2017

Design, Synthesis and Biological Screening of Novel Cucsinspired Estrone Analogues Towards Treatment of Hepatocellular Carcinoma

Mater Hussen Mahnashi
South Dakota State University

DESIGN, SYNTHESIS AND BIOLOGICAL SCREENING OF NOVEL CUCS-

INSPIRED ESTRONE ANALOGUES TOWARDS TREATMENT OF

HEPATOCELLULAR CARCINOMA

By

MATER HUSSEN MAHNASHI

A dissertation submitted in partial fulfillment of the requirements for the

Doctoral of Philosophy

Major in Chemistry

South Dakota State University

2017

DESIGN, SYNTHESIS AND BIOLOGICAL SCREENING OF NOVEL CUCS-INSPIRED ESTRONE ANALOGUES TOWARDS TREATMENT OF HEPATOCELLULAR CARCINOMA

MATER HUSSEN MAHNASHI

This dissertation is approved as a creditable and independent investigation by a candidate for Doctor of Philosophy in Chemistry degree and is acceptable for meeting the dissertation requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Fathi T. Halaweish, Ph.D.
Dissertation Advisor   Date

Douglas Raynie, Ph.D./
Head, Department of Chemistry   Date
& Biochemistry

Dean, Graduate School   Date

This dissertation is dedicated first of all to ALLAH who gave me the strength, guide me to the right path and help me to accomplish my dream.

Second, this dissertation is dedicated to my family, who has always stand for me and supported me.

My Parents

I could never have accomplished my dream without your prayers, supports and continuous encouragement. Thank you for everything and for your love.

My Advisor

To Dr. Fathi Halaweish, who is always support and guide me. Thank you so much.

My Wife

To the women whom I love, respect and believe in, thank you so much for every single thing that you have done to me. If I want to list all of these things, it will take the whole thesis. Thank you so much.

My Friends

To the people who are always with me and make me laugh at the sad time, support me at hard time. Thank you so much.

ACKNOWLEDGEMENTS

CONTENTS

ABBREVIATIONS

ADME	Absorption, Distribution, Metabolism and Excretion
EGFR	Epidermal Growth Factors
ERK	Extracellular Signal-Related Kinase
FDA	Food and Drug Administration
HBV	Hepatitis B virus
HCC	Hepatocellular Carcinoma
HCV	Hepatitis C virus
HTS	High Throughput Screening
LC	Lead Compound
MAPK	Mitogen Activated Protein Kinase
MDR	Multidrug Resistance
MEK	MAPK/ERK kinase
MRP	Multidrug resistance protein
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NMR	Nuclear Magnetic resonance
SAR	Structure Activity Relationship
STAT	Signal Transducer and Activator of Transcription

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

DESIGN, SYNTHESIS AND BIOLOGICAL SCREENING OF NOVEL CUCS-

INSPIRED ESTRONE ANALOGUES TOWARDS TREATMENT OF

HEPATOCELLULAR CARCINOMA

MATER MAHNASHI

2017

Cucurbitacins (CUCS) are natural products with highly oxygenated tetracyclic triterpenes produced mostly by Cucurbitaceae family plant. They are known for their therapeutic efficiency with different biological activities, such as anti-inflammatory, hepatoprotective and anti-cancer. Hepatocellular carcinoma (HCC) is the third leading cause of death worldwide. Previous reports have shown the ability of CUCS to inhibit the growth of hepatocellular carcinoma cell lines (HepG-2) significantly. Structural activity relationship studies suggested the potential of the 23, 24 enone side chain of CUCS to bind to the Epidermal Growth Factor Receptor (EGFR). Due to the limited quantities of CUCS upon isolation and the challenges of total synthesis of CUCS, therefore estrone skeleton were used as a starting scaffold to synthesize CUCS-inspire estrone analogues (CIEA) targeting HCC. Molecular docking study of cucurbitacin-inspired estrone analogs was conducted using 1M17 (EGFR receptor) co-crystallized with Erolitinib (known EGFR HCC anti-cancer chemotherapeutic drug) and several analogs were identified from the docking study and were processed for multiple steps organic synthesis. Novel CUCUS-inspired estrone analogs with aliphatic enone side chain such as **MMA102**, **MMA132** were synthesized by installing the CUCS side chain at C17 of estrone scaffold. In addition, various CUCUS-inspired estrone analogs with different aliphatic, aromatic and

heterocyclic pharmacophores at C-3, C-16 and C-25 were synthesized. The novel analogs showed a comparable affinity to EGFR receptor based on the docking study and improved binding through hydrophobic filling of the binding of EGFR pocket and hydrogen bonding interactions. Cell proliferation inhibition assay results demonstrated the ability of analogues **MMA102** and **MMA132** to inhibit HCC cell line (HepG2) proliferation with IC_{50} 3μM and 2μM, respectively in comparison to Erlotinib IC_{50} (of 25 μM). Western blot experiments proved that compounds **MMA132** has the ability to bind to the EGFR-TK and inhibit its phosphorylation by 90%. Flow-cytometry/cell cycle arrest study showed that our lead compound **MMA132** induce a significant change in G1/S phases at different IC_{50} concentrations (1μM, 2μM and 3μM). The novel synthesized cucurbitacin-inspired estrone analogs showed a significant 12 times more cytotoxicity than standard chemotherapy and bind through hydrogen bonding to the same amino acids that Erlotinib binds to in the 1M17 EGFR pocket. Compound **MMA311**, which consist of 2-bromothiophene enone side chain along with methoxy at C-3 and double bond at C-16 and C-17, showed a significant and potent anti-proliferation activity with IC_{50} value of 0.7 μM. Due to the high hydrophobic character of the first set of compounds and the expected metabolism of estrone by O-dealkylation of methoxy group at C-3, which may trigger estrone side effects, C-3 methoxy group were substituted with sulfamoyl moiety to improve the pharmacokinetic profile of the synthesized analogues. Several CIEA analogs that contain sulfamoyl group at C-3 and various aliphatic, aromatic and heterocyclic enone side chains at C-17 were synthesized and biologically tested as EGFR inhibitors. Compounds that contain heterocyclic enone side chains at C-17 along with sulfamoyl moiety or hydroxyl group at C-3 such as **MMA297, MMA314, MMA313** and **MMA312** showed an outstanding cytotoxicity with

IC$_{50}$ 1 µM, 1.5 µM, 9 µM and 8 µM; respectively, in comparison to the Erlotinib IC$_{50}$ of 25 µM. The potential of our novel CIEA to overcome cancer resistance to current chemotherapy was explored and identified their ability to decrease the drug resistance by inhibiting MRP1 utilizing high-content based assay in presence of calcein-AM as MRP1 substrate. Compounds **MMAmix, MMA242, MMA132, MMA335, MMA337** and **MMA320** showed potential inhibitory activity on **MRP1** with inhibitory activity of 70%, 63%, 46.2%, 46%, 30% and 22%; respectively, in comparison to MK-571, which known MRP1 inhibitor.

Our study demonstrated the design, synthesis of novel CIEA analogs of potent anti-proliferation/anticancer activities toward hepatocellular carcinoma and potential application to overcome cancer resistance to current chemotherapeutic agents.

Chapter One

General Introduction and Background

1.1   Drug Discovery:

Drug discovery is a convoluted procedure that consumes much time and money, taking an average time of 10-15 years and cost of around $ 1 billion to get a new drug in the market. This complex procedure needs a great number of steps, skills and technologies such as chemoinformatics, molecular modeling, chemical synthesis, toxicity screening and clinical studies [1]. Typically, the drug discovery pipeline starts from a disease description, followed by molecular targets identification and validation, thus utilizing high throughput screening (HTS) and/or Insilco–invitro screening to identify hit compounds. Finally, hit compounds will be optimized to find the lead compound followed by animal studies to identify the compounds pharmacokinetics, and finally the clinical trials to test the drug before it reaches the pharmaceutical market (Fig.1.1) [1, 2]. All of these steps will be completed sequentially so, if any one of these steps delayed, the whole process is slowed [2]. A strategy known as High Throughput Screening (HTS) has adopted by pharmaceutical industries to identify and develop a hit compounds which apply insilco molecular modeling and computer assisted drug design (CADD) using scoring function and docking (**Fig.1.1**). All of these processes used by pharmaceutical industries to speed the drug discovery process.

HTS used by pharmaceutical companies to screen 1 to 5 million compounds in few weeks which is a very coasty process. For example, an HTS of 1 million compounds might cost an average of $ 500,000–1,000,000 [3]. However, the efficacy or toxicity problems

associated with many compounds are the reason behind the fail test at the last stages of clinical trials [4, 5].
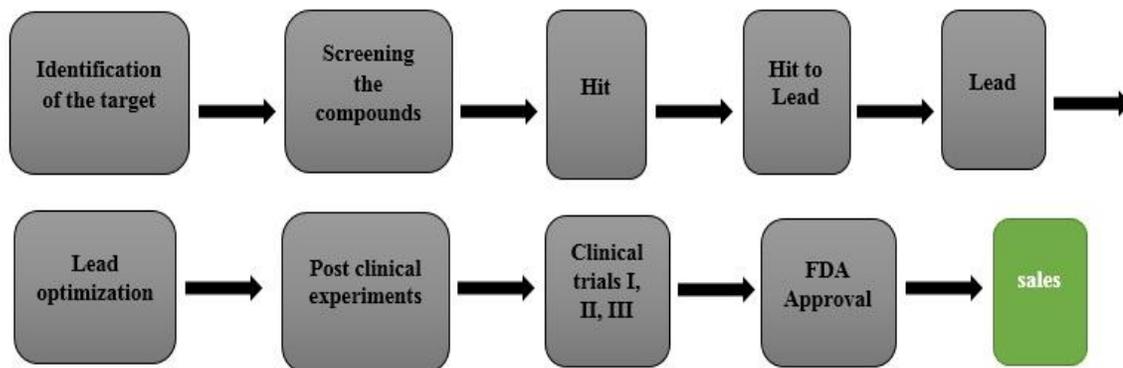
Figure 1.1 Drug discovery processes (10-15 years/ 1 billion dollars).

In addition, the physic-chemical characteristics such as, absorption, distribution, metabolism and elimination (ADME) could play an important role in the efficacy and safety of a new drug [6-8]. Therefore, to eliminate some of these issues, scientists create a virtual library of thousands of chemical compounds that can undergo virtual insilco screening to obtain hit compounds prior to the bench work, which includes synthesis and biological evaluation, to find potential drug candidates [8-10].

Virtual library generation in the drug discovery field is a time- and effort-saving step, where different factors can be chosen during the design process of the library. The process of designing the virtual library considers the possible synthetic approaches which can be utilized on a chemical structure, including the capability of installing diverse functional groups (pharmacophores) on the parent compound using viable chemical reactions [9]. One of the common approaches used in the design stage is the concepts of

bioisostrism, which is a planning technique used in medicinal chemistry field to manage systematic molecular modifications aiming. The process of bioisostrism refers to substituting chemical functional groups with others similar in their physical or chemical properties, which generally result in similar or better biological activity. The bioisostrism molecular modification process uses a lead compound (LC) in which its chemical structure, mechanism of action, drug-protein interaction and important pharmacophores are known. Additionally, the concept of bioisostrism is utilized to either improve biological activity or to decrease the adverse side effects and toxicity of a compound [11].

At some point, the basic steps towards drug design should use either structure-based or ligand-based drug design concepts to create a chemical database in order to define the desired scaffold or skeleton. The lead compounds (LC's) will undergo additional optimization to find the hit compounds, followed by the development of a route to synthesize a high consensus score chemical scaffold with the designed functional groups, which is then used for high throughput screening and biological activity evaluation.

In conclusion, drug discovery is a long-term process used to find the lead compounds, which will go through different clinical tests and biological evaluation,s as shown in Figure 1.2.

1.2 Natural Products and Drug Discovery:

Natural products, biologicals, total synthetic or vaccines are the main sources of small organic molecule drugs [12]. Since ancient times, natural products have played an important role as a resource for medicine; for example, herbs can be chewed to relieve pain

or wrapped around wounds to heal them. Using natural products to treat diseases and injuries is now known as folk medicine [13, 14]. Lately, significant advances in biological screening techniques have allowed the study of the biological mechanisms and chemical profiles of living systems, which has encouraged researchers to investigate the pharmacological effects of natural compounds. Studies are done to clarify the natural products' synergistic impacts and their clinical effects on the individual body, which could help provide novel curative approaches to different diseases [15].
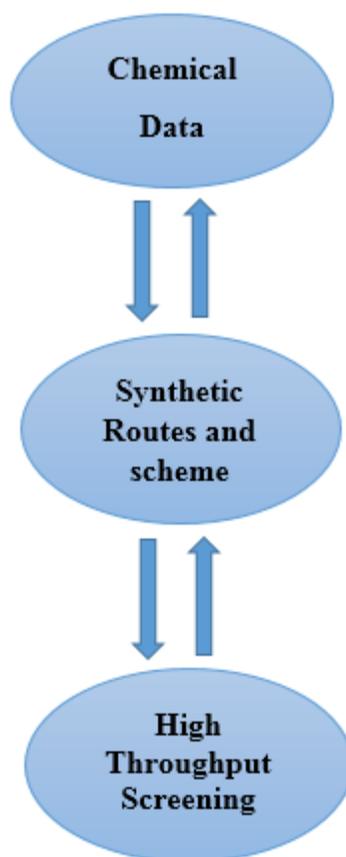
Figure 1.2 Process of drug design.

Chain et al, 2006 clarified the criteria for evaluating natural products as three steps: 1) the rate of display of novel compounds of wide structure diversity to be used as model and pharmacophore for systematic synthetic modification, total synthesis and semi-synthetic; 2) different diseases healed or prevented by these compounds; and 3) their number of use for treatment of diseases can be valued by pharmacoeconomic analysis adjusting the number of prescriptions and the performance of the drugs. Currently, 56% of the prescribed drugs in the United States are related to natural products [16].

A recent study ranked all drugs derived from natural products among the top 35 worldwide best ethical drugs sellers [17]. Additoinally, Newman et al, 2012 proved the advantages and biological efficacy of the natural products approved by the Food and Drug Administration (FDA) between 1981 and 2010 (Chart 1.1) [18]. Natural products have been used as platform for drug discovery for many available drugs, such as anti-epileptics, anti-angina, anti-obesity, anticancer, anti-viral, anti-migraine and bronchodilators.

Newman et al, 2012 additionally attracted attention to the significance of different types of synthetic approaches utilizing natural products, such as total synthesis of a natural product and semi-synthesis of compounds containing functional groups (parmacophore) imitating natural products. This clarifies the role of nature in inspiring future organic chemists to prepare novel organic compounds that can provide better or similar activity as the natural compound but with fewer side effects [18].

Biological	B
Natural Product	N
Semi-synthetic modification to Natural Product	ND
Totally Synthetic	S
Totally Synthetic but with natural product pharmacophore	S*
Natural Mimic	NM
Vaccines	V
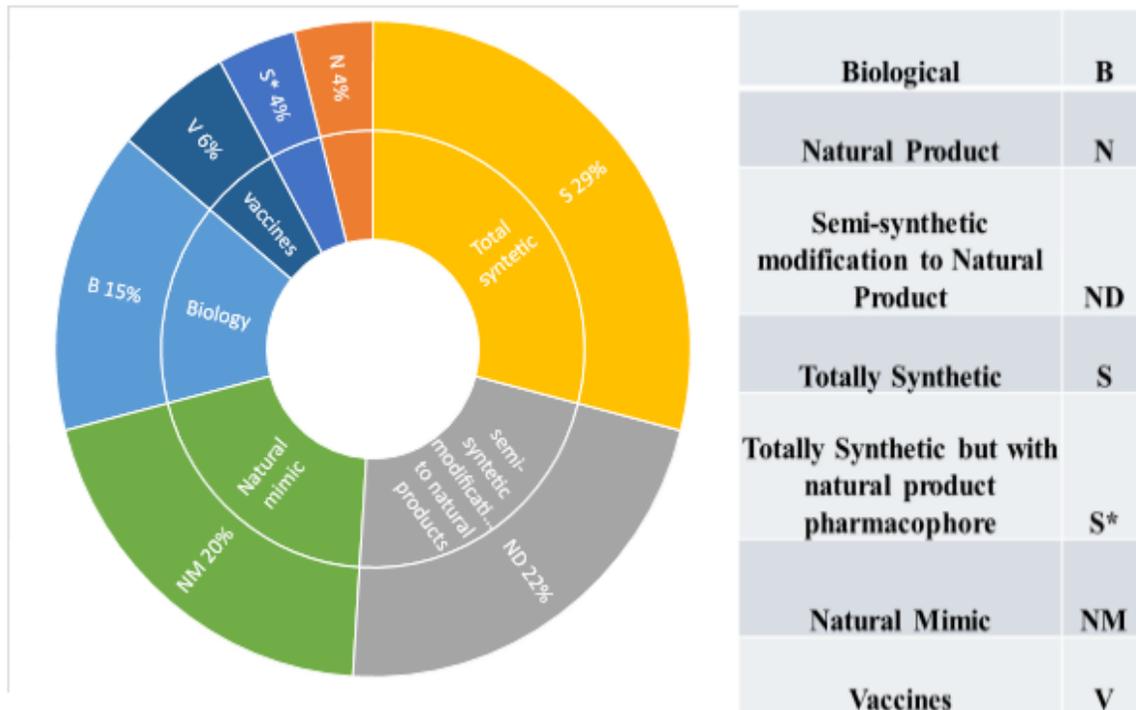
Chart 1.1 All New Chemical Entities between 1981 -2012 which approved by FDA.

One of the better known examples of a natural product that has been used as a model for a total or semi-synthetic drug is Taxol. Another example used as model for a partially synthetic approach is cucurbitacins, a natural product which shows anti-cancer, hepatoprotictive and anti-inflammatory activities.

1.3 Cucurbitacins:

Cucurbitacins are natural compounds extracted mostly from the plant of *cucurbitaceae* family such as *Ecballium Elaterium*, *Genystlus keithii, Cayaponia Tayuya, Citrillus Colocynthic, Trichosanthes Kirilowii* and *Ecballium Elaterium*. In addition, they

can be found in other plant families, such as *Scraphulaiaceae and Cruciferae* [19]. The *Cucurbitaceae* family was first used in folk medicine because of its biological significance as an anti-inflammatory. According to various studies, the cucurbitacins plants were used for the first time in Asia, the Middle East and other parts of the world as herbal remedies. According to various studies, the cucurbitacins plants were used for the first time in Asia, the Middle East and other parts of the world as herbal remedies. The *Cucurbita* plants were used in ancient herbal medicine to treat chronic hepatitis, liver cirrhosis, jaundice, dyspepsia and cancer [19-23].

The first isolation of cucurbitacins happened in 1831 usign *Ecballium Elaterium,* but this did not show up in the market as a biologically active product until the 1960s. There are twelve different types of cucurbitacins, as follows: A, B, C, D, and E, and so forth to T. Additionally, hundreds of cucurbitacin derivatives have been synthesized and identified [22].

Cucurbitacins can be isolated from different parts of the plant, including the seed, roots, rhizomes and aerial parts of *Cantaloupe, Watermelon, Pumpkin, Honeydew Melon, Spaghetti Squash* and *Crenshaw Melon* [23]. Various reports have proven the medicinal activity of the cucurbitacins and their clinical effects. Some of the cucrbitacins are still currently used as a treatment in some parts such as Asia, Africa and South America. Many biological impacts have been connected with cucurbitacins and their glycosylic derivatives [23-26]. Recently, cucurbitacins B,C, Q, and E showed antiproliferctive activity on

different cancer cell lines, such as HepG2, MDA-MB-468, MCF-7 and A549. Animal models and their apoptosis stimulating activity occurred through the inhibition of the Janus kinase (JAK), signaling marker and activator transcription3 (STAT3) signaling [19, 24-26].

Cucurbitacins R and 23, 24-dihydrocucurbitacin B have been isolated from the roots of *Cayaponia Tayuya* by Recio et al [27]. These two cucurbitacins showed several useful activities, such as anti-arthritic, anti-allergic and anti-inflammatory activity, both in vitro and in vivo. This was due to their capabilities to block the expression of tumor necrosis factor-Alpha (TFN-α) in macrophages and in lymphocytes, as well as their involvement with the action of the nuclear factors [22, 27]. The activity of cucurbitacins as hepatoprotictive and anti-proliferative agents has been proved by Bartalis and Halaweish on two different cell lines, HepG2 (hepatocellular carcinoma cell line) and HSC-T6 (Immortal hepatic stellate cell lines) [28].

Kintak et al. 2010 showed that cucurbitacin B has an inhibitory activity on ERK/RAF/STAT3/MEK and RAS signaling pathway [29]. Subsequently, there is a huge need to synthesize cucurbitacins because of their outstanding biological results, and also their very low concentrations in natural sources. Several trials have been conducted trying to synthesize cucurbitacins, but all of them failed to provide a final compound [30-32]. The varieties of cucurbitacins functional groups or pharmacophores, as well as cucurbitacins'

main structural complexity, are the main reasons that totally synthesizing cucurbitacins has been difficult to achieve.

1.3.1 Physical Characteristics of Cucurbitacins:

Cucurbitacins are crystalline materials at room temperature which typically absorb ultraviolet light at the range between 228 and 234 nm. Petroleum ether, benzene, chloroform, ethel acetate, ethanol and methanol are the common solvents used to dissolve cucurbitacins, since they are insoluble in ether and poorly soluble in water [33]. All of these properties are reasonable because of the hydrophobic properties of the chemical structure of cucurbitacins [26]. Mdavi et al. proved that both cucurbitacin B and I are soluble and can be delivered in vitro and in vivo utilizing two different polymeric micelles, developed [poly (ethyleneoxide)-block-poly (Ecaprolactone) (PEO-b-PCL). By using a polyemiric micellar, the in vitro anti-cancer and STAT3 antagonist activity in the melanoma cell line (B16.F10) demonstrated a similar performance to the free drug using the same cell line. A mouse melanoma tumors study in vivo demonstrated similar anticancer activity against the B16.F10 cancer cell line using intratumoral administration. This yielded an insignificant drug level in the animal plasma while preserving high drug levels in the tumor. Some research has indicated potential ways to deliver cucurbitacins, such as polymeric micelles [34].

1.3.2   Chemistry of Cucurbitacins:

Cucurbitacins are tetracyclic triterpenes characterized by highly oxygenated functional groups. They possess 30 carbon atoms on their general structure, known as 19-(10 ➔ 9-β)-abeo-5 α-lanostane (9-β-methyl-19-nor-lanosta-5-ene). Cucurbitacins have more than 100 different structures including glycosyl cucurbitacins and hexanol cucurbitacins (**Fig.1.3**) [19, 22, 23, 35].

Cucurbitacins E and B are the most common in the cucubitacea plant family and their biosynthetic pathways have been documented in several studies. Since cucurbitacins are found in common plants, some of them are originated via enzymatic reactions during the plant's natural processes such as growth. For example, cucurbitacins A,D, C, F, G and H can be found via biotransformation of cucurbitanic B, whereas cucurbitacins I, K, J and L can be formed via biotransformation of cucrbitacin E [22]. However, due to the complexity of the stereochemistry of these natural analogues, organic synthesis has not yet been established.

The cucurbitacins four-ring system resembles the structure of a steroid ring [36]. Two common features mark the difference between cucurbitacins and steroids structures, with cucurbitacins having a gem-dimethyl group at C-4 and a C-10 methyl in C-9. Cucurbitacins' main structures share common features, such as: 1) the double bond between C-6 and C-5; 2) a high level of oxidation due to the substitutions of many carbons (C-16, C-11, C-3, C-2) by oxygen atoms; 3) the presence of hydroxyl groups, α at C-16 and β at C-20 and C-25; and 4) α-β-unsaturated ketone in the side chain located at C-22, C-23 and

C-24. As shown in Figure 1.3, an aromatic ring is a common feature in some of the cucurbitacin derivatives, such as Fevicordin A [37, 38]. Additionally, cucurbitacins can be found as free glycone structures or glycosidic structures by a β-linkage to the hydroxyl moiety from monoside at C-2, C-3 and C-25 or from bidesmosides at C-26 or C-27 [39].

Different studies have demonstrated that the biological effects of cucurbitacins are due to the presence of a triterpenoid structure which resembles steroids. Additionally, multiple studies on different cell lines have investigated the structure activity relationship (SAR) of cucurbitacins. First, the presence of α-β-unsaturated ketone of the side chain at C-22 and C-24 is very significant for cucurbitacin's biological activities; cucurbitacin B showed more activity than dihyedro-cucurbitacin B and cucurbitacin D showed more activity than that of dihydero-cucurbitacin D in U-937 cell lines which are histiocytic lymphoma cell lines in different intervals by about 1000 fold. Second, the toxicity of the glycosides of cucurbitacin E and I is less than that of aglecones, while the cucurbitacin biological activity increased by 1000 fold in the presence of 25-acetoxy moiety. For example, the IC50 of cucurbitacin B is less than that of cucurbitacin D in U-937 histiocytic lymphoma cell lines at different test intervals [39].

1.4 Cucurbitacins as Potential Drug Candidates Targeting Different Molecular Targets:

1.4.1 Cucurbitacins Activity as Anti-inflammatory Compounds:

Cucurbitacins have been proven to be anti-inflammatory compounds via targeting many biological targets to decrease the inflammation. The anti-inflammatory activity of cucurbitacins is due to the inhibition of cyclooxygenase (COX) enzymes, particularly cyclooxygenase 2 (COX-2). In comparing the non-steroidal anti-inflammatory drugs (NSAIDs) and cucurbitacins, the inhibitory rate for COX-2 by cucurbitacins is less than the NSAIDs; for example, at 100mM concentrated cucurbitacins B, D, E and I showed inhibitory rates of 32, 29, 35 and 27%, respectively, compared to the COX-2 of NSAIDs such as Refeoxib, Ibuprofen and Naproxen. This study also proved the selectivity of cucurbitacins towards COX-2, since they do not show any activity on the cyclooxygenase-1 (COX-1) enzyme [27].

Additionally, other studies detected that 23, 24-dihydrocucurbitacin B particularly have a significant inhibitory activity against COX-2 without reducing its expression. Also, cucurbitacin R demonstrated noticeable inhibition in the proteins and mRNA for inducible NO synthase (iNOS) by preventing the activation of nuclear factor-KB (NF-KB) [27].

Cucurbitacins isolated from the roots of C. tayuya such as cucurbitacin R and 23, 24-dihydrocucurbitacin B displayed anti-arthritic, anti-allergic and anti-inflammatory activities in both in vivo and in vitro tests due to their efficacy to suppress tumor necrosis

factor (TNF-α) expression in macrophages and lymphocytes. The in vivo studies demonstrated the capability of both cucurbitacins to minimize the expression of nitric-oxide cyclooxygenase-2 and synthease-2 (pro-inflammatory enzymes) [36].

Cucurbitacin

Steroidal Scofold

Cucurbitacin	R_1	R_2	R_3	R_4	R_5	$\Delta^{1,2}$	$\Delta^{23,24}$
A	OH	=O	MeOH	H	AC	-	+
B	OH	=O	CH₃	H	AC	-	+
C	H	OH	MeOH	H	AC	-	+
D	OH	=O	CH₃	H	OH	-	+
E	OH	=O	CH₃	H	AC	+	+
F	OH	OH	CH₃	H	OH	-	+
H	OH	=O	CH₃	OH	OH	-	-
I	OH	=O	CH₃	H	OH	+	+
J	OH	=O	CH₃	OH	OH	+	-

Figure 1.3 General structures of cucurbitacin with various functional groups and steroid.

1.4.2   Cucurbitacins Effect on Filamentous-Actin:

Some studies demonstrated the activity of cucurbitacins E, I, and B on the cytoskeleton, particularly on F-actin. Another study proved that some cucurbitacins derivatives including cucurbitacin E, stimulate actin cytoskeleton disturbance [40]. This disturbance connected with the effect of cucurbitacins on the actin-proliferative action in prostate cancer cell lins. One of the astonishing attributes of cucurbitacin E is that it has a selective inhibition on F-actin depolymerization, but not on monomeric globular G-actin, by forming a covalent bond with CYS 257 amino acid residue [41].

Recent studies on cucurbitacins D, E, and I were conducted utilizing breast cancer cell line (MCF-7) to detect their activities on actin-based structures and cell viability suggested that cytotoxic characteristics of cucurbitacins might be largely separated from their effects on cytoskeleton rearrangement [42]. In conclusion, cucurbitacins have demonstrated strong effects on the cell cytoskeleton by essentially simulating actin through the stimulation of depolymerization and aggregation.

1.4.3   Mitogen Activated Protein Kinase (MAPK) Pathway Activation by Cucurbitacins:

One of the key parts in the cellular process is the MAPK pathway, which has signaling transducing cascade including Ras/B-Raf/MEK/ERK **(Fig.1.6)**. Very few research studies have addressed the potential biological activities of cucurbitacins targeting the MAPK pathway. Chean et al. stated that cucurbitacin B inhibits the STAT3 and

RAS/B-Raf/MEK/ERK cell downstream signaling cascade using the K562 leukemia cell line [29]. Utilizing computational semi-flexible molecular docking, MTT cell viability assay and binding immune assay, Salama and Halaweish demonstrated the ability of several types of cucrbitacins to target MAPK signaling pathway using mutant B-Raf cell lines [43].

1.4.4 Cucurbitacins as Potential Modulator for Epidermal Growth Factor Receptors (EGFR):

EGFR, in human cancer has been demonstrated to be involved in the mutation and deletion of the cell upstream and downstream targets, which makes EGFR a promising biological target for different types of human cancer [44]. Hollbro et al proved that ErbB receptors, a member of the GFR family, are very important targets for the treatment of different kinds of cancer [45]. Liovet et al. showed that activation of EGFR and its downstream cascade has increased the survival rate of patients in the late stages of hepatocellular carcinoma [46].

In conclusion, cucurbitacins have demonstrated wide and broad biological activities toward various molecular targets, including anti-inflammatory agents, hepatoprotective, anti-cancer agents and anti-virus agents. This makes cucurbitacins promising drug candidates for various diseases and health problems. Additionally, cucurbitacins have been proven to have effective inhibitory activity against various cancer cell lines, such as hepatocellular carcinoma cell line (HepG2), colon cancer cell line (HCT-116), lung cancer

cell lines (NCI-H460 and A-549), leukemia cell line (U-937), breast cancer cell line (MCF-7) and melanoma cell line (B16/F10).

Understanding the genetic changes in HCC and its molecular targets is essential for designing potential anti-cancer drugs.

1.5   Hepatocellular Carcinoma (HCC):

Hepatocellular carcinoma (HCC) is classified as one of the highest cause of mortality and the third causing of death worldwide [47]. Asia and Africa have the highest reports of HCC, with high incidence among men, while HCC is somewhat rare in the United States compare to the other countries. HCC is a continuous and slowly progressing disease that is generally associated with other factors, such as cirrhosis, hepatitis C virus (HCV), hepatitis B virus (HBV) and toxin/ environmental disorders (obesity, diabetes and alcoholic consumption) [48].

1.5.1   Etiology of HCC:

HCC influences all groups of people worldwide. However, regional differences and ethnicity are also a significant influence on the prevalence and etiology of HCC incidence [49, 50]. HCC is a continuous and slowly progressing disease that is generally associated with other factors such as cirrhosis, hepatitis C virus (HCV), hepatitis B virus (HBV) and toxin/ environmental disorders (obesity, diabetes and alcoholic consumption) [43, 49, 51].

An additional factor that can lead to HCC is gender, with males having a higher prevalence for HCC than females [49].

1.5.1.1 Hepatitis B Virus (HBV):

HBV is one of the Hepadnaviridae family, which consider as covered DNA virus. It is a very common viral disease in that 2 billion people are HBV carriers and about 320,000 people die annually due to complications of HBV infection. Asia and Africa have the highest reports of HBV incidence worldwide[52]. This disease is very contagious, so it can infect other individuals in several ways— prenatally, percutaneouslly, and sexually [52, 53]. Reverse transcriptase (RT) enzyme is responsible for the virus genome replication [49, 53], but other proteins are also involved in the progression of HBV, such as capside proteins, common as hepatitis B core antigen (HBcAg), and other enveloped proteins connected with endoplasmic reticulum (ER) (HBV-1). Approximately one-third to one-half of HBV-associated deaths are due to the progression of HCC [49, 53]. The influence of the HBV infection on the progression of HCC is connected with the high incidence of HCC in the infected individual, with a high level of HBV DNA in their biological serum [54].

Various studies have proven the ability of HBV to transform into HCC. First, the genome integration of HBV is related to the host DNA microdeletion, which has the ability to target cancer-related proteins including platelet-derived growth factor receptor-B

(PDGFR-B), telomerase reverse transcriptase (TERT), and mitogen activated protein kinase-I (MAPK-I) [55, 56]. Second, an alteration in the expression of growth-control proteins such as Ras, Raf, MEK, ERK, JNK and tyrosine kinases takes place through activation of HBX transcription activity. Third, card inactivate and bind in invitro to the tumor oppressor P53, which accelerates the cellular proliferation and survival [49, 57, 58]. Genetic validation of HBx as a hepatocarcinoginic has been proven in transgenic mice, resulting in 90% transformation to HCC [59].

1.5.1.2 Hepatitis C Virus (HVC):

HCV is a member of Flaviviridee family that is a single-stranded RNA non-cytoplasmic type virus. It is a contagious virus that can infect other individuals by direct blood contact, perinatal from mother to fetus, and in rare cases, by sexual intercourse. China, Egypt and Pakistan have the highest number of HCV cases worldwide [49, 60, 61]. Approximately 170 million individuals are HCV infected worldwide [61]. Six HCV genotypes (1-6) have been isolated, with genotypes 1-3 being very common worldwide. Genotype 4 is common in Egypt and the Middle East, genotype 5 is prevalent in South Africa and genotype 6 is widespread in China [61, 62]. HCV has been shown to have biological uniqueness compared to HBV, which is an associated hepatocarcinogenesis. First, HCV has a better tendency (60-80%) to produce chronic infections compared to HBV, which has only a 10% tendency. That tendency is connected to HCV's ability to generate a fast rate of replication errors that can cause immune evasion [49, 63, 64]. Second, HCV has a 10-20 times higher tendency to initiate liver cirrhosis compared to

HBV, with 10% of the patients infected by HCV incubating liver cirrhosis after 10 years of contagion. The high connection of HCV with cirrhosis is a critical correlation in developing HCC. Finally, since HCV is an RNA virus, it cannot combine into a host genome [63].

1.5.1.3 Alcohol Consumption:

Alcohol consumption is one of the significant causal factors for HCC. Chronic alcohol consumption causes pre-inflammatory cytokines by activating and disturbing the monocyte which leads to an abnormal evaluation circulating endotoxin concentration. This can cause the activation of Kupffer cells which emit various cytokines and chemokines, such as prostaglandin E_2, TNFα, 1L6 and interleukin-1B (IL1B). These bring about hepatocyte damage [65]. In the case of chronic ethanol toxicity, the hepatocyte demonstrates a high sensitivity to the TNFα cytotoxicity effect, which is a clear sign of chronic hepatocyte disturbance, activation of stellate cell, liver cirrhosis, and eventually, HCC [66].

In addition, alcohol can destroy the liver via a process called oxidative stress. An increase in the isoprstane level is a sign of alcohol hepatitis, which indicates lipid peroxidation [67]. Three explanations can be given for the connection between oxidative stress and the liver damage that leads to HCC. First, oxidative stress elevates the progression of cirrhosis and fibrosis, which are considered the main causes of HCC [68]. Second, oxidative stress resulting from ethanol toxicity shows a relevant effect on the HCC-signaling cascade, such as a decrease in the tyrosine phosphorylation of the signaling

transducer and the activator of transcription I (STATI) [69]. Third, oxidative stress might increase the rate of telomere shortening which may interrupt the DNA replication process and cause HCC [70].

1.5.1.4 liver Cirrhosis:

Liver cirrhosis consider as the main reason for HCC development, which mainly caused by the most two common livers viral infection, HBV and HCV [71, 72]. Table 1.1 shows the annual percentage of individual who develop HCC from cirrhosis caused by liver viruses or liver complications [71].

While these results show HCC development via liver cirrhosis caused by different liver diseases, the development of HCC may also include additional factors [73]. The development of cirrhosis usually comes when patients have chronic liver disease over a period of years, which can be distinguished by a decrease in hepatocyte proliferation as a sign of liver damage. This causes an increase in the fibrous tissue and a disturbance of liver cells that leads to the development of liver cancer [71, 74]. Various studies have been conducted to determine the exact mechanism associated with HCC development from liver cirrhosis, but only a few possible mechanisms have been proposed, including micro- and macro-environmental changes that induce cellular proliferation and telomese dysfunction [75].

Table 1.1 Annual percentage of HCC incidences caused by liver cirrhosis.

Disease	Annual percentage %
HBV	1-15
HCV	1-8
Liver disease due to alcohol consumption	1
Liver disease due to non-alcoholic reasons	2.6

1.5.1.5 Aflatoxin B$_1$ and Environmental Factors:

Aflatoxin B1 (AFB1) is a mycotoxin found in Aspergillus fungus that infects different foods, such as peanuts and corn. AFB1 has been proven as carcinogenic based on several animal studies [75, 76]. When AFB1 is absorbed in the blood circulation, it is activated by metabolism to AFB1-exo-8, 9-epoxide. The active metabolite binds to DNA and damages it, so it can be distinguished by the activation of the P53 protein [77]. This DNA damage has been distinguished in 30%- 60% of HCC patients in AFB1 epidemic regions [78]. AFB1 as a risk factor for HCC development was supported by several studies done between 1992-2007. These studies were conducted on an aflatoxin metabolite developed from urine assays in order to find signs of aflatoxin-damaged DNA [75].

Environmental disorders such as obesity, diabetes mellitus and tobacco have also been studied and proven as risk factors for HCC development [75]. A study on 900,000 people with obesity over a period of 16 years found that the incidence of hepatocarcinoma death was five times higher than that of non-obese people [79]. Diabetes mellitus,

specifically type-2, is another factor for HCC development via non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepaties (NASH). The connection between type-2 diabetes mellitus and HCC has been studied in several countries, with several studies showing a significant connection between HCC and type-2 diabetes [75]. Another significant factor involved in HCC development is cigarette smoking. Numerous studies on HCC patients demonstrated positive and negative associations between tobacco and HCC [80].

1.5.2 Diagnoses and Treatments of HCC:

1.5.2.1 Diagnosis of HCC:

HCC can be diagnosed by various tests, including imaging, histology and serological tests. Imaging tests, such as ultrasound (US), computerized tomography (CT) and magnetic resonance imaging (MRI) can be used to detect hepatic nodules. The molecular biomarker test for the diagnosis of HCC is now becoming popular. With this test, the three types of alfa-fetoprotein (AFP-1L1, AFP-L2 and AFP-L3) are the most popular molecular biomarkers for the diagnosis of HCC [47, 81, 82]. The biomedical problems associated with the AFP molecular biomarker are that it can be found in high levels in some non-malignant cells such as with pregnancy and severe liver diseases. In addition, it can be found in various levels in different types of HCC [81, 83]. Figure 1.4 demonstrates the diagnostic strategy of the American Association for the Study of Liver Diseases (AASLD).

1.5.2.2 Treatments of HCC:

Various treatment choices are available for HCC, with the type of treatment depending on factors such as the condition of liver function, the HCC level, the type of local medical resources available and the prevalence of other chronic diseases [47]. The treatment of HCC can be divided into two types— curative and palliative. The curative options for HCC treatment include ablation, surgical resection and liver transplantation; these usually provide a high percentage of treatment response which increases the survival rate. On the other hand, the palliative options of HCC treatment, such as chemotherapy and radiotherapy, do not tend to provide a high rate of response compared to the curative options, but they can improve the survival rate of HCC patients in general [84].
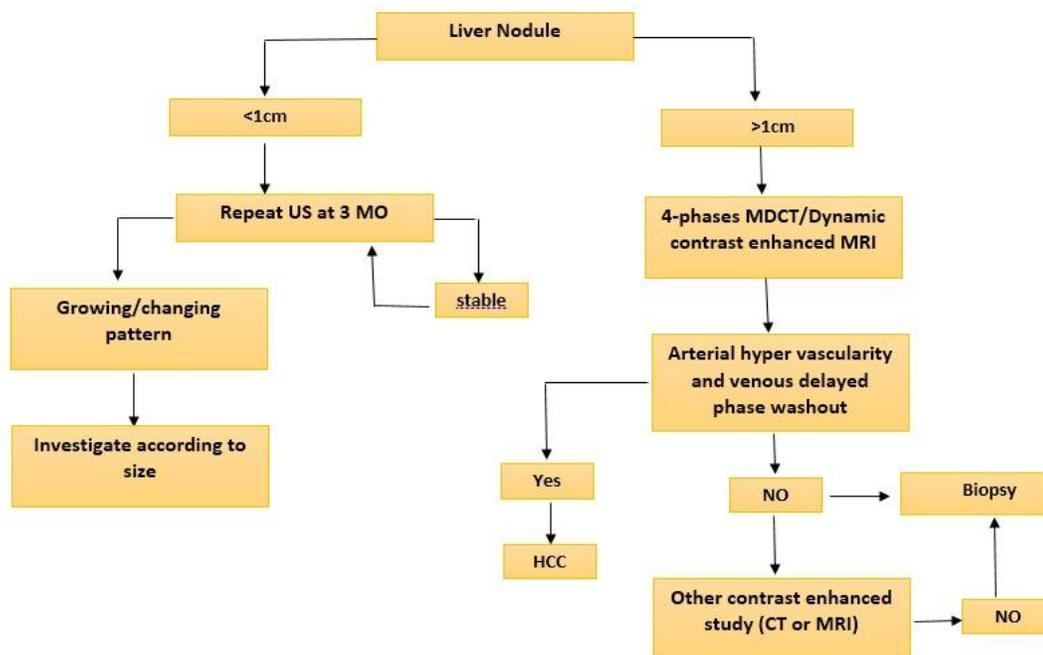
Figure 1.4 Diagnosis strategy of the American Association for the study of liver diseases (AASLD) (reprinted from [81]).

1.5.2.2.1 Surgical Resection:

HCC patients with a non-cirrhotic liver, an early stage of HCC and good liver function are good candidates for the resection surgery. The process of liver resection consists of removing the specific part of the liver that has the tumor mass, along with a small range of liver tissues around the mass, leaving the healthy part of the liver to renew the whole organ. Among all the HCC treatments choices, surgical resection for the early stage of HCC is considered the best choice because it provides complete extirpation of the tumor mass and allows for liver function regeneration [81]. However, surgical resection has some limitations for curing HCC. First, some clinical tests should be made on the liver to ensure that the remaining part of the liver has the ability to renew the liver function. For example, HCC patients with liver cirrhosis are not suitable for surgical resection. Second, removing the liver tumor mass will not remove the tumor completely, which will increase the chances of generating a *de novo* primary tumor mass. The statistical studies indicate that 75-80% of HCC patients who have had the tumor mass removed will survive [85]. Third, the chance of treating HCC by surgical resection will be small if there are several tumor masses in the liver. In addition, HCC diagnosed at the late stages with the association of liver cirrhosis will make surgical resection an impractical option [86].

1.5.2.2.2 Liver Transplantation:

Liver transplantation is considered one of the best curative treatment options for HCC patients, especially in the early stages, since it removes the whole liver including its tumor masses. It thus provides a solution for the underlying cirrhosis [87]. The main

problem with liver transplantation is the spread of the tumor to another organ. Mazzafero et al, 1996 developed the Milan Criteria, which are criteria used to distinguish HCC patients who are suitable for liver transplantation. These criteria include: a solitary tumor < 5cm in diameter, total of 3 lesions <3cm in diameter, and no spread of the tumor to the other organs or vessels. The Milan Criteria increased the survival rate to 80% and returning survival rate to 83% after liver transplantation [81, 87]. The limited number of liver donors is the major issue with this process [47].

1.5.2.2.3 Ablation Therapy:

Ablation therapy provides local management of the HCC cancer cells with only a small influence on the neighboring cells and other hepatic tissues. Ablation therapy is a convenient technique for HCC patients who are still in the early stage and do not have a match for liver resection or liver transplantation. The three types of ablation therapy include radiofrequency ablation (RFA), percutaneous ethanol injection (PEI) and microwave ablation (MWA). RFA uses an electrical current with a high frequency to deliver heat to the liver tissues and cause coagulative necrosis. PEI mainly causes liver cell dehydration, which is responsible for the death of the exposed liver tissues, also produces coagulative necrosis. Unlike RFA, MWA generates heat by applying an electromagnetic source to the liver tissues that can cause the tumor masses to die [47, 81].

1.5.2.2.4 Chemotherapy:

The name chemotherapy refers to a treatment using drugs or medicines. Treatment of HCC with chemotherapy agents has not been very promising to date, since they have

not increased the survival rate. The major reason for this is multidrug resistance associated with most of the drugs available [88]. Chemotherapy treatment is suitable for the patients who are in advanced stages of HCC and do not fulfill the criteria for surgical resection or liver transplantation [81]. Several goals are possible when using chemotherapy to treat HCC patients, including curing the cancer, slowing cancer growth, and treating cancer symptoms [89]. In general, four anti-cancer drugs types are available, including alkylating agents, antineoplastic agents, intercalating agents and molecular target anti-cancer agents. However, the Food and Drug Administration has approved only a few drugs for the treatment of HCC (Sorefenib and Erlotinib), which are considered molecular target anti-cancer agents. [90].

1.5.2.2.5 Radiotherapy:

Radiotherapy is another option for the treatment of HCC. This therapy can achieve necrosis by killing the tumor cells in a small liver tumor mass. However, radiotherapy has some risks for the patients, such as causing abdominal injuries or extensive hepatitis. For these reasons, the use of this technique for treating HCC is very limited [88].

1.5.2.3 Resistance of the Chemotherapeutic Agents:

The main available treatments for cancer are chemotherapeutic agents which can either stop or slow the abnormal fast division of the cells. The major problem associated with the available chemotherapeutic agents is drug resistance, which occurs with 30%-80% of cancer patients. Cancer cells such as HCC cells produces drug resistance via three

mechanisms: multidrug resistance (MDR), P-glycoprotein drug resistance (P-gp) and multidrug resistance protein (MRP). All of them are activated and effective in the HCC treatment process [94, 95]. Therefore, there is an urgent need to find a new agent to overcome chemotherapeutic resistance.

1.5.3 Molecular Targets for Treatment of HCC:

HCC can be cured by either liver transplantation or surgical resection if discovered in its early stages. However, most of the HCC cases are discovered in the late stages, which decreases the general survival rate [91]. The HCC patients who are suitable for curative treatments represent approximately 15% of those diagnosed [92]. The essential reason for this disappointing survival rate is the lack of effective chemotherapeutic agents that can cure HCC in its late stages. A large number of studies have demonstrated that only 10%-20% of HCC patients respond to the chemotherapy treatments, with toxicity and cellular resistance to available chemotherapeutic agents being the main obstacle to successful treatment [93]. In conclusion, searching for new active molecular target anti-cancer agents for HCC has become a popular area for research due to the urgent need to overcome the toxicity and cellular resistance problems. In addition, identifying molecular targets is essential for the discovery of a therapeutic treatment that overcomes the resistance and decreases the side effects.

1.5.3.1 Epidermal Growth Factor:

Epidermal growth factor receptor (EGFR) tyrosine kinase (TK) is one of the tyrosine kinases that has been studied as a promising target for the treatment of different

carcinomas, including HCC [96]. EGFR, which is also known as ErbB1, is a member of a family of growth factor receptors that also includes ErbB2, ErbB3 and ErbB4 **(Fig. 1.5)**. Paracrine or juxtacrine extracellular ligand binding, such as epidermal growth factor (EGF), and transforming growth factor (TGF)-α stimulate the EGFR which leads to hetro- or homo-dimerization and conformational change that activates the tyrosine kinase and allows autophosphorylation. When the phosphorylation occurs, a number of signaling pathways activate, leading to cancer cell invasions, proliferation, metastasis, inhibitory of apoptosis and angiogenesis **(Fig. 1.5).**
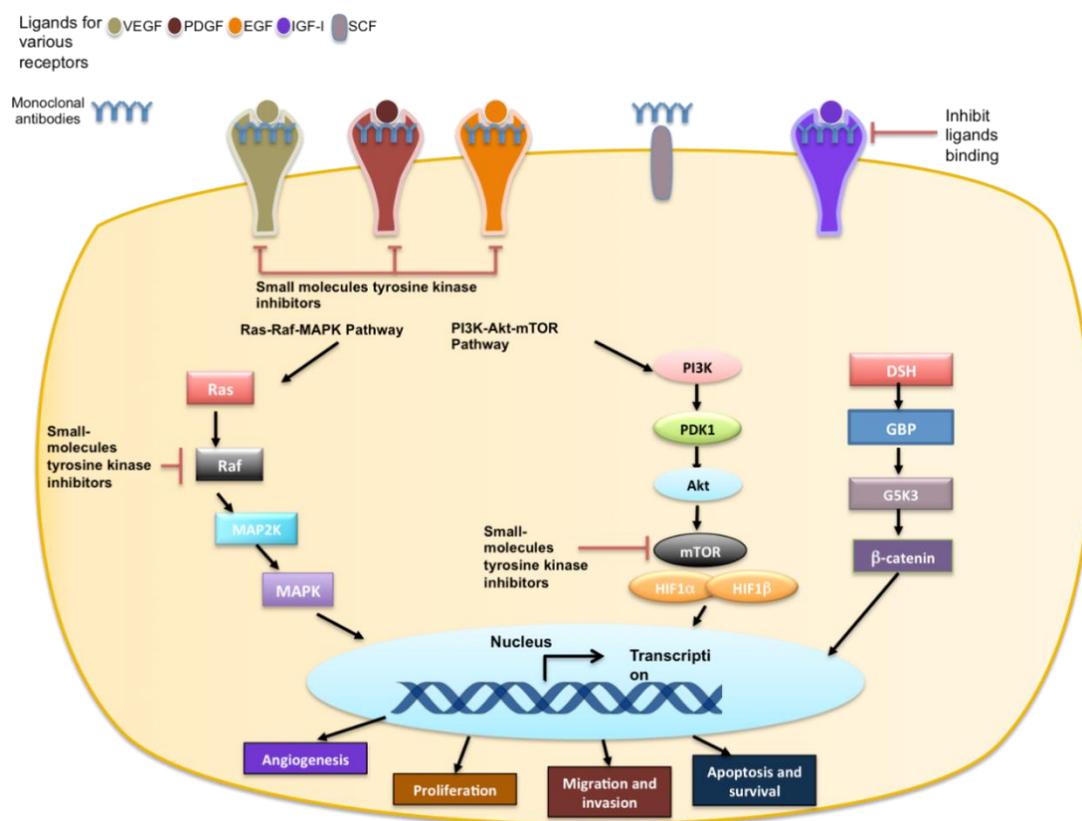
Figure 1.5 Potential molecular targets for treatment of HCC.

Therefore, inhibition of the EGFR-TK signaling cascades provides a potential approach for the treatment of hepatocellular carcinoma (HCC). In the signaling cascade, either upstream or downstream targets of EGFR can be used for the treatment of any cancer in general and HCC specifically [97-100]. Different organic compounds, such erlotinib and lapatinib **(Fig. 1.6),** have shown potent inhibitory activity against EGFR by inhibiting its phosphorylation; these are known chemotherapeutic agents for the treatment of HCC (**Fig 1.6**). Erlotinib (Tarceva) is a very active and selective inhibitor of the EGRF-TK protein. It has an advantage over most of the anti-cancer agents in that it can be taken orally, inhibits cellular proliferation and causes cell cycle arrest at the G1 phase. In addition, it has been approved as an active drug for pancreatic and lung cancer by the FDA but is still in Phase II clinical trials as an anti-cancer drug for HCC [101].

1.5.3.2 Vascular Endothelial Growth Factor (VEGF):

HCC is considered a hypervascular cancer, which has a large amount of tumor vascularity. VEGF is connected to the angiogenesis of various cancer types and is a significant part of HCC angiogenesis [102], enhancing the progression of HCC and its resistance [97]. HCC and all other tumor masses need blood vessels to survive and enlarge, so these blood vessels are considered abnormal since they are responsible for the high fluid pressure inside the tumor mass. Therefore, VEGF targeting agents may cause a decrease in the tumor vessels' supplies and their sinuosity, which leads to a decrease in the internal pressure of the tumor mass. All of these processes will lead to normal blood vessels [103]. Gerber et al, 2005 proved the ability of anti-VEGF drugs, in combination with other anti-

cancer agents, to cause a fast decrease in the internal vessel pressure of the tumor mass, which resulted into faster targeting of the agents to the tumor mass, a reduction in the tumor size and an increase in the survival rate [104]. Different agents have been designed to treat HCC by targeting VEGF or VEGFR. Some of them have been proven effective, such as Sorafenib **(Fig. 1.6),** and some are still in clinical trials to verify their pharmacokinetic profiles. Sorafenib is the first molecular target drug approved by the Food and Drug Administration (FDA) for the treatment of HCC. It is considered a multi-kinase agent that stops tumor cell proliferation by inhibiting different molecular targets, including VEGFR and PDGFR tyrosin kinases, which produces an anti-angiogenic effect. In addition, it targets the downstream cascades such as the Raf/MEK/ERK signaling pathway **(Fig. 1.6)** [91, 97, 105].
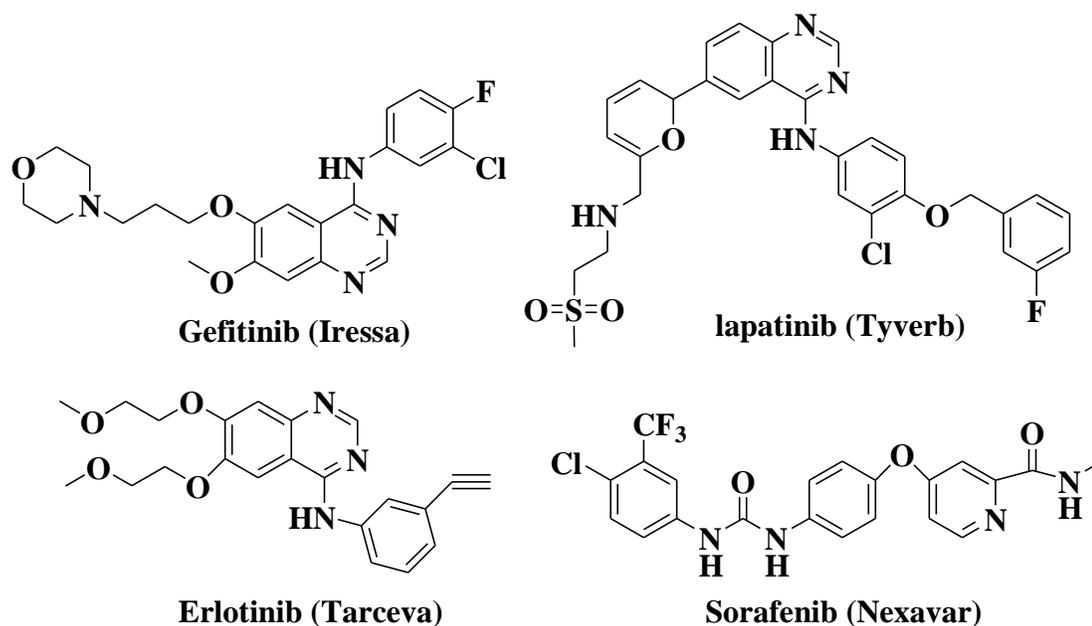
Figure 1.6 Structures of known Chemotherapeutics that target EGFR for treatment of HCC.

1.5.3.3 Mitogen Activated Protein Kinase (MAPK) Pathway:

MAPK includes, in its downstream, four main kinases— Ras, Raf, MEK and ERK **(Fig. 1.5)**— which communicate with each other by phosphorylation. They are responsible for cell division, growth and regulation. These downstream proteins are connected to the upstream receptors such as PDEFR, EGFR, and VEGFR **(Fig. 1.5)** [105, 106]. So the MAPK pathway is an essential player in the growth and survival of HCC cells, which makes it a promising target for the treatment of HCC [107, 108]. ABT-100 is one of the anti-cancer agents in phase II clinical trials that inhibits the farnsylation process of the protein Ras by inhibiting the enzyme farnesyl transferase, which leads to a decrease in tumor cell growth [105, 109]. The family of the protein Raf includes three members: A-Raf, B-Raf, C-Raf. Hyperactivation of C-Raf (wild type) in various cancer types, including HCC, was the only reported one, which makes it a valuable target for treating HCC [110]. Sorefineb is an approved HCC chemotherapeutic agent that inhibits B-Raf, C-Raf, FGFR, PDGFR and VEGFR [111]. The family of the protein MEK includes two subunits, MEK1 and MEK2. Huynh et al reported that overexpression of MEK1 and MEK2 lead to an activation of ERK1 and ERK2; in the case of HCC, this occurs in different percentages. Additionally, it has been proven in in vitro studies that the addition of a MEK1 or MEK2 inhibitor to HepG2 or Hep3B HCC cell lines will inhibit the autophosphorylation and cause cell apoptosis. MEK inhibitors, including include Selunetinib, RDEA119 and ASCO2010, are still in phase II clinical trials, [112].

1.5.3.4 PI3K/Akt/mTOR Pathway:

PI3K/Akt/mTOR pathway is a downstream signaling cascade that connected to the upstream receptors of the cell (**Fig. 1.5**). Cellular activation of the protein PI3K will automatically activate the Akt protein kinase and result in number of phosphorylation, which leads to cellular proliferation and apoptosis. In the case of HCC, this pathway will be superexpressed, making it a good target for the treatment. mTOR is a protein that is responsible for the regulation of the cellular translation process. This process includes various numbers of initiation factors, such as 40s ribosomal, protein S6 kinase and 4E-binding protein-1, which are involved in the mRNA translation process during cell proliferation and angiogenesis. All of these features make mTOR an interesting target for the treatment of HCC [113]. The immunosuppressive and antibiotic agent Rapamycin is known as an mTOR inhibitor and has demonstrated inhibition activity against HCC cell lines [105, 114].

In general, to find new agents that have the ability to selectively target the biological molecular targets, understanding the nature and behaviors of these molecular targets with the designed compounds is essential. One of the known techniques to assist with the design and to predict the simulation of the designed compounds inside the molecular targets is molecular modeling.
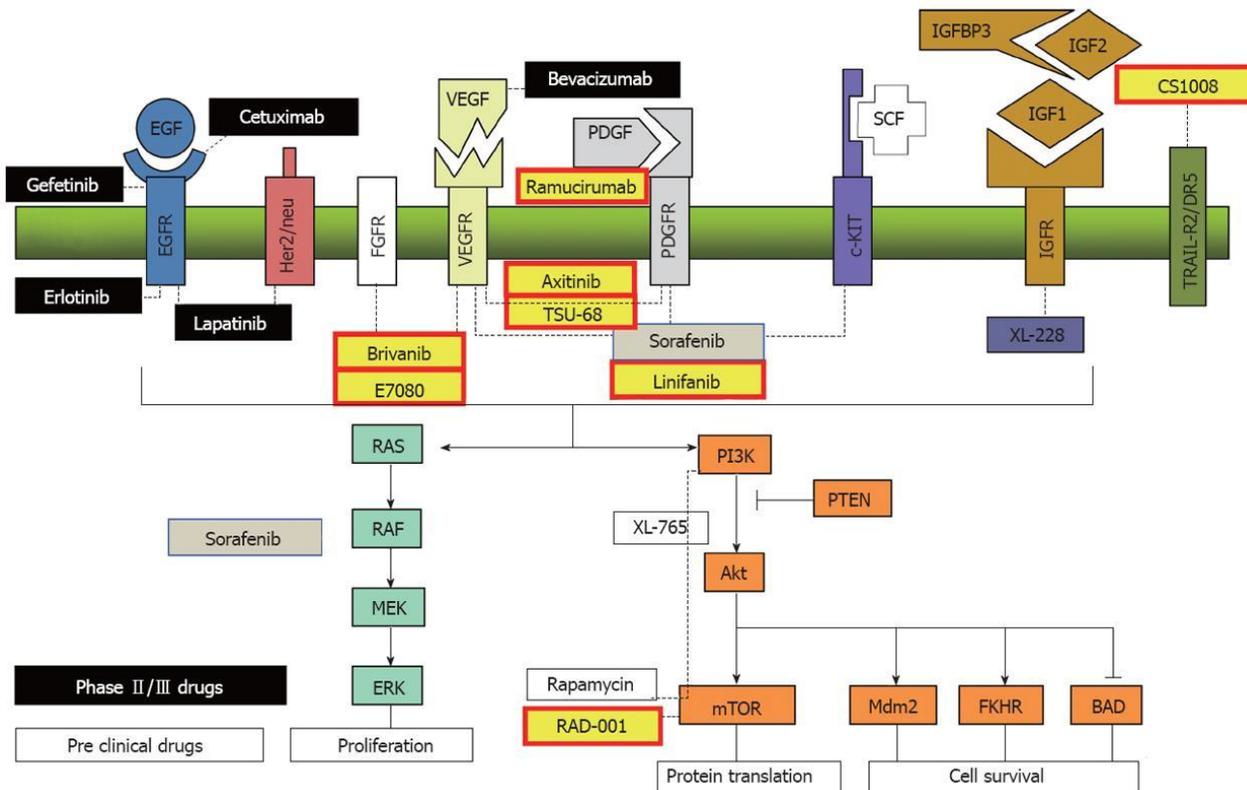
Figure 1.7 Approved and clinical trials agents for treatment of HCC. (Copied from [106]).

1.6   Molecular Modeling:

1.6.1   Introduction:

Research and development to discover a new drug in the pharmaceutical industries requires much time and a high cost, so new approaches are always needed to reduce the time and cost. Molecular modeling is one means to provide a fast prediction of the proposed compounds' behavior in the molecular biological target with a low cost [115]. Molecular modeling is used to calculate the action and characteristics of each atom and molecule of

the structure with a logical conclusion of the synthetic strategies and verification of the mechanistic approaches by predicting the patterns of transition states (TS). In addition, the molecular docking technique can be utilized to calculate the binding affinity between the compounds and the molecular biological targets, such as proteins, receptors and enzymes, in order to determine the potential drug candidates and hit compounds [115, 116]. Molecular modeling has a number of programs and applications that can predict the physical and chemical properties of a compound, such as absorption, distribution, metabolism and excretion (ADME) [117].

The drug discovery process has improved since the use of computational approaches began in the 1980s [116, 118]. Several computational methods have been involved in the process of drug discovery, such as structure-based drug design (SBDD) and ligand-based drug design (LBDD) [118]. SBDD utilizes the 3D structure and biological data of the protein of interest to determine the best compounds (usually small organic molecules) that bind to the selected protein. The compounds that show a high binding affinity to the targeted protein are called a "hit" which will be screened in a fast biological screening system, called high throughput screening (HTS), to find the lead compound (LC). On the other hand, LBDD uses known active compounds (known drugs) and then identifies their pharmacophores as the origin for the lead compound determination [119]. LBDD is considered a significant approach in the drug discovery process; however, it needs a large group of known ligands for which biological tests are handy. Thus, SBDD considered easier to utilize for the process of drug discovery [118].

Molecular docking is a process where small organic compounds are docked against a specific molecular active site of a protein (target). This is considered a very helpful process in the computational methodology of drug discovery. The main goal of the molecular docking process is to find the appropriate conformational structure that binds to the active site of the protein and score a high binding affinity. Therefore, the docking process is a powerful technique because it can provide a pattern of set compounds with specific pharmacophores based on their interaction with the targeted protein. This may lead to the discovery of promising novel candidates. The process of docking begins with the building of a database of molecules and a structure of targeted proteins in order to detect the compound with a high binding affinity to the protein. The main obstacle involved in this process is comparing the resulting binding affinity of the proposed molecules toward a specific protein with the binding affinity of known standards; this is called relative binding affinity. In general, molecular docking may take a long time to determine the hit molecules which can be optimized to a lead compound (LC). However, docking is a safe financial technique that protects the research in case of a failure to discover the lead compound; the research can then return to the database to find another hit molecule [120].

