T.C.

YEDİTEPE UNIVERSITY

INSTITUTE OF HEALTH SCIENCE

DEPARTMENT OF PHARMACEUTICAL CHEMISTRY

DEVELOPMENT OF A SERIES OF 5-SUBSTITUTED 6-FLUORO INDOLE DERIVATIVES AND EVALUATION OF THEIR BIOLOGICAL ACTIVITY

DOCTOR OF PHILOSOPHY THESIS

GÜLAY YELKEN DEMİREL, MSc.

İSTANBUL, 2023

T.C.

YEDİTEPE UNIVERSITY

INSTITUTE OF HEALTH SCIENCE

DEPARTMENT OF PHARMACEUTICAL CHEMISTRY

DEVELOPMENT OF A SERIES OF 5-SUBSTITUTED 6-FLUORO INDOLE DERIVATIVES AND EVALUATION OF THEIR BIOLOGICAL ACTIVITY

DOCTOR OF PHILOSOPHY THESIS

GÜLAY YELKEN DEMİREL, MSc.

Supervisor

Prof Dr. Filiz Esra Önen Bayram

İSTANBUL, 2023

THESIS APPROVAL FORM

Institute : Yeditepe University Institute of Health Sciences

Programme : Pharmaceutical Chemistry, PhD

Title of the Thesis : Development of a series of 5-substituted 6-fluoro indole derivatives and evaluation of their biological activity

Owner of the Thesis : Gülay Yelken Demirel, MSc.

Examination Date : 19/12/2023

This study have approved as a Doctorate Thesis in regard to content and quality by the Jury.

	Title, Name-Surname (Institution)
Chair of the Jury:	Prof. Dr. Hülya Akgün / Yeditepe University
Supervisor:	Prof Dr. Filiz Esra Önen Bayram / Yeditepe University
Member/Examiner:	Prof. Dr. Meriç Köksal Akkoç/ Yeditepe University
Member/Examiner:	Prof. Dr. Mert Ülgen/ Acibadem University
Member/Examiner:	Dr.Tuğçe Özyazici / Institute of Health Sciences

APPROVAL

This thesis has been deemed by the jury in accordance with the relevant articles of Yeditepe University Graduate Education and Examinations Regulation and has been approved by Administrative Board of Institute with decision dated ..... and numbered .....

Prof. Dr. Bayram YILMAZ

Director of Institute of Health Sciences

BEYAN

Bu tezin kendi çalışmam olduğunu, planlamasından yazımına kadar hiçbir aşamasında etik dışı davranışımın olmadığını, tezdeki bütün bilgileri akademik ve etik kurallar içinde elde ettiğimi, tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara kaynak gösterdiğimi ve bu kaynakları kaynaklar listesine aldığımı, tez çalışması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını beyan ederim.

19.12.2023

Gülay Yelken Demirel

DECLARATION

I hereby declare that this thesis is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree except where due acknowledgment has been made in the text.

19.12.2023

Gülay Yelken Demirel

DEDICATION

I dedicate this thesis to my son Poyraz Demirel.

ACKNOWLEDGEMENT

I would like to present my gratitude to my supervisor Prof. Dr. Esra Önen Bayram for being such a great mentor and guiding me, and she always believed that I could succeed.

I would like to thank Dean of Faculty of Pharmacy at Yeditepe University Prof. Dr. Meriç Köksal Akkoç for her support in all aspects. I would like to thank Prof. Dr. Hülya Akgün and all my professors that I took courses with during my doctoral education. I would also like to express my gratitude to all technical and academic staff, especially Sedat Kale who is Research Assistant at Department of Pharmaceutical Chemistry at Yeditepe University for valuable contribution for every respect.

This thesis subject is also a collaboration project with Chemo Iberica Pharmaceutical Company, Italy. I would also like to thank to Angel Calderon who has been working API - Corporate R&D Director at Chemo Pharmaceuticals giving to collaborate with Chemo and Yeditepe University; to Giuseppe Barreca who has been working as R&D Director at Chemo Pharmaceuticals for spectral analyses. The starting material, chemicals used during this thesis activities are supplied by Chemo Iberica, and the spectral analyzes such as NMR, LC-MS are carried out by the company.

I would further like to extend my thanks to Prof. Dr. Mert Ülgen Acıbadem University for his support in the thesis monitoring committee, and Dr. İrem Durmaz Şahin at Koç University for biological activity studies.

I sincerely appreciate my parents for their encouragement and would especially like to thank my husband Cem Demirel for his continuous support during my studies.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

APT	Attached Proton Test
CAS	Chemical Abstracts Service
CF	Cystic Fibrosis
CFTR	Cystic Fibrosis Transmembrane Conductance Regulator
CHOL	Cholangiocarcinoma
COAD	Colon adenocarcinoma
COSY	Correlated Spectroscopy
DMSO	Dimethyl sulfoxide
ESCA	Esophageal Carcinoma
FDC	Fixed Dosed Combination
FT-IR	Fourier-transform infrared
GSH	Antioxidant glutathione
HPLC	High Performance Liquid Chromatography
HSQC	Heteronuclear Single Quantum Coherence
IR	Infrared spectra
KICH	Kidney Chromophobe
KPLS	Kernel-based Partial Least Squares Regression
LAML	Acute myeloid leukemia
LC-MS	Liquid chromatography–mass spectrometry
LIHC	Liver hepatocellular carcinoma
LUAD	Lung adenocarcinoma
LUSC	Lung squamous cell carcinoma
MLR	Multiple Linear Regression
MS	Mass Spectrum Analyses
NMR	Nuclear magnetic resonance
NSAID	Nonsteroidal anti-inflammatory drug
OVSAHA	Human High-Grade Serous Ovarian Cancer Cell Line
OVCAR-3	Human Ovarian Cancer Cell Line
PCR	Principal Components Regression
PLS	Partial Least Squares Regression
PRSS1	Cationic trypsinogen gene
Q^2	Coefficient of determination

QSAR	Quantitative Structure Activity Relationship
R_f	Retention Factor
RMSE	Root Mean Square Error
ROS	Reactive oxygen species
RP	Recursive Partitioning
SD	Standard Deviation
SKCM	Skin cutaneous melanoma
SPINK1	Serine protease inhibitor Kazal-type 1
SRB	Sulforhodamine B
STAD	Stomach adenocarcinoma
TLC	Thin Layer Chromatograph
UV	Ultraviolet

ABSTRACT

Yelken Demirel, G. Development of a series of 5-substituted 6-fluoro indole derivatives and evaluation of their biological activity, Yeditepe University Institute of Health Science, PhD Thesis on Pharmaceutical Chemistry Program, Istanbul, 2023.

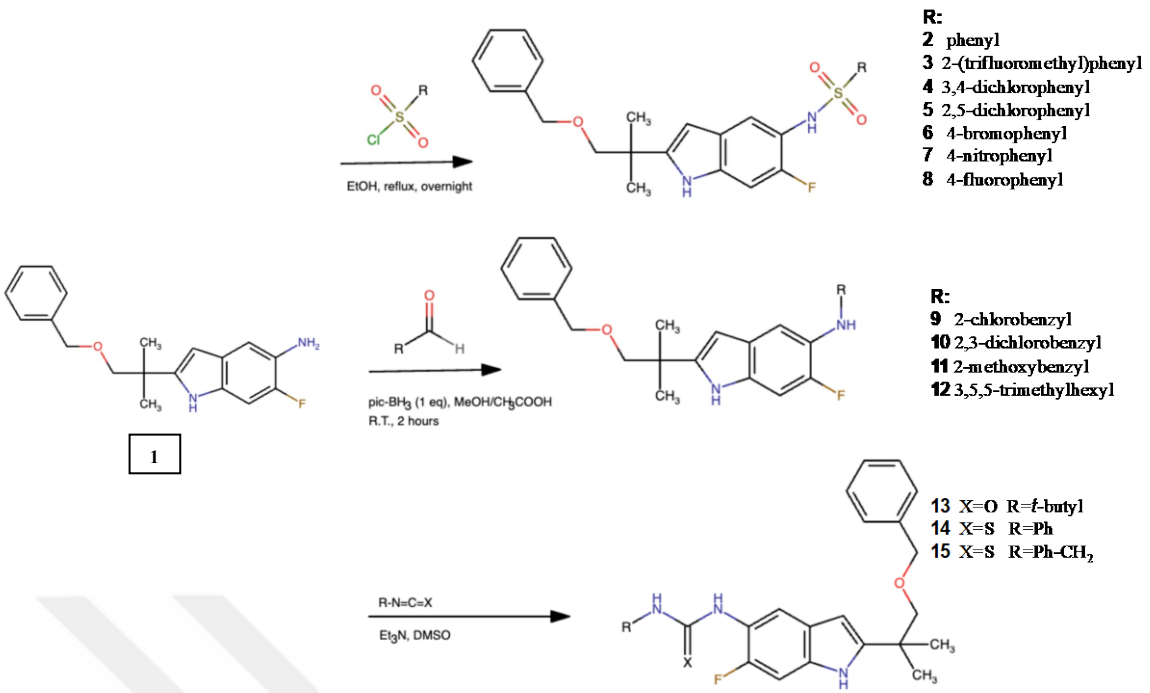
This study focused on the development of a quantitative structure activity relationship (QSAR) model to predict the cystic fibrosis transmembrane conductance regulator (CFTR) binding activity of a series of 5-substituted 6-fluoro indole derivatives, structures similar to the Tezacaftor molecule, an approved CFTR modulator drug. Also, synthetic routes for the preparation of these molecules were explored and their antiproliferative activity, due to the potential link between CFTR and cancer, was also aimed to be assessed.

