2023

Design, Synthesis, and Biological Evaluation of Novel Steroidal Analogs: Potential Anticancer Agents

Trevor Ostlund
South Dakota State University, trevor.ostlund@sdstate.edu

DESIGN, SYNTHESIS, AND BIOLOGICAL EVALUATION OF NOVEL STEROIDAL ANALOGS: POTENTIAL ANTICANCER AGENTS

BY

TREVOR RYAN OSTLUND

A dissertation submitted in partial fulfillment of the requirements for the

Doctor of Philosophy

Major in Chemistry

South Dakota State University

2023

DISSERTATION ACCEPTANCE PAGE
Trevor Ostlund

This dissertation is approved as a creditable and independent investigation by a candidate for the Doctor of Philosophy 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 Halaweish
Advisor Date

Douglas Raynie
Department Head Date

Nicole Lounsbery, PhD
Director, Graduate School Date

This dissertation is dedicated to a number of people who have inspired and encouraged me throughout my academic career.

To my advisor Dr. Halaweish, you have gained my admiration through your continued support, mentorship, and help throughout my life. I will never forget it.

To my parents, who have offered their support and gratitude throughout my life and put me where I am today.

To my friends and colleagues, who have helped me through this journey and given me advice when times were tough.

ACKNOWLEDGEMENTS

The author would like to pass appreciation to those who supported and participated in these projects. I would like to acknowledge and thank my advisor Dr. Fathi Halaweish for his unwavering support and guidance. Dr. Halaweish has mentored me from the beginning of my research career and has taught me how to be a better researcher, mentor, and leader. I would like to thank the Department of Chemistry and Biochemistry for their support through my academic career and allowing me to be a part of this journey. I would also like to thank my graduate committee members Dr. Cheng Zhang, Dr. Rachel Willand-Charnley, and Dr. Robert McTaggart for their helpful and constructive criticism. I want to give a special thanks to all those in the Halaweish research group past and present (Mater, Mahrous, Khaled, Sara, Faez, John, Felix, Saad, and Kakan) for their discussions and help as well as all the undergraduates and colleagues that have passed along under my guidance and assisted me along the way. I also want to thank my close friends for their help through my degree. Finally, I have to thank my parents and family who have supported my dream every step of the way.

CONTENTS

ABBREVIATIONS

ABC	ATP Binding Cassette
ADME	Adsorption, Distribution, Metabolism, and Excretion
Akt	Protein Kinase B
BCRP	Breast Cancer Resistance Protein
CADD	Computer Aided Drug Design
COX	Cyclooxygenase
DDQ	2,3-dichloro-5,6-dicyano-1,4-benzoquinone
DMSO	Dimethylsulfoxide
EGFR	Epidermal Growth Factor Receptor
ERK	Extracellular signal-related Kinase
FRED	Fast Rigid Exhaustive Docking
GPER	G-Protein coupled Estradiol Receptor
HTS	High Throughput Screening
JAK	Janus Kinases
MAPK	Mitogen Activated Protein Kinase
MMGBSA	Molecular Mechanics with Generalized Born and Surface Area solvation.
MRP	Multidrug Resistance Protein
MS	Mass Spectroscopy
mTOR	mechanistic Target of Rapamycin
NF-κB	Nuclear Factor Kappa B
NMR	Nuclear Magnetic Resonance
NOS	Nitric Oxide Synthase
NSCLC	Non-Small Cell Lung Cancer
PDAC	Pancreatic Ductal Adenocarcinoma

P-gP	P-glycoprotein
PI3K	Phosphoinositide 3-Kinase
QSAR	Quantitative Structure Activity Relationships
RAF	Rapidly Accelerated Fibrosarcoma
RAS	Rat Sarcoma Virus
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
STAT	Signal Transducer and Activator of Transcription
THF	Tetrahydrofuran
TKI	Tyrosine Kinase Inhibitor
TLC	Thin Layer Chromatography

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

DESIGN, SYNTHESIS, AND BIOLOGICAL EVALUATION OF NOVEL STEROIDAL ANALOGS: POTENTIAL ANTICANCER AGENTS

TREVOR RYAN OSTLUND

2023

Cancer is a malicious disease with low survivability rates primarily due to lack of reliable treatment options. Treatments are often non-specific for patients, and cancer resistance to treatment is becoming more prevalent as fast as new therapies are created. Previous studies in our laboratory have focused on the Epidermal Growth Factor Receptor (EGFR) and its downstream kinase family members as a key target for cancer. This cascade controls cell proliferation, anti-apoptotic effects, cell regulation, and other survival factors. Additionally, ABC transporters, responsible for efflux of chemotherapeutic material from cells, are key targets to combat drug resistance. We first evaluated a set of previously synthesized triazole-estradiol analogs using ovarian (A2780) and colorectal (HT-29) cancer cell lines. Colorectal cancer is the second deadliest cancer and is often difficult to treat. Current chemotherapies limit survivability rates to on average an extension of only 6-8 months. Ovarian cancer is the deadliest reproductive cancer and is known as a silent killer due to its difficulty in detection and ability to resist drug action. We found that lead analog Fz25 shows low micromolar cytotoxicity in both lines (15.29 ± 2.19 μM and 15.98 ± 0.39 μM respectively). Inhibition studies in both cell lines show inhibition against various proteins in the EGFR pathway, namely EGFR (23%), STAT3 (65%), ERK (27%), and mTOR (26%). Fz25 was also studied against ABC Transporters and was found to inhibit P-glycoprotein up to 50% as the known inhibitor and Breast Cancer Resistance Protein better than the known inhibitor, suggesting it plays a role in combating drug

resistance. Additional studies were performed to monitor the binding effects of Fz25 using molecular dynamic simulations and molecular docking and were found to compliment biological findings.

In a second study, we wanted to focus on design of a new analog for treating pancreatic cancer. Pancreatic cancer, typically in the form of pancreatic ductal adenocarcinoma (PDAC), has a low 5-year survivability rate of only 5%, and a new treatment option is needed. Previous studies in our lab prompted the creation of twelve novel heterocyclic estrone analogs that were synthesized and evaluated on two pancreatic cancer cell lines AsPC-1 and Panc-1. Lead analog TR26 had an IC_{50} of 10.16 ± 0.83 μM in Panc-1 as compared to Gemcitabine standard with an IC_{50} above 50 micromolar. TR26 was also shown to arrest cells in the G0/G1 phase of the cell cycle by 12%. Additionally, TR26 was able to inhibit ERK by 25% at twice its IC_{50} value, and molecular dynamic studies performed suggested that key hydroxyl residues in its structure are important for binding. These heterocyclic analogs were studied further by members in our group and were found to be cytotoxic to non-small cell lung cancer cells (down to 15.7μM) as well as triple negative breast cancer cells (down to 3.18 μM). They were also additionally able to restore sensitivity of vincristine, an anticancer drug, to a resistant cell line. Finally, molecular dynamic simulations were performed to validate biological data; this showed TR26 binding to ERK complemented our findings with its 80% inhibition *in*-vitro.

These heterocyclic estrone analogs proved moderately effective against pancreatic cancer analogs *in-vitro*, and modification with C-11 oxygenation increases their effectiveness in another study from our laboratory. Using Bioisosterism we theorize that the installation of a nitrogenous group at this carbon will increase their effectiveness due

to their increased hydrogen bonding capabilities and similar atomic properties. Molecular modeling was performed with 650+ analogs created and docked within various EGFR pathway proteins; in specific those of the ERK/MAPK pathway and STAT3 pathway. Various C-11 amine analogs showed promising binding activities *in-silico* as opposed to known pancreatic cancer standard Gemcitabine, as well as to previously synthesized analogs tested against pancreatic cancer (2-5 fold better). This prompted their synthetic creation. Optimization of the C-11 ketone intermediate was first performed to limit the use of harsh chemicals which cause lower yields, such as hydroboration agents, as well as those that produce multiple by-products. Many iterations were then performed to obtain the C-11 amine from the C-11 ketone, and eventually it was obtained through installation of a benzylamine intermediate followed by hydrogenation. We ultimately propose a full-scale synthesis for C-11 amino analogs with heterocyclic side chains as new lead analogs to eventually be tested in pancreatic cancer.

Chapter One

General Introduction

1.1 Drug Discovery and Drug Development

Drug discovery is an arduous and time-consuming process that requires on average 10-12 years to complete for a single new drug, with costs upward of $1-3 billion (Figure 1.1). The procedure requires the use of cheminformatics, drug modeling, chemical synthesis, toxicity screenings, clinical trials, and much more[1]. This process typically begins with identifying the target disease followed by researching the molecular targets of the disease. There are a multitude of ways to approach finding a way to treat a disease, but often this arises in the form of small molecules, a species that due to its low molecular weight can enter a cell and perform a function[2]. Usually, tens of thousands of molecules begin the process to act as a potential treatment option for a disease. These initial molecules can be of various types, including current drugs to be repurposed for a different effect[3,4], molecules that have similar cores with slight modifications, and completely new molecules altogether.

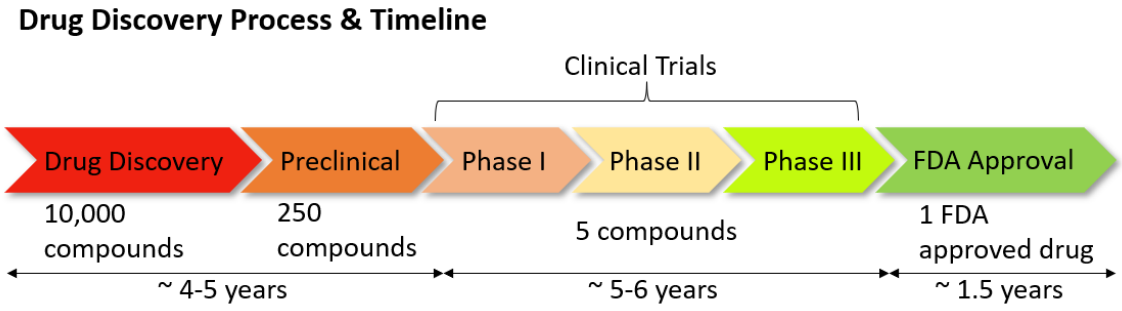


Figure 1.1 Timeline and process for Drug Discovery to FDA approval of a drug.

There exist numerous ways to sort out the large list of molecules into those that will be bioactive against a disease. Typically this is in the form of high throughput screening (HTS) or bioinformatics and *in-silico* modeling[5,6]. HTS is a technique in which a large number of drugs are screened biologically in an automated process to determine if they will be useful for further study[7,8]. In most cases a large number of compounds will first be processed in multi-well plates, upward of 1586 wells per plate, to allow for faster screening times. A primary screen can be done that will elicit information about potential hits, wherein those hits will further be evaluated for screening to determine their mechanism of action. This process allows for detection of compounds that will be most beneficial for a disease, or in some cases what pharmacophores of those compounds, similar functional groups within a molecule, will be beneficial for biological activity.

Upon identification of a lead compound that shows activity, the process usually moves into lead-to-hit optimization[9]. This is often done with *in-silico* computer modeling or dynamics which can help predict the effectiveness of supposed drugs without having to put them inside of a living organism. There are often two ways to perform molecular modeling: using a structure-based approach or a ligand based approach. Structure based approaches rely on the disease target to be known, typically a macromolecule such as DNA or a protein. If the molecular characteristics of the macromolecule are known, then a drug can be designed from these characteristics. This often means making a drug that will have a specific structure to fit into the correct dimensions of the macromolecule. This process is usually better for lead optimization, but often is much harder as the target structure must be known and the creation of the drug is often more tedious[10,11]. Ligand based design is typically much more common due to the ease at which it can be completed. As compounds

are found to be bioactive, as in HTS, certain functional groups are taken from them and modified in order to increase the effect on the disease[12]. This modification can be performed multiple times over until a finalized lead compound is developed. This is also sometimes termed as Quantitative Structure Activity Relationships (QSAR); the idea that certain functional groups will behave the same in the same system.

After discovering lead compound(s) they can either be purchased (if applicable) or synthesized directly followed by identification using spectroscopic techniques such as mass spectroscopy (MS) and nuclear magnetic resonance (NMR) followed by testing *in-vitro* on the disease. A highlight of this testing process includes identification of key features that cannot be predicted with computer simulation, including mechanism of action, metabolic studies, and binding energies. Successful *in-vitro* studies eventually lead to *in-vivo* studies where further metabolic properties and cytotoxic effects can be measured within a biological system before leading to clinical trials, and eventually possible FDA approval.

Worth noting is that researchers who create drugs have come across many rules that are typical of most drugs. These rules often follow under the ADME properties (Adsorption, Distribution, Metabolism, Excretion). The most famous of which is the Lipinski rule of five in which a drug should have: no more the five hydrogen bond donors, no more than ten hydrogen bond acceptors, a molecular mass less than 500 Daltons, and an octanol-water partition coefficient (LogP) less than five. These features likely help to ensure the drug will be properly absorbed into the body, localize to its specific site, metabolize in a time-efficient manner, and be excreted so as to not induce toxicity.

1.2 Natural Products and their Role in Drug Discovery

Some of the most common therapeutics are those made from or derived from nature and used as drugs[13,14,15]. The natural world provides plants and animals all they need to overcome life's hardships, so it seems only fitting science utilizes these therapeutics for human benefit. A review by Newman et al. (2020) showed that over the last four decades of the almost 1900 drugs approved, 23.5% were of natural origin, either being natural products themselves, or a derivative of a natural product. A further 11% were also natural product inspired drugs[15].

1.3 Importance of Cucurbitacins as a potential Therapeutic

Cucurbitacins are tetracyclic triterpenes found in some members of the family *Cucurbitaceae*, such as pumpkins, gourds, cucumbers, and melons[16]. Eighteen cucurbitacins have been identified and are designated from A-T. The Cucurbitacins themselves can be found in many areas of the plant, including the roots or leaves, but are often found more in the fruit and seeds of the plants[17]. The taste of cucurbitacin is quite bitter, often seen as a taste deterrent in the animal kingdom, owing to its toxic properties. However, the compounds themselves have many anti-inflammatory and oncological properties.

1.3.1 Chemistry of Cucurbitacins

Cucurbitacins (Figure 1.2) follow the general steroid structure with three fused cyclohexane rings (A, B, and C) and one fused cyclopentane ring (ring D). Some major differences between cucurbitacins and common steroids include the geminal dimethyl group at C-4, the methyl group at C-9, a ketone at C-11, and an alcohol at C-16. Additionally, there exists variation within the cucurbitacin groups as well. Most of the

cucurbitacins contain an α-hydroxy ketone at C-2/C-3 with a select few having either a diol (Cuc's F, O, P, and Q) or a diosphenol moiety (Cuc's E, J, and L). The other main variation exists in the C-17 side chain of the molecule. Normally, there exists an α,β-unsaturated ketone, but this unsaturation is lost in Cuc's H, L, P, and R. Additionally, a handful of cucurbitacins (A, B, C, E, Q) have an acetylated alcohol at C-25. Several cucurbitacins can also be glycosylated in nature as well, such as cucurbitacin E glycoside. Typically, this occurs at the C-2 hydroxyl group[18].

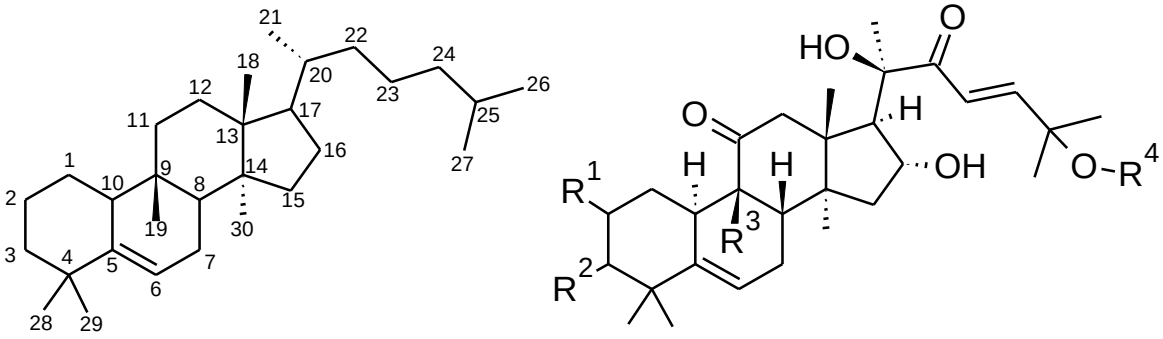


Figure 1.2 Cucurbitacin numbering scheme (left) and general cucurbitacin structure (right).

1.3.2 Pharmacological Activity of Cucurbitacins

Cucurbitacins have a plethora of pharmacological properties which have been investigated over the last few decades[18]. Their highly oxygenated and unsaturated systems provide them to be good anti-inflammatory molecules which were used as traditional medicines. Cucurbitacin R was found to inhibit nitric oxide creation by nitric oxide synthases (NOS), a key promotor in inflammation response[19]. Park et al (2004) describe the generation of nitric oxide primarily found from iNOS (inducible nitric oxide synthase) or its upstream activator nuclear factor κB (NF-κB). Cucurbitacin R in increasing doses

was found to lower nitrate concentrations *in-vitro* as well as decrease iNOS and NF-κB protein expression at 80 μM[19].

Cucurbitacin B has also been shown to have anti-inflammatory activity. In two studies it was found to prevent exudate leakage in cells by 50% and 70% at nontoxic levels[20, 21]. Exudate is commonly known as any fluid that is filtered through the circulatory system that exists from a wound and acts as a visual cue for inflammation. Likewise, Peters et al. (1999) also describes the ability for cucurbitacin B to reduce the amount of Prostaglandin E_2 in exudate by 40.7%. Prostaglandins are key mediators in inflammation response and are typically overexpressed in inflamed areas of the body. Other key players in inflammation include cyclooxygenase enzymes I and II (COX-1 or 2). Cucurbitacin E has been shown to modulate the activity of both COX-1 and COX-2 as well as decrease nitric oxide concentration at non-cytotoxic levels, similar to Cucurbitacin R[22].

Cucurbitacin D shows bioactivity against multiple targets. In a study by Hall et al. (2015) Cucurbitacin D prevented the ability of heat shock proteins to form complexes needed for chaperoning in oncogenic studies[84]. Additionally, in hepatocellular carcinoma, Cucurbitacin D showed ability to be hepatoprotective against carbon tetrachloride toxicity[85,86]. Moreover, Cucurbitacin D has also shown promise targeting the mitogen activated protein kinase pathway[87] that we will discuss in detail below. As a whole, cucurbitacins seem poised to be potential therapeutics for several diseases.

1.3.3 Cucurbitacin Anti-proliferation Activity in Cancer

Recently, cucurbitacins have been increasingly utilized as anticancer species in many cancers including breast, lung, liver, leukemia, ovarian, colorectal, and pancreatic cancers[23-27]. Cucurbitacins have widely been used as potential anticancer therapies and

have been explored in a large number of studies. From a review by Hussain et al (2019) Cucurbitacins B, D, E, I, F, O, P, Q, and R all show anticancer effects in various cell lines[57]. Primarily antiproliferation activity of cucurbitacins is from cell cycle arrest, induced apoptosis, or by inhibition of cell migration. They have also been demonstrated to interfere with signaling cascades involved in tumor growth, such as the Epidermal Growth Factor Receptor (EGFR) pathway[18]. EGFR is a receptor tyrosine kinase responsible for promoting cell proliferation and anti-apoptotic effects, and as such is considered a proto-oncogene. The protein itself exists as a monomer in its unbound form and has two natural ligands: Epidermal growth factor (EGF) and transforming growth factor alpha (TFG-α). Upon binding (Figure 1.3), EGFR dimerizes with autophosphorylation on specific tyrosine residues. This phosphorylation leads to downstream signaling of a variety of pathways[28,29]. Anticancer activity towards EGFR usually results in prevention of EGF binding or in inhibition of the downstream signaling proteins.

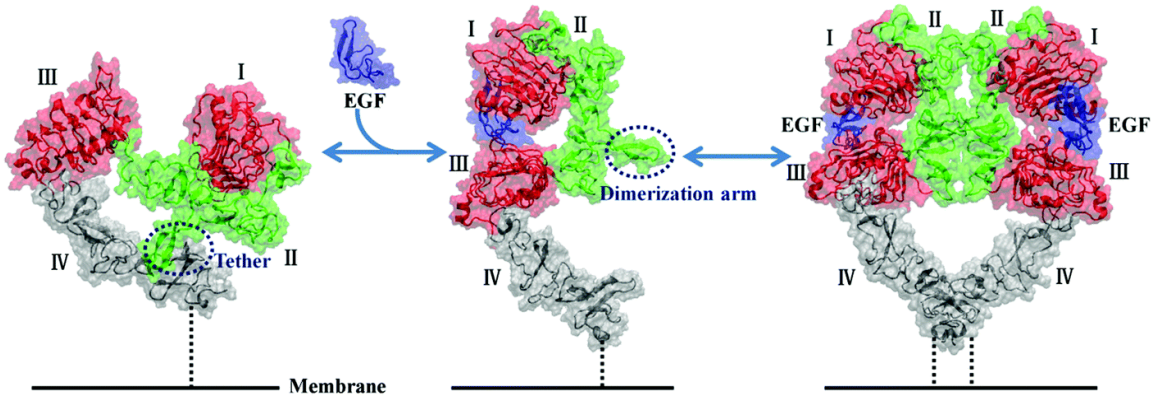


Figure 1.3 EGFR dimerization occurs upon binding of natural ligand EGF. The tether arm is untethered and becomes the eventual spot for dimerization, leading to autophosphorylation of key tyrosine residues in the cytoplasmic domain (not pictured). Reprinted from Shao et al (2019)[30].

EGFR inhibition has effects on three primary signaling cascades (Figure 1.4). The first and most targeted of which is the Raf/Ras/MEK/ERK pathway. This pathway is responsible for gene expression, cell cycle regulation, and anti-apoptotic measures, and exists as a common cancer therapeutic target[31]. Another pathway is that of Janus Kinases (JAK's) and Signal Transducer and Activator of Transcription (STAT) proteins. Specifically, the Jak/STAT3 pathway is one responsible for tumor progression, metastasis, and cell proliferation[32]. Cucurbitacin Q has been shown to act as a selective inhibitor of STAT3 and can also suppress tumor growth with single digit micromolar activities[88]. The final pathway is the PI3K/Akt/mTOR pathway, responsible for cell cycle regulation[33]. It has previously been shown that Cucurbitacin D inhibits the phosphorylation of PI3K, Akt, and mTOR in prostate cancer[26] as well as STAT3 phosphorylation in cervical cancer[25]. Additionally, Cucurbitacin B was able in ovarian cancer to inhibit phosphorylation of STAT3 and ERK[24], suggesting that Cucurbitacins play a role in inhibiting all three pathways, and that each of these pathways are intrinsically related signaling wise. It's also worth noting that the α,β-unsaturated group at C-23 and C-24 is thought to play a role in cytotoxicity of the cucurbitacin molecules as a potent Michael acceptor species[34].

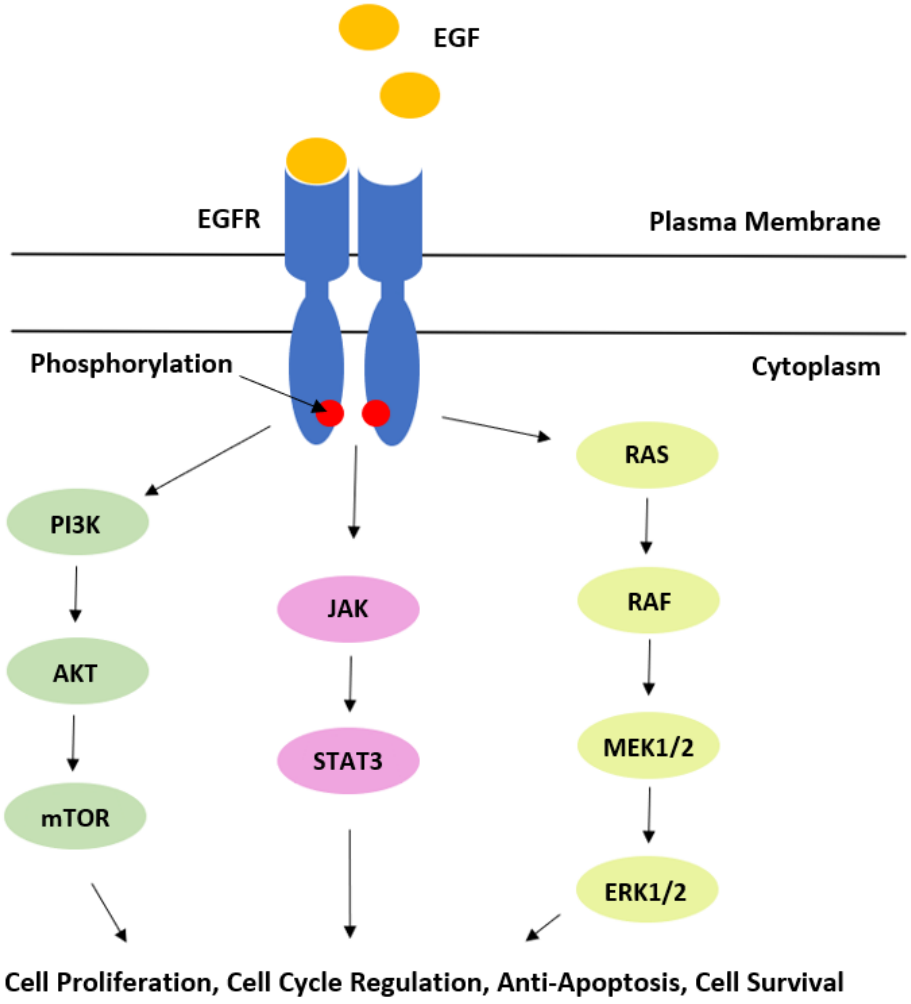


Figure 1.4 Epidermal Growth Factor Receptor Cascade. Binding of natural ligand EGF causes dimerization leading to autophosphorylation and downstream signaling to three primary pathways responsible for cancer progression.

1.3.4 Mechanism of Action for Anti-cancer Activity

1.3.4.1 Epidermal Growth Factor Receptor (EGFR) and Cancer

EGFR (ErbB1) is part of a family of four members of the ErbB family of receptors including ErbB1, ErbB2, ErbB3, and ErbB4. Each is a tyrosine kinase receptor responsible for downstream signaling pathways that are a common focus in cancer studies. EGFR has

found to be overexpressed in a number of different cancers, including head and neck, ovarian, cervical, bladder, esophageal, breast, colorectal, gastric, endometrial, and non-small cell lung cancer[35], the former half of which likely play a key role in patient outlook. Overactivation of the EGFR can result in uncontrolled cell division, a key feature of cancer, and inhibition of EGFR remains a powerful target for cancer therapeutics.

Currently there exists a number of therapies for cancer using EGFR inhibitors. These inhibitors are also known as tyrosine kinase inhibitors (TKI's). First round EGFR TKI's Erlotinib (Tarceva) and Geftinib (Iressa) are quinazoline based inhibitors. These inhibitors bind to the ATP binding site of EGFR and prevent dimerization and downfield signaling[36]. Primarily, they are used for Non-Small Cell Lung Cancer (NSCLC), breast, or pancreatic cancer. Often however, these first round TKI's have led to resistance. The most common resistance, T790M mutation, results in increased binding of ATP to the active site of EGFR due to a conformation change in the pocket and is found in over half of NSCLC patients[37]. Other minor types of resistance can come from MET amplification, or overactivation of ErbB2/ErbB3[38]. Another fairly common occurrence is that of mutations in exons 18-21 of the EGFR intracellular domain. Primarily these are exon 19 deletions or exon 21-point mutations which can lead to less effectiveness against first generation TKI's[39].

Second and third generation EGFR TKI's have been and are continually under development for treatment of cancer in which first generation EGFR TKI's become resistant. Most second-generation inhibitors often target the inhibition of multiple proteins in the ErbB family in hopes of a synergistic inhibition effect. Second generation TKI afatinib is a dual inhibitor of EGFR and ErbB2[40] with successful treatment in NSCLC against an L858R mutation, but with the T790M mutation afatinib shows less

improvement[41]. An additional second TKI, dacomitinib, binds irreversibly to EGFR, ErbB2, and ErbB4[42] as well as showing activity in NSCLC against the T790M in clinical trials[43]. A third-round inhibitor specifically targeting the T790M mutation is Osimertinib. It is also active against exon 19 deletions and is approved for patients with blood tests indicating either mutation. Even so, there have been increasing studies indicating resistance to third generation TKI's including Osimertinib, in which the C797S mutation is prevalent[44]. As such fourth generation TKI's are currently in development.

Mutations are often the crux of chemotherapy resistance. The high variability from patient-to-patient results in less useful chemotherapies over time and the future of chemotherapy will have to combat this. One of the likeliest methods to combat this in future studies is genetic testing. Predisposition to a certain disease or certain mutations leading to different biological actions are difficult to identify in current literature, and as genetic testing advancements occur this will likely promote more personalized therapies for individuals. Other key developments will likely come in personalized doses. With additional clinical testing using a wider array of patients we can discover how much of a chemotherapeutic is needed depending on the extent of the disease. These personalized therapies could likely become the future of chemotherapy regiments.

1.3.4.2 Rat Sarcoma (RAS) and Cancer

Rat Sarcoma Virus (RAS) was a gene discovered in the 1960's from rats due to sarcomas developed in newborn rats. The first two discovered RAS genes, Harvey sarcoma and Kirsten sarcoma coined the name[45]. RAS proteins are responsible for signal transduction by way of their GTPase domain, which binds guanosine nucleotides responsible for activation of the gene and leads to downstream activators. The three most

well-known oncogenes from RAS are KRAS, HRAS, and NRAS due to their prevalence in cancer. Typically, mutations in these RAS genes result in preferred GTP binding (activation) and overactivation of downstream signaling proteins[46].

Chemotherapies in this area are sparse in recent literature. Cetuximab, a monoclonal antibody given primarily for colorectal cancer and head and neck cancer, is an EGFR inhibitor that primarily acts on KRAS wild type genes[89]. More recently, in a study by Hong et al. (2020), the first KRAS mutant chemotherapy was studied in NSCLC[90]. It was found to inhibit KRAS mutants with the G12C mutation and was approved by the FDA for use in NSCLC in 2021. Ultimately, the more we learn about these specific mutants and their prevalence the more targeted therapy can begin. Likely this will be in the form of targeting specific binding sites of mutated proteins to develop selective inhibitors.

1.3.4.3 Rapidly Accelerated Fibrosarcoma (RAF) and Cancer

Activation of the GTPase RAS results in phosphorylation of Rapidly Accelerated Fibrosarcoma (RAF) and subsequent downstream signaling proteins. There are three members of this protein family, termed A-RAF, B-RAF, and C-RAF. The most well-known, due to its oncogenic behavior, is that of B-RAF. Over 30 mutants exist of B-RAF, and at least 18 of these result in increased kinase activity[47] leading to increased cell signaling responsible for cancer development and tumorigenesis. A number of B-RAF inhibitors do exist for B-RAF mutations, namely the V600E mutation[48]. Of these, sorafenib is one commonly used for treating advanced hepatocellular carcinoma and acts as an inhibitor of the Ras/Raf/Mek/ERK pathway.

1.3.4.4 Mitogen Activated Protein Kinases (MAPK's) and Cancer

Mitogen Activated Protein Kinases (MAPK's) are a family of proteins responsible for activation of a number of pathways that regulate cell proliferation, mitosis, apoptosis, and cell survival. They are activated by a number of factors including mitogens, heat shock proteins, and cytokines, as well as upstream proteins RAS and RAF. MEK1 and MEK2 (MAP2K1 and MAP2K2) are dual specificity MAPK's which act as phosphorylating proteins for the traditional set of MAPK's known as Extracellular signal-related kinases (ERK's). MEK1/2 activation from upstream kinases activates ERK1 and ERK2 which activate transcription factors in the cytoplasm of cells to control cell transcription, differentiation, apoptosis, and gene expression[49,50].

1.3.4.5 Signal Transducer and Activator of Transcription 3 (STAT3) and Cancer

Signal Transducers and Activators of Transcription (STAT) are a family of intracellular proteins responsible for mediating cellular responses from other cytokines and growth factors, particularly Janus Kinases (JAK), which act as signaling proteins for STAT family members. There are 7 STAT family members which cover a range of biological functions including cell differentiation, proliferation, and apoptosis[51]. Of these 7 members, STAT3 is the most widely known for its role as a potential oncogene. Upregulation of upstream proteins typically causes overactivation of STAT3. This increase in STAT3 activation can induce transformation[52] and interaction with a number of other proteins, such as chaperone proteins, nuclear translocation proteins, and growth factors[53]. As such the JAK/STAT pathway has remained as a target for potential treatment options in cancer.

1.3.4.6 Mechanistic Target of Rapamycin (mTOR) and Cancer

The Mechanistic Target of Rapamycin (mTOR) is a serine/threonine kinase from the family of Phosphatidylinositol-3-kinases. mTOR itself comprises a subunit of two larger complexes, mTOR complex 1 and mTOR complex 2, sometimes termed as mTORC1/2[54]. Complex 1's function is primarily that of regulation of protein translation for newly dividing cells[55] and is regulated by rapamycin, mechanical stimuli, oxidative stress, but is not frequently involved with growth factors. mTORC2 however, is responsive to growth factors, including its upstream regulator Akt, which regulates cell survival and is often overexpressed in many cancers[56].

1.3.5 Cancer Therapeutics and Targeted Therapies

As we've mentioned, each of these downstream proteins, as well as EGFR itself, provide opportunities for targeted inhibition, typically associated with specific patients and their cancer. These targeted therapies are becoming increasingly more common as new technology is being developed to determine gene expressions and to critically evaluate biological targets. Future development of chemotherapies likely lies in using a specific treatment option for each individual patient depending on their case. This could involve treatment at the site of the cancer, limiting side effects of the treatment, and prolonging patient life.

1.4 Structure Activity Relationships in Drug Development

Designing a molecule for bioactive intentions, i.e., binding to a protein, is often a lengthy and arduous process. Computer Aided Drug Design (CADD) is sometimes useful for *in-silico* methods in which screening of molecules through computer programs, such as molecular docking or molecular dynamics, can elicit important information about the

interaction between small molecule and protein. These two techniques will be discussed in detail in the following chapters, but another method used involves the idea of structure activity relationships (SAR), which were briefly aforementioned. When a small molecule effectively elicits a response biologically that is wanted, it becomes reasonable that similar molecules in design will yield the same response. Often times, large libraries are generated in hopes of finding a compound with increased bioactivity. A common example of SAR is switching a sulfur atom for an oxygen atom or vice versa. The two possess similar chemical profiles in their two lone pairs and p^4 orbital notation, with minor differences in size, electronegativity, and other properties. Aromatic groups are also often switched in hopes of creating new bonding patterns or altering the size of the molecule to fit in the binding pocket. These methods of switching out similar atoms are known as Bioisosterism. The result is typically a slight change in toxicity or bioavailability or altering the lifetime of the compound.

A further design option for small molecules exists in the form of fragment-based design. This can either be ligand-based or structure-based but relies on the knowledge that certain fragments of a molecule are responsible for bioactivity of the drug. The fragments that are needed for bioactivity are kept, and the connecting structures are altered for one or more reasons. One of these reasons is to increase or decrease rigidity of the compound. This can be done to fit the pocket of the target more uniformly or can allow for flexibility in the compound. Often this is done by changing aliphatic chains to rings or vice versa. Another option is to create new fragments that can be attached to existing fragments in hope of increasing binding to the target, and therefore increasing bioactivity. Often times the ability to identify these fragments, also known as chromophores, leads to the creation of new

molecules that are easier to synthesize when the original compound is either time-consuming to create or economically unpractical. These SAR ultimately give us the ability to create new lead compounds that we can create.

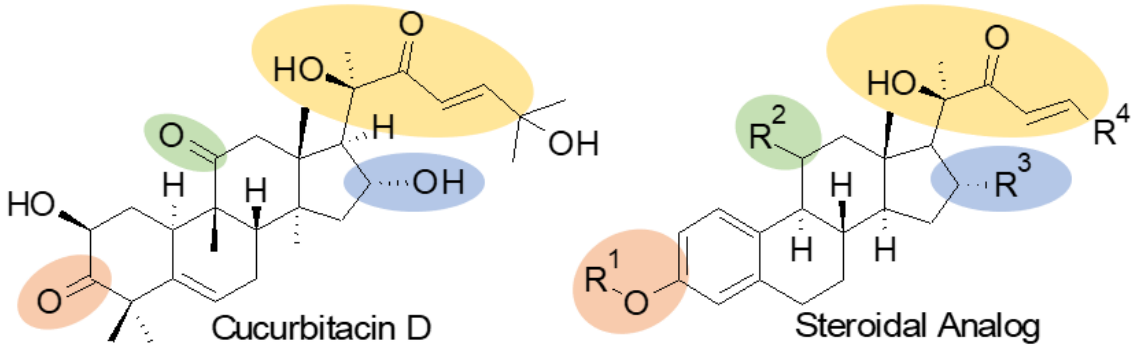
1.4.1 Cucurbitacins and Steroidal Derivatives for Cancer Therapy

Cucurbitacins, highly oxygenated triterpenes of the cucurbitane structure, have been utilized in medicine since the 1960's, and have become prominent for their anticancer activity[58], Cucurbitacins B and E are considered precursors for all other cucurbitacins, differing only in their saturation between C1-C2, with Cuc E containing the diosphenol moiety. Their biological activity in cancer is attributed to the α,β-unsaturated ketone group from C22-C24[59], with cucurbitacins B, D, E, and I demonstrating potent anticancer activity in breast, liver, cervical, ovarian, prostate, pancreatic, leukemia, and a number of other cell lines[24-27, 60-63]. The major problem with cucurbitacins as anticancer however is their availability and difficulty to synthesize due to the complexity of the structure. Cucurbitacins have biosynthetically been created, but a full-scale chemical synthesis of cucurbitacins has not been achieved to date. Couple that with the quantity of naturally produced cucurbitacins being quite low in most plants, the cucurbitacins themselves are difficult and expensive to obtain in a pure form. We mentioned previously the ability to copy chromophores from active structures, and the following discussion highlights the creation of new analogs in which bioactive chromophores from cucurbitacins are taken and used in a steroidal carrier to create novel compounds.

1.4.2 Cucurbitacin Inspired Estrone Analogs

To overcome the aforementioned pitfalls, our laboratory group has used structure activity relationships from the cucurbitacin core structure (Figure 1.5) and designed various

analogs using the tetracyclic steroid Estrone as a cheap and versatile core structure. C-3 of the structure is previously oxygenated in Estrone, and has been functionalized with varying electron donating and withdrawing groups to modulate the polarity of the molecule. Cucurbitacins are not aromatic in ring A, but this added aromaticity allows for easier installation of functional groups along ring A. C2 and C4, the ortho positions to the phenol moiety allows for installation using aromatic substitution reactions, and is currently being studied. The key feature of cucurbitacins, the C22-C24 α,β-unsaturated ketone, was installed in a 6 step process, with various R^4 groups being added using different aldol condensation reactions with differing aldehydes[64]. The result was a comparable structure with biological activity similar to that of cucurbitacins.



R^1	R^2	R^3
-H	-OH	-OH
$-CH_3$	=O	=O
$-SO_2NH_2$	-N-R	-Br

Figure 1.5 Depiction of the core components of the cucurbitacin structure that have been kept and modified in novel synthetic steroidal analogs used in anticancer studies in the Halaweish lab. R^4 is often aromatic in nature, but analogs do exist with the 2,2-dimethyl-2-hydroxy core.

Promising results from our first study successfully installing the C-17 side chain prompted us to explore functionalize in other parts of the steroid core, primarily in rings A, C, and D, and at the enone side chain. The enone side chain itself is a potent Michael acceptor, and can be metabolized by the body to a species with decreased activity. Studies were performed to attach asides or thiophenols to C-24 of the enone side chain in a Michael reaction, and results indicated low micromolar activity in melanoma cells[65,66]. Ring D modification, typically carbon 16-modification, was explored in two studies. Unsaturation of the C-16-C-17 bond results in a slight conformation change of the steroid structure, which performs well in hepatocellular carcinoma[67]. Another study was performed to mimic the C-16 functionality of cucurbitacins by implementing various oxygenated derivatives[68]. Biological screening was performed in MCF-7 breast cancer cells and found low micromolar/high nanomolar activities for various analogs. Most recently we have begun functionalizing rings A and C of the steroid structure. Ring A functionalization is currently being performed at C-2 and C-4 with fluorine functionality, while ring C has recently been functionalized with oxygenated derivatives at C-11 to mimic that of cucurbitacins[69]. The latter details low micromolar activity in pancreatic cancer, and will be discussed in detail in chapter 4. Ultimately, this work describes the further functionalization of the steroidal core. Our goal is to continue to functionalize the steroid to increase cytotoxicity towards cancer.

1.5 Drug Resistance in Cancer

Resistance to chemotherapies from cancer cells is one of the main complications that arise in patient treatment. Typically this results in an eventual reoccurrence of the disease and mortality of the patient. There are two main types of resistance: intrinsic resistance, or

those developed prior to treatment, and acquired resistance, which is acquired through or after treatment. Those with intrinsic resistance to treatment are typically those with weak immune responses in general, the elderly for example, or those that fail to elicit a T-cell response[70]. Other mechanisms could be due to altered expression or function of a target based on genetic factors or drug-drug interactions[71,72]. Acquired resistance on the other hand arises from a number of factors including mutations of the target active site, changes in tumor environment, tumor heterogeneity, drug inactivation, or drug efflux, among others[73]. Of the mentioned types of acquired drug resistance, this work will focus on drug efflux.

1.5.1 ATP-Binding Cassette Transporters as a target

Drug efflux is a result of membrane-bound pumps that efflux anticancer material (and other substrates) from cells resulting in decreased accumulation of drug in the cell, and overall lower efficacy. The majority of this research is focused on the ATP-Binding Cassette (ABC) transporter family. ATP importers are primarily found in prokaryotes, while exporters are found throughout the eukaryotic kingdom. This superfamily of transporters typically has one to two transmembrane domains and an ATP binding domain intracellularly that provides the energy needed for translocation of material across the cell membrane[74]. Upon ATP binding to the nucleotide binding domain, the two domains close together, altering the transmembrane domains to allow for access of the substrate to be exported from the cell. Substrates for most ABC transporters typically include sugars, steroids, amino acids, phospholipids, peptides, proteins, antibiotics, and xenobiotics[75].

To date there have been 48 ABC transporter genes encoded in the human genome[76], categorized into seven subfamilies ABCA-ABCG. Over half of these transport lipid

materials, and mutations in them cause 12 diseases in humans, highlighting their important function in the body[77]. The most studied ABC transporter, and first to be discovered, is ABCB1, also termed P-Glycoprotein (P-gP). It is composed of two nucleotide binding domains and two transmembrane domains. P-gP is highly expressed in cancerous tissue, but also in normal tissue, including on the surface of hepatocytes in the liver and renal tubules in the kidneys, as well as endothelium cells in the brain and intestinal tract[78]. Substrates of P-gP range from small (250 Da) to relatively large (1200 Da) small molecules, but most are lipophilic in nature, ranging from lipids to steroids[78]. Included in this list are a number of chemotherapeutics, including paclitaxel, vinblastine, doxorubicin, and etoposide[72], suggesting that P-gP plays a major role in drug resistance studies.

Owing to its name, multidrug resistance protein (MRP) is another family of ABC transporter (ABCC) which was discovered a few years after P-gP as a set of cells that were resistant to a number of drugs and subsequently termed multidrug resistance gene 1 (MDR1)[79]. The family consists of 13 proteins discovered (ABCC1-ABCC13), and 9 of which have been classified of MRP's colloquially termed MRP1-MRP9, the former most of which has been the most talked about in recent literature[80]. MRP's are structurally similar to P-gP including their transmembrane domains and nucleotide binding domains, but the number of each can vary, particularly the transmembrane domains. A key part of MRP's is that they export a number of important endogenous species, as well as xenobiotics. We mentioned that MRP1 is the most studied of the ABCC family, partially due to it being discovered first, but has since been found in a number of cancers including lung, breast, prostate, and neuroblastoma[81]. MRP1, despite its similarity to P-gP, extrudes a wide range of materials from cells, some of which overlap with substrates of P-gP.

Primarily the substrates of MRP1 are large hydrophobic molecules that diffuse through the cell membrane[81]. Increased drug resistance is thus associated with MRP1, and has become a target in recent studies for drug resistance.

The final transporter of interest is that of ABCG2, also termed Breast cancer resistance protein (BCRP). ABCG2 was discovered by Doyle et al (1998) after noticing two breast cancer lines had developed resistance to a number of chemotherapeutic drugs. They tested for P-gP and MRP1, but found that an additional resistance method was likely at play, and was so termed the Breast cancer resistance protein[82]. The structure of ABCG2 is similar of that to P-gP and MRP1, with transmembrane domains and nucleotide binding domains responsible for binding of ATP, which are conserved throughout most of the family. ABCG2 expression is found minimally in normal tissue within the brain, prostate, GI tract, ovaries, and liver[83], and is likely to play a role in cancer resistance similar to its two predecessors.

Resistance to chemotherapies highlights this work, as we explore the ability for steroidal drugs created in our laboratory to act on the three previously mentioned ABC transporters. We hypothesize that these are responsible for drug resistance of a number of chemotherapies, and inhibition of these transporters may play a key role in a new therapeutic option.

1.6 Project Objectives

Previous studies in our laboratory have focused on new synthetic pathways for creating analogs for various cancers. This project's focus is broader, and aims at optimizing different steroidal analogs in order to improve their efficacy and determine if they could overcome cancer resistance. The objectives include areas of *in-silico* modeling work,

molecular dynamic calculations, synthetic optimization and compound identification, biological screening in a variety of different cancers, and target identification through biological assays. Thus, the main objectives of this project are as follows:

- Biologically screen a set of novel triazole-estradiol analogs in ovarian and colorectal cancer targeting EGFR proteins and ABC transporters responsible for drug resistance.
- Optimize triazole-estradiol analogs using molecular docking and synthesize new promising analogs for testing in triple negative breast cancer.
- Synthesize and evaluate biologically a set of heterocyclic estrone analogs in pancreatic cancer, as well as highlight their use in other cancers.
- Optimize heterocyclic estrone analogs using *in-silico* modeling for pancreatic cancer analog development.
- Perform synthetic optimization of a set of novel C-11 nitrogenous estrone analogs targeting pancreatic cancer.

1.7 References

1. DiMasi, J. A.; Grabowski, H. G.; Hansen, R. W., Innovation in the pharmaceutical industry: New estimates of R&D costs. *J Health Econ* **2016,** *47*, 20-33.

2. National Cancer Institute. National Insitutes of Health definition of small molecule. https://www.cancer.gov/publications/dictionaries/cancer-terms/def/small-molecule-drug (accessed 12/20/2021).

3. Sleire, L.; Forde, H. E.; Netland, I. A.; Leiss, L.; Skeie, B. S.; Enger, P. O., Drug repurposing in cancer. *Pharmacol Res* **2017,** *124*, 74-91.

4. Cha, Y.; Erez, T.; Reynolds, I. J.; Kumar, D.; Ross, J.; Koytiger, G.; Kusko, R.; Zeskind, B.; Risso, S.; Kagan, E.; Papapetropoulos, S.; Grossman, I.; Laifenfeld, D., Drug repurposing from the perspective of pharmaceutical companies. *Br J Pharmacol* **2018,** *175* (2), 168-180.

5. Mavromoustakos, T., Methodologies and applied strategies in the rational drug design. *Curr Med Chem* **2011,** *18* (17), 2516.

6. Mandal, S.; Moudgil, M.; Mandal, S. K., Rational drug design. *Eur J Pharmacol* **2009,** *625* (1-3), 90-100.

7. Mayr, L. M.; Bojanic, D., Novel trends in high-throughput screening. *Curr Opin Pharmacol* **2009,** *9* (5), 580-8.

8. Blay, V.; Tolani, B.; Ho, S. P.; Arkin, M. R., High-Throughput Screening: today's biochemical and cell-based approaches. *Drug Discov Today* **2020,** *25* (10), 1807-1821.

9. Keseru, G. M.; Makara, G. M., Hit discovery and hit-to-lead approaches. *Drug Discov Today* **2006,** *11* (15-16), 741-8.

10. Anderson, A. C., The process of structure-based drug design. *Chem Biol* **2003,** *10* (9), 787-97.

11. Sledz, P.; Caflisch, A., Protein structure-based drug design: from docking to molecular dynamics. *Curr Opin Struct Biol* **2018,** *48*, 93-102.

12. Acharya, C.; Coop, A.; Polli, J. E.; MackKerell Jr., A. D., Recent Advances in Ligand-Based Drug Design: Relevance and Utility of the Conformationally Sampled Pharmacophore Approach. *Curr Comput Aided Drug Des* **2011,** *7* (1), 10-22.

13. Atanasov, A. G.; Zotchev, S. B.; Dirsch, V. M.; International Natural Product Sciences, T.; Supuran, C. T., Natural products in drug discovery: advances and opportunities. *Nat Rev Drug Discov* **2021,** *20* (3), 200-216.

14. Thomford, N. E.; Senthebane, D. A.; Rowe, A.; Munro, D.; Seele, P.; Maroyi, A.; Dzobo, K., Natural Products for Drug Discovery in the 21st Century: Innovations for Novel Drug Discovery. *Int J Mol Sci* **2018,** *19* (6).

15. Newman, D. J.; Cragg, G. M., Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J Nat Prod* **2020,** *83* (3), 770-803.

16. American Chemical Society Molecule of the Week. Cucurbitacins. https://www.acs.org/content/acs/en/molecule-of-the-week/archive/c/cucurbitacins.html (accessed 12/29/21).

17. Matsuo, K.; DeMilo, A. B.; Schroder, R. F. W.; Martin, P. A. W., Rapid High-Performace Liquid Chromatography Method To Quantitate Elaterinide in Juice and Reconstituted REsidues from a Bitter Mutant of Hawkesbury Watermelon. *J Agric Food Chem* **1999,** *47*, 2755-2759.

18. Kaushik, U.; Aeri, V.; Mir, S. R., Cucurbitacins - An insight into medicinal leads from nature. *Pharmacogn Rev* **2015,** *9* (17), 12-18.

19. Park, C. S.; Lim, H.; Han, K. J.; Baek, S. H.; Sohn, H. O.; Lee, D. W.; Kim, Y. G.; Yun, H. Y.; Baek, K. J.; Kwon, N. S., Inhibition of nitric oxide generation by 23,24-dihydrocucurbitacin D in mouse peritoneal macrophages. *J Pharmacol Exp Ther* **2004,** *309* (2), 705-10.

20. Yesilada, E.; Tanaka, S.; Sezik, E.; Tabata, M., Isolation of an anti-inflammatory principle from the fruit juice of *Ecballium Elaternium*. *J Nat Prod* **1988,** *51* (3), 504-508.

21. Peters, R. R.; Saleh, T. F.; Lora, M.; Patry, C.; de Brum-Fernandes, A. J.; Farias, M. R.; Ribeiro-do-Valle, R. M., Anti-inflammatory effects of the products from *Wilbrandia ebracteata* on carrageenan-induced pleurisy in mice. *Life Sci* **1999,** *64* (26), 2429-2437.

22. Abdelwahab, S. I.; Hassan, L. E.; Sirat, H. M.; Yagi, S. M.; Koko, W. S.; Mohan, S.; Taha, M. M.; Ahmad, S.; Chuen, C. S.; Narrima, P.; Rais, M. M.; Hadi, A. H., Anti-inflammatory activities of cucurbitacin E isolated from Citrullus lanatus var. citroides: role of reactive nitrogen species and cyclooxygenase enzyme inhibition. *Fitoterapia* **2011,** *82* (8), 1190-7.

23. Hussain, H.; Green, I. R.; Saleem, M.; Khattak, K. F.; Irshad, M.; Ali, M., Cucurbitacins as Anticancer Agents: A Patent Review. *Recent Pat Anticancer Drug Discov* **2019,** *14* (2), 133-143.

24. El-Senduny, F. F.; Badria, F. A.; El-Waseef, A. M.; Chauhan, S. C.; Halaweish, F., Approach for chemosensitization of cisplatin-resistant ovarian cancer by cucurbitacin B. *Tumour Biol* **2016,** *37* (1), 685-98.

25. Sikander, M.; Hafeez, B. B.; Malik, S.; Alsayari, A.; Halaweish, F. T.; Yallapu, M. M.; Chauhan, S. C.; Jaggi, M., Cucurbitacin D exhibits potent anti-cancer activity in cervical cancer. *Sci Rep* **2016,** *6*, 36594.

26. Sikander, M.; Malik, S.; Chauhan, N.; Khan, P.; Kumari, S.; Kashyap, V. K.; Khan, S.; Ganju, A.; Halaweish, F. T.; Yallapu, M. M.; Jaggi, M.; Chauhan, S. C., Cucurbitacin D Reprograms Glucose Metabolic Network in Prostate Cancer. *Cancers (Basel)* **2019,** *11* (3).

27. Sikander, M.; Malik, S.; Khan, S.; Kumari, S.; Chauhan, N.; Khan, P.; Halaweish, F. T.; Chauhan, B.; Yallapu, M. M.; Jaggi, M.; Chauhan, S. C., Novel Mechanistic Insight into the Anticancer Activity of Cucurbitacin D against Pancreatic Cancer (Cuc D Attenuates Pancreatic Cancer). *Cells* **2019,** *9* (1).

28. Jorissen, R., Epidermal growth factor receptor: mechanisms of activation and signaling. *Experimental Cell Research* **2003,** *284* (1), 31-53.

29. Kim, E. S.; Khuri, F. R.; Herbst, R. S., Epidermal growth factor receptor biology. *Curr Opin Oncol* **2001,** *13*, 506-513.

30. Shao, Q.; Zhu, W. Ligand binding effects on the activation of the EGFR extracellular domain. *Phys. Chem. Chem. Phys.* **2019**, *21*, 8141-8151.

31. Chang, F.; Steelman, L. S.; Lee, J. T.; Shelton, J. G.; Navolanic, P. M.; Blalock, W. L.; Franklin, R. A.; McCubrey, J. A., Signal transduction mediated by the Ras/Raf/MEK/ERK pathway from cytokine receptors to transcription factors: potential targeting for therapeutic intervention. *Leukemia* **2003,** *17* (7), 1263-93.

32. Johnson, D. E.; O'Keefe, R. A.; Grandis, J. R., Targeting the IL-6/JAK/STAT3 signaling axis in cancer. *Nat Rev Clin Oncol* **2018,** *15* (4), 234-248.

33. King, D.; Yeomanson, D.; Bryant, H. E., PI3King the Lock: Targeting the PI3K/Akt/mTOR Pathway as a Novel Therapeutic Strategy in Neuroblastoma. *J Pediatr Hematol Oncol* **2015,** *37* (4), 245-251.

34. Amslinger, S., The tunable functionality of alpha,beta-unsaturated carbonyl compounds enables their differential application in biological systems. *ChemMedChem* **2010,** *5* (3), 351-6.

35. Nicholson, R. I.; Gee, J. M. W.; Harper, M. E., EGFR and cancer prognosis. *Eur J Cancer* **2001,** *37*, S9-S15.

36. Cataldo, V. D.; Gibbons, D. L.; Perez-Soler, R.; Quintas-Cardama, A., Treatment of Non-Small-Cell Lung Cancer with Erlotinib or Gefitinib. *N Engl J Med* **2011,** *364* (10), 947-955.

37. Nagano, T.; Tachihara, M.; Nishimura, Y., Mechanism of Resistance to Epidermal Growth Factor Receptor-Tyrosine Kinase Inhibitors and a Potential Treatment Strategy. *Cells* **2018,** *7* (11).

38. Hsu, J. L.; Hung, M. C., The role of HER2, EGFR, and other receptor tyrosine kinases in breast cancer. *Cancer Metastasis Rev* **2016,** *35* (4), 575-588.

39. Xu, C. W.; Lei, L.; Wang, W. X.; Lin, L.; Zhu, Y. C.; Wang, H.; Miao, L. Y.; Wang, L. P.; Zhuang, W.; Fang, M. Y.; Lv, T. F.; Song, Y., Molecular Characteristics and Clinical Outcomes of EGFR Exon 19 C-Helix Deletion in Non-Small Cell Lung Cancer and Response to EGFR TKIs. *Transl Oncol* **2020,** *13* (9), 100791.

40. Yu, H. A.; Riely, G. J., Second Generation Epidermal Growth Factor Receptor Tyrosine Kinase Inhibitors in Lung Cancer. *J Natl Compr Canc Netw* **2013,** *11* (2), 161-169.

