These structures refer as heavy chain (55 kDa) and light chain (25 kDa). The foreign object

of the immune system called antigen. The antibody used paratope to bind with epitope of antigen. The antibody can recognize the many structures of antigen including linear structure and conformational structure. The structure of antibody is similar to Y shape and has 150 kDa. The structure of antibody is composed of heavy and light chain. These structures are held together by disulfide and non-covalent bond. According to the function of antibody, there are two important domains such as fragment crystallizable (Fc) and fragment for antigen binding (Fab). The Fc is the domain that contains both of heavy and light chain. The function of Fc domain is responsible for the biological activity of antibody. This activity is the binding of involved cells such as phagocytes, mast cell, and basophils. Moreover, this domain also activate the complement system for eliminate the pathogen. The Fab domain involved with antigen binding activity of antibody. This domain consists of hypervariable region or complementarity-determining regions (CDRs) and framework residues. There are three regions in CDRs regions including CDR1, CDR2, and CDR3. Each CDR contains 10 amino acids. The CDR regions provide specificity of the antibody. For framework residues, they contain intervening residues between CDR regions. These residues are invariant and show little difference in amino acid resides between chains. The antibody is a useful tool in molecular biology. It can be used as a diagnostic antibody and therapeutic antibody. The antibody can recognize antigen of interest with different specificity. The large number of antibody that provide a different specificities and epitope affinities are known as polyclonal antibodies (PAbs). The PAbs are made by several different immune cells. For antibodies that provide specificity to one epitope and made by identical immune cell are known as monoclonal antibodies (mAb). According to the application of antibodies, there are many formats of antibodies such as immunoglobulin, variable domain of heavy chain (VH), variable domain of light chain (VL), single chain variable domain antibody fragment (ScFv), variable domain of *camelidae* (V_HH or nanobodies), and variable domain of cartilaginous fish (V-NAR) (Kontermann, 2010).

There are many methods for antibody construction including *in vivo* technique and *in vitro* technique. The hybridoma technique was first developed in 1975 by Kohler and Milstein (An, 2010). This technique is used to produce antibody by using

tumor cells exploitation. The hybridoma technique is *in vivo* technique for production of antibody. The tumor cells, myeloma, is fused with spleen cells that challenge with relevant antigen. The fusion cells, hybridoma, multiply rapidly produce a large amount of antibodies that against antigen of interest (Figure 3).

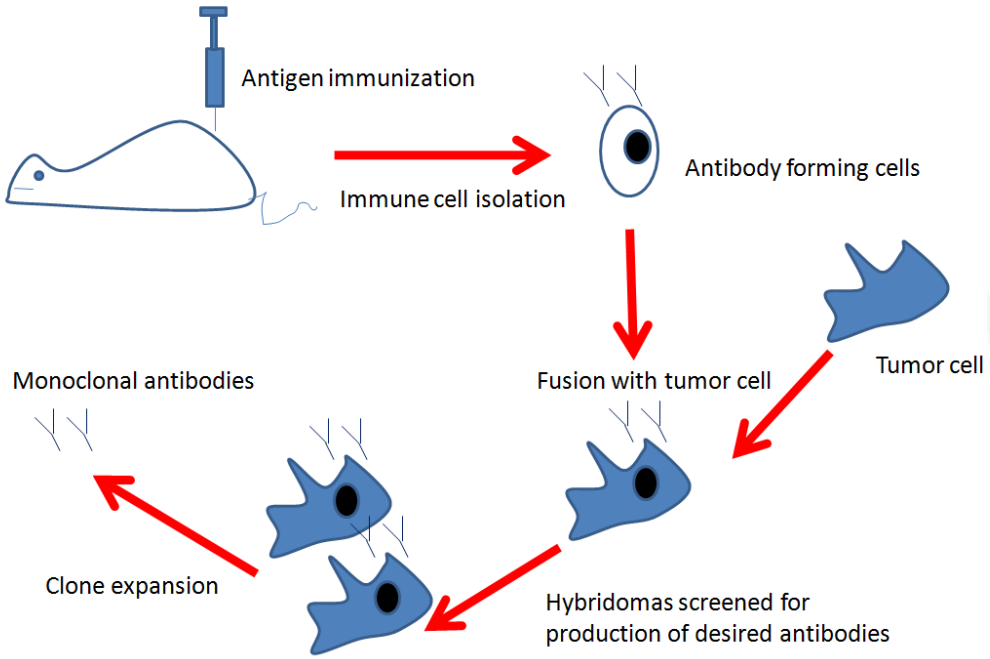


Figure 3 Production of monoclonal antibody by using hybridoma technology.

The recombinant technology is *in vitro* technique that applied for antibody production and antibody selection. For the antibody production, the gene encoded antibody is amplified and introduced into the host cell including prokaryotic and eukaryotic cell. According to antibody selection, the antibody can be isolated from recombinant antibody libraries *in vitro* by using the platform selection. There are several display methods that used for selection process including ribosome and mRNA display, microbial cell display, and phage display.

The phage display was firstly introduced by Smith (1985). Phage display is the one of the method for proteins expression, including antibodies, proteins, or peptides on the surface of filamentous bacteriophages (or simply called phages) of *E. coli, e.g.* M13 and fd. The protein that expressed on the surface of phage is used as the library for antibodies selection. The selection process called bio-panning. The bio-panning can be adapted to many specific conditions that cannot perform in most other display platform, including selection on whole cell, tissue, and even in animal with high affinity and specificity. The essential discoveries for antibody phage display technology are composed of two important steps. First step is the demonstration of foreign DNA inserted into filamentous phage gene III (g3) that is expressed as a fusion protein. Second step is the successful expression of functional antibody fragment in periplasmic space of *E. coli* (Skerra and Pluckthun, 1988). The periplasmic space is located between the two cell membranes of the bacterium. The advantage of antibodies that located in periplasmic space is the correctly folding. By recombinant DNA technology, the large repertoire of antibody, estimated about 10^{11}-10^{12} diversity in human's body, can fusion with filamentous phage gene III protein (g3p) and subsequently displayed on the surface of phage particles in the form of a heavy chain antibody fragment (VH/V_HH). The VH/V_HH consists of variable domain of heavy chain. Antigens of interest can be used to select antigen-specific phages from the *in vitro* antibody repertoire (Bio-panning). The selected phage can be used to produce VH/V_HH molecule for further used (Clackson .*et al.*, 1991).

There are many advantages of monoclonal antibody from ScFv and VH/V_HH. Monoclonal antibody from ScFv and VH/V_HH provide a large number of antibody genes, effective selection of the correct gene from gene pool, and improvement in the affinity and specificity of a selected antibody fragment. In addition its can be saved disadvantage from hybridoma technique such as unstable hybridoma lines or primary fusion cultures, by rescuing the antibody gene, low affinity to low immugenicity substances, hapten, self-antigen and high toxicity substance by using monoclonal antibody from ScFv and VH/V_HH. The important advantage of VH/V_HH is penetrating activity into the binding pocket. Due to the size of structure, VH/V_HH can

reached into the binding pocket that never be reached by conventional monoclonal antibody or ScFv antibody (Desmyter *et al.*, 2002).

The gene pool of antibodies can be accessed by using the polymerase chain reaction (PCR). The using primer complement with the termini of the desired antibody gene and can be used to amplify the portion of the inner DNA. The amplicon is not only amplified from individual gene but also entire of gene families. The amplicon of heavy and light chain are linked into a single protein by using a peptide of 15-18 amino acid namely peptide linker. Four numbers of glycine and one serine are composed of the repeating unit of peptide linker. The constructed antibodies gene is incorporated into expression vector. The diversity of bacterial antibodies depends on the efficiency of transformation method. The best diversity is about 10^{10} different antibody genes.

6. Camelid immunoglobulins

Camel immunoglobulins are devoid of light chain and are composed of heavy chain homodimer. These called heavy chain antibody (HCAbs), are expressed after V-D-J rearrangement and required dedicated constant γ-gene. Since the binding of antigen of HCAbs, it comprise in one single domain, refer to as variable domain of heavy chain of HCAbs (V_HH) or nanobody (Nb). The heavy chain of HCAbs is composed of three instead of four globular domains. The two constant domains are highly homologous to Fc domain of classical antibody. The CH1 domain is missing in HCAbs. Moreover, the Fab is reduced to single variable domain in HCAbs. This variable domain is referred to V_HH. The VH and V_HH domains adopt an immunoglobulin (Ig) fold of two beta sheets, one of four strands and one of five. Within this fold, the three hypervariable regions, clustering at one end of the domain, are located in loops that connect the beta strands where they participate in antigen recognition. The alignment of the VH and V_HH amino acid sequences immediately reveals that the first and third hypervariable regions of the V_HH are more extended than those of VH's, and in dromedary but not in llama, the CDR1 and CDR3 often contain a Cysteine. From crystal structures of V_HH we know that, the Cysteine form a

disulfide bond that possibly assists in shaping the loop structure. Another striking distinction between VH and V_HH occurs in the framework (FR)-2 region. This region is normally highly conserved in VH domains and populated by hydrophobic amino acids that serve as an anchoring place for the VL domain. In contrast the corresponding region in a V_HH contains more hydrophilic amino acids, which explains the absence of VL association and the soluble behavior of V_HH as a single-domain entity. The V_HH has lager paratope size than classical antibody due to the extension of CDR1 and CDR3 loop.

The dromedary genome contains a cluster of VH and V_HH germline genes, followed by a pool of D genes, a pool of J genes and the C genes including μ, γ, ε, α the dromedary genome contains multiple Cg genes, some of which are dedicated to constitute the conserved part of the H-chain of HCAbs whereas other Cg genes are employed to produce the H chain of classical Abs The dromedary genome contains about 50 VH and 40 V_HH germline genes. The V_HH germline genes have their Cysteine encoded in the CDR1 region, and the characteristic amino acid substitutions within the V_HH-FR-2 region.

7. Antibodies phage display system and bio-panning

7.1 Structure of filamentous phage

The bacteriophages or phages in short**,** consist of 11 genes (Figure 4). The important structures of phage are capsid and coat protein. The phages express 2,700 copies of gene 8 protein (g8p or pVIII). This protein is a major capsid protein. In addition, the phage express 3-5 copies of gene 3 protein (g3p or pIII) that is a one of three minor coat protein on tip of phages. The phages infect a variety of gram-negative bacteria by using pili as a receptor such as F pili of phage strain M13, f1, and fd.

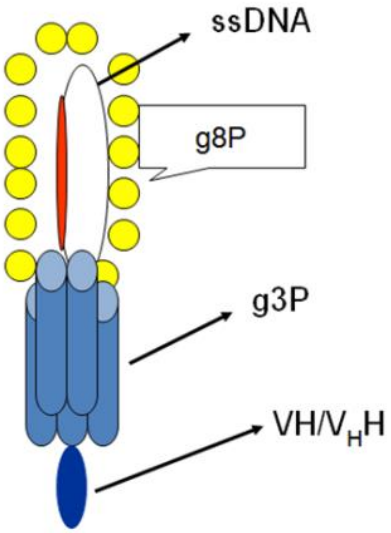


Figure 4 Structure of filamentous phage containing VH/V_HH. The blue eclipses represent the structure of VH/V_HH.

7.2 Life cycle

The life cycle of phage starts from the attachment of phage g3p to the f pilus of the male *E. coli* (e.g., *E. coli* TG1) without the lytic infection. After the attachment, the circular single-stranded DNA (ssDNA) of phage is converted to double-stranded plasmid like replicative form (RF) by using host's DNA replication machinery. The RF undergoes rolling circle replication to make ssDNA and also serves as template for g3p and g8p production. Phage progeny is assembled a protein coat by using gene from helper phage. After phage assembly process, phage extrude through bacterial membrane. In recombinant phage, the g3p fusion protein displays on the tip of M13 phage (Figure 5).

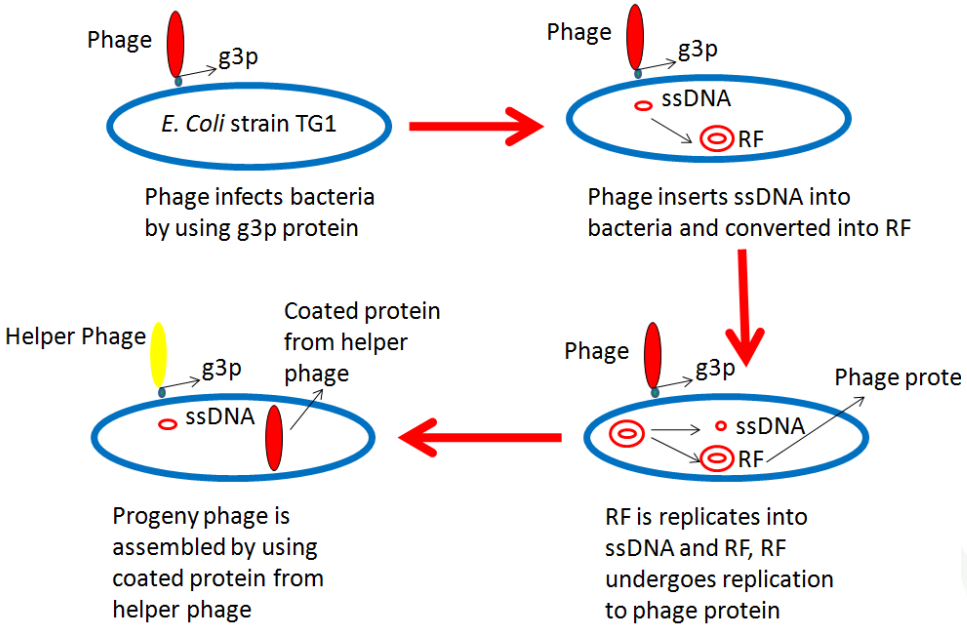


Figure 5 Life cycle of filamentous phage. Phage uses the host DNA replication machinery to produce progeny phage.

7.3 Phagemid cloning vector

Phagemid is the most popular vector for phage display. It is a hybrid of phage and plasmid vector. The one of all DNA in phagemid vector derived from plasmid-like RF of M13. M13 phage contains 2 forms of DNA including ssDNA and double-stranded DNA (plasmid-like RF). The ssDNA is prepared from phage medium and used for DNA sequencing. Double strand DNA is isolated from infected bacteria and used for cloning of target fragment. Phagemid consists of original of replication for both M13 phage and *E. coli*. In addition, it contains multiple cloning site and antibiotic-resistance gene. The example of phagemid cloning vector is pCANTAB5E (Figure 6). However, phagemid cloning vector lack of non-structural and structural gene is required for a generating complete phage. Therefore, progeny phage requires all structural protein from helper phage to generate complete phage. The helper phage is phage that contains a slightly defect at origin of replication (such as M13KO7 and VSCM13). The process that added helper phage into phagemid for structural protein adding is called phage rescue. The ration of polypeptide-pIII fusion protein:wild type

pIII may range between 1:9 and 1:1000 depending on the type of phagemid, growth condition, the nature of polypeptide fused to pIII, and proteolytic cleavage of antibody-pIII fusions.

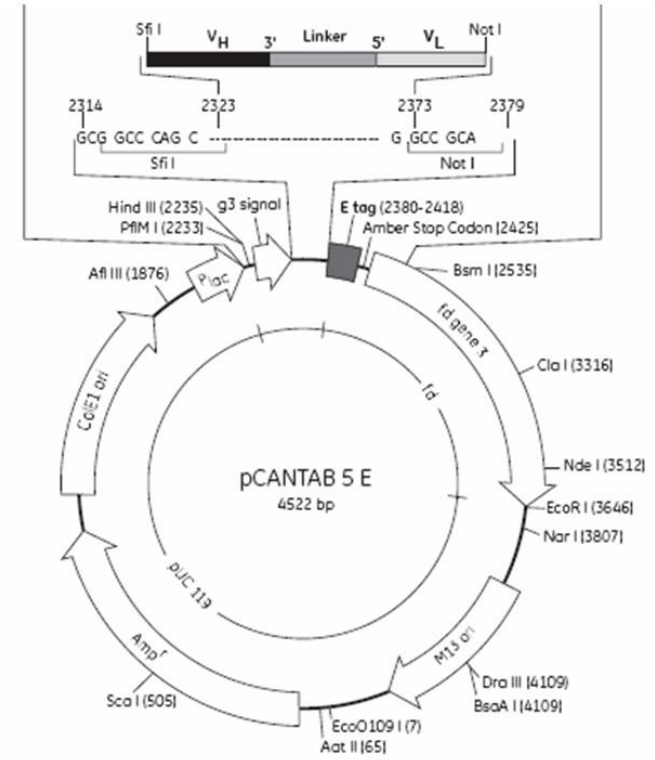


Figure 6 Circle map of phagemid cloning vector, pCANTAB5E (GE healthcare, USA)

7.4 Types of phage display library

7.4.1 Non-immune or single pot library

The single pot library is constructed from naïve animals. This library is used to produce monoclonal antibodies against antigen especially with low immunogenic antigen. In addition, the library is used to against high toxicity and self-antigen that cannot perform in conventional technique (Nissim *et al.*, 1994). Constructed small-size human single-pot library at $3x10^7$ different antibody clones and used to produce antibodies against self-antigen, hapten, and foreign antigen. Moreover, constructed a large size library (1.4×10^{10} clones) from over 40 non-immunized human donors and produced antibodies with binding affinities (*Ka*) in the low nanomolar range (Vaughan *et al.*, 1996).

7.4.2 Immune library

The immune library is constructs from immunize animal or human, that is a different point from single pot library. The characteristic of immune library including the enrichment of antigen-specific antibodies, the affinity maturation of some antibodies resulted from immune responses, and class-switching of antibody isotype which is the characteristic of the memory immune response. The example of immune library such as constructed human library from vaccinated donor, the library has been used for against human B cell lymphoma (Shen *et al.*, 2007). In addition, there is the production of immune library from animal such as mice (Andersen *et al.*, 1996), chicken (Yamanaka *et al.*, 1996), and rabbits (Lang *et al.*, 1996).

7.4.3 Synthetic library

The synthetic library is constructed from mutation of comple--mentary determining regions (CDRs) by using mutagenesis or PCR based technique. The advantage of this library is including, obtain vast diversity of library. There are six CDRs is used to mutate especially CDR3 of the heavy chain (VH-CDR3) that is the most central to the antigen-binding site of all CDRs and is the most diversity loop in composition and length (Chothia and Lesk, 1987).

7.5 Construction of VH/V_HH phage

The VH/V_HH is constructed from dromedary (*Camelus dromedarius*). The structure of VH/V_HH is contained only VL region. The gene fragment encoded VH/V_HH is amplified by using human-immunoglobulin specific primer. The variable region of heavy chain (VH) -encoding DNA sequences is amplified from cDNA of B-cell. The cDNA is reverses transcribed from mRNA of B lymphocytes. The VH/V_HH coding-DNA sequence is inserted into phagemid vector such as pCANTAB5E. The recombinant *VH/V_HH*/phagemid vectors are introduced into competent *E. coli*. by using $CaCl_2$ transformation or electroporation. The ligation and transformation step influences the size of the library. The antibody phage display library is prepared from *E. coli.* transformants by using helper phage (M13VCS or KO7). This process is known as phage rescue, to yield recombinant phages which display VH/V_HH antibody fragment as a fusion to one of the phage coat protein. The quality of VH/V_HH phage depend on number of clone that contain *VH/V_HH* insert, number of clone expressing phages carrying VH/V_HH, and number of clone expressing soluble VH/V_HH. The quality of VH/V_HH phage can be accessed by a variety method including PCR screening, Western blot analysis, and DNA fingerprinting.

8. Selection of antigen-specific antibodies from the phage display library of antibody: "bio-panning"

Bio-panning process is used to enrich antigen specific to phage clone. The library of antibody is incubated with immobilize or mobilize antigen of interest. After incubation step, the unbound phage is removed with washing buffer. The remaining library which specifically binds the antigen is eluted with elution buffer (Figure 7). The selected antibodies are introduces into *E. coli* for specific phage amplification. The bio-panning should be performed in one time. In fact, there is the binding limitation of specific phage then several rounds bio-panning are necessary. There are four types of bio-panning.

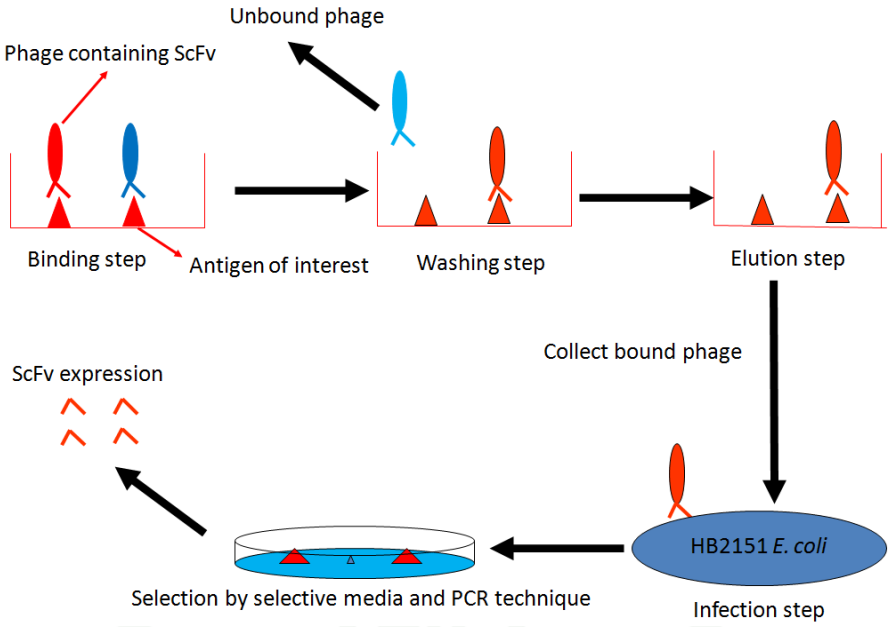


Figure 7 Process of bio-panning for ScFv selection.

8.1 Immobilize bio-panning

The bio-panning using immobilize antigen is used to select specific phage library that recognizes immobilize antigen. The interest antigen is trapped by affinity column (Clackson *et al.*, 1991; McCafferty *et al.*, 1991), plastic surface such as

immunotubes (Maxisorb® tubes; Nalge Nunc Intl., Naperville, IL), enzyme link immunosorbant assay (ELISA plate) (Kang *et al.*, 1991; Marks *et al.*, 1991) or BIA® core sensor chips (Malmborg *et al.*, 1996). The important aspect of immobilize bio-panning is conformation of immobilize antigen. The conformation of antigen is associated with the phage antibody recognition. The specific phage can be eluted with acidic solution (HCl or glycine buffer) (Kang *et al.*, 1991; Roberts *et al.*, 1992), basic solution (triethylamine) (Marks *et al.*, 1991), enzyme cleavage of protease site constructed between the antibody and g3p (Ward *et al.*, 1996), and competition with excess antigen (Clackson *et al.*, 1991).

8.2 Antigen in solution bio-panning

The advantage of antigen in solution bio-panning is accurate quantification of the antigen used during selection (Hawkins *et al.*, 1992) and consequently enhance the ability to use lower concentration of the antigen to favor selection of high affinity phage antibodies. The conformational epitope of antigen can be recognized by phage library in this bio-panning. The biotinylated antigen and allows binding with library. The specific antibodies are recovered with avidin or streptavidin coated paramagnetic beads. Disadvantage of this bio-panning is only isolation of anti-streptavidin antibodies. However, this problem can be solved by a depletion step using streptavidin-coated bead.

8.3 Cell bio-panning

The phage library selection on cell surface marker is performed on monolayer of adherent cell or cell suspension. Unbound phage can be washed by rinsing tissue culture flask (monolayer) or centrifugation (cell suspension). To optimize specificity of antibodies, a simultaneous positive and negative selection may be applied (Kruif *et al.*, 1995). In this approach, the antigen-negative cells "absorber" is used to exclude other surface marker-specific phage in competitive manner. The target cells can be isolated from the absorber by labeling specific antibodies and subsequently sorting by Fluorescence-activated cell sorting (FACS).

8.4 *In vivo* bio-panning

This technique is using direct injection of phage repertoires into animal. After injection, tissue are collected and examined for phage bound to tissue specific endothelial cell markers as was demonstrated for peptide phage. Phage-display peptide to selective vascular beds *in vivo* was first isolated by Pasqualini and Ruoslahti (1996). The advantages of this technique are better specificity of phage library to an intact target and the elimination of phage library that recognize ubiquitous plasma and cell surface protein in an inherent blocking step.

9. Production of soluble VH/V_HH molecule

Antigen-specific phage is introduced into specific strain of *E. coli* such as HB2151 and TG1 *E. coli*. This transformant *E. coli* produces soluble antibodies which conjugated with E-Tag protein at C-terminal region. Characteristic of expressed protein is depends on isolated clone. The expressed protein can be secreted into various parts of bacterial cells including, cytoplasm and periplasm of bacterial cell. The expressed antibodies can be purified by using general protein purification including ion-exchange chromatography, gel filtration chromatography, and affinity chromatography (Anti-E Tag column).

10. **Evolution of therapeutic antibody**

Therapeutic antibody is used for cure a wide range of diseases. Therapeutic antibody is antibody that can bind specifically to target cells. Moreover, this antibody can stimulate patient's immune system to attack those cells. There are many types of therapeutic antibody including polyclonal antibody, monoclonal antibody, single chain Fv, and etc. In 1890, first therapeutic antibody was introduced by Berhring (1890). This antibody is polyclonal antibody that against tetanus and diphtheria toxin. However, this antibody has a disadvantage including batch to batch variation and virus contamination risk. In 1975, hybridoma technology was used to produce monoclonal antibody. The advantage of monoclonal antibody is high specificity to

antigen. However, monoclonal antibody can increase immunogenicity after repetitive long term application. Although, the monoclonal antibody is specific to antigen of interest, the monoclonal antibody can induce hypersensitivity reaction. Chimeric antibody was used to solve hypersensitivity problem. The chimeric antibody is immunoglobulin that have different source of antigen binding site such as murine. This antigen binding site is graft onto human constant region. The Chimeric antibody was introduced by Sahagan (1985). However, there are the disadvantages of chimeric antibody such as labor intensive, not always feasible, and hypersensitivity risk. In 1991, antibody phage display library was used as a therapeutic antibody by Clarkson (1991). however, this technology provide high affinity antibody without hypersensitivity reaction, and it require development of affinity maturation. Another type of antibody display technology is antibody ribosomal display technology. Ribosomal display is cell-free system, offering advantages of screening larger libraries and continuously expending new diversity during selection. Moreover, production of displayed derived antibodies can be achieved by choosing one of a variety of prokaryotic and eukaryotic cell-based expression system. A ribosomal display technology was introduced by Hanes (1998).

In 1991, there are two therapeutic antibody technologies were introduced, affibodies and nanobodies. Affibodies is small alpha helical domain of SPA (Staphylococcal surface display of immunoglobulin A). This technology was initiated by Hansson (1991). For nanobodies, this antibody is single domain antibody from camel heavy chain antibody. Nanobodies were introduced by Muyldermans (1991). In 2001, Anticalins was first initiated by Skerra (2001). Anticalins is a mimetic antibody. These antibodies are artificial proteins that are able to bind the antigen including protein or small molecule. They are derived from human lipocalins which are family of natural binding protein. The anticalins were smaller than monoclonal antibody about eight times, they can use in lieu of monoclonal antibody. Designed ankyrin repeat protein (DARPINs) are antibody mimetic proteins that have high affinity target protein binding. They derived from ankyrin, adaptor proteins that attach the integral membrane protein to the spectrin-actin based membrane skeleton. The DARPINs were introduced by Kawe (2008). The newest antibodies mimetic technology is

monobodies that initiated in 2007 by Olson (2007). Monobodies also known as acnectin. They are smaller than immunoglobulin G for fifteen times. The structure of monobodies is base on human fibronectin that similar to antibody variable domain.

11. Tyrosine kinase assay

The activity of tyrosine kinase is important for a variety of cellular pathways and processes. The function of tyrosine kinase enzyme is transfer the phosphate group from adenosine triphosphate (ATP) to the target protein. This action can activate of inactivate the cellular function. This activity can be investigated by using many techniques such as scintillation proximity assay (SPA), homogenous time-resolved fluorescence resonance energy transfer (HTR-FRET), fluorescence polarization (FP), and fluorescence spectroscopy technique (Sills *et al.*, 2002). For the Fluorescence spectroscopy technique, the Antibody Beacon™ Tyrosine kinase assay kit (A-35725) detects the complex of small molecule tracer ligand that exhibits quenched fluorescence when bound to anti-phosphotyrosine antibodies. The Antibody Beacon™ Tyrosine kinase assay kit (A-35725) exploits the affinity of phosphopeptide. This peptide provides a high affinity of binding and displaces the small-molecule tracer ligand. After the release of small-molecule tracer ligand, the ligand increases the fluorescence intensity. This assay provides many advantages including real-time monitoring of tyrosine kinase assay, using unlabeled peptide substrates, and detection the efficiency of tyrosine kinase inhibitor. The principle of this assay was shown in Figure 8.

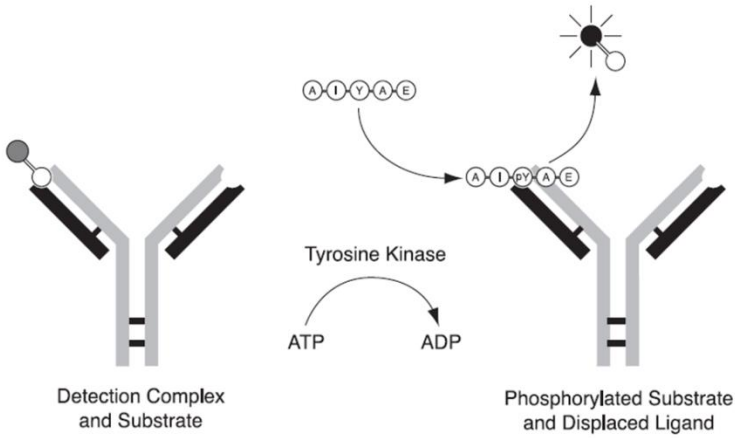


Figure 8 Antibody Beacon tyrosine kinase reaction assay scheme. The detection complex is composed of antibody and small-molecule tracer. After the presence of phosphotyrosine containing peptide that resulting from the reaction of tyrosine kinase, the ligand is rapidly displaced and there is an increase in fluorescence intensity.