A predictive QSAR model has been developed using regression-based learning techniques as Kernel-based Partial Least Squares regression (KPLS), providing satisfying statistical results as R^2=0.8865, Q^2= 0.7142, and RMSE (root mean square error) = 0.3985. Promising pEC50 values were predicted from the developed model since the obtained values were found to be between 4.79 and 5.11 for the all of the fifteen tested 5-substituted 6-fluoro indole derivatives.

To achieve the synthesis of these molecules various sulfonyl chlorides, aldehydes, isocyanates, and isothiocyanates were coupled with 2-[1-(benzyloxy) -2-methylpropan-2-yl] -6-fluoro-1H-indole-5-amine. Yet only some sulfonamide structures were isolated purely. The structures of synthesized compounds were elucidated using ^{1}H-NMR, ^{13}C APT-NMR, COSY-NMR, HSQC-NMR, and HPLC-UV-MS techniques.

Finally, the synthesized compounds were tested for their in vitro anticancer activity. Their cytotoxicity was evaluated using the NCI-Sulforhodamine B assay on ovarian OVSAHO and OVCAR-3 cancer cell lines. None of the synthesized 5-substituted 6-fluoro indole derivatives was found to exhibit notable cytotoxicity.

Keywords: QSAR, cystic fibrosis, CFTR modulators, cancer, indole, tezacaftor


EtOH, reflux, overnight
pic-BH3 (1 eq), MeOH/CH3COOH
R.T., 2 hours
R-N=C=X
Et3N, DMSO
1
R:
2 phenyl
3 2-(trifluoromethyl)phenyl
4 3,4-dichlorophenyl
5 2,5-dichlorophenyl
6 4-bromophenyl
7 4-nitrophenyl
8 4-fluorophenyl
R:
9 2-chlorobenzyl
10 2,3-dichlorobenzyl
11 2-methoxybenzyl
12 3,5,5-trimethylhexyl
13 X=O R=t-butyl
14 X=S R=Ph
15 X=S R=Ph-CH2

ÖZET

Yelken Demirel, G. Bir dizi 5-sübstitüe 6-floro indol türevlerinin geliştirilmesi ve biyolojik aktivitelerinin değerlendirilmesi, Yeditepe Üniversitesi Sağlık Bilimleri Enstitüsü, Farmasötik Kimya Programı Doktora Tezi, İstanbul, 2023.

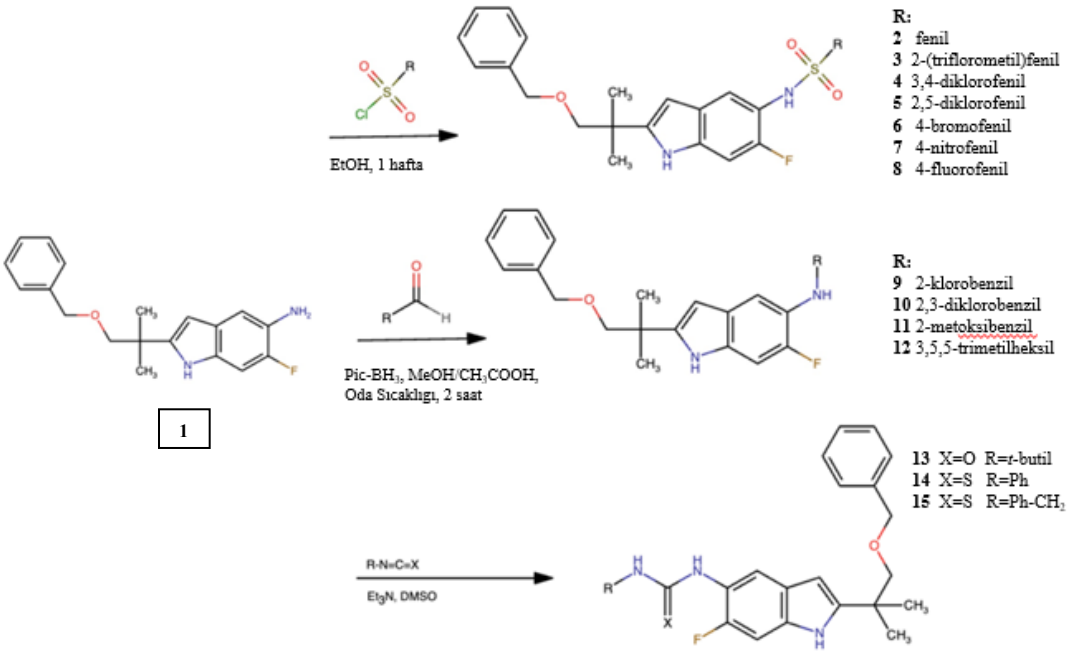
Bu çalışmanın amacı, bir kistik fibrozis transmembran regülatör proteini (CFTR) olarak onaylanmış Tezacaftor molekülüne benzer yapıda tasarlanmış, bir seri 5- sübstitüe 6-floro indol türevlerinin CFTR bağlanma aktivitelerini tahmin etmek için kantitatif bir yapı aktivite ilişkisi (QSAR) modelinin geliştirilmesidir. Ayrıca, bu moleküller için sentez yolları araştırılmış ve CFTR ile kanser arasındaki olası ilişkisi nedeniyle antiproliferatif aktivitelerinin değerlendirilmesi de amaçlanmıştır.

En küçük kareler yöntemi (KPLS) analiz metodu kullanılarak tahmine dayalı bir QSAR modeli geliştirilmiş; R^2=0.8865, Q^2= 0.7142 ve Kök Ortalama Kare Hatası (RMSE) = 0.3985 istatistiksel sonuçları elde edilmiştir. Test edilen on beş 5- sübstitüe 6-floro indol türevlerden gelecek vadeden pEC50 değerleri (4.79-5.11) elde edilmiştir.

Sentezde çeşitli sülfonil klorür, aldehit, izosiyanat ve izotiyosiyanatların 2-(1-(benziloksi)-2-metilpropan-2-il)-6-floro-5-amino-1H-indol-5-amin ile birleştirilerek sülfonamid, sekonder amin, üre ve tiyoüre türevleri oluşturulması hedeflenmistir. Sentezlenen bileşiklerin yapıları ise; ^{1}H-NMR, ^{13}C APT-NMR, COSY-NMR, HSQC-NMR ve HPLC-UV-MS analizleri ile aydınlatılmıştır.

In-vitro antikanser aktivite ve sitotoksite çalışmaları, yumurtalık OVSAHO ve OVCAR-3 kanser hücreleri üzerinde NCI-sülforodamin B yöntemi kullanarak değerlendirilmiştir. Sentezlenen 5- sübstitüe 6-floro indol türevleri herhangi bir sitotoksisite göstermemiştir.

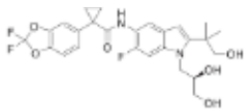
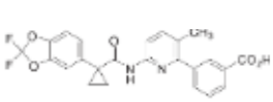
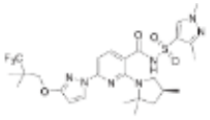
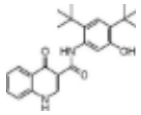
Anahtar kelimeler: QSAR, kistik fibroz, CFTR modülatörleri, kanser, indol, tezacaftor


EtOH, 1 hafta
R:
2 fenil
3 2-(triflorometil)fenil
4 3,4-diklorofenil
5 2,5-diklorofenil
6 4-bromofenil
7 4-nitrofenil
8 4-fluorofenil
Pic-BH3, MeOH/CH3COOH,
Oda Sıcaklığı, 2 saat
R:
9 2-klorobenzil
10 2,3-diklorobenzil
11 2-metoksibenzil
12 3,5,5-trimetilheksil
1
R-N=C=X
Et3N, DMSO
13 X=O R=t-butil
14 X=S R=Ph
15 X=S R=Ph-CH2

1. INTRODUCTION AND AIM

Cystic Fibrosis (CF) is an autosomal recessive disease with serious, chronically debilitating morbidities and high premature mortality. CF is a condition caused by genetic mutations in the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) gene, leading to the absence or reduced function of the CFTR protein on cell surfaces. The CFTR protein serves as a chloride channel in the epithelial cells, which are responsible for regulating the balance of salt and water absorption and secretion. When chloride transport is disrupted in various organs, it results in the development of multiple health issues associated with CF. In individuals affected by CF, the defective CFTR protein leads to the accumulation of thick and sticky mucus specifically in the bronchi of the lungs [1]. Additionally, it causes problems such as impaired pancreatic function, reduced absorption of nutrients in the intestines, reproductive difficulties, and increased concentration of chloride in sweat. The primary source of illness and death among CF patients is lung disease [2]. For the last decade, with the growing number of people with CF disease, drugs known as CFTR modulators target the basic defect in CF have become available. These drugs are under development, however, only four of them are registered up until today and are used either alone or as combination [3]. The four compounds that are approved for this purpose are Ivacaftor, Lumacaftor, Tezacaftor, and Elaxacaftor[4].