1.6.2   Molecular Docking Approaches:

Two main approaches have been developed to conduct molecular docking: 1) the simulation approach, which is based on a computer simulation to calculate the energy differences during the compound-protein docking process; 2) the shape complementarity

approach, which uses a specific application to estimate the compound-protein surface complementarity [121, 122].

1.6.3 Molecular Docking Types:

Molecular docking consists of three different types based on the docked molecules and the protein of interest. First, flexible molecular docking includes different conformers of the ligands, as well as the targeted proteins, to find the best ligand-protein conformers complex. This type relies on the induced-fit model [121, 123, 124]. The second type is semi-flexible molecular docking, in which different conformers of the compounds will be generated and docked on a rigid protein to find the best ligand conformer to create a stable complex with the protein. This type of molecular docking is the most popular type among all molecular docking types [121]. Third, in rigid molecular docking, both the compounds and the targeted proteins are rigid, so the concept of key and lock will be applied [121].

1.6.4 Limitation of Molecular Docking:

Some noticeable points in the molecular docking process, from both site ligand and protein, may cause limitations for the molecular docking process. These include conformational changes in both the ligands and targeted proteins that occur during the binding, the accuracy of the X-ray crystallography of the protein structure crystalized with the known ligand, entropic and enthalpic influence, the reality that the whole biological system can be represented by a single protein and the effect of metals and water molecules in the interaction between the ligand and protein [115, 119]. Multi-step approaches have

been used in the docking process to overcome all of these limitations. The steps begin with putting a small organic molecule (ligand) into the active site of a protein that mimics a favorable interaction with the protein. Then by evaluating the interaction between the ligand and the protein, their source of binding affinity can be determined; this is called scoring.

1.6.5   Overview of Docking Scoring:

The interaction between the protein and the ligand is the main output of the process and takes into consideration multiple molecular factors, such as H-bonds, hydrophobic interactions, overlay with the co-crystalized ligand, and types of amino acids involved in the interaction. During the docking, the protein can be represented at three different levels: atomic, surface and grid [125, 126]. In regard to the ranking steps of molecular docking, the atomic representation and potential energy function are relatively easy to determine due to the intricacy of computer calculations at the atomic interaction level. Protein-protein molecule docking can utilize surface-dependent molecular docking for their receptor representation. Various docking programs using potential energy grid functions are available to perform the energy calculation. Van der waals and electrostatic potential can be saved as grid points using grid-based docking [127, 128]. Ligand and protein flexibilities can be measured by utilizing algorithm search procedures. Three methods have been developed for measuring the ligand flexibility simulation— systemic and random methods. Molecular dynamics simulation is the cornerstone for simulation ligand flexibility search methods [129]. However, simulation search methods have a disadvantage, in that they

cannot pass high energy barriers during the measurement of flexibility in a short time. The ligand flexibility systemic search mainly discovers all possible classes of reedom in a ligand. However, an alternative method for a systemic ligand flexibility search starts by generating a library of compounds and creating different conformers for each one of the compounds. By utilizing this approach, the library of different ligand conformers will be docked against the biological target at one time, which reduces the time and efforts of this step. Lastly, the random ligand flexibility search method allows the ligand to change to different conformers and mimic the degrees of ligand freedom. Genetic algorithm and Mento Carlo are the most popular random ligand flexibility search methods [129].

The ligand flexibility search method is considered more advanced than the protein flexibility search method [125]. In general, search methods measure the flexibility in a particular part of the protein, usually the side chain, which is accomplished by various methods such as rotomer libraries, molecular dynamic, protein ensemble grids and Monte Carlo [126, 130, 131].

1.6.6 Types of Docking Scoring Functions:

The main goal of docking scoring is to calculate the energy difference (ΔG) of the binding of a ligand-protein complex. Generally, various attempts have been made to accomplish the ΔG calculation in a short time, since it is considered a time- and money-consuming process. The three types of docking scoring functions are called empirical-based, force-field based and knowledge-based [125]. Empirical-based scoring functions

rely on the binding energy of the ligand-protein complex that can be measured by the addition of single unrelated terms; the coefficients of the terms are measured from a regression examination utilizing experimental measurements of binding energy and known data from an X-ray crystallography (3D structure). In the force-field scoring function, the sum of the internal-ligand energy and protein-ligand energy is the main consideration. The majority of force-field scoring functions do not count the various conformer structures of one protein, which make the internal protein energy measurement useless for most purposes [125, 128]. The force-field scoring functions are dependent on various groups of force-field parameters, including AMBER, Tripos force field, OPLS-AA and MMFF94. Electrostatic and Van der Waals energy interactions are the parameters that describe the protein-ligand interactions. The electrostatic expression is described by Coulomb equations that explain the force interaction between two differently charged particles, while the Van der Waals interaction expression is explained by Lennrd-Jones potential functions, where scoring functions are responsible for different results [125, 128]. Knowledge-based docking scoring functions were invented to ignore binding energy and instead lean on experimental X-ray crystallography (3D structure) [125]. The complex interaction of a ligand-protein is designed utilizing an uncomplicated atomic interaction system.

The computational effort is the distinguishing characteristic of these scoring functions, which facilitate the process of scoring huge libraries in a short time. However, their use relies on known information from small numbers of ligand-protein complex structures. The numerous interactions that rule ligand-protein interaction formation cannot be managed by a single scoring function. Therefore, consensus scoring functions are very

common, since they acquire information from various scores, which minimizes the errors of a single score and enhances the chance of finding the hits [125, 132]. However, if the terms in the various scoring functions are safely matched, it may destroy the balance and increase the errors.

In conclusion, molecular modeling is a helpful technique to find potential candidates to be synthesized and structurally optimized. These candidates can then be followed by biological evaluation to provide the clinical form of the synthesized compounds.

1.7   Biological Evaluation of Synthesized Compounds:

All new compounds should be biologically screened on various biological assays. Biological evaluation can be conducted in animals, isolated organs, lower organisms, or ex-vivo and in-vivo screening assays. Some of these assays, such as whole animal evaluations, are very expensive to use, especially in the early stages of biological screening, compared to less expensive biological assays such as cultured cell lines known as in-vitro assays [133].

The most commonly used in-vitro biological assay is cytotoxicity. This assay measures the lowest concentration of a particular compound that can inhibit the growth of 50% of a particular cell line (IC_{50}). The cell growth inhibition can be tested by one of the

following tests: cellular biomass assay, membrane integrity assays, day binding assays and metabolic impairment assays.

To measure the dissolution of enzyme activity or the concentration metabolites, a metabolic impairment assay can be used by adding a toxin. The MTT assay mainly checks the capability of live cells to convert by reduction the yellow 3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyltrazollium bromide salts, known as MTT salt, to its purple formazan state utilizing a mitochondrial oxidoreductase enzyme [134].

The ligand-binding biological assay is a very common assay in the drug discovery process. This assay can be conducted utilizing Enzyme Linked Immunoassays (ELISA) which simply confirm the binding between the protein and the ligand [135, 136]. In order to pinpoint a particular protein in the cell, the target specific assay in a cell-based ELISA kit can be utilized to identify this particular protein. This is considered a very powerful assay compared to other assays, such as the western blot.

In conclusion, biological evaluation is one of the cornerstones of the drug discovery process because it confirms the hypothesis using molecular modeling studies. In addition, it helps to identify the significant pharmacophores for further chemical optimization in the synthetic scheme. All of these connected steps are very helpful in order to develop active promising candidates for future drugs.

1.8   Project Objectives:

This project concentrated on using cucurbitacins as a natural product that has inhibitory activity toward the epidermal growth factor receptor (EGFR) to develop drug candidates for the treatment of hepatocellular carcinoma (HCC). Few studies have defined the activity of cucurbitacins toward EGFR in HCC cell lines, such as the HepG2 cell line. In addition, no cucurbitacin-like compounds have been synthesized and biologically evaluated targeting HCC. The main objectives of this project are:

A- Design of novel CUCS-inspired estrone analogues targeting epidermal growth factors receptor (EGFR) using molecular modeling studies.

B- Synthesis of novel CUCS-inspired estrone analogues targeting epidermal growth factors receptor (EGFR).

C- Biological screening of the synthesized novel CUCS-inspired estrone analogues targeting epidermal growth factors receptor (EGFR).

D- Inhibitory activity screening of CUCUS-inspired estrone on MRP1 in order to overcome the cancer resistance for current chemotherapeutic drugs.

E- Structure activity relationships (SAR) studies on the synthesized novel CUCS-inspired estrone analogues.

Docking studies were performed on cucurbitacins structure to identify the significant pharmacophores in order to design novel CUCS-inspired estrone analogues targeting epidermal growth factors receptor (EGFR). The synthetic route was designed and outlined by using hybrid structural design methods in which the cucurbitacin side chain

was installed in the steroidal skeleton structure, then structural modifications were conducted on the C-17 and C-17, C-3 and B/C juncture of the steroidal skeleton structure. In addition, a series of various aldehydes were installed on the side chain as part of the chemical design using Aldol condensation reactions. The biological screening tests were used to analyze the synthesized compounds to be evaluated for further optimizations targeting EGFR.

1.9   References:

1. *Schmid*, E.F. and D.A. Smith, *Pharmaceutical R&D in the spotlight: why is there still unmet medical need?* Drug discovery today, 2007. **12**(23): p. 998-1006.

2. *Augen*, J., *The evolving role of information technology in the drug discovery process.* Drug Discovery Today, 2002. **7**(5): p. 315-323.

3. Schnecke, V. and J. Boström, *Computational chemistry-driven decision making in lead generation.* Drug discovery today, 2006. **11**(1): p. 43-50.

4. Milne, G.M., . *Pharmaceutical productivity—the imperative for new paradigms.* Annual Reports in Medicinal Chemistry, 2003. **38**: p. 383-396.

5. Lipinski, C.A., et al., *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings.* Advanced drug delivery reviews, 2012. **64**: p. 4-17.

6. Alavijeh, M. and A. Palmer, *The pivotal role of drug metabolism and pharmacokinetics in the discovery and development of new medicines.* IDrugs: the investigational drugs journal, 2004. **7**(8): p. 755-763.

7. Anand, S.A.A., et al., *Synthesis, structure prediction, pharmacokinetic properties, molecular docking and antitumor activities of some novel thiazinone derivatives.* New Journal of Chemistry, 2015. **39**(9): p. 7120-7129.

8. Leeson, P.D. and B. Springthorpe, *The influence of drug-like concepts on decision-making in medicinal chemistry.* Nature Reviews Drug Discovery, 2007. **6**(11): p. 881-890.

9. Fox, S., et al., *High-throughput screening: update on practices and success.* Journal of biomolecular screening, 2006. **11**(7): p. 864-869.

10. Williams, A.J., *A perspective of publicly accessible/open-access chemistry databases.* Drug discovery today, 2008. **13**(11): p. 495-501.

11. Ali, G., et al., *Input of Isosteric and Bioisosteric Approach in Drug Design.* Journal of the Chemical Society of Pakistan, 2014. **36**(1): p. 150-169.

12. Newman, D.J., G.M. Cragg, and K.M. Snader, *Natural products as sources of new drugs over the period 1981-2002.* Journal of natural products, 2003. **66**(7): p. 1022-1037.

13. Harvey, A.L., R. Edrada-Ebel, and R.J. Quinn, *The re-emergence of natural products for drug discovery in the genomics era.* Nature Reviews Drug Discovery, 2015. **14**(2): p. 111-129.

14. Ji, H.F., X.J. Li, and H.Y. Zhang, *Natural products and drug discovery.* EMBO reports, 2009. **10**(3): p. 194-200.

15. Koehn, F.E. and G.T. Carter, *The evolving role of natural products in drug discovery.* Nature reviews Drug discovery, 2005. **4**(3): p. 206-220.

16. Chin, Y.-W., et al., *Drug discovery from natural sources.* The AAPS journal, 2006. **8**(2): p. E239-E253.

17. Graul, A.I., *The year's new drugs.* Drug News Perspect, 2001. **14**: p. 12-31.

18. Newman, D.J. and G.M. Cragg, *Natural products as sources of new drugs over the 30 years from 1981 to 2010.* Journal of natural products, 2012. **75**(3): p. 311-335.

19. Miro, M., *Cucurbitacins and their pharmacological effects.* Phytotherapy research, 1995. **9**(3): p. 159-168.

20. Ram, V.J. and A. Goel, *Past and present scenario of hepatoprotectants.* Current medicinal chemistry, 1999. **6**(3): p. 217-254.

21. Han, T., et al., *Preventive effects of cucurbitacin B on experimental hepatitis and cirrhosis.* Chung-hua I Hsueh Tsa Chih (Beijing), 1979. **59**: p. 206-209.

22. Chen, X., et al., *Biological activities and potential molecular targets of cucurbitacins: a focus on cancer.* Anti-cancer drugs, 2012. **23**(8): p. 777-787.

23. Fuller, R.W., et al., *Cucurbitacins: differential cytotoxicity, dereplication and first isolation from Gonystylus keithii.* Journal of natural products, 1994. **57**(10): p. 1442-1445.

24. Huang, Y., et al., *Complement-inhibiting cucurbitacin glycosides from Picria fel-terrae.* Journal of natural products, 1998. **61**(6): p. 757-761.

25. Lee, D.H., G.B. Iwanski, and N.H. Thoennissen, *Cucurbitacin: ancient compound shedding new light on cancer treatment.* The Scientific World Journal, 2010. **10**: p. 413-418.

26. Alghasham, A.A., *Cucurbitacins–a promising target for cancer therapy.* International journal of health sciences, 2013. **7**(1).

27. Recio, M.C., et al., *Anti-inflammatory activity of two cucurbitacins isolated from Cayaponia tayuya roots.* Planta medica, 2004. **70**(05): p. 414-420.

28. Bartalis, J. and F.T. Halaweish, *In vitro and QSAR studies of cucurbitacins on HepG2 and HSC-T6 liver cell lines.* Bioorganic & medicinal chemistry, 2011. **19**(8): p. 2757-2766.

29. Chan, K.T., et al., *Cucurbitacin B inhibits STAT3 and the Raf/MEK/ERK pathway in leukemia cell line K562.* Cancer letters, 2010. **289**(1): p. 46-52.

30. de Reinach-Hirtzbach, F. and G. Ourisson, *Synthese de la chaine laterale des cucurbitacines.* Tetrahedron Letters, 1973. **14**(16): p. 1363-1366.

31. Kolaczkowski, L. and W. Reusch, *Total synthesis of tetracyclic triterpenes. 1. The (.+-.)-5-epi-euphane ring system.* The Journal of Organic Chemistry, 1985. **50**(24): p. 4766-4771.

32. Jung, M.E. and R.M. Lui, *Studies toward the Total Syntheses of Cucurbitacins B and D.* The Journal of organic chemistry, 2010. **75**(21): p. 7146-7158.

33. Gry, J., I. Søborg, and H.C. Andersson, *Cucurbitacins in plant food*. 2006: Nordic Council of Ministers.

34. Molavi, O., et al., *Polymeric micelles for the solubilization and delivery of STAT3 inhibitor cucurbitacins in solid tumors.* International journal of pharmaceutics, 2008. **347**(1): p. 118-127.

35. Chen, J.C., et al., *Cucurbitacins and cucurbitane glycosides: structures and biological activities.* Natural product reports, 2005. **22**(3): p. 386-399.

36. Bernard, S.A. and O.A. Olayinka, *Search for a novel antioxidant, anti-inflammatory/analgesic or anti-proliferative drug: Cucurbitacins hold the ace.* Journal of Medicinal Plants Research, 2010. **4**(25): p. 2821-2826.

37. Achenbach, H., U. Hefter-Bübl, and M.A. Constenla, *Fevicordin A and fevicordin A glucoside, novel norcucurbitacins from Fevillea cordifolia.* Journal of the Chemical Society, Chemical Communications, 1987(6): p. 441-442.

38. Valente, L.M., et al., *New norcucurbitacin and heptanorcucurbitacin glucosides from Fevillea trilobata.* Journal of natural products, 1993. **56**(10): p. 1772-1778.

39. Matsuda, H., et al., *Cucurbitane-type triterpenes with anti-proliferative effects on U937 cells from an egyptian natural medicine, Bryonia cretica: structures of new triterpene glycosides, bryoniaosides A and B.* Chemical and Pharmaceutical Bulletin, 2010. **58**(5): p. 747-751.

40. Duncan, K.L., et al., *Cucurbitacin E-induced disruption of the actin and vimentin cytoskeleton in prostate carcinoma cells.* Biochemical pharmacology, 1996. **52**(10): p. 1553-1560.

41. Sörensen, P.M., et al., *The natural product cucurbitacin E inhibits depolymerization of actin filaments.* ACS chemical biology, 2012. **7**(9): p. 1502-1508.

42. Gabrielsen, M., et al., *Cucurbitacin covalent bonding to cysteine thiols: the filamentous-actin severing protein Cofilin1 as an exemplary target.* Cell Communication and Signaling, 2013. **11**(1): p. 1.

43. Ahmed, M.S. and F.T. Halaweish, *Cucurbitacins: potential candidates targeting mitogen-activated protein kinase pathway for treatment of melanoma.* Journal of enzyme inhibition and medicinal chemistry, 2014. **29**(2): p. 162-167.

44. Bollée, G., et al., *Epidermal growth factor receptor promotes glomerular injury and renal failure in rapidly progressive crescentic glomerulonephritis.* Nature medicine, 2011. **17**(10): p. 1242-1250.

45. Holbro, T. and N.E. Hynes, *ErbB receptors: directing key signaling networks throughout life.* Annu. Rev. Pharmacol. Toxicol., 2004. **44**: p. 195-217.

46. Llovet, J.M. and J. Bruix, *Molecular targeted therapies in hepatocellular carcinoma.* Hepatology, 2008. **48**(4): p. 1312-1327.

47. Bellissimo, F., et al., *Diagnostic and therapeutic management of hepatocellular carcinoma.* World journal of gastroenterology, 2015. **21**(42): p. 12003.

48. Altekruse, S.F., K.A. McGlynn, and M.E. Reichman, *Hepatocellular carcinoma incidence, mortality, and survival trends in the United States from 1975 to 2005.* Journal of clinical oncology, 2009. **27**(9): p. 1485-1491.

49. Farazi, P.A. and R.A. DePinho, *Hepatocellular carcinoma pathogenesis: from genes to environment.* Nature Reviews Cancer, 2006. **6**(9): p. 674-687.

50. Cokkinides, V., et al., *American cancer society: Cancer facts and figures 2005.* Atlanta. American Cancer Society, 2005.

51. Badvie, S., *Hepatocellular carcinoma.* Postgraduate Medical Journal, 2000. **76**(891): p. 4-11.

52. Lok , A.S.F., *Chronic Hepatitis B.* New England Journal of Medicine, 2002. **346**(22): p. 1682-1683.

53. Lavanchy, D., *Hepatitis B virus epidemiology, disease burden, treatment, and current and emerging prevention and control measures.* Journal of viral hepatitis, 2004. **11**(2): p. 97-107.

54. Chen, C.-J., et al., *Risk of hepatocellular carcinoma across a biological gradient of serum hepatitis B virus DNA level.* Jama, 2006. **295**(1): p. 65-73.

55. Tokino, T. and K. Matsubara, *Chromosomal sites for hepatitis B virus integration in human hepatocellular carcinoma.* Journal of virology, 1991. **65**(12): p. 6761-6764.

56. Murakami, Y., et al., *Large scaled analysis of hepatitis B virus (HBV) DNA integration in HBV related hepatocellular carcinomas.* Gut, 2005. **54**(8): p. 1162-1168.

57. Feitelson, M.A., et al., *Genetic mechanisms of hepatocarcinogenesis.* Oncogene, 2002. **21**(16): p. 2593-2604.

58. Ueda, H., et al., *Functional inactivation but not structural mutation of p53 causes liver cancer.* Nature genetics, 1995. **9**(1): p. 41-47.

59. Yu, D.-Y., et al., *Incidence of hepatocellular carcinoma in transgenic mice expressing the hepatitis B virus X-protein.* Journal of hepatology, 1999. **31**(1): p. 123-132.

60. Bartosch, B., *Hepatitis B and C viruses and hepatocellular carcinoma.* Viruses, 2010. **2**(8): p. 1504-1509.

61. Modi, A. and T. Liang, *Hepatitis C: a clinical review.* Oral diseases, 2008. **14**(1): p. 10-14.

62. Martins, T., J.L. Narciso-Schiavon, and L.d.L. Schiavon, *Epidemiology of hepatitis C virus infection.* Revista da Associação Médica Brasileira, 2011. **57**(1): p. 107-112.

63. Rehermann, B. and M. Nascimbeni, *Immunology of hepatitis B virus and hepatitis C virus infection.* Nature Reviews Immunology, 2005. **5**(3): p. 215-229.

64. Weiner, A., et al., *Persistent hepatitis C virus infection in a chimpanzee is associated with emergence of a cytotoxic T lymphocyte escape variant.* Proceedings of the National Academy of Sciences, 1995. **92**(7): p. 2755-2759.

65. McClain, C.J., et al., *Monocyte activation in alcoholic liver disease.* Alcohol, 2002. **27**(1): p. 53-61.

66. Hoek, J.B. and J.G. Pastorino, *Ethanol, oxidative stress, and cytokine-induced liver cell injury.* Alcohol, 2002. **27**(1): p. 63-68.

67. McClain, C.J., et al., *S-Adenosylmethionine, cytokines, and alcoholic liver disease.* Alcohol, 2002. **27**(3): p. 185-192.

68. Campbell, J.S., et al., *Platelet-derived growth factor C induces liver fibrosis, steatosis, and hepatocellular carcinoma.* Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(9): p. 3389-3394.

69. Osna, N.A., D.L. Clemens, and T.M. Donohue, *Ethanol metabolism alters interferon gamma signaling in recombinant HepG2 cells.* Hepatology, 2005. **42**(5): p. 1109-1117.

70. Kurz, D.J., et al., *Chronic oxidative stress compromises telomere integrity and accelerates the onset of senescence in human endothelial cells.* Journal of cell science, 2004. **117**(11): p. 2417-2426.

71. Sanyal, A.J., S.K. Yoon, and R. Lencioni, *The etiology of hepatocellular carcinoma and consequences for treatment.* The oncologist, 2010. **15**(Supplement 4): p. 14-22.

72. Gurtsevitch, V., *Human oncogenic viruses: hepatitis B and hepatitis C viruses and their role in hepatocarcinogenesis.* Biochemistry (Moscow), 2008. **73**(5): p. 504-513.

73. Gomaa, A.I., et al., *Hepatocellular carcinoma: epidemiology, risk factors and pathogenesis.* World J Gastroenterol, 2008. **14**(27): p. 4300-4308.

74. Caillot, F., et al., *Transient and etiology-related transcription regulation in cirrhosis prior to hepatocellular carcinoma occurrence.* World J Gastroenterol, 2009. **15**(3): p. 300-309.

75. El–Serag, H.B. and K.L. Rudolph, *Hepatocellular carcinoma: epidemiology and molecular carcinogenesis.* Gastroenterology, 2007. **132**(7): p. 2557-2576.

76. Humans, I.W.G.o.t.E.o.C.R.t., *IARC monographs on the evaluation of carcinogenic risks to humans. Ingested nitrate and nitrite, and cyanobacterial peptide toxins.* IARC monographs on the evaluation of carcinogenic risks to humans/World Health Organization, International Agency for Research on Cancer, 2010. **94**: p. v.

77. Garner, R.C., E.C. Miller, and J.A. Miller, *Liver microsomal metabolism of aflatoxin B1 to a reactive derivative toxic to Salmonella typhimurium TA 1530.* Cancer research, 1972. **32**(10): p. 2058-2066.

78. Bressac, B., et al., *Selective G to T mutations of p53 gene in hepatocellular carcinoma from southern Africa.* Nature, 1991. **350**(6317): p. 429-431.

79. Calle, E.E., et al., *Overweight, obesity, and mortality from cancer in a prospectively studied cohort of US adults.* New England Journal of Medicine, 2003. **348**(17): p. 1625-1638.

80. Evans, A.A., et al., *Eight-year follow-up of the 90,000-person Haimen City cohort: I. Hepatocellular carcinoma mortality, risk factors, and gender differences.* Cancer Epidemiology Biomarkers & Prevention, 2002. **11**(4): p. 369-376.

81. Grandhi, M.S., et al., *Hepatocellular carcinoma: From diagnosis to treatment.* Surgical Oncology, 2016. **25**(2): p. 74-85.

82. Chen, J., et al., *Screening for liver cancer: results of a randomised controlled trial in Qidong, China.* Journal of Medical Screening, 2003. **10**(4): p. 204-209.

83. Di Bisceglie, A.M., et al., *Serum alpha-fetoprotein levels in patients with advanced hepatitis C: results from the HALT-C Trial.* Journal of hepatology, 2005. **43**(3): p. 434-441.

84. Llovet, J.M., A. Burroughs, and J. Bruix, *Hepatocellular carcinoma.* The Lancet. **362**(9399): p. 1907-1917.

85. Schwartz, M., S. Roayaie, and M. Konstadoulakis, *Strategies for the management of hepatocellular carcinoma.* Nature clinical practice Oncology, 2007. **4**(7): p. 424-432.

86. Poon, R.T.P. and S.T. Fan, *Hepatectomy for hepatocellular carcinoma: patient selection and postoperative outcome.* Liver transplantation, 2004. **10**(S2).

87. Mazzaferro, V., et al., *Liver transplantation for the treatment of small hepatocellular carcinomas in patients with cirrhosis.* New England Journal of Medicine, 1996. **334**(11): p. 693-700.

88. Bruix, J., *Treatment of hepatocellular carcinoma.* Hepatology, 1997. **25**(2): p. 259-262.

89. society, A.c. *A guide to chemotherapy* 2015.

90. Zhang, J., et al., *Erlotinib for advanced hepatocellular carcinoma. A systematic review of phase II/III clinical trials.* Saudi Medical Journal, 2016. **37**(11): p. 1184-1190.

91. Deng, G.-L., S. Zeng, and H. Shen, *Chemotherapy and target therapy for hepatocellular carcinoma: New advances and challenges.* World J Hepatol, 2015. **7**(5): p. 787-798.

92. Finn, R.S., *Development of molecularly targeted therapies in hepatocellular carcinoma: where do we go now?* Clinical Cancer Research, 2010. **16**(2): p. 390-397.

93. Yeo, W., et al., *A randomized phase III study of doxorubicin versus cisplatin/interferon α-2b/doxorubicin/fluorouracil (PIAF) combination chemotherapy for unresectable hepatocellular carcinoma.* Journal of the National Cancer Institute, 2005. **97**(20): p. 1532-1538.

94. Ng, I.O., et al., *Expression of P-Glycoprotein in Hepatocellular Carcinoma.* American journal of clinical pathology, 2000. **113**(3): p. 355-363.

95. Park, J.-G., et al., *MDR1 gene expression: its effect on drug resistance to doxorubicin in human hepatocellular carcinoma cell lines.* Journal of the National Cancer Institute, 1994. **86**(9): p. 700-705.

96. Chattopadhyay, D., D. Manas, and H. Reeves, *The development of targeted therapies for hepatocellular cancer.* Current pharmaceutical design, 2007. **13**(32): p. 3292-3300.

97. Furuse, J., *Growth factors as therapeutic targets in HCC.* Critical reviews in oncology/hematology, 2008. **67**(1): p. 8-15.

98. Liu, Y., et al., *Blockage of epidermal growth factor receptor by quinazoline tyrosine kinase inhibitors suppresses growth of human hepatocellular carcinoma.* Cancer letters, 2007. **248**(1): p. 32-40.

99. Yarden, Y. and M.X. Sliwkowski, *Untangling the ErbB signalling network.* Nature reviews Molecular cell biology, 2001. **2**(2): p. 127-137.

100. Kudo, M., *Signaling pathway and molecular-targeted therapy for hepatocellular carcinoma.* Digestive diseases, 2011. **29**(3): p. 289-302.

101. Thomas, M.B., et al., *Phase 2 study of erlotinib in patients with unresectable hepatocellular carcinoma.* Cancer, 2007. **110**(5): p. 1059-1067.

102. Yamaguchi, R., et al., *Expression of vascular endothelial growth factor in human hepatocellular carcinoma.* Hepatology, 1998. **28**(1): p. 68-77.

103. Jain, R.K., *Normalization of tumor vasculature: an emerging concept in antiangiogenic therapy.* Science, 2005. **307**(5706): p. 58-62.

104. Gerber, H.-P. and N. Ferrara, *Pharmacology and pharmacodynamics of bevacizumab as monotherapy or in combination with cytotoxic therapy in preclinical studies.* Cancer research, 2005. **65**(3): p. 671-680.

105. Thomas, M., *Molecular targeted therapy for hepatocellular carcinoma.* Journal of gastroenterology, 2009. **44**(19): p. 136-141.

106. Kudo, M., *Signaling pathway/molecular targets and new targeted agents under development in hepatocellular carcinoma.* World Journal of Gastroenterology : WJG, 2012. **18**(42): p. 6005-6017.

107. Toyoda, M., et al., *Increased activity and expression of MAP kinase in HCC model rats induced by 3'-methyl-4-dimethylamino-azobenzene.* Journal of hepatology, 1999. **31**(4): p. 725-733.

108. De Yun Feng, H.Z., Y. Tan, and R.X. Cheng, *Citation of This Article.* World J Gastroenterol, 2001. **7**(1): p. 33-36.

109. Carloni, V., F. Vizzutti, and P. Pantaleo, *Farnesyltransferase inhibitor, ABT-100, is a potent liver cancer chemopreventive agent.* Clinical cancer research, 2005. **11**(11): p. 4266-4274.

110. Hwang, Y.H., et al., *Over-expression of c-raf-1 proto-oncogene in liver cirrhosis and hepatocellular carcinoma.* Hepatology research, 2004. **29**(2): p. 113-121.

111. Wilhelm, S.M., et al., *Preclinical overview of sorafenib, a multikinase inhibitor that targets both Raf and VEGF and PDGF receptor tyrosine kinase signaling.* Molecular cancer therapeutics, 2008. **7**(10): p. 3129-3140.

112. Huynh, H., et al., *Over-expression of the mitogen-activated protein kinase (MAPK) kinase (MEK)-MAPK in hepatocellular carcinoma: its role in tumor progression and apoptosis.* BMC gastroenterology, 2003. **3**(1): p. 1.

113. Adjei, A.A. and M. Hidalgo, *Treating cancer by blocking cell signals.* Journal of Clinical Oncology, 2005. **23**(23): p. 5279-5280.

114. Sieghart, W., et al., *Mammalian target of rapamycin pathway activity in hepatocellular carcinomas of patients undergoing liver transplantation.* Transplantation, 2007. **83**(4): p. 425-432.

115. Csermely, P., et al., *Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review.* Pharmacology & therapeutics, 2013. **138**(3): p. 333-408.

116. Kuntz, I.D., et al., *A geometric approach to macromolecule-ligand interactions.* Journal of molecular biology, 1982. **161**(2): p. 269-288.

117. Ekins, S., J. Mestres, and B. Testa, *In silico pharmacology for drug discovery: applications to targets and beyond.* British journal of pharmacology, 2007. **152**(1): p. 21-37.

118. Zhong, S., A.T. Macias, and A.D. MacKerell, *Computational identification of inhibitors of protein-protein interactions.* Current topics in medicinal chemistry, 2007. **7**(1): p. 63-82.

119. Deng, Z., C. Chuaqui, and J. Singh, *Knowledge-based design of target-focused libraries using protein-ligand interaction constraints.* Journal of medicinal chemistry, 2006. **49**(2): p. 490-500.

120. Shoichet, B.K., et al., *Lead discovery using molecular docking.* Current opinion in chemical biology, 2002. **6**(4): p. 439-446.

121. Mehrotra, S.A.a.R., *An overview of Molecular Docking.* JSM Chemistry, 2016. **4**(2): p. 1024.

122. Mukesh, B. and K. Rakesh, *Molecular docking: a review.* Int J Res Ayurveda Pharm, 2011. **2**: p. 746-1751.

123. Huang, S.-Y. and X. Zou, *Advances and challenges in protein-ligand docking.* International journal of molecular sciences, 2010. **11**(8): p. 3016-3034.

124. Mangoni, M., D. Roccatano, and A. Di Nola, *Docking of flexible ligands to flexible receptors in solution by molecular dynamics simulation.* Proteins: Structure, Function, and Bioinformatics, 1999. **35**(2): p. 153-162.

125. Kitchen, D.B., et al., *Docking and scoring in virtual screening for drug discovery: methods and applications.* Nature reviews Drug discovery, 2004. **3**(11): p. 935-949.

126. Halperin, I., et al., *Principles of docking: An overview of search algorithms and a guide to scoring functions.* Proteins: Structure, Function, and Bioinformatics, 2002. **47**(4): p. 409-443.

127. Sperandio, O., et al., *Receptor-based computational screening of compound databases: the main docking-scoring engines.* Current Protein and Peptide Science, 2006. **7**(5): p. 369-393.

128. Kroemer, R.T., *Structure-based drug design: docking and scoring.* Current Protein and Peptide Science, 2007. **8**(4): p. 312-328.

129. Brooijmans, N. and I.D. Kuntz, *Molecular recognition and docking algorithms.* Annual review of biophysics and biomolecular structure, 2003. **32**(1): p. 335-373.

130. Hart, T.N. and R.J. Read, *A multiple‐ start Monte Carlo docking method.* Proteins: Structure, Function, and Bioinformatics, 1992. **13**(3): p. 206-222.

131. Leach, A.R., *Ligand docking to proteins with discrete side-chain flexibility.* Journal of molecular biology, 1994. **235**(1): p. 345-356.

132. Charifson, P.S., et al., *Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins.* Journal of medicinal chemistry, 1999. **42**(25): p. 5100-5109.

133. McCauley, J., A. Zivanovic, and D. Skropeta, *Bioassays for anticancer activities.* Metabolomics Tools for Natural Product Discovery: Methods and Protocols, 2013: p. 191-205.

134. Berridge, M.V., P.M. Herst, and A.S. Tan, *Tetrazolium dyes as tools in cell biology: new insights into their cellular reduction.* Biotechnology annual review, 2005. **11**: p. 127-152.

135. Gan, S.D. and K.R. Patel, *Enzyme immunoassay and enzyme-linked immunosorbent assay.* Journal of Investigative Dermatology, 2013. **133**(9): p. 1-3.

136. Aydin, S., *A short history, principles, and types of ELISA, and our laboratory experience with peptide/protein analyses using ELISA.* Peptides, 2015. **72**: p. 4-15.

Chapter Two

Design of CUCUS-Inspired Estrone Analogs (CIEA) Targeting Epidermal Growth Factor Receptor (EGFR)

2.1   Introduction:

Natural products are known as one of the major source for drug design and treatment of cancer. Cucurbitacins is a group of steroidal-triterpene tetracyclic natural products, which reported for their anti-cancer activities [1, 2]. Although cucurbitacins have reported for their potent activities, synthesis of these compounds is challenging due to the complexity of the carbon skeleton and functionalities of these compounds. Current studies in Halaweish's group to study cucurbitacins targeting epidermal growth factor receptor (EGFR) in unclear and limited [3].

Molecular modeling and docking methods were utilized to find potential affinity between cucurbitacins and EGFR along with downstream proteins cascade including Ras, Raf, MEK and ERK (**Fig. 1.5**). In addition, additional studies including cytotoxicity, western blot and ELISA were used to confirm the molecular docking studies results. Cucurbitacins proved to have activities against different cancer cell lines; however, their activities don't show specificity or selectivity toward their biological targets [4]. Gastrointestinal toxicity is one of the side effects involved with cucurbitacins subjections due to their cellular activations [5]. Potent cytotoxicity of cucurbitacins in in-vivo model toward renal carcinoma demonstrated narrow safety and have been withdrawn from

preclinical studies due to their fatal activity [6]. The broad biological activities of cucurbitacins, non-selectivity and toxicities are due to their complex chemical structures.

Model the bioisosters of these compounds to mimic the structure functionalities on steroidal skeleton. Cucurbitacins inspired estrone compounds was accomplished in Halaweish's group targeting melanoma [15]. The success in Halaweish's group of utilizing the molecular docking inspired us to model novel analogues to target hepatocellular carcinoma (HCC).

Due to the similarity in chemical structure, particularly the cyclopentane and the four-ring system, redesigning and mimicking cucurbitacins utilizing steroids may improve their biological activities and selectivities. Cucurbitacins side chain, which contain α-β-unsaturated ketone, is significant pharmacophore for their biological activities. Also, the presence of C-16 hydroxyl group increased the chance of forming H-bond with C-24 ketone which may enhance the electrophilisity of the α-β-unsaturated ketone [7,8].

Studying and testing cucurbitacins as promising EGFR inhibitors is not favorable sometimes due to availability limitation, has encouraged us to identify the most significant pharmacophores of cucurbitacins in order to synthetically modify their main skeleton to increase the selectivity toward EGFR and minimized their undesirable side effects [14]. Few researches and studies on the structure modifications of cucurbitacins have been conducted including quantitative structure activity relationship (QSAR) studies of semi synthetic of cucurbitacins by Bartalis and Halaweish [8], partial synthesis of cucurbitacin B and D by Jung [9], hexanorcucurbitacin compounds by Ryu [10] and synthesis of dihydrocucurbitacin B by Lang [11]. However, some of these synthetic studies are

impracticable for certain wanted cucurbitacins pharmacophores to be modified for certain biological activities.

In order to understand and test the potential interactions between the designed ligands and their biological targets, molecular modeling considered a precious technique for this study. Precise ligand conformation is highly recommended during drug design to perform an accurate structure-based drug design and virtual screening in order to have full understand of ligand-target complex interactions through reading the energy parameter [12, 13].

The biological activity of any molecule should rely on one conformation that possesses the lowest energy conformation which can be distinguished among all other low-energy conformers by using Cambridge crystallographic database (**Fig. 2.1**). All of these signs promoted us to build an assumption that utilizing steroidal skeleton to install cucurbitacins pharmacophores to inhibit EGFR could be a potential assumption as shown in **figure 2.2.**
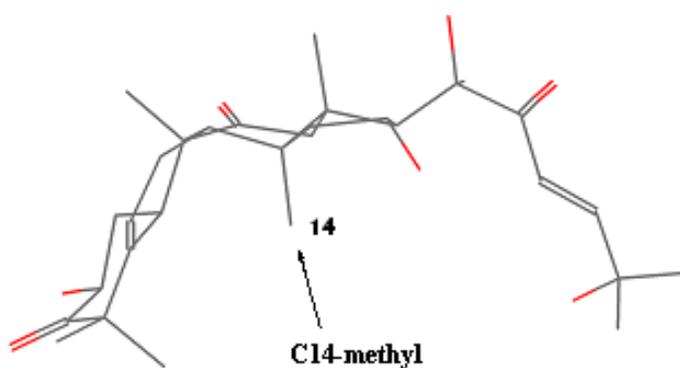
Figure 2.1 X-ray crystallography of cucurbitacin D showing the potential functional groups for the binding with biological targets.

Figure2.2 General structures of Cucurbitacins and steroid.

Cucurbitacin	R_1	R_2	R_3	R_4	R_5	$\Delta^{1,2}$	$\Delta^{23,24}$
A	OH	=O	MeOH	H	AC	-	+
B	OH	=O	CH_3	H	AC	-	+
C	H	OH	MeOH	H	AC	-	+
D	OH	=O	CH_3	H	OH	-	+
E	OH	=O	CH_3	H	AC	+	+
F	OH	OH	CH_3	H	OH	-	+
H	OH	=O	CH_3	OH	OH	-	-
I	OH	=O	CH_3	H	OH	+	+
J	OH	=O	CH_3	OH	OH	+	-

Assigning the significant pharmacophores of the steroids is complicated mission. The B and C rings of the steroids have trans configuration but it is cis configuration at cucrbitacins B and C rings which should be taken in the consideration during the designing process [14, 15].

Bartalis et. al 2011 showed that possessing α-β-unsaturated ketone moiety of the cucurbitacins is essential for the biological activity [8]. In addition, Matsuda et al. 2010 demonstrated the significant of α-β-unsaturated ketone moiety of the cucurbitacins scaffolds at C-22 and C-24 for their biological activities when he found that some cucurbitacins-like structures that has no α-β-unsaturated ketone such as hexanor compounds lost their biological activities [7, 15]. Furthermore, Rodrigues et al, 2016 showed that one sixth of all known natural compounds contain α-β-unsaturated carbonyls [16]. Which prove the biological significant of α-β-unsaturated ketone at the cucurbitacins side chain.

Estrogens including estriol, estradiol and estrone are group of popular steroids that naturally present in women and men and they involved in different biological processes [17]. In addition, estrogens are involved in the human body with hormonal activities and development of estrogen-dependent breast cancer [17, 18]. The ability of estrogens and their derivatives to influence various biological activities and their few adverse effects make them very interested analogues to be modified for other biological interests. For these reasons, estrone was chosen as starting material for installing the cucurbitacins α-β-unsaturated ketone side chain with different functional groups to see their effects on the binding affinity toward the biological targets particularly EGFR.
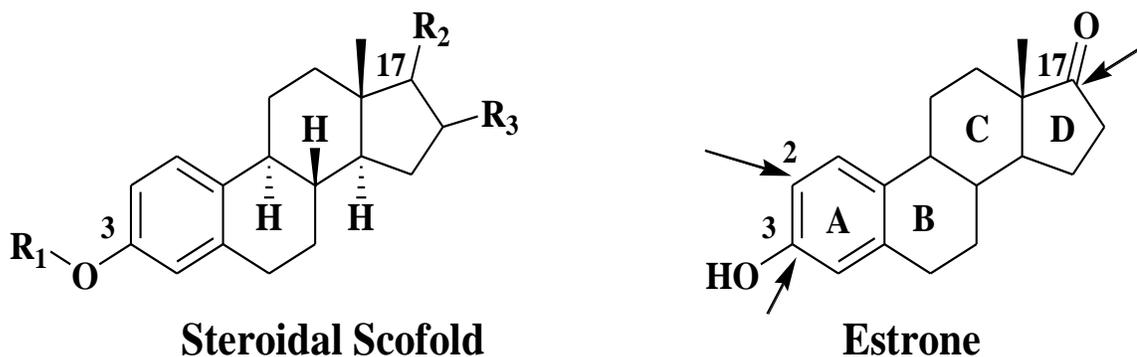
Figure 2.3 Significant positions for biological activities in estrone main structure.

Modifications of steroids skeleton have got the interest of organic chemist due to number of reasons 1) steroidal skeleton provides important and significant chemical challenges; 2) modification in their main structure may provide an interested biological changes as these compounds may act as pharmacophore carrier; 3) their total synthesis is very interested for chemist which may need to find new synthetic approaches for further structural modifications [19]. Serious of estrone modified compounds have been developed based on the fact that small structure modifications on estrone skeleton scaffold may lead to large variations in the biological activities [20]. Estrone structural modifications can be categorized into two types 1) structural modifications by adding new chemical moieties to the estrone scaffold; 2) structural modifications on the steroids complicated ring system by either addition of heteroatoms to one of the steroids carbons, or by chemical modifications in one of the ring system (**Figure 2.3**) [20, 21]. The most popular type is the first one since it provides an overall synthetic tasks that is needed to combine various heteroatoms with the steroids skeleton structure.

Various pharmacophores of natural compounds such as cambretastatin A4, gallic acid and chelcones were installed into the estrone skeleton structure, which is known to possess estrogenic activity and anti-cancer activities targeting different biological targets [22, 23]. However, this hybrid structure modification is not always providing the desired biological activity. Another examples are derivatives of 2-methoxyestradiol such as alkoxy and benzyloxy methylestradiol which proved as anti-breast cancer agents through stabilizing the microtubule without having estrogenic activity as side effect [24]. Because of the low bioavailability of methoxyestradiol, number of modified estradiol in different positions such as C-17, C-16, C-3, and C-2 have been conducted to improve their bioavailability as well as their biological activities [24-27].

Small modifications at C-2 of estrone main structure by installing sulfamate moiety change the biological properties of the estrone dramatically by blocking the estrogenic activity and performing anti-peroflorative activity in breast cancer cells [28]. Bodnar et al, 2016 proved that triazol substitution at C-3 position of estrone enhanced the biological activity as anti-cancer with $IC_{50}=0.3-0.9\mu M$ [27]. Ahmed et al. 2014 demonstrated that substitution on C-17 of the estrone skeleton structure showed a potent inhibitory activity of MAPK pathway toward treatment of melanoma [15]. Therefore, by utilizing molecular modeling, series of modified estrone at C-25, C-17, C-16, C-11 and C-3 positions were designed and developed to target Epidermal Growth Factor Receptor (EGFR) toward treatment of Hepatocellular Carcinoma (HCC).

Estron   **Cucurbitacin D**
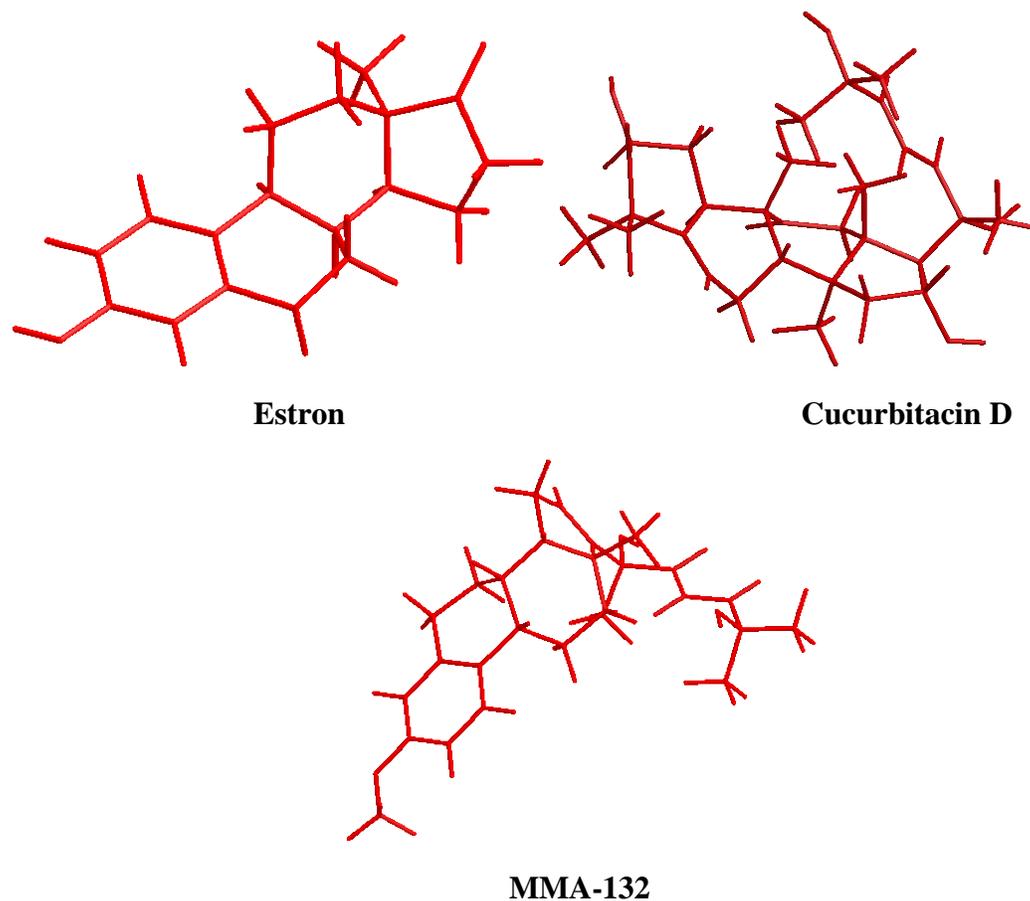
MMA-132

Figure 2.4 3-D structure demonstrated the conformations of Estrone, Cucurbitacin D and MMA-132 analog.

2.1.1   Molecular Modeling and Rational of Inhibitor Design:

2.1.1.1 Protein Kinase:

Protein kinases are the most common family among human enzymes representing 1.7% of the total proteins, which involve in the regulation of several of biological functions in the cell via autophosphorylation including angiogenesis, cell growth and differentiation, which make them an interested target for treatment of cancer [30].

Protein kinase transfer the ATP phosphate group to the amino acids residues of the protein such as theronin, tyrosin and serin. It consists of two main domains, cyctain-rich (CR) domain and ligand domain (LD). In order to design a protein kinase inhibitor, deep understanding of signaling pathway mechanism of interactions should be known, which categorized the protein kinases inhibitors into three categories. 1) Inhibitors that directly bind to the ATP binding pocket of the protein kinase, which usually consist of H-bond donors and acceptors to form at least single H-bond such as erlotinib, fasudil and gefitinib [32]. 2) Inhibitors that bind with inactive protein kinase ATP binding pocket which consist of extra hydrophobic site which increase the selectivity to this site compare the first type such as sorefinib, nilotinib and lapatinib. 3) Allosteric protein kinase inhibitors and they don't compete with ATP to bind to its binding site.

Briefly, most of the protein kinase inhibitors that approved as anti-cancer are ATP competitive inhibitors and form at least one H-bond with protein amino acid residues [29-33].

2.1.1.2 PI3K/Akt/mTOR and RAS/Raf/MEK/ERK Proteins Cascades:

PI3K/Akt/mTOR is significant downstream proteins cascade which can be activated by various cellular activators that lead to number of cellular processes including cell-growth, cell-survival, translation, angiogenesis and transcription. Different types of cancer lead to disturbance in PI3K/Akt/mTOR pathway which make them promising targets for designing promising inhibitors as ant-cancer agents.

Known Raf inhibitor proved to be one of the ATP competitive inhibitors performing various kind of interactions with hydrophobic binding site such as PLX-4720 which has been advanced by Plexxikon to possess high selectivity to B-Raf and high oral bioavailability. Other B-Raf inhibitors includes PLX-4032 (Vemurfenib) and Sorefinib. Various MEK inhibitors have been discovered such as Selumetinib, U0126, PD184352 and PD0325901. MEK inhibitors bind allosterically to the non-ATP binding site of MEK protein which increase their specificity and selectivity. MEK1/2 crystal structures demonstrated that they have outstanding hydrophobic binding pocket beside the ATP-binding pocket [34, 35]. Some other compounds such as resorcyclic inhibitors which demonstrated MEK inhibitory activity and other kinases as well. The presence of enone-based pharmacophore in these inhibitors as Micheal acceptor is the base of their broad biological activities [36]. Our hypothesis of development of esterone to possess anti-cancer activities is by installing the cucurbitacins enone side chain which consist of α-β-unsaturated ketone at C-22 and C-24 of the estrone main scaffold.

2.1.1.3 Compounds with α-β-unsaturated Carbonyl (Micheal Acceptor):

Michael acceptor or α-β-unsaturated carbonyl pharmacophores considered as one of the most reactive functional groups in natural compounds. Recent review in nature chemistry showed that 1/6 of the all known natural compounds contain α-β-unsaturated carbonyl groups. It is a question mark about weather adding Michael acceptor pharmacophores to design a potential candidates targeting specific biological target is valuable or not since these groups are very reactive electrophiles and have broad biological

activities because they can react with various number of nucleophiles in the body such as DNA, RNA and proteins which lead to unwanted side effects. However, Michael acceptor functional groups demonstrated various beneficial activities such as anti-oxidants by trapping thiol and radical scavenger [16, 37].

Therefore, a virtual library of estrone modified structures at different positions including C-3, C-11, C-16, C-17, C-24 and C-25 were generated containing cucurbitacins enone side chain , known EGFR inhibitors and cucurbitacins to be investigated as anti-hepatocellular carcinoma agents utilizing OpenEye® Scientific software including Omega, FRED and VIDA. The designing strategies were divided into two different series using molecular modeling. First series is to assemble cucurbitacins enone side chain at C-17 of estrone with different aliphatic and aromatic moieties at C-25, double bond at C-16 and C-17 and methoxy moiety at C-3. The last functional groups is to block the estrogenic activity of the estrone. Second series is to install cucurbitacins enone side chain at C-17 of estrone with different aliphatic and aromatic moieties at C-25 and sulfamyl or hydroxyl groups at C-3.

2.2   Methods of Molecular Modeling:

All computer works were conducted on Gateway Computer with Windows XP. OpenEye® software, semi flexible molecular docking program, were utilized to conduct the docking process. It consist of different applications such as OMEGA, FRED and VIDA and each one of them is directly related to the other functionally as they will be described in detail in next paragraphs.
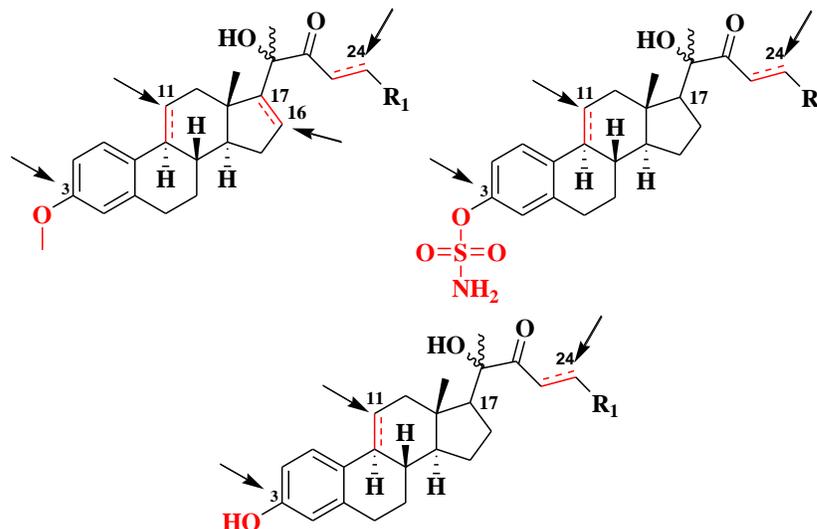
R1= C(CH$_3$)$_2$ OH, P-PhF, P-PhCl, P-PhBr
P-PhCF$_3$, P-PhNO$_2$, 5-Bromo-2-thiophene,
5-Bromo-2-furan.

Figure 2.5 Proposed modified estrone structures.

2.2.1 2-D and 3-D Structures Molecular Modeling:

A virtual library of 200 CUCUS-inspired estrone analogues, which contain cuucurbitacins enone side chain assembled in C-17 of estrone skeleton, other related derivatives, known EGFR inhibitors and different cucurbitacins as standards were generated utilizing Chem office 2012, and using MMFF94 calculation for energy minimization in order to acquire similar structural confirmation mimic to natural 3-D structure for each compound.

2.2.2 Utilizing OMEGA to Generate Conformers:

The energy minimized compounds were converted into *.pdb forms and then all of the *.pdb files collected as one *.pdb file to be utilized as an input parameter for Omega

calculations. Omega created different conformers of each single compound in the virtual library using MMFF94 calculation parameters so as to perform ligand-protein flexible docking. Some of Omega settings were modified from the default settings such as highest number of output conformers 400 (GP-NUM-OUT-CONFS), rejecting conformers that has energy differences compare to standards minimum > 0.5 Kal/mol (GP-ENERGY-WINDO) and choosing the conformers with lowest energy from the final calculations (GP-SELECT-RANDOM false). Furthermore, increase the number of free rotatable bonds in each ligand to be 30 (GP-MAX-ROTORS) so as to create different conformers for all molecules in our library [38].

2.2.3 Preparation of the Receptor Utilizing FRED:

PDB format of the receptors were downloaded from the protein data bank (PDB) database (EGFR PDB ID # 1M17, RAS PDB ID # 4DTS, RAF PDB ID # 3OMV, MEK PDB # 4AN2, ERK PDB ID # 2OJJ) **(Fig.2.6)**. These proteins were prepared utilizing FRED (Receptor preperation program of OpenEye® software). For the preparation of the proteins, graphical interface was used as part of FRED make receptor application. There are three main components of FRED make receptor application including mode selection, mode control and 3D viewing window [39].

The process of receptor preparation started by opening the receptor of interest into the FRED make receptor application which convert the .pdb file into 3-D view of the receptor chains including its binding ligands and co-factors. Then the next step is pointing

to the binding pocket and the co-crystalized ligand binds to it. Then mode selection window of the application will allow the generation of the grid box which should be in specific size not exceeding 60,000 A°. However, if a large grid box is required to accommodate (larger than 60.000 A°) then the box can be divided to two sites, otherwise the docking process will not be conducted smoothly. Final step is the mode selection window which is the one responsible for specifying the pocket shape so the docking program can recognize the inner and outer contour [39]. After conducting all of the previous steps, the protein should be prepared for executive molecular docking.

Newman et al, 2012 additionally attract attention to the significant of possessing different types synthetic approaches utilizing the natural products such as, total synthesis of natural product and semi-synthesis of compounds contain functional groups (parmacophore) mimic from natural products, which clarify the role of the nature to inspire future organic chemist to prepare novel organic compounds, that can give better or the same activity as the natural compound with less side effects [18].

Consensus score of the docking study can be calculated by various scoring functions at the final stage such as shapegauss, chemgauss3, oechemscore, screenscore and PLP. All the functional commands for OpenEye® FRED can be found in this link www.eyesopen.com/products. 3D structure of the docked ligands with protein can be visualized utilizing OpenEye® VIDA application where 3D picture of the ligand-protein complex can be taken [40].
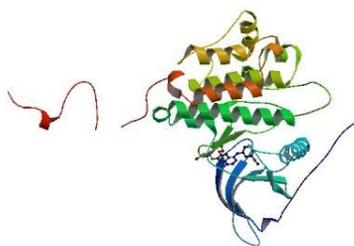
PDB ID: 1M17

Epidermal Growth Factor Receptor tyrosine kinase domain with 4-anilinoquinazoline inhibitor erlotinib.

PDB ID: 4DST
Small-molecule ligands bind to a distinct pocket in Ras and inhibit SOS-mediated nucleotide exchange
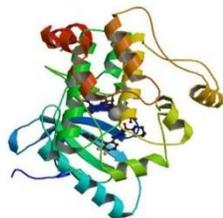
PDB ID: 3OMV

Crystal structure of c-Raf
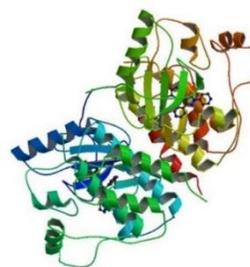
PDB ID: 4AN2

Crystal structures of human MEK1 with carboxamide-based allosteric inhibitor XL518 (GDC-0973), or related analogs.
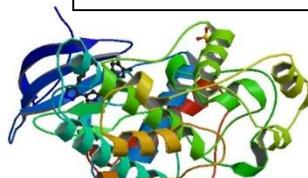
PDB ID: 2OJJ

Crystal structure of ERK2 in complex with (S)-N-(1-(3-chloro-4-fluorophenyl)-2-hydroxyethyl)-4-(4-(3-chlorophenyl)-1H-pyrazol-3-yl)-1H-pyrrole-2-carboxamide.

Figure 2.6 3-D structure of all proteins and their standards inhibitors (downloaded from PDB).

2.3   Results and Discussions:

The prepared virtual library, which include estrone modified structures, cucurbitacins and known receptors inhibitors, were docked on five different proteins starting from the upstream cellular receptor EGFR and going down to the cellular downstream proteins pathway including Ras, Raf, MEK and ERK. Novel visual compounds of the CUCUS-inspired estrone analogues were designed in systemic way by assembling the cucurbitacins pharmacophores I n the estrone scaffold. Specifically, five main positions of the estrone skeleton were modified. First, installing cucurbitacins enon side chain at C-17 with different functional groups at C-25. Second, substituting C-3 hydroxyl group with methoxy or sulfamoyl moieties. Third, modification at $\Delta^{9,11}$ position of the estrone core structure. Finally, installing various functional groups at C-25 of the enone side chain **(Fig. 2.5).**

An average of 40-55 modified estrone showed lower consensus scores (better binding affinity) toward the targeted receptors than the known inhibitors such as erlotinib, cobimetinib and sorefinib.