According to fluorescence response, the complex of phosphorylated substrate and displaced ligand exhibited a 4-flod enhancement over the fluorescence of the antibody Beacon complex in buffer alone.

12. Protein structure

The structure of protein can be divided in four levels of structure. The primary structure refers to amino acid linear sequence of polypeptide chain; this structure is held together with covalent or peptide bond. There are two terminals of structure including amino terminus (N-terminus) and carboxyl terminus (C-terminus). The secondary structure refers to highly local sub-structure; this structure is held together with covalent bond and hydrogen bond (H-bond). There are two types of secondary structure including alpha helix and beta strand (Pauling *et al.*, 1951). The alpha helix and beta strand are held together with the intra chain H-bond and inter chain H-bond respectively. Moreover, several sequential secondary structures can form a supersecondary structure. The example of supersecondary structure includes beta

hairpin and helix-turn-helix motif. The tertiary structure refers to three dimensional structure of single protein molecule; this structure is held together with non-specific hydrophobic interaction, water bridges, salt bridges, and disulfide bond. The last level is quaternary structure; this structure refers to three dimensional structure of multi-subunit protein. The stabilized interaction of this structure is the same interaction with tertiary structure. The levels of protein structure were shown in Figure 9.

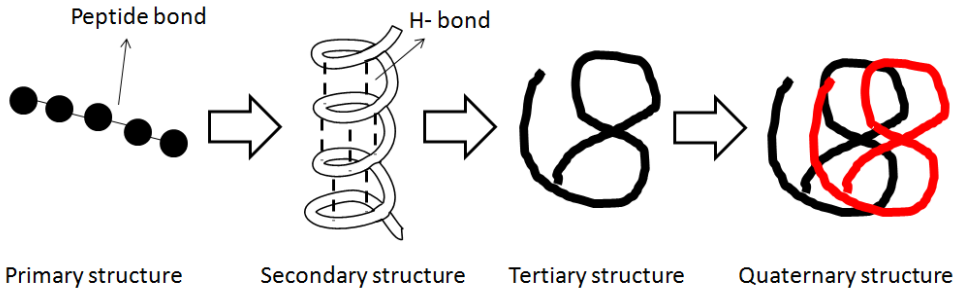


Figure 9 Four levels of protein structure: the black bead in primary structure refers as amino acid residue. The primary structure is held together with peptide. The secondary structure was shown in white helix and held together with H-bond. For the tertiary structure, the non-specific interaction is used for stabilizing the single protein molecule (black line). The quaternary structure is used the same interaction with tertiary structure; the black and red line represent the single domain of protein molecule.

According to different level of protein structure, there are many appropriate methods for structural determination.

12.1 Primary structure

The interaction which involved with primary structure is peptide bond. The method that used to solve the primary structure is protein sequencing. The protein sequences are achieved by using mass spectrometry and Edman degradation reaction. Mass spectrometry is an analytical technique that measured mass to charge ratio of charged particle. For Edman degradation, the amino acid at terminal residue is labeled

and cleaved from the peptide without disrupting the peptide bond between other amino acid residues.

12.2 Secondary structure

The interactions that build secondary structure are hydrogen bonds between amino acid residues including inter and intra chain of protein. Intra-chain interaction is involved with helix structure. There are many types of helix structure including 2_7, 3_{10}, alpha helix, and π helix. For inter-chain interaction, the beta-pleated sheet structure is involved with this interaction. There are two types of beta-sheet including parallel and anti-parallel. Moreover, the secondary structure is built into higher structure called motif. The motif structure can build by using random coil structure. The example of motif structure is helix-turn-helix. Method that used for investigate secondary structure is experimental method and computational method. The experimental methods for studying protein structure are protein nuclear magnetic resonance NMR, circular dichroism (CD), x-ray crystallography. For the computational methods, energy landscape of protein folding and homology modeling are used as studying methods.

12.3 Tertiary structure

Tertiary structure is spatial arrangement of amino acid that is far apart in the linear sequence as well as those residues that are adjacent. This structure involve with interaction of amino acid in three dimensions. The three dimensional structure is maintained not only by hydrophobic interactions, but also by electrostatic force, hydrogen bonding, and covalent disulfide bond. The method that used for determine tertiary structure is similar to investigation method of secondary structure.

12.4 Quaternary structure

Quaternary structure differs from other structures because it containing more than one polypeptide chain, it exhibits a fourth level of protein structure. This

level of structure refers to the spatial arrangement of the polypeptide subunits and the interaction between them. The involved interaction may be covalent link (disulfide bonds) or non-covalent interaction such as electro static forces, hydrogen bonding, and hydrophobic interaction. For structural investigation method, the experimental such as X-ray crystallography, protein NMR, and circular dichorism (CD) are used. The computational structural method, energy landscape of protein folding and homology modeling are used as studying methods.

In conclusion, the primary protein structure is investigated at peptide bond, meaning protein sequencing. For secondary structure, tertiary structure, and quaternary structure, the interaction of these levels of protein structures are involved protein folding.

13. Molecular interactions of protein structures

The interactions between molecules are used when the more than one molecule are held together. There are two major types of molecular interaction including covalent and non-covalent interactions.

13.1 Covalent interaction

The covalent interaction is involved with primary structure and tertiary structure of protein via peptide bond and disulfide bond respectively. The covalent bond is the interaction that exploit the sharing a pair of electrons between two neighboring atoms. The stabilization of covalent interaction depends on the proper number of electron in electron shell.

13.1.1 Peptide bond

A peptide bond is the covalent bond between the carboxyl group of one molecule reacts with amino group of other molecule and releasing the water molecule. This bond also provides the resonance effect that resulting from partial

double bond of peptide bond. The resonance of the peptide bond is the distribution of electron between carbonyl oxygen, carbonyl carbon, and amide nitrogen. Moreover the resonance effect of peptide bond also provides the increase of polarity of peptide bond. This bond can be broken by amide hydrolysis and release 2-4 kcal/mol. The length of the peptide bond is 1.33 Å (Martin, 1998).

13.1.2 Disulfide bond

A disulfide bond is the covalent bond that occurs when there is the cysteine in molecule. The thiol (or namely sulfhydryl) group of cysteine two molecules provide an oxidized S-S linkage. The most stabilized angle of this bond is either +90° or -90°. This bond can be broken by reductants such as beta-mercaptoethanol and release 60 kcal/mol (Thornton, 1981). According to the property of disulfide bond, the molecules with this bond are highly reactive molecule. The length of the this bond is 2.05 Å

13.2 Non-covalent interaction

A non-covalent interaction or called weak interaction is the interaction between ions, molecules, and part of molecules. This interaction important for stabilized the structure of molecules and weak enough to be broken and reforming in the dynamic of molecular interplay. This interaction is 10-100 times weaker than a covalent bond. There are five non-bonding interactions including electrostatic interaction, dipole interaction, van der waals interaction, hydrogen bonding, and hydrophobic interaction. A non-covalent interaction is an interaction that important for interaction among bio-molecules because of the distance of interaction.

13.2.1 Electrostatic interaction

Electrostatic interaction or called Coulomb interaction, ionic interaction, charge-charge interaction. This interaction occurs between electrically charged particle. The character of electrostatic interaction can be explained via the

Coulomb's law states. This states that *the magnitude of the electrostatic force of interaction between two point charges is directly proportional to the scalar multiplication of the magnitude of charges and inversely proportion to the square of distance between them.* The equation of states was shown below

$$E_{(elec)} = \frac{q_i \cdot q_j}{4\pi \cdot \varepsilon \cdot r_{ij}}$$

$E_{(elec)}$ represents the strength of electrostatic interaction (kJ/mol) between two charges, q_i and q_j represent the quantity of charges and divided by the distance between them r_{ij} and the dielectric constant ε of the medium. According to the equation, the magnitude of electrostatic force depends on the distance between two molecules and type of charges. Electrostatic interaction shows no involvement of direction of molecule.

13.2.2 Dipole interaction

Dipole is a separation of positive and negative charge. The molecules that involved with dipole interaction have no net charge and provide the distribution of charge. Any configuration of charge is called dipole moment. The direction and distance of involved molecules are important for dipole interaction. Moreover, a dipole molecule is often linear rather than spherical in shape and its dipole moment is a directional interaction that depends on the distribution of electron. There are four types of dipole interaction including charge-dipole interaction, dipole-dipole interaction, charge-induce dipole interaction and dipole-induce dipole interaction.

13.2.2.1 Charge-dipole interaction

This interaction occurs between charge molecule and dipole molecules. The dipole molecule attracts the charge molecule because of the

distribution of its electron. An example of charge-dipole interaction is the interaction between the NH_3^+ group of the arginine/lysine side chain and the O atom of a water molecule. The strength of charge-dipole interaction is proportional to $1/r^2$

13.2.2.2 Dipole-dipole interaction

This interaction occurs between two water molecules or polar side chain amino acid and water molecules. This interaction involves with the dielectric constant of involved molecules. This interaction is an important interaction for the folding of protein causing from the involvement of water molecule. The strength of dipole-dipole interaction is proportional to $1/r^3$

13.2.2.3 Charge-induced dipole interaction

This interaction occurs between non-polar molecule and charge molecule. The charge molecule induces the dipole moment of non-polar molecule. The strength of charge-induced dipole interaction is proportional to $1/r^4$. An example of charge-induced dipole interaction is the interaction between the NH_3^+ group of arginine/lysine and the phenyl ring of phenylalanine (cation-π interaction).

13.2.2.4 Dipole-induced dipole interaction

This interaction occurs between non-polar molecule and polar molecule. The polar molecule induces the dipole of non-polar molecule. The strength of dipole-induced dipole interaction is proportional to $1/r^5$. An example of dipole-induced dipole interaction is the interaction between the H-atoms of a water molecule and the phenyl ring of phenylalanine.

13.2.3 Van der Waals force

Van der Waals force or London dispersion force occurs between non-polar molecules without dipole moment. The fluctuation of electron

initiates the temporally polarization and attracts the involved molecules. The equation of Van der Waals interactions are represented by a potential energy as a function of a distance *r* that includes both the attractive and repulsive forces. This energy is known as the Lennard-Jones 6,12 potential.

$$E_{(vdW)} = 4\varepsilon \cdot \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]$$

The ε and σ are characteristic energy and bond-length parameters which depend on types of interacting atoms. As the equation, the size and volume of molecule influenced Van der Waals force. The large molecule provides more polarization than small molecule due to the area of fluctuation of electron. Example of Van der Waals contacts are the interaction between among amino acids with aliphatic side chains and the interaction between two phenyl rings of phenylalanine/tyrosine/tryptophan (π-π interaction).

13.2.4 Hydrogen bonds

Hydrogen bonds occur between hydrogen atom and electronegative atom such as nitrogen, oxygen, and fluorine. After the covalently linked of hydrogen atom and electronegative atom, decentralizes the electron density cloud of hydrogen atom was occur and induce the hydrogen atom becomes partially positive. When a group with this positive charge density (called hydrogen donor, D) attracts lone pair electrons on another strong electronegative atom (called hydrogen acceptor, A), then hydrogen bonding occurs. Because of a large charge density of H, making it being strongly positive, hydrogen bonding can be described as a consequence of the electrostatic interaction between the partial positive charge of H and the negative charge of A, and is usually denoted as $^{\delta-}D - H^{\delta+} \cdots {}^{\delta-}A$. This interaction can occur between molecules (intermoleculary), or within the different part of a single molecule (intramoleculary). The strength of hydrogen bonds depend on electronegativity of acceptor, electonegativity of donor, angle of involved atoms,

and distance of involved atoms. The figure of angle and distance of hydrogen bond were shown in Figure 10.

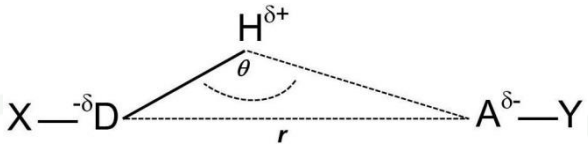


Figure 10 Configuration of hydrogen bonds. X, D, H, A, and Y represent binding atom of hydrogen donor atom, hydrogen donor atom, hydrogen atom, hydrogen acceptor atom, and binding atom of hydrogen acceptor atom. θ and r represents angle and distance of interacting atoms.

The distance r usually ranges from 2.5 - 3.5 Å and angle θ ranges from 140-180°. As the character of hydrogen bonds, these bonds are a very important in biological molecule.

13.2.5 Hydrophobic interactions

Hydrophobic interactions occur between water and hydrophobes (low water soluble molecule). The hydrophobes are non-polar groups that no react with water molecule. Thermodynamic can explain the formation of hydrophobic interactions. When hydocphobes come together and interact with each other, the enthalpy is increased resulting from the broken of hydrogen bonds. Moreover, the broken of hydrogen bonds are the cause of increase of entropy. As the thermodynamic equation, the negative Gibb's free energy indicates that the interactions are spontaneous reaction. The hydrophobicity of any molecule (amino acid) can be a measurement of the change in ΔG^0. Hydrophobic interactions are relatively stronger than other weak interaction. The strength of hydrophobic interactions depends on three factors including temperature, number of carbons on hydrophobes, and the shape of hydrophobes. For temperature factor, the increase of temperature also increases the strength of interaction. However, the improper temperature, hydrophobic

interactions will be broken. Molecules with greatest number of carbons on hydrophobes have strongest hydrophobic interactions. For the shape of hydrophobes, the aliphatic and linear carbon chains provide the strongest hydrophobic interactions.

14. Bioinformatics

Bioinformatics or computational biology is the application of computer sciences and allied technology that used for answer the question of biology. This term is first introduced by Paulien Hogeweg and Ben Hesper in 1978 (Hogeweg, P. and B. Hesper, 1978) for the study of information processes in biotic systems. The commonly software that used for bioinformatics including Java, XML, Perl, C, C++, Python, R, MySQL, SQL, CUDA, MATLAB, and Microsoft Excel. There are two interfaces of software tools such as command line interface (CLI) and graphic user interface (GUI). The bioinformatics tools range from open source bioinformatics software, licensed bioinformatics software and web service software. There are two main topics about computational approach. First topic is nucleotide that including DNA and RNA. Second topic is Protein. The computational approach for nucleotide including, DNA and RNA mapping and analyzing, DNA and RNA sequence alignment, DNA and RNA modeling. Computational approach for protein is similar to nucleotide. Moreover, molecular interaction is used for bioinformatics. The molecular interaction study is an essential component of many chemical, physical, and biological studies. This study is together growth with protein structural study techniques including X-ray crystallography, NMR, and electron microscope. Moreover, the capacity of computer plays an important role for the growth of this study. The molecular interaction study is used molecular modeling, molecular docking and molecular dynamic simulation as tools. The molecular interaction studies are the studying interactions among proteins, ligands and peptides.

15. Protein modeling

Protein modeling or structure prediction is the computational method for prediction of protein structure. The computational method for protein modeling provide many advantage over structural information from X-ray crystallography and nuclear magnetic resonance (NMR) such as time consuming, technical complexity, and size consideration. There are four methods for protein modeling such as homology modeling or comparative modeling, fold recognition or threading method, *ab intio* method that utilize knowledge-based information, and *ab intio* method without the aid of knowledge-based information.

15.1 Homology modeling or comparative modeling

Homology modeling or comparative modeling predicts the structure base on the functional relationship of homologous proteins. The homologous proteins share similar structures, resulting from the evolutionary conservation of protein structure. According to amino acid sequence alignment, homology modeling provides most accuracy when sequence alignment is most identity. However, there is a large structural bias towards the template structure. The limitation of this technique is errors in sequence alignment rather than errors in structural prediction (Zhang and Skolnick, 2005). The examples of homology modeling software include MODELLER and SWISS-MODEL.

15.2 Fold recognition or threading method

Fold recognition or threading method, the method that scans the amino acid sequence of interest against the database of solved structure. The scoring function is used to access the compatibility of the sequence to the structure. The difference between fold recognition and homology modeling involves the compatibility analysis. Fold recognition is known as 3D-1D fold recognition due to its compatibility analysis between 3D structure of solved structure and 1D structure of sequence of interest (Bowie *et al.*, 1991). According to the evaluating method, this

method is performed by using inverse folding search method, thus prediction harder than homology modeling technique. Moreover, this technique is suitable for low identity amino acid. The examples of protein threading software include HHpred, RAPTOR, phyre, MUSTER, and SPARKX.

15.3 *ab intio* protein modeling

ab intio protein modeling or *de novo* protein structure prediction, the method that used to build the three dimensional structure of protein from amino acid sequence. The computer attempts to construct the three dimensional structure of protein by mimic protein folding method and stochastic method. The evaluation of folded protein depends on the free energy of structure. This technique require the vast computational resource, thus only can be predicted a tiny protein. According to the computational requirement, this technique is suitable for distributed computer project such as Folding@home, Rosetta@home, and The human proteome folding project (Dill *et al.*, 2007).

16. Molecular docking

Molecular docking is a key tool in structural molecular biology and computer-assisted drug design (Morris and Wilby, 2008). There are two aims of molecular docking: accurate structural modeling and correct the prediction of activity. Indeed, the molecular docking is used computer to predict orientation and pose of one molecule to other molecule, and form a stable complex. The principles of molecular docking calculation are based on *lock and key* model and *'induce fit'* models. The descriptions of these principles are rigid body docking and flexible docking respectively. The major protocol of molecular docking is composed of generation the many docked structures or searching space of possible poses and conformations, filtration of the given structures, are and evaluation the structures by using scoring function. For There are two influence factors about docking calculation such as compatibility of geometry and chemistry of involved molecules. The involved docking molecules can be divided into receptor and ligand. The receptor molecule

refers as the molecules that provide the binding area or binding pocket for ligand. Basic algorithms for molecular docking include rigid and flexible docking algorithms. For the flexible docking, the computer generates the probability of conformational changes of involved molecules. There are three types of flexible docking algorithms such as systematic search algorithms, stochastic search algorithms, and deterministic search algorithms. The scoring function that used to determine the generated poses includes knowledge-base methods, force field base scoring method, and empirical method. The molecular docking can be applied to binding site identification, drug discovery via *virtual screening technique,* and prediction of molecular function. Moreover, the molecular docking also applied with *bioremediation* (Suresh *et al.*, 2008). Bioremediation is the use of microorganism metabolite to remove pollutants. For bioremediation, the molecular docking is used to determine properties and capacity of microorganism enzyme to bind with pollutants. There are many success stories of molecular docking for computational drug design (Kolb *et al.*, 2009).

16.1 Rigid body docking

The rigid body docking is the molecular docking that treats both ligand and receptor target as the rigid bodies. There are two factors that influent with searching space including poses and conformations. For the calculation of rigid body docking, there are six degree of freedom of relative position and orientation. Moreover, there is no the calculation of rotation bond including the bond angles, bond lengths and torsion angles. Although this docking method provides many advantages such as very fast calculation, lack of realistic consideration of bond angles, bond lengths and torsion angles. The first docking program was introduced by Kuntz in 1982 (Kuntz *et al.*, 1982). Next time, the other rigid docking programs for receptor and small molecule were launched including CLIX (Lawrence and Davis, 1992), FLOG (Miller *et al.*, 1994), FRED (McGann *et al.*, 2003) and MS-DOCK (Sauton *et al.*, 2008). The example of rigid body molecular docking program for investigate the protein-protein interaction are including ZDOCK (Chen and We, 2002), GRAMM (Vakser, 1997) and also protein-DNA/RNA rigid docking studies using HADDOCK (Dominguez *et al.*, 2003).

16.2 Flexible body docking

The flexible body docking is the molecular docking that treats both ligand and receptor target as the flexible bodies or treats one ligand or receptor target as the flexible body. There are three factors that influent with searching space including poses, conformations, and consideration of bond. For the calculation of flexible body docking, there are six degree of freedom of relative position and orientation like rigid body docking. In addition, this docking method also calculates bond consideration of receptor and ligand. . An example of this docking method is the incremental construction algorithms, also known as *'the anchor-and-grow'* (Ewing *et al.*, 2001; Kramer *et al.*, 1999). The ligand is divided into rigid region, *'anchor'*, and flexible parts, *'side chains'*. The anchor is the first part of ligand that docked into binding pocket. After the docking of anchor, the flexible parts of ligand were added into anchor and perform structure determination. The best structure is used as a ligand for next docking calculation. According to searching algorithms, they can be divided into stochastic search algorithms and Deterministic search algorithms. For the flexible docking, there are many software were launched including DOCK (Ewing *et al.*, 2001), FlexX Placement of medium-sized molecular fragments into active sites of proteins., ADAM (Mizutani *et al.*, 1994), Hammerhead (Welch *et al.*, 1996), and SLIDE (Zavodszky *et al.*, 2002).

16.2.1 Stochastic search algorithms

Stochastic search algorithms are the non-deterministic finding algorithms. These algorithms generate the structure by changing one degree of freedom of the system at a time. The term of stochastic is based on the theory of probability. For stochastic search algorithms, it includes Monte Carlo method and evolutionary algorithms.

16.2.1.1 Monte Carlo methods

The Monte Carlo methods are algorithms that rely on repeated random sampling to generate their structures. The Monte Carlo methods were first coined in 1940s. The Monte Carlo casino is used as a name of this method. This method is useful for calculate the system with many coupled degree of freedom. The molecular docking application, the structure is generally minimized and determined result is based on a Boltzmann probability. The repeating cycles is moving on until reach the setting number of pose. There are many programs that used Monte Carlo methods such as Prodock, ICM, MCDOCK, PRO_LEADS and AutoDock.

16.2.1.2 Evolutionary algorithms

The evolutionary algorithms are the population-based search algorithms that inspired by biological evolution such as reproduction, mutation, recombination, and selection. The different of evolutionary algorithms and biological evolution is genotype-phenotype distinction. The artificial evolution lacks of complex process that involved with the expression of phenotype. There are many evolutionary techniques *e.g.* genetic algorithms, genetic programming, and mimetic algorithms. The most popular type of evolutionary algorithms is genetic algorithms. For genetic algorithms, the data structures are kept as a binary code (0 and1). The algorithms evolved the data structure rely on the possible solution of genetic operators (mutation rates, crossover rate and number of evolutionary rounds). These algorithms select the fitness structures and combine them together. The combination structures are determined by using scoring function. As the evolutionary protocol, the fitness value and the final results are depend on the structural evolution. The advantage of these algorithms is provides an easy ascent toward a global optimum. Moreover, these algorithms also provide multi-modal optimization (Hiroyasu *et al.*, 2011). The molecular docking program with evolutionary algorithms includes AutoDock and GOLD.

16.2.2 Deterministic search algorithms

The deterministic search algorithms are algorithms that provide same output resulting from particular input. These algorithms are predictable algorithms. The example of deterministic approach was shown below.

$$2 + 2(x) = 4$$

$$X = 1$$

These algorithms provide an ascent toward a local optimum. According to the obtained structure, they are local optimum structure or global optimum structure of search space. . Examples of these algorithms are energy minimization and molecular dynamics simulations. These two approaches cannot cross easily energy barriers. Commonly, the expected poses restricted to local search around the stating pose. Therefore, these algorithms appropriate for the refinement step of intermediate pose.

16.2.2.1 Scoring functions

One of most crucial aspect of molecular docking is how to evaluate the binding score for docking complex, referred as scoring functions problem. According to scoring functions, the method that used to discriminate the correct docking complex structure from incorrect ones. The ideal scoring function should provide accuracy of binding affinity calculation and ranking. The parameters that influence with scoring functions include free energy perturbation (FEP) and thermodynamic integration (TI). For scoring functions calculation, time consuming and accuracy are the important concerning when deal with large datasets. Thus, the selection of appropriate class of scoring function is important for docking calculation. There are three major classes of scoring function including forcefield-based scoring functions, empirical-based scoring functions, and knowledge-based scoring functions.

(a) Force field-based scoring functions

Force field-based scoring function or first-principle scoring functions is calculated based on term from molecular mechanics forcefield which examined only non-bonded interaction energies (Van der Waals and electrostatic potentials) between receptor and ligand. Moreover, the intramolecular energies of each involved molecule also frequency include for calculation. These scoring functions use the atomic parameters from molecular mechanics forcefield including AMBER and CHARMM, force fields in order to estimate the binding affinity. The examples of forcefield-based scoring function include D-score, G-score, GoldScore, DOCK, and AutoDock.

(b) Empirical-based scoring functions

Empirical-based scoring functions or Regression-based approaches are fit to reproduce the experimental data including binding energies, binding poses, and sum of several parameterized functions. These scoring functions were first introduced by Böhm (1992). The design of these scoring functions based on the experimental data can be predicted by the sum of individual correlation terms. The determination of scoring functions was obtained from regression analysis of experimentally binding energies and crystallographic structures data. The advantage of these scoring functions is fast and manipulated in various docking algorithms. The examples of empirical-based scoring function include LUDI, F-score, ChemScore, SCORE, Fresno, and X-SCORE.

(c) Knowledge-based scoring functions

Knowledge-based scoring functions is calculated based on the statistics information and reproduce experimental structure rather than binding energies of complex. For calculation protocol, the computer generates the complex by using relative simple atomic interaction-pair potential from crystallographic protein-ligand complex in Protein Data Bank (PDB). According to

these scoring functions, the binding affinity depends on information implicitly that encoded in set of interactions that found in the database. The advantage of these scoring functions is computational simplicity which permits screening of large compound database. The examples of knowledge-based scoring function include PMF, DrugScore, and SMoG.

17. Antibody-antigen docking

The Antibody-antigen docking or protein-protein docking is a computational technique that used to investigate the interaction between antibody and antigen. Given data from docking analysis are including docking pose, docking energy, and interacting residue. There are two algorithms for docking such as rigid body docking and flexible body docking. As the complexity of antibody structure, the rigid body docking is proper method for antibody-antigen docking because this method provides computational simplicity and extremely fast method. The protein-protein docking techniques are available as standalone computing unit and web-based computing unit. For standalone computing, there are many programs for docking such as 3D-Dock, HEX, GRAMM, PPD, DOT, BIGGER, MERL, DARWIN, and ZDOCK. In addition, the programs that are designed for protein-ligand docking may be useable in protein-protein docking. The examples of these programs are DOCK, AutoDOCK, and FlexX.

18. Virtual screening

Virtual screening (VS) or *in silico* screening is the computational screening technique that used in drug discovery research (Rollinger *et al.*, 2008a). This screening technique deals with a large set of chemical structures (also called library). Nowadays, there is enormous chemical space (search space), over 10^{60} conceivable compounds (Bohacek *et al.*, 1996). The examples of libraries are NCI, Chemiebase, and ACD. There are two types of virtual screening including ligand-based and structure-based virtual screening (Morris *et al.*, 2009). For the ligand-based, this technique involved with pharmacophore, an abstract description of molecular feature

which are impact for molecular recognition of a ligand by a receptor. The ligand based virtual screening find the biological molecules which can bind tightly with target molecules. The major type of ligand-based virtual screening is the 2D and 3D similarity searching. For structure-based virtual screening, this technique involved with the docking of candidate ligand into the target of interest. The major type of structure-based virtual screening is the molecular docking. This method required the detail of 3D structure of target protein binding site to arrange the ligand compound by their likelihood. There are major attractions of virtual screening including; less time consuming, less cost, and computational simplicity, when compared with the conventional high-throughput screening (HTS) which has to perform in the test tube or well-plated

MATERIALS AND METHODS

Materials for experiment 1

1. Chemicals

1.1 Antibody Beacon™ Tyrosine Kinase Assay Kit (A-35725): Invitrogen, USA

1.1.1 Oregon Green 488 ligand, Antibody Beacon reagent for phosphotyrosine detection (Component A), 200 µL of a 2.5 µM solution in 1X reaction buffer

1.1.2 Anti-phosphotyrosine antibody, P-Tyr-100 (Component B), 200 µL of a 5 µM (0.75 mg/ml) solution in pH 7.4 buffer containing 5 mM sodium azide

1.1.3 2X Tyrosine kinase reaction buffer (Component C), 25 mL of 100 mM Tris-HCl, 20 mM $MgCl_2$, 2 mM EGTA, 0.02% Brij 35, pH 7.5 (at 22°C)

1.1.4 Tyrosine kinase substrate # 1 (Component D), 400 µL of a 10 mg/ml poly (Glu:Tyr) solution, 4:1 ratio, in dH_2O

1.1.5 Dithiothreitol (DTT) (Component F), 31 mg

1.1.6 Adenosine triphosphate (ATP) (Component G), 200 µL of a 100 mM solution in buffer

1.1.7 Reference phosphotyrosine-containing peptide (Component H), 50 µl of a 500 µM solution of phospho-pp60 c-src (521.533) in dH_2O; sequence = TSTEPQpYQPGENL; MW = 1543.7

1.2 2–mercaptoethanol: Merck, Germany

1.3 Acrylamide: Bio Basic, Canada

1.4 Adenosine triphosphate (ATP): Fermentus, USA

1.5 Agar bacteriologico americano: Pronadisa, Spain

1.6 Ammonium persulfate: Ajax Finechem, Australia

1.7 Antibiotics (ampicillin)

1.8 Bis-acrylamide: Bio Basic, Canada

1.9 Bovine serum albumin (BSA): Fluka Biochemika, USA

1.10 Coomassic brilliant blue R-250: Bio Basic, Canada

1.11 D-glucose

1.12 Epidermal Growth Factor Receptor (EGFR) human:Sigma-Adrich, USA.

1.13 GeneJet™ Gel extraction kit: Fermentus, USA

1.14 Isopropyl β-D-1-thiogalactopyranoside (IPTG): Fermentus, USA

1.15 Magnesium chloride ($MgCl_2$): Ajax Finechem, Australia

1.16 Methanol: Ajax Finechem, Australia

1.17 Plasmids (pCANTAB5E): Amersham-Pharmacia biotech, USA

1.18 Restriction enzyme (*Sfi*I and *Not*I): Fermentus, USA recombinant protein of human HER2/ErbB2 kinase: Sino Biological Inc. , China

1.19 Recombinant protein of human HER2/ErbB2 kinase: Sino Biological Inc. , China

1.20 Sodium chloride: J. T. Baker, Malaysia

1.21 Sodium dodecyl sulfate (SDS): Bio Basic, USA

1.22 Tetramethylethylenediamine (TEMED): Bio Basic, USA

1.23 Tris (hydroxymethyl) aminomethane: Research Organic, USA

1.24 Triton® X-100: USB, USA

1.25 Tryptone type-I: Himedia, India

1.26 Tween® 20: USB, USA

1.27 Yeast extract powder: Himedia, India

2. Instruments

2.1 Autopipette: Gilson, Germany

2.2 Balance (4 digits): Denver balance

2.3 Biosafety Cabinet Class II: Microtech, USA

2.4 Centrifuge: Hermle refrigerate centrifuge model Z383K, Hermle Labortechnik GmbH, Germany

2.5 Fluorescence spectroscopy: LS-50B Luminascence Spectrosopy, Perkin Elmer, USA

2.6 Freezer -20 °C

2.7 Heat Box: D1100, Labnet, USA

2.8 Hybond™-ECL™ Nitrocellulose membrane (Amersham Biosciences, UK)

2.9 Incubator 26°C: Model i250, Accuplus, Thailand

2.10 Microplate reader: TECAN, Austria

2.11 Nano drop UV/Visible spectrophotometer ND-1000: Nanodrop Technologies, USA

2.12 Oven: Fisher Scientific, USA

2.13 PCR Thermal cycler TC-28/4 : Lio-Labinter, USA

2.14 pH meter: Seven Easy pH meter, Mettler toledo, USA

2.15 Power supply: AE 8750, ATTO, Japan

2.16 Refrigerator 4°C

2.17 SDS-PAGE electrophoresis: Bio-RAD CA, USA

2.18 Shaker VS-8480SFN : Meditop, England

2.19 Sonicator

2.20 UV detector: ECONO UV monitor, Bio-Rad, USA

2.21 ELISA plate (E.I.A/R.I.A. 8 Well strip Flat bottom): Corning, NY USA

Methods for experiment 1

1. ErbB2-TK immobilization

The purified recombinant protein of human HER2/ErbB2 (Lys676-Val1255) kinase was purchased from Sino Biological Inc. (Beijing, P.R. China) (Catalog number 10004-H20B1). The ErbB2-TK recombinant protein was analyzed by using SDS-PAGE. The ErbB2-TK recombinant protein was used as antigen to select the antibody from phage display library. This phage display library was obtained from collaborators (Department of Parasitology, Faculty of Medicine Siriraj Hospital, Mahidol University). For antigen immobilzation process, the 1 μg of recombinant protein was dissolved in 50 μl of coating buffer (Carbonate-bicarbonate buffer, pH 9.6). The mixture was immobilized on an ELISA plate. The plate was incubated in moist chamber at 37ºC for 18 hours. After incubation period, the mixture was discarded. Four hundred milliliter of blocking buffer (5% skim milk in PBS) was added into the well and incubated in moist chamber for 30 minutes. The blocking buffer was discarded and washed the plate with PBS-T (phosphate buffer saline containing 0.1% of Tween 20). Four hundred milliliters of PBS-T was added into the well and incubated for 5 minutes as a washing step. After incubation period, PBS-T was discarded. This step was repeated for 3 times.