Table 1.1 The chemical structures of CFTR Modulator Drugs

Tezacfator	Lumacaftor	Elexacaftor	Ivacaftor
			
$C_{26}H_{27}N_2F_3O_6$	$C_{26}H_{27}N_2F_3O_6$	$C_{26}H_{34}F_3N_7O_4S$	$C_{24}H_{28}N_2O_3$
936727-05-8	1152311-62-0	2216712-66-0	873054-44-5
Corrector	Corrector	Corrector	Potentiator

Cancer is the disease of our age with high mortality rates, resulting in a significant health and economic burden worldwide [5]. There will be thirteen million cancer related deaths with twenty-two million new cancers yearly by 2030 [6]. Recent studies show that

CF patients have more tendency to develop cancer even if the mechanism linking dysfunctional CFTR to carcinogenesis has not been fully clarified yet [7].

Moreover, CFTR plays a significant role in the initiation and progression of various cancers by controlling cell proliferation, metastasis, invasion, and programmed cell death. However, the diverse patterns of CFTR gene expression in different cancer types pose challenges in considering CFTR as a universal biomarker for all cancers. Studies have revealed that most cancers exhibit elevated CFTR gene expression in tumor tissues compared to normal tissues, except for CHOL (Cholangiocarcinoma), ESCA (Esophageal Carcinoma), KICH (Kidney chromophobe), LAML (Acute myeloid leukemia), SKCM (Skin cutaneous melanoma), and STAD (Stomach adenocarcinoma). Consequently, CFTR's potential as a prognostic marker has been identified in nine specific cancer types, where CFTR expression levels play a crucial role in predicting the effectiveness of immune checkpoint blockade therapy and forecasting patient outcomes. Hence, CFTR has the potential to emerge as a novel biomarker for predicting prognosis and assessing immune cell infiltration in cancer, particularly in the context of immune checkpoint suppression therapy. Further research will likely shed light on the future utility of CFTR as a valuable clinical tool in cancer management [8]. Given the role CFTR plays in the development and progression of various cancers, it is reasonable to consider the application of CFTR inhibitors in cancer treatment therapies.

The quantitative structure-activity relationship (QSAR) is a mathematical model for quantitative analysis to predict biological and physicochemical properties of molecules and it is based on application of machine learning techniques to datasets. Molecular descriptors are used to create a training set, afterwards, the training set of molecules is used to design a mathematical model for QSAR technique. Such mathematical models are referred to as a predictive model, which aims to predict the investigated property for new compounds. Several studies shows that it is also a significant object for predicting CFTR activities as well [9,10]. Based on the chemical structure and activity profile of F508del-CFTR correctors, QSAR models developed, and novel molecular descriptors found.

Therefore, in this thesis we aimed to develop a QSAR model to predict the molecular activities of some 5-substituted 6-fluoro indole derivatives, structures that were derived from the Tezacaftor CFTR modulator drug that also presents in its structure a 6-fluorindole

moiety. Structure that were evaluated *in silico* for their binding property were also aimed to be synthesized and tested for their cytotoxic activities on some cancer lines. These structures that were aimed to be evaluated with the developed model and subsequently be synthesized from various sulphonamides, secondary amine structures, urea and thiourea derivatives are presented in the following figure (Figure 1.1).

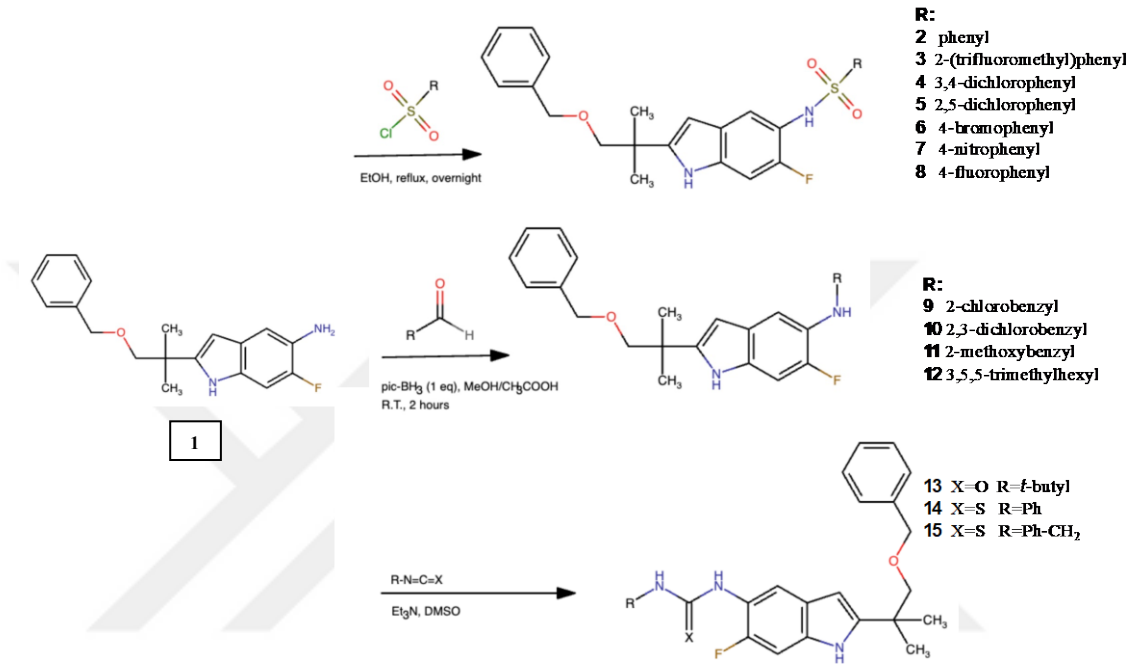


Figure 1. 1 Structure and synthesis pathways of of targeted compounds

2. GENERAL INFORMATION

2.1 General Information for Cystic Fibrosis Transmembrane Conductance Regulator Proteins

2.1.1 The CF Disease

CF is an inherited condition characterized by the presence of various mutations in the gene responsible for producing the CFTR protein. CFTR serves as an ion channel, facilitating the movement of chloride and sodium ions across cell membranes. The CFTR protein is primarily active in epithelial cells found in organs like the lungs, pancreas, liver, digestive system, and reproductive tract. Changes in the CFTR gene can lead to alterations in protein production, folding, or function, disrupting the normal flow of fluids and ions across cell membranes. Consequently, individuals with CF produce thick and sticky mucus that obstructs the ducts in affected organs. This obstructed mucus increases the likelihood of complications such as infections, lung damage, inadequate pancreatic function, and malnutrition [11].

Worldwide approximately 80.000 people in 2019 have CF and 70% of all patients are diagnosed by the age of two. Symptoms vary from patient to patient. General symptoms are excessive coughing, wheezing, bronchitis, chronic sinusitis, asthma, salty sweat/skin, clubbed finger, etc [11].

CF is caused by mutations in the CF gene. These mutations lead to defects in the CFTR protein. The most frequent mutation which is found in about 70% of CF patients, results in deletion of phenylalanine at position 508 of the CFTR amino acid sequence, denoted as F508del-CFTR [11]. As a result of these defects, the CFTR proteins do not work the way they should. This disruption is illustrated in Figure 2.1.

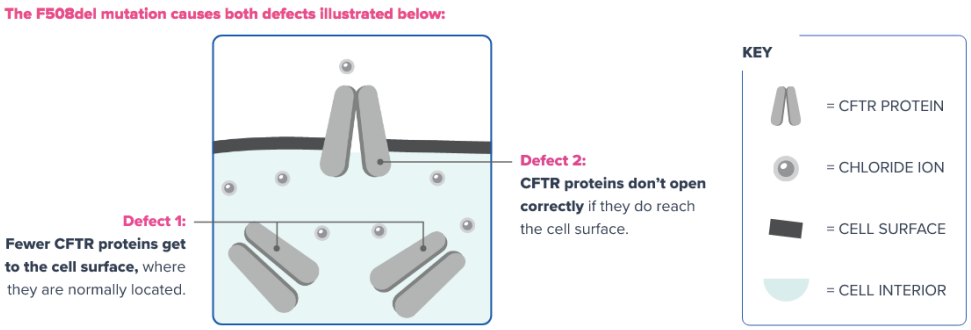


Figure 2.1 F508del mutation defect illustration [27]

The use of CFTR modulators, such as the approved molecules Tezacaftor and Elexacaftor, assists the defective proteins in positioning correctly on the cell surface, enabling them to fulfill their role as ion channels (Figure 2.2).

