ZONGULDAK BÜLENT ECEVİT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

INVESTIGATION OF *IN SILICO* ANTICANCER POTENTIAL of SELECTED NATURAL PRODUCTS POTENT MODULATORS of EPIDERMAL GROWTH FACTOR RECEPTOR

DEPARTMENT OF BIOMEDICAL ENGINEERING

MASTER OF SCIENCE

ÖZLEM AYDIN

NOVEMBER 2023

ZONGULDAK BÜLENT ECEVİT UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

INVESTIGATION OF *IN SILICO* ANTICANCER POTENTIAL of SELECTED NATURAL PRODUCTS POTENT MODULATORS of EPIDERMAL GROWTH FACTOR RECEPTOR

DEPARTMENT OF BIOMEDICAL ENGINEERING

MASTER OF SCIENCE THESIS

Özlem AYDIN

ADVISOR: Prof. Dr. İdris ARSLAN

ZONGULDAK
November 2023

APPROVAL OF THE THESIS:

The thesis entitled "Investigation of *In Silico* Anticancer Potential of Selected Natural Products Potent Modulators of Epidermal Growth Factor Receptor" and submitted by Özlem AYDIN has been examined and accepted by the jury as a Master of Science thesis in Department of Biomedical Engineering, Graduate School of Natural and Applied Sciences, Zonguldak Bülent Ecevit University.13/11/2023

Advisor: Prof. Dr. Idris ARSLAN .....
Zonguldak Bülent Ecevit University, Faculty of Science, Department of Molecular Biology and Genetics

Member: Prof. Dr. Ergin YILMAZ .....
Zonguldak Bülent Ecevit University, Faculty of Engineering, Department of Biomedical Engineering

Member: Assistant Prof. Dr. Veli BAYSAL .....
Bartın University, Faculty of Engineering, Architecture and Design, Department of Computer Engineering

---

Approved by the Graduate School of Natural and Applied Sciences. ,.../.../2023

Prof. Dr. Fikret GÖLGELEYEN
Director

"With this thesis it is declared that all the information in this thesis is obtained and presented according to academic rules and ethical principles. Also as required by academic rules and ethical principles all works that are not result of this study are cited properly."

Özlem AYDIN

ABSTRACT

Master of Science

INVESTIGATION OF *IN SILICO* ANTICANCER POTENTIAL of SELECTED NATURAL PRODUCTS POTENT MODULATORS of EPIDERMAL GROWTH FACTOR RECEPTOR

Özlem AYDIN

Zonguldak Bülent Ecevit University
Graduate School of Natural and Applied Sciences
Department of Biomedical Engineering

Thesis Advisor: Prof. Dr. Idris Arslan
Nowember 2023, 47 pages

Cancer can be shortly described as a disorder in multicellular organisms in which some of the cells grow abnormally and spread to other organs of the body called metastasis. It can embark on almost any tissue in the human body. Normally, human cells grow and multiply by means of tightly controlled cell division to generate new cells as the body requires them. Actually, metastasis makes up the main cause of death for approximately more than >90% of patients with cancer.

Drug discovery dates back to the ancient times of human civilization. Historically, treatments and/or new medicines were generally discovered by chance or through observation of their therapeutic effect on disease phenotypes, either directly in humans or in models of disease. However, modern drug discovery research depends on pre-clinical (basic research) and clinical trials and includes expensive and complicated processes.

ABSTRACT (continued)

Natural products are a great group of secondary metabolites produced by living organisms to optimize nature. Natural sources such as terrestrial higher plants, fungi, bacteria, and marine organisms synthesize the metabolites to adapt the their microenvironments. To date, many FDA approved drugs have been discovered for therapeutic use from natural sources, including Taxol (paclitaxel), vincristine/vinblastine (anticancer), fingolimod (multiple sclerosis), cyclosporine and doxorubicin (immunosuppressant). Nowadays, naturally occurring compounds are a great reservoir for drug discovery and design.

In this MSc thesis, we have focused on potential modulatory activities of triterpenoid-based natural products and their derivatives by Click Chemistry towards epidermal growth factor receptor (EGFR) protein through computed-aided screening.

Our findings revealed that azole derivatives of betulinic acid and ursolic acid showed significantly binding affinity on EGFR protein by molecular docking.

Keywords: Epidermal growth factor receptor (EGFR), triterpenoid, hybrid derivatives, Click Chemistry, molecular docking, free binding of energy.

ÖZET

Yüksek Lisans Tezi

INVESTIGATION OF *IN SILICO* ANTICANCER POTENTIAL of SELECTED NATURAL PRODUCTS POTENT REGULATORS of EPIDERMAL GROWTH FACTOR RECEPTOR

Özlem AYDIN

Zonguldak Bülent Ecevit Üniversitesi
Fen Bilimleri Enstitüsü
Biyomedikal Mühendisliği Anabilim Dalı

Tez Danışmanı: Prof. Dr. İdris ARSLAN
Kasım 2023, 47 sayfa

Kanser hastalığı günümüz dünyadasındaki en ölümcül hastalıkların başında gelmektedir. Kanser vücudumuzda bulunan hücrelerin kontrolsüz bir şekilde bölünmesini ve hücre bölünmesi dinamiklerinin bu bölünmeyi kontrol edemeyişini ifade etmektedir. Kontrolsüz bölünen hücrelerin vücuda yayılarak başka organlarda bölünmeye devam etmesi metastaz olarak adlandırılmaktadır. Aslında kanser ölümlerinin >90% oranında metastazdan kaynaklanmaktadır.

Hücre bölünmesini control eden çok sayıda inta- ve ekstrasellular ajan vardır. Büyüme faktörleri tıpkı hormonlar gibi hücreler için sinyal taşıyan aracı moleküllerdir. Epidermal Büyüme Faktörü Reseptörü proteini hücre yüzeyinde bulunan ve hücre bölünmesini kontrol eden ağa sinyal gönderen önemli bir proteindir.

ÖZET (devam ediyor)

Bazı kanser türlerinde EGFR proteinin aşırı ifade edildiği ve kanseri tetikleyen mekanizmalardan birisi olduğuna dair güçlü kanıtlar vardır. Kanser tedavisinde kullanılan afitinib ve gefitinib ilaçlar FDA tarafından onaylanmış ve EGFR hedefli ilaçlardır.

Doğal bileşikler canlı organizmalarca üretilen etkin ve biyoaktif ajanladır. Birçok ilaç tasarımında doğal ilaçlardan faydalanılmıştır.

Bu tez çalışmasında; triterpenoid bileşiklerden olan ve üzerinde çok sayıda bilimsel çalışma yapılan betulinik asit ve ursolik asit bileşiklerinin doğal ve Klik Kimyası yöntemiyle sentezlenmiş kimi türevlerinin olası EGFR düzenleyici etkisi araştırılmıştır. Araştırmada; moleküler kenetlenme ve serbest bağlanma enerjisine bağlı affinitenin hesaplanmasında kullanılan bilgidayar algoritmaları kullanılmıştır. Çalışmada elde edilen sonuçlara göre tüm seçilen ligandlar EGFR proteinine farklı oranlardaki güçlükte bağlanabilmiştir.

Anahtar Kelimeler: Epidermal büyüme faktör reseptörü (EGFR), triterpenoid, hibrit türevler, Klik kimya, serbest bağlanma enerjisi, affinite, moleküler kenetlenme

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my MSc supervisor Professor Idris Arslan for his guidance, creativity, support, and scientific integrity. His knowledge of natural products and medicinal chemistry and dedication to the advancement of the research field will always be a role model for me and light my way of life.

In addition to my advisor, I would like to thank my mother, father, and sister, I have been blessed with encouragement and support from them. They deserve so much acknowledgement. Thanks.

I want to also acknowledge my friends, Ezgi DİNÇ, Pelin YILMAZ, Murat KARA and Ebru Büşra YILMAZ for their generous assistance and support.

I want to thank Professor Ergin YILMAZ, Professor Rahime Seda TIĞLI AYDIN, Professor Muhammet Samet KILIÇ, Professor Ceren KAYA, Research Assistant Tuğba PALABAŞ, and Research Assistant Esra TOPRAK for their supports as a lecturer both my BSc and MSc and giving me opportunities for preparation this MSc thesis.

I would like to express my deep gratitude to my managers Mr. Erkan Onur KÖSELI and Mrs. Fatma Gülcan EKEN at the workplace where I work.

TABLE OF CONTENTS

Page

APPROVAL OF THE THESIS: ii

ABSTRACT iii

ÖZET v

ACKNOWLEDGEMENTS vii

TABLE OF CONTENTS ix

LIST OF FIGURES xi

LIST OF TABLES xiii

LIST OF APPENDICES xv

LIST OF SYMBOLS AND ABBREVIATIONS xvii

CHAPTER 1 INTRODUCTION 1

1.1 CANCER 1

1.2 CANCER STATISTICS 2

1.3 GROWTH FACTORS AND EPIDERMAL GROWTH FACTOR RECEPTOR 4

1.4 DRUG DISCOVERY & DESIGN APPROACHES 7

1.5 VIRTUAL SCREENING AND MOLECULAR DOCKING 9

1.6 SELECTED LIGAND MOLECULES 10

CHAPTER 2 MATERIAL AND METHOD 17

2.1 SELECTION OF LIGANDS AND TARGET PROTEIN 17

2.2 RECEPTOR PREPARATION AND MOLECULAR DOCKING 17

2.3 DRUG-LIKENESS AND TOXICITY PREDICTIONS 18

2.4 CALCULATION OF INHIBITION CONSTANT (K_i) 18

TABLE OF CONTENTS (continued)

Page

CHAPTER 3 RESULTS AND CONCLUSION 21

3.1 RESULTS 21
3.2 DRUG LIKENESS PREDICTION STUDIES 31
3.3 CONCLUSION 32

REFERENCES 35
CURRICULUM VITAE 47

LIST OF FIGURES

No Page

Figure 1.1 Metastasis is primary tumour cells' spread to other organs by blood circulation 2
Figure 1.2 A superposition of similar interfaces observed in crystal structures of the ERBB kinases, including EGFR, ERBB2 (HER2) and ERBB4 (HER4). The protein chains are colored from blue to red from N to C terminus. The kinase at the top of each dimer activates the kinase at the bottom of each dimer 5
Figure 1.3 EGFR-mediated signaling pathways 6
Figure 1.4 Chemical structure of FDA-approved gefitinib and afatinib drugs 6
Figure 1.5 A schematic view of Click Chemistry approach 9
Figure 1.6 3D structure of epidermal growth factor receptor (PDB code:1MOX) 10
Figure 1.7 A brief overview of hybrid drug designing approach 12
Figure 1.8 Chemical structure of betulinic acid(3β-hydroxy-lup-20(29)-en-28-oic acid) 13
Figure 1.9 Chemical structure of betulinic acid analogue 1 (ligand 2)(3[1N(2-cyanophenyl)-1H-1,2,3-triazol-4 yl] methyloxy betulinic acid) 13
Figure 1.10 Chemical structure of betulinic acid analogue 2 (ligand 3) (3[1N(5-hydroxy-naphth-1yl)-1H-1,2,3-triazol-4 yl] methyloxy betulinic acid) 13
Figure 1.11 Chemical structure of ursolic acid (ligand 4)(3*β*-hydroxyurs-12-en-28-oic acid) 14
Figure 1.12 Chemical structure of ursolic acid analogue 1 (ligand 5) (ursolic acid-1-phenyl-1H-[1,2,3]triazol-4 yl *o*-bromo methylester) 15
Figure 1.13 Chemical structure of ursolic acid analogue 2 (ligand 6) (ursolic acid-1-phenyl-1H-[1,2,3]triazol-4 yl *o*-methoxy methylester) 15
Figure 1.14 C-3, C-28, and C-29 functionalities on molecular structure enable to chemical transformation of betulinic acid 16
Figure 3.1 Chemical interactions between amino acids of tarheted protein (1MOX) and ligand 1 (betulinic acid). 24
Figure 3.2 Chemical interactions between amino acids of targeted protein and ligand 2 (betulinic acid analogue 1). 25
Figure 3.3 Chemical interactions between amino acids of targeted protein and ligand 3 (betulinic acid analogue 2). 26
Figure 3.4 Chemical interactions between amino acids of targeted protein and ligand 4 (ursolic acid). 27
Figure 3.5 Chemical interactions between amino acids of targeted protein and ligand 5 (ursolic acid analogue 1). 28
Figure 3.6 Chemical interactions between amino acids of targeted protein and ligand 6 (ursolic acid analogue 2). 29
Figure 3.7 Chemical interactions between amino acids of targeted protein and control group (afatinib). 30
Figure 3.8 Chemical interactions between amino acids of targeted protein and control group (gefitinib) 30
Figure A.1 SwissADME results of ligand 1 (betulinic acid) 41
Figure A.2 SwissADME results of ligand 2 (betulinic acid analogue 1). 42

LIST OF FIGURES (continued)

No Page

Figure A.3 SwissADME results of ligand 3 (betulinic acid analogue 2). 43
Figure A.4 SwissADME results of ligand 4 (ursolic acid). 44
Figure A.5 SwissADME results of ligand 5 (ursolic acid analogue 1) 45
Figure A.6 SwissADME results of ligand 6 (ursolic acid analogue 1). 46

LIST OF TABLES

No Page

Table 1.1. Cancer statistics for the most common cancers in the world in 2020 (both sexes) .. 3
Table 1.2. Country-based cancer statistics for the most common cancers in the world in 2020 4
Table 3.1 Free binding energy, inhibition constant, interaction number and interaction with amino acid values of EGFR and ligands 22
Table 3.2 Smiles format of ligands tested by SwissADME algorithm 23
Table 3.3 Lipinski's rule for ligands tested by SwissADME algorithm 23

LIST OF APPENDICES

Page

APPENDICES SWISSADME RESULTS OF LIGANDS 41

LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOLS

μ	: Micro
M	: Molar
n	: Nano
p	: piko

ABBREVIATIONS

CD95	**:** A prototypical death receptor that regulates tissue homeostasis
CuAAC	**:** Cu-catalyzed Huisgen 1,3-dipolar cycloaddition between alkynes and azides
Da	**:** Dalton
EGF	**:** Epidermal Growth Factor
EGFR	**:** Epidermal Growth Factor Receptor
K_i	**:** Inhibition constant
Log*P*	**:** Partition coefficient
NFκB	**:** Nuclear factor kappa B
NSCLC	**:** Non-small cell lung carcinoma
p53	**:** Tumor protein p53
pAKT	**:** Phosphorilated AKT a serine/threonine protein kinase
PDB	**:** Protein Data Bank
ROS	**:** Reactive oxygen species
TGFα	**:** Transforming Growth Factor α
WHO	**:** World Health Organization

CHAPTER 1

INTRODUCTION

1.1 CANCER

The term 'cancer' states the uncontrolled growth of abnormal cells in the human body. A malignant neoplasm is another term for a cancerous tumour. The term "neoplasm" refers to an abnormal growth of tissue. The word "malignant" means the tumour is cancerous and is likely to spread beyond its point of origin. Cancer cells from a primary tumour may spread to other organs and form new tumours which is called metastasis. Metastasis is often more dangerous than the primary tumour that gives rise to them and responsible for 90% of all cancer deaths. The life threat begins when cancerous cells spread to other tissues of the body like lung, liver, bone, and breast. Secondary cancer growths that spread through the body are generally viewed as its "deadly offspring" (Seyfried and Huysentruyt 2013). In spite of the metastasis process being the key cause of failure of cancer therapy and mortality, the pathophysiology of metastasis is highly complex and not entirely understood. Although the most significant diagnostic and treatment advances, cancer disease is one of the leading causes of death among both men and women around the world.

Basically, human cancer is closely related to certain alterations in several oncogenes and tumour suppressor genes which occur at several stages, from carcinogenesis to tumour growth, progression, and metastasis (Lee 2010, Zhu, Liu and Zhou 2010). *Proto*-oncogenes are genes which normally control cell division and growth; molecular alterations, including gain-of-function mutations, amplification, and overexpression, can trigger the activation of oncogenes which cause uncontrolled cell division and consequently cancer (Berger at al. 2016, Sabapathy 2015). On the contrary, tumor suppressor genes are "good" genes which normally decelerate cell division and growth. If there is any mutation in these genes such as loss-of-function mutations and the deletion of tumor suppressor genes occur, and then tumor suppressors become inactive. As a result of this, cell division and growth get out of control, resulting in cancerous cells (Fritsche and Knopf 2017, Velez-Cruz and Johnson 2017). Thus, the

equilibrium between the oncogenes and tumor suppressor genes seriously controls the initiation and progression of cancer. To suppress the oncogenes and induce tumor suppressor genes are indispensable approaches for preventing or treating malignancies.