2. Selection antibody against ErbB2-TK by using bio-panning technique

The twenty-five microliter of VH/V_HH library was diluted with PBS-T to make 25% v/v final concentration. The 100 μl of diluted phage (~1 x 10^{11} colony forming units; cfu) was added into the ErbB2-TK coated well and incubated in moist chamber at 37ºC for 2 hours. After incubation period, the mixture was collected as unbound phage. The following step is a washing step. The well was extensively washed with PBS-T. The washing step was repeated for ten times.

3. Transformation of *E. coli* strain HB2151 with bounded phage

To collect the phage binding ErbB2-TK, forty microliters of log phase *E. coli* strain HB2151 was added into the ELISA well. The ELISA well was incubated in moist chamber at 37ºC for 30 minutes. After incubation period, bacterial was added into 900 microliters of Luria-Bertani (LB) medium containing ampicillin. The medium was incubated at 37°C for 30 minutes. After incubation period, the cell pellet will be collected by centrifugation. The cell pellet was spread into LB broth containing ampicillin and glucose (LB-AG) agar plate. The agar plate was incubated at 37ºC for 18 hours.

4. Screening of VH/V_HH-expressing HB2151 *E. coli* clone

The VH/V_HH-expressing HB2151 *E. coli* clone was determined by using polymerase chain reaction technique (PCR). The bacterial colonies were used as a template for PCR analysis. The oligonucleotide primers used in the PCR for *VH/V_HH* amplification were: pCANTAB5-*R1* primer (5′-CCA TGA TTA CGC CAA GCT TTG GAG CC-3′) that bound specifically to pCANTAB5E vector closed to the 5′ end of *Sfi*I restriction endonuclease site, and a reverse nucleotide primer designated pCANTAB5-*R2* primer (5′-GCT AGA TTT CAA AAC AGC AGA AAG G-3′) that bound specifically to pCANTAB5E vector closed to the 3′ end of *Not*I restriction endonuclease site. The PCR mixture containing 8.4 µl of ultrapure distilled water (UDW), 1.25 µl of PCR buffer (10 x *Taq* DNA polymerase buffer with KCl), 18.75 µM of $MgCl_2$, 0.1 µM of dNTPs mix (2.5 mM for each), 0.5 µl of 3’ forward primer, 0.5 µl of 5’ reverse primer, and 0.25 u of *Taq* DNA polymerase. The mixture was performed PCR under the following condition. The condition is 94ºC for 5 minutes, 30 cycles of 94ºC for 1 minute, 50ºC for 1 minute, and 72ºC for 90 seconds then 72ºC for 10 minutes and hold at 4ºC. The PCR product was analyzed by agarose gel electrophoresis and ethidium bromide staining. The expected length of *VH/V_HH* was 600 base pairs.

5. Agarore gel electrophoresis

Agarose gel electrophoresis was used to analyze the PCR product. The 1% agarose gel was prepared by adding 0.3 g of agarose powder into 30cc of 1x TAE buffer. The mixture was completely dissolved by microwave oven heating. It was heated until the mixture became clear and homogeneous. The mixture was allowed to cool to approximately 50-60ºC before poured it into the gel tray. The gel was allowed to harden for 45 minutes. PCR product was mixed with 10x sample buffer (0.25% bromophenolblue, 0.25% xylenecyanole FF, and 50% glycerol) at 1x final concentration. The mixture was filled to the gel and electrophoresis was performed at constant 100V until the bromophenol blue dye front had arrived about 2/3 of the gel length. The gel was incubated with ethidium bromide buffer (0.5µg /ml ethidium bromide in 1x TAE buffer) for 15 minutes. After incubation, the gel was removed excess ethidium bromide by incubate with 1 x TAE for 10 minutes. The nucleic acid pattern on the gel was directly visualized under ultraviolet (UV) light.

6. Expression of VH/V_HH

The *E. coli* strain HB2151 that provided *VH/V_HH* was picked and inoculated into 10 ml of LB-AG (100 µg/ml of ampicillin and glucose) broth at 37ºC for 18 hours. The five-hundred microliter of incubated culture was inoculated into fresh LB-A broth and incubated at 37ºC, 250 rpm in shaking incubator until optical density at 600 nm equal to 0.5. The Isopropyl-1-thio-β-D-galactopyranoside (IPTG) was added into the broth to final concentration at 1 mM. The broth was incubated at 37ºC, 250 rpm in shaking incubator for 5 hours. The *E. coli* cell was collected by centrifugation at 4,000 x *g* for 10 minutes at 4ºC. The whole cell lysate was prepared by sonication. The bacterial cell pellet was sonicted by using ultrasonic-homogenizer LABSONIC® P (Satorius AG, Germany). The amplitude and the cycle of sonication are 30% and 0.5 cycles respectively. The sonication step is performed for 10 minutes on ice-bath. After sonication, whole cell lysate was centrifuged at 12,000 x *g* for 10 minutes at 4ºC. The supernatant was collected as soluble part. The pellet was collected as insoluble part and solubilized by boiling with 10% SDS. Both soluble and insoluble

parts were subjected to SDS-PAGE, Western blot analysis, and affinity column purification.

7. SDS-polyacrylamide gel electrophoresis

The SDS-polyacrylamide gel electrophoresis was used Laemmli buffer system. A 0.75 mm-thick polyacrylamide gel was prepared on 10x10.5 cm glass plates. A 12% separating gel solution containing 0.65 ml of acrylamide stock solution (46.5% w/v acrylamide and 3% w/v N, N-methylene bis-acrylamide) was mixed with 1.25 ml of 1.5 M Tris-HCl/SDS buffer, pH 8.8, 1.52 ml of deionized distilled water (DDW), 0.05 ml of 10% SDS, 25 µl of 10 % ammonium persulfate and 2.5 µl of N, N, N', N'-tetramethylenediamine (TEMED). The mixture was filled into setting glass plates. Thereafter, the top of the gel was overlaid with the water and left it to polymerize for 1 hour. Following complete polymerization, the water was removed by rinsing. The 4% stacking gel containing 2 ml of acrylamide stock solution (46.5% w/v acrylamide and 3% w/v N, N-methylene bis-acrylamide), 1.25 ml of 0.5 M Tris-HCl/SDS buffer, pH 8.8, 3.05 ml of DDW, 0.05 ml of 10% SDS, 25 µl of 10 % ammonium persulfate and 5 µl of TEMED was poured on top of the resolving gel. The gel was allowed to polymerize for 45 minutes before the comb was removed. Protein sample was prepared under reducing and non-reducing condition. Protein samples were mixed at a ratio of 1:2 with tris-tricine sample buffer (0.1M Tris base, 24% v/v glycerol, 2% w/v SDS, 2% v/v β-mercaptoethanol). The mixture was boiled before loaded into 1 µg/well. Prestained SDS-PAGE standards Broad Range (Bio-RAD CA, USA) was run in parallel with the protein samples as a size reference. Electrophoresis was performed by using electrophoresis buffer (0.1 M Tris, 0.1 M Tricine, 0.1% SDS, pH 8.25) and The Mini-PROTEAN® 3 cell electrophoresis system (Bio-RAD, CA, USA) at constant current 20 mA per plate. Electrophoresis was stopped when the tracking dye reached the bottom of the gel. The gel was removed from the electrophoresis cassette. Following step is gel incubation with transfer buffer (25mM Tris, 192 mM glycine and 20% (v/v) methanol) at 25°C with gentle agitating for 5 minutes.

8. Western blot analysis

In this experiment, the soluble and insoluble proteins on SDS-PAGE gel were transferred onto Hybond™-ECL™ Nitrocellulose membrane (Amersham Biosciences, UK) (NC membrane). The protein transfer experiment was performed at constant 100 V for 90 minutes. The transferred proteins were stained with Ponceau S solution (Sigma, USA) by briefly soaking the blotting membrane in the solution ad washing with distilled water until the background became clear. After Ponceau S staining, the empty sites on the NC membrane were blocked by soaking in a blocking buffer [3 % BSA in 0.01 M phosphate buffer saline (PBS), pH 7.4] at 25°C with gentle agitating for 1 hour. The nitrocellulose blot was then washed 3 times for at least 15 minutes with washing buffer [phosphate buffer saline, pH 7.4 containing 0.05% Tween-20 (PBST)] to remove the excess BSA. After washing, the NC membrane was reacted with 1:3,000 diluted HRP-conjugated Nickel probe with gentle agitating at 25°C for 1 hour. After washing, the color reaction was equilibrated in 1/15 M PB, pH 7.6 and the color was developed by 2, 6-dichlorophenolindophenol. The reaction was stopped by washing the membrane with DW and the membrane was allowed to air-dry.

9. Immunostaining of VH/V_HH antibody

The immunostaining of VH/V_HH antibody is used to determine the expression of the antibody by using the specificity of mouse anti-E tag monoclonal antibody (Amersham Biosciences, USA). The binding of VH/V_HH antibody and mouse anti-E-tag was revealed by using goat anti-mouse immunoglobulins-HRP conjugate (Southern Biotech, USA), and substrate, respectively. In this experiment, the NC membrane containing VH/V_HH antibody was incubated with blocking buffer (3% skim milk in PBS pH 7.4) for 30 minutes. The membrane was washed three times with washing buffer (0.05% of Tween 20 in PBS pH 7.4 [PBS-T]). After washing step, the membrane was incubated with 1 ml of VH/V_HH solution containing 500 µl of whole cell lysate and 500 µl of PBS-T was added and incubated at 25ºC for 1 hour. The membrane was washed three times with washing buffer and incubated with

horseradish peroxidase (HRP)-conjugated anti-E tag diluted in PBS-T at dilution 1:400. After incubation step at 25ºC for 1 hour, the membrane was washed three times with PBS-T. The membrane was incubated with 1/15 PB at 25ºC for 10 minutes. The membrane was incubated with mixture of 2, 6-dichlorophenolindophenol (DCIP) and hydrogen peroxide for 5 minutes until color was observed. The excess color was removed by soaking with DW and the membrane was allowed to air-dry.

10. Indirect ELISA

The one microgram of ErbB2-TK recombinant protein was coated on an ELISA plate (E.I.A/R.I.A. 8 Well strip Flat bottom Corning, NY USA) as previously described in bio-panning step. For elimination of non-specific binding signal, the one microgram of skim milk was coated on an ELISA plate as previously described in bio-panning step. After coating step, the unbound protein was discarded and the 300 µl of blocking buffer (3% skim milk in PBS pH 7.4) was added into the well and incubated at 25ºC for 30 minutes. After taking out blocking buffer, the well was washed three times with washing buffer (0.05% of Tween 20 in PBS pH 7.4 [PBS-T]). After washing step, the 100 µl of cell lysate from VH/V_HH-expressing HB2151 *E. coli* and HB2151 *E. coli* cell lysate protein were used as test group and control group, respectively. The antibody solution containing 50 µl of whole cell lysate and 50 µl of PBS-T was added and incubated at 25ºC for 1 hour. After incubation, the well was washed for three times with PBS-T. The ErbB2-TK-bound antibody was detected by using horseradish peroxidase (HRP)-conjugated anti-E tag diluted in PBS-T at dilution 1:5,000. After incubation step at 25ºC for 1 hour, the well was washed three times with PBS-T. Thereafter washing step, 50 µl of ABTS (2, 2 azino-bis[3-ethylbenzthiazoline-6-suulphonic acid]) solution (ZYMED CA, USA) was added into the well and incubated at 25ºC for 30 minutes with light protection. The ELISA signal was measured at OD 405 nm by using ELISA reader (Multiskan® EX, Labsystem).

11. Tyrosine kinase activity inhibiting assay

Antibody Beacon™ Tyrosine Kinase Assay Kit (A-35725): Invitrogen, USA was used to assess the tyrosine kinase activity. In this experiment, the VH/V_HH antibodies were tested the inhibition activity to inhibit the TK activity of EGFR-TK. The experimental process was performed as a description of manufacturer. The reaction mixture of 50 µl contained 150 nM Anti-phosphotyrosine antibody, 75 nM Oregon Green 488 ligand, 0.6 mg/ml poly(Glu:Tyr), various [ATP] = 0.025 – 0.8 mM. Before determine the activity of TK, the reaction mixture was incubated with PBS, 3µg of irrelevant VH/V_HH antibody, and VH/V_HH antibody against ErbB2-TK at 25ºC for 15 minutes. The PBS, non-specific VH/V_HH antibody, and VH/V_HH antibody against ErbB2-TK were used as negative control, non-specific control (irrelevant), and test group, respectively. After the incubation period, the mixtures were added into the reaction mixture. Before measured the TK activity, one hundred and twenty-five microliters of 1X kinase buffer plus 2 mM DTT were added into the mixture solution. Fluorescence was measured on an automated 96-well infinite M200 Pro™ plate reader (Tecan, Crailsheim, Germany) using an excitation wavelength of 485 nm and an emission wavelength of 535 nm.

12 Plasmid extraction and DNA sequencing

The VH/V_HH expressing bacterial colony that provided the TK inhibiting activity was subjected to plasmid extraction. The plasmid extraction was performed by using alkaline lysis method. The bacterial colony was inoculated into 3 ml of LB-AG broth and incubated at 37°C with shaking at 250 rpm for 18 hours. The culture was centrifuged at 4,000x *g*, 25°C for 15 minutes and the bacterial cells in the pellet were resuspended in 200 µl of solution I (50 mM glucose, 25 mM Tris-HCl, pH 8.0, 10 mM EDTA) and the solution was transfer into 1.5 ml tube. The 400 µl of solution II (0.1 N NaOH, 1% SDS) was added into the tube and mixed by inverting. The tube was incubated on ice for 5 minutes and 300 µl of solution III (3 M potassium acetate, pH 5.2) was added into the tube. The tube was incubated on ice for 5 minutes and the tube was spun at 12,000 x *g* for 5 minutes at 25ºC. The supernatant was transferred to

1.5 ml tube and 0.6 volume of isopropanol was added to the supernatant. The mixture was mixed by inverting and the mixture was spun at 12,000 x *g* for 5 minutes at 4ºC. The supernatant was discarded and 70% ethanol was added into the tube and the mixture was spun at 12,000 x *g* for 5minutes at 4ºC. Pellet was dried and the pellet was resuspended with 20 µl of DW containing RNaseA (10 µg/ml). The plasmid DNA concentration was determined by measuring the OD at $A_{260\ nm.}$ The plasmids DNA of screened clones were subjected to DNA sequencing

Materials for experiment 2

1. Computing Resources

1. Laboratory of Protein Engineering and Modeling, Department of Biochemistry, Faculty of Science, Kasetsart University (Linux PC, AMD Athlon ™ 64 X2 Dual Core Processor 4400+)

2. Personal computer (Window 7 32 bit PC, intel™ core i3 Processor)

3. Computing Programs

   3.1 AutoDock 4.0
   3.2 Discovery Studio 2.5
   3.3 Pymol v0.99
   3.4 PROCHECK v.3.5.4.

Methods for experiment 2

In this study, the antibody-antigen complexes were determined the binding interaction by using computational methods. There are three antibody-antigen complexes including ScFv against tetrodotoxin, VH/V_HH against RNA dependent RNA polymerase of hepatitis C virus, and VH/V_HH against tyrosine kinase activity of ErbB2. These computational methods contain molecular modeling and molecular docking.

1. Molecular modeling

The homology modeling was used as a tool for molecular modeling. All processes of homology modeling were calculated by using Discovery Studio 2.5 program (Accelrys Inc., CA, USA). The amino acid sequences of antibodies were subjected into BLAST search analysis. The homolog sequence with high percent

identity that provides 3D structure was used as a template for homology modeling. The amino acid sequence of antibody and its template were subjected into sequence alignment. The result from sequence alignment was used as an initial data for homology modeling. The best model from homology modeling was accessed the quality of structure by using Ramachandran plot. The plot was generated by using PROCHECK v.3.5.4.

2. Molecular docking

The molecular docking of antibody and antigen requires an appropriate algorithm due to type of antigen. There are two types of antigen including macromolecule and small molecule. In this study, both types of antigen were investigated the binding interaction such as tetrodotoxin (small molecule), and RNA dependent RNA polymerase of hepatitis C virus (macromolecule). The Autodock4 was used to study the complex of antibody and tetrodotoxin. In addition, the ZDOCK and RDOCK were used to study the complex of antibody and RNA dependent RNA polymerase of hepatitis C virus.

2.1 AutoDock4

The 2D structure of tetrodotoxin was obtained from PubChem (Compound ID: 107878) and set as the ligand for docking calculation. The 2D structure of tetrodotoxin was converted to 3D structure by using Discovery Studio 2.5 program (Accelrys Inc., CA, USA). For ScFv antibodies, the constructed HuScFv clone s16 and s35 were obtained from homology modeling and set as macromolecule. The rotational bonds of antibody were regarded as being rigid, while the rotational bonds of the ligands were set as flexible. Subsequently, apply the hydrogen atoms, the Kollman united-atom charges and the solvent parameter to the antibody. The grid box was covered the antibody structure and the dimensions of the box were set as 110, 110, and 110 as well as the grid points spacing was 0.375 Å. The grid affinity maps for A (aromatic carbon), C, HD, N, NA (hydrogen-bond-accepting N), OA (hydrogen-bond-accepting O), and SA (hydrogen-bond-accepting S) atom types were calculated

by using AutoGrid 4.0. The Lamarckian genetic algorithm search parameter was activated to simulate protein-ligand docking with 100 trial runs. The population size was set as 150. Docking conformations were clustered using a tolerance of 2.0 Å. The docking results were sorted by the lowest binding energy (AutoDock4 score). In addition, the interacting residues of antibodies complexes were determined by using ligand interactions scripts that embedded on Discovery Studio 2.5 program (Accelrys Inc., CA, USA).

2.2 ZDOCK and RDOCK calculation

The processes of protein-protein docking are consisted of molecular docking and refinement of docking structure. All docking processes were calculated by using ZDOCK and RDOCK module that embedded on Discovery Studio 2.5 program (Accelrys Inc., CA, USA). The 3D structure of antibodies and their antigens including tetrodotoxin, RNA dependent RNA polymerase of hepatitis C virus, and VH/V_HH against tyrosine kinase activity of ErbB2 were subjected into protein-protein docking. In ZDOCK process, there is no flexible body setting because the ZDOCK are rigid-body docking method. For docking parameter setting, the antibodies and their antigens were set as ligand and receptor respectively. The angular step size for rotational sampling of the ligand orientations were set as 15. All docking processes have no receptor and ligand blocked residues. After ZDOCK calculation, the set of docked protein poses that generated by ZDOCK were subjected to energy optimization and structure refinement by using RDOCK calculation. The given poses from ZDOCK were used as an initial data RDOCK calculation. Before running RDOCK calculation, both antibody and their antigen were applied CHARMm (Chemistry at HARvard Macromolecular Mechanics) force field. Like ZDOCK calculation, the antibodies and their antigens were set as ligand and receptor respectively. The ZDOCK poses with high ZDCOK scores, high Density values, and low cluster numbers were selected by using 3D point plot. The selected ZDOCK poses were used as input poses for RDOCK calculation. The rest of parameters were set as default values. All Figures of protein structures were produced by using Pymol v0.99.

Materials for experiment 3

1. Computing Resources

1. Laboratory of Protein Engineering and Modeling, Department of Biochemistry, Faculty of Science, Kasetsart University (Linux PC AMD Athlon ™ 64 X2 Dual Core Processor 4400+)

2. Computing Programs

2.1. AutoDock 4.0
2.2. Pymol v0.99
2.3. SiMMap server

Methods for experiment 3

Virtual screening (VS) is a computational drug discovery technique (Rollinger *et al.*, 2008b). This method provides lower cost and used time for working, while searching for novel potential drug-liked compounds (Schneider, 2002). This method requires the detail of the 3D structure of the binding site of the target protein to arrange the ligand compounds by their likelihood. There are two categories of screening technique such as ligand-based and receptor-based. For this research perform under receptor-based virtual screening with National cancer institute (NCI) database and Chembridge Diverset™.

1. Virtual screening procedures for ErbB2-TK

The Virtual screening was calculated by using AutoDock program V.4.0. The NCI diversity was obtained from http://zinc.docking.org/browse/catalog/all. In addition, ChemBridge Diverset™ was obtained from ChemBridge Corporation. Further details are available at www.chembridge.com. For the 3D structure of kinase

domain of human HER2 (erbB2), it was obtained from protein data bank (PDB entry 3pp0).

Before VS calculation, the VS parameter was validated by using self-docking calculation. The self-docking calculation is the docking calculation of the co-crystallized ligand into the co-crystallized receptor. To determine the self-docking calculation, the root mean square deviation (RMSD) value of self-docking pose was compared with RMSD of the co-crystallized ligand. The 3D structure of ErbB2 model was set as the target protein to screen the ligands from NCI diversity. The docking process consisted of sampling the co-ordinate space around the binding site and scoring each possible ligand pose, which was then taken as the predicted binding mode for each compound. For parameter setting, the rotational bonds of all proteins were regarded as being rigid, while the rotational bonds of the ligands were treated as flexible in a Python script (prepare_ligand4.py). This script is embedded in the AutoDock program as script commands. The hydrogen atoms, the Kollman united-atom charges and the solvent parameter were applied to the proteins, and the set-up process for the grid was performed in AutoDockTools v.1.5.2. A grid box was built with twelve different atom types (A = aromatic carbon, C = aliphatic carbon, H = hydrogen, O = oxygen, S = sulfur, N = nitrogen, P = phosphorus, I = iodine, F = iron, f = fluorine, b = bromine, and c = chlorine) via AutoGrid calculations. The grid space was set at 0.375 Å. The grid box layout is 23.49*-0.369*53.848 points in x*y*z axis from center of macromolecule. AutoDock calculation was performed 50 times per ligand. All of the processes were controlled by in-house scripts. The program was used the genetic algorithm (GA) for sampling ligand flexibility and a force field for scoring the function. The Lamarckian genetic algorithm search parameter was activated to simulate protein-ligand docking with 50 trial runs. The population size was set to 150. Docking conformations were clustered using a tolerance of 2.0 Å by employing the clustering python script (summarize_results4.py). The docking results were sorted by the lowest binding energy (AutoDock4 score). The interactions between ErbB2-TK and ligand were observed by using Discovery Studio 2.5 program (Accelrys Inc., CA, USA)

RESULTS AND DISCUSSION

Results and discussion for experiment 1

1. Selection antibody against ErbB2-TK by using bio-panning technique

One microgram of recombinant human HER2/ErbB2 (Lys676-Val1255) kinase protein was coated on an ELISA plate (E.I.A/R.I.A. 8 Well strip Flat bottom Corning, NY USA) as previously described. The coated protein was successfully detected by using mouse anti-glutathione S-transferase (GST) tag monoclonal antibody. After detection with mouse anti-GST tag monoclonal antibody, the 5.9 $x10^8$ cfu of phage library was added into the well and performed bio-panning process as previously described.

2. Screening of HB2151 *E. coli* clone carrying *VH/V_HH* gene

In bio-panning experiment, there are twenty-eight HB2151 *E. coli* colonies from LB-AG agar plate. All of bacterial colonies of HB2151 *E. coli* were subjected to analyze the *VH/V_HH* gene by using PCR technique. The primer R1 and R2 were used as primer for amplification of *VH/V_HH* gene. The PCR products of *VH/V_HH* gene were shown in Figure 11 and 12. There are 15 colonies that gave DNA band at size 600 bp as a positive clone indicating that these *E. coli* carried *VH/V_HH* gene. The *E. coli* carried *VH/V_HH* gene was subjected to express the VH/V_HH antibody**.** The percentage of positive colonies from LB-AG agar plate was 53.57%.

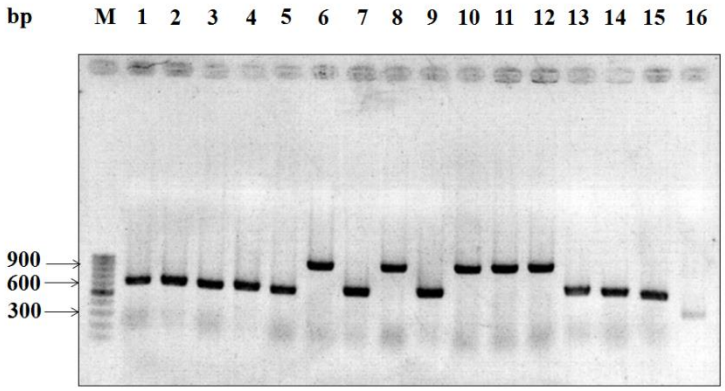


Figure 11 Agarose gel electrophoresis of PCR product was shown in invert color. Lane M, 100 bp DNA ladder. Lane 1-16 showed the PCR product of selected colonies. The PCR products were analyzed on 1% agarose gel. The expected size of *VH/V$_H$H* gene is 600 bp.

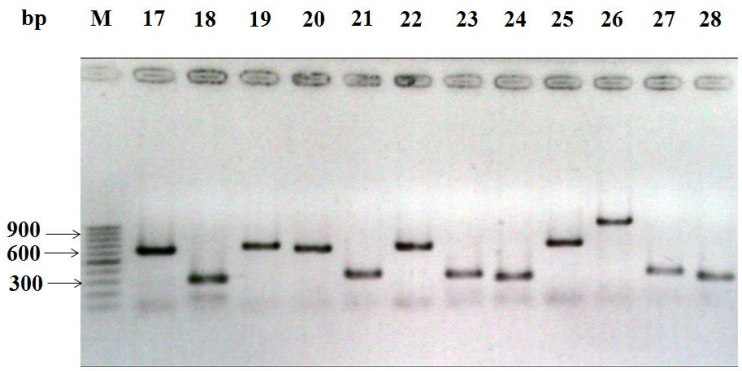


Figure 12 Agarose gel electrophoresis of PCR product was shown in invert color. Lane M, 100 bp DNA ladder. Lane 17-28 showed the PCR product of selected colonies. The PCR products were analyzed on 1% agarose gel. The expected size of *VH/V$_H$H* gene is 600 bp.

3. Screening of VH/V_HH-expressing HB2151 *E. coli* clone by using Western blot analysis

As the result of agarose gel electrophoresis, the bacterial colonies clone number 1, 2, 3, 4, 5, 7, 9, 13, 14, 15, 17, 19, 20, 22, and 25 provided a positive DNA band at 600 bp. This result indicated the bacterial colonies carrying *VH/V_HH* genes. The HB2151 *E. coli* colonies carrying *VH/V_HH* genes were determined by using small scale protein expression as previously described. For small scale protein expression, the bacterial cultures were incubated at 37ºC, 250 rpm in shaking incubator until optical density at 600 nm equal to 0.5. After the incubation period, the IPTG was added into cultures and incubated at 37ºC, 250 rpm for 5 hours. The whole cell lysate was prepared and subjected to SDS-PAGE and Western blot analysis. The Western blot analysis of HB2151 *E. coli* colonies carrying *VH/V_HH* genes were shown in Figure 13 and 14. The expected size of VH/V_HH antibody was 15-20 kDa. There are 9 colonies that gave protein band at size 15-20 kDa as a positive clone indicating that these contained VH/V_HH antibody. The VH/V_HH antibodies were subjected to determine the binding activity by using ELISA technique**.** The percentage of positive colonies that provided VH/V_HH antibody was 64.2%. As the variation size of protein band, the variation was resulting from the variation number of amino acid residues in Complementary-determining region (CDR).