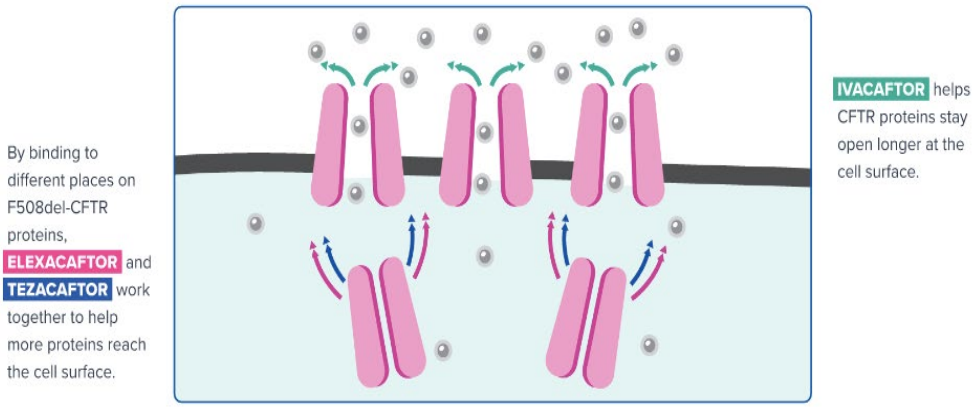


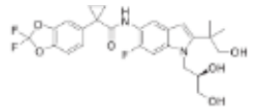
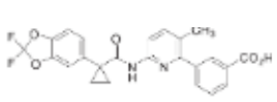
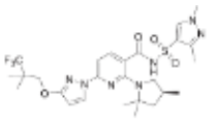
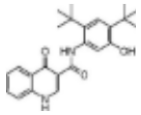
Figure 2.2 Three components function on F508del-CFTR proteins [12]

2.1.2 CFTR Modulators

CFTR modulators are a class of molecules that target the defective CFTR protein. These modulators work at the cellular level, aiding in either the correct folding and placement of the CFTR protein on the cell surface, or in enhancing the function of the CFTR protein once it has reached the cell surface [13]. There are, to date, four molecules that have been approved in the US and in Europe to manage CF [14]. These molecules are namely, Tezacaftor, Lumacaftor, Elexacaftor and Ivacaftor. The structure of these molecules and their mechanism of action are summarized in Table 2.1. Tezacaftor, Lumacaftor, and Elexacaftor are showing their activity by correcting the folding of the defective CFTR protein and supporting its transportation to the cell surface. These molecules are thus referred

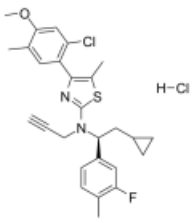
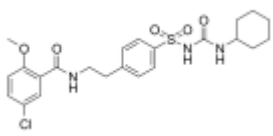
as corrector molecules [15]. Ivacaftor, whereas, is known as a potentiator since it improves the protein's ion transporting property. The molecule has been described to interact with the CFTR protein at the cell surface, particularly at the ATP-binding sites, enhancing thus the flow of more chloride and bicarbonate ions to out of the cell [16].

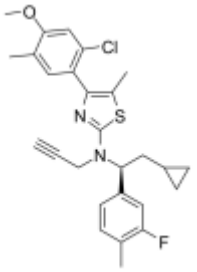
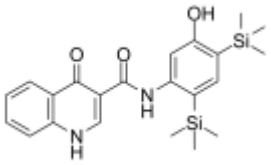
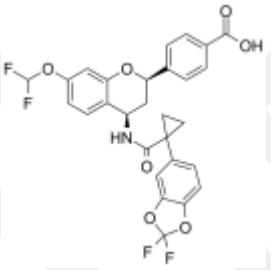
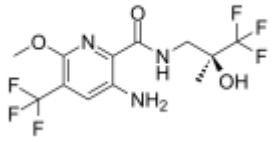
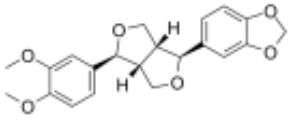
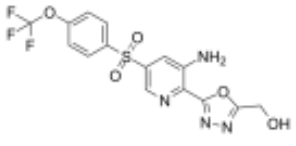
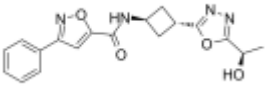
Table 2.1 FDA approved CFTR Modulator Drugs [17]

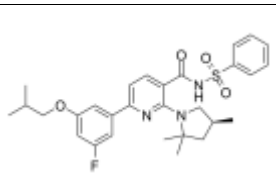
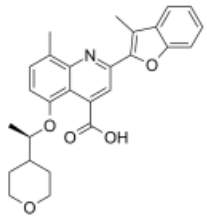
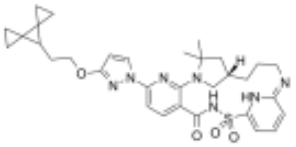
Tezacfator	Lumacaftor	Elexacaftor	Ivacaftor
			
$C_{26}H_{27}N_2F_3O_6$	$C_{26}H_{27}N_2F_3O_6$	$C_{26}H_{34}F_3N_7O_4S$	$C_{24}H_{28}N_2O_3$
936727-05-8	1152311-62-0	2216712-66-0	873054-44-5
Corrector	Corrector	Corrector	Potentiator

In addition to these well-known CFTR modulators, several compounds that have been identified as CFTR inhibitors, agonists, antagonists, activators, and modulators to their potential therapeutic benefits. A list of these molecules presents in Table 2.2.

Table 2.2 CFTR Drugs [18]

Chemical Compound	Chemical Structure	Mode of Action
Crinecerfont hydrochloride		CRF1 receptor antagonist
Glibenclamide		Inhibits P-glycoprotein

Chemical Compound	Chemical Structure	Mode of Action
Crinecerfont		CRF1 receptor antagonist
Dirocaftor		CFTR potentiator
Galicaftor		CFTR corrector
Icenticaftor		CFTR potentiator
Kobusin		CFTR activator
Navocaftor		CFTR modulator
Nesolicaftor		CFTR modulator

Chemical Compound	Chemical Structure	Mode of Action
Olacaftor		CFTR modulator
Posenacaftor		CFTR modulator
Vanzacaftor		CFTR modulator

2.1.3 CFTR and Cancer

Recent studies have suggested a potential link between CFTR and cancer. The inherent susceptibility of CF patients was described many years ago. The first report from CF patients that developed leukemia and Wilms' tumors dates from 1969 and the first pancreatic adenocarcinoma in 27 years old CF patients diagnosed in 1982. After that, the researchers focused on the relation between cancer and CF and several studies reported as cancer in patient with CF, mostly carcinomas of the digestive tract such as pancreatic, ileal, adenocarcinomas, and leukemia [19].

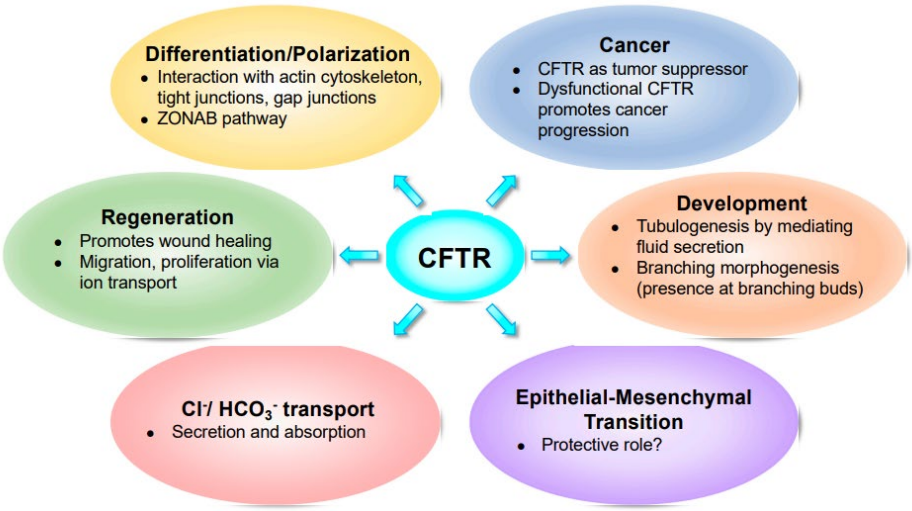


Figure 2.3 Several impacts of CFTRs [19]

Several studies have investigated the relationship between CFTR and various types of cancer, including colorectal, pancreatic, and breast cancer. Some studies have suggested that CFTR mutations may increase the risk of certain cancers, while others have suggested that CFTR dysfunction may play a role in cancer progression and metastasis [20]. One proposed mechanism for the role of CFTR in cancer involves the regulation of epithelial-mesenchymal transition (EMT), a process by which cells lose their epithelial characteristics and acquire a more mesenchymal phenotype, allowing them to migrate and invade other tissues. CFTR has been shown to regulate EMT by modulating the activity of other proteins involved in the process [21]. Another proposed mechanism involves the regulation of cellular metabolism. CFTR has been shown to play a role in the regulation of cellular energy metabolism [22].

Pancreatic Cancer

Chloride channels including CFTR, play a critical role in homeostasis of gastrointestinal tract with the functions involving osmoregulation, transport of major ions across epithelia, cellular glucose metabolism, tissue inflammation, cellular autophagy and protein turnover, migration of cells, innate and adaptive immune responses, programmed cell death, cell to cell interactions, membrane potential, mitochondrial function and related oxidative stress, polarity of cells, microbiota composition, cellular pH, mucus secretion and intestinal stem cell regulation. CFTR gene is known to be expressed in pancreatic ducts and promotes the dilution of pancreatic juice. Thus, the lack of CFTR function causes

concentrated pancreatic secretions, and CF disease certainly affects the pancreas and histological lesions appear from the early stage of life as the first year [23]. It is shown that mutations in CFTR gene are associated with pancreatic disease and demonstrate higher risk of pancreatic cancer with the mechanism likely being secondary to ion regulator dysfunction in pancreatic ductal cells impairing drainage. Many patients who are 23 to 39 years with CF disease have developed pancreatic adenocarcinoma [24,25].