41. Li, D.; Ambrogio, L.; Shimamura, T.; Kubo, S.; Takahashi, M.; Chirieac, L. R.; Padera, R. F.; Shapiro, G. I.; Baum, A.; Himmelsbach, F.; Rettig, W. J.; Meyerson, M.; Solca, F.; Greulich, H.; Wong, K. K., BIBW2992, an irreversible EGFR/HER2 inhibitor highly effective in preclinical lung cancer models. *Oncogene* **2008,** *27* (34), 4702-11.

42. Reckamp, K. L.; Giaccone, G.; Camidge, D. R.; Gadgeel, S. M.; Khuri, F. R.; Engelman, J. A.; Koczywas, M.; Rajan, A.; Campbell, A. K.; Gernhardt, D.; Ruiz-Garcia, A.; Letrent, S.; Liang, J.; Taylor, I.; O'Connell, J. P.; Janne, P. A., A phase 2 trial of dacomitinib (PF-00299804), an oral, irreversible pan-HER (human epidermal growth factor receptor) inhibitor, in patients with advanced non-small cell lung cancer after failure of prior chemotherapy and erlotinib. *Cancer* **2014,** *120* (8), 1145-54.

43. Engelman, J. A.; Zejnullahu, K.; Gale, C. M.; Lifshits, E.; Gonzales, A. J.; Shimamura, T.; Zhao, F.; Vincent, P. W.; Naumov, G. N.; Bradner, J. E.; Althaus, I. W.; Gandhi, L.; Shapiro, G. I.; Nelson, J. M.; Heymach, J. V.; Meyerson, M.; Wong, K. K.; Janne, P. A., PF00299804, an irreversible pan-ERBB inhibitor, is effective in lung cancer models with EGFR and ERBB2 mutations that are resistant to gefitinib. *Cancer Res* **2007,** *67* (24), 11924-32.

44. Patel, H.; Pawara, R.; Ansari, A.; Surana, S., Recent updates on third generation EGFR inhibitors and emergence of fourth generation EGFR inhibitors to combat C797S resistance. *Eur J Med Chem* **2017,** *142*, 32-47.

45. Malumbres, M.; M., B., RAS oncogenes: the first 30 years. *Nat Rev Cancer* **2003,** *3*, 7-13.

46. Prior, I. A.; Lewis, P. D.; Mattos, C., A comprehensive survey of Ras mutations in cancer. *Cancer Res* **2012,** *72* (10), 2457-67.

47. Wan, P. T. C.; Garnett, M. J.; Roe, S. M.; Lee, S.; Niculescu-Duvaz, D.; Good, V. M.; Jones, C. M.; Marshall, C. J.; Springer, C. J.; Barford, D.; Marais, R., Mechanism of Activation of the RAF-ERK Signaling Pathway by Oncogenic Mutations of B-RAF. *Cell* **2004,** *116*, 855-867.

48. Khan, P. S.; Rajesh, P.; Rajendra, P.; Chaskar, M. G.; Rohidas, A.; Jaiprakash, S., Recent advances in B-RAF inhibitors as anticancer agents. *Bioorg Chem* **2022,** *120*, 105597.

49. Wee, P.; Wang, Z., Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. *Cancers (Basel)* **2017,** *9* (5).

50. Guo, Y. J.; Pan, W. W.; Liu, S. B.; Shen, Z. F.; Xu, Y.; Hu, L. L., ERK/MAPK signalling pathway and tumorigenesis. *Exp Ther Med* **2020,** *19* (3), 1997-2007.

51. Lim, C. P.; Cao, X., Structure, function, and regulation of STAT proteins. *Mol Biosyst* **2006,** *2* (11), 536-50.

52. Turkson, J.; Jove, R., STAT proteins: novel molecular targets for cancer drug discovery. *Oncogene* **2000,** *19*, 6613-6626.

53. Laudisi, F.; Cherubini, F.; Monteleone, G.; Stolfi, C., STAT3 Interactors as Potential Therapeutic Targets for Cancer Treatment. *Int J Mol Sci* **2018,** *19* (6).

54. Wullschleger, S.; Loewith, R.; Hall, M. N., TOR signaling in growth and metabolism. *Cell* **2006,** *124* (3), 471-84.

55. Hay, N.; Sonenberg, N., Upstream and downstream of mTOR. *Genes Dev* **2004,** *18* (16), 1926-45.

56. Kim, L. C.; Cook, R. S.; Chen, J., mTORC1 and mTORC2 in cancer and the tumor microenvironment. *Oncogene* **2017,** *36* (16), 2191-2201.

57. Hussain, H.; Green, I. R.; Saleem, M.; Khattak, K. F.; Irshad, M.; Ali, M., Cucurbitacins as Anticancer Agents: A Patent Review. *Recent Pat Anticancer Drug Discov* **2019,** *14* (2), 133-143.

58. Chen, X.; Bao, J.; Guo, J.; Ding, Q.; Lu, J.; Huang, M.; Wang, Y., Biological activities and potential molecular targets of cucurbitacins: a focus on cancer. *Anticancer Drugs* **2012,** *23* (8), 777-87.

59. Chen, J. C.; Chiu, M. H.; Nie, R. L.; Cordell, G. A.; Qiu, S. X. J. N. p. r., Cucurbitacins and cucurbitane glycosides: structures and biological activities. *Nat Prod Rep* **2005**, *22* (3), 386-399.

60. Chan, K. T.; Li, K.; Liu, S. L.; Chu, K. H.; Toh, M.; Xie, W. D., Cucurbitacin B inhibits STAT3 and the Raf/MEK/ERK pathway in leukemia cell line K562. *Cancer Lett* **2010,** *289* (1), 46-52.

61. Ge, W.; Chen, X.; Han, F.; Liu, Z.; Wang, T.; Wang, M.; Chen, Y.; Ding, Y.; Zhang, Q., Synthesis of Cucurbitacin B Derivatives as Potential Anti-Hepatocellular Carcinoma Agents. *Molecules* **2018,** *23* (12).

62. Wakimoto, N.; Yin, D.; O'Kelly, J.; Haritunians, T.; Karlan, B.; Said, J.; Xing, H.; Koeffler, H. P., Cucurbitacin B has a potent antiproliferative effect on breast cancer cells in vitro and in vivo. *Cancer Sci* **2008,** *99* (9), 1793-7.

63. Zhang, T.; Li, J.; Dong, Y.; Zhai, D.; Lai, L.; Dai, F.; Deng, H.; Chen, Y.; Liu, M.; Yi, Z., Cucurbitacin E inhibits breast tumor metastasis by suppressing cell migration and invasion. *Breast Cancer Res Treat* **2012,** *135* (2), 445-58.

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

65. Ahmed, M. S.; El-Senduny, F.; Taylor, J.; Halaweish, F. T., Biological screening of cucurbitacin inspired estrone analogs targeting mitogen-activated protein kinase (MAPK) pathway. *Chem Biol Drug Des* **2017,** *90* (3), 478-484.

66. Kopel, L. C.; Ahmed, M. S.; Halaweish, F. T., Synthesis of novel estrone analogs by incorporation of thiophenols via conjugate addition to an enone side chain. *Steroids* **2013,** *78* (11), 1119-25.

67. Mahnashi, M.; Elgazwi, S. M.; Ahmed, M. S.; Halaweish, F. T., Cucurbitacins inspired organic synthesis: Potential dual inhibitors targeting EGFR - MAPK pathway. *Eur J Med Chem* **2019,** *173*, 294-304.

68. Alsayari, A.; Kopel, L.; Ahmed, M. S.; Pay, A.; Carlson, T.; Halaweish, F. T., Design, synthesis, and biological evaluation of steroidal analogs as estrogenic/anti-estrogenic agents. *Steroids* **2017,** *118*, 32-40.

69. Alseud, K, M. Design, Synthesis, and Biological Screening of Novel Cucs-Inspired Estrone Analogs Towards Treatment of Pancreatic Adenocarcinoma. Ph.D. Dissertation. South Dakota State University, Brookings, SD 2019.

70. Kelderman, S.; Schumacher, T. N.; Haanen, J. B., Acquired and intrinsic resistance in cancer immunotherapy. *Mol Oncol* **2014,** *8* (6), 1132-9.

71. Nikolaou, M.; Pavlopoulou, A.; Georgakilas, A. G.; Kyrodimos, E., The challenge of drug resistance in cancer treatment: a current overview. *Clin Exp Metastasis* **2018,** *35* (4), 309-318.

72. Wang, X.; Zhang, H.; Chen, X., Drug resistance and combating drug resistance in cancer. *Cancer Drug Resist* **2019,** *2,* 141-160.

73. Housman, G.; Byler, S.; Heerboth, S.; Lapinska, K.; Longacre, M.; Snyder, N.; Sarkar, S., Drug resistance in cancer: an overview. *Cancers (Basel)* **2014,** *6* (3), 1769-92.

74. Locher, K. P., Review. Structure and mechanism of ATP-binding cassette transporters. *Philos Trans R Soc Lond B Biol Sci* **2009,** *364* (1514), 239-45.

75. Gottesman, M. M.; Ambudkar, S. V., Overview: ABC Transporters and Human Disease. *J Bioenerg Biomembr* **2001,** *33* (6), 453-458.

76. Dean, M., The Genetics of ATP-Binding Cassette Transporters. *Methods Enzymol* **2005,** *400,* 409-429.

77. Tarling, E. J.; de Aguiar Vallim, T. Q.; Edwards, P. A., Role of ABC transporters in lipid transport and human disease. *Trends Endocrinol Metab* **2013,** *24* (7), 342-50.

78. Lin, J. H.; Yamazaki, M., Role of P-Glycoprotein in Pharmacokinetics. *Clin Pharmacokinet* **2003,** *42* (1), 59-98.

79. Fojo, A. T.; Udea, K.; Slamon, D. J.; Poplack, D. G.; Gottesman, M. M.; I., P., Expression of a multridrug-resistance gene in human tumors and tissues. *Proc Natl Acad Sci* **1987,** *84*, 265-269.

80. Wang, J. Q.; Yang, Y.; Cai, C. Y.; Teng, Q. X.; Cui, Q.; Lin, J.; Assaraf, Y. G.; Chen, Z. S., Multidrug resistance proteins (MRPs): Structure, function and the overcoming of cancer multidrug resistance. *Drug Resist Updat* **2021,** *54*, 100743.

81. Munoz, M.; Henderson, M.; Haber, M.; Norris, M., Role of the MRP1/ABCC1 multidrug transporter protein in cancer. *IUBMB Life* **2007,** *59* (12), 752-7.

82. Doyle, L. A.; Yang, W.; Abruzzo, L. V.; Krogmann, T.; Gao, Y.; Rishi, A. K.; Ross, D. D., A multidrug resistance transporter from human MCF-7 breast cancer cells. *Proc Natl Acad Sci* **1998,** *95*, 15665-15670.

83. Robey, R. W.; Polgar, O.; Deeken, J.; To, K. W.; Bates, S. E., ABCG2: determining its relevance in clinical drug resistance. *Cancer Metastasis Rev* **2007,** *26* (1), 39-57.

84. Hall, J. A.; Seedarala, S.; Rice, N.; Kopel, L.; Halaweish, F.; Blagg, B. S., Cucurbitacin D Is a Disruptor of the HSP90 Chaperone Machinery. *J Nat Prod* **2015,** *78* (4), 873-9.

85. Bartalis, J.; Halaweish, F. T., In vitro and QSAR studies of cucurbitacins on HepG2 and HSC-T6 liver cell lines. *Bioorg Med Chem* **2011,** *19* (8), 2757-66.

86. Arjaibi, H. M.; Ahmed, M. S.; Halaweish, F. T., Mechanistic investigation of hepato-protective potential for cucurbitacins. *Medicinal Chemistry Research* **2017,** *26* (7), 1567-1573.

87. Ahmed, M. S.; Halaweish, F. T., Cucurbitacins: potential candidates targeting mitogen-activated protein kinase pathway for treatment of melanoma. *J Enzyme Inhib Med Chem* **2014,** *29* (2), 162-7.

88. Sun, J.; Blaskovich, M. A.; Jove, R.; Livingston, S. K.; Coppola, D.; Sebti, S. M., Cucurbitacin Q: a selective STAT3 activation inhibitor with potent antitumor activity. *Oncogene* **2005,** *24* (20), 3236-45.

89. Vincenzi, B.; Schiavon, G.; Silletta, M.; Santini, D.; Tonini, G., The biological properties of cetuximab. *Crit Rev Oncol Hematol* **2008,** *68* (2), 93-106.

90. Hong, D. S.; Fakih, M. G.; Strickler, J. H.; Desai, J.; Durm, G. A.; Shapiro, G. I.; Falchook, G. S.; Price, T. J.; Sacher, A.; Denlinger, C. S.; Bang, Y. J.; Dy, G. K.; Krauss, J. C.; Kuboki, Y.; Kuo, J. C.; Coveler, A. L.; Park, K.; Kim, T. W.; Barlesi, F.; Munster, P. N.; Ramalingam, S. S.; Burns, T. F.; Meric-Bernstam, F.; Henary, H.; Ngang, J.; Ngarmchamnanrith, G.; Kim, J.; Houk, B. E.; Canon, J.; Lipford, J. R.; Friberg, G.; Lito, P.; Govindan, R.; Li, B. T., KRAS(G12C) Inhibition with Sotorasib in Advanced Solid Tumors. *N Engl J Med* **2020,** *383* (13), 1207-1217.

Chapter Two

Triazole-Estradiol Analogs: A Potential Cancer Therapeutic Targeting Drug Resistance in Multiple Cancers

Abstract

1,2,3-triazoles have continuously shown effectiveness as biologically active systems towards various cancers, and when used in combination with steroid skeletons as a carrier, which can act as a drug delivery system, allows for a creation of a novel set of analogs that may be useful as a pharmacophore leading to a potential treatment option for cancer. A common molecular target for cancer inhibition is that of the Epidermal Growth Factor Receptor/Mitogen Activated Protein Kinase pathways, as inhibition of these proteins is associated with a decrease in cell viability. Estradiol-Triazole analogs were thus designed using a molecular modeling approach to combat these proteins. Thirteen of the high scoring analogs were then synthesized and tested *in-vitro* on an ovarian cancer cell line (A2780) and colorectal cancer cell line (HT-29). The most active compound, Fz25, shows low micromolar activity in both the ovarian (15.29 ± 2.19 μM) and colorectal lines (15.98 ± 0.39 μM). Mechanism of action studies proved that Fz25 moderately arrests cells in the G1 phase of the cell cycle, specifically inhibiting STAT3 in both cell lines. Additionally, Fz57 shows activity in the colorectal line (24.19 ± 1.37 μM). Inhibition studies in both cell lines show inhibition against various proteins in the EGFR pathway, namely EGFR (23%), STAT3 (65%), ERK (27%), and mTOR (26%). To further study their effects as potential therapeutics, Fz25 and Fz57 were studied against ABC transporters involved in drug efflux, namely P-glycoprotein and Breast Cancer Resistance Protein, which are associated with drug resistance. Ultimately Fz25 was able to inhibit P-glycoprotein by 5x the untreated

control, and half as well as the known inhibitor. Additional studies included molecular docking and dynamics, which identified that the hydrophobic nature of Fz25 and Fz57 lead to their binding affinities to respective proteins.

*Note, the research in this chapter was a part of the dissertation, but is published as such according to Ostlund et al 2022[52].

2.1 Introduction

1,2,3-Triazoles are an important chromophore which have gained popularity in organic synthesis for the development of many medicinal drugs due to their simplicity to make using click chemistry[1]. The triazole nitrogen allows for increased hydrogen bonding possibilities as well as an increase in hydrophilicity, an important drug ADME property. Triazoles are aplenty in medicine, often acting as cytochrome P450 inhibitors, anticancer, anti-HIV, and antifungals[2]. They are generally metabolized by glucuronidation, but can be quite resistant to metabolism when in close proximity to a substituted carbon system[3], providing attractive benefits for drugs containing this moiety. Specifically, 1,2,3-triazoles show high promise as anticancer. Numerous studies have been performed in the last decade alone describing the anticancer effect that triazoles possess, as well as their structure activity relationships[4-7].

Two of the deadliest cancers: colorectal cancer and ovarian cancer; give patients that are diagnosed with bleak outlooks for recovery. Colorectal cancer remains the second deadliest cancer, with incidence projections in 2020 at 127,000 and a death toll projected to be 49,000[8]. Colorectal cancer primarily starts as a noncancerous polyp, eventually leading to high dysplasia and developing into invasive cancer[9]. If caught early enough the disease can usually be treated, most commonly by resection of the polyp, but if left undetected survivability drops dramatically. A number of different drugs have been developed for treating colorectal cancer. Wolpin et al. (2008) discusses six, three of which act as DNA/RNA damaging drugs, and three of which are growth factor inhibitors[10]. Most notably however is that these drugs typically have a low survivability rate. The DNA/RNA inhibition drugs Irinotecan, Capecitabine, and Oxaliplatin have multi-year overall

survivability rates (1.5-2.5 years)[11-12], while the growth factor inhibitors Cetuximab, Bevacizumab, and Panitumumab have much lower survivability rates (6 months to 2 years)[13-15]. This provoked us to look more into the growth factor pathway in hopes of finding a new inhibitor which could increase overall median survivability.

Ovarian cancer is the deadliest reproductive system cancer and in 2020 there were expected to be 21,750 new cases alone and about 14,000 deaths. The disease accounts for 5% of all cancer deaths in females but has a high 5-year survivability rate of greater than 90% when detected early[16-17]. Ovarian cancer is known as a "silent killer" because, when diagnosed, the disease usually has already spread into the peritoneum through the detachment of spheroids. The spheroids are resistant to common therapeutic drugs such as cisplatin. This resistance could be in part due to the low penetration rate of the drug through the spheroids, hypoxic conditions, and the low proliferating cells inside the spheroid[18,19]. Additional resistance could come from overexpression of drug-efflux proteins, enhanced anti-apoptotic proteins, or elevated DNA damage tolerance[20,21].

As previously mentioned, growth factors remain a primary target. The Epidermal Growth Factor Receptor (EGFR) pathway is responsible for many cellular functions including cell growth, transcription, survival, and apoptosis[22]. Specifically, Signal Transducer and Activator of Transcription 3 (STAT3) and Extracellular-Regulated Kinase (ERK 1/2) are two proteins that are valuable targets. STAT3 has been found to be overexpressed in ovarian cancer due to its role in tumor growth and viability[23], while ERK has a positive role in cell proliferation and survival[24]. The Phosphoinositide 3-kinase (PI3K), Protein Kinase B (Akt) and the mechanistic Target of Rapamycin (mTOR) pathway is also highly important in colorectal cancer, with all three proteins showing

overregulation[25]. Therefore, growth receptors are an ideal target for inhibition in hopes to combat cancer. There exists a plethora of recent literature in which these pathways are studied in hopes of finding a molecular target to inhibit a variety of these proteins[26-28]. It is typically found that these inhibitors often inhibit specific proteins in the pathway but can promote others which leads to a resistance that eventually develops. We in turn also wish to look into the effects that an inhibitor may play on drug efflux proteins, specifically P-glycoprotein (P-gp) and Breast Cancer Resistance Protein (BCRP, also known as ABCG2). Drug efflux proteins are commonly associated with drug resistance due to their ability to efflux inhibitors from cells[29,30]. Therefore, if an inhibitor could also slow down the activation of these proteins as well as the growth factor proteins, it may act as a promising anticancer agent.

Steroids analogs have long been thought to be a potent source of clinical significance pertaining to cancer. Fulvestrant, Estramustine, and Exemestane are a few oncological steroids analogs that play a clinical role in breast and prostate cancer by protein inhibition or degradation of cell structure. Additionally, a number of steroidal analogs developed in our laboratory also possess anticancer properties[34-38]. Steroids analogs possess key features sought in drug candidates, including their lipophilicity and bioactive nature in the body. It is thus likely that the steroid skeleton can act as a carrier for a bioactive conformer. As such we utilized the steroid skeleton as a carrier for the triazole moiety so that it may have better pharmacokinetic properties. Additionally, the anti-inflammatory properties of steroids may play a role in oncology[31,32]. Overall, these features provide steroids as ideal skeletons for drug development. As previously mentioned, 1,2,3-triazoles also possess anticancer activities, and so a novel idea was formed to incorporate the two together. Hybrid

anticancer agents have been studied greatly, with a review performed by Fortin et al. (2013)[33]**.** In specific the authors describe so-called "Combi-molecules" acting as a combination of two molecules for synergistic effects. Specifically there describes a section on conjugation of small molecules to steroid nuclei as a potent source for carrying pharmacophores through the body.

The use of steroids as a drug carrier in intestinal cancer cell lines is an approach not often taken as most steroids are metabolized quickly in the intestinal tract. A major review was published in 2015 however that describes the impact that steroids have on the GI tract[40]. Bouguen et al. (2015) describe the production of glucocorticoids and the regulation of steroids in the intestinal system and found that these are often prone to strong immune responses. Of particular interest to this study is a section on colorectal cancer in which estrogens played a key response in cell apoptosis by binding to estrogen receptor β[41]. The progression of disease is thought to be due to regulation of key steroidal enzymes and provides a source that steroids may exhibit a key role in cancer progression in areas of the body where steroids may be aplenty.

Our research group has previously synthesized many steroid analogs by installation of side chain to carbon-17 of the steroid scaffold[34-38]. These analogs have shown competitive inhibitory effects against members of both the EGFR and MAPK pathways. A novel set of Triazole-Estradiol analogs has been synthesized in our lab previously by Alotaibi et al. (2017) and has shown inhibitory effects in colorectal cancer within the EGFR pathway[39]. Due to the potency of triazole analogs in various cancers we wished to see how an analog containing the triazole moiety attached to the steroid skeleton would act against ovarian and colorectal cancers. This prompted us to create a study wherein we investigated the

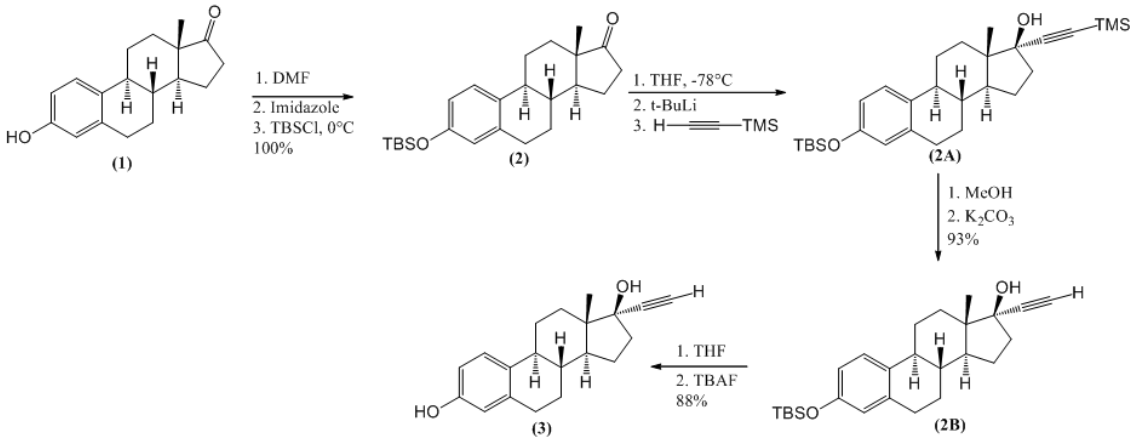
binding of these Estradiol-Triazole analogs toward proteins involved in the EGFR pathway using an *in-silico* approach through the OpenEye® software. We then discuss the synthesis of these analogs, followed by their effectiveness in ovarian and colorectal cancer cell lines. Additionally, we provide a plausible target for the analogs through cell cycle analysis and In-cell western assay studies. Finally, we look into the ability for analogs to inhibit drug efflux proteins P-gp and BRCP to determine if the analog may be of interest in drug resistant studies.

2.2 Results and Discussion

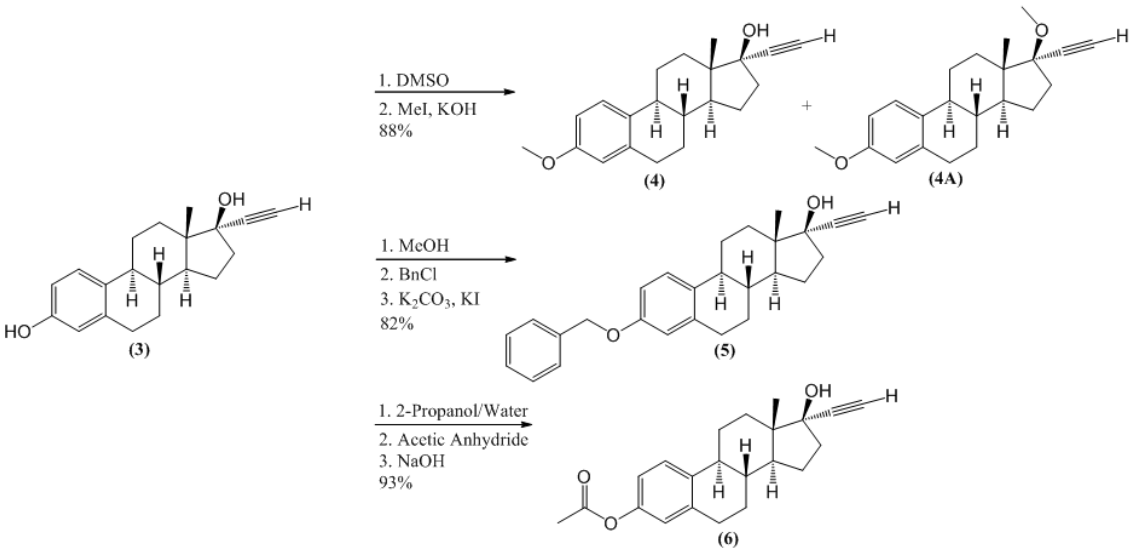
2.2.1 Synthetic Strategy and Design of Triazole-Estradiol Analogs

Novel triazole-estradiol analogs were previously synthesized using various reactions from our laboratory group as well as according to known chemical reactions[36,39]. Synthetic schemes and final analog structures will be reported here, but the original synthesis and identification of analogs can be viewed in Alotaibi et al (2017)[39]. Design and synthesis of the 1,2,3-triazole moiety as a biological pharmacophore that the steroid analog could carry was priority. As such it was decided that carbon-17 of the steroid structure would be a prime location for adaptation to triazole functionality as this position provides anticancer effects in our previous studies[34-38]. Analogs were synthesized from alkyne intermediate (**Scheme 2.1**). Estrone (**1**) was protected with TBS protecting group at carbon-3 to yield TBS-protected Estrone (**2**) which then underwent nucleophilic attack by alkyne[42]. TBS protection allows for increased lipophilicity of the analog which makes solubility of all reagents easier for the proceeding reactions as well as provides less side reaction. Alkylation of carbon-17 of the steroid (**2** to **2A)** primarily yields (*R*)-estradiol intermediate. This is due to the methyl group at carbon-13 promoting attack of the alkyne from beneath

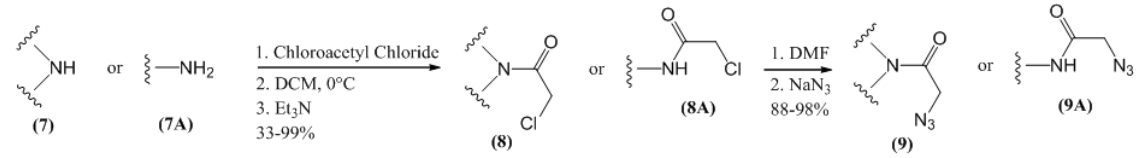
the steroid. This can further be corroborated as the R-isomer is solely obtained as reported[43,44]. TMS protected alkyne (**2A**) was first deprotected to yield primary alkyne (**2B**) followed by deprotection of TBS group with fluoride anion to yield intermediate estradiol with alkyne functionality (**3**). It is also worth noting that this dual deprotection can likely be completed in a single step with 2 equivalents of the fluoride anion. From this point carbon-3 of the steroidal structure was either left as a phenol species, or substitution was performed to add a methyl, benzyl, or acetyl group to the phenol (**Scheme 2.2**). During methylation, primary product will be species **4** as the phenolic hydrogen is more acidic; however methylation of the c-17 alcohol may occur if left for excess time or with excess methylating reagent. Azide side chains were created according to **Scheme 2.3.** Mono- or disubstituted amines were first acylated to yield their corresponding chloride derivatives, followed by substitution of the chlorine with sodium azide. The azide substitution requires the use of rapid spinning as the sodium azide is left undissolved but proceeds with great yields due to increased surface contact. Finally, cycloaddition of alkyne (**3, 4, 5, or 6)** with respective azide species was performed with Copper (II) Sulfate catalyst to yield triazole-estradiol analogs (**Scheme 2.4**). The cycloaddition requires the use of Copper (I), but is often made in-situ with the use of the reducing agent sodium ascorbate[45].



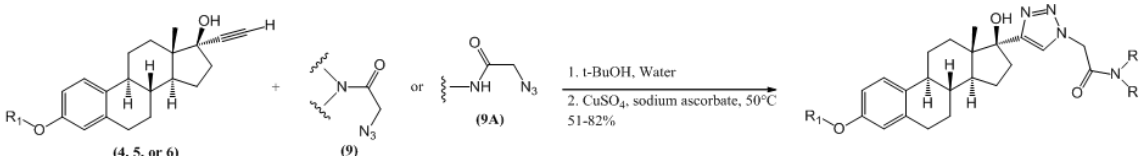
Scheme 2.1. Synthetic route for alkyne intermediate (**3**).



Scheme 2.2. Synthetic route for carbon-3 additions (**4, 5, and 6)**.



Scheme 2.3. Synthetic route for azide side chains of likeness to **9** or **9A**.



Scheme 2.4. Synthetic route for synthesis of Estradiol-Triazole analogs via cycloaddition of alkyne with azide.

Ultimately 14 Triazole-Estradiol analogs were synthesized, and the structures can be viewed in Figure 2.1.

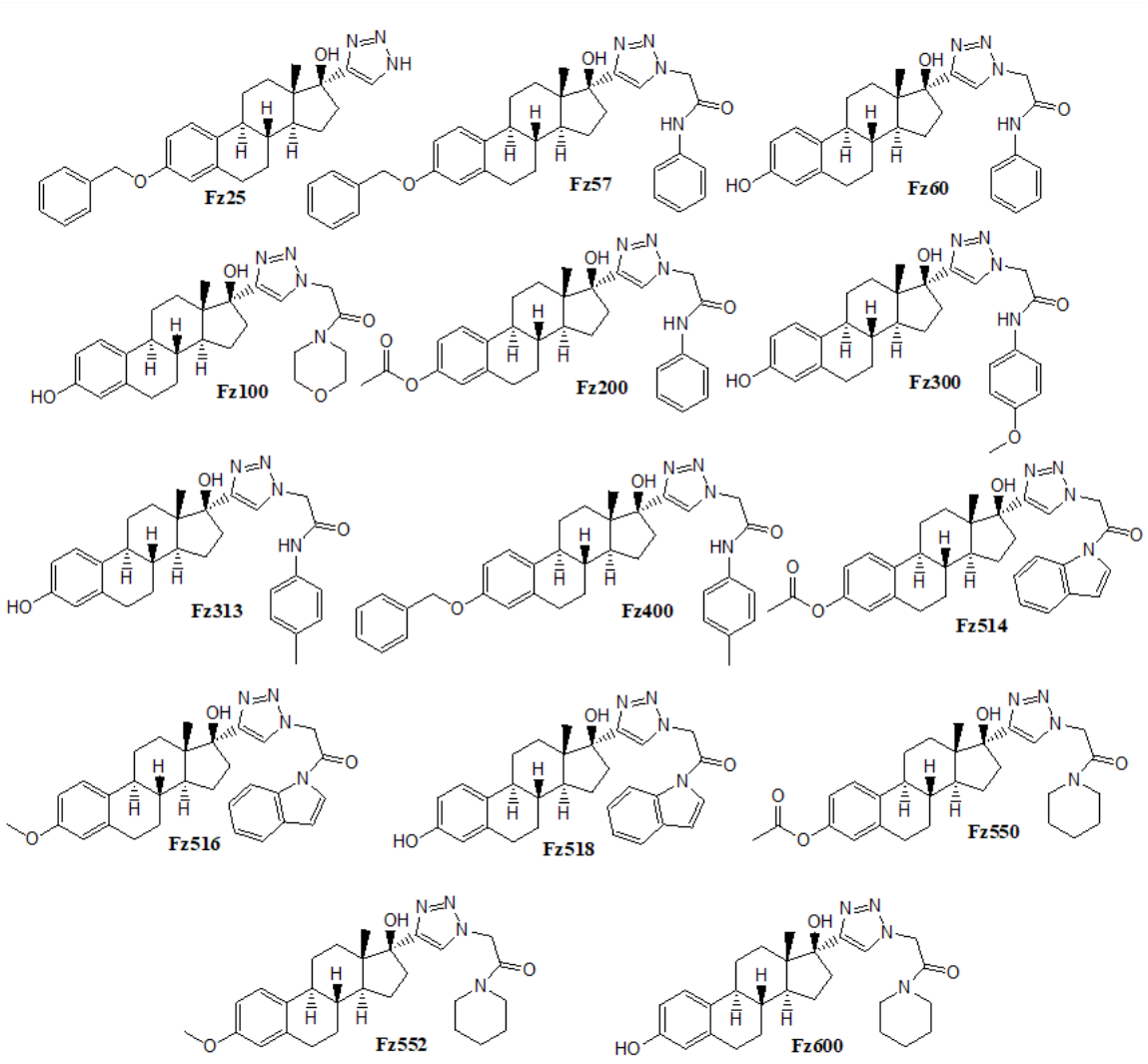


Figure 2.1 Synthesized Triazole-Estradiol analogs used for biological evaluation.

2.2.2 Biological Evaluation of Triazole-Estradiol Analogs in Ovarian and Colorectal Cancer

Cytotoxicity Assay

Of the fourteen Estradiol-Triazole analogs synthesized and tested, three showed measurable anti-proliferation activity against ovarian carcinoma cells (A2780). IC_{50} values and analog functionalization are shown in Table 2.1. Fz25, Fz313, and Fz400 show measurable anti-proliferation activity with Fz25 having an IC_{50} of 15.29 ± 2.19 µM. In the

colorectal cells (HT-29), Fz25, Fz57, and Fz400 show measurable anti-proliferation activity with Fz25 also showing the highest ability to inhibit cell viability. Interestingly, Fz25 possesses a single hydrogen as the R^2 side chain (Figure 2.1) whereas the other analogs have much larger side chains at this position, possibly suggesting that the size of Fz25 (lack of a triazole side chain) appears to be key for fitting into a specific protein active site, specifically STAT3 (see figure 2.9). We also point out the C-3 benzyl group as a functional group of likely importance, which we see later in molecular docking studies. It is worth noting that Fz313 and Fz57 show measurable cytotoxicity in the ovarian and colorectal cell lines respectively, but not the other. No mechanism of action studies were performed on these analogs in both cell lines to identify the cause of this behavior, but it may be plausible that one cell line has overexpression of certain proteins more than the other, which may lead to higher cytotoxic values between different cell types. Ultimately, because of its ability to limit proliferation of cancerous cells, Fz25 underwent further testing in both cell lines to determine its mechanism of action by utilization of cell cycle analysis and In-cell western assay. Fz57 also underwent testing in the cell cycle analysis and In-cell western assay due to its cytotoxic effects in colorectal cells.

Table 2.1 IC_{50} for synthesized Triazole-Estradiol analogs.

Analog ID	R_1	R_2	IC_{50} (A2780) (μM)	IC_{50} (HT-29) (μM)
Fz25	Benzyl	H	15.29 ± 2.19	15.98 ± 0.39
Fz57	Benzyl	N-Phenylethanamide	> 50	24.19 ± 1.37
Fz60	H	N-Phenylethanamide	> 50	>50
Fz100	H	4-Acetylmorpholine	> 50	>50
Fz200	Acetyl	N-Phenylethanamide	> 50	>50
Fz300	H	N-(4-Methoxyphenyl) ethanamide	> 50	>50
Fz313	H	N-(4-Methylphenyl) ethanamide	37.86 ± 4.03	>50
Fz400	Benzyl	N-(4-Methylphneyl) ethanamide	43.60 ± 1.38	28.98 ± 1.15
Fz514	Acetyl	N-Acetylindole	> 50	>50
Fz516	Methyl	N-Acetylindole	> 50	>50
Fz550	Acetyl	N-Acetylpiperidine	> 50	>50
Fz552	Methyl	N-Acetylpiperidine	> 50	>50
Fz600	H	N-Acetylpiperidine	> 50	>50

Cell Cycle Analysis

To help determine its mechanism of action, Fz25 was incubated with ovarian carcinoma cells and DNA content was measured by flow cytometry to identify in what cycle of cell replication inhibition was occurring. Using DMSO as a standard (0.1%), Fz25 was incubated with cells at the IC_{50} concentration as well as at twice the IC_{50} concentration (Figure 2.2). Analysis was performed in quadruplicate. There was a slight increase (3%) in G1 percentage at twice the IC_{50} as well as a decrease overall in G2 phase percentage (1.7%).

The decrease in G1 phase percentage of cells at the IC_{50} is interesting and is likely due to an increase in apoptotic cells in the trials as we also see a decrease in G2 phase percentage. The overall cell cycle analysis in the ovarian cancer line is inconclusive however, as no significant rise above the standard error in G1 phase is seen.

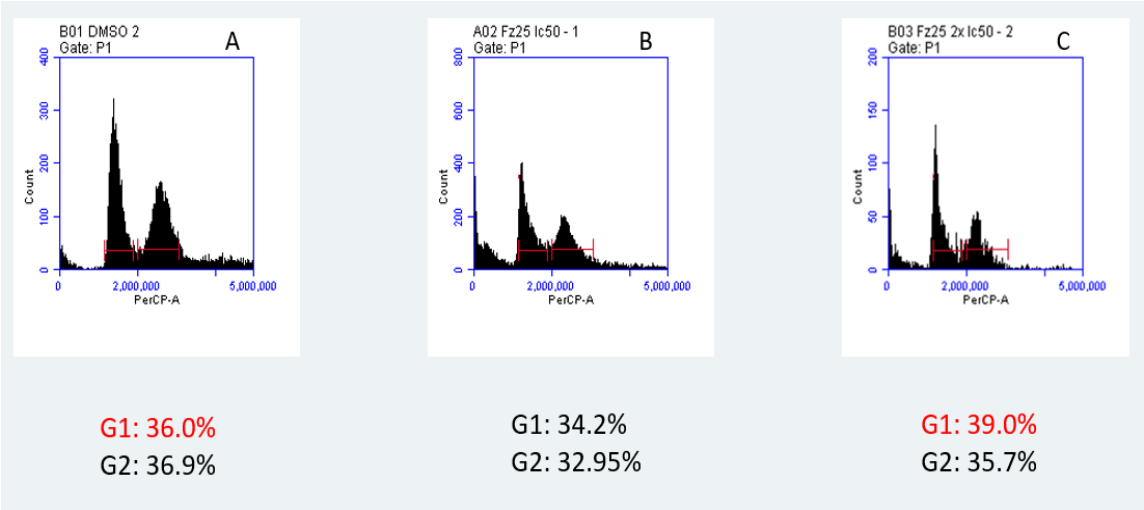


Figure 2.2. Flow cytometry analysis of the cell cycle of ovarian carcinoma cells. A) DMSO control. B) Fz25 at 15.29 µM. C) Fz25 at 30.58 µM. G1 and G2 phases are represented as an average of four trials.

In the colorectal cell line, Fz25 and Fz57 were analyzed according to the same protocol. Fz25 shows a moderate rise in G1 levels, while Fz57 shows a significant 3.2-fold increase in G1 levels as compared to the control (Figure 2.3). Interestingly, Fz25 shows a very minor increase when it comes to an increase in G1 phase accumulation in both ovarian and colorectal cells. As it will be mentioned later, Fz25 shows significant effects against multiple proteins, and would be expected to have a large G1 phase arrest. A rise in G1 phase concentration is typical of a deficit in growth factors and cell regulators [**47**], which helps elucidate that Fz25 and Fz57 may interact with growth factor receptors as previously

hypothesized. We theorize that perhaps Fz25 may be cell cycle non-specific, but more testing would likely be necessary to confirm this. In any case an increase in G1 arrest probes us to look into protein inhibition studies using an In-cell western assay.

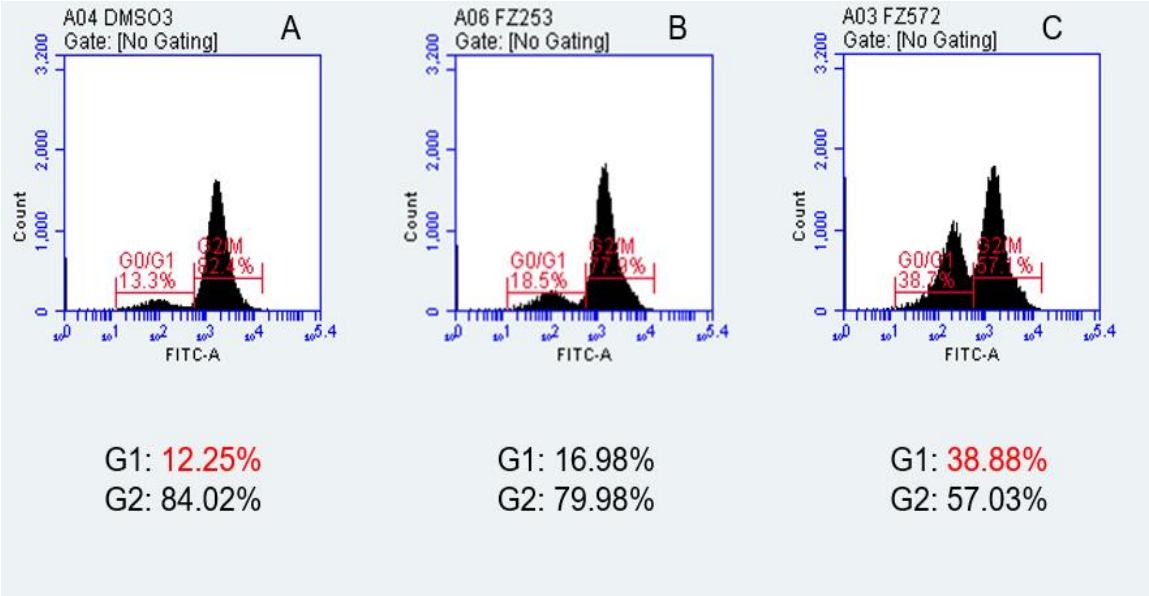


Figure 2.3. Flow cytometry analysis of the cell cycle of colorectal cells. A) DMSO control. B) Fz25 at 15.98 µM. C) Fz57 at 24.19 µM. G1 and G2 phases are represented as an average of four trials.

In-cell western assay

Fz25 was tested in the ovarian cell line against four common proteins within the EGFR cascade, namely EGFR itself, ERK, STAT3, and mTOR. These proteins were chosen due to their inhibition in various studies related to ovarian and colorectal cancers, as well as overexpression of the proteins in one or the other[22-25,37,38]. Concentrations were chosen based on the IC_{50} value of the analog used at values of 0.25, 0.5, 1, and 2x the IC_{50} value. To be considered significant, average inhibition had to have a p-value less than or equal to 0.05. As we saw that Fz25 and Fz57 show a rise in G1 phase cells, it corroborates our

theory that signaling proteins and cell proliferation proteins, of which the previously four are grouped as, are prime targets for inhibition studies. Of the four proteins analyzed, EGFR showed only moderate inhibition studies, being inhibited by Fz25 slightly up to 15 µM, and up to 25% at 30 µM as seen in Figure 2.4. For STAT3 we saw even better inhibition values; 43% and 66% respectively for the IC_{50} (15.29 µM) and twice the IC_{50} (30.58 µM). Fz25 and Fz57 were also analyzed in the colorectal cell line against the same four proteins. Figure 2.5 sees Fz25 showing significant activity against STAT3 and mTOR at twice the IC_{50} (31.96 µM), and in STAT3 specifically showing activity at half the IC_{50} (7.99 µM). As Fz25 shows significant activity against STAT3 in dual cell lines, it becomes obvious that the interaction of Fz25 with the STAT3 protein is important on a molecular level. This led us to utilize a molecular docking study in depth on STAT3 with Fz25 bound that is discussed later. Indeed, it seems that the JAK/STAT3 pathway is responsible for cytotoxic effects that Fz25 shows, and the probing of this pathway is worth investigating more in future studies.

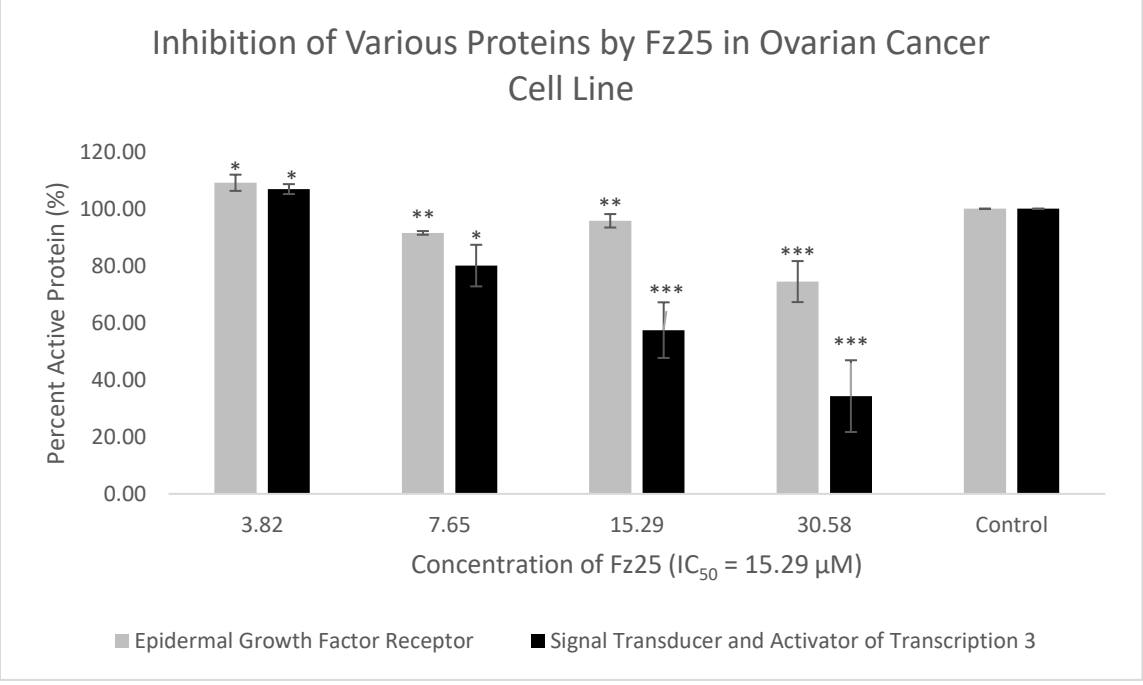


Figure 2.4. In-cell western data for Fz25 in ovarian cell line A2780 expressed as a percentage of the active forms of Epidermal Growth Factor Receptor (EGFR) and Signal Transducer and Activator of Transcription 3 (STAT3). Level of Significance: *not significant, ** p≤0.05, *** p≤0.01.

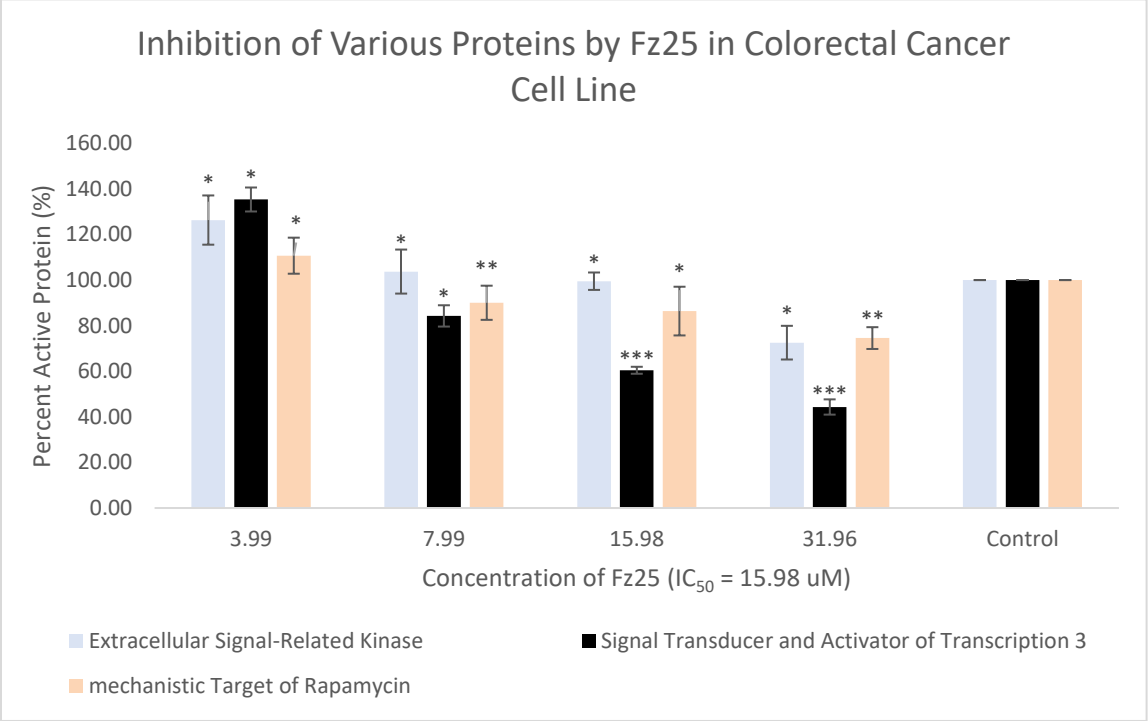


Figure 2.5. In-cell western data for Fz25 in colorectal cell line HT-29 expressed as a percentage of the active forms of Extracellular Signal-Related Kinase (ERK), Signal Transducer and Activator of Transcription 3 (STAT3), and the mechanistic Target of Rapamycin (mTOR). Level of Significance: *not significant, ** p≤0.05, *** p≤0.01.

Further inhibition studies were performed for Fz57 on the colorectal cell line. Results showed measurable EGFR and ERK inhibition activity. At twice the IC_{50} concentration, Fz57 showed around 10% inhibition of EGFR and ERK. This inhibition is low in correspondence with the large G1 phase arrest. The cytotoxicity then may be due to

inhibition of various proteins in the ERK pathway, such as RAS, RAF, or MAPK, as it seems that the inhibition of ERK alone may not be responsible for the cytotoxic effects of Fz57. Indeed, the ERK/RAS/MAPK pathway has shown previous success against analogs with steroidal skeletons as carriers of active groups[34,37,38], and begs the question if this pathway provides the key for cell death. Additionally, there exists the possibility that a G-protein coupled estradiol receptor (GPER) may play a role in cytotoxicity due to its coupling with EGFR. GPER activation thus exists as a plausible mechanism for tumor suppression[46,47]; however, as seen in figure 1.4, the scope of this project is focused on downstream effects of EGFR signaling to support our hypothesis. It's also worth noting that the analogs in this manuscript are inhibitors of EGFR proteins and may not activate GPER. Nonetheless this seems to warrant future pharmacokinetic studies as well as lead-hit optimization of these triazole analogs.

Inhibition fluorescence screening

As we have elucidated the cytotoxicity of Fz25 and Fz57 in the selected cancer cell lines, it begs the question if Fz25 and Fz57 will play a role in drug resistance. Efflux transporters can extrude a wide range of xenobiotics including anticancer agents and have been widely associated with tumor resistance in cancer. The activity of drug efflux proteins, particularly P-glycoprotein (P-gp), have been reported to mediate the resistance of ovarian tumors allowing them to evade the chemotherapy-induced cytotoxicity of various anticancer agents[48-50]. Therefore, drug efflux transporters serve as attractive targets for novel drug strategies against drug resistance in cancer. To investigate the effects Fz25 and Fz57 have on efflux proteins, two fluorescent substrates, calcein-AM and JC-1, were used

to report the activities of P-gp and Breast Cancer Resistance Protein (BCRP) transporters respectively via an image-based fluorescence inhibition assay. The strategy for the assay was to evaluate fluorescence accumulation as a measure of transporter activity in the presence or absence of inhibition. From the results (Figure 2.6), fluorescent substrates accumulation was depleted in P-gp or BCRP untreated (no inhibition) cells which demonstrates the active efflux of substrates out of the cells. In contrast, cells treated with positive inhibitor controls Reversan, Verapamil, or KO143, which are known inhibitors of P-gp and BCRP, showed improvement in intracellular fluorescence accumulation. Fz25 and Fz57 likewise also show moderate inhibition of calcein efflux. Fz25 was more than half (~55%) of that exhibited by the positive control in P-gp overexpressing cells. However, in BCRP overexpressing cells, Fz25 and Fz57 show no significant increase in BCRP inhibition. As expected, fluorescence accumulation was higher for control cells due to no transporter activity. The results suggest that Fz25 and Fz57 both exert inhibitory effects on P-gp, but neither shows it against BCRP. This screening was performed by Dr. Jennifer Kyerematen in collaboration work.

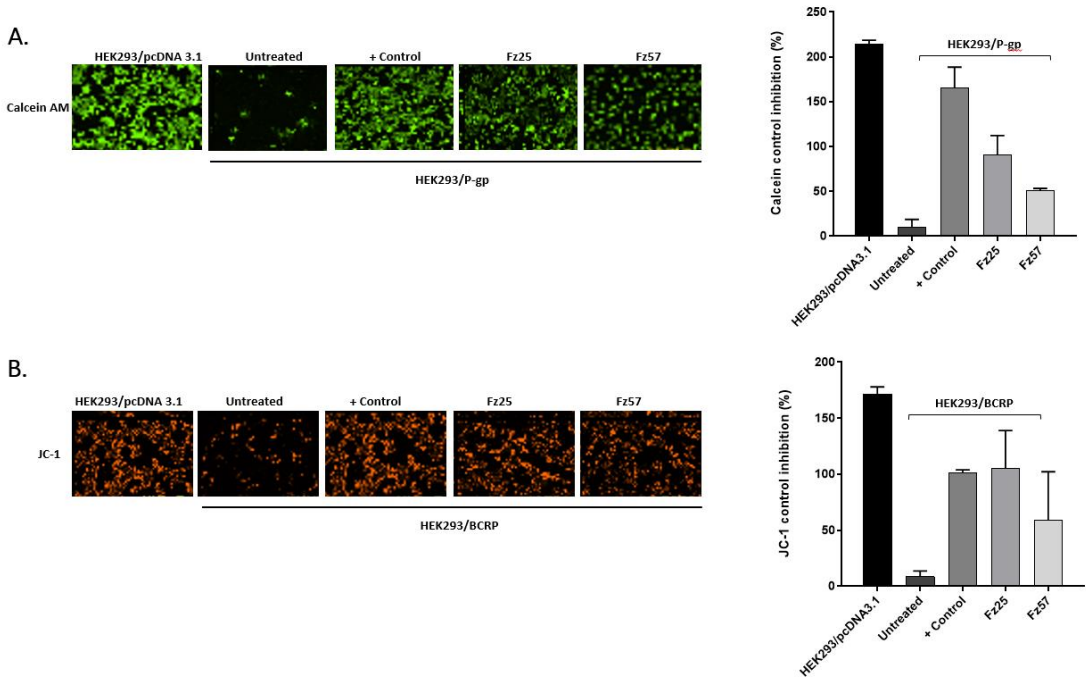


Figure 2.6. Imaging-based detection of fluorescent substrate accumulation. (A) HEK293/pcDNA3.1 was treated with DMSO (0.2% final concentration) and HEK293/P-gp cells treated with DMSO (0.2% final concentration) (untreated) or positive control (15 µM reversan) or test compounds (10 µM Fz25 or Fz57) and calcein AM (0.25 µM). (B) HEK293/pcDNA3.1 was treated with DMSO (0.2% final concentration) and HEK293/BCRP cells treated with DMSO (0.2% final concentration) (untreated) or positive control, KO143 (10 µM) or test compounds (10 µM Fz25 or Fz57) and JC-1 (1 µM).

Fluorescence accumulation assay

To validate further Fz25 and Fz57 inhibition on P-gp and BCRP, we conducted flow cytometry-based accumulation assays using the same fluorescent reporters. Similarly, Fz25 and Fz57 enhanced intracellular fluorescence accumulation levels of both calcein and JC-1 (Figure 2.7) as consistent with the results from the fluorescence inhibition assay.