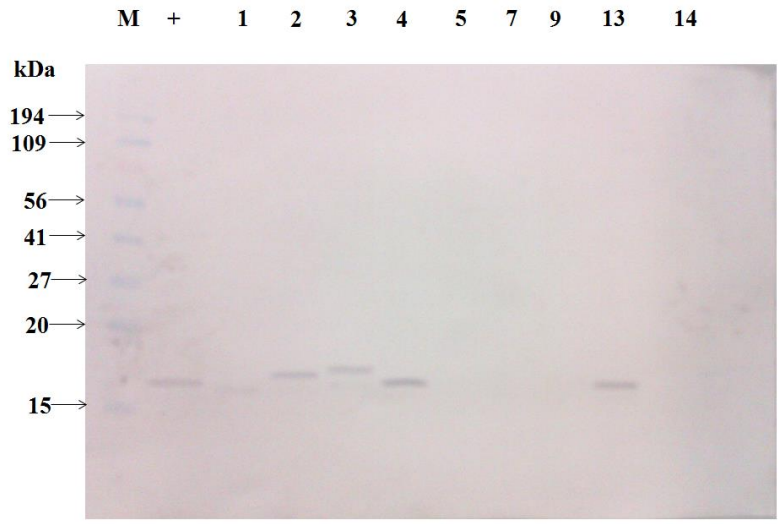


Figure 13 Western blot analysis of VH/V_HH-expressing HB2151 *E. coli* clone, the protein from VH/V_HH-expressing HB2151 *E. coli* clone was separated in 12% polyacrylamide gel under reducing condition. The antibodies were transferred onto nitrocellulose (NC) membrane and incubated with mouse anti-E tag monoclonal antibody. Lane M showed the pre-stained protein molecular weight maker. Lane positive (+) shown the VH/V_HH antibody. Lane number 1, 2, 3, 4, 5, 7, 9, 13, and 14 showed the protein from VH/V_HH-expressing HB2151 *E. coli* clone.

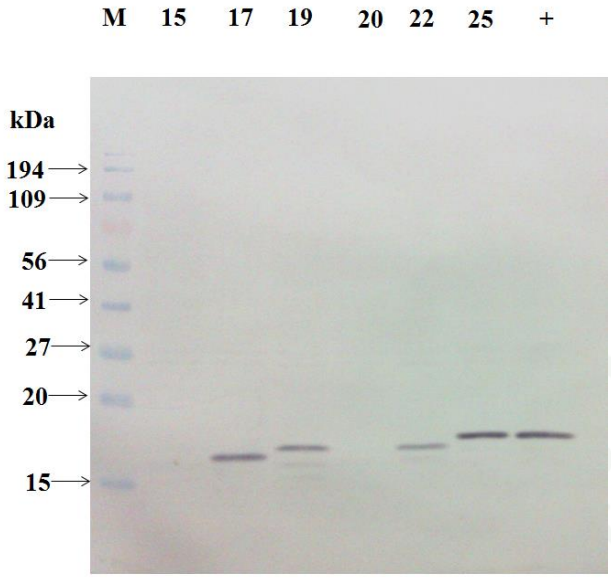


Figure 14 Western blot analysis of VH/V_HH-expressing HB2151 *E. coli* clone, the protein from VH/V_HH-expressing HB2151 *E. coli* clone was separated in 12% polyacrylamide gel under reducing condition. The antibodies were transferred onto NC membrane and incubated with mouse anti-E tag monoclonal antibody. Lane M showed the pre-stained protein molecular weight maker. Lane positive (+) shown the VH/V_HH antibody. Lane number 15, 17, 19, 20, 22, and 25 showed the protein from VH/V_HH-expressing HB2151 *E. coli* clone.

4. Specificity determination of VH/V_HH antibodies against ErbB2-TK by using indirect ELISA technique

According to the concerning aspect of antibody, there are two aspects of antibody against antigen of interest. The two aspects are binding property of antibody and inhibiting property, respectively. To determine the binding property of antibody, the indirect ELISA technique is a proper technique because it simply and accuracy technique. The VH/V_HH antibody from HB2151 *E. coli* colony number 2, 3, 4, , 13, 17, 19, 22, and 25 were subjected to determine the binding affinity by using indirect ELISA technique. The cell lysate of HB2151 *E. coli* was used as a negative control.

The amount of VH/V_HH antibody from HB2151 *E. coli* and negative control were 5 μg. The optical density (OD) of VH/V_HH antibody against ErbB2-TK was shown in Figure 15. The VH/V_HH antibody number 4, 17, 22, and 25 were selected to determine the binding interaction by using computational approaches.

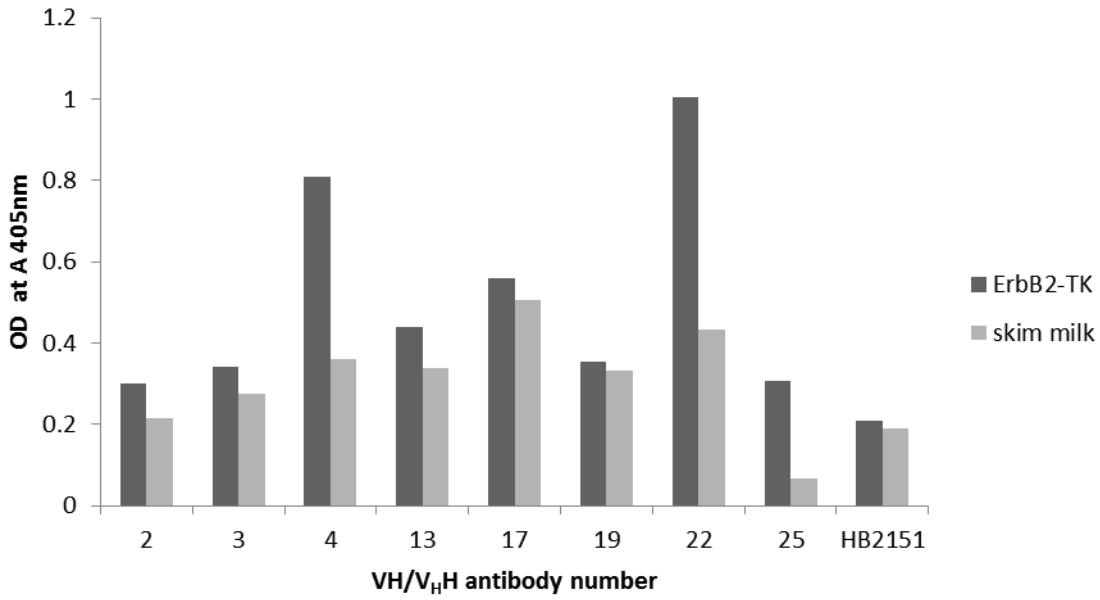


Figure 15 Indirect ELISA for specificity determination of the binding between different VH/V_HH antibody and ErbB2-TK. The binding signal of antibody and skim milk was used to determine the non-specific binding of antibody.

5. Specificity determination of VH/V_HH antibodies against EGFR-TK by using indirect ELISA technique

The evidence for a dual binding mode of VH/V_HH antibody was determined by using indirect ELISA technique. The VH/V_HH antibody against ErbB2-TK from HB2151 *E. coli* colony number 2, 3, 4, 5, 13, 17, 19, 22, and 25 were subjected to determine the binding affinity by using indirect ELISA technique. The cell lysate of HB2151 *E. coli* was used as a negative control. The amount of VH/V_HH antibody from HB2151 *E. coli* and negative control were 5 μg. The optical density (OD) of

VH/V_HH antibody against ErbB2-TK was shown in Figure 16. The VH/V_HH antibody number 4 and 22 were selected to determine inhibiting function and binding interaction by using TK activity inhibiting assay and computational approaches, respectively.

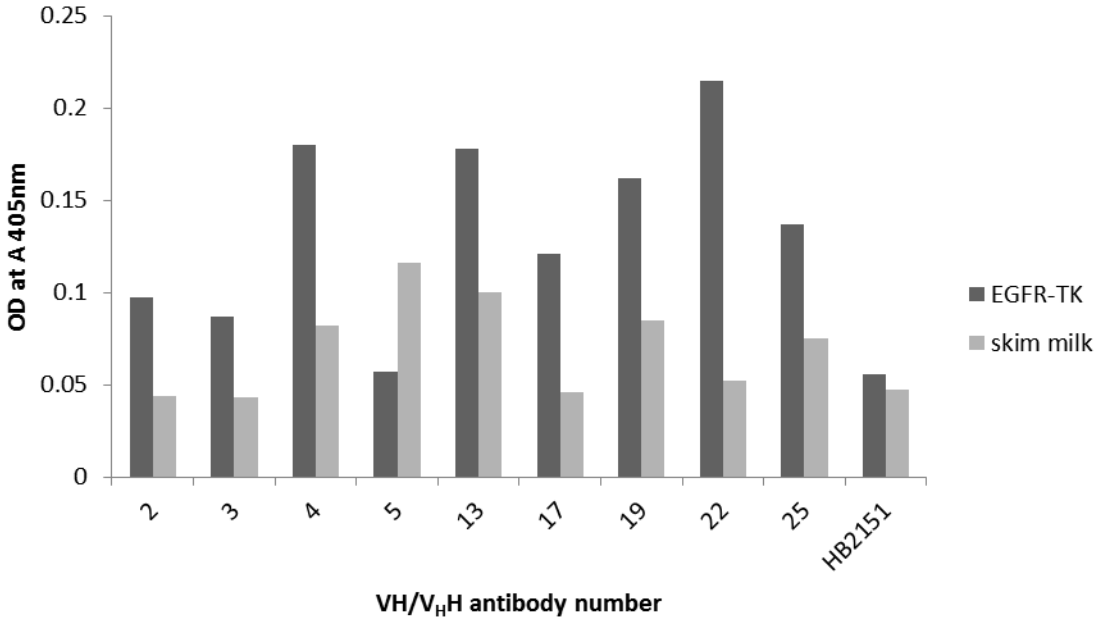


Figure 16 Indirect ELISA for specificity determination of the binding between different VH/V_HH antibody and EGFR-TK. The binding signal of antibody and skim milk was used to determine the non-specific binding of antibody.

The indirect ELISA technique provided the many advantages including maximum immune reactivity and sensitivity. In addition, the ELISA technique can wash away nonspecifically binding. As the advantage point of indirect ELISA, this technique is a powerful tool for measuring antigen of interest within crude preparation. Then, the ELISA signal of crude VH/V_HH antibody, ErbB2-TK, and EGFR-TK can be used as reliable data.

6. Specificity determination of VH/V_HH antibodies against EGFR-TK by using TK activity inhibiting assay

Although the binding property of VH/V_HH antibody against ErbB2-TK was determined by using indirect ELISA technique, this technique cannot determine the inhibiting property of antibody and ErbB2-TK. Thus, the TK activity inhibiting assay is an appropriate technique to investigate the inhibiting property of the antibody. Unfortunately, the ErbB2-TK recombinant protein was not available for this experiment. The EGFR-TK was used instead. According to the sequence and structural similarities in the TK domain of ErbB2 and EGFR, an antibody that designed for inhibiting ErbB2-TK may be used for inhibiting EGFR-TK. In this experiment, the VH/V_HH antibody against ErbB2-TK from HB2151 *E. coli* colony number 4 and 22 were subjected to determine the inhibiting property by using TK activity inhibiting assay. Although, VH/V_HH antibody against ErbB2-TK from HB2151 *E. coli* colony number 25 had the difference in the ELISA signal of EGFR-TK and skim milk, the ELISA signal of colony number 25 did not higher than ELISA signal of HB2151. The VH/V_HH antibody against ErbB2-TK from HB2151 *E. coli* colony number 25 was not selected for TK activity inhibiting assay. There are three candidate group of experiment including EGFR-TK, EGFR-TK with irrelevant antibody, and EGFR-TK inhibiting antibody. The inhibiting property of antibodies against EGFR-TK was reported as a co arbitrary unit of inhibition (Figure 17).

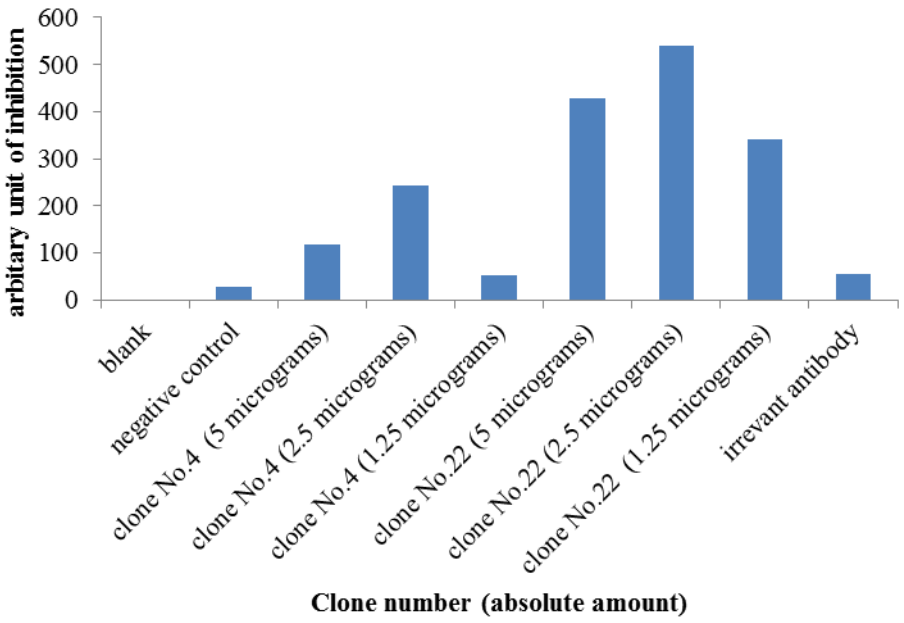


Figure 17 Inhibition of different clones of VH/V_HH antibodies against EGFR-TK was shown in arbitrary unit of inhibition. The blank group represents reaction mixture without antibody. The negative control represents reaction mixture with cell lysate of HB2151 *E. coli*. The irrelevant group represents reaction mixture with irrelevant antibody (VH/V_HH antibodies against ErbB2-TK clone No. 2). The antibody groups represent reaction mixture with inhibiting antibody in different absolute amount.

The result from TK activity inhibiting assay reveal that, the VH/V_HH antibody against ErbB2-TK clone number 4 and 22 can inhibit the TK activity of EGFR-TK when compared with irrelevant antibody. In addition, the antibody clone number 4 and 22 had highest arbitrary unit of inhibition. The absolute amount of antibody clone number 4 and 22 that provided the highest arbitrary unit of inhibition was 2.5 μg. Although, the irrelevant group revealed lowest arbitrary unit of inhibition, it still showed the arbitrary unit of inhibition. The remaining of arbitrary unit of inhibition may be resulted from protein contamination. However, the signal from contaminated protein did not higher than the lowest signal from test groups. Then, the result from TK activity inhibiting assay can be used to determine the inhibiting activity of

antibody clone number 4 and 22. As the effect of absolute amount and arbitrary unit of inhibition, the absolute amount of VH/V_HH antibodies were 2.5 μg provided the maximum 2.5 μg. This effect may be resulted from concentration of antibody.

7. DNA sequencing of HB2151 *E. coli* clone carrying *VH/V_HH* gene

According to ELISA experiment, there are 4 clones of VH/V_HH antibody that shown binding property against ErbB2-TK and EGFR-TK. The DNA of VH/V_HH antibody clone number 4, 17, 22, and 25 were subjected to DNA sequencing experiment. The plasmids of HB2151 *E. coli* clone carrying *VH/V_HH* gene were subjected to DNA sequencing. The DNA sequencing was used pCANTAB5-*R1* primer (5′-CCA TGA TTA CGC CAA GCT TTG GAG CC-3′) and pCANTAB5-*R2* primer (5′-GCT AGA TTT CAA AAC AGC AGA AAG G-3′). The DNA sequences of VH/V_HH antibody clone number 4, 17, 22, and 25 were shown the restriction site of restriction endonuclease (*Sfi*I and *Not*I). These DNA sequences of antibody clone were subjected to deduce amino acid. The presence of restriction sites of restriction endonuclease (*Sfi*I and *Not*I) in antibody sequence confirmed the precision of DNA sequencing. The deduced amino acid of VH/V_HH antibody clone number 4, 17, 22, and 25 were used as initial data for computational analysis. The restriction sites of *Sfi*I and *Not*I were shown in figure 18.

The DNA sequencing results of VH/V_HH antibody clone number 4, 17, 22, and 25 were used as initial data for VH and V_HH identification. According to the various types of camelid antibody, the conventional antibody and heavy-chain antibody, the VH/V_HH antibody should be classified by using the characterization of VH and V_HH antibodies. The structures of conventional antibody and heavy-chain antibody were shown in figure 19. The V_HH from heavy chain antibody provided mutations of hydrophobic amino acid at residue number 42, 49, 50, and 52. These residues were located at framework 2 (FR2). These mutations were including V42F/Y, G49E, L50R/C, and W52G/L. The amino acid sequence analysis revealed that clone number 4, 22, and 25 were VH antibodies, while clone number 17 was V_HH antibody. The amino acid sequence alignment of antibody clone number 4, 17, 22,

and 25 was shown in figure 20. The mutations of hydrophobic amino acid at residue number 42, 49, 50, and 52 were shown in figure 21.

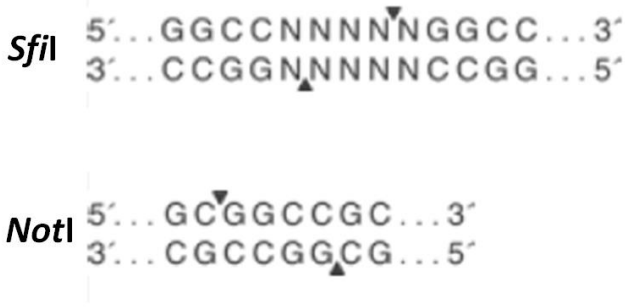


Figure 18 The figure showed the restriction sites of *Sfi*I and *Not*I

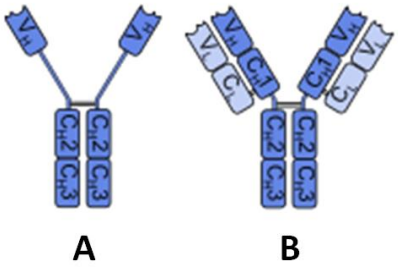


Figure 19 The figure showed structure of heavy-chain antibody and conventional antibody. The figure “A” represented the structure of heavy-chain antibody. The figure “B” represented the structure of conventional antibody.

Figure 20 The figure showed amino acid sequence alignment of antibody clone number 4, 17, 22, and 25. The non-matching, weak, strong, and identical residues were colored in white, pale-blue, medium-blue, blue, and dark-blue, respectively.

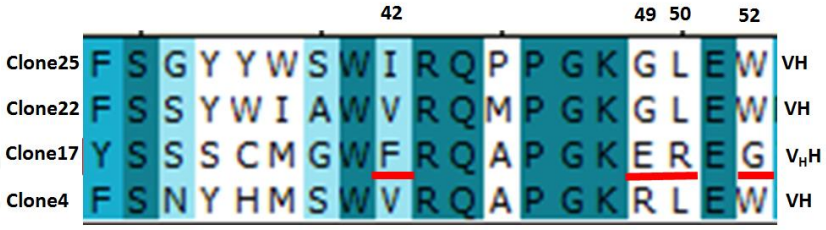


Figure 21 The figure showed the mutation residues of antibody clone number 17 compared with antibody clone number 4, 22, and 25. The mutated amino acid residue number 42, 49, 50, and 52 were located at FR2. The mutated residues of clone number 17 were labeled in red line.

The DNA sequences of VH/V_HH antibody clone number 4, 17, 22, and 25 were shown the restriction site of restriction endonuclease (*Sfi*I and *Not*I). These DNA sequences of antibody clone were subjected to deduce amino acid. The presence of restriction sites of restriction endonuclease (*Sfi*I and *Not*I) in antibody sequence confirmed the precision of DNA sequencing. The deduced amino acid of VH/V_HH

antibody clone number 4, 17, 22, and 25 were used as initial data for computational analysis.

Results and discussion for experiment 2

1. Molecular modeling and molecular docking of ScFv against tetrodotoxin

The tetrodotoxin (TTX) or tetrodox is a potent neurotoxin that has found in many species of fish including pufferfish, porcupinefish, and ocen sunfish. For this toxin, the antidote is not available. Tetrodotoxin blocked the action potential by binding to the voltate-gate resulting in nerve conducting failure. The 3D structure models of ScFv clone s16 and s35 were constructed by using structures of heavy and light chain variable domain of anti-DNA binding antibody (PDB entry 2GKI) as a template. For ScFv clone s16, the percentage of identity sequence and similarity sequence are 61.7% and 75.8%, respectively. For ScFv clone s35, the percentage of identity sequence and similarity sequence are 62.7% and 75.1%, respectively. The structures of ScFv clone s16 and ScFv clone s35 were assessed by using Ramachandran plot. The Ramachandran plot of ScFv clone s16 revealed the percentage of residue in core region, allowed region, generously allowed region, and disallowed region are 88.2%, 8.8%, 2.0%, and 1% respectively. For ScFv clone s35, the percentage of residue in core region, allowed region, and generously allowed region are 89.7%, 9.8%, and 0.5% respectively. ScFv clone s35 have no residue in disallowed region. The assessment of antibodies structures revealed that both of antibodies structures are appropriate structures for further analysis because they provide acceptable number of residue in disallowed region.

The structure of ScFv clone s16 and ScFv clone s35 were performed molecular docking with structure of tetrodotoxin was obtained from PubChem (Compound ID: 107878). The docking complexes of ScFv clone s16 and ScFv clone s35 were shown in Figure 22 and 23, respectively.

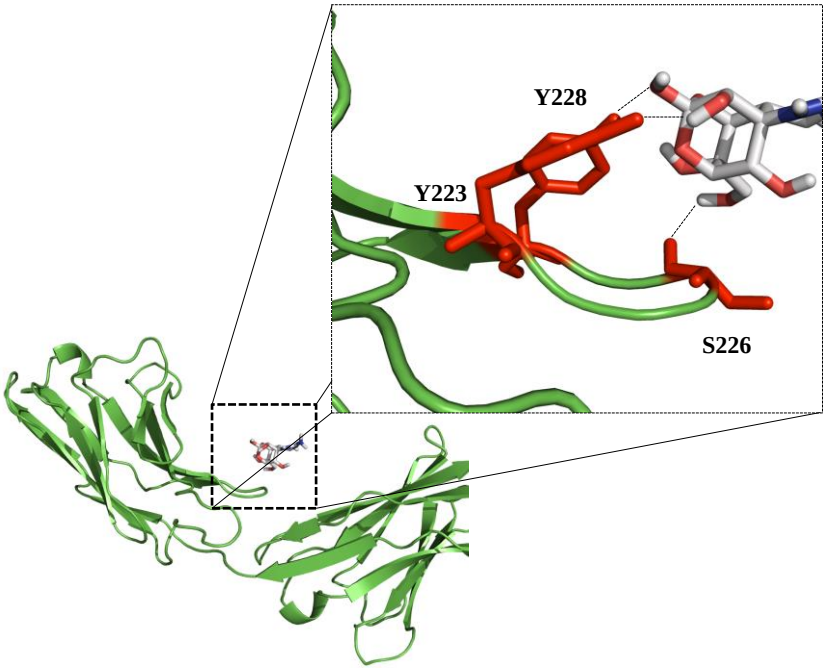


Figure 22 Docked complex between ScFv clone s16 and tetrodotoxin. The green ribbon and white stick refer as ScFv clone s16 and tetrodotoxin respectively. The detailed-view represented the interaction between ScFv clone s16 and tetrodotoxin; the dash-line refers as the H-bond between ScFv clone s16 and tetrodotoxin.

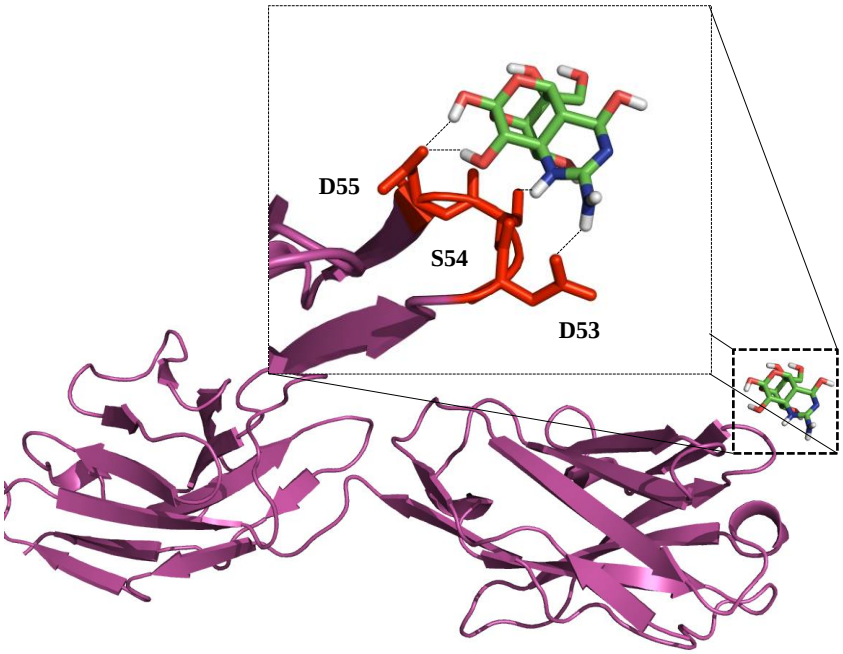


Figure 23 Docked complex between ScFv clone s35 and tetrodotoxin. The purple ribbon and green stick refer as ScFv clone s35 and tetrodotoxin respectively. The detailed-view represented the interaction between ScFv clone s35 and tetrodotoxin; the dash-line refers as the H-bond between ScFv clone s35 and tetrodotoxin.

The docking results reveal the calculated binding energy and binding residues. The binding energy of ScFv clone s16 and ScFv clone s35 are -3.23 kcal/mol and -3.79 kcal/mol respectively. As the calculated docking energy, the interaction between antibodies and tetrodotoxin are spontaneous reaction due to minus binding energy. The binding residue analysis of ScFv clone s16 complex revealed that the tetrodotoxin was interacted with Y223, S226, and Y228 that are residues of light chain. For binding residue analysis of ScFv clone s35 complex, the tetrodotoxin was interacted with D53, S54, and D55 that are residues of heavy chain. The prediction of binding interaction of antibodies revealed that the ScFv clone s16 and ScFv clone s35 may be used to inhibit the function of tetrodotoxin because the functional groups of toxin were interacted with the antibodies. Although, the molecular docking is the prediction

technique to determine the inhibiting activity of the antibodies, there are many supportive studies that confirmed the inhibiting activity of the antibodies such as tetrodotoxin bioactivity in nerve cell based and mouse bio-assays (Chulanetra *et al.*, 2012).

2. Molecular modeling and molecular docking of VH/V_HH against RNA dependent RNA polymerase of hepatitis C virus

Hepatitis C virus is a positive-sense single strand RNA virus that is a cause of hepatitis C in human. The RNA dependent RNA polymerase involved with replication of this virus. NS5B is RNA dependent RNA polymerase of hepatitis C virus. Nowadays, the 3D structure of NS5B is not available. The 3D structure model of NS5B was constructed by using structures of HCV NS5B RNA polymerase in complex with novel class of dihydropyrone-containing inhibitor (PDB entry 2HAI) as a template. For VH/V_HH against RNA dependent RNA polymerase of hepatitis C virus, there are four antibodies including VH9, VH13, V_HH6, and V_HH24. The 3D structure model of VH9 was constructed by using structures of humanized antibody C25 Fab fragment (PDB entry 2GCY) as a template. The percentage of identity sequence and similarity sequence are 75.6% and 83.7%, respectively. The 3D structure model of VH13 was constructed by using structures of humanized antibody C25 Fab fragment (PDB entry 2H32) as a template. The percentage of identity sequence and similarity sequence are 80.3% and 89.3%, respectively. The 3D structure model of V_HH6 was constructed by using structures of antibody heavy chain variable domain (VH)-P8 (PDB entry 1VHP) as a template. The percentage of identity sequence and similarity sequence are 66.1% and 76.6%, respectively. The 3D structure model of V_HH24 was constructed by using structures of single domain camelid antibody (PDB entry 1F2X) as a template. The percentage of identity sequence and similarity sequence are 75.8% and 78.9%, respectively. The structures of VH9, VH13, V_HH6, and V_HH24 were assessed by using Ramachandran plot. The Ramachandran plot of VH9 revealed the percentage of residue in core region, allowed region, generously allowed region, and disallowed region are 42.5%, 48.1%, 7.5%, and 1.9% respectively.

Figure 24 Docked complex between VH9 and NS5B. There are three domains of NS5B. Thumb, palm, and finger domains are colored in green, blue, and red. The VH9 is colored in purple. The binding residues of NS5B that interacted with VH9 are colored in yellow.

For VH13, the percentage of residue in core region, allowed region, are 90.4%, 9.6%, respectively. There is no residue in generously allowed region and disallowed region. The Ramachandran plot of V_HH6 revealed the percentage of residue in core region, allowed region, generously allowed region, and disallowed region are 54.5%, 34.7%, 7.9%, and 3.0% respectively. For V_HH24, the percentage of residue in core region, allowed region, and generously allowed region are 92.3%, 5.8%, and 1.9%, respectively. There is no residue in disallowed region. The assessment of antibodies structures revealed that all of antibodies structures are appropriate structures for further analysis because they provide acceptable number of residue in disallowed region.

The structure of antibodies clone VH9, VH13, V_HH6, and V_HH24 were performed molecular docking with structure of NS5B. According to the structure of NS5B, there are three important domains such as thumb, palm, and finger domain.

These domains are involved with the binding of RNA duplex. The docking complexes of antibodies clone VH9, VH13, V_HH6, and V_HH24 were shown in Figure 24, 25, 26, and 27, respectively.

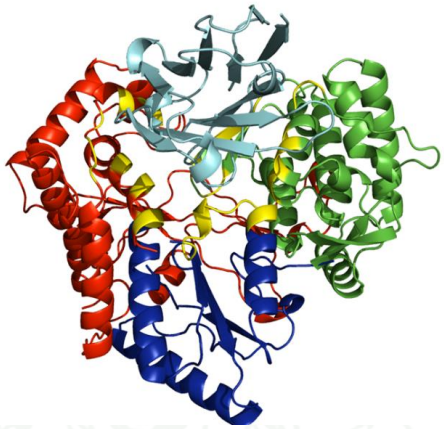


Figure 25 Docked complex between VH13 and NS5B. There are three domains of NS5B. Thumb, palm, and finger domains are colored in green, blue, and red. The VH13 is colored in cyan. The binding residues of NS5B that interacted with VH13 are colored in yellow.