All of the designed virtual compounds were divided into two main groups: 1) modified estrone compounds with cucurbitacins enone side chain at C-17 of the estrone scaffold, double bond at C-16 and C-17 position, methoxy group at C-3 and different aliphatic and aromatic functional groups at C-25 **(Fig.2.7)**; 2) modified estrone analogues with cucurbitacins side chain at C-17 of the estrone skeleton structure, sulfamoyl or hydroxyl groups at C-3 and various aliphatic and aromatic functional groups at C-25 **(Fig.2.8 and Fig.2.9).**

The rational of building these different groups of estrone derivatives is to compare between different enon side chain by adding various functional groups at C-25 and their effect on the biological activities (to be discussed in Ch.3 and Ch.4).
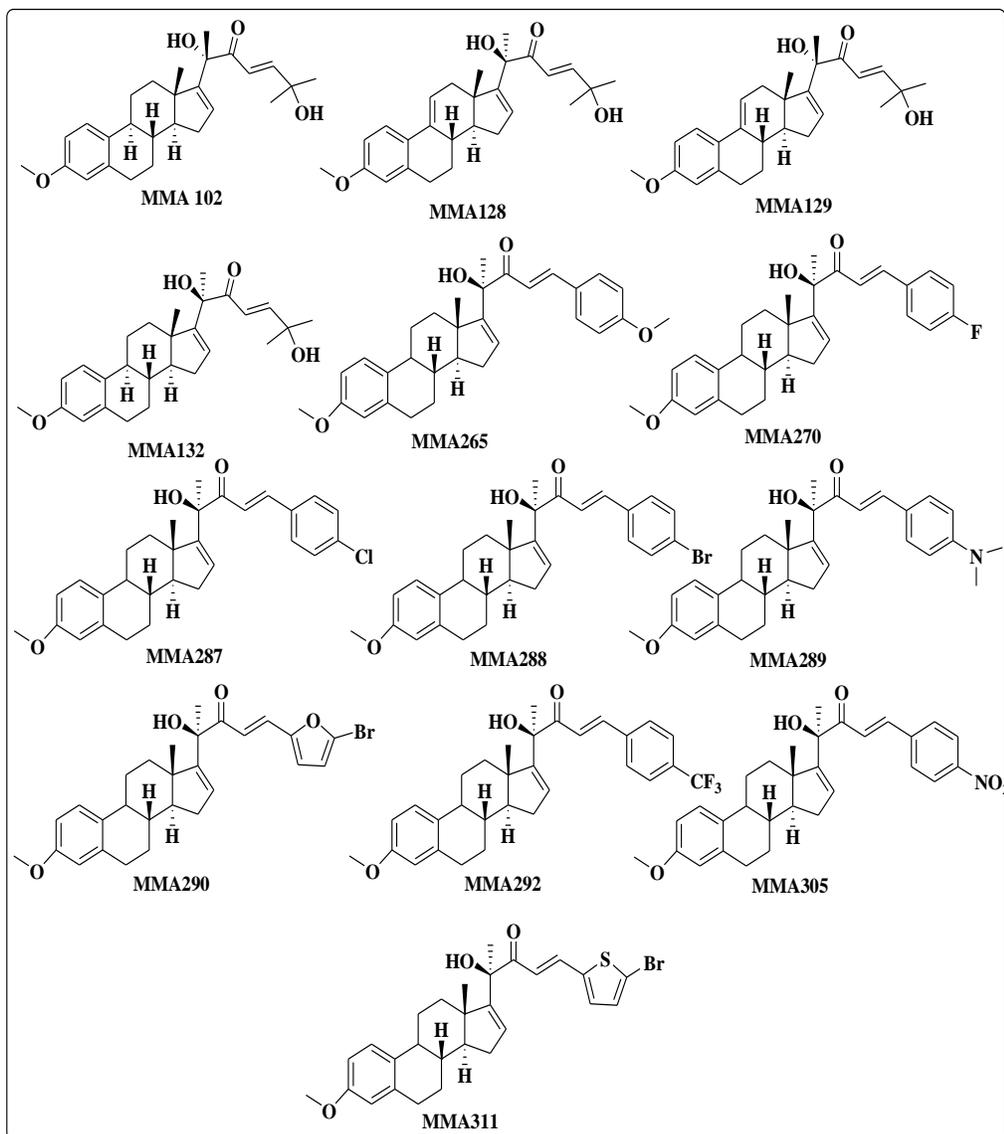
Figure 2.7 First set of synthesized MMA analogues that contain methoxy at C-3, double bond at C-16 and C-17 and various aliphatic and aromatic functional groups at C-25.
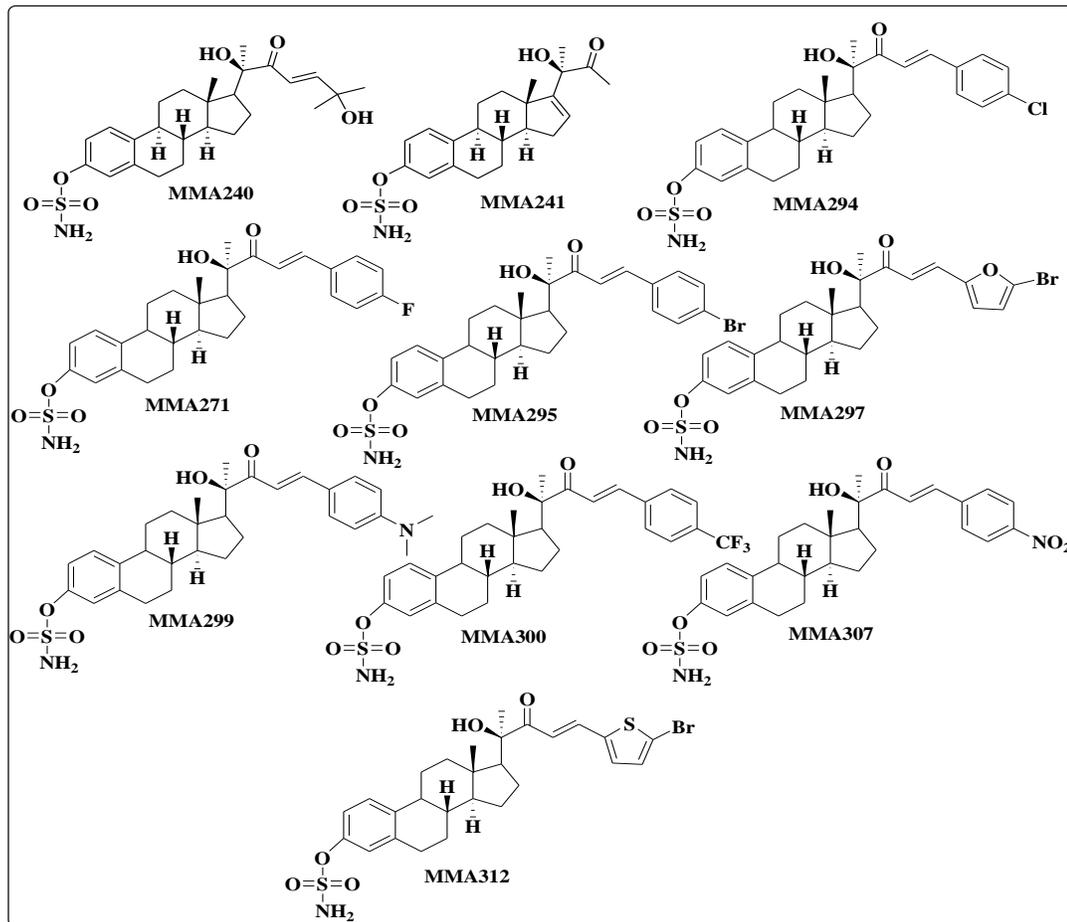
Figure 2.8 Second set of synthesized MMA analogues that contain sulfamoyl at C-3 and various aliphatic and aromatic functional groups at C-25.
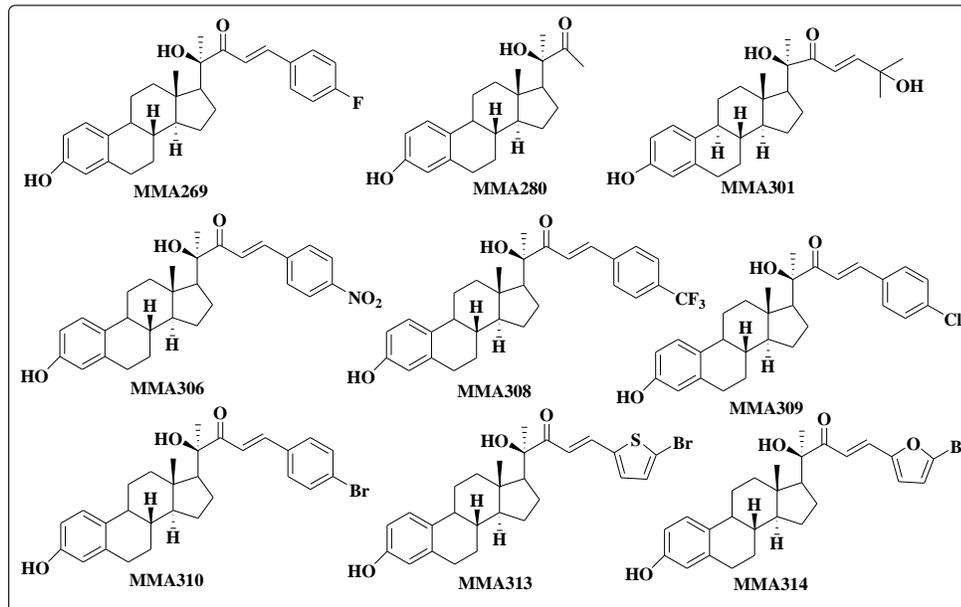
Figure 2.9 Third set of synthesized MMA analogues that contain hydroxyl at C-3 and various aliphatic and aromatic functional groups at C-25.

2.3.1 Results of the Molecular Modeling of CIEA on EGFR:

The results of molecular docking on the crystal structure of EGFR showed an outstanding binding affinity with the CUCUS-inspired estrone analogues containing various functional groups compare to the known EGFR inhibitor, Erlotinib. Modified estrone at C-17 with isopropanol enon side chain, methoxy group at C-3 and double bond at C-16 and C-17 position such as **MMA102, MMA128, MMA129** and **MMA132** demonstrated various binding mods with EGFR binding pocket. **MMA102** and **MMA132** are diastereomers to each other and showed varieties in the binding mode with the receptor. **MMA132**, which possess the stereochemistry of cucurbitacin D side chain, showed an outstanding binding mode with EGFR by forming H-bond with MET:769:A, which is same amino acids residues that erlotinib binds to in EGFR to induce anti-cancer activity by H-bonding with the same amino acid MET:769-A; also **MMA132** perform hydrophobic

interactions with the amino acids residues inside the binding pocket. While MMA102 which has the opposite stereochemistry of cucurbitacin D demonstrated less binding affinity toward the EGFR binding site only with a hydrophobic interaction mode with the EGFR binding pocket. This result proved the significant of assembling the enone side chain with the stereochemistry of cucurbitacin D **(Fig.2.10).**
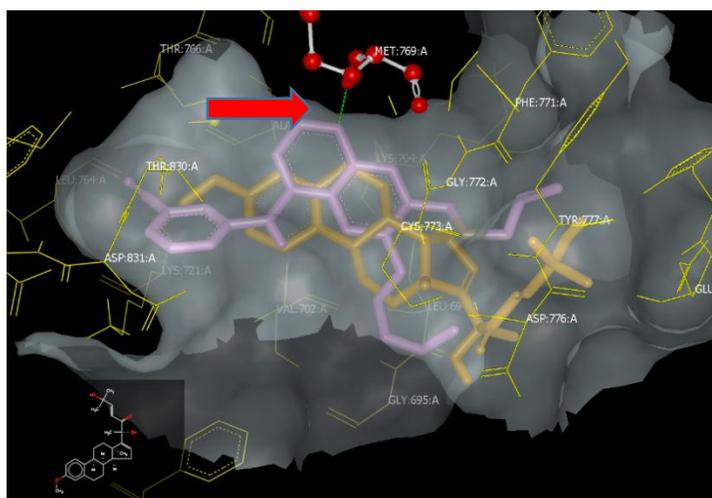
A
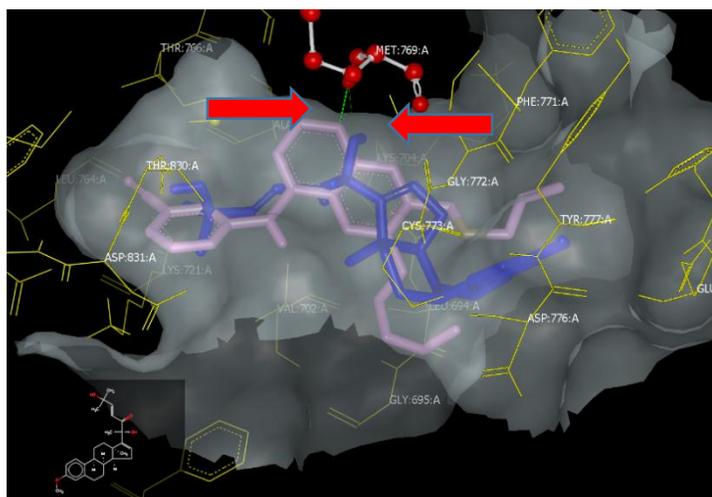
B

Figure 2.10 Visual representation of A) MMA-102 (orange) B) MMA-132 (blue) in the EGFR ATP-binding site along with Erlotinib (purple).

In addition, the presence of C-16 and C-17 double bond plays a significant role for the binding affinity which corresponds to our group publication before [14]. On the other hand, CUCUS-inspired estrone analogues with aromatic enone side chain such as **MMA265, MMA270, MMA271, MMA287, MMA288, MMA289, MMA290, MMA292, MMA305** and **MMA311** showed good binding affinity with H-bonds and hydrophobic interactions with the crystal structure of EGFR. Compounds with strong electron withdrawing groups at C-25 of the enone side chain such as **MMA290, MMA292** and **MMA31**1 showed both H-bonds and hydrophobic mode of interactions **(Fig.2.11)**; while compounds with electron donating groups at C-25 of the enone side chain such as **MMA-265** showed only hydrophobic-hydrophobic interaction with amino acid residues of the EGFR **(Fig2.12)**.
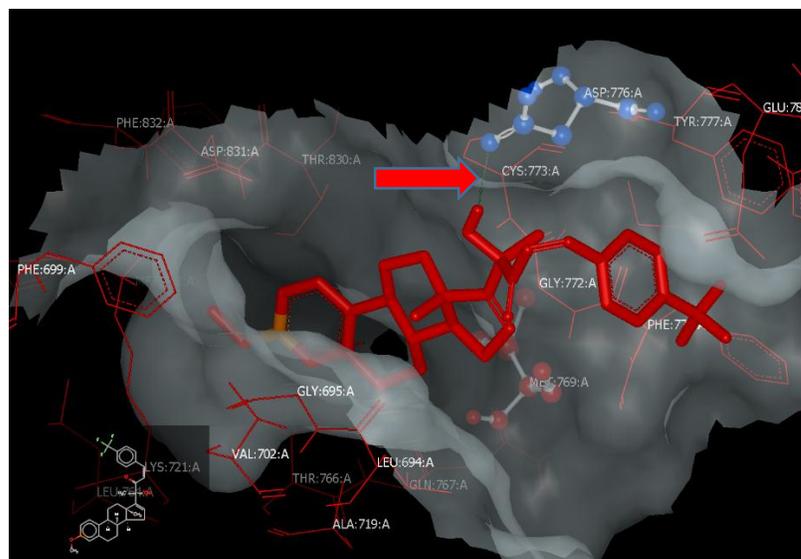
Figure 2.11 Visual representation of MMA-292 (Red) in the crystal structure of EGFR.
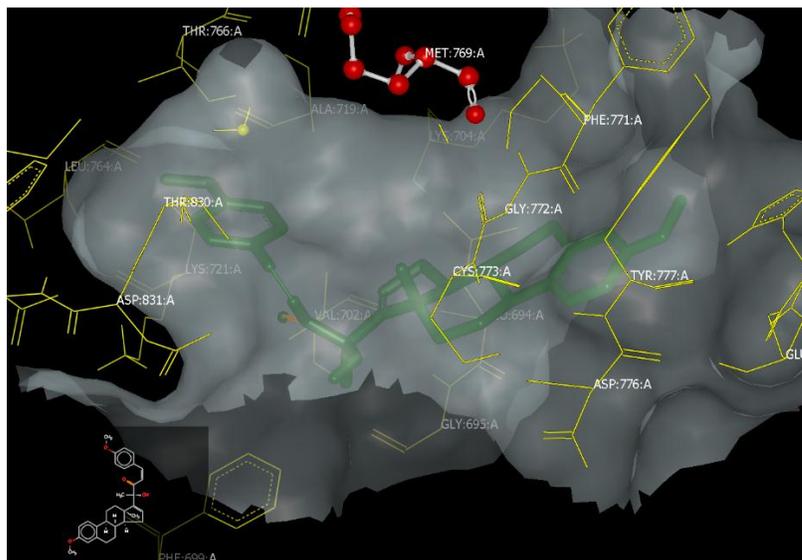
Figure 2.12 Visual representation of MMA-265 (Green) in the crystal structure of EGFR.

Second group of modified estrone that contain enone side chain at C-17, sufamoyl or hydroxyl groups at C-3 and different aliphatic and aromatic functional groups at C-25, such as **MMA-240, MMA-241, MMA-294, MMA-295, MMA-297, MMA-300, MMA-301, MMA-306, MMA-307, MMA-308, MMA-309, MMA-310, MMA-312, MMA-313** and **MMA-314**, docking calculations showed strong binding affinities with the EGFR binding site with H-bonds and hydrophobic interactions with the amino acid residues at the binding site. Compounds with different enone side chains at C-17, sulfamoyl moiety at C-3 and different functional groups at C-25 demonstrated outstanding binding affinity with the EGFR binding pocket by forming both H-bonds and hydrophobic interactions, while analogues with same functionalities at C-17 but with hydroxyl group at C-3 showed less binding affinity and lost their ability of forming H-bonds and only demonstrated hydrophobic interactions. These results revealed the significant of the presence of

sulfamoyl moiety at C-3 for the binding affinity toward the binding pocket of the 3D structure of EGFR **(Fig.2.13).**

In conclusion, CUCUS-inspired estrone analogues with cucurbitacin D enone side chain at C-17, double bond at C-16 and C-17 and methoxy group at C-3 such as MMA-132 are showing the highest binding affinity toward the crystal structure of the EGFR binding pocket with H-bonds and hydrophobic interactions. On the other side, estrone derivatives with cucurbitacins enone side chain at C-17 and sulfamoyl group at C-3 demonstrating significant binding mode comparet to the known EGFR inhibitor (Erlotinib).
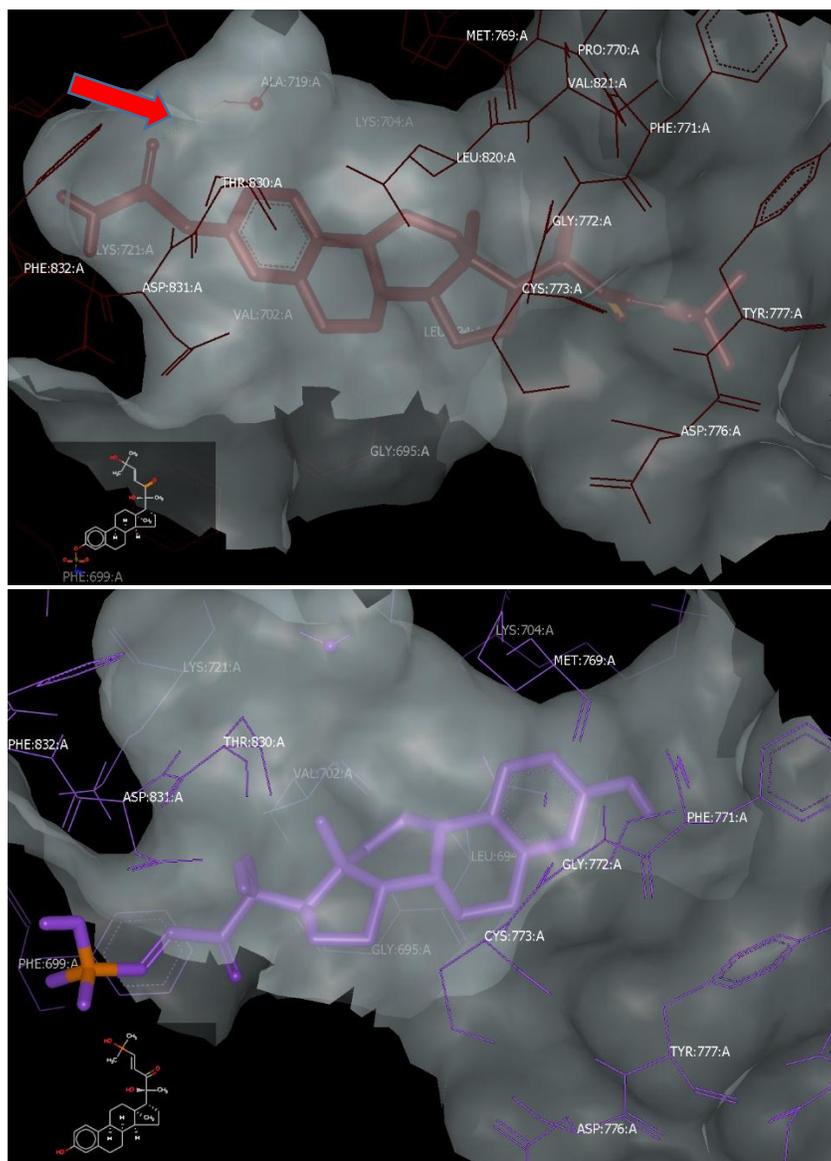
Figure 2.13 Visual representation of A) MMA-240 (Red) B) MMA-301 (purple) in the EGFR ATP-binding site.

2.3.2 Results of Molecular Docking of CIEA on Ras:

CUCUS-inspired estrone analogues with various functional groups demonstrated special binding affinity with the 3D structure binding pocket of Ras protein. In particular, estrone cmodified compounds with cucurbitacins enone side chain at C-17, double bond at C-16 and C-17, methoxy group at C-3 and various aliphatic and aromativ functional groups

at C-25 such as **MMA102, MMA128, MMA129, MMA132, MMA265, MMA270, MMA271, MMA287, MMA288, MMA289, MMA290, MMA292, MMA305** and **MMA311** showed strong binding affinities toward the targeted pocket in Ras protein. The presence of Cucurbitacin D enone side chain stereochemistry or the opposite stereochemistry in some derivatives such as **MMA102** and **MMA132** did not play role in affecting the binding affinity since both compounds demonstrated very good binding mode with Ras along with H-bons and hydrophobic interactions as shown in **figure 2.14.**
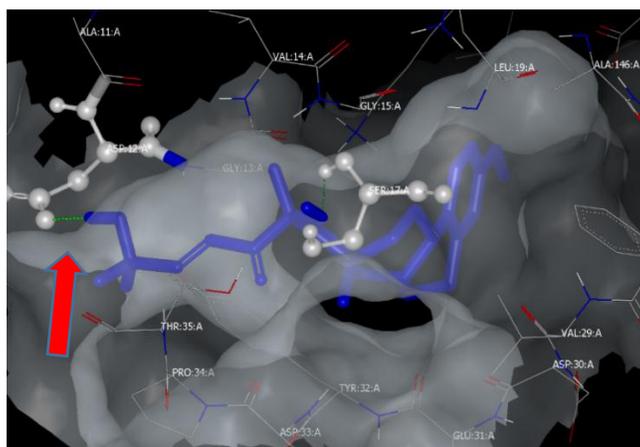
A

Figure 2.14 Visual representation of A) MMA-102 (blue) B) MMA-132 (purple) in the Ras binding site.
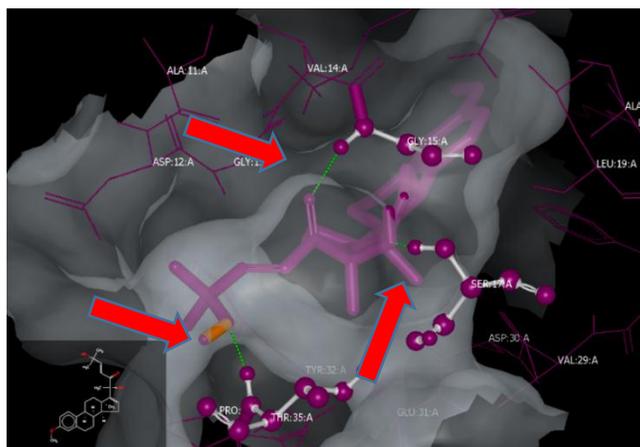
B

Estrone derivatives with aromatic functional groups at C-25 of enone side chain, methoxy at C-3 and double bond at C-16 and C-17 positions such as **MMA265, MMA270, MMA271, MMA271, MMA287, MMA288, MMA289, MMA290, MMA292, MMA305** and **MMA311** showed slight decrease in the binding affinity toward the Ras crystal structure in comparison to the aliphatic enone side chain such as **MMA-102, MMA-128, MMA129** and **MMA132** which has higher binding affinity **(Fig.2.15).**
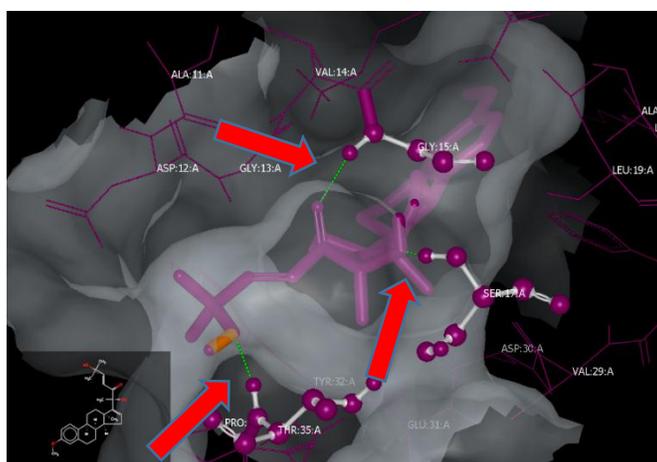
A
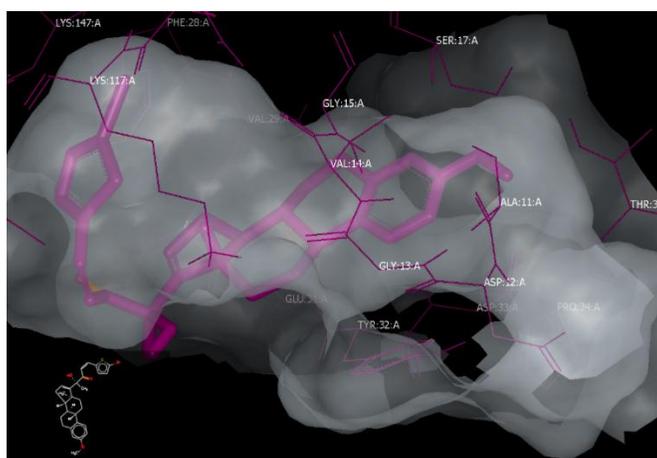
B

Figure 2.15 Visual representation of A) MMA-132 (purple) B) MMA-311 (purple) in the Ras binding site.

On the other hand, CUCUS-inspired estrone analogues with cucurbitacins enone side chain at C-17 of the estrone skeleton structure, sulfamoyl or hydroxyl groups at C-3

and various aliphatic and aromatic moieties at C-25 such as **MMA240, MMA241, MMA294, MMA295, MMA297, MMA300, MMA301, MMA306, MMA307, MMA308, MMA309, MMA310, MMA312, MMA313** and **MMA314** demonstrated lower binding affinity toward the 3D structure of Ras compare to that of first group discussed before. Even though most of these analogues don't show high binding affinity but they show both H-bonds and hydrophobic interactions with the binding pocket of Ras **(Fig.2.16),** which may demonstrate significant biological activities as Ras inhibitors.
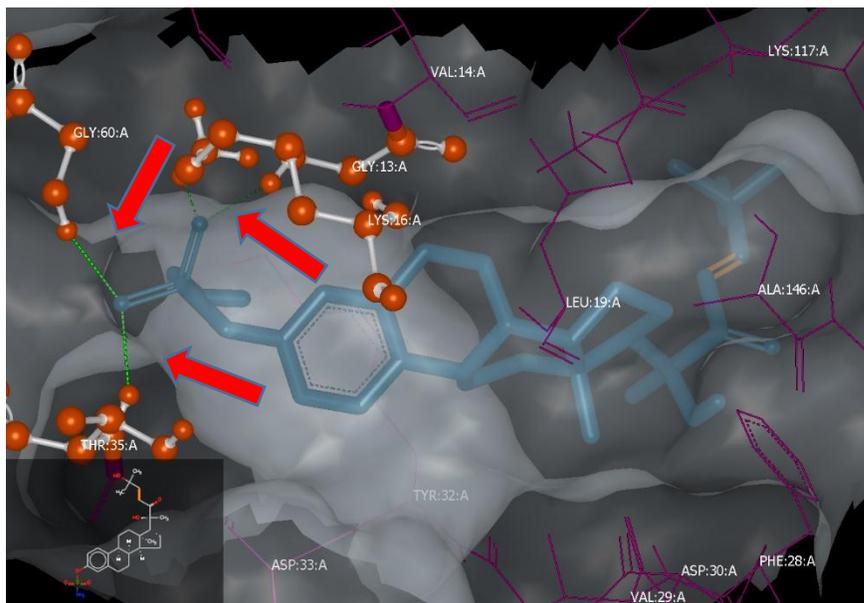
Figure 2.16 Visual representation of MMA-240 (Blue) at the binding pocket of Ras.

2.3.3 Results of Molecular Modeling of CIEA on Raf:

Molecular docking calculations of CUCUS-inspired estrone analogues with Raf protein showed varieties of binding modes. Cucurbitacins enone side chain at C-17 of estrone analogues scaffold, methoxy at C-3, double bond at C-16 and C-17 and various aliphatic and aromatic moieties at C-25 showed different binding modes such as **MMA132,**

which possess the stereochemistry of cucurbitacin D side chain, has hydrophobic interaction with Raf binding pocket, while **MMA102**, which has the opposite stereochemistry of cucurbitacin D, which show H-bonding with the same targeted protein **(Fig.2.17).**

A
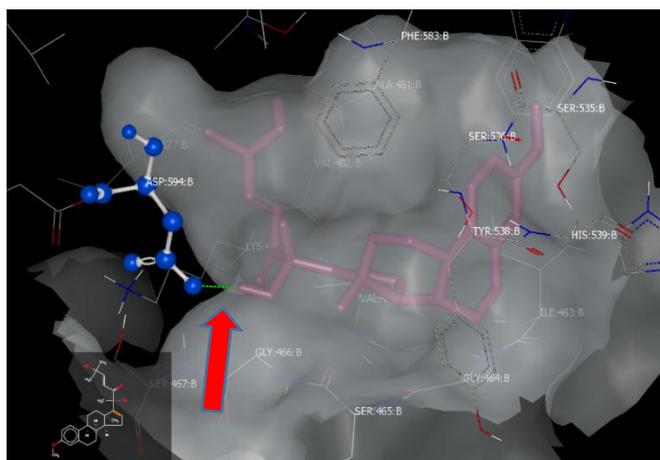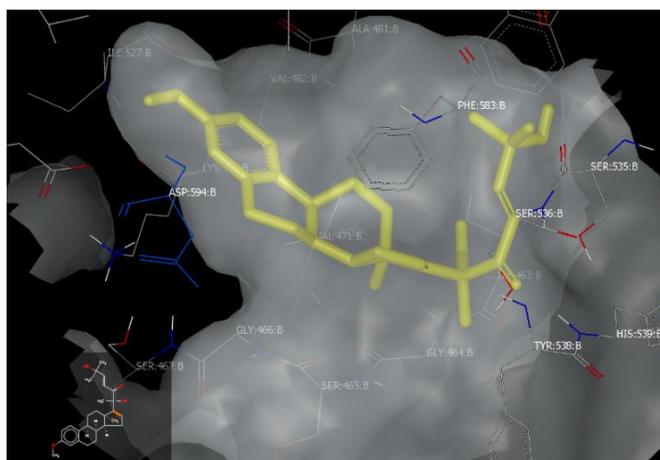
Figure 2.17 Visual representation of A) MMA-102 (purple) B) MMA-132 (yellow) in the Raf binding pocket.

B

In addition, compounds with aromatic functional groups at C-25 of the enone side chain such as **MMA265, MMA270, MMA271, MMA271, MMA287, MMA288, MMA289, MMA290, MMA292, MMA305** and **MMA311** demonstrated various binding

affinities toward the targeted protein, Raf. For example**, MMA292**, which contain P-trifluromethyl benzene at C-25 of the enone side chain, showed significant binding affinity toward the crystal structure of Raf compare to MMA-311, which has bromothiophene at C-25 of the enone side chain **(Fig.2.18)**.
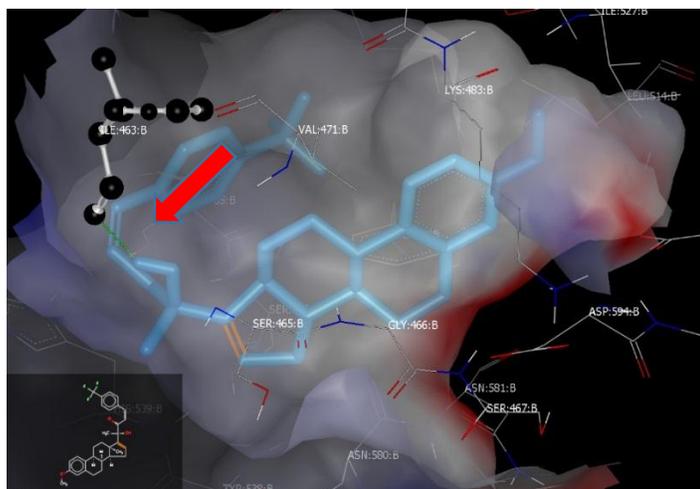
A

Figure 2.18 Visual representation of A) MMA-292 (blue) B) MMA-311 (green) in the Raf binding pocket.
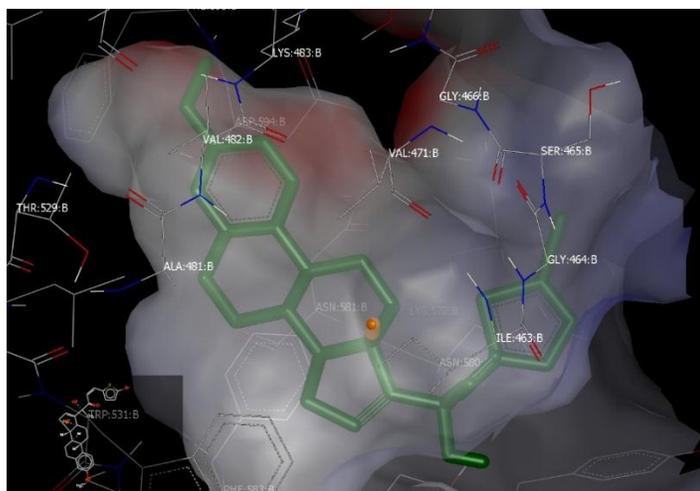
B

For the second group of CUCUS-inspired estrone analogues that contain cucurbitacins enone side chain at C-17 of the estrone skeleton structure, sulfamoyl or hydroxyl groups at C-3 and different aliphatic or aromatic groups at C-25 of the enone side

chain such as **MMA240, MMA241, MMA294, MMA295, MMA297, MMA300, MMA301, MMA306, MMA307, MMA308, MMA309, MMA310, MMA312, MMA313** and **MMA314**, the molecular docking study revealed that they demonstrated different levels of binding mods toward the targeted Raf protein. For example, compounds that possess strong electronwithdrawing aromatic functional groups at C-25 of the estrone enone side chain such as P-trifluromethyl benzene in **MMA300,** showed lower level binding affinity compare to the other compounds that have electron donating groups at the same position such as **MMA267** which has P-methoxy benzene at C-25 **(Fig.2.19)**.

A
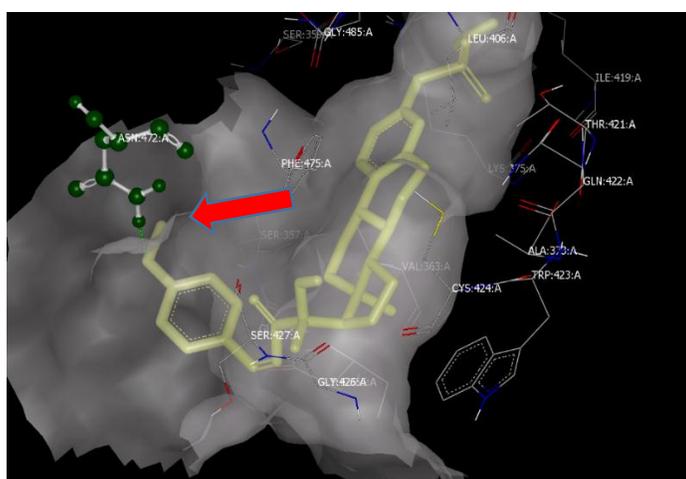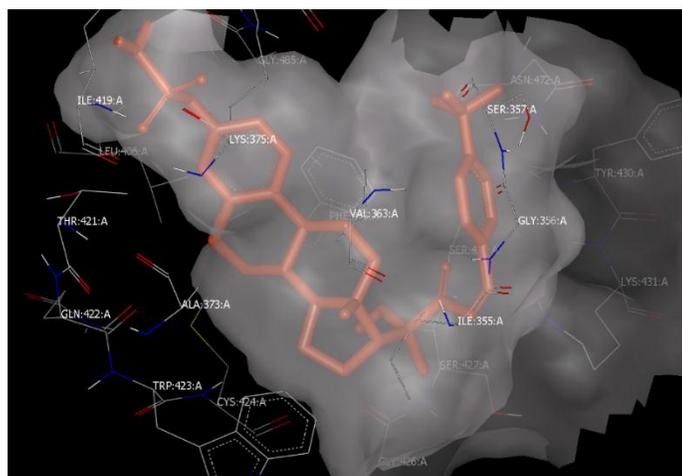
Figure 2.19 Visual representation of A) MMA-267 (yellow) B) MMA-300 (orange) in the Raf binding pocket.

B

2.3.4   Results of Molecular Docking of CIEA on MEK:

The first set of CUCUS-inspired estrone analogues that consist cucurbitacins enone side chain at C-17 of the estrone scaffold, methoxy at C-3, double bond at C-16 and C-17 and different functional groups at C-25 including **MMA102, MMA128, MMA129, MMA132, MMA265, MMA270, MMA271, MMA287, MMA288, MMA289, MMA290, MMA292, MMA305** and **MMA311** showed strong binding affinities toward MEK protein. Stereochemistry of the enone side chain play an important role in the binding affinity; for example, analogue **MMA132,** which contains the stereochemistry of cucurbitacin D side chain, showed less binding affinity toward MEK with hydrophobic interaction compare to analogue **MMA102,** which has the reversed stereochemistry of cucurbitacin D side chain, that demonstrated an outstanding binding affinity with H-bonds and hydrophobic interactions with MEK amino acid residues at the binding pocket (**Fig.2.20**).

On the other hand, the estrone analogues with C-17 enone side chain, sulfamoyl or hydroxyl groups at C-3 and number of aliphatic and aromatic functional groups at C-25 showed structural specificity correlation with binding affinity; for example, derivatives with aliphatic enone side chain showed higher binding affinities with H-bonds and hydrophobic interactions such as **MMA240** in comparison to the other derivatives that have aromatic enone side chain such as **MMA314 (Fig.2.21**).
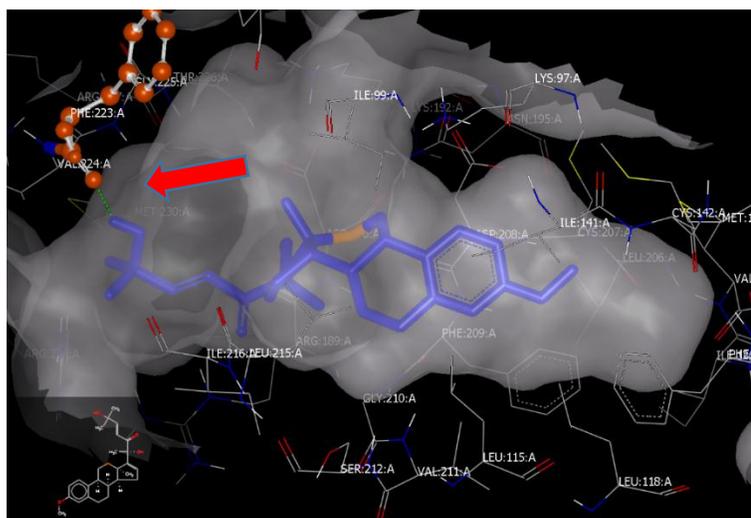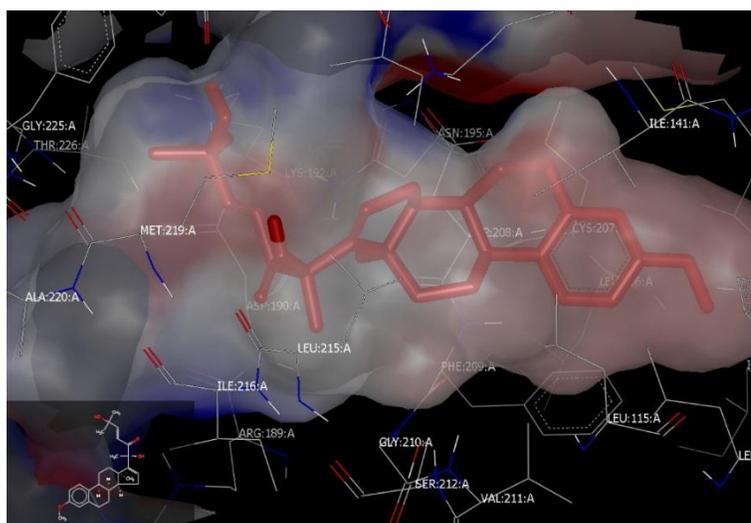
Figure 2.20 Visual representation of A) MMA-102 (blue) B) MMA-132 (red) in the MEK binding pocket.
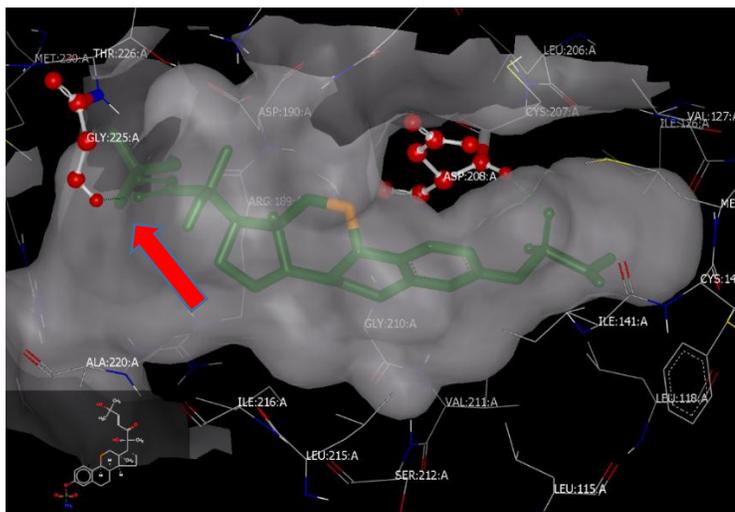
Figure 2.21 Visual representation of A) MMA-240 (green) B) MMA-314 (purple) in the MEK binding pocket.
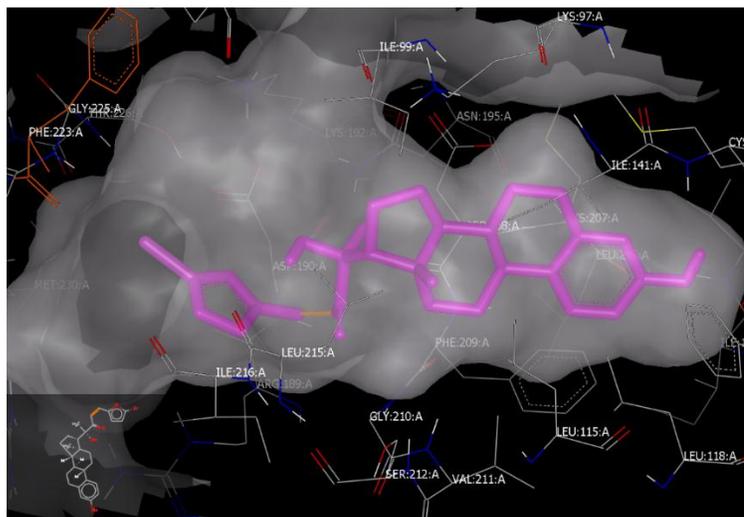
In addition, the presence of sulfamoyl moiety at C-3 of the esterone scaffold along with the enone side chain at C-17 of **MMA300** showed a significant increase for the binding affinity by forming H-bonds and hydrophobic interactions with the crystal structure of MEK binding pocket compare to the derivatives that have hydroxyl group at C-3 along with the enone side chain such as **MMA308 (Fig.2.22)**.
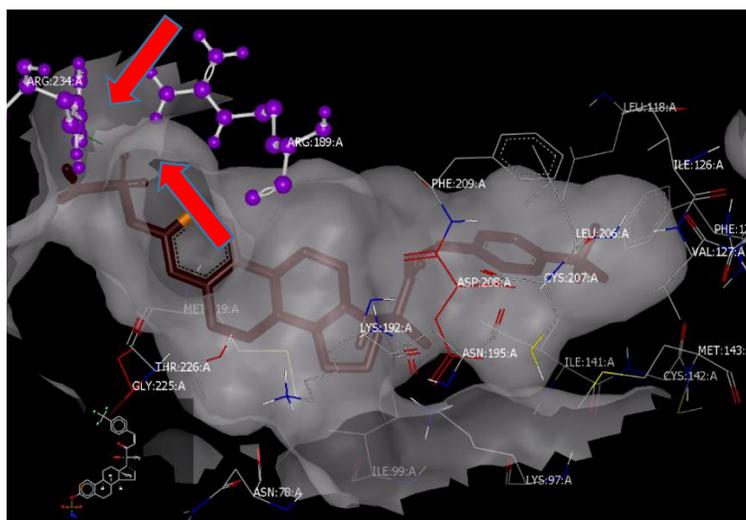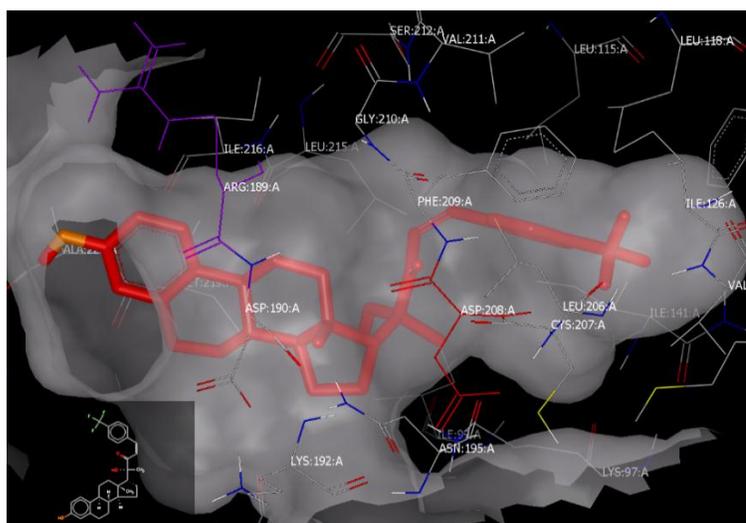
Figure 2.22 Visual representation of A) MMA-300 (brown) B) MMA-308 (red) in the MEK binding pocket.

2.3.5   Results of Molecular Docking with ERK:

The molecular docking study of CUCUS-inspired estrone analogues that contain cucurbitacins enone side chain at C-17 of the estrone scaffold, methoxy at C-3, different aliphatic and aromatic moieties at C-25 and double bond at C-16 and C-17 such as **MMA102, MMA128, MMA129, MMA132, MMA265, MMA270, MMA271, MMA287, MMA288, MMA289, MMA290, MMA292, MMA305** and **MMA311**

showed strong binding affinities toward ERK 3D structure which is consistent to their structure; for example, compound **MMA132**, which possess the stereochemistry of cucurbitacin D side chain, showed very high binding affinity toward the ERK binding site compare to analogue **MMA102**, which has the opposite stereochemistry of the side chain which proved the important of the stereochemistry choice for the structure design **(Fig.2.23)**.

A
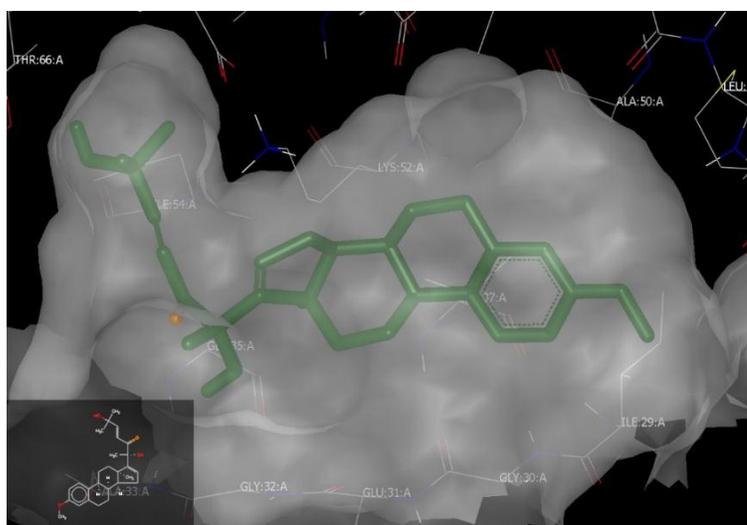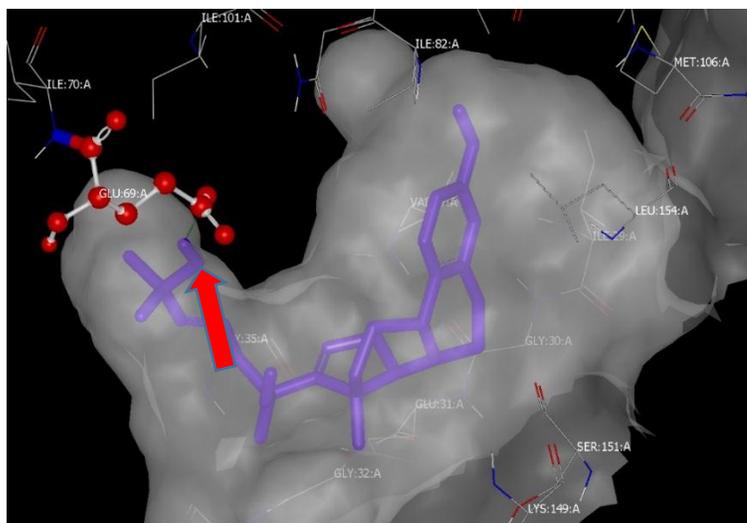
Figure 2.23 Visual representation of A) MMA-102 (green) B) MMA-132 (purple) in the ERK binding pocket.

B

Very distinguishable results of the molecular docking calculations that all of the derivatives with aromatic enone side chain substituents at C-25 form H-bonds with the same amino acid *(MET:106:A)* at the ERK binding site; while all aliphatic enone side chain derivatives at C-25 don't bind to that particular amino acid *(MET:106:A),* which may play role for their biological activities **(Fig.2.24).**
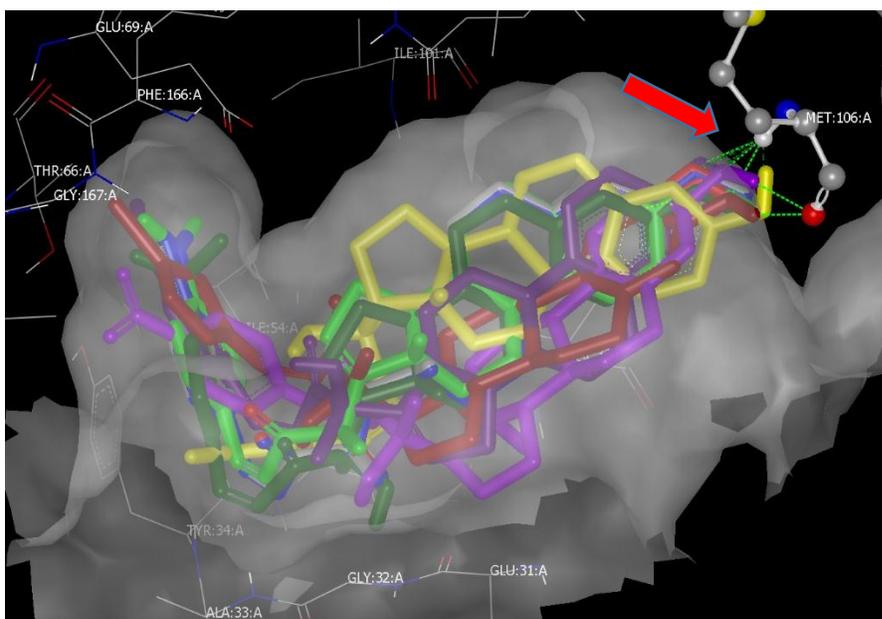
Figure 2.24 Visual representation of Various Modified estrone with aromatic functional groups at C-25.

CUCUS-inspired estrone analogues along with cucurbitacins enone side chain at C-17 of the estrone scaffold, different aliphatic and aromatic substituents at C-25 and sulfamoyl or hydroxyl groups at C-3 such as **MMA240, MMA241, MMA294, MMA295, MMA297, MMA300, MMA301, MMA306, MMA307, MMA308, MMA309, MMA310, MMA312, MMA313** and **MMA314** proved by molecular docking studies to have some distinguishable results. For example, compound **MMA240**, which contain in its

structure cucurbitacin D enone side chain at C-17 of the estrone scaffold along with sulfamoyl group at C-3 showed high binding affinity toward the ERK binding pocket compare to **MMA301**, which has hydroxyl group at C-3 instead of sulfamoyl, by forming H-bonds with *MET:106:A* and *ARG:65:A* along with hydrophobic interaction with other amino acid residues inside the binding pocket **(Fig.2.25),** which prove the importance of sulfamoyl for the binding mods.
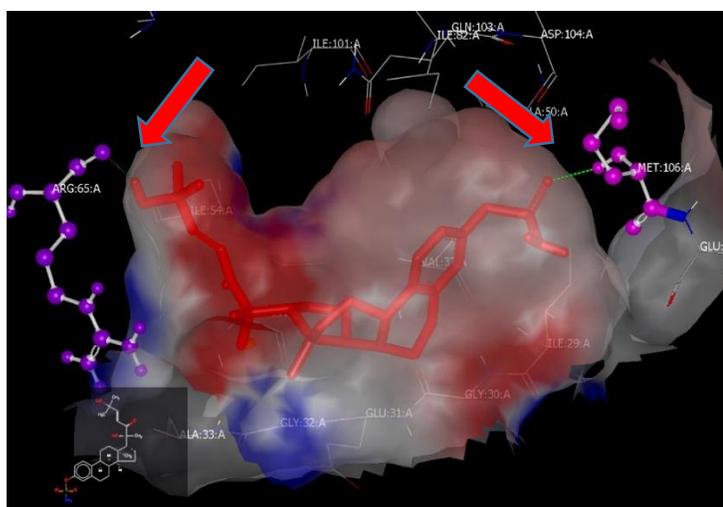
A
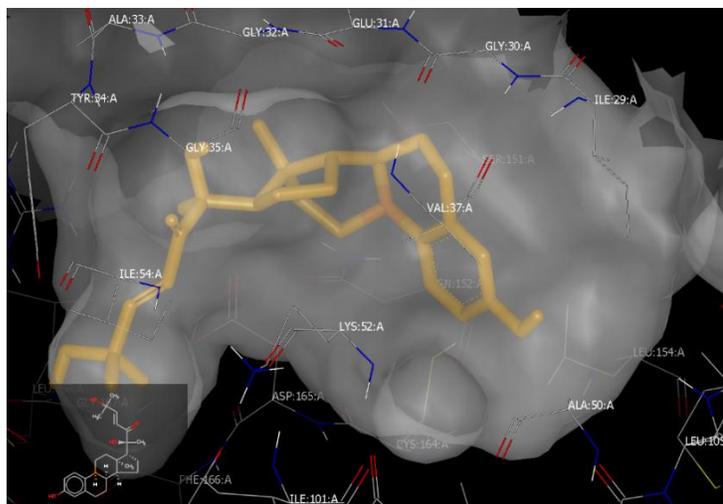
B

Figure 2.25 Visual representation of A) MMA-240 (red) B) MMA-301 (yellow) in the ERK binding pocket.

2.4   Conclusion:

Molecular modeling for the drug design process is very powerful technique for drug discovery since it can calculate how the small organic compounds will behave in the biological system before spending millions of dollars without knowing their behaviors. In our molecular docking study, we considered the pharmacophores of cucurbitacins structures to be assembled on the estrone main skeleton targeting several EGFR proteins biological targets aiming to discover drug candidates for the treatment of HCC.

The molecular docking study demonstrated the significance of CUCUS-inspired esterone analogues with cucurbitacins enone side chain at C-17 of the estrone main structure to bind to the EGFR. So, analogues with cucurbitacin D enone side chain at C-17 (with the right stereochemistry) such as **MMA132** showed significant binding affinity toward EGFR binding pocket compare to the opposite stereochemistry as in **MMA102**. In addition, CUCUS-inspired esterone analogues with cucurbitacins enone side chain at C-17 and sulfamoyl moiety at C-3 of the estrone such as **MMA240** proved by molecular docking to demonstrate significant binding modes with the 3D crystal structure of EGFR.

All of the EGFR downstream pathway proteins such as Ras, Raf, MEK and ERK were presented in this molecular docking study, which showed the importance of the presence of cucurbitacins enone side chain at C-17 for the binding affinity with the targeted

proteins. Also, it revealed the significant of other functional groups such as double bond at C-16 and C-17, sulfamoyl group at C-3 and hydroxyl group at C-3.

Therefore, the results of molecular docking calculations proved the possibility of designing drug candidates for the inhibition of EGFR and all downstream proteins (Ras, Raf, MEK and ERK) by assembling specific functional groups such as cucurbitacins enone side chain at C-17, methoxy, sulfamoyl, hydroxyl at C-3 and various aliphatic and aromatic functional groups at C-25.

These results guided our tension to build our synthetic schemes for assembling the cucrbitacins enone side chain at C-17 of the estrone main skeleton, installing different functional groups at C-3 and assembling various aliphatic and aromatic groups at C-25.

2.5   References:

1. Yar Saglam, A., et al., *Treatment with cucurbitacin B alone and in combination with gefitinib induces cell cycle inhibition and apoptosis via EGFR and JAK/STAT pathway in human colorectal cancer cell lines.* Human & experimental toxicology, 2016. **35**(5): p. 526-543.

2. Silva, I.T., et al., *In vitro and in vivo antitumor activity of a novel semisynthetic derivative of cucurbitacin B.* PloS one, 2015. **10**(2): p. e0117794.

3. Chan, K.T., et al., *Cucurbitacin B inhibits STAT3 and the Raf/MEK/ERK pathway in leukemia cell line K562.* Cancer letters, 2010. **289**(1): p. 46-52.

4. Kaushik, U., V. Aeri, and S.R. Mir, *Cucurbitacins–An insight into medicinal leads from nature.* Pharmacognosy reviews, 2015. **9**(17): p. 12.

5. Puri, R., et al., *Gastrointestinal toxicity due to bitter bottle gourd (Lagenaria siceraria)— a report of 15 cases.* Indian Journal of Gastroenterology, 2011. **30**(5): p. 233.

6. Mertins, S.D., et al., *Screening for and identification of novel agents directed at renal cell carcinoma.* Clinical cancer research, 2001. **7**(3): p. 620-633.

7. Matsuda, H., et al., *Cucurbitane-type triterpenes with anti-proliferative effects on U937 cells from an egyptian natural medicine, Bryonia cretica: structures of new triterpene glycosides, bryoniaosides A and B.* Chemical and Pharmaceutical Bulletin, 2010. **58**(5): p. 747-751.

8. Bartalis, J. and F.T. Halaweish, *In vitro and QSAR studies of cucurbitacins on HepG2 and HSC-T6 liver cell lines.* Bioorganic & medicinal chemistry, 2011. **19**(8): p. 2757-2766.

9. Jung, M.E. and R.M. Lui, *Studies toward the Total Syntheses of Cucurbitacins B and D.* The Journal of organic chemistry, 2010. **75**(21): p. 7146-7158.

10. Ryu, S.Y., et al., *Cytotoxicity of cucurbitacins in vitro.* Archives of Pharmacal Research, 1995. **18**(1): p. 60-61.

11. Lang, K.L., et al., *Synthesis and cytotoxic activity evaluation of dihydrocucurbitacin B and cucurbitacin B derivatives.* Bioorganic & medicinal chemistry, 2012. **20**(9): p. 3016-3030.

12. Csermely, P., et al., *Structure and dynamics of molecular networks: a novel paradigm of drug discovery: a comprehensive review.* Pharmacology & therapeutics, 2013. **138**(3): p. 333-408.

13. Shoichet, B.K., et al., *Lead discovery using molecular docking.* Current opinion in chemical biology, 2002. **6**(4): p. 439-446.

14. Ahmed, M.S. and F.T. Halaweish, *Cucurbitacins: potential candidates targeting mitogen-activated protein kinase pathway for treatment of melanoma.* Journal of enzyme inhibition and medicinal chemistry, 2014. **29**(2): p. 162-167.

15. Ahmed, M.S., L.C. Kopel, and F.T. Halaweish, *Structural Optimization and Biological Screening of a Steroidal Scaffold Possessing Cucurbitacin‐Like Functionalities as B‐Raf Inhibitors.* ChemMedChem, 2014. **9**(7): p. 1361-1367.

16. Rodrigues, T., et al., *Counting on natural products for drug design.* Nature chemistry, 2016. **8**(6): p. 531-541.

17. Cornil, C.A., G.F. Ball, and J. Balthazart, *The dual action of estrogen hypothesis.* Trends in neurosciences, 2015. **38**(7): p. 408-416.

18. Russo, J. and I.H. Russo, *The role of estrogen in the initiation of breast cancer.* The Journal of steroid biochemistry and molecular biology, 2006. **102**(1): p. 89-96.

19. Huisman, H., *Approaches to total synthesis of heterocyclic steroidal systems.* Angewandte Chemie International Edition in English, 1971. **10**(7): p. 450-459.

20. Ibrahim-Ouali, M., *Recent advances in oxasteroids chemistry.* Steroids, 2007. **72**(6): p. 475-508.

21. Yeung, Y.-Y., R.-J. Chein, and E. Corey, *Conversion of Torgov's synthesis of estrone into a highly enantioselective and efficient process.* Journal of the American Chemical Society, 2007. **129**(34): p. 10346-10347.

22. Saxena, H.O., et al., *Synthesis of chalcone derivatives on steroidal framework and their anticancer activities.* Steroids, 2007. **72**(13): p. 892-900.

23. Parihar, S., et al., *Gallic acid based steroidal phenstatin analogues for selective targeting of breast cancer cells through inhibiting tubulin polymerization.* Steroids, 2012. **77**(8): p. 878-886.

24. Kumar, B.S., et al., *Synthesis of 2-alkoxy and 2-benzyloxy analogues of estradiol as anti-breast cancer agents through microtubule stabilization.* European journal of medicinal chemistry, 2014. **86**: p. 740-751.

25. Leese, M.P., et al., *Structure–activity relationships of C-17 cyano-substituted estratrienes as anticancer agents.* Journal of medicinal chemistry, 2008. **51**(5): p. 1295-1308.

26. Jourdan, F., et al., *Structure–activity relationships of C-17-substituted estratriene-3-O-sulfamates as anticancer agents.* Journal of medicinal chemistry, 2011. **54**(13): p. 4863-4879.

27. Bodnár, B., et al., *Synthesis and Biological Evaluation of Triazolyl 13α-Estrone–Nucleoside Bioconjugates.* Molecules, 2016. **21**(9): p. 1212.

28. Leese, M.P., et al., *A-ring-substituted estrogen-3-O-sulfamates: potent multitargeted anticancer agents.* Journal of medicinal chemistry, 2005. **48**(16): p. 5243-5256.

29. Mowafy, S., et al., *Toward discovery of mutant EGFR inhibitors; Design, synthesis and in vitro biological evaluation of potent 4-arylamino-6-ureido and thioureido-quinazoline derivatives.* Bioorganic & medicinal chemistry, 2016. **24**(16): p. 3501-3512.

30. MendFsorsohn, J. and J. Baselga. *Epidermal growth factor receptor targeting in cancer*. in *Seminars in oncology*. 2006. Elsevier.

31. Ahire, V., et al., *Designing inhibitors for EGFR to improve anti-cancer therapy: An in silico approach.* European Journal of Biotechnology and Bioscience, 2014. **2**(5): p. 09-14.

32. Noble, M.E., J.A. Endicott, and L.N. Johnson, *Protein kinase inhibitors: insights into drug design from structure.* Science, 2004. **303**(5665): p. 1800-1805.

33. Yarmoluk, S., A.Y. Nyporko, and V. Bdzhola, *Rational design of protein kinase inhibitors.* Biopolymers and Cell, 2013. **29**(4): p. 339-347.

34. Chappell, W.H., et al., *Ras/Raf/MEK/ERK and PI3K/PTEN/Akt/mTOR inhibitors: rationale and importance to inhibiting these pathways in human health.* Oncotarget, 2011. **2**(3): p. 135-164.

35. Asati, V., D.K. Mahapatra, and S.K. Bharti, *PI3K/Akt/mTOR and Ras/Raf/MEK/ERK signaling pathways inhibitors as anticancer agents: structural and pharmacological perspectives.* European journal of medicinal chemistry, 2016. **109**: p. 314-341.

36. Dinkova- Kostova, A.T. and P. Talalay, *Direct and indirect antioxidant properties of inducers of cytoprotective proteins.* Molecular nutrition & food research, 2008. **52**(S1).

37. Talalay, P., M.J. De Long, and H.J. Prochaska, *Identification of a common chemical signal regulating the induction of enzymes that protect against chemical carcinogenesis.* Proceedings of the National Academy of Sciences, 1988. **85**(21): p. 8261-8265.

38. Jug, G., M. Anderluh, and T. Tomašič, *Comparative evaluation of several docking tools for docking small molecule ligands to DC-SIGN.* Journal of molecular modeling, 2015. **21**(6): p. 164.

39. McGann, M., *FRED and HYBRID docking performance on standardized datasets.* Journal of computer-aided molecular design, 2012. **26**(8): p. 897-906.

40. McGann, M., *FRED pose prediction and virtual screening accuracy.* Journal of chemical information and modeling, 2011. **51**(3): p. 578-596.

Chapter Three

Design, Synthesis, and Biological Screening of Novel CUCS-Inspired Estrone

Analogues (CIEA) towards Treatment of Hepatocellular Carcinoma

3.1   Introduction:

Natural products, biological compounds, total synthesis or vaccines are the main sources of small organic molecules used as drugs [1]. Since the ancient times, natural products play an important role as resource of medicines. For example, some herbes used to be chew to decrease the pain and some of them used to be wrapped around wounds to heal it. Using natural products to treat diseases and injuries known as folk medicine [2]. Lately, the large improve in developing materials to study the biological mechanism of all new chemical entities, encouraged researchers to investigate more in the pharmacological effects of natural compounds in order to clarify their synergistic impact and their clinical effects on the individual body. Natural compounds could be provide novel curative approaches toward a variety of diseases [3].

Cucurbitacins (CUCS) **(Fig. 3.1)** are natural products extracted from plant of Cucurbitacea family such as *Gonystylus keithii*, *Cayaponia tayu*, and *Citrillus colcynte*. Cucurbitacins plant family first used in folk medicine due its biological significant as anti-inflammatory agents. There are different types of cucurbitacins range as following A, B, C, D, and E, to T. Additionally, hundreds of cucurbitacins derivatives have been

synthesized and identified some of them listed in **Fig. 3.1** [4]. They have been used for treatment of different diseases such as chronic hepatitis, liver cirrhosis, jaundice, dyspepsia, inflamation and cancer [5]. Recent studies have shown a significant activities of CUCS as potential candidates for treatment of hepatocellular carcinoma (HCC) [6]. Current knowledge of molecular targets and singling pathways of different types of cancers provide a clear understanding of tumor cell regulation, which in turn paved the way to design promising potential drug candidates.