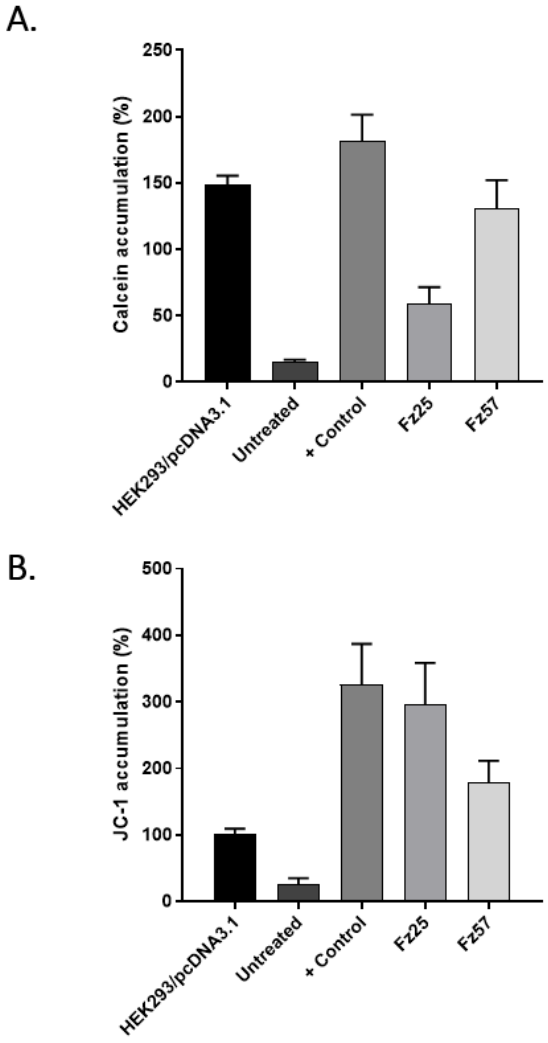


Figure 2.7. Flow cytometry-based confirmation of inhibition (A). HEK293/pcDNA3.1 was treated DMSO (0.2% final concentration) DMSO (0.2% final concentration) while HEK293/P-gp cells were treated with DMSO (0.2% final concentration) (untreated) or positive control (10 µM verapamil) or test compounds (10 µM Fz25 or Fz57) for 10min and then calcein AM (0.25 µM) at 37ºC for 30 min. (B) HEK293/pcDNA3.1 was treated with DMSO (0.2% final concentration) while HEK293/BCRP cells were treated with DMSO (0.2% final concentration) (untreated) or positive control KO143 (10 µM) or test

compounds (10 µM Fz25 or Fz57) at 37ºC for 30 min and JC-1 (0.25 µM) for 2 h . Collated results from two independent experiments done in duplicates presented as Mean ± SD.

2.2.3 Confirmation of Biological Activity with Molecular Modeling

Due to the *in-vitro* cytotoxicity studies and inhibition studies, Fz25 was studied further in depth using an *in-silico* modeling study to determine the binding effects of Fz25 against EGFR and STAT3. Fz25 bound in EGFR (Figure 2.8) shows a tight fit within the active site of the protein, owing to its absence of a side chain on the triazole moiety that we mentioned earlier. Likely binding affinity and inhibition comes from electrostatic and hydrophobic interactions. It can be seen that a particularly negative electrostatic potential exists at GLU-738 and ASP-831. These electrostatic interactions may play a role in how this binding occurs, particularly ion-dipole interactions that may occur between GLU-738 and Nitrogen-1 of the triazole group and between ASP-831 and Nitrogen 2 and 3 of the triazole moiety. This triazole interaction helps affirm our hypothesis that the triazole acts as a bioactive pharmacophore. In STAT3 (Figure 2.9), Fz25 shows an additional hydrogen bond between the triazole group and the protein backbone between CYS-418 and GLY-419 of 1.7 Å. There also appears to be some additional electrostatic interaction in the form of an ion-dipole interaction between the 17-methoxy species and ARG-379 (3.82 Å). Likely these two factors account for the binding affinity of Fz25 and its inhibition against STAT3.

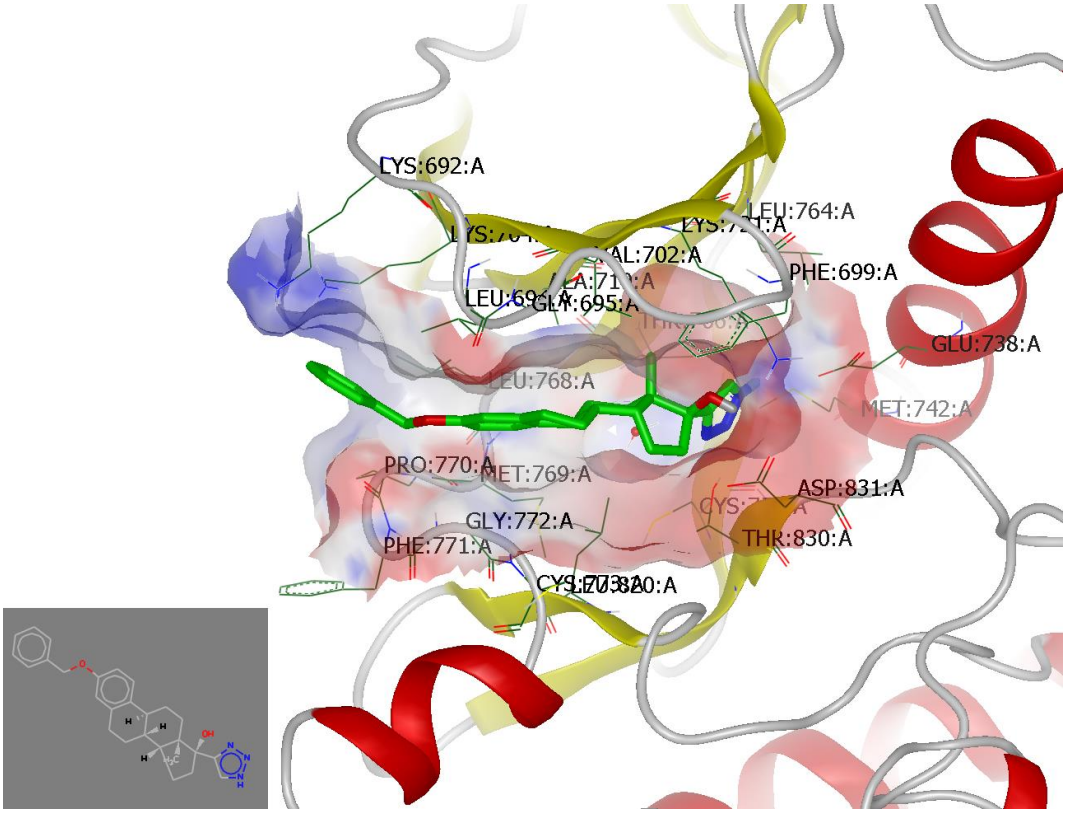


Figure 2.8. Molecular modeling image displaying Fz25 bound within epidermal growth factor receptor (EGFR) active site. The binding of Fz25 to EGFR is likely due to hydrophobic and electrostatic interactions. The triazole moiety is in close proximity to GLU-738 and ASP-831, displaying the electrostatic interactions present.

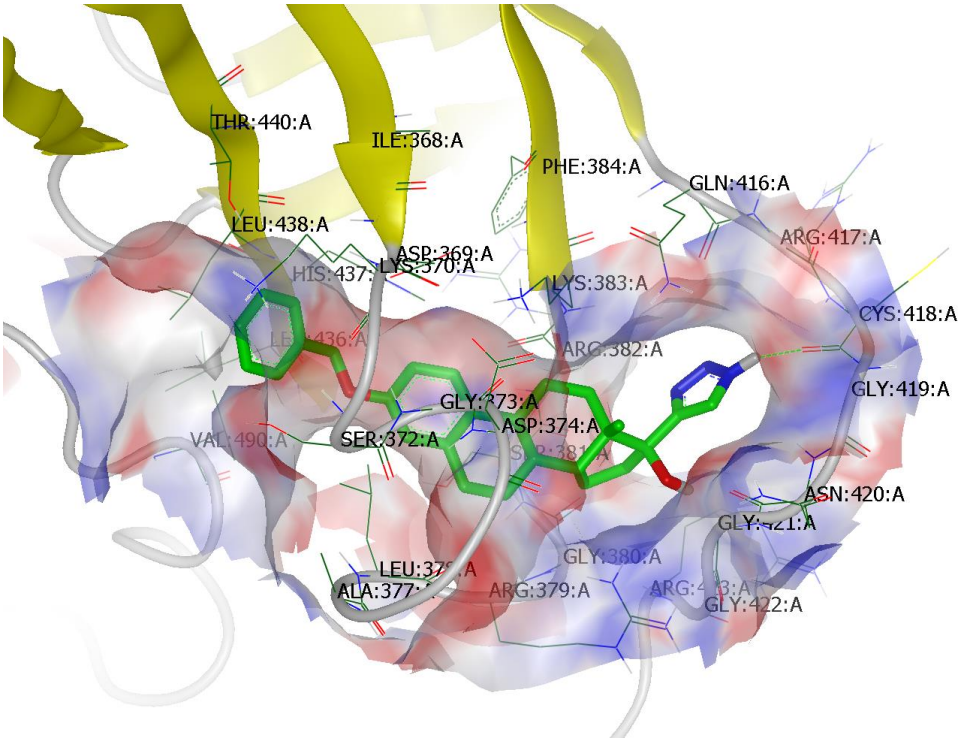
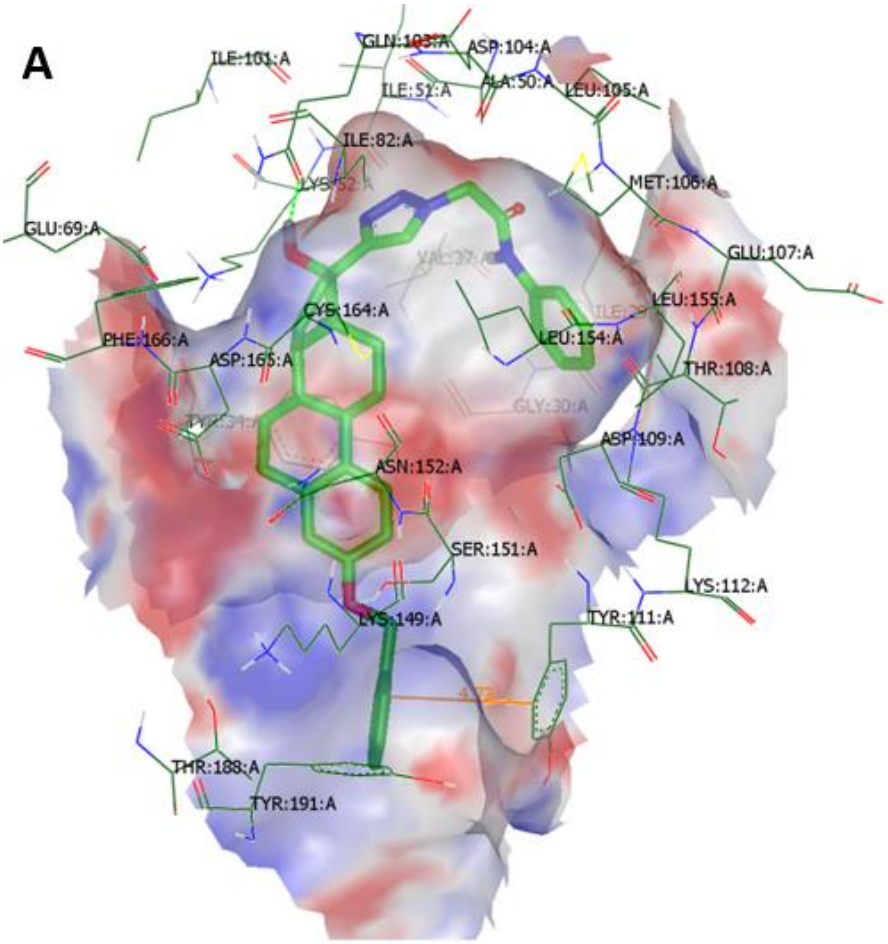


Figure 2.9. Molecular modeling image displaying Fz25 bound within signal transducer and activator of transcription 3 (STAT3) active site. Hydrogen bonding between triazole hydrogen and backbone amide of CYS-418 seen on right.

We also investigated the binding effects of Fz57 *in-silico* within the ERK pocket. As seen from a top-down view in Figure 2.10A, Fz57 fits well into the closed pocket area of ERK. It also has pi-pi stacking at a distance of 4.73 Å between the benzyl group at carbon 3 and TYR-111. One of the major hydrogen bonds can also be seen between the carbon 17 hydroxyl group and the carbonyl of GLN-103. A side view of the docking interaction (Figure 2.10B) provides another view of the binding interaction, with additional views at 2 more hydrogen bonds. One is between the carbon 3 oxygen and SER-151, and the other between the Triazole amide carbonyl and the backbone N-H of MET-106. These interactions together provide insight as to why Fz57 is an inhibitor of ERK.

The conclusion of the molecular modeling study shows that indeed the triazole functionality, the triazole ring itself, and functionality at C3 of the steroid nucleus have definitive binding to proteins in the EGFR cascade. This reinforces our hypothesis that the steroid can act as a carrier molecule for the triazole moiety.



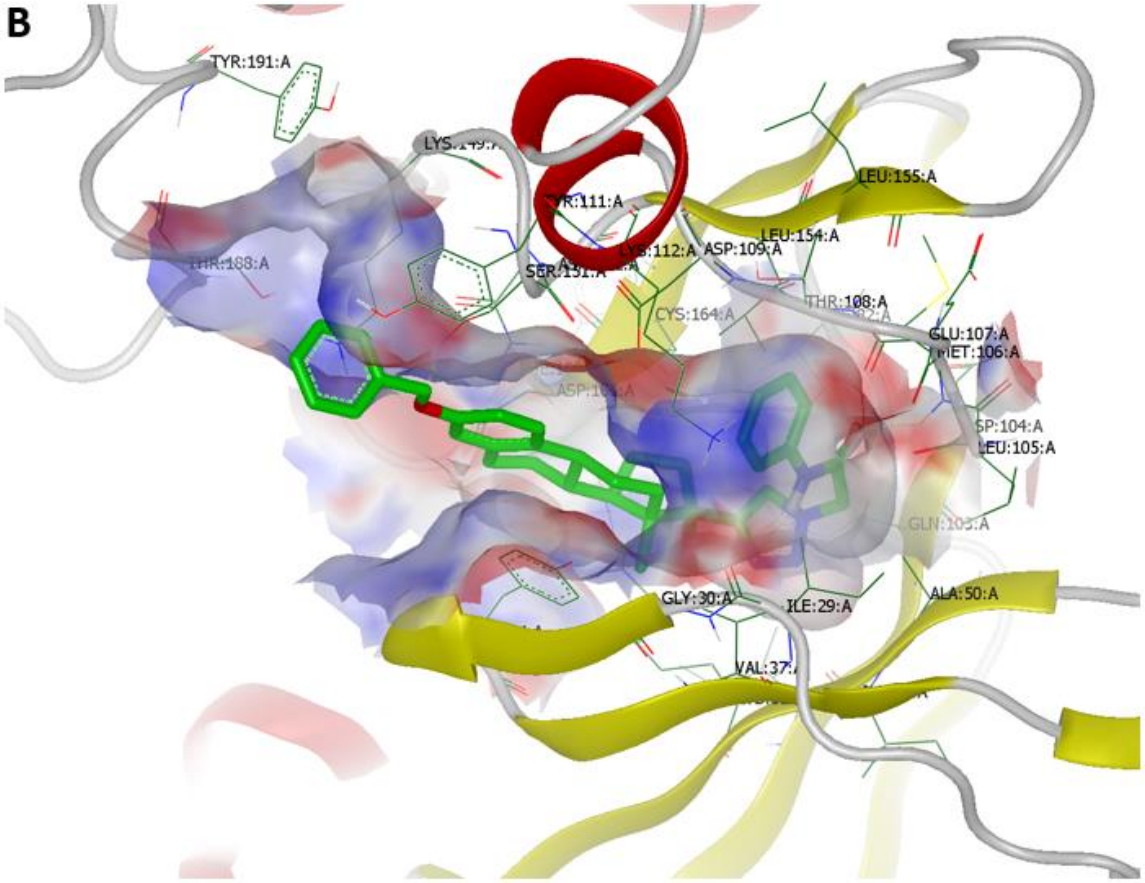


Figure 2.10. A) Top-down view of Fz57 bound to ERK. Pi-Pi interaction seen between TYR-111 and Benzyl side chain at a distance of 4.73 Å. B) Side view of Fz57 bound to ERK. Hydrogen bonding seen between C-17 hydroxyl group with GLN-103, C-3 oxygen with SER-151, and Amide carbonyl with MET-106.

2.2.4 Molecular Dynamics Study

To validate our findings from the *in-vitro* and *in-silico* data we wished to conduct molecular dynamic simulations using Schrödinger's Maestro application. We explored the simulation events between Fz25 and STAT3 as well as Fz57 within the ERK protein. During the molecular dynamic simulation of Fz25 within STAT3 (pdb code: 6tlc) it was determined that Fz25 was not stable in the STAT3 pocket, moving around throughout the

simulation. We can tell due to three key factors. First, The RMSD of Fz25 during the first 50 nanoseconds of the simulation is low, but increases from 6 to 40 Å after 50 ns for the remainder of the simulation (figure 2.11A). This high of an RMSD is notorious with an unbound ligand. Second, we can see by looking at the protein ligand contacts (figure 2.11B) that the contacts of Fz25 throughout the simulation are less than 0.15% of the time, indicating very little stable contacts. Finally, this is corroborated in figure 2.11C where the contact lifetimes are monitored over the duration of the simulation. Very little contacts are stable throughout the simulation, suggesting that Fz25 is not stable in the STAT3 pocket.

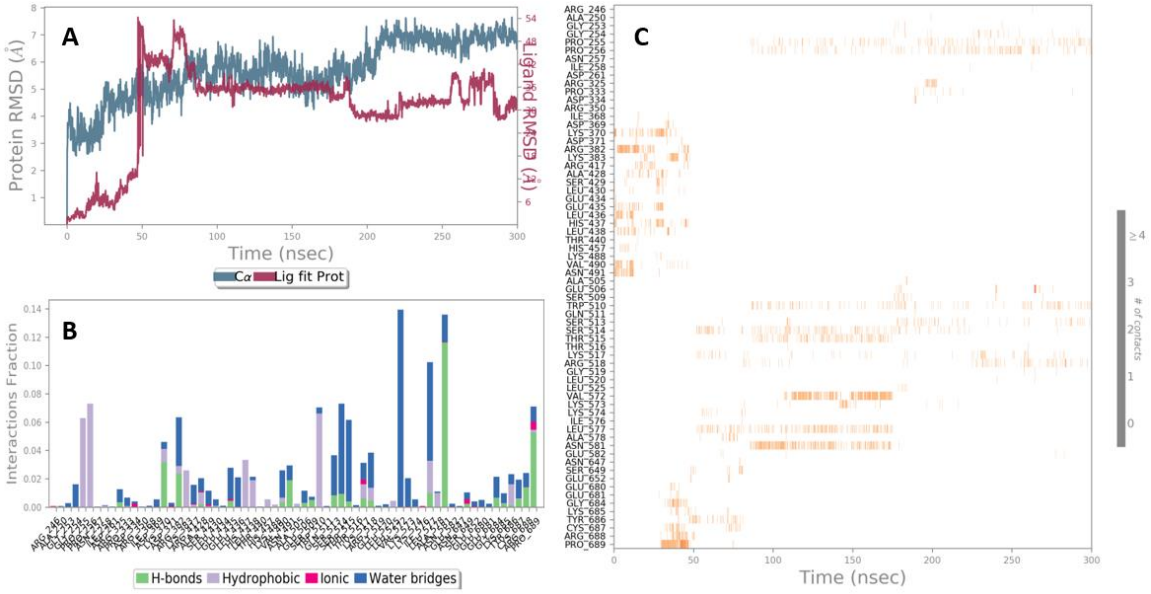


Figure 2.11. Molecular dynamic simulation results of Fz25 within the STAT3 pocket. A) RMSD of both Fz25 and STAT3 during the simulation. B) Contact interaction between Fz25 and STAT3 within the simulation. C) Duration of contacts between Fz25 and STAT3 over the timeline of the simulation.

There thus exists a discrepancy between the biological results and dynamic simulation results of Fz25 binding to STAT3. We ran the molecular dynamic simulation twice, and

identified similar results between each, and as such we have a few probable ideas to suggest that the molecular dynamic simulation for STAT3 is an outlier. The first suggestion we thought of was that the solvation energy for STAT3 might be very high, thus the release of water from the binding of the ligand may produce a significant amount of energy, but the MMGBSA data suggests that there actually is more van der Waals and lipophilic energy culminating to the $\Delta G_{binding}$ values. We thus dismissed this idea, and looked at the more likely possibility that perhaps the force field is not estimating this protein well. We found a similar example from Roos et al. 2019 in which OPLS-2005 has poor abilities to accurately determine torsional and partial charge parameters in some proteins[51]. They suggest a newer version, OPLS-3e, in which free energy perturbations were much closer to the experimental results. We wish to explore this option in future studies to obtain a higher degree of accuracy in our molecular dynamic simulations.

When we performed molecular docking studies with Fz57 in the ERK pocket, we found pi-cation interactions to exist and possibly play a key role in binding. Molecular dynamic simulation of Fz57 within the ERK pocket found similar results, albeit with different amino acids. Figure 2.12A details the contacts between Fz57 and ERK. Many hydrophobic interactions exist as we previously thought, but when we look at the visualized contacts in figure 2.12B it's determined that a cation-pi interaction between LYS-52 and the Aromatic ring A of the steroidal core of Fz57 is responsible for most of the binding interactions. While these results vary from the molecular docking results, the pi interactions seem to be a key player in binding for Fz57 with ERK.

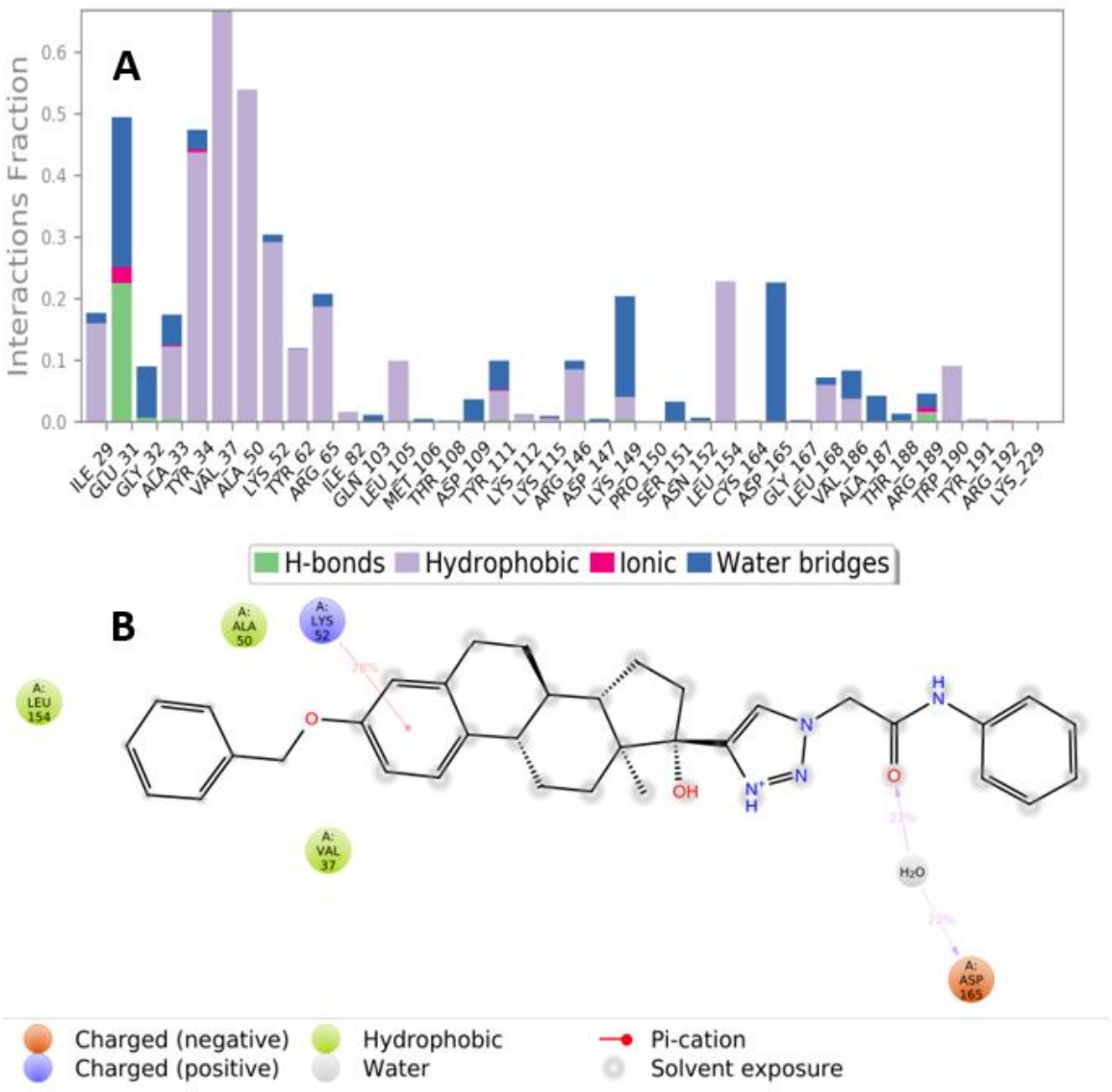


Figure 2.12. Molecular dynamic simulation results from Fz57 bound in ERK. A) Protein-ligand contacts suggest high hydrophobic contacts of Fz57 with ERK. B) Protein-ligand contacts are primarily hydrophobic in nature, with a cation-pi interaction of LYS-52 with Fz57.

A final confirmation that Fz57 binds to ERK in a hydrophobic manner can be seen through the MMGBSA results (Figure 2.13). The results indicate a $\Delta G_{binding}$ of -72.1

kJ/mole, with the primary amount of that binding energy arising due to lipophilic and van der Waals interactions.

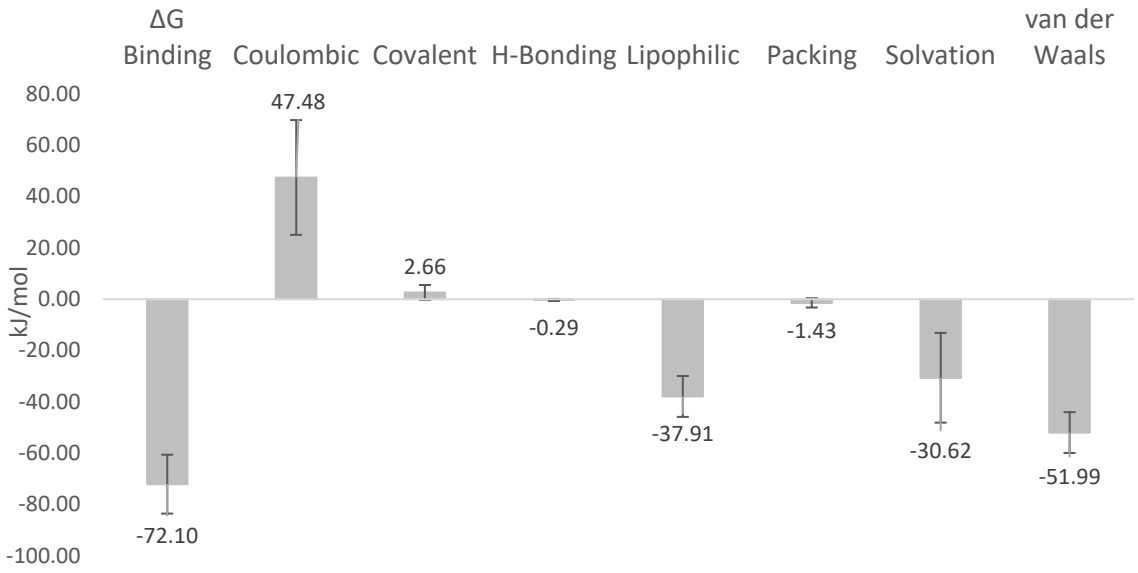


Figure 2.13. MMGBSA binding energy results from Fz57 simulation with ERK. Results are an average of the entire simulation length with values taken every 30 ns.

2.3 Materials and Methods

2.3.1 Biological Evaluation

Cell Lines and Culture

Calcein-AM was procured from Corning Life Sciences (Corning, NY). Reversan were purchased from Sigma-Aldrich (St. Louis, MO). JC-1 and KO143 were purchased from Life Technologies (Eugene, OR) and Tocris Bioscience (Bristol, UK) respectively. The ovarian cancer cell line (A2780) and colorectal cancer cell line (HT-29) were obtained from American Type Culture Collection (ATCC). The ovarian cells were maintained in RPMI-1640 (Thermoscientific™ Hyclone™) supplemented with 10% fetal bovine serum (FBS) (Thermoscientific™ Hyclone™) and 1% penicillin (100 IU/mL)/streptomycin (100

µg/mL) (Corning™ Cellgro™). The colorectal cells were maintained in Dulbecco's modified eagle medium (DMEM) and supplemented likewise in FBS and antibiotics. HEK293/pcDNA3.1, HEK293/P-gp and HEK293/BCRP cells were kind gifts from Dr. Suresh V. Ambudkar (NIH, Bethesda, MD). HEK293 cell lines were cultured in Dulbecco's modified Eagle's medium (DMEM) (GE Healthcare, Marlborough, MA) supplemented with 10% (v/v) fetal bovine serum (FBS) (HycloneTM, GE Healthcare Life Sciences). All cells were maintained at 37°C with 5% CO_2.

Cytotoxicity Assay

The cell lines were seeded in a 96-well plate at $3*10^6$ cells/mL (100 µL/well). A serial dilution of all Estradiol-Triazole analogs were added after overnight incubation of the cells at 37°C and 5% CO_2. DMSO was used as a control (0.1 %). The cells were incubated with the analogs for 48 hrs. followed by addition of 15 µL of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) (5 mg/mL PBS) to each well and incubation for another 2 hours. The formazan crystals that formed were solubilized with 100 µL DMSO. The absorbance was measured at 570 nm by Biotek plate reader.

Cell Cycle Analysis

Cells were seeded at $2.5*10^5$ cells/mL in a 6-well plate (2 mL/well) and allowed to adhere overnight at 37°C and 5% CO_2. The cells were incubated with different concentrations of hit compounds (Fz25 and Fz57) at their respective IC_{50} values and a DMSO control for 24 hours. The cells were then washed twice with ice-cold 1X PBS and collected after trypsinization. The cell pellet was washed two times with ice-cold 1X PBS and fixed with ice-cold 70% ethanol overnight at -4°C. When ready for analysis, the cells were centrifuged. Then they were washed once with ice-cold PBS followed by a second

wash with ice-cold PBS-2% FBS. The cell pellet was re-suspended in 500 µL propidium iodide (PI)/RNase (BD Biosciences, BDB550825) staining solution for 15 min at room temperature (RT) in the dark and analyzed within 1 hour by BD C6 Accuri flow cytometer (BD Biosciences, San Jose, CA).

In-Cell Western Assay

Cells were seeded as described above then prepared for In-cell western assay according to Rockland protocol[53]. After seeding, hit compounds (Fz25 or Fz57, one per plate) was added to wells in various concentrations (1/4, 1/2, 1, and 2 times the IC_{50}) along with a DMSO control at the same concentration as twice the IC_{50} and incubated for 24 hours. Rockland protocol was then followed to fix, permeabilize, and block cells. Primary antibodies in both phosphorylated and non-phosphorylated forms were added for EGFR, mTOR, STAT3, and ERK. GAPDH was also used as a control for non-phosphorylated forms. Secondary antibodies used were Mouse IgG (H&L Antibody Dylight) 680 Conjugated with fluorochrome and Rabbit IgG (H&L Antibody DyLight) 800 Conjugated with fluorochrome to detect both forms of the proteins simultaneously. The plate was read on LiCOR Odyssey® imager. Images were processed with ImageJ software[54]. Inhibition of a protein is taken as a percentage of the phosphorylated versus non-phosphorylated forms of each protein that was expressed. Values are taken from quadruplicate testing.

Fluorescence Inhibition Screening Assay

Fluorescence inhibition screening assay was performed on HEK293/P-gp cells using calcein AM as a fluorescent substrate and reversan as a positive control inhibitor. For HEK293/BCRP cells, JC-1 was used as a fluorescent substrate and KO143 as a positive control inhibitor. HEK293/pcDNA3.1 cell line, which contains an empty pcDNA3.1 vector

(no P-gp or BCRP expression), was used as a control cell line. HEK293/P-gp and HEK293/BCRP at 8 $\times 10^3$ and 4×10^4 cell densities respectively were seeded in a 96-well optical-bottom plates (100 μL/well) with polymer base (Thermo Fisher Scientific, Waltham, MA) coated with poly-D-lysine (Sigma-Aldrich, St. Louis, MO) and incubated overnight. Control cell lines were also seeded at the respective cell concentrations. Prior to drug treatment, the culture medium in each well was replaced with 80 μL serum-free medium. For drug treatment, Fz25 or Fz57 (10 μM final concentration) along with DMSO control (0.2% final concentration), Reversan (10 μM final concentration), or KO143 (10 μM final concentration) were added. After incubation for 30 min, P-gp cells were treated with calcein AM (0.25 μM final concentration), and JC-1 (1 μM final concentration) was added to BCRP cells. Both cells were incubated for an additional 1 hour. Following the incubation period, media containing treatments were removed. Cells were then suspended in 100 uL PBS containing 10 mM HEPES and 4.5% glucose after washing once with PBS prior to analysis.

An ImageXpress Micro XLS Widefield High-Content Analysis System (Molecular Devices, Sunnyvale, CA) equipped with a 60× objective with 0.70 numerical aperture was used to acquire fluorescent images. Images for calcein AM were detected using an FITC filter with excitation and emission wavelengths of 482/35 nm and 536/40 nm respectively and exposure time for 100 ms. JC-1 was viewed in a Texas red channel with excitation of 562/40 nm and emission achieved at 624/40 nm. Fluorescent images were quantified according to a custom application module that identifies fluorescent objects against a background fluorescent intensity and applies segmentation masks per parameters set for object size. Experiments were conducted at two independent times.

Flow-Cytometry Fluorescence Accumulation Assay

HEK293/P-gp or HEK293/BCRP along with HEK293/pcDNA3.1 cells were prepared in serum free DMEM culture medium at a density of 7×10^5 cells/ml. Cells were treated with test compounds (Fz25 or Fz57 at 10 μM final concentration), DMSO (0.2% final concentration) or positive controls (verapamil or KO143 at 10 μM final concentration) and incubated at 37℃ for 10 min. After treatment, P-gp cells were incubated with 0.25 μM calcein-AM at 37℃ for 30 min. For BCRP, cells were incubated with JC-1 (0.25 μM final concentration) at 37℃ for 2 hours after treatments with test compounds and controls. The reaction was stopped by addition of 3 ml ice-cold PBS. Cells were then pelleted and washed three times with ice-cold PBS. Prior to analysis, cells were fixed with 1% paraformaldehyde (in cold PBS) and kept on ice. Samples were analyzed on a BD Accuri C6 flow cytometer (BD Biosciences, San Jose, CA) equipped with 488 nm argon laser. Calcein-AM and JC-1 were detected with 533/30 nm and 585/40 nm band filters respectively. Fluorescence intensity values were collated as the mean of 10,000 events. Two independent experiments with duplicate treatments were performed.

2.3.2 Molecular Modeling

A Library of over 700 analogs were virtually screened through an *in-silico* molecular docking process. OpenEye® Scientific Software and ChemOffice were utilized in the docking process. Analog structures were first designed using ChemDraw software according to their stereochemistry followed by 3-Dimension optimization in Chem3D using the energy minimization function Merck Molecular Force Field (MMFF94) calculation[55]. This force field calculation utilizes a maximum number of iterations that is set at 500, with an RMS gradient of 0.100. The dielectric constant of the analog is not

changed during this calculation, and for every analog is left to be assumed as 1.000 with no cut-off method required. This force field calculation utilizes location and identity of the atoms to predict the molecule's cumulative potential energy[55]. Each analog can be minimized in the span of a few seconds. Minimized energy conformers were then used in a docking process utilizing Omega[56,57] and FRED (Fast Rigid Exhaustive Docking)[58] modules along with the crystal structures for the proteins as followed: EGFR (PDB ID: 1M17), ERK (PDB ID: 2OJG), CDK2 (PDB ID: 3QL8), mTOR (PDB ID: 4JSV), STAT3 (PDB ID: 1BG1), Akt (PDB ID: 3MV5), PI3K (PDB ID: 3L54), ERα (PDB ID: 1A52), and ERβ (PDB ID: 3OLS). Crystal structures are downloaded from PDB without alteration using MakeReceptor program following software protocols[56-59]. Any bound ligands and water molecules are removed from the crystal structure in this program, and then the binding pocket is identified using a cavity detection. This docking is done in the absence of water molecule due to lower computational time and allow for more flexibility for the analog-protein interaction. Finally, any part of the protein that is not near the binding pocket can be removed from the program to lower computational time when performing docking. FRED is then utilized to dock the analogs within the proteins, which spans anywhere from 5 minutes to 10 hours depending on the size of the protein and analog library[56-59]. OpenEye measures binding affinity using a relative value known as a consensus score. Consensus scores are output numerical values that take into account many analog-protein binding affects, including steric interactions, hydrogen bonding, lipophilic and hydrophilic interactions, electrostatics, and shape-based interactions. The lower the consensus score, the better the binding to the protein[59].

2.3.3 Molecular Dynamics

Analogs Fz25 and Fz57 underwent dynamic simulation in two respective proteins: STAT3 and ERK, according to their preference in the biological data. The analogs were drawn using ChemDraw with respective stereochemical features, minimized in energy using Chem3D's MMFF94 energy minimization, and saved in a .pdb format. Proteins were downloaded in .pdb format from Protein Data Bank using the following codes: 6TLC (STAT3) and 2OJG (ERK). Structural alterations were first performed on the protein in PyMOL to remove excess chains and ligands from downloaded files. Swiss PDB viewer was then used to minimize proteins with GROMOS 43B1 energy minimization function. To create the protein-ligand complexes, PyRx-vina software was utilized for molecular docking, followed by conversion of the protein-ligand complex to .pdb file using PyMOL. Molecular dynamic simulations were then performed using Schrodinger Desmond Maestro software[60]. Pre-processing of the protein-ligand complex was performed to set correct bond orders, add hydrogens, and detect disulfide bonds. Secondly, a refining step was performed on the prepared structured file in order to analyze and change protonation and tautomeric states of variable residues by maestro's protein preparation wizard. Thirdly, A structural optimization step was performed to optimize hydrogen bond network of the protein-ligand complex. Energy minimization was then performed to relieve any bond strain or steric hindrance, followed by introduction of a water model using TIP3P 3-site models. Next, a force field was incorporated using OPLS-2005 force field function to estimate forces between atoms within molecules. Ionization models were set for pH equal to 7.4 to mimic the human body and to distribute positive or negative counter ions to neutralize the system. Upon completion of the pre-processing steps for the complex, molecular dynamic

simulation was performed using a 300 ns simulation event with 3,000 frames. After the simulation was completed, Simulation event analysis was utilized to calculate the RMSD, RMSF, Radius of Gyration, Intra- and Intermolecular hydrogen bonding events for both the protein and ligand, including the complex. Finally, simulation interaction diagrams were created to display the average interactions of the protein and ligand over the entire simulation event. Additionally, an MMGBSA was performed using the completed trajectory to calculate total binding free energies for the protein-ligand complexes.

2.4 Optimization of Triazole-Estradiol Analogs

Follow ups of this study have already begun, including optimization of the triazole-estradiol analogs previously synthesized. Currently three new analogs have been synthesized, all containing the benzyl side chain at C3 of the analog structure, as all of these analogs showed measurable IC_{50} values in first round studies. One of the compounds, Fz515, with an indole side chain off the triazole ring, shows top binding scores among synthesized analogs in BCRP, MRP1, and P-glycoprotein, suggesting it may pose as a potent inhibitor of ABC transporters responsible for drug resistance in cancer. The other two analogs are modulations of Fz57, containing a fluorine or trifluoromethyl group at the para position of the amide phenyl group. These are being tested in breast cancer, and are showing comparable micromolar activities as of writing.

2.5 Conclusions

Analogs with 1,2,3-triazole groups continue to show promising results in many different cancers. The combination of bioactive 1,2,3-triazoles with a steroidal skeleton carrier shows many promising features, including as a standalone or possible combination therapy for ovarian and colorectal cancers. Additionally, we found that these analogs partially

inhibit the drug efflux protein P-glycoprotein which gives rise to the idea that functionality in this compound may lead to future drugs with increased role in true drug resistance studies. In order to find a novel analog that may inhibit cancer cell growth and drug resistance proteins, we employed an *in-silico* modeling approach to find estradiol-triazole analogs that bind strongly to proteins in the EGFR/MAPK pathways. Promising analogs were synthesized and tested *in-vitro* on an ovarian cancer and colorectal cancer cell line to determine cytotoxicity and mechanism of action. Two analogs, Fz25 and Fz57, showed measurable inhibition of EGFR and ERK protein, while Fz25 had a significant inhibitory effect on STAT3. It's possible the analogs themselves are non-specific as binding occurs to multiple proteins, which can be a problem *in-vivo*, but this is a focus for future pharmacokinetic studies. Currently we've developed an analog that may act on more than one or have downstream effects of protein inhibition for proteins involved in the EGFR cascade. These triazole analogs as a whole continue to show promise as inhibitors for growth factor proteins, and the inclusion of a triazole moiety on the steroid scaffold strengthens the results we have seen in the past in our research group. The takeaways from this study show that the 1,2,3-triazole moiety is incredibly important as a structure activity relationship for drug candidates processed in the future. It's our hope that this study allows for a more focused design of triazole candidates for future oncology studies.

To further investigate Fz25 and Fz57 as potential drug candidates, we studied their ability to inhibit drug efflux proteins. The role of drug efflux proteins in limiting the bioavailability of many anticancer drugs is one of the key studied mechanisms of tumor resistance. For the improvement in anticancer therapy and for many drug molecules targeting tumor resistance, it is provident to test their interaction with efflux transporters.

The essence is to prevent situations where a drug candidate is a substrate of the transporter and thereby rendered non-beneficial or ineffective due to active efflux from target sites. Using calcein-AM and JC-1 as fluorescence reporters of P-gp and BCRP activity, fluorescence inhibition and accumulation studies showed Fz25 and Fz57 interfered with P-gp efflux activity through enhancing intracellular accumulation levels of its substrates. The results suggest that Fz25 and Fz57 have an inhibitory effect on P-gp and thus could be used as a starter compound for further studies of this type of analog to combat drug resistance. Finally, we investigated Fz25 and Fz57 in molecular docking and dynamic simulations, and determined that hydrophobic effects of both molecules play a role in their binding. Ultimately, resistance studies are ongoing, and in the future, we wish to investigate the role of Fz25 and Fz57 on a true resistant cell line, either as a standalone therapy, or in conjugation with a chemotherapeutic agent in hopes of finding a synergistic effect. Additionally, it is clear that more modification of the structure and computational studies/molecular dynamics need to be performed to find better candidates.

2.6 References

1. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B., A Stepwise Huisgen Cycloaddition Process: Copper(I)-Catalyzed Regioselective "Ligation" of Azides and Terminal Alkynes. *Angew. Chem.* **2002,** *114*, 2708-2711.

2. Dheer, D.; Singh, V.; Shankar, R., Medicinal attributes of 1,2,3-triazoles: Current developments. *Bioorg Chem* **2017,** *71*, 30-54.

3. Dalvie, D. K.; Kalgutkar, A. S.; Khojasteh-Bakht, S. C.; Obach, R. S.; O'Donnell, J. P., Biotransformation Reactions of Five-Membered Aromatic Heterocyclic Rings. *Chem. Res. Toxicol* **2002,** *15* (3), 269-299.

4. Duan, Y. C.; Ma, Y. C.; Zhang, E.; Shi, X. J.; Wang, M. M.; Ye, X. W.; Liu, H. M., Design and synthesis of novel 1,2,3-triazole-dithiocarbamate hybrids as potential anticancer agents. *Eur J Med Chem* **2013,** *62*, 11-9.

5. Xu, Z.; Zhao, S. J.; Liu, Y., 1,2,3-Triazole-containing hybrids as potential anticancer agents: Current developments, action mechanisms and structure-activity relationships. *Eur J Med Chem* **2019,** *183*, 111700.

6. Penthala, N. R.; Madhukuri, L.; Thakkar, S.; Madadi, N. R.; Lamture, G.; Eoff, R. L.; Crooks, P. A., Synthesis and anti-cancer screening of novel heterocyclic-(2H)-1,2,3-triazoles as potential anti-cancer agents. *Med. Chem. Commun.* **2015,** *6*, 1535-1543.

7. Prachayasittikul, V.; Pingaew, R.; Anuwongcharoen, N.; Worachartcheewan, A.; Nantasenamat, C.; Prachayasittikul, S.; Ruchirawat, S.; Prachayasittikul, V., Discovery

of novel 1,2,3-triazole derivatives as anticancer agents using QSAR and in silico structural modification. *Springerplus* **2015,** *4*, 571.

8. Rahib, L.; Smith, B. D.; Aizenberg, R.; Rosenzweig, A. B.; Fleshman, J. M.; Matrisian, L. M., Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* **2014,** *74* (11), 2913-21.

9. Markowitz, S. D.; Bertagnolli, M. M., Molecular origins of cancer: Molecular basis of colorectal cancer. *N Engl J Med* **2009,** *361* (25), 2449-60.

10. Wolpin, B. M.; Mayer, R. J., Systemic treatment of colorectal cancer. *Gastroenterology* **2008,** *134* (5), 1296-310.

11. Fujita, K.; Kubota, Y.; Ishida, H.; Sasaki, Y., Irinotecan, a key chemotherapeutic drug for metastatic colorectal cancer. *World J Gastroenterol* **2015,** *21* (43), 12234-48.

12. Comella, P., A review of the role of capecitabine in the treatment of colorectal cancer. *Ther Clin Risk Manag* **2007,** *3* (3), 421-431.

13. Jonker, D. J.; O'Callaghan, C. J.; Karapetis, C. S.; Zalcberg, J. R.; Tu, D.; Au, H.-J.; Berry, S. R.; Krahn, M.; Price, T.; Simes, R. J.; Tebbutt, N. C.; Hazel, G.; Wierzbicki, R.; Langer, C.; Moore, M. J., Cetuximab for the Treatment of Colorectal Cancer. *N Engl J Med* **2007,** *357*, 2040-2048.

14. Grothey, A.; Sugrue, M. M.; Purdie, D. M.; Dong, W.; Sargent, D.; Hedrick, E.; Kozloff, M., Bevacizumab beyond first progression is associated with prolonged overall

survival in metastatic colorectal cancer: results from a large observational cohort study (BRiTE). *J Clin Oncol* **2008,** *26* (33), 5326-34.

15. Battaglin, F.; Puccini, A.; Ahcene Djaballah, S.; Lenz, H. J., The impact of panitumumab treatment on survival and quality of life in patients with RAS wild-type metastatic colorectal cancer. *Cancer Manag Res* **2019,** *11*, 5911-5924.

16. Siegel, R. L.; Miller, K. D.; Jemal, A., Cancer statistics, 2016. *CA Cancer J Clin* **2016,** *66* (1), 7-30.

17. Torre, L. A.; Trabert, B.; DeSantis, C. E.; Miller, K. D.; Samimi, G.; Runowicz, C. D.; Gaudet, M. M.; Jemal, A.; Siegel, R. L., Ovarian cancer statistics, 2018. *CA Cancer J Clin* **2018,** *68* (4), 284-296.

18. Iwanicki, M. P.; Davidowitz, R. A.; Ng, M. R.; Besser, A.; Muranen, T.; Merritt, M.; Danuser, G.; Ince, T.; Brugge, J. S. Ovarian cancer spheroids use myosin-generated force to clear the mesothelium. *Cancer Discov* **2011**, *1*(2), 144–157.

19. Nederman, T.; Acker, H.; Carlsson, J. Penetration of substances into tumor tissue: a methodological study with microelectrodes and cellular spheroids. *In Vitro* **1983**, *19* (6), 479–488.

20. Rebucci, M.; Michiels, C. Molecular aspects of cancer cell resistance to chemotherapy. *Biochem Pharmacol* **2013**, *85* (9), 1219–1226.

21. Stewart, D. J. Mechanisms of resistance to cisplatin and carboplatin. *Crit Rev Oncol Hematol* **2007**, *63* (1), 12–31.

22. Wee, P.; Wang, Z. Epidermal Growth Factor Receptor Cell Proliferation Signaling Pathways. *Cancers (Basel)* **2017,** *9* (5).

23. Duan, Z.; Foster, R.; Bell, D. A.; Mahoney, J.; Wolak, K.; Vaidya, A.; Hampel, C.; Lee, H.; Seiden, M. V. Signal transducers and activators of transcription 3 pathway activation in drug-resistant ovarian cancer. *Clin Cancer Res* **2006,** *12* (17), 5055-5063.

24. Steinmetz, R.; Wagoner, H. A.; Zeng, P.; Hammond, J. R.; Hannon, T. S.; Meyers, J. L.; Pescovitz, O. H. Mechanisms regulating the constitutive activation of the extracellular signal-regulated kinase (ERK) signaling pathway in ovarian cancer and the effect of ribonucleic acid interference for ERK1/2 on cancer cell proliferation. *Mol Endocrinol* **2004,** *18* (10), 2570-2582.

25. Johnson, S. M.; Gulhati, P.; Rampy, B. A.; Han, Y.; Rychahou, P. G.; Doan, H. Q.; Weiss, H. L.; Evers, B. M., Novel expression patterns of PI3K/Akt/mTOR signaling pathway components in colorectal cancer. *J Am Coll Surg* **2010,** *210* (5), 767-76, 776-8.

26. Gasparri, M. L.; Bardhi, E.; Ruscito, I.; Papadia, A.; Farooqi, A. A.; Marchetti, C.; Bogani, G.; Ceccacci, I.; Mueller, M. D.; Panici, P. B., PI3K/AKT/mTOR Pathway in Ovarian Cancer Treatment: Are we on the Right Track? *Geburtshilfe und Frauenheilkunde* **2017,** *77* (10), 1095-1103.

27. Glaysher, S.; Bolton, L. M.; Johnson, P.; Atkey, N.; Dyson, M.; Torrance, C.; Cree, I. A., Targeting EGFR and PI3K pathways in ovarian cancer. *Br J Cancer* **2013,** *109* (7), 1786-94.

28. Wen, W.; Wu, J.; Liu, L.; Tian, Y.; Buettner, R.; Hsieh, M. Y.; Horne, D.; Dellinger, T. H.; Han, E. S.; Jove, R.; Yim, J. H., Synergistic anti-tumor effect of combined inhibition of EGFR and JAK/STAT3 pathways in human ovarian cancer. *Mol Cancer* **2015,** *14*, 100.

29. Osa-Andrews B, Tan KW, Sampson A and Iram SH. Development of Novel Intramolecular FRET-Based MRP1 Biosensors to Identify New Substrates and Modulators. *Pharmaceutics*, **2018** Oct 13;10(4).

30. Molinski SV, Bozóky Z, Iram SH and Ahmadi S. Biophysical Approaches facilitate Computational Drug Discovery for ATP-Binding Cassette Proteins. *Int. J. Med. Chem.* **2017**:1529402.

31. Gupta, A.; Kumar, B. S.; Negi, A. S., Current status on development of steroids as anticancer agents. *J Steroid Biochem Mol Biol* **2013,** *137*, 242-270.

32. Lossignol, D., A little help from steroids in oncology. *J Transl Int Med* **2016,** *4* (1), 52-54.

33. Fortin, S.; Berube, G., Advances in the development of hybrid anticancer drugs. *Expert Opin Drug Discov* **2013,** *8*, 1029-1047.

34. Ahmed, M. S.; Kopel, L.; Halaweish, F. Structural optimization and biological screening of a steroidal scaffold possessing cucurbitacin-like functionalities as B-Raf inhibitors. *ChemMedChem* **2014**, *9*, 1361-1367.

35. Alsayari, A.; Kopel, L.; Ahmed, M. S.; Pay, A.; Carlson, T.; Halaweish F. Design, synthesis, and biological evaluation of steroidal analogs as estrogenic/anti-estrogenic agents. *Steroids* **2017,** *188*, 32-40.

36. Kopel, L.; Ahmed, M. S.; Halaweish. F. T. Synthesis of novel estrone analogs by incorporation of thiophenols via conjugate addition to an enone side chain. *Steroids* **2013**, *78*, 1119-1125.

37. Ahmed, M. S.; El-Senduny, F.; Taylor, J.; Halaweish, F. T., Biological screening of cucurbitacin inspired estrone analogs targeting mitogen-activated protein kinase (MAPK) pathway. *Chem Biol Drug Des* **2017,** *90* (3), 478-484.

38. Mahnashi, M.; Elgazwi, S. M.; Ahmed, M. S.; Halaweish, F. T., Cucurbitacins inspired organic synthesis: Potential dual inhibitors targeting EGFR - MAPK pathway. *Eur J Med Chem* **2019,** *173*, 294-304.

39. Alotaibi, F. Design, Synthesis and Biological Evaluation of Novel Estradiol-Triazole Analogs Targeting Epidermal Growth Factor Receptors in Colorectal Cancer. MS Thesis, South Dakota State University, Brookings, SD, 2017.

40. Bouguen, G.; Dubuquoy, L.; Desreumaux, P.; Brunner, T.; Bertin, B., Intestinal steroidogenesis. *Steroids* **2015**, *103*, 64-71.

41. Wilkins, H. R.; Doucet, K.; Duke, V.; Morra, A.; Johnson, N., Estrogen prevents sustained COLO-205 human colon cancer cell growth by inducing apoptosis, decreasing c-myb protein, and decreasing transcription of the anti-apoptotic protein bcl-2. *Tumor Biol* **2010**, *31*, 16-22.

42. Trost, B. M.; Weiss, A. H. The Enantioselective Addition of Alkyne Nucleophiles to Carbonyl Groups. *Adv Synth Catal*. **2009**, *351*, 7-8.

43. Boivin, R. P.; Luu-The, V.; Lachance, R.; Labrie, F.; Poirier, D., Structure-Activity Relationships of 17α-Derivatives of Estradiol as Inhibitors of Steroid Sulfatase. *J. Med. Chem.* **2000,** *43*, 4465-4478.

44. de la Torre, M. C.; Asenjo, M.; Ramírez-López, P.; Sierra, M. A., The Reversible Nicholas Reaction in the Synthesis of Highly Symmetric Natural-Product-Based Macrocycles. *European Journal of Organic Chemistry* **2015,** *2015* (5), 1054-1067.

45. Huisgen, R., Szeimies, G., and Möbius, L. *Chem. Ber*. **1967**, *100*, 2494.

46. Ignatov, T.; Modl, S.; Thulig, M.; Weibenborn, C.; Treeck, O.; Ortmann, O.; Zenclussen, A.; Dan Costa, S.; Kalinski, T.; Ignatov, A., GPER-1 acts as a tumor supressor in ovarian cancer. *J. Ovarian Res.* **2013,** *6* (51), 1-10.

47. Liu, Q.; Chen, Z.; Jiang, G.; Zhou, Y.; Yang, X.; Huang, H.; Liu, H.; Du, J.; Wang, H., Epigenetic down regulation of G protein-coupled estrogen receptor (GPER) functions as a tumor suppressor in colorectal cancer. *Mol Cancer* **2017,** *16* (1), 87.

48. Ding, J.; Li, M.; Deng, L.; Li, T., Study on Biological Characteristics and Mechanism of Paclitaxel Induced Drug Resistance in Endometrial Carcinoma Cells. *Biomed Res Int* **2018,** *2018*, 8372085.

49. El-Araby, M.E., et al., Molecular Mimics of Classic P-Glycoprotein Inhibitors as Multidrug Resistance Suppressors and Their Synergistic Effect on Paclitaxel. PLoS One, 2017. 12(1): p. e0168938.

50. Guo, W., et al., Mitochondria P-glycoprotein confers paclitaxel resistance on ovarian cancer cells. Onco Targets Ther, 2019. 12: p. 3881-3891.

51. Roos, K.; Wu, C.; Damm, W.; Reboul, M.; Stevenson, J. M.; Lu, C.; Dahlgren, M. K.; Mondal, S.; Chen, W.; Wang, L.; Abel, R.; Friesner, R. A.; Harder, E. D., OPLS3e: Extending Force Field Coverage for Drug-Like Small Molecules. *J Chem Theory Comput* **2019,** *15* (3), 1863-1874.

52. Ostlund, T.; Alotaibi, F.; Kyerematen, J.; Halaweish, H.; Kasten, A.; Iram, S.; Halaweish, F., Triazole-estradiol analogs: A potential cancer therapeutic targeting ovarian and colorectal cancer. *Steroids* **2022**, *177*, 108950.