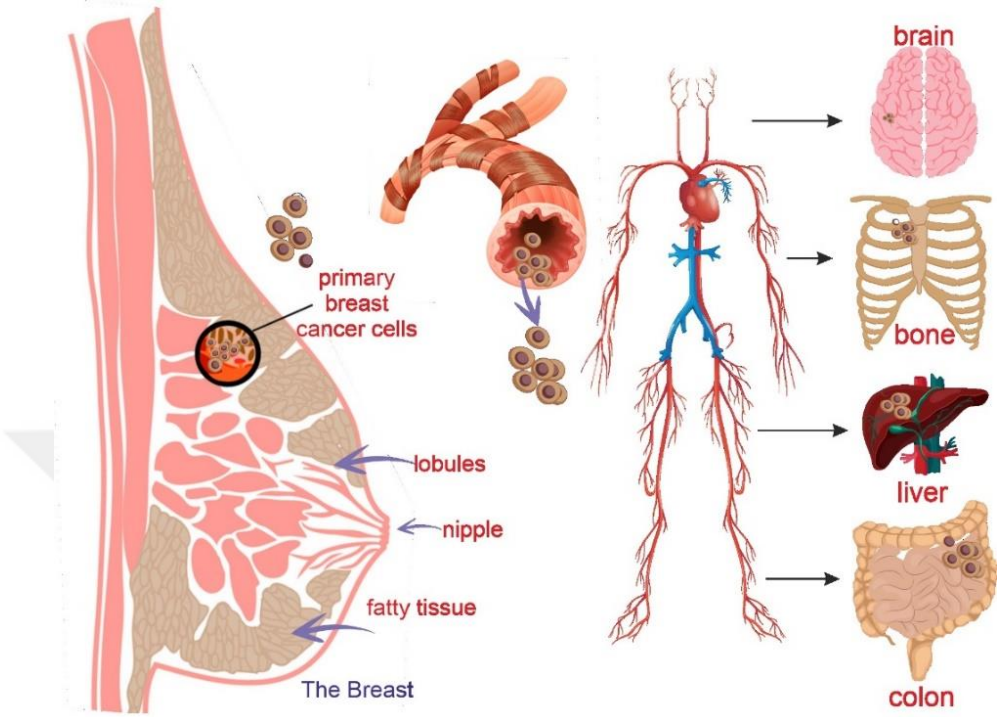

Figure 1.1 Metastasis is primary tumour cells' spread to other organs by blood circulation

1.2 CANCER STATISTICS

Türkiye is suffering from cancer disease, and this adversely affects the people's lives. Across the globe, cancer remains a leading cause of death, with a still rising incidence. Researchers have recently estimated 18.1 million cancer cases all over the world in 2020 year. Of these, approximately 9.3 million cases were reported to be in men and 8.8 million in women (URL-1). Globally, the most common cancers are listed in Table 1.1.

As it can easily be seen in data, the cancer types of breasts and lung were the most common cancers worldwide which constitute 12.5% and 12.2% of the total number of new cases in 2020. Colorectal cancer was reported to be the third most common cancer with 1.9 million new cases in 2020, contributing 10.7% of new cases.

Doubtlessly, prevention of cancer is one of the most critical public health challenges of the 21st century. In addition to action by individuals, achieving healthy patterns of diet and sustained

physical activity over the life course needs concerted and integrated action from all sectors of society, such as civil society, private sector, and health and decision-makers (URL-1).

Table 1.1. Cancer statistics for the most common cancers in the world in 2020 (both sexes)

Rank	Cancer	New cases in 2020	% of all cancer
1	Breast	2,261,419	12.5
2	Lung	2,206,771	12.2
3	Colorectal	1,931,590	10.7
4	Prostate	1,414,259	7.8
5	Stomach	1,089,103	6.0
6	Liver	905,677	5.0
7	Cervix uteri	604,127	3.3
8	Oesophagus	604,100	3.3
9	Thyroid	586,202	3.2
10	Bladder	573,278	3.2
11	Non-Hodgkin lymphoma	544,352	3.0
12	Pancreas	495,773	2.7
13	Leukaemia	474,519	2.6
14	Kidney	431,288	2.4
15	Corpus uteri	417,367	2.3
16	Lip, oral cavity	377,713	2.1
17	Melanoma of skin	324,635	1.8
18	Ovary	313,959	1.7
19	Brain, central nervous system	308,102	1.7
20	Larynx	184,615	1.0
21	Multiple myeloma	176,404	1.0
22	Nasopharynx	133,354	0.7
23	Gallbladder	115,949	0.6
24	Oropharynx	98,412	0.5
25	Hypopharynx	84,254	0.5
26	Hodgkin lymphoma	83,087	0.5
27	Testis	74,458	0.4
28	Salivary glands	53,583	0.3
29	Vulva	45,240	0.3
30	Penis	36,068	0.2
31	Kaposi sarcoma	34,270	0.2
32	Mesothelioma	30,870	0.2
33	Vagina	17,908	0.1

According to World Health Organization (WHO), approximately between 30% and 50% of cancer deaths could be prevented by avoiding key risk factors and implementing present evidence-mediated prevention approaches. The cancer burden could also be decreased by early detection of cancer and management of patients with cancer. Prevention of cancer and decreasing number of patients also support the most cost-effective long-term strategy for cancer therapy.

Growth factors in the human body are hormone-like molecules which can include both cell division and migration of normal cells. Even though numerous growth factors were isolated up till now, their physiological and pathological functions in the body are still debating (Uribe et al 2021).

Table 1.2. Country-based cancer statistics for the most common cancers in the world in 2020

Country	New Case	Average annual case rate per 100.000
All over the world	19.292.789	201
China	4.568.754	204.8
USA	2.281.658	362.2
India	1.324.413	97.1
Japan	1.028.658	285.1
Germany	628.519	313.2
Brasil	592.212	215.4
Russian	591.371	234.3
France	467.965	341.9
United Kingdom	457.960	319.9
Italy	415.269	292.6
Indonesia	396.914	141.1
Spain	282.421	277.2
Canada	274.364	348
TÜRKİYE	233.834	231.5
South Korea	230.317	242.7
Polond	204.575	267.3
Australia	200.021	452.4
Mexico	195.499	140.4
Thailand	190.636	164
Vietname	182.563	159.7
Pakistan	178.388	110.4
Ukraine	162.594	212.8
Bangladesh	156.775	106.2
Philippinnes	153.751	162
Egypt	134.632	159.4

Table 1.2 outlines the number of new cases and case frequency by country in 2020. According to statistics, Türkiye ranks 50th in terms of case cancer frequency, that is, the number of new cancers per 100 thousand people annually (URL-1, URL-2).

1.3 GROWTH FACTORS AND EPIDERMAL GROWTH FACTOR RECEPTOR

The epidermal growth factor receptor (EGFR) belongs to the ErbB family of receptor tyrosine kinases including Erbb1/EGFR/Her1, ErbB2/Her2, ErbB3 and ErbB4 that generate homo- and heterodimeric *trans*-membrane receptors. ErbB family of receptor tyrosine kinases contain the

domains such as an extracellular ligand-binding, a lipid-soluble transmembrane and an intracellular protein tyrosine kinase catalytic (Olayioye et al. 2000).

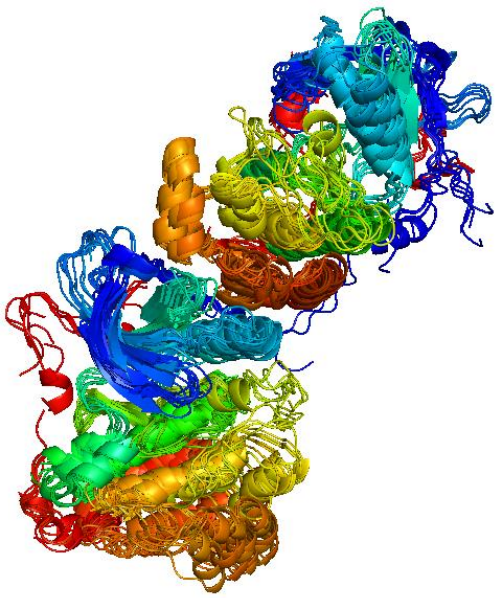

Figure 1.2 A superposition of similar interfaces observed in crystal structures of the ERBB kinases, including EGFR, ERBB2 (HER2) and ERBB4 (HER4). The protein chains are colored from blue to red from N to C terminus. The kinase at the top of each dimer activates the kinase at the bottom of each dimer (Zhang et al. 2008)

EGFR is involved in cell signaling pathways and plays a vital role in the regulation of cell proliferation, survival and differentiation. EGFR mutations usually result in higher expression that induces cancer progression. Evidence suggests that EGFR up-regulates cancer progressions, including in non-small-cell lung cancer, metastatic colorectal cancer, glioblastoma, head and neck cancer, pancreatic cancer, and breast cancer (Ali and Wendt 2017; Wee et al. 2017).

Oncogenic mutant versions of EGFR have been reported to mimic the ligand-activated wild type receptors. Nonetheless, even though the mutated EGFR forms of tumors are catalytically active and transforming, their tyrosine phosphorylation status is remarkably lower when compared to ligand-activated wild type normal form (Shtiegman et al. 2007).

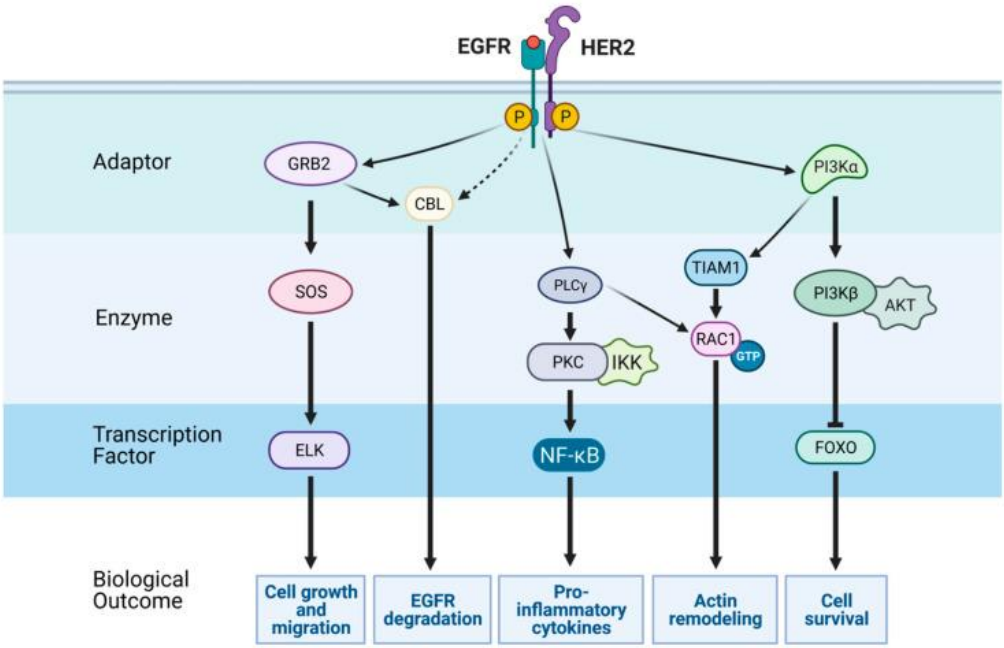

Figure 1.3 EGFR-mediated signaling pathways (Uribe et al. 2021).

Hence, specific inhibition of EGFR is one of the key targets for cancer therapy. So far, many pharmaceutical drugs have been designed to target this strategic oncogenic driver. It is well-documented that downstream signaling from EGFRs progresses through tyrosine phosphorylation action and a positive correlation exists between EGFR and *proto*-oncogenes (Yarden and Sliwkowski 2001). To date, several EGFR inhibitor-based drugs have been approved by FDA, including gefitinib **1**, afatinib **2,** as well as erbitux a chimeric (mouse/human) monoclonal antibody for various cancer chemotherapy (Cohen et al. 2003, Dungo and Keating 2013, Mazzarella et al. 2018).

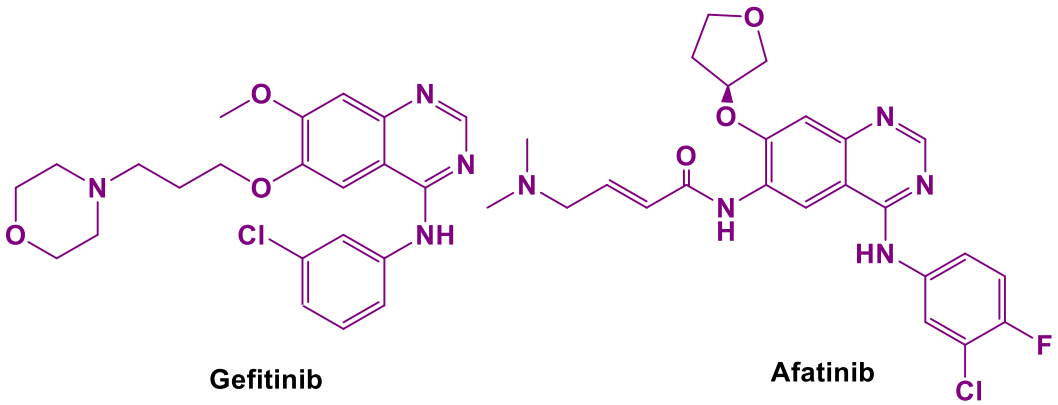

Figure 1.4 Chemical structure of FDA-approved gefitinib and afatinib drugs

EGFR frequently becomes in an auto-inhibited monomeric form under resting status, but ligand binding offers a conformation poised to form either homodimers or heterodimers including HER2 or other binding receptors. However, activation of the receptor often depends on the generation of an asymmetric dimer of kinase domains where one of the domains allosterically activates the other one (Kovacs et al. 2015). Similarly, in the asymmetric dimer, the kinase domain's Clobe of the activator kinase contacts with the *N*-lobe of the receiver kinase, which catalytically stimulates the latter. Afterward, the receiver kinase trans-phosphorylates certain tyrosine residues of the activator kinase. The newly phosphorylated tyrosines act a role as attachment sites for various adaptors like GRB2, protoplasmic enzymes such as phospholipase C-ɣ or specific factors attending transcription regulation like STAT3.

Additionally, aforementioned signaling effectors and adaptor proteins attach activated receptors directly or indirectly to canonical intracellular pathways, as well as to the endocytic machinery, which desensitizes active receptors. Interestingly, various tyrosine docking sites of EGFR can bind over one adaptor or effector. (Watermann et al. 2002).

Afatinib breaks up signaling by means of the EGFR in target cells and is often administered to treat non-small cell lung carcinoma (NSCLC) (Minkovsky and Berezov 2008, Spreitzer 2008). It is fundamentally used to treat cases of NSCLC that harbour mutations in the EGFR gene (Vavalà 2017) and one of the World Health Organization's List of Essential Medicines (WHO 2021).

Gefitinib is a medicine sold under the brand name Iressa and used to treat breast, lung and other soft tissue malignancies. Since gefitinib is the first selective EGFR blocker, like afatinib, it is only effective for mutated and overexpressed EGFR cancers (Pao et al. 2004).

1.4 DRUG DISCOVERY & DESIGN APPROACHES

Drug discovery dates back to the early years of human civilization. Historically, treatments and/or new medicines were generally discovered by chance or through observation of their therapeutic effect on disease phenotypes, either directly in humans or in models of disease. However, modern drug discovery research depends on *pre*-clinical (basic research) and clinical trials and includes expensive and complicated processes.

The starting research, often occurring in academic institutes, produces data to make a hypothesis that the up-regulation or down-regulation of a targeted protein or biochemical pathway will cause a therapeutic effect in a disease. The result of this trial is the selection of a target which might need further validation before progression into the lead discovery phase to advocate a drug discovery effort. During lead discovery, an intensive search uncovers to find a drug-like small molecule or biologics, often termed a drug candidate, that will progress into preclinical, and if successful, into clinical development and finally be a marketed medicine (Hughes et al. 2011).

Natural products are secondary metabolites synthesized to optimize nature by different living organisms such as terrestrial higher plants, fungi, bacteria, and marine organisms. So far, many FDA approved drugs have been discovered for therapeutic use from natural sources, including Taxol (paclitaxel), vincristine/vinblastine (anticancer), fingolimod (multiple sclerosis), cyclosporine and doxorubicin (immunosuppressant) (Newman and Cragg 2020).

Since the most of natural product are complex molecules with various stereo centers, sp^3 carbon, and labile functionalities, the *de novo* synthesis of naturally occurring compounds or their analogues generally require complicated synthetic approaches.

'Click Chemistry' is an immensely powerful technique for the fast, cost-effective, and influential covalent integration of molecular functionalities which has great potential for use in binding between organic molecules such as nucleic acids, lipids, proteins, phytochemicals.

Click Chemistry, particularly the Cu(I)-catalyzed Huisgen 1,3-dipolar cycloaddition between alkynes and azides (CuAAC) has been gained a popularity by researchers as versatile players in the modification of molecules particularly complex natural in order to enhance the pharmacological or biological properties of them (Zhang at al. 2021). Click Chemistry strategies are summarized in Figure 1.4.