Figure 26 Docked complex between V_HH6 and NS5B. There are three domains of NS5B. Thumb, palm, and finger domains are colored in green, blue, and red. The V_HH6 is colored in orange. The binding residues of NS5B that interacted with V_HH6 are colored in yellow.

Figure 27 Docked complex between V_HH24 and NS5B. There are three domains of NS5B. Thumb, palm, and finger domains are colored in green, blue, and red. The V_HH24 is colored in light grey. The binding residues of NS5B that interacted with V_HH24 are colored in yellow.

The docking results reveal the calculated binding energy and binding residues. The binding energy of VH9, VH13, V_HH6, and V_HH24 were -9.21, -10.43, -14.63, and -30.0 kcal/mol, respectively. As the calculated docking energy, the interaction between antibodies and NS5B are spontaneous reaction due to the value of binding energy is minus value. The binding residue analysis of all antibodies revealed that the antibodies interacted with thumb, palm, and finger domain of NS5B. Moreover, the interaction of NS5B and RNA duplex was also investigated by using molecular docking technique. The docking complex of RNA duplex and NS5B was shown in Figure 28.

Figure 28 Docked complex between RNA duplex and NS5B. There are three domains of NS5B. Thumb, palm, and finger domains are colored in green, blue, and red. The RNA duplex is colored in warm pink. The binding residues of NS5B that interacted with RNA duplex are colored in yellow.

From docking result, the RNA duplex was interacted with thumb, palm, and finger domains of NS5B. This docking result revealed that the binding regions of RNA duplex like the binding regions of antibodies against NS5B. Therefore, the VH9, VH13, V_HH6, and V_HH24 may be used as the occupied molecules that inhibited function of NS5B. There are many studies that confirmed the inhibition activity of antibodies against NS5B such as mimotope mapping and inhibition assays (Thueng-In *et al.*, 2012).

3. Molecular modeling and molecular docking of VH/V_HH against tyrosine kinase domain of ErbB2-TK

The computational approaches are used to study the interactions of antibodies against tyrosine kinase domain of ErbB2 (ErbB2-TK). These antibodies were taken from bio-panning technique. Before running the computational approaches, the antibody clones were determined the binding property and inhibitory activity against ErbB2-TK by using ELISA and TK assay. There are four clones of VH/V_HH antibody including clone 4, 17, 22, and 25 that provided these activity against ErbB2-TK. These antibody clones were used to perform molecular modeling and molecular docking calculation. The 3D structure model of VH/V_HH clone 4 was constructed by using structures of humanized antibody C25 Fab fragment (PDB entry 2GCY) as a template. The percentage of identity sequence and similarity sequence are 26% and 29%, respectively. The 3D structure model of VH/V_HH clone 17 was constructed by using structures of single domain camelid antibody CAB-CA05 (PDB entry 1F2X) as a template. The percentage of identity sequence and similarity sequence are 76.7% and 79.8%, respectively. The 3D structure model of VH/V_HH clone 22 was constructed by using structures of pre-B cell receptor (PDB entry 2H32) as a template. The percentage of identity sequence and similarity sequence are 28.2% and 31.6%, respectively. The 3D structure model of VH/V_HH clone 25 was constructed by using structures of FAB fragment from a human I_GM cold agglutinin (PDB entry 1DN0) as a template. The percentage of identity sequence and similarity sequence are 30.5% and 32.6%, respectively. The constructed structures of VH/V_HH antibody including clone 4, 17, 22, and 25 were assessed by using Ramachandran plot. The Ramachandran plot of VH/V_HH clone 4 revealed the percentage of residue in core region, allowed region, and generously allowed region are 88.0%, 11.0%, and 1.0%, respectively. There is no residue in disallowed region. The Ramachandran plot of VH/V_HH clone 17 revealed the percentage of residue in core region, allowed region, and disallowed region are 94.3%, 4.8%, and 1.0% respectively. There is no residue in generously allowed region. The Ramachandran plot of VH/V_HH clone 22 revealed the percentage of residue in core region, allowed region, and disallowed region are 90.4%, 8.7%, and 1.0% respectively. There is no residue in generously allowed

region. The Ramachandran plot of VH/V_HH clone 25 revealed the percentage of residue in core region, allowed region, generously allowed region, and disallowed region are 88.9%, 8.3%, 0.9%, and 1.9% respectively. The Ramachandran plots of all antibodies were shown in appendix. The assessment of antibodies structures revealed that all of antibodies structures are appropriate structures for further analysis because they provide acceptable number of residue in disallowed region.

The structure of VH/V_HH antibody clone 4, 17, 22, and 25 antibodies were performed molecular docking with structure of ErbB2-TK. The crystal structure of kinase domain of human HER2 (erbb2) was obtained from protein data bank (PDB entry 3pp0). This ERrb2-TK and antibodies were set as receptor and ligand for molecular docking. The docking complexes of VH/V_HH antibodies clone 4, 17, 22, and 25 were shown in Figure29, 30, 31 and 32, respectively

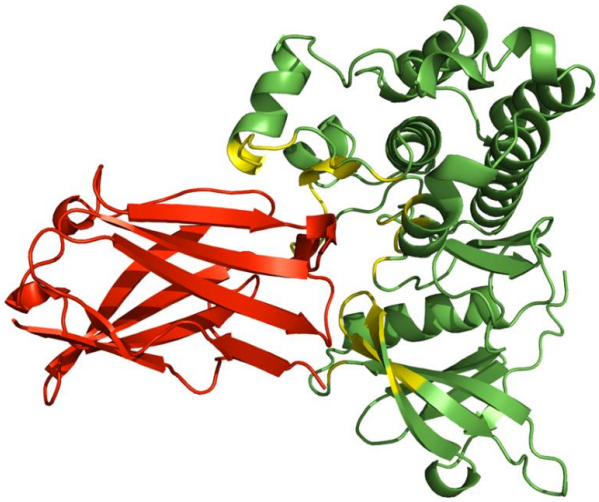


Figure 29 Docked complex between VH/V_HH antibody clone 4 and ErbB2-TK. VH/V_HH antibody clone 4 and ErbB2-TK are colored in red and green. The binding residues of ErbB2-TK that interacted with VH/V_HH antibody clone 4 are colored in yellow. These interacted residues are located in glycine-rich nucleotide phosphate binding loop and activation loop of ErbB2-TK.

Figure 30 Docked complex between VH/V_HH antibody clone 17 and ErbB2-TK. VH/V_HH antibody clone 17 and ErbB2-TK are colored in cyan blue and green. The binding residues of ErbB2-TK that interacted with VH/V_HH antibody clone 17 are colored in yellow. These interacted residues are located in surface loop of ErbB2-TK.

Figure 31 Docked complex between VH/V_HH antibody clone 22 and ErbB2-TK. VH/V_HH antibody clone 22 and ErbB2-TK are colored in orange and green. The binding residues of ErbB2-TK that interacted with VH/V_HH antibody clone 22 are colored in yellow. These interacted residues are located in glycine-rich nucleotide phosphate binding loop, DFG motif, and activation loop of ErbB2-TK.

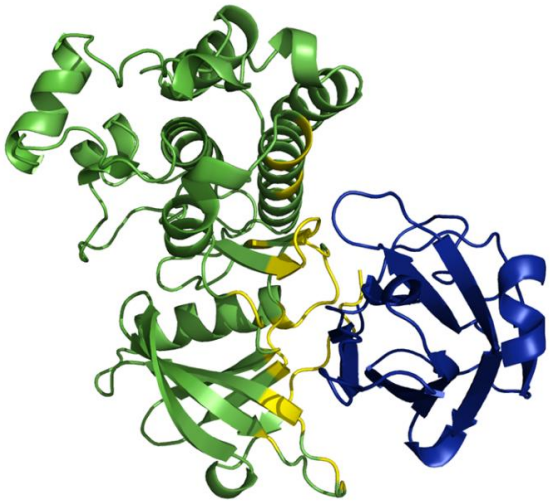


Figure 32 Docked complex between VH/V_HH antibody clone 25 and ErbB2-TK. VH/V_HH antibody clone 25 and ErbB2-TK are colored in blue and green. The binding residues of ErbB2-TK that interacted with VH/V_HH antibody clone 25 are colored in yellow. These interacted residues are located in N-terminal residue of ErbB2-TK.

The docking results reveal the calculated binding energy and binding residues. The binding energy of VH/V_HH antibodies clone 4, 17, 22, and 25 were -5.12, -14.71, -23.84, and -6.55 kcal/mol, respectively. As the calculated docking energy, the interaction between antibodies and ErbB2-TK are spontaneous reactions resulting from the value of binding energy. The binding residue analysis of all antibodies revealed that the VH/V_HH antibodies clone 4, 17, 22, and 25 interacted with a different epitope of ErbB2-TK as described above. As the docking model, VH/V_HH antibodies clone 4, 17, and 25 did not bind directly to catalytic loop. However, these antibodies may interfere with the activity of ErbB2-TK. The inhibiting mechanism of antibodies may result from the possession at the binding residues of ErbB2-TK.

Moreover, the VH/V_HH antibodies clone 4, 17, 22, and 25 were also investigated TK activity inhibiting assay with EGFR-TK. The result of TK activity inhibiting assay showed the inhibitory function of VH/V_HH antibodies clone 4 and

clone 22. These VH/V_HH antibodies clone 4 and clone 22 were performed molecular docking with structure of EGFR-TK. The crystal structure of epidermal growth factor receptor tyrosine kinase domain with 4-anilinoquinazoline inhibitor erlotinib was obtained from protein data bank (PDB entry 1M17). This EGFR-TK and antibodies were set as receptor and ligand for molecular docking. Before docking calculation, the structure of EGFR-TK was prepared by removing the 4-anilinoquinazoline inhibitor erlotinib. The docking complexes of VH/V_HH antibodies clone 4 and 22 were shown in Figure 33 and 34, respectively.

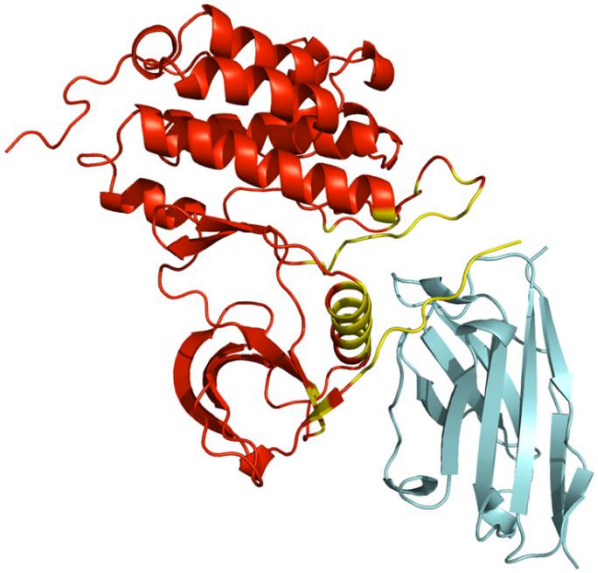


Figure 33 Docked complex between VH/V_HH antibody clone 4 and EGFR-TK. VH/V_HH antibody clone 4 and EGFR-TK are colored in red and blue. The binding residues of EGFR-TK that interacted with VH/V_HH antibody clone 4 are colored in yellow. These interacted residues are located in α helix C, DFG motif, and activation loop of EGFR.

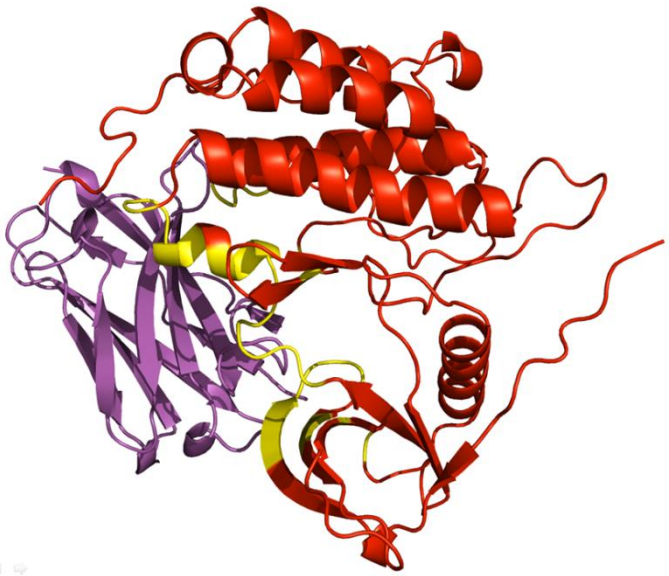


Figure 34 Docked complex between VH/V_HH antibody clone 22 and EGFR-TK. VH/V_HH antibody clone 22 and EGFR-TK are colored in red and purple. The binding residues of EGFR-TK that interacted with VH/V_HH antibody clone 22 are colored in yellow. These interacted residues are located in glycine-rich nucleotide phosphate binding group and catalytic loop of EGFR-TK.

The indirect ELISA and TK activity inhibiting assay were used to prove the binding property and inhibiting property, respectively. The docking calculation of EGFR-TK and the antibodies were used to confirm the indirect ELISA and TK activity inhibiting assay. For binding property, the binding energy of VH/V_HH antibodies clone 4 and 22 were -10.58 and -14.87 kcal/mol, respectively. The calculated binding energy of VH/V_HH antibodies reveled that the binding affinity of clone 4 lower than binding affinity of clone 22. As the indirect ELISA, the OD of VH/V_HH antibodies clone 4 and 22 were 1.005 and 0.808, respectively. The results from docking calculation and indirect ELISA were showed the correlated binding affinity of VH/V_HH antibodies clone 4 and 22. For inhibiting property, there are two evidences of docking calculation that confirmed the result of TK activity inhibiting assay. First, the minus value of binding energy correlated with spontaneous reactions

of antibodies and EGFR-TK. Second, EGFR-TK binding residues correlated with percentage of TK inhibiting activity. For antibodies clone 4, the antibody bound with α helix C, DFG motif, and activation loop of EGFR-TK. The interaction of antibodies clone 4 against α helix C, DFG motif, and activation loop of EGFR-TK may interfere the interaction of ATP and EGFR-TK. However, the antibodies clone 4 cannot occupy the ATP binding site (figure 33). For antibodies clone 22, the antibody bound with glycine-rich nucleotide phosphate binding group and catalytic loop of EGFR-TK. The interaction of antibodies clone 22 against glycine-rich nucleotide phosphate binding group and catalytic loop of EGFR-TK may interfere the interaction of ATP and EGFR-TK. The binding pose of antibody clone 22 and EGFR-TK conceal the key ATP binding site (figure34). According to the binding pose of antibody clone number 4 and 22, the antibody clone number 22 should be provided the higher percentage of TK inhibiting activity than antibody clone number 4. The result of TK activity inhibiting assay revealed that the antibody clone number 22 provided the higher percentage of TK inhibiting activity than antibody clone number 4. According to the size of antibody, there is no VH/V_HH antibody which can penetrate into the ATP binding site of ErbB2-TK and EGFR-TK. There are two proposed mechanisms of antibody against ErbB2-TK and EGFR-TK including competitive antibody and non-competitive antibody. The meaning of competitive antibody is the antibody which conceals the binding site of antigen of interest. The examples of competitive antibody against ErbB2-TK were VH/V_HH antibody clone number 4 and 22. For competitive antibody against EGFR-TK, The example was VH/V_HH antibody clone number 4. On contrary, the meaning of non-competitive antibody is the antibody which the binding of antibody did not conceal the binding interface. The examples of competitive antibody against EGFR-TK were VH/V_HH antibody clone number 4 and 22. For competitive antibody against EGFR-TK, The example was VH/V_HH antibody clone number 22. For non-competitive antibody against EGFR-TK, The example was VH/V_HH antibody clone number 4.

In conclusion, the computational approaches including molecular modeling and molecular docking provide the useful information to determine the property of antibody. The molecular modeling and molecular docking confirm the binding activity and inhibiting activity of antibodies. The information from computational

analysis correlated with experimental data such as indirect ELISA and TK inhibiting activity assay. Moreover, the computational approaches also provide important information such as binding energy, binding pose, and binding residue.

Results and discussion for experiment 3

1. Virtual screening of ErbB2-TK against NCI database

In self-docking calculation, the co-crystallized ligand (2-2{-[4-9{5-chloro-6-[3-9trifluoromethyl) phenoxy] pyridine-3-yl} amino0-5H-pyrrolo [3, 2d] pyrimidin-5-yl] ethoxy}) ethanol) was docked into the ErbB2-TK. The result of self-docking analysis revealed that the RMSD of the co-crystallized ligand is 1.76 when compared with 3D structure from PDB database (Figure 35). The calculated binding energy of the co-crystallized ligand is -9.11 kcal/mol.

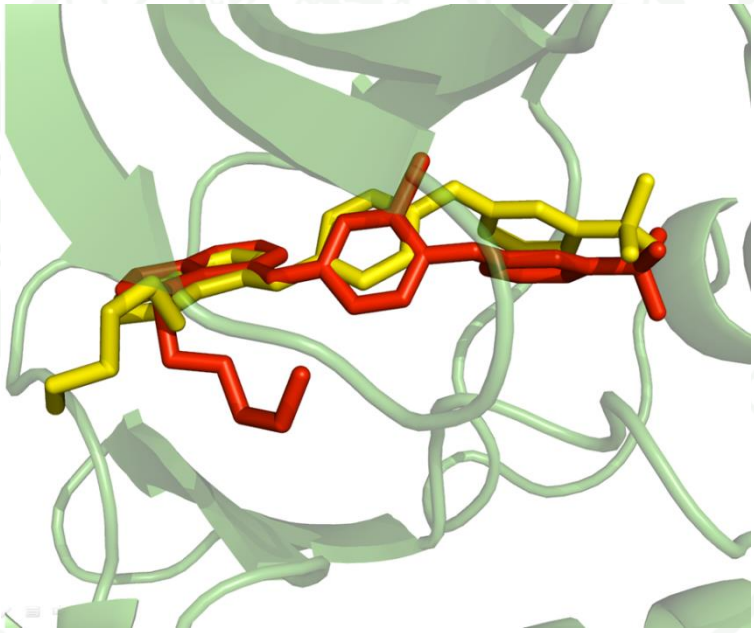


Figure 35 Superimposition of the co-crystallized ligand with self-docked ligand. The structure of the co-crystallized ligand and self-docked ligand were colored in red and yellow, respectively. The structure of ErbB2-TK and its ligand were presented in cartoon and stick, respectively.

The self-docking calculation revealed that the validated parameters provide verisimilitude of structure. As the self-docking result, the validated parameter appropriate for VS calculation. The NCI database composed of 1990 compounds.

This database was used to perform VS with ErbB2-TK. There are 45 compounds of NCI database that provide calculated binding energy lower than calculated binding energy of the co-crystallized ligand. The list of fist 10 compounds was shown in Table 3.

Table 3 The first 10 compounds from NCI database that provide calculated binding energy lower than calculated binding energy of the co-crystallized ligand.

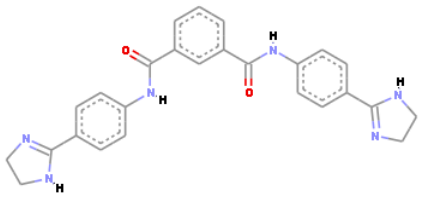
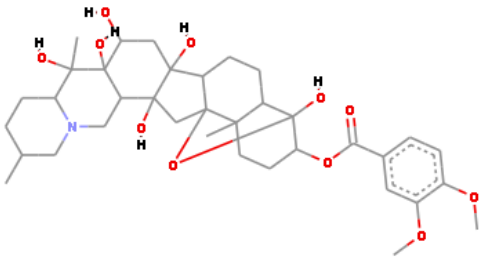
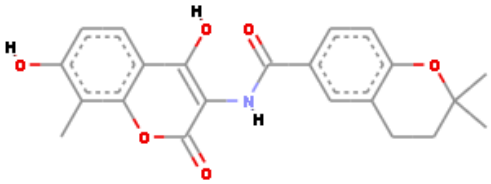
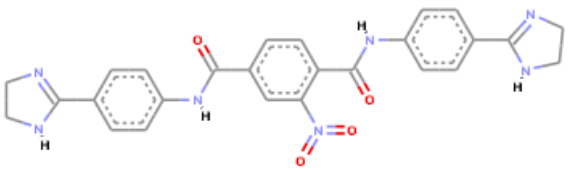
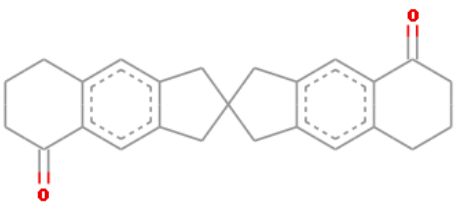
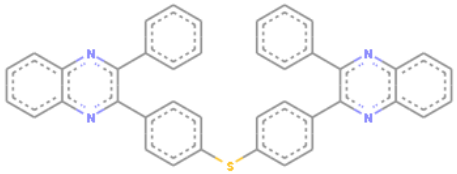
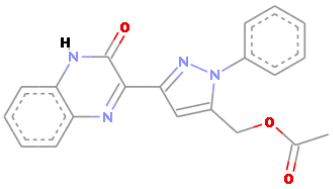
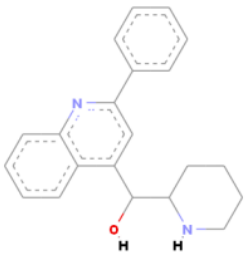
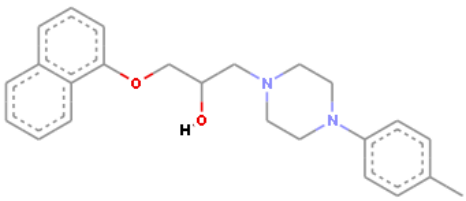
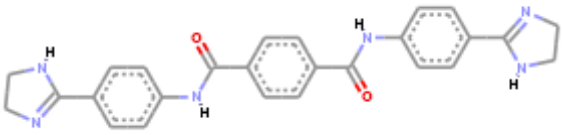
NSC no.	Chemical structure (2D)	Calculated binding energy (kcal/mol)
67436		-9.91
7524		-9.86
5157		-9.83
60340		-9.82
670283		-9.79

Table 3 (Continued)

NSC no.	Chemical structure (2D)	Calculated binding energy (kcal/mol)
293778		-9.76
339161		-9.76
13316		-9.75
91402		-9.75
60339		-9.73

The first 10 docked poses were further analyzed binding interaction by using *Ligand Interaction function* that embedded on Discovery Studio 2.5 program (Accelrys Inc., CA, USA). This function is used to investigate H-bond between receptor and ligand. The binding complexes of first 10 docked poses were shown in Figure 36-45.

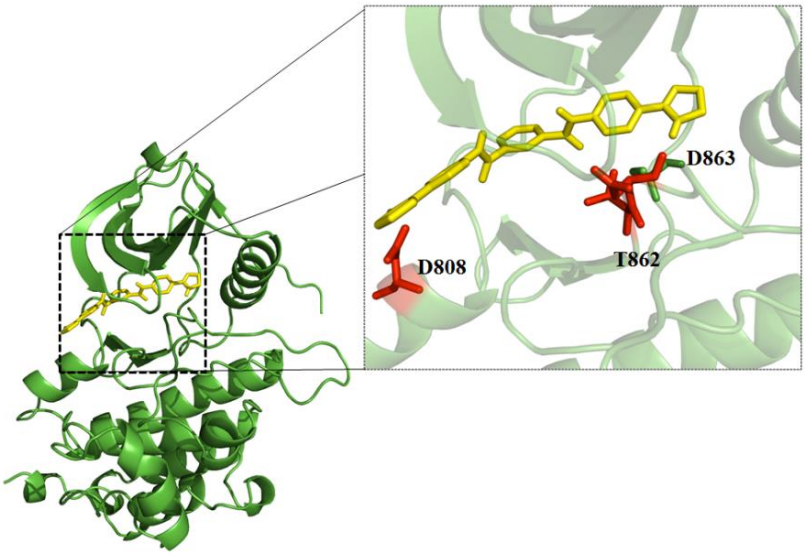


Figure 36 Docked complex between ErbB2-TK and ligand NSC no. 67436. The green ribbon and yellow stick refer as ErbB2-TK and ligand NSC no. 67436 respectively. The detailed-view represented the interaction between ErbB2-TK and ligand NSC no. 67436. The binding interactions reveal the binding residues of ligand NSC no. 67436 are the sugar pocket, phosphate-binding region, and DFG motif of ErbB2-TK.

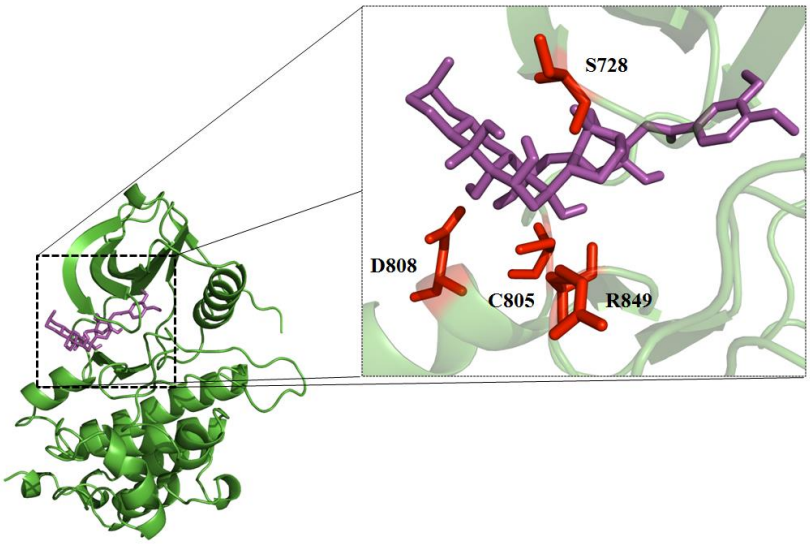


Figure 37 Docked complex between ErbB2-TK and ligand NSC no. 7524. The green ribbon and purple stick refer as ErbB2-TK and ligand NSC no. 7524 respectively. The detailed-view represented the interaction between ErbB2-TK and ligand NSC no. 7524. The binding interactions reveal the binding residues of ligand NSC no. 7524 are the sugar pocket of ErbB2-TK.

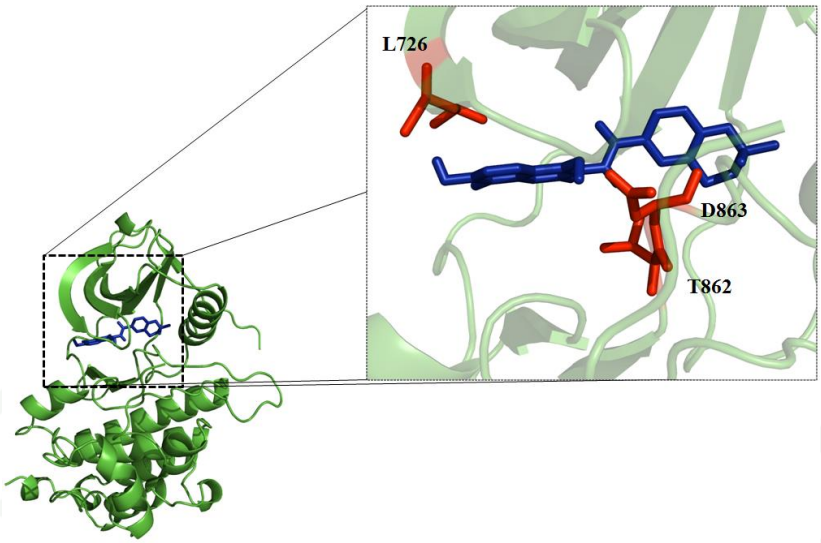


Figure 38 Docked complex between ErbB2-TK and ligand NSC no. 5157. The green ribbon and blue stick refer as ErbB2-TK and ligand NSC no. 5157 respectively. The detailed-view represented the interaction between ErbB2-TK and ligand NSC no. 5157. The binding interactions reveal the binding residues of ligand NSC no. 5157 are the adenine region, phosphate-binding region, and DFG motif of ErbB2-TK.

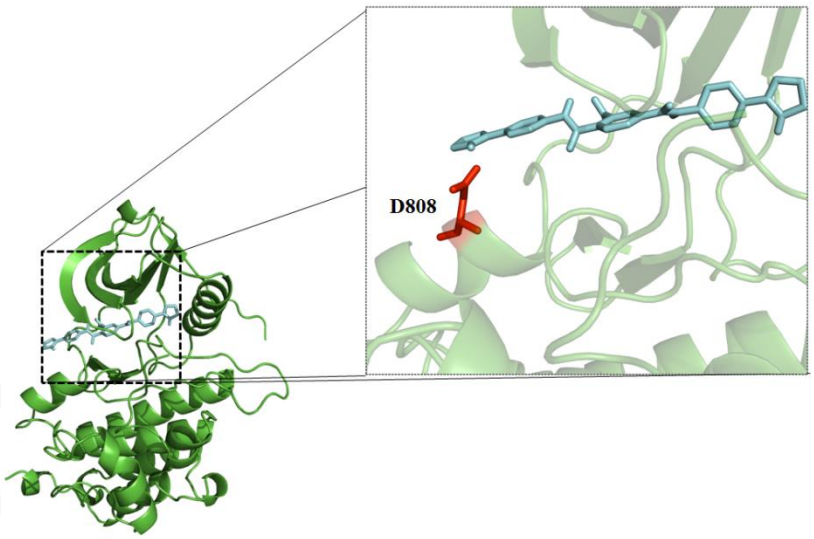


Figure 39 Docked complex between ErbB2-TK and ligand NSC no. 60340. The green ribbon and cyan stick refer as ErbB2-TK and ligand NSC no. 60340 respectively. The detailed-view represented the interaction between ErbB2-TK and ligand NSC no. 60340. The binding interactions reveal the binding residues of ligand NSC no. 60340 are the sugar pocket of ErbB2-TK.