The role of CFTR is still poorly understood, however, in 1996 the discovery of pancreatitis associated cationic trypsinogen mutations showed that trypsinogen plays a central role in the pathogenesis of human pancreatitis. In Figure 2.4, the normal pancreas trypsin shown as left side that prematurely activated within the pancreas is inhibited by SPINK1 and in the second line by trypsin and mesotrypsin preventing autodigestion. On the other hand, the inherited pancreatitis mutations in PRSS1 or SPINK1 as right side lead to an imbalance of proteases and their inhibitors resulting in autodigestion.

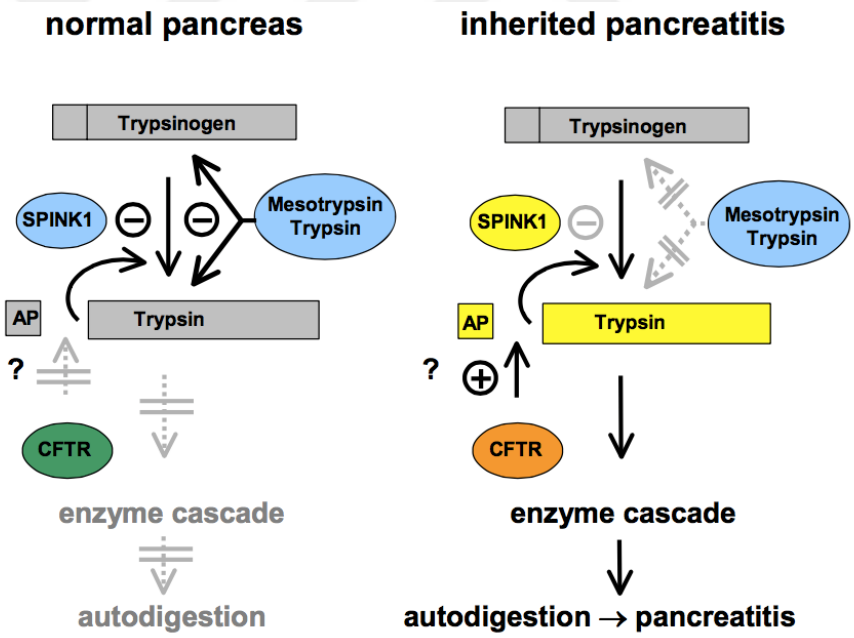


Figure 2.4 Model of inherited pancreatitis [26]

Intestinal Cancer

Recent epidemiological and clinical studies show that CF patients are at high risk for developing tumors in the colon especially due to CFTR's essential role in ion and acid-base homoeostasis with promoting chloride and bicarbonate secretion in the gastrointestinal tract. BL Than et al. conducted a study with mutant mice that carried intestinal specific knock-out of CFTR to investigate the effects of CFTR dysregulation on GI cancer. CFTR mutant mice developed significantly more tumors in the whole small intestine and colon as well, and CFTR deficiency alone caused the development of intestinal tumors in higher than 60% of in mutant mices which are approximately 1 years old mice. The study indicates that CFTR is a tumor suppressor gene in the intestinal tract [27, 28].

Due to causing biallelic inactivating mutations in CFTR gene, CF categorized as a familial colorectal cancer syndrome. CF patients are potentially candidates to early and aggressive colorectal tumor development. Patricia Scott *et. al* has shown a study that they carried out in with endoscopic screening that adenomas developed half of CF patients by the age of forty, aggressive advanced adenomas developed 25% CF patients and some of CF patients have already advanced to adenocarcinomas. Figure 2.5 shows CFTR deficiency leads to increased cellular response to Wnt/ β-catenin signaling which is a hallmark of colorectal cancer and cellular proliferation [29, 30].

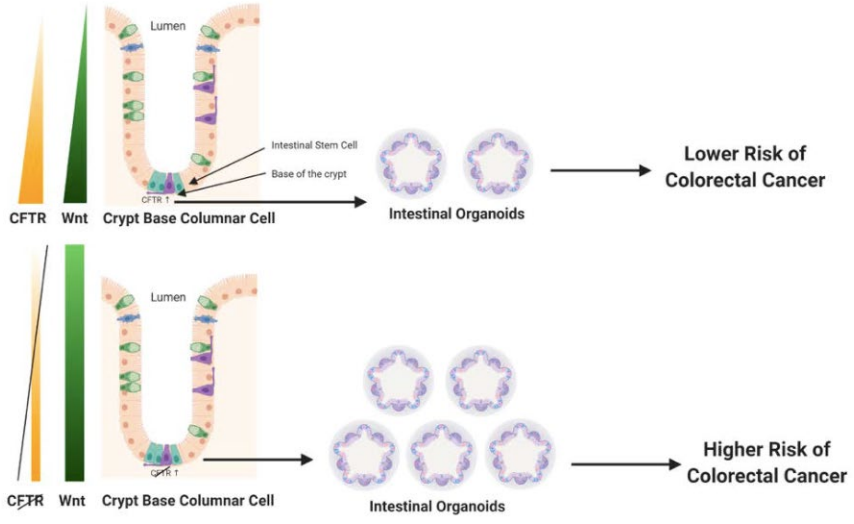


Figure 2.5 The relation between CFTR deficiency and colorectal cancer [30]

In the context of colon cancer, the deficiency of CFTR has been found to contribute to the improved survival of cancer cells when they experience oxidative stress, as depicted

in Figure 2.6. Cancer cells are particularly vulnerable to oxidative stress due to the heightened generation of reactive oxygen species (ROS), which can trigger cell death. However, in the case of colon cancer cells, CFTR deficiency acts as a protective mechanism by enhancing the retention of the antioxidant glutathione (GSH) within the mitochondria. The increased levels of GSH effectively lower ROS levels, subsequently reducing oxidative stress and the occurrence of apoptosis, or programmed cell death. Thus, through this mechanism, CFTR deficiency may facilitate the survival of colorectal cancer (CRC) cells [31].

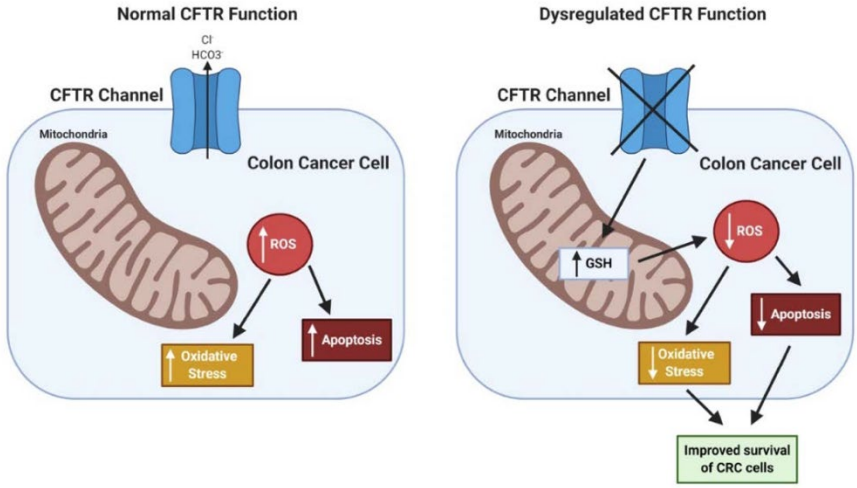


Figure 2.6 Colon cancer cells behaviors under oxidative stress with dysregulated CFRT function [31]

Cervical Cancer

Cervical cancer ranks as the third most prevalent cancer among women worldwide. Recent research has shed light on the pivotal role of CFTR in maintaining cellular equilibrium by regulating fluid balance and composition. CFTR gene expression can be influenced by estrogen and contributes to the production of cervical mucus and tubal secretions within the female reproductive system. CFTR and cervical cancer are closely linked to chronic inflammation, and investigations have highlighted the potential clinical significance of the CFTR gene in cervical cancer. Xue Peng et al. conducted a study aiming to explore the relationship between CFTR and the progression and prognosis of cervical cancer. By examining CFTR expression levels in various cervical tissues and cell lines, they observed a gradual increase in both CFTR mRNA and protein expression from normal to precancerous and cervical cancer tissues. Moreover, CFTR expression showed significant correlations with tumor stage, histological grades, lymphatic metastasis, vascular invasion,

interstitial invasive depth, tumor size, and HPV infection. The study suggests that CFTR expression level closely correlates with clinical and pathological factors, making CFTR a potential valuable biomarker in cervical cancer [32,33].

2.2. *In Silico* Modelling and QSAR

In silico models also known as computational non-testing methods represent a fast and reliable alternative approach to *in vivo* and *in vitro* methods. The main idea behind these approaches is the similarity principle: structurally similar compounds should have similar biological properties. It refers to computer-aided models that investigate pharmacological hypotheses using methods such as databases, data analysis tools, homology models, machine learning, pharmacophores, quantitative structure-activity relationships (QSAR), and network analysis tools. The three main families of *in silico approaches* are read-across (automated tools for data gap filling or for analogue search), QSAR, and expert systems (built based on expert knowledge). Data generated *via in silico* models can be applied to many domains including chemicals, pharmaceuticals, biocides, cosmetics, medical devices etc.