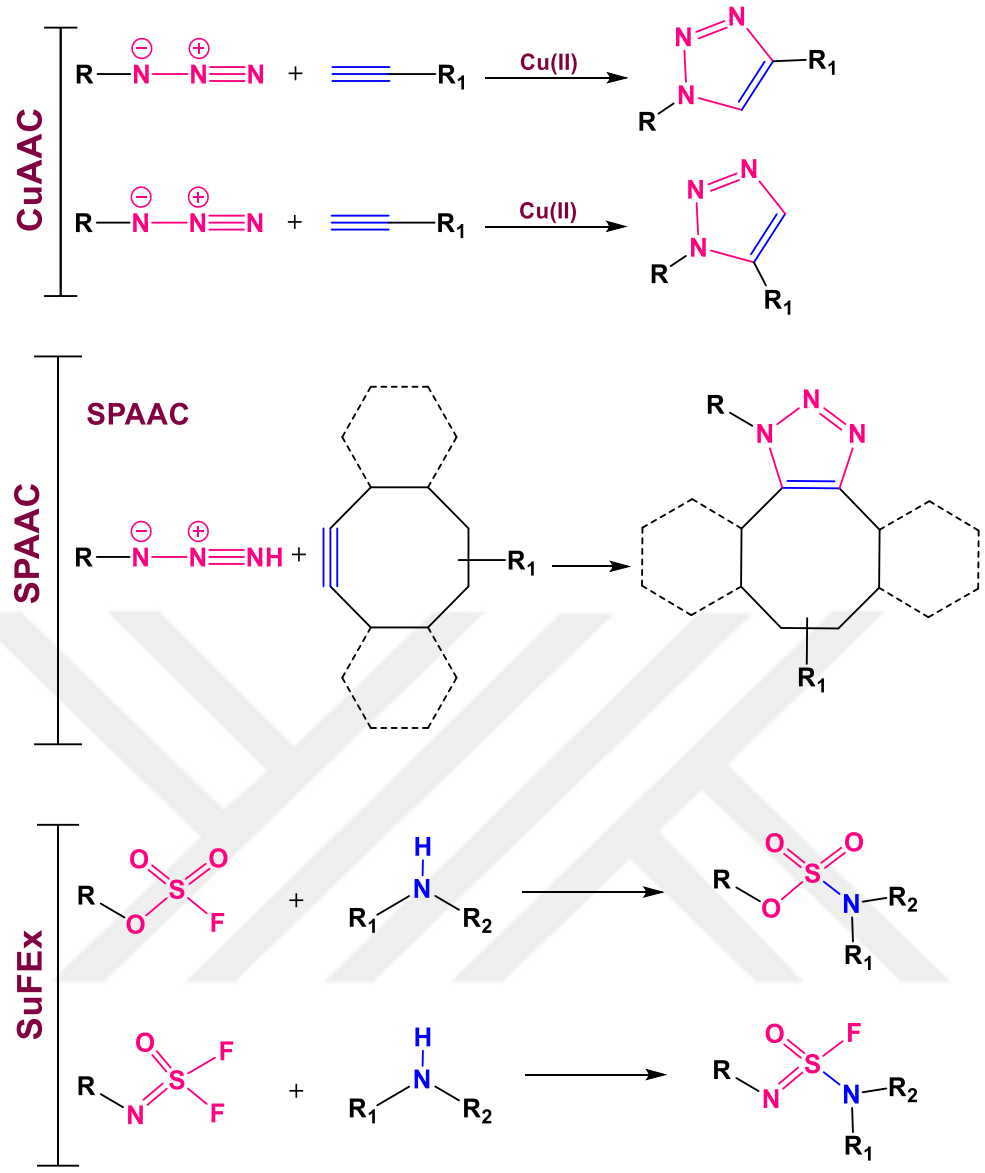

Figure 1.5 A schematic view of Click Chemistry approach

Pharmacological properties of natural products and their triazole analogues obtained by the Click Chemistry approach has been gaining popularity.

1.5 VIRTUAL SCREENING AND MOLECULAR DOCKING

Computer-aided screening approach is able to use to model the interaction between a ligand and a target protein at the atomic level to characterize the behaviour of ligands in the active site of target proteins and to elucidate basic biochemical processes as well (McConkey at al. 2002).

'Computer-aided virtual screening' is one of the most widely used research methods *in silico* drug discovery, where scientists screen a large library of ligands for interesting hits, in a simulation known as 'molecular docking'. Even though molecular docking is a time-consuming and computationally intensive simulation, it screens candidate compounds more efficiently than that of physical evaluation.

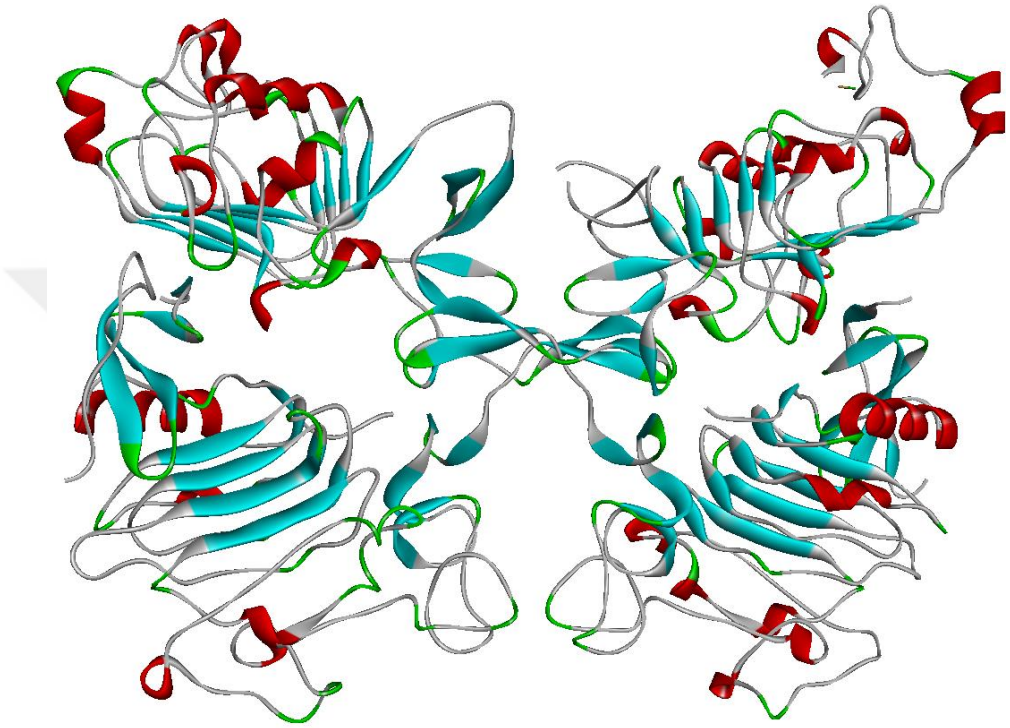

Figure 1.6 3D structure of epidermal growth factor receptor (PDB code:1MOX)

The molecular structure and functional confirmation of the EGFR system has intrigued clinicians and researchers for more than four decades. Biochemically, crystal structure of a truncated EGFR extracellular domain bound to Transforming Growth Factor α (TGFα) (Figure 1.5) was reported to easily bind a family of highly variable ligands (Garrett et al. 2002).

1.6 SELECTED LIGAND MOLECULES

Betulinic acid (**1**) (3*β*-hydroxy-lup-20(29)-en-28-oic acid) is a naturally occurring pentacyclic lupane-type triterpenoid compound commonly found in higher plants such as the birch tree (*Betula* sp., Betulaceae), eucalyptus and plane trees. It was first isolated from Gratiola officinalis at the beginning of the 20th century by Dr. Retzlaff (Retzlaff 1902). Additionally,

betulinic acid was reported to be present in several plant species of the genera Ziziphus (Rhamnaceae), Syzygium (Myrtaceae), Diospyros (Ebenaceae), and Paeonia (Paeoniaceae). It is often found in the plant kingdom as a free sapogenin or as glycosylated analogues (Moghaddam et al. 2012). Betulinic acid is a secondary metabolite with a potential of inducing apoptosis in a variety of malignancies and displays significant selectivity for tumor cells over non-transformed cells (Gao vd. 2011). Shen and colleagues (2019) showed that betulinic acid mediated cytotoxicity in multiple myeloma cells by apoptosis, S-phase arrest, mitochondrial membrane collapse, and overwhelming reactive oxygen species (ROS) accumulation. Furthermore, betulinic acid-mediated apoptosis was partially reversed when the reactive oxygen species scavenger *N*-acetyl cysteine remarkably abated elevated reactive oxygen species (Shen et al. 2019). Additionally, it also inhibits the protein specificity transcription factors which induces the stress kinases p38, functions as potent blocker of mammalian type I DNA topoisomerase, nuclear factor kappa B (NFκB) and exhibits apoptosis in a p53- and CD95- independent manner (Tan vd. 2003; Kasperczyk vd. 2005).

Despite outstanding antitumor activities, further clinical progression of betulinic acid is highly blocked due to its poor aqueous solubility and pharmacokinetic features. Moreover, it possesses weak metabolic stability of the compound getting metabolized causing low plasma concentration which restricts the druggability of betulinic acid. Both poor permeability and aqueous solubility suggests that betulinic acid might be incompatible with oral administration. Hence, researchers have focused on modification of betulinic acid on the C-3 and/or C-28 positions to circumvent the limitations and enhance the water solubility and possibly permeability.

The synthetical hybrid analogue of betulinic acid 3[1*N*(2-cyanophenyl)-1*H*-1,2,3-triazol-4 yl] methyloxy betulinic acid (**2**) (CBA) has been reported to induce G0/G1 cell cycle arrest and increase expression of sub-G0 DNA fraction and annexin V binding of the cells besides imparting the typical surface features of cell death (Majeed et al. 2016).

The hybrid strategy was reported to show distinct adventage over the conventional approach of co-administration of single hit molecules owing to better long-term prognosis and reduced toxicity (Choudhary et al. 2018).

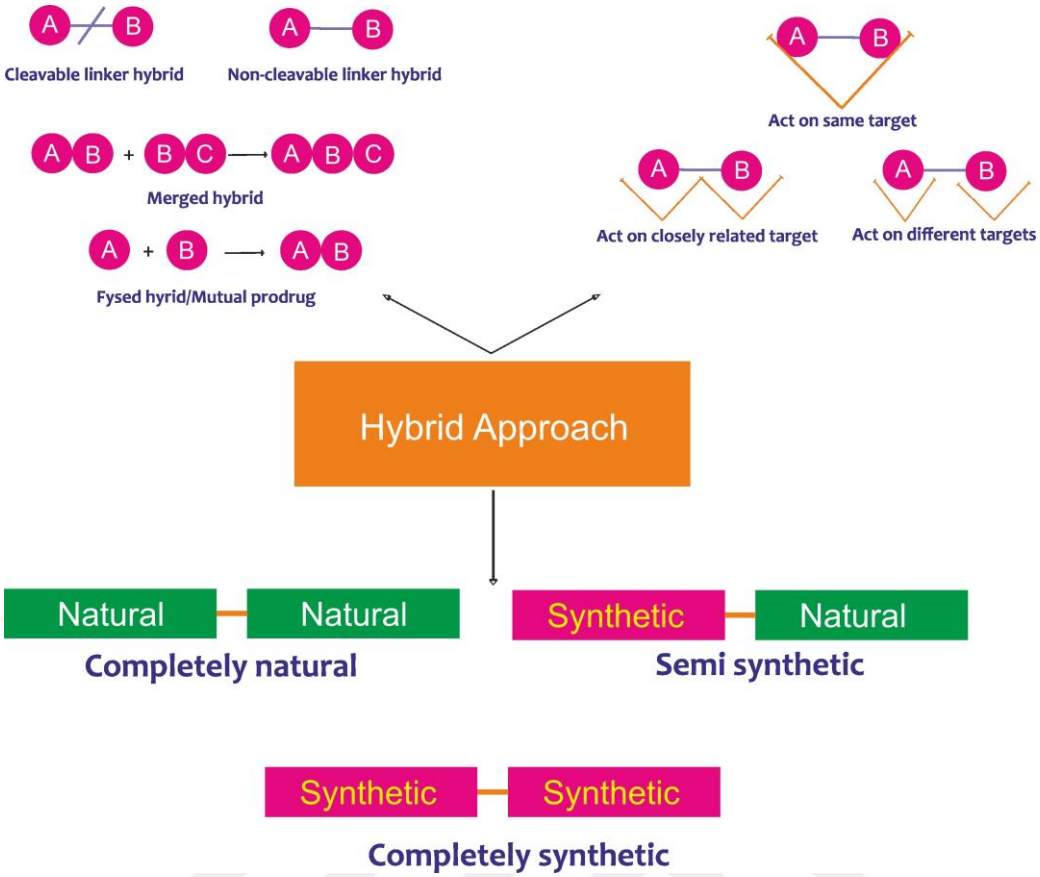

Figure 1.7 A brief overview of hybrid drug designing approach (Choudhary et al. 2018)

Majeed and colleagues (2014) has successfully synthesized a potential hybrid molecule 3[1*N*(5-hydroxy-naphth-1yl)-1*H*-1,2,3-triazol-4 yl] methyloxy betulinic acid (**3**) with significant inhibition of cancer cell growth. This hybrid molecule has been also reported to decrease the expression of phosphotidylinositol-3 kinase (PI3K) p110α and p85α and leading to remarkably downregulation of pAKT and of NFκB using human leukemia and breast cancer cells as *in vitro*.

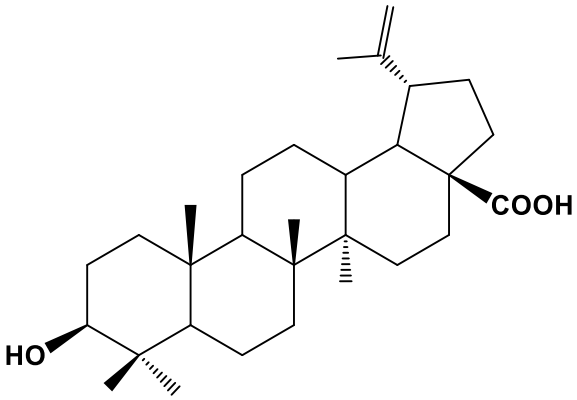

Figure 1.8 Chemical structure of betulinic acid(3β-hydroxy-lup-20(29)-en-28-oic acid)

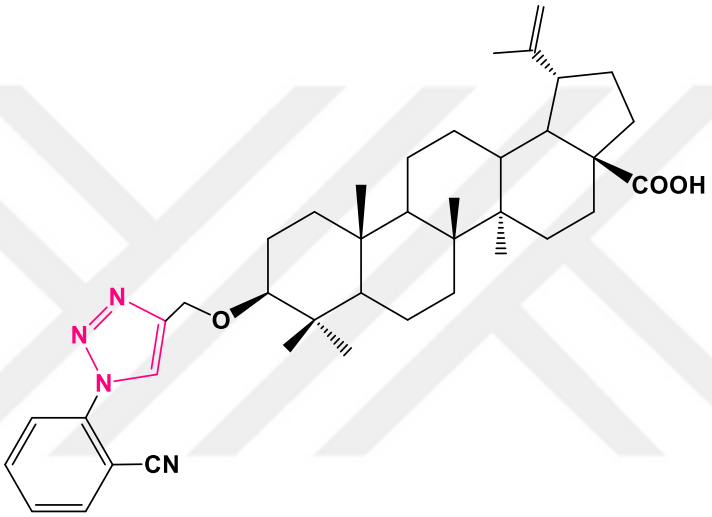

Figure 1.9 Chemical structure of betulinic acid analogue 1 (ligand 2)(3[1N(2-cyanophenyl)-1H-1,2,3-triazol-4 yl] methyloxy betulinic acid)

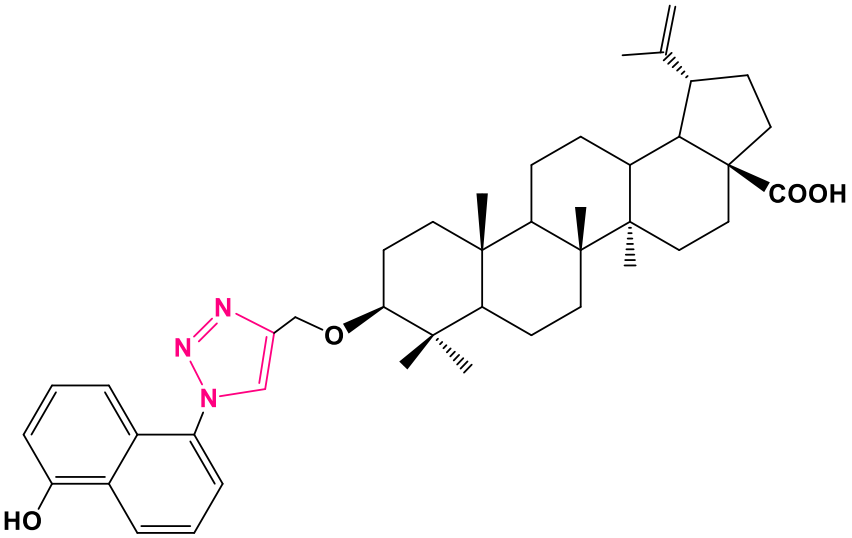

Figure 1.10 Chemical structure of betulinic acid analogue 2 (ligand 3) (3[1N(5-hydroxy-naphth-1yl)-1H-1,2,3-triazol-4 yl] methyloxy betulinic acid)

Ursolic acid (**4**) (3β-hydroxyurs-12-en-28-oic acid) is a pentacyclic triterpene acid characterized in the epicuticular waxes of apples in 1920s and commonly distributed in medical herbs and edible plants, like wax like coating of apples and other fruits and in herbs and spices like rosemary and thyme as well (Cha et al., 1998, Sohn et al., 1995). Ursolic acid was reported to reduce further damage to neurons and helped rebuild the protective sheaths covering neurons, apparently by inhibiting Th17 immune cells and triggering precursor cells that mature into myelin-sheath-making cells, termed oligodendrocytes in mice with chronic multiple sclerosis (MS) (Zhang et al. 2020). Similarly, ursolic acid was reported to enhance domoic acid-induced cognitive deficits in mice (Wu et al. 2013). Furthermore, it attenuates lipopolysaccharide-induced cognitive deficits in mouse brain through suppressing p38/NF-κB mediated inflammatory pathways (Wang et al. 2011). Mechanistic studies displayed that ursolic acid inhibits the cell cycle progression in the G1 phase and its treatment results in the triggering of apoptosis as determined by DNA fragmentation assay (Hsu et al. 2004).