53. Rockland antibodies and assays. In-Cell Western Protocol. https://rockland-inc.com/uploadedFiles/Support/Protocols/In_cell_Western_Protocol_v3.pdf.

54. Rasband, W. S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, https://imagej.nih.gov/ij/, 1997-2018.

55. Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J Comput Chem* **1996**, *17*, 490-519.

56. Hawkins, P.; Skillman, A.; Warren, G.; Ellingson, B.; Stahl, M. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J Chem Inf Model* **2010**, *50*, 572-584.

57. OMEGA 2.5.1.4: OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.

58. McGann, M. FRED pose prediction and virtual screening accuracy. *J Chem Inf Model* **2011**, *51*, 578-596.

59. OpenEye Toolkits. OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.

60. *Desmond* version 2.2. Schrödinger. 120 West 45th Street, 17th Floor. New York, NY, 2009.

Chapter Three

Synthesis and Biological Evaluation of Novel Heterocyclic Estrone Analogs Targeting Pancreatic Cancer

Abstract

Pancreatic Ductal Adenocarcinoma (PDAC) is the fourth leading cause of cancer death in the United States with a low 5-year survivability rate of only 8%. Thus, a new treatment option is needed for pancreatic cancer. Twelve heterocyclic estrone analogs were synthesized and characterized by spectroscopic techniques before undergoing biological evaluation in two pancreatic cancer cell lines: AsPC-1 and Panc-1. In AsPC-1, TR33 had an IC_{50} of 38.37 ± 1.76 μM, while in Panc-1 TR26 had an IC_{50} of 10.16 ± 0.83 μM. Both analogs underwent further analysis in their respective cell lines. Cell cycle analysis showed that TR26 arrests cells in the G0/G1 phase of the cell cycle, a key feature highlighting a deficit in growth factors. In-cell western protein analysis assay confirmed these results in which extracellular signal-related kinase (ERK) was inhibited by both analogs at twice their IC_{50} value. These heterocyclic analogs were studied further by members in our group and were found to be cytotoxic to non-small cell lung cancer cells (down to 15.7μM) as well as triple negative breast cancer cells (down to 3.18 μM). They were also additionally able to restore sensitivity of vincristine, an anticancer drug, to a resistant cell line. Finally, molecular dynamic simulations were performed to validate biological data; this showed TR26 binding to ERK complemented our findings with its 25% inhibition *in-vitro*. Ultimately, these heterocyclic analogs provide new insights into pharmacophores useful for future synthesis of estrone analogs as potential anticancer options.

3.1 Introduction

Pancreatic Ductal Adenocarcinoma (PDAC) is the fourth leading cause of cancer death in the United States with over 90% of pancreatic cancer cases being a form of PDAC. The overall 5-year survivability rate is one of the lowest in all cancers at around 8%[1,2]. Additionally, pancreatic cancer is expected to grow to be the second deadliest cancer by the year 2030[3] due to the lack of a true reliable treatment option. Tumors develop in the pancreatic ducts responsible for carrying digestive enzymes, most often in the head of the pancreas (Figure 3.1). The low survivability rate is primarily due to the difficulty with diagnosing the disease. By the time the disease is typically diagnosed it is often too late, and the cancer has spread to other parts of the body, owing to its aggressive and evasive nature[4]. The disease itself is divided into four stages labeled from stage I to stage IV used by the American Joint Committee on Cancer. This system is traditionally known as the TNM system which measures the extent of the tumor, the lymph nodes the disease has spread to, as well as any metastatic sites of the disease[5]. Stage I involves tumors remaining within the pancreas, while stage II cancers have begun to spread to 3 or less nearby lymph nodes, but with no metastasis to other organs. Stage III pancreatic cancer is used for sufficiently large tumors that have spread to more than 4 lymph nodes, but without metastatic sites, while stage IV tumors have metastasized.

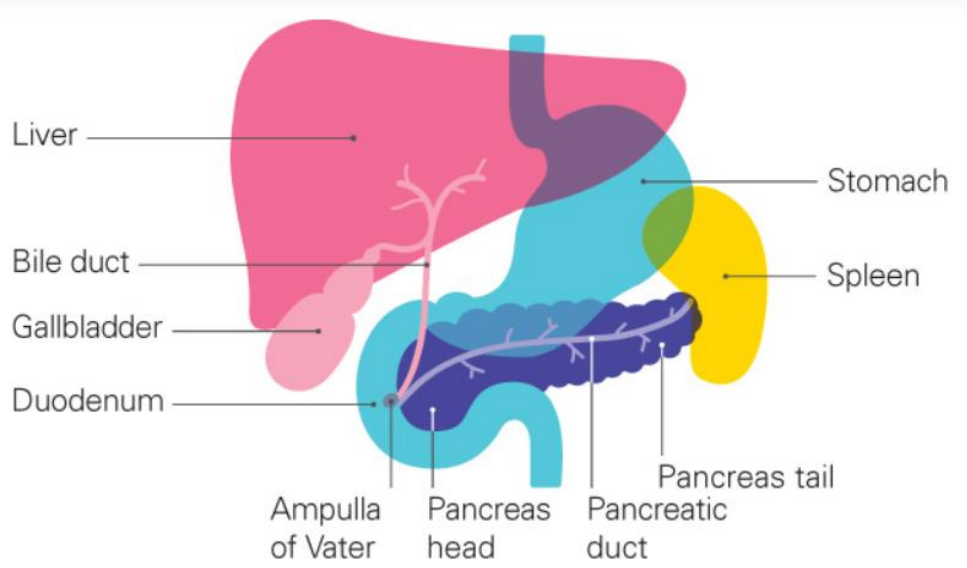


Figure 3.1. Diagram of the pancreas surrounded by adjoining organs. Adapted from pancreaticcancer.org[6].

Surgical resection of the pancreas is preferred in cases where it is an option, typically only in early stages of the disease before it has spread around surrounding veins and surgery risks the life of the patient. It was previously mentioned that most pancreatic cancer occurs in the head of the pancreas (Figure 3.1), and surgical resection of this type is often done with the Whipple procedure (pancreaticoduodenectomy) wherein the head of the pancreas, the gallbladder, and the duodenum are removed, and the stomach and remainder of the pancreas is attached to the jejunum of the small intestine (Figure 3.2). This procedure however typically leads to complications in around half of patients[8] and is often given with neoadjuvant therapy[9]. Cancer in the body or tail of the pancreas is often treated with distal pancreatectomy in which the portion of the distal pancreas containing the cancer, and often the spleen, are removed[10]. This surgery can usually be done laparoscopically, but

complications do still arise in a number of patients. Often however the cancer itself when detected has spread to other parts of the surrounding anatomy or metastasized, leading to surgical resection not being possible.

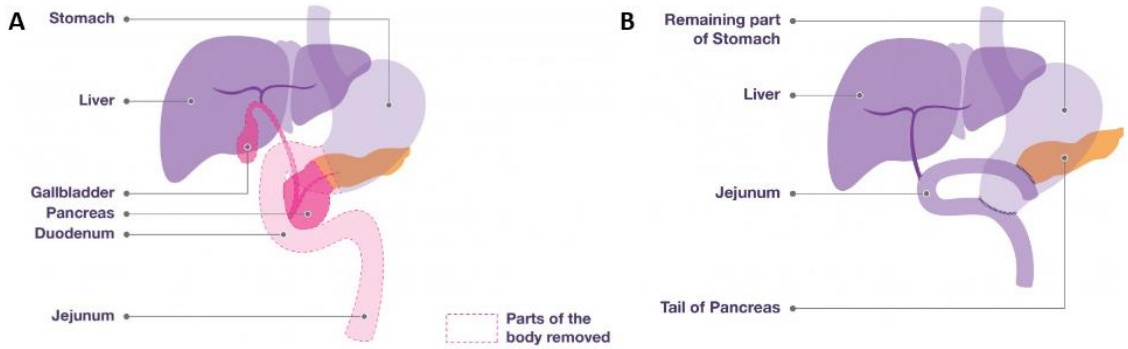


Figure 3.2. A Whipple procedure (pancreaticoduodenectomy) in which A) the highlighted portions of the pancreas, duodenum, and gallbladder are removed and B) the reattachment of organs after surgery. Adapted from pancreaticcanceraction.org[7].

Radiation therapy for PDAC is used, but typically not as a sole therapy. In most cases it is used as a neoadjuvant or adjuvant therapy with resection or chemotherapies[11]. Use with surgical resection is generally performed if it is determined that the radiation may shrink the tumor to a size that can be resected easier. Use with chemotherapy is much more common; but in recent studies is used less often due to low beneficial effects[11]. As such the standard line of treatment for PDAC exists as chemotherapy. A number of chemotherapies exist for PDAC, including FOLFIRINOX (Folinic acid, 5-fluorouracil, Irinotecan, and Oxaliplatin), Gemcitabine, and 5-fluorouracil[12] (**Figure 3.3**). In all cases, the aforementioned therapies act primarily against DNA replication and synthesis. 5-

fluorouracil, an anticancer drug used in multiple cancers, has been around for 60+ years. It acts on the ability for DNA synthesis to produce thymine, and lack of thymine results in cell death. This therapy is still used for pancreatic cancer, albeit more commonly in combination therapies due to developed drug resistance[13]. An alternative is Gemcitabine, a cytosine ribonucleotide mimic, which is still used today in chemotherapy regimens. Gemcitabine acts as cytosine in DNA synthesis, and halts DNA replication due to masked chain termination[14], a process in which proofreading enzymes are unable to remove Gemcitabine and the strand cannot be synthesized. Numerous reports exist for survivability studies with Gemcitabine use. In most cases these survivability's are less than a year in length[15,16]. Most treatments today include Gemcitabine treatment in combination with 5-fluorouracil or paclitaxel. FOLFIRINOX is primarily used for advanced pancreatic cancer treatment today as a combination therapy for metastatic pancreatic cancer. A phase three clinic trial found that median survivability rates of FOLFIRINOX improved to 11.1 months compared to 6.8 months with Gemcitabine[16]. It has been found that FOLFIRINOX does include more adverse effects as opposed to other treatments, suggesting a more permanent treatment option is needed.

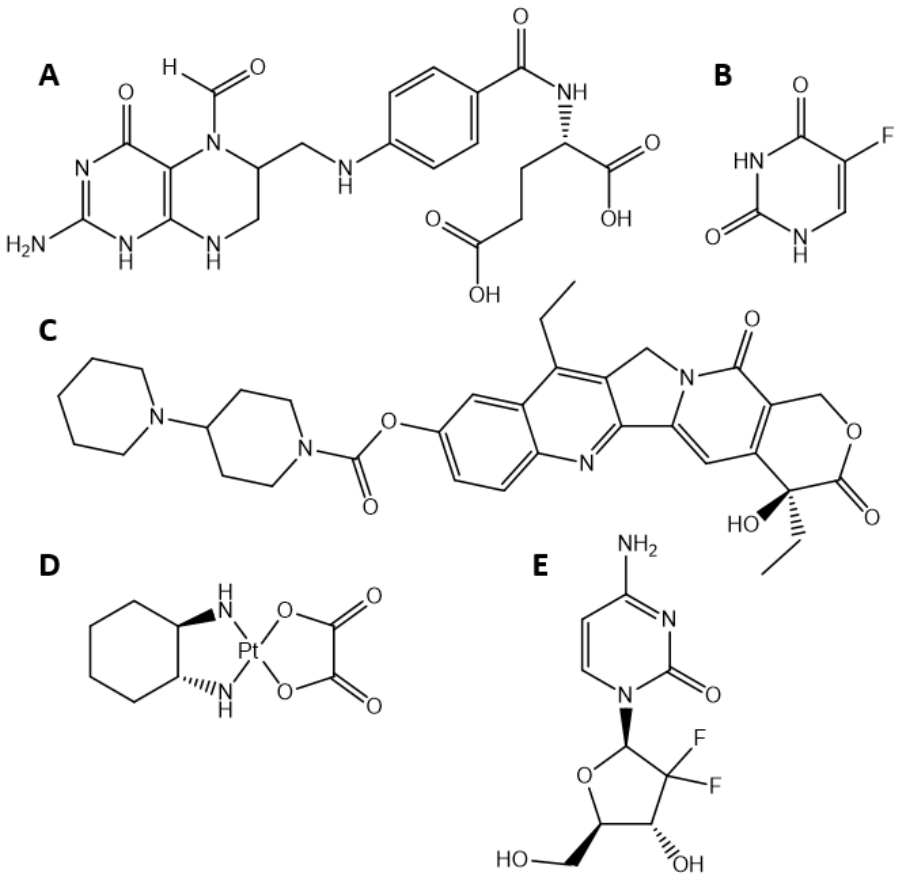


Figure 3.3 Structures of pancreatic chemotherapies. FOLFIRINOX (mixture of A-D): A) Folinic acid, B) 5-Flurouracil, C) Irinotecan, D) Oxaliplatin. E) Gemcitabine. B) 5-Fluorouracil (as a standalone).

A new therapy for pancreatic cancer is needed, either as a standalone or combination therapy. One such set of compounds focuses on combating the Epidermal Growth Factor Receptor (EGFR), a protein tyrosine kinase responsible for cellular functions that is commonly overexpressed in many cancers. Inhibition with tyrosine kinase inhibitors is a pleasing choice due to EGFR's role in pancreatic cancer[17,18]. Many downstream proteins in the EGFR cascade are also responsible for tumor growth and maintenance, and as such

are common targets for cancer therapy. Multiple studies from our laboratory have determined that estrone analogs with an α-hydroxy methyl ketone moiety at the C-17 position possess key anti-cancer properties associated with this pathway among a number of cancers including melanoma, hepatocellular carcinoma, pancreatic cancer, breast cancer, and ovarian cancer[19-23]. Herein we report the synthesis and characterization of a set of heterocyclic estrone analogs and their biological activity towards two pancreatic cancer cell lines, AsPC-1 and Panc-1.

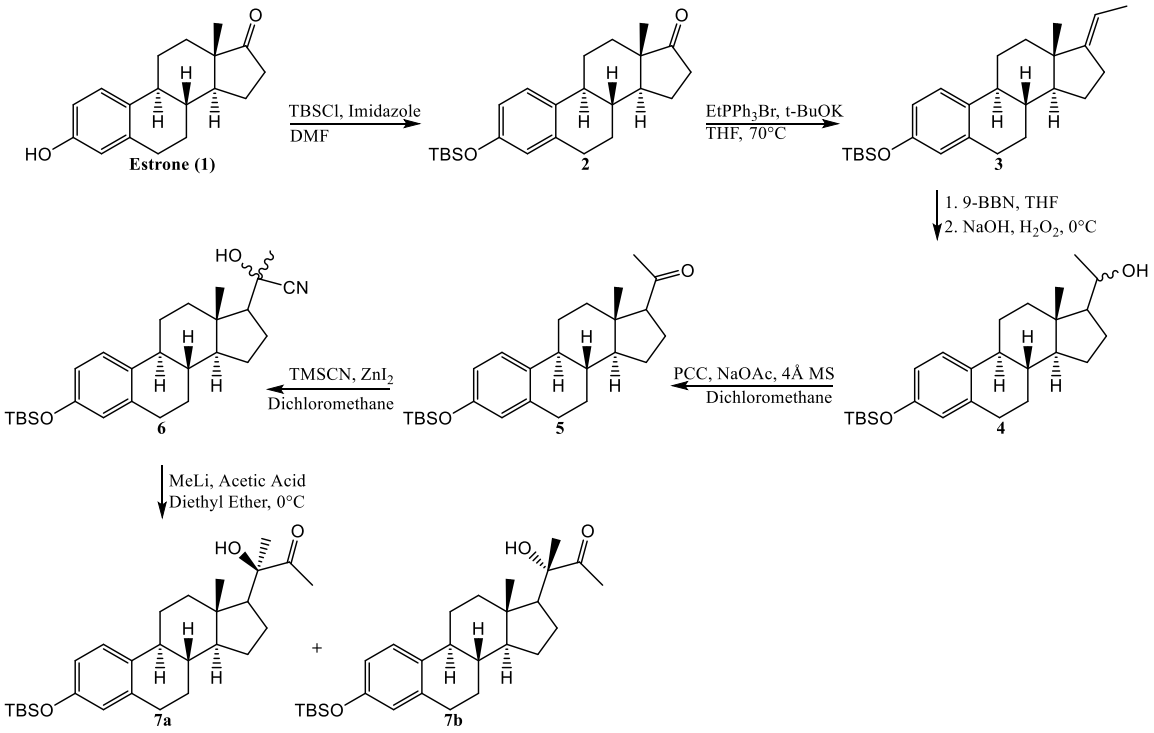
3.2 Results and Discussion

3.2.1 Synthetic Strategy

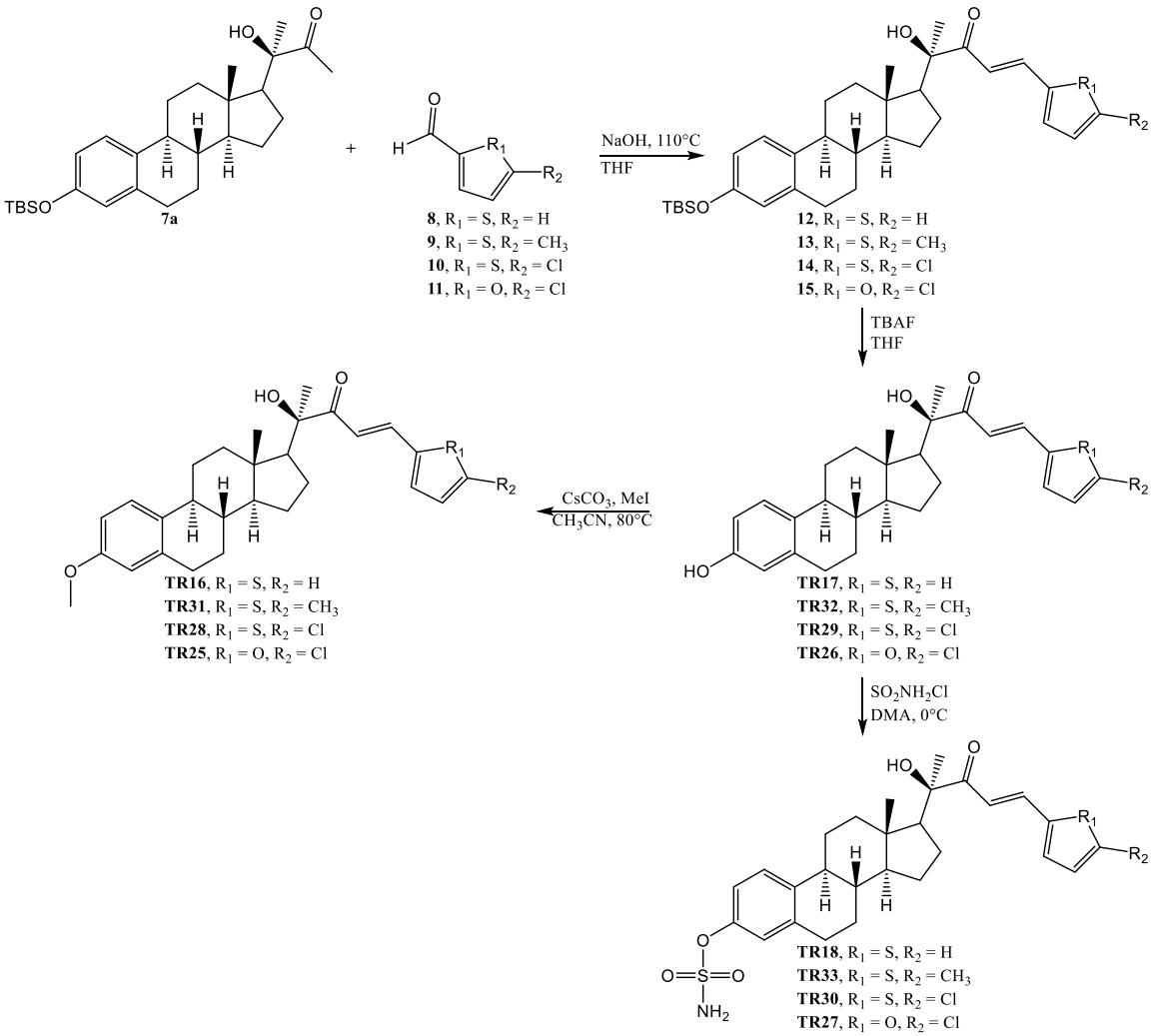
We have previously reported the synthesis of an α-hydroxyl methyl ketone[19-21] simulating that of the Cucurbitacin structure containing an α,β unsaturated side chain[27]. The scheme was modified to create new analogs containing heterocyclic furan and thiophene side chains via aldol condensation. Modification was also performed at C3 to vary polarity of the compound in hope to obtain analogs with increased cytotoxicity due to an increase in hydrophilicity (**Scheme 3.1**).

Estrone (**1**) is modified in six synthetic steps to achieve α-hydroxyl methyl ketone (**7a**) as previously reported[21]. The R-enantiomer is obtained in roughly a 3:1 yield over the S-enantiomer due to the nearby methyl group attached to C13. In an aldol condensation reaction, various aldehydes are reacted with ketone (**7a**) in addition with sodium hydroxide in the presence of THF for 20 minutes in a vial at 110℃ to form α,β enones (**12-15**). Additional THF was added to the reaction to avoid drying during reflux. Previous reports from our laboratory have utilized lithium diisopropyl amide for this aldol condensation due

to its selectivity. However, due to the single alpha-proton used in the aldol reaction, a weaker base can be used along with more favorable reaction conditions to achieve the same result. The reaction is also completed within 20 minutes as opposed to overnight conditions. TBAF was added to enones (**12-15**) in THF to remove protecting group and produce four phenolic estrone analogs (**TR17, TR26, TR29, TR32)**. Deprotected enone was reacted with cesium carbonate and iodomethane in acetonitrile to afford four methoxy species (**TR16, TR25, TR28, TR31**). Finally, four sulfonamide species are formed (**TR18, TR27, TR30, TR33**) when reacted with sulfonyl chloride in the presence of DMA and deprotected enones (**Scheme 3.2**). Due to the hygroscopic nature of sulfonyl chloride, this reagent should be weighed, and the reaction performed, under an inert atmosphere to obtain greater yields.



Scheme 3.1. Synthetic route for α-hydroxyl methyl ketone.



Scheme 3.2. Synthesis of Analogs TR16-TR18 and TR25-TR33.

3.2.2 Biological Evaluation of Heterocyclic Analogs in Pancreatic Cancer Cell Lines AsPC-1 and Panc-1

The 12 synthesized compounds were screened for cytotoxicity against two pancreatic cancer cell lines: AsPC-1 and Panc-1 (**Table 3.1**). Interestingly, no common theme between analogs is seen among the two cell lines used in this study. This could be due to

altered mechanisms of action among the analogs themselves, but generally we find that more polar species, those with phenolic or sulfonamide C3 substituents and heterocyclic chlorine side chains, are more apt to show cytotoxicity.

Table 3.1 IC_{50} results of synthesized heterocyclic estrone in AsPC-1 and Panc-1 Cells by MTT Cell Viability Assay

Compound	AsPC-1 IC_{50} (μM)	Panc-1 IC_{50} (μM)
TR16	>50	38.16 ± 3.28
TR17	>50	41.70 ± 2.49
TR18	48.40 ± 0.45	52.96 ± 5.29
TR25	40.17 ± 0.46	>50
TR26	>50	10.16 ± 0.83
TR27	>50	>50
TR28	>50	33.35
TR29	>50	39.80
TR30	>50	>50
TR31	>50	>50
TR32	46.50 ± 0.87	>50
TR33	38.37 ± 1.76	>50
Gemcitabine	>50	>50

To further investigate the mechanism of action of these analogs, we examined the most effective analogs in both cell lines (TR33 for AsPC-1 and TR26 for Panc-1) and performed

cell cycle analysis to investigate how these analogs affect the cell cycle. The analogs performance on different parts of the cell cycle can help determine mechanistic properties of the analog and how it is interacting with various proteins. For example, G0/G1 phase arrest can sometimes be attributed to a deficit of growth factors such as Akt and mTOR as well as inactivation of cyclin dependent kinases D and E that control cell transcription[24,25]. G2/M phase inhibition is more typical of DNA damage to the cell, wherein cyclin B kinases are responsible for cell cycle checkpoints to allow for mitosis[26]. **Figure 3.4** outlines the results for flow cytometry analysis of TR33 in AsPC-1 cells. We see a minor G2 phase inhibition but not significant. However, when we look at TR26 in Panc-1 we see a significant G1 phase inhibition at both the IC_{50} concentration and twice the IC_{50} concentration (**Figure 3.5**). This indicates that growth factors likely play a role in the inhibitory function of TR33, and is explored further using the in-cell western assay.

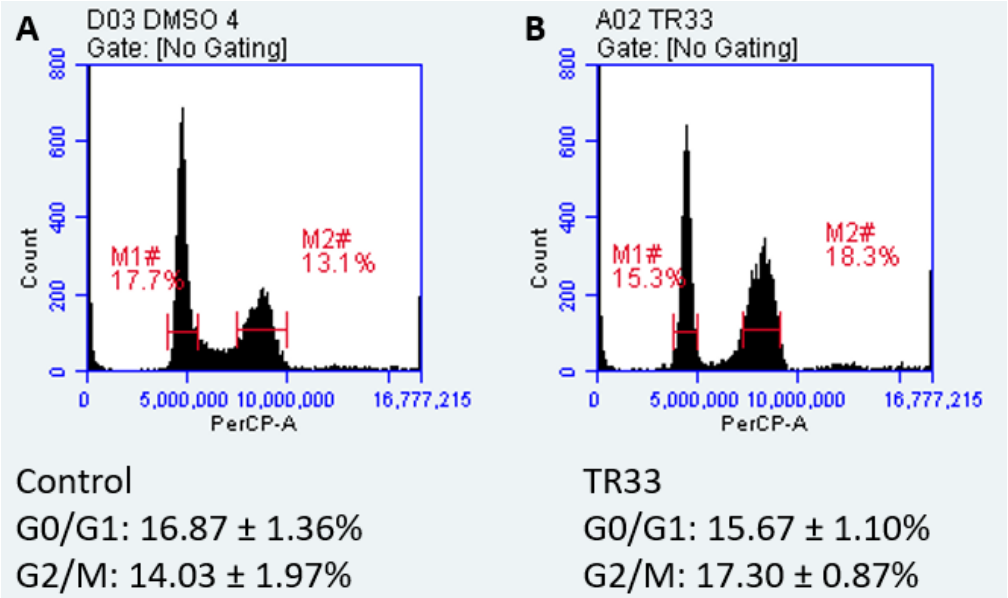


Figure 3.4 Cell cycle analysis histograms detailing A) DMSO Control and B) TR33 for AsPC-1 cell line.

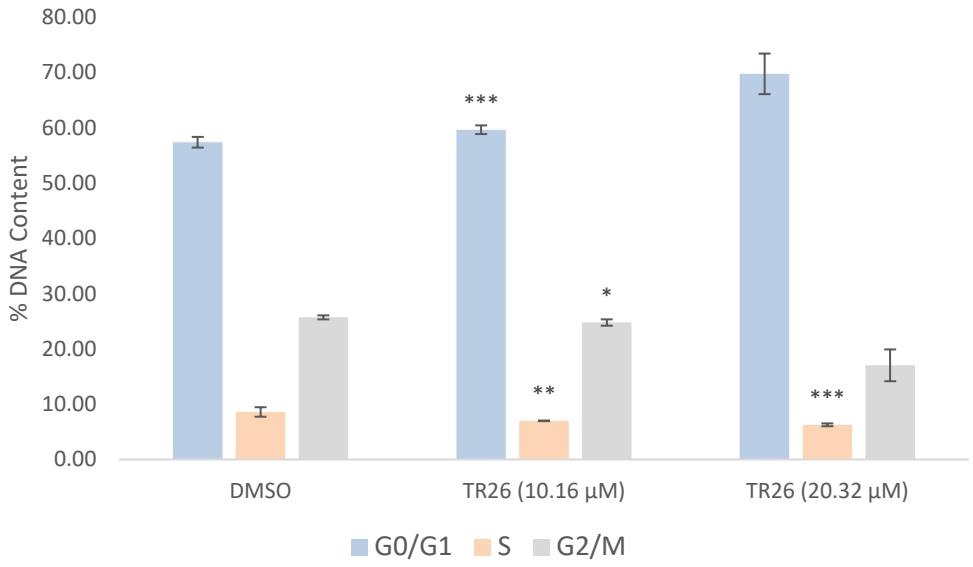


Figure 3.5 Bar graph detailing numerical cell cycle analysis data for TR26 in Panc-1 cell line. Statistical analysis was performed with a Welch's T-test. Levels of Significance: * $p < 0.05$, ** $p < 0.025$, *** $p < 0.01$, **** $p < 0.005$.

To explore the mechanistic effects of TR26 and TR33 we utilized in in-cell western protein analysis assay to determine if certain growth factor proteins were inhibited by either analog. One protein from each of the three major EGFR downstream signaling pathways (mTOR, ERK, and STAT3) and EGFR itself (see Figure 1.4) were utilized for this analysis. These proteins were chosen based on inhibition results from previous analogs in our laboratory[20,22,23]. TR26 and TR33 both show significant inhibitions at twice their IC_{50} values towards ERK (**Figure 3.6** and **Figure 3.7**). Additional inhibition also occurs for TR26 towards EGFR and for TR33 towards mTOR. The latter is a serine/threonine protein kinase which regulates cell growth, cell proliferation, cell motility, cell survival, protein

synthesis and transcription. These results suggest that analogs of this nature can be lead drug candidates for potential treatment of pancreatic cancer.

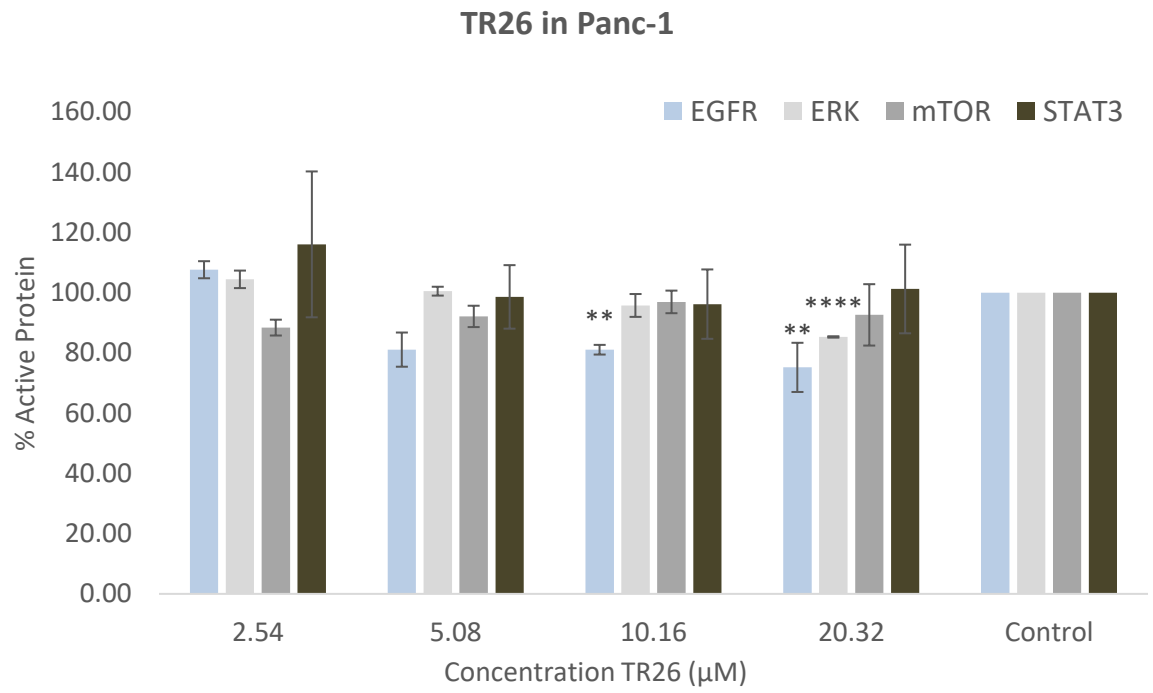


Figure 3.6 In-cell western data for TR26 in Panc-1 cell line expressed as a percentage of the active forms of Epidermal Growth Factor Receptor (EGFR), Extracellular Signal-Related Kinase (ERK), mechanistic Target of Rapamycin (mTOR), and Signal Transducer and Activator of Transcription 3 (STAT3). Statistical analysis was performed with a student's t-test. Levels of Significance: * $p < 0.05$, ** $p < 0.025$, *** $p < 0.01$, **** $p < 0.005$.

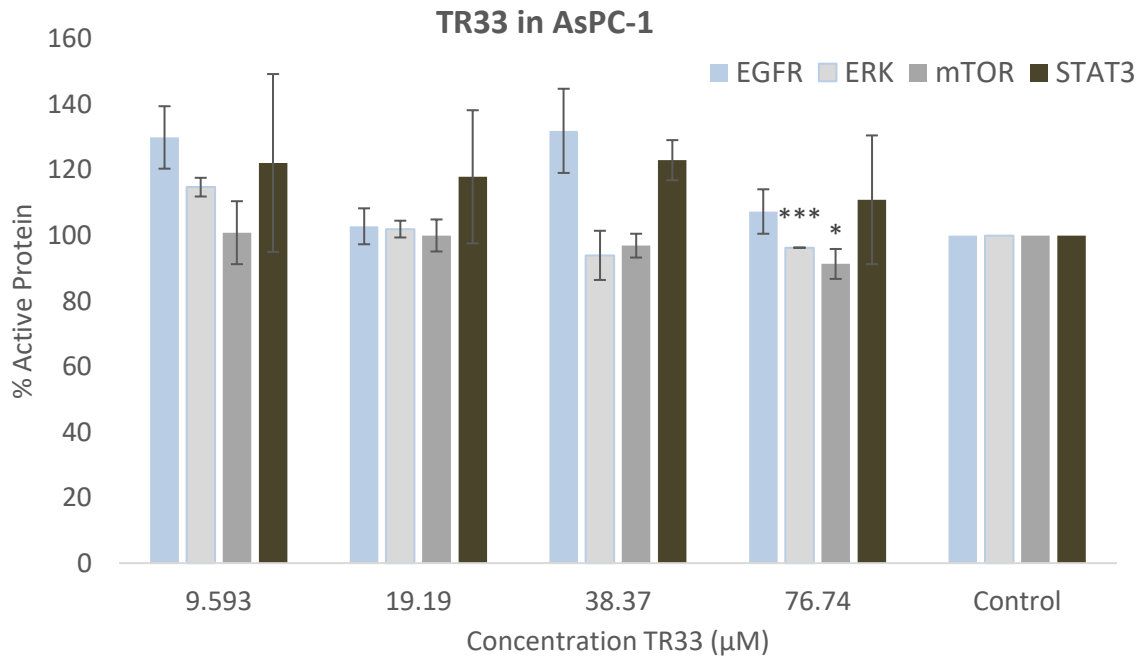


Figure 3.7 In-cell western data for TR33 in AsPC-1 cell line expressed as a percentage of the active forms of Epidermal Growth Factor Receptor (EGFR), Extracellular Signal-Related Kinase (ERK), mechanistic Target of Rapamycin (mTOR), and Signal Transducer and Activator of Transcription 3 (STAT3). Statistical analysis was performed with a student's t-test. Levels of Significance: * $p < 0.05$, ** $p < 0.025$, *** $p < 0.01$, **** $p < 0.005$.

3.2.3 Molecular Dynamic Simulations

To confirm biological activities of lead compounds, we investigated the roles of TR26 and TR33 binding to specific molecular targets proteins identified earlier in our in-cell western analysis, we explored *in-silico* molecular dynamic simulations to model how TR26 and TR33 bind to inhibited proteins seen in the in-cell western assay. TR33 shows

significant inhibition (up to 10%) at twice its IC_{50} value for ERK and mTOR. For ERK, we found that TR33 moves in the pocket throughout the simulation, leading to little contact (less than half the simulation) and high RMSD calculations (up to 10Å). In mTOR, we see similar results, as TR33 is not stabilized in the pocket of mTOR during any of the simulation. We have seen these problems before in previous simulations (See chapter 2) in which the OPLS_2005 force field sometimes fails to incorrectly determine torsional and charge parameters for some large proteins. We fear this may have happened to mTOR, as it was a much larger protein than the others tested.

For TR26, we see more compelling data. We first ran TR26 in EGFR, and found that contact was made with both the phenolic group and the C-17 hydroxyl group with amino acids THR-766 and CYS-773 respectively, but these contacts are minimal during the simulation (Figure 3.8). When we explore the RMSD calculations for this binding, we see that EGFR shows a large (15Å) RMSD value for the simulation (Figure 3.9). This is not unexpected, as EGFR contains a binding arm for its dimerization in amino acids 270-300 that is very flexible. The movement of the protein then suggests as to why binding in the simulation is not as good as we hoped. For TR26 in ERK we see high binding during the simulation for again the phenolic group and the C-17 hydroxyl group. Binding occurs for more than 90% of the simulation in case of the phenolic group to amino acids GLN-103 and ASP-104, while the C-17 hydroxyl group binds during 47% of the simulation to GLU-31 (Figure 3.10).

It is worth discussing as to why TR26 has such a low IC_{50} value but binding to EGFR and ERK proteins remains low, despite good binding parameters by the dynamic simulations. It is likely that TR26's cytotoxicity is due to a combinatorial effect, that is,

binding to multiple proteins or affecting multiple proteins at the same time, and is something that cannot be modeled using simulations and tested with single protein studies. This warrants future study of course, and we hope to explore that more soon.

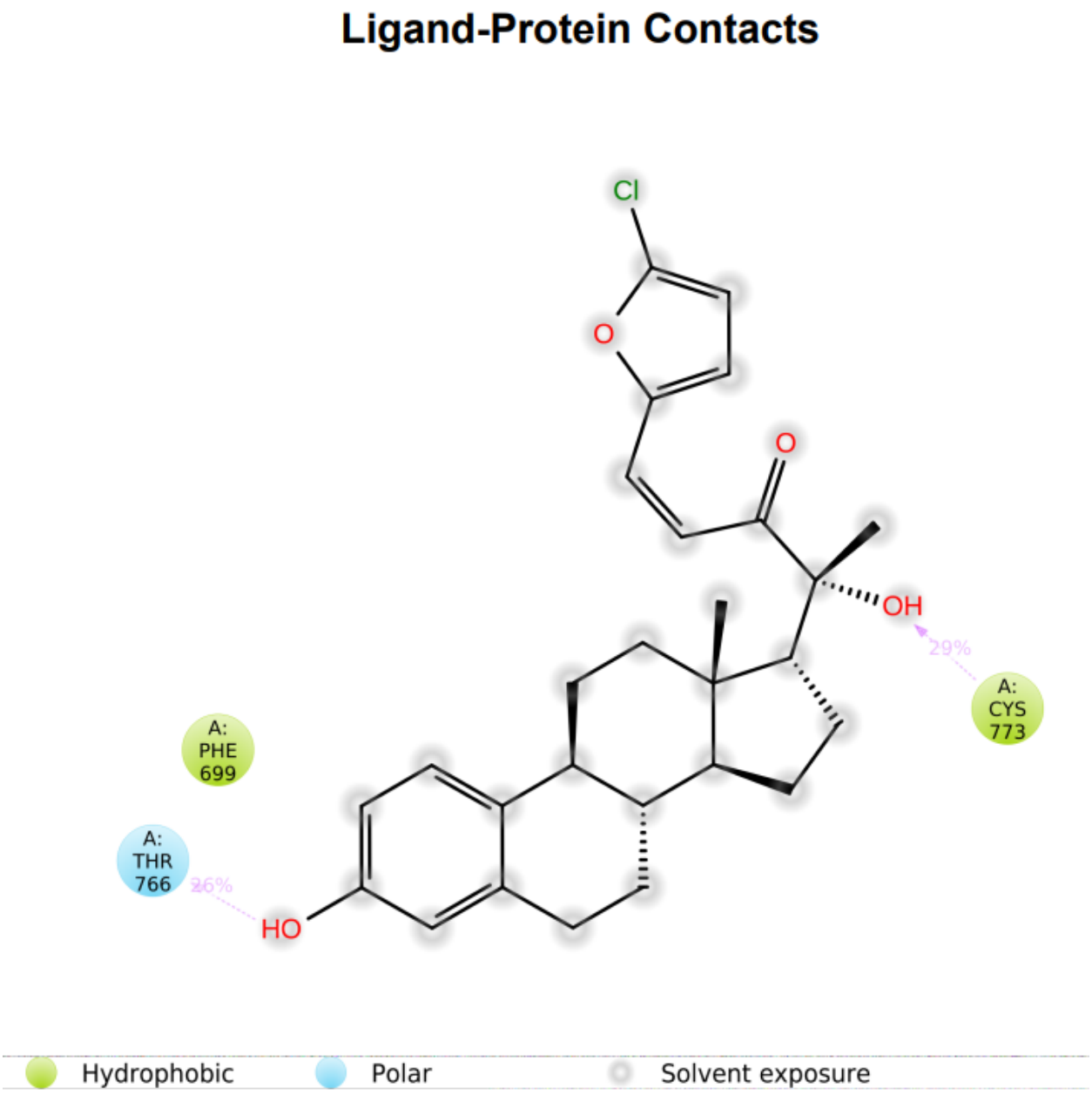


Figure 3.8 TR26 binding interactions to amino acids in EGFR binding pocket. Percentages are indicated as amount of binding during the whole of the simulation time.

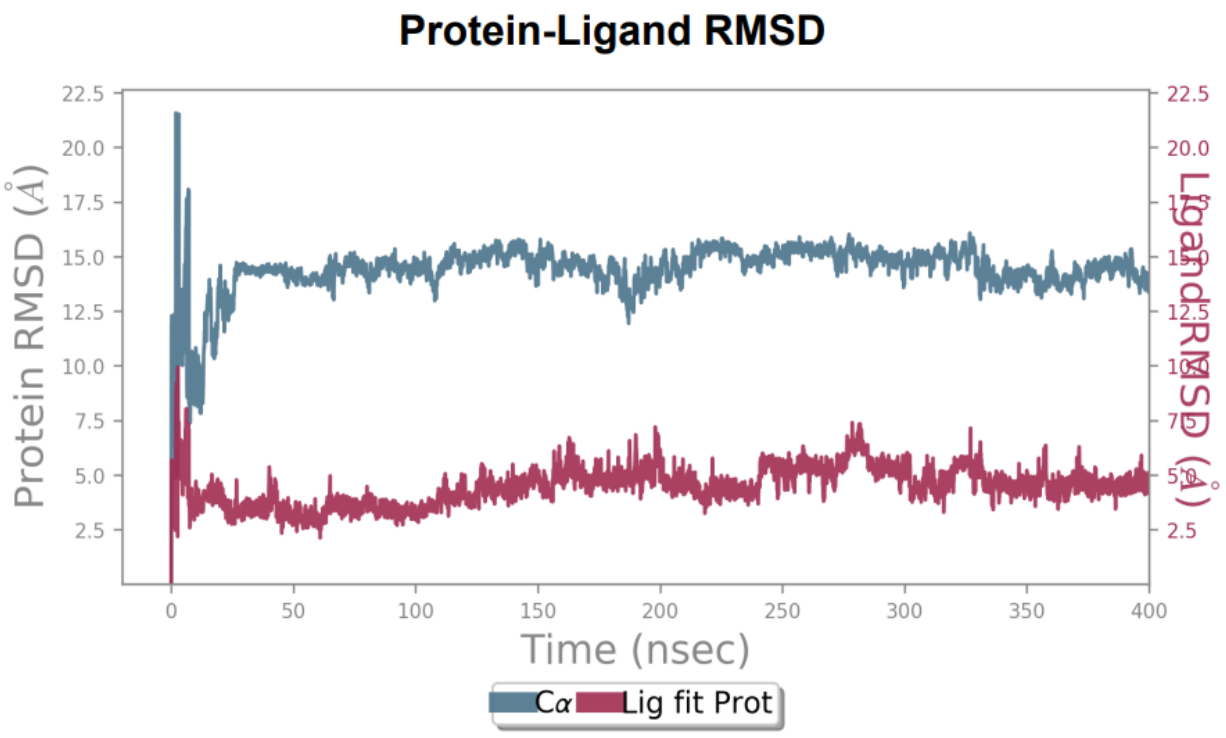


Figure 3.9. RMSD calculations for TR26 binding to EGFR. Blue shows the protein RMSD calculations while red shows the ligand RMSD calculations. RMSD measures the movement of atoms in respect to the reference frame zero.

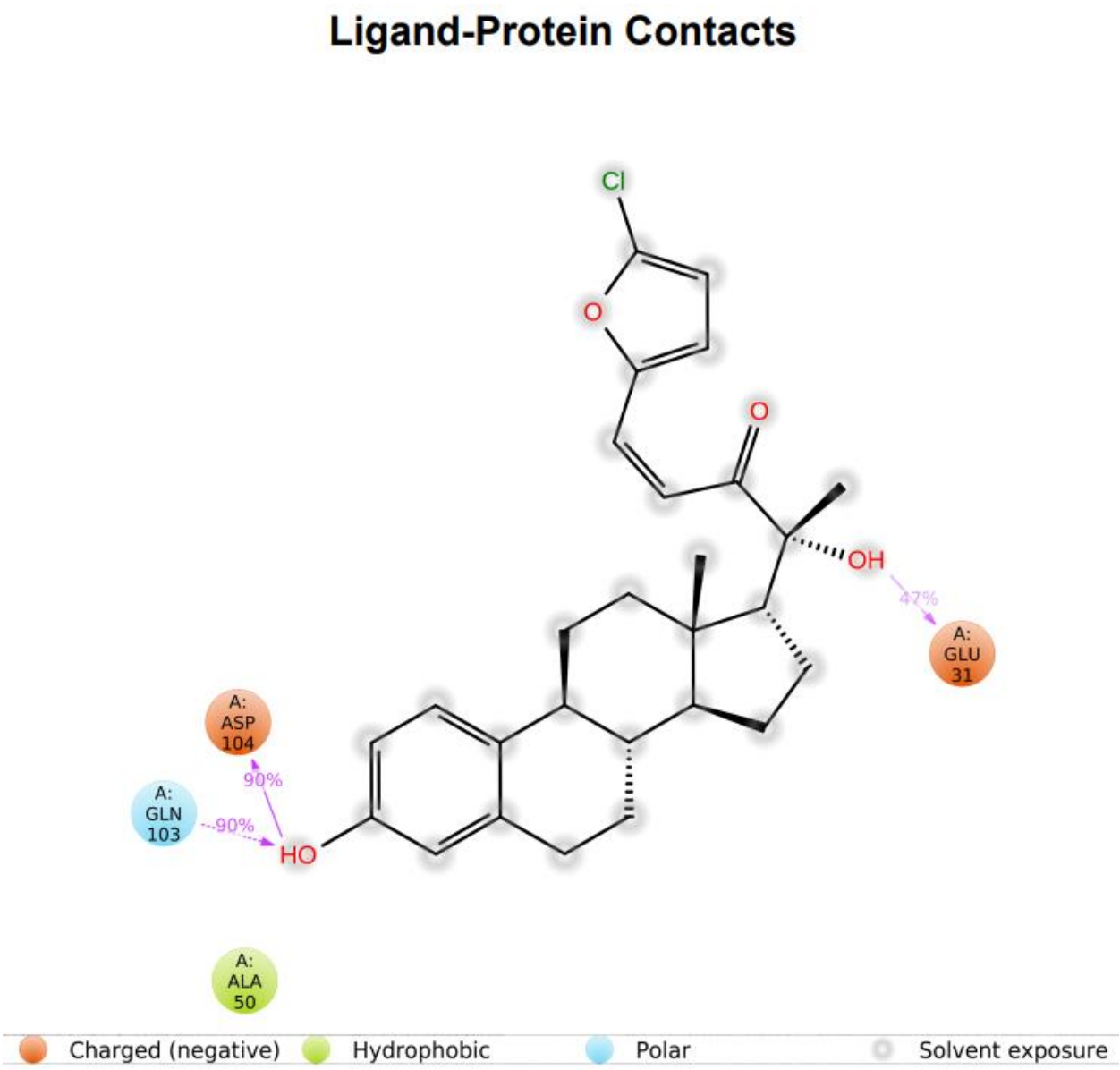


Figure 3.10 Protein-ligand contacts for TR26 in the ERK active site. Percentages are indicated as amount of binding during the whole of the simulation time.

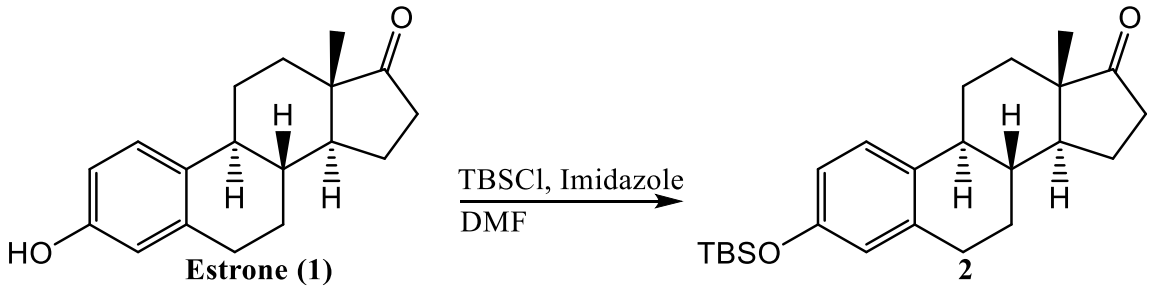
3.3 Materials and Methods

3.3.1 Chemical Synthesis

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 an over overnight before conducting chemical reactions. All chemical reactions requiring anhydrous environments utilized nitrogen gas applied at the reaction time. 1H and 13C-NMR spectra were acquired on a Bruker AVANCE-600 MHz NMR spectrometer in deuterated solvents such as CDCl3, Acetone-d6, or DMSO-d6 using solvent residual peak as the internal standard. Reporting of coupling constants is in Hz and the signal multiplicities are reported as singlet (s), doublet (d), triplet (t), quartet (q), doublet of doublets (dd), doublet of triplets (dt), triplet of doublets (td), triplet of triplets (tt), doublet of quartets (dq), quartet of triplets (qt), doublet of doublet of doublets (ddd), doublet of doublet of triplets (ddt), doublet of doublet of doublet of doublets (dddd), doublet of doublet of doublet of doublet of doublets (ddddd), multiplet (m), or broad (br). TLC analysis was performed using precoated silica gel PE sheets. Products were purified via column chromatography and using silica gel 40–63 μm (230–400 mesh), normal phase preparative TLC plates. TLC plates were visualized by ultraviolet at 254 nm. TLC plates were stained by Iodine, Vanillin, and Ceric Ammonium Molybdate (CAM) stain.

3-OTBS Protected Estrone (2)



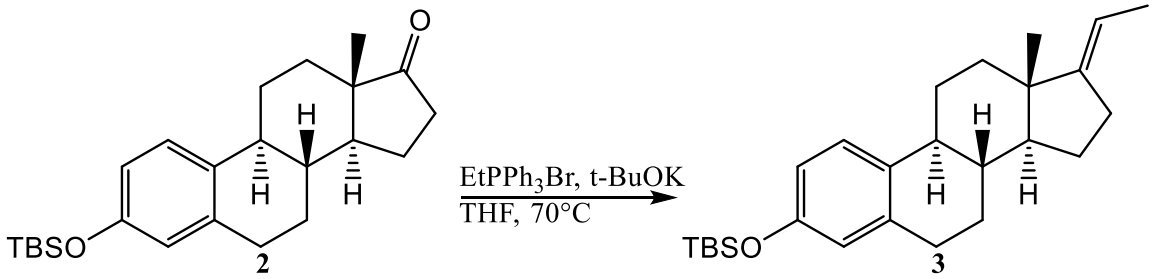
Estrone (1) **2**

5.0 grams (18.5 mmol) of Estrone was dissolved in 70 mL of Dimethylformamide in a 500 mL round bottom flask at room temperature. After dissolution, 3.46 grams (50.875 mmol) was added to the flask. Finally, 4.16 grams (27.75 mmol) of tert-Butyldimethylsilyl chloride was added for protection of compound (1) overnight. 75 mL of distilled water was then added to remove the Dimethylformamide from the mixture. Extraction of the reaction mixture was then performed three times using 75 mL aliquots of ethyl acetate. The organic fractions were collected, dried over anhydrous sodium sulfate, and concentrated by rotary evaporator. The crude product was purified using column chromatography with a mobile phase mixture of 7:3 hexane: ethyl acetate to yield compound 3-OTBS protected Estrone (2) in 92% yield.

^{1}H 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, CDCl3) 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.

Alkene (3)

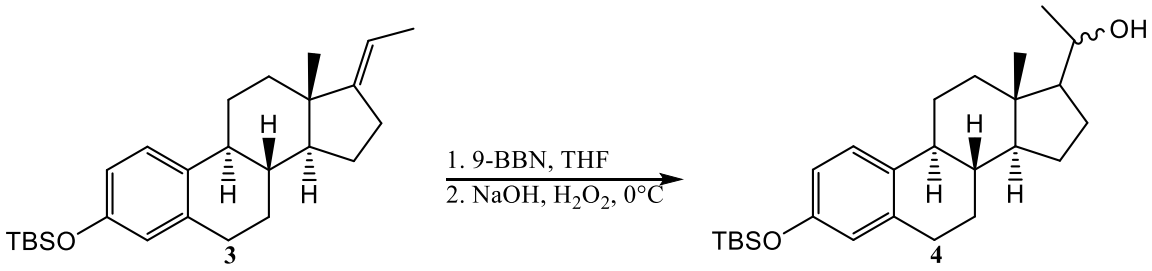


To form the ylide, Ethyltriphenylphosphonium bromide (3 eq) is first dissolved in 0.5 M tetrahydrofuran (1 molar equivalent) in a 500 mL round bottom flask. Potassium tert-butoxide was used as the strong base in 2.7 equivalents to the starting material. It was added in 2 portions slowly to the reaction mixture under nitrogen gas. Upon addition the solution changed to a bright orange color. The contents of the flask were left to mix at room temperature for one hour. 6.57 grams of compound (1) was dissolved in 77 mL of THF and upon dissolution the mixture was added to the ylide. This reaction was left to reflux in an oil bath at approximately 70°C for 5 hours. After completion, the reaction mixture was quenched with ammonium chloride and extracted three times with 75 mL aliquots of ethyl acetate. The organic layers were collected, dried over anhydrous sodium sulfate, and concentrated under vacuo. Column chromatography was performed to purify the product in a 9:1 mobile phase mixture of hexane: ethyl acetate to obtain pure Alkene (3) in 72% yield.

^{1}H NMR (400 MHz, Chloroform-d) δ 7.19 (dd, J = 8.5, 1.1 Hz, 1H), 6.68 (dd, J = 8.4, 2.6 Hz, 1H), 6.62 (d, J = 2.6 Hz, 1H), 5.22 (dddd, J = 9.2, 7.1, 4.6, 2.1 Hz, 1H), 2.88 (dq, J = 6.8, 4.3, 2.7 Hz, 2H), 2.52 – 2.25 (m, 6H), 2.01 – 1.94 (m, 1H), 1.85 – 1.73 (m, 6H), 1.67 – 1.58 (m, 2H), 1.53 – 1.31 (m, 7H), 1.05 (d, J = 2.7 Hz, 10H), 0.98 (s, 3H), 0.26 (s, 6H).

^{13}C NMR (101 MHz, CDCl3) δ 153.32, 150.32, 137.88, 133.31, 126.10, 119.98, 117.16, 113.43, 55.31, 44.62, 43.92, 38.38, 37.31, 31.50, 29.76, 27.64, 26.95, 25.78, 24.22, 18.23, 17.02, 13.22, -4.32.

Alcohol (4)

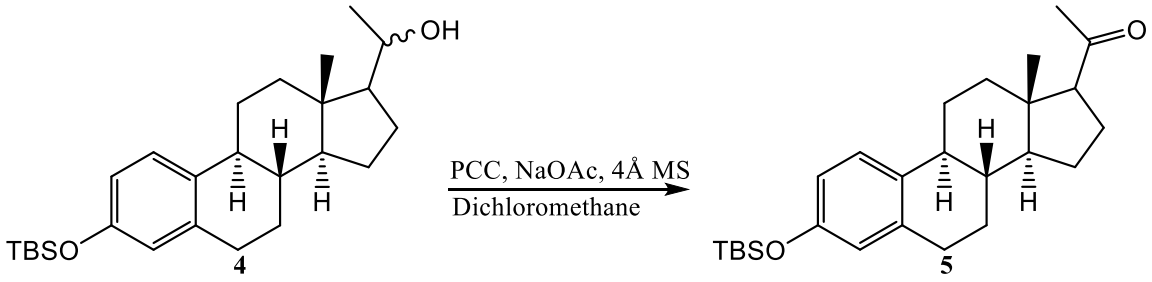


9.19 grams (23.16 mmol) of compound 3 was added to a 1 L round bottom flask and held under nitrogen gas. 3.88 equivalents of 9-Borabicyclo [3.3.1] nonane was administered to the flask dropwise, and the reaction mixture was left to stir at room temperature for 18 hours. After 18 hours the flask was removed and cooled in an ice bath. 3.88 equivalents of a 10% sodium hydroxide solution was added slowly to the uncovered flask. After addition, 3.88 equivalents of cold 30% hydrogen peroxide was also added slowly to the uncovered flask. The reaction contents were then left to mix for one hour without being covered. The crude material was then extracted three times with 75 mL aliquots of ethyl acetate. The organic fractions were collected, dried by anhydrous sodium sulfate, and concentrated under vacuo. A column was performed in a mobile phase mixture of 8:2 hexane: ethyl acetate to obtain diastereomers of alcohol (4) in 78% yield.