Hepatocellular carcinoma (HCC) classified as one of the highest cause of mortality and the third causing of death worldwide [7]. Asia and Africa have the highest report of HCC with high incidence between men. HCC was somewhat rare in the United States compare to the other countries; however, in the last recent decades it became one of the main cause of death in the United States. HCC is a continuous and slowly progressing disease that is generally associated with other factors such as cirrhosis, hepatitis C virus (HCV), hepatitis B virus (HBV) and toxin/ environmental disorders (obesity, diabetes and alcoholic consumption) [8]. There are different therapeutical options for HCC such as local ablation therapy, surgical resection and liver transplantation. However, these options are not applicable for late diagnostic patients. Chemotherapeutic drugs such as Erlotinib, Soreftinib are common treatments for HCC. However, drugs resistance and undesirable side effects are the most common problems associated with these drugs [9]. Therefore, there is an urgent need to find a new drug candidate to overcome these problems.

Two main complex mechanisms for HCC molecular pathogenesis; 1) mutation, which happen in some tumor suppress genes or oncogenes; 2) some diseases or metabolic disorder such as hepatitis infection, metabolic effects (such as obesity, insulin resistance, type-2 diabetes), toxin (such as alcohol) that cause tissue damage which lead to cirrhosis [10, 11]. Both of these mechanism have been connected with irregularity in different cell signaling pathways that continue the process of carcinogenic results. From a therapeutic view, all of these signaling pathways are very significant in order to treat HCC. Thus, growth factors-mediated, angiogenic signaling, epidermal growth factor (EGFR), vascular endothelial growth factor (VEGF), and hepatocyte growth factor (HGF), insulin-like growth factor (IGF), platelet-derived growth factor (PDGF), and the mitogen activated protein kinase (MAPK) pathways are the most distinguished targets for treating HCC due to their noticeable overexpression during the disease [10].

Epidermal growth factor receptor (EGFR) tyrosine kinase (TK) **(Fig. 1.5)** is one of the tyrosine kinase that have been studied as a promising target for the treatment of different carcinoma including HCC [12]. EGFR, which also known as ErbB1, is a member of family of growth factor receptors including ErbB2, ErbB3 and ErbB4. Paracrine or juxtacrine extracellular ligand binding such as epidermal growth factor (EGF) and transforming growth factor (TGF)-α stimulate the EGFR, which lead to hetro- or homo dimerization and conformational charge that activate the tyrosine kinase and allow autophosphorylation [13-15]. When the phosphorylation occurs, number of signaling pathways activated leading to cancer cell invasions, proliferation, metastasis, inhibitory of

apoptosis and angiogenesis [14-16]. Therefore, inhibition of EGFR-TK signaling cascades provides an approach for the treatment of hepatocellular carcinoma (HCC).

Cucurbitacin	R$_1$	R$_2$	R$_3$	R$_4$	R$_5$	$\Delta^{1,2}$	$\Delta^{23,24}$
A	OH	=O	MeOH	H	AC	-	+
B	OH	=O	CH$_3$	H	AC	-	+
C	H	OH	MeOH	H	AC	-	+
D	OH	=O	CH$_3$	H	OH	-	+
E	OH	=O	CH$_3$	H	AC	+	+
F	OH	OH	CH$_3$	H	OH	-	+
H	OH	=O	CH$_3$	OH	OH	-	-
I	OH	=O	CH$_3$	H	OH	+	+
J	OH	=O	CH$_3$	OH	OH	+	-

Figure 3.1 Structures of different types of Cucurbitacins and Starting Material Estrone.

CUCS demonstrated a wide range of biological activities due to their cytotoxicity on cancer cells and their potency on different biological pathways [17]. These

biological activities nominate CUCS as a potential drug that targeting multiple types of cancer. In order to avoid the undesirable adverse effects of the natural products and increase their selectivity, structural modifications to their structure may improve their efficiency [18]. Therefore, the identification of the most significant pharmacophore of the natural products is essential in order to conduct structural modifications.

Based on twenty-four different CUCS, Van Dang et. al. build some structural relationships of CUCS structures and determined the most significant pharmacophores [19]. They found that the presence of α-β-unsaturated ketone of the side chain, the free 16α-OH, and the different functionalities on C-3 of the ring A are critical for the biological activity [19, 20]. Therefore, maintaining the CUCS's side chain during the modifications is essential for the biological activity. In addition, installing different functional groups in various positions of the tetracyclic core structure of Estrone will enhance the biological activity of the synthesized compounds.

CUCS has tetracyclic moiety, they are very similar to that of steroids (**Fig. 3.1**). However, they are different from each other in the fact that C-10 methyl is located at C-9, possess a gem-dimethyl group at C-4 and the configuration of ring B and ring C [21]. Due to the similarity between the core structures of CUCS and steroids, the concept of hybrid drug design were used to install the essential pharmacophore of the CUCS into the steroid structure as promising alternative for the complicated functionalized structure of the CUCS [22, 23]. Specifically, using the estrone skeleton as a starting material to install different

functional moieties including the CUCS side chain and other functionalities has been done by Ahmed et. al [22]. Furthermore, adding various moieties to the estrone skeleton structure at C-3, such as methoxy and hydroxyl groups, beside the CUCS side chain at C-17 in the presence of double bond at C16-C17 **(Fig. 3.2)** proved to enhance the biological activity of these series of compounds. The presence of the double bond at C16-C17 changed the conformation of estrone CUCS-like compounds which improve its binding affinity towards the EGFR-TK [24]. Supporting to this hypothesis comes from the capability of estrone derivatives involved in other biological process a voiding its side effects as estrogen treatment [25-27].

To find hit compounds and to develop candidates targeting the EGFR-TK pathways, the concept of bioisosterism was utilized to systematically design and install CUCS functionality on the steroidal core structure and to build a virtual library of 900 compounds, then molecular docking was conducted for these virtual analogues against EGFR-TK [5]. The four structural positions of estrone skeleton investigated were; 1) combination of multiple substituted enone side chain at C-17; 2) Modification of C-16, C-17; 3) Functionalization on C-3 of the phenol ring and 4) Installation of aliphatic, aromatic and heterocyclic functional groups at C-25 of the enone side chain **(Fig. 3.2)**. OpenEye® Scientific software were used for the molecular modeling studies including fast exhaustive docking (FRED), Omega, and VIDA.
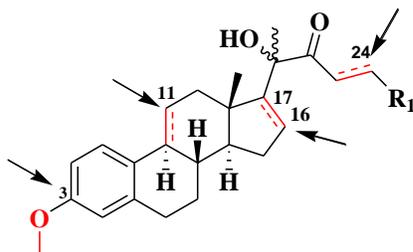
R1= C(CH$_3$)$_2$ OH, P-PhMeO, P-PhF, P-PhCl, P-PhBr, P-PhCF$_3$, P-PhNO$_2$, 5-Bromo-2-thiophene, 5-Bromo-2-furan.

Figure 3.2 Proposed modified estrone structure.

3.2   Results and Discussions:

3.2.1   Results of Molecular Docking of CIEA on EGFR:

Molecular docking data for the estrone CUCS-inspired analogues containing olefin at C-16, C-17, cucurbitacins enone side chain with α,β-unsaturated ketones functional group and methoxy group at C-3 showed an outstanding binding affinity towards the crystal structure of EGFR, when compared with the standard EGFR inhibitors such as Erlotinib, which has H-bond with **MET:796:A** and this H-bond is responsible for its biological activity along with hydrophobic interaction as anti-cancer agent **(Fig. 3.3)**. CUCS side chain at C-17 of the estrone scaffold, substitutions at C-3 and double bond at C-16 and C-17 of the CUCS-inspired analogues such as **MMA102** and **MMA132** demonstrated a very promising binding affinity by making a hydrophobic interaction as in **MMA102,** which has the opposite stereochemistry of cucurbitacin D side chain **(Fig. 3.3)** with amino acids residues of the crystal structure of **EGFR** binding pocket; while analogue **MMA132,** which possess the exact stereochemistry of the side chain of cucurbitacin D, demonstrated an outstanding binding affinity through hydrophobic interaction with amino acids residues

of the EGFR along H-bond with **MET:796:A**, which is the same amino acid that Erlotinib make H-bond with and responsible for its anti-cancer activity **(Fig. 3.3)**. These results showed the significant of possessing the cucurbitacin D stereochemistry for the enone side chain at C-17 of the estrone scoffold, which make analogue **MMA132** a very promising candidate to be synthesized and biologically tested to confirm its biological activity as an anti-cancer candidate.

On the other hand, CUCUS-inspired estrone analogues with aromatic and heterocyclic enone side chain that possess the stereochemistry of cucurbitacin D side chain such as **MMA265, MMA270, MMA287, MMA288, MMA289, MMA290, MMA292, MMA305** and **MMA311 (Fig. 3.2)** showed promising binding affinity with H-bonds and hydrophobic interactions with the binding site of the crystal structure of EGFR. Compounds with strong electron withdrawing groups at C-25 of the enone side chain such as **MMA290, MMA292** and **MMA311** showed both H-bonds and hydrophobic modes of interactions **(Fig. 3.4)**; while molecules with electron donating groups at C-25 of the enone side chain such as **MMA265** showed only hydrophobic-hydrophobic interaction with amino acid residues of the EGFR crystal structure **(Fig. 3.5)**. CIEA possess cucurbitacin aromatic enon side chain with the opposite stereochemistry of cucurbitacin D such as **MMA316, MMA317, MMA318, MMA319, MMA320, MMA321, MMA330** and **MMA334 (Fig. 3.2)** demonstrated similar binding affinities to the estrone derivatives that have similar aromatic functional groups at the enone side chain but with stereochemistry

of the cucurbitacin D. All of the top consensus score compounds will be chemically synthesized in order to be evaluated biologically as anti-cancer candidates.
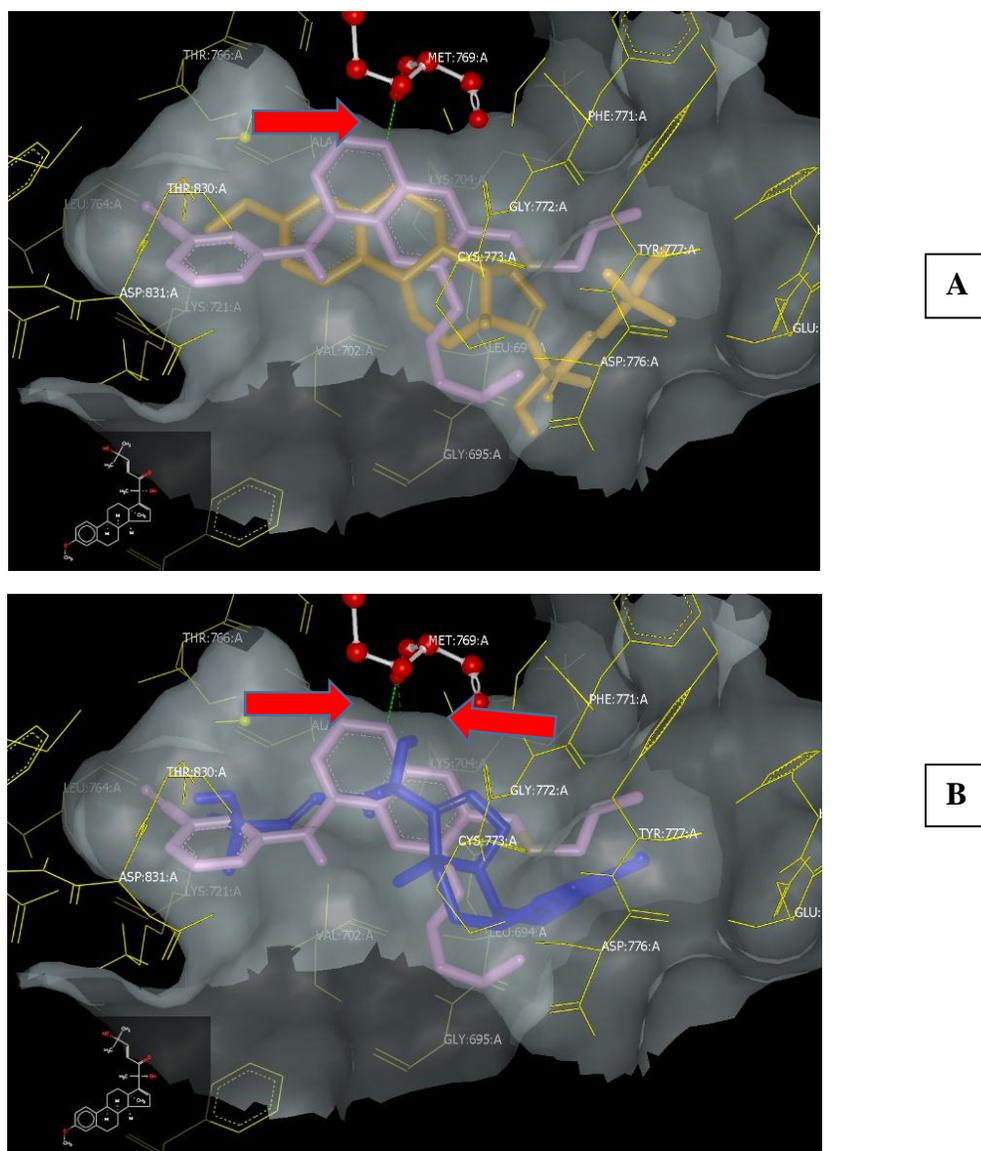
Figure 3.3 Visual representation of A) MMA-102 (orange) B) MMA-132 (blue) in the EGFR ATP-binding site along with Erlotinib (purple).
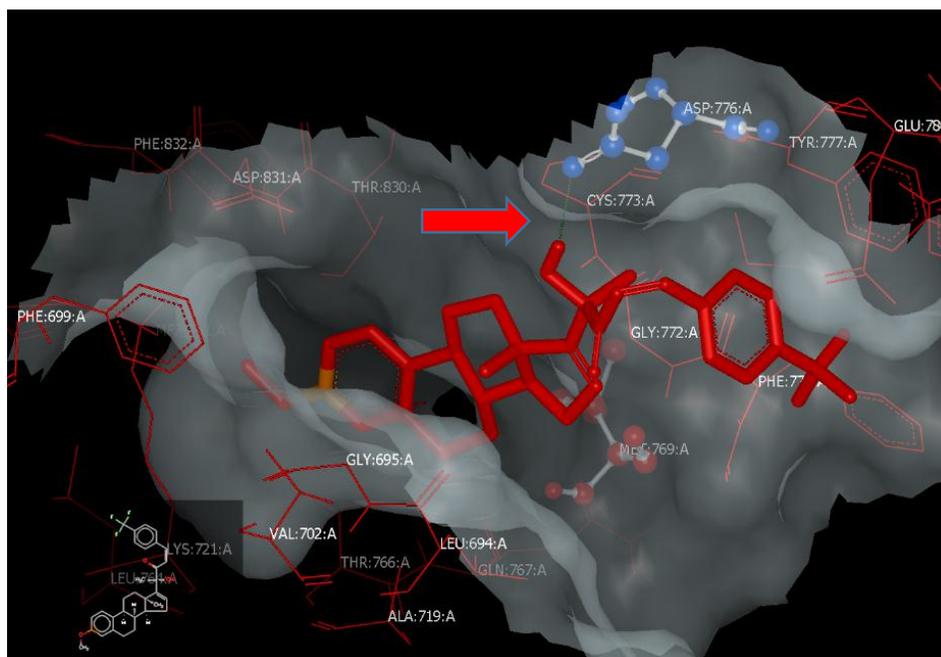
Figure 3.4 Visual representation of MMA-292 (Red) in the crystal structure of EGFR.
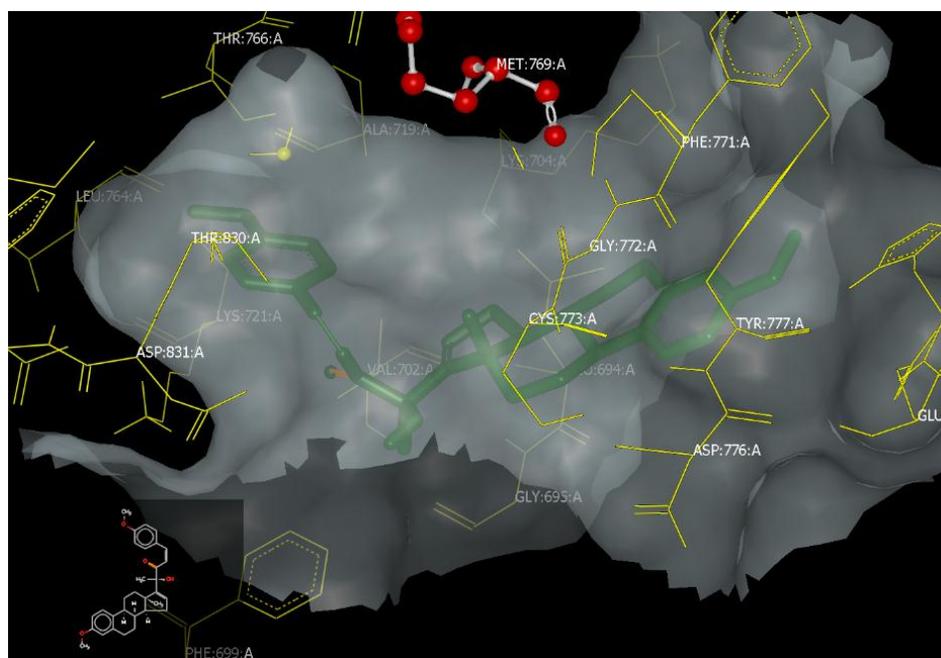
Figure 3.5 Visual representation of MMA-265 (Green) in the crystal structure of EGFR.

3.2.2   Synthesis of MMA Analogues:

To explore these novel CUCS-inspired estrone analogues (CIEA), installation of cucurbitacins α,β-unsaturated ketone side chain with the double bond in C-16, C-17 position into the estrone skeleton structure were synthesized at C-17 **(Fig. 3.1)**. Number of synthetic approaches were investigated in order to accomplish this task. First approach was started with installing p-toluenesulfonyl hydrazide in C-17 of the estrone structure followed by Shipiro reaction **(scheme 3.1)** and adding two equivalents of strong base such as n-butyllithium to form nucleophilic center followed by the addition of various electrophiles such as paraformaldhehyde, dimethyl formamide, N-methoxy-N-methylacetamide and 2,3-butanedione [28, 29] as shown in **scheme 3.2**. This approach was used to save steps from the original route, which will be discussed in the next paragraph, but unfortunately, several electrophiles gave a very small yield which make this approach not convenient to start with and built the whole scheme in **(scheme 3.1)**.
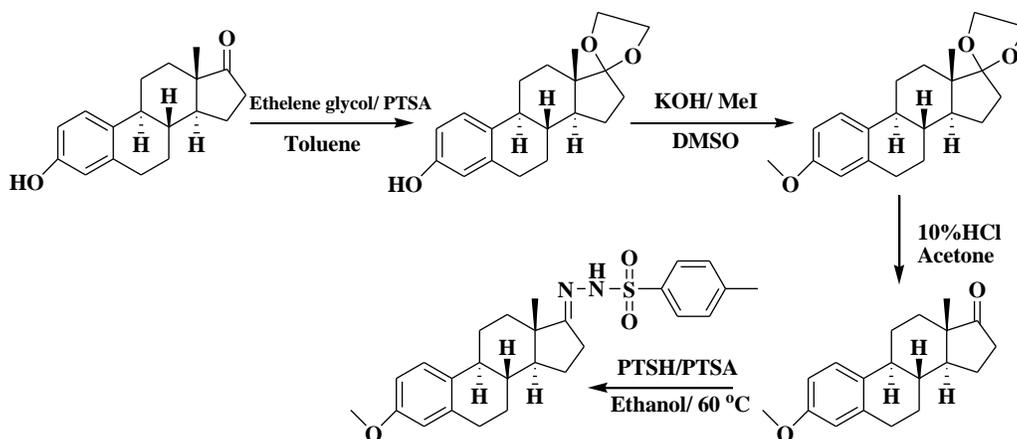
Scheme 3.1 The approach of Shapiro reaction.
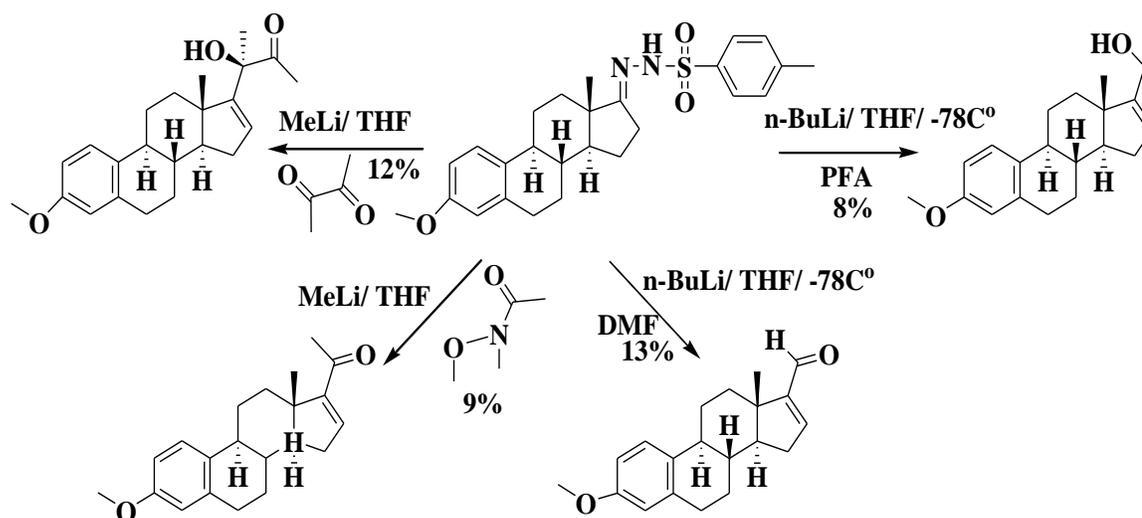
Scheme 3. 2 Shapiro reaction followed by nucleophilic reactions with various electrophiles.

Then synthesis of **MMA** analogues started again with protection of ketone at C-17 of commercially available estrone (**Fig. 3.1**) with ketal group in order to methylate hydroxyl group at C-3 using methyl iodide in presence of potassium hydroxide to provide compound **2** [30, 31], followed by deprotection of the ketal group under acidic condition to produce methoxy estrone **3** [31]. This followed by addition of TMSCN in presence of ZnI$_2$ was added to **3** to form cyanohydrin intermediate **4,** followed by elimination reaction using POCl$_3$ and DBU as a base to obtain double bond in C-16, C-17 position **5** [32]. Nucleophilic substitutions reaction of nitrile group using methyl lithium was used to provide methyl ketone **6** [33]. Adding catalytic zinc iodide followed by TMSCN to ketone **6** to form (**1:1**) diasteriomers mixture of cyanohydrin **7** [22]. It was very challenging to separate the diasteriomers mixture of **7** using silica gel chromatography which have the same RF on the TLC plate. Then by treatment of diasteriomers mixture of cyanohydrin **7**

with methyl lithium resulted in formation of required diasteriomers mixture (1:1) of hydroxyl methyl ketone **8 (MMA225)** and **8`** [22]. Column chromatography was used to separate the diasteriomers mixture **8** and **8`** to obtain pure seperated **8** and **8` (Scheme 3.3)**. The configuration of **8 (MMA225)** were confirmed using X-ray crystallography **(Fig. 3.5)**. At this stage of the synthesis, Aldol condensation reaction used to install aliphatic and aromatic aldehydes into hydroxyl ketone **8** and **8`**, with in situ elimination in order to provide the required enone. In attempting to install CUCS-side chain, a three synthetic reactions steps were used to prepare aldehyde **14** which started by the protection hydroxyl group of methyl 2-hydroxyisobutyrate **11** using tert-butyldimtylsillyl chloride (TBSCl). The synthesized ester **12** was treated by reducing agent diisobutylaluminum hydride (DIBAL-H) to form alcohol **13**. Control oxidizing agent tetrapropyl ammonium perrutherate (TPAP) and N-methylmorphline N-oxide (NMO) were utilized to prepare the desired aldehyde **14** [5] **(Scheme 3.4)** . Aldol optimized reaction condition of aldehyde **14** and hydroxyl ketone **8** and **8`** separately in presence of lithium diisopropylamine (LDA) at -78°C to room temperature to give enone **9** and **9`**. Final compounds **MMA102 and MMA132** were obtained by the addition of tetrabutyl ammonium fluoride (TBAF) in THF **(schemes 3.5 and 3.6)** [5].
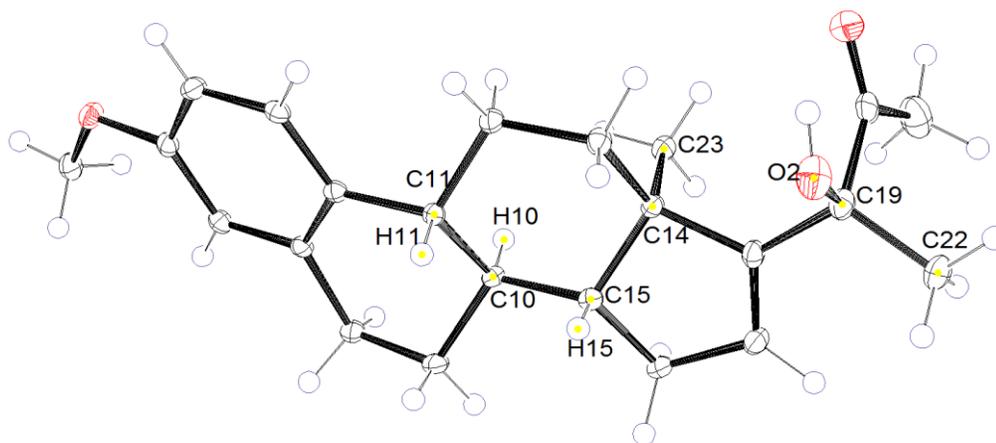
Figure 3.6 X-ray crystal of compound **8** which confirm the right stereochemistry to install Cucurbitacin D enone side chain.

Further derivativesation into hydroxyl ketone **8** started with the assembly of $\Delta^{9,11}$ olefin at the B/C ring ligament to gain the pseudo-*cis* configuration by adding 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) in methanol resulted in $\Delta^{9,11}$ hydroxyl ketone **15** and **15`**. The produced intermediate $\Delta^{9,11}$ hydroxyl ketone **15** and **15`** processed through the same Aldol condensation reaction with aldehyde **14** to produce in $\Delta^{9,11}$ enone **16** followed by the deprotection of TBSCl using TBAF to obtain $\Delta^{9,11}$ enone **MMA128** **(scheme 3.7)** [5].

Different aromatic and heterocyclic aldehydes have been reacted with hydroxyl ketone **8** and **8`** to produce different aromatic enone side chains that have two different stereochemistry of each analogue such as para-methoxybenzene **(MMA265** and **MMA333)**, para-flurobenzene **(MMA279** and **MMA334),** para-chlorobenzene **(MMA287** and **MMA316),** para-bromobenzene **(MMA288** and **MMA319),** para-

dimethylaminobenzene (**MMA289** and **MMA317**), 5-bromofuran (**MMA290** and **MMA318**), para-trifluromethylbenzene (**MMA292** and **MMA320**), para-nitrobenzene (**MMA305** and **MMA321**) and 5-bromothiophene (**MMA311** and **MMA330**) (**scheme 3.8**). All aromatic α,β-unsaturated enone side chain increased the hydrophobicity of the CUCS-inspired analogues so, it can be compared biologically with more polar first synthesized set of compounds **MMA102, MMA128,** and **MMA132**.
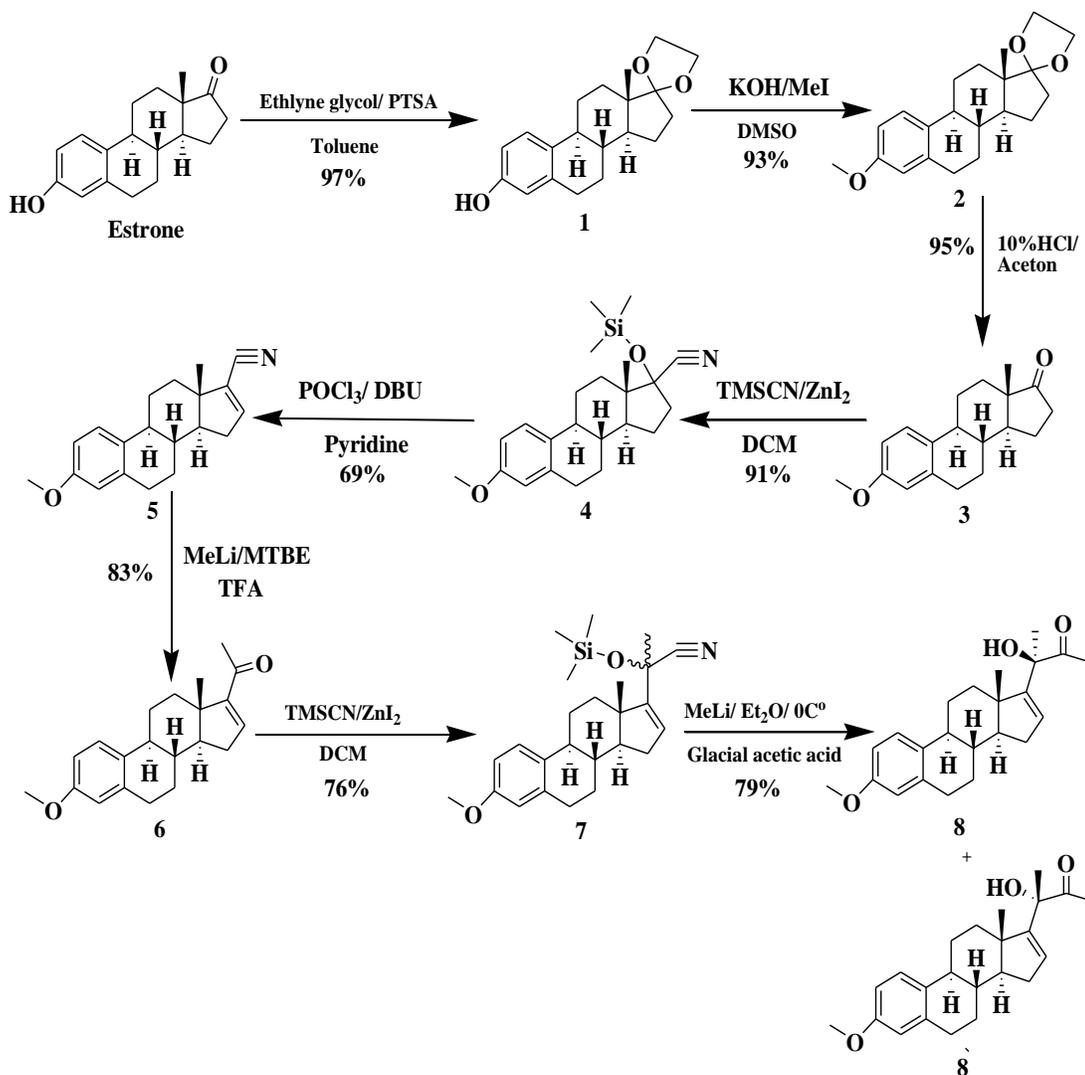
Scheme 3.3 Synthesis of Disteriomers 8 and 8`.
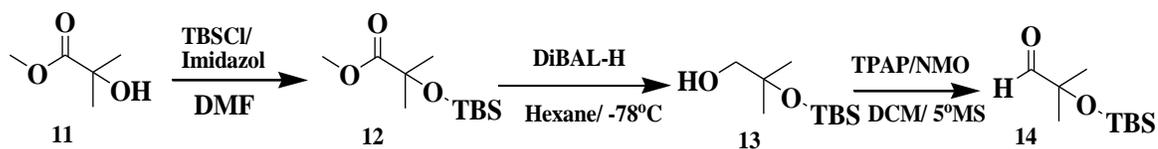
Scheme 3.4 Synthesis of Aldehyde 14.
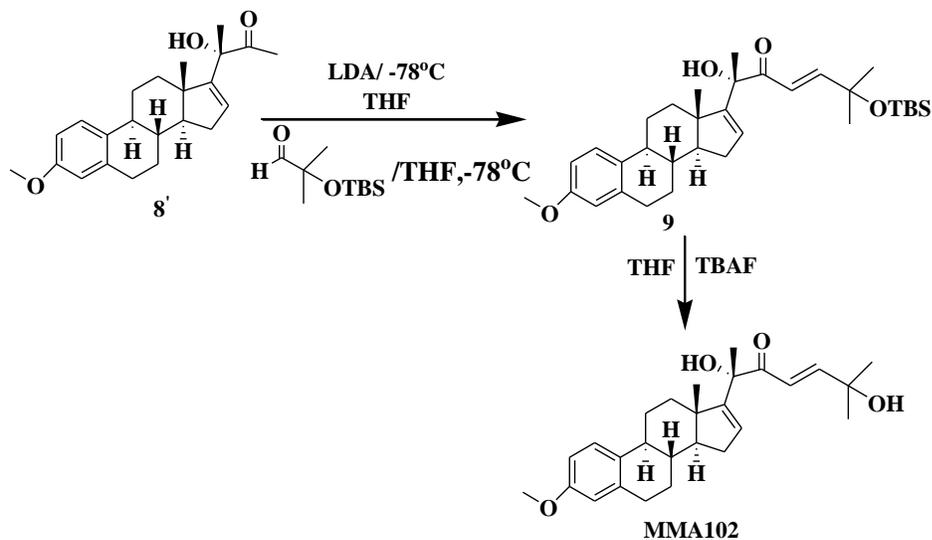
Scheme 3.5 Synthesis of MMA102.
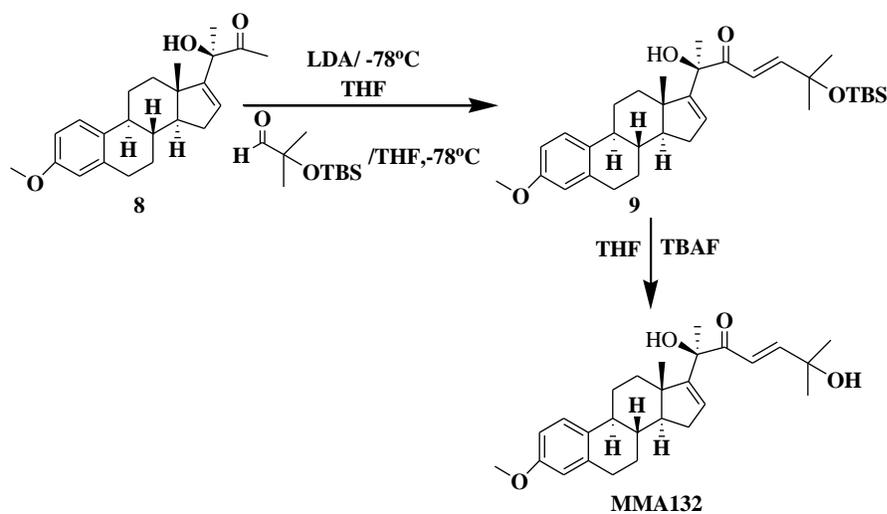
Scheme 3.6 Synthesis of MMA132.
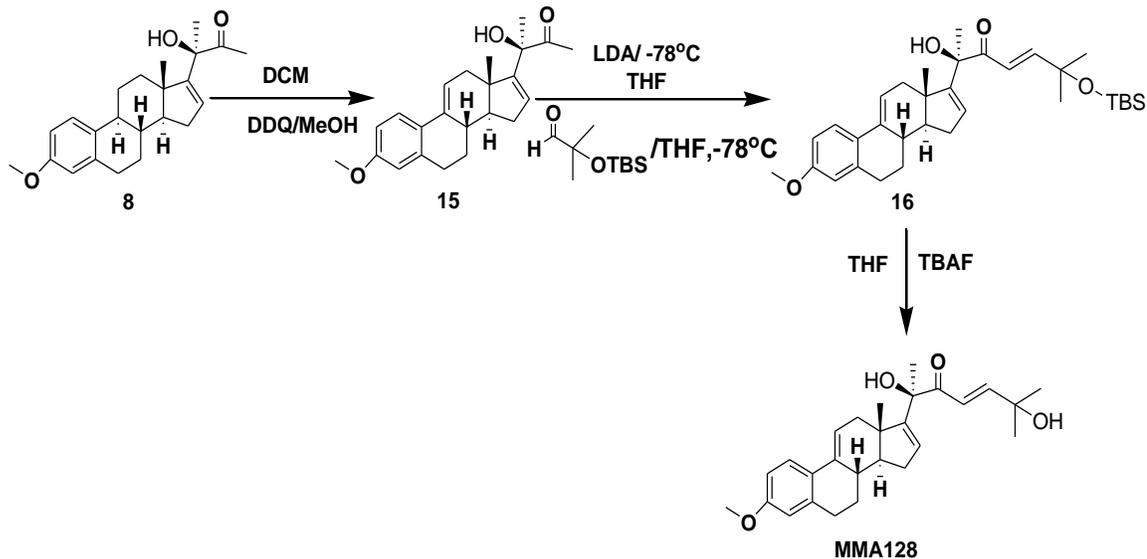
Scheme 3.7 Synthesis of MMA128.
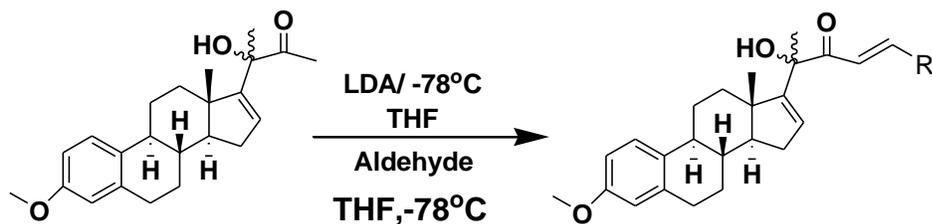
R= P-PhMeO, P-PhF, P-PhCl, P-PhBr, P-N (CH$_3$)$_2$Ph, P-PhCF$_3$, P-PhNO$_2$, 5-Bromo-2-

thiophene, 5-Bromo-2-furan.

Scheme 3.8 Synthesis of various aromatic enone side chains.

3.2.3 Biological Evaluations of the CIEA for the Treatment of HCC :

In vitro biological evaluation of CIEA was conducted to study their ability to target

and inhibit EGFR-TK as anti-cancer candidates. First, MTT cell viability assay was used

to measure the cytotoxicity and anti-proliferative activity of the synthesized compounds.

Analogues that contain aromatic enone side chains such as **MMA265, MMA270, MMA287, MMA288, MMA292, MMA305, MMA316, MMA319, MMA320, MMA321, MMA333 and MMA334** showed strong cytotoxicity with IC_{50} values $16\ \mu M$, $7\ \mu M$, $25\ \mu M$, $32 \mu M$, $2\ \mu M$, $29 \mu M$, *NA*, NA, $3\ \mu M$, $20 \mu M$, $2\ \mu M$ and $8\ \mu M$; respectively toward hepatocellular carcinoma cell line (HepG2) in comparison to the known EGFR inhibitor, Erlotinib, which has IC_{50} value of $25\ \mu M$ **(Table 3.1)**. While CIEA that contain heterocyclic enone side chains such as **MMA290, MMA311, MMA318** and **MMA330** showed variety of cytotoxicity on HepG2 cell line based on their stereoisomer; for example, compound **MMA290** and **MMA311**, which possess cucurbitacin D side chain stereoisomer, showed potent cytotoxicity with IC_{50} values $6\ \mu M$ and $0.7\ \mu M$, respectively. On the hand, the same compounds but with the opposite stereoisomer such as **MMA318** and **MMA330** showed significant change in their cytotoxicity since analogue **MMA318** lost its cytotoxicity completely and compound **MMA330** showed IC_{50} of $11 \mu M$ compare to $0.7\ \mu M$ in **MMA311**. These results demonstrated the importance of possessing the exact stereochemistry of cucurbitacin D enone side chain for antiproliferation activity of these isomers. Meanwhile, compounds **MMA102, and MMA1320,** which contain isopropanol enone side chains of cucurbitacin D, demonstrated outstanding cytotoxicity with IC_{50} values $3 \mu M$, and $2 \mu M$, respectively **(Table 3.1)**. All of the previous IC_{50} results were compared to Erlotinib (current chemotherapeutic drug for treatment of HCC), which has an IC_{50} value of $25\ \mu M$ **(Table 3.1)**. Second, for more and deep understanding of the anti-proliferation mechanism of the potent analogues, **MMA102** and **MMA132**, western blot analysis was conducted **(Fig. 3.8)**. The western blot data indicate the potential of compound

MMA132 to bind to the upstream signaling pathway through the inhibition of EGFR-TK

and inhibit its phosphorylation at $2 \, \mu M$ after 48 hours incubation compare to the loading

control (Actin) and Erlotinib as positive control as shown in Figure 3.8. Third, cell cycle

arrest experiment using flowcytometry was conducted on **MMA132** at various

concentration (1/2 IC$_{50}$, IC$_{50}$, 2 IC$_{50}$), which showed its ability to induce G1/S phases cell

cycle arrest as shown in **Figure 3.9**. The promising data of compound **MMA132**

demonstrated the importance of the presence of the CUCS-side chain at C-17, double

bound in C-16, C-17 and C-3 methoxy group. Docking results showed that compound

MMA132 has H-bond with **MET:796:A (Fig. 3.3),** the same as the standard Erlotinib.

This observation in addition to biological results demonstrated the importance of

possessing the CUCS-side chain configuration to gain a promising binding affinity inside

the pocket and inhibit the targeted receptor, EGFR. On the other hands, compounds that

contain aromatic and heterocyclic enone side chains showed a wide range of cytotoxicity

toward HepG2 cell line. The wide range of cytotoxicity can be attributed to several factors

including stereochemistry of the enone side chain, types of the pharmacophores at enone

side chain and the electron withdrawing or donating functional groups of the enone side

chain that effect the electrophilicity of Michael acceptor at enone side chain.

Table 3.1 MTT cell viability results in HepG2 cell line.

Compound	IC$_{50}$ (μM)	Compound	IC$_{50}$ (μM)
Erlotinib	25 μM	MMA305	29 μM
MMA102	3 μM	MMA311	0.7 μM
MMA132	2 μM	MMA316	NA
MMA225	25 μM	MMA318	NA
MMA265	16 μM	MMA319	NA
MMA270	7 μM	MMA320	3 μM
MMA287	25 μM	MMA321	20 μM
MMA288	32 μM	MMA330	11 μM
MMA290	6 μM	MMA333	2 μM
MMA292	2 μM	MMA334	8 μM

Figure 3.7 Chart represent the ability of CIEA to inhibit the growth of HepG2 cell line.

(Hepatocellular Carcinoma cell line).
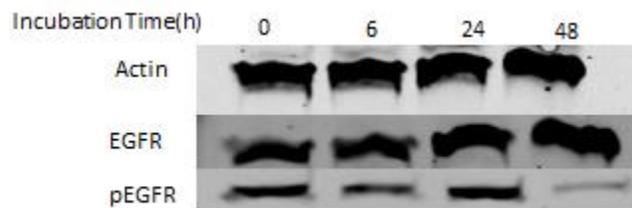
Figure 3.8 Analogue MMA132 manage to inhibit p-EGFR after 48h incubation at IC50 3 μM.
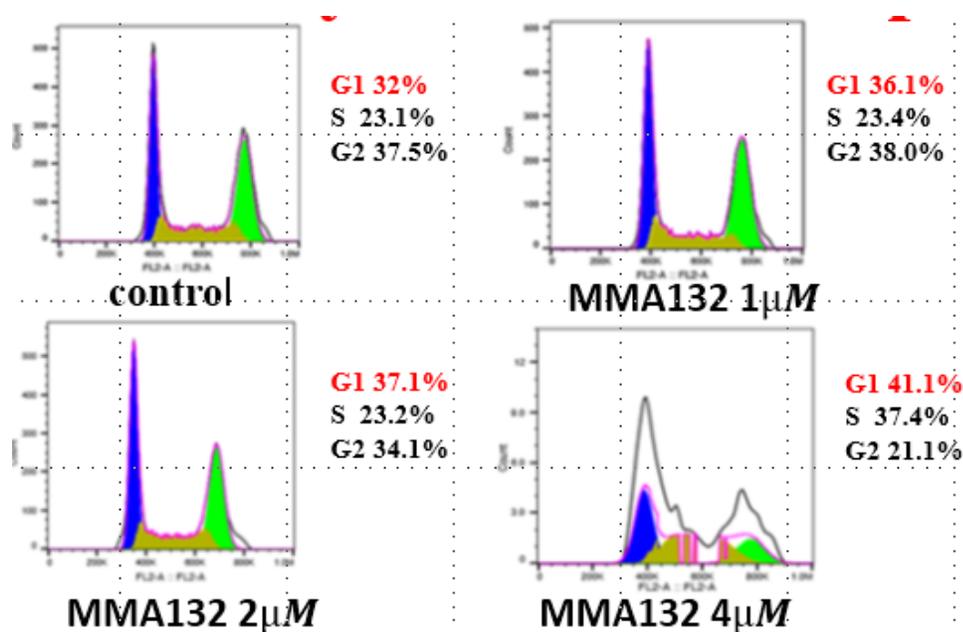
Figure 3.9. Cell cycle arrest analysis for analogue MMA132 showed induction activity for G1/S phases.

3.3 Conclusion:

In conclusion, CUCS-inspired analogues were designed using molecular modeling to mimic cucurbitacins structure more specifically its side chain, which is the most important pharmacophore for its biological activities and installed it into the estrone

skeleton structure at C-17 in the presence of double bond at C-16, C-17 position and methoxy group at C-3 which significantly improved their activity and the selectivity. Assortment of enone side chain were installed at the estrone skeleton structure such as isopropanol, para-flourobenzene, para-chlorobenzene, para-bromobenzene, para-methoxybenzene, para-trifluromethyl benzene, para-nitrobenzene, 5-bromofuran and 5-bromothiophene along with methoxy at C-3 and alkene at C16, C17. The cytotoxicity results showed that **MMA102** and **MM132** are the most potent analogues among the compunds that contain aliphatic enone side chains with IC$_{50}$ 3 μM and 2 μM respectively. On the other hand compound MMA311 which contain thiophene enone side chain at C-25 along with methoxy group at C-3 showed an outstanding IC$_{50}$ value of 0.7μM. Further biological experiments including western blot showed that compounds **MMA102** and **MMA132** have the ability to bind to the EGFR-TK and inhibit its phosphorylation by 90%. Additionally, cell cycle arrest has been conducted on the same analogues and demonstrated that they have the ability to arrest the cell cycle at G1/S phases.

3.4   Experimental section:

3.4.1   General:

All chemicals and solvents (ACS grades) were provided from Fisher Scientific or Sigma Aldrich and used without any additional purification. All glassware were cleaned, washed and dried in oven for overnight before conducting chemical reactions requires anhydrous environment and nitrogen gas applied at the reaction time. Pre-coated silica gel

PE plates were used to analyze the reaction condition and UV-light were also used at 254 or 365 to visualize the chemical reactions spots. All synthetic intermediate and final compounds were purified using column chromatography with 230^x400 mesh silica gel. ^{1}H and ^{13}C NMR spectra were using Bruker AVANCE-400 MHZ and 600 MHZ NMR spectrometer, in $CDCl_3$ and D-acetone. NMR chemical shifts were presented in δ(PPM) using residual solvent peaks as standards ($CDCl_3$, 7.26 (H), 77.16 (C)). High resolution mass (HRMS) was gained using thermofinnigan MAT 95XL mass spectrometer at Buffalo mass spectroscopy facility. X-ray crystallography were conducted in University of South Dakota on compound 8 using Bruker APEX[II] diffractometer.

3.4.2 Protected Estrone:
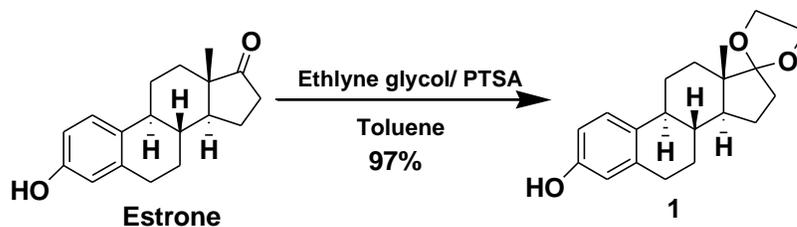
To a stirred solution of estrone (5g, 18.49mmole) in 308 ml of toluene, ethylene glycol (4.52ml, 81.35mmole) was poured in one portion. Then para-toluene sulfonic acid (0.21g, 1.10 mM) was added. The reaction mixture was allowed tom be stirred and reflux for 8 hours utilizing Dean-Stark apparatus to prevent water from going back to the reaction mixture. Then the reaction was cooled to the room temperature followed by the addition of $NaHCO_3$. Ethyl acetate (3^x50) was used to extract the aqueous level, then the organic layer dried over sodium sulfate anhydrous and evaporated by vacuo to provide white solid of the

protected estrone 1 (5.65g, 97%). The crude material was used in the next step without purification. ^{1}H NMR proved >95% purity.

^{1}H NMR (400 MHz, CDCl3) δ 7.20 (d, J = 8.6 Hz, 1H, -CH(1)), 6.72 (dd, J = 8.6, 2.8 Hz, 1H, -CH (2)), 6.62 (d, J = 2.7 Hz, 1H, -CH (4)), 3.93 (m, 4H, 2X -CH$_2$ (20, 21)), 2.88 (m, 2H), 2.29 (m, 2H, CH$_2$ (11)), 2.03 (m, 1H), 1.84 (m, 4H), 1.49 (m, 6H), 0.92 (s, 3H).

^{13}C NMR (100 MHz, CDCl3) δ 157.5 (C-3), 138.1 (C-5), 132.9 (C-10), 126.3 (C-1), 118.9 (C-17), 113.8 (C-4), 111.5 (C-2), 65.4 (C-20), 64.7 (C-21), 49.46 (C-13), 46.16 (C-14), 43.92 (C-9), 38.56 (C-8), 34.33 (C-16), 30.72 (C-6), 29.68 (C-11), 26.67 (C-7), 26.05 (C-12), 22.48 (C-15), 14.4(C-18).

3.4.3   Methoxy Protected Estrone 2:
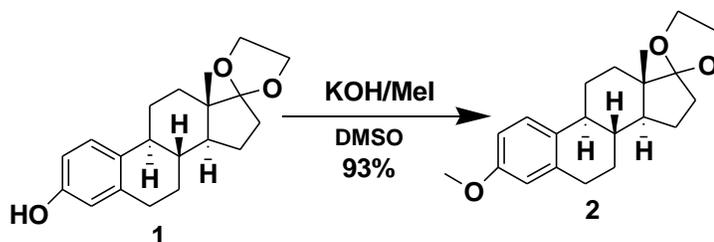
Protected estrone 1 (5.65g, 17.97 mmole) was dissolved in DMSO (90 ml), followed by the addition of crushed granulate potassium hydroxide (4.032g, 71.88 mmole), then methyl iodide (2.24g, 35.94 mmole) was added to the reaction mixture. The reaction was stirred at room teprature for 2 hours and quenched by the addition of water (400 ml) to be stirred for 15 minutes. Methylene dichloride (3 ×100 ml) was used to extract the

aqueous layer. The organic layer was dried over sodium sulfate anhydrous and concentrated in Vacuo to obtained the methoxy protected estrone 2 (5.5g, 93%) as a white powder. The crude material was used in the next step without purification.

1H NMR (400 MHz, CDCl3) δ 7.20 (d, J = 8.6 Hz, 1H, -CH(1)), 6.72 (dd, J = 8.6, 2.8 Hz, 1H, -CH(2)), 6.62 (d, J = 2.7 Hz, 1H, -CH(4)), 3.93 (m, 4H, 2 X -CH2 (20, 21)), 3.77 (s, 3H, -OCH3 (3)), 2.88 (m, 2H), 2.29 (m, 2H), 2.03 (m, 1H), 1.84 (m, 4H), 1.49 (m, 6H), 0.92 (s, 3H).

13C NMR (100 MHz, CDCl3) δ 157.5 (C-3), 138.0 (C-5), 132.5 (C-10), 126.3 (C-1), 118.9 (C-17), 113.9 (C-4), 111.6 (C-2), 65.4 (C-20), 64.7(C-21), 55.8 (-OCH3), 49.46 (C-13), 46.16 (C-14), 43.92 (C-9), 38.56 (C-8), 34.33 (C-16), 30.72 (C-6), 29.68 (C-11), 26.67 (C-7), 26.05 (C-12) , 22.48 (C-15), 14.43 (C-18).

3.4.4 Methoxy Estrone 3:
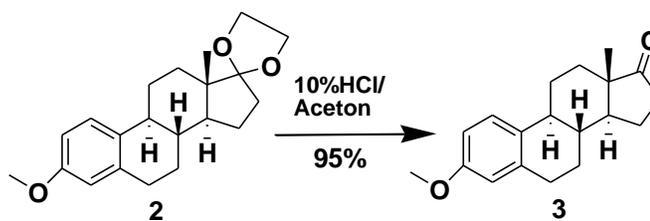
To stirred solution of methoxy protected estrone 2 (5.5g, 16.745mmole) in acetone (140ml), 10% HCl (140ml) was added. The reaction mixture allowed to stirred at room temperature for overnight. The reaction was quenched by dropwise addition of NaHCO3 to

balance the acidity of the mixture. Ethyl acetate (3 ˣ 50) was used to extract the aqueous

level, and then the organic layer dried over sodium sulfate anhydrous and evaporated by

vacuo. The crude product was purified using silica gel column chromatography (20% ethyl

acetate in hexane) to provide methoxy estrone 3 (4.52g, 95%) as white solid.

¹H NMR (400 MHz, CDCl3) δ 7.20 (d, J = 8.6 Hz, 1H, -**CH (1)**), 6.72 (dd, J = 8.6,

2.8 Hz, 1H, -**CH (2)**), 6.62 (d, J = 2.7 Hz, 1H, -**CH (4)**), 3.77 (s, 3H, -**OCH3 (3)**), 2.88 (m,

2H), 2.29 (m, 2H), 2.03 (m, 1H), 1.84 (m, 4H), 1.49 (m, 6H), 0.92 (s, 3H).

¹³C NMR (100 MHz, CDCl3) δ 220.8 (C-17), 157.5 (C-3), 138.0 (C-5), 132.5 (C-

10), 126.3 (C-1), 113.9 (C-4), 111.6 (C-2), 55.8 (-OCH3), 49.46 (C-13), 46.16 (C-14),

43.92 (C-9), 38.56 (C-8), 34.33 (C-16), 30.72 (C-6), 29.68 (C-11), 26.67 (C-7), 26.05 (C-

12), 22.48 (C-15), 14.43 (C-18).

4.4.5 Cyanohydrin:
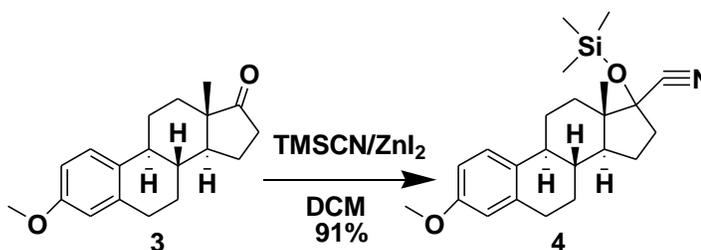
To a stirred solution of 3 (6g, 21.09mmole) in a dry dichloromethane (DCM) (42.2

ml), zinc iodide (0.6g, 1.87mmole) were added followed by the addition of trimethylsillyl

cyanide (TMSCN) (11.8ml, 94.94mmole). The reaction were allowed to be stirred for 3h at the room temperature then concentrated in vacuo. A mixture of water/ethyl acetate (1:1) ration were added to the concentrated slurry, then the aqueous layer was extracted using ethyl acetate (3^x50ml), dried over sodium sulfate anhydrous (Na_2SO_4) and concentrated under vacuo. The crude product was purified using silica gel column chromatography (20% ethyl acetate in hexane) to give (7.2g, 89%) of compound 4 as a white solid.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.04 (dd, *J* = 8.6, 1.0 Hz, 1H), 6.56 (dd, *J* = 8.6, 2.8 Hz, 1H), 6.47 (d, *J* = 2.8 Hz, 1H), 3.61 (s, 3H), 2.70 (m, 2H), 2.36 – 2.18 (m, 2H), 2.10 (td, *J* = 10.9, 10.4, 4.2 Hz, 1H), 1.91 – 1.58 (m, 5H), 1.48 (m, 1H), 1.39 – 1.27 (m, 3H), 1.27 – 1.13 (m, 1H), 0.72 (d, *J* = 8.8 Hz, 1H), 0.68 (s, 2H), 0.11 (s, 9H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 156.26, 136.44, 130.77, 125.08, 121.07, 112.52, 110.26, 80.39, 53.86, 47.04, 46.95, 42.04, 38.04, 36.75, 32.00, 28.44, 25.90, 25.00, 21.76, 10.97.

3.4.6 Alkene 5:

To a stirred solution of 4 (5g, 13.047mmole) in dry pyridine (125ml), phosphorus oxychloride (POCl₃) (8.65ml, 92.83mmole) and 1, 8-Diazabicyclo[20]undec-7-ene (DBU) (8.8ml, 58.7115mmole). The mixture were heated to the reflux temperature and stirred for 28h, then cooled to the room temperature. The black mixture was spilled into cooled mixture of (1:1) 5% hydrochloric acid (75ml) and ethyl acetate (75ml). The aqueous phase was extracted with ethyl acetate (3ˣ50ml), dried with sodium sulfate anhydrous (Na_2SO_4), and concentrated under vacuo. The crude material was purified using silica gel chromatography (10% ethyl acetate in hexane) to yield (2.4g, 64.05%) as a white solid.

¹H NMR (400 MHz, Chloroform-*d*) δ 7.24 (dd, 1H), 6.78 – 6.59 (m, 3H), 3.74 (s, 3H), 2.97 (m, 2H), 2.40 (m, 2H), 2.31 – 2.11 (m, 2H), 2.04 (m, 1H), 1.89 (m, 1H), 1.72 – 1.50 (m, 3H), 1.51 – 1.35 (m, 1H), 1.34 – 1.23 (m, 1H), 0.92 (s, 3H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 157.68, 147.36, 137.59, 132.07, 127.56, 126.07, 115.90, 113.96, 111.56, 55.24, 55.22, 48.51, 44.15, 37.24, 34.13, 32.73, 29.59, 27.74, 26.28, 16.47.

3.4.7 Methyl ketone 6:

In a 2-nacked flask, methyl lithium (MeLi) (12.8ml, 20.4654mmole) were added in a dropwise matter to a stirred solution of alkene 5 (3g, 10.2327mmole) in methyl tert-butyl ether (MTBE) (55ml) at 0°C. the reaction stirred at 0°C for 1h, then the reaction cooled down to -78 °C to quench it with trifluroacetic acid (2.554ml, 33.38mmole) in one portion followed by the addition of 10%H_3PO_4 (26ml). The mixture was allowed to stir for 30 min. the mixture was poured into 50ml ethyl acetate and 25ml 10% H_3PO_4. The aqueous layer was washed with ethyl acetate (3×50ml). The organic extract was washed with 250ml water, 15ml 1M Na_2CO_3, then 250ml water, and 100ml water. The organic extract was dried over sodium sulfate anhydrous (Na_2SO_4), filtrated, concentrated under vacuo. The resulted product purified by silica gel column chromatography (10% ethyl acetate in hexane) to give (2.97g, 93.6%) of methyl ketone 6 as a white solid.

¹H NMR (400 MHz, Chloroform-*d*) δ 7.24 (m, 1H), 6.77 – 6.66 (m, 3H), 3.76 (s, 3H), 2.98 (m, 2H), 2.56 (m, 2H), 2.44 – 2.30 (m, 2H), 2.27 (s, 3H), 2.19 – 2.04 (m, 1H), 1.96 – 1.85 (m, 1H), 1.72 – 1.60 (m, 3H), 1.56 (m, 1H), 1.45 (m, 1H), 1.26 (m, 1H), 0.91 (s, 3H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 196.78, 157.48, 155.52, 144.38, 137.73, 132.75, 126.13, 113.87, 111.41, 55.61, 55.20, 46.52, 44.26, 37.04, 34.84, 32.03, 29.67, 27.81, 27.19, 26.50, 15.97.

3.4.8 Cyanohydrin 7:
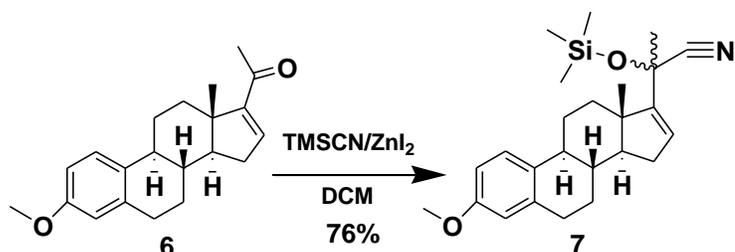
To a stirred solution of methyl ketone 6 (3g, 9.67mmole) in a dry dichloromethane (DCM) (20ml), zinc iodide (0.1g, 0.3mmole) were added followed by the addition of trimethylsillyl cyanide (TMSCN) (1.6ml, 12.6mmole). The reaction were allowed to be stirred for 3h at the room temperature then concentrated in vacuo. A mixture of water/ethyl acetate (1:1) ration were added to the concentrated slurry, then the aqueous layer was extracted using ethyl acetate (3^{x}50ml), dried over sodium sulfate anhydrous (Na_2SO_4) and concentrated under vacuo. The crude product was purified using silica gel column chromatography (10% ethyl acetate in hexane) to give a mixture (1:1) diasteriomers of Cyanohydrin 7 (3.1g, 78.3%).

^{1}H NMR (400 MHz, Chloroform-d) δ 7.33 (m, 1H), 6.86 – 6.68 (m, 2H), 6.04 (m, 1H), 3.92 (s, 3H), 3.00 (m, 2H), 2.52 – 2.26 (m, 2H), 2.18 – 1.88 (m, 7H), 1.88 – 1.62 (m, 2H), 1.60 – 1.39 (m, 2H), 1.33 – 1.08 (m, 3H), 1.08 – 0.93 (m, 1H), 0.42 (s, J = 2.9 Hz, 9H).

13C NMR (101 MHz, Chloroform-*d*) δ 156.07, 152.88, 136.22, 131.08, 126.35, 124.52, 120.07, 112.40, 109.95, 68.03, 67.21, 56.16, 53.58, 46.13, 45.60, 42.50, 35.70, 34.16, 33.55, 33.28, 30.23, 29.35, 29.02, 28.21, 26.16, 24.98, 23.90, 21.29, 15.90, 12.80.

3.4.9 Hydroxyl methyl ketone 8 and 8`:
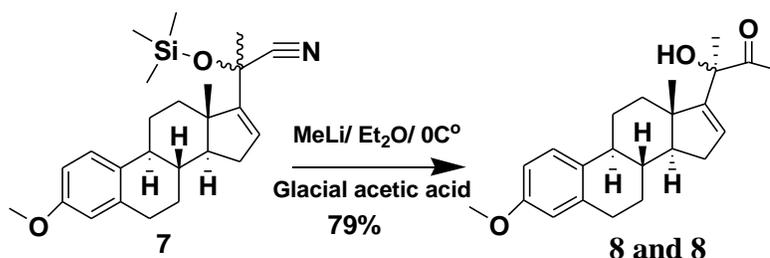
To a stirred solution of diasteriomers 7 (2.5g, 6.108mmole) in a dry ether (18ml), methyl lithium (MeLi) (11.5ml, 18.324mmole) was added in dropwise at 0^0C. The reaction mixture was allowed to stir for 2h at 0^0C, then the reaction quenched by adding glacial acetic acid (2.3ml) in one portion at 0^0C and allowed to stir for 30 min. at 0^0C. Sodium bicarbonate solution was added to neutralize the acidic mixture. Dichloromethane (DCM) was used to extract the aqueous layer (3^x50ml), dried under sodium sulfate anhydrous (Na$_2$SO$_4$), then concentrated under vacuo. The resulted diasteriomers were purified by silica gel column chromatography (100% hexane, 5% ethyl acetate in hexane and 10% ethyl acetate in hexane) to give the two separated diasteriomers hydroxyl methyl ketone 8 and 8`.