Rashid and colleagues (2013) have designed a series of ursolic acid-1-phenyl-1*H*-[1,2,3]triazol-4 yl methylester congeners in an attempt to develop potent antitumor agents by click chemistry approaches. Analogs with *o*-bromo (**5**), and *o*-methoxy (**6**) substituents on aromatic rings were reported to be the most promising anticancer compounds (Rashid et al. 2013). Ursolic acid analogues were detected to be the most potent among the whole series with IC_{50} value less than 0.1 mM towards two of the malignancies tested, MCF-7 and THP-1.

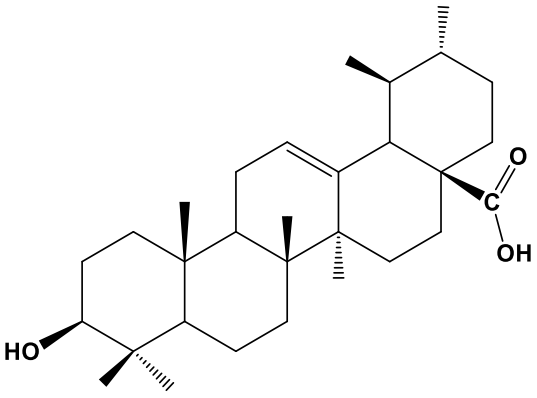

Figure 1.11 Chemical structure of ursolic acid (ligand **4**)(3*β*-hydroxyurs-12-en-28-oic acid)

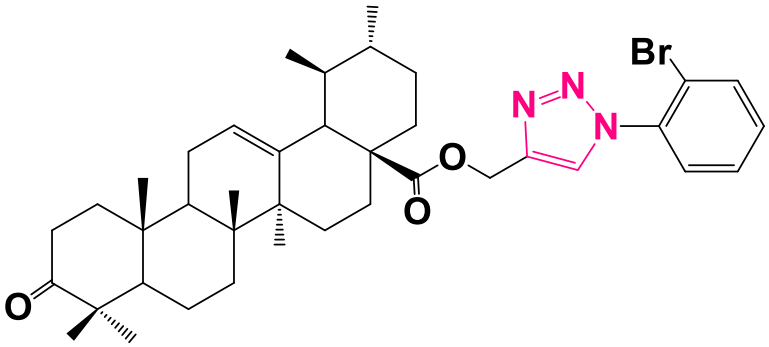

Figure 1.12 Chemical structure of ursolic acid analogue 1 (ligand **5**) (ursolic acid-1-phenyl-1H-[1,2,3]triazol-4 yl *o*-bromo methylester)

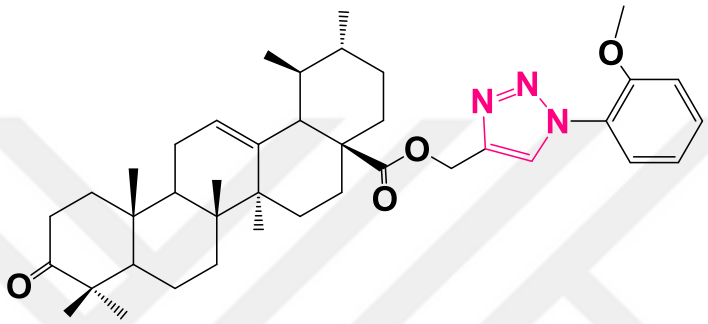

Figure 1.13 Chemical structure of ursolic acid analogue 2 (ligand 6) (ursolic acid-1-phenyl-1H-[1,2,3]triazol-4 yl *o*-methoxy methylester)

Triterpenes allow easy chemical modification since they contain different functional groups attached to their carbon skeleton structure. These modifications can offer different options to find out the ideal drug candidate for better efficacy, lesser toxicity. For example, the betulinic acid molecule can be modified from different groups in the C-3 hydroxyl group, C-28 carboxyl group and C-29 alkene structure (Figure 1.13).

Majeed and colleagues (2013) synthesized betulinic acid triazole derivatives at C-3 position of ring A through the Click Chemistry approach.

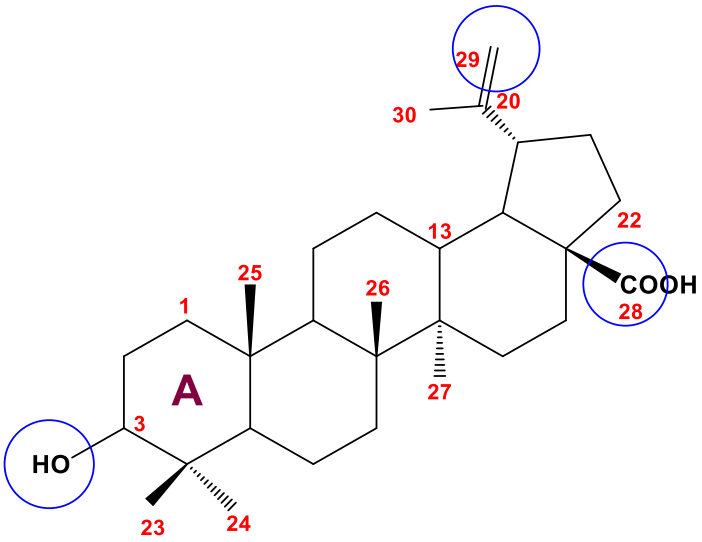

Figure 1.14 C-3, C-28, and C-29 functionalities on molecular structure enable to chemical transformation of betulinic acid

The overall aim of the MSc thesis was to investigate *in silico* the potent modulatory effects on epidermal growth factor receptor protein (EGFR) of both natural products (betulinic and ursolic acid) and semi-synthetic hybrids (hybrid derivatives) by computer-aided drug development methods have proved helpful in the early stages of research into new pharmacological therapies. To the best of our knowledge, there are no reports revealing the molecular interaction among betulinic and ursonic acid and/or their derivatives by Click Chemistry technique and EGFR protein.

CHAPTER 2

MATERIAL AND METHOD

2.1 SELECTION OF LIGANDS AND TARGET PROTEIN

Three-dimensional crystal structure of human epidermal growth factor receptor (residues 1-501) in complex with TGF-α (EGFR) (PDB ID:1MOX) (Garrett et al. 2002) was downloaded from the RCSB PDB library. Ligands were selected from the various databases as a result of literature survey. Chemical structures of natural products were drawn by ChemDraw toolbar and then converted into MOL file format. A powerful open-source chemical toolbox Open Babel was used to convert the '.mol' files to '.pdb' files. All ligands and water molecules were removed from the 3D protein structures via the Discovery Studio Visualizer 4.1 Client program, and then loop segments were completed.

Water molecules were carefully excluded as criteria for preparing the protein structure, while Gasteiger charges and essential hydrogen atoms were also included (Yan et al. 2014). Figure 1.5 shows a representative scheme of hit protein structure.

2.2 RECEPTOR PREPARATION AND MOLECULAR DOCKING

The 2D chemical structure of each ligand was drawn through the ChemDraw Professional 16.0.1 (Perkin Elmer). Initially, the 2D structures were saved in the file format of '.mol2', and then converted to the .pdb file format (3D) using the The Open Babel Graphical User Interface: The Open Source Chemistry Toolbox.

Molecular docking simulations to determine the molecular interaction of represented ligands with the EGFR protein at the active site were performed using AutoDock 4.2.6, an automated docking tool (Morris et al. 2009). During the molecular docking, both Gasteiger partial charges and polar hydrogen atoms were incorporated into the three-dimensional structure of the EGFR protein. All protein structures and selected ligands were converted into the file format of '.pdbqt' for further analysis on tool interface. Afterward, the grid size for simulation of EGFR and

ligands (**1**-**6**) were set at 60×60×60 points, followed by 0.375 Å spacing centred. Grid centres x (-5.323), y (66.115), z (30.596). However, minimum coordinates in grid = (-16.573, 54.865, 19.346) and maximum coordinates in grid = (5.927, 77.365, 41.846) were recorded.

The grid boxes were determined according to the amino acid active sites on EGFR protein. Additionally, the Lamarckian Genetic Algorithm 4.2 was applied in the docking analysis, while the protein macromolecules were kept rigid throughout the docking simulation. The genetic algorithm runs were set at 400, while default settings were maintained for the other parameters for docking analyses. The best protein-ligand conformations were selected from the AutoDock 4.2 scoring function which ranked the results according to their estimated free binding of energy (kcal/mol). However, the inhibition constant (K_i value) (nM, μM or pM) which is an indication of how potent an inhibitor for each ligand was also recorded. All docking results were analyzed using the Discovery Studio Visualizer 4.1 client. Getitinib and afatinib have been set as a control group.

Free Energy of Binding (kcal/mol) was calculated the following equation,

Estimated Free Energy of Binding = [(1)+(2)+(3)-(4)]

(1) Final Intermolecular Energy
(2) Final Total Internal Energy
(3) Torsional Free Energy
(4) Unbound System's Energy

2.3 DRUG-LIKENESS AND TOXICITY PREDICTIONS

Lipinski's rule of five was applied to predict the drug-likeness property of ligands which is widely used to determine the consistency of orally active drugs (Lipinski 2004). In the present study, the SwissADME predictor (Daina et al. 2017) was also used for drug-likeness and toxicity predictions.

2.4 CALCULATION OF INHIBITION CONSTANT (K_i)

The inhibitor constant, K_i, states an indication of how potent an inhibitor is and the concentration required to produce half maximum inhibition.

The inhibitory constant is a type of equilibrium dissociation constant (K_d) of a chemical reaction which constitutes the equilibrium binding affinity for a ligand which decreases the activity of

its binding partner. K_i constitutes the concentration at which the blocker ligand occupies at least 50% of the receptor sites when no competing ligand is present.

The inhibition constant values of all ligands tested were calculated from the binding energy (ΔG) using the formula:

$K_i = \exp(\Delta G/RT)$,

where R is the universal gas constant (1.985×10^{-3} kcal mol^{-1} K^{-1}) and

T is the temperature (298.15 K).

CHAPTER 3

RESULTS AND CONCLUSION

3.1 RESULTS

The aim of the MSc thesis was to investigate the potent modulatory activities of betulinic acid and ursolic acid and their hybrid derivatives on epidermal growth factor receptor protein (EGFR) by computer-aided drug development methods have proved helpful in the early stages of research into new pharmacological therapies. To the best of our knowledge, there are no reports revealing the molecular interaction among betulinic and ursonic acid and/or their derivatives by Click Chemistry technique and EGFR protein.

Free binding energy values (kcal/mol), inhibition constant (K_i values nM, pM or pM), and interaction with amino acids pertaining to target protein of EGFR and ligand compounds (**1**-**6**) is shown in Table 3.1. Natural and their hybrid compounds displayed different levels of convenient interactions with target protein. Herein, the low free binding energy between ligand and hit molecule states the degree of binding affinity and strong interaction.

As it can be easily seen in Table 3.1, our findings revealed that 5 and 6 had the highest binding affinity towards EGFR with a reported value of -16.07and -15.73kcal/mol, respectively. The free binding energies between afatinib – gefitinib and EGFR were found to be -9.95 and -9.63 kcal/mol, respectively. The lowest binding affinity was detected to be -7.64 kcal/mol between ligand 4 and EGFR. Interestingly, the binding affinity of 2, 3, 5, and 6 towards EGFR were found higher than both positive control groups of afatinib and gefitinib. Both betulinic acid and ursolic acid derivatives showed higher binding affinity towards EGFR than that of the original compound, suggesting that Click Chemistry analogues are more efficient to interact with amino acid residue than the original molecule.

When compared to naturally occurring products, semi-synthetic hybrid analogues are higher binding affinities to hit molecules.

Table 3.1 Free binding energy, inhibition constant, interaction number and interaction with amino acid values of EGFR and ligands

Target protein	Ligands		Free binding energy (kcal/mol)	Inhibition constant (K_i values)	Interaction number	Interaction with amino acid
1MOX	Control	Afatinib	-9.95	50.76 nM	20	His A:280, Pro A:242, Gly A:281, Pro A:241, Thr A:239, Cys A:240, Ser B:262, Tyr A:246, Leu A:245, Cys B:204, Phe B:230, Met A:253, Pro: 242, Gly B:281, His 280, Leu B:243, Met A:244, Met B:244, Leu A:243
		Gefitinib	-9.63	87.62 nM	10	Tyr B:246, Leu B:245, Thr A:239, Pro B:242, Gly B:281, Ser B:262, Cys B:240, His B:280, His A:280
	1		-8.51	573.53 nM	5	Glu221, Pro242, Lys237, Pro241, Thr239
	2		-10.12	37.97 nM	5	Lys A:229, Pro A:242, Pro B:241, Pro B:242, His B:280
	3		-11.94	1.31 nM	7	Pro B:242, Lys A:229, His B:280, Leu B:243, Cys B:240, Pro A:242, Pro B:241
	4		-7.64	2.51 μM	6	Pro A:242, Thr A:239, Glu B:221, Lys B:237, Pro B:242, Pro B:241
	5		-16.07	1.66 pM	9	Cys A:240, Ser A:262, Leu A:245, Tyr A:246, His A:280, Pro A:242, His B:280, Pro B:242, Cys B:240
	6		-15.73	2.96 pM	10	Cys B:240, Thr A:239, Pro B:242, His A:280, Pro A:241, Pro A:242, Tyr A:246, Leu A:245, Ser A:262, His B:280

Table 3.2 Smiles format of ligands tested by SwissADME algorithm

Ligands	Smiles Format
1	C1[C@@H](C(C2[C@](C1)(C1[C@@](CC2)([C@]2(C(CC1)C1[C@@](CC2)(CCC1C(=C)C)C(=O)O)C)C)C)(
2	C1[C@@H](C(C2[C@](C1)(C1[C@@](CC2)([C@]2(C(CC1)C1[C@@](CC2)(CC[C@H]1C(=C)C)C(=O)O)C)C)C)(C)C)OCc1nnn(c1)c1ccccc1C#N
3	C1[C@@H](C(C2[C@](C1)(C1[C@@](CC2)([C@]2(C(CC1)C1[C@@](CC2)(CC[C@H]1C(=C)C)C(=O)O)C)C)C)(C)C)OCc1nnn(c1)c1c2c(ccc1)c(ccc2)O
4	C1[C@@H](C(C2[C@](C1)(C1[C@@](CC2)([C@]2(C(=CC1)C1[C@](CC2)(C(=O)O)CC[C@H]([C@@H]1C)C)C)C)C)(C)C)O
5	C1C(=O)C(C2[C@](C1)(C1[C@@](CC2)([C@]2(C(=CC1)C1[C@](CC2)(C(=O)OCc2cn(nn2)c2ccccc2Br)CC[C@H]([C@@H]1C)C)C)C)C)(C)C
6	C1C(=O)C(C2[C@](C1)(C1[C@@](CC2)([C@]2(C(=CC1)C1[C@](CC2)(C(=O)OCc2cn(nn2)c2ccccc2OC)CC[C@H]([C@@H]1C)C)C)C)C)(C)C