^{1}H NMR (400 MHz, Chloroform-d) δ 6.92 (dd, J = 8.5, 1.0 Hz, 1H), 6.42 (dd, J = 8.4, 2.7 Hz, 1H), 6.36 (d, J = 2.6 Hz, 1H), 3.65 (tt, J = 8.0, 3.8 Hz, 1H), 3.53 (dt, J = 8.3, 6.1 Hz, 1H), 2.65 – 2.59 (m, 2H), 1.08 (t, J = 5.3 Hz, 8H), 0.79 (s, 12H), 0.73 – 0.65 (m, 5H), 0.51 (s, 3H), 0.24 (s, 6H).

^{13}C NMR (101 MHz, CDCl3) δ 153.28, 137.86, 133.22, 126.05, 119.96, 117.12, 72.15, 58.62, 55.36, 43.79, 42.08, 39.03, 38.39, 31.64, 27.45, 25.74, 22.70, 18.19, 14.17, 12.63, -4.37.

Ketone (5)



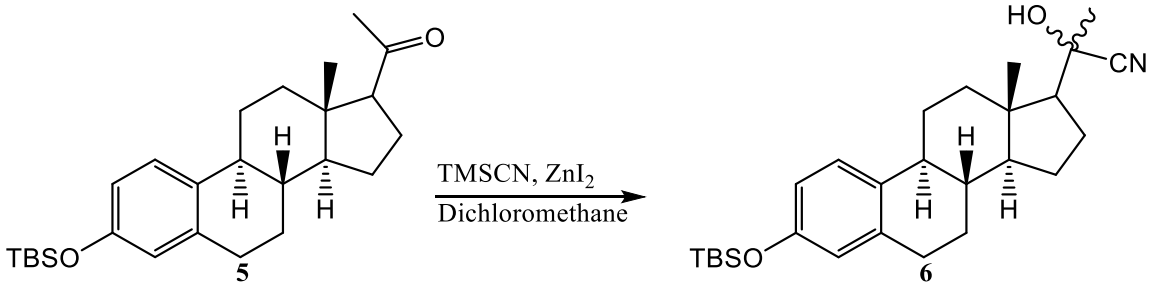
7.50 grams of mixed isomer alcohol (4) was added to a 1 L round bottom flask and dissolved in 0.2 M dichloromethane (1 Molar equivalent). 4Å Molecular sieves were then added to the flask to ensure complete anhydrous conditions of the solution. The weight of sieves needed was equal to the mass needed of the sodium acetate (4.8 equivalents). After addition of molecular sieves, the sodium acetate itself was added to the flask. Finally, the flask was put on ice, and 2 equivalents of pyridinium chlorochromate were added. The reaction was stirred at room temperature for four hours. The reaction mixture was purified

using a dry silica column with a mobile phase ratio of 9:1 hexane: ethyl acetate to yield ketone (5) in 67% yield.

^{1}H NMR (400 MHz, Chloroform-d) δ 6.93 (dd, J = 8.6, 1.1 Hz, 1H), 6.43 (dd, J = 8.5, 2.6 Hz, 1H), 6.37 (d, J = 2.6 Hz, 1H), 3.93 (q, J = 7.1 Hz, 4H), 1.97 (d, J = 5.5 Hz, 4H), 1.86 (s, 6H), 1.07 (t, J = 7.1 Hz, 7H), 0.80 (d, J = 3.4 Hz, 12H), 0.47 (s, 3H), 0.24 (s, 6H).

^{13}C NMR (101 MHz, CDCl3) δ 209.58, 153.34, 137.76, 132.78, 126.03, 119.97, 117.17, 63.84, 60.39, 55.70, 44.42, 41.92, 38.67, 31.50, 27.15, 25.64, 21.02, 18.16, 14.18, 13.43, -4.41.

Nitrile (6)



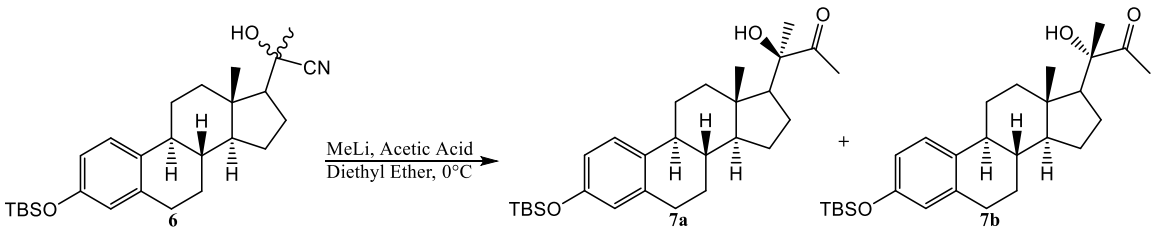
5.0 grams (12.1 mmol) of ketone (5) was dissolved in 24.2 mL of 0.5 M dichloromethane in a 500 mL round bottom flask. 0.116 grams (0.36 mmol) of zinc iodide catalyst was quickly added upon dissolution of starting material. Finally, 1.97 mL (15.74 mmol) of trimethylsilyl cyanide was injected via syringe into the flask in a closed fume hood. The reaction was run for 3 hours. Water was added to the crude mixture followed by extraction with 3x 50 mL aliquots of ethyl acetate. The organic fractions were collected, dried with anhydrous sodium sulfate, and were concentrated in vacuo. The crude product was purified

using an 8:2 mobile phase mixture of hexane: ethyl acetate via column chromatography to yield nitrile (6) in 81% yield.

^{1}H NMR (400 MHz, Chloroform-d) δ 6.92 (dd, J = 8.5, 1.0 Hz, 1H), 6.42 (dd, J = 8.4, 2.6 Hz, 1H), 6.36 (d, J = 2.7 Hz, 1H), 2.61 (dd, J = 6.7, 3.0 Hz, 1H), 1.86 – 1.81 (m, 2H), 1.43 (d, J = 4.3 Hz, 4H), 0.80 (d, J = 3.0 Hz, 12H), 0.72 – 0.65 (m, 3H), 0.06 (s, 12H).

^{13}C NMR (101 MHz, CDCl3) δ 151.71, 136.18, 131.56, 124.46, 120.58, 118.35, 115.57, 70.83, 58.93, 53.32, 42.18, 36.51, 35.66, 30.04, 29.23, 26.01, 24.15, 21.10, 19.49, 16.60, 12.58, -0.00, -0.23, -5.95.

Hydroxyl Methyl Ketone (7a + 7b)



5.02 grams (9.7 mmol) of nitrile (6) was dissolved in 28.7 mL of diethyl ether (0.34 M) and cooled to 0°C. 18.31 mL (29.3 mmol) of methyllithium was then added dropwise to the reaction mixture. After 2 hours the reaction was again cooled back down to 0°C, and 3.66 grams (63.5 mmol) of glacial acetic acid was introduced to the reaction mixture to promote formation of ketone. The reaction was left for a final 30 minutes. Due to the formation of a chiral center, there are two diastereomers formed with the dextrorotatory product formed in a 3:1 yield. Excess aqueous sodium bicarbonate was used to neutralize any remaining acid, and the mixture was extracted with 3x 50 mL aliquots of

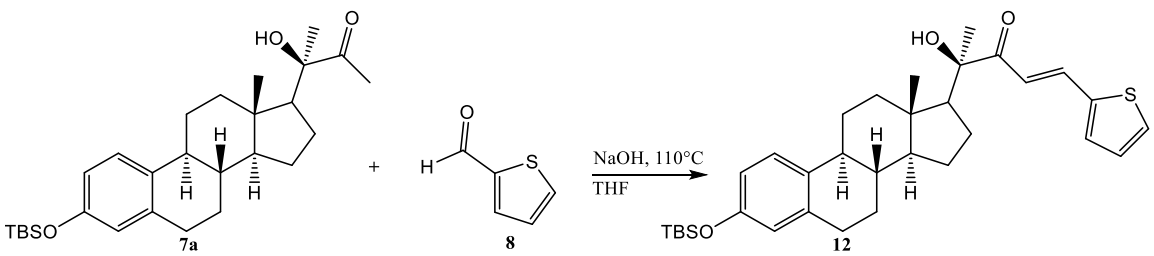
dichloromethane. The organic fractions were dried with anhydrous sodium sulfate and concentrated in vacuo.

Crude product was purified using a long column and a 97:3 mobile phase mixture of hexane to ethyl acetate to begin, with slow addition of a more polar mobile phase until reaching a 95:5 ratio. To verify separation of enantiomers thin layer chromatography was performed 4x each in a 9:1 hexane to ethyl acetate mixture. 600 mg of α-hydroxyl methyl ketone (7a) was collected along with 2.50 grams of mixed enantiomer compound for a yield of compound (7b) in 69.5% yield. This process was repeated until a substantial amount of 7a was collected.

^{1}H NMR (400 MHz, Chloroform-d) δ 6.93 (dd, J = 8.5, 1.2 Hz, 1H), 6.44 – 6.40 (m, 1H), 6.36 (d, J = 2.6 Hz, 1H), 4.96 (qt, J = 7.1, 2.1 Hz, 1H), 2.71 – 2.59 (m, 2H), 1.59 – 1.48 (m, 5H), 0.79 (d, J = 0.8 Hz, 12H), 0.73 – 0.71 (m, 3H), 0.11 (s, 6H).

^{13}C NMR (101 MHz, CDCl3) 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.

Condensation with 2-thiophenecarboxaldehyde (12)

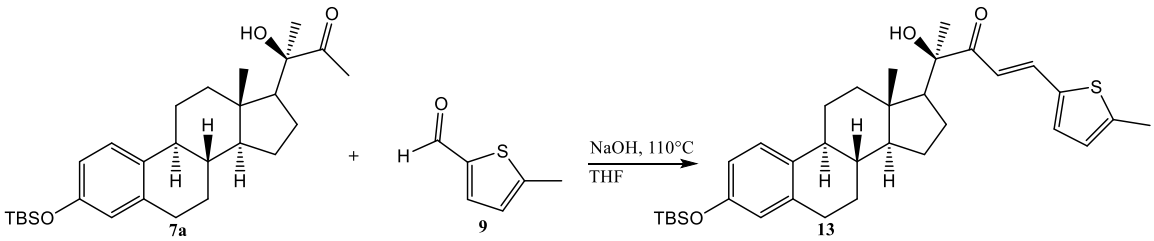


400 mg (0.88 mmol) of compound (7a), 102 mg (2.55 mmol) of sodium hydroxide, 147 mg (1.32 mmol) of 2-thiophenecarboxaldehyde (8), and 1.04 mL (0.85 M) of THF were added to a vial and heated to 110°C for 20 minutes. Additional THF was added as needed to keep reagents in solution during heating. TLC in 8:2 mixture of hexane to ethyl acetate was performed to identify reaction completion. Approximately 25 mL of distilled water was added to the solution to dissolve any remaining sodium hydroxide in the solution. Extraction was performed using 3x 25 mL aliquots of ethyl acetate. The organic layer was collected and dried with anhydrous sodium sulfate. The sodium sulfate was filtered off, and the organic mixture was concentrated in vacuo. The crude product was purified using column chromatography with a mobile phase ratio of 9:1 hexane to ethyl acetate. Compound (12) was achieved in 82.6% yield.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.88 (d, *J* = 15.2 Hz, 1H), 7.37 (d, *J* = 5.1 Hz, 1H), 7.29 (s, 1H), 7.09 – 7.00 (m, 2H), 6.72 (d, *J* = 15.2 Hz, 1H), 6.56 (dd, *J* = 8.4, 2.8 Hz, 1H), 6.48 (d, *J* = 2.7 Hz, 1H), 4.18 (s, 1H), 2.72 (q, *J* = 4.6, 3.3 Hz, 2H), 2.27 – 2.08 (m, 5H), 1.76 (d, *J* = 9.0 Hz, 2H), 1.48 (d, *J* = 3.3 Hz, 4H), 1.24 – 1.15 (m, 8H), 0.88 (s, 3H), 0.05 (s, 6H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 201.70, 153.46, 139.73, 138.25, 138.20, 132.83, 132.66, 129.52, 128.50, 126.44, 117.16, 115.30, 112.69, 55.65, 55.12, 44.32, 43.80, 40.64, 38.12, 29.73, 29.64, 27.57, 26.67, 24.30, 23.65, 22.05, 13.65, -4.37.

Condensation with 5-methyl-2-thiophenecarboxaldehyde (13)

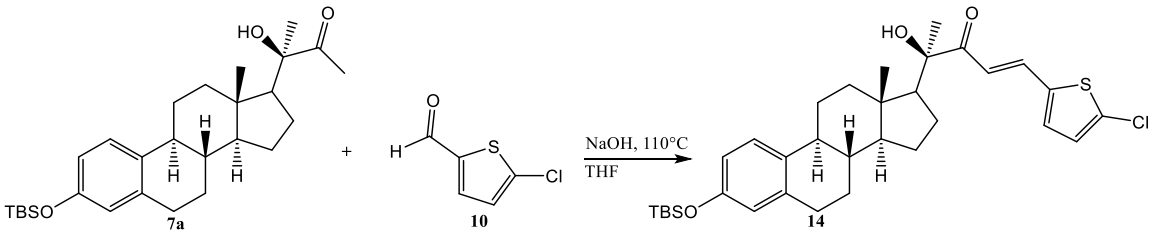


400 mg (0.88 mmol) of compound (7a), 166 mg (1.32 mmol) of 5-methyl-2-thiophenecarboxaldehyde (9), and 102 mg (2.55 mmol) of sodium hydroxide were added to a scintillation vial. 1.04 mL (0.85 M) of THF was added to assist in dissolution. The vial was then heated for approximately 20 minutes on a hotplate at 110°C. Additional THF was added in excess to keep an aqueous mixture while heating. TLC in 8:2 hexane to ethyl acetate confirmed completion of the reaction. 25 mL of distilled water was added to dissolve any remaining base, and the organic layer was extracted with 3x 25 mL aliquots of ethyl acetate. Anhydrous sodium sulfate was added in excess to remove any water from the organic layer and was subsequently filtered off. Crude product was concentrated in vacuo and purified using a mobile phase mixture of 9:1 hexane to ethyl acetate. Compound (13) was achieved in 58.6% yield.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.35 (d, *J* = 15.2 Hz, 1H), 7.05 (dd, *J* = 8.5, 1.0 Hz, 1H), 6.76 (d, *J* = 15.2 Hz, 1H), 6.58 (d, *J* = 3.5 Hz, 1H), 6.53 (dd, *J* = 8.4, 2.8 Hz, 1H), 6.45 (d, *J* = 2.8 Hz, 1H), 6.21 (d, *J* = 3.5 Hz, 1H), 2.70 (ddd, *J* = 9.5, 7.0, 4.5 Hz, 1H), 1.95 (s, 1H), 1.80 – 1.70 (m, 2H), 1.55 – 1.43 (m, 8H), 1.15 (td, *J* = 9.8, 7.7, 3.8 Hz, 5H), 0.85 (s, 4H), 0.76 (ddddd, *J* = 10.3, 9.2, 6.1, 3.1, 2.0 Hz, 5H), 0.07 (s, 6H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 201.71, 153.48, 150.67, 140.17, 138.23, 132.63, 130.30, 126.44, 119.05, 116.14, 115.29, 112.69, 109.78, 55.64, 54.95, 44.33, 43.80, 40.59, 38.13, 29.64, 27.57, 26.66, 24.23, 23.65, 22.03, 21.10, 14.21, 13.64, -4.25.

Condensation with 5-chloro-2-thiophenecarboxaldehyde (14)

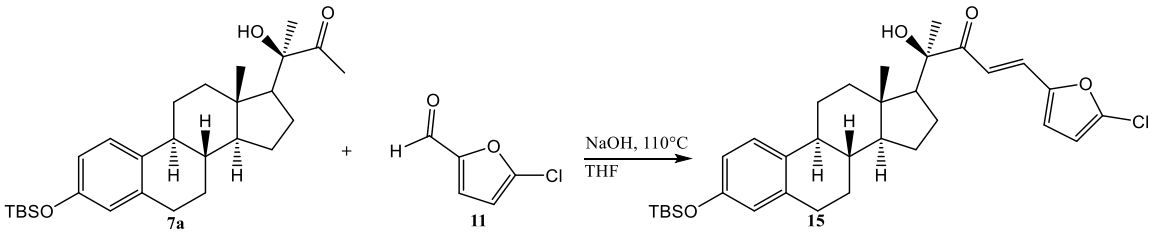


400 mg (0.88 mmol) of compound (7a) was dissolved in 1.04 mL (0.85 M) of THF in a vial. 102 mg (2.55 mmol) of sodium hydroxide was added to the vial as well as 193.5 mg (1.32 mmol) of 5-chloro-2-thiophenecarboxaldehyde (10). The vial was heated at 110°C for approximately 20 minutes. Approximately 2 mL of THF was added during heating to ensure reagents remained in solution. After 20 minutes, TLC, using 8:2 ratio of hexane to ethyl acetate, was performed to confirm completion of reaction. 25 mL of distilled water was added to the mixture to remove any remaining base. The organic layer was then collected using extraction with 3x 25 mL aliquots of ethyl acetate. Anhydrous sodium sulfate was added to the organic layer to remove any remaining water and was subsequently removed via filtration. The organic layer was then concentrated in vacuo and purified using column chromatography. A 9:1 hexane to ethyl acetate mobile phase mixture was used until product was eluted from the column to afford compound (14) in 38.0% yield.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.78 (d, *J* = 2.1 Hz, 1H), 7.65 (d, *J* = 15.5 Hz, 1H), 7.41 (dd, *J* = 8.5, 2.1 Hz, 1H), 7.08 (d, *J* = 8.4 Hz, 1H), 6.84 (dd, *J* = 12.1, 3.5 Hz, 2H), 6.56 (dd, *J* = 8.4, 2.7 Hz, 1H), 6.48 (d, *J* = 2.7 Hz, 1H), 4.18 (s, 1H), 4.05 (q, *J* = 7.2 Hz, 1H), 3.87 (s, 3H), 1.98 (s, 1H), 1.49 (s, 4H), 1.21 – 1.16 (m, 10H), 0.89 (s, 5H), 0.03 (s, 6H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 201.76, 157.98, 153.52, 144.06, 138.22, 132.60, 130.44, 128.31, 126.42, 117.16, 115.29, 112.52, 111.81, 60.55, 56.45, 55.67, 55.08, 53.46, 44.34, 43.82, 40.65, 38.13, 31.95, 29.39, 27.58, 24.33, 23.64, 22.05, 21.09, 13.66, -4.42.

Condensation with 5-chloro-2-furaldehyde (15)



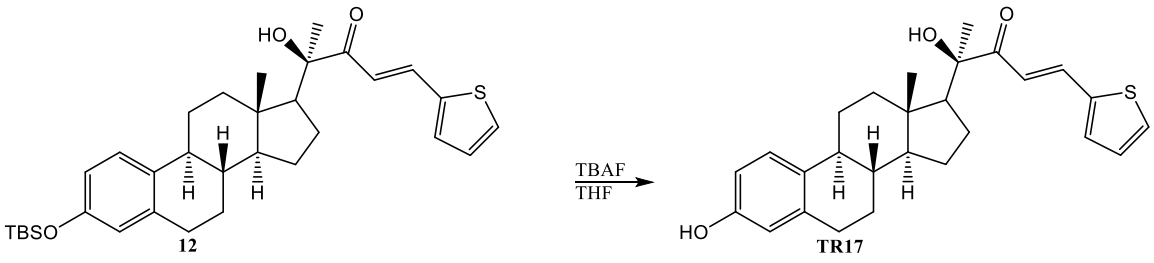
400 mg (0.88 mmol) of compound (7a) was dissolved in 1.04 mL (.85 M) of THF in a vial. 102 mg (2.55 mmol) of sodium hydroxide solid was added to the vial as well as 172 mg (1.32 mmol) of 5-chloro-2-furaldehyde (11). The vial was heated at 110°C for approximately 20 minutes. 2 mL of THF was added during heating to ensure reagents remained in solution. After 20 minutes TLC using 8:2 ratio of hexane to ethyl acetate was performed to confirm completion of reaction. 25 mL of distilled water was added to the mixture to remove any remaining base. The organic layer was then collected using extraction with 3x 25 mL aliquots of ethyl acetate. Anhydrous sodium sulfate was added

to the organic layer to remove any remaining water and was subsequently removed via filtration. The organic layer was then concentrated in vacuo and purified using column chromatography. A 9:1 hexane to ethyl acetate mixture was used as the mobile phase until product was eluted from the column. Compound (15) was achieved in 70.1% yield.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.76 (d, *J* = 15.1 Hz, 1H), 7.06 – 7.01 (m, 2H), 6.64 (d, *J* = 3.7 Hz, 1H), 6.58 – 6.52 (m, 2H), 6.45 (d, *J* = 2.7 Hz, 1H), 4.24 (s, 1H), 2.41 (s, 3H), 1.94 (s, 1H), 1.72 (ddt, *J* = 8.6, 6.4, 3.3 Hz, 2H), 1.43 (d, *J* = 4.6 Hz, 4H), 1.16 (d, *J* = 6.6 Hz, 8H), 0.84 (d, *J* = 2.8 Hz, 4H), 0.10 (s, 6H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 201.71, 153.66, 145.49, 138.75, 138.17, 137.79, 133.85, 132.46, 127.06, 126.40, 115.75, 115.37, 112.77, 55.67, 55.25, 44.29, 43.82, 40.65, 38.15, 31.63, 27.60, 26.68, 24.33, 22.70, 15.99, 14.18, 13.64, -4.34.

TR17



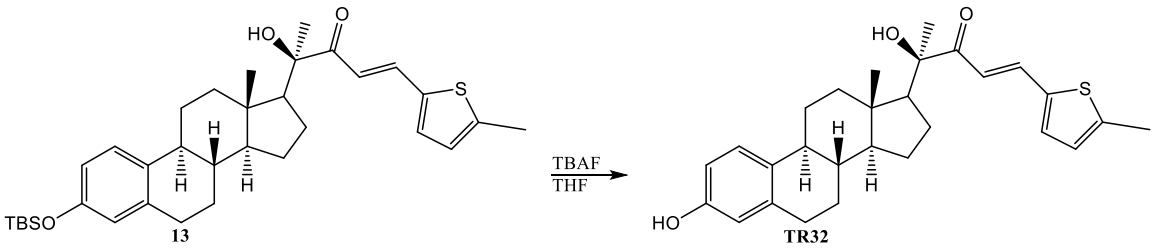
400 mg of compound (12) was dissolved in 11.25 mL (0.0645 M) of THF solution. After dissolution 2.25 mL (2.25 mmol) of TBAF was added to the reaction mixture. The reaction was left to proceed at room temperature for 18+ hours. The reaction was quenched with 25 mL of distilled water. Extraction of crude products was performed using 3x 25 mL aliquots

of ethyl acetate and distilled water. Anhydrous sodium sulfate was added in excess to remove any remaining water in the organic layer and was subsequently filtered off. Crude product was concentrated in vacuo and purified using column chromatography. To begin, a mobile phase mixture of 500 mL of an 8:2 ratio hexane to ethyl acetate was used, followed by 300 mL of a 7:3 ratio of hexane to ethyl acetate. Finally, 300 mL of a 6:4 ratio of hexane to ethyl acetate was used to elute the material. Final Product TR17 was collected in 76.0% yield.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.88 (d, *J* = 15.2 Hz, 1H), 7.37 (d, *J* = 5.1 Hz, 1H), 7.29 (s, 1H), 7.09 – 7.00 (m, 2H), 6.72 (d, *J* = 15.2 Hz, 1H), 6.56 (dd, *J* = 8.4, 2.8 Hz, 1H), 6.48 (d, *J* = 2.7 Hz, 1H), 5.16 (s, 1H), 4.18 (s, 1H), 2.72 (q, *J* = 4.6, 3.3 Hz, 2H), 2.27 – 2.08 (m, 5H), 1.76 (d, *J* = 9.0 Hz, 2H), 1.48 (d, *J* = 3.3 Hz, 4H), 1.24 – 1.15 (m, 8H), 0.88 (s, 3H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 201.70, 153.46, 139.73, 138.25, 138.20, 132.83, 132.66, 129.52, 128.50, 126.44, 117.16, 115.30, 112.69, 55.65, 55.12, 44.32, 43.80, 40.64, 38.12, 29.73, 29.64, 27.57, 26.67, 24.30, 23.65, 22.05, 13.65.

TR32

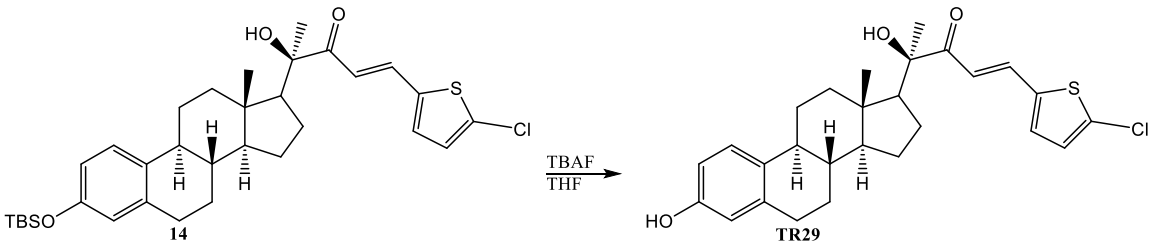


290 mg (0.515 mmol) of compound (13) was dissolved in 8 mL (0.0645 M) THF. 1.6 mL (1.6 mmol) of TBAF was added to the dissolved starting material and the reaction was left for 18+ hours. The reaction was quenched with 25 mL of distilled water. Extraction of crude products was performed using 3x 25 mL aliquots of ethyl acetate and distilled water. Anhydrous sodium sulfate was added in excess to remove any remaining water in the organic layer and was subsequently filtered off. Crude product was concentrated in vacuo and purified using column chromatography. To begin a mobile phase mixture of approximately 500 mL 85:15 hexane to ethyl acetate was used. Mobile phase was then strengthened to an 8:2 ratio for another 300 mL, and finally strengthened again to a 7:3 ratio for 200 mL until TR32 was eluted in 96% yield.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.35 (d, *J* = 15.2 Hz, 1H), 7.05 (dd, *J* = 8.5, 1.0 Hz, 1H), 6.76 (d, *J* = 15.2 Hz, 1H), 6.58 (d, *J* = 3.5 Hz, 1H), 6.53 (dd, *J* = 8.4, 2.8 Hz, 1H), 6.45 (d, *J* = 2.8 Hz, 1H), 6.21 (d, *J* = 3.5 Hz, 1H), 2.70 (ddd, *J* = 9.5, 7.0, 4.5 Hz, 2H), 1.95 (s, 1H), 1.80 – 1.70 (m, 2H), 1.55 – 1.43 (m, 8H), 1.15 (td, *J* = 9.8, 7.7, 3.8 Hz, 5H), 0.85 (s, 4H), 0.76 (ddddd, *J* = 10.3, 9.2, 6.1, 3.1, 2.0 Hz, 5H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 201.71, 153.48, 150.67, 140.17, 138.23, 132.63, 130.30, 126.44, 119.05, 116.14, 115.29, 112.69, 109.78, 55.64, 54.95, 44.33, 43.80, 40.59, 38.13, 29.64, 27.57, 26.66, 24.23, 23.65, 22.03, 21.10, 14.21, 13.64.

TR29

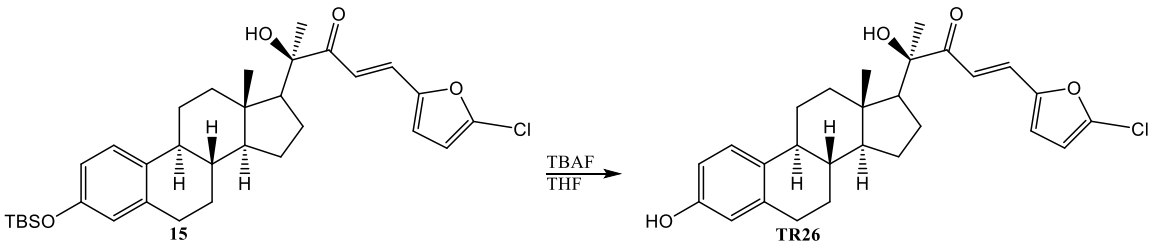


200 mg (0.334 mmol) of compound (14) was dissolved in 5.2 mL (0.0645 M) of THF solution. After dissolution 1.0 mL (1.01 mmol) of TBAF was added to the solution. The reaction mixture was left at room temperature for 18+ hours. The reaction was quenched with 25 mL of distilled water. Extraction of crude product was performed using 3x 25 mL aliquots of ethyl acetate. The organic layer was collected and dried using excess anhydrous sodium sulfate. After drying the sodium sulfate was filtered off, and the crude product was concentrated in vacuo. Crude product was purified in two mobile phase ratios. First 500 mL of an 8:2 ratio of hexane to ethyl acetate was used followed by 400 mL of a 7:3 ratio of hexane to ethyl acetate. Final product TR29 was collected in 98.5% yield.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.78 (d, *J* = 2.1 Hz, 1H), 7.65 (d, *J* = 15.5 Hz, 1H), 7.41 (dd, *J* = 8.5, 2.1 Hz, 1H), 7.08 (d, *J* = 8.4 Hz, 1H), 6.84 (dd, *J* = 12.1, 3.5 Hz, 2H), 6.56 (dd, *J* = 8.4, 2.7 Hz, 1H), 6.48 (d, *J* = 2.7 Hz, 1H), 5.21 (s, 1H), 4.18 (s, 1H), 4.05 (q, *J* = 7.2 Hz, 1H), 3.87 (s, 3H), 1.98 (s, 1H), 1.49 (s, 4H), 1.21 – 1.16 (m, 10H), 0.89 (s, 5H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 201.76, 157.98, 153.52, 144.06, 138.22, 132.60, 130.44, 128.31, 126.42, 117.16, 115.29, 112.52, 111.81, 60.55, 56.45, 55.67, 55.08, 53.46, 44.34, 43.82, 40.65, 38.13, 31.95, 29.39, 27.58, 24.33, 23.64, 22.05, 21.09, 13.66.

TR26

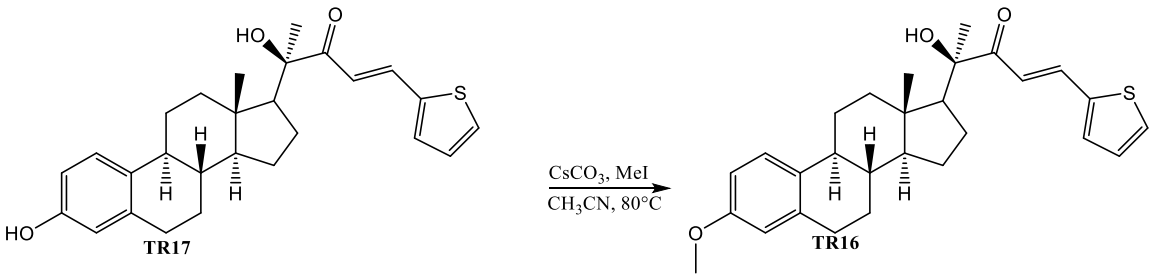


350 mg (0.617 mmol) of compound (15) was dissolved in 10 mL (0.0645 M) of THF solution. After dissolution 1.9 mL (1.9 mmol) of TBAF was added to the solution. The reaction mixture was left at room temperature for 18+ hours. The reaction was quenched with 25 mL of distilled water. Extraction of crude product was performed using 3x 25 mL aliquots of ethyl acetate. The organic layer was collected and dried using excess anhydrous sodium sulfate. After drying the sodium sulfate was filtered off, and the crude product was concentrated in vacuo. Crude product was purified using multiple mobile phase ratios. To begin, 500 mL of an 8:2 ratio of hexane to ethyl acetate was used. Next 300 mL of a 75:25 mixture was used, and finally 500 mL of a 6:4 mixture was used. Final Product TR26 was collected in 81.9% yield.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.76 (d, *J* = 15.1 Hz, 1H), 7.06 – 7.01 (m, 2H), 6.64 (d, *J* = 3.7 Hz, 1H), 6.58 – 6.52 (m, 2H), 6.45 (d, *J* = 2.7 Hz, 1H), 5.79 (s, 1H), 4.24 (s, 1H), 2.41 (s, 3H), 1.94 (s, 1H), 1.72 (ddt, *J* = 8.6, 6.4, 3.3 Hz, 2H), 1.43 (d, *J* = 4.6 Hz, 4H), 1.16 (d, *J* = 6.6 Hz, 8H), 0.84 (d, *J* = 2.8 Hz, 4H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 201.71, 153.66, 145.49, 138.75, 138.17, 137.79, 133.85, 132.46, 127.06, 126.40, 115.75, 115.37, 112.77, 55.67, 55.25, 44.29, 43.82, 40.65, 38.15, 31.63, 27.60, 26.68, 24.33, 22.70, 15.99, 14.18, 13.64.

TR16

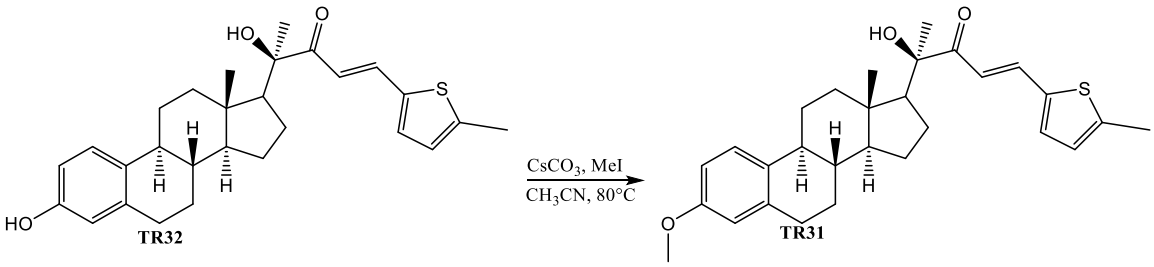


100 mg (0.229 mmol) of compound TR17 was dissolved in 0.557 mL of acetonitrile. After dissolution 82 mg (0.230 mmol) of anhydrous cesium carbonate was added to the reaction mixture and left to react for 10 minutes at room temperature. Finally, 0.07 mL of methyl iodide was added, and the reaction was refluxed at 80°C for approximately 4 hours. After 4 hours the acetonitrile was evaporated off, and the reaction mixture was quenched with approximately 25 mL of distilled water to dissolve any remaining cesium carbonate. The organic layer was then collected by extracting the mixture with 3 x 25 mL aliquots of ethyl acetate. Excess anhydrous sodium sulfate was added to the organic layer to remove any water and is then filtered off. Crude product was concentrated in vacuo and then purified using column chromatography. A mobile phase mixture of 8:2 hexane to ethyl acetate was used until product eluted from the column. Final Product TR16 was collected in 62.1% yield.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.94 (d, *J* = 15.2 Hz, 1H), 7.67 (d, *J* = 5.1 Hz, 1H), 7.29 (s, 1H), 7.09 – 7.00 (m, 2H), 6.72 (d, *J* = 15.2 Hz, 1H), 6.56 (dd, *J* = 8.4, 2.8 Hz, 1H), 6.48 (d, *J* = 2.7 Hz, 1H), 5.16 (s, 1H), 4.18 (s, 1H), 2.72 (q, *J* = 4.6, 3.3 Hz, 2H), 2.27 – 2.08 (m, 5H), 1.76 (d, *J* = 9.0 Hz, 2H), 1.55 (s, 1H), 1.48 (d, *J* = 3.3 Hz, 4H), 1.24 – 1.15 (m, 8H), 0.88 (s, 3H).

^{13}C NMR (151 MHz, $CDCl_3$) δ 201.34, 157.38, 138.46, 137.95, 137.20, 134.02, 132.66, 132.15, 127.79, 126.19, 117.24, 113.62, 111.43, 79.13, 55.68, 55.24, 55.13, 44.34, 43.83, 40.66, 38.15, 29.84, 27.62, 26.66, 24.34, 23.65, 22.05, 13.62.

TR31



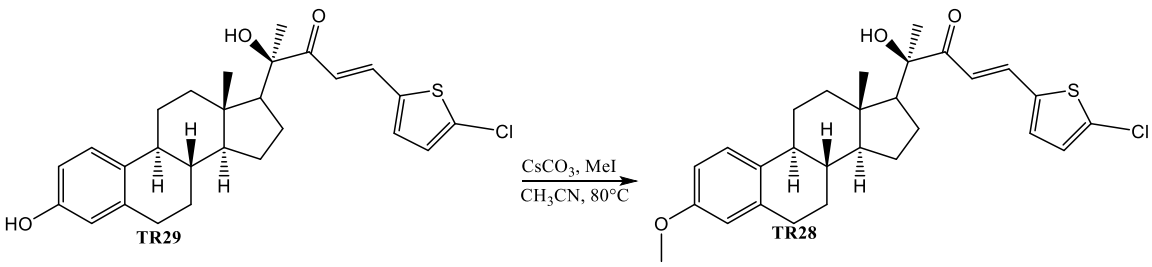
100 mg (0.236 mmol) of compound TR32 was dissolved in 0.574 mL of acetonitrile. After dissolution 84 mg (0.238 mmol) of anhydrous cesium carbonate was added to the reaction mixture and left to react for 10 minutes at room temperature. 0.07 mL of methyl iodide was then added, and the reaction mixture was refluxed at 80°C for approximately 4 hours. After 4 hours the acetonitrile was evaporated off, and the reaction mixture was quenched with approximately 25 mL of distilled water to dissolve any remaining cesium carbonate. The organic layer was then collected by extracting the mixture with 3 x 25 mL aliquots of ethyl acetate. Excess anhydrous sodium sulfate was added to the organic layer to remove any water and then filtered off. Crude product was then concentrated in vacuo and purified using column chromatography. A mobile phase mixture of 8:2 hexane to ethyl acetate ratio was used to purify crude product in 54.2% yield of final product TR31.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.94 (d, *J* = 15.2 Hz, 1H), 7.67 (d, *J* = 5.1 Hz, 1H), 7.29 (s, 1H), 7.09 – 7.00 (m, 2H), 6.72 (d, *J* = 15.2 Hz, 1H), 6.56 (dd, *J* = 8.4, 2.8 Hz, 1H),

6.48 (d, *J* = 2.7 Hz, 1H), 5.16 (s, 1H), 4.18 (s, 1H), 2.72 (q, *J* = 4.6, 3.3 Hz, 2H), 2.64 (s, 3H), 2.27 – 2.08 (m, 5H), 1.76 (d, *J* = 9.0 Hz, 2H), 1.48 (d, *J* = 3.3 Hz, 4H), 1.24 – 1.15 (m, 8H), 0.88 (s, 3H).

^{13}C NMR (151 MHz, $CDCl_3$) δ 201.34, 157.38, 138.46, 137.95, 137.20, 134.02, 132.66, 132.15, 127.79, 126.19, 117.24, 113.62, 111.43, 79.13, 64.10, 55.68, 55.24, 55.13, 44.34, 43.83, 40.66, 38.15, 29.84, 27.62, 26.66, 24.34, 23.65, 22.05, 13.62.

TR28



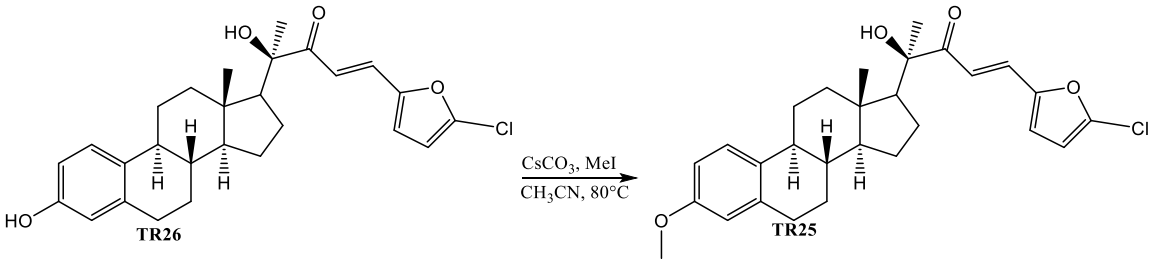
50 mg (.106 mmol) of compound TR29 was dissolved in 0.1 mL of acetonitrile. After dissolution 40.5 mg (.11 mmol) of anhydrous cesium carbonate was added to the reaction mixture and left for 10 minutes at room temperature. Finally, 0.03 mL of methyl iodide was added to the reaction mixture. The reaction then refluxed at 80°C for 4 hours. After 4 hours the acetonitrile was evaporated off, and the reaction mixture was quenched with approximately 25 mL of distilled water to dissolve any remaining cesium carbonate. The organic layer was then collected by extracting the mixture with 3 x 25 mL aliquots of ethyl acetate. Excess anhydrous sodium sulfate was added to the organic layer to remove any water and then filtered off. Crude product was then concentrated in vacuo and purified using column chromatography. A mobile phase mixture of 8:2 hexane to ethyl acetate was

used until all product was eluted from the column. Final product TR28 was collected in 58.4% yield.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.94 (d, *J* = 15.2 Hz, 1H), 7.67 (d, *J* = 5.1 Hz, 1H), 7.29 (s, 1H), 7.09 – 7.00 (m, 2H), 6.72 (d, *J* = 15.2 Hz, 1H), 6.56 (dd, *J* = 8.4, 2.8 Hz, 1H), 6.48 (d, *J* = 2.7 Hz, 1H), 5.16 (s, 1H), 4.18 (s, 1H), 2.72 (q, *J* = 4.6, 3.3 Hz, 2H), 2.27 – 2.08 (m, 5H), 1.76 (d, *J* = 9.0 Hz, 2H), 1.48 (d, *J* = 3.3 Hz, 4H), 1.24 – 1.15 (m, 8H), 0.88 (s, 3H).

^{13}C NMR (151 MHz, $CDCl_3$) δ 201.34, 157.38, 138.46, 137.95, 137.20, 134.02, 132.66, 132.15, 127.79, 126.19, 117.24, 113.62, 111.43, 79.13, 55.68, 55.24, 55.13, 44.34, 43.83, 40.66, 38.15, 29.84, 27.62, 26.66, 24.34, 23.65, 22.05, 13.62.

TR25



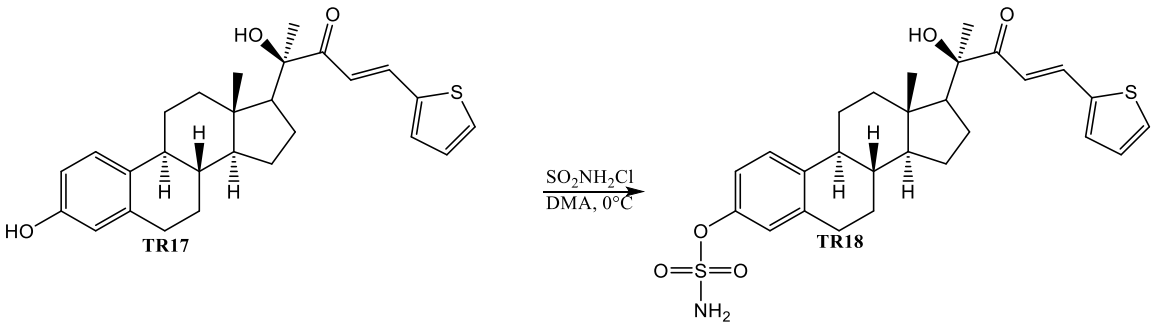
100 mg (0.22 mmol) of compound TR26 was dissolved in 0.374 mL of acetonitrile. After dissolution 55 mg (0.23 mmol) of anhydrous cesium carbonate was added to the reaction mixture and left for 10 minutes at room temperature. Finally, 0.05 mL of methyl iodide was added to the reaction mixture, and the flask was left to reflux for approximately 4 hours at 80°C. After 4 hours the acetonitrile was evaporated off, and the reaction mixture was

quenched with approximately 25 mL of distilled water to dissolve any remaining cesium carbonate. The organic layer was then collected by extracting the mixture with 3 x 25 mL aliquots of ethyl acetate. Excess anhydrous sodium sulfate was added to the organic layer to remove any water and then filtered off. Crude product was then concentrated in vacuo and purified using column chromatography. A mobile phase mixture of 8:2 hexane to ethyl acetate was used until the product was eluted. Final product TR25 was collected in 44.7% yield.

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.88 (d, *J* = 15.2 Hz, 1H), 7.37 (d, *J* = 5.1 Hz, 1H), 7.29 (s, 1H), 7.09 – 7.00 (m, 2H), 6.72 (d, *J* = 15.2 Hz, 1H), 6.56 (dd, *J* = 8.4, 2.8 Hz, 1H), 6.48 (d, *J* = 2.7 Hz, 1H), 5.16 (s, 1H), 4.18 (s, 1H), 2.72 (q, *J* = 4.6, 3.3 Hz, 2H), 2.27 – 2.08 (m, 5H), 1.76 (d, *J* = 9.0 Hz, 2H), 1.48 (d, *J* = 3.3 Hz, 4H), 1.24 – 1.15 (m, 8H), 0.88 (s, 3H).

^{13}C NMR (151 MHz, $CDCl_3$) δ 201.25, 157.64, 138.58, 137.92, 137.11, 134.08, 132.67, 132.11, 127.97, 126.33, 117.28, 113.60, 111.35, 82.14, 55.71, 55.48, 55.11, 44.34, 43.87, 40.69, 38.02, 29.88, 27.61, 26.78, 24.23, 23.77, 22.01, 13.57.

TR18

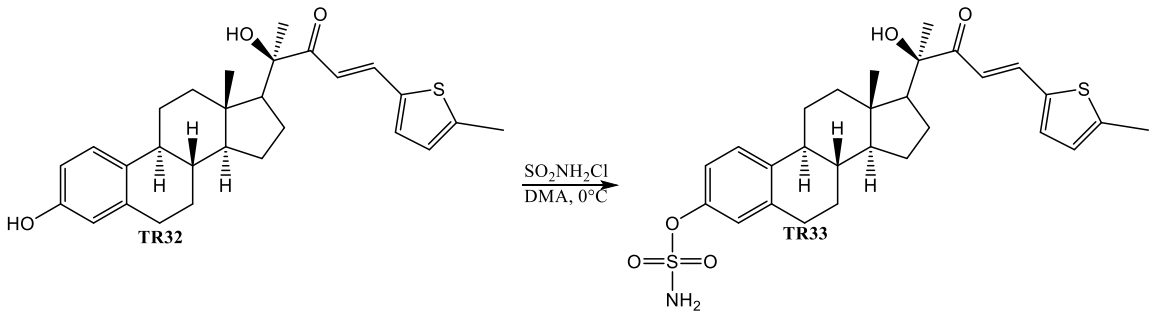


0.546 mL (6.214 mmol) of N-N-dimethylacetamide was separated into two 0.273 mL aliquots. One aliquot was used to dissolve 100 mg (0.239 mmol) of compound TR17. After dissolution the mixture was brought to 0°C. The remaining aliquot of solvent was used to dissolve 83 mg (0.717 mmol) of Sulfamoyl chloride in a vial. The Sulfamoyl chloride mixture was then added to the reaction mixture on ice and left 18+ hours overnight for the reaction to proceed. After reaction completion the mixture was quenched with 25 mL of distilled water and extracted with 3x 25 mL aliquots of ethyl acetate. The organic layer was collected, and anhydrous sodium sulfate was added to remove any remaining water from the layer. The solid was filtered off and the liquid was concentrated in vacuo. The compound was purified using column chromatography. Multiple mobile phases were used to purify this compound. First 500 mL of an 8:2 ratio of hexane to ethyl acetate was used. Next 300 mL of a 7:3 mixture of hexane to ethyl acetate, and finally 200 mL of 6:4 hexane to ethyl acetate to elute the rest of the product. Final product TR18 was collected in 27.3% yield.

^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.88 (d, *J* = 15.2 Hz, 1H), 7.38 (dt, *J* = 5.1, 1.0 Hz, 1H), 7.30 (dd, *J* = 3.8, 1.1 Hz, 1H), 7.23 (d, *J* = 8.5 Hz, 1H), 7.01 (ddd, *J* = 12.1, 6.9, 3.2 Hz, 2H), 6.95 (d, *J* = 2.6 Hz, 1H), 6.71 (d, *J* = 15.3 Hz, 1H), 5.06 (d, *J* = 18.1 Hz, 2H), 2.81 – 2.76 (m, 2H), 2.28 – 2.14 (m, 3H), 1.81 – 1.76 (m, 2H), 1.60 – 1.49 (m, 4H), 1.45 (s, 3H), 1.23 – 1.15 (m, 5H), 0.87 (d, *J* = 1.4 Hz, 3H), 0.82 – 0.76 (m, 2H).

^{13}C NMR (151 MHz, $CDCl_3$) δ 201.60, 147.88, 139.76, 139.71, 138.96, 138.25, 132.90, 129.55, 128.53, 126.74, 121.93, 118.94, 117.09, 79.08, 55.64, 55.09, 44.20, 43.98, 40.52, 37.66, 29.54, 27.27, 26.46, 24.32, 23.65, 22.02, 13.60.

TR33

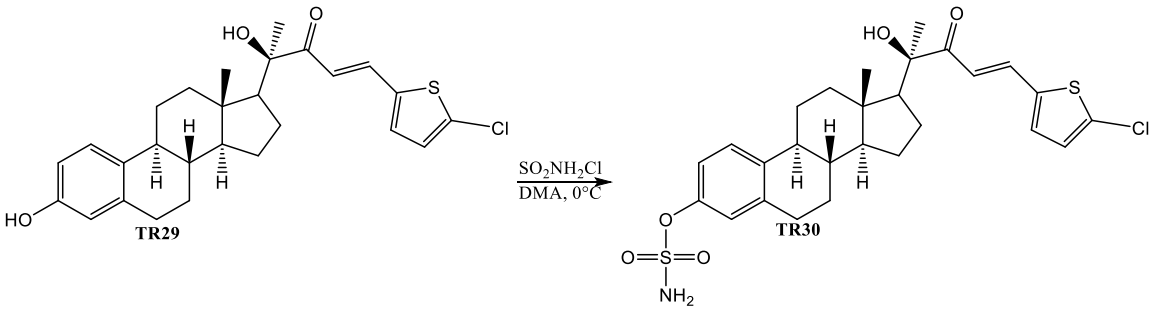


0.533 mL (6.136 mmol) of N,N-dimethylacetamide was separated into two 0.267 mL aliquots. One aliquot was used to dissolve 100 mg (0.236 mmol) of compound TR32. After dissolution the mixture was brought to 0°C. The remaining aliquot of solvent was used to dissolve 82 mg (0.708 mmol) of Sulfamoyl chloride in a vial. The Sulfamoyl chloride mixture was then added to the reaction mixture on ice and left 18+ hours overnight for the reaction to proceed. After reaction completion the mixture was quenched with 25 mL of distilled water and extracted with 3x 25 mL aliquots of ethyl acetate. The organic layer was collected, and anhydrous sodium sulfate was added to remove any remaining water from the layer. The solid was filtered off and the liquid was concentrated in vacuo. The compound was purified using column chromatography. A mobile phase mixture of 500 mL of 7:3 hexane to ethyl acetate ratio was used at first followed by 200 mL of a 6:4 mixture. Final product TR33 was collected in 42.2% yield.

^{1}H NMR (400 MHz, Acetone-d_6) δ 7.68 (d, J = 15.3 Hz, 1H), 7.25 – 7.18 (m, 2H), 6.94 – 6.87 (m, 5H), 6.73 (dt, J = 3.6, 1.2 Hz, 1H), 4.13 (s, 1H), 2.38 (s, 3H), 2.26 – 2.08 (m, 4H), 1.94 – 1.91 (m, 2H), 1.77 (s, 1H), 1.58 – 1.50 (m, 4H), 1.37 (d, J = 2.8 Hz, 3H), 1.22 – 1.14 (m, 5H), 0.83 (s, 3H).

^{13}C NMR (151 MHz, $CDCl_3$) δ 201.60, 147.88, 139.76, 139.71, 138.96, 138.25, 132.90, 129.55, 128.53, 126.74, 121.93, 118.94, 117.09, 79.08, 64.85, 55.64, 55.09, 44.20, 43.98, 40.52, 37.66, 29.54, 27.27, 26.46, 24.32, 23.65, 22.02, 13.60.

TR30



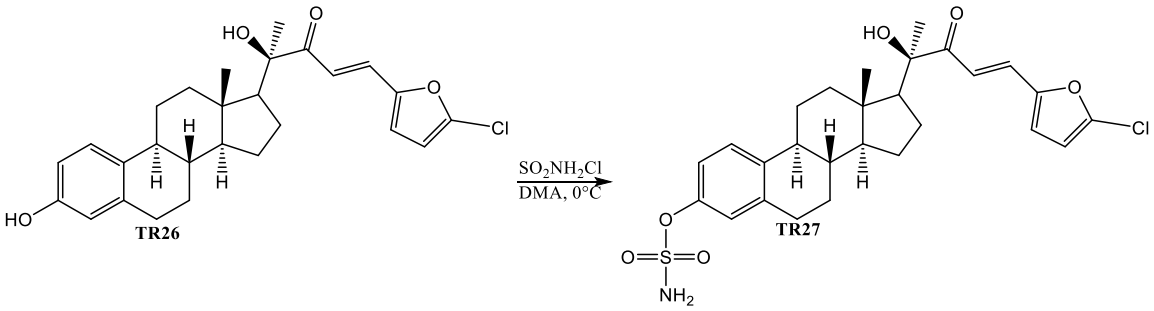
0.480 mL (5.512 mmol) of N,N-dimethylacetamide was separated into two 0.240 mL aliquots. One aliquot was used to dissolve 100 mg (0.212 mmol) of compound TR29. After dissolution the mixture was brought to 0°C. The remaining aliquot of solvent was used to dissolve 73.6 mg (0.636 mmol) of Sulfamoyl chloride in a vial. The Sulfamoyl chloride mixture was then added to the reaction mixture on ice and left 18+ hours overnight for the reaction to proceed. After reaction completion the mixture was quenched with 25 mL of distilled water and extracted with 3x 25 mL aliquots of ethyl acetate. The organic layer was collected, and anhydrous sodium sulfate was added to remove any remaining water from the layer. The solid was filtered off and the liquid was concentrated in vacuo. The compound was purified using column chromatography. Multiple mobile phase mixtures were required to purify this product. First 600 mL of an 8:2 hexane to ethyl acetate mixture was used, followed by 400 mL of a 7:3 hexane to ethyl acetate mixture, and finally with

300 mL of a 6:4 hexane to ethyl acetate mixture. Final product TR30 was isolated in 5.6% yield.

^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.72 (dd, *J* = 15.2, 0.6 Hz, 1H), 7.21 (d, *J* = 1.1 Hz, 1H), 7.08 (d, *J* = 3.9 Hz, 1H), 7.02 (dd, *J* = 8.5, 2.7 Hz, 1H), 6.96 (d, *J* = 2.5 Hz, 1H), 6.85 (d, *J* = 4.0 Hz, 1H), 6.60 (d, *J* = 15.2 Hz, 1H), 5.81 (s, 2H), 2.80 – 2.77 (m, 2H), 2.00 (s, 9H), 1.44 (s, 3H), 1.18 (d, *J* = 2.0 Hz, 4H), 0.87 (s, 3H), 0.82 – 0.77 (m, 3H).

^{13}C NMR (151 MHz, $CDCl_3$) δ 201.33, 170.94, 148.02, 139.35, 138.69, 138.45, 137.40, 132.43, 127.76, 126.57, 122.07, 119.11, 117.18, 79.09, 55.63, 55.10, 44.19, 43.97, 40.50, 38.10, 37.70, 35.27, 29.51, 27.29, 24.31, 21.52, 13.57.