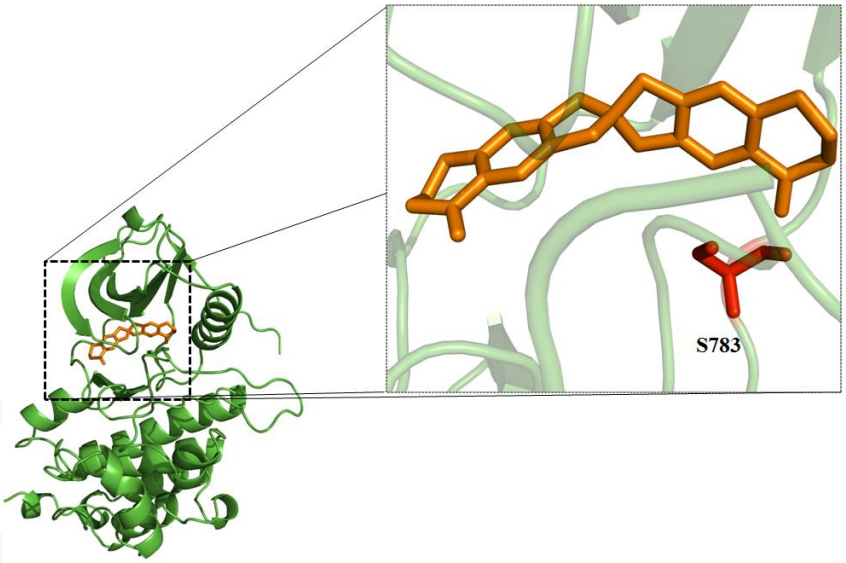


Figure 40 Docked complex between ErbB2-TK and ligand NSC no. 670283. The green ribbon and orange stick refer as ErbB2-TK and ligand NSC no. 670283 respectively. The detailed-view represented the interaction between ErbB2-TK and ligand NSC no. 670283. The binding interactions reveal the binding residues of ligand NSC no. 60340 are the phosphate-binding region of ErbB2-TK.

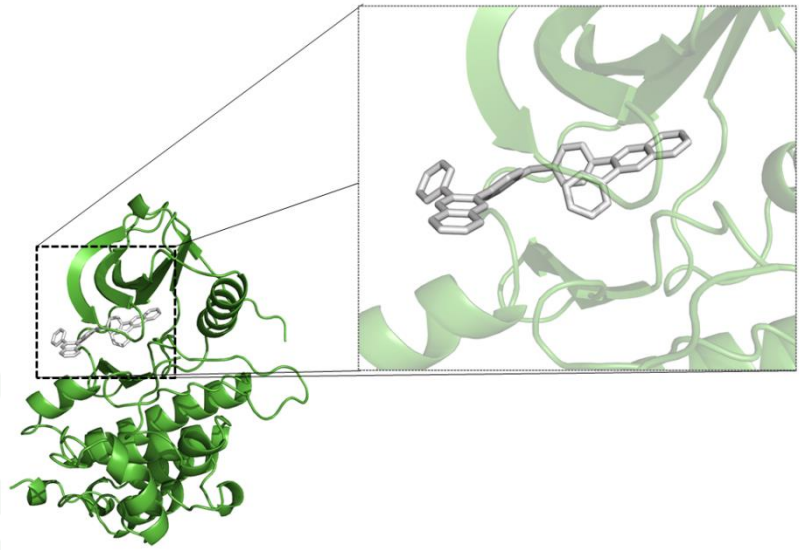


Figure 41 Docked complex between ErbB2-TK and ligand NSC no. 293778. The green ribbon and white stick refer as ErbB2-TK and ligand NSC no. 293778 respectively. The detailed-view represented the interaction between ErbB2-TK and ligand NSC no. 293778. There is no H-bond between ErbB2-TK and of ligand NSC no. 293778. However, the ligand NSC no. 293778 was occupied the ATP binding site.

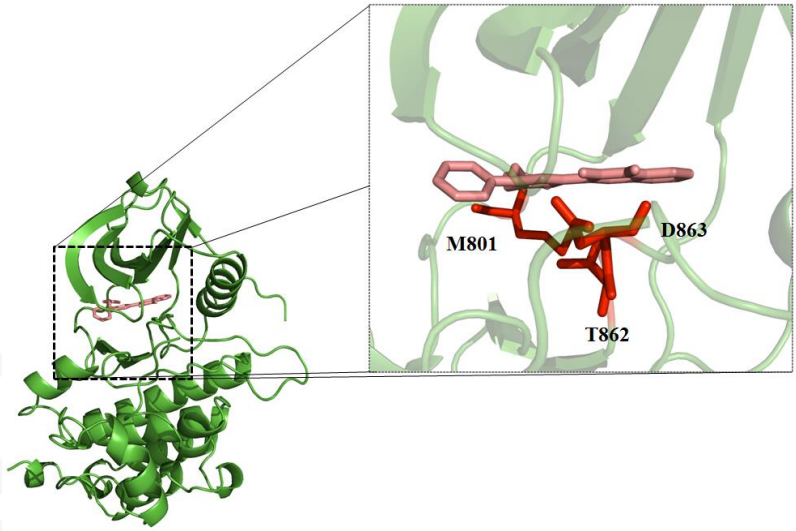


Figure 42 Docked complex between ErbB2-TK and ligand NSC no. 339161. The green ribbon and salmon stick refer as ErbB2-TK and ligand NSC no. 339161 respectively. The detailed-view represented the interaction between ErbB2-TK and ligand NSC no. 339161. The binding interactions reveal the binding residues of ligand NSC no. 339161 are adenine region, phosphate-binding region, and EFG motif of ErbB2-TK.

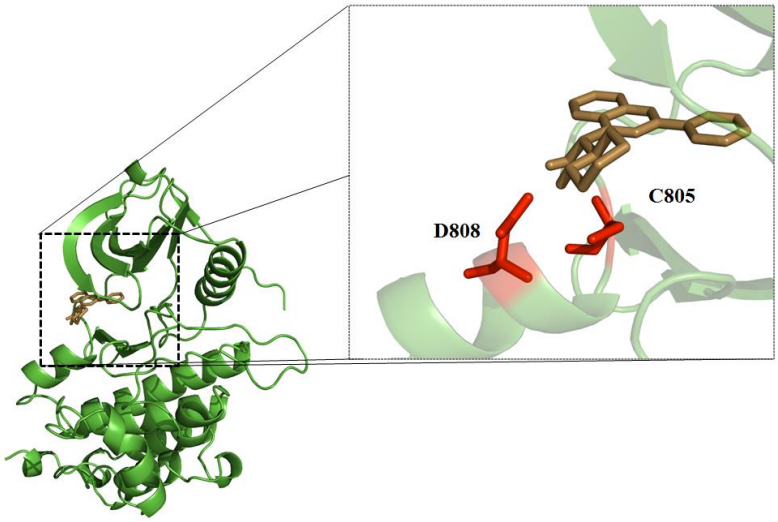


Figure 43 Docked complex between ErbB2-TK and ligand NSC no. 13316. The green ribbon and brown stick refer as ErbB2-TK and ligand NSC no. 13316 respectively. The detailed-view represented the interaction between ErbB2-TK and ligand NSC no. 13316. The binding interactions reveal the binding residues of ligand NSC no. 13316 are sugar pocket of ErbB2-TK.

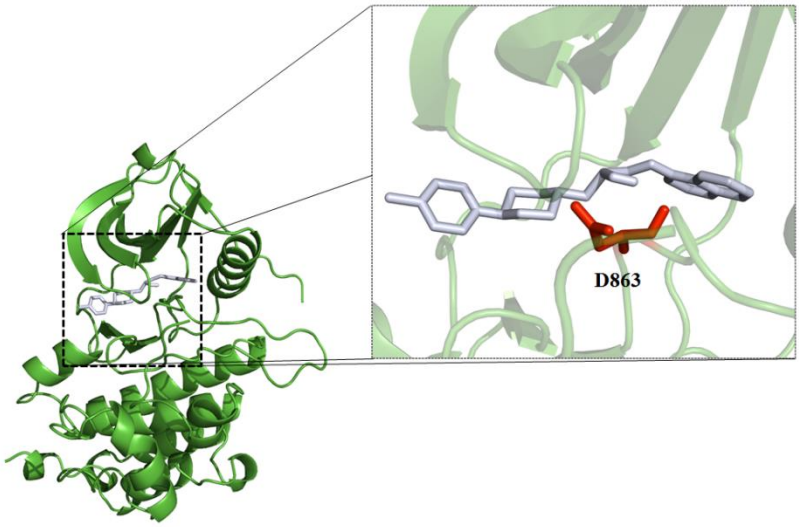


Figure 44 Docked complex between ErbB2-TK and ligand NSC no. 91402. The green ribbon and light blue stick refer as ErbB2-TK and ligand NSC no. 91402 respectively. The detailed-view represented the interaction between ErbB2-TK and ligand NSC no. 91402. The binding interactions reveal the binding residues of ligand NSC no. 91402 are DFG motif of ErbB2-TK.

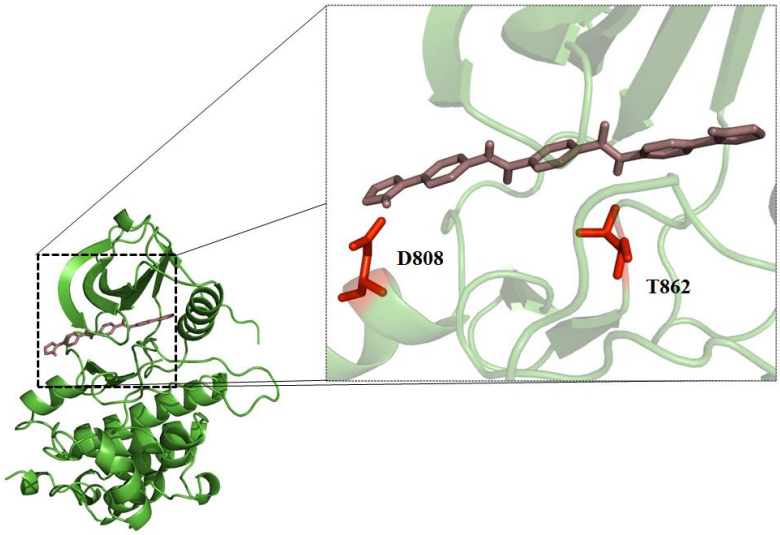
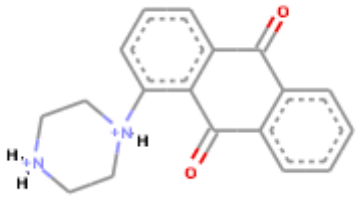
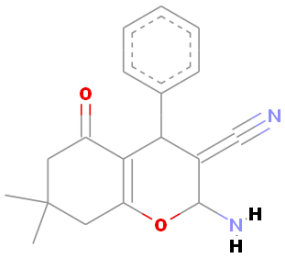
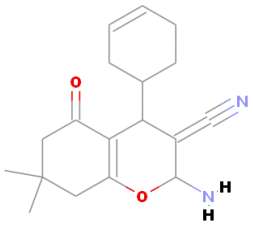
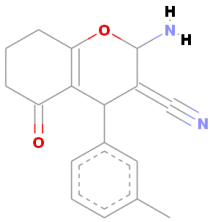
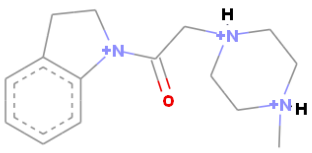


Figure 45 Docked complex between ErbB2-TK and ligand NSC no. 60339. The green ribbon and dark brown stick refer as ErbB2-TK and ligand NSC no. 60339 respectively. The detailed-view represented the interaction between ErbB2-TK and ligand NSC no. 60339. The binding interactions reveal the binding residues of ligand NSC no. 60339 are sugar pocket and phosphate-binding region of ErbB2-TK.

2. Virtual screening of ErbB2-TK against ChemBridge Diverset™

The validated parameters from self-docking were used as parameter for VS against ChemBridge Diverset™. In this study, the 1690 compounds of ChemBridge Diverset™ were used as ligand for VS calculation. There is no compound from ChemBridge Diverset™ that provides the lower binding energy than binding energy of co-crystallized molecule. However, the first 10 compounds that provide lowest binding energy were selected to investigate the binding interaction. The list of fist 10 compounds was shown in Table 4.

Table 4 The first 10 compounds from ChemBridge Diverset™ that provide calculated binding energy lower than calculated binding energy of the co-crystallized ligand.

ChemBridge Diverset™ no.	Chemical structure (2D)	Calculated binding energy (kcal/mol)
2965		-8.69
146		-8.17
1904		-8.15
1023		-8.08
3073		-8.06

The first 10 docked poses were further analyzed binding interaction by using *Ligand Interaction function* that embedded on Discovery Studio 2.5 program (Accelrys Inc., CA, USA). This function is used to investigate H-bond between receptor and ligand. The binding complexes of first 10 docked poses were shown in Figure 46-55.

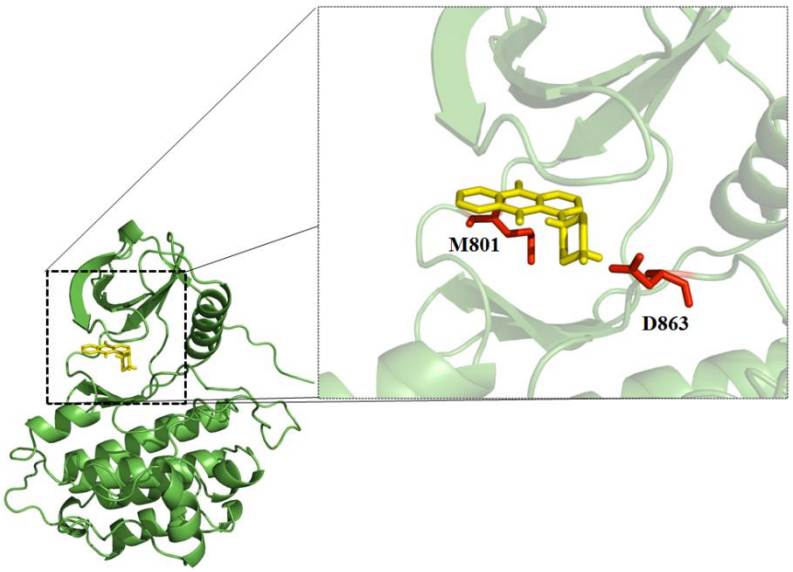


Figure 46 Docked complex between ErbB2-TK and ligand ChemBridge Diverset™ no. 2965. The green ribbon and yellow stick refer as ErbB2-TK and ChemBridge Diverset™ no. 2965 respectively. The detailed-view represented the interaction between ErbB2-TK and ChemBridge Diverset™ no. 2965. The binding interactions reveal the binding residues of ligand ChemBridge Diverset™ no. 2965 are adenine region and DFG motif of ErbB2-TK.

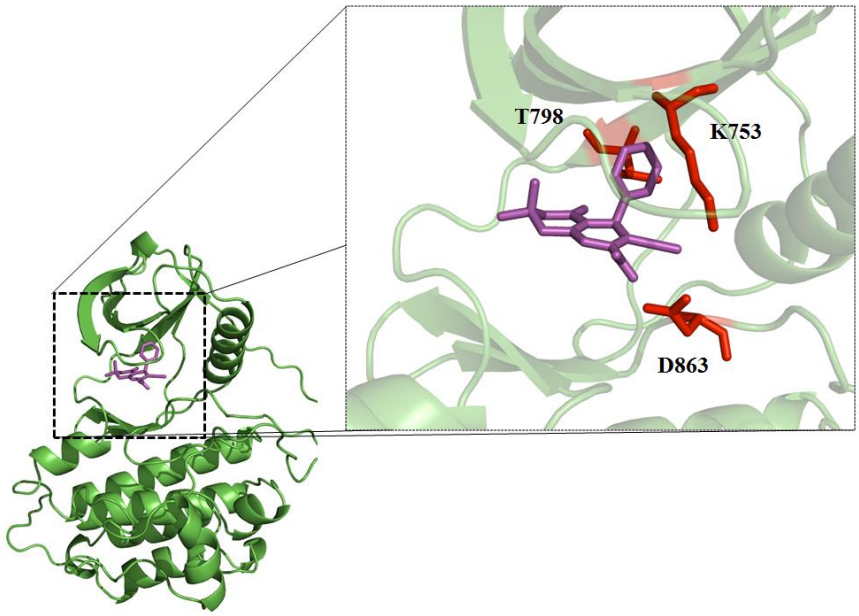


Figure 47 Docked complex between ErbB2-TK and ligand ChemBridge Diverset™ no. 146. The green ribbon and purple stick refer as ErbB2-TK and ChemBridge Diverset™ no. 146 respectively. The detailed-view represented the interaction between ErbB2-TK and ChemBridge Diverset™ no. 146. The binding interactions reveal the binding residues of ligand ChemBridge Diverset™ no. 146 are minor pocket, gate keeper region, and DFG motif of ErbB2-TK.

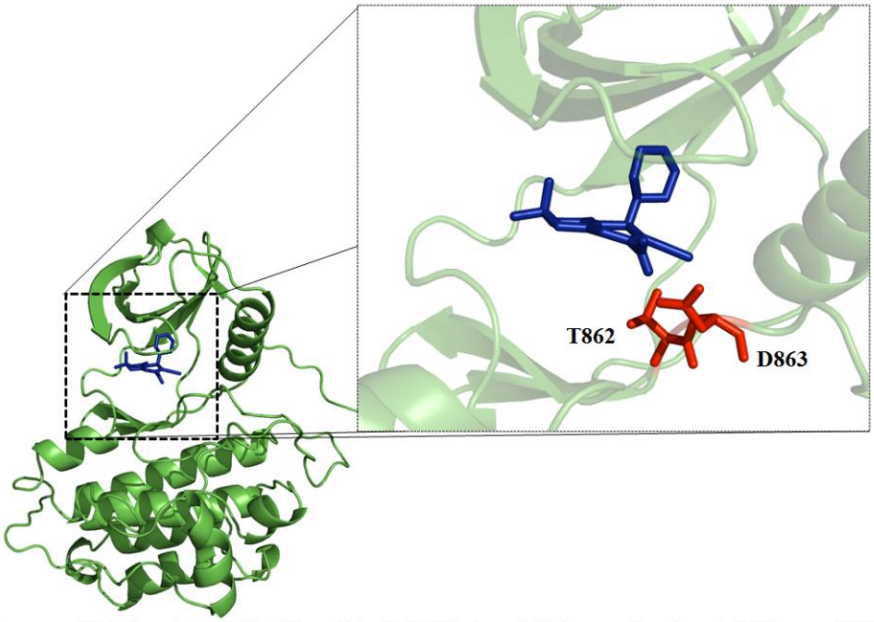


Figure 48 Docked complex between ErbB2-TK and ligand ChemBridge Diverset™ no. 1904. The green ribbon and blue stick refer as ErbB2-TK and ChemBridge Diverset™ no. 1904 respectively. The detailed-view represented the interaction between ErbB2-TK and ChemBridge Diverset™ no. 1904. The binding interactions reveal the binding residues of ligand ChemBridge Diverset™ no. 1904 are DFG motif and phosphate-binding region of ErbB2-TK.

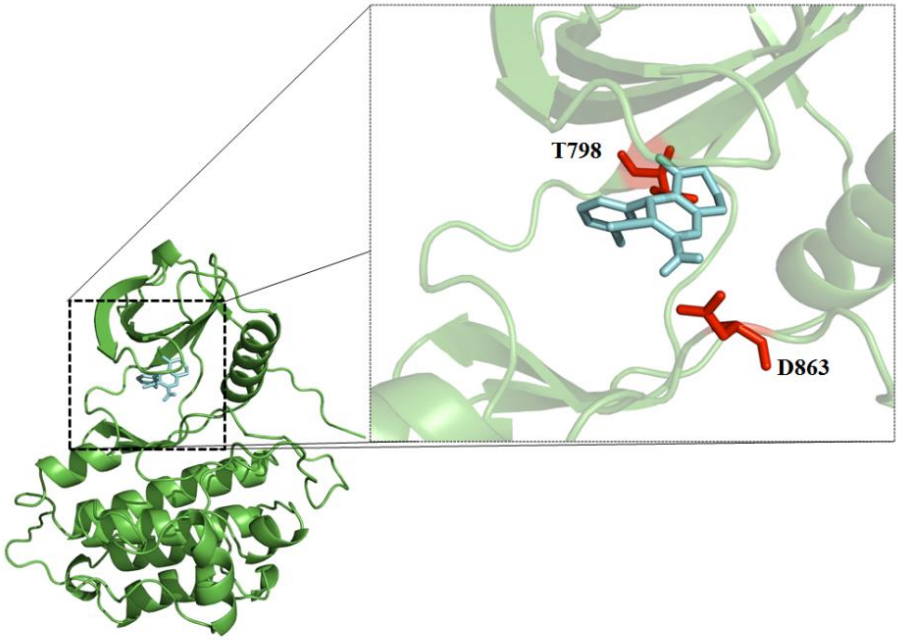


Figure 49 Docked complex between ErbB2-TK and ligand ChemBridge Diverset™ no. 1023. The green ribbon and cyan stick refer as ErbB2-TK and ChemBridge Diverset™ no. 1023 respectively. The detailed-view represented the interaction between ErbB2-TK and ChemBridge Diverset™ no. 1023. The binding interactions reveal the binding residues of ligand ChemBridge Diverset™ no. 1023 are gate keeper region and DFG motif of ErbB2-TK.

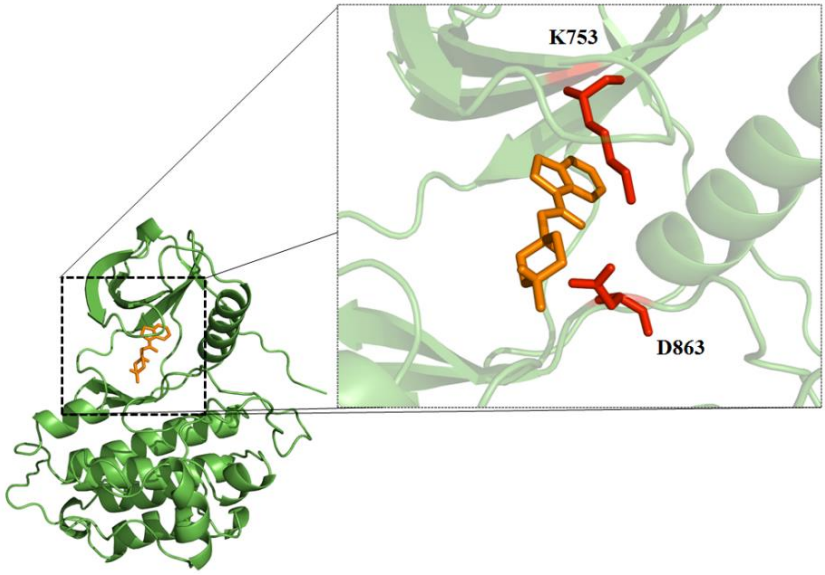


Figure 50 Docked complex between ErbB2-TK and ligand ChemBridge Diverset™ no. 3073. The green ribbon and orange stick refer as ErbB2-TK and ChemBridge Diverset™ no. 3073 respectively. The detailed-view represented the interaction between ErbB2-TK and ChemBridge Diverset™ no. 3073. The binding interactions reveal the binding residues of ligand ChemBridge Diverset™ no. 3073 are minor pocket region and DFG motif of ErbB2-TK.

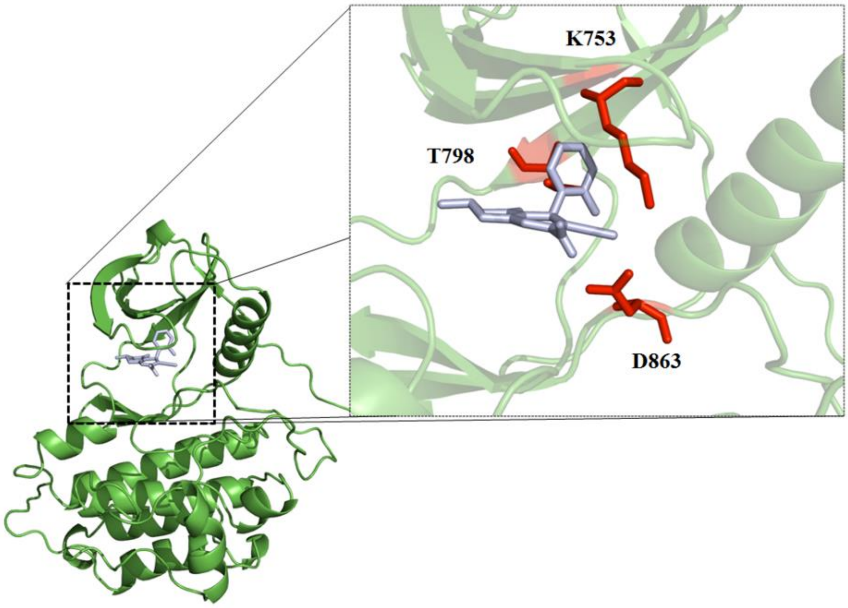


Figure 51 Docked complex between ErbB2-TK and ligand ChemBridge Diverset™ no. 1014. The green ribbon and light blue stick refer as ErbB2-TK and ChemBridge Diverset™ no. 1014 respectively. The detailed-view represented the interaction between ErbB2-TK and ChemBridge Diverset™ no. 1014. The binding interactions reveal the binding residues of ligand ChemBridge Diverset™ no. 1014 are minor pocket region, gate keeper region, and DFG motif of ErbB2-TK.

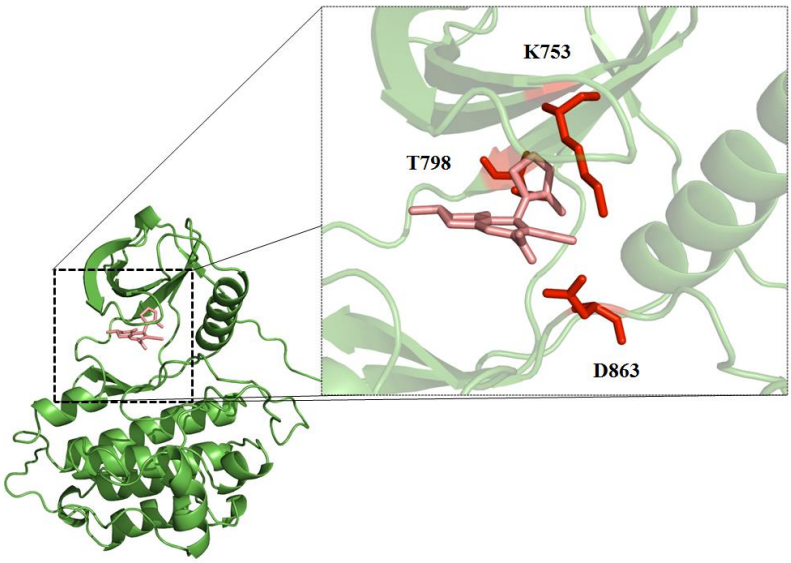


Figure 52 Docked complex between ErbB2-TK and ligand ChemBridge Diverset™ no. 1048. The green ribbon and salmon stick refer as ErbB2-TK and ChemBridge Diverset™ no. 1048 respectively. The detailed-view represented the interaction between ErbB2-TK and ChemBridge Diverset™ no. 1048. The binding interactions reveal the binding residues of ligand ChemBridge Diverset™ no. 1048 are minor pocket region, gate keeper region, and DFG motif of ErbB2-TK.

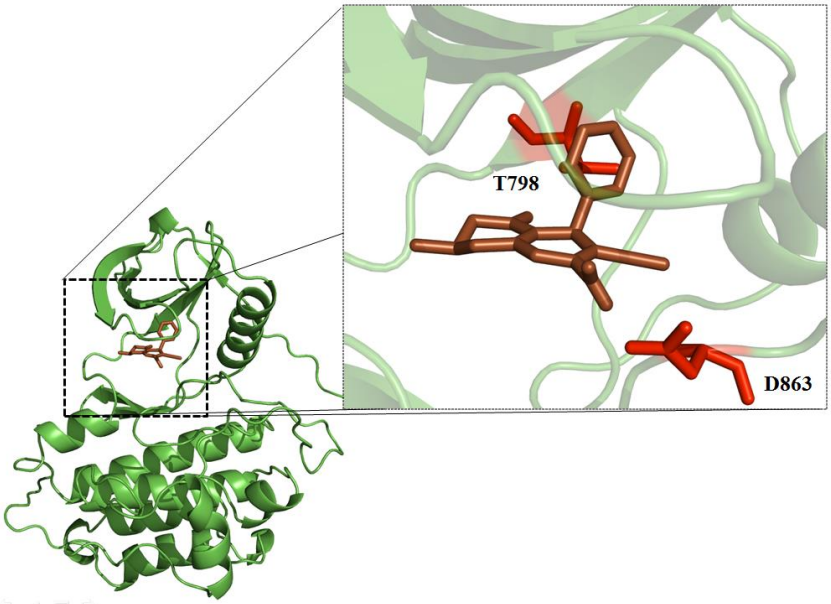


Figure 53 Docked complex between ErbB2-TK and ligand ChemBridge Diverset™ no. 1018. The green ribbon and brown stick refer as ErbB2-TK and ChemBridge Diverset™ no. 1018 respectively. The detailed-view represented the interaction between ErbB2-TK and ChemBridge Diverset™ no. 1018. The binding interactions reveal the binding residues of ligand ChemBridge Diverset™ no. 1018 are gate keeper region and DFG motif of ErbB2-TK.