Table 3.3 Lipinski's rule for ligands tested by SwissADME algorithm

Ligands	Molecular formula	Molecular weight (g/mol)	Hydrogen bond donor	Hydrogen bond acceptor	Log *P*	Molar refractivity	Rules satisfied
1	$C_{30}H_{48}O_3$	456.70	2	3	6.14	136.91	4/5
2	$C_{40}H_{54}O_3$	638.88	1	6	7.05	185.76	3/5
3	$C_{43}H_{57}O_4$	679.93	2	6	7.82	200.58	3/5
4	$C_{30}H_{48}O_3$	456.70	2	3	5.87	136.91	4/5
5	$C_{39}H_{52}BrN_3O_3$	690.75	0	5	7.59	187.38	3/5
6	$C_{40}H_{55}N_3O_4$	641.88	0	6	6.97	186.17	3/5

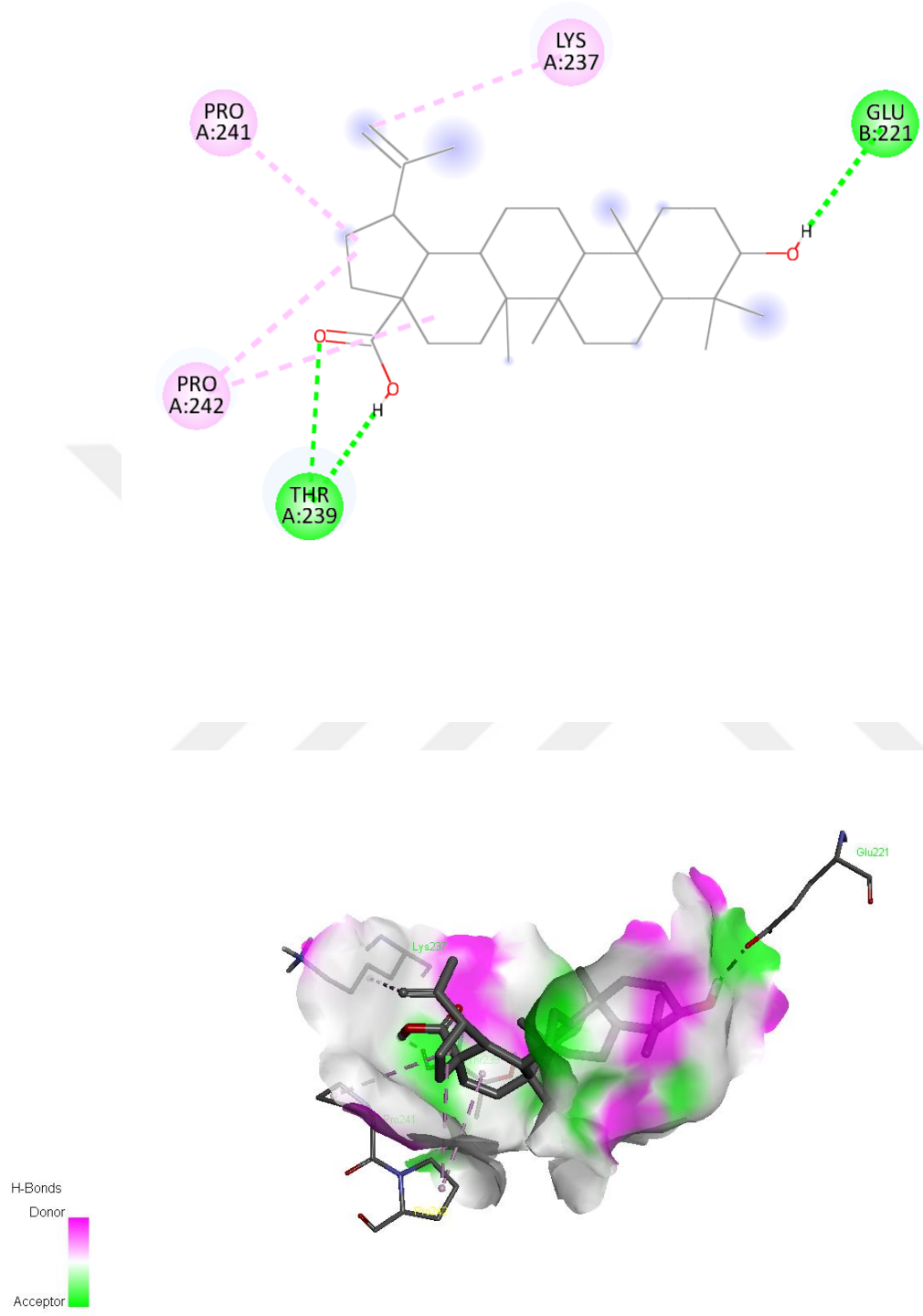

Figure 3.1 Chemical interactions between amino acids of tarheted protein (1MOX) and ligand 1 (betulinic acid).

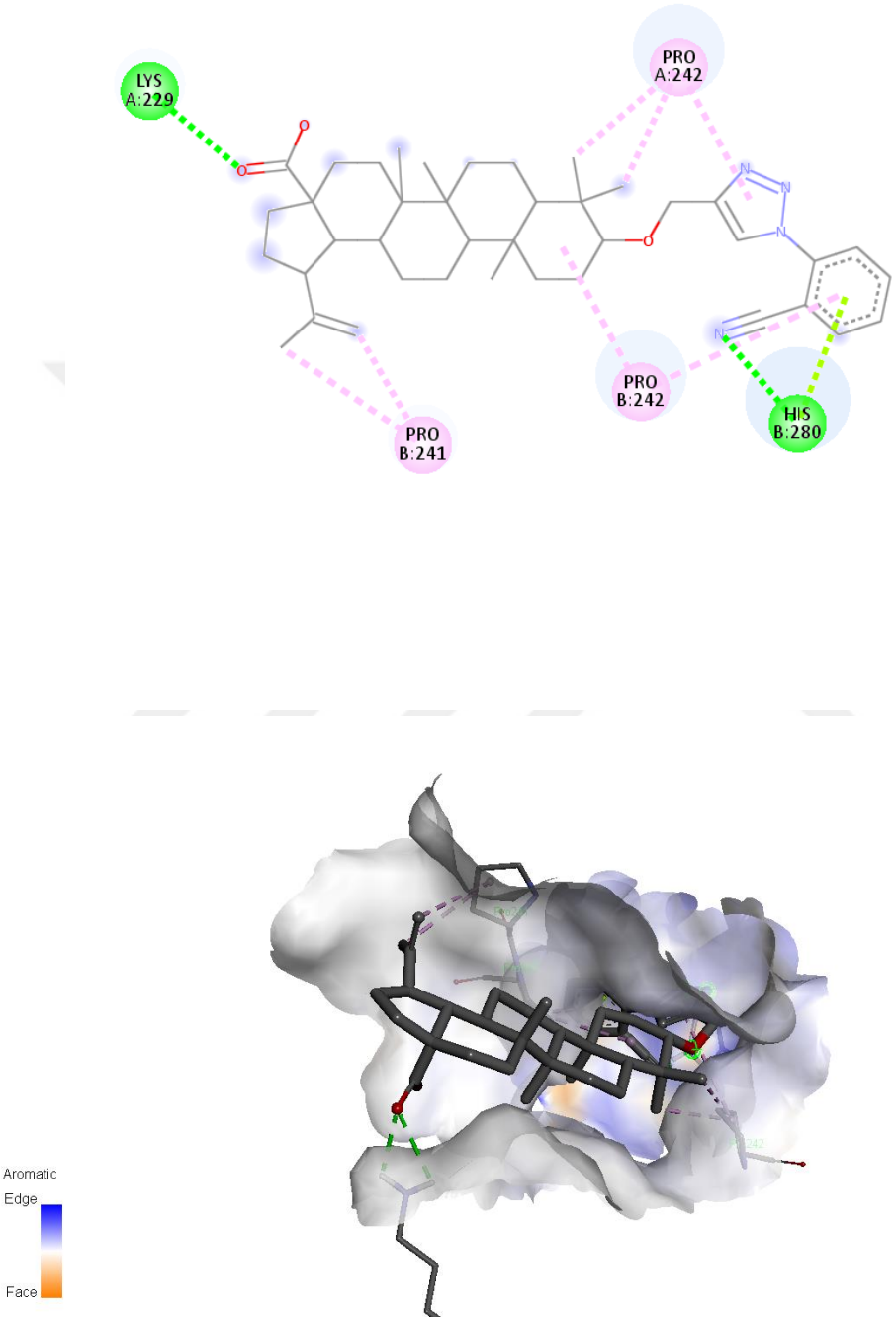

Figure 3.2 Chemical interactions between amino acids of targeted protein and ligand 2 (betulinic acid analogue 1).

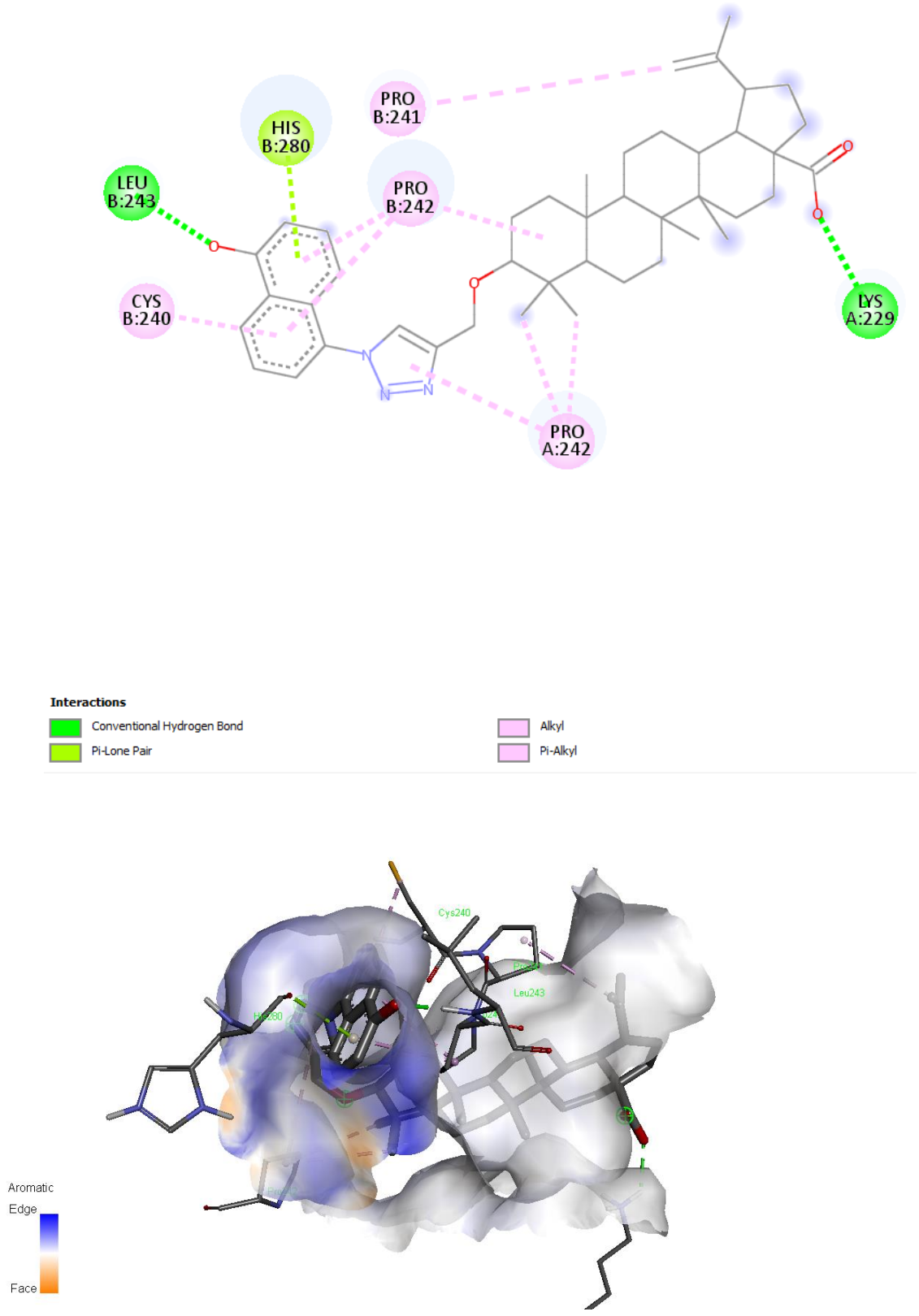

Figure 3.3 Chemical interactions between amino acids of targeted protein and ligand 3 (betulinic acid analogue 2).

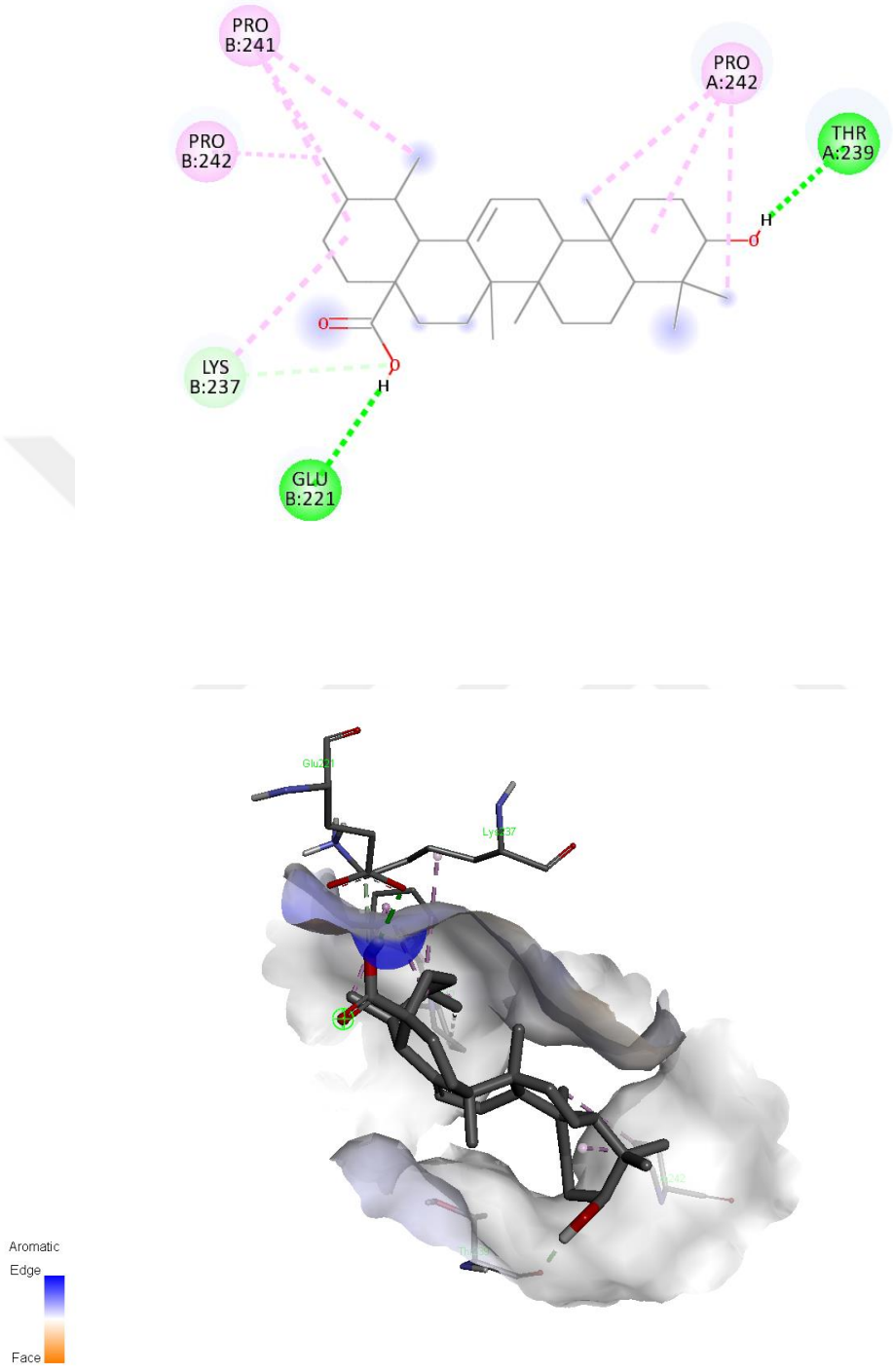

Figure 3.4 Chemical interactions between amino acids of targeted protein and ligand 4 (ursolic acid).