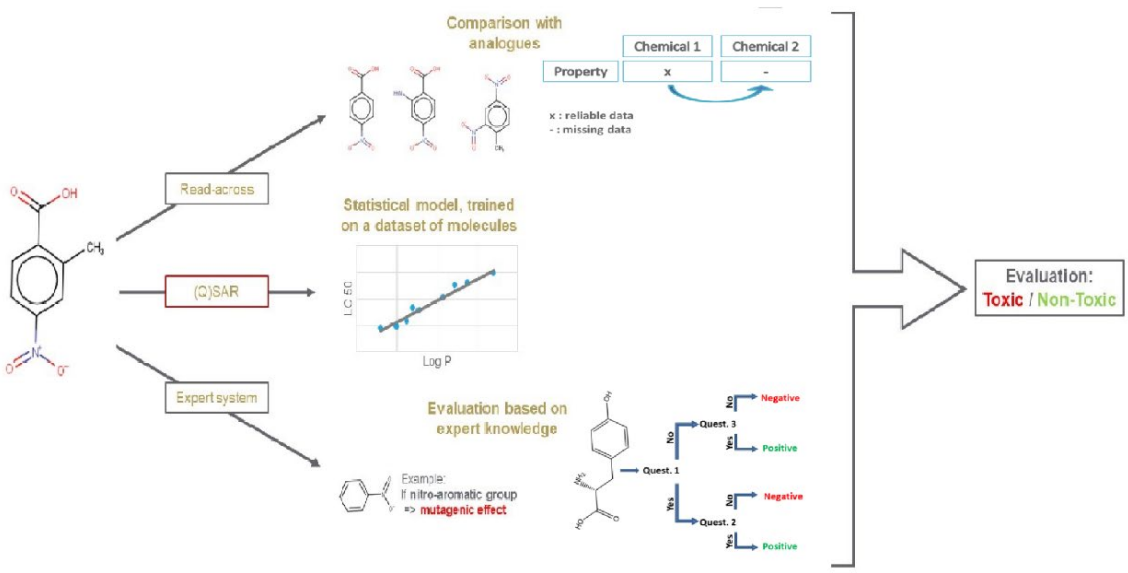


Figure 2.7 *In silico* Models

QSAR is a mathematical model for quantitative analysis to predict biological and physicochemical properties of molecules. It is based on the application of machine learning techniques to datasets with experimental or calculated properties. To execute this

quantitative analysis numerical features of molecules known as molecular descriptors are used to create a training set, afterwards, the training set of molecules is used to design a mathematical model for QSAR technique. Such mathematical models are referred to as a predictive model, which aims to predict the investigated property for new compounds.

2.2.1 History and Development of QSAR

QSAR is one of the most important computational techniques for ligand-based drug design, which can predict the correlation between the structural and bioactive properties of compounds.

Crum-Brown and Fraser presented an equation in 1868 that is regarded as the first comprehensive definition of a quantitative structure-activity relationship (QSAR) [34]. They discovered through their studies of several alkaloids that the permanently charged quaternary ammonium compounds formed by alkylation of the basic nitrogen atoms differed markedly from the basic amines in their biological effects. They suggested a mathematical equation to show that physiological activity is a function of the chemical structure.

In 1899, Meyer and Overton separately observed linear connections between narcotic activity and lipophilicity according to the oil-water partition coefficients, which measures lipophilicity. After that, Fuhner discovered that narcotic activities rise in a geometric progression within homologous series, providing the first proof that group contributions to biological activity levels are additive. Numerous other investigations that employed diverse lipophilicity metrics to characterize a range of non-specific biological activities corroborated this finding. Ferguson provided a thermodynamic explanation of such general structure-activity connections that also explained for observed biological activity levels above a particular range of lipophilicity [35].

QSAR technology quickly advanced from the mid-fifties on. For instance, Bruice, Kharasch, and Winzler developed an equation that demonstrated the quantitative contributions of various thyroid hormone analogues to their biological activity values [36].

In 1962 the first link between the toxicity of substituted benzoic acids and the electronic constants of their substituents was derived by Hansen. The association of the biological activity of phenoxy acetic acids with partition coefficients, Corwin Hansch's first

QSAR study, was published in the same year [34].

In 1964 the contemporary QSAR approach was created, and the time had come for broader explanations of how to treat structure-activity interactions quantitatively. Two publications were independently published: one by Free and Wilson on ´A mathematical contribution to structure activity studies´ and the other by Hansch and Fujita on ´A method for the correlation of biological activity and chemical structure´. Both contributions served as the foundation for the two new approaches to quantitative structure-activity relationships that would subsequently be known as Free Wilson analysis (a regression analysis approach using only indicator variables of different structural features) and Hansch analysis (correlates biological activity values with physicochemical properties by linear, linear multiple, or nonlinear regression analysis), accordingly [37,38].

2.2.2 Descriptors

QSAR models rely on molecular descriptors which are molecule´s structure and physicochemical properties [39]. Commonly used descriptors include molecular weigth, partition coefficient (log P), and other characteristics relevant to its biological and chemical activity. However, there is no single best descriptor that suits for all datasets.

Figure 2.8 Commonly used molecular descriptors in QSAR [40]

Descriptors can be categorized as 1D, 2D, 3D etc. Roberto T. [41] et al., defined descriptors in five different groups.

1D Descriptors

These descriptors are such as molecular weigth, melting point, and log P which can be calculated from directly the molecular structure.

2D Descriptors

These descriptors are such as topoligical indices and atom pairs considered the connectivity between atoms and their stereo chemistry.

3D Descriptors

These descriptors are three dimensional structure of molecules such as molecular shape, surface area, and electrostatic properties which are important for molecular modelling of complex compounds.

Quantum Chemical Descriptors

These descriptors involve quantum mechanical calculations and provide information such as electronic structure and reactivity of a molecule.

Fingerprint Descriptors

These descriptors are also known as structural keys or molecular fingerprints which are similarity-based encode whether presence or absence of spesific chemical structures within a molecule. There are several fingerprint types used in QSAR Modelling and most using models are given in Table 2.3 [42].

Table 2.3 Fingerprints utilized in QSAR models

Fingerprint Type	Description
Dendritic	Linear and branched fragments
Linear	Linear fragments and ring closures
MOLPRINT2D	A radial-like fingerprint that encodes atom environments using lists of atom types located at different topological distances
Radial	Fragments that grow radially from each atom. Also known as extended connectivity fingerprints.

Dendritic

A structure is broken down into fragments made up of linear paths and intersections of linear paths, with a default maximum of five bonds per path, to encode both linear and branching properties. No special handling of rings is applied to dendritic fingerprints.

Linear

All linear pathways in a structure that have up to a user-specified number of bonds, by default, this number is seven are found, and a hashing operation is used to create an integer bit address from a string-based description of each linear fragment. The default maximum path is increased from 7 to 14 bonds for ring closure solely to enhance the description of rings without drastically increasing the number of fragments. This only generates a small portion of the pieces that one would obtain if one enumerated all pathways comprising 0-14 bonds, but it effectively encodes information within and around most ring systems.

MOLPRINT2D

Each heavy atom in a structure has an environment made up of all other heavy atoms that are two bonds or less away from it. A bit in the fingerprint is derived from a tabular data structure that specifies the sorts of atoms detected at one and two bond distances and stores the core heavy atom type. Each item on the list is represented by a string of the form Type-freq(Type)-d, where freq(Type) represents the frequency with which a specific atom type is discovered at a specified distance from the center atom, denoted by d. Before being hashed to an integer that sets a bit in the fingerprint, these terms are sorted by distance, kind, and lastly, hash.

Radial

Radial fingerprints, often referred to as extended connectivity fingerprints (ECFPs or FCFPs), is created by repeatedly expanding a set of pieces radially from each heavy atom. By hashing a description of the atoms and bonds within the fragment and the bonds connecting it to the surrounding area, a variant of the Morgan algorithm is used to map each chemically unique fragment to a unique numeric value.

Data bases play a crucial role in data analyses and computational chemistry

modelling creation. Some data bases are summarized in Table 2.4.

Table 2.4 Databases for QSAR Modelling

Databases	**Sources**
Bioactivity data	ChEMBL, PubChem, BindingDB
Chemical database	ZINC, ChemSpider, GDB-17, Reaxys
Biological database	PDB, UniProt

2.2.3 Machine Learning Methods

A variety of regression-based and machine learning techniques are available to create linear and nonlinear QSAR models aiming biological endpoint prediction. A summary of machine learning statistical methods employed by QSAR is given Table 2.5.

Table 2.5 QSAR Machine Learning Methods [39]

Method	**Description**	**Supported variable types**	
		Dependent(y)	**Independent(x)**
MLR	Best subset multiple linear regression	Continuous	Descriptors
PLS	Partial least squares regression	Continuous	Descriptors
PCR	Principal components regression	Continuous	Descriptors
KPLS	Kernel-based PLS	Continuous	Descriptors, fingerprint
Bayes	Naive Bayes classification	Categorical	Descriptors, fingerprint
RP	Ensemble recursive partitioning	Categorical	Descriptors

MLR

The multiple linear regression technique (MLR) [43] is one of the easiest methods for developing calibration models since it uses a straight line to assess the relationship between a quantitative dependent variable and two or more independent variables.