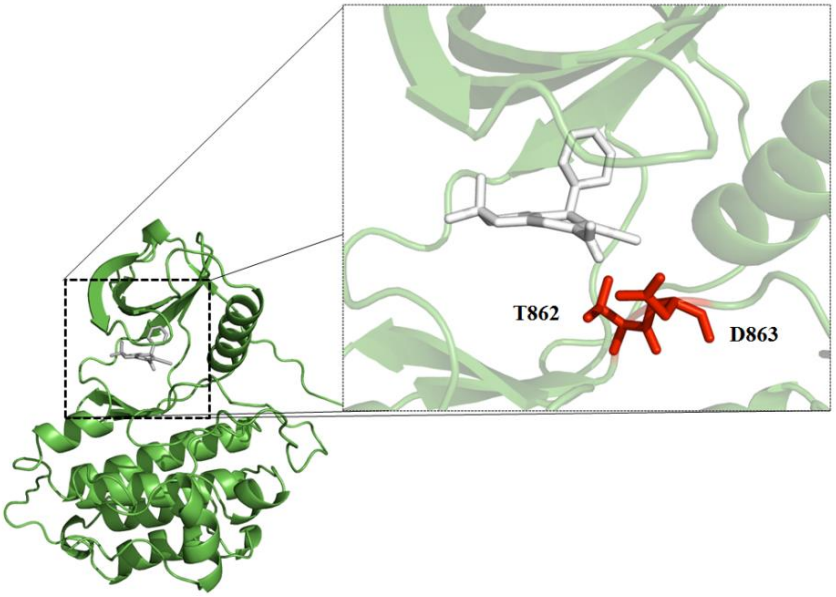


Figure 54 Docked complex between ErbB2-TK and ligand ChemBridge Diverset™ no. 1902. The green ribbon and white stick refer as ErbB2-TK and ChemBridge Diverset™ no. 1902 respectively. The detailed-view represented the interaction between ErbB2-TK and ChemBridge Diverset™ no. 1902. The binding interactions reveal the binding residues of ligand ChemBridge Diverset™ no. 1902 are phosphate-binding region and DFG motif of ErbB2-TK.

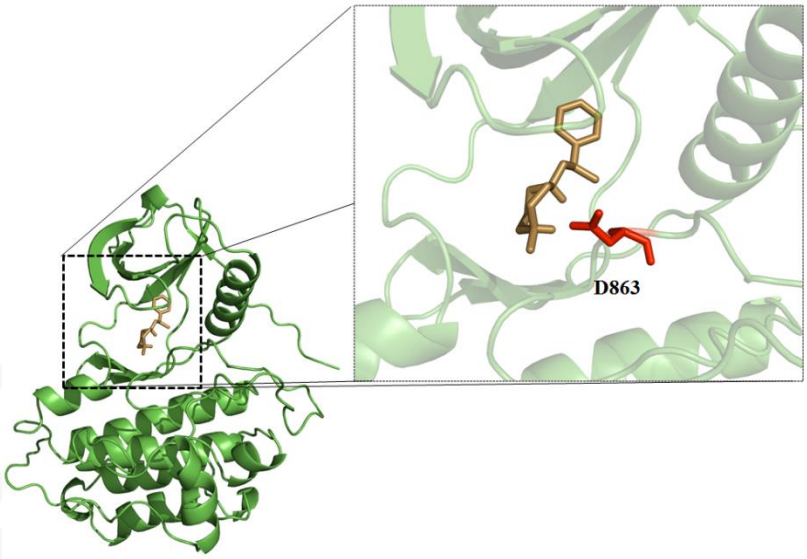


Figure 55 Docked complex between ErbB2-TK and ligand ChemBridge Diverset™ no. 1550. The green ribbon and sand stick refer as ErbB2-TK and ChemBridge Diverset™ no. 1550 respectively. The detailed-view represented the interaction between ErbB2-TK and ChemBridge Diverset™ no. 1550. The binding interactions reveal the binding residues of ligand ChemBridge Diverset™ no. 1550 are DFG motif of ErbB2-TK.

CONCLUSION

The experiments were divided in three parts including screening antibody against ErbB2-TK by using bio-panning technique, computational study of antibody against ErbB2-TK, and screening of lead compound against ErbB2-TK by using virtual screening technique. In antibody screening experiment, the ErbB2-TK and VH/V_HH antibody library were used as an antigen and antibody library, respectively. In bio-panning experiment, there are 28 HB2151 *E. coli* colonies that growth and colonized in LB-AG agar plate. These colonies were selected and subjected to PCR analysis by using primer R1 and R2. There are 15 bacterial colonies carrying *VH/V_HH* genes including colony number 2, 3, 4, 5, 7, 9, 13, 14, 15, 17, 19, 20, 22, and 25. The expected size of *VH/V_HH* gene is 600 bp. The following process was determined the expression of VH/V_HH antibody by using Western blot analysis. There are 8 colonies of HB2151 *E. coli* that shown the expression of VH/V_HH antibody including colony number 4, 17, 22, and 25. These VH/V_HH antibodies were determined the binding and inhibiting properties by using indirect ELISA technique and TK inhibiting activity assay. The indirect ELISA technique was used to determine the binding activity of VH/V_HH antibody against ErbB2-TK and EGFR-TK. The colony number 4, 17, 22, and 25 showed the high indirect ELISA signal against ErbB2-TK. Moreover, the VH/V_HH antibodies against ErbB2-TK were subjected to determine the binding activity against EGFR-TK by using indirect ELISA technique. As the indirect ELISA experiment, the VH/V_HH antibodies against ErbB2-TK colony number 4 and 22 showed the high indirect ELISA signal against EGFR-TK. Subsequently, VH/V_HH antibodies against ErbB2-TK colony number 4 and 22 were subjected to determine the TK inhibiting activity assay. The TK inhibiting activity assay of antibody from colony number 4 and 22 showed the inhibiting properties against EGFR-TK. In addition, the VH/V_HH antibodies were prepared in a different degree of antibody titer. The antibody that provided the highest arbitrary unit of inhibition was VH/V_HH antibody colony number 22 at reciprocal titer of 2. The plasmid of bacterial colonies carrying *VH/V_HH* genes including colony number 4, 17, 22, and 25 were subjected to perform DNA sequencing analysis.

For computational study of antibody against ErbB2-TK, the DNA sequencing of bacterial colonies carrying *VH/V_HH* genes including colony number 4, 17, 22, and 25 were deduced to amino acid sequence. These amino acid sequences were used as initial data to model the 3D of protein structure by using homology modeling technique. Before the docking calculation, the computational methodology was validated by using antibody against tetrodotoxin and RNA dependent RNA polymerase of hepatitis C virus. The method validation showed that, the validated methodology provided the precise docking results that correlated with following data. This data was including mimotope mapping and inhibiting assay. According to the method validation, this methodology was proper for docking calculation. The structures of VH/V_HH antibody clone number 4, 17, 22, and 25 were performed molecular docking against structure of ErbB2-TK. The binding energy of all docking calculation revealed that, the binding reactions were spontaneous reaction resulting from the minus values of the binding energy. The VH/V_HH antibody clone number 4 was reacted with glycine-rich nucleotide phosphate binding loop and activation loop of ErbB2-TK. For VH/V_HH antibody clone number 17, the antibody was reacted with surface loop of ERBB2-TK. The VH/V_HH antibody clone number 22 was reacted with glycine-rich nucleotide phosphate binding loop, DFG motif, and activation loop of ErbB2-TK. The VH/V_HH antibody clone number 25 was reacted with N-terminal residue of ErbB2-TK. As the TK activity inhibiting assay of VH/V_HH antibodies and EGFR-TK, the VH/V_HH antibody clone number 4 and 22 can inhibit the TK activity of EGFR-TK. These VH/V_HH antibodies are potential dual-inhibitor. The structure of VH/V_HH antibody clone number 4 and 22 were docked with EGFR-TK. The VH/V_HH antibody clone number 4 was reacted α helix C, DFG motif, and activation loop of EGFR-TK. For VH/V_HH antibody clone number 17, the antibody was reacted with glycine-rich nucleotide phosphate binding group and catalytic loop of EGFR-TK. The molecular docking results confirmed the results from indirect ELISA technique and TK activity inhibiting assay.

For screening of lead compound against ErbB2-TK by using virtual screening technique, the structure of ErbB2-TK was set as the receptor to screen the lead compound from two libraries including NCI database and ChemBridge Diverset™.

The experiment was calculated by using AutoDock 4.0 program. Before the computational analysis, the computational methodology was validated by using self-docking calculation with the co-crystallized ligand. The calculated binding energy of co-crystallized ligand is -9.11 kcal/mol. For first 10 compounds from NCI database, the range of calculated binding energy was between -9.91 and -9.73. The binding residues compounds from NCI database were involved with the activity of ErbB2. From this experiment, the 10 compounds from NCI database represented the potential competitive inhibiting activity. For first 10 compounds from ChemBridge Diverset™, the range of calculated binding energy was between -8.69 and -7.84. The calculated binding energy of compounds from NCI database had lower binding energy than that of co-crystallized ligand. The calculated binding energy of co-crystallized ligand is lower than that of compounds from ChemBridge Diverset™. However, the binding residues compounds from ChemBridge Diverset™ were involved with the activity of ErbB2. From this experiment, the 10 compounds from ChemBridge Diverset™ represented the potential competitive inhibiting activity.

LITERATURES CITED

Alroy, I. and Y. Yarden. 1997. The Erbb signaling network in embryogenesis and oncogenesis: signal diversification through combinatorial ligand-receptor interactions. **FEBS Lett.** 410 (1): 83-86.

An, Z. 2010. Monoclonal antibodies - a proven and rapidly expanding therapeutic modality for human diseases. **Protein Cell.** 1 (4): 319-330.

Andersen, P.S., A. Stryhn, B.E. Hansen, L. Fugger, J. Engberg and S. Buus. 1996. A recombinant antibody with the antigen-specific, major histocompatibility complex-restricted specificity of T cells. **Proc. Natl. Acad. Sci. USA.** 93 (5): 1820-1824.

Arteaga, C.L. 2003. Erbb-targeted therapeutic approaches in human cancer. **Exp. Cell. Res.** 284 (1): 122-130.

Blair, J.A., D. Rauh, C. Kung, C.H. Yun, Q.W. Fan, H. Rode, C. Zhang, M.J. Eck, W.A. Weiss and K.M. Shokat. 2007. Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. **Nat. Chem. Biol.** 3 (4): 229-238.

Bohacek, R.S., C. McMartin and W.C. Guida. 1996. The art and practice of structure-based drug design: a molecular modeling perspective. **Med. Res. Rev.** 16 (1): 3-50.

Bohm, H.J. 1992. Ludi: rule-based automatic design of new substituents for enzyme inhibitor leads. **J. Comput. Aided. Mol. Des.** 6 (6): 593-606.

Bowie, J.U., R. Luthy and D. Eisenberg. 1991. A method to identify protein sequences that fold into a known three-dimensional structure. **Science** 253 (5016): 164-170.

Brewer, M.R., S.H. Choi, D. Alvarado, K. Moravcevic, A. Pozzi, M.A. Lemmon and G. Carpenter. 2009. The juxtamembrane region of the Egf receptor functions as an activation domain. **Mol. Cell.** 34 (6): 641-651.

Carter, P., L. Presta, C.M. Gorman, J.B. Ridgway, D. Henner, W.L. Wong, A.M. Rowland, C. Kotts, M.E. Carver and H.M. Shepard. 1992. Humanization of an anti-P185her2 antibody for human cancer therapy. **Proc. Natl. Acad. Sci. USA.** 89 (10): 4285-4289.

Casalini, P., M.V. Iorio, E. Galmozzi and S. Menard. 2004. Role of her receptors family in development and differentiation. **J. Cell. Physiol.** 200 (3): 343-350.

Chavanayarn, C., J. Thanongsaksrikul, K. Thueng-In, K. Bangphoomi, N. Sookrung and W. Chaicumpa. 2012. Humanized-single domain antibodies (Vh/Vhh) that bound specifically to naja kaouthia phospholipase A2 and neutralized the enzymatic activity. **Toxins** 4 (7): 554-567.

Chen, J., J.B. Long, A. Hurria, C. Owusu, R.M. Steingart and C.P. Gross. 2012. Incidence of heart failure or cardiomyopathy after adjuvant trastuzumab therapy for breast cancer. **J. Am. Coll. Cardiol.** 2 (12): 4777-4778.

Chen, R. and Z. Weng. 2002. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. **Proteins** 47 (3): 281-294.

Chothia, C. and A.M. Lesk. 1987. Canonical structures for the hypervariable regions of immunoglobulins. **J. Mol. Biol.** 196 (4): 901-917.

Chulanetra, M., K. Bangphoomi, N. Sookrung, J. Thanongsaksrikul, P. Srimanote, Y. Sakolvarvaree, K. Choowongkomon and W. Chaicumpa. 2012. Human Scfv that block sodium ion channel activity of tetrodotoxin. **Toxicon** 59 (2): 272-282.

Clackson, T., H.R. Hoogenboom, A.D. Griffiths and G. Winter. 1991. Making antibody fragments using phage display libraries. **Nature** 352 (6336): 624-628.

Coussens, L., T.L. Yang-Feng, Y.C. Liao, E. Chen, A. Gray, J. McGrath, P.H. Seeburg, T.A. Libermann, J. Schlessinger, U. Francke, A.Levinson and A. Ullrich. 1985. Tyrosine kinase receptor with extensive homology to Egf receptor shares chromosomal location with neu oncogene. **Science** 230 (4730): 1132-1139.

Desmyter, A., S. Spinelli, F. Payan, M. Lauwereys, L. Wyns, S. Muyldermans and C. Cambillau. 2002. Three camelid Vhh domains in complex with porcine pancreatic alpha-amylase. inhibition and versatility of binding topology. **J. Biol. Chem.** 277 (26): 23645-23650.

Dill, K.A., S.B. Ozkan, T.R. Weikl, J.D. Chodera and V.A. Voelz. 2007. The protein folding problem: when will It Be Solved? **Curr. Opin. Struct. Biol.** 17 (3): 342-346.

Dominguez, C., R. Boelens and A.M. Bonvin. 2003. Haddock: a protein-protein docking approach based on biochemical or biophysical information. **J. Am. Chem. Soc.** 125 (7): 1731-1737.

Efferth, T. 2012. Signal transduction pathways of the epidermal growth factor receptor in colorectal cancer and their inhibition by small molecules. **Curr. Med. Chem.** 11.

Ellis, S.E., G.F. Newlands, A.J. Nisbet and J.B. Matthews. 2012. Phage-display library biopanning as a novel approach to identifying nematode vaccine antigens. **Parasite Immunol.** 34 (5): 285-295.

Ewing, T.J., S. Makino, A.G. Skillman and I.D. Kuntz. 2001. Dock 4.0: Search Strategies for Automated Molecular Docking of Flexible Molecule Databases. **J. Comput. Aided. Mol. Des.** 15 (5): 411-428.

Fiszman, G.L. and M.A. Jasnis. 2011. Molecular mechanisms of trastuzumab resistance in Her2 overexpressing breast cancer. **Int. J. Breast. Cancer.** 35 (2): 182-186.

Garrett, T.P., N.M. McKern, M. Lou, T.C. Elleman, T.E. Adams, G.O. Lovrecz, M. Kofler, R.N. Jorissen, E.C. Nice, A.W. Burgess and C.W. Ward. 2003. The crystal structure of a truncated Erbb2 ectodomain reveals an active conformation, poised to interact with other Erbb receptors. **Mol. Cell.** 11 (2): 495-505.

Gavin, A.C., M. Bosche, R. Krause, P. Grandi, M. Marzioch, A. Bauer, J. Schultz, J.M. Rick, A.M. Michon, C.M. Cruciat, M. Remor, C. Hőfer, M. Schelder, M. Brajenovic, H. Ruffer, A. Merino, K. Klein, M. Hudak, D. Dickson, T. Rudi, V. Gnau, A. Bauch, S. Bastuck, B. Huhse, C. Leutwein, M.A. Heurtier, R.R. Copley, A. Edelmann, E. Querfurth, V. Rybin, G. Drews, M. Raida, T. Bouwmeester, P. Bork, B. Saraphin, B. Kuster, G. Neubauer and G. Superti - Furga. 2002. Functional organization of the yeast proteome by systematic analysis of protein complexes. **Nature** 415 (6868): 141-147.

Gottesman, M.M. 2002. Mechanisms of cancer drug resistance. **Annu. Rev. Med.** 53: 615-627.

Haugsten, E.M., A. Wiedlocha, S. Olsnes and J. Wesche. 2010. Roles of fibroblast growth factor receptors in carcinogenesis. **Mol. Cancer Res.** 8 (11): 1439-1452.

Hawkins, R.E., S.J. Russell and G. Winter. 1992. Selection of phage antibodies by binding affinity. mimicking affinity maturation. **J. Mol. Biol.** 226 (3): 889-896.

Hiroyasu, T., Y. Miyabe and H. Yokouchi. 2011. Training data selection method for prediction of anticancer drug effects using a genetic algorithm with local search. **Conf. Proc. IEEE Eng. Med. Biol. Soc.** 124-128.

Hogeweg, P. and B. Hesper. 1978. Interactive instruction on population interactions. **Comput. Biol. Med.** 8 (4): 319-327.

Holbro, T., G. Civenni and N.E. Hynes. 2003. The Erbb receptors and their role in cancer progression. **Exp. Cell Res.** 284 (1): 99-110.

Ishikawa, T., M. Seto, H. Banno, Y. Kawakita, M. Oorui, T. Taniguchi, Y. Ohta, T. Tamura, A. Nakayama, H. Miki, H. Kamiguchi, T. Tanaka, N. Habuka, S. Sagabe and J. Yano. 2011. Design and synthesis of novel human epidermal growth factor receptor 2 (Her2)/epidermal growth factor receptor (Egfr) dual inhibitors bearing a pyrrolo[3,2-D]pyrimidine scaffold. **J. Med. Chem.** 54 (23): 8030-8050.

Kang, A.S., C.F. Barbas, K.D. Janda, S.J. Benkovic and R.A. Lerner. 1991. Linkage of recognition and replication functions by assembling combinatorial antibody fab libraries along phage surfaces. **Proc. Natl. Acad. Sci. USA.** 88 (10): 4363-4366.

Klement, G.L., D. Goukassian, L. Hlatky, J. Carrozza, J.P. Morgan and X. Yan. 2012. Cancer therapy targeting the Her2-Pi3k pathway: potential impact on the heart. **Front. Pharmacol.** 3 (113): 27.

Kolb, P., R.S. Ferreira, J.J. Irwin and B.K. Shoichet. 2009. Docking and chemoinformatic screens for new ligands and targets. **Curr. Opin. Biotechnol.** 20 (4): 429-436.

Kontermann, R.E. 2010. Alternative antibody formats. **Curr. Opin. Mol. Ther.** 12 (2): 176-183.

Kramer, B., M. Rarey and T. Lengauer. 1999. Evaluation of the flexx incremental construction algorithm for protein-ligand docking. **Proteins** 37 (2): 228-241.

Kruif, J., L. Terstappen, E. Boel and T. Logtenberg. 1995. Rapid selection of cell subpopulation-specific human monoclonal antibodies from a synthetic phage antibody library. **Proc. Natl. Acad. Sci. USA.** 92 (9): 3938-3942.

Kuntz, I.D., J.M. Blaney, S.J. Oatley, R. Langridge and T.E. Ferrin. 1982. A geometric approach to macromolecule-ligand interactions. **J. Mol. Biol.** 161 (2): 269-288.

Lang, I.M., C.F. Barbas, 3rd and R.R. Schleef. 1996. Recombinant rabbit Fab with binding activity to type-1 plasminogen activator inhibitor derived from a phage-display library against human alpha-granules. **Gene** 172 (2): 295-298.

Lawrence, M.C. and P.C. Davis. 1992. Clix: A search algorithm for finding novel ligands capable of binding proteins of known three-dimensional structure. **Proteins** 12 (1): 31-41.

Lin, L. and T.G. Bivona. 2012. Mechanisms of resistance to epidermal growth factor receptor inhibitors and novel therapeutic strategies to overcome resistance in Nsclc patients. **Chemother. Res. Pract.** 817297 29.

Malmborg, A.C., M. Duenas, M. Ohlin, E. Soderlind and C.A. Borrebaeck. 1996. Selection of binders from phage displayed antibody libraries using the biacore biosensor. **J. Immunol. Methods.** 198 (1): 51-57.

Marks, J.D., H.R. Hoogenboom, T.P. Bonnert, J. McCafferty, A.D. Griffiths and G. Winter. 1991. By-passing immunization human antibodies from V-gene libraries displayed on phage. **J. Mol. Biol.** 222 (3): 581-597.

Martin, R.B. 1998. Free energies and equilibria of peptide bond hydrolysis and formation. **Biopolymers** 45 (5): 351-353.

McCafferty, J., R.H. Jackson and D.J. Chiswell. 1991. Phage-enzymes: expression and affinity chromatography of functional alkaline phosphatase on the surface of bacteriophage. **Protein Eng.** 4 (8): 955-961.

McGann, M.R., H.R. Almond, A. Nicholls, J.A. Grant and F.K. Brown. 2003. Gaussian docking functions. **Biopolymers** 68 (1): 76-90.

Miles, D., G. von Minckwitz and A.D. Seidman. 2002. Combination versus sequential single-agent therapy in metastatic breast cancer. **Oncologist** 6: 13-19.

Miller, M.D., S.K. Kearsley, D.J. Underwood and R.P. Sheridan. 1994. Flog: a system to select 'Quasi-Flexible' ligands complementary to a receptor of known three-dimensional structure. **J. Comput. Aided. Mol. Des.** 8 (2): 153-174.

Milstein, C. 1999. The hybridoma revolution: an offshoot of basic research. **Bioessays** 21 (11): 966-973.

Mizutani, M.Y., N. Tomioka and A. Itai. 1994. Rational automatic search method for stable gocking models of protein and ligand. **J. Mol. Biol.** 243 (2): 310-326.

Morris, G.M. and M.L. Wilby. 2008. Molecular docking. **Methods Mol. Biol.** 443: 365-382.

_____, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell and A.J. Olson. 2009. Autodock4 and Autodocktools4: Automated Docking with Selective Receptor Flexibility. **J. Comput. Chem.** 30 (16): 2785-2791.

Murray, S., V. Karavasilis, M. Bobos, E. Razis, S. Papadopoulos, C. Christodoulou, P. Kosmidis and G. Fountzilas. 2012. Molecular predictors of response to tyrosine kinase inhibitors in patients with non-small-cell lung cancer. **J. Exp. Clin. Cancer Res.** 31 (1): 77.

Nielsen, D.L., M. Andersson and C. Kamby. 2009. Her2-targeted therapy in breast cancer. monoclonal antibodies and tyrosine kinase inhibitors. **Cancer Treat. Rev.** 35 (2): 121-136.

Nissim, A., H.R. Hoogenboom, I.M. Tomlinson, G. Flynn, C. Midgley, D. Lane and G. Winter. 1994. Antibody fragments from a 'Single Pot' phage display library as immunochemical reagents. **EMBO. J.** 13 (3): 692-698.

Osterhaus, A.D. and A.G. UytdeHaag. 1985. Current developments in hybridoma technology. **Tijdschr. Diergeneeskd.** 110 (20): 835-839.

Oxnard, G.R. 2012. Strategies for overcoming acquired resistance to epidermal growth factor receptor-targeted therapies in lung cancer. **Arch. Pathol. Lab. Med.** 136 (10): 1205-1209.

Pasqualini, R. and E. Ruoslahti. 1996. Organ targeting in vivo using phage display peptide libraries. **Nature** 380 (6572): 364-366.

Pauling, L., R.B. Corey and H.R. Branson. 1951. The structure of proteins; two hydrogen-bonded helical configurations of the polypeptide chain. **Proc. Natl. Acad. Sci. USA.** 37 (4): 205-211.

Quirke, P., A. Pickles, N.L. Tuzi, O. Mohamdee and W.J. Gullick. 1989. Pattern of expression of C-Erbb-2 oncoprotein in human fetuses. **Br. J. Cancer.** 60 (1): 64-69.

Roberts, B.L., W. Markland, K. Siranosian, M.J. Saxena, S.K. Guterman and R.C. Ladner. 1992. Protease inhibitor display M13 phage: selection of high-affinity neutrophil elastase inhibitors. **Gene** 121 (1): 9-15.

Rollinger, J.M., H. Stuppner and T. Langer. 2008a. Virtual screening for the discovery of bioactive natural products. **Prog. Drug. Res.** 65 (211): 213-249.

_____, T.M. Steindl, D. Schuster, J. Kirchmair, K. Anrain, E.P. Ellmerer, T. Langer, H. Stuppner, P. Wutzler and M. Schmidtke. 2008b. Structure-based virtual screening for the discovery of natural inhibitors for human rhinovirus coat protein. **J. Med. Chem.** 51 (4): 842-851.

Rowinsky, E.K. 2004. The Erbb family: targets for therapeutic development against cancer and therapeutic strategies using monoclonal antibodies and tyrosine kinase inhibitors. **Annu. Rev. Med.** 55: 433-457.

Sauton, N., D. Lagorce, B.O. Villoutreix and M.A. Miteva. 2008. Ms-dock: accurate multiple conformation generator and rigid docking protocol for multi-step virtual ligand screening. **BMC Bioinformatics** 9: 184.

Schneider, G. 2002. Trends in virtual combinatorial library design. **Curr. Med. Chem.** 9 (23): 2095-2101.

Shen, Y., X. Yang, N. Dong, X. Xie, X. Bai and Y. Shi. 2007. Generation and selection of immunized fab phage display library against human B cell lymphoma. **Cell. Res.** 17 (7): 650-660.

Sills, M.A., D. Weiss, Q. Pham, R. Schweitzer, X. Wu and J.J. Wu. 2002. Comparison of assay technologies for a tyrosine kinase assay generates different results in high throughput screening. **J. Biomol. Screen.** 7 (3): 191-214.

Skerra, A. and A. Pluckthun. 1988. Assembly of a functional immunoglobulin Fv fragment in *Escherichia Coli*. **Science** 240 (4855): 1038-1041.

Smith, G.P. 1985. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. **Science** 228 (4705): 1315-1317.

Stamos, J., M.X. Sliwkowski and C. Eigenbrot. 2002. Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor. **J. Biol. Chem.** 277 (48): 46265-46272.

Suresh, P.S., A. Kumar, R. Kumar and V.P. Singh. 2008. An in silico [correction of insilico] approach to bioremediation: laccase as a case study. **J. Mol. Graph. Model.** 26 (5): 845-849.

Thornton, J.M. 1981. Disulphide bridges in globular proteins. **J. Mol. Biol.** 151 (2): 261-287.

Thueng-In, K., J. Thanongsaksrikul, P. Srimanote, K. Bangphoomi, O. Poungpair, S. Maneewatch, K. Choowongkomon and W. Chaicumpa. 2012. Cell penetrable humanized-Vh/V(H)H that inhibit RNA dependent RNA polymerase (Ns5b) of Hcv. **PLoS. One.** 7 (11): 8.

Turner, N. and R. Grose. 2010. Fibroblast growth factor signalling: from development to cancer. **Nat. Rev. Cancer.** 10 (2): 116-129.

Tzahar, E., J.D. Moyer, H. Waterman, E.G. Barbacci, J. Bao, G. Levkowitz, M. Shelly, S. Strano, R. Pinkas-Kramarski, J.H. Pierce, G.C. Andrews and Y. Yarden. 1998. Pathogenic poxviruses reveal viral strategies to exploit the Erbb signaling network. **EMBO. J.** 17 (20): 5948-5963.

_____, R. Pinkas-Kramarski, J.D. Moyer, L.N. Klapper, I. Alroy, G. Levkowitz, M. Shelly, S. Henis, M. Eisenstein, B.J. Ratzkin, M. Sela, G.C. Andrews and Y. Yarden. 1997. Bivalence of Egf-like ligands drives the Erbb signaling network. **EMBO. J.** 16 (16): 4938-4950.

Uetz, P., L. Giot, G. Cagney, T.A. Mansfield, R.S. Judson, J.R. Knight, D. Lockshon, V. Narayan, M. Srinivasan, P. Pochart and et al. 2000. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae.* **Nature** 403 (6770): 623-627.

Vakser, I.A. 1997. Evaluation of gramm low-resolution docking methodology on the hemagglutinin-antibody complex. **Proteins** 1: 226-230.

Vaughan, T.J., A.J. Williams, K. Pritchard, J.K. Osbourn, A.R. Pope, J.C. Earnshaw, J. McCafferty, R.A. Hodits, J. Wilton and K.S. Johnson. 1996. Human antibodies with sub-nanomolar affinities isolated from a large non-immunized phage display library. **Nat. Biotechnol.** 14 (3): 309-314.

Ward, R.L., M.A. Clark, J. Lees and N.J. Hawkins. 1996. Retrieval of human antibodies from phage-display libraries using enzymatic cleavage. **J. Immunol. Methods.** 189 (1): 73-82.

Welch, W., J. Ruppert and A.N. Jain. 1996. Hammerhead: fast, fully automated docking of flexible ligands to protein binding sites. **Chem. Biol.** 3 (6): 449-462.

Wood, E.R., A.T. Truesdale, O.B. McDonald, D. Yuan, A. Hassell, S.H. Dickerson, B. Ellis, C. Pennisi, E. Horne, K. Lackey, K.J. Alligood, D.W. Rusnak, T.M. Gilmer and L. Shewchuk. 2004. A unique structure for epidermal growth factor receptor bound to Gw572016 (lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells. **Cancer Res.** 64 (18): 6652-6659.

Xu, G., M.C. Abad, P.J. Connolly, M.P. Neeper, G.T. Struble, B.A. Springer, S.L. Emanuel, N. Pandey, R.H. Gruninger, M. Adams, S. Moreno-Mazza, A.R. Fuentes-Pesquera and S.A. Middleton. 2008a. 4-amino-6-arylamino-pyrimidine-5-carbaldehyde hydrazones as potent Erbb-2/Egfr dual kinase inhibitors. **Bioorg. Med. Chem. Lett.** 18 (16): 4615-4619.

_____, L.L. Searle, T.V. Hughes, A.K. Beck, P.J. Connolly, M.C. Abad, M.P. Neeper, G.T. Struble, B.A. Springer, S.L. Emanuel, R.H. Gruninger and et al. 2008b. Discovery of novel 4-amino-6-arylaminopyrimidine-5-carbaldehyde oximes as dual inhibitors of Egfr and Erbb-2 protein tyrosine kinases. **Bioorg. Med. Chem. Lett.** 18 (12): 3495-3499.

Yamanaka, H.I., T. Inoue and O. Ikeda-Tanaka. 1996. Chicken monoclonal antibody isolated by a phage display system. **J. Immunol.** 157 (3): 1156-1162.

Yeon, C.H. and M.D. Pegram. 2005. Anti-Erbb-2 antibody trastuzumab in the treatment of Her2-amplified breast cancer. **Invest. New Drugs.** 23 (5): 391-409.