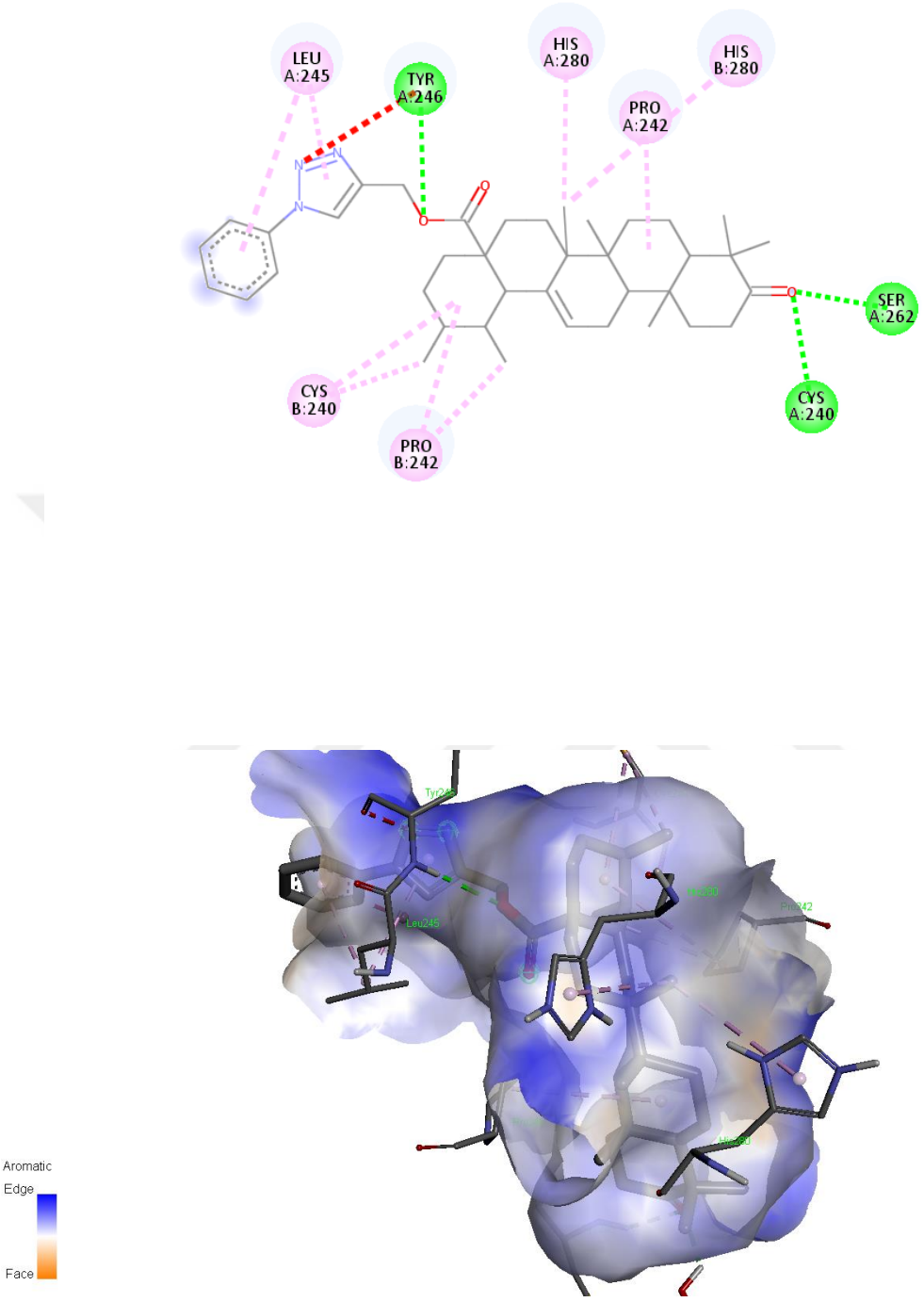

Figure 3.5 Chemical interactions between amino acids of targeted protein and ligand 5 (ursolic acid analogue 1).

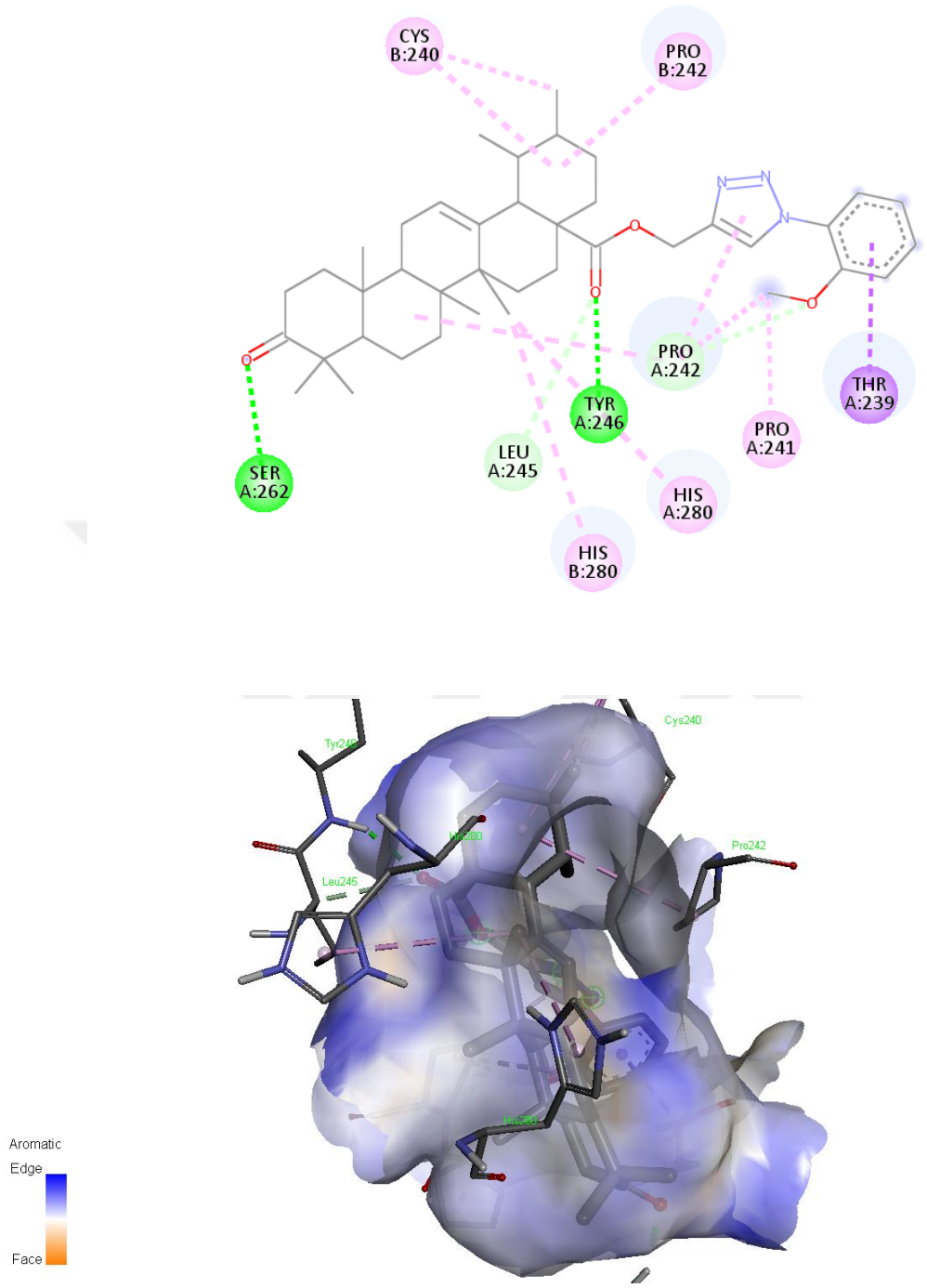

Figure 3.6 Chemical interactions between amino acids of targeted protein and ligand 6 (ursolic acid analogue 2).

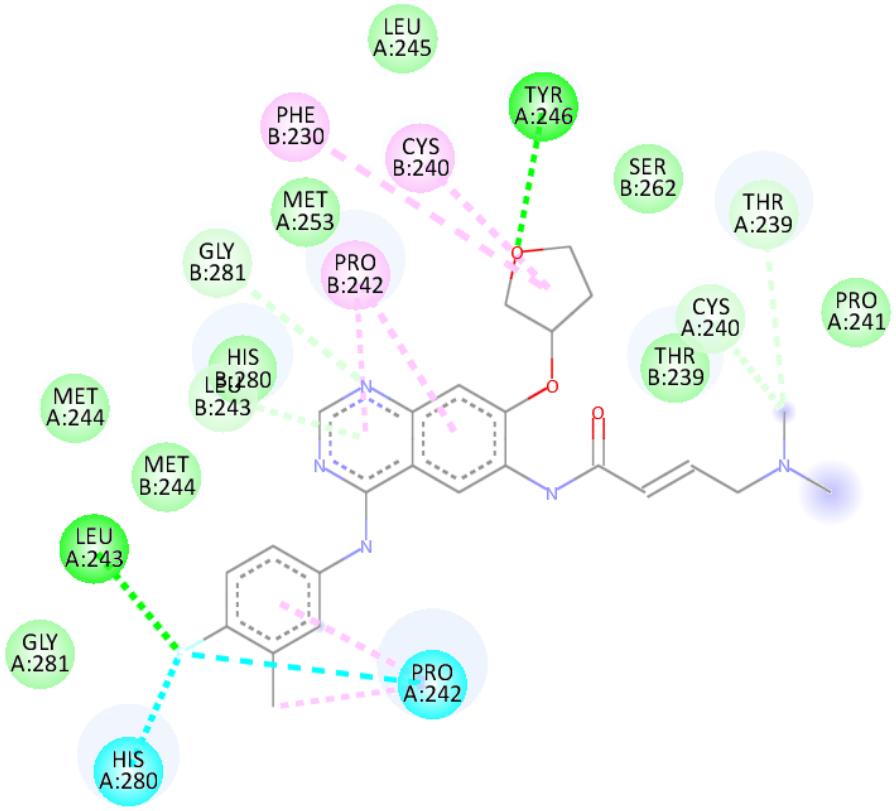

Figure 3.7 Chemical interactions between amino acids of targeted protein and control group (afatinib).

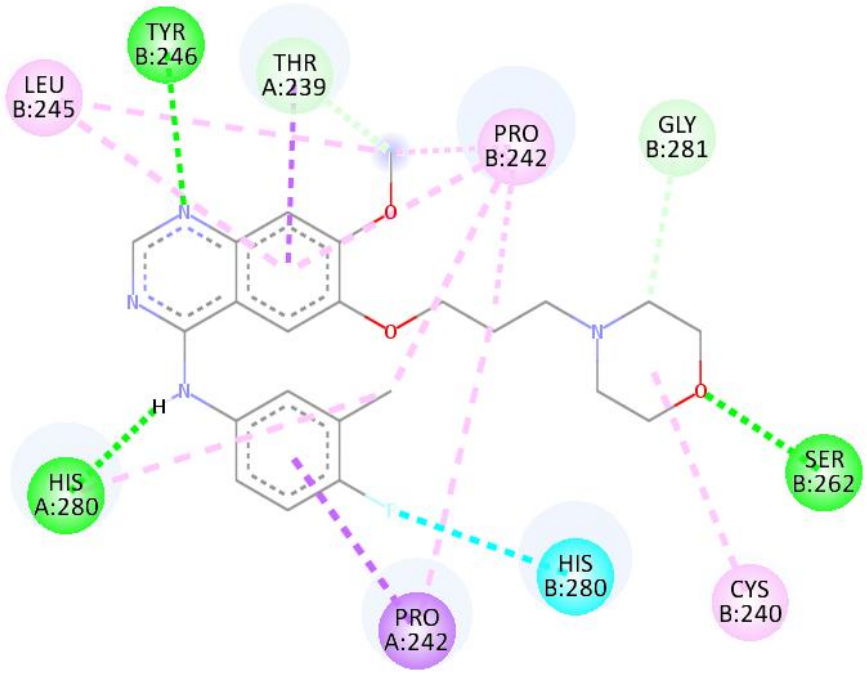

Figure 3.8 Chemical interactions between amino acids of targeted protein and control group (gefitinib)

Ligand 5 an ursolic acid derivative, had the lowest K_i value with a value of 1.66 pM, while 1 had the highest K_i value with a 573.53 nM. However, ligand 3 had an inhibition constant value of 1.31 nM, and displayed interactions Pro A:242, Lys A:229, His A:280, Leu B:243, Gly B:281, Lys B:260, Lys B:237, Pro B:241, Ser B:262, Lys B:237 amino acid residues of EGFR protein. Betulinic acid showed the lowest number of interactions (5) with the amino acids: Glu221, Pro242, Lys237, Pro241, Thr239.

Figures 3.1 – 3.6 contain the information on interactions between ligands and macromolecules in which amino acids that interact with ligands are summarized.

The hybrid strategy was reported to show distinct advantage over the conventional approach of co-administration of single hit molecules owing to better long-term prognosis and reduced toxicity (Choudhary et al. 2018). Lately, molecular docking has been commonly used in research studies to analyse protein-ligand interactions, clarify the ligand-binding mechanisms, and provide insights into the most stable ligand-protein interaction complexes (Hassan et al. 2020).

3.2 DRUG LIKENESS PREDICTION STUDIES

The term 'drug-like' states a potential and often means a type of mode of administration. Test ligands' drug-likeness potentials were calculated based on Lipinski's rules and Ghose filter using the SwissADME algorithm (URL-3) (Ghose et al. 1999, Lipinski 2004). Lipinski's rules state that physicochemical parameter regarding drug candidate molecules should not violate more than one of the following rules: molecular mass has to be less than 500 Da, LogP value (high lipophilicity) that is less than five, less than 5 hydrogen bond donors, less than 10 hydrogen bond acceptors associated with 90% of orally active drugs which have achieved phase II clinical status (Lipinski 2004). Additionally, Ghose and colleagues added the molar refractivity on the rules which should be between 40 and 130 (Ghose et al. 1999).

As it can be clearly seen in Table 3.3, ligands (1-6) were detected to satisfy most of the rules and qualified to be oral drugs according to Lipinski's rules. Molecular weights of betulinic acid and ursolic acid are less than 500 Da, while all hybrid compounds are higher than 500.

LogP is the log of the partition coefficient of a solute between octanol and water, at near infinite dilution. Lipinski's rule logP value must not be greater than 5. Our virtual screening results

showed that all ligand compounds had a higher logP values and then this rule seems to be violated for all ligands tested. However, rules satisfied for all ligands are 3/5 or 4/5.

3.3 CONCLUSION

Throughout history, natural sources have been a rich reservoir for conventional drug discovery and design with a great variety of chemical structures and fascinating biological activities. However, compounds obtained from natural sources such as higher-plants, bacteria, fungi and marine organisms display the features which are low toxicity, better efficacy and easy elimination from the body. Nonetheless, there are some restrictions on their druggability potential such as drug delivery, solubility, and stability.

So far, plant-based anticancer drugs were reported to be restricted by their low water-solubility. These are often hydrophobic in nature and require different solvents to formulate the dosage that also generate severe toxicity. Therefore, it is extremely crucial to design drugs with desired properties such as natural, efficient, and welfare and possess the useful characteristics including cost-effectiveness and targeted delivery. Unfortunately, discovery of selectively targeting drugs is still slow and failure rate is so high, especially in the metastasized cancer. In the current scenario, single target drug therapies have failed to achieve wanted biological actions particularly against complex disorders.

Within the sphere of cancer, a number of important new commercialized drugs have been obtained from natural sources either by structural modification of natural compounds or by the synthesis of new compounds using natural compounds as model (Fidoc vd. 2004).

The effect of epidermal growth factor-driven signalling in the pathogenesis and progression of cancer has been long demonstrated. However, different mechanisms might contribute to amplifying the signal pathways driven by epidermal growth factors such as overexpression of receptor proteins on the surface of cancer cells which can enhance their sensitivity. Nowadays, it is well-known that certain *proto*-oncogenes code for proteins which can be growth factors, growth factor receptors, or a partner of intracellular signal transduction pathways, suggesting that growth factors contribute to tumorigenesis via different mechanisms.

Recently, structural hybridization of natural products to improve the desired biological properties such as stability, target specificity, better efficacy, and less toxicity has been reported (Majeed et al. 2013). Triterpenoids can easily be esterified from hydroxy moiety attached to

backbone and double bond functionality which could be employed to synthesize more stable and specific hybrid molecules.

Here, our findings strongly support the hypothesis that naturally occurring products could modulate the EGFR protein involved in tumorigenesis.

Results revealed that natural products or semi-synthetic hybrid compounds might be key molecules for cancer therapy. Our findings also displayed that the triazole group has a critical role for molecular interactions between hit molecule and ligand. Herein, we also recommend that ligands should be further investigated *in vitro* and *in vivo* models to clarify their potentials.

REFERENCES

Ali R and Wendt M K (2017) The paradoxical functions of EGFR during breast cancer progression. *Signal Transduction and Targeted Therapy*, 2, 16042.

Berger T, Saunders M E and Mak T W (2016) Beyond the oncogene revolution: four new ways to combat cancer. *Cold Spring Harb Symp Quant Biol.* 81: 85–92.

Cha H J, Park M T, Chung H Y, Kim N D, Sato H, Seiki M and Kim K W (1998) Ursolic acid-induced down-regulation of MMP-9 gene is mediated through the nuclear translocation of glucocorticoid receptor in HT1080 human fibrosarcoma cells *Oncogene*, 16: 771-778.