TR27



0.528 mL (6.058 mmol) of N,N-dimethylacetamide was separated into two 0.264 mL aliquots. One aliquot was used to dissolve 100 mg (0.233 mmol) of compound TR26. After dissolution the mixture was brought to 0°C. The remaining aliquot of solvent was used to dissolve 81 mg (0.699 mmol) of Sulfamoyl chloride in a vial. The Sulfamoyl chloride mixture was then added to the reaction mixture on ice and left 18+ hours overnight for the reaction to proceed. After reaction completion the mixture was quenched with 25 mL of

distilled water and extracted with 3x 25 mL aliquots of ethyl acetate. The organic layer was collected, and anhydrous sodium sulfate was added to remove any remaining water from the layer. The solid was filtered off and the liquid was concentrated in vacuo. The compound was purified using column chromatography. A mobile phase mixture of 600 mL of 8:2 hexane to ethyl acetate ratio was used at first followed by 300 mL of a 7:3 mixture. Final product TR27 was collected in 27.9% yield.

^{1}H NMR (600 MHz, Chloroform-*d*) δ 7.80 (d, *J* = 15.1 Hz, 1H), 7.34 (d, *J* = 8.6 Hz, 2H), 7.07 (s, 2H), 6.85 (d, *J* = 3.6 Hz, 1H), 6.48 (d, *J* = 7.3 Hz, 1H), 6.43 (d, *J* = 3.6 Hz, 1H), 5.33 (s, 1H), 4.91 (s, 2H), 3.52 (s, 3H), 2.19 (s, 3H), 2.08 (s, 2H), 1.81 (s, 3H), 1.29 (s, 7H), 0.97 (d, *J* = 3.5 Hz, 3H), 0.69 (s, 3H).

^{13}C NMR (151 MHz, $CDCl_3$) δ 201.33, 170.94, 148.02, 139.35, 138.69, 138.45, 137.40, 132.43, 127.76, 126.57, 122.07, 119.11, 117.18, 82.45, 55.63, 55.10, 44.19, 43.97, 40.50, 38.10, 37.70, 35.27, 29.51, 27.29, 24.31, 21.52, 13.57.

3.3.2 Biological Evaluations

Pancreatic Cell Lines and Culture

Two pancreatic cell lines were used in these experiments: AsPC-1 and Panc-1 obtained from American Type Culture Collection (ATCC). The cells were maintained in Dulbecco's Modified Eagle's Medium (DMEM, Thermoscientific™ Hyclone™) supplemented with 10% fetal bovine serum (FBS) (Thermoscientific™ Hyclone™) and 1% penicillin (100 IU/mL)/streptomycin (100 µg/mL) (Corning™ Cellgro™). Cells were maintained at 37°C with 5% CO_2.

Cytotoxicity Assay

The cell lines were seeded in a 96-well plate at $3*10^6$ cells/mL (100 µL/well). A serial dilution of all heterocyclic estrone analogs were added after overnight incubation of the cells at 37°C and 5% CO_2. DMSO was used as a control (0.1 %). The cells were incubated with the analogs for 48 hrs. followed by addition of 20 µL of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl tetrazolium bromide (MTT) (5 mg/mL PBS) to each well and incubation for another 2 hours. The formazan crystals that formed were solubilized with 100 µL DMSO. The absorbance was measured at 570 nm by Biotek plate reader.

Cell Cycle Analysis

Cells were seeded at $5*10^5$ cells per well in a 6-well plate (3 mL/well) and allowed to adhere overnight at 37°C and 5% CO_2. The cells were incubated with different concentrations of hit compounds (TR33 or TR26) at their respective IC_{50} values and a DMSO control for 48 hours. The cells were then washed twice with ice-cold 1X PBS and collected after trypsinization. The cell pellet was washed two times with ice-cold 1X PBS and fixed with ice-cold 70% ethanol overnight at -4°C. When ready for analysis, the cells were centrifuged. Then they were washed once with ice-cold PBS followed by a second wash with ice-cold PBS-2% FBS. The cell pellet was re-suspended in 500 µL propidium iodide (PI)/RNase (BD Biosciences, BDB550825) staining solution for 15 min at room temperature (RT) in the dark and analyzed within 1 hour by BD C6 Accuri flow cytometer (BD Biosciences, San Jose, CA).

In-Cell Western Assay

Cells were seeded as described above at $1*10^5$ cells/well then prepared for In-cell western assay according to Rockland protocol[28]. After seeding, hit compounds (TR33 or

TR26) were added to wells in various concentrations (1/4, 1/2, 1, and 2 times the IC_{50}) along with a DMSO control at the same concentration as twice the IC_{50} and incubated for 24 hours. Rockland protocol was then followed to fix, permeabilize, and block cells. Primary antibodies in both phosphorylated and non-phosphorylated forms were added for EGFR, mTOR, STAT3, and ERK. GAPDH was also used as a control for non-phosphorylated forms. Secondary antibodies used were Mouse IgG (H&L Antibody Dylight) 680 Conjugated with fluorochrome and Rabbit IgG (H&L Antibody DyLight) 800 Conjugated with fluorochrome to detect both forms of the proteins simultaneously. The plate was read on LiCOR Odyssey® imager. Images were processed with ImageJ software[29]. Inhibition of a protein is taken as a percentage of the phosphorylated versus non-phosphorylated forms of each protein that was expressed. Values are taken from quadruplicate testing.

3.3.3 Molecular Dynamics

Analogs TR26 and TR33 underwent dynamic simulation in respective proteins that they inhibited during in-cell western studies: EGFR and ERK for TR26 and ERK and mTOR for TR33. The analogs were drawn using ChemDraw with respective stereochemical features, minimized in energy using Chem3D's MMFF94 energy minimization, and saved in a .pdb format. Proteins were downloaded in .pdb format from Protein Data Bank using the following codes: 1M17 (EGFR), 2OJG (ERK), and 4JSV (mTOR). Structural alterations were first performed on the protein in PyMOL to remove excess chains and ligands from downloaded files. Swiss PDB viewer was then used to minimize proteins with GROMOS 43B1 energy minimization function. To create the protein-ligand complexes, PyRx-vina software was utilized for molecular docking,

followed by conversion of the protein-ligand complex to .pdb file using PyMOL. Molecular dynamic simulations were then performed using Schrodinger Desmond Maestro software[30]. Pre-processing of the protein-ligand complex was performed to set correct bond orders, add hydrogens, and detect disulfide bonds. Secondly, a refining step was performed on the prepared structured file in order to analyze and change protonation and tautomeric states of variable residues by maestro's protein preparation wizard. Thirdly, A structural optimization step was performed to optimize hydrogen bond network of the protein-ligand complex. Energy minimization was then performed to relieve any bond strain or steric hindrance, followed by introduction of a water model using TIP3P 3-site models. Next, a force field was incorporated using OPLS-2005 force field function to estimate forces between atoms within molecules. Ionization models were set for pH equal to 7.4 to mimic the human body and to distribute positive or negative counter ions to neutralize the system. Upon completion of the pre-processing steps for the complex, molecular dynamic simulation was performed using a 300 ns simulation event with 3,000 frames. After the simulation was completed, Simulation event analysis was utilized to calculate the RMSD, RMSF, Radius of Gyration, Intra- and Intermolecular hydrogen bonding events for both the protein and ligand, including the complex. Finally, simulation interaction diagrams were created to display the average interactions of the protein and ligand over the entire simulation event. Additionally, an MMGBSA was performed using the completed trajectory to calculate total binding free energies for the protein-ligand complexes.

3.4 Additional Data

It is worth noting the synthesized heterocyclic estrone analogs above have also been studied in brief by other colleagues wherein this data has not been previously reported. Additional cytotoxicity data for these analogs has been performed in a non-small cell lung cancer line NCIH-226, as well as in two triple negative breast cancer lines in an unpublished study by Dr. Felix Acheampong. The results for this data can be seen in **Table 3.2**. In particular we see TR26 and TR33, the two analogs that showed activity in pancreatic cancer, also show activity in all three of the additional cell lines they were tested in. We also see these analogs perform well in triple negative breast cancer cell lines, suggesting they may be useful as a precursor analog to a future hit compound.

Table 3.2 IC_{50} results of synthesized heterocyclic estrone analogs on lung cancer cell line NCIH-226 and breast cancer cell lines MDA-MB-231 and MDA-MB-468

Compound	NCIH-226 IC_{50} (μM)	MDA-MB-231 IC_{50} (μM)	MDA-MB-468 IC_{50} (μM)
TR16	>50	>50	>50
TR17	>50	>50	9.89 ± 0.37
TR18	>50	34.19 ± 4.19	5.93 ± 0.46
TR25	>50	>50	>50
TR26	40.52 ± 0.78	49.60 ± 0.40	13.71 ± 1.87
TR27	>50	>50	>50
TR28	>50	>50	8.92 ± 0.22
TR29	>50	>50	9.37 ± 0.37
TR30	>50	>50	3.18 ± 0.12
TR31	50.55 ± 0.98	>50	>50
TR32	>50	8.12 ± 0.18	9.30 ± 0.06
TR33	15.70 ± 0.53	10.73 ± 0.93	11.13 ± 0.11

In an additional experiment, the heterocyclic analogs were utilized by Dr. Jennifer Kyerematen in an unpublished study to determine their ability to reverse drug resistance of H69AR cells, which highly express multidrug resistance proteins, to vincristine, an anticancer drug. Cytotoxicity of the analogs were first determined in the H69AR cell line (**Table 3.3**). They were then utilized at non-cytotoxic concentrations along with vincristine to determine if they could reverse the drug resistance of H69AR cells towards vincristine

(**Table 3.4**). In H69 cells, vincristine shows low nanomolar cytotoxicity, but in the drug resistant cell line H69AR this is 14-fold worse at 16.29 nM. MK571, a known multidrug resistance protein inhibitor, is able to fully restore cytotoxicity of vincristine. We note that most of the heterocyclic analogs are able to restore some amount of the cytotoxicity of vincristine, further increasing the reliability of these analogs to act as a precursor analogs to further hit compounds which will be discussed in chapter 4.

Table 3.3 IC_{50} results of synthesized heterocyclic estrone analogs on H69AR cells

Analog ID	IC_{50} (H69AR) (μM)
TR17	44.10 ± 0.79
TR18	9.98 ± 0.19
TR25	41.19 ± 1.26
TR28	43.15 ± 0.61
TR29	1.70 ± 0.030
TR30	24.42 ± 2.36
TR32	10.86 ± 0.27
TR33	6.90 ± 2.02

Table 3.4 Effect of heterocyclic analogs on the IC_{50} values of vincristine in H69 and H69AR cells

Cell line/Treatment	IC_{50}[a] (nM)	
	Vincristine	Fold resistance[b]
H69	1.15 ± 0.09	1
H69AR	16.29 ± 3.08	14.19
H69AR + MK571 [25 μM]	1.01 ± 0.02	0.88
H69AR + TR17 [20 μM]	5.16 ± 0.53	4.50
H69AR + TR18 [5 μM]	3.21 ± 0.19	2.80
H69AR + TR25 [20 μM]	25.42 ± 3.54	22.15
H69AR + TR28 [10 μM]	5.39 ± 0.35	4.70
H69AR + TR29 [15 μM]	25.05 ± 0.08	21.82
H69AR + TR30 [5 μM]	2.99 ± 1.47	2.60
H69AR + TR32 [25 μM]	8.59 ± 0.93	7.49
H69AR + TR33 [5 μM]	9.02 ± 0.70	7.86

[a]Mean ± SEM of at least three independent experiments performed in triplicates.

[b] Fold resistance determined by dividing the IC_{50} value for each treatment by the IC_{50} value of H69 with vincristine alone.

3.5 Conclusions

Twelve heterocyclic estrone analogs were synthesized for testing against pancreatic cancer. Modifications were made at the C3 position to vary polarity of the molecule. Biological evaluation of the molecules was performed using an MTT assay against two pancreatic cell lines: AsPC-1 and Panc-1. The analog with the lowest IC_{50} in both cell lines

was used for further analysis. TR26 had an IC_{50} of 10.16 ± 0.83 µM in Panc-1 and TR33 had an IC_{50} of 38.37 ± 1.76 µM in AsPC-1. Cell cycle analysis showed that TR26 had a significant G0/G1 phase arrest, indicating a deficit of growth factors. In-cell western protein analysis assay confirmed that TR26 and TR33 both inhibit ERK at twice their IC_{50} value. These compound also showed a significant cell proliferation inhibition when studied in lung cancer and breast cancer in our group and demonstrated that these heterocyclic analogs are key precursor molecules for optimized analogs that may be potent pancreatic cancer drugs. They were able to have low micromolar cytotoxicity in both lung and breast cancer, as well as were able to moderately reverse drug resistance to vincristine. Molecular dynamic simulations were performed to confirm biological evaluations of hit compounds and found that TR26 binding to ERK was complimentary to its biological activity, while binding in TR33 was not as clear, possibly due to limitations to dynamic software. Ultimately, we've determined that these heterocyclic analogs act as key anticancer agents to pancreatic cancer, and show cytotoxicity to breast cancer and lung cancer as well. This suggests that they, and future development of analogs of their type, could be useful as a potential option for anticancer therapy.

3.6 References

1. Hezel, A. F.; Kimmelman, A. C.; Stanger, B. Z.; Bardeesy, N.; Depinho, R. A., Genetics and biology of pancreatic ductal adenocarcinoma. *Genes Dev* **2006**, *20* (10), 1218-49.

2. Siegel, R. L.; Miller, K. D.; Jemal, A., Cancer statistics, 2018. *CA Cancer J Clin*, **2018**, *68* (1), 7-30.

3. Rahib, L.; Smith, B. D.; Aizenberg, R.; Rosenzweig, A. B.; Fleshman, J. M.; Matrisian, L. M., Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res* **2014,** *74* (11), 2913-21.

4. Maitra, A.; Hruban, R. H., Pancreatic cancer. *Annu Rev Pathol* **2008,** *3*, 157-188.

5. American Joint Committee on Cancer. Exocrine Pancreas. In: AJCC Cancer Staging Manual. 8th ed. New York, NY: Springer; 2017:337.

6. Pancreatic Cancer UK. https://www.pancreaticcancer.org.uk/information/just-diagnosed-with-pancreatic-cancer/what-is-the-pancreas/ (accessed 10/30/22).

7. Pancreatic Cancer Action. https://pancreaticcanceraction.org/about-pancreatic-cancer/treatment-options-for-pancreatic-cancer/treating-pancreatic-cancer/surgery-for-operable-pancreatic-cancer/whipples-procedure-and-pylorus-preserving-pancreatoduodectomy-pppd/ (accessed 10/30/22).

8. Karim, S. A. M.; Abdulla, K. S.; Abdulkarim, Q. H.; Rahim, F. H., The outcomes and complications of pancreaticoduodenectomy (Whipple procedure): Cross sectional study. *Int J Surg* **2018,** *52*, 383-387.

9. De Dosso, S.; Siebenhuner, A. R.; Winder, T.; Meisel, A.; Fritsch, R.; Astaras, C.; Szturz, P.; Borner, M., Treatment landscape of metastatic pancreatic cancer. *Cancer Treat Rev* **2021,** *96*, 102180.

10. Loos, M.; Kleeff, J.; Friess, H.; Buchler, M. W., Surgical treatment of pancreatic cancer. *Ann N Y Acad Sci* **2008,** *1138*, 169-80.

11. Venkatesulu, B. P.; Hsieh, C. E.; Sanders, K. L.; Krishnan, S., Recent advances in radiation therapy of pancreatic cancer. *F1000Res* **2018,** *7*.

12. Mizrahi, J. D.; Surana, R.; Valle, J. W.; Shroff, R. T., Pancreatic cancer. *The Lancet* **2020,** *395* (10242), 2008-2020.

13. Wang, W. B.; Yang, Y.; Zhao, Y. P.; Zhang, T. P.; Liao, Q.; Shu, H., Recent studies of 5-fluorouracil resistance in pancreatic cancer. *World J Gastroenterol* **2014,** *20* (42), 15682-90.

14. Barton-Burke, M., Gemcitabine: A pharmacologic and clinical overview. *Cancer Nurs* **1999,** *22* (2), 176-183.

15. Muler, J. H.; McGinn, C. J.; Normolle, D.; Lawrence, T.; Brown, D.; Hejna, G.; Zalupski, M. M., Phase I trial using a time-to-event continual reassessment strategy for dose escalation of cisplatin combined with gemcitabine and radiation therapy in pancreatic cancer. *J Clin Oncol* **2004,** *22* (2), 238-43.

16. Conroy, T.; Desseigne, F.; Ychou, M.; Bouche, O.; Guimbaud, R.; Becouarn, Y.; Adenis, A.; Raoul, J. L.; Gourgou-Bourgade, S.; de la Fouchardiere, C.; Bennouna, J.; Bachet, J. B.; Khemissa-Akous, F.; Pere-Verge, D.; Delbaldo, C.; Assenat, E.; Chauffert,

B.; Michel, P.; Montoto-Grillot, C.; Ducreux, M., FOLFIRINOX versus Gemcitabine for Metastatic Pancreatic Cancer. *N Engl J Med* **2011,** *364* (19), 1817-1825.

17. Oliveira-Cunha, M.; Newman, W. G.; Siriwardena, A. K., Epidermal growth factor receptor in pancreatic cancer. *Cancers (Basel)* **2011**, *3* (2), 1513-1526.

18. Morgan, M. A.; Parsels, L. A.; Kollar, L. E.; Normolle, D. P.; Maybaum, J.; Lawrence, T. S., The combination of epidermal growth factor receptor inhibitors with gemcitabine and radiation in pancreatic cancer. *Clin Cancer Res* **2008,** *14* (16), 5142-5149.

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

20. Ahmed, M. S.; El-Senduny, F.; Taylor, J.; Halaweish, F. T., Biological screening of cucurbitacin inspired estrone analogs targeting mitogen-activated protein kinase (MAPK) pathway. *Chem Biol Drug Des* **2017,** *90* (3), 478-484.

21. Kopel, L. C.; Ahmed, M. S.; Halaweish, F. T., Synthesis of novel estrone analogs by incorporation of thiophenols via conjugate addition to an enone side chain. *Steroids* **2013,** *78* (11), 1119-25.

22. Mahnashi, M.; Elgazwi, S. M.; Ahmed, M. S.; Halaweish, F. T., Cucurbitacins inspired organic synthesis: Potential dual inhibitors targeting EGFR - MAPK pathway. *Eur J Med Chem* **2019,** *173*, 294-304.

23. Alseud, K, M. Design, Synthesis, and Biological Screening of Novel Cucs-Inspired Estrone Analogs Towards Treatment of Pancreatic Adenocarcinoma. Ph.D. Dissertation. South Dakota State University, Brookings, SD 2019.

24. Cell Signaling Technology, G1/S Checkpoint. https://www.cellsignal.com/pathways/g1-s-checkpoint-pathway. (accessed 11/10/22).

25. Dubravka, D.; Scott, D. W., Regulation of the G1 phase of the mammalian cell cycle. *Cell Res* **2000,** *10*, 1-16.

26. Cuddihy, A. R.; O'Connell, M. J., Cell-Cycle Responses to DNA Damage in G2. *Int Rev Cytol* **2003,** *222*, 99-140.

27. Ahmed, M. S.; Halaweish, F. T. Cucurbitacins: potential candidates targeting mitogen activated protein kinase for treatment of melanoma. *J Enzyme Inhib Med Chem*, **2014**, *29*, 162-167.

28. Rockland antibodies and assays. In-Cell Western Protocol. https://rockland-inc.com/uploadedFiles/Support/Protocols/In_cell_Western_Protocol_v3.pdf.

29. Rasband, W. S., ImageJ, U. S. National Institutes of Health, Bethesda, Maryland, USA, https://imagej.nih.gov/ij/, 1997-2018.

30. *Desmond* version 2.2. Schrödinger. 120 West 45th Street, 17th Floor. New York, NY, 2009.

Chapter Four

Design and Synthetic Optimization of Novel 11-amino based Estrone Analogs

Abstract

Pancreatic cancer continues to have abysmal outlooks for patients, and a new therapy is needed to combat this aggressive and difficult to treat disease. First round heterocyclic estrone analogs proved moderately effective against pancreatic cancer analogs *in-vitro*, and modification with C-11 oxygenation increases their effectiveness. Using Bioisosterism we theorize that the installation of a nitrogenous group at this carbon will increase their effectiveness due to their increased hydrogen bonding capabilities and similar atomic properties. Molecular modeling was performed with 650+ analogs created and docked within various EGFR pathway proteins; in specific those of the ERK/MAPK pathway and STAT3 pathway. Various C-11 amine analogs showed promising binding activities *in-silico* as opposed to known pancreatic cancer standard Gemcitabine, as well as to previously synthesized analogs tested against pancreatic cancer (2-5 fold better). This prompted their synthetic creation. Optimization of the C-11 ketone intermediate was first performed to limit the use of harsh chemicals which cause lower yields, such as hydroboration agents, as well as those that produce multiple by-products. Many iterations were then performed to obtain the C-11 amine from the C-11 ketone, and eventually it was obtained through installation of a benzylamine intermediate followed by hydrogenation. Further aspects of this project entail the discussion for full scale synthesis.

4.1 Introduction

As we previously mentioned, pancreatic cancer is the fourth deadliest cancer in the United States with 5 year survival rates being among the worst of all cancers at 8%[1,2]. Difficulty in detection of the disease along with the speed at which it develops owes to its aggressive nature[3]. Treatment options for pancreatic cancer depend on the extent of the disease, but most commonly include surgical resection[4,5], radiotherapy[6], and chemotherapy[7]. Surgical resection is often difficult and has a number of complications, while chemotherapy regiments do improve survivability of patients, albeit only by 6-12 months on average[8,9].

New therapy options are needed for pancreatic cancer, and one of the strongest candidates exists as tyrosine kinase inhibitors (TKI's) responsible for inhibiting growth factor receptors like the epidermal growth factor receptor (EGFR). Treatment options including EGFR inhibitors have been widely studied in pancreatic cancer[10,11] but often succumb to eventual drug resistance. Three generations of TKI's have been developed, but resistance by various mutations, exon changes, and protein overexpression has become commonplace in all of them[12-15], and a new treatment option is needed to combat this resistance.

We've previously discussed the synthesis of heterocyclic estrone analogs (Figure 4.1A) combating pancreatic cancer and their potential usefulness as a possible precursor analog. They demonstrated micromolar cytotoxicity against pancreatic cancer cell lines AsPC-1 and Panc-1 while also displaying the ability to moderately reverse drug resistance of vincristine, validating their potential as a precursor. A study conducted by Alseud (2019)[16] demonstrated a second series of pancreatic cancer analogs with further modification of

estrone structure at C11 of the steroid core (Figure 4.1B) have significantly better IC_{50} values against pancreatic cell lines AsPC-1, Panc-1, and BxPC-3. Analogs with oxygenated derivatization have improved mechanism of actions against EGFR and downstream proteins, and are also cytotoxic in 3D cell culture assays.

As we saw increased cytotoxicity with 11-oxygenated derivatives we theorized that perhaps 11-nitrogenous derivatives (Figure 4.1C) may also increase cytotoxicity due to increased hydrogen bonding capabilities of amines versus hydroxyl groups as well as increased bioavailability with a suggested lower LogP value. In specific we apply the idea of Bioisosterism, in which atoms or groups of atoms with similar chemical properties will behave in similar manners. We use this rationale to suggest that the oxygen vs nitrogen debate has merit and rationale; and its simplicity warrants these facts. Their similar van der Waals radius, valency, and spatial arrangement provide key parameters for successful Bioisosterism as shown in many works[46]. This work reports the molecular modeling study of 11-nitrogenous heterocyclic estrone analogs and the optimization towards their synthesis. We report the high binding scores of 11-nitrogenous analogs, as well as the difficulty in placing the 11-amino group on the steroid core during our optimization process.

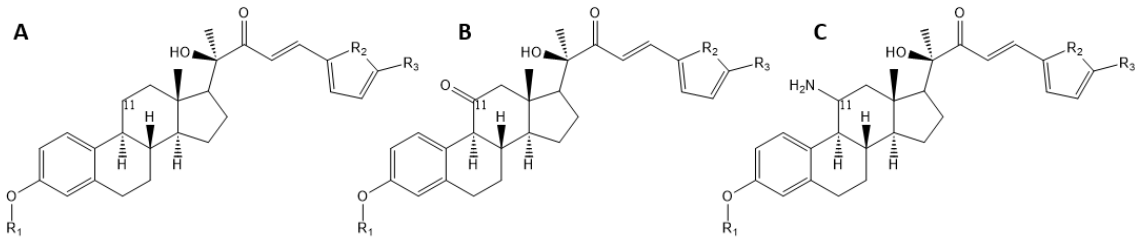


Figure 4.1 Generalized structures of optimized analogs for use in pancreatic cancer. A) First-round analogs with no C-11 modification. B) Second-round analogs with oxygenated C-11. C) Third-round analogs with nitrogenated C-11.

4.2 Molecular Modeling Design Strategy

We've briefly discussed previously the use of molecular modeling as a predictive and conformation-based tool to measure the ability of a ligand binding to a protein. Herein we discuss the greater role molecular modeling has to play in the design of new analogs expanded on thoughts from the introduction chapter. Molecular modeling itself is useful in providing expanded insights into how a ligand will interact with a protein of interest. Two of the most common types of molecular modeling and drug design is that of structure-based drug design and ligand-based drug design. Structure-based drug design implies the identification of a target with a known structure, typically the 3D crystal structure of a protein in relation to molecular modeling. Structural alterations are made to an existing ligand to create new ligands with specific electrostatic and stereospecific properties pertaining to the binding site. Upon identification of a lead, the process begins anew with further optimization[17]. Ligand-based drug design typically involves the use of structure activity relationships (SAR) to modulate the structure of a ligand in absence of a known target structure[18]. Known ligands that bind to the target are modulated using existing pharmacophores in hopes of finding increasing binding. The latter is useful to us in this study to modulate the existing steroid structure.

There are three main types of molecular modeling in which ligands are docked inside of a protein. The first uses flexible docking in which the protein and the ligand have

multiple conformations in hopes of finding the best fit for binding[19]. This is often computer intensive however due to the need to allow for protein and ligand movement during the simulation. Semi-flexible docking involves a rigid protein and the use of flexible ligand conformations to modulate the movement of the ligand[20]. The third type of docking is rigid docking, in which both protein and ligand are rigid structures. Often however this type does not allow for accurate representation of molecular binding. Due to the large number of molecules that need to be looked at and limitations in computer speed, usually the semi-flexible option is best and will be discussed moving forward.

Our molecular modeling strategy utilizes a variation of the semi-flexible docking output, in which protein receptors are kept rigid and conformers of different ligands are used to measure binding. We also use an energy optimization to create the lowest energy ligands. Analogs are first designed using ChemDraw® 2D visualization software. Analogs are drawn with varying functionalities and conformations and each given a unique identifier. To model these for docking, analogs are taken into ChemBio3D Ultra 12.0 to visualize their 3D aspects[21]. ChemBio3D saves the 3D coordinates of each atom and uses them within the docking system. To better describe the likelihood of a molecule in the biological system, an energy optimization is performed to model the ligand. A Merck Molecular Force Field (MMFF94) calculation is used to lower energy of the molecule[22]. This force field calculation utilizes location and identity of the atoms to predict the molecule's cumulative potential energy over a number of iterations until an RMS gradient of 0.1000 or less is achieved. Bond movement, torsional strain, angle strain, and non-bonded interactions are all taken into account among other items. The molecular modeling simulation is performed using OpenEye® Scientific Software programs OMEGA[23,24],

which account for conformation generations, and FRED[25] (fast rigid exhaustive docking). Crystal structures of proteins are generated using MakeReceptor® Program[26].

4.2.1 Molecular Modeling Scoring Functions

Outputs of the molecular modeling software are known as scoring functions. OpenEye® toolkits utilize five scoring functions known as shapegauss, chemgauss3, oechemscore, screenscore, and plp[26]. The score of a ligand is the best score of any pose of said ligand. Shapegauss utilizes a shape based score which focuses on the active site pocket shape and atom shapes. Scores are higher when atoms are touching but not overlapping, suggesting interaction between atoms. Chemgauss3 uses gaussian potentials to measure shape, hydrogen bonding between ligand and protein, hydrogen bonding with solvent, and any metal chelating factors. Oechemscore measures interactions regarding lipophilicity, hydrogen bonding, metal chelators, clashing, and rotatable bonds. Screenscore measures rotatable bonds and hydrogen bonding. Plp is a pairwise function that measure hydrogen bonding and shape characteristics between the ligand and pocket. The cumulative scores are output for each ligand, and the results themselves are trivial on their own. To combat this, FRED outputs a consensus score. This score is a numerical value that takes into account each factor previously mentioned. The lower this score, the better binding affinity of the ligand to the protein.

Due to the large success of the 11-oxygenated analogs, we attempted to design 11-nitrogenous derivatives in sequential libraries to determine how they bind to key EGFR pathway proteins. We herein report the results of those docking studies, the lead 11-amino analogs, and the optimization for the synthesis of these analogs.

4.3 Results and Discussion

4.3.1 Molecular Modeling

A virtual library of 465 estrone analogs were designed with modifications at C3, C11, C16, and the heterocyclic side chain (Figure 4.2). Additionally, Gemcitabine, the chemotherapy for pancreatic cancer, as well as cucurbitacins A, B, D, and E, were analyzed as modeling standards. Modifications at C3 were kept from the previous heterocyclic estrone analogs in chapter 3, as well as the side chain modifications. The additional changes were made in two arrangements. We modified C16 due to the results seen in Alsayri et al (2017) in which estrogenic effects were modulated with C16 modifications[27]. Finally, modifications were made at C11 to see the difference between the three generations of pancreatic analogs. It must be noted that aside from a general 11-amino group there are likely additional pharmacophores that may be potent analogs for modeling study. We concede this fact and discuss the creation of roughly 200 new analogs with C11 modifications. Analogs possessing both mono- and di- alkyl substrates on the amino group were created ranging from methyl- to butyl- groups. We also looked at the creation of 11-acetamide and 11-nitrile species to see if adding some carbon structure between C11 and the amine group placed a role in its binding.

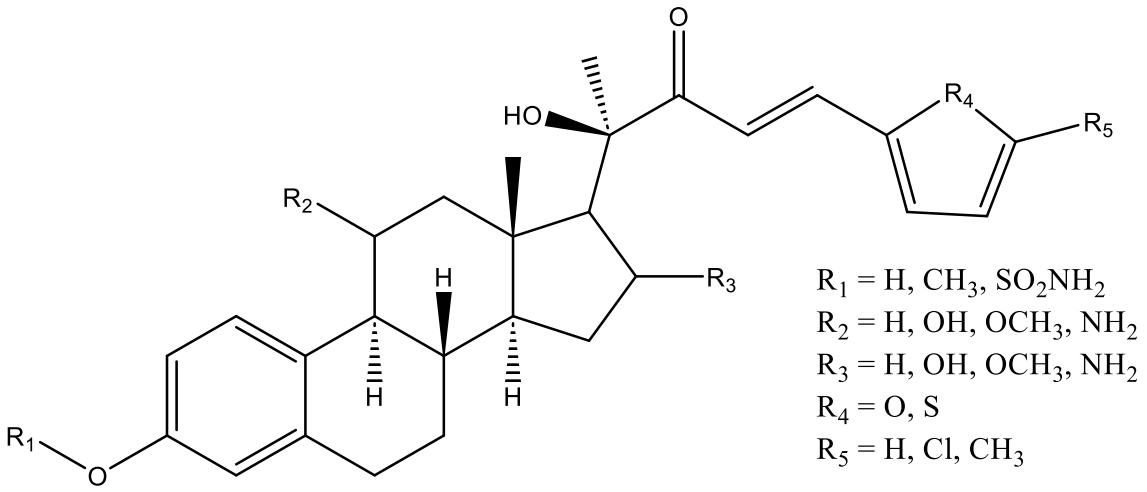


Figure 4.2. Structure of C11 and C16 substituted third-round heterocyclic estrone analogs designed for modeling in EGFR proteins.

The aforementioned 650 estrone analogs were docked within 11 proteins of the EGFR pathway: EGFR (PDB ID: 1M17), Pi3K (3L54), Akt (3MV5), mTOR (4JSV), JAK1 (4EI4), JAK2 (6VNK), STAT3 (1BG1), RAS (1X1R), RAF (5FD2), MEK (3W8Q), and ERK (2OJG) along with heat shock proteins 70 (4IO8) and 90 (5CF0). Consensus scores for the top 100 analogs were analyzed along with the standards Cuc A, B, D, E and gemcitabine for each protein, and some interesting results were determined. Two sets of analogs showed generally good binding scores across the proteins of interest and are worth discussing. The first set (Figure 4.3A) have phenolic moieties and the C11 amine, with variation in C16 and the side chain. The second set (Figure 4.3B) also have the phenolic group, suggesting its role in binding is critical, along with a ethyl substituent on the C11 amine. The structures of these analogs can be viewed in Figures 4.4 and 4.5, and the standards in Figure 4.6.

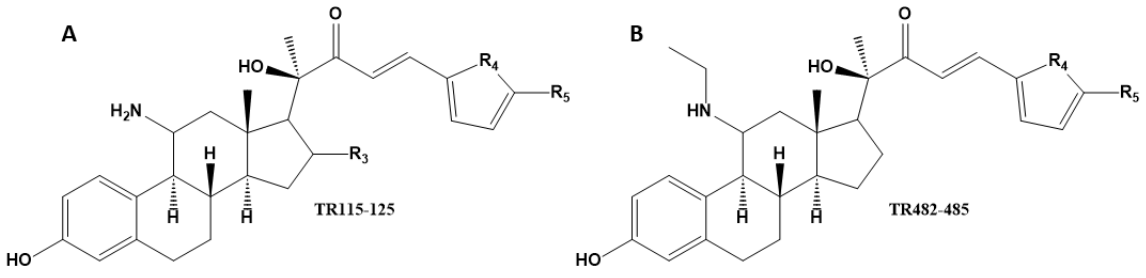


Figure 4.3 A) General structure of TR115-125 and B) General structure of TR482-485. These were the top compounds from the molecular modeling study.

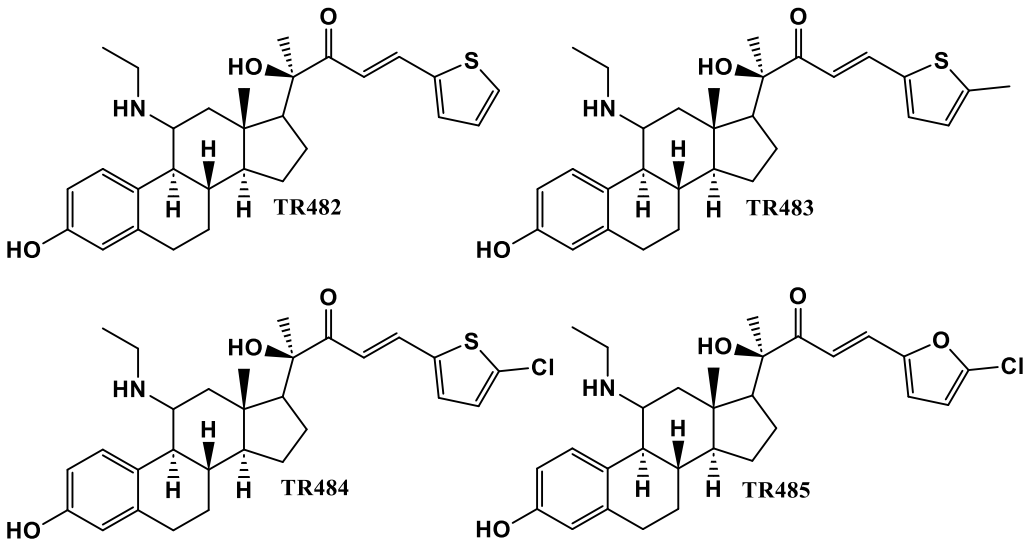


Figure 4.4 Structures of TR482-TR485.

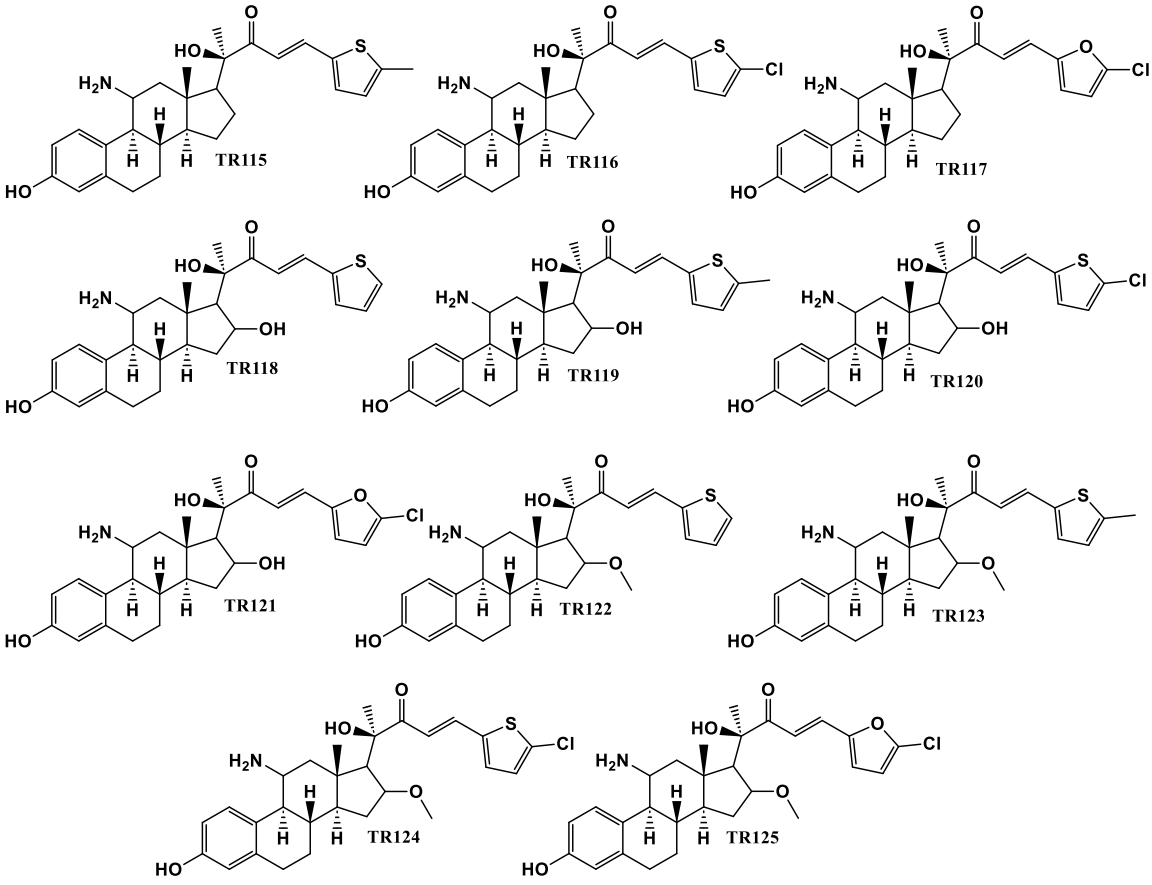


Figure 4.5 Structures of TR115-125.

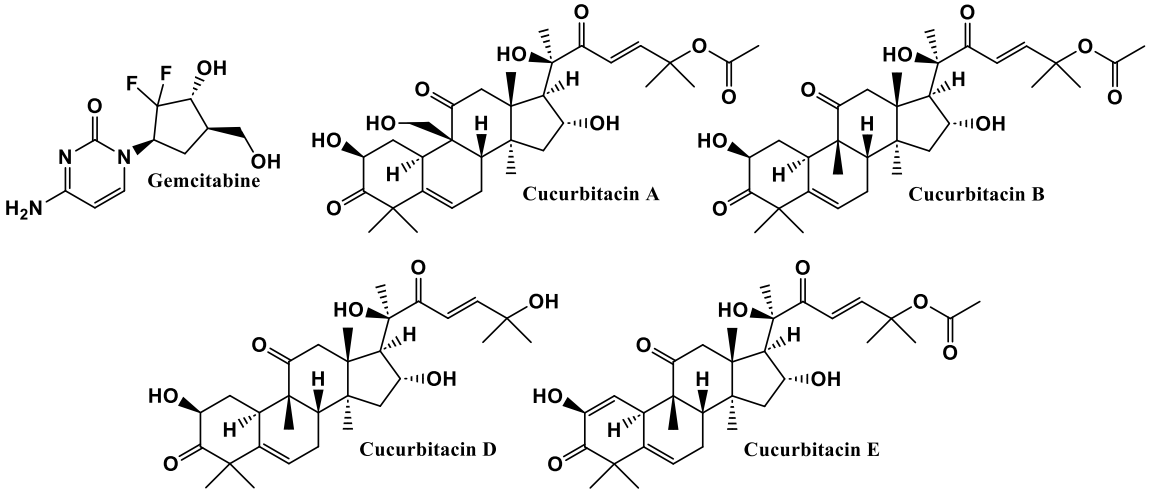


Figure 4.6 Structures of docking standards Gemcitabine and Cucurbitacins A,B, D, and E.

Consensus scores for the above compounds can be seen in Table 4.1 To better represent the data for the C11 amine analogs, we have only included scores for the top 100 analogs in each protein. The standards scores are represented regardless. Scores highlighted in red are the top scoring analog in that protein. Ultimately, we see that phenolic C3 analogs are of preference in most proteins, likely due to hydrogen bonding capability by the phenolic hydrogen. For example, in MEK, we see this hydrogen bond to MET-146, and in ERK we see it to ASP-104 and MET-106. Additionally, the amine groups are favorable over the oxygenated groups in almost all proteins tested, furthering the discussion that these analogs are worth synthesizing. We wish to highlight the results for a few specific proteins. In ERK, all C11 amine analogs listed in the table aside from TR485 score in the top 30 of analogs tested. In MEK, TR115-TR125 are all scoring in the top 25 of analogs tested. In RAS, TR118-TR121, which all possess the C16 hydroxyl group, are in the top 5 analogs tested. When we look into the docking for these analogs, we see that all four possess a key hydrogen bond between the 11-amino group and the Val-24 backbone amide (Figure 4.7). Finally, in mTOR, the four ethylamine analogs TR482-TR485 are the top 4 scoring analogs in the protein. Analyzing the side chain preference is needed as well, but there seems to be no significant difference seen between any of the analogs tested in most proteins. We can see that specific analogs behave very well in specific proteins, suggesting that any of the abovementioned analogs may show biological activity when tested. As most of them show high scores in multiple proteins, it suggests their synthesis is warranted.

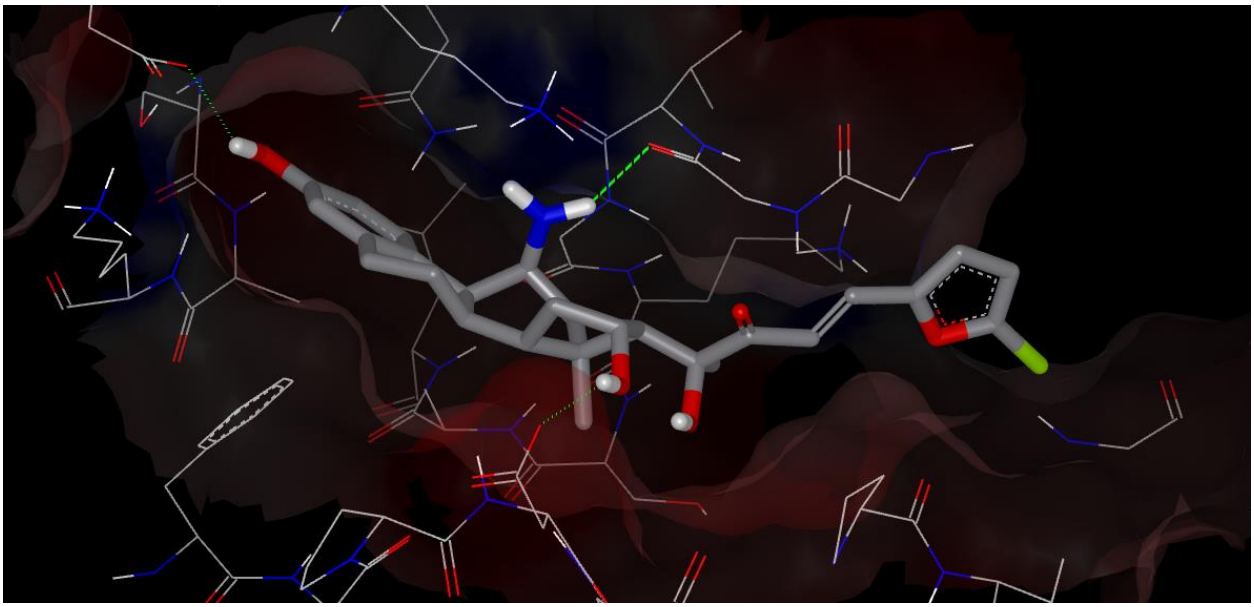


Figure 4.7 TR121 docked in RAS active site. Hydrogen bonding exists for TR118-121 between 11-amino group and Val-24 amide backbone. Additional hydrogen bonding occurs between phenolic group and SER-156 and C-17 hydroxyl group with ASP-41. Hydrogen bonds shown in green.

To explore the binding effects of some of these analogs, we wish to highlight the docked images within protein active sites. Figure 4.8 highlights TR115 in the MEK active site. MEK inhibition has shown to be promising for estrone analogs in previous studies in our laboratory[28,29]. We can notice a few key features that explain why TR115 has a low consensus score. First, phenolic hydrogen bonding to MET-146 backbone, as well as a hydrogen bond from the 11-amino group to SER-194. Both of these touch on key events we discussed earlier in which the phenolic group and the 11-amino group are important for binding. Additionally, we also see that the oxygenated groups of the C17 side chain, particularly the α-hydroxyl methyl ketone moiety, have electrostatic interactions with the lysine rich MEK active site.

Table 4.1 Consensus scores for high binding 11-amino analogs in various targets. Only scores for top 100 analogs are shown.

Analog ID	AKT	RAF	EGFR	ERK	HSP70	MEK	RAS	mTOR	Pi3K
TR115	183	-	-	132	99	51	411	419	434
TR116	266	471	340	146	178	51	-	-	-
TR117	-	-	-	189	126	151	340	485	-
TR118	266	-	466	119	163	133	90	414	227
TR119	320	-	390	55	161	70	18	-	-
TR120	287	-	-	59	-	53	69	-	139
TR121	-	-	394	193	102	124	55	378	298
TR122	103	-	-	51	274	139	394	299	-
TR123	72	456	474	34	374	161	318	-	-
TR124	145	-	-	38	-	130	319	356	-
TR125	179	-	-	16	-	131	-	445	-
TR482	122	48	416	76	-	-	197	4	85
TR483	52	25	169	147	106	-	426	37	82
TR484	90	239	214	234	388	-	-	39	51
TR485	-	202	93	423	-	-	221	63	270
Gemcit.	930	1284	1225	1011	729	928	896	885	949
Cuc A	861	886	1111	865	295	1129	1067	879	1244
Cuc B	907	542	1143	939	539	807	937	802	802
Cuc D	1026	867	1080	912	420	668	986	899	873
Cuc E	1023	811	1048	837	513	685	992	984	621

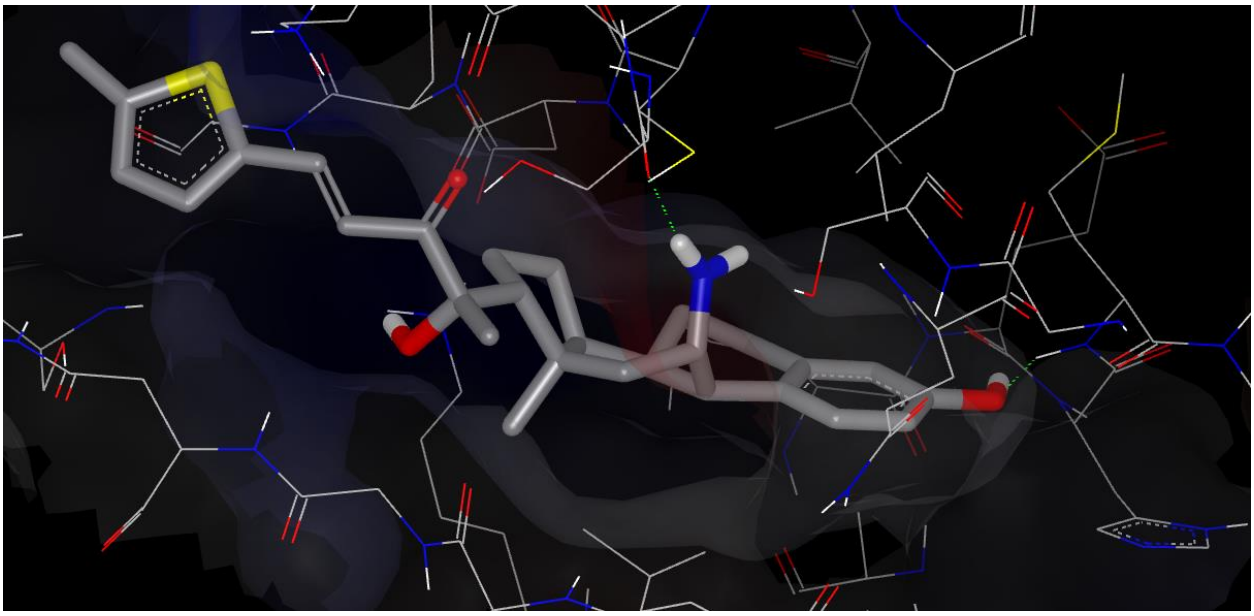


Figure 4.8 TR115 docked in MEK active site. Hydrogen bonding occurs between 11-amino group and SER-194 and between phenolic group and MET-146. Hydrogen bonds shown in green.

4.3.2 Synthetic Strategy and Optimization

As we've seen the potential usefulness of these 11-amine analogs, their synthetic creation was seen as imperative. We will discuss the literature below for the installation of the 11-nitrogenous group, but a retrosynthetic analysis for the entire molecule provides a more apt discussion (Figure 4.9). Regardless of the installation of a C16 group, the final step(s) of the synthesis would be addition of the heterocyclic group via aldol condensation and any deprotection steps. For simplicity, we will assume the C11 is a primary amine, but could be ethylated to form TR482-TR485 by simple substitution. Before these final steps, creation of the α-hydroxyl methyl ketone as reported[30,31] would be performed from the C17

ketone. For adding the C16 functional group in TR118-TR125 we would follow the report from Alsayari et al. (2017) for installing hydroxyl groups form the C17 ketone[27]. This could be protected with a TBDMS group easily if installed.

The most important step, the installation of the amine, was theorized to be formed from either the 11-hydroxy or 11-keto variant seen in Alseud et al. (2019)[16]. This task was increasingly more monumental of a feat then we initially endeavored to think as the synthesis on an estrone core had not been performed before, and the synthetic optimization for synthesizing this amine will be discussed in great detail. The 11-hydroxy/11-keto species synthesized in Alseud et al. (2019), while easily enough to synthesize, provided low yields due to hydroboration steps, and utilized chemicals with side products difficult to remove. The retrosynthetic plan to form this 11-hydroxy/11-keto group was first revisited from the readily available starting material estrone.

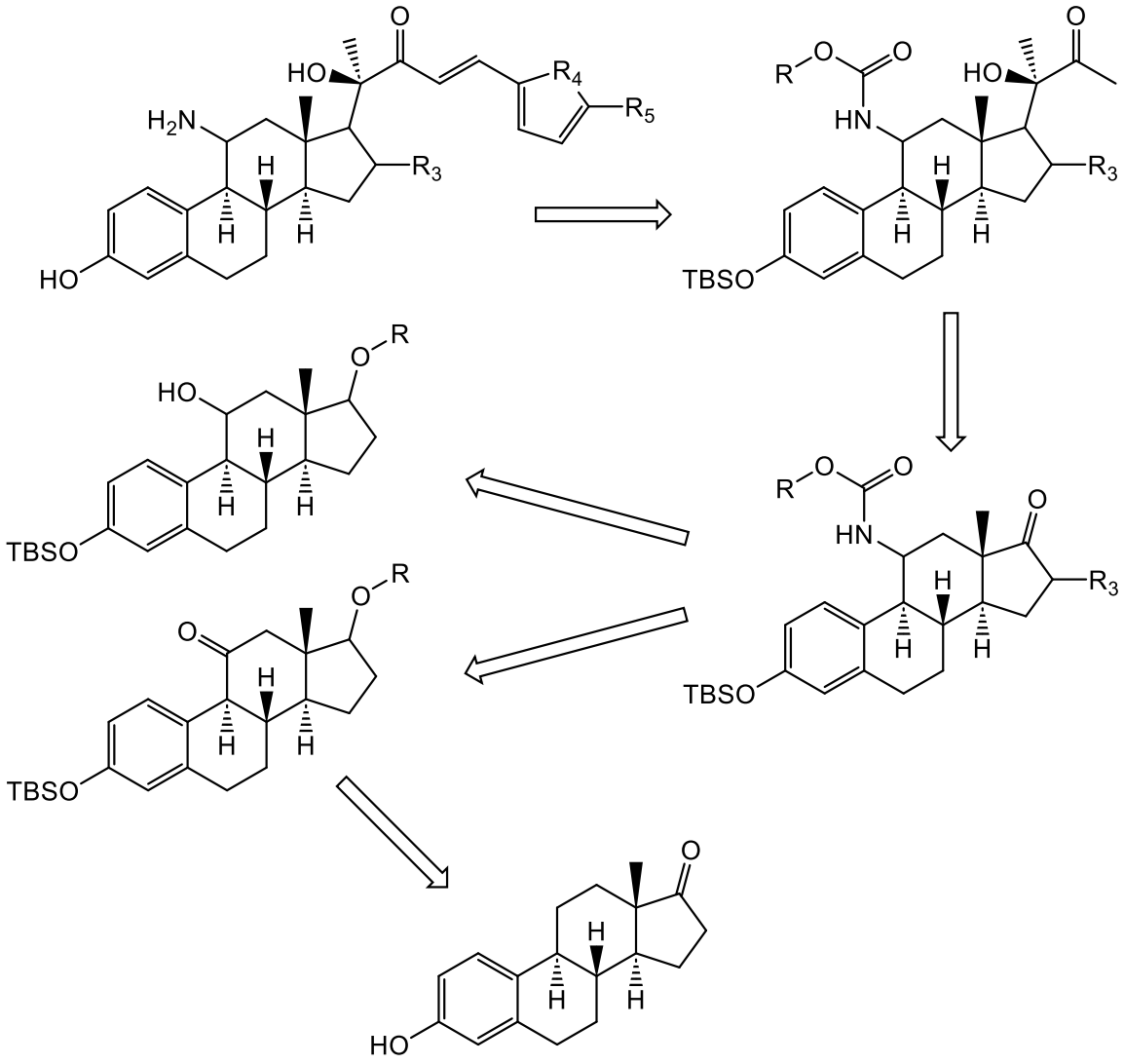
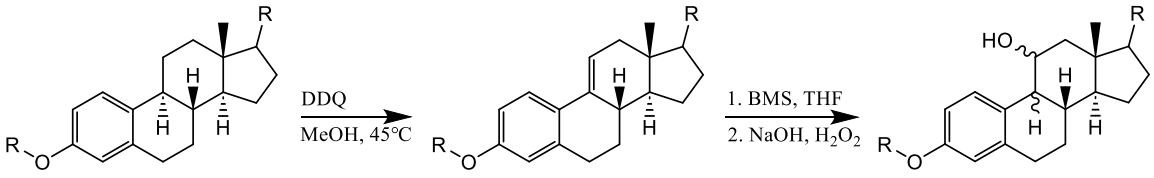


Figure 4.9 Retrosynthetic plan for heterocyclic C-11 amino estrone analogs.

Revisiting 11-hydroxy and 11-keto synthesis

Current synthesis of the 11-hydroxy intermediate exists as seen in Scheme 4.1. Estrone derivatives are oxidized selectively at the Δ9,11 position with 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) followed by hydroboration oxidation at position 11 to yield the 11-hydroxy steroid[32]. This 11-hydroxy steroid can then easily be oxidized to 11-keto

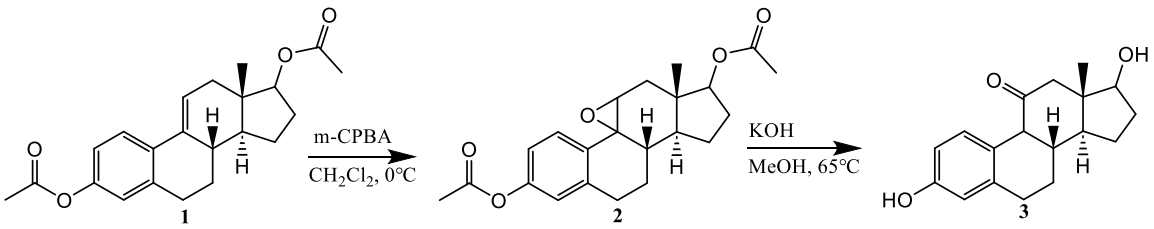
groups. The DDQ reaction is quite a unique one, as it allows for selective oxidation of the steroid structure due to the phenolic group conjugated with the Δ9,11 position. The mechanism for this is thought to be due to hydride transfer from the benzylic hydrogen on carbon 9 with following rearrangement of the benzene ring[33,34]. Ultimately, the ring is reformed and the double bond is placed. The drawbacks of this reaction lie in the side-products formed, namely the reduced form of DDQ, known as DDH. Soxhlet extraction is typically performed to remove DDH[32], but even so not all of the DDH is typically removed, and the following reactions typically garner many side-reactions. The sequential reaction, hydroboration-oxidation, is well known for causing competing reactions due to oxidation at competing sites of the borane species. The resulting yield for the enantiomers is around 60-70% in good cases, and even less so when a pure product is needed. As such, we looked into new ways to create the 11-hydroxy species and 11-keto species that removed one or both of these reactions.