Yu, D. and M.C. Hung. 2000. Overexpression of Erbb2 in cancer and Erbb2-targeting strategies. **Oncogene** 19 (53): 6115-6121.

Yun, C.H., T.J. Boggon, Y. Li, M.S. Woo, H. Greulich, M. Meyerson and M.J. Eck. 2007. Structures of lung cancer-derived Egfr mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity. **Cancer Cell.** 11 (3): 217-227.

_____, K.E. Mengwasser, A.V. Toms, M.S. Woo, H. Greulich, K.K. Wong, M. Meyerson and M.J. Eck. 2008. The T790m mutation in Egfr kinase causes drug resistance by increasing the affinity for ATP. **Proc. Natl. Acad. Sci. USA.** 105 (6): 2070-2075.

Zavodszky, M.I., P.C. Sanschagrin, R.S. Korde and L.A. Kuhn. 2002. Distilling the essential features of a protein surface for improving protein-ligand docking, scoring, and virtual screening. **J. Comput. Aided. Mol. Des.** 16 (12): 883-902.

Zhang, W., L.P. Stabile, P. Keohavong, M. Romkes, J.R. Grandis, A.M. Traynor and J.M. Siegfried. 2006. Mutation and polymorphism in the Egfr-Tk domain associated with lung cancer. **Journal of thoracic oncology: official publication of the International Association for the Study of Lung Cancer.** 1 (7): 635-647.

Zhang, X., K.A. Pickin, R. Bose, N. Jura, P.A. Cole and J. Kuriyan. 2007. Inhibition of the Egf receptor by binding of Mig6 to an activating kinase domain interface. **Nature** 450 (7170): 741-744.

Zhang, Y. and J. Skolnick. 2005. The protein structure prediction problem could be solved using the current Pdb library. **Proc. Natl. Acad. Sci. USA.** 102 (4): 1029-1034.

_____, Z. Yan, A. Farooq, X. Liu, C. Lu, M.M. Zhou and C. He. 2004. Molecular basis of distinct interactions between Dok1 Ptb domain and tyrosine-phosphorylated Egf receptor. **J. Mol. Biol.** 343 (4): 1147-1155.

APPENDICES

Appendix A

Reagents and buffers for agarose gel electrophoresis

Reagents and buffers for agarose gel electrophoresis

1. Gel loading buffer (loading dye)

The loading dye buffer composed of 0.25% bromphenol blue, 0.25% xylene cyanol, 30% glycerol and 35 ml of ultrapure distilled water. The loading dye solution was kept at 4°C.

2. Tris acetate buffer (50x TAE)

The stock 50x TAE was prepared by dissolved 242 grams of Tris-base in 500 ml of distilled water. After the ingredient was completely dissolved, 57.1 ml of concentrate glacial acetic acid and 100 ml of 0.5 M EDTA, pH 8.0, were added into the solution. The final volume was adjusted to 1,000 ml by distilled water. The 50x TAE was stored at 25°C. The 1x working solution was freshly prepared by diluting the stock 50x TAE buffer with distilled water.

3. Working (1x TAE)

Twenty milliliter of 50X TAE was added to 980 ml of UDW. This solution can be reused three times.

Appendix B

Media for *E. coli* culture

Media for *E. coli* culture

1. LB broth

This medium consists of:

Bacto tryptone	10	g
Bacto yeast extract	5	g
NaCl	5	g

The broth was prepared by completely dissolving all of the above reagents in DW to final volume of one liter. The preparation was sterilized by autoclaving. The broth was stored at 4°C.

2. LB agar plates

This medium consists of:

Bacto tryptone	10	g
Bacto yeast extract	5	g
NaCl	5	g
Agar	15	g

The agar was prepared by dissolving all of the above reagents in DW to final volume of one liter and then sterilized by autoclaving. The solution was poured into petridishes (25 ml/plate), and allowed for solidification at 25°C and stored at 4°C.

3. LB-ampicillin broth (100 μg ampicillin/ml)

The medium was prepared by mixing 400 μl of 250 mg/ml sterilized ampicillin with 1 liter of the autoclaved LB-broth. The medium was then mixed by swirling.

4. LB-ampicillin agar (100 μg ampicillin/ml)

One liter of LB agar was prepared, autoclaved and cooled down to 55°C. Then 400 μl of 250 mg/ml sterilized ampicillin was added to the agar. After gently mixing by swirling, the preparation was poured into petridishes (25 ml/plate). The plates were stored at 4°C.

Appendix C

Reagents for extraction of plasmid DNA from *E. coli*

Reagents for extraction of plasmid DNA from *E. coli*

1. Tris-HCl, pH 8.0 stock solution (2 M)

The solution was prepared by dissolving 242.28 g of Tris base in a volume of UDW. The pH was adjusted to 8.0 with HCl; UDW was added to a final volume of one liter and the preparation was autoclaved.

2. Glucose stock solution (2M)

The solution was prepared by dissolving 36 g of glucose in 100 ml of UDW. The solution was steriled by filtering and it was stored at -20°C until use.

3. EDTA stock solution (0.5 M)

The solution was prepared by dissolving 18.6 g of EDTA Na_2 $2H_2O$ in a volume of UDW. The pH was adjusted to 8.0 with NaOH, added UDW to final volume of 100 ml and autoclaved.

4. Solution I (25 M Tris pH 8.0, 50 mM glucose, 10 mM EDTA)

2.0 M Tris-HCl, pH 8.0	1.25	ml
2.0 M glucose	2.5	ml
0.5 M EDTA pH 8.0	200	µl

All components were dissolved with UDW to final volume of 100 ml.

5. Solution II (0.1 N NaOH, 1% SDS) (freshly prepared)

The solution was freshly prepared by mixing one milliliter of 1 N NaOH with 0.5 ml of 20% SDS and then UDW was added to bring the final volume to 10 ml.

6. Solution III (2.7 M potassium acetate, pH 4.8)

The solution was prepared by dissolving 26.5 g of potassium acetate in a volume of UDW. The solution was adjusted to pH 4.8 with glacial acetic acid and UDW was added to bring the final volume to 100 ml.

Appendix D

Reagents for SDS-PAGE

Reagents for SDS-PAGE

1. Sample buffer (SDS reducing buffer)

The sample buffer was prepared as a stock solution by combination the following ingredients:

0.5 M Tris-HCL, pH 6.8	1.0	ml
Glycerol	2.0	ml
SDS (10% solution)	1.6	ml
0.05% Bromophenol blue	0.2	ml
2-6-mercaptoethanol	0.4	ml
UDW	2.8	ml

This mixture was stored at 25°C in small aliquots. One part of sample was diluted with equal part of the sample buffer and heated at 100°C for 4 min before loading into gel.

2. Tris-HCl (1.5 M, pH 8.8)

To prepare this solution, 18.15 g of Tris base was dissolved in 50 ml of UDW, then the pH was adjusted to 8.8 with 1 N HCl. The final volume was brought up to 100 ml with UDW. The solution was filtered through sterile a 0.2 μm membrane. This stock solution was stored at 4°C until use for preparing a working solution.

3. Tris-HCl (0.5 M, pH 6.8)

To prepare this solution, 6.05 g of Tris base was dissolved in 50 ml of UDW, then the pH was adjusted to 6.8 with 1 N HCl. The final volume was brought up to 100 ml with UDW. The solution was filtered through a sterile 0.22 μm membrane. This stock solution was stored at 4°C.

4. Sodium dodecyl sulfate (10% SDS; w/v)

This solution was prepared by dissolving 10 g of SDS in 100 ml of UDW.

5. Ammonium persulfate (10%; w/v)

This solution was prepared just before use by dissolving 50 mg of ammonium persulfate (Bio-Rad) in 0.5 ml of UDW.

6. Separating gel (12%)

Polyacrylamide separating gel (12%) was prepared by mixing the following ingredients together:

UDW	3.35	ml
1.5 M Tris-HCl, pH 8.8	2.5	ml
10% SDS solution	100	µl
30% Acrylamide/Bis, 29:1 ratio solution (BioRad)	4	ml

The reagents were gently mixed and degassed under a vacuum for at least 5 min. The polymerization was initiated by adding 50 µl of the 10% ammonium persulfate (freshly prepared) and 5 µl of TEMED (Bio-Rad). The gel was poured into the casting apparatus, over-layered with UDW and allowed to polymerize for at least 20 min at 25°C.

7. Stacking gel (4%)

The stacking gel (4%) was prepared by mixing the following reagents:

UDW	6.0	ml
0.5 M Tris-HCl	2.5	ml
SDS (10% solution w/v)	0.1	ml
30% Acrylamide/Bis, 29:1 ratio solution (BioRad)	4	ml

All reagents were mixed gently and degassed under a vacuum for 15 min, and then 50 µl of freshly prepared 10% ammonium persulfate and 10 µl of TEMED were subsequently added, respectively. After complete mixing and degassing, the upper portion of the gel polymerized in the casting apparatus was rinsed with UDW, the comb was inserted between the glass plates over the polymerized separating gel. The stacking gel was poured and allowed to polymerize for at least 45 min at 25°C before use.

8. Electrode (running) buffer (pH 8.3; 10 x)

The buffer contained the following reagents: 30.3 g of Tris base; 142.9 g of glycine and 10 g of SDS. The buffer was prepared by dissolving all of the above reagents in a volume of UDW. After all ingredients were dissolved, the volume was made up to one liter with UDW. The buffer was stored at 4°C until use for preparing a working electrode (running) buffer.

9. Working electrode (running) buffer (1x)

One hundred ml of the 10x electrode buffer was diluted with 900 ml of UDW. Each preparation of the working running buffer was used for only one electrophoretic run.

10. Coomassie Brilliant Blue stain

10.1 Coomassie® Brilliant Blue R-250 dye solution

Coomassie® Brilliant Blue R-250 dye (USB Corporation, USA) (2.5 g) was dissolved in 454 ml of absolute methanol before 92 ml of glacial acetic acid and 454 ml of UDW were added. This dye was filtered through a Whatman No. 1 paper and kept at 25°C until use.

10.2 Destaining solution

10.2.1 High methanol destain solution

The solution was prepared by mixing 75 ml of glacial acetic acid, 454 ml of methanol and 25 ml of glycerol together. UDW was added to make 1,000 ml. The solution was kept at 25°C until use.

10.2.2 Standard (low-methanol) destain solution

The solution was prepared by mixing 75 ml of glacial acetic acid, 50 ml of methanol and 25 ml of glycerol together. UDW was added to make 1,000 ml. The solution was kept at 25°C until use.

Appendix E

Reagents for Western blot analysis

Reagents for Western blot analysis

1. Transfer buffer [25mM Tris, 192 mM glycine and 20% (v/v) methanol]

To prepare 1,000 ml of this buffer, 3.03 g of Tris base and 14.4 g of glycine were dissolved in 800 ml of UDW. Subsequently, 200 ml of methanol was added to yield 20% (v/v).

2. Phosphate buffered saline (0.01 M PBS, pH 7.4)

This solution was prepared by dissolving 1.22 g of anhydrous Na_2HPO_4, 0.17 g of anhydrous NaH_2PO_4 and 8.77 g of NaCl in 1 liter of DW. The pH of this solution was adjusted to 7.4 with 1 N HCl.

3. Phosphate buffer (1/15 M PB, pH 7.6)

The buffer was prepared by dissolving 0.06 g of NaH_2PO_4 and 0.47 g of Na_2HPO_4 in 57.7 ml of UDW. The pH of this solution was adjusted to 7.6 with 1 N HCl.

4. Washing buffer (0.05% Tween-20 in PBS, pH 7.4; PBST)

This solution was prepared by adding 0.5 ml of Tween-20 in one liter of 0.01 M PBS (pH 7.4) and mixed well.

5. Blocking solution (3% BSA, in PBS, pH 7.4)

The solution was prepared by dissolving 3 g of bovine serum albumin (BSA) in 100 ml of 0.01 M PBS, pH 7.4.

6. Diluent solution (0.2% BSA, 0.2% gelatin in PBS, pH 7.4)

The solution was prepared by dissolving 0.2 g of BSA and 0.2 g of gelatin in 100 ml of 0.01 M PBS, pH 7.4.

7. Substrate solution

The solution was freshly prepared by dissolving 0.02 g of 2, 6-dichloro-phenol indophenol (Sigma Chemical Co.) in 10 ml of 1/15 M PB (pH 7.6). Ten μl of 30% H_2O_2 was added immediately before use.

Appendix F

Reagents for ELISA

Reagents for ELISA

1. Coating buffer (Carbonate-bicarbonate buffer, pH 9.6)

The buffer was prepared by dissolving 1.26 g of $NaHCO_3$ in 300 ml of DW; then the pH was adjusted to 9.6 with 0.05 Na_2CO_3 (0.53 g in 100 ml of DW).

2. Phosphate buffered saline (0.01 M PBS, pH 7.4)

The solution was prepared by dissolving 1.22 g of anhydrous Na_2HPO_4, 0.17 g of anhydrous NaH_2PO_4 and 8.77 g of NaCl in one liter of DW. The pH of this solution was adjusted to 7.4 with 1 N HCl.

3. Washing solution (PBST)

Washing solution (PBST) was prepared by mixing Tween-20 in PBS, pH 7.4 to a 0.05 % concentration.

4. Blocking solution

The solution was prepared by dissolving 1 g of BSA in 100 ml of 0.01 M PBS, pH 7.4.

5. Diluents solution

The diluents was prepared by dissolving 0.2 g of BSA and 0.2 g of gelatin in 100 ml of 0.01 M PBS, pH 7.4.

6. Substrate buffer (0.1 M citrate buffer, pH 4.5)

The buffer was prepared by dissolving 14.7 g of trisodium citrate ($Na_3C_6H_5O_7.H_2O$) in DW. The volume was made up to 500 ml after the pH was adjusted to 4.5 with 1M HCl.

7. Substrate solution

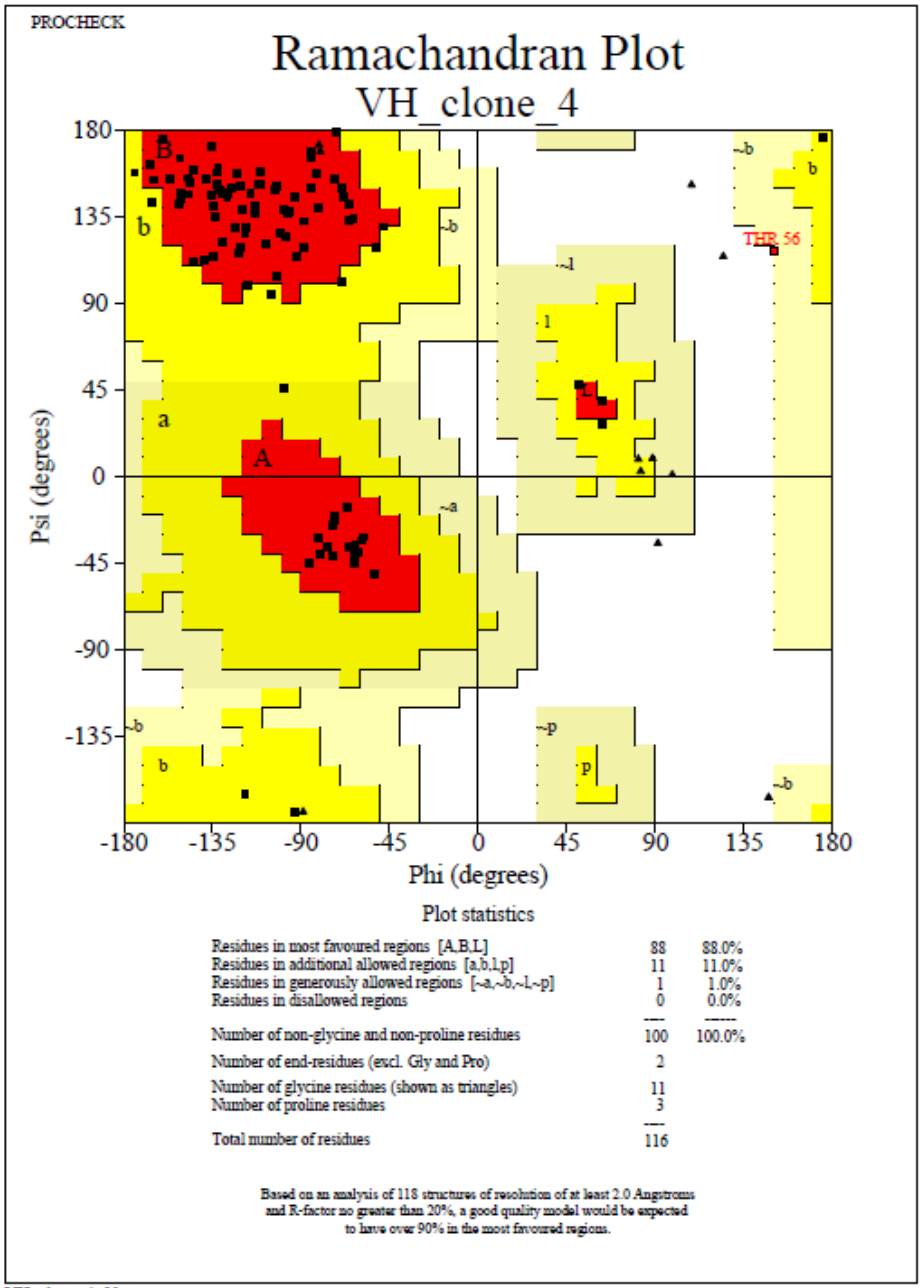
The substrate solution consisted of 0.05% 1, 4-*p*-phenylenediamine-dihydrochloride (PPD) (Sigma Chemical Co.) in citrate buffer, pH 4.5 and 0.01% of 30% H_2O_2. This solution was prepared freshly before use and always protected from light.

8. Stop solution (1N NaOH)

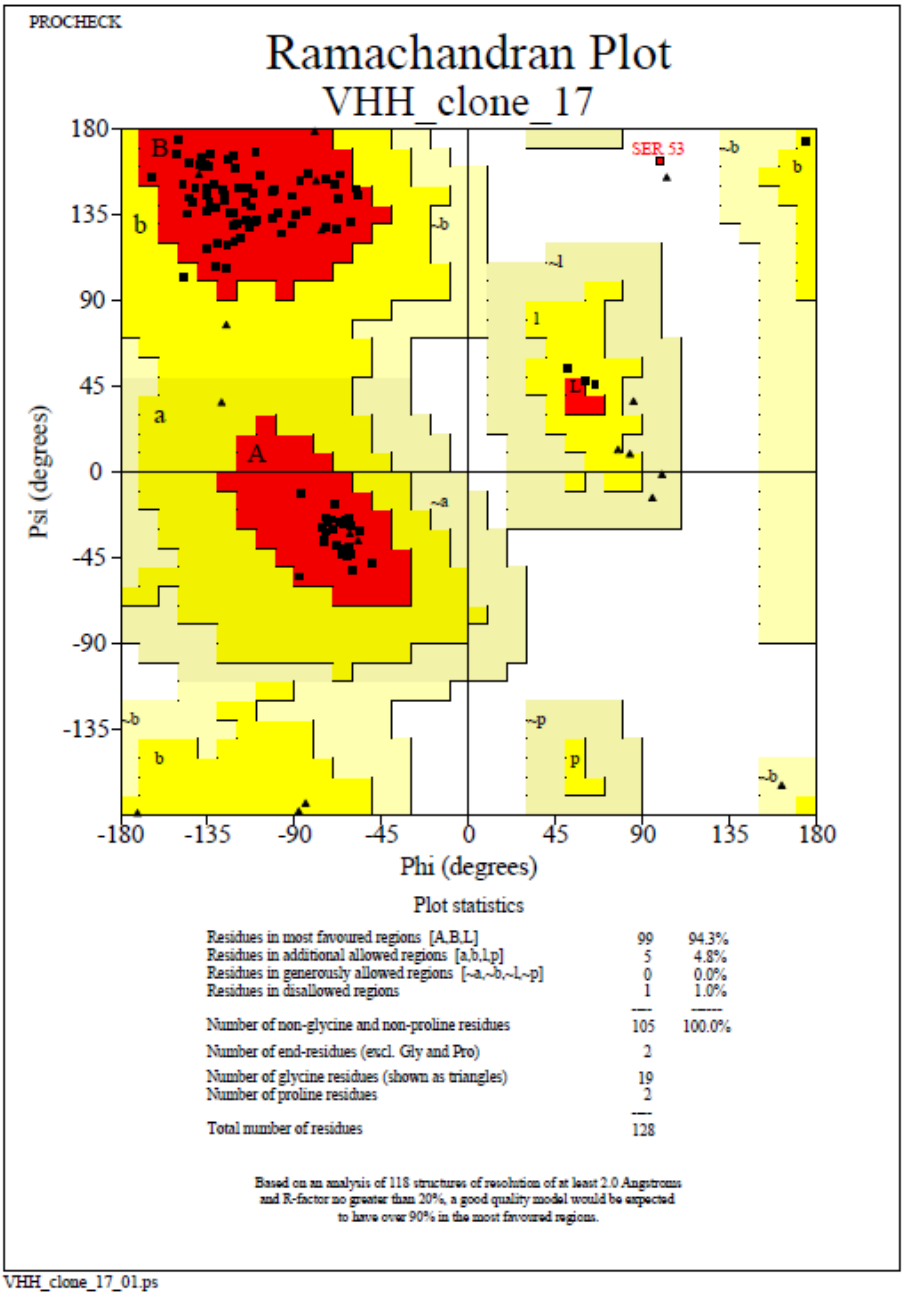
The solution was prepared by dissolving 20 g of NaOH in 500 ml of UDW.

Appendix G

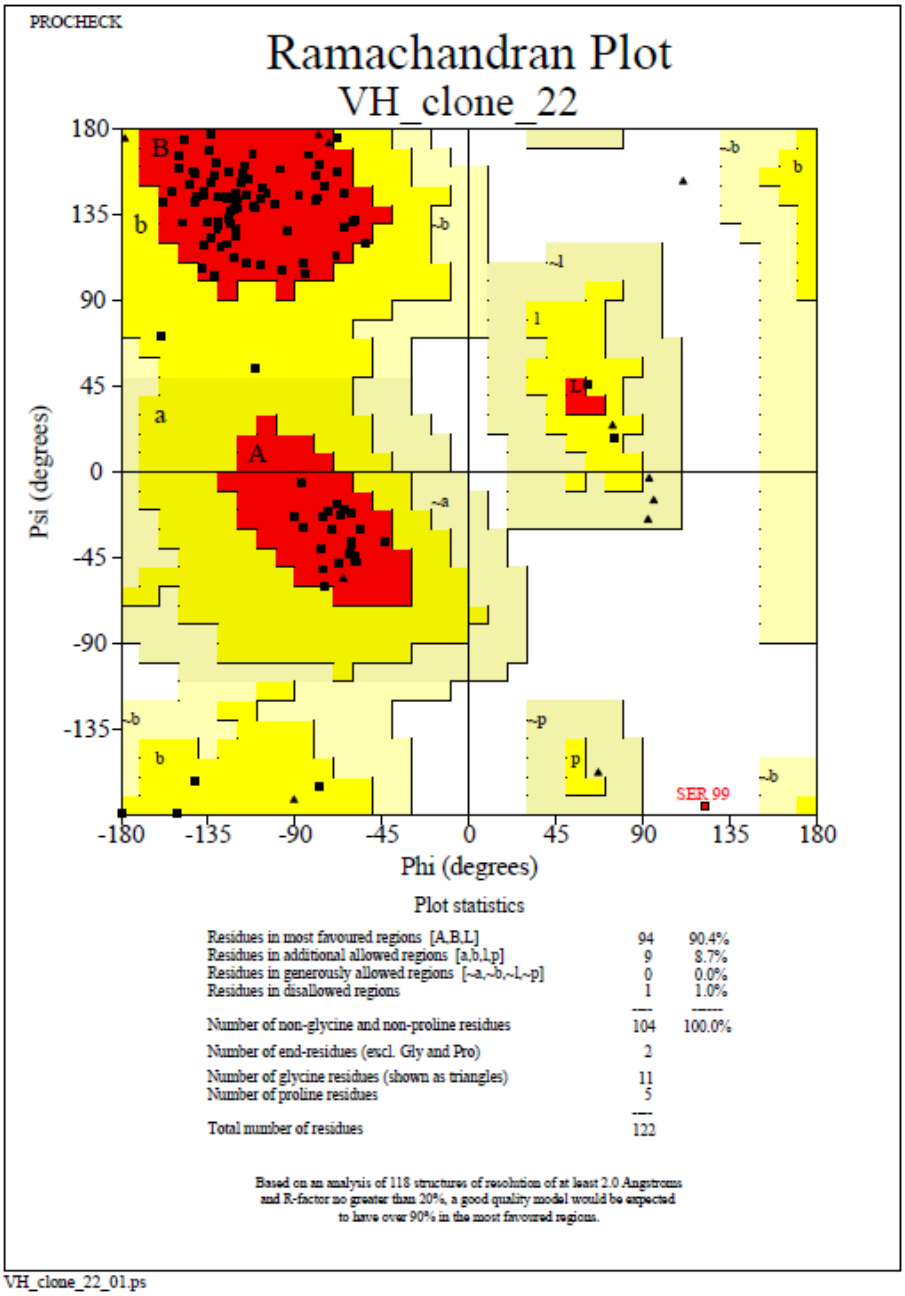
Ramachandran plot of antibodies against ErbB2-TK



VH_clone_4_01.ps



VHH_clone_17_01.ps

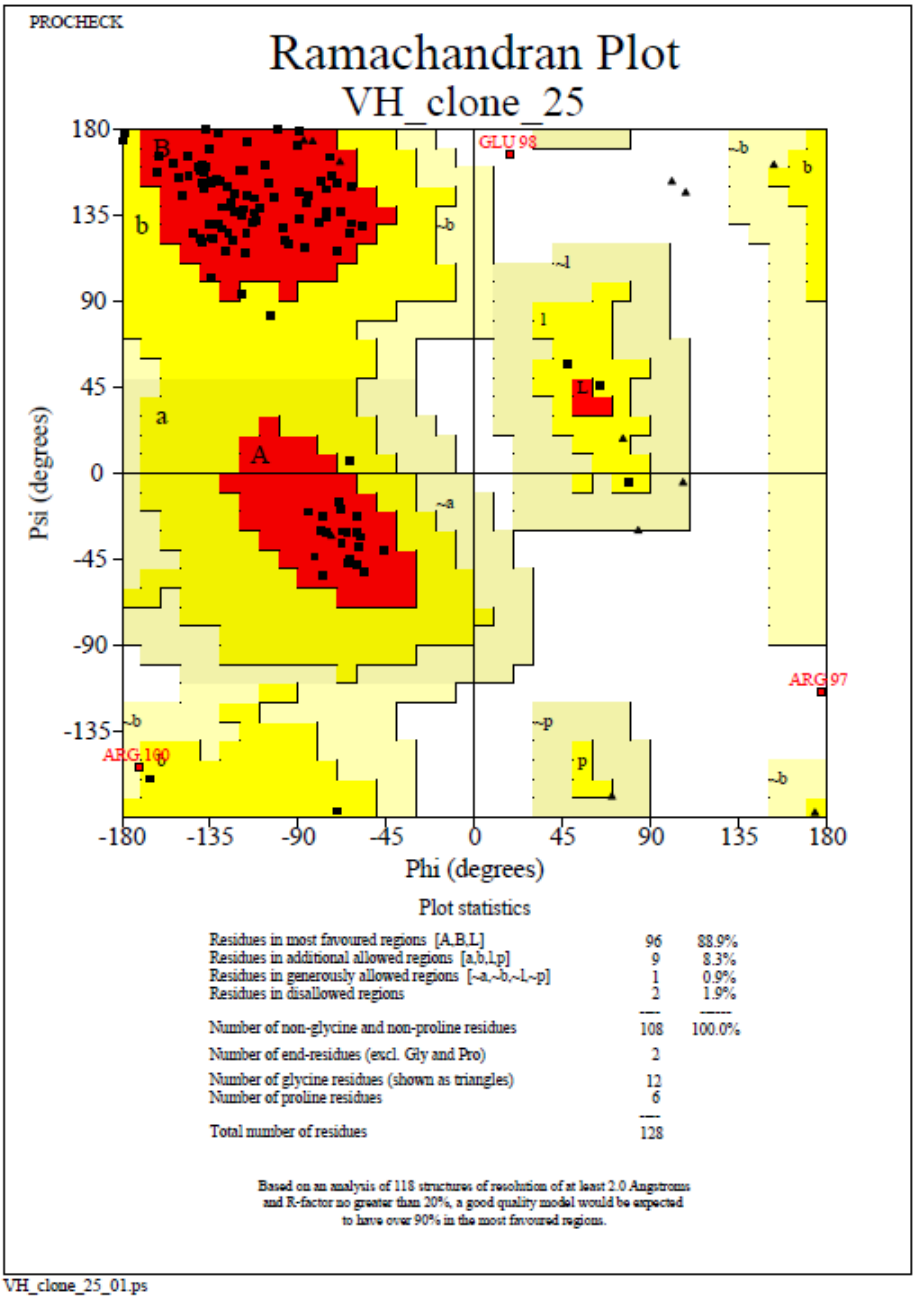


Plot statistics

Residues in most favoured regions [A,B,L]	94	90.4%
Residues in additional allowed regions [a,b,l,p]	9	8.7%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	1	1.0%
Number of non-glycine and non-proline residues	104	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	11	
Number of proline residues	5	
Total number of residues	122	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

VH_clone_22_01.ps



Plot statistics

Residues in most favoured regions [A,B,L]	96	88.9%
Residues in additional allowed regions [a,b,l,p]	9	8.3%
Residues in generously allowed regions [~a,~b,~l,~p]	1	0.9%
Residues in disallowed regions	2	1.9%
Number of non-glycine and non-proline residues	108	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	12	
Number of proline residues	6	
Total number of residues	128	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

VH_clone_25_01.ps

CIRRICULUM VITAE

NAME : Mr. Kunan Bangphoomi

BIRTH DATE : November 2, 1979

BIRTH PLACE : Bangkok, Thailand

EDUCATION :

YEAR	INSTITUTE	DEGREE
2004	Kasetsart University	D.V.M.
2008	Kasetsart University	M.Sc.

WORK PLACE **:** Faculty of Science, Kasetsart University