Compound 8 ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.15 (m, 1H), 6.70 (m, 1H), 6.62 (m, 1H), 5.90 (m, 1H), 4.25 (s, 1H), 3.76 (s, 3H), 2.87 (m, 2H), 2.31 – 2.14 (m, 3H),

2.10 – 1.99 (m, 2H), 1.91 (m, 2H), 1.66 – 1.53 (m, 3H), 1.53 (s, 3H), 1.48 – 1.33 (m, 1H), 1.31 – 1.12 (m, 2H), 1.01 (s, 3H), 0.98 – 0.79 (m, 1H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 211.7, 157.45, 155.56, 137.86, 132.73, 128.85, 126.04, 113.86, 111.39, 79.55, 57.49, 55.20, 47.79, 44.07, 37.09, 34.48, 31.05, 29.70, 27.64, 26.27, 25.31, 23.33, 17.24.

Compound 8` ¹H NMR (400 MHz, Chloroform-*d*) δ 7.28 (m, 1H), 6.69 (m, 1H), 6.62 (m, 1H), 5.89 (m, 1H), 4.08 (s, 1H), 3.76 (s, 3H), 2.96 – 2.85 (m, 2H), 2.85 – 2.78 (m, 2H), 2.37 – 2.15 (m, 3H), 2.08 – 1.76 (m, 3H), 1.76 – 1.47 (m, 3H), 1.47 – 1.34 (m, 2H), 1.30 – 1.08 (m, 3H), 1.06 – 0.85 (m, 3H), 0.85 – 0.80 (m, 1H). **¹³C NMR** (101 MHz, Chloroform-*d*) δ 210.93, 157.40, 155.48, 137.78, 132.75, 128.71, 126.01, 113.81, 111.32, 80.04, 56.61, 55.16, 48.09, 43.82, 37.16, 36.13, 31.01, 29.62, 27.59, 26.49, 24.88, 23.69, 16.93.

3.4.10 Δ9,11 hydroxyl methyl ketone 15 and 15`:
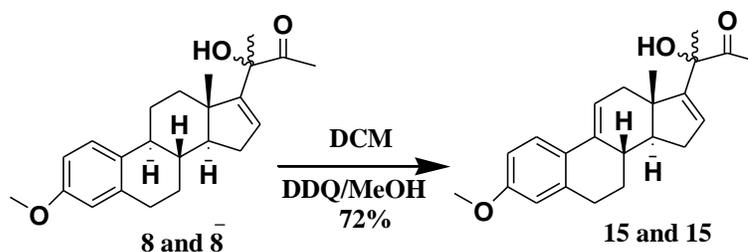
To a stirred solution of hydroxyl methyl ketone 8 and 8` (1g, 2.823mmole) in a dry dichloromethane (DCM) (14.1ml), a solution of 2,3-Dichloro-5,6-dicyano-1,4-

benzoquinone (DDQ) (0.83ml, 3.6699mmole) dissolved in methanol (31ml) was added at 0^0C under nitrogen. The reaction was raised to the room temperature and allowed to be stirred for 1h. The resulted crude material was concentrated under vacuo, then silica gel column chromatography were used to purify the crude material (10% ethyl acetate in hexane) to provide $\Delta^{9,11}$ hydroxyl methyl ketone 15 and 15`(0.82g, 82.4%) as a white material.

Compound 15 ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.49 (m, 1H), 6.69 − 6.49 (m, 2H), 6.25 (m, 1H), 5.73 (m, 1H), 4.42 (s, 1H), 3.70 (s, 3H), 2.80 − 2.68 (m, 2H), 2.23 − 2.17 (m, 3H), 2.14 (m, 1H), 2.12 − 1.99 (m, 1H), 1.99 − 1.85 (m, 1H), 1.55 (m, 3H), 1.43 (m, 1H), 1.35 − 1.24 (m, 1H), 1.18 (m, 2H), 1.03 (s, 3H), 0.79 (m, 1H).

^{13}**C NMR** (101 MHz, Chloroform-*d*) δ 211.13, 158.33, 153.91, 138.72,135.69, 128.77, 127.45,125.16, 117.25, 113.30, 112.61, 79.73, 55.22, 54.01, 46.13, 37.14, 36.62, 31.89, 29.96, 28.40, 25.59, 23.19, 17.15.

3.4.11 Ester 12:
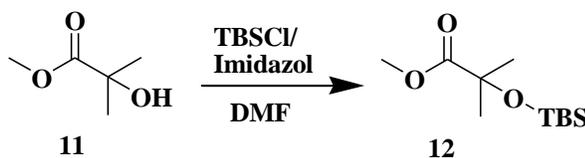
To a stirred solution of commercially available methyl2-hydroxy isobutyrate 11 (5g, 42.3mmole) in dimethyl formmide (DMF) (12.5ml), tert-butylsilyl chloride (TBSCl) (7.65g, 50.75mmole) were added followed by the addition of imidazole (7.45g, 110mmole). The reaction was stirred for 24h, then solution of sodium bicarbonate (25ml)

was added. The aqueous layer was extracted using ethyl acetate (3ˣ50ml), dried over

sodium sulfate anhydrous (Na_2SO_4), filtrated, and concentrated under vacuo. The crude

product was purified using silica gel column chromatography (10% ethyl acetate in hexane)

to give ester 12 (7.2g, 73.2%) as a yellow oil.

¹H NMR (400 MHz, CDCl3): d=3.60(s, 3H), 1.34 (s, 6H), 0.80 (s, 9 H), 0.08 ppm

(s, 6H);

¹³C NMR (100 MHz, CDCl3): d=175.8, 74.5, 51.4, 28.4, 25.5, 17.9, 3.2 ppm.

3.4.12 Alcohol 13:
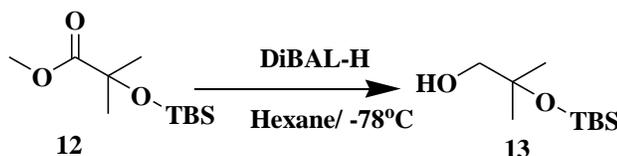
To a stirred solution of ester 12 (7.2g, 30.98mmole) in hexane (98ml),

Diisobutylaluminium hydride (DiBAL-H) (56.8ml, 68.16mmole) was added at -78°C in a

dropwise matter. The reaction stirred for 30 min. at 0⁰C, then for 20 min. at the room

temperature. The reaction then cooled back to -78°C, followed by the addition of solution

of potassium tartrate (45ml). The reaction was allowed to stir for overnight. The aqueous

phase was extracted using ethyl acetate (3ˣ50ml), dried over sodium sulfate anhydrous

(Na_2SO_4), filtrated, and concentrated under vacuo. Silica gel column chromatography was

used to purify the crude material (20% ethyl acetate in hexane) to give alcohol 13 (5.75g,

90.8%) as colorless liquid.

¹H NMR (400 MHz, CDCl3): d=3.19 (s, 2H), 2.31 (s, 1 H), 1.11 (s, 6 H), 0.76 (s, 9 H), 0.08 ppm (s, 6H);

¹³C NMR (100 MHz, CDCl3): d=76.3, 74.4, 28.5, 28.0, 20.3, 0.3 ppm.

3.4.13 Aldehyde 14:
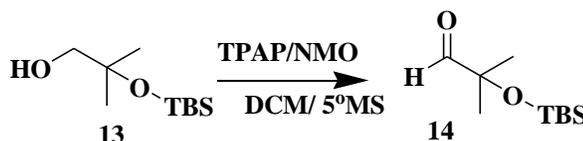
To a stirred solution of alcohol 13 (2g, 9.7858mmole) in a dry dichloromethane (DCM) (98ml), active molecular sieves (4°A, 2.3g) was added, then the reaction mixture was allowed to stir for 5 min. followed by the addition on N-Methylmorpholine N-oxide (NMO) (2.3g, 19.57mmole), then Tetrapropylammonium perruthenate (TPAP) (0.34g, 0.98mmole) was added at 0°C. The reaction stirred at 0°C for 2h, then at room temperature for 24h. Pad of silica gel were used to filtrate the reaction mixture using diethyl ether as eluting solvent. The filtrated material concentrated under vacuo to provide aldehyde 14 (1.45g, 73.3%) as colorless oil.

¹H NMR (400 MHz, CDCl3): d=9.44 (s, 1H), 1.17 (s, 6 H), 0.79 (s, 9H), 0.08 ppm (s, 6H).

¹³C NMR (100 MHz, CDCl3): d=206.4, 80.3, 28.0, 27.2, 20.4, 0.3 ppm.

3.4.14 Protected enone 9 and 9`:
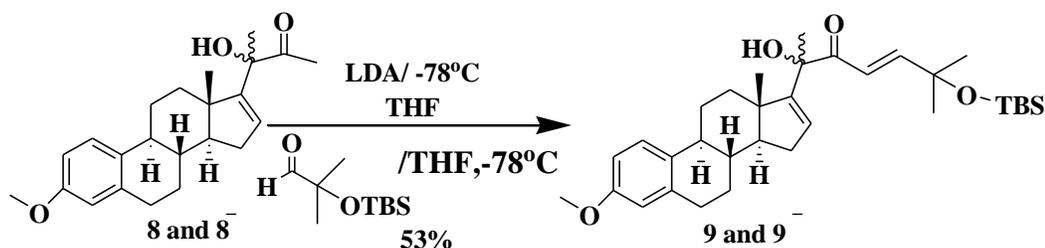
To a stirred solution of hydroxyl ketone 8 and 8 (0.5g, 1.4115mmole) in a dry tetrahydrofuran (THF) (3ml), lithiumdiisopropyl amine (LDA) (2.5ml, 5.08mmole) was added in a dropwise at -78°C the reaction mixture was allowed to stir at -78°C for 1 h, then the solution of aldehyde 14 (0.6g, 2.823mmole) in THF (18.8ml) was added to the reaction mixture at -78°C. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of ammonium chloride (NH$_4$Cl) (25ml). The aqueous layer was extracted using ethyl acetate (3^{x}50ml), dried over sodium sulfate anhydrous (Na$_2$SO$_4$), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to provide protected enone 9 and 9` (0.44g, 57.9%).

Compound 9 ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.13 − 6.92 (m, 2H), 6.70 − 6.54 (m, 2H), 6.52 (s, 1H), 5.82 (m, 1H), 4.32 (s, 1H), 3.65 (d, *J* = 4.2 Hz, 3H), 2.76 (m, 2H), 2.20 − 2.05 (m, 1H), 1.97 (m, 1H), 1.86 − 1.76 (m, 1H), 1.81 − 1.68 (m, 1H), 1.54 − 1.39 (m, 2H), 1.37 (d, *J* = 9.8 Hz, 3H), 1.33 − 1.19 (m, 6H), 1.23 − 1.03 (m, 4H), 0.93 (s, 2H), 0.91 − 0.81 (m, 1H), 0.86 − 0.73 (m, 12H), 0.78 − 0.70 (m, 1H), 0.11 (s, 3H).

^{13}**C NMR** (101 MHz, Chloroform-*d*) δ 203.82, 159.48, 158.50, 157.15, 139.80, 134.70, 131.27, 128.01, 120.90, 115.87, 113.42, 80.52, 76.84, 59.54, 57.12, 49.55, 46.26, 39.10, 36.14, 33.07, 31.88, 31.72, 29.70, 28.37, 28.26, 27.88, 27.79, 27.75, 27.24, 20.22, 20.05, 19.98, -1.47.

Compound 9 **¹H NMR** (400 MHz, Chloroform-*d*) δ 7.19-7.07 (m, 2H), 6.94-6.92 (m, 2H), 6.65 (s, 1H), 5.90 – 5.77 (m, 1H), 4.26 (s, 1H), 3.67 (d, *J* = 4.2 Hz, 3H), 2.87 – 2.71 (m, 2H), 2.22 (m, 1H), 2.19 (m, 1H), 2.15 (m, 1H), 1.62 (m, 1H), 1.55 – 1.42 (m, 2H), 1.41 (d, *J* = 9.8 Hz, 3H), 1.31 – 1.21 (m, 6H), 1.21 – 1.12 (m, 4H), 0.93 (m, 2H), 0.91 – 0.81 (m, 1H), 0.88 – 0.75 (m, 12H), 0.80 – 0.74 (m, 1H), 0.12 (s, 3H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 203.55, 159.45, 157.93, 157.44, 139.86, 134.92, 131.22, 128.12, 121.43, 115.91, 113.41, 81.27, 75.61, 58.83, 57.20, 50.10, 45.99, 39.28, 38.38, 33.08, 31.92, 31.76, 29.74, 28.61, 28.41, 28.07, 27.90, 27.82, 27.12, 20.28, 20.12, 19.24, -1.43.

3.4.15 MMA102 and MMA 132:

9 and 9   **MMA102 and MMA132**
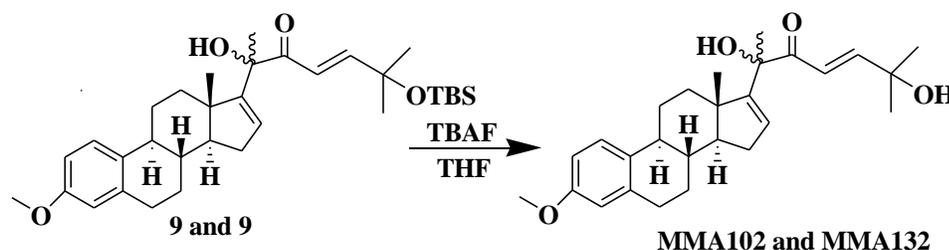
To a stirred solution of protected enone 9 and 9` (0.4g, 0.742mmole) in tetrahydrofuran (THF) (11.5ml), tetrabutyl ammonium fluoride (TBAF) (2.3ml, 2.3mmole) was added and stirred for 6h. Ammonium chloride (NH₄Cl) solution was added to quench the reaction. Ethyl acetate (3ˣ50ml), dried over sodium sulfate anhydrous (Na₂SO₄), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to provide MMA102 and MMA132 (0.28g, 88.88%).

Compound MMA102 ¹H NMR (400 MHz, Chloroform-*d*) δ 7.14 – 6.98 (m, 2H), 6.65 – 6.45 (m, 2H), 5.83 (m, 2H), 4.02 (s, 1H), 3.67 (s, 3H), 2.77 (m, 2H), 2.26 – 2.06 (m, 2H), 1.97 – 1.70 (m, 3H), 1.61 (m, 1H), 1.45 (s, 3H), 1.42 (m, 6H), 1.24 – 1.07 (m, 4H), 0.85 – 0.74 (m, 2H), 0.65 (s, 3H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 201.30, 157.39, 155.23, 154.75, 137.86, 132.85, 129.19, 126.07, 119.36, 113.83, 111.36, 79.17, 71.18, 56.77, 55.21, 48.08, 43.87, 37.21, 36.20, 31.08, 29.67, 29.45, 29.31, 27.61, 26.51, 24.90, 17.36. **HR-FT-MS calcd for C₂₇H₃₆O₄Na₁ 447.2506 found 447.24952.**

Compound MMA132 ¹H NMR (400 MHz, Chloroform-*d*) δ 7.13 – 7.01 (m, 2H), 6.65 – 6.50 (m, 2H), 5.84 (m, 2H), 4.02 (s, 1H), 3.67 (s, 3H), 2.77 (m, 2H), 2.22 – 1.88 (m, 2H), 1.78 (m, 3H), 1.64 (m, 1H), 1.57 (s, 3H), 1.36 (m, 6H), 1.28- 1.15 (m, 4H), 1.08 (m, 2H), 0.98 (s, 3H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 201.55, 157.38, 155.40, 154.86, 137.88, 132.75, 129.45, 126.01, 118.76, 113.81, 111.39, 78.61, 71.18, 57.40, 55.19, 47.71, 44.07, 37.06, 34.29, 31.10, 29.70, 29.47, 29.42, 27.59, 26.23, 25.18, 17.27. **HR-FT-MS calcd for C₂₇H₃₆O₄Na₁ 447.2506 found 447.24952.**

3.4.16 $\Delta^{9,11}$ protected enone 16:
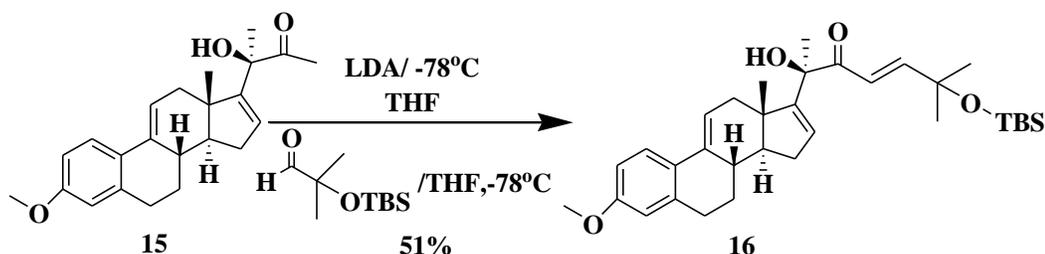
To a sirred solution of $\Delta^{9,11}$ hydroxyl ketone 15 (1g, 2.84mmole) in a dry tetrahydrofuran (THF) (5.7ml), lithiumdiisopropyl amine (LDA) (5.112mlml, 10.224mmole) was added in a dropwise at -78°C the reaction mixture was allowed to stir at -78°C for 1 h, then the solution of aldehyde 14 (1.15, 5.68mmole) in THF (37.8ml) was added to the reaction mixture at -78°C. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of ammonium chloride (NH4Cl) (25ml). The aqueous layer was extracted using ethyl acetate 3ˣ50ml), dried over sodium sulfate anhydrous (Na2SO4), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to provide $\Delta^{9,11}$ protected enone 16.

¹H NMR (600 MHz, Chloroform-*d*) δ 7.43 (d, *J* = 8.8 Hz, 1H), 6.98 (d, *J* = 15.2 Hz, 1H), 6.72 – 6.63 (m, 1H), 6.66 – 6.61 (m, 1H), 6.54 (dd, *J* = 14.5, 3.2 Hz, 1H), 5.94 (dt, *J* = 5.2, 2.3 Hz, 1H), 5.85 (dd, *J* = 3.3, 1.6 Hz, 1H), 4.37 (s, 1H), 3.71 (s, 3H), 2.91 – 2.71 (m, 2H), 2.26 (ddt, *J* = 13.9, 7.1, 3.6 Hz, 1H), 2.23 – 2.15 (m, 1H), 2.09 – 1.92 (m, 2H), 1.91 (dt, *J* = 17.4, 3.2 Hz, 1H), 1.58 – 1.48 (m, 1H), 1.46 – 1.33 (m, 2H), 1.34 – 1.20

(m, 4H), 1.23 – 1.12 (m, 3H), 1.15 – 1.04 (m, 2H), 0.95 (s, 3H), 0.93 – 0.77 (m, 9H), 0.80

– 0.71 (m, 6H).

13C NMR (151 MHz, Chloroform-*d*) δ 203.89, 160.41, 158.95, 155.34 (d, *J* = 26.4

Hz), 139.50, 137.55, 131.35, 129.65, 127.20, 120.62, 119.70, 115.42, 114.71, 80.74, 75.69,

57.33, 56.15, 47.97, 38.91, 38.66, 33.96, 32.03, 31.96, 31.90, 30.48, 27.98 – 27.82 (m),

27.56, 20.25 (d, *J* = 5.7 Hz), 19.23, -0.04 (t, *J* = 7.5 Hz).

3.4.17 MMA 128 (Δ9,11 OH enone):
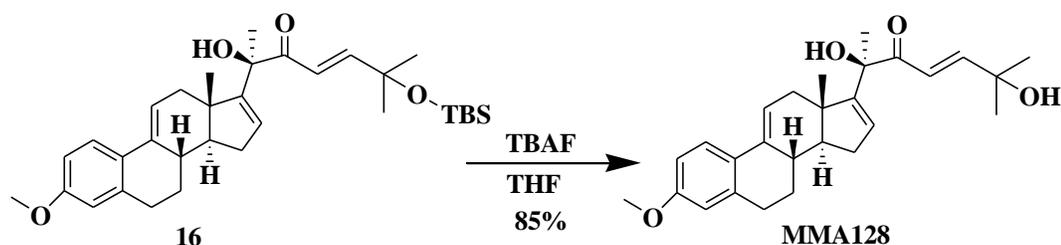
To a stirred solution of $\Delta^{9,11}$ protected enone 16 and 16 (0.25g, 0.466mmole) in

tetrahydrofuran (THF) (7.2ml), tetrabutyl ammonium fluoride   (TBAF) (1.44ml,

1.44mmole) was added and stirred for 6h. ammonium chloride (NH₄Cl) solution was added

to quench the reaction. Ethyl acetate (3ˣ50ml), dried over sodium sulfate anhydrous

(Na₂SO₄), filtrated, and concentrated under vacuo. Silica gel column chromatography was

used to purify the crude material (10% ethyl acetate in hexane) to provide MMA128

(0.162g, 82.35%).

1H NMR (400 MHz, Chloroform-*d*) δ 7.11 – 7.02 (m, 2H), 6.63 – 6.47 (m, 2H),

5.82 (m, 2H), 5.72 (m, 1H), 4.01 (s, 1H), 3.64 (s, 3H), 2.73 (m, 1H), 2.20 – 1.86 (m, 2H),

1.78 (m, 2H), 1.64 (m, 1H), 1.57 (s, 3H), 1.36 (m, 6H), 1.28- 1.15 (m, 3H), 1.08 (m, 2H), 0.98 (s, 3H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 201.53, 157.37, 155.38, 154.83, 137.87, 132.74, 129.43, 125.92, 118.74, 113.80, 111.37, 78.59, 71.16, 57.39, 55.17, 47.72, 44.05, 37.04, 34.27, 31.07, 29.68, 29.45, 29.40, 27.57, 26.21, 25.15, 17.24. **HR-FT-MS calcd for C$_{27}$H$_{34}$O$_4$Na$_1$ 447.25181 found 447.24621.**

3.4.18 Para-methoxyphenyl enone MMA265 and MMA333:

To a sirred solution of hydroxyl ketone 8 and 8` (0.3g, 0.847mmole) in a dry tetrahydrofuran (THF) (1.7ml), lithiumdiisopropyl amine (LDA) (1.5ml, 3.05mmole) was added in a dropwise at -78°C the reaction mixture was allowed to stir at -78°C for 1 h, then the solution of Para-anisaldehyde (0.206ml, 1.7mmole) in THF (11.33ml) was added to the reaction mixture at -78°C. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of ammonium chloride (NH$_4$Cl) (25ml). The aqueous layer was extracted using ethyl acetate (3^{x}50ml), dried over sodium sulfate anhydrous (Na$_2$SO$_4$), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl

acetate in hexane) to provide Para-methoxy benzene enone MMA265 and MMA333 (0.26g, 65%).

Compound 265 ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.75 – 7.59 (m, 1H), 7.49 – 7.37 (m, 2H), 7.15 – 7.01 (m, 1H), 6.86 – 6.74 (m, 3H), 6.57 (dd, *J* = 8.6, 2.8 Hz, 1H), 6.56 – 6.47 (m, 1H), 5.86 (m, 1H), 4.30 (s, 1H), 3.72 (s, 3H), 3.64 (s, 3H), 2.84 – 2.66 (m, 2H), 2.23 – 2.08 (m, 4H), 2.00 – 1.72 (m, 3H), 1.64 – 1.55 (m, 3H), 1.48 (s, 3H), 1.35 – 1.25 (m, 1H), 1.24 – 1.08 (m, 3H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 200.61, 161.99, 157.43, 155.97, 144.24, 137.83, 132.89, 130.47, 128.74, 127.08, 126.07, 117.20, 114.47, 113.85, 111.35, 79.03, 60.42, 55.43, 55.17, 48.13, 43.90, 37.24, 36.26, 31.97, 29.76, 29.71, 29.69, 27.67, 26.56, 25.26, 21.06. **HR-FT-MS calcd for $C_{31}H_{36}O_4Na_1$ 495.2506 found 495.24909.**

Compound MMA333 ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.74 – 7.57 (m, 1H), 7.47 – 7.35 (m, 2H), 7.13 – 6.99 (m, 1H), 6.84 – 6.72 (m, 3H), 6.55 (dd, *J* = 8.6, 2.8 Hz, 1H), 6.54 – 6.45 (m, 1H), 5.84 (m, 1H), 4.28 (s, 1H), 3.70 (s, 3H), 3.62 (s, 3H), 2.82 – 2.64 (m, 2H), 2.21 – 2.06 (m, 4H), 1.98 – 1.70 (m, 3H), 1.62 – 1.53 (m, 3H), 1.46 (s, 3H), 1.32 – 1.23 (m, 1H), 1.22 – 1.06 (m, 3H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 200.59, 161.97, 157.41, 155.95, 144.22, 137.81, 132.87, 130.45, 128.72, 127.06, 126.05, 117.18, 114.45, 113.83, 111.33, 79.01, 60.40, 55.41, 55.15, 48.11, 43.88, 37.22, 36.24, 31.95, 29.74, 29.69, 29.67, 27.65, 26.54, 25.24, 21.04. **HR-FT-MS calcd for $C_{31}H_{36}O_4Na_1$ 495.2506 found 495.24883.**

3.4.19 Para-flurophenyl enone MMA270 and MMA334:
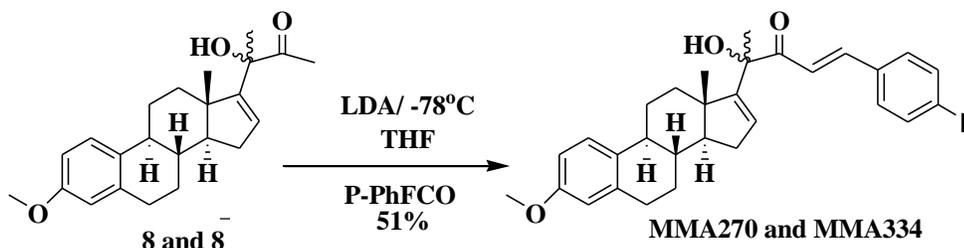
To a stirred solution of hydroxyl ketone 8 and 8` (0.4g, 1.13mmole) in a dry tetrahydrofuran (THF) (2.3ml), lithiumdiisopropyl amine (LDA) (2.034ml, 4.068mmole) was added in a dropwise at -78°C the reaction mixture was allowed to stir at -78°C for 1 h, then the solution of Para-fluorobenzaldehyde (0.242ml, 2.26mmole) in THF (15ml) was added to the reaction mixture at -78°C. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of ammonium chloride (NH_4Cl) (25ml). The aqueous layer was extracted using ethyl acetate (3^x50ml), dried over sodium sulfate anhydrous (Na_2SO_4), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to provide Para-flurobenzene enone MMA270 and MMA334 (0.32g, 61.5%).

Compound MMA270 ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.70 (m, 1H), 7.54 – 7.35 (m, 2H), 7.10 (m, 1H), 7.05 – 6.96 (m, 2H), 6.87 (m, 1H), 6.69 – 6.51 (m, 2H), 5.91 (m, 1H), 4.22 (s, 1H), 3.68 (s, 3H), 2.79 (m, 2H), 2.29 (s, 1H), 2.25 – 2.01 (m, 3H), 1.93 – 1.76 (m, 3H), 1.76 – 1.37 (m, 3H), 1.37 – 1.36 (m, 1H), 1.36 – 1.22 (m, 2H), 1 1.04 – 0.75 (m, 1H), 0.67 (s, 3H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 200.56, 157.43, 155.69, 143.10, 137.99, 137.84, 132.84, 130.64, 130.60, 130.55, 129.05, 126.07, 119.29, 116.32, 116.10, 113.85, 111.37, 79.20, 55.21, 53.48, 48.13, 43.88, 37.23, 36.23, 31.14, 29.67, 27.64, 26.54, 25.10, 17.49. **HR-FT-MS calcd for C₃₀H₃₃O₃F₁Na₁ 483.2306 found 483.22909.**

Compound MMA334 ¹H NMR (400 MHz, Chloroform-*d*) δ 7.68 (m, 1H), 7.52 – 7.33 (m, 2H), 7.08 (m, 1H), 7.03 – 6.94 (m, 2H), 6.85 (m, 1H), 6.67 – 6.49 (m, 2H), 5.89 (m, 1H), 4.20 (s, 1H), 3.66 (s, 3H), 2.77 (m, 2H), 2.27 (s, 1H), 2.23 – 1.99 (m, 3H), 1.91 – 1.74 (m, 3H), 1.74 – 1.35 (m, 3H), 1.35 – 1.34 (m, 1H), 1.34 – 1.20 (m, 2H), 1.02 – 0.73 (m, 1H), 0.65 (s, 3H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 200.54, 157.41, 155.67, 143.08, 137.97, 137.82, 132.82, 130.62, 130.58, 130.53, 129.03, 126.05, 119.25, 116.30, 116.08, 113.83, 111.35, 79.18, 55.19, 53.46, 48.11, 43.86, 37.21, 36.21, 31.12, 29.65, 27.62, 26.52, 25.08, 17.47. **HR-FT-MS calcd for C₃₀H₃₃O₃F₁Na₁ 483.2306 found 483.22909.**

3.4.20 Para-chlorophenyl enone MMA287 and MMA316:

To a stirred solution of hydroxyl ketone 8 and 8` (0.3g, 0.847mmole) in a dry tetrahydrofuran (THF) (1.7ml), lithiumdiisopropyl amine (LDA) (1.5ml, 3.05mmole) was added in a dropwise at -78°C the reaction mixture was allowed to stir at -78°C for 1 h, then the solution of Para-anisaldehyde (0.206ml, 1.7mmole) in THF (11.33ml) was added to the reaction mixture at -78°C. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of ammonium chloride (NH$_4$Cl) (25ml). The aqueous layer was extracted using ethyl acetate (3^{x}50ml), dried over sodium sulfate anhydrous (Na$_2$SO$_4$), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to obtained P-chlorobenzene enone MMA287 and MMA316 (0.25g, 62%).

Compound 287 ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.80 (m, 1H), 7.56 – 7.51 (m, 2H), 7.44 – 7.38 (m, 2H), 7.21 (m, 1H), 7.03 (m, 1H), 6.73 (m, 1H), 6.66 (m, 1H), 6.03 (m, 1H), 4.29 (s, 1H), 3.80 (s, 3H), 2.97 – 2.84 (m, 2H), 2.36 – 2.25 (m, 2H), 2.10 – 2.00 (m, 1H), 1.97 – 1.87 (m, 1H), 1.77 (m, 1H), 1.64 (s, 3H), 1.55 – 1.41 (m, 2H), 1.37 – 1.25 (m, 2H), 0.94 (m, 2H), 0.79 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 200.54, 157.45, 155.66, 142.92, 137.84, 136.86, 132.83, 129.77, 129.30, 129.12, 126.07, 120.00, 113.86, 111.38, 79.24, 56.83, 55.21, 48.14, 43.88, 37.24, 36.23, 31.96, 31.15, 29.74, 29.67, 29.40, 27.64, 26.54, 25.07, 17.51. **HR-FT-MS calcd for C$_{30}$H$_{33}$O$_3$Cl$_1$Na$_1$ 499.2010 found 499.19958.**

Compound 316 ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.78 (d, m, 1H), 7.54 – 7.49 (m, 2H), 7.42 – 7.36 (m, 2H), 7.19 (m, 1H), 7.01 (m, 1H), 6.71 (m, 1H), 6.64 (m, 1H), 6.01 (m, 1H), 4.27 (s, 1H), 3.78 (s, 3H), 2.95 – 2.82 (m, 2H), 2.34 – 2.23 (m, 2H), 2.08 – 1.98 (m, 1H), 1.95 – 1.85 (m, 1H), 1.75 (m, 1H), 1.62 (s, 3H), 1.53 – 1.39 (m, 2H), 1.35 – 1.23 (m, 2H), 0.92 (m, 2H), 0.77 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 200.52, 157.43, 155.64, 142.90, 137.82, 136.84, 132.81, 129.75, 129.28, 129.10, 126.05, 119.98, 113.84, 111.36, 79.22, 56.81, 55.19, 48.12, 43.86, 37.22, 36.21, 31.94, 31.13, 29.72, 29.65, 29.38, 27.62, 26.52, 25.05, 17.49. **HR-FT-MS calcd for $C_{30}H_{33}O_3Cl_1Na_1$ 499.2010 found 499.19943.**

3.4.21 Para-bromophenyle enone MMA288 and MMA319:

To a stirred solution of hydroxyl ketone 8 and 8` (0.4g, 1.13mmole) in a dry tetrahydrofuran (THF) (2.3ml), lithiumdiisopropyl amine (LDA) (2.034ml, 4.068mmole) was added in a dropwise at -78°C the reaction mixture was allowed to stir at -78°C for 1 h, then the solution of Para-flurobenzaldehyde (0.242ml, 2.26mmole) in THF (15ml) was added to the reaction mixture at -78°C. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of ammonium chloride (NH4Cl) (25ml). The aqueous layer was extracted using ethyl acetate

(3ˣ50ml), dried over sodium sulfate anhydrous (Na_2SO_4), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to provide para-bromobenzene enone MMA288 and MMA319 (0.35g, 59%).

Compound MMA288 ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.70 – 7.61 (m, 1H), 7.53 – 7.42 (m, 2H), 7.41 – 7.32 (m, 2H), 7.15 – 7.08 (m, 1H), 6.93 (m, 1H), 6.63 (m, 1H), 6.55 (d, 1H), 5.91 (m, 1H), 4.16 (s, 1H), 3.74 (s, 3H), 2.86 – 2.74 (m, 2H), 2.27 – 2.15 (m, 2H), 2.03 (s, 1H), 1.98 – 1.89 (m, 1H), 1.86 – 1.76 (m, 1H), 1.65 (s, 3H), 1.51 (d, 2H), 1.47 – 1.31 (m, 2H), 0.93 – 0.75 (m, 2H), 0.67 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 200.6, 157.5, 155.6, 142.9, 133.3, 132.9, 132.4, 129.8, 129.3, 129.1, 126.1, 125.3, 120.1, 113.9, 111.4, 79.3, 55.2, 53.4, 48.2, 43.9, 37.3, 36.3, 31.2, 29.8, 29.6, 27.8, 26.6, 25.2, 25.01, 17.6. **HR-FT-MS calcd for $C_{30}H_{33}O_3Br_1Na_1$ 543.1505 found 543.15269.**

Compound MMA319 ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.68 – 7.59 (m, 1H), 7.51 – 7.40 (m, 2H), 7.39 – 7.30 (m, 2H), 7.13 – 7.06 (m, 1H), 6.91 (m, 1H), 6.61 (m, 1H), 6.53 (d, 1H), 5.89 (m, 1H), 4.14 (s, 1H), 3.72 (s, 3H), 2.84 – 2.72 (m, 2H), 2.25 – 2.13 (m, 2H), 2.01 (s, 1H), 1.96 – 1.87 (m, 1H), 1.84 – 1.74 (m, 1H), 1.63 (s, 3H), 1.49 (d, 2H), 1.45 – 1.29 (m, 2H), 0.91 – 0.73 (m, 2H), 0.65 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 200.4, 157.3, 155.4, 142.7, 133.1, 132.7, 132.2, 129.6, 129.1, 128.99, 127.99, 125.1, 119.99, 113.7, 111.2, 79.1, 54.99, 53.2, 48.02, 43.7, 37.1, 36.1, 31.01, 29.6, 29.4, 27.6, 26.4, 25.03, 24.99, 17.4. **HR-FT-MS calcd for $C_{30}H_{33}O_3Br_1Na_1$ 543.1505 found 543.14872.**

3.4.22 2-bromofuran enone MMA290 and MMA318:

8 and 8

MMA290 and MMA318

To a stirred solution of hydroxyl ketone 8 and 8` (0.4g, 1.13mmole) in a dry tetrahydrofuran (THF) (2.3ml), lithiumdiisopropyl amine (LDA) (2.034ml, 4.068mmole) was added in a dropwise at -78°C the reaction mixture was allowed to stir at -78°C for 1 h, then the solution of Para-flurobenzaldehyde (0.242ml, 2.26mmole) in THF (15ml) was added to the reaction mixture at -78°C. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of ammonium chloride (NH4Cl) (25ml). The aqueous layer was extracted using ethyl acetate (3ˣ50ml), dried over sodium sulfate anhydrous (Na2SO4), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to provide para-bromobenzene enone MMA290 and MMA318 (0.31g, 53.7%).

Compound MMA290 ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.47 (m, 1H), 7.21 (m, 1H), 6.88 (m, 1H), 6.73 (m, 1H), 6.67 (m, 2H), 6.48 (m, 1H), 6.04 – 6.00 (m, 1H), 4.33 (s, 1H), 3.83 (s, 3H), 2.97 – 2.84 (m, 2H), 2.31 (m, 2H), 2.11 – 1.98 (m, 1H), 1.99 – 1.88 (m, 1H), 1.76 (m, 1H), 1.64 – 1.55 (m, 3H), 1.53 – 1.41 (m, 2H), 1.36 – 1.26 (m, 2H), 0.98 – 0.83 (m, 2H), 0.78 (m , 3H).

¹³C NMR (151 MHz, Chloroform-*d*) δ 200.53, 157.43, 155.42, 153.03, 137.88, 132.90, 129.22, 128.92, 126.35, 126.07, 119.01, 117.47, 114.76, 113.85, 111.37, 79.17, 56.78, 55.21, 48.13, 43.89, 37.24, 36.29, 31.16, 29.69, 27.65, 26.56, 25.06, 17.48. **HR-FT-MS calcd for C₂₈H₃₁O₄Br₁Na₁ 533.1298 found 533.12803.**

Compound MMA318 ¹H NMR (600 MHz, Chloroform-*d*) δ 7.45 (m, 1H), 7.19 (m, 1H), 6.86 (m, 1H), 6.71 (m, 1H), 6.65 (m, 2H), 6.46 (m, 1H), 6.02 – 5.98 (m, 1H), 4.31 (s, 1H), 3.81 (s, 3H), 2.95 – 2.82 (m, 2H), 2.29 (m, 2H), 2.09 – 1.96 (m, 1H), 1.97 – 1.86 (m, 1H), 1.74 (m, 1H), 1.62 – 1.53 (m, 3H), 1.51 – 1.39 (m, 2H), 1.34 – 1.24 (m, 2H), 0.96 – 0.81 (m, 2H), 0.76 (m , 3H).

¹³C NMR (151 MHz, Chloroform-*d*) δ 200.51, 157.41, 155.40, 153.01, 137.86, 132.88, 129.20, 128.90, 126.33, 126.05, 118.99, 117.45, 114.74, 113.83, 111.35, 79.15, 56.76, 55.19, 48.11, 43.87, 37.22, 36.27, 31.14, 29.67, 27.63, 26.54, 25.04, 17.46. **HR-FT-MS calcd for C₂₈H₃₁O₄Br₁Na₁ 533.1298 found 533.12807.**

3.4.23 Para-trifluromethyl phenyl enone MMA292 and MMA320:

To a stirred solution of hydroxyl ketone 8 and 8` (0.3g, 0.847mmole) in a dry tetrahydrofuran (THF) (1.7ml), lithiumdiisopropyl amine (LDA) (1.5ml, 3.05mmole) was added in a dropwise at -78°C the reaction mixture was allowed to stir at -78°C for 1 h, then the solution of Para-anisaldehyde (0.206ml, 1.7mmole) in THF (11.33ml) was added to the reaction mixture at -78°C. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of ammonium chloride (NH_4Cl) (25ml). The aqueous layer was extracted using ethyl acetate (3^x50ml), dried over sodium sulfate anhydrous (Na_2SO_4), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to obtained para-trifluromethyl benzene enone MMA292 and MMA320 (0.21g, 61.7%).

Compound MMA292 ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.77 (m, 1H), 7.62 – 7.53 (m, 4H), 7.05 – 6.98 (m, 2H), 6.59 (m, 1H), 6.53 (m, 1H), 5.97 – 5.93 (m, 1H), 4.32 (s, 1H), 3.67 (m, 3H), 2.88 – 2.71 (m, 2H), 2.21-2.12 (m, 2H), 2.02 (m, 1H), 1.81 (m, 1H), 1.53 (m, 1H), 1.49 – 1.37 (m, 3H), 1.30 (m, 1H), 1.11 (m, 1H), 0.97 (s, 3H), 0.90 – 0.71 (m, 4H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 200.66, 157.42, 155.15, 142.93, 137.83, 137.67, 132.74, 129.72, 128.80, 126.00, 125.98, 125.95, 125.93, 121.08, 113.82, 111.36, 78.80, 57.47, 55.17, 47.85, 43.93, 37.09, 34.48, 31.96, 31.23, 29.74, 29.69, 27.59, 26.27, 25.26, 17.21. **HR-FT-MS calcd for $C_{31}H_{33}O_3F_3Na_1$ 533.2274 found 533.22689.**

Compound MMA320 ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.75 (m, 1H), 7.60 – 7.51 (m, 4H), 7.03 – 6.96 (m, 2H), 6.57 (m, 1H), 6.51 (m, 1H), 5.95 – 5.91 (m, 1H), 4.30 (s, 1H), 3.65 (m, 3H), 2.86 – 2.76 (m, 2H), 2.19-2.10 (m, 2H), 2.01 (m, 1H), 1.79 (m, 1H), 1.51 (m, 1H), 1.47 – 1.35 (m, 3H), 1.28 (m, 1H), 1.09 (m, 1H), 0.95 (s, 3H), 0.88 – 0.69 (m, 4H). **HR-FT-MS calcd for $C_{31}H_{33}O_3F_3Na_1$ 533.2274 found 533.22544.**

^{13}C NMR (151 MHz, Chloroform-*d*) δ 200.64, 157.40, 155.13, 142.91, 137.81, 137.65, 132.72, 129.70, 128.78, 125.98, 125.96, 125.93, 125.91, 121.06, 113.80, 111.34, 78.878, 57.45, 55.15, 47.83, 43.91, 37.07, 34.46, 31.94, 31.21, 29.72, 29.67, 27.57, 26.25, 25.24, 17.19. **HR-FT-MS calcd for $C_{31}H_{33}O_3F_3Na_1$ 533.2274 found 533.22544.**

3.4.24 Para-nitrophenyl enone MMA305 and MMA321:

To a stirred solution of hydroxyl ketone 8 and 8` (0.4g, 1.13mmole) in a dry tetrahydrofuran (THF) (2.3ml), lithiumdiisopropyl amine (LDA) (2.034ml, 4.068mmole) was added in a dropwise at -78ºC the reaction mixture was allowed to stir at -78ºC for 1 h, then the solution of Para-flurobenzaldehyde (0.242ml, 2.26mmole) in THF (15ml) was added to the reaction mixture at -78ºC. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of

ammonium chloride (NH$_4$Cl) (25ml). The aqueous layer was extracted using ethyl acetate (3^{x}50ml), dried over sodium sulfate anhydrous (Na$_2$SO$_4$), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to provide para-nitrobenzene enone MMA305 and MMA321 (0.29g, 52.7%).

Compound MMA305 ^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.16 (m, 2H), 7.77 (m, 1H), 7.63 (m, 2H), 7.14 (s, 1H), 7.10 – 6.98 (m, 2H), 6.61 – 6.54 (m, 1H), 6.51 (d, 1H), 5.97 (d, 1H), 4.26 (s, 1H), 3.66 (s, 3H), 2.81- 2.75 (m, 2H), 2.22-2.12 (m, 2H), 2.01 (m, 1H), 1.81 (d, 1H), 1.50 – 1.37 (m, 1H), 1.34 – 1.15 (m, 3H), 1.10 (m, 1H), 0.96 (s, 3H), 0.87- 0.78 (m, 4H).

^{13}**C NMR** (101 MHz, Chloroform-*d*) δ 200.47, 157.42, 154.97, 148.75, 141.76, 140.37, 137.81, 132.64, 129.99, 129.26, 125.98, 124.22, 122.60, 113.80, 111.37, 78.91, 57.50, 55.16, 47.85, 43.93, 37.07, 34.69, 34.56, 31.62, 31.26, 29.67, 27.59, 26.27, 25.31, 25.20, 22.69, 17.21, 14.18. **HR-FT-MS calcd for C$_{30}$H$_{33}$O$_5$N$_1$Na$_1$ 510.2251 found 510.22450.**

Compound MMA321 ^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.14 (m, 2H), 7.75 (m, 1H), 7.61 (m, 2H), 7.12 (s, 1H), 7.08 – 6.96 (m, 2H), 6.59 – 6.52 (m, 1H), 6.49 (d, 1H), 5.95 (d, 1H), 4.24 (s, 1H), 3.64 (s, 3H), 2.79- 2.73 (m, 2H), 2.20-2.10 (m, 2H), 2.01 (m, 1H), 1.79 (d, 1H), 1.48 – 1.35 (m, 1H), 1.32 – 1.13 (m, 3H), 1.08 (m, 1H), 0.94 (s, 3H), 0.85- 0.76 (m, 4H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 200.45, 157.40, 154.95, 148.73, 141.74, 140.35, 137.79, 132.62, 129.97, 129.24, 125.96, 124.20, 122.58, 113.78, 111.35, 78.89, 57.48, 55.14, 47.83, 43.91, 37.05, 34.67, 34.54, 31.60, 31.24, 29.65, 27.57, 26.25, 25.29, 25.18, 22.67, 17.19, 14.16. **HR-FT-MS calcd for C₃₀H₃₃O₅N₁Na₁ 510.2251 found 510.22454.**

3.4.25 5-bromothiphene enone MMA311 and MMA330:

To a stirred solution of hydroxyl ketone 8 and 8` (0.3g, 0.847mmole) in a dry tetrahydrofuran (THF) (1.7ml), lithiumdiisopropyl amine (LDA) (1.5ml, 3.05mmole) was added in a dropwise at -78°C the reaction mixture was allowed to stir at -78°C for 1 h, then the solution of Para-anisaldehyde (0.206ml, 1.7mmole) in THF (11.33ml) was added to the reaction mixture at -78°C. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of ammonium chloride (NH₄Cl) (25ml). The aqueous layer was extracted using ethyl acetate (3ˣ50ml), dried over sodium sulfate anhydrous (Na₂SO₄), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl

acetate in hexane) to obtained 5-bromothiphene enone MMA311 and MMA330 (0.35g, 78.5%).

Compound MMA311 ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.86 (m, 1H), 7.19 – 7.10 (m, 2H), 7.08 (m, 1H), 6.77 – 6.69 (m, 2H), 6.66 (m, 1H), 6.03 (m, 1H), 4.47 (s, 1H), 3.80 (s, 3H), 2.98 – 2.84 (m, 2H), 2.39 – 2.29 (m, 2H), 2.07 (m, 1H), 1.94 (m, 1H), 1.70 – 1.51 (m, 3H), 1.44 (m, 1H), 1.38 – 1.19 (m, 4H), 1.08 (s, 3H), 0.98 – 0.88 (m, 3H), 0.89 (s, 1H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 200.39, 157.42, 155.26, 141.39, 137.85, 136.24, 132.97, 132.82, 131.45, 129.46, 126.02, 117.97, 117.15, 113.83, 111.36, 78.54, 57.42, 55.17, 47.81, 43.93, 37.13, 36.70, 34.25, 31.22, 27.61, 26.80, 25.34, 17.22. **HR-FT-MS calcd for C$_{28}$H$_{31}$O$_3$Br$_1$Na$_1$S$_1$ 549.1069 found 549.10931.**

Compound MMA330 ^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.84 (m, 1H), 7.17 – 7.08 (m, 2H), 7.06 (m, 1H), 6.75 – 6.67 (m, 2H), 6.64 (m, 1H), 6.01 (m, 1H), 4.45 (s, 1H), 3.78 (s, 3H), 2.96 – 2.82 (m, 2H), 2.37 – 2.27 (m, 2H), 2.05 (m, 1H), 1.92 (m, 1H), 1.68 – 1.49 (m, 3H), 1.42 (m, 1H), 1.36 – 1.17 (m, 4H), 1.06 (s, 3H), 0.96 – 0.86 (m, 3H), 0.87 (s, 1H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 200.37, 157.40, 155.24, 141.37, 137.83, 136.22, 132.95, 132.80, 131.43, 129.44, 126.01, 117.95, 117.13, 113.81, 111.34, 78.52, 57.40, 55.15, 47.79, 43.91, 37.11, 36.68, 34.23, 31.19, 27.60, 26.79, 25.32, 17.21. . **HR-FT-MS calcd for C$_{28}$H$_{31}$O$_3$Br$_1$Na$_1$S$_1$ 549.1069 found 549.1019.**

3.5   References:

1. Newman, D.J. and G.M. Cragg, *Natural Products as Sources of New Drugs over the Last 25 Years⊥.* Journal of natural products, 2007. **70**(3): p. 461-477.

2. Ji, H.F., X.J. Li, and H.Y. Zhang, *Natural products and drug discovery.* EMBO reports, 2009. **10**(3): p. 194-200.

3. Koehn, F.E. and G.T. Carter, *The evolving role of natural products in drug discovery.* Nat Rev Drug Discov, 2005. **4**(3): p. 206-220.

4. Chen, X., et al., *Biological activities and potential molecular targets of cucurbitacins: a focus on cancer.* Anticancer Drugs, 2012. **23**.

5. Ahmed, M.S., L.C. Kopel, and F.T. Halaweish, *Structural Optimization and Biological Screening of a Steroidal Scaffold Possessing Cucurbitacin-Like Functionalities as B-Raf Inhibitors.* ChemMedChem, 2014. **9**(7): p. 1361-1367.

6. Ahmed, M.S. and F.T. Halaweish, *Cucurbitacins: potential candidates targeting mitogen-activated protein kinase pathway for treatment of melanoma*, in *J Enz Inhib Med Chem*. 2013.

7. Bellissimo, F., et al., *Diagnostic and therapeutic management of hepatocellular carcinoma.* World journal of gastroenterology, 2015. **21**(42): p. 12003.

8. Altekruse, S.F., K.A. McGlynn, and M.E. Reichman, *Hepatocellular Carcinoma Incidence, Mortality, and Survival Trends in the United States From 1975 to 2005.* Journal of Clinical Oncology, 2009. **27**(9): p. 1485-1491.

9. Qian, L., et al., *Matrine derivative WM130 inhibits hepatocellular carcinoma by suppressing EGFR/ERK/MMP-2 and PTEN/AKT signaling pathways.* Cancer Letters. **368**(1): p. 126-134.

10. Whittaker, S., R. Marais, and A. Zhu, *The role of signaling pathways in the development and treatment of hepatocellular carcinoma.* Oncogene, 2010. **29**(36): p. 4989-5005.

11. Ohata, K., et al., *Hepatic steatosis is a risk factor for hepatocellular carcinoma in patients with chronic hepatitis C virus infection.* Cancer, 2003. **97**(12): p. 3036-3043.

12. Chattopadhyay, D., D.M. Manas, and H.L. Reeves, *The Development of Targeted Therapies for Hepatocellular Cancer.* Current Pharmaceutical Design, 2007. **13**(32): p. 3292-3300.

13. Liu, Y., et al., *Blockage of epidermal growth factor receptor by quinazoline tyrosine kinase inhibitors suppresses growth of human hepatocellular carcinoma.* Cancer Letters, 2007. **248**(1): p. 32-40.

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

15. Kudo, M., *Signaling Pathway and Molecular-Targeted Therapy for Hepatocellular Carcinoma.* Digestive Diseases, 2011. **29**(3): p. 289-302.

16. Furuse, J., *Growth factors as therapeutic targets in HCC.* Critical Reviews in Oncology/Hematology, 2008. **67**(1): p. 8-15.

17.   Gabrielsen, M., et al., *Cucurbitacin covalent bonding to cysteine thiols: the filamentous-actin severing protein Cofilin1 as an exemplary target.* Cell Communication and Signaling, 2013. **11**(1): p. 58.

18.   Radjasa, O.K., et al., *Highlights of marine invertebrate-derived biosynthetic products: Their biomedical potential and possible production by microbial associants.* Bioorganic & Medicinal Chemistry, 2011. **19**(22): p. 6658-6674.

19.   Dang, G.v., B.M. Rode, and H. Stuppner, *Quantitative electronic structure-activity relationship (QESAR) of natural cytotoxic compounds: maytansinoids, quassinoids and cucurbitacins.* European Journal of Pharmaceutical Sciences, 1994. **2**(5): p. 331-350.

20.   Matsuda, H., et al., *Cucurbitane-Type Triterpenes with Anti-proliferative Effects on U937 Cells from an Egyptian Natural Medicine, <i>Bryonia cretica</i>: Structures of New Triterpene Glycosides, Bryoniaosides A and B.* Chemical and Pharmaceutical Bulletin, 2010. **58**(5): p. 747-751.

21.   Miró, M., *Cucurbitacins and their pharmacological effects.* Phytotherapy Research, 1995. **9**(3): p. 159-168.

22.   Kopel, L.C., M.S. Ahmed, and F.T. Halaweish, *Synthesis of novel estrone analogs by incorporation of thiophenols via conjugate addition to an enone side chain.* Steroids, 2013. **78**(11): p. 1119-1125.

23.   Parihar, S., et al., *Gallic acid based steroidal phenstatin analogues for selective targeting of breast cancer cells through inhibiting tubulin polymerization.* Steroids, 2012. **77**(8–9): p. 878-886.

24.   Bunyathaworn, P., et al., *Further study on synthesis and evaluation of 3, 16, 20-polyoxygenated steroids of marine origin and their analogs as potent cytotoxic agents.* Steroids, 2010. **75**(6): p. 432-444.

25.   Tietze, L.F., et al., *A Novel Approach in Drug Discovery: Synthesis of Estrone–Talaromycin Natural Product Hybrids.* Chemistry – A European Journal, 2000. **6**(20): p. 3755-3760.

26.   Sinha, S., et al., *A Lipid-Modified Estrogen Derivative that Treats Breast Cancer Independent of Estrogen Receptor Expression through Simultaneous Induction of Autophagy and Apoptosis.* Molecular Cancer Research, 2011. **9**(3): p. 364-374.

27.   Mueck, A.O. and H. Seeger, *2-Methoxyestradiol—Biology and mechanism of action.* Steroids, 2010. **75**(10): p. 625-631.

28.   Yajima, A. and K. Mori, *Synthesis and Absolute Configuration of (−)-Phytocassane D, a Diterpene Phytoalexin Isolated from the Rice Plant, Oryza sativa.* European Journal of Organic Chemistry, 2000. **2000**(24): p. 4079-4091.

29.   Shi, G. and Y. Xu, *A convenient synthesis of 1, 1-difluoroallenes from trifluoromethylketones via the shapiro reaction pathway.* Journal of Fluorine Chemistry, 1989. **44**(1): p. 161-166.

30.   Heretsch, P., S. Rabe, and A. Giannis, *A Biomimetic Approach to C-nor-D-homo-Steroids.* Journal of the American Chemical Society, 2010. **132**(29): p. 9968-9969.

31.   Montenegro, H.E., et al., *Two Versatile and Parallel Approaches to Highly Symmetrical Open and Closed Natural Product-Based Structures.* Chemistry – A European Journal, 2010. **16**(12): p. 3798-3814.

32.   Paquette, L.A., M.E. Okazaki, and J.C. Caille, *A formal total synthesis of (.+-.)-laurenene.* The Journal of Organic Chemistry, 1988. **53**(3): p. 477-481.

33.   Kevin R. Campos, B.H., *Process For Making Lactam Tachkinin Receptor Antagonists* in *US patent* 2009: United States

Chapter Four

Design, Synthesis and Biological Screening of CUCUS-inspired Estrone Analogues with Sulfamoyl Moiety at C-3 Targeting Epidermal Growth Factor toward Treatment of Hepatocellular Carcinoma

4.1   Introduction:

The study of molecular docking provided background information about the behavior of the CUCUS-inspired estrone analogues in the binding pocket of the EGFR and pointed out at the analogues that will be synthesized and then biologically screened. Number of research groups have done chemical modification on the main structure of cucurbitacins such as installing tosyl functional groups in the ring A, followed by addition of halogens [1, 2] and alkylation reaction [3, 4]. Various synthetic approaches were conducted to totally synthesized cucurbitacins but all of these trays were failed to provide fully synthesized cucurbitacin [5, 6]. However, there are some successful studies on modification of cucurbitacins including quantitative structure activity relationship (QSAR) study of semi synthetic cucurbitacin by Bartalis and Halaweish [4], partial synthesis of cucurbitacin B and D by Jung [7], synthesis of hexanorcucurbitacin compounds by Ryu [8] and synthesis of dihydrocucurbitacin B by Lang [1]. However, some of these synthetic paths are impracticable for cucurbitacins pharmacophores to be modified for biological activities.

There have been few studies and researches trying to install the cucurbitacins enone side chain into the C-17 of the steroid main structure **(Fig4.1)**. However, assembling the enone side chain is challenging mission because of interfering functional groups on the main skeleton which may affect each other during the chemical reactions. Ourisson et al. 1973 using Grinard reagent, tetrhydro pyranyl ether, tried to install cucurbitacin I side chain into C-17 of tetracyclic steroid but he ended up with racemic mixture of S- and R-isomers at C-20 [9]. An attempting to modify lanastrol, which is tetracyclic triterpins, Paryzek et al. 1981 resulted in the synthesis of different analogue which is 4, 4, 14α-trimethyl-19(10-8β) abeo-steroids with side chain at C-2 and C-5 [10]. In another study Levy et al. tried to modify 3β-acetoxy-lanost-9 chemically to cucurbitacin structure; however, his attempts did not provide the expected results and instead they got organo-tin structure [11].

The main structure of cucurbitacin which known as [19(10→9β) abeo-10α-lanost-5-en][12] is similar in the structure to that of steroids in the fact of the presence of four rings system. However, they are structurally different from each other in the ring system that C-19 methyl group of cucurbitacin is at C-9 and ring C and B fusion. So, estrone scaffold was selected to install the cucurbitacin enone side chain at C-17 **(Fig.4.1)**. In chapter 3, we discussed the importance of substitution at C-3 of estrone structure for the biological activities, particularly the presence methoxy at C-3 [13]. However, these estrone derivatives with cucurbitacin enone side chain at C-17 and methoxy at C-3 are slightly non-polar as proved in our group previously [14]. In addition, metabolism of the methoxy group at C-3 by the CYP450 O-dealkylation might retrieve the estrogenic activity of the estrone as feminizing horman. So, substituting the methoxy group at C-3 with more polar groups

such as sulfamoyl will increase the polarity of the synthesized compounds and in the same time it will prevent the expected metabolism of the C-3 methoxy group. All of these structural modifications lead to dramatically change in the pharmacokinetic and biological activities of the CUCUS-inspired estrone analogues.

Cucurbitacin

Steroidal Scofold

Cucurbitacin	R_1	R_2	R_3	R_4	R_5	$\Delta^{1,2}$	$\Delta^{23,24}$
A	OH	=O	MeOH	H	AC	-	+
B	OH	=O	CH_3	H	AC	-	+
C	H	OH	MeOH	H	AC	-	+
D	OH	=O	CH_3	H	OH	-	+
E	OH	=O	CH_3	H	AC	+	+
F	OH	OH	CH_3	H	OH	-	+
H	OH	=O	CH_3	OH	OH	-	-
I	OH	=O	CH_3	H	OH	+	+
J	OH	=O	CH_3	OH	OH	+	-

Figure 4.1 general structure of cucurbitacin and estrone.

Installation of sulfamoyl moiety into the estrone skeleton has got the attention due to its ability to change the biological activities. Leese et al proved that estratrine sulfmate compounds improved the anti-cancer activities and increased the oral bioavailability [15]. In addition, other studies demonstrated the ability of estrone derivatives with sulfamoyl group at C-3 to have effect on different cancer cell lines such as breast cancer cell line [16], and effect on pharmacokinetic of the modified estrone structure [17]. All of these studies and results proved that installing sulfamoyl functional group into C-3 of the estrone scaffold along with the cucurbitacin enone side chain at C-17 will enhance the anti-cancer activities as well as the pharmacokinetic properties of the CUCUS-inspired estrone analogues.
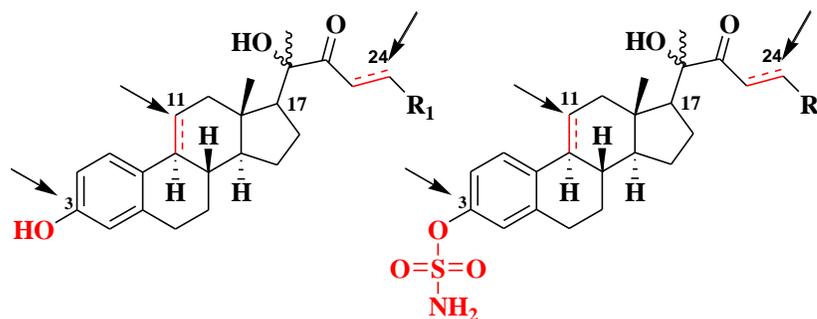
R1= C(CH$_3$)$_2$ OH, P-PhF, P-PhCl, P-PhBr
P-PhCF$_3$, P-PhNO$_2$, 5-Bromo-2-thiophene,
5-Bromo-2-furan.

Figure 4.2 proposed positions for estrone structure modification.

In order to find the hit analogues to identify the lead compounds (LDs) targeting EGFR, bioisosterism process are used to design systematically the CUCS-inspired estrone analogues. Therefore, a virtual library of 900 compounds were created including esterone derivatives with cucurbitacins enone side chain at C-17 and sulfamoyl or hydroxyl groups at C-3, known EGFR inhibitors a nd different types of cucurbitacins, to be involved in the molecular docking studies [18]. Estrone main structure were studied and investigated in three different positions 1) assembling of cucurbitacin enone side chain at C-17, 2) installation of sulfamoyl or hydroxyl moieties at C-3 and 3) installation of various aliphatic and aromatic substituents at C-25 (**Fig.4.2**). OpenEye® scientific software were utilized for the molecular docking studies including fast executive docking (FRED), OMEGA and VIDA.

4.2   Results and Discussion:

4.2.1   Results of Molecular Docking with EGFR:

The molecular docking calculations of CUCUS-inspired estrone analogues that consist of cucurbitacins enone side chain at C-17, sulfomyl or hydroxyl groups at C-3 and different aliphatic and aromatic functional groups at C-25, such as **MMA-240, MMA-242, MMA-294, MMA-295, MMA-297, MMA-300, MMA-301, MMA-306, MMA-307, MMA-308, MMA-309, MMA-310, MMA-312, MMA-313 AND MMA-314 (Fig. 2.8 and 2.9**), showed an outstanding binding mode toward the EGFR binding site compare to the known EGFR inhibitor, such as Erlotinib, by making H-bonds and hydrophobic interactions with the amino acids residues of the EGFR binding pocket (**Fig. AP-3**).