PLS

The most often used chemometric techniques for calibrating models is partial least squares regression (PLS) [39]. It uses a linear approach to analyze data. It is employed in the analysis of various data because of how straightforward and minimal the calculations

are.

PCR

The most sophisticated method for multivariate analysis of spectral data is principal component regression (PCR) [44]. This strategy improves forecast accuracy by reducing the number of variables and eliminating noise.

KPLS

Kernel-based Partial least squares regression (KPLS) [45,46] can use scores to quantitatively describe how similar or dissimilar components are. KPLS, a nonlinear Gaussian kernel method, has a better capacity for prediction than traditional statistical techniques. By using a kernel, a nonlinear function of the scalar products, it extends partial least-squares regression by adding some nonlinearity to the scalar products of the X variables used in the regression.

Regression Model Analyses

To describe the model performance in different ways and below criteria should be used together.

The Root mean square error (RMSE) is calculated accordingly to the expression. Low RMSE values indicate that the model fits the data well and has more precise prediction. The lowest RMSE value can be used in selecting one prediction model over another. Generally, 0.4-0.5 values are acceptable.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (\tilde{y}_i - y_i)^2}{N}}$$

The coefficient of determination (Q^2) is often reported in QSAR publications complementary to r^2, it is calculated as follows: Q^2 can be interpreted as the percentage of

the variance in the property, explained by the model. It should be closed to 1 which is generally preferred >0.6.

$$q^2 = 1 - \frac{RMSE}{\sigma(y)} = 1 - \frac{\sum_{i=1}^{N} (y_i - \tilde{y}_i)^2}{\sum_{i=1}^{N} (y_i - E(y))^2}$$

The standard Deviation is a measure of how well a regression line fits the data. Higher SD values signify that more data points are further away from the mean.

The model Score is the performance of a given model on its training and test sets is used to compute a score that assesses the quality of the model.

$$score_M = accuracy_{test} \cdot (1.0 - |accuracy_{train} - accuracy_{test}|)$$

Graphical representation of score relation between accuracy of the model is given Figure 2.9. The z-axis represents the modelling score based on the accuracy of the training and test set predictions.

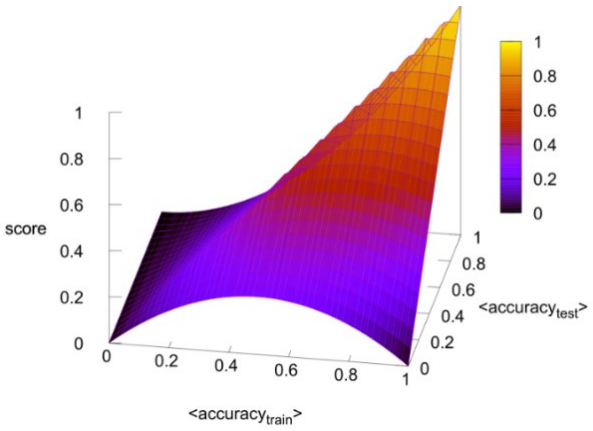


Figure 2.9 The scoring function of QSAR modelling [36]

Bayes

The Naive Bayes technique often employs discovered statistics to assess the stability of a ligand-protein complex by treating distributions of the parameters and predictions as

random variables.

RP

Unlike other machine learning methods, ensemble recursive partitioning (RP) [47] uses a decision tree to generate several descriptor planes.

2.2.4 Auto-QSAR

To achieve productivity through automation, several trends in drug discovery and QSAR modeling have come together. In recent years, increasing amounts of public data have become accessible, including data relevant to the pharmaceutical industry in repositories such as ChEMBL, PubChem, and ChemSpider [48]. Auto-QSAR, an automated machine-learning application, enables all machine learning models to build predictive QSAR models quickly and accurately. Starting with version 2016-1 of the Schrödinger Suite, the technique is encoded into a computer program that is a component of the Schrödinger Suite. The Auto-QSAR workflow is represented in Figure 2.10.

Auto-QSAR models can help selection ideal combinations of descriptors and learning methods. In this purpose, a model score is calculated to determine the model' s quality based on how well it performed on the training and test sets.

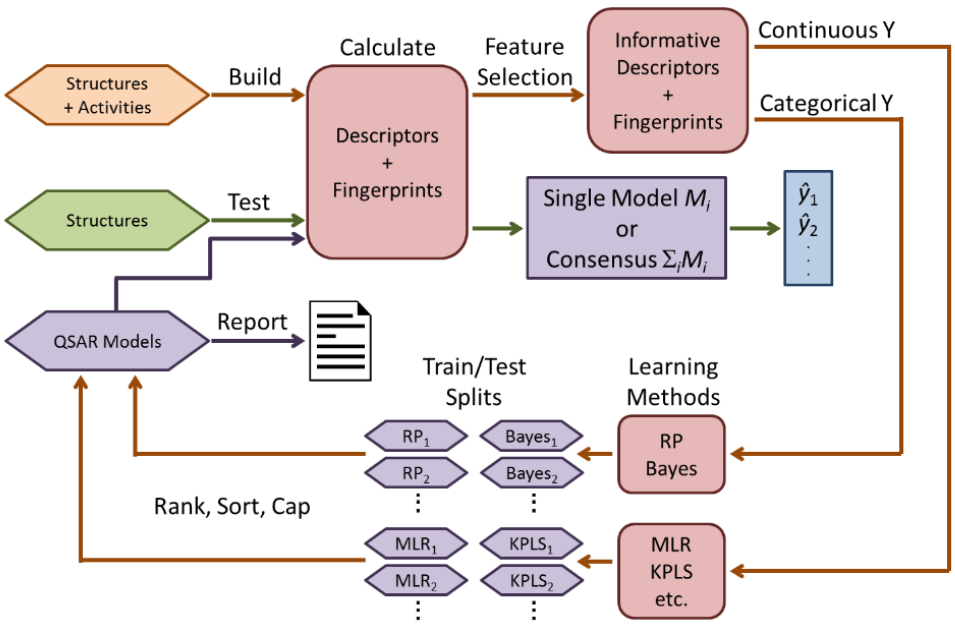


Figure 2.10 Schematic workflow of the Auto-QSAR work [49]

2.2.5 QSAR Studies Approach of CFTR Modulators

QSAR is an important object for predicting CFTR activities. Very recent studies published on developing QSAR modelling to speed up the discovery process of novel potent CFTR modulators. Based on the chemical structure and activity profile of F508del-CFTR correctors, Giada R. [9,10] et al. created a QSAR model and identified novel molecular descriptors that explain the SAR of their dataset. The first study was based on different descriptors referring to two-dimensional (2D) and three-dimensional (3D) parameters (three hundred molecular descriptors), with a dataset containing 75 compounds and 306 molecular descriptors using partial least-squares (PLS) multivariate analysis by means of MOE software. Indole-containing compounds, pyrazolquinoline, thienopyrane, cyanoquinoline and amino aryl thiazoles scaffolds showed a variable CFTR potentiator profile given in Figure 2.11. A highly predictive QSAR model has been developed (r^2=0.90, RMSE=0.347, r^2=0.86) with different 2D and 3D descriptors, PLS analyses via MOE software, identify pEC50 values (4.25-9.26).

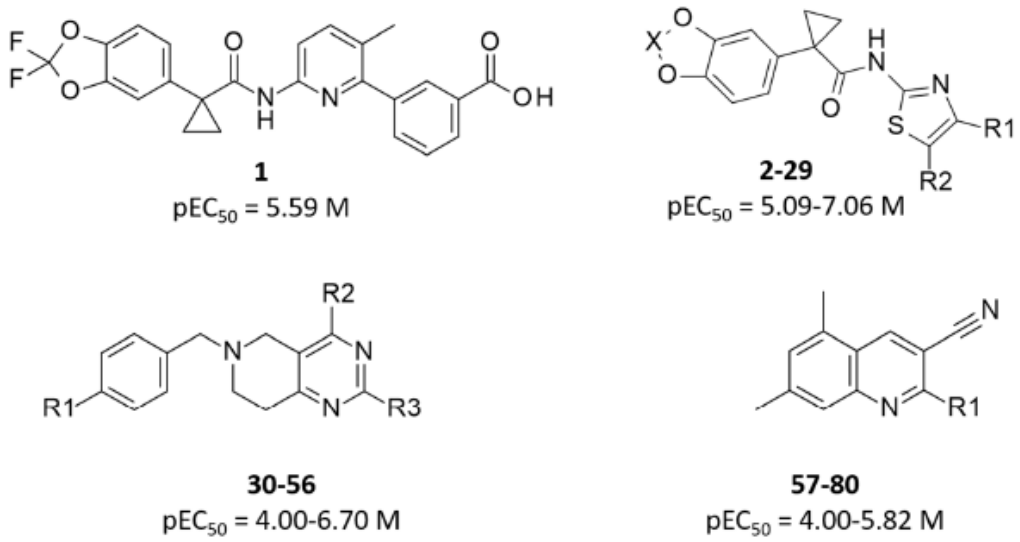


Figure 2.11 pEC50 values of the investigated F508del-CFTR correctors [9]

In the second study, 302 molecular descriptors (3D and 2D) were calculated by means of MOE software with the same statistical approach which used to predict new derivatives. In addition, the QSAR analysis module of the software MOE was used to generate the final mathematical models with the PLS multivariate analyses method. Sixty correctors were included in the training set for model generation and the others in the test set for model validation. The study group obtained the biological activity range for the

investigated F508del-CFTR correctors given in Figure 2.12. The pEC50 value´ s of Tezacaftor and Ivacaftor given as > 5 and 5.2 respectively in the European Medicines Agency (EMA) public assessment report published by [50].