Scheme 4.1 Current synthesis for 11-hydroxy steroids using DDQ and hydroboration-oxidation.

To negate the poor yields and side-products of hydroboration oxidation, we conducted experiments using the Prilezhaev reaction, in which a peroxy acid is used in conjunction with a double bond to yield an epoxide. The resulting epoxide (2), when generated on the

Δ9,11 double bond, can be opened using potassium hydroxide (Scheme 4.2) to yield the corresponding 11-keto steroid (3)[35]. The resulting reaction must be performed with a free phenolic anion, which can be achieved using an acetoxy- protected phenol on the steroid. This mechanism is discussed in detail in scheme 4.4. The reaction(s) themselves result in a cumulative 40% yield, albeit with less side products. Likely this is due to the steric hindrance of the double bond being used for the peroxy acid (as it does as well in hydroboration oxidation). The oxidation however occurs almost selectively at C-11 due to the mechanism of the reaction.

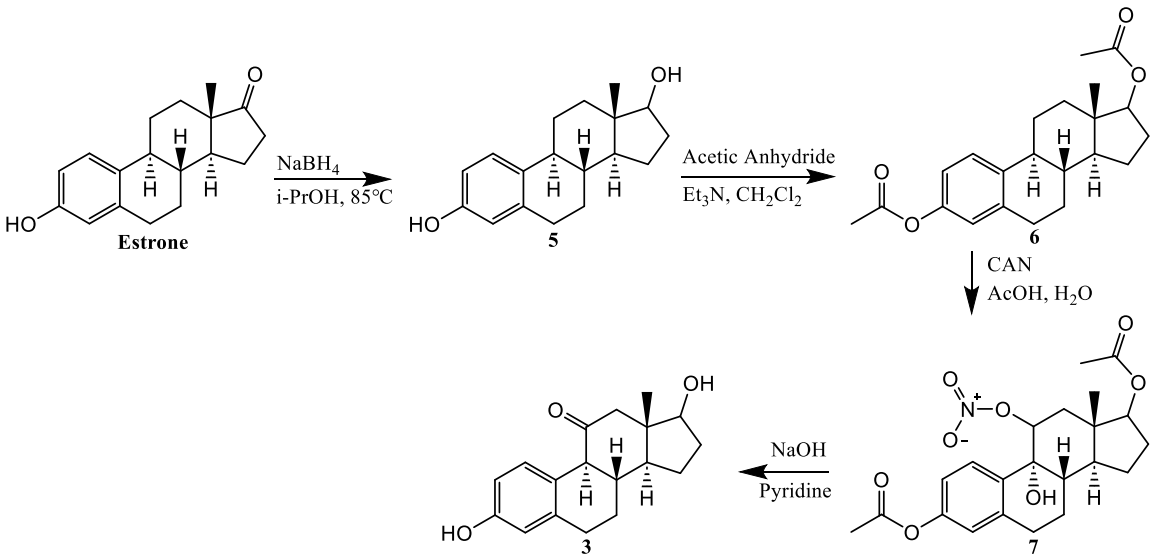


Scheme 4.2 Reaction scheme detailing the formation of 11-keto steroids using a Prilezhaev reaction followed by base-catalyzed epoxide opening.

We wanted to see if we could prevent the use of both the DDQ and hydroboration oxidation, and to do so came across an article detailing the same drawbacks we pose here. They discuss the use of 11-substitued nitrates and their reactivity to form 11-keto steroids[36]. Scheme 4.3 details the reactions performed to achieve this 11-keto steroid without using either DDQ or hydroboration oxidation. Similarities appear between this and the previous epoxide reaction, in which the free phenolic anion must be present for the reaction to proceed. In order to perform this set of reactions, C17 must first be protected. However,

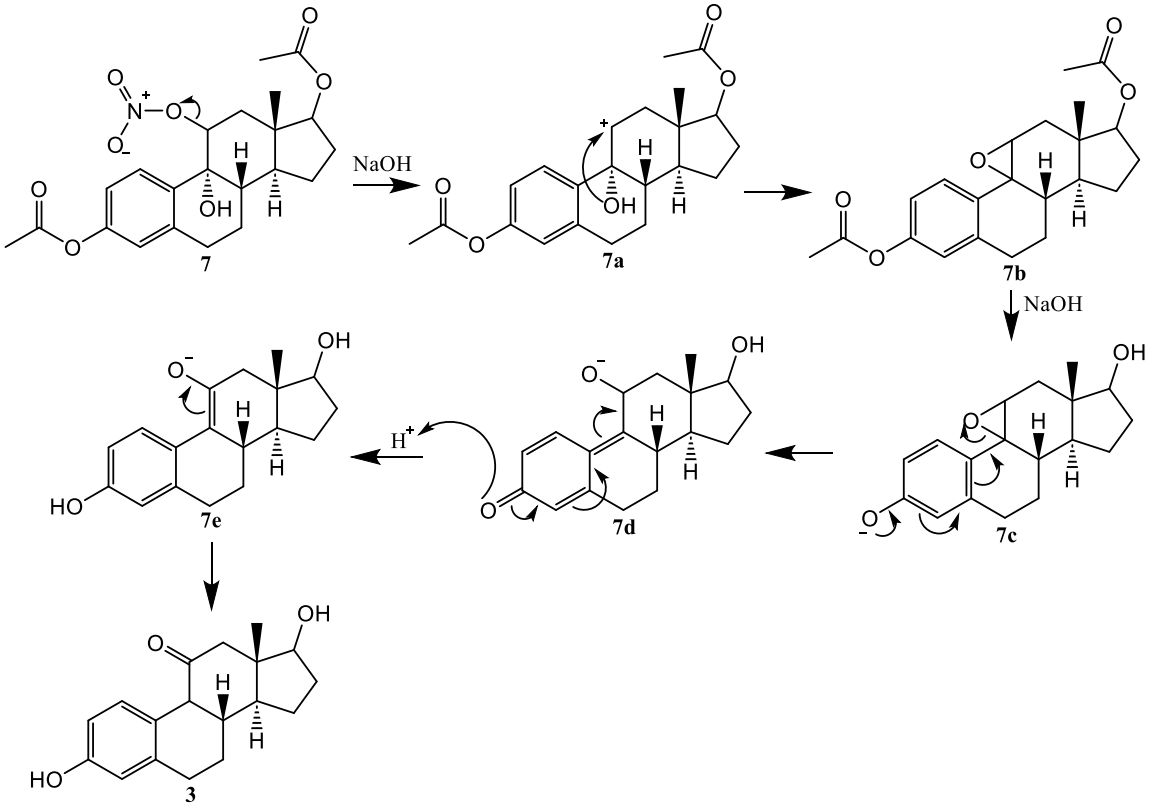
simple protection of the ketone via ketal protecting group is not reliable, as acid is used in the scheme, and would cleave the ketal, rendering a diketo- product that would be difficult to selectively target. The result is then reduction of the C17 ketone (4), followed by protection using the same acetoxy group as the phenolic group to yield (6). This reaction can be performed for both alcohols with 2 equivalents of acetic anhydride and enough reaction time.

The diacetyl steroid (6) is then reacted with cerium ammonium nitrate (CAN) in the presence of acetic acid to form the 9-hydroxy-11-nitroso intermediate (7). Base catalyzed reaction is then performed to form the epoxide *in-situ*, followed by opening of the epoxide in a one-pot reaction to form the 11-keto species (8). The mechanism of this reaction can be seen in Scheme 4.4. Ultimately, the entire reaction scheme provides yields similar to that of hydroboration-oxidation (50-65%) with either reactions, providing an alternative to form 11-keto steroids.



Scheme 4.3 Reaction scheme detailing the formation of 11-keto steroids via cerium ammonium nitrate induced epoxidation followed by base-catalyzed opening and oxidation.

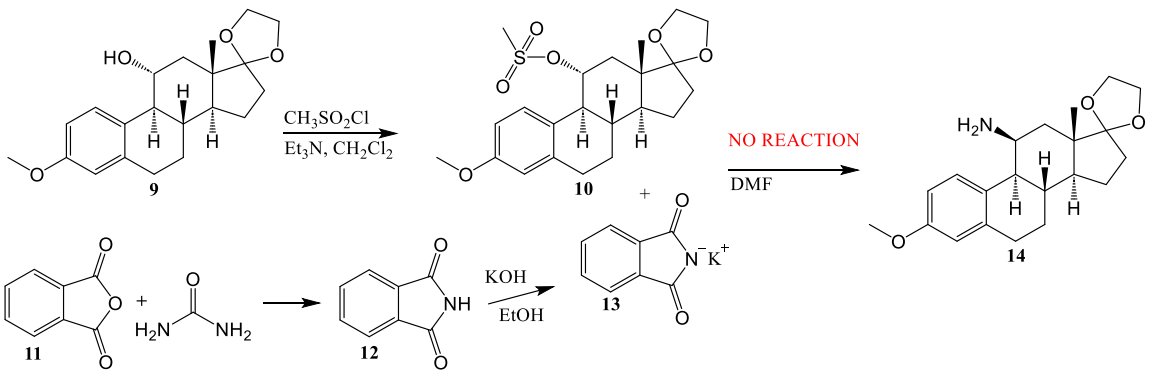
The mechanism for the formation of 11-keto steroids (8) from the 9-hydroxy-11-nitroso intermediate (7) is as follows: The nitroso group is removed with sodium hydroxide, followed by attack of the resulting carbanion (7a) with the 9-hydroxy group to yield the epoxide (7b). Further amounts of sodium hydroxide cleave the acetoxy groups at C3 and C17 to yield phenolic anion (7c). This anion undergoes rearrangement to form the intermediate cyclohexadienone (7d). This intermediate rearomatizes to form the enol (7e) which can then rearrange to form the resulting ketone (8) in scheme 4.4



Scheme 4.4 Mechanism of action for formation of 11-keto steroids from 9-hydroxy-11-nitroso steroids.

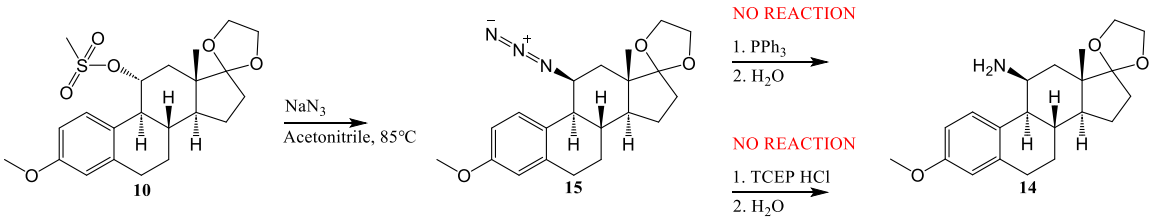
Optimization for 11-amino installation

Upon optimizing the installation of the 11-oxygenated systems, we wished to begin the installation of the nitrogen nucleus onto the C11 carbon. This underwent many trials, most of which were unsuccessful, but warrant discussion. We first tried a modified Gabriel synthesis (Scheme 4.5). To complete this, the C11 hydroxyl species (9) needs to be converted to a potential leaving group. This was successfully performed using installation of a mesylate with methanesulfonyl chloride to yield (10). To perform the Gabriel synthesis, phthalic anhydride (11) was reacted with urea to create potassium phthalimide (13) upon basification with potassium hydroxide. The phthalimide was finally reacted with the mesylate steroid (10) in the presence of DMF, but only yielded the starting material after several attempts of forcing the substitution. This reaction is likely unsuccessful due to the steric hindrance of both the leaving group and the phthalimide nucleophile. Approach of the phthalimide must occur from the top of the steroid, which will clash with the C13 methyl group.



Scheme 4.5 Modified Gabriel synthesis for installation of C11 amine.

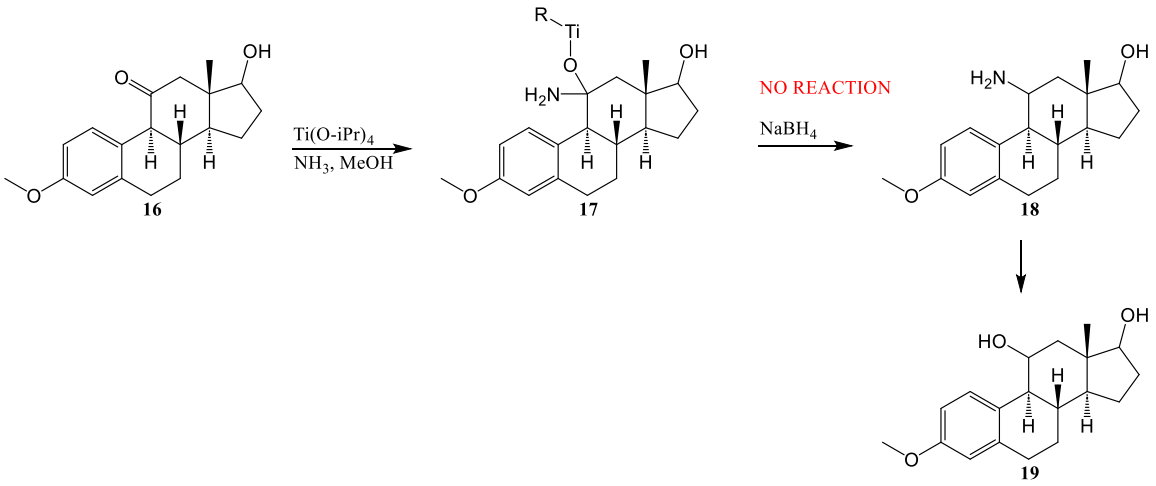
To combat the bulkiness of the phthalimide group, we thought of using a smaller nitrogen nucleophile, such as an azido group (Scheme 4.6). Addition of sodium azide with vigorous stirring refluxed under acetonitrile affords azide (15) from mesylate (10). Reduction of the azide was first explored using a standard Staudinger reaction with triphenylphosphine. This was ultimately unrealistic due to steric strain of the phosphine coordinating with the azide group. The bulkiness of the phosphine promoted the idea of using a smaller phosphine such as tris(2-carboxyethyl)phosphine hydrochloride (TCEP HCl)[37]. We saw no conversion to amine with this new phosphine, and we theorize this may be due to solubility issues with the TCEP. It's possible that this reaction could be pursued using a phase-transfer catalyst, but this optimization was ultimately not explored due to finding of an alternative option for amine introduction that will be discussed below.



Scheme 4.6 Conversion of mesylate to azide, followed by unsuccessful Staudinger reduction.

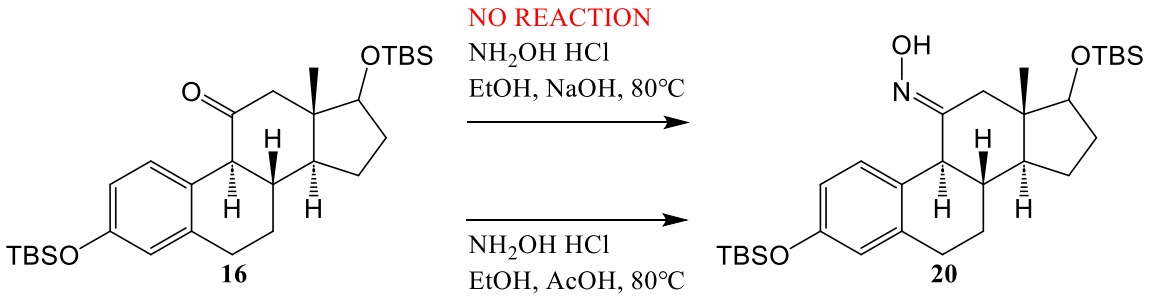
We began to switch gears to try and use the 11-keto group as the promotor for nitrogen introduction. The 11-keto group provides superior ability for nitrogen introduction due to

the lack of a leaving group needed for the nitrogen to attack the C11 carbon. By using a Lewis acid, we can promote the electrophilic nature of the C11 carbon, and subsequently utilize smaller nitrogen nucleophiles to attack the carbon itself. We first tried this using direct amination of the C11 ketone (Scheme 4.7). Direct amination of ketone (16) was performed with a methanolic ammonia solution and titanium isopropoxide, followed by reduction of the intermediate with sodium borohydride according to Miriyala et al. (2004)[38]. When we analyzed the product upon completion of the borohydride step, we determined that we did not isolate the amine, but rather the 11-hydroxyl variant (19), suggesting that the ammonia did not attach to the ketone, and the borohydride reduced the ketone directly. This could be due to the strength of the Lewis acid, or perhaps the bulkiness of it, and we suggest that monitoring the reaction during its progression to determine the mechanistic parameters would be useful.



Scheme 4.7 Scheme for unsuccessful direct amination of C11 ketone with methanolic ammonia.

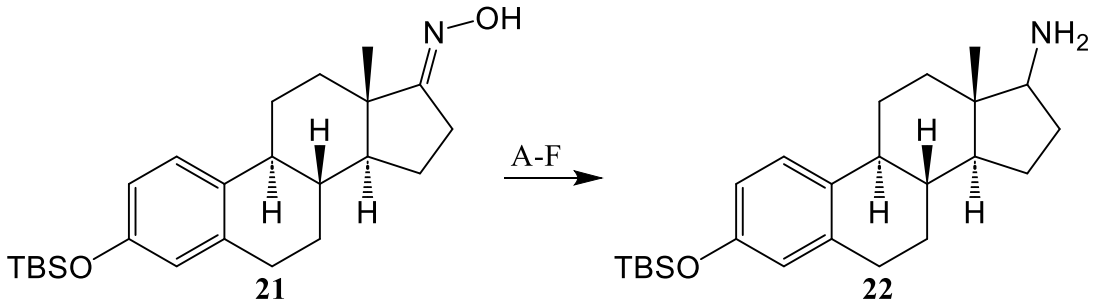
Our final attempts at synthesizing the C11 amine group rested on the discovery of an article in which 11-keto steroids were converted into 11-ketoxime[39]. We tried to reproduce this on ketone (16) under basic conditions with no luck. We theorized that under acidic conditions, this would likely be more viable by increasing the electrophilicity of the C11 carbon, and this indeed worked to yield ketoxime (20) as seen in Scheme 4.8.



Scheme 4.8 Conversion of 11-ketone into 11-ketoxime via hydroxylamine hydrochloride in the presence of acid.

To protect our valuable C-11 ketoxime intermediate, we began performing the following ketoxime reduction trials on the C17 ketone, as this was readily converted to ketoxime under the same conditions seen in Scheme 4.8 above. We herein report a discussion of six attempted reagents for reduction of 17-ketoxime to 17-amine (Scheme 4.9). First, reduction of oximes is reported in Abiraj et al. (2003) to occur using zinc dust in the presence of ammonium formate in a quick and easy manner[40]. It is worth noting their oxime starting materials all had conjugated systems attached to the oxime carbon, suggesting that the stability of the intermediate during reduction may play a role in why our reaction did not work under these conditions. Another condition showed the use of a

simple sodium reduction in the presence of isopropanol[39]. We optimized this reaction more than the others due to success of the reaction on a simple cyclohexanone structure, but were unable to repeat our results on the steroid even after leaving the reaction for longer or increasing the equivalency of the sodium. We ponder that maybe our sodium was too oxidized, or that our alcohol was not fully dry after distillation, and maybe the hydrogen was not being formed *in-situ*. We additionally explored the use of direct hydrogenation using palladium on carbon[44], but this saw no conversion to amine either. We theorize this may be due to the carbon load of the palladium we possessed, and that the palladium was not being activated as it should be. We'll see this again in the benzylamine example below.



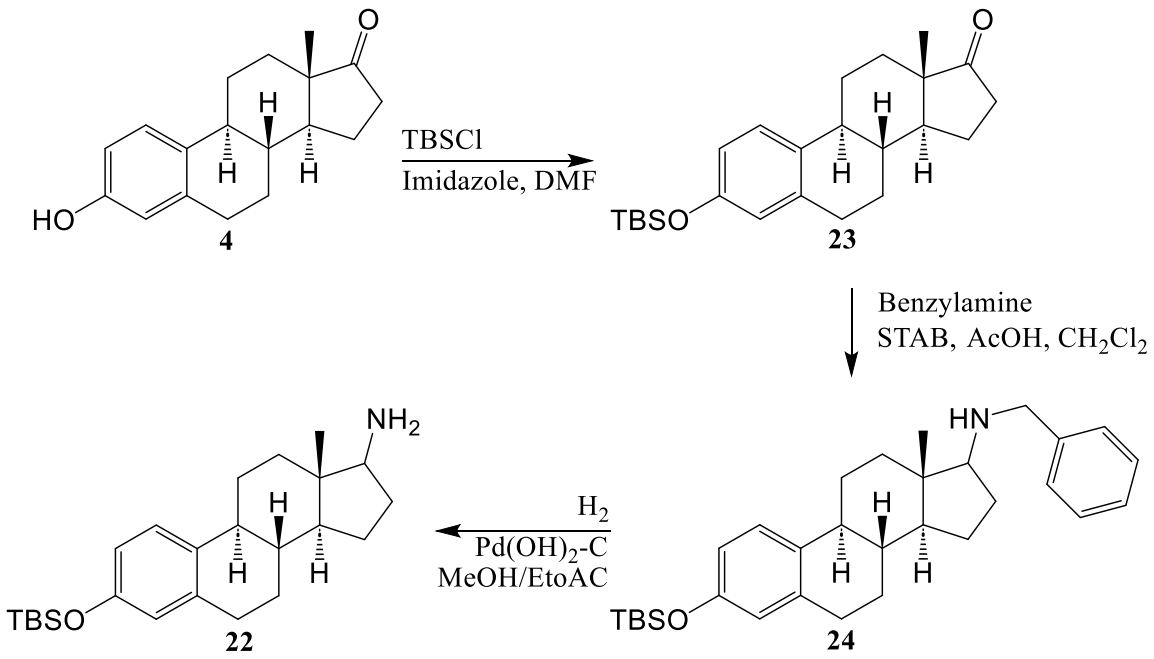
Scheme 4.9 Conversion of 17-ketoxime to 17-amine with various reagents. A) $Zn/HCOONH_4$, MeOH, 65°C; B) Na, 2-propanol, 85°C; C) $NaBH_4$, MeOH, -30°C, $NiCl_2 \cdot 6H_2O$; D) $TiCl_3$, $NaCNBH_3$, MeOH; E) $LiAlH_4$, THF, 70°C; F) H_2, Pd/C, MeOH.

The other three reagents we tried for oxime reductions were all hydride reagents. First, sodium borohydride was utilized along with a nickel chloride hexahydrate catalyst[41]. We do indeed see the formation of the black nickel-boride catalyst upon addition of the two species, but reduction does not take place on a quantitative scale. Our theorize is that the

conditions may not be fully anhydrous, and optimization of this reaction is imperative to yield success. Another article by Leeds et al. (1988) suggests the use of titanium chloride as a Lewis acid with reduction by sodium cyanoborohydride[42]. This reaction was again unsuccessful, and we theorize that the solubility of this reagent may play a role as it is in aqueous solution. We also tried a slightly stronger hydride catalyst in fear that our hydride was not strong enough. Lithium aluminum hydride has been shown to reduce oximes on steroids[43], but we question the ability of the ability for these hydrides to reduce the oxime, due to the less polar nature of the C-N vs C-O bond. Ultimately however, we found our answer in the form of a benzylamine reduction.

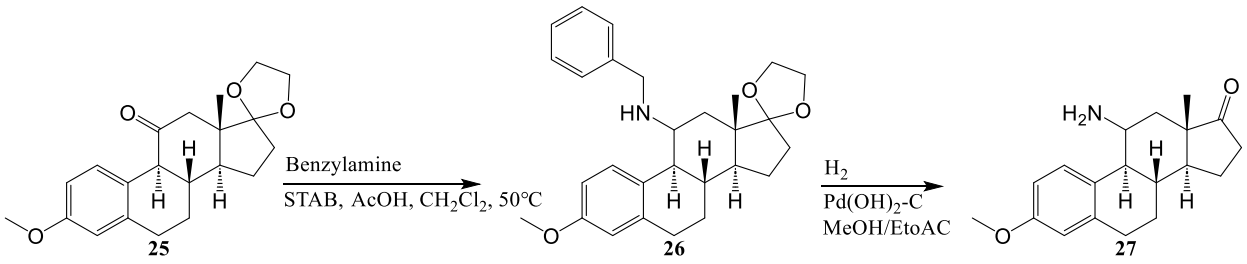
Benzylamine reduction to amine

A report by Mostafa et al. (2015) details the use of benzylamines to introduce amine groups onto estrone analogs[45]. This was of interest to us due to the ease at which it could be implemented if successful. We abandoned the idea of the oxime reduction and followed the simple three step procedure outlined in Scheme 4.10 to install amine at the C17 position. Protection of estrone (4) with TBDMS yields (23) in quantitative yields. Reaction with benzylamine in the presence of acid and sodium triacetoxyborohydride yields benzylamine (24). The hydrogenation of the benzylamine required some optimization, as the initial report by Mostafa et al. (2015) described the use of palladium hydroxide as the catalytic reducing agents. We contacted Mostafa after this reaction did not work, and proceeded to determine that the reagent must contain some carbon loading, and after utilizing palladium hydroxide on carbon, found conversion to the amine (22) within 24 hours.



Scheme 4.10 Synthesis of C17 amine using benzylamine reduction.

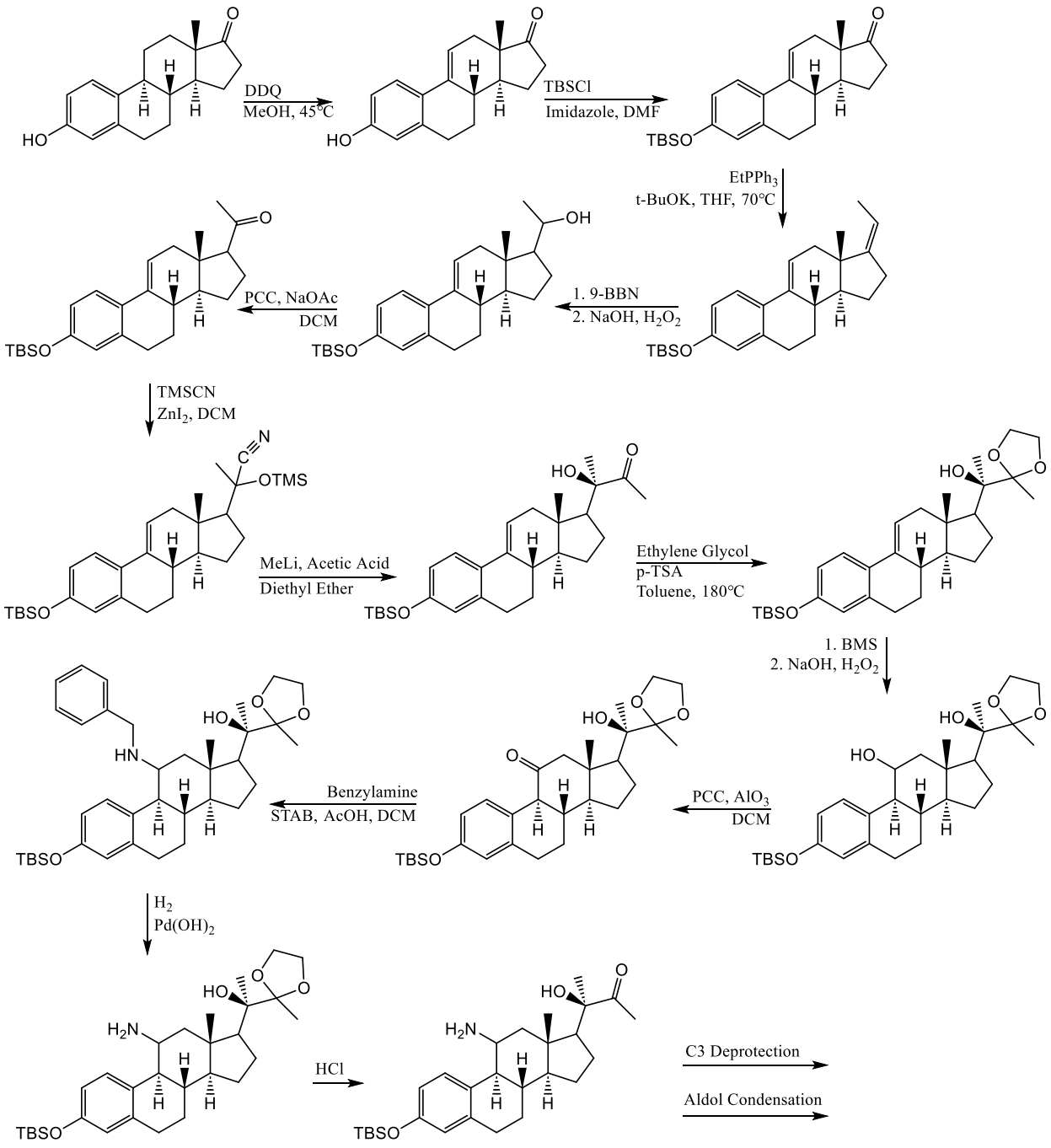
We proceeded to synthesize this amine at the C11 position from a small amount of sample that we had remaining from one of our previous trials (Scheme 4.11). The 11-keto steroid (25) we had available was reacted in the same way with benzylamine, acetic acid, and sodium triacetoxyborohydride. We found incomplete conversion to the benzylamine, likely due to the C13 methyl group and C1 hydrogen atom hindering attack by the nitrogen, and had to add an additional equivalent of benzylamine to push the reaction forward along with heating. The hydrogenation of benzylamine (26) to amine (27) was also slower, and required an additional 48 hours of vigorous stirring and an additional hydrogen balloon to succeed. Additionally, the reaction seemed to cleave the ketal protecting group at C17, affording the respective ketone group.



Scheme 4.11 Synthesis of C11 amine from C11 ketone.

Outstanding work for heterocyclic analog synthesis

Successful synthesis of the C11 amine indicates that a total synthesis for our previously mentioned heterocyclic analogs (TR115-TR125, TR482-TR485) is still possible. The resulting synthetic pathway would likely involve synthesis of the side chain at C17 first, protection of the resulting α-hydroxyl methyl ketone, followed by synthesis of the amino group, and ultimate aldol condensation (Scheme 4.12). Note the 11-keto group likely needs to be synthesized in this matter due to the protected phenolic group preventing rearrangement of electrons in the benzene ring to allow for epoxidation as shown in figure 4.4.



Scheme 4.12 Proposed synthetic pathway for C11 amino based heterocyclic estrone analogs.

4.4 Materials and Methods

4.4.1 Molecular Modeling Methods

A Library of over 650 analogs were virtually screened through an *in-silico* molecular docking process. OpenEye® Scientific Software and ChemOffice were utilized in the docking process. Analog structures were first designed using ChemDraw software according to their stereochemistry followed by 3-Dimension optimization in Chem3D using the energy minimization function Merck Molecular Force Field (MMFF94) calculation[22]. This force field calculation utilizes a maximum number of iterations that is set at 500, with an RMS gradient of 0.100. The dielectric constant of the analog is not changed during this calculation, and for every analog is left to be assumed as 1.000 with no cut-off method required. This force field calculation utilizes location and identity of the atoms to predict the molecule's cumulative potential energy[22]. Minimized energy conformers were then used in a docking process utilizing Omega[23,24] and FRED (Fast Rigid Exhaustive Docking)[25] modules along with the crystal structures for the proteins as followed: EGFR (PDB ID: 1M17), Pi3K (3L54), Akt (3MV5), mTOR (4JSV), JAK1 (4EI4), JAK2 (6VNK), STAT3 (1BG1), RAS (1X1R), RAF (5FD2), MEK (3W8Q), ERK (2OJG) HSP70 (4IO8), and HSP90 (5CF0). Crystal structures were downloaded from PDB without alteration using MakeReceptor program following software protocols[23-26]. Any bound ligands and water molecules are removed from the crystal structure in this program, and then the binding pocket is identified using a cavity detection. This docking is done in the absence of water molecule due to lower computational time and allow for more flexibility for the analog-protein interaction. Finally, any part of the protein that is not near the binding pocket can be removed from the program to lower computational time when

performing docking. FRED is then utilized to dock the analogs within the proteins, which spans anywhere from 5 minutes to 10 hours depending on the size of the protein and analog library[23-26]. OpenEye measures binding affinity using a relative value known as a consensus score. Consensus scores are output numerical values that take into account many analog-protein binding affects, including steric interactions, hydrogen bonding, lipophilic and hydrophilic interactions, electrostatics, and shape-based interactions. The lower the consensus score, the better the binding to the protein[26].

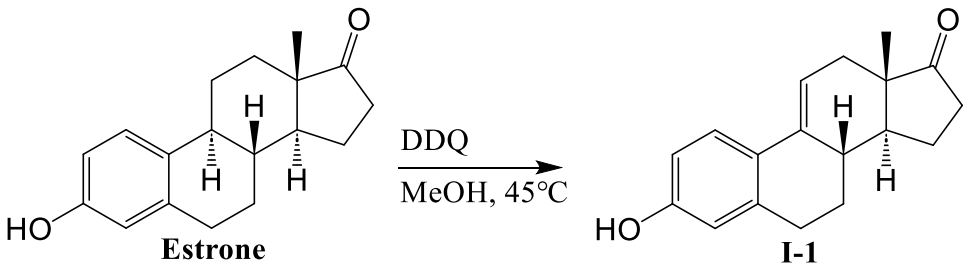
4.4.2 Chemical Synthesis

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 an over overnight before conducting chemical reactions. All chemical reactions requiring anhydrous environments utilized nitrogen or argon gas applied at the reaction time. ^{1}H and ^{13}C-NMR spectra were acquired on a Bruker AVANCE-600 MHz NMR spectrometer in deuterated solvents such as $CDCl_3$, Acetone-d6, or DMSO-d6 using solvent residual peak as the internal standard. Reporting of coupling constants is in Hz and the signal multiplicities are reported as singlet (s), doublet (d), triplet (t), quartet (q), doublet of doublets (dd), doublet of triplets (dt), triplet of doublets (td), triplet of triplets (tt), doublet of quartets (dq), quartet of triplets (qt), doublet of doublet of doublets (ddd), doublet of doublet of triplets (ddt), doublet of doublet of doublet of doublets (dddd), doublet of doublet of doublet of doublet of doublets (ddddd), multiplet (m), or broad (br). Note that NMR spectra are only provided for successful reactions and not for unsuccessful

optimization reactions. TLC analysis was performed using precoated silica gel PE sheets. Products were purified via column chromatography and using silica gel 40–63 lm (230–400 mesh), normal phase preparative TLC plates. TLC plates were visualized by ultraviolet at 254 nm. TLC plates were stained by Iodine, Vanillin, and Ceric Ammonium Molybdate (CAM) stain.

Intermediate-1

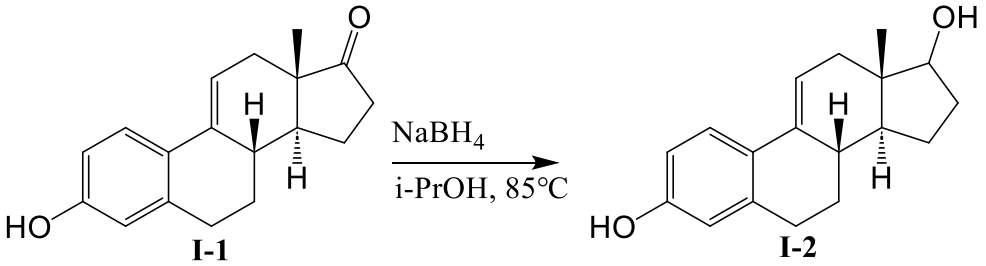


Estrone (5g, 18.49 mmole, 1 eq) was added to a 1L round bottom flask with 600 mL of methanol at 45°C. 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) (6.3 g, 27.74 mmole, 1.5 eq) was added upon estrone dissolution, and was left to reflux for 4 hours with stirring. Following completion of the reaction, the methanol was evaporated, and the product was separated from the crude solid using Soxhlet extraction with 500 mL of dichloromethane to yield 4.65 g of Intermediate-1 (93.7%).

^{1}H NMR (600 MHz, Acetone-d6) δ 7.42 (m, 1H), 6.57 (m, 1H), 6.48 (m, 1H), 6.02 (m, 1H), 2.74 (m, 4H), 2.52 (m, 1H), 2.44 (m, 1H), 2.41 (d, J=7.9 Hz, 1H), 2.16 (m, 3H), 2.1 (m, 1H), 2.04 (m, 3H), 1.57 (m, 2H), 1.26 (m, 1H), 0.81 (s, 3H).

^{13}C NMR (151 MHz, Acetone-d6) δ 220.30, 156.24, 150.78, 137.13, 135.21, 125.06, 115.17, 114.76, 113.78, 47.00, 45.43, 37.68, 35.66, 33.55, 29.21, 27.29, 23.80, 22.01, 14.17, 11.93.

Intermediate-2



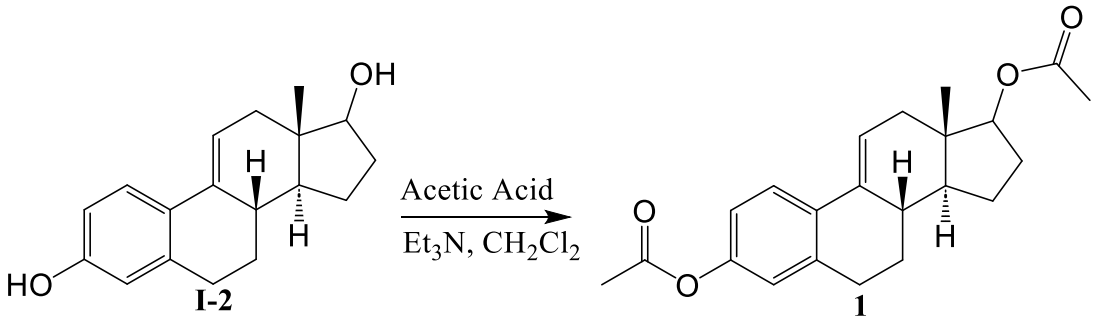
I-1 **I-2**

Intermediate-1 (0.8 g, 2.98 mmole, 1 eq) was added to a 100 mL round bottom flask along with 25 mL of isopropanol and heated to 85°C with stirring. Sodium borohydride (0.496 g, 13.12 mmole, 4.4 eq) was then added to the reaction mixture, and left to reflux for 2 hours. Upon completion of the reaction, the isopropanol was removed under vacuo followed by addition of 20 mL of water. The crude product was then extracted with 3x20 mL of ethyl acetate, dried with anhydrous sodium sulfate, filtered, and concentrated under vacuo. The crude product was then purified by column chromatography with 9:1 hexane: ethyl acetate to yield intermediate-2 in 96.2% yield.

^{1}H NMR (600 MHz, Chloroform-d) δ 7.44 (d, J=8.7 Hz, 1H), 6.62 (dd, J=8.7, 3 Hz, 1H), 6.54 (d, J=2.6 Hz, 1H), 6.10 (m, 1H), 2.94 (s, 1H), 2.86 (s, 2H), 2.81 (dd, J=5.3, 2.3 Hz, 1H), 2.5 (d, J=17.4 Hz, 1H), 2.28 (t, J=3.2(x2) Hz, 2H), 2.18 (m, 3H), 2.12 (s, 1H), 1.63 (m, 2H), 1.42 (dd, J=12.7, 5.1 Hz, 1H), 0.92 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-d) δ 154.65, 137.38, 135.43, 127.55, 125.14, 119.99, 118.32, 116.79, 79.83, 47.86, 46.26, 38.21, 36.65, 36.24, 33.93, 31.60, 29.72, 27.81, 26.59, 22.55, 19.65, 14.51.

Compound 1

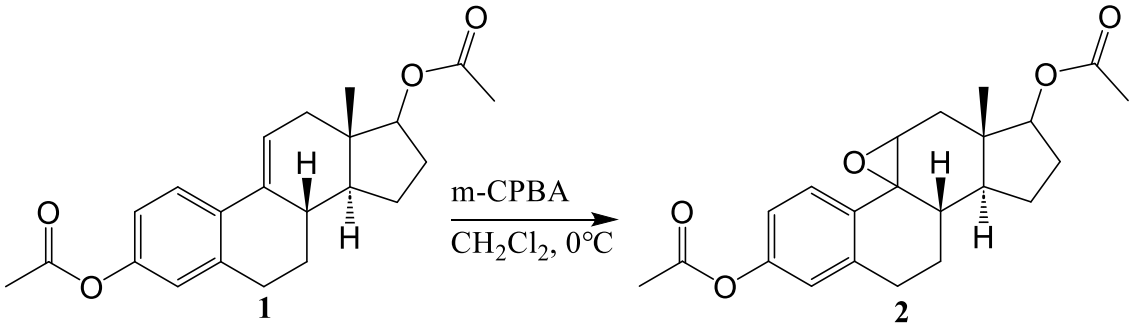


Estradiol intermediate-2 (0.775 g, 2.87 mmole, 1 eq) was added to a 50 mL round bottom flask with 15 mL of dichloromethane. Argon gas was added to the flask, followed by addition of triethylamine (1.2 mL, 8.598 mmole, 3 eq) and acetic acid (0.81 mL, 8.598 mmole, 3 eq). The reaction was stirred overnight until completion. 20 mL of water was added to the reaction mixture, followed by extraction with 3x20 mL of ethyl acetate. The crude product was dried with anhydrous sodium sulfate, filtered off, and concentrated in vacuo. Crude material was purified using column chromatography with an 8:2 hexane: ethyl acetate mixture to yield compound 1 in 38.3% yield.

^{1}H NMR (600 MHz, Chloroform-d) δ 7.27 (m, 1H), 6.83 (dd, J = 8.5, 2.8 Hz, 1H), 6.79 (d, J=2.6 Hz, 1H), 5.98 (m, 1H), 4.69 (dd, J=9.4, 7.9 Hz, 1H), 2.86 (m, 2H), 2.88 (m, 4H), 2.23 (m, 1H), 2.09 (s, 3H), 2.06, (s, 3H), 1.88 (m, 2H), 1.74 (m, 1H), 1.55 (m, 2H), 1.48 (m, 2H), 1.40 (m, 2H), 1.28 (m, 4H), 0.83 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-d) δ 174.93, 170.00, 150.62, 148.42, 138.26, 138.06, 126.43, 121.54, 118.61, 115.23, 81.77, 50.07, 44.17, 43.22, 38.49, 36.69, 30.41, 29.57, 27.08, 26.19, 23.16, 21.18, 11.12.

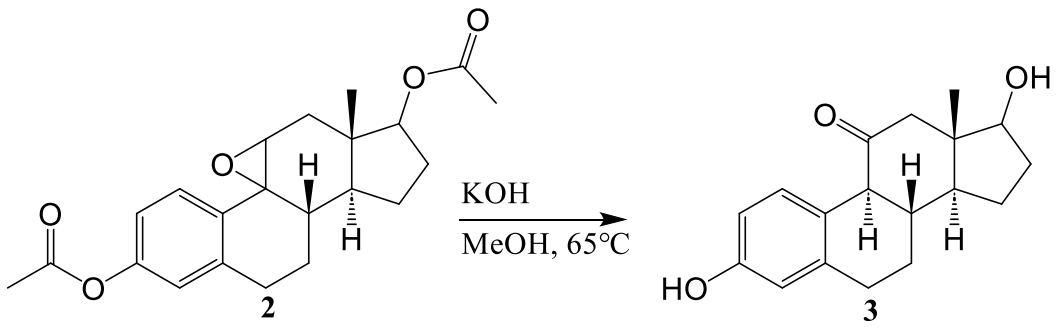
Compound 2



1 **2**

To a cold (0°C) solution of the diacetate (0.39 g, 1.1 mmole, 1 eq) in 10 mL of dichloromethane was added dropwise over 30 minutes m-chloroperoxybenzoic acid (0.2 g, 1.16 mmole, 1.06 eq) dissolved in 5 mL of dichloromethane. Reaction completion was monitored by TLC and complete after 6 hours. 20 mL of sodium bicarbonate was added to the mixture, followed by extraction of the crude product with 3x20 mL of ethyl acetate. The crude solution was dried with anhydrous sodium sulfate, filtered off, and concentrated in vacuo. The crude product was then utilized without purification in the next step.

11-keto-estradiol (3)

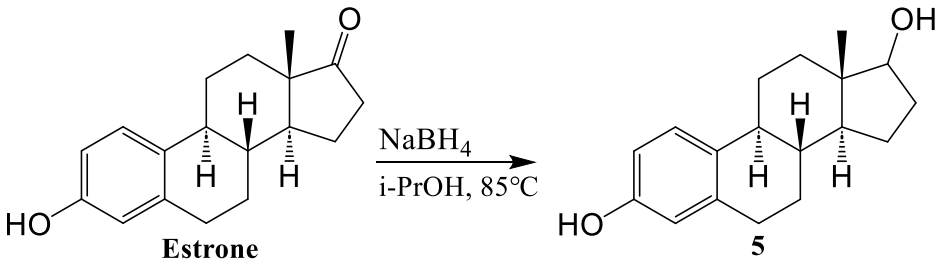


Epoxide 2 (0.3 g, 0.810 mmol, 1 eq) in methanol (20 mL) was refluxed with a 5% methanolic potassium hydroxide solution (5 mL) for 30 minutes. After cooling, the mixture was neutralized with a 6N HCl solution followed by extraction with dichloromethane (3x20 mL). Crystallization of the product was induced following addition of a hexane: ethyl acetate mixture (5:5 v/v) to yield 0.1 g of 11-keto-estradiol (3) in 43% yield.

^{1}H NMR (600 MHz, Acetone-d6) δ 9.00 (br, 1H), 7.05 (d, J=8.7 Hz, 1H), 6.51 (dd, J=8.3, 3.0 Hz, 1H), 6.44 (d, J =3.0 Hz, 1H), 4.51 (dd, J=4.2, 0.8 Hz, 1H), 3.53 (td, J=8.6(x2), 4.3 Hz, 1H), 2.71 (m, 2H), 2.24 (m, 1H), 2.06 (m, 1H), 1.89 (m, 1H), 1.84 (dt, J=12.6, 3.5(x2) Hz, 1H), 1.77 (ddt, J=12.3, 5.6, 2.5(x2) Hz, 1H), 1.59 (m, 1H), 1.38 (m, 1H), 1.30 (td, J=12.7(x2), 3.4 Hz, 2H), 1.24 (m, 2H), 1.17 (m, 1H), 1.1 (m, 1H), 0.67 (s, 3H).

^{13}C NMR (151 MHz, Acetone-d6) δ 209.11, 154.86, 130.64, 130.41, 126.02, 114.89, 112.68, 80.05, 49.50, 43.51, 42.79, 38.68, 36.57, 29.88, 29.14, 26.93, 26.06, 22.77, 11.25.

Compound 5



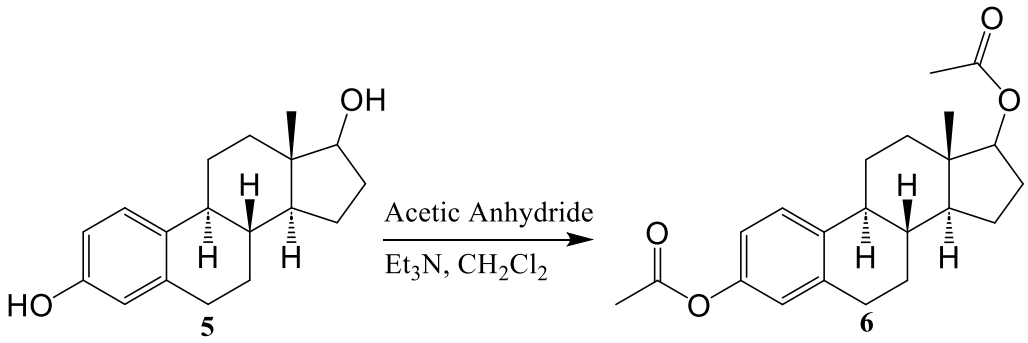
Estrone **5**

Estrone (2.5g, 9.25 mmole, 1 eq) was added to a 250 mL round bottom flask along with 85 mL of isopropanol and heated to 85°C with stirring. Sodium borohydride (1.536 g, 40.59 mmole, 4.4 eq) was then added to the reaction mixture, and left to reflux for 2 hours. Upon completion of the reaction, the isopropanol was removed under vacuo followed by addition of 40 mL of water. The crude product was then extracted with 3x40 mL of ethyl acetate, dried with anhydrous sodium sulfate, filtered, and concentrated under vacuo. The crude product was then purified by column chromatography with 9:1 hexane: ethyl acetate to yield compound 5 in 94.7% yield.

^{1}H NMR (600 MHz, Chloroform-d) δ 7.44 (d, J=8.7 Hz, 1H), 6.62 (dd, J=8.7, 3 Hz, 1H), 6.54 (d, J=2.6 Hz, 1H), 2.94 (s, 1H), 2.86 (s, 2H), 2.81 (dd, J=5.3, 2.3 Hz, 1H), 2.5 (d, J=17.4 Hz, 1H), 2.28 (t, J=3.2(x2) Hz, 2H), 2.18 (m, 3H), 2.12 (s, 1H), 1.63 (m, 2H), 1.42 (dd, J=12.7, 5.1 Hz, 1H), 0.92 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-d) δ 154.65, 137.38, 127.55, 125.14, 118.32, 116.79, 79.83, 47.86, 46.26, 38.21, 36.65, 36.24, 33.93, 31.60, 29.72, 27.81, 26.59, 25.63, 22.55, 19.65, 14.51.

Compound 6

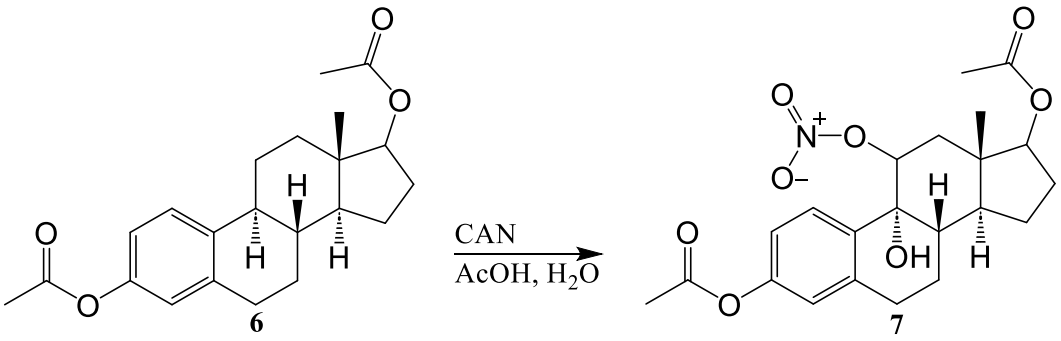


Compound 5 (2.386 g, 8.76 mmole, 1 eq) was added to a 250 mL round bottom flask with 50 mL of dichloromethane. Argon gas was added to the flask, followed by addition of triethylamine (3.7 mL, 26.28 mmole, 3 eq) and acetic acid (1.5 mL, 26.28 mmole, 3 eq). The reaction was stirred overnight until completion. 40 mL of water was added to the reaction mixture, followed by extraction with 3x40 mL of ethyl acetate. The crude product was dried with anhydrous sodium sulfate, filtered off, and concentrated in vacuo. Crude material was purified using column chromatography with an 8:2 hexane: ethyl acetate mixture to yield compound 6 in 93.4% yield.

^{1}H NMR (600 MHz, Chloroform-d) δ 7.27 (m, 1H), 6.83 (dd, J = 8.5, 2.8 Hz, 1H), 6.79 (d, J=2.6 Hz, 1H), 4.69 (dd, J=9.4, 7.9 Hz, 1H), 2.86 (m, 2H), 2.88 (m, 4H), 2.23 (m, 1H), 2.09 (s, 3H), 2.06, (s, 3H), 1.88 (m, 2H), 1.74 (m, 1H), 1.55 (m, 2H), 1.48 (m, 2H), 1.40 (m, 2H), 1.28 (m, 4H), 0.83 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-d) δ 174.93, 170.00, 150.62, 138.26, 126.43, 121.54, 118.61, 115.23, 81.77, 50.07, 44.17, 43.22, 38.49, 36.69, 30.41, 29.57, 27.08, 26.19, 23.16, 21.18, 11.12.

Compound 7



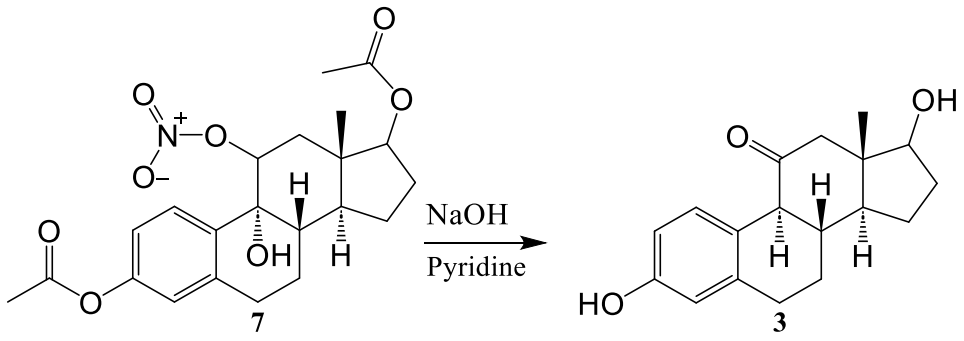
6 7

To a solution of compound 6 (1 g, 2.805 mmole, 1 eq) in 42 mL of acetic acid was added dropwise a solution of cerium ammonium nitrate (7.07 g, 12.9 mmole, 4.6 eq) in 5.6 mL of water. The solution was stirred under argon gas for 2 hours before being diluted by half with water and extracted with 3x40 mL ethyl acetate. The organic fractions were combined, washed with 30 mL sodium bicarbonate saturated solution, dried over anhydrous sodium sulfate, filtered, and concentrated in vacuo. The crude product was isolated using a 7:3 hexane to ethyl acetate mixture to yield compound 7 in 49.3% yield.

^{1}H NMR (600 MHz, Acetone-d6) δ – 7.38 (d, J=8.7 Hz, 1H), 6.95 (m, 1H), 6.91 (d, J=3.0 Hz, 1H), 5.88 (t, J=3(x2) Hz, 1H), 4.79 (dd, J=9.4, 7.2 Hz, 1H), 4.58 (s, 1H), 2.93 (m, 2H), 2.87 (m, 1H), 2.28 (m, 1H), 2.2 (t, J=2.8(x2) Hz, 2H), 2.08 (m, 1H), 1.98 (m, 5H), 1.63 (m, 1H), 1.55 (m, 1H), 1.09 (s, 1H), 1.0 (s, 3H).

^{13}C NMR (151 MHz, Acetone-d6) δ 171.06, 169.74, 151.29, 140.45, 136.30, 127.36, 123.38, 120.59, 83.74, 83.16, 71.64, 43.99, 42.66, 39.24, 38.77, 37.76, 36.84, 28.20, 25.72, 24.23, 23.54, 19.64, 14.61.

11-keto estradiol (3)

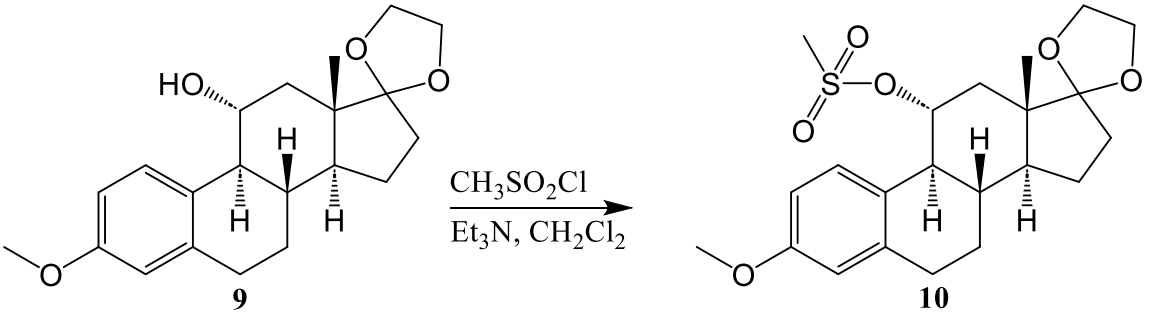


To a solution of nitrate 7 (0.6 g, 1.38 mmole, 1 eq) in 15 mL of pyridine under argon gas was added 1 mL of a 20% aqueous sodium hydroxide solution (4.9 mmole). The reaction was stirred at room temperature whereby the reaction turned yellow after 20 minutes, then turns a dark purple color after 2 hours and sodium nitrate is precipitated. After 3 hours of stirring and monitoring by TLC, the reaction was diluted with 30 mL of water, acidified using 10% aqueous HCl until neutral, followed by extraction with 3x30 mL of ethyl acetate. The organic fractions were collected, washed with a saturated salt solution, and dried with anhydrous sodium sulfate. The solid was filtered off, and the fractions concentrated in vacuo. The product was isolated to yield 11-keto-estradiol (3) in 97.2% yield.