Analogues that contain cucurbitacin enone side chain at C-17 along with sulfamoyl moiety at C-3 such as **MMA-240** and **MMA242** showed a distinguishable binding affinity toward the EGFR binding site by associating in H-bond with **ALA:719:A** inside the EGFR pocket along with hydrophobic interactions with other amino acids inside the same binding pocket; while analogue with the same enone side chain at C-17 but with hydroxyl group at C-3 instead of sulfamoyl as in **MMA301** demonstrated less binding mode with only hydrophobic interactions **(Fig.4.3).** This outstanding result proved that the presence of sulfmoyl moiety at C-3 along with the cucurbitacin enone side chain is essential for the binding affinity toward the EGFR binding pocket.
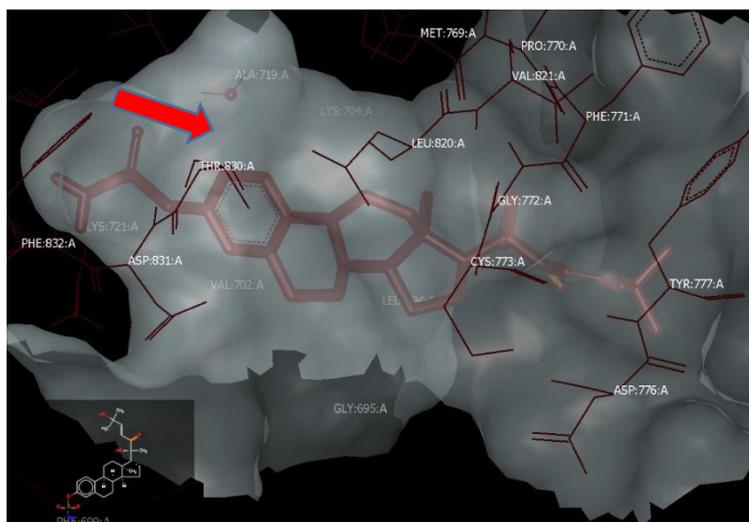
Figure 4.3 Visual representation of A) MMA-240 (Red) B) MMA-301 (purple) in the EGFR ATP-binding site.
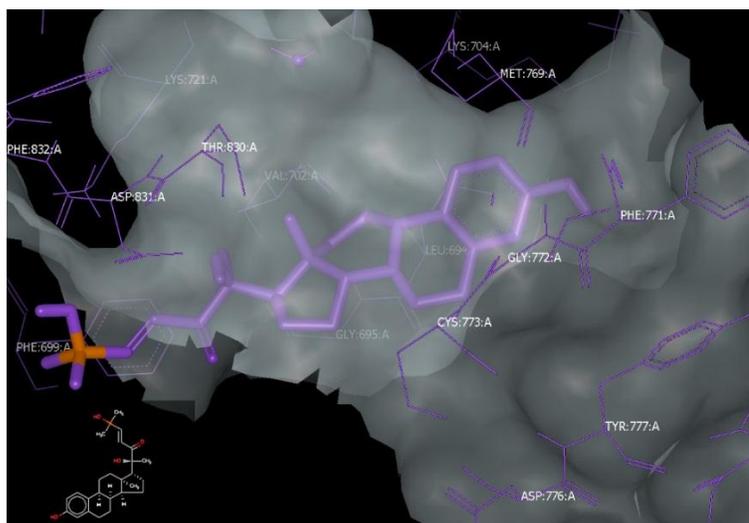
On the other hand, compounds that contain cucurbitacin enone side chain at C-17 along with different aromatic substituents at C-25 and sulfamoyl group at C-3 such as **MMA294, MMA295, MMA297, MMA300, MMA307** and **MMA312** showed better binding mode toward the EGFR binding pocket by making H-bonds and hydrophobic interactions compare to the analogues that contain the same aromatic enone side chain but

with hydroxyl group at C-3 such as **MMA269, MMA306, MMA309, MMA310, MMA313** and **MMA314** which only showed hydrophobic interactions with the amino acids of the EGFR binding pocket **(Fig.4.4).**

In conclusion, CUCUS-inspired estrone analogues that contain cucurbitacin enone side chain at C-17 along with aliphatic and aromatic substituents at C-25 and sulfamoyl moiety at C-3 demonstrated significant binding mode toward the 3D crystal structure of EGFR binding side by making H-bonds and hydrophobic interactions. In addition, the presence of hydroxyl group at C-3 of the estrone scaffold along with cucurbitacin enone side chain at C-17 showed lower binding affinity toward the same binding site of the EGFR. All of these signs and results demonstrated the important of assembling of the sulfamoyl group at C-3 for the binding affinity and may be for the biological activities. The results of the molecular docking was the core to move to the synthesis of the most promising analogues to be investigated biologically in next step.
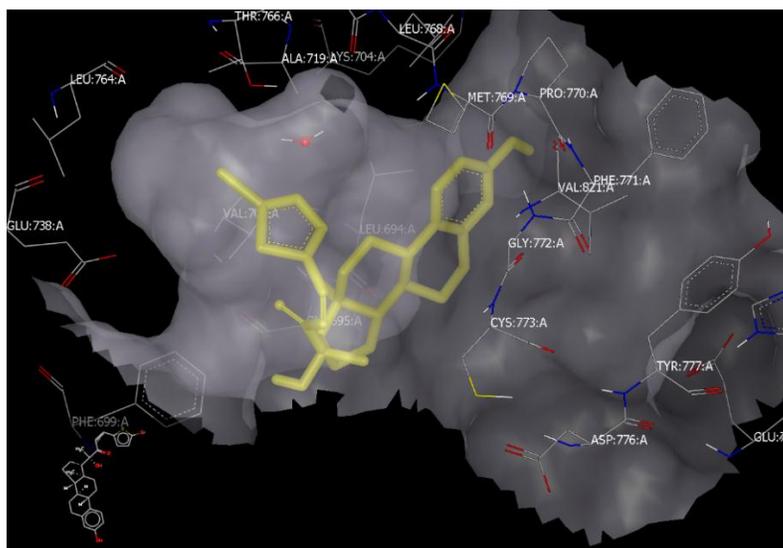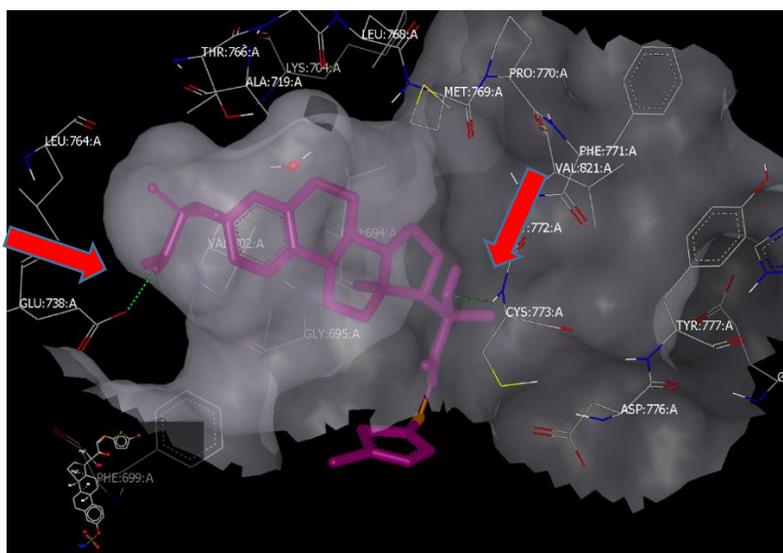
Figure 4.4 Visual representation of A) MMA-312 (purple) B) MMA-313 (yellow) in the EGFR ATP-binding site.

4.2.2   Synthesis of MMA Analogues:

In order to investigating the biologically activities of CUCUS-inspired estrone analogues (CIEA), chemical reactions were used to assemble the cucurbitacins enone side chain at C-17 along with sulfamoyl or hydroxyl groups at C-3 of the estrone skeleton

(**Fig.4.1**). The challenges on conducting these chemical reactions were started by making the decision what step should the sulfamoyl group at C-3 to be installed during the process of assembling the cucurbitacin enone side chain at C-17. Due to the diversity of the chemical reactions that will be used to install the cucurbitacin enone side chain into the estrone scaffold, the concept of trial and error were used to optimaize the complete synthetic schemes. In the beginning, installation of sulfamoyl functional group at C-3 of the estrone structure were done utilizing sulfamoyl chloride in the presence of N, N-dimethylacetamide to provide 3-sulfamoyl estrone (**Scheme 4.1**) [19]. However, the resulted 3-sulfamoyl estrone was found to be polar and cannot be used for the next chemical reactions to install cucurbitacin enone side chain. Therefore, protection of hydroxyl group at C-3 of the estrone main structure with tert-butyldimethylsilyl chloride were carried out first in the presence of imidazole and dimethyl formamide to provide estrone tert-butyldimethylsilyl ether **1** [20]. Wittig reaction was utilized on estrone tert-butyldimethylsilyl ether 1 to prepare alkene **2** by reacting the phosphonium yield with C-17 ketone [21].
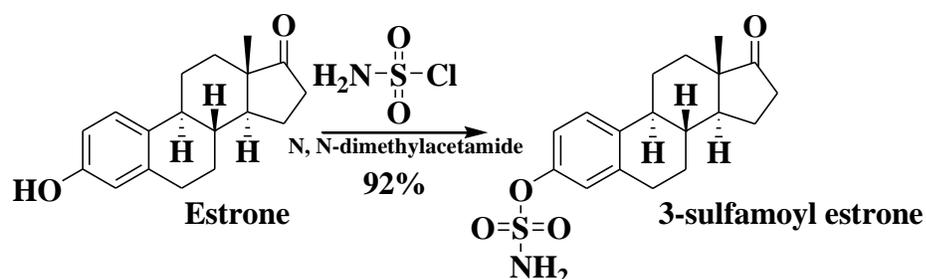
Scheme 4.1 Synthesis of 3-sulfamoyl estrone.

Alcohol 3 was produced by hydroboration reaction on alkene **2** using 9-Borabicyclo[3.3.1]nonane (9-BBN) in presence of tetrahydrofuran (THF) as solvent for overnight followed by the drop wise addition of sodium hydroxide (NaOH) and hydrogen peroxide (H_2O_2) [14]. This was followed by oxidation of alcohol 3 utilizing pyridinum chlorochromate (PCC) to give ketone 4 in 91% yield [22, 23]. Treatment of ketone 4 with trimethylsilyl cyanide (TMSCN) in the presence of zinc iodide (ZnI_2) as Lewis acid to activate the ketone to provide the desired disteriomer cyanohydrin 5 [24]. Ahmed et al. was the first to form cyanohydrin using OTMS in this scaffold [14]. Addition of methyl lithium to cyanohydrin 5 in the presence of ether as solvent to produce the desired disteriomer α-hydroxyl ketone 6 **(Scheme 4.2)** [14]. The 3-D structure of the Compound **MMA242** were confirmed using x-ray crystallography as shown in **Figure 4.5**, which demonstrated the R-configuration.
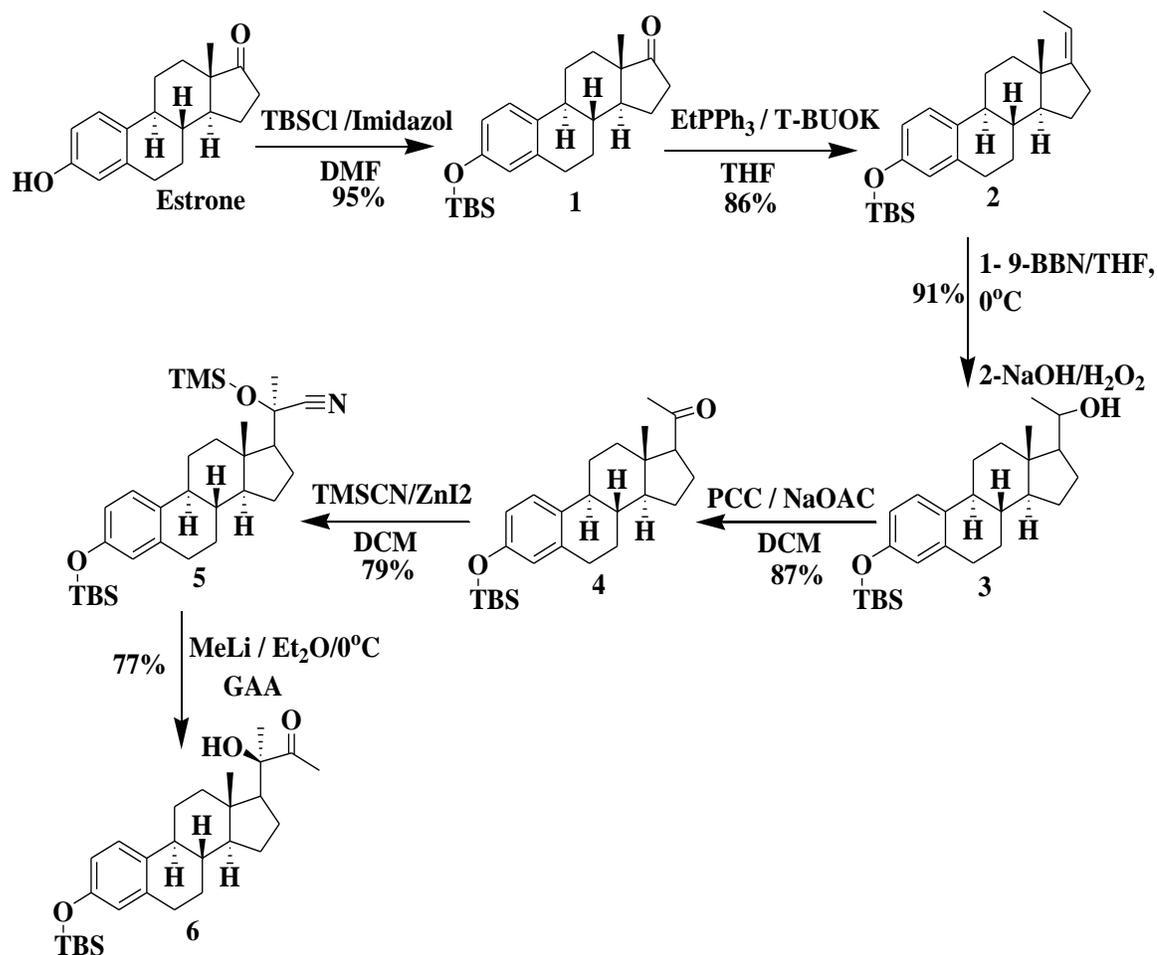
Scheme 4.2 Synthesis of α-hydroxyl ketone **6**.
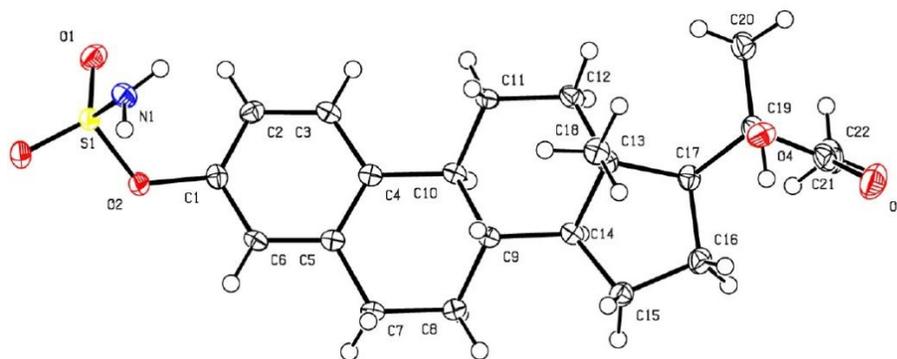
Figure 4.5 X-ray crystallography of α-hydroxyl ketone **MMA242**.

α-Hyd roxyl ketone **6** was utilized as a starting material to install various functional groups including cucurbitacin D enone side chain at C-17 with different aromatic moieties at C-25 and sulfamoyl at C-3. First, α-hydroxyl ketone **6** was used as starting material for Aldo condensation reaction with the prepared aldehyde 9 to prepare compound **10**, which is protected estrone with protected enone side chain [14]; followed by deportation of protected hydroxyl groups at C-3 and C-25 using tetra-n-butylammonium fluoride (TBAF) to produce estrone derivative **MMA301** with hydroxyl group at C-3 and cucurbitacin D enone side chain at C-17 [25]. Compound **MMA240** was prepared to by adding sulfamoyl chloride to compound **MMA301** in presence of N, N-dimethyl acetamide as solvent; sulfamoyl chloride selectively reacted with phenolic hydroxyl group as shown in **scheme 4.3** [19]. Aldehyde **9** was prepared in our group previously as shown in **scheme 4.4** [14].
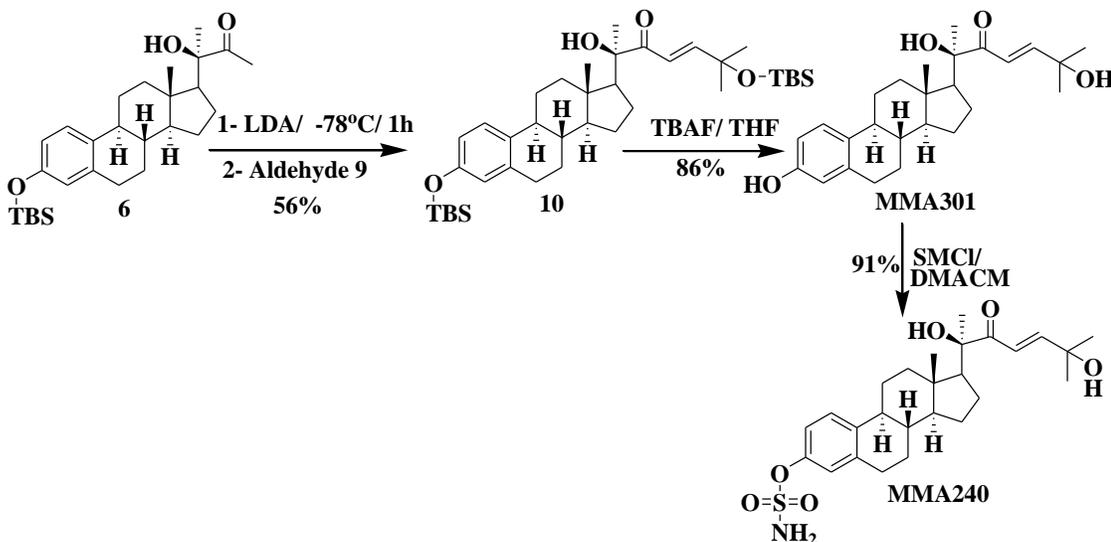
Scheme 4.3 Synthesis of compounds **MMA301** and **MMA240**.
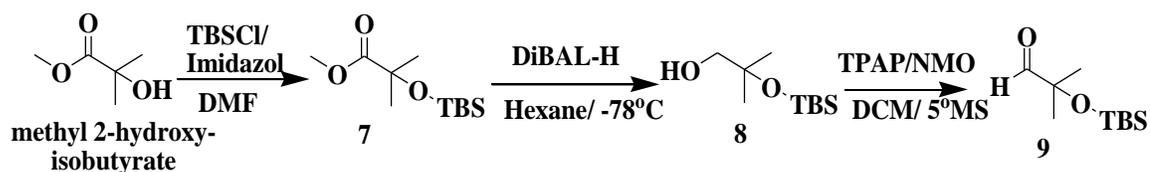
Scheme 4.4 Synthesis of Aldehyde **9**.

In addition, α-hydroxyl ketone **6** was used to prepare analogue **MMA280** by deportation of TBS-protecting group utilizing tetra-n-butylammonium fluoride (TBAF) [25]; followed by preparing compound **MMA242** by adding sulfamoyl chloride in presence of N, N-dimethylacetamide as solvent **(Scheme 4.5)** [19].
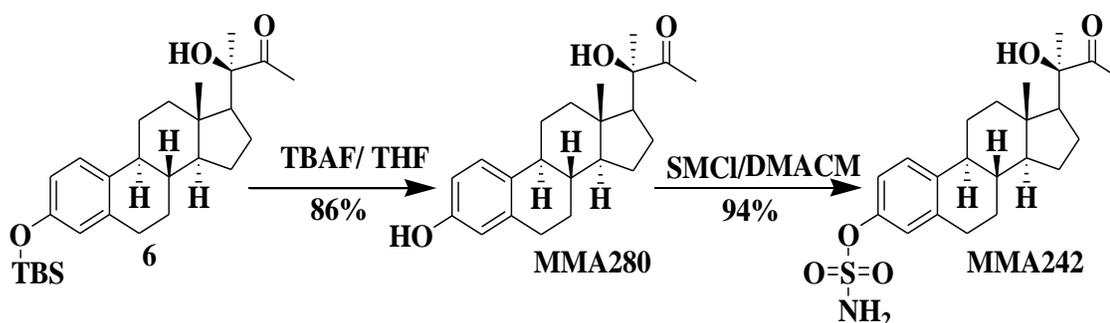
Scheme 4.5 synthesis of compounds **MMA280** and **MMA242**.

Various number of aromatic enone side chains, such as para-methoxybenzene enone side chain along with hydroxyl group at C-3 as in **MMA268** or sulfamoyl group at C-3 as in **MMA267**, para-flurobenzene enone side chain along with hydroxyl group at C-3 as in **MMA269** or sulfamoyl group at C-3 as in **MMA271**, para-chlorobenzene enone side chains along with hydroxyl group at C-3 as in **MMA309** or sulfamoyl group at C-3 as

Figure 4.6 Synthesis of estrone derivatives with various aromatic enone side chain along with hydroxyl or sulfamoyl groups at C-3.
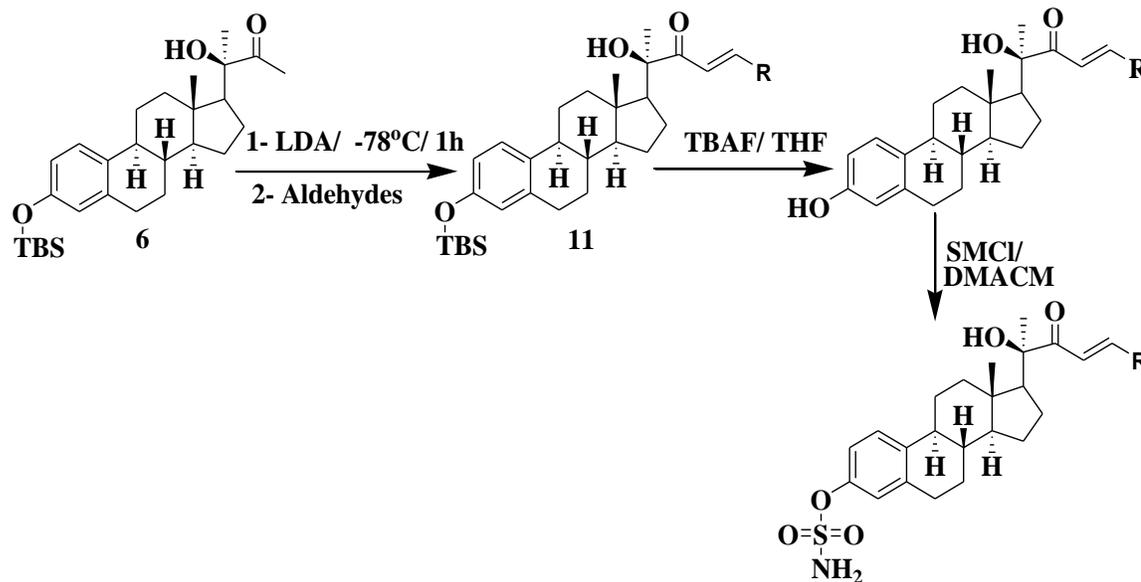
Aldehyde	Compound 12	Compound 13
R= p-methoxybenzaldehyde	MMA268	MMA267
R= p-flurobenzaldehyde	MMA269	MMA271
R= p-chlorobenzaldehyde	MMA309	MMA294
R= p-bromobenzaldhyde	MMA310	MMA295
R= p-trifluromethyl - benzaldhyde	MMA308	MMA300
R= p-nitrobenzaldehyde	MMA306	MMA307
R= 5-bromo-2-furan aldehyde	MMA297	MMA314
R= 5-brom-2-thiophene aldehyde	MMA313	MMA312

in **MMA294**, para-bromobenzene enone side chain along with hydroxyl group at C-3 as in **MM310** or sulfamoyl group at C-3 as in **MMA295**, 5-bromofuran enone side chain along with hydroxyl group at C-3 as in **MMA297** or sulfamoyl group at C-3 as in **MMA314**, para-trifluromethyl benzene along with hydroxyl group at C-3 as in **MMA308** or sulfamoyl group at C-3 as in **MMA300**, para-nitrobenzene enone side chain along with hydroxyl group at C-3 as in **MMA306** or sulfamoyl group at C-3 as in **MMA307** and 5-bromothiphene enone side chain along with hydroxyl group at C-3 as in **MM313** or sulfamoyl or sulfamoyl group at C-3 as in **MMA312**, were assembled in the C-17 of α-hydroxyl ketone 6 using Aldol condensation reaction to form the related compound of 11 followed by the deportaction reaction of TBS group with TBAF [25] followed by addition of sulfamoyl chloride **(Scheme 4.6)** [19]**.**

4.2.3 Biological Evaluation of the Synthesized Compounds:

The biological studies for the synthesized compounds started by in vitro study to identify their capability to inhibit the EGFR. MTT cell viability assay were utilized to find the cytotoxicity of the synthesized compounds against the hepatocellular carcinoma cell line (HepG2) and that expressed as IC_{50} (inhibition concentration). Compound **MMA242**, which contain α-hydroxyl ketone on C-17 along with sulfamoyl group at C-3, showed potent inhibitory activity toward the HepG2 cell line with IC_{50} value of 3μM compare to the known EGFR inhibitor, Erlotinib with 25 μM (Table 4.1). Compounds with aliphatic enone side chain at C-17 along with sulfamoyl at C-3 such as **MMA240** demonstrated modest inhibitory activity with IC_{50} value of 8μM; while the same compound but with

hydroxyl group at C-3 instead of sulfamoyl such as MMA301 lost their cytotoxicity completely. These results verify the importance of possessing sulfamoyl group at C-3, which we confirmed at the molecular docking study **(Fig. 4.3)**. Analogues that contain cucurbitacin enone side chain at C-17 with various aromatic and heterocyclic functional groups at C-25 along with sulfamoyl moiety at C-3 such as **MMA267, MMA271**, **MMA294, MMA295, MMA297, MMA300**, **MMA307** and **MMA312** demonstrated various cytotoxicity levels with IC_{50} values 12 µM, 8 µM, 11.5 µM, NA, 1 µM, 3 µM, 10 µM, 8 µM; respectively **(Table 4.1)**. On the other hand, compounds with same enone side chains at C-17 but with hydroxyl group at C-3 Such as **MMA268, MMA269**, **MMA306, MMA308, MMA309, MMA310**, **MMA313** and **MMA314** showed various cytotoxicity with IC_{50} values 7 µM, 2 µM, 8 µM, 8 µM, NA, 6 µM, 9 µM and 1.5 µM; respectively **(Table 4.1).** All of the previous results proved the biological importance of possessing the sulfamoyl moiety at the C-3 of the estrone main structure along with various aliphatic, aromatic and heterocyclic functionalities at the enone side chain compare to the first set synthesized compounds that contain methoxy group at C-3 **(Table 3.1)**.

Table 4.1 IC$_{50}$ values of the synthesized compounds.

Compound	IC$_{50}$ (μM)	Compound	IC$_{50}$ (μM)
Erlotinib	25 μM	MMA310	6 μM
MMA240	8 μM	MMA295	NA
MMA301	NA	MMA297	1 μM
MMA242	3 μM	MMA314	1.5 μM
MMA267	12 μM	MMA300	3 μM
MMA268	7 μM	MMA308	8 μM
MMA269	2 μM	MMA306	8 μM
MMA271	8 μM	MMA307	10 μM
MMA309	NA	MMA313	9 μM
MMA294	11.5 μM	MMA312	8 μM

Figure 4.7 Chart represent the ability of CIEA to inhibit the growth of HepG2 cell line.

4.3   Conclusion:

CUCUS-inspired estrone analogues were structurally designed utilizing molecular docking technique by assembling the   biologically important pharmacophores such as cucurbitacin's enone side chain at C-17, various aromatic and heterocyclic substituents at C-25 and sulfamoyl or hydroxyl groups at C-3 to be explored biologically as anti-cancer candidates toward the treatment of Hepatocellular carcinoma (HCC) through the inhibition of EGFR. Various aliphatic, aromatic and heterocyclic enone side chains were installed chemically on the estrone main skeleton such as isopropanol, para-methoxybenzene, para-flurobenzene,   para-chlorobenzene,   para-bromobenzene,   5-bromofuran,   para-

trifluromethyl, para-nitrobenzene, 5-bromothiphene along with sulfamoyl or hydroxyl groups at C-3 (**Fig. 4.6**). The biological evaluation demonstrated that analogues that contain cucurbitacin enone side chain at C-17 of the estrone along with sulfamoyl group at C-3 such as **MMA240** showed a modest inhibitory activity; while the same compound but with hydroxyl group at C-3 such as **MMA301** lost their biological activity. On the other hand, compounds with various aromatic and heterocyclic enone side chains at C-17 showed wide range of cytotoxicity compare to the known EGFR inhibitor, erlotinib. However, installing sulfamoyl group at C-3 of the CUCS-inspired estrone analogues (CIEA) proved to increase the cytotoxicity in comparison to the same analogues but with methoxy group at C-3. Compound **MMA292** showed the most potent cytotoxicity in HepG2 with IC50 value $1\mu M$.

4.5   Experimental section:

4.5.1   General:

All chemicals and solvents (ACS grades) were provided from Fisher Scientific or Sigma Aldrich and used without any additional purification. All glassware were cleaned, washed and dried in oven for overnight before conducting chemical reactions requires anhydrous environment and nitrogen gas applied at the reaction time. Pre-coated silica gel PE plates were used to analyze the reaction condition and UV-light were also used at 254 or 365 to visualize the chemical reactions spots. All synthetic intermediate and final compounds were purified using column chromatography with 230ˣ400 mesh silica gel. ^{1}H and ^{13}C NMR spectra were using Bruker AVANCE-400 MHZ and 600 MHZ NMR

spectrometer, in CDCl₃ and D-acetone. NMR chemical shifts were presented in δ(PPM) using residual solvent peaks as standards (CDCl₃, 7.26 (H), 77.16 (C)). High resolution mass (HRMS) was gained using thermofinnigan MAT 95XL mass spectrometer at Buffalo mass spectroscopy facility. X-ray crystallography were conducted in University of South Dakota on compound 8 using Brouker APEX[II] diffractometer.

4.5.2   3-Sulfamoyl Estrone:
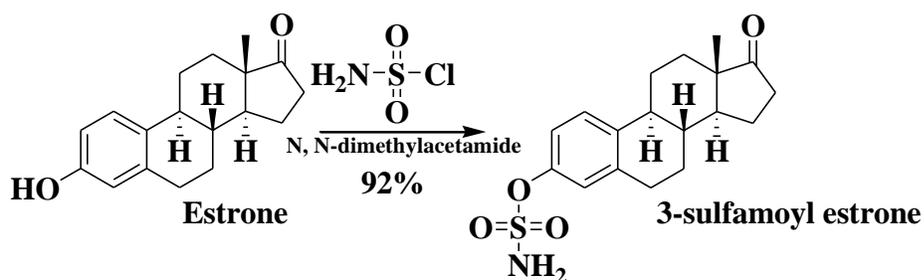
Estrone (0.2g, 0.74mmole) were dissolved in N, N-dimethyacetamide (1.8 ml) then cooled down to 0 °C; followed by addition of sulfamoyl chloride (0.26g, 2.22mmole). the reaction mixture was allowed to be stirred and warm to the room temperature for 18h. amixture of ethyl acetate (30ml) and water (30ml) was added to the reaction mixture and the organic layer was separated, washed by water (3×25ml) followed by the addition of brain (30ml). Sodium sulfate anhydrous (Na₂SO₄) to dry the organic layer, filtered and concentrated in vacuo. Column chromatography (30% ethyl acetate in hexane) to provide white solid of 3-Sulfamoyl Estrone (0.24g, 92%).

¹H NMR (400 MHz, Acetone-d_6) δ 7.35 (dd, J = 8.6, 1.1 Hz, 1H), 7.07 (dd, J = 8.5, 2.6 Hz, 1H), 7.06 − 7.02 (m, 2H), 3.03 − 2.84 (m, 2H), 2.57 − 2.37 (m, 2H), 2.30 (m, 2H), 2.14 − 1.98 (m, 4H), 1.86 (m, 1H), 1.76 − 1.34 (m, 5H), 0.90 (s, 3H).

¹³C NMR (101 MHz, Acetone-d_6) δ 219.62, 149.44, 139.22 (d, J = 6.2 Hz), 127.46, 123.03, 120.25, 51.07, 48.40, 45.01, 38.90, 36.11, 32.56, 30.06, 27.02, 26.56, 22.16, 14.15.

4.5.3 Estrone tert-butyldimethylsilyl ether 1:
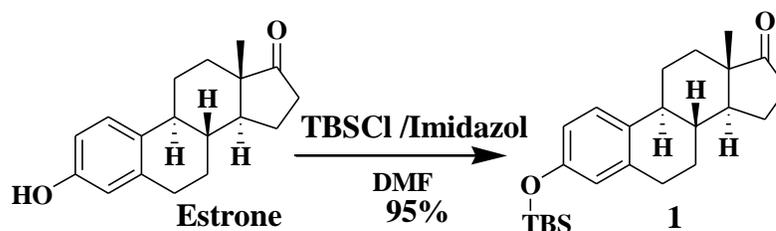
To a stirred solution of estrone (5 g, 18.5 mmole) in DMF (75 ml); followed by the addition of imidazole (3.5 g, 50.85 mmole), tert-butyldimethylsilyl chloride (TBSCl) (4.2 g, 27.75 mmole). The reaction stirred at room temperature for 24h. The reaction solvent were evaporated to provide brown oil. Column chromatography (40% ethyl acetate in hexane) were used to purify the crude material which give white solid of Estrone tert-butyldimethylsilyl ether 1 (6.75 g, 95%).

¹H NMR (400 MHz, Chloroform-d) δ 6.91 (dd, J = 8.6, 1.0 Hz, 1H), 6.47 − 6.33 (m, 2H), 2.72 − 2.59 (m, 3H), 2.28 (dd, J = 18.7, 8.7 Hz, 1H), 2.16 (dt, J = 14.0, 3.7 Hz,

1H), 2.08 – 1.65 (m, 5H), 1.48 – 1.14 (m, 7H), 0.79 (s, 9H), 0.75 – 0.65 (m, 4H), 0.00 (s, 6H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 219.97, 153.47, 137.57, 132.47, 126.17, 120.02, 117.35, 50.44, 47.98, 44.04, 38.34, 35.87, 31.65, 29.54, 26.62, 25.85 21.63, 18.20, 13.90, -4.30.

4.5.4   Alkene 2:
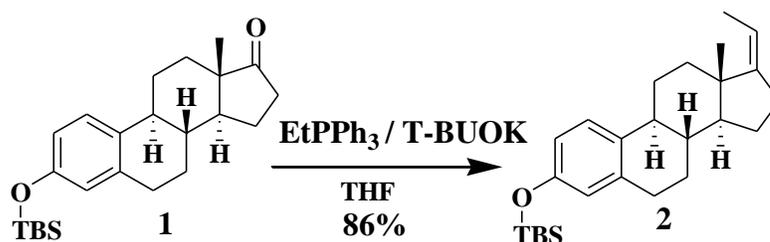
First, the ylide was prepared by adding potassium tert-butoxide in different portions (5.85 g, 52.13 mmole) to the solution of ethyl triphenylphosphonium bromide (20.73 g, 55.85 mmole) in THF (93 ml) at room temperature for 1h. Then solution of Estrone tert-butyldimethylsilyl ether 1 (6.5g, 18.617 mmole) in THF (37 ml) was added to the first mixture and allowed to be stirred for 6h at 70 ^{0}C. The reaction mixture was allowed to cool down to the room temperature, solution of saturated ammonium chloride (NH$_4$Cl) was added to the reaction mixture followed by extraction of the aqueous layer using ethyl acetate (3×100 ml), brain, dried over sodium sulfate anhydrous (Na$_2$SO$_4$), filtrated and concentrated in vacuo. The purification of the crude material was purified by silica gel column chromatography (20% ethyl acetate in hexane) to obtained alkene 2 (5.8 g, 85%).

¹H NMR (400 MHz, Chloroform-*d*) δ 7.13 − 7.04 (m, 1H), 6.62 − 6.49 (m, 2H), 5.12 (m, 1H), 2.87 − 2.69 (m, 3H), 2.45 − 2.12 (m, 3H), 1.95 − 1.82 (m, 1H), 1.77 − 1.60 (m, 5H), 1.58 − 1.42 (m, 1H), 1.42 − 1.21 (m, 3H), 1.09 − 0.99 (m, 1H), 0.96 (s, 9H), 0.86 (m, 4H), 0.16 (s, 6H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 153.38, 150.26, 137.87, 133.35, 126.17, 120.04, 117.25, 113.53, 55.40, 44.69, 44.00, 38.48, 37.41, 31.58, 29.86, 27.76, 27.05, 25.89, 24.33, 18.30, 17.12, 13.33, -4.22.

4.5.5 Alcohol 3:
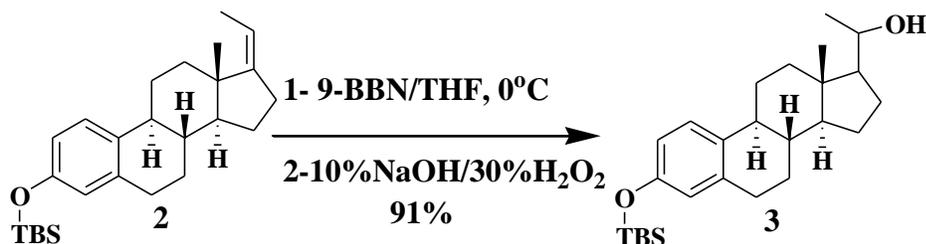
To a solid alkene 2 (5.5 g, 13.865 mmole), 9-BBN (0.5 M in THF, 110.9 ml, 55.46 mmole) was added at room temperature. The reaction mixture was allowed to be stirred for 20h, then cooled down to 0 °C; followed by the addition of 105ml of 10%NaOH and 180ml of 30% H_2O_2 sequentially in drop-wise matter. The reaction mixture then stirred for 1h at 0 °C, then stirred for 1h at room temperature. Ethyl acetate (3×100 ml) was utilized to extract the organic layer, then the collected organic layers washed with 100 ml of saturated sodium thiosulfate, dries over sodium sulfate anhydrous (Na_2SO_4) and concentrated in vacuo. Column chromatography was used to purify the crude material (20% ethyl acetate in hexane) to provide alcohol 3 (5.25 g, 91%).

¹H NMR (400 MHz, Chloroform-*d*) δ 6.94 (m, 1H), 6.47 – 6.32 (m, 2H), 3.56 (m, 1H), 2.75 – 2.52 (m, 2H), 2.25 – 21.98 (m, 1H), 1.79 – 1.44 (m, 4H), 1.40 – 1.12 (m, 4H), 1.14 – 1.01 (m, 6H), 0.98 (d, 3H), 0.82 (s, 9H), 0.61 (s, 3H), 0.02 (s, 6H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 153.25, 137.80, 133.41, 126.15, 119.97, 117.17, 70.45, 58.63, 55.07, 43.93, 42.84, 41.96, 40.05, 38.67, 29.78, 27.92, 27.24, 26.67, 24.31, 23.79, 18.23, 12.56, -4.27.

4.5.6   Ketone 4:
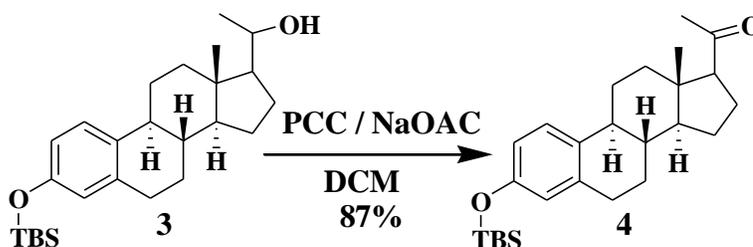
The produced alcohol 3 (5 g, 12.01 mmole) was dissolved in DCM (60.3 ml); followed by the addition of 4 ^{0}A powdered molecular sieves (4.73g), sodium acetate (NaOAc) (4.73 g, 57.6 mmole) and Pyridinium chlorochromate (PCC) (5.2 g, 24.02 mmole). The reaction mixture was allowed to be stirred at room temperature for 2h. Ether was added to the reaction mixture, then filtered over silica gel pad using ether to elute the material. The collected material then concentrated in vacuo. The resulted crude material was purified by column chromatography (20% ethyl acetate in hexane) to give ketone 4 (4.3 g, 87%).

¹H NMR (400 MHz, Chloroform-*d*) δ 6.91 (m, 1H), 6.46 – 6.32 (m, 2H), 2.70 – 2.53 (m, 2H), 2.39 (m, 1H), 2.17 – 1.98 (m, 2H), 1.94 (s, 4H), 1.76 – 1.02 (m, 10H), 0.79 (s, 9H), 0.45 (s, 3H), 0.00 (s, 6H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 209.20, 153.37, 137.74, 132.81, 126.06, 120.01, 117.23, 63.83, 55.72, 44.39, 43.74, 39.04, 38.73, 31.52, 29.67, 27.77, 26.65, 25.79, 24.20, 22.91, 18.21, 13.48, -4.31.

4.5.7 Cyanohydrin 5:
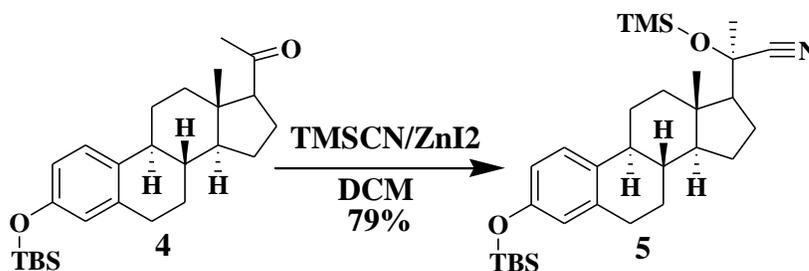
To attired solution of ketone 4 (4 g, 9.7 mmole) in DCM (19.4 ml), zinc iodide (ZnI₂) (0.25 g, 0.78 mmole) was added, followed by the addition of trimethylsilyl cyanide (TMSCN) (2.4 g, 24.25 mmole) in one portion. The reaction mixture then allowed to be stirred at the room temperature for 3h. Then the solvent was evaporated to produce a slurry mixture which was mixed with ethyl acetate/water (2:1, 100 ml). The aqueous layer was extracted using ethyl acetate (3×100 ml), then the extracted organic layer dried over sodium sulfate anhydrous (Na₂SO₄) and concentrated in vacou. Column chromatography was used to purify the crude material (20% ethyl acetate in hexane) to give cyanohydrin 5 (3.95 g, 79%).

¹H NMR (400 MHz, Chloroform-*d*) δ 6.95 (m, 1H), 6.50 – 6.37 (m, 2H), 2.76 – 2.58 (m, 2H), 2.17 – 1.94 (m, 3H), 1.87 (m, 1H), 1.78 – 1.50 (m, 4H), 1.46 (m, 3H), 1.42 – 0.96 (m, 5H), 0.85 (s, 9H), 0.78 – 0.73 (m, 1H), 0.14 (s, 9H), 0.10 (m, 3H), 0.05 (s, 6H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 151.67, 136.05, 131.49, 124.44, 120.43, 118.32, 115.56, 70.77, 58.90, 53.28, 42.24, 42.15, 38.40, 36.53, 33.11, 30.06, 29.21, 28.03, 26.13, 26.01, 24.88, 24.16, 23.72, 23.26, 22.27, 21.12, 19.49, 16.57, 12.61, 11.23, -5.94.

4.5.8 α-hydroxyl ketone 6:
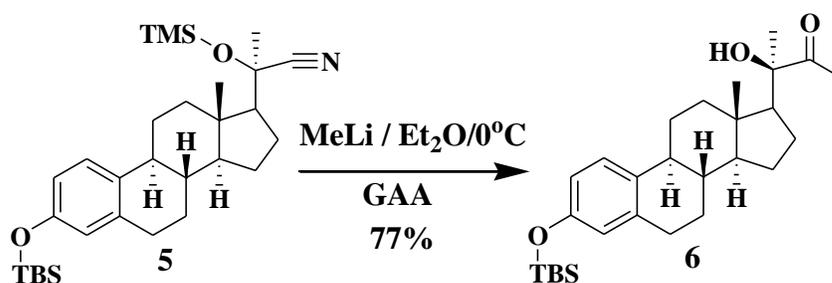
To a stirred solution of cyanohydrin 5 (3.5 g, 6.88 mmole) in a dry ether (20ml), methyl lithium (MeLi) (1.6 M in ether, 17.2ml, 27.52 mmole) was added in dropwise at 0^0C. The reaction mixture was allowed to stir for 2h at 0^0C, then the reaction quenched by adding glacial acetic acid (2.9 ml) in one portion at 0^0C and allowed to stir for 30 min at 0^0C. Sodium bicarbonate solution was added to neutralize the acidic mixture. Dichloromethane (DCM) was used to extract the aqueous layer (3×50ml), dried under sodium sulfate anhydrous (Na_2SO_4), then concentrated under vacuo. The resulted diasteriomers were purified by silica gel column chromatography (100% hexane, 5% ethyl acetate in hexane and 10% ethyl acetate in hexane) to give the α-hydroxyl ketone 6 (2.4 g, 77%).

¹H NMR (400 MHz, Chloroform-*d*) δ 6.99 – 6.87 (m, 1H), 6.47 – 6.34 (m, 2H),

3.79 (s, 1H), 2.73 – 2.52 (m, 3H), 2.16 – 2.04 (m, 2H), 2.02 (s, 3H), 1.74 – 1.57 (m, 2H),

1.55 – 1.31 (m, 3H), 1.27 (s, 3H), 1.26 – 0.93 (m, 5H), 0.79 (s, 9H), 0.74 (s, 3H), 0.00 (s,

6H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 211.71, 153.32, 137.78, 133.10, 126.05,

119.99, 117.18, 80.12, 55.77, 55.14, 44.26, 43.88, 40.72, 38.08, 29.68, 27.72, 26.61, 25.81,

24.67, 23.77, 23.32, 22.15, 18.22, 13.59, -4.29.

4.5.9   Ester 7:

**methyl2-hydroxy
isobutyrate**

7

To a stirred solution of commercially available methyl2-hydroxy isobutyrate (5g,

42.3mmole) in dimethyl formmide (DMF) (12.5ml), tert-butylsilyl chloride (TBSCl)

(7.65g, 50.75mmole) were added followed by the addition of imidazole (7.45g,

110mmole). The reaction was stirred for 24h, then solution of sodium bicarbonate (25ml)

was added. The aqueous layer was extracted using ethyl acetate (3ˣ50ml), dried over

sodium sulfate anhydrous (Na_2SO_4), filtrated, and concentrated under vacuo. The crude

product was purified using silica gel column chromatography (10% ethyl acetate in hexane)

to give ester 7 (7.2g, 73.2%) as a yellow oil.

¹H NMR (400 MHz, CDCl3): d=3.60(s, 3H), 1.34 (s, 6H), 0.80 (s, 9 H), 0.08 ppm (s, 6H);

¹³C NMR (100 MHz, CDCl3): d=175.8, 74.5, 51.4, 28.4, 25.5, 17.9, _3.2 ppm.

4.5.10 Alcohol 8:
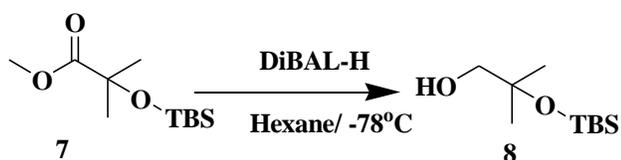
To a stirred solution of ester 7 (7.2g, 30.98mmole) in hexane (98ml), Diisobutylaluminium hydride (DiBAL-H) (56.8ml, 68.16mmole) was added at -78°C in a dropwise matter. The reaction stirred for 30 min. at 0^0C, then for 20 min. at the room temperature. The reaction then cooled back to -78°C, followed by the addition of solution of potassium tartrate (45ml). The reaction was allowed to stir for overnight. The aqueous phase was extracted using ethyl acetate (3^x50ml), dried over sodium sulfate anhydrous (Na_2SO_4), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (20% ethyl acetate in hexane) to give alcohol 8 (5.75g, 90.8%) as colorless liquid.

¹H NMR (400 MHz, CDCl3): d=3.19 (s, 2H), 2.31 (s, 1 H), 1.11 (s, 6 H), 0.76 (s, 9 H), 0.08 ppm (s, 6H).

¹³C NMR (100 MHz, CDCl3): d=76.3, 74.4, 28.5, 28.0, 20.3, 0.3 ppm.

4.5.11 Aldehyde 9:
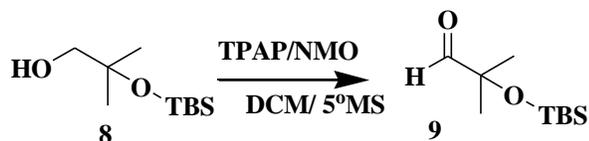
To a stirred solution of alcohol 8 (2g, 9.7858mmole) in a dry dichloromethane (DCM) (98ml), active molecular sieves (4°A, 2.3g) was added, then the reaction mixture was allowed to stir for 5 min. followed by the addition on N-Methylmorpholine N-oxide (NMO) (2.3g, 19.57mmole), then Tetrapropylammonium perruthenate (TPAP) (0.34g, 0.98mmole) was added at 0°C. The reaction stirred at 0°C for 2h, then at room temperature for 24h. Pad of silica gel were used to filtrate the reaction mixture using diethyl ether as eluting solvent. The filtrated material concentrated under vacuo to provide aldehyde 9 (1.45g, 73.3%) as colorless oil.

1H NMR (400 MHz, CDCl3): d=9.44 (s, 1H), 1.17 (s, 6 H), 0.79 (s, 9H), 0.08 ppm (s, 6H).

13C NMR (100 MHz, CDCl3): d=206.4, 80.3, 28.0, 27.2, 20.4, 0.3 ppm.

4.5.12 Protected estrone with protected enone side chain 10:
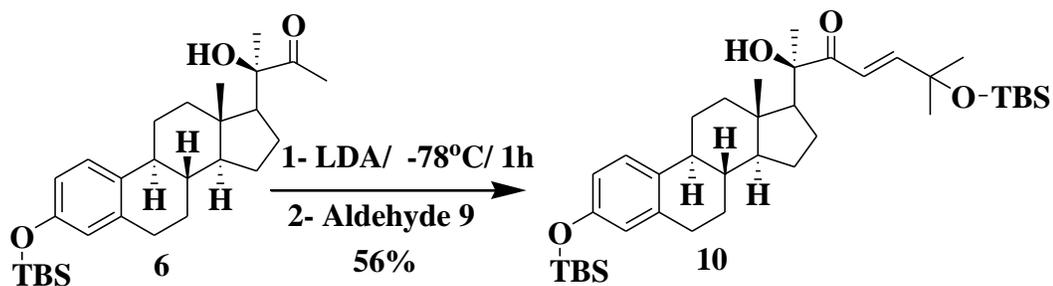
To a stirred solution of α-hydroxyl ketone 6 (0.4g, 0.879 mmole) in a dry tetrahydrofuran (THF) (5.85 ml), lithiumdiisopropyl amine (LDA) (1.6 ml, 3.154 mmole) was added in a dropwise at -78oC the reaction mixture was allowed to stir at -78oC for 1 h, then the solution of aldehyde 9 (0.354 g, 1.75 mmole) in THF (11.5 ml) was added to the reaction mixture at -78oC. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of ammonium chloride (NH4Cl) (25ml). The aqueous layer was extracted using ethyl acetate (3x50ml), dried over sodium sulfate anhydrous (Na2SO4), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to provide protected estrone with protected enone side chain 10 (0.31 g, 56%).

^{1}H NMR (600 MHz, Chloroform-*d*) δ 6.96 – 6.92 (m, 1H), 6.89 (m, 1H), 6.55 (m, 1H), 6.43 (m, , 1H), 3.96 (s, 1H), 2.62 (m, 2H), 2.18 – 1.97 (m, 2H), 1.71 – 1.60 (m, 2H), 1.52 – 1.31 (m, 3H), 1.29 (s, 2H), 1.20 – 1.16 (m, 6H), 1.13 – 0.97 (m, 6H), 0.79 (s, 9H), 0.76 (m, 9H), 0.00 (s, 6H), -0.04 – -0.11 (m, 12H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 204.55, 159.22, 156.58, 155.34, 139.88, 135.17, 128.07, 126.60, 122.03, 120.34, 119.20, 81.15, 75.63, 57.97, 56.92, 55.48, 46.30, 46.00, 42.79, 40.14, 32.01, 31.96, 31.72, 30.69, 29.71, 29.10, 28.02 – 27.77 (m), 26.24, 25.75, 23.97, 20.36, 20.25, 15.63, 0.00, -2.31.

4.5.13 Compound MMA301:

To a stirred solution of protected enone 10 (0.3 g, 0.47 mmole) in tetrahydrofuran (THF) (2.5ml), tetrabutyl ammonium fluoride (TBAF) (1 M in THF, 2.914 ml, 2.914 mmole) was added and stirred for 6h. Ammonium chloride (NH$_4$Cl) solution was added to quench the reaction. Ethyl acetate (3^{x}50ml), dried over sodium sulfate anhydrous (Na$_2$SO$_4$), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to provide Compound MMA301 (0.165 g, 86%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.13 – 6.96 (m, 2H), 6.64 – 6.44 (m, 2H), 5.82 (m, 1H), 5.21 (s, 1H), 4.01 (s, 1H), 2.75 (m, 2H), 2.32 (s, 1H), 2.25 – 2.06 (m, 3H), 1.96 – 1.69 (m, 3H), 1.61 (m, 2H), 1.44 (s, 3H), 1.41 (m, 6H), 1.23 – 1.06 (m, 4H), 0.85 – 0.72 (m, 2H), 0.64 (s, 3H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 201.36, 155.37, 151.23, 146.75, 137.73, , 129.04, 125.07, 113.73, 111.28, 80.17, 75.25, 56.89, 55.07, 48.02, 43.86, 37.22, 36.18, 31.01, 29.58, 29.39, 29.34, 27.57, 26.59, 24.88, 17.28. **HR-FT-MS calcd for C$_{26}$H$_{36}$O$_4$Na$_1$ 412.56163 found 412.56159.**

4.5.14 Compound MMA240:

Compound MMA301 (0.1g, 0.24 mmole) was dissolved in N, N-dimethyacetamide (0.6 ml) then cooled down to 0 0C; followed by addition of sulfamoyl chloride (0.08 g, 0.72 mmole). The reaction mixture was allowed to be stirred and warm to the room temperature for 18h. Admixture of ethyl acetate (30ml) and water (30ml) was added to the reaction mixture and the organic layer was separated, washed by water (3×25ml) followed by the addition of brain (30ml). Sodium sulfate anhydrous (Na2SO4) to dry the organic layer, filtered and concentrated in vacuo. Column chromatography (30% ethyl acetate in hexane) to provide compound MMA240 (0.102 g, 91%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.19 – 6.98 (m, 2H), 6.66 – 6.46 (m, 2H), 5.85 (m, 1H), 4.06 (s, 1H), 2.77 (m, 2H), 2.35 (s, 1H), 2.27 – 2.08 (m, 3H), 1.99 – 1.72 (m, 3H), 1.63 (m, 2H), 1.45 (s, 3H), 1.43 (m, 6H), 1.25 – 1.08 (m, 4H), 0.87 – 0.74 (m, 2H), 0.66 (s, 3H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 215.15, 151.37, 148.18, 146.21, 137.32, , 130.07, 125.88, 113.97, 111.49, 80.17, 75.25, 56.89, 55.07, 48.02, 43.86, 37.22, 36.18, 31.01, 29.58, 29.39, 29.34, 27.57, 26.59, 24.88, 17.28. **HR-FT-MS calcd for** $C_{26}H_{37}N_1O_6S_1Na_1$ **491.64006 found 491.64001.**

4.5.15 Compound MMA280:
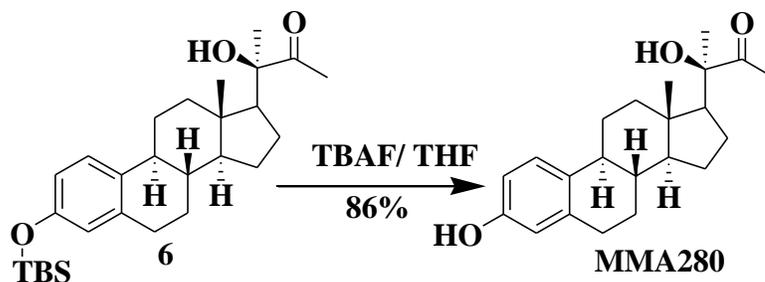
To a stirred solution of α-hydroxyl ketone 6 (0.3 g, 0.657 mmole) in tetrahydrofuran (THF) (2.2ml), tetrabutyl ammonium fluoride (TBAF) (1 M in THF, 2.0367 ml, 2.0367 mmole) was added and stirred for 6h. Ammonium chloride (NH$_4$Cl) solution was added to quench the reaction. Ethyl acetate (3^{x}50ml), dried over sodium sulfate anhydrous (Na$_2$SO$_4$), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to provide compound MMA280 (0.195 g, 86%).

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.98 (s, 1H), 7.09 − 7.02 (m, 1H), 6.60 − 6.47 (m, 2H), 4.06 (s, 1H), 3.22 (m, 2H), 2.82 − 2.62 (m, 3H), 2.23 (m, 1H), 2.20 (s, 3H), 2.11 (m, 1H), 2.06 (s, 1H), 2.02 (m, 1H), 1.90 (m, 1H), 1.86 − 1.74 (m, 1H), 1.73 − 1.44 (m, 2H), 1.40 (s, 3H), 1.37 − 1.09 (m, 3H), 0.88 (s, 3H).

^{13}C NMR (101 MHz, Acetone-d_6) δ 213.19, 155.95, 138.35, 132.03, 126.99, 116.00, 113.63, 81.00, 56.27, 56.03, 44.74, 44.68, 41.24, 39.24, 28.52, 27.51, 25.19, 24.43, 24.11, 22.78, 14.09.

4.5.16 Compound MMA242:
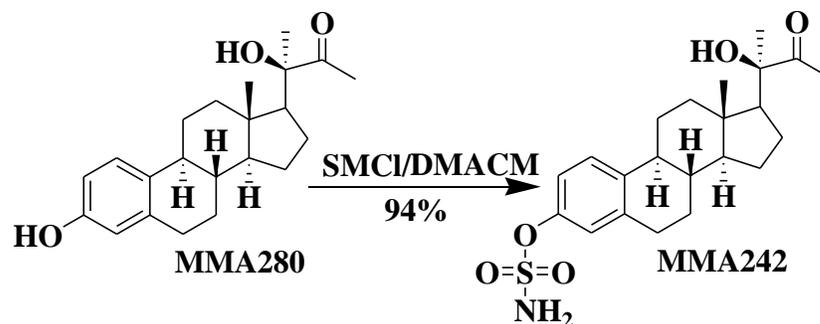
Compound MMA280 (0.1g, 0.3 mmole) was dissolved in N, N-dimethyacetamide (0.7 ml) then cooled down to 0 0C; followed by addition of sulfamoyl chloride (0.1 g, 0.88 mmole). The reaction mixture was allowed to be stirred and warm to the room temperature for 18h. Admixture of ethyl acetate (30ml) and water (30ml) was added to the reaction mixture and the organic layer was separated, washed by water (3×25ml) followed by the addition of brain (30ml). Sodium sulfate anhydrous (Na_2SO_4) to dry the organic layer, filtered and concentrated in vacuo. Column chromatography (30% ethyl acetate in hexane) to provide compound MMA242 (0.116 g, 94%).

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.32 (m, 1H), 7.07 (m, 1H), 7.01 (m, 3H), 4.09 (s, 1H), 2.85 (m, 2H), 2.37 – 2.26 (m, 4H), 2.23 (s, 3H), 2.00 – 1.81 (m, 3H), 1.80 – 1.48 (m, 3H), 1.43 (s, 3H), 1.40 – 1.12 (m, 4H), 0.91 (s, 3H).

^{13}C NMR (101 MHz, Acetone-d_6) δ 213.26, 149.32, 139.73, 139.18, 127.36, 123.02, 120.22, 56.22, 55.95, 44.81, 44.66, 41.09, 38.73, 30.63, 29.48, 28.17, 27.28, 25.21, 24.44, 24.17, 22.78, 14.08.

4.5.17 general procedure for preparing protected estrone with various aromatic enone side chain at C-17 (compound 11):
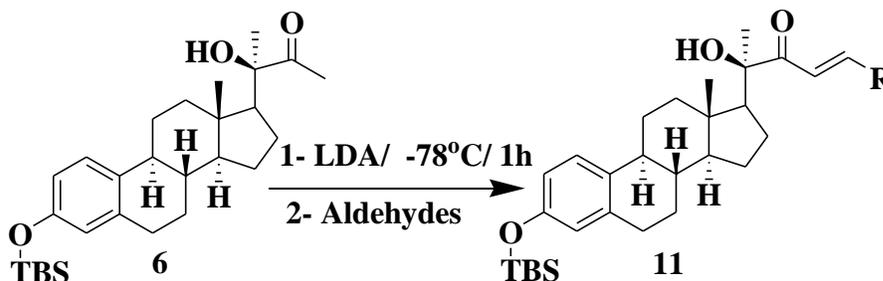
To a stirred solution of α-hydroxyl ketone 6 (0.4g, 0.876 mmole) in a dry tetrahydrofuran (THF) (1.75 ml), lithiumdiisopropyl amine (LDA) (1.6 ml, 3.2 mmole) was added in a dropwise at -78oC the reaction mixture was allowed to stir at -78 °C for 1 h, then the solution of aldehydes (1.75 mmole) in THF (0.15 M) was added to the reaction mixture at -78 °C. The reaction was allowed to stir and warm to the room temperature for 24h. The quench of the reaction was completed by the addition of ammonium chloride (NH$_4$Cl) (25ml). The aqueous layer was extracted using ethyl acetate (3^{x}50ml), dried over sodium sulfate anhydrous (Na$_2$SO$_4$), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to purify the crude material (10% ethyl acetate in hexane) to provide protected estrone with various aromatic enone side chain at C-17 (compound 11).

TBS-protected Para-methoxyphenyl ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.63 (m, 1H), 7.38 (m, 2H), 6.94 (m, 1H), 6.80 − 6.70 (m, 3H), 6.47 − 6.34 (m, 2H), 4.09 (s, 1H), 3.65 (s, 3H), 2.72 − 2.53 (m, 2H), 2.23 − 1.96 (m, 3H), 1.76 − 1.59 (m, 2H), 1.57 − 1.39 (m, 3H), 1.36 (s, 3H), 1.34 − 0.96 (m, 6H), 0.79 (m, 12H), 0.00 (m, 6H). **^{13}C NMR** (101 MHz, Chloroform-*d*) δ 201.84, 162.03, 153.31, 145.52, 137.86, 133.18, 130.53,

127.03, 126.04, 119.99, 117.15, 115.98, 114.48, 79.02, 55.79, 55.46, 55.21, 44.32, 43.91, 40.76, 38.11, 29.66, 27.68, 26.64, 25.78, 24.48, 23.70, 22.06, 18.22, 13.69, -4.32.

TBS-protected Para-Flurophenyl ^{1}**H NMR** (400 MHz, Chloroform-*d*) δ 7.62 (m, 1H), 7.46 – 7.37 (m, 2H), 6.97 – 6.87 (m, 3H), 6.80 (m, 1H), 6.46 – 6.34 (m, 2H), 3.98 (s, 1H), 2.61 (m, 2H), 2.20 – 1.96 (m, 2H), 1.75 – 1.59 (m, 3H), 1.57 – 1.38 (m, 3H), 1.33 – 0.92 (m, 3H), 0.79 (m, 12H), 0.75 – 0.61 (m, 6H), 0.00 (s, 6H). ^{13}**C NMR** (101 MHz, Chloroform-*d*) δ 201.75, 153.33, 144.34, 137.82, 133.11, 130.65 (d, *J* = 8.4 Hz), 126.02, 120.00, 118.19, 117.16, 116.32, 116.10, 79.19, 55.80, 55.12, 44.37, 43.92, 40.77, 38.10, 29.86 – 29.55 (m), 29.42, 27.68, 26.62, 25.76, 24.36, 23.69, 22.08, 18.21, 14.18, 13.69, -4.35.

TBS-protected Para-chlorophenyl ^{1}**H NMR** (400 MHz, Chloroform-*d*) δ 7.59 (m, 1H), 7.35 (m, 2H), 7.26 – 7.16 (m, 2H), 6.93 (m, 1H), 6.84 (m, 1H), 6.49 – 6.33 (m, 2H), 3.96 (s, 1H), 2.73 – 2.52 (m, 2H), 2.19 – 1.95 (m, 3H), 1.66 (m, 2H), 1.44 (m, 3H), 1.37 (s, 3H), 1.30 – 0.98 (m, 6H), 0.79 (m, 12H), 0.00 (s, 6H). ^{13}**C NMR** (101 MHz, Chloroform-*d*) δ 201.76, 153.33, 144.18, 137.83, 136.90, 133.11, 132.81, 129.84, 129.31, 126.04, 120.01, 118.93, 117.17, 79.26, 55.78, 55.08, 44.38, 43.90, 40.75, 38.10, 29.66, 27.68, 26.62, 25.79, 24.35, 23.70, 22.10, 18.23, 13.72, -4.31.