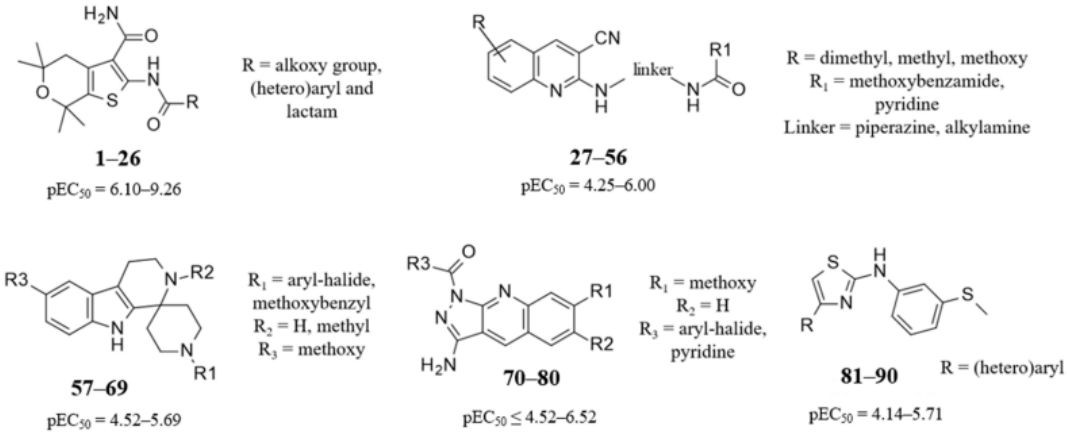


Figure 2.12 Chemical structures and biological activity range of the investigated F508del-CFTR correctors

2.3 Indole Chemistry

2.3.1 History

Indoles are heterocyclic organic compounds that have a bicyclic structure, consisting of a benzene ring fused to a pyrrole ring. They have diverse biological and chemical properties, making them useful in a variety of applications, including pharmaceuticals, agrochemicals, and materials science.

The first synthesis of indole was reported by Adolf von Baeyer in 1866, who synthesized it from oxindole by a pyrolytic method using zinc dust [51].

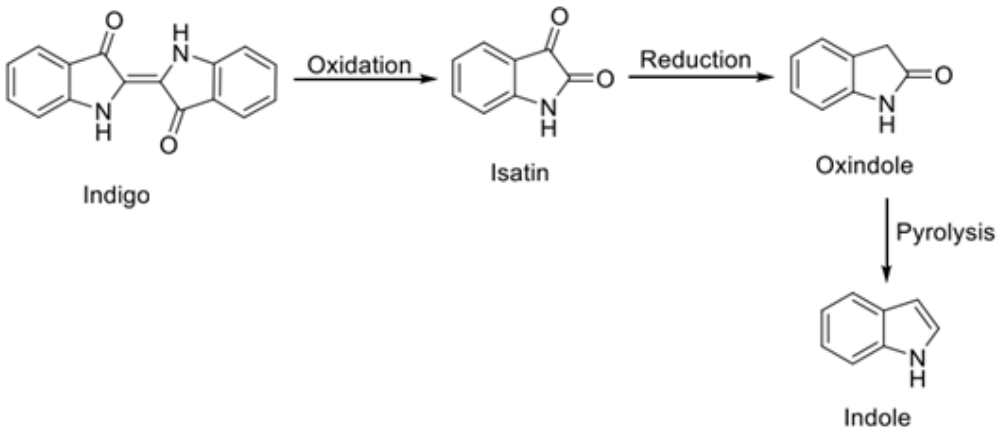


Figure 2.13 First Indole synthesis method

In 1896, Arthur Lapworth discovered that indole could be obtained by treating ortho-nitro cinnamic acid with zinc and hydrochloric acid [52]. However, the commercial production of indole did not begin until the early 20th century, when it was produced on a large scale by BASF and other chemical companies.

Some of natural indole derivatives are essentially important such as tryptophan, melatonin, serotonin etc. presented [53] in Figure 2.14.

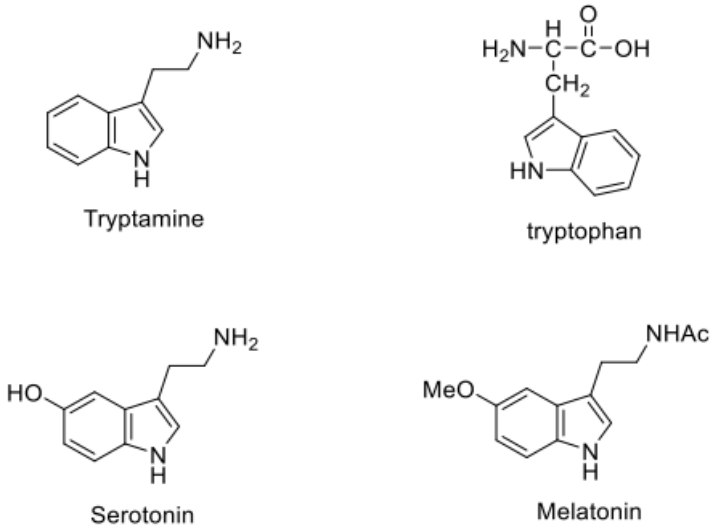


Figure 2.14 Some Important Natural Indole Derivatives

Indoles have been found to have a variety of biological activities, including antimicrobial, anticancer, and anti-inflammatory properties [54]. Many indole-containing

compounds have been developed as pharmaceuticals, including the antipsychotic drug clozapine, the anti-inflammatory drug indomethacin, and the anticancer drug vincristine [50]. Indoles are also used as intermediates in the synthesis of other compounds, such as dyes and agrochemicals.

2.3.2 Chemical Properties

The chemical structure of indole otherwise called as 1H-benzopyrrole is presented in Figure 2.15. The numbering of the atoms starts from nitrogen. It is a colorless crystalline solid and melts at 52 °C, soluble in most organic solvents such as alcohol, benzene and ether and can be recrystallized from water. Indole has a planar structure, and its benzene and pyrrole rings are coplanar, giving the molecule its characteristic shape. Indole is a relatively reactive molecule due to the presence of the nitrogen atom in the pyrrole ring. The nitrogen atom is capable of participating in various chemical reactions, such as nucleophilic substitution and electrophilic addition reactions.

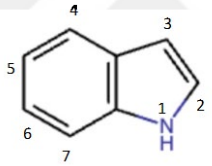


Figure 2.15 The molecular structure of Indole

Indole structure has 10 π electrons and two of them are coming from the lone pair of nitrogen and eight of them from carbon-carbon double bonds (C=C). It has resonance forms due to the delocalization of lone pair around the ring. The resonance energy of the structure is around 48 kcal/ mol and very weak acid with pKa= -3.63.

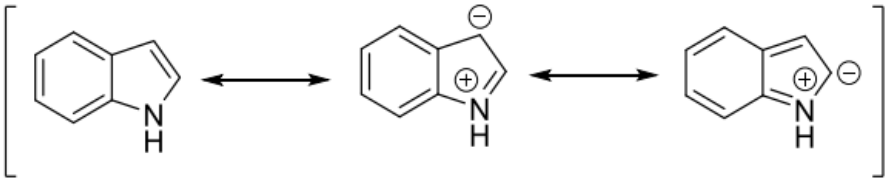


Figure 2.16 Resonance forms of 1H-indole structure

There are several indole isomers such as 3H-Indole, 2H-Indole and 1H-Isoindole are presented in Figure 2.17. Because of its higher reactivity, 1H-indole structure is the most well-known indole isomer [55].

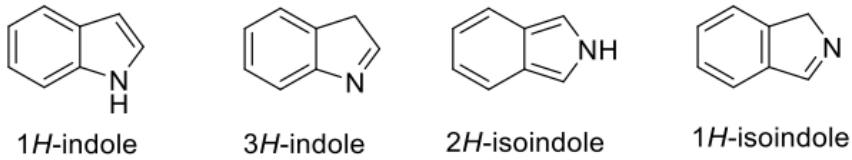


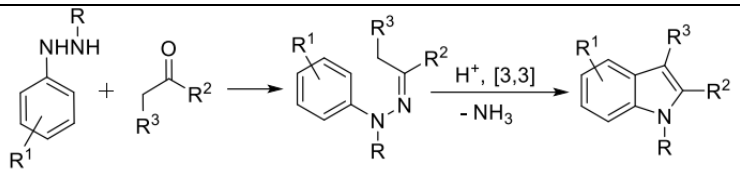
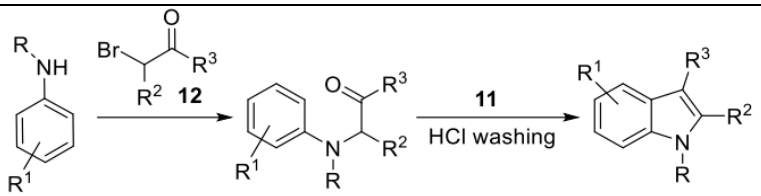
Figure 2.17 Indole Isomer