^{1}H NMR (600 MHz, Acetone-d6) δ 9.00 (br, 1H), 7.05 (d, J=8.7 Hz, 1H), 6.51 (dd, J=8.3, 3.0 Hz, 1H), 6.44 (d, J =3.0 Hz, 1H), 4.51 (dd, J=4.2, 0.8 Hz, 1H), 3.53 (td, J=8.6(x2), 4.3 Hz, 1H), 2.71 (m, 2H), 2.24 (m, 1H), 2.06 (m, 1H), 1.89 (m, 1H), 1.84 (dt, J=12.6, 3.5(x2) Hz, 1H), 1.77 (ddt, J=12.3, 5.6, 2.5(x2) Hz, 1H), 1.59 (m, 1H), 1.38 (m, 1H), 1.30 (td, J=12.7(x2), 3.4 Hz, 2H), 1.24 (m, 2H), 1.17 (m, 1H), 1.1 (m, 1H), 0.67 (s, 3H).

^{13}C NMR (151 MHz, Acetone-d6) δ 209.11, 154.86, 130.64, 130.41, 126.02, 114.89, 112.68, 80.05, 49.50, 43.51, 42.79, 38.68, 36.57, 29.88, 29.14, 26.93, 26.06, 22.77, 11.25.

Mesylate (10)

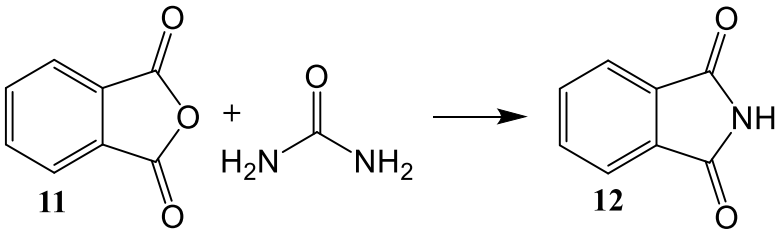


To a solution of compound 9 (0.18 g, 0.523 mmole, 1 eq) in 5 mL of dichloromethane was added first under slow addition triethylamine (0.146 mL, 1.046 mmole, 2 eq) and left to stir for 10 minutes. Then, very slow addition of methanesulfonyl chloride (0.061 mL, 0.784 mmole, 1.5 eq) was performed under careful stirring. The reaction was stirred and monitored by TLC and completed after 2 hours. Extraction of the crude product was performed by adding 10 mL of water followed by extraction with 3x10 mL of ethyl acetate. The organic fractions were collected, dried over anhydrous sodium sulfate, filtered, and concentrated in vacuo. The crude product was purified using a 95:5 hexane to ethyl acetate mixture with gradually increasing polarity to yield mesylate 10 in 85.6% yield.

^{1}H NMR (600 MHz, Chloroform-d) δ 7.08 (d, J=7.9 Hz, 1H), 6.82 (dd, J=8.5, 2.8 Hz, 1H), 6.68 (d, J=3 Hz, 1H), 5.36 (m, 1H), 4.29 (t, J=6.2(x2) Hz, 2H), 3.81 (s, 3H), 3.76 (s, 3H), 2.76 (t, 7.4(x2) Hz, 2H), 2.51 (m, 1H), 2.01 (m, 2H), 1.92 (m, 2H), 1.86 (m, 2H), 1.77 (m, 2H), 1.52 (m, 1H), 1.41 (m, 1H), 1.25 (m, 2H), 0.89 (s, 3H).

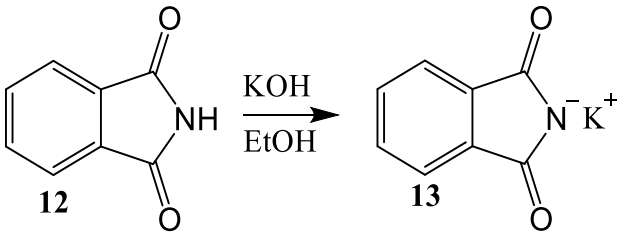
^{13}C NMR (151 MHz, Chloroform-d) δ 158.76, 139.50, 134.69, 128.71, 125.35, 113.93, 112.36, 69.80, 65.30, 64.53, 55.22, 49.65, 46.91, 39.92, 38.13, 37.31, 34.22, 29.42, 28.17, 26.20, 22.22, 14.66.

Phthalimide (12)



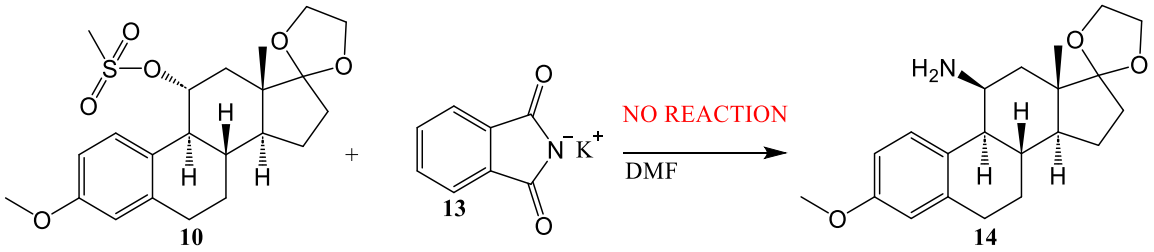
Phthalimide 12 was prepared by reacting phthalic anhydride (11) (10 g, 67.5 mmole, 1 eq) with urea (3.3 g, 54.9 mmole, 0.8 eq). The two solids were first mixed together and crushed using a mortar and pestle and then transferred to a 500 mL round bottom flask. The flask is then heated until 135℃ and left to heat until the solids melt, bubble, and turn into a solid component. The heat was turned off, followed by addition of 150 mL of water to dissolve the solid impurities. The crude product was filtered, and was then recrystallized from ethanol to yield phthalimide (12) in 91.5% yield.

Potassium phthalimide (13)



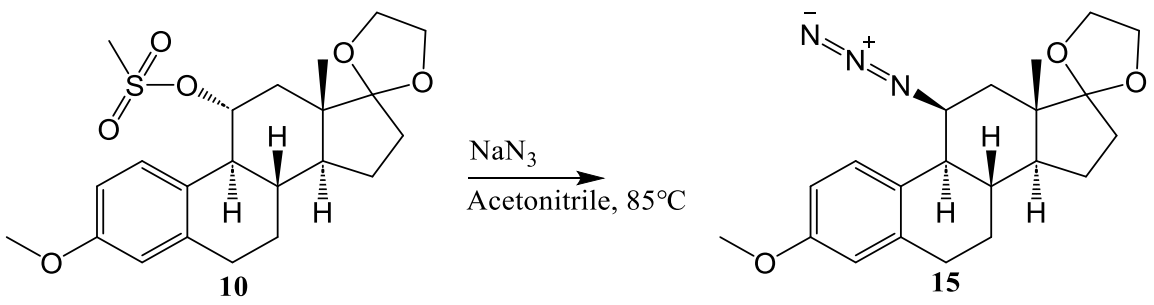
1.4 g of potassium hydroxide (24.9 mmole, 1.2 eq) was added to a 250 mL beaker with 50 mL of methanol. This was stirred under heat at 85°C to further dissolution of the salt, and upon dissolution 3 g of phthalimide (20.4 mmole, 1 eq) was added and the reaction stirred for 1 hour. The solid salt was then filtered to yield potassium phthalimide (13) in 86.5% yield.

Compound 14 (Potassium phthalimide)



Mesylate 10 (0.02 g, 0.047 mmole, 1 eq) was added to a 5 mL round bottom flask with 0.1 mL of dimethylformamide. Then, potassium phthalimide (0.008 g, 0.047 mmole, 1 eq) was added to the reaction mixture, and left to stir for 24 hours at room temperature. However, after 24 hours at room temperature, the reaction had not proceeded at all. Upon isolation via extraction with ethyl acetate, the starting material remained.

Azide (15)

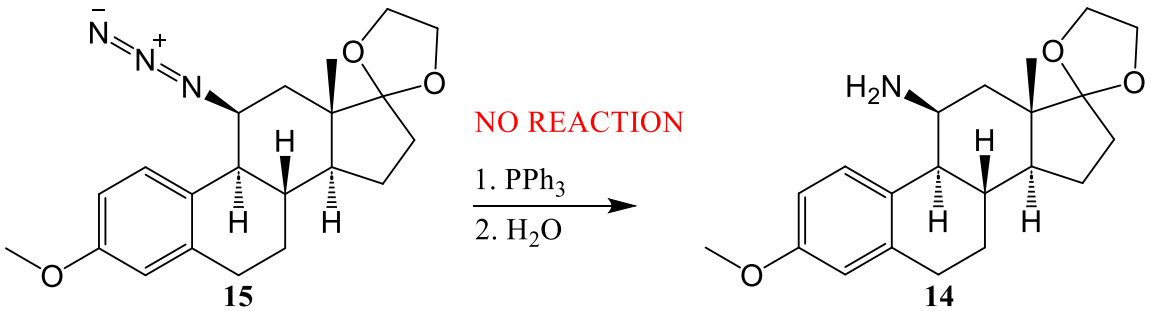


Mesylate (10) (0.08 g, 0.19 mmole, 1 eq) was added to a 25 mL round bottom flask with sodium azide (0.03 g, 0.456 mmole, 2.4 eq) under vigorous stirring with 2.4 mL carbon tetrachloride. The reaction was left at room temperature to stir under argon for 24 hours. Upon completion of the reaction, 5 mL of water was added to the reaction mixture, followed by extraction of the compound with 3x10 mL ethyl acetate. The organic fractions were combined, dried over anhydrous sodium sulfate, filtered, and concentrated in vacuo. The azide (15) was afforded in 88.4% yield and used as such in the next reaction.

^{1}H NMR (600 MHz, Chloroform-d) δ 7.66 (m, 1H), 7.3 (d, J=8.3 Hz, 1H), 7.22 (d, J=8.3 Hz, 1H), 5.2 (m, 4H), 4.36 (m, 1H), 3.58 (m, 1H), 2.35 (m, 1H), 2.31 (s, 3H), 1.89 (s, 3H), 1.7 (m, 1H), 1.64 (ddt, J=11.9, 7.6, 3.7(x2) Hz, 2H), 1.56 (m, 2H), 1.39 (m, 2H), 1.32 (m, 1H), 1.16 (m, 2H), 1.11 (s, 3H).

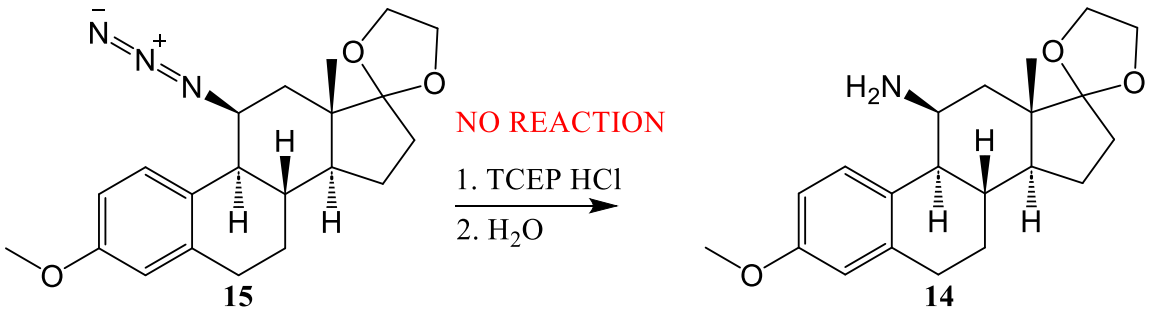
^{13}C NMR (151 MHz, Chloroform-d) δ 146.27, 144.31, 135.34, 134.71, 130.24, 129.65, 127.39, 67.68, 60.09, 53.58, 32.14, 31.49, 25.47, 25.25, 24.71, 23.65, 23.19, 21.30, 20.70, 13.99.

Compound 14 (Triphenylphosphine)



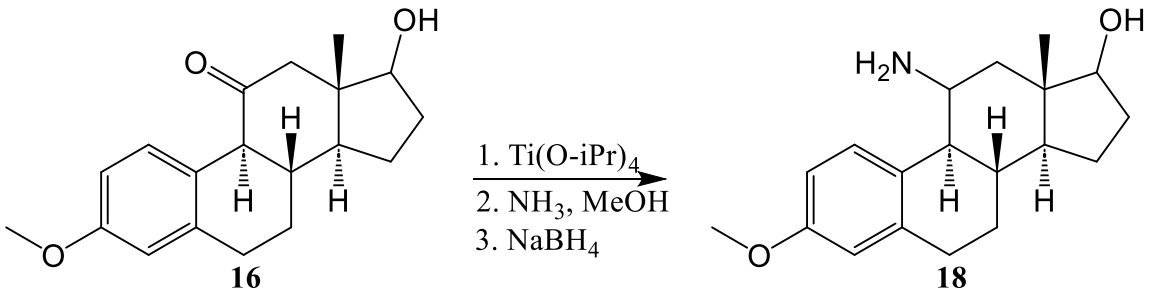
Azide (15) (0.06 g, 0.162 mmole, 1 eq) was added to a 5 mL round bottom flask with 0.6 mL carbon tetrachloride and 1.8 mL dimethylformamide. Then, 0.178 g of triphenylphosphine (0.68 mmole, 4.2 eq) was added to the reaction mixture and left to stir for 3 hours before checking by thin layer chromatography. No intermediate was seen on the thin layer, and was therefore left to stir for an additional 24 hours. After this, still no intermediate was seen on the thin layer, and the reaction was stopped.

Compound 14 (TCEP)



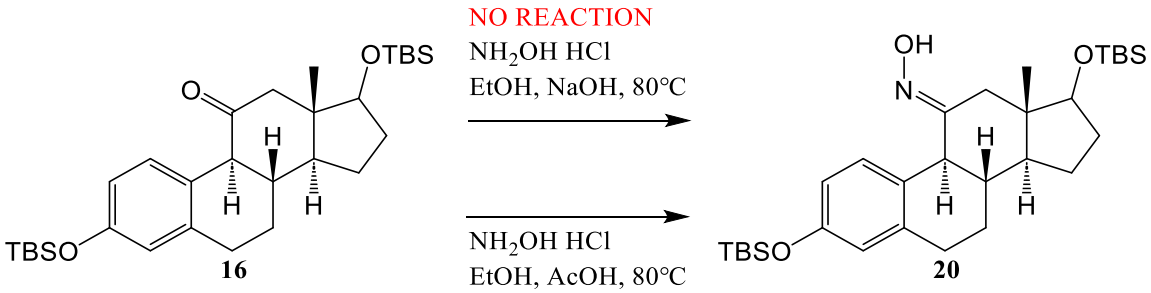
Azide (15) (0.3 mg, 0.0024 mmole, 1 eq) was added to a 5 mL round bottom flask with 200 microliter of phosphate buffer (200 mM, pH = 7.5). Then, tris(2-carboxyethyl) phosphine hydrochloride (7 mg, 0.024 mmole, 10 eq) was added to the reaction mixture and monitored by TLC. After 30 minutes there showed no movement of reaction completion, and after 4 hours there was still no product that had been formed, suggesting problems with dissolution.

Compound 18



Compound (16), which was afforded from a colleague, was added (0.431 g, 1.44 mmole, 1 eq) to a 10 mL round bottom flask followed by addition of 0.85 mL (2.87 mmole, 2 eq) of titanium isopropoxide. This was left to stir for 5 minutes followed by addition of a 7N methanolic ammonia solution (1.1 mL, 7.17 mmole, 5 eq). The mixture was stirred under argon gas for 6 hours, followed by addition of sodium borohydride (0.081 g, 2.15 mmole, 1.5 eq) for 3 hours. The reaction mixture showed movement of the starting material by thin layer chromatography so the reaction was stopped. 10 mL of water was added to the mixture, and the compound was extracted with 3x10 mL of ethyl acetate. Organic fractions were collected, dried over anhydrous sodium sulfate, filtered, and concentrated in vacuo. NMR was performed on the crude compound, and it was identified that compound (18) was not formed, but instead compound 19, the 11-hydroxy variant, had been formed.

Oxime (20)

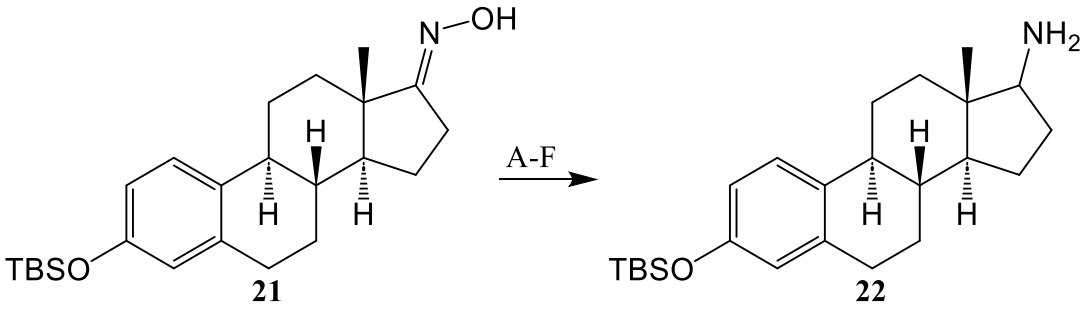


Initially, ketone (16) (0.4 g, 1.397 mmole, 1 eq) was dissolved in 10 mL of methanol and supplemented with 1.795 g of hydroxylamine hydrochloride (25.84 mmole, 18.5 eq) and 1.865 g of sodium hydroxide (33.25 mmole, 23.8 eq). This reaction was refluxed for 5 days at 80°C. Monitoring of the reaction by thin layer chromatography yielded no results, so the reaction was stopped. We extracted the unreacted starting material with ethyl acetate and proceeded to repeat the reaction, but instead of sodium hydroxide utilized acetic acid until the solution had turned a pH of 4. Monitoring of the reaction showed complete removal of the starting material after 24 hours of refluxing. The crude material was thus extracted with 3x20 mL of ethyl acetate and dried over anhydrous sodium sulfate. The solid was filtered off, and the organic fractions were concentrated in vacuo. The crude material was isolated using a hexane and ethyl acetate mixture (8:2 v/v) to yield oxime 20 in 78.2% yield.

^{1}H NMR (600 MHz, Acetone-d6) δ 9.0 (s, 1H), 7.05 (d, J=8.3 Hz, 1H), 6.53 (dd, J=8.3, 3 Hz, 1H), 6.46 (m, 1H), 3.4 (m, 1H), 2.75 (m, 2H), 2.52 (m, 2H), 2.35 (m, 3H), 2.17 (m, 1H), 1.93 (ddt, 12.9, 3.5(x2) Hz, 1H), 1.84 (m, 2H), 1.68 (m, 1H), 1.53 (m, 1H), 1.39 (m, 3H), 1.24 (m, 1H), 1.18 (m, 1H), 0.87 (s, 3H).

^{13}C NMR (151 MHz, Acetone-d6) δ 167.89, 154.96, 137.03, 130.11, 125.93, 114.92, 112.73, 82.45, 59.74, 52.49, 51.14, 43.58, 43.33, 37.86, 34.25, 29.98, 29.07, 27.77, 26.84, 25.91, 24.85, 22.49, 17.29.

Amine (22)



Condition A

Oxime (21) (0.0825 g, 0.156 mmole, 1 eq) was added to a 25 mL round bottom flask and supplemented with 10 mL of methanol. Ammonium formate (0.039 g, 0.623 mmole, 4 eq) was added to the flask along with 0.0204 g of zinc powder (0.311 mmole, 2 eq). The mixture was refluxed at 65℃ and monitored by thin layer chromatography. After 24 hour the thin layer had not changed, and the reaction was extracted with ethyl acetate. NMR of crude product showed oxime starting material.

Condition B

Sodium (0.115 g, 5 mmole, 20 eq) was added in small portions to a refluxing solution of oxime (21) (0.1 g, 0.25 mmole, 1 eq) in 20 mL of 2-propanol at 85℃. The reaction was refluxed for 2 hours, and after 2 hours additional 2-propanol was added to remove any remaining unreacted sodium. The mixture was concentrated in vacuo, and water was added to dissolve the sodium isopropoxide side product. The mixture was extracted with dichloromethane (3x30 mL) and washed with a brine solution (30 mL). The organic fractions were collected, dried over anhydrous sodium sulfate, filtered, and concentrated in vacuo. NMR of the crude product however yielded oxime starting material.

Condition C

Oxime (21) (0.1 g, 0.25 mmole, 1 eq) was dissolved in 10 mL of methanol in a 25 mL round bottom flask. Nickel chloride hexahydrate (0.119 g, 0.5 mmole, 2 eq) was added to the flask and the flask was then put on dry ice until -30°C. Sodium borohydride (0.0945 g, 2.5 mmole, 10 eq) was added to the flask in small portions over 30 minutes. The reaction produced a black nickel borohydride precipitate and was left to stir under argon gas overnight. After 24 hours, the black solid was filtered off, and the mixture concentrated in vacuo. The crude solid was run on NMR and yielded the oxime starting material.

Condition D

Titanium trichloride (20% aqueous solution, 0.16 mL, 0.55 mmole, 2.2 eq) was added to a solution of oxime (21) (0.1 g, 0.25 mmole, 1 eq) and sodium cyanoborohydride (0.0471 g, 0.75 mmole, 3 eq) in 15 mL of methanol over 1 hour. The reaction was stirred under argon gas for 24 hours. The reaction was diluted with 20 mL of water, and was extracted with 3x20 mL of ethyl acetate. The organic fractions were combined, dried over anhydrous sodium sulfate, filtered, and concentrated in vacuo. The crude material was run in NMR and the oxime starting material remained.

Condition E

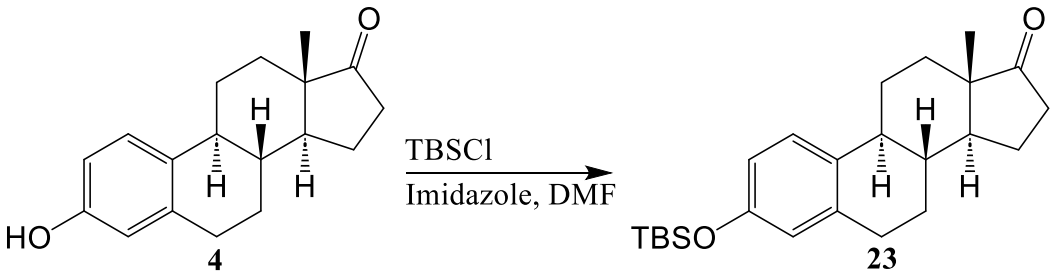
Oxime (21) (0.1 g, 0.25 mmole, 1 eq) was added to a 25 mL round bottom flask in 10 mL of THF. Lithium aluminum hydride (0.1 g, 2.625 mmole, 10.5 eq) was added slowly to the flask and the reaction stirred under argon gas. The mixture was then refluxed at 70°C overnight. The solvent was removed under vacuo and 15 mL of water was added to the crude solid. The material was extracted with 3x15 mL of ethyl acetate and dried over

anhydrous sodium sulfate. The organic fractions were filtered and concentrated. NMR of the crude product showed oxime starting material.

Condition F

A solution of oxime (21) (0.1 g, 0.25 mmole, 1 eq) and 10% Pd-C (m/m, 10 mg) in 10 mL of methanol was stirred under hydrogen gas at atmospheric pressure for 2 days. Monitoring of the reaction by thin layer chromatography showed no evolution of amine product after the full 2 days. The reaction was left for a further 2 days to no avail.

Protected Estrone (23)



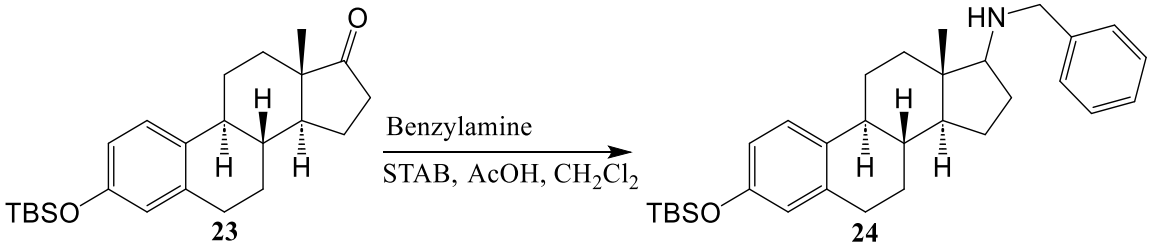
Estrone (4) (2 g, 7.407 mmole, 1 eq) was added to a 100 mL round bottom flask with 35 mL of Dimethylformamide. First, Imidazole (1.387 g, 20.369 mmole, 2.75 eq) was added to the reaction flask and stirred for 10 minutes under argon gas. Then, tert-butyldimethylsilyl chloride (1.675 g, 11.11 mmole, 1.5 eq) was added to the reaction flask and stirred for 4 hours. Evolution of the product is seen upon solidifying of the mixture. The reaction was stopped, and 40 mL of water was added to the reaction mixture. The mixture was then extracted with 3x40 mL of ethyl acetate, followed by collection of the organic fractions. The organic fraction was then washed with a further 500 mL of water to

remove any excess DMF in the solution, followed by drying of the organic fraction with anhydrous sodium sulfate. The solution was filtered, and concentrated in vacuo. Nearly pure protected estrone (23) was collected in 98.8% yield.

^{1}H NMR (600 MHz, Chloroform-d) δ 7.44 (d, J=8.7 Hz, 1H), 6.62 (dd, J=8.7, 3 Hz, 1H), 6.54 (d, J=2.6 Hz, 1H), 2.94 (s, 1H), 2.86 (s, 2H), 2.81 (dd, J=5.3, 2.3 Hz, 1H), 2.5 (d, J=17.4 Hz, 1H), 2.28 (t, J=3.2(x2) Hz, 2H), 2.18 (m, 3H), 2.12 (s, 1H), 1.63 (m, 2H), 1.42 (dd, J=12.7, 5.1 Hz, 1H), 0.92 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-d) δ 220.56, 154.65, 137.38, 127.55, 125.14, 118.32, 116.79, 47.86, 46.26, 38.21, 36.65, 36.24, 33.93, 31.60, 29.72, 27.81, 26.59, 25.63, 22.55, 19.65, 14.51.

Benzylamine (24)



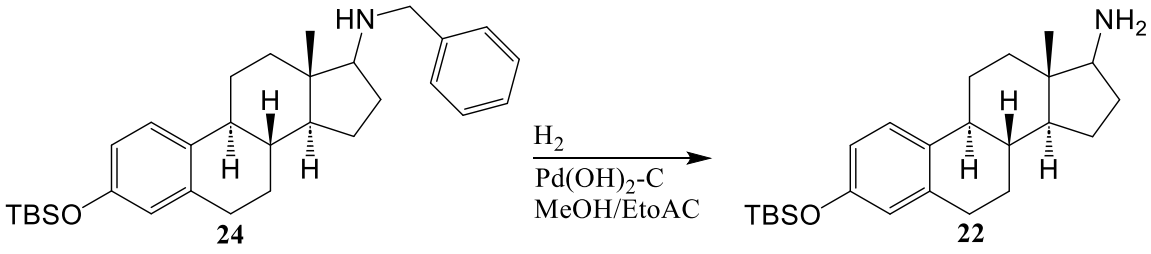
Protected estrone (23) (2.815 g, 7.319 mmole, 1 eq) was added to a 250 mL round bottom flask along with 80 mL of dichloromethane. Next, benzylamine (3.18 mL, 29.129 mmole, 3.98 eq) and glacial acetic acid (1.256 mL, 21.957 mmole, 3 eq) were added to the reaction mixture. After 10 minutes, sodium triacetoxyborohydride (3.102 g, 14.638 mmole, 2 eq) was added to the mixture, and was stirred at room temperature for 48 hours.

Monitoring of the reaction was performed by thin layer chromatography. After completion of the reaction, 50 mL of water was added to the reaction mixture, and the crude material was extracted with 3x50 mL of dichloromethane. The organic fractions were collected, dried over anhydrous sodium sulfate, filtered, and concentrated in vacuo. Crude material was purified using a hexane ethyl acetate mixture (9:1 v/v) to yield benzylamine (24) in 89.6% yield.

^{1}H NMR (600 MHz, Chloroform-d) δ 7.16 (m, 2H), 7.12 (m, 2H), 7.04 (m, 1H), 6.92 (d, J=8.7 Hz, 1H), 6.42 (dd, J=8.5, 2.8 Hz, 1H), 6.36 (d, J=3 Hz, 1H), 3.65 (m, 2H), 2.61 (m, 2H), 2.47 (t, J=8.7(x2) Hz, 1H), 2.08 (m, 1H), 1.97 (m, 1H), 1.66 (m, 1H), 1.51 (m, 1H), 1.31 (m, 2H), 1.23 (m, 2H), 1.14 (m, 3H), 1.07 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-d) δ 153.27, 141.19, 137.84, 133.24, 128.29, 127.96, 126.74, 126.10, 119.94, 117.12, 68.34, 52.79, 52.35, 44.11, 43.21, 38.80, 38.16, 29.71, 27.51, 26.48, 25.73, 23.58, 18.17, 11.97.

Amine (22)



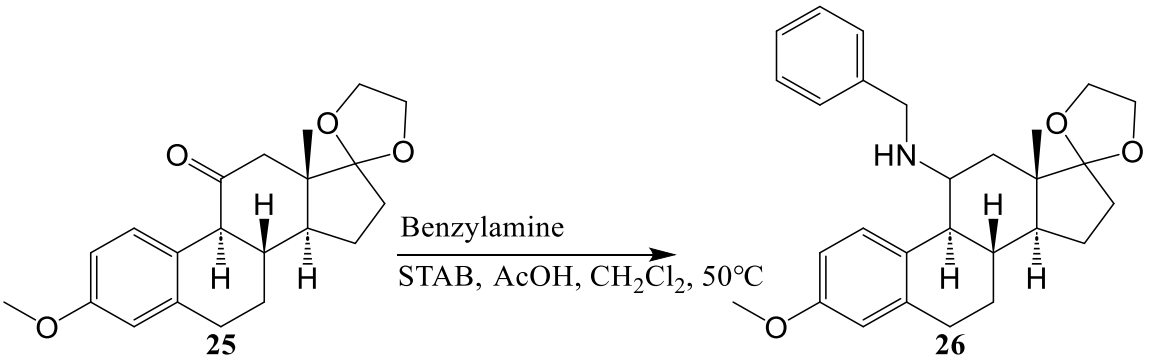
3.12 grams of benzylamine (24) (6.55 mmole, 1 eq) was added to a 250 mL round bottom flask with a 5:1 v/v mixture of methanol and ethyl acetate (120 mL). Next, 15%

w/w Palladium hydroxide on carbon was added to the mixture (0.468 g) and the reaction stirred under hydrogen gas at atmospheric pressure for 72 hours. After completion of the reaction by thin layer chromatography, the palladium was filtered off through celite, and the mixture concentrated in vacuo to afford pure amine (22) in 91.0% yield.

^{1}H NMR (600 MHz, Chloroform-d) δ 6.93 (d, J=8.3 Hz, 1H), 6.42 (m, 1H), 6.37 (d, J=3 Hz, 1H), 2.62 (m, 2H), 2.56 (t, J=8.7(x2) Hz, 1H), 2.11 (dd, J=13.6, 3 Hz, 1H), 1.95 (m, 3H), 1.68 (m, 2H), 1.53 (m, 1H), 1.28 (m, 2H), 1.23 (m, 1H), 1.14 (m, 2H), 1.06 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-d) δ 153.28, 137.90, 133.16, 126.10, 119.97, 117.14, 62.83, 52.07, 50.37, 44.04, 42.92, 39.04, 36.73, 31.18, 29.66, 27.47, 25.73, 23.36, 18.18, 11.12.

Benzylamine (26)



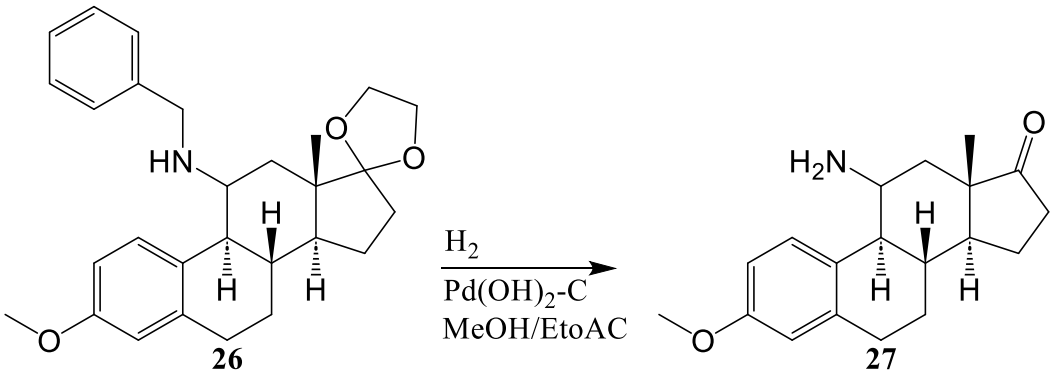
Ketone (25) (0.077 g, 0.225 mmole, 1 eq) was dissolved in 10 mL of dichloromethane followed by addition of benzylamine (0.1 mL, 0.895 mmole, 3.98 eq) and glacial acetic acid (0.04 mL, 0.675 mmole, 3 eq). After 10 minutes, sodium triacetoxyborohydride (0.095 g, 0.45 mmole, 2 eq) was then added to the reaction mixture and the reaction left at room

temperature for 24 hours. Monitoring of the reaction by thin layer showed no conversion to product, so we added an additional equivalent of acetic acid and benzylamine, as well as heating to 50℃ for 48 hours. This saw conversion of the starting material to product. Water (20 mL) was added to the reaction mixture followed by extraction with 3x20 mL of ethyl acetate. The organic fractions were collected, dried over anhydrous sodium sulfate, filtered, and concentrated in vacuo. Crude product was purified using a 9:1 hexane to ethyl acetate mixture to yield benzylamine (26) in 31.78% yield.

^{1}H NMR (600 MHz, Chloroform-d) δ 7.27 (m, 2H), 7.21 (m, 3H), 6.88 (d, J=8.7 Hz, 1H), 6.65 (dd, J=8.7, 3 Hz, 1H), 6.58 (d, J=3 Hz, 1H), 3.95 (m, 4H), 3.75 (s, 3H), 3.66 (s, 3H), 3.57 (d, J=2.3 Hz, 1H), 2.53 (d, J=4.5 Hz, 1H), 2.74 (d, J=3.4 Hz, 1H), 2.05 (m, 1H), 2.00 (dd, J=13.8, 2.5 Hz, 1H), 1.91 (t, J=10.4(x2) Hz, 2H), 1.83 (d, J=6.4 Hz, 1H), 1.74 (d, J=2.6 Hz, 2H), 1.68 (m, 3H), 1.44 (m, 2H), 1.35 (s, 3H).

^{13}C NMR (151 MHz, Chloroform-d) δ 157.44, 140.73, 139.90, 128.92, 128.55, 128.17, 126.74, 126.54, 119.56, 113.86, 112.38, 65.20, 64.53, 61.86, 55.14, 54.07, 52.15, 50.97, 49.58, 46.37, 34.29, 34.07, 31.19, 30.18, 26.99, 22.71, 22.16, 16.51, 14.14.

Amine (27)

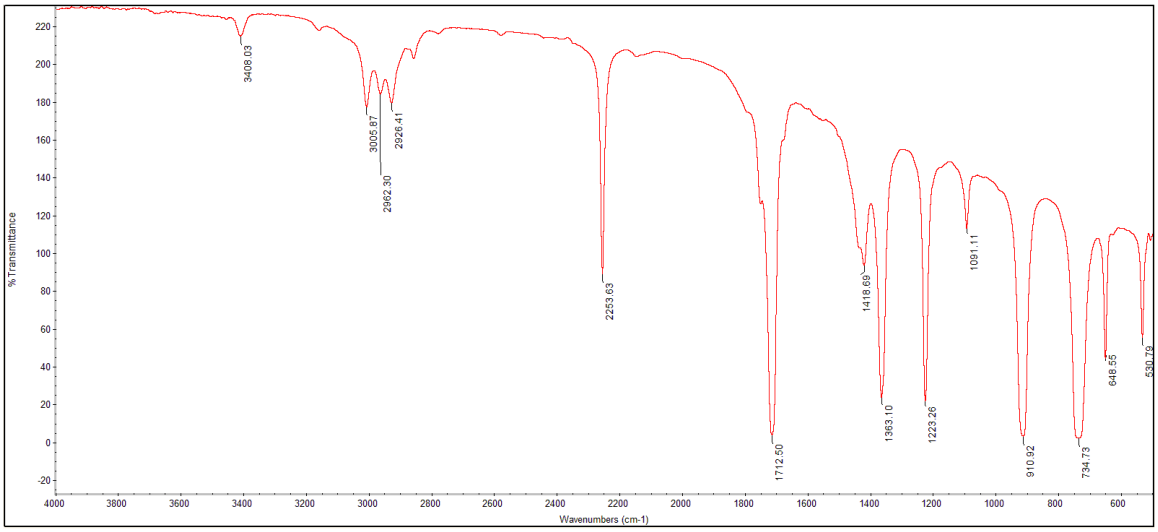


Benzylamine (26) (0.031 g, 0.0715 mmole, 1 eq) was dissolved in a 5:1 v/v mixture of methanol and ethyl acetate (6 mL). Palladium hydroxide on carbon (15% w/w, 5 mg) was added to the reaction mixture, and the reaction was stirred for 24 hours under hydrogen gas at atmospheric pressure. The reaction was monitored by thin layer chromatography and the product had not been completely formed after 24 hours. An additional balloon of hydrogen was added and stirred for another 48 hours to afford conversion to amine. The palladium was filtered off through celite, and the mixture concentrated in vacuo to yield crude amine (27).

^{1}H NMR (600 MHz, Chloroform-d) δ 7.06 (d, J=8.7 Hz, 1H), 6.73 (dd, J=8.7, 3 Hz, 1H), 6.62 (d, J=3 Hz, 1H), 5.12 (s, 2H), 3.78 (s, 3H), 2.84 (m, 2H), 2.74 (m, 2H), 2.61 (m, 1H), 2.35 (m, 4H), 2.04 (m, 4H), 1.92 (m, 4H), 1.26 (s, 3H).

IR (NaCl) ν_{max}/cm^{-1} 3510 and 3408 (N-H), 1712 (C=O), 1363 (C-O), and 1223 (C-N).

MS Calculated for $C_{19}H_{25}NO_2$ 299.41 found 299.80



4.5 Conclusions

Pancreatic cancer continues to be a detriment to those involved, and this work explored potential therapeutics that were created in our previous chapter. Further optimization of heterocyclic estrone analogs have shown promising results that functionalization of position 11 of the steroid structure yields increased binding to EGFR pathway proteins. Analogs with C-11 oxygenated modifications show improved biological activity[16], and prompted us to design C-11 nitrogenous analogs using an *in-silico* modeling process. We created 650+ analogs and docked them within EGFR pathway proteins and found that key C-11 amino heterocyclic estrone analogs showed promise in a number of proteins tested, including RAS, ERK, and MEK. This prompted us to optimize their synthetic creation. We first reoptimized the synthesis of 11-keto derivatives using new methods to avoid the use of low yield hydroboration reactions and difficult to separate reagents. This entailed using an epoxidation method and the uniqueness of the neighboring phenolic group to form the 11-keto groups through an electron rearrangement process. We then proceeded to optimize the synthesis of the C-11 amine analog. Initial studies focused on using the previously synthesized 11-hydroxyl intermediate with nitrogen nucleophiles. We found that phthalimides were too bulky to attach, and reduction of azido materials proved difficult to reduce due to steric hindrance or solubility issues. Direct amination was also unable to provide the amine, as we saw direct reduction from 11-ketone to 11-hydroxyl. Oxime reductions were also difficult in reducing, and we theorize this may be due to instability of the reduction intermediates. Further optimization however proved fruitful, as we successfully installed a benzylamine group to the 11-ketone and were able to reduce it

using hydrogenation to yield our 11-amino intermediate. Further work entails a full scale synthesis for the 11-amino heterocyclic analogs.

4.6 References

1. Hezel, A. F.; Kimmelman, A. C.; Stanger, B. Z.; Bardeesy, N.; Depinho, R. A., Genetics and biology of pancreatic ductal adenocarcinoma. *Genes Dev* **2006**, *20* (10), 1218-49.

2. Siegel, R. L.; Miller, K. D.; Jemal, A., Cancer statistics, 2018. *CA Cancer J Clin*, **2018**, *68* (1), 7-30.

3. Maitra, A.; Hruban, R. H., Pancreatic cancer. *Annu Rev Pathol* **2008,** *3*, 157-188.

4. Karim, S. A. M.; Abdulla, K. S.; Abdulkarim, Q. H.; Rahim, F. H., The outcomes and complications of pancreaticoduodenectomy (Whipple procedure): Cross sectional study. *Int J Surg* **2018,** *52*, 383-387.

5. De Dosso, S.; Siebenhuner, A. R.; Winder, T.; Meisel, A.; Fritsch, R.; Astaras, C.; Szturz, P.; Borner, M., Treatment landscape of metastatic pancreatic cancer. *Cancer Treat Rev* **2021,** *96*, 102180.

6. Venkatesulu, B. P.; Hsieh, C. E.; Sanders, K. L.; Krishnan, S., Recent advances in radiation therapy of pancreatic cancer. *F1000Res* **2018,** *7*.

7. Mizrahi, J. D.; Surana, R.; Valle, J. W.; Shroff, R. T., Pancreatic cancer. *The Lancet* **2020,** *395* (10242), 2008-2020.

8. Muler, J. H.; McGinn, C. J.; Normolle, D.; Lawrence, T.; Brown, D.; Hejna, G.; Zalupski, M. M., Phase I trial using a time-to-event continual reassessment strategy for dose escalation of cisplatin combined with gemcitabine and radiation therapy in pancreatic cancer. *J Clin Oncol* **2004,** *22* (2), 238-43.

9. Conroy, T.; Desseigne, F.; Ychou, M.; Bouche, O.; Guimbaud, R.; Becouarn, Y.; Adenis, A.; Raoul, J. L.; Gourgou-Bourgade, S.; de la Fouchardiere, C.; Bennouna, J.; Bachet, J. B.; Khemissa-Akous, F.; Pere-Verge, D.; Delbaldo, C.; Assenat, E.; Chauffert, B.; Michel, P.; Montoto-Grillot, C.; Ducreux, M., FOLFIRINOX versus Gemcitabine for Metastatic Pancreatic Cancer. *N Engl J Med* **2011,** *364* (19), 1817-1825.

10. Oliveira-Cunha, M.; Newman, W. G.; Siriwardena, A. K., Epidermal growth factor receptor in pancreatic cancer. *Cancers (Basel)* **2011**, *3* (2), 1513-1526.

11. Morgan, M. A.; Parsels, L. A.; Kollar, L. E.; Normolle, D. P.; Maybaum, J.; Lawrence, T. S., The combination of epidermal growth factor receptor inhibitors with gemcitabine and radiation in pancreatic cancer. *Clin Cancer Res* **2008,** *14* (16), 5142-5149.

12. Nagano, T.; Tachihara, M.; Nishimura, Y., Mechanism of Resistance to Epidermal Growth Factor Receptor-Tyrosine Kinase Inhibitors and a Potential Treatment Strategy. *Cells* **2018,** *7* (11).

13. Hsu, J. L.; Hung, M. C., The role of HER2, EGFR, and other receptor tyrosine kinases in breast cancer. *Cancer Metastasis Rev* **2016,** *35* (4), 575-588.

14. Xu, C. W.; Lei, L.; Wang, W. X.; Lin, L.; Zhu, Y. C.; Wang, H.; Miao, L. Y.; Wang, L. P.; Zhuang, W.; Fang, M. Y.; Lv, T. F.; Song, Y., Molecular Characteristics and Clinical Outcomes of EGFR Exon 19 C-Helix Deletion in Non-Small Cell Lung Cancer and Response to EGFR TKIs. *Transl Oncol* **2020,** *13* (9), 100791.

15. Patel, H.; Pawara, R.; Ansari, A.; Surana, S., Recent updates on third generation EGFR inhibitors and emergence of fourth generation EGFR inhibitors to combat C797S resistance. *Eur J Med Chem* **2017,** *142*, 32-47.

16. Alseud, K, M. Design, Synthesis, and Biological Screening of Novel Cucs-Inspired Estrone Analogs Towards Treatment of Pancreatic Adenocarcinoma. Ph.D. Dissertation. South Dakota State University, Brookings, SD 2019.

17. Ferreira, L. G.; Dos Santos, R. N.; Oliva, G.; Andricopulo, A. D., Molecular docking and structure-based drug design strategies. *Molecules* **2015,** *20* (7), 13384-421.

18. Acharya, C.; Coop, A.; Polli, J. E.; MackKerell Jr., A. D., Recent Advances in Ligand-Based Drug Design: Relevance and Utility of the Conformationally Sampled Pharmacophore Approach. *Curr Comput Aided Drug Des* **2011,** *7* (1), 10-22.

19. Bursulaya, B. D.; Totrov, M.; Abagyan, R.; Brooks III, C. L., Comparative study of several algorithms for fleixble ligand docking. *J Comput Aided Mol Des* **2003,** *17*, 755-763.

20. Totrov, M.; Abagyan, R., Flexible ligand docking to multiple receptor conformations: a practical alternative. *Curr Opin Struct Biol* **2008,** *18* (2), 178-184.

21. Cousins, K. R., Computer review of ChemDraw Ultra 12.0. ACS Publications: 2011.

22. Halgren, T. A. Merck Molecular Force Field. I. Basis, Form, Scope, Parameterization, and Performance of MMFF94. *J Comput Chem* **1996**, *17*, 490-519.

23. Hawkins, P.; Skillman, A.; Warren, G.; Ellingson, B.; Stahl, M. Conformer generation with OMEGA: algorithm and validation using high quality structures from the Protein Databank and Cambridge Structural Database. *J Chem Inf Model* **2010**, *50*, 572-584.

24. OMEGA 2.5.1.4: OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.

25. McGann, M. FRED pose prediction and virtual screening accuracy. *J Chem Inf Model* **2011**, *51*, 578-596.

26. OpenEye Toolkits. OpenEye Scientific Software, Santa Fe, NM. http://www.eyesopen.com.

27. Alsayari, A.; Kopel, L.; Ahmed, M. S.; Pay, A.; Carlson, T.; Halaweish, F. T., Design, synthesis, and biological evaluation of steroidal analogs as estrogenic/anti-estrogenic agents. *Steroids* **2017,** *118*, 32-40.

28. Ahmed, M. S.; El-Senduny, F.; Taylor, J.; Halaweish, F. T., Biological screening of cucurbitacin inspired estrone analogs targeting mitogen-activated protein kinase (MAPK) pathway. *Chem Biol Drug Des* **2017,** *90* (3), 478-484.

29. Mahnashi, M.; Elgazwi, S. M.; Ahmed, M. S.; Halaweish, F. T., Cucurbitacins inspired organic synthesis: Potential dual inhibitors targeting EGFR - MAPK pathway. *Eur J Med Chem* **2019,** *173*, 294-304.

30. Ahmed, M. S.; Kopel, L.; Halaweish, F. Structural optimization and biological screening of a steroidal scaffold possessing cucurbitacin-like functionalities as B-Raf inhibitors. *ChemMedChem* **2014**, *9*, 1361-1367.

31. Kopel, L.; Ahmed, M. S.; Halaweish. F. T. Synthesis of novel estrone analogs by incorporation of thiophenols via conjugate addition to an enone side chain. *Steroids* **2013**, *78*, 1119-1125.

32. Stephan, E.; Zen, R.; Authier, L.; Jaouen, G., Improved synthesis of a protected 11-oxoestrone. *Steroids* **1995,** *60*, 809-811.

33. Batista, V. S.; Crabtree, R. H.; Konezny, S. J.; Luca, O. R.; Praetorius, J. M., Oxidative functionalization of benzylic C–H bonds by DDQ. *New Journal of Chemistry* **2012,** *36* (5).

34. Tanemura, K.; Nishida, Y.; Suzuki, T., 2,3-Dichloro-5,6-dicyano-p-benzoquinone (DDQ) as the useful synthetic reagent. *Bulletin of the Nippon Dental University, General Education* **2011,** *40*, 31-34.

35. Palomino, E.; Heeg, M. J.; Horwitz, J. P.; Polin, L.; Brooks, S. C., Skeletal Conformations and Receptor Binding of Some 9,11-Modified Estradiols. *J Steroid Biochem Molec Biol* **1994,** *50*, 75-84.

36. Golubovskaya, L. E.; Minailova, O. N.; Rzheznikov, V. M., A simple method of obtaining 11-keto-9β-estra-1,3,5(10)-trienes: potential reactants for the synthesis of steroidal antigestagens. *Pharm Chem J* **2002,** *36* (9), 507-509.

37. Nepomniaschiy, N.; Grimminger, V.; Cohen, A.; DiGiovanni, S.; Lashuel, H. A.; Brik, A., Switch Peptide via Staudinger Reaction. *Org Lett* **2008,** *10* (22), 5243-5246.

38. Miriyala, B.; Bhattacharyya, S.; Williamson, J. S., Chemoselective reductive alkylation of ammonia with carbonyl compounds: synthesis of primary and symmetrical secondary amines. *Tetrahedron* **2004,** *60* (7), 1463-1471.

39. Slavikova, B.; Bujons, J.; Matyas, L.; Vidal, M.; Babot, Z.; Kristofikova, Z.; Sunol, C.; Kasal, A., Allopregnanolone and pregnanolone analogues modified in the C ring: synthesis and activity. *J Med Chem* **2013,** *56* (6), 2323-36.

40. Abiraj, K.; Gowda, D. C., Zinc/ammonium formate: a new facile system for the rapid and selective reduction of oximes to amines. *J Chem Research* **2003**, 332-334.

41. Szendi, Z.; Dombi, G.; Vincze, I., Steroids, LIII: New routes to aminosteroids. *Monatshefte fur Chemie - Chemical Monthly* **1996,** *127*, 1189-1196.

42. Leeds, K., A mild single-step reduction of oximes to amines. *Synth Commun* **1988,** *18* (8), 777-782.

43. Hernando, E.; Villalva, J.; Martínez, Á. M.; Alonso, I.; Rodríguez, N.; Gómez Arrayás, R.; Carretero, J. C., Palladium-Catalyzed Carbonylative Cyclization of Amines via γ-C(sp3)–H Activation: Late-Stage Diversification of Amino Acids and Peptides. *ACS Catalysis* **2016,** *6* (10), 6868-6882.

44. Dugar, S.; Schreiner, F. G.; Mahajan, D.; Sharma, A.; Patil, I. R.; Kuila, B. Novel compounds of 11β-hydroxy steroids for use in mitochondria biogenesis and diseases associated with mitochondria dysfunction or depletion. International Patent WO 2014/115167 A2, July 31, 2014.

45. Mostafa, Y. A.; Kralt, B.; Rao, P. P.; Taylor, S. D., A-ring substituted 17beta-arylsulfonamides of 17beta-aminoestra-1,3,5(10)-trien-3-ol as highly potent reversible inhibitors of steroid sulfatase. *Bioorg Med Chem* **2015,** *23* (17), 5681-92.

46. Patani, G. A.; LaVoie, E. J. Bioisosterism: A Rational Approach in Drug Design. *Chem Rev* **1996**, *96*, 3147-3176.

Chapter Five

General Conclusions and Future Directions

In this work, we describe three main projects to combat the poor prognosis of cancer patients by design, creation, and testing of novel estrone analogs against both epidermal growth factor receptor proteins and ABC transporter proteins responsible for dug resistance. Our first project details the use of triazole analogs on a steroid scaffold to target colorectal and ovarian cancer, two of the deadliest cancers. The triazole moiety provides key benefits for anticancer analogs including ease of synthesis and increased bioavailability within the body. This moiety was attached to a steroid scaffold, and we investigated the biological activity of these previously synthesized analogs in both cancers. The most active compound, Fz25, shows low micromolar activity in both the ovarian (15.29 ± 2.19 μM) and colorectal lines (15.98 ± 0.39 μM). Mechanism of action studies proved that Fz25 moderately arrests cells in the G1 phase of the cell cycle, specifically inhibiting STAT3 in both cell lines up to 65%. Ultimately Fz25 was also able to inhibit ABC transporter P-glycoprotein, which is involved in drug resistance, half as well as the known inhibitor, and Breast Cancer Resistance Protein better than the known inhibitor. Additional studies performed included molecular docking and dynamics, which identified that the key aspects of Fz25, including its benzyl group and free triazole hydrogen, are key for binding.

In our second project, we investigated the creation of novel estrone analogs with heterocyclic side chains as a potential pancreatic cancer therapeutic. Ultimately, twelve heterocyclic estrone analogs were synthesized and tested for their biological activity against two pancreatic cancer cell lines. In AsPC-1, TR33 had an IC_{50} of 38.37 ± 1.76 μM,

while in Panc-1 TR26 had an IC_{50} of 10.16 ± 0.83 μM. It was determined using cell cycle analysis that these analogs arrest cells in the G0/G1 phase of the cell cycle, a key feature highlighting a deficit in growth factors. In-cell western protein analysis assay confirmed these results in which extracellular signal-related kinase (ERK) was inhibited by both analogs at twice their IC_{50} value (up to 25%). These analogs were also investigated by colleagues and showed promising potential to reverse drug resistance to vincristine in a combination study, and additionally showed low micromolar activities in both a lung cancer and triple negative breast cancer cell line. Molecular dynamic simulations were also performed to confirm biological evaluations of hit compounds and found that TR26 binding to ERK was complimentary to its biological activity with key hydroxyl residues responsible for hydrogen bonding.

Further optimization of these heterocyclic estrone analogs have shown promising results that functionalization of position 11 of the steroid structure yields increased binding to EGFR pathway proteins. Analogs with C-11 oxygenated modifications show low micromolar activities and prompted us to design C-11 nitrogenous analogs using an *in-silico* modeling process. We created 650+ analogs and docked them within EGFR pathway proteins and found that key C-11 amino heterocyclic estrone analogs showed promise in a number of proteins tested, including RAS, ERK, and MEK. This prompted us to optimize their synthetic creation. We first reoptimized the synthesis of 11-keto derivatives using new methods to avoid the use of low yield hydroboration reactions and difficult to separate reagents. We then proceeded to optimize the synthesis of the C-11 amine analog. Initial studies focused on using the previously synthesized 11-hydroxyl intermediate with nitrogen nucleophiles. We found that phthalimides were too bulky to attach, and reduction

of azido materials proved difficult to reduce due to steric hindrance or solubility issues. Direct amination was also unable to provide the amine, as we saw direct reduction from 11-ketone to 11-hydroxyl. Oxime reductions were also difficult in reducing, and we theorize this may be due to instability of the reduction intermediates. Further optimization however proved fruitful, as we successfully installed a benzylamine group to the 11-ketone and were able to reduce it using hydrogenation to yield our 11-amino intermediate. Further work entails a full-scale synthesis for the 11-amino heterocyclic analogs. We hope that these heterocyclic analogs with 11-amine groups will provide increased bioavailability in the body and act better than their heterocyclic counterparts.

Future Directions

Due to the success our triazole analogs show against ABC Transporters, we wish to explore the creation of second and third round analogs with potentially more bioactivity. We have begun this process by synthesizing three new triazole analogs with benzyl side chains at C3 and addition of fluorine groups on the side chains. Testing has begun in a triple negative breast cancer line, which has shown preliminary data that Fz25 is still our hit compound and should be used for design of third round analogs. For our heterocyclic analogs, we know that the complete synthesis of analogs such as TR115-TR125 is needed now that the optimization for C-11 amine groups has been completed. Full synthesis of these analogs will be a hefty task, but their creation and biological investigation is needed. Upon successful synthesis, biological evaluation will be performed in pancreatic cancer cell lines to determine their effectiveness *in-vitro*.