TBS-protected Para-bromophenyl ^{1}**H NMR** (400 MHz, Chloroform-*d*) δ 7.57 (m, 1H), 7.39 – 7.31 (m, 2H), 7.31 – 7.24 (m, 2H), 6.93 (m, 1H), 6.85 (d, *J* = 15.6 Hz, 1H), 6.48 – 6.35 (m, 2H), 3.95 (s, 1H), 2.71 – 2.52 (m, 2H), 2.21 – 1.95 (m, 3H), 1.73 – 1.60 (m, 2H), 1.54 – 1.39 (m, 3H), 1.37 (s, 3H), 1.33 – 1.01 (m, 6H), 0.79 (m, 12H), 0.00 (s, 6H). ^{13}**C NMR** (101 MHz, Chloroform-*d*) δ 201.77, 153.33, 144.25, 137.83, 133.22,

133.16, 132.27, 130.02, 126.04, 125.32, 120.00, 119.02, 117.17, 79.27, 55.77, 55.06, 44.38, 43.90, 40.75, 38.09, 29.65, 27.67, 26.62, 25.79, 24.34, 23.70, 22.09, 18.36 18.23, 13.72, -4.31.

TBS-protected Para-trifluromethylphenyl ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.65 (m, 1H), 7.50 (m, 4H), 6.94 (m, 2H), 6.48 – 6.35 (m, 2H), 3.89 (s, 1H), 2.61 (m, 2H), 2.21 – 1.94 (m, 3H), 1.75 – 1.60 (m, 2H), 1.45 (m, 3H), 1.39 (s, 3H), 1.32 – 0.99 (m, 6H), 0.79 (s, 12H), 0.00 (s, 6H). **^{13}C NMR** (101 MHz, Chloroform-*d*) δ 201.72, 153.34, 143.60, 137.75 (d, *J* = 14.0 Hz), 133.06, 132.44, 128.75, 125.99 (d, *J* = 6.2 Hz), 120.84, 120.00, 117.17, 79.40, 55.77, 55.04, 44.41, 43.89, 40.74, 38.09, 29.63, 27.66, 26.60, 25.75, 24.27, 23.67, 22.10, 18.20, 13.70, -4.36.

TBS-protected Para-nitrophenyl ^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.08 (m, 2H), 7.65 (d, 1H), 7.61 – 7.54 (m, 2H), 7.10 – 6.91 (m, 2H), 6.49 – 6.33 (m, 2H), 3.81 (s, 1H), 2.71 – 2.54 (m, 2H), 2.21 – 1.96 (m, 3H), 1.67 (m, 2H), 1.61 – 1.41 (m, 3H), 1.39 (s, 3H), 1.01 – 0.83 (m, 6H), 0.79 (s, 12H), 0.00 (s, 6H). **^{13}C NMR** (151 MHz, Chloroform-*d*) δ 201.50, 153.35, 148.78, 142.31, 140.38, 137.74, 132.97, 129.15, 125.97, 124.17, 122.40, 119.97, 117.15, 79.51, 55.78, 55.05, 44.43, 43.89, 40.73, 38.09, 34.60, 31.96, 29.75, 29.32, 29.07, 27.65, 26.58, 25.27, 24.22, 23.65, 22.11, 20.66, 18.72, 18.15, 14.08, 13.66, 11.39, -4.44.

TBS-protected 2-Bromofuran ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.27 (d, 1H), 6.94 (d, 1H), 6.70 (d, , 1H), 6.50 – 6.35 (m, 3H), 6.26 (d, 1H), 3.98 (s, 1H), 2.72 – 2.49 (m, 2H), 2.22 – 1.95 (m, 3H), 1.77 – 1.59 (m, 2H), 1.56 – 1.39 (m, 3H), 1.35 (s, 3H), 1.32 – 0.96 (m, 6H), 0.78 (d, 12H), 0.00 (s, 6H). **^{13}C NMR** (101 MHz, Chloroform-*d*) δ 201.61,

153.29, 153.03, 137.83, 133.18, 130.01, 126.37, 126.05, 119.98, 119.08, 117.15, 116.41, 114.78, 79.17, 55.70, 54.90, 44.35, 43.87, 40.65, 38.11, 29.66, 27.67, 26.62, 25.77, 24.31, 23.70, 22.06, 18.21, 13.69, -4.32.

TBS-protected 2-Bromothiophene ^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.62 (d, 1H), 6.93 (d, 1H), 6.89 (m, 1H), 6.85 (m, , 1H), 6.52 (m, 1H), 6.46 – 6.34 (m, 2H), 3.92 (s, 1H), 2.70 – 2.51 (m, 2H), 1.70 – 1.58 (m, 3H), 1.56 – 1.37 (m, 2H), 1.33 (s, 3H), 1.08 (q, , 3H), 0.78 (m, 12H), 0.73 – 0.62 (m, 6H), 0.00 (s, 6H). **^{13}C NMR** (101 MHz, Chloroform-*d*) δ 201.40, 153.32, 141.35, 137.82, 136.97, 133.13, 132.94, 131.41, 126.04, 119.99, 117.58, 117.16, 117.06, 79.13, 55.74, 55.13, 44.35, 43.87, 40.71, 38.10, 29.66, 27.67, 26.62, 25.77, 24.35, 23.69, 22.09, 18.22, 13.68, -4.33.

4.5.18 Compound MMA268:

To a stirred solution of desired compound 11 (0.3 g, 0.5218 mmole) in tetrahydrofuran (THF) (2.6 ml), tetrabutyl ammonium fluoride (TBAF) (1 M in THF, 1.62 ml, 1.62 mmole) was added and stirred for 6h. Ammonium chloride (NH4Cl) solution was added to quench the reaction. Ethyl acetate (3×50ml), dried over sodium sulfate anhydrous (Na2SO4), filtrated, and concentrated under vacuo. Silica gel column chromatography was

used to purify the crude material (10% ethyl acetate in hexane) to provide **Compound MMA268** (0.21 g, 87.5%).

¹H NMR (400 MHz, Chloroform-*d*) δ 7.72 (m, 1H), 7.45 (m, 2H), 7.01 (m, 1H), 6.88 – 6.78 (m, 3H), 6.60 – 6.45 (m, 2H), 4.34 (s, 1H), 3.72 (s, 3H), 2.67 (m, 2H), 2.34 – 1.99 (m, 3H), 1.95 (m, 2H), 1.47 (m, 3H), 1.53 – 1.42 (m, 1H), 1.24 – 1.10 (m, 5H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 202.01, 162.10, 153.91, 145.96, 138.06, 132.21, 130.65, 126.98, 126.36, 115.84, 115.45, 114.50, 112.88, 79.32, 55.70, 55.45, 55.26, 44.31, 43.84, 40.69, 38.19, 32.00, 29.74, 29.68, 27.65, 26.71, 24.34, 23.67, 22.77. **HR-FT-MS calcd for C₃₀H₃₇O₄ 461.2686 found 461.26819.**

4.5.19 Compound MMA267:
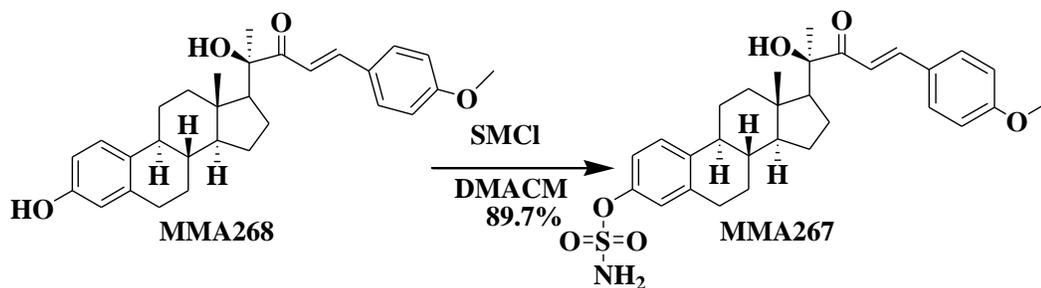
Compound MMA268 (0.2g, 0.434 mmole) was dissolved in N, N-dimethyacetamide (1.05 ml) then cooled down to 0 ⁰C; followed by addition of sulfamoyl chloride (0.15 g, 1.302 mmole). The reaction mixture was allowed to be stirred and warm to the room temperature for 18h. Admixture of ethyl acetate (30 ml) and water (30 ml) was added to the reaction mixture and the organic layer was separated, washed by water (3×25 ml) followed by the addition of brain (30 ml). Sodium sulfate anhydrous (Na₂SO₄) to dry

the organic layer, filtered and concentrated in vacuo. Column chromatography (30% ethyl acetate in hexane) to provide **compound MMA267** (0.21 g, 89.7%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.75 (m, 1H), 7.58 – 7.44 (m, 2H), 7.26 (s, 1H), 7.05 – 6.92 (m, 2H), 6.91 – 6.83 (m, 3H), 4.93 (s, 2H), 4.21 (s, 1H), 3.91 – 3.75 (m, 3H), 2.34 – 2.13 (m, 2H), 2.11 (s, 3H), 1.97 (m, 2H), 1.68 – 1.50 (m, 3H), 1.48 (s, 1H), 1.21 – 1.17 (m, 5H), 0.89 (s, 3H), 0.86 – 0.74 (m, 3H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 202.03, 162.113, 148.88, 145.94, 138.03, 132.19, 130.62, 126.95, 126.35, 115.85, 115.43, 114.52, 112.86, 79.37, 55.69, 55.43, 55.24, 44.29, 43.82, 40.67, 38.17, 32.06, 29.73, 29.66, 27.62, 26.70, 24.32, 23.65, 22.75. **HR-FT-MS calcd for $C_{30}H_{37}O_6N_1Na_1S_1$ 562.2234 found 562.22222.**

4.5.20 Compound MMA269:
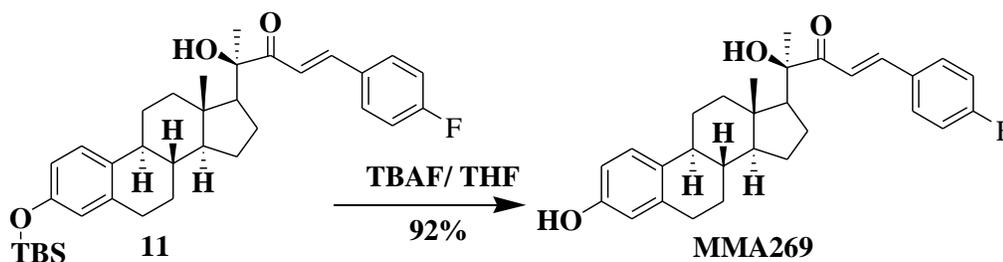
To a stirred solution of desired compound 11 (0.3 g, 0.533 mmole) in tetrahydrofuran (THF) (2.6 ml), tetrabutyl ammonium fluoride (TBAF) (1 M in THF, 1.65 ml, 1.65 mmole) was added and stirred for 6h. Ammonium chloride (NH_4Cl) solution was added to quench the reaction. Ethyl acetate (3×50ml), dried over sodium sulfate anhydrous (Na_2SO_4), filtrated, and concentrated under vacuo. Silica gel column chromatography was

used to purify the crude material (10% ethyl acetate in hexane) to provide **Compound MMA269** (0.22 g, 92%).

¹H NMR (400 MHz, Chloroform-*d*) δ 7.86 (m, 1H), 7.71 – 7.61 (m, 2H), 7.22 – 7.10 (m, 3H), 7.05 (m, 1H), 6.70 (m, 1H), 6.62 (m, 1H), 6.33 – 6.14 (m, 1H), 4.39 (s, 1H), 2.92 – 2.75 (m, 2H), 2.42 – 2.16 (m, 4H), 1.98 – 1.82 (m, 2H), 1.63 (s, 3H), 1.52 – 1.20 (m, 5H), 1.01 (s, 3H), 0.92 (m, 3H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 201.97, 153.65, 144.79, 138.15, 132.41, 130.81, 130.72, 130.53, 130.50, 126.41, 118.04, 116.35, 116.13, 115.42, 112.83, 79.51, 55.69, 55.14, 44.35, 43.83, 40.66, 38.14, 29.72, 29.65, 27.61, 26.68, 24.23, 23.65, 22.09. **HR-FT-MS calcd for C₂₉H₃₃O₃F₁Na₁ 471.2306 found 471.23231.**

4.5.21 Compound MMA271:
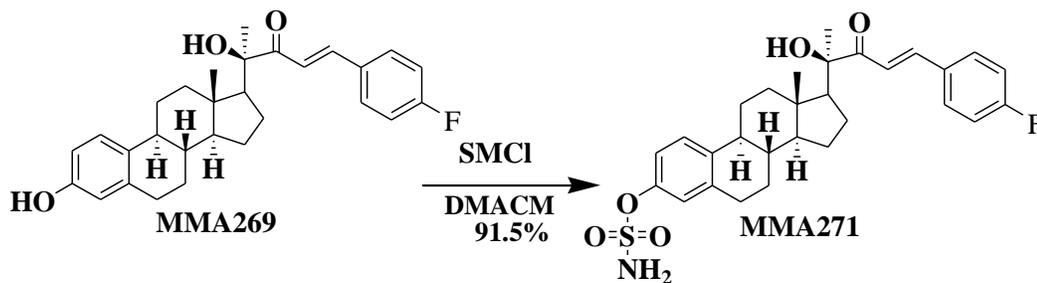
Compound MMA269 (0.2g, 0.446 mmole) was dissolved in N, N-dimethyacetamide (1.1 ml) then cooled down to 0 °C; followed by addition of sulfamoyl chloride (0.2 g, 1.338 mmole). The reaction mixture was allowed to be stirred and warm to the room temperature for 18h. Admixture of ethyl acetate (30 ml) and water (30 ml) was added to the reaction mixture and the organic layer was separated, washed by water (3×25 ml) followed by the addition of brain (30 ml). Sodium sulfate anhydrous (Na₂SO₄) to dry

the organic layer, filtered and concentrated in vacuo. Column chromatography (30% ethyl acetate in hexane) to provide **compound MMA271** (0.215 g, 91.5%).

^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.73 (m, 1H), 7.58 − 7.46 (m, 2H), 7.25 − 7.18 (m, 1H), 7.02 (m, , 3H), 6.95 (m, 1H), 6.89 (m, 1H), 5.14 (s, 2H), 4.11 (s, 1H), 2.83 − 2.73 (m, 2H), 2.33 − 2.12 (m, 3H), 1.97 (s, 1H), 1.92 − 1.74 (m, 3H), 1.72 − 1.50 (m, 2H), 1.47 (s, 3H), 1.44 − 1.24 (m, 3H), 1.23 − 1.16 (m, 5H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 201.75, 147.90, 144.57, 139.64, 138.90, 130.71, 130.65, 126.69, 121.93, 118.95, 116.30, 116.16, 79.19, 55.69, 55.06, 44.22, 43.99, 40.54, 37.66, 29.72, 29.51, 27.27, 26.45, 24.30, 23.64, 22.71, 22.02. **HR-FT-MS calcd for $C_{29}H_{34}O_5F_1Na_1S_1$ 550.2034 found 550.20497.**

4.5.22 Compound MMA309:
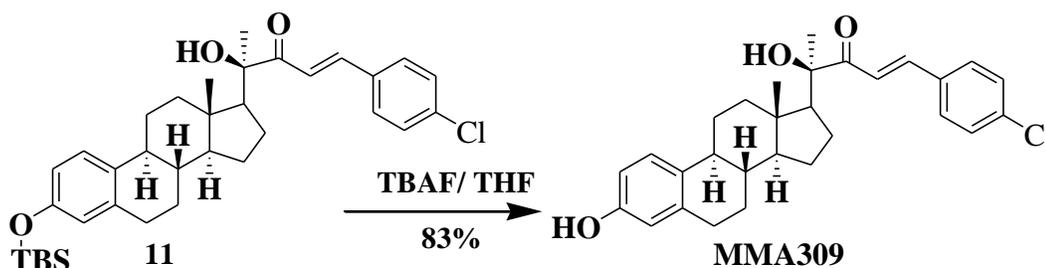
To a stirred solution of desired compound 11 (0.3 g, 0.518 mmole) in tetrahydrofuran (THF) (2.6 ml), tetrabutyl ammonium fluoride (TBAF) (1 M in THF, 1.6 ml, 1.6 mmole) was added and stirred for 6h. Ammonium chloride (NH_4Cl) solution was added to quench the reaction. Ethyl acetate (3×50ml), dried over sodium sulfate anhydrous (Na_2SO_4), filtrated, and concentrated under vacuo. Silica gel column chromatography was

used to purify the crude material (10% ethyl acetate in hexane) to provide **Compound MMA309** (0.2 g, 83%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.71 (m, 1H), 7.57 − 7.42 (m, 2H), 7.38 − 7.28 (m, 2H), 7.08 (m, 1H), 6.94 (m, 1H), 6.63 − 6.44 (m, 2H), 4.95 (s, 1H), 4.11 (s, 1H), 2.83 − 2.63 (m, 3H), 2.33 − 2.17 (m, 3H), 2.11 (s, 3H), 1.86 − 1.71 (m, 2H), 1.67 − 1.50 (m, 3H), 1.49 (s, 3H), 1.42 − 1.09 (m, 5H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 201.80, 153.41, 144.30, 138.24, 136.95, 132.74, 132.63, 129.83, 129.30, 126.42, 118.83, 115.27, 112.66, 79.31, 55.69, 55.07, 44.35, 43.82, 40.67, 38.10, 30.98, 29.61, 27.56, 26.65, 24.27, 23.64, 22.04. **HR-FT-MS calcd for C$_{29}$H$_{33}$O$_3$Cl$_1$Na$_1$ 487.2010 found 487.20045.**

4.5.23 Compound MMA294:
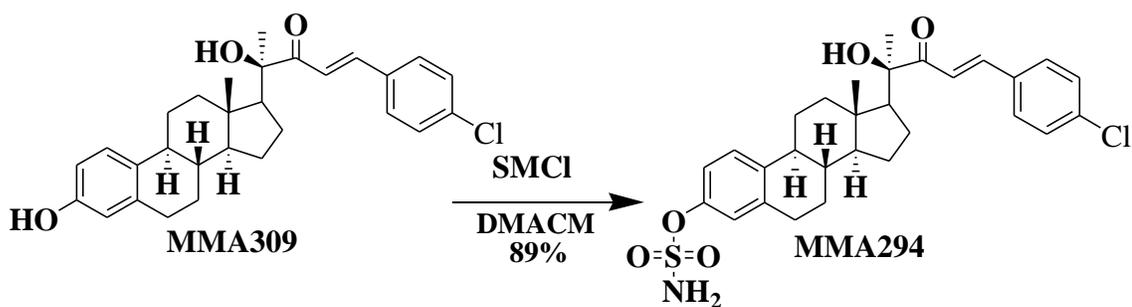
Compound MMA309 (0.2g, 0.43 mmole) was dissolved in N, N-dimethyacetamide (1.1 ml) then cooled down to 0 ^{0}C; followed by addition of sulfamoyl chloride (0.15 g, 1.3 mmole). The reaction mixture was allowed to be stirred and warm to the room temperature for 18h. Admixture of ethyl acetate (30 ml) and water (30 ml) was added to the reaction mixture and the organic layer was separated, washed by water (3×25

ml) followed by the addition of brain (30 ml). Sodium sulfate anhydrous (Na$_2$SO$_4$) to dry the organic layer, filtered and concentrated in vacuo. Column chromatography (30% ethyl acetate in hexane) to provide **compound MMA294** (0.209 g, 89%).

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.86 (m, 2H), 7.50 (m, 3H), 7.35 (m, 1H), 7.12 – 7.00 (m, 3H), 5.13 (m, 2H), 4.29 (m, 1H), 2.35 (m, 3H), 2.29 – 2.22 (m, 3H), 2.12 (s, 3H), 1.71 (m, 2H), 1.62 – 1.57 (m, 3H), 1.43 (m, 3H), 1.39 – 1.19 (m, 5H).

^{13}C NMR (101 MHz, Acetone-d_6) δ 203.04, 149.36, 143.75, 139.68, 139.16, 136.79, 134.54, 131.22, 129.95, 127.32, 122.99, 121.30, 120.18, 80.17, 56.30, 55.61, 44.89, 44.80, 41.01, 38.77, 32.68, 28.14, 27.26, 24.83, 24.38, 23.39, 22.80. **HR-FT-MS calcd for C$_{29}$H$_{34}$O$_5$N$_1$Cl$_1$Na$_1$S$_1$ 566.1738 found 566.17551.**

4.5.24 Compound MMA310:
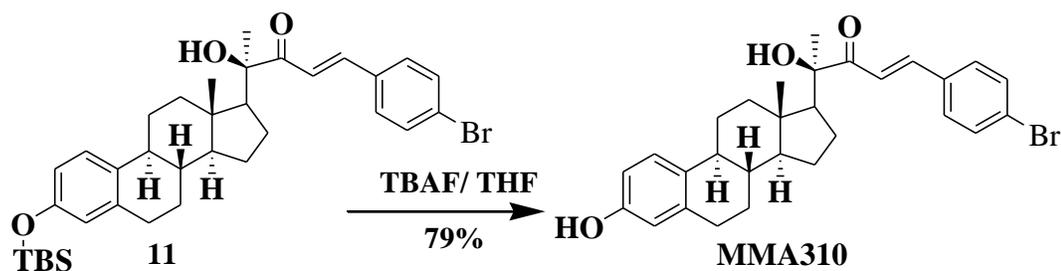
To a stirred solution of desired compound **11** (0.3 g, 0.48 mmole) in tetrahydrofuran (THF) (2.4 ml), tetrabutyl ammonium fluoride (TBAF) (1 M in THF, 1.5 ml, 1.5 mmole) was added and stirred for 6h. Ammonium chloride (NH$_4$Cl) solution was added to quench the reaction. Ethyl acetate (3×50ml), dried over sodium sulfate anhydrous (Na$_2$SO$_4$), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to

purify the crude material (10% ethyl acetate in hexane) to provide **Compound MMA310** (0.194 g, 79%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.69 (m, , 1H), 7.51 – 7.36 (m, 4H), 7.07 (d, *J* = 8.5 Hz, 1H), 6.96 (m, 1H), 6.63 – 6.44 (m, 2H), 5.17 (s, 1H), 4.12 (s, 1H), 2.85 – 2.64 (m, 3H), 2.34 – 2.17 (m, 3H), 2.11 (s, 3H), 1.86 – 1.70 (m, 2H), 1.67 – 1.50 (m, 3H), 1.49 (s, 3H), 1.41 – 1.10 (m, 5H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 201.84, 153.46, 144.43, 138.22, 133.15, 132.57, 132.27, 130.02, 126.42, 125.37, 118.92, 115.30, 112.69, 79.37, 55.69, 55.06, 44.35, 43.82, 40.67, 38.11, 31.61, 30.99, 29.63, 27.56, 26.65, 24.25, 23.64, 22.06. **HR-FT-MS calcd for C$_{29}$H$_{33}$O$_3$Br$_1$Na$_1$ 531.1505 found 531.15002.**

4.5.25 Compound MMA295:
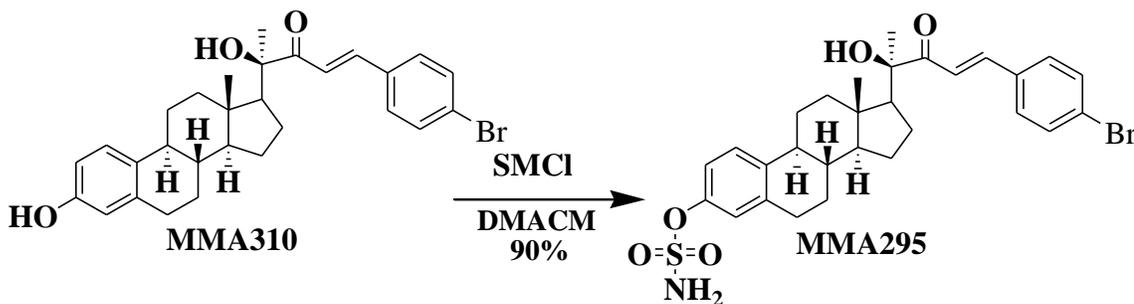
Compound MMA310 (0.1 g, 0.2 mmole) was dissolved in N, N-dimethyacetamide (0.5 ml) then cooled down to 0 ^{0}C; followed by addition of sulfamoyl chloride (0.07 g, 0.6 mmole). The reaction mixture was allowed to be stirred and warm to the room temperature for 18h. Admixture of ethyl acetate (30 ml) and water (30 ml) was added to the reaction mixture and the organic layer was separated, washed by water (3×25 ml) followed by the

addition of brain (30 ml). Sodium sulfate anhydrous (Na$_2$SO$_4$) to dry the organic layer, filtered and concentrated in vacuo. Column chromatography (30% ethyl acetate in hexane) to provide **compound MMA295** (0.104 g, 90%).

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.69 (m, 2H), 7.54 (m, 3H), 7.47 (m, , 1H) 7.08 – 6.92 (m, 3H), 5.16 (s, 2H), 4.09 (s, , 1H), 2.41 (s, 3H), 2.30 – 2.25 (m, 3H), 2.25 – 2.12 (s, 3H), 1.84 – 1.70 (m, 2H), 1.58 (m, 3H), 1.47 (m, 3H), 1.27 – 1.16 (m, 5H).

^{13}C NMR (101 MHz, Acetone-d_6) δ 203.09, 149.36, 143.84, 139.70, 139.17, 134.93, 132.95, 131.40, 127.32, 125.20, 122.98, 121.42, 120.18, 80.18, 56.37, 55.47, 44.89, 44.85, 41.01, 38.78, 32.66, 28.11, 27.26, 24.80, 24.36, 23.37, 22.79. **HR-FT-MS calcd for C$_{29}$H$_{34}$O$_5$N$_1$Br$_1$Na$_1$S$_1$ 610.1233 found 610.12377.**

4.5.26 Compound MMA308:
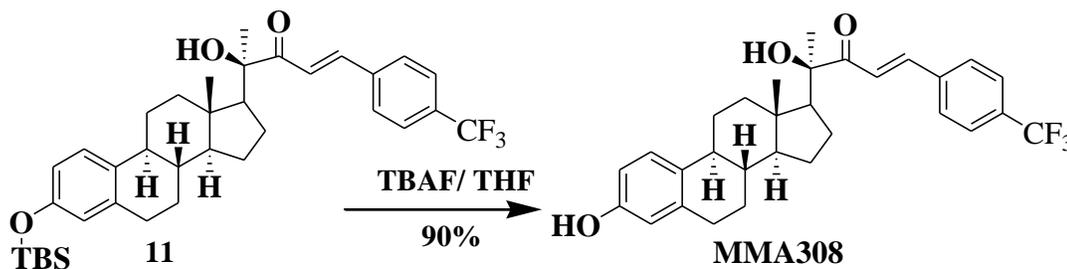
To a stirred solution of desired compound 11 (0.3 g, 0.49 mmole) in tetrahydrofuran (THF) (2.5 ml), tetrabutyl ammonium fluoride (TBAF) (1 M in THF, 1.519 ml, 1.519 mmole) was added and stirred for 6h. Ammonium chloride (NH$_4$Cl) solution was added to quench the reaction. Ethyl acetate (3×50ml), dried over sodium sulfate anhydrous (Na$_2$SO$_4$), filtrated, and concentrated under vacuo. Silica gel column chromatography was

used to purify the crude material (10% ethyl acetate in hexane) to provide **Compound MMA308** (0.23 g, 90%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.73 (m, 1H), 7.57 (m, 4H), 7.07 – 6.99 (m, 2H), 6.58 – 6.42 (m, 2H), 5.54 (s, 1H), 4.07 (s, 1H), 2.86 – 2.62 (m, 3H), 2.55 (m, 3H), 2.28 – 2.01 (m, 3H), 1.85 – 1.66 (m, 2H), 1.66 – 1.51 (m, 3H), 1.51 – 1.45 (m, 3H), 1.38 – 1.02 (m, 5H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 201.83, 153.55, 143.86, 138.18, 137.61, 132.48, 128.79, 126.40, 125.94, 120.73, 115.35, 112.75, 79.59, 55.68, 55.06, 44.39, 43.81, 40.65, 38.12, 31.96, 29.62, 27.57, 26.65, 24.17, 23.64, 22.73, 22.10. **HR-FT-MS calcd for $C_{30}H_{33}O_3F_3Na_1$ 521.2274 found 521.22942.**

4.5.27 Compound MMA300:
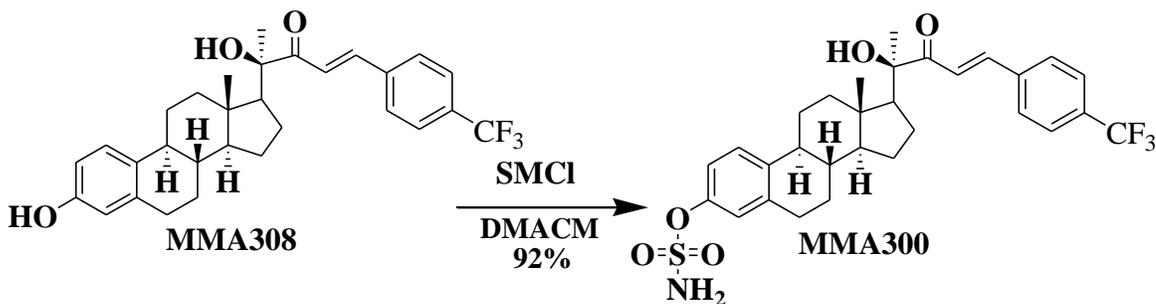
Compound MMA308 (0.2 g, 0.4 mmole) was dissolved in N, N-dimethyacetamide (1 ml) then cooled down to 0 ^{0}C; followed by addition of sulfamoyl chloride (0.14 g, 1.2 mmole). The reaction mixture was allowed to be stirred and warm to the room temperature for 18h. Admixture of ethyl acetate (30 ml) and water (30 ml) was added to the reaction mixture and the organic layer was separated, washed by water (3×25 ml) followed by the

addition of brain (30 ml). Sodium sulfate anhydrous (Na_2SO_4) to dry the organic layer, filtered and concentrated in vacuo. Column chromatography (30% ethyl acetate in hexane) to provide **compound MMA300** (0.214 g, 92%).

^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.77 (m, 1H), 7.68 – 7.57 (m, 4H), 7.24 (m, 1H), 7.03 (s, 2H), 6.95 (m, 1H), 4.97 (s, 2H), 4.01 (s, 1H), 2.79 (m, 3H), 2.56 (m, 3H), 2.33 – 2.13 (m, 3H), 1.86 – 1.73 (m, 2H), 1.69 – 1.52 (m, 3H), 1.50 (s, 3H), 1.20 – 1.17 (m, 5H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 201.66, 147.90, 143.76, 139.66, 138.92, 137.62, 132.45, 132.23, 128.75, 126.71, 121.90, 120.71, 118.92, 79.36, 55.69, 55.02, 44.28, 43.99, 40.55, 37.66, 30.96, 29.25, 27.26, 26.45, 24.26, 23.64, 22.71. **HR-FT-MS calcd for $C_{30}H_{34}O_5N_1F_3Na_1S_1$ 600.2002 found 600.20229.**

4.5.28 Compound MMA306:
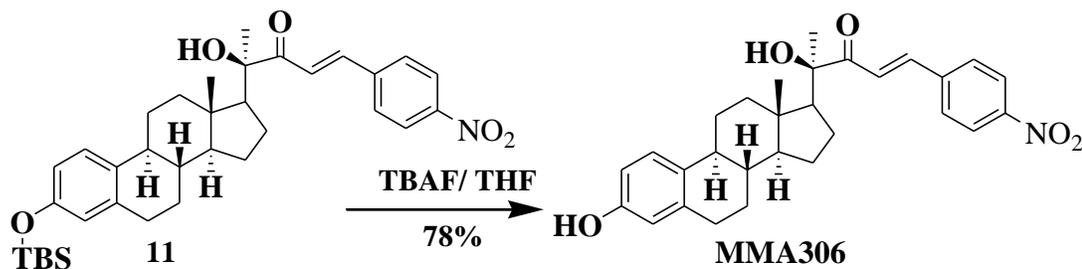
To a stirred solution of desired compound **11** (0.3 g, 0.51 mmole) in tetrahydrofuran (THF) (2.5 ml), tetrabutyl ammonium fluoride (TBAF) (1 M in THF, 1.6 ml, 1.6 mmole) was added and stirred for 6h. Ammonium chloride (NH_4Cl) solution was added to quench the reaction. Ethyl acetate (3×50ml), dried over sodium sulfate anhydrous (Na_2SO_4), filtrated, and concentrated under vacuo. Silica gel column chromatography was used to

purify the crude material (10% ethyl acetate in hexane) to provide Compound MMA306 (0.187 g, 78%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.30 – 8.13 (m, 2H), 7.77 (d, *J* = 15.7 Hz, 1H), 7.72 – 7.65 (m, 2H), 7.15 – 7.04 (m, 2H), 6.60 – 6.46 (m, 2H), 5.02 (s, 1H), 3.95 (s, 1H), 2.85 – 2.63 (m, 3H), 2.58 (m, 3H), 2.33 – 2.04 (m, 3H), 1.86 – 1.72 (m, 2H), 1.72 – 1.54 (m, 3H), 1.51 (d, *J* = 2.8 Hz, 3H), 1.42 – 1.04 (m, 5H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 201.59, 153.46, 148.78, 142.48, 140.35, 138.20, 132.51, 129.21, 126.40, 124.22, 122.34, 115.28, 112.68, 79.60, 55.68, 55.02, 44.42, 43.80, 40.65, 38.10, 31.94, 29.60, 27.55, 26.63, 24.21, 23.64, 22.11. **HR-FT-MS calcd for $C_{29}H_{33}O_5N_1Na_1$ 498.2251 found 498.22654.**

4.5.29 Compound MMA307:
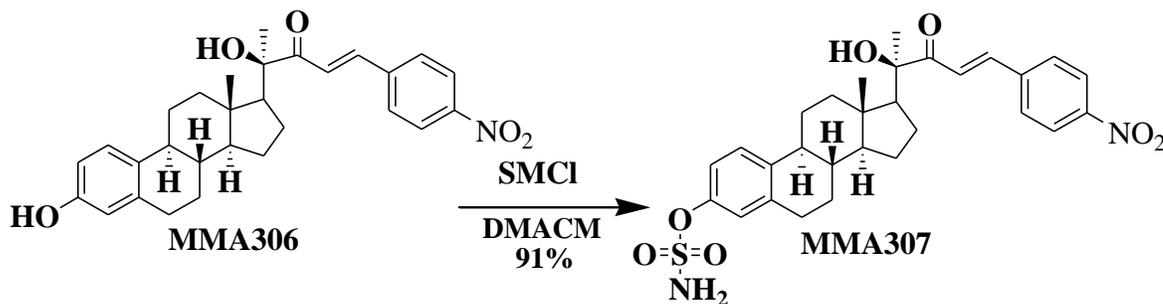
Compound **MMA306** (0.15 g, 0.135 mmole) was dissolved in N, N-dimethyacetamide (0.8 ml) then cooled down to 0 ^{0}C; followed by addition of sulfamoyl chloride (0.12 g, 0.945 mmole). The reaction mixture was allowed to be stirred and warm to the room temperature for 18h. Admixture of ethyl acetate (30 ml) and water (30 ml) was added to the reaction mixture and the organic layer was separated, washed by water (3×25

ml) followed by the addition of brain (30 ml). Sodium sulfate anhydrous (Na_2SO_4) to dry the organic layer, filtered and concentrated in vacuo. Column chromatography (30% ethyl acetate in hexane) to provide compound **MMA307** (0.16 g, 91%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 8.36 – 8.19 (m, 2H), 7.82 (d, *J* = 15.7 Hz, 1H), 7.79 – 7.71 (m, 2H), 7.20 – 7.09 (m, 2H), 6.68 – 6.56 (m, 2H), 5.12 (s, 2H), 4.13 (s, 1H), 3.99 (s, 1H), 2.89 – 2.65 (m, 3H), 2.62 (m, 3H), 2.35 – 2.10 (m, 3H), 1.92 – 1.79 (m, 2H), 1.75 – 1.58 (m, 3H), 1.62 (d, *J* = 2.8 Hz, 3H), 1.45 – 1.10 (m, 5H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 201.12, 149.46, 146.78, 142.31, 140.12, 138.22, 132.43, 129.09, 126.36, 124.17, 122.29, 115.19, 112.57, 79.52, 55.61, 54.96, 44.38, 43.67, 40.60, 38.01, 31.88, 29.60, 27.48, 26.58, 24.17, 23.59, 22.03. **HR-FT-MS calcd for $C_{29}H_{34}O_7N_2Na_1S_1$ 577.1979 found 577.19790.**

4.5.30 Compound MMA297:
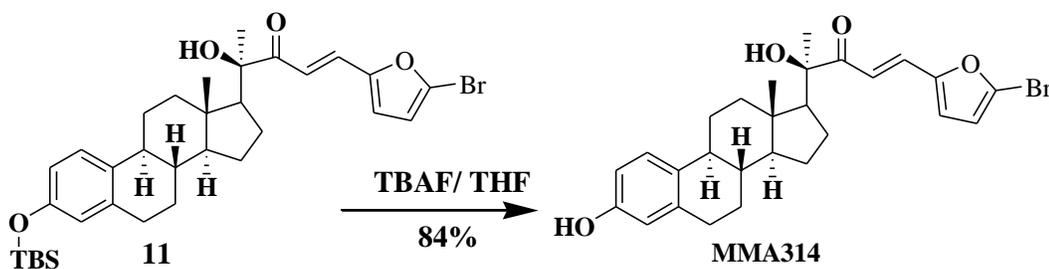
To a stirred solution of desired **compound 11** (0.3 g, 0.184 mmole) in tetrahydrofuran (THF) (1 ml), tetrabutyl ammonium fluoride (TBAF) (1 M in THF, 0.6 ml, 0.6 mmole) was added and stirred for 6h. Ammonium chloride (NH_4Cl) solution was added to quench the reaction. Ethyl acetate (3×50ml), dried over sodium sulfate anhydrous (Na_2SO_4), filtrated, and concentrated under vacuo. Silica gel column chromatography was

used to purify the crude material (10% ethyl acetate in hexane) to provide **Compound MMA297** (0.204 g, 84%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.38 (m, 1H), 7.07 (m, , 1H), 6.80 (m, 1H), 6.57 (m, 2H), 6.49 (d, *J* = 2.7 Hz, 1H), 6.38 (d, *J* = 3.4 Hz, 1H), 5.46 (s, 1H), 4.17 (s, 1H), 2.88 − 2.64 (m, 3H), 2.22 (m, 3H), 2.12 (m, 3H), 1.87 − 1.69 (m, 2H), 1.54 (m, 3H), 1.48 (s, 3H), 1.40 − 1.05 (m, 5H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 201.72, 153.52, 152.98, 138.20, 132.57, 130.18, 126.47, 126.42, 119.24, 116.29, 115.30, 114.81, 112.71, 79.31, 55.63, 54.92, 44.33, 43.80, 40.58, 38.13, 30.99, 29.64, 27.58, 26.66, 24.22, 23.65, 22.04. **HR-FT-MS calcd for $C_{27}H_{31}O_4Br_1Na_1$ 521.1298 found 521.12905.**

4.5.31 Compound MMA314:
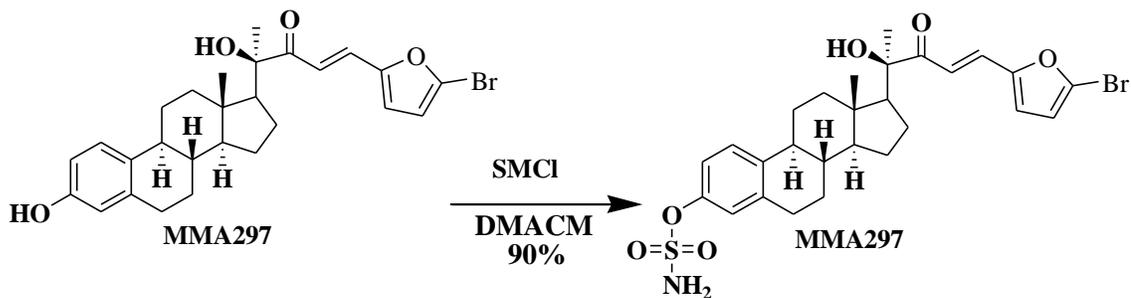
Compound MMA297 (0.2 g, 0.4 mmole) was dissolved in N, N-dimethyacetamide (1 ml) then cooled down to 0 ^{0}C; followed by addition of sulfamoyl chloride (0.14 g, 1.2 mmole). The reaction mixture was allowed to be stirred and warm to the room temperature for 18h. Admixture of ethyl acetate (30 ml) and water (30 ml) was added to the reaction mixture and the organic layer was separated, washed by water (3×25 ml) followed by the

addition of brain (30 ml). Sodium sulfate anhydrous (Na_2SO_4) to dry the organic layer, filtered and concentrated in vacuo. Column chromatography (30% ethyl acetate in hexane) to provide **compound MMA314** (0.208 g, 90%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.56 – 7.45 (m, 3H), 7.39 – 7.27 (m, 1H), 7.11 (m, 2H), 6.90 (m, 1H), 5.28 (s, 2H), 4.24 (s, 1H), 2.98 (m, 3H), 2.61 (m, 3H), 2.41 – 2.22 (m, 3H), 2.02 – 1.84 (m, 2H), 1.66 (m, 3H), 1.56 (m, 3H), 1.23 – 1.09 (m, 5H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 201.52, 153.01, 147.92, 139.86, 138.85, 130,15, 126.68, 126.44, 122.10, 119.21, 118,98, 116.17, 114.77, 79.14, 55.70, 55.02, 44.21, 43.99, 34.45, 34.23, 32.78, 31.98, 29.76, 28.01, 26.80, 26.49, 22.74. . **HR-FT-MS calcd for $C_{27}H_{32}O_6N_1Br_1Na_1S_1$ 600.1026 found 600.10038.**

4.5.32 Compound MMA313:
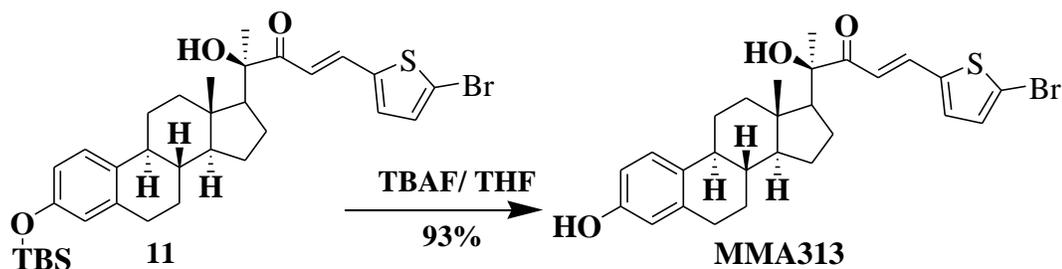
To a stirred solution of desired compound 11 (0.3 g, 0.476 mmole) in tetrahydrofuran (THF) (2.4 ml), tetrabutyl ammonium fluoride (TBAF) (1 M in THF, 1.47 ml, 1.47 mmole) was added and stirred for 6h. Ammonium chloride (NH_4Cl) solution was added to quench the reaction. Ethyl acetate (3×50ml), dried over sodium sulfate anhydrous (Na_2SO_4), filtrated, and concentrated under vacuo. Silica gel column chromatography was

used to purify the crude material (10% ethyl acetate in hexane) to provide **Compound MMA313** (0.229 g, 93%).

¹H NMR (400 MHz, Chloroform-*d*) δ 7.86 (d, *J* = 15.3 Hz, 1H), 7.21 − 7.03 (m, 3H), 6.83 − 6.57 (m, 3H), 6.17 (s, 1H), 4.32 (s, 1H), 2.95 − 2.73 (m, 3H), 2.48 (m, 3H), 2.24 (m, 3H), 1.88 (dd, *J* = 12.2, 6.4 Hz, 2H), 1.82 − 1.62 (m, 3H), 1.60 (d, *J* = 3.3 Hz, 3H), 1.51 − 1.19 (m, 5H).

¹³C NMR (101 MHz, Chloroform-*d*) δ 201.59, 153.63, 141.28, 138.12, 137.36, 133.19, 132.41, 131.47, 126.40, 117.40, 117.32, 115.42, 112.83, 79.44, 55.65, 55.19, 44.34, 43.80, 40.63, 38.15, 32.00, 29.66, 27.62, 26.68, 24.22, 23.67, 22.11. **HR-FT-MS calcd for C₂₇H₃₁O₃Br₁Na₁S₁ 537.1069 found 537.10627.**

4.5.33 Compound MMA312:
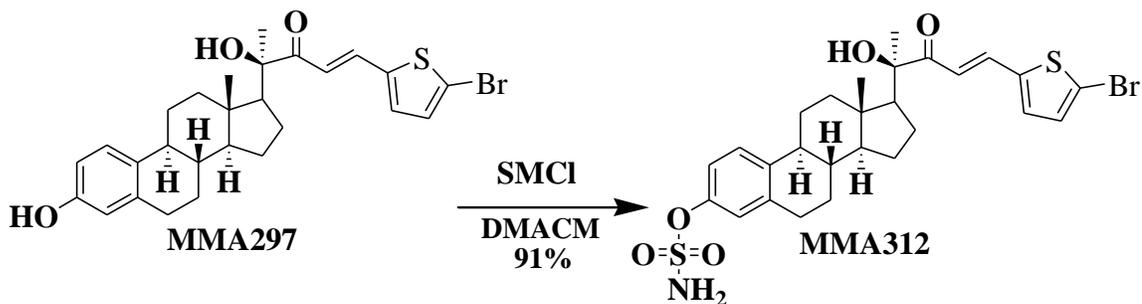
Compound MMA313 (0.2 g, 0.388 mmole) was dissolved in N, N-dimethyacetamide (1 ml) then cooled down to 0 ⁰C; followed by addition of sulfamoyl chloride (0.134 g, 1.164 mmole). The reaction mixture was allowed to be stirred and warm to the room temperature for 18h. Admixture of ethyl acetate (30 ml) and water (30 ml) was added to the reaction mixture and the organic layer was separated, washed by water (3×25 ml) followed by the addition of brain (30 ml). Sodium sulfate anhydrous (Na₂SO₄) to dry the organic layer, filtered and concentrated in vacuo.

Column chromatography (30% ethyl acetate in hexane) to provide **compound MMA312** (0.21 g, 91%).

^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.97-7.80 (m, 1H), 7.38-7.27 (m, 2H), 7.09 (m, 3H), 6.73 (m, 1H), 5.36 (s, 2H), 4.18 (s, 1H), 2.82 (m, 3H), 2.59 (m, 3H), 2.35 – 2.17 (m, 3H), 1.90 – 1.79 (m, 2H), 1.54 (m, 3H), 1.47 (m, 3H), 1.13 – 0.89 (m, 5H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 201.43, 147.88, 141.25, 139.63, 138.88, 137.28, 133.19, 131.47, 126.69, 121.97, 118.98, 117.34, 117.22, 79.17, 55.63, 55.08, 44.19, 43.95, 40.49, 37.65, 31.96, 29.74, 27.27, 26.45, 24.25, 23.64, 22.74. . **HR-FT-MS calcd for C$_{27}$H$_{32}$O$_5$N$_1$Br$_1$Na$_1$S$_2$ 616.0797 found 616.07782.**

4.6   References:

1. Lang, K.L., et al., *Synthesis and cytotoxic activity evaluation of dihydrocucurbitacin B and cucurbitacin B derivatives.* Bioorganic & medicinal chemistry, 2012. **20**(9): p. 3016-3030.

2. Lang, K.L., et al., *Chemical modification produces species-specific changes in cucurbitacin antifeedant effect.* Journal of agricultural and food chemistry, 2013. **61**(23): p. 5534-5539.

3. Bartalis, J. and F.T. Halaweish, *Relationship between cucurbitacins reversed-phase high-performance liquid chromatography hydrophobicity index and basal cytotoxicity on HepG2 cells.* Journal of Chromatography B, 2005. **818**(2): p. 159-166.

4. Bartalis, J. and F.T. Halaweish, *In vitro and QSAR studies of cucurbitacins on HepG2 and HSC-T6 liver cell lines.* Bioorganic & medicinal chemistry, 2011. **19**(8): p. 2757-2766.

5. Schlegel, W., A. Melera, and C. Noller, *Reduction and Oxidation Products of Cucurbitacin B1.* The Journal of Organic Chemistry, 1961. **26**(4): p. 1206-1210.

6. Chenera, B., *Studies toward the total synthesis of cucurbitacins*. 1984.

7. Jung, M.E. and R.M. Lui, *Studies toward the Total Syntheses of Cucurbitacins B and D.* The Journal of organic chemistry, 2010. **75**(21): p. 7146-7158.

8. Ryu, S.Y., et al., *Cytotoxicity of cucurbitacins in vitro.* Archives of Pharmacal Research, 1995. **18**(1): p. 60-61.

9. de Reinach-Hirtzbach, F. and G. Ourisson, *Synthese de la chaine laterale des cucurbitacines.* Tetrahedron Letters, 1973. **14**(16): p. 1363-1366.

10.   Paryzek, Z. and R. Wydra, *Tetracyclic triterpenes. Part VI [1]. The synthesis of 4, 4, 14α-Trimethyl-19 (10→ 9β) abeo-steroids. Synthons for the preparation of cucurbitacins.* Steroids, 1981. **38**(2): p. 141-148.

11.   Levy, E. and D. Lavie, *Attempted skeletal rearrangements in the lanostane series.* Israel Journal of Chemistry, 1970. **8**(4): p. 677-684.

12.   Braun, M., *Fundamentals and transition-state models. Aldol additions of group 1 and 2 enolates.* Modern aldol reaction, 2003. **1**: p. 1-4.

13.   Bodnár, B., et al., *Synthesis and Biological Evaluation of Triazolyl 13α-Estrone– Nucleoside Bioconjugates.* Molecules, 2016. **21**(9): p. 1212.

14.   Kopel, L.C., M.S. Ahmed, and F.T. Halaweish, *Synthesis of novel estrone analogs by incorporation of thiophenols via conjugate addition to an enone side chain.* Steroids, 2013. **78**(11): p. 1119-1125.

15.   Leese, M.P., et al., *Structure–activity relationships of C-17 cyano-substituted estratrienes as anticancer agents.* Journal of medicinal chemistry, 2008. **51**(5): p. 1295-1308.

16.   MacCarthy-Morrogh, L., et al., *Differential effects of estrone and estrone-3-O-sulfamate derivatives on mitotic arrest, apoptosis, and microtubule assembly in human breast cancer cells.* Cancer research, 2000. **60**(19): p. 5441-5450.

17.   Ireson, C., et al., *Pharmacokinetics and efficacy of 2-methoxyoestradiol and 2-methoxyoestradiol-bis-sulphamate in vivo in rodents.* British journal of cancer, 2004. **90**(4): p. 932-937.

18.   Ahmed, M.S., L.C. Kopel, and F.T. Halaweish, *Structural Optimization and Biological Screening of a Steroidal Scaffold Possessing Cucurbitacin‐Like Functionalities as B‐Raf Inhibitors.* ChemMedChem, 2014. **9**(7): p. 1361-1367.

19.   Stander, A., F. Joubert, and A. Joubert, *Docking, synthesis, and in vitro evaluation of antimitotic estrone analogs.* Chemical biology & drug design, 2011. **77**(3): p. 173-181.

20.   Roy, S., et al., *17β-estradiol-linked nitro-L-arginine as simultaneous inducer of apoptosis in melanoma and tumor-angiogenic vascular endothelial cells.* Molecular pharmaceutics, 2011. **8**(2): p. 350-359.

21. Dauben, W.G. and J.J. Takasugi, *Organic reactions at high pressure. Wittig reaction of hindered ketones with nonstabilized ylides.* Tetrahedron letters, 1987. **28**(38): p. 4377-4380.

22. Piancatelli, G., A. Scettri, and M. D'auria, *Pyridinium chlorochromate: a versatile oxidant in organic synthesis.* Synthesis, 1982. **1982**(04): p. 245-258.

23. Tu, Y., et al., *Synthesis of 1, 2: 4, 5‑Di‑O‑Isopropylidene‑D‑Erythro‑2, 3‑Hexodiulo‑2, 6‑Pyranose. A Highly Enantioselective Ketone Catalyst for Epoxidation.* Organic syntheses, 2003: p. 1-8.

24. Greenlee, W. and D. Hangauer, *Addition of trimethylsilyl cyanide to α-substituted ketones: Catalyst efficiency.* Tetrahedron Letters, 1983. **24**(42): p. 4559-4560.

25. Crouch, R.D., *Selective deprotection of silyl ethers.* Tetrahedron, 2013. **69**(11): p. 2383-2417.

Chapter Five

Design, Synthesis and Biological Evaluation of CUCS-Inspired Estrone Analogues

Targeting Multidrug Resistance Protein 1 (MRP1)

5.1 Introduction:

Cancer is a serious disease that rank on the second position among diseases that lead to death, which found to be behind the death of one person in every five or six people in the western countries. It is also responsible for 8.2 million of deaths around the world. Every year more than 14.1 million cancer cases revealed mainly in the developing countries. In addition, people who live with cancer count more than 32.5 million [1, 2].

Cancer is known as serious genatic disorder when the cells become abnormally fast divided. So, if the abnormal divided cells untreated, it will affect the other tissues of the body and lead to death. The main available treatments of cancer are chemotherapeutic agents which can either stop or slow the abnormal fast division of the cells [3, 4]. The major problem associated with the available chemotherapeutic agents is the drug resistance which involved with 30%- 80% of cancer patients [5]. Therefore, there is an urgent need to find a new agent to overcome the chemotherapeutic resistance.

Drug resistance in cancer cells is not specific for one anticancer drug, but the whole available chemotherapeutic agents within the same family can be influenced by the same mechanism. Some cancer cells that developed drug resistance can be resistance to other type of drugs that different in their mechanism of action and structure. This process defined as multidrug resistance (MDR). This phenomenon might clarify the failure drug combination to overcome the cancer cell resistance [3]. There are two main clinical times of MDR; first one happen at the time of treatment and the second is already present at the time of the diagnosis [6].

Various biological reactions represent the frist defense for the cells include:

- Activation of cellular elimination process, cellular uptake process and metabolic reactions to inactivate the drug process inside the cell, all of these processes will cause the decrease of the chemotherapy concentration inside the cell **(Fig.5.1)** [7-9].

  - Changes of the drug delivery to the targeted tissues because of different reasons including poor pharmacokinetics profile of the drug such as absorption, distribution, metabolism and excretion (ADME) [10, 11].

  - Enhance the process of DNA repair [9, 11].

  - Structural modification of the targeted tissues [9, 11].

All the previous process were extracellular factors that increase the cell anticancer drug resistance. However, the cellular factors that affect the drug presence inside the cell play the main role for the cell resistance to anticancer drugs through biochemical changes

in the tumor cells and this process include transport-based MDR known as ATP-binding cassette (ABC).
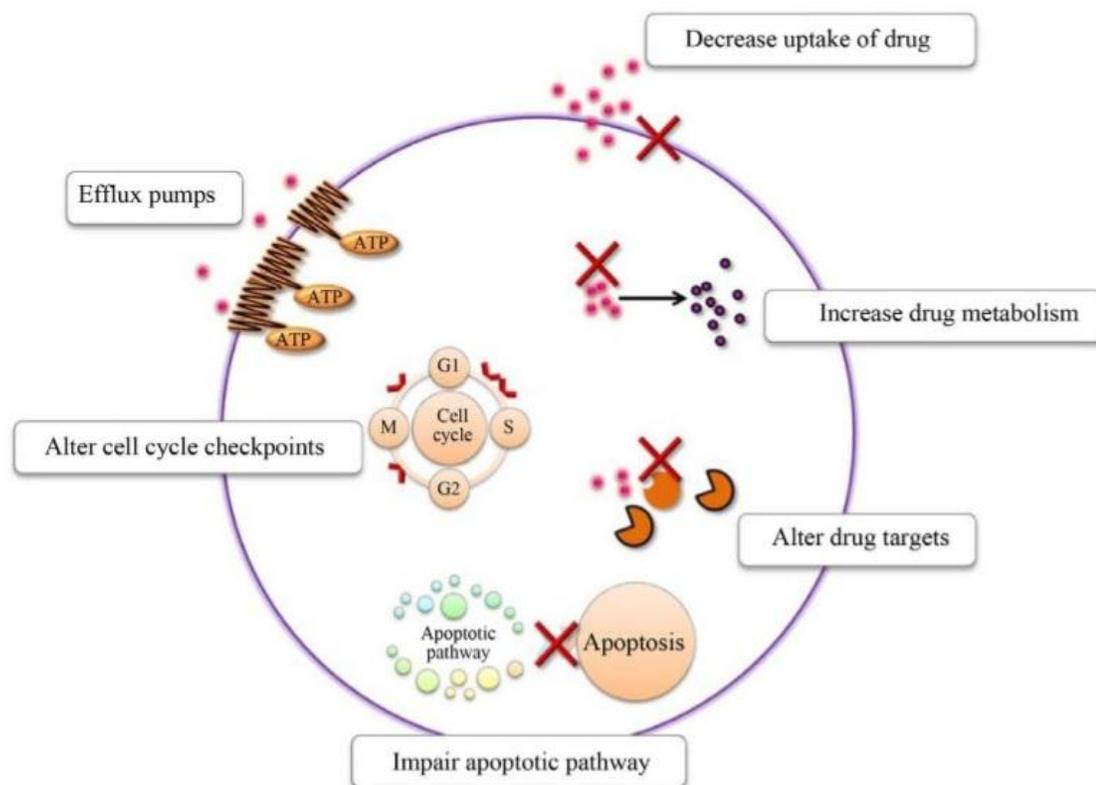
Figure 5.1 Different mechanisms of cell development of drug resistance (copied from [3]).

Different transporter proteins are located in the lipophilic membrane of the cell, which play significant role in the pharmacodynamics and pharmacokinetics of various drugs. Therefore, enormous studies on the transporters have been conducted to identify their locations, functions, structures, selectivity and cellular distribution [3, 12]. Cellular MDR decrease the intracellular concentration of the drugs by enhance the ATP-dependent

efflux pumps, which is one of the family membrane of ATP-binding cassette (ABC) [13]. ATP-binding cassette is the largest transmembrane protein family that demonstrated wide range of specificity. There are 49 known human ABC genes, which are categorized into 7 subfamilies starting with ABC and end with ABCG and the classification were relay on the arranging of the domain and similarity of the sequence [9, 14].

Products resulted from the metabolic reactions, lipids and various types of chemotherapeutic agents are pumped out of the cells through ATB-transporter utilizing ATP-energy dependent movement processes [15-17]. The chemotherapeutic agents that commonly affected by MDR process are hydrophobic containing drugs, natural products such as docetaxel and paclitaxel, anthracyclines (daunorubicin, doxorubicin), anti-microtubule alkaloids (vincristine), antimetabolic agents (6-mercaptpurin, methotrexate, gemcitabine, fluorouracil), natural compounds of epipodophyllotoxin (teniposide and etoposide) and RNA elongation inhibitors (actinomycin-D) [3, 18, 19].

MRP1 (Multidrug Resistance Protein 1) is one of the ABC-transporter protein family specifically belong to the C subfamily that consist of 12 proteins [19, 20]. **MRP1** functionally categorized as an ATP-dependent protein [19]. They can be found in various types of human cells and tissues with distinguishable level in placenta, lung, macrophages, kidney, cardiac muscle, testis and skeletal muscle [21]. **MRP1** is responsible for low intracellular concentration of many hydrophobic anticancer drugs by ATP-dependent efflux mechanism, which lead to improve the cell resistance toward these drugs [17, 22].

For example, various number of hydrophobic and cationic chemotherapeutic agents such as anthracycline (epirubicin, daunorubicin, doxorubicin), methotrexate, mitoxantrone and flutamide have been resisted by **MRP1** ATP-binding efflux [17]. In addition, some phase II metabolic products such as glutathione, sulfate and glucuronide conjugates are preferable substrates of **MRP1** [23]. Therefore, discovery of new agents that has the ability to inhibit the **MRP1** and lead to increase the intracellular concentration of the chemotherapeutic agents to perform their anticancer activity without being pumped out of the cell.

Natural products possess many pharmacophores that provide fertile base for the chemists to discover new analogues targeting various molecular biology [10]. Natural products were extensively studied as precious source for drug design and discovery [24]. Cucurbitacins is one of the natural products that is known of their broad biological activities such as anticancer, hepatoprotective and antiviral activities such as chronic hepatitis [25] **(Fig5.2)**. in addition, cucurbitacins proved to be potent on various molecular pathways [26]. Therefore, cucurbitacins have been chosen to be studied as anti-**MRP1** agents to overcome the drug resistance. However, the limitation of cucurbitacins as natural products was one of the obstacles that redirect our intention to conduct the chemical synthesis of and/or cucurbitacin-inspired estrone analogues. Due to the varieties of functional groups in the cucurbitacins, the total synthesis of cucrbitacins s almost practically challenging.

A study conducted on 24 different types of cucurbitacins, Van Dang identified all the pharmacophores that responsible for different biological activities [27]. In the result of

their studies they demonstrated that α-β-unsaturated ketone of the enone side chain, C-3 functionality and C-16 functional groups are significant for the biological activities [27, 28]. Therefore, possessing these pharmacophores during the chemical synthesis are very important.

Figure 5.2 General structures of cucurbitacins and steroid.

Cucurbitacin	R_1	R_2	R_3	R_4	R_5	$\Delta^{1,2}$	$\Delta^{23,24}$
A	OH	=O	MeOH	H	AC	-	+
B	OH	=O	CH_3	H	AC	-	+
C	H	OH	MeOH	H	AC	-	+
D	OH	=O	CH_3	H	OH	-	+
E	OH	=O	CH_3	H	AC	+	+
F	OH	OH	CH_3	H	OH	-	+
H	OH	=O	CH_3	OH	OH	-	-
I	OH	=O	CH_3	H	OH	+	+
J	OH	=O	CH_3	OH	OH	+	-

The similarity between cucurbitacins general structure and steroids moieties is very close, mainly in the tetracyclic system **(Fig5.2)**. However, the main differences between the two structures are in the presence of aromatic ring of the steroids which is aliphatic ring in the cucurbitacins, also the presence of gem-dimethyl group at C-4 [29] **(Fig 5.2)**. Due to the similarity between the main skeletons of estrone and cucurbitacins, the main pharmacophores of cucurbitacins were chemically installed into the steroid skeleton to be synthesized and biologically evaluated [30, 31].

Utilizing estrone as main skeleton to assemble the cucurbitacin's enone side chain has been done by Ahmed et al. [32]. However, further structural modifications on estrone like cucurbitacin analogues were performed such as assembling various aliphatic and aromatic enone side chains at C-17, modification on C-16 and substituting the hydroxyl group at C-3 with methoxy and sulfamoyl groups to investigate their biological activities as anti-**MRP1** agents **(Fig.5.3)**. The ability of estrone derivatives to dramatically possess a new biological activity along with losing the estrogenic activity as a side effect was the motive to use the estrone as starting scaffold to install various cucurbitacins pharmacophores [33-35].
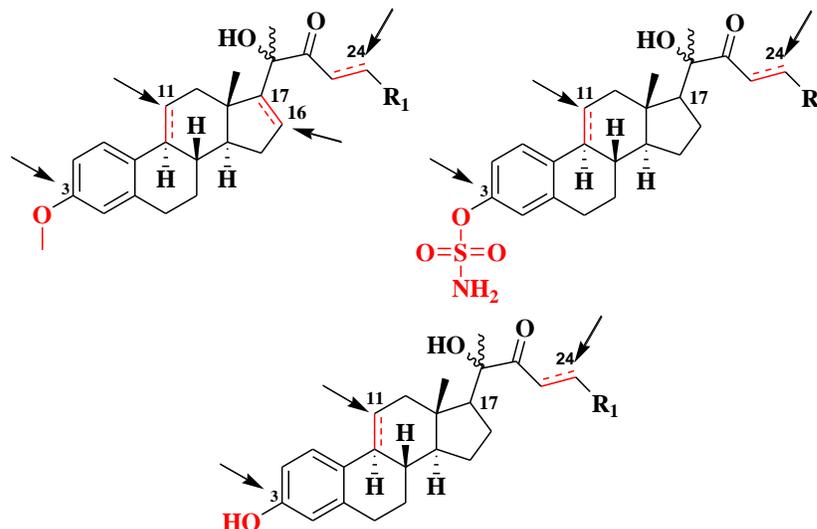
R1= C(CH$_3$)$_2$ OH, P-PhF, P-PhCl, P-PhBr
P-PhCF$_3$, P-PhNO$_2$, 5-Bromo-2-thiophene,
5-Bromo-2-furan.