Choudhary S, Singh P K, Verma H and Singh H (2018) Success stories of natural product-based hybrid molecules for multifactorial diseases. *European Journal of Medicinal Chemistry*, 151: 62-97.

Cohen M H, Williams G A, Sridhara R, Chen G and Pazdur R (2003) FDA drug approval summary: gefitinib (ZD1839) (Iressa) tablets. Oncologist. 8(4):303-6.

Daina A, Michielin O and Zoete V (2017) SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Scientific Reports. 7:1–13.

Dungo R T and Keating G M (2013) Afatinib: first global approval. Drugs 73(13):1503-15.

Fidock D A, Rosenthal P J, Croft S L, Brun R and Nwaka S (2004) Antimalarial drug discovery: efficacy models for compound screening. *Nat Rev Drug Discov*, 3: 509–520.

Ghose A K, Viswanadhan V N and Wendoloski J J (1999) A knowledge–based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases. *J Comb Chem*, 1(1): 55–56.

Hsu Y L, Kuo P L and Lin C C (2004) Proliferative inhibition, cell-cycle dysregulation, and induction of apoptosis by ursolic acid in human non-small cell lung cancer A549 cells. *Life Science*, 75(19): 2303-2316.

Gao Y, Jia Z, Kong X, Li Q, Chang DZ, Wei D, Le X, Suyun H, Huang S, Wang L and Xie K (2011) Combining betulinic acid and mithramycin a effectively suppresses pancreatic cancer by inhibiting proliferation, invasion, and angiogenesis. *Cancer Res* 2011; 71: 5182–5193.

Fritsche MK and Knopf A (2017) The tumor suppressor p53 in mucosal melanoma of the head and neck. *Genes*, 8 (12): 384.

Hughes J P, Rees S, Kalindjian S B and Philpott K L (2011) Principles of early drug discovery. *Br J Pharmacol.* 162 (6): 1239-49.

Kasperczyk H, La Ferla-Brühl K, Westhoff MA, Behrend L, Zwacka RM, Debatin KM and Fulda S (2005) Betulinic acid as new activator of NF-κB: molecular mechanisms and implications for cancer therapy. *Oncogene,* 24: 6945–6956.

Kovacs E, Zorn J A, Huang Y, Barros T and Kuriyan J A (2015) Structural Perspective on the Regulation of the Epidermal Growth Factor Receptor. *Annu. Rev. Biochem.* 84: 739–764.

Lee E Y and Muller W J (2010) Oncogenes and tumor suppressor genes. *Cold Spring Harb Perspect Biol*. 2(10):a003236.

Margold M and Jansson K N (2012) Evaluation of Data Sources for Mapping Glacial Meltwater Features. *International Journal of Remote Sensing*, 33 (8): 2355-2377.

Majeed R, Hussain A, Sangwan P L, Chinthakindi P K, Khan I, Sharma P R, Koul S, Saxena A K and Hamid A (2016) PI3K target based novel cyano derivative of betulinic acid induces its signalling inhibition by down-regulation of pGSK3β and cyclin D1 and potentially checks cancer cell proliferation. *Mol Carcinog.* 55(5):964-76.

Majeed R, Hamid A, Sangwan P L, Chinthakindi P K, Koul S, Rayees S, Singh G, Mondhe D M, Mintoo M J, Singh S K, Rath S K and Saxena A K (2014) Inhibition of phosphotidylinositol-3 kinase pathway by a novel naphthol derivative of betulinic acid induces cell cycle arrest and apoptosis in cancer cells of different origin. *Cell Death Dis.* 9;5(10):e1459.

Majeed R, Sangwan PL, Chinthakindi PK, Khan I, Dangroo NA, Thota N, Hamid A, Sharma PR, Saxena AK and Koul S. (2013) Synthesis of 3-O-propargylated betulinic acid and its 1,2,3-triazoles as potential apoptotic agents, *European Journal of Medicinal Chemistry*, 63:-792.

Mazzarella L, Guida A and Curigliano G (2018) Cetuximab for treating non-small cell lung cancer. *Expert Opin Biol Ther.* 18(4):483-493.

REFERENCES (continued)

McConkey BJ, Sobolev V and Edelman M. (2002) The performance of current methods in ligand-protein docking. *Current Science,* 83:845–855.

Minkovsky N and Berezov A (2008) BIBW-2992, a dual receptor tyrosine kinase inhibitor for the treatment of solid tumors. *Current Opinion in Investigational Drugs*, 9 (12): 1336–46.

Moghaddam M G, Ahmad J B H and Samzadeh-Kermani A (2012) Biological activity of betulinic acid: a review. *Phamacol Pharm*, 2012; 3: 119–123.

Lipinski C A (2004) Lead-and drug-like compounds: the rule-of-five revolution. *Drug Discovery Today: Technologies.* 1(4):337–341.

Newman D J and Cragg G M (2020) Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* 83, 770–803.

Olayioye M A, Neve R M, Lane H A and Hynes N E (2000) The ErbB signaling network: receptor heterodimerization in development and cancer. *EMBO J*, 19, 3159-3167.

Pao W, Miller V, Zakowski M, Doherty J, Politi K, Sarkaria I, Singh B, Heelan R, Rusch V, Fulton L, Mardis E, Kupfer D, Wilson R, Kris M and Varmus H (2004) EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. *Proceedings of the National Academy of Sciences of the United States of America*. 101 (36): 13306–11.

Rashid S, Dar B A, Majeed R, Hamid A and Bhat B A (2013) Synthesis and biological evaluation of ursolic acid-triazolyl derivatives as potential anti-cancer agents. *Eur J Med Chem*, 2013 66:238-45.

Retzlaff F and Ueber H G (1902) Ueber herba gratiolae. *Arch. Pharm*. 1902, 240, 561–568.

Sabapathy K (2015) The contrived mutant p53 oncogene - beyond loss of functions. Front *Oncol*. 5:276.

Seyfried T N and Huysentruyt L C (2013) On the origin of cancer metastasis. *Crit Rev Oncog*. 18(1-2):43-73.

Shen M, Hu Y, Yang Y, Wang L, Yang X, Wang B and Huang M (2019) Betulinic Acid Induces ROS-Dependent Apoptosis and S-Phase Arrest by Inhibiting the NF-κB Pathway in Human Multiple Myeloma. *Oxid Med Cell Longev*. 9 5083158.

REFERENCES (continued)

Shtiegman K, Kochupurakkal B S, Zwang Y, Pines G, Starr A, Vexler A, Citri A, Katz M, Lavi S, Ben-Basat Y, Benjamin S, Corso S, Gan J, Yosef R B, Giordano S and Yarden Y (2007) Defective ubiquitinylation of EGFR mutants of lung cancer confers prolonged signaling. *Oncogene*, 26: 6968–6978.

Sohn K H, Lee H Y, Chung H Y, Young H S, Yi S Y and Kim K W (1995) Anti-angiogenic activity of triterpene acids. *Cancer letters*, 94: 213-218.

Spreitzer H (2008). Neue Wirkstoffe – Tovok. *Österreichische Apothekerzeitung* (in German) (10/2008): 498.

Tan Y, Yu R and Pezzuto J M (2003) Betulinic acid-induced programmed cell death in human melanoma cells involves mitogen-activated protein kinase activation. *Clin Cancer Res* 9: 2866–2875.

Uribe M L, Marocco I and Yarden Y (2021) EGFR in Cancer: signaling mechanisms, drugs, and acquired resistance. *Cancers*, 13:2748.

Vavalà T (2017) Role of afatinib in the treatment of advanced lung squamous cell carcinoma. *Clinical Pharmacology*, 9: 147–157.

Velez-Cruz R and Johnson D G (2017) The retinoblastoma (RB) tumor suppressor: pushing back against genome instability on multiple fronts. *Int J Mol Sci.* 18(8):1776.

Wang Yong-Jian, Lu Jun, Wu Dong-mei, Zheng Zi-hui, Zheng Yuan-Lin, Wang Xiao-hui, Ruan Jie, Sun Xiao and Shan Qun (2011) "Ursolic acid attenuates lipopolysaccharide-induced cognitive deficits in mouse brain through suppressing p38/NF-κB mediated inflammatory pathways". *Neurobiology of Learning and Memory*. 96 (2): 156–165.

Waterman H, Katz M, Rubin C, Shtiegman K, Lavi S, Elson A, Jovin T and Yarden Y (2002) A mutant EGF-receptor defective in ubiquitylation and endocytosis unveils a role for Grb2 in negative signaling. *EMBO J.* 21, 303–313.

Wee P and Wang Z (2017) Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. *Cancers.* 2017; 9(5):52.

World Health Organization (2021). World Health Organization model list of essential medicines: 22nd list (2021). Geneva: World Health Organization.

Wu Dong-mei, Lu Jun, Zhang Yan-qiu, Zheng Yuan-lin, Hu Bin, Cheng Wei, Zhang Zi-feng and Li Meng-qiu (2013) Ursolic acid improves domoic acid-induced cognitive deficits in mice. *Toxicology and Applied Pharmacology*, 271 (2): 127–136.

REFERENCES (continued)

Yarden Y and Sliwkowski M X (2001) Untangling the ErbB signalling network. *Nat Rev Mol Cell Biol* 2:127-137.

Yan J, Zhang G, Pan J and Wang, Y (2014) A-Glucosidase inhibition by luteolin: Kinetics, interaction and molecular docking. *International Journal of Biological Macromolecules*, 64, 213–223.

Zhang X, Gureasko J, Shen K, Cole PA and Kuriyan J (2006). An allosteric mechanism for activation of the kinase domain of epidermal growth factor receptor. *Cell*, 125 (6):1137-1149.

Zhang X, Zhang S, Zhao S, Wang X, Liu B and Xu H (2021) Click Chemistry in Natural Product Modification. *Front Chem.* 17;9: 774977.

Zhang Y, Li X, Ciric B, Curtis M T, Chen W J, Rostami A and Zhang G X (2020) A dual effect of ursolic acid to the treatment of multiple sclerosis through both immunomodulation and direct remyelination". *Proceedings of the National Academy of Sciences*, 117 (16): 9082–9093.

Zhu K, Liu Q and Zhou Y (2015) Oncogenes and tumor suppressor genes: comparative genomics and network perspectives. *BMC Genom.* 16(suppl 7):S8.

URL-1 < *www.wcrf.org/cancer-trends/worldwide-cancer-data*>, Last Visit: 05.09.2023.

URL-2 <*www.drozdogan.com/turkiye-kanser-istatistikleri-2020*>, Last Visit: 05.09.2023.

URL-3 < *http://www.swissadme.ch/index.php* >, Last Visit: 05.09.2023.

APPENDICES

Appendix A: SwissADME Results of Ligands

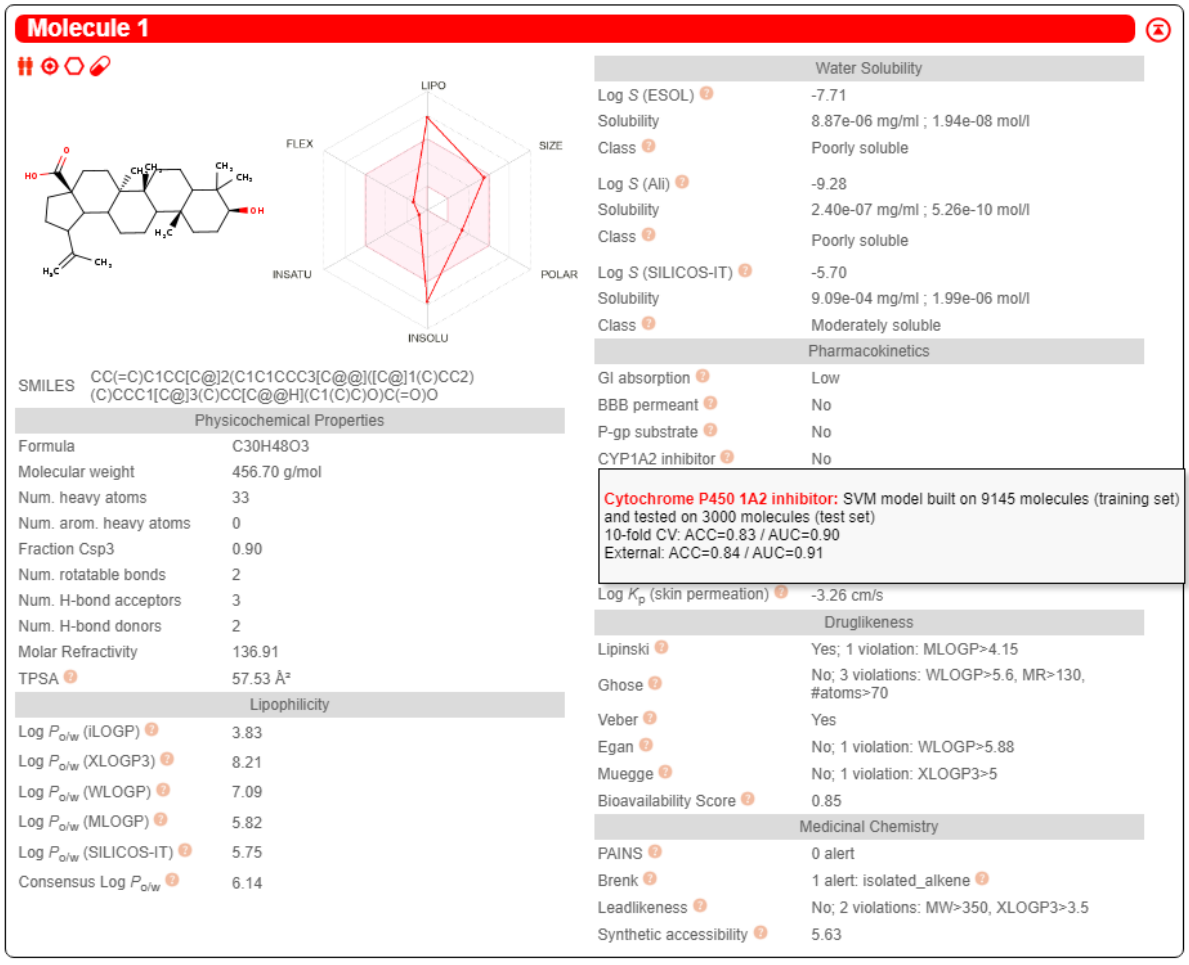

Figure A.1 SwissADME results of ligand **1** (betulinic acid)

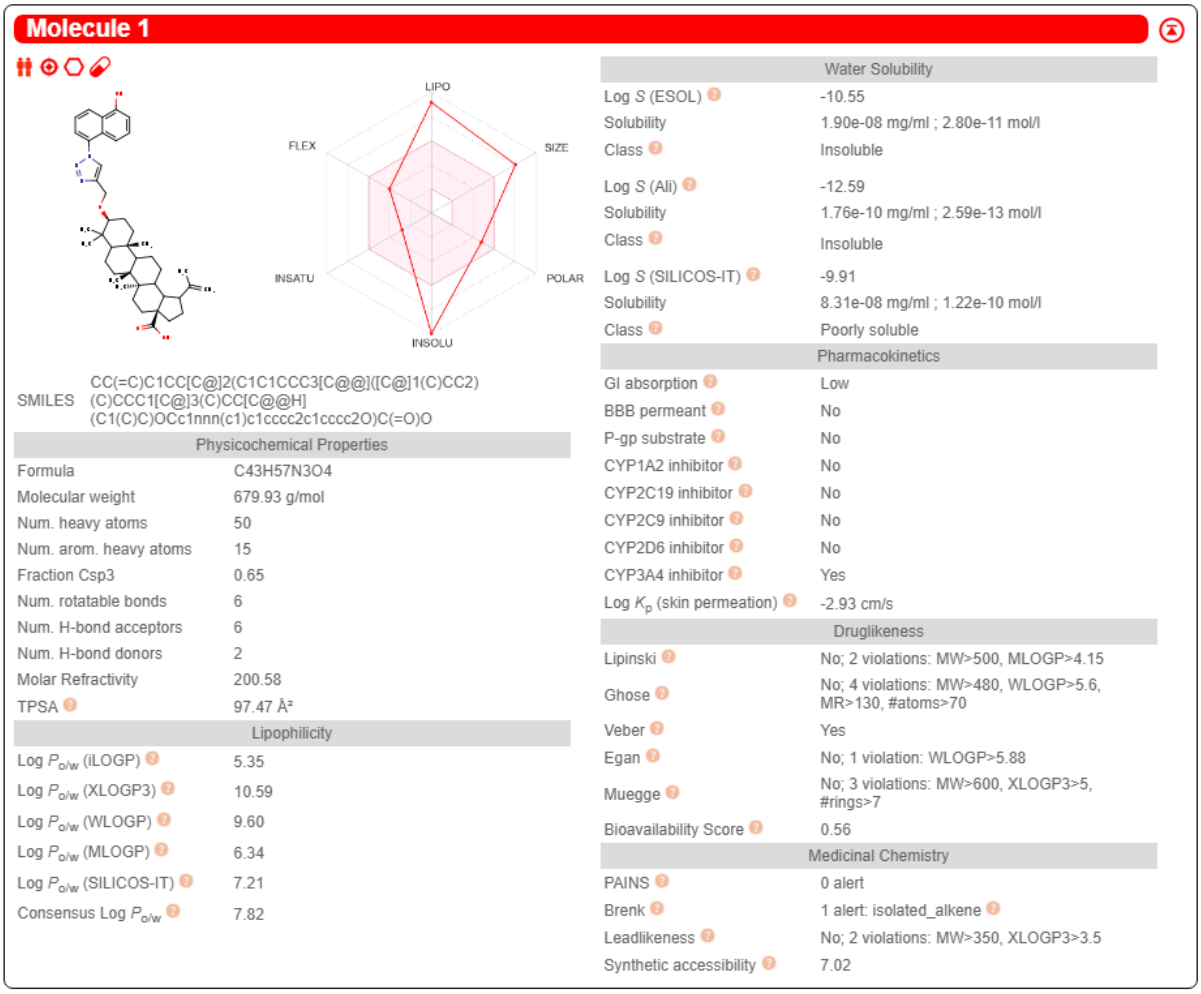

Figure A.2 SwissADME results of ligand **2** (betulinic acid analogue 1).