Figure 5.3 Proposed modified estrone structures.

To prove the ability of CUCUS-inspired estrone analogues to bind to MRP1, molecular docking study were conducted. First, a virtual library of 900 compounds were prepared using the concept of bioisosterism to have a systematic way of designing the structures and these include esterone derivatives with cucurbitacins enone side chain at C-17, various aliphatic and aromatic groups at C-25, double bond at C-16 and C-17 and hydroxyl, methoxy and sulfamoyl groups at C-3, known **MRP1** inhibitors, cucurbitacins. Second, OpenEye® scientific software were utilized for the semi-flexible molecular docking studies including fast executive docking (FRED), OMEGA and VIDA.

5.2 Materials and Methods:

5.2.1 General:

Calcein-AM was acquired from Corning life science. Doxorubicin, Adenosine triphosphate (ATP), Adenosine monophosphate (AMP), estradiol, thiazolyl blue tetrazolium bromide (MTT) and poly-D-lysin were purchased from sigma Aldrich. MK-571 was purchased from Cayman chemical. H69AR cells were provided from ATCC.

5.2.2 Inhibition screening of MRP1 with florescent accumulation of calcein-Am in H69AR cell line:

The inhibition screening of MRP1 were completed in H69AR cells in the presence of known MRP1 inhibitor as positive control (MK-571) and fluorescent substrate (Calcein-AM). The inhibition screening of the CIEA were conducted in 96-well Optical-Bottom plates. The seeding of the cells started by the injection of 6×10^4 cells per well in 100 µL of cultural medium then to be incubated for 24 h. before the injection of the CIEA, cultural medium was substituted with 80 µL of serum-free medium, followed by the addition of 10 µL of the tested compounds (10 µM). the negative control, 0.2% of DMSO in medium, and the positive control, 50 µM of MK-571, were injected in the 96-well plates. Calcein-Am were injected (10 µL, 0.25 µM) after 30 min and the cells were incubated for 1 h. then the cells were washed two times with 100 µL PBS that contain 10 mM HEPES and 4.5% glucose. The fluorescent measurements of the calcein-AM level were performed utilizing MetaXpress software [39].

5.2.3 Data analysis of the screened compounds:

The metaXpress software automatically measure the fluorescent intensity in every single well to obtain the % inhibition of MRP1 [39]. The inhibition activity of the injected compounds on MRP1 (Calcein-AM efflux) was calculated using:

$$\%inhibition = \frac{X_T - X_{calcein-AM}}{X_{MK-571} - X_{calcein-AM}} \times 100,$$

X = average florescent intensity,   **T** = tested compounds.

5.3   Results and Discussion:

5.3.1   Results of Molecular Docking with Homology Structure of MRP1:

     To conduct the molecular docking, the crystal structure of the targeted protein should be used. Since there is no crystal structure for the **MRP1** available in the protein data bank (PDB) [36], the homology structure of **MRP1** was prepared using Swiss-Model Workspace **(Fig.5.4)**. Homology modeling is commonly utilized in various studies including molecular docking of virtual library. It consists of four essential steps; first, template recognition of the protein structure; second, alignment between the structure of the template and the sequence of targeted protein; third, building the model; fourth, general evaluation of the resulted homology structure. The repetition of all of these steps is recommended until a satisfaction reached [37, 38].

The molecular docking studies of the CUCUS-inspired estrone analogues that contain the cucurbitacins enone side chains at C-17 along with various aliphatic and aromatic moieties at C-25, functionalities at C-16 and C-17 and hydroxyl, methoxy and sulfamoyl groups at C-3 on the homology structure of **MRP1** showed distinguishable results of binding affinities to the **MRP1**. Several analogues showed better binding modes compare to the known **MRP1** inhibitors such as brobenecide, **MK-571** and JS-2190 (**Table AP-3**). Compounds that contain aromatic (hydrophobic) enone side chains at C-17, double bond at C-16 and C-17 and methoxy group at C-3 such as **MMA265, MMA270, MMA287, MMA288, MMA289, MMA290, MMA292, MMA305** and **MMA311** showed significant binding affinity toward the 3D structure of the MRP1 by making H-bonds and hydrophobic interactions compare to the analogues that contain aliphatic enone side chains such as **MMA102 and MMA132 (Fig.5.5)**. These results are lined up with the nature of the **MRP1,** which is hydrophobic membrane protein and bind to substrates which are mostly hydrophobic ligands [35].

In addition, stereochemistry played a significant role in the determination of binding affinity. For example, compounds that possess cucurbitacin D stereochemistry with aromatic enone side chain at C-17, double bond at C-16 and C-17 and methoxy group at C-3 such as **MMA265, MMA270, MMA287, MMA288, MMA289, MMA290, MMA292, MMA305** and **MMA311** showed better binding affinity along with H-bonds and hydrophobic interactions compare to the compounds that possess the opposite

stereochemistry with the same functional groups such as **MMA316, MMA317, MMA318, MMA319, MMA320,MMA321,MMA330 and MMA334 (Fig.5.6)**.
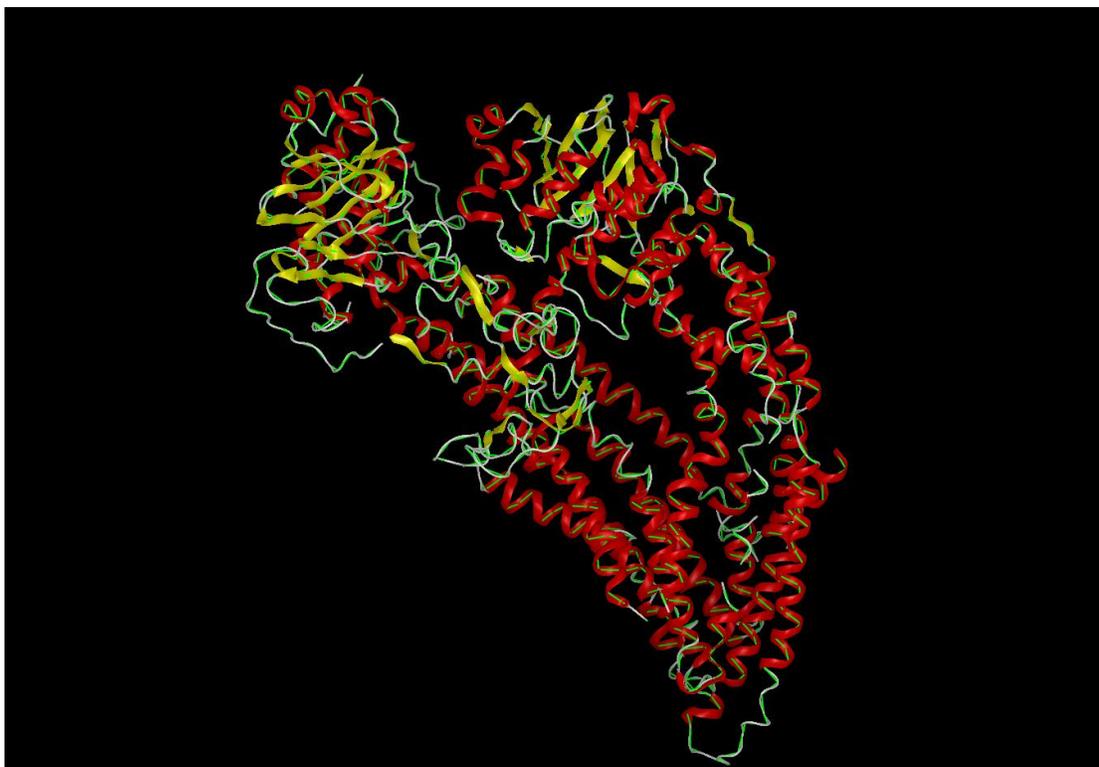
Figure 5.4 Homology structure of MRP1.

On the other hand, the presence of sulfamoyl moiety at C-3 of the estrone skeleton along with aromatic enone side chains at C-17 such in **MMA271, MMA294, MMA295, MMA297, MMA300** and **MMA312** showed significant binding affinity with H-bonds and hydrophobic interactions inside the binding site of the MRP1 compare to the analogues that have the same functional groups but with hydroxyl group at C-3 instead of sulfamoyl

such as **MMA269, MMA306, MMA308, MMA309, MMA310, MMA313** and **MMA314 (Fig.5.7).**

Finally, compound **MMA242** which contain α-hydroxyl ketone at C-17 and sulfamoyl moiety at C-3 demonstrated an outstanding binding mode with H-bonds and hydrophobic interactions with the amino acids inside the binding pocket of the MRP1 crystal structure compare to known inhibitors of MRP1 such as **MK-571 (Fig.5.8).**

In conclusion, CUCUS-inspired estrone analogues that contain aromatic enone side chain at C-17 demonstrated an outstanding binding affinity toward the binding site of the 3D structure of **MRP1** compare to the same analogues but with aliphatic enone side chain at C-17. In addition, stereochemistry played a significant role for the binding affinity; for example, analogues that contain the stereochemistry of cucurbitacin D enone side chain demonstrated better binding mode compare to the same analogues but with opposite stereochemistry. Finally, the presence of sulfamoyl moiety at C-3 of the estrone derivatives along with the aromatic enone side chains proved to improve the binding mode toward the crystal structure of **MRP1** compare to the same compounds but with hydroxyl group at C-3. All of the molecular docking calculations and results were the bases to decide the analogues that should be chemically synthesized and tested in different biological assays as **MRP1** inhibitors to overcome the resistance of the anticancer drugs.
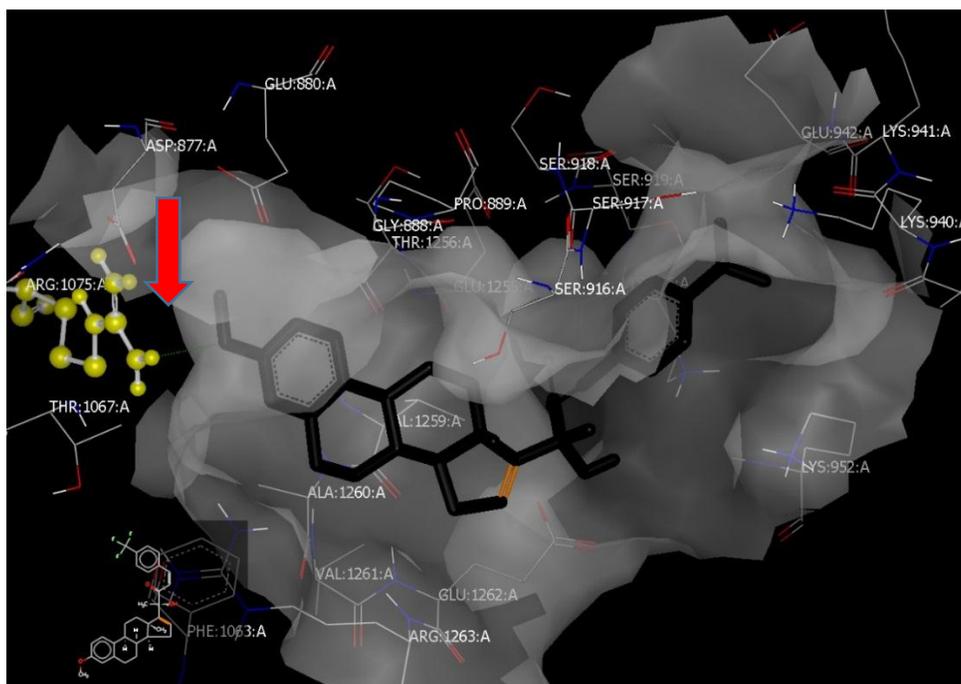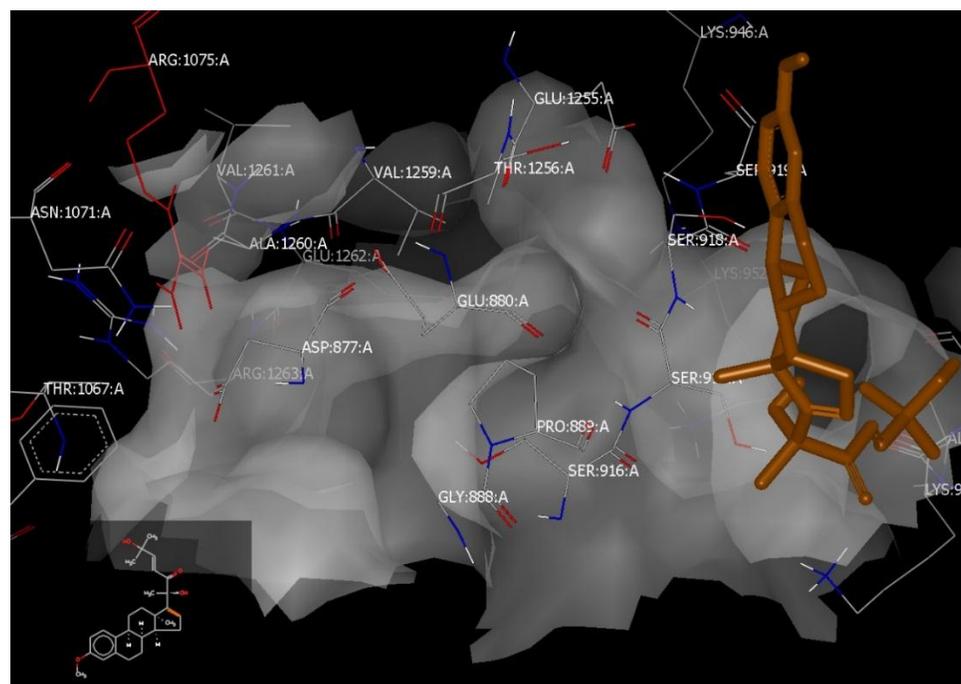
Figure 5. 5 Visual representation of A) MMA292 (black) B) MMA132 (brown) in the MRP1 binding site.

Figure 5.6 Visual representation of A) MMA292 (purple) B) MMA320 (green) in the MRP1 binding site.

5.3.2 Study of synthesized CUCUS-Inspired Estrone Analogues Targeting MRP1:

To explore the inhibitory activity of CUCUS-inspired estrone analogues as **MRP1** inhibitor, three sets of CUCUS-inspired estrone compounds which were synthesized using

various chemical reactions reported in the third and fourth chapters. First set of CUCUS-inspired estrone derivatives contain the cucurbitacins enone side along with aliphatic and aromatic functional groups at C-17, double bond at C-16 and C-17, and methoxy group at C-3 (**Fig.5.9**). Second set of analogues consist of cucurbitacin enone side chain along with aliphatic and aromatic moieties at C-17 and sulfamoyl group at C-3 (**Fig.5.10**). Third set of synthesized compounds contain the cucurbitacin enone side chain along with aliphatic and aromatic moieties at C-17 and hydroxyl group at C-3 (**Fig.5.11**).
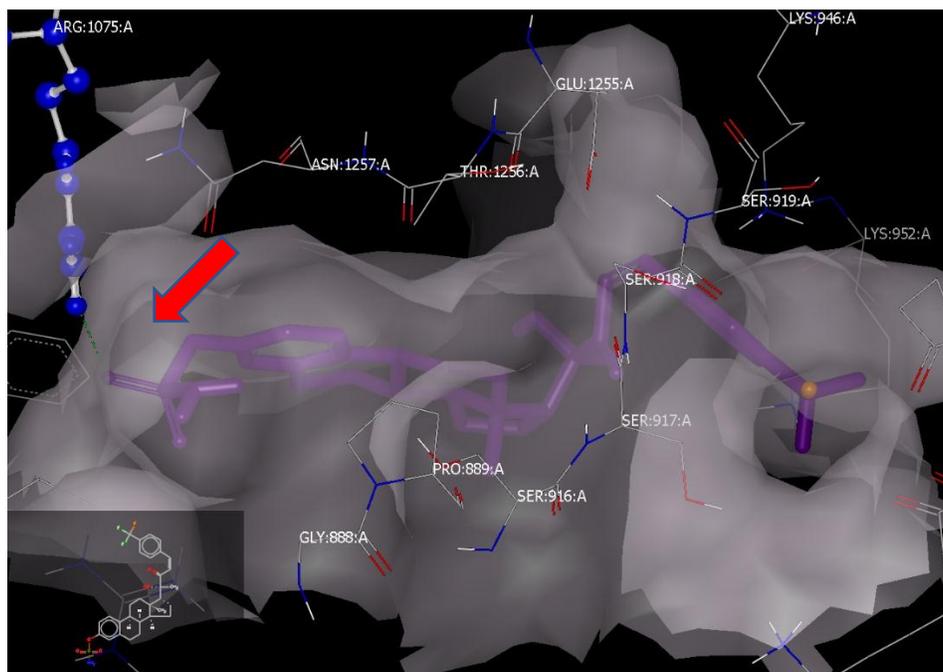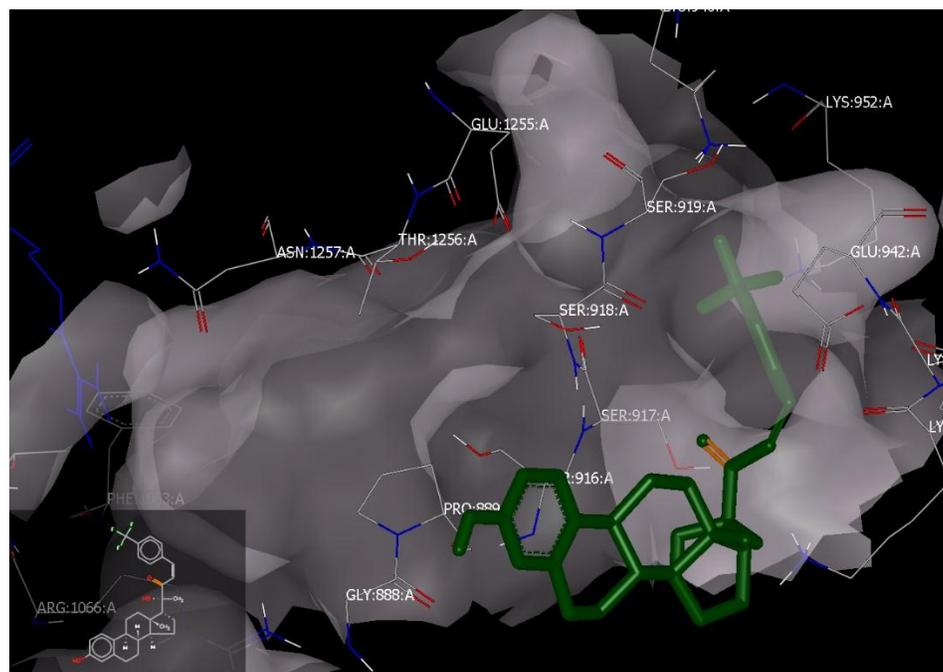
Figure 5.7 Visual representation of A) MMA300 (purple) B) MMA308 (green) in the MRP1 binding site.
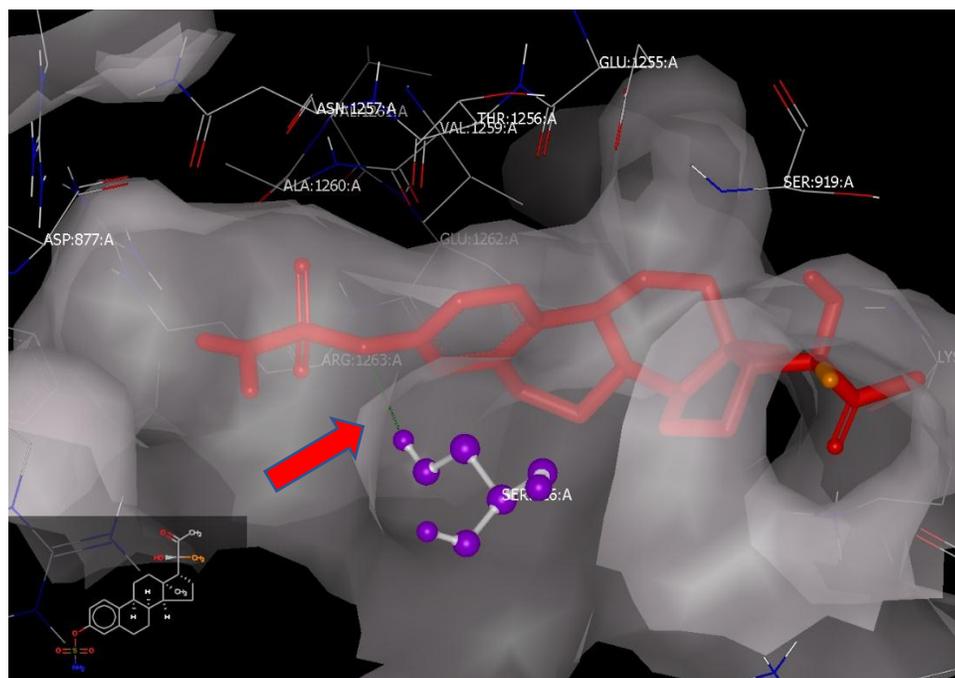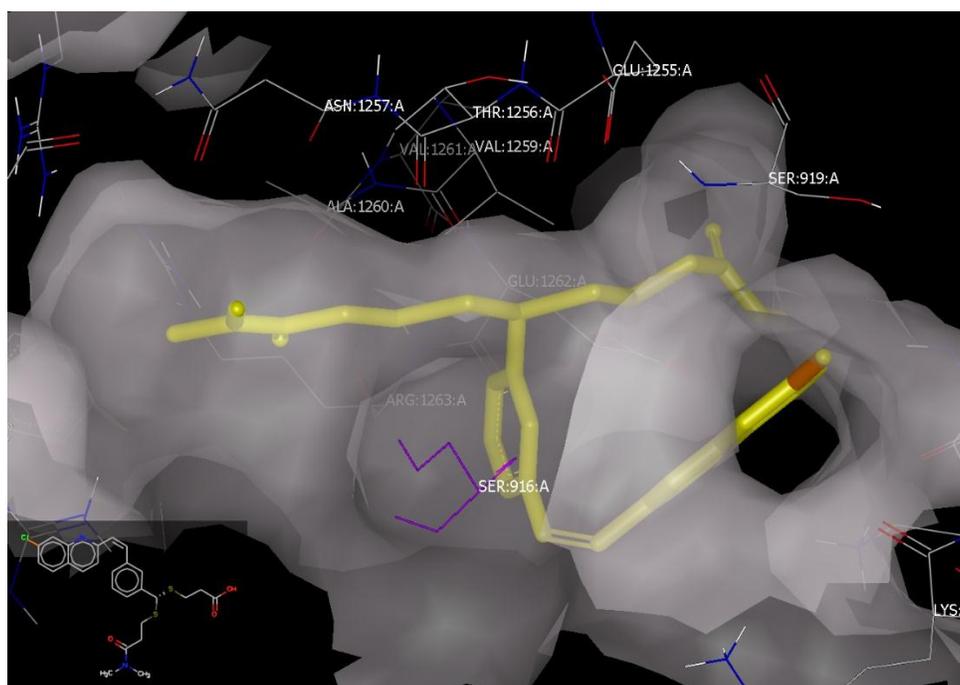
Figure 5.8 Visual representation of A) MMA242 (red) B) MK-571 (yellow) in the MRP1 binding site.
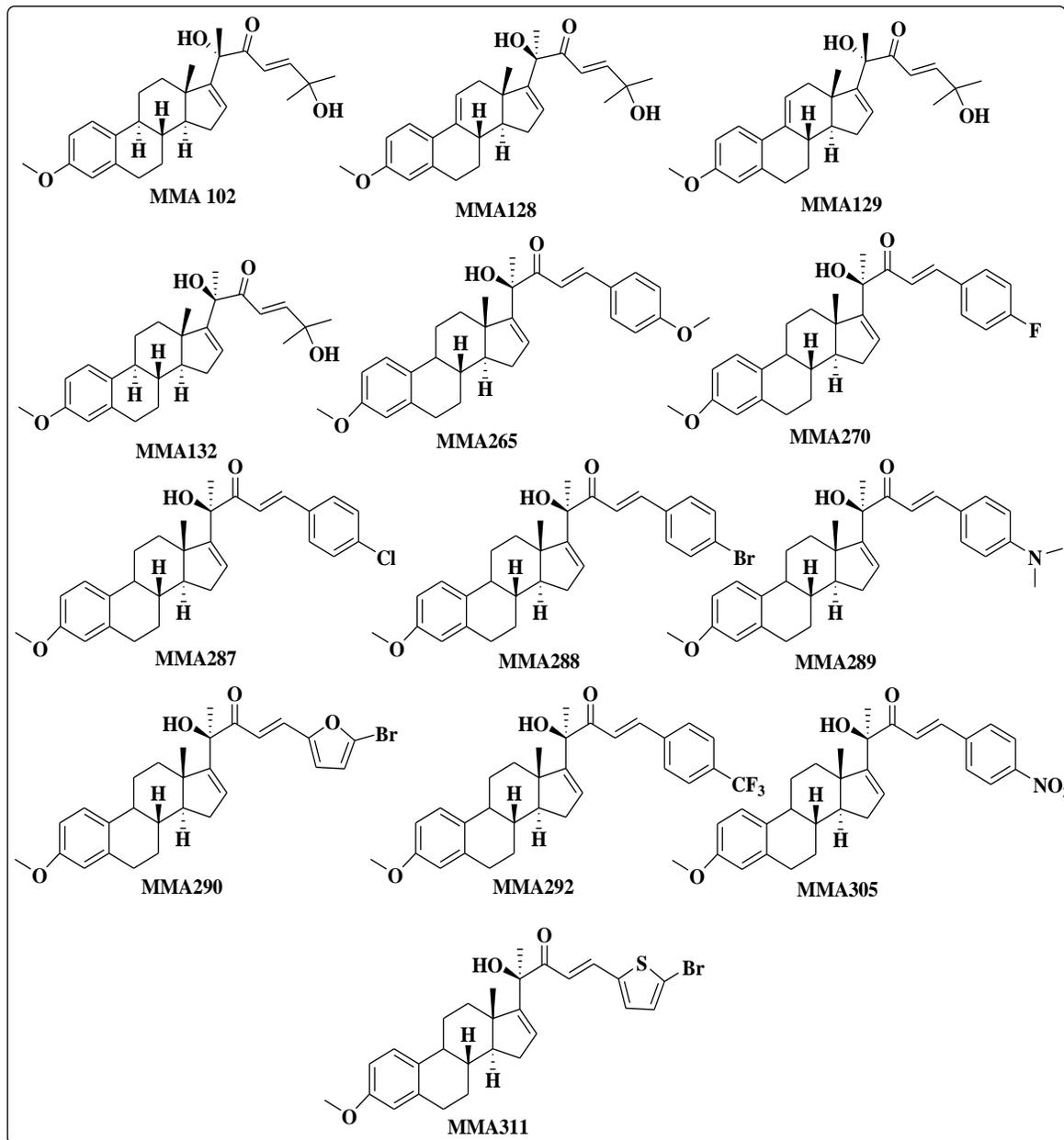
Figure 5.9 First set of synthesized MMA analogues.
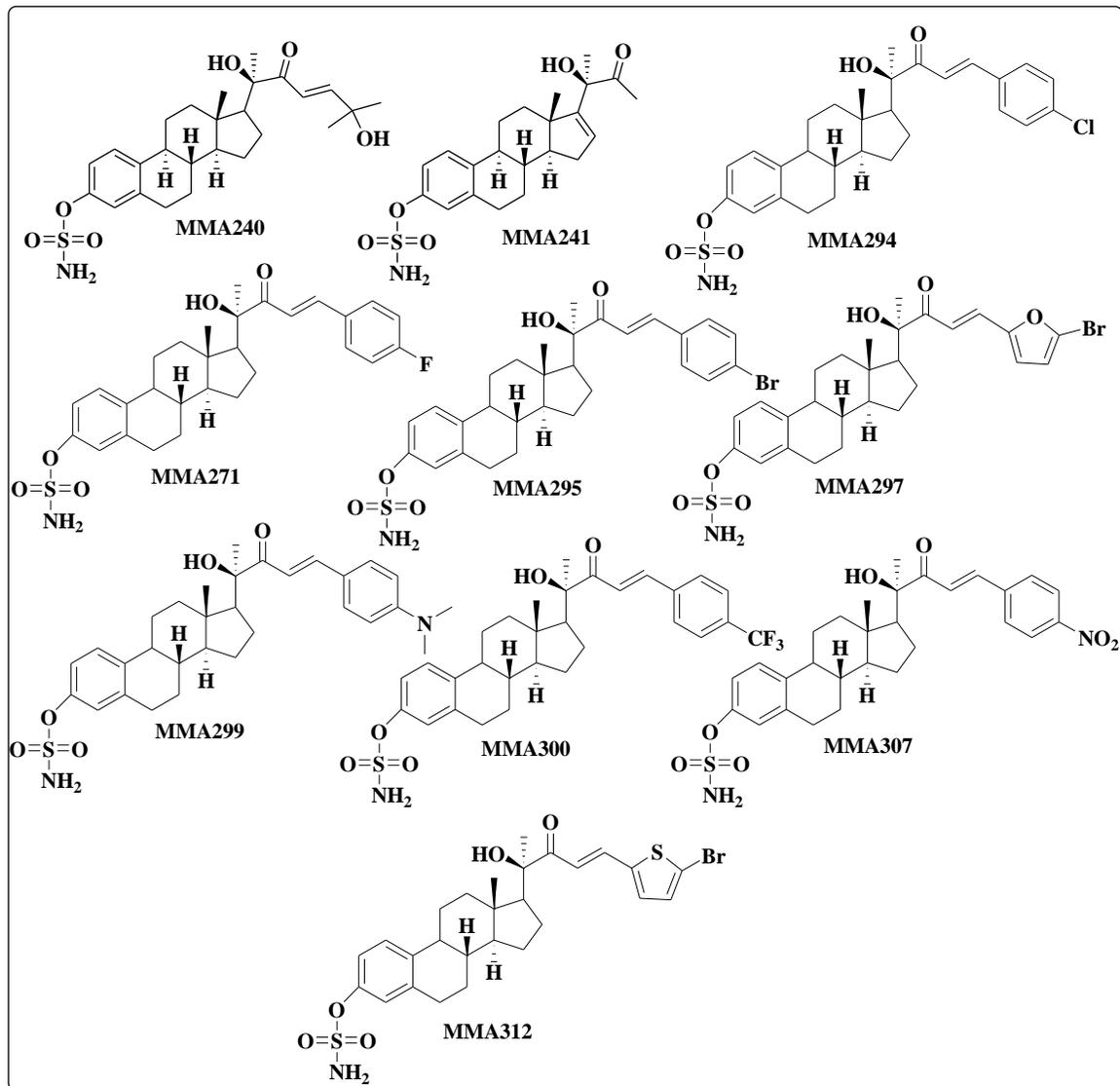
Figure 5.10 Second set of synthesized MMA analogues.
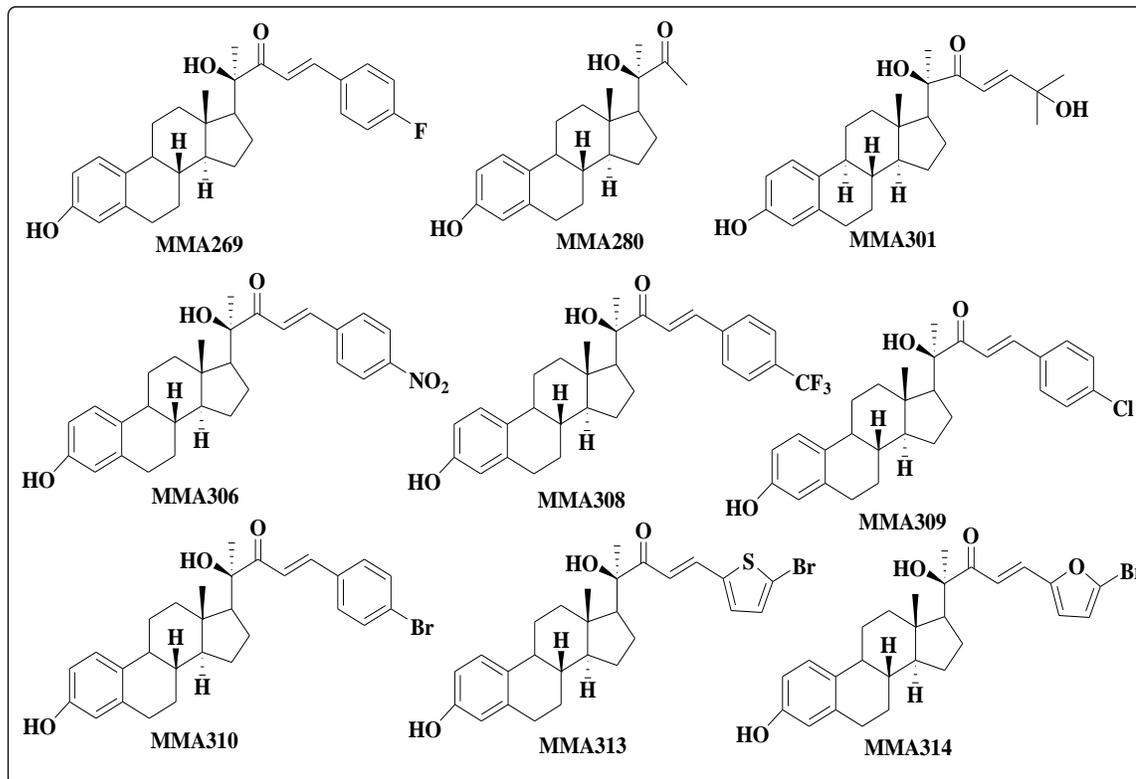
Figure 5.11 Third set of synthesized MMA analogues.

5.1.1 Biological Evaluation of CIAE towards MRP1:

High-content imaging based assay were utilized on the CIEA to detect their inhibitory

activities on **MRP1**. The screening were conducted on H69AR cells in the presence of

Calcein-Am as **MRP1** substrate. **MK-571**, which is known inhibitor of **MRP1**, was used

as standard to compare the % inhibition of CIEA. Compounds **MMAmix, MMA242,**

MMA132, MMA335, MMA337 and **MMA320** showed potential inhibitory activity on

MRP1 (Table 5.2). All of these results were compared to the known **MRP1** inhibitor, **MK-**

571 as shown in figure 5.12. Compound **MMAmix** and **MMA132**, which possess the

enone side chain, proved to be a potential candidate as **MRP1** inhibitor with 70%, 46.2%

inhibitory activity; respectively. Compound **MMA242**, which contain sulfamoyl group at C-3 along with hydroxyl methyl ketone at C-17, demonstrated a very significant inhibitory activity for **MRP1** with 63% inhibition, which demonstrated the important of possessing the sulfamoyl moiety. On the other hand, compounds **MMA320** and **MMA337**, which contain hydrophobic functional groups at C-17, showed moderate inhibitory activities for **MRP1**.

Table 5.2. Inhibitory effect of CIEA to MRP1 activities (% calcein accumulation inhibition) in comparison of standard MRP1 inhibitor MK-571.

Compound ID Provided	Average % Control Inhibition
MK-571	100
MMAMIX	70.2
MMA242	62.6
MMA132	46.2
MMA335	46
MMA337	30.2
MMA320	21.8
MMA326	18
MMA328	17.8
MMA318	17.4
MMA315	16.3
MMA324	15.6
MMA321	15.3
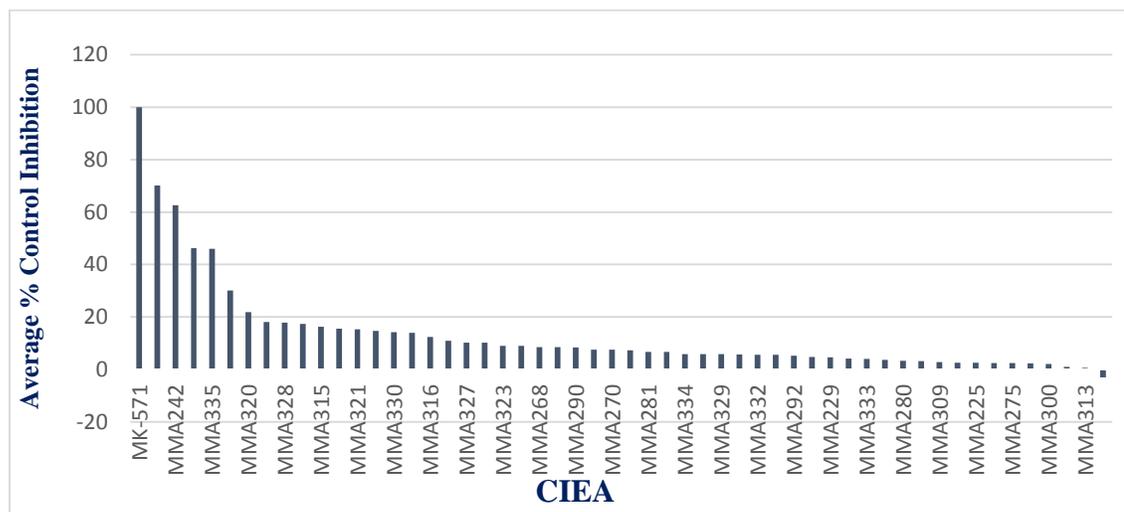
Figure 5.12. Chart represent the inhibitory activities of CIEA for MRP1.

5.2   Conclusion:

CUCUS-inspired estrone analogues were structurally designed utilizing molecular docking technique by assembling the biologically important pharmacophores such as cucurbitacin's enone side chain at C-17, various aromatic and heterocyclic substituents at C-25 and sulfamoyl or hydroxyl groups at C-3 to be explored biologically as anti-cancer candidates toward the treatment of Hepatocellular carcinoma (HCC) through the inhibition of EGFR. Various aliphatic, aromatic and heterocyclic enone side chains were installed chemically on the estrone main skeleton such as isopropanol, para-methoxybenzene, para-flurobenzene, para-chlorobenzene, para-bromobenzene, 5-bromofuran, para-trifluromethyl, para-nitrobenzene, 5-bromothiphene along with sulfamoyl, hydroxyl and methoxy groups at C-3. The biological evaluation using high-content based assay demonstrated that compounds **MMAmix** and **MMA132**, which possess the enone side

chain, proved to be a potential candidate as **MRP1** inhibitor with 70%, 46.2% inhibitory activity; respectively.

5.3   References:

1. Stewart, B. and C.P. Wild, *World cancer report 2014.* 2014.

2. Society, A.C., *Cancer Treatment & Survivorship Facts & Figures.* 2017.

3. Eid, S.Y., et al., *Natural products modulate the multifactorial multidrug resistance of cancer.* Pharmacology & Pharmacy, 2015. **6**(03): p. 146.

4. Altieri, F., et al., *DNA damage and repair: from molecular mechanisms to health implications.* Antioxidants & redox signaling, 2008. **10**(5): p. 891-938.

5. Velingkar, V. and V. Dandekar, *Modulation of P-glycoprotein mediated multidrug resistance (MDR) in cancer using chemosensitizers.* International Journal of Pharmaceutical Sciences and Research, 2010. **2**: p. 104-111.

6. Baguley, B.C., *Multiple drug resistance mechanisms in cancer.* Molecular biotechnology, 2010. **46**(3): p. 308-316.

7. Shen, D.-w., I. Pastan, and M.M. Gottesman, *Cross-resistance to methotrexate and metals in human cisplatin-resistant cell lines results from a pleiotropic defect in accumulation of these compounds associated with reduced plasma membrane binding proteins.* Cancer research, 1998. **58**(2): p. 268-275.

8. Shen, D.W., et al., *Decreased accumulation of [14c] carboplatin in human cisplatin-resistant cells results from reduced energy-dependent uptake.* Journal of cellular physiology, 2000. **183**(1): p. 108-116.

9. Housman, G., et al., *Drug resistance in cancer: an overview.* Cancers, 2014. **6**(3): p. 1769-1792.

10. Jain, R.K., *Delivery of molecular and cellular medicine to solid tumors.* Advanced drug delivery reviews, 2001. **46**(1): p. 149-168.

11. Teodori, E., et al., *The functions and structure of ABC transporters: implications for the design of new inhibitors of Pgp and MRP1 to control multidrug resistance (MDR).* Current drug targets, 2006. **7**(7): p. 893-909.

12. Litman, T., et al., *From MDR to MXR: new understanding of multidrug resistance systems, their properties and clinical significance.* Cellular and Molecular Life Sciences, 2001. **58**(7): p. 931-959.

13. Holohan, C., et al., *Cancer drug resistance: an evolving paradigm.* Nature Reviews Cancer, 2013. **13**(10): p. 714-726.

14. Dean, M., Y. Hamon, and G. Chimini, *The human ATP-binding cassette (ABC) transporter superfamily.* Journal of lipid research, 2001. **42**(7): p. 1007-1017.

15. Juliano, R.L. and V. Ling, *A surface glycoprotein modulating drug permeability in Chinese hamster ovary cell mutants.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 1976. **455**(1): p. 152-162.

16. Ueda, K., et al., *Expression of a full-length cDNA for the human" MDR1" gene confers resistance to colchicine, doxorubicin, and vinblastine.* Proceedings of the National Academy of Sciences, 1987. **84**(9): p. 3004-3008.

17. Cole, S.P., *Targeting multidrug resistance protein 1 (MRP1, ABCC1): past, present, and future.* Annual review of pharmacology and toxicology, 2014. **54**: p. 95-117.

18. Ambudkar, S.V., et al., *Biochemical, cellular, and pharmacological aspects of the multidrug transporter 1.* Annual review of pharmacology and toxicology, 1999. **39**(1): p. 361-398.

19. Keppler, D., *Multidrug resistance proteins (MRPs, ABCCs): importance for pathophysiology and drug therapy*, in *Drug Transporters*. 2011, Springer. p. 299-323.

20. Slot, A.J., S.V. Molinski, and S.P. Cole, *Mammalian multidrug-resistance proteins (MRPs).* Essays in biochemistry, 2011. **50**: p. 179-207.

21. Deeley, R.G., C. Westlake, and S.P. Cole, *Transmembrane transport of endo-and xenobiotics by mammalian ATP-binding cassette multidrug resistance proteins.* Physiological reviews, 2006. **86**(3): p. 849-899.

22. Zhou, S.-F., et al., *Substrates and inhibitors of human multidrug resistance associated proteins and the implications in drug development.* Current medicinal chemistry, 2008. **15**(20): p. 1981-2039.

23. Burg, D., et al., *Inhibition of the multidrug resistance protein 1 (MRP1) by peptidomimetic glutathione-conjugate analogs.* Molecular pharmacology, 2002. **62**(5): p. 1160-1166.

24. Patridge, E., et al., *An analysis of FDA-approved drugs: natural products and their derivatives.* Drug discovery today, 2016. **21**(2): p. 204-207.

25. Lipinski, C.A., et al., *Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings.* Advanced drug delivery reviews, 1997. **23**(1-3): p. 3-25.

26. Gabrielsen, M., et al., *Cucurbitacin covalent bonding to cysteine thiols: the filamentous-actin severing protein Cofilin1 as an exemplary target.* Cell Communication and Signaling, 2013. **11**(1): p. 58.

27. van Dang, G., B.M. Rode, and H. Stuppner, *Quantitative electronic structure-activity relationship (QESAR) of natural cytotoxic compounds: maytansinoids, quassinoids and cucurbitacins.* European journal of pharmaceutical sciences, 1994. **2**(5-6): p. 331-350.

28. Matsuda, H., et al., *Cucurbitane-type triterpenes with anti-proliferative effects on U937 cells from an egyptian natural medicine, Bryonia cretica: structures of new triterpene glycosides, bryoniaosides A and B.* Chemical and Pharmaceutical Bulletin, 2010. **58**(5): p. 747-751.

29. Miro, M., *Cucurbitacins and their pharmacological effects.* Phytotherapy research, 1995. **9**(3): p. 159-168.

30. Kopel, L.C., M.S. Ahmed, and F.T. Halaweish, *Synthesis of novel estrone analogs by incorporation of thiophenols via conjugate addition to an enone side chain.* Steroids, 2013. **78**(11): p. 1119-1125.

31. Parihar, S., et al., *Gallic acid based steroidal phenstatin analogues for selective targeting of breast cancer cells through inhibiting tubulin polymerization.* Steroids, 2012. **77**(8): p. 878-886.

32. Ahmed, M.S., L.C. Kopel, and F.T. Halaweish, *Structural Optimization and Biological Screening of a Steroidal Scaffold Possessing Cucurbitacin‐ Like Functionalities as B‐ Raf Inhibitors.* ChemMedChem, 2014. **9**(7): p. 1361-1367.

33. Tietze, L.F., et al., *A novel approach in drug discovery: synthesis of estrone–talaromycin natural product hybrids.* Chemistry–A European Journal, 2000. **6**(20): p. 3755-3760.

34. Sinha, S., et al., *A lipid-modified estrogen derivative that treats breast cancer independent of estrogen receptor expression through simultaneous induction of autophagy and apoptosis.* Molecular Cancer Research, 2011. **9**(3): p. 364-374.

35. Mueck, A. and H. Seeger, *2-Methoxyestradiol—biology and mechanism of action.* Steroids, 2010. **75**(10): p. 625-631.

36. bank, P.d., http://www.rcsb.org/pdb/home/home.do.

37. Arnold, K., et al., *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.* Bioinformatics, 2006. **22**(2): p. 195-201.

38. Martí-Renom, M.A., et al., *Comparative protein structure modeling of genes and genomes.* Annual review of biophysics and biomolecular structure, 2000. **29**(1): p. 291-325.

39. Peterson, B.G., et al., *High-content screening of clinically tested anticancer drugs identifies novel inhibitors of human MRP1 (ABCC1).* Pharmacological Research, 2017. **119**: p. 313-326.

Chapter Six

General Conclusion and Future Trends

Our objectives and goals of this study were accomplished through design, synthesis and biological evaluation of novel hit candidates as anticancer agents. Various cucurbitacins structures were studied and reviewed as promising anticancer natural products utilizing significant number of literature reviews and articles in order to determine essential pharmacophores that responsible for their anticancer activities. Studies in our group by Ahmed et al have determined the significant pharmacophores including α-β-unsaturated enone side chain, functionalities at C-3, functional groups at C-11 and C16. However, cucurbitacins have very weak selectivity and specificity due to the presence of multiple functional groups which targets different biological targets. Therefore, there is need to design novel compounds based on the cucurbitacin main structure to overcome the side effects of the available cucurbitacins.

The first step to find the lead compounds was the design step utilizing the concepts of hybrid structural design between the main scaffold of the estrone to assemble the cucurbitacins pharmacophores and investigate their anticancer activities. The concept of bioisosterism was used to install different aliphatic and aromatic moieties at C-3, C-16 and C-25 of the designed analogues along with the cucurbitacins pharmacophores. Three

different objectives were outlined in this research to optimize the biological activities of CIEA as anticancer agents.

First objective, CUCS-inspired estrone analouges were designed by assembling the α-β-unsaturated enone side chain of the cucurbitacins at C-17 of the estrone main structure along with various aliphatic and aromatic moieties at C-25, methoxy at C-3 and double bond at C-16 and C-17. The molecular docking results reveled the significant of possessing the α-β-unsaturated enone side chain of cucurbitacin D with the exact stereochemistry at C-18 as in **MMA132**, which demonstrated an outstanding binding affinity toward the crystal structure of **EGFR** compare to known inhibitor such as Erlotinib. Compound **MMA132** found to have H-bond with **MET:769:A** which is the same H-bond with erlotinib that responsible for its anticancer activity.; while the same compound but with opposite stereochemistry at C-18 as in **MMA102** did not show H-bond inside the binding pocket of the crystal structure of **EGFR**. Compounds with various aromatic groups at C-25 of the enone side chain showed different binding modes with the 3D structure of **EGFR** compare to erlotinib. The top consensus scored analogues were chemically synthesized utilizing different chemical reactions and their potential biological activity towards HCC were investigated. Biological assays including **MTT** cell viability assay, western blot and flowcytometry were used to study the mechanism of the anticancer activities of the synthesized compounds. Compound **MMA132** and **MMA102** showed a significant cytotoxicity with IC$_{50}$ values 2μM and 3μM, respectively. Compounds that contain aromatic and heterocyclic enone side chains such as **MMA265**, **MMA270, MMA287,**

MMA288, MMA290, MMA292, MMA305, MMA311, MMA316, MMA318, MMA319, MMA320, MMA321, MMA330, MMA333 and **MMA334** showed various cytotoxicity toward **HepG2** with IC$_{50}$ values 16 µM, 7 µM, 25 µM, 32 µM, 6 µM, 2 µM, 29 µM, 0.7 µM, NA, NA, NA, 3 µM, 20 µM, 11 µM, 2 µM and 8 µM; respectively. Western blot experiment was used to detect the antiproliferative mechanism of the potent compounds, the results of the Western blot showed the ability of **MMA132** to inhibit the phosphorylated EGFR after 48 h of incubation at 3 µM. Finally, flowcytometry assay showed induction activity at G1/S phases for compound **MMA132**.

To overcome the higher hydrophobic characters of the first set of CIEA compounds and expected O-demethylation of C-3 methoxy group which may trigger the estrogenic activity of the CUCUS-inspired estrone analogues side effects, sulfamoyl pharmacophore was assembled at C-3 to improve the pharmacokinetic profile of CIEA. Molecular docking study for the estrone analogs that contain cucurbitacin D enone side chain at C-17 along with different aliphatic and aromatic moieties at C-25 and sulfamoyl and hydroxyl groups at C-3 were conducted to predict the binding to molecular target (**EGFR**). The results of the molecular docking showed that compound with the sulfamoyl group at C-3 of the estrone scaffold along with various aliphatic and aromatic enone side chain showed significant binding affinity with H-bonds and hydrophobic interactions in comparison to the same compounds but with hydroxyl group instead of sulfamoyl at C-3. The results revealed the importance of assembling the sulfamoyl group along with the different enone side chains at the estrone scaffold for the binding affinity with the EGFR which may

enhance the biological activity. Different chemical reactions were utilized to install these functionalities at the estrone main structure. All synthesized analogues were biologically screened to verify their anticancer activity. Cytotoxicity of all synthesized compounds on hepatocellular carcinoma cell line such as HepG2 were conducted using MTT cell viability assay. The IC_{50} values of the synthesized analogues including **MMA240, MMA301, MMA242, MMA267, MMA268, MMA269, MMA271, MMA309, MMA294, MMA310, MMA295, MMA297, MMA314, MMA300, MMA308, MMA306, MMA307, MMA313** and **MMA312** were 8 µM, NA, 3 µM, 12 µM, 7 µM, 2 µM, 8 µM, NA, 11.5 µM, 6 µM, NA, 1 µM, 1.5 µM, 3 µM, 8 µM, 8 µM, 10 µM, 9 µM and 8 µM; respectively.

Finally, our objective to study the potential of synthesized CIEA to overcome the resistance of cancer to current chemotherapy, we examine the inhibitory activity of synthesized analogs on MRP1 which is one of the known protein that responsible for multidrug resistance. Molecular docking study between the homology structure of MRP1 and CIEA revealed that the presence of hydrophobic moieties in the enone side chain at C-17 such as **MMA265, MMA270, MMA287, MMA288, MMA290, MMA292, MMA305** and **MMA311** showed enhanced binding affinity in comparison to the aliphatic enone side chain such as **MMA102** and **MMA132**. In addition, the presence of sulfamoyl group at C-3 along with hydrophobic (aromatic) enone side chain is essential for the binding affinity. Biological activity of the synthesized compounds as MRP1 inhibitors were verified using high-content imaging based assay. Compounds **MMAmix, MMA242, MMA132,**

MMA335, MMA337 and **MMA320** showed potential inhibitory activity on **MRP1 (Table 5.2)** with % inhibition 70%, 63%, 46.2%, 46%, 30% and 22%; respectively, in comparison to MK-571, which known MRP1 inhibitor.

Our study demonstrated the design, synthesis of novel CIEA analogs of potent anti-proliferation/anticancer activities toward hepatocellular carcinoma and potential application to overcome cancer resistance to current chemotherapeutic agents.

Appendix

Protocol of the Molecular Modeling

I-   Steps of designing the virtual library:

The first step of designing the virtual library was the drawing of the structure of the compounds utilizing Chem Draw software. Then all compounds in the virtual library were energy minimized using chem3-D software, specifically MMFF94 application to obtain the relative crystal structure of each compound. Every single compound was saved in pdb format. After that, all created pdb files were combined into single pdb file. All of these processes were conducted utilizing the command prompt as following start menu> command prompt, then the address where the pdb file of the ligands were copied to be used to combined into single pdb file as following:

cd address where the pdb file are located (enter)

Then using command, copy *.pdb name.pdb (enter), which tells the computer to combined all the pdb files to be in one pdb file. The created pdb file can be visualized utilizing notepad as shown in figure AP-1.

II-   Creating Different Conformers of each Ligands:

OMEGA application, particularly MMFF94 force filed was used to create different conformers of each ligand so to be in the fast rigid exhaustive docking (FRED). The resulted file of this process will be in this format name.gz.pdb. this process were done following these steps:

a- Copy the combined file of ligand pdb.

b- Past it where the OMEGA application is.

c- The application license should be available at the same place.

d- The following command was used to generate the conformers "omega2.2.1 –in name.pdb –out newname.gz.pdb –includeinput –warts (Enter). The created file is the one that should be used in the FRED calculation.

Figure AP-1 created pdb file using command prompt.

III- **Preparation of the targeted receptor:**

FRED make receptor were utilized to prepare the targeted receptor. By starting the program, clicking on file to import the pdb file of the receptor that downloaded from the protein data bank to show the chains of the receptor, water molecules, binding ligands and co-factors. Choosing the ligand that co-crystalized with the targeted protein help to generate the grid box for the docking calculations, since it binds to the binding pocket. The size of the created box cannot exceed 50,000- 60,000 A° and that can be controlled by the

mods controls. Next step is to generate the shape of the grid box by choosing the shape mode. The created shape assign the shape of the active site of the receptor which will be used during the molecular docking. The active site consist of two contours outer and inner contours. Finally, the prepared receptor should be saved as .oeb.gz format (Fig.AP-2).
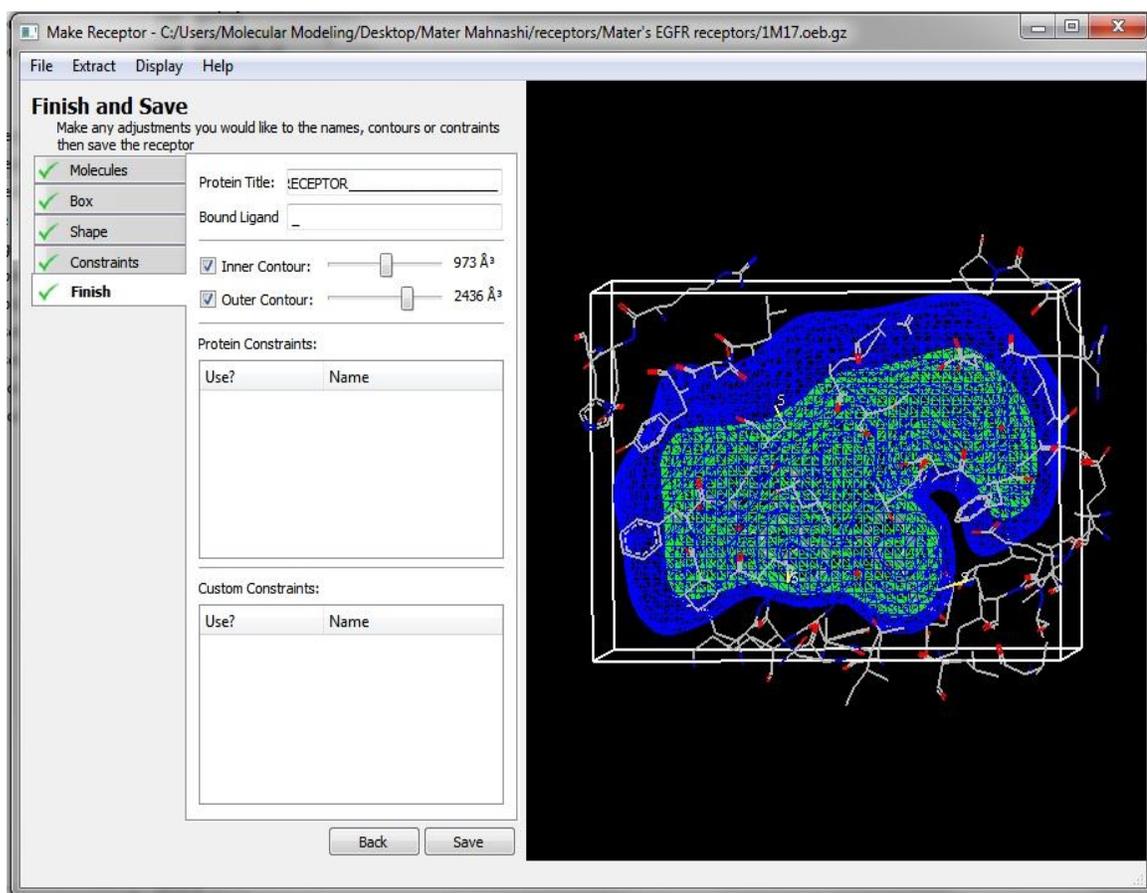
Figure AP-2 Sample of the prepared grid box.

IV- Fast Rigid Executive Docking (FRED) Calculations:

In order to conduct the FRED calculations, three files should be available at where the FRED application is located. First, the file of the fred license; second, the file that contain the different conformers of each ligands, which generated by OMEGA name.gz.pdb; third, the prepared rigid receptor "receptor.oeb.gz". then certain command

used to start the FRED calculations "fred.exe –rec receptor.oeb.gz –dbase name.gz.pdb –prefix newname shapegauss –chemgauss3 –oechemscore –screenscore –plp –hitlist_size 5000 (enter)". The result of FRED calculations will be in consensus score, which calculate the binding affinity between the ligands and the targeted receptors. The lower the consensus score the better the binding affinity (Fig.AP-3).

Table AP-3 sample of consensus score of the FRED docking.

VIDA Name	VIDA ID	PLP	Chemgauss3	OEChemscore	Screenscore	Consensus Score
MMA292_6	2	-48.8383	-62.3556	-41.8789	-126.747	11
MMA305_71	3	-53.1802	-58.3102	-43.6724	-113.016	14
MMA266_109	4	-55.3548	-52.8859	-45.5678	-122.833	18
MMA311_103	5	-46.8621	-58.0439	-39.7369	-113.75	29
MMA128_44	6	-42.578	-55.8818	-43.6541	-116.397	31
MMA291_69	7	-44.6388	-54.3626	-39.1158	-118.789	36
MMA262_67	8	-44.5989	-60.0755	-41.0308	-102.036	43
MMA290_34	9	-44.039	-57.059	-39.6733	-108.785	43
MMA282_93	10	-43.5212	-52.4658	-43.2122	-110.732	45
MMA303_45	11	-43.5525	-51.6006	-45.7898	-105.967	48
MMA270_125	12	-43.5141	-63.5376	-39.6715	-104.555	49
MMA281_159	13	-45.0053	-48.3764	-43.1412	-116.666	50
MMA271_33	14	-41.0513	-54.1366	-40.7371	-107.453	54
MMA265_143	15	-42.9178	-53.0514	-41.1596	-101.578	64
MMA289_39	16	-42.6201	-52.9413	-41.639	-95.7639	74
MMA267_23	17	-39.3882	-53.0557	-37.9022	-107.407	77
MMA296_154	18	-42.6266	-54.3291	-34.3674	-111.931	77
MMA243_20	19	-37.7833	-53.6937	-39.8163	-104.531	78

MMA312_105	20	-45.2513	-48.6491	-37.2378	-105.689	80
MMA285_15	21	-38.3901	-45.4425	-41.2682	-118.057	97
MMA288_52	22	-39.5374	-54.1164	-35.5891	-97.6439	100
MMA284_3	23	-40.9979	-53.7463	-34.8896	-96.2328	104
MMA102_159	24	-35.2947	-45.2269	-39.7102	-109.6	132
MMA100_33	25	-35.7384	-51.3318	-42.029	-83.6088	136
MMA129_35	26	-40.7517	-45.4562	-39.0489	-92.2197	138
MMA101_33	27	-35.7384	-51.3318	-42.029	-83.6088	140
MMA287_27	28	-33.7545	-52.434	-36.1576	-96.1252	142
MMA261_191	29	-43.7523	-41.8876	-34.9793	-93.2034	161
MMA297_168	30	-37.1344	-50.7964	-33.1599	-94.6202	161
MMA279_5	31	-36.2629	-52.6934	-33.4031	-87.0033	165

V-   VIDA visualization of the Docking Results:

VIDA application can be utilized to visualize and show 3D structure of the docked ligands inside the targeted receptor (Fig.AP-4). In addition, by choosing the "Data analysis", all the docking numbers can be shown as in figure AP-3. Snapshot of the best binding ligands with the receptor can be taken to visualize the behavior of the synthesized ligands with the targeted receptor.
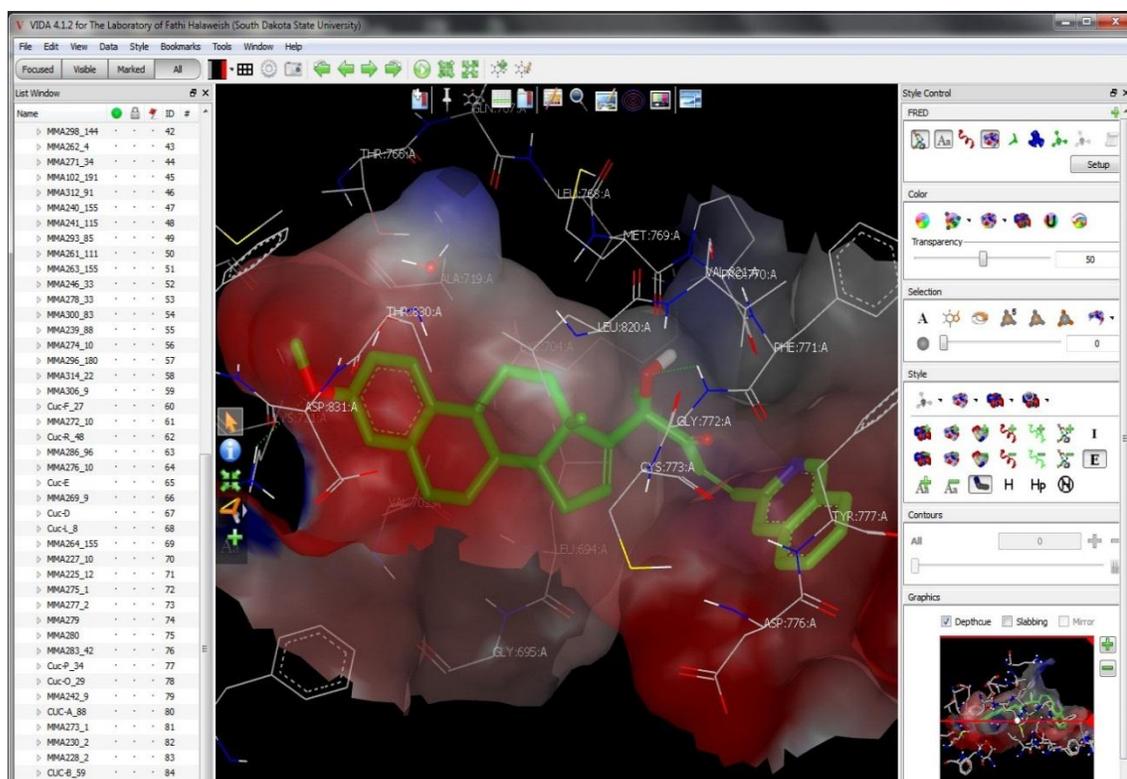
Figure AP-4 Sample of VIDA Visual Representation.