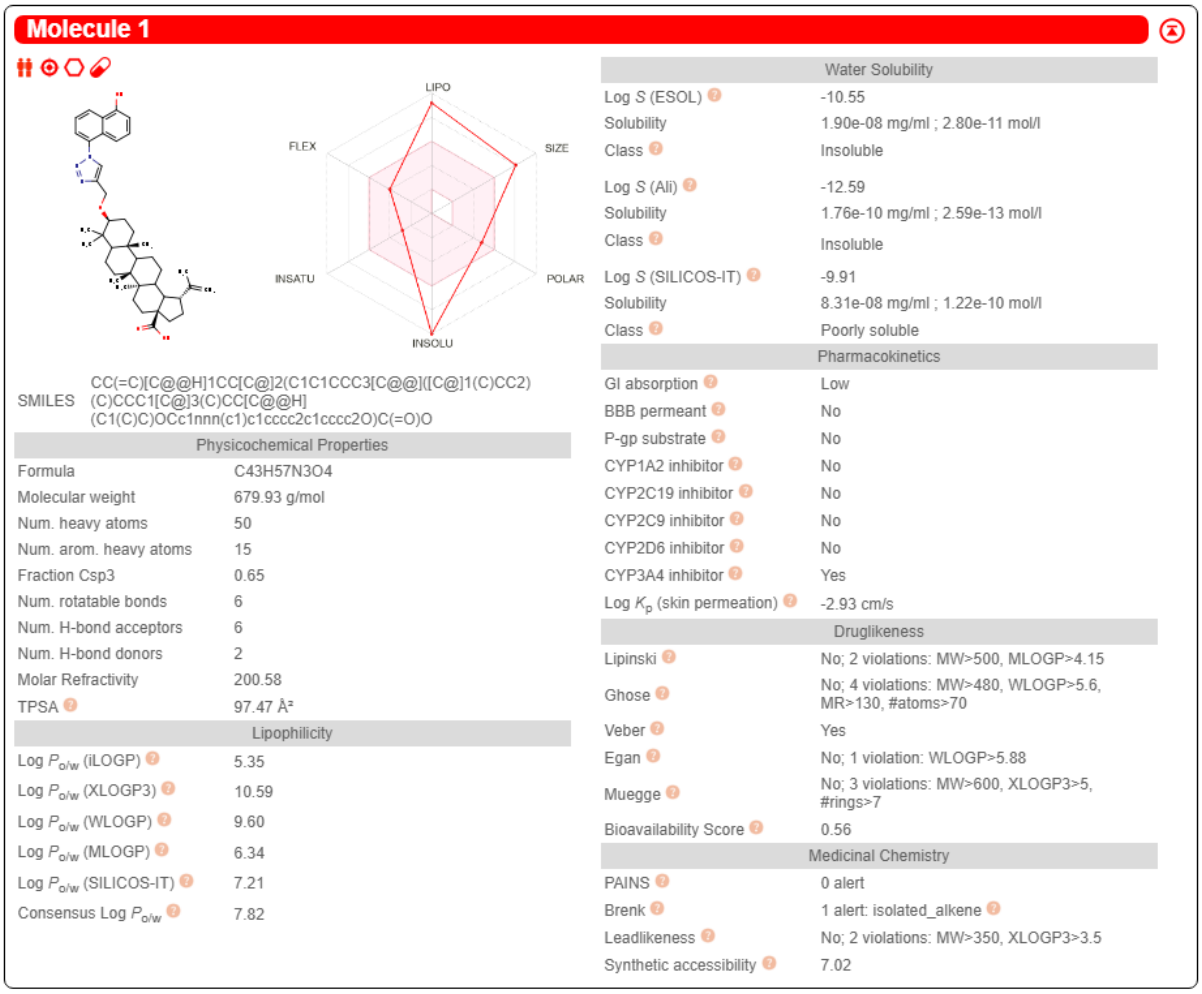

Figure A.3 SwissADME results of ligand **3** (betulinic acid analogue 2).

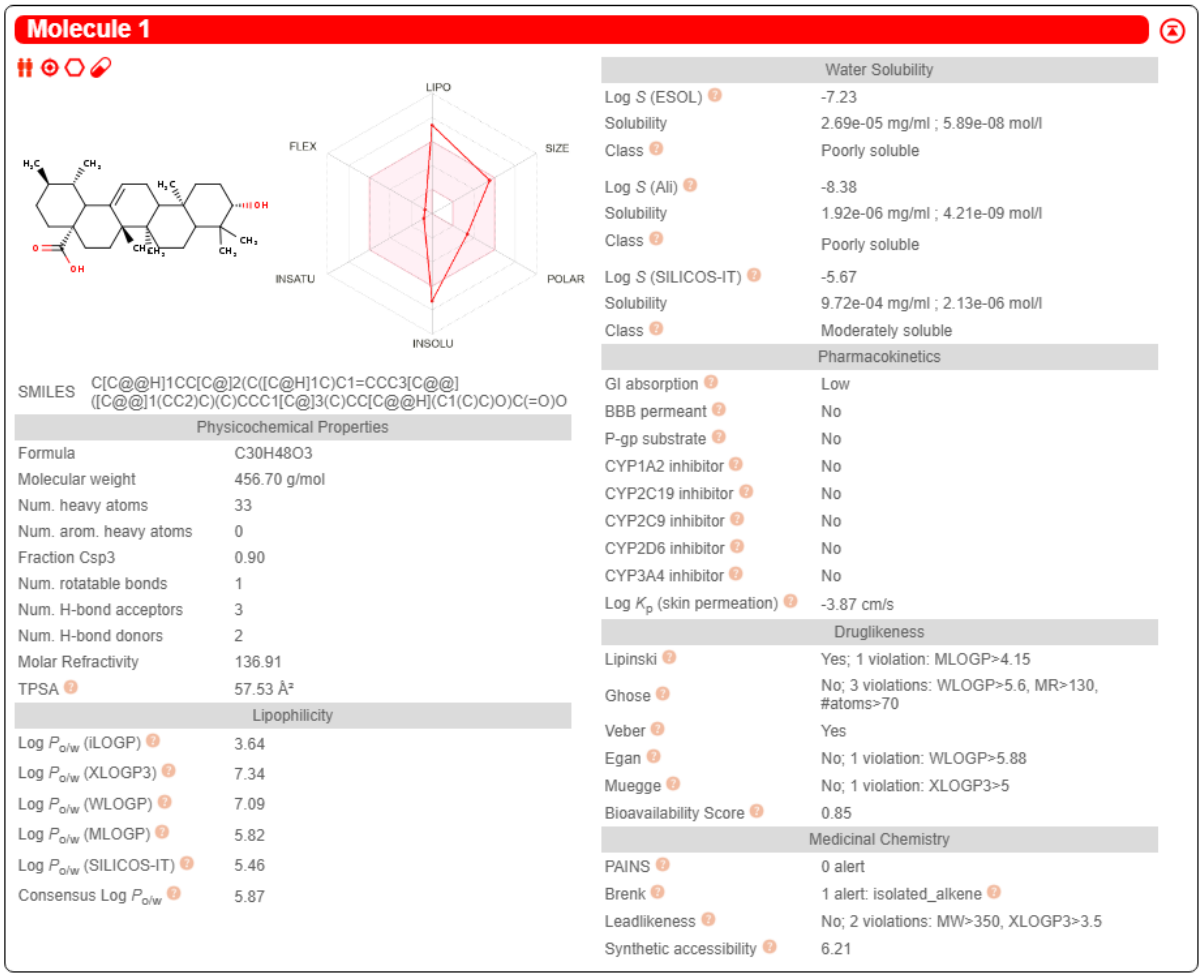

Figure A.4 SwissADME results of ligand **4** (ursolic acid)**.**

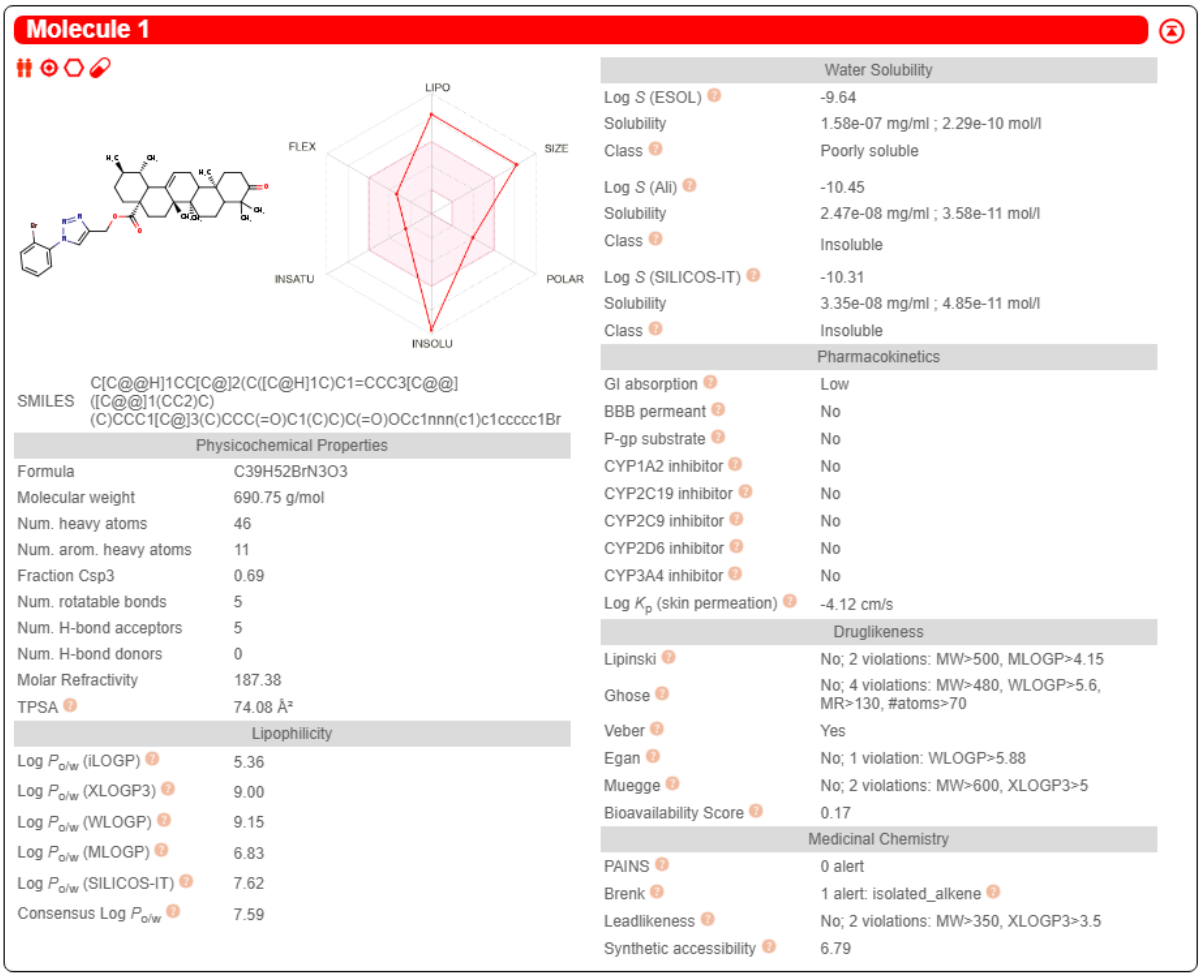

Figure A.5 SwissADME results of ligand **5** (ursolic acid analogue **1**) **.**

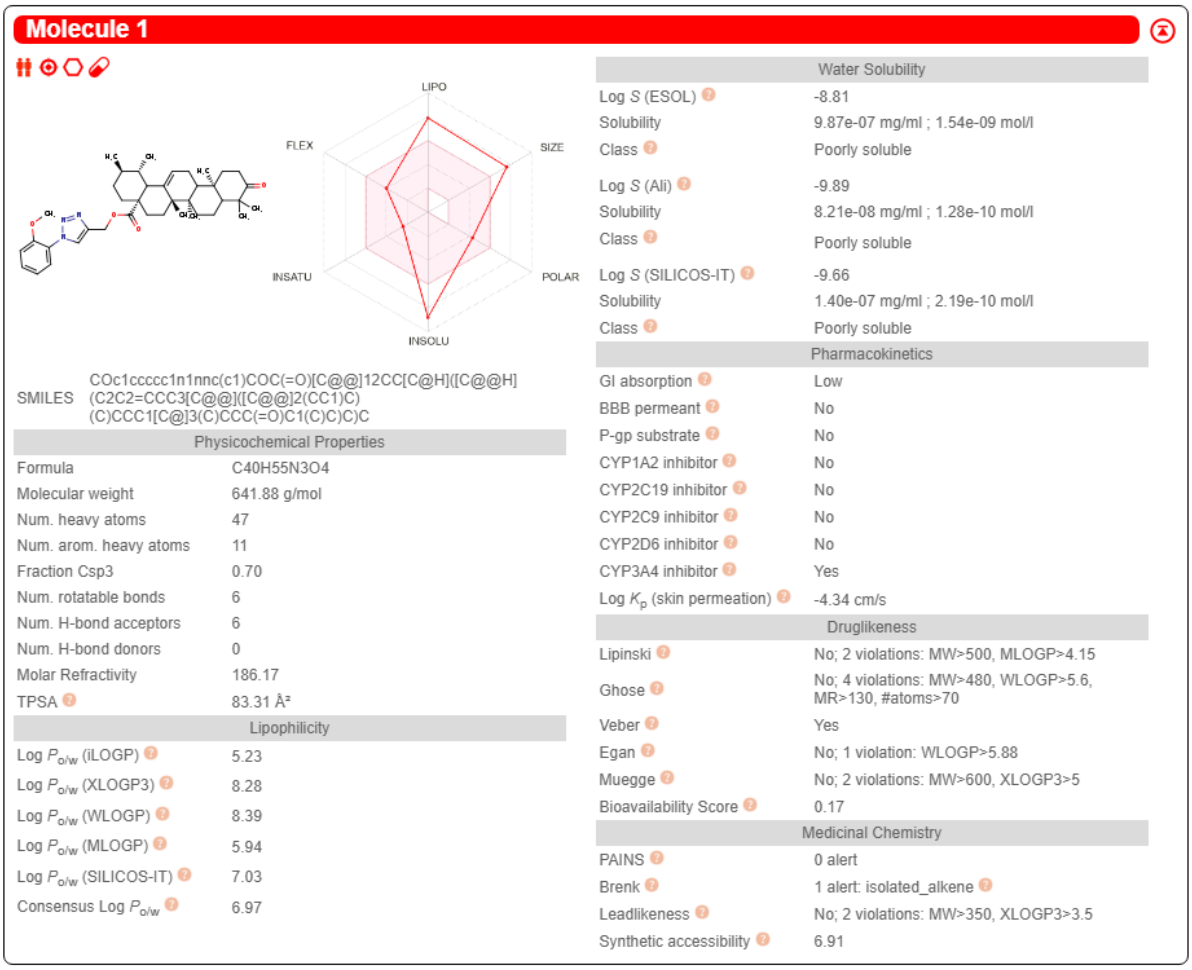

Figure A.6 SwissADME results of ligand **6** (ursolic acid analogue **1**)**.**

CURRICULUM VITAE

Özlem Aydın, She graduated from Beşikdüzü IMKB Anatolian Teacher High School between 2009-2013. Ms. Aydin completed her undergraduate education (BSc) at the Department of Biomedical Engineering, Zonguldak Bülent Ecevit University in 2020. Currently, she has been studying her master's degree in the same department. She speaks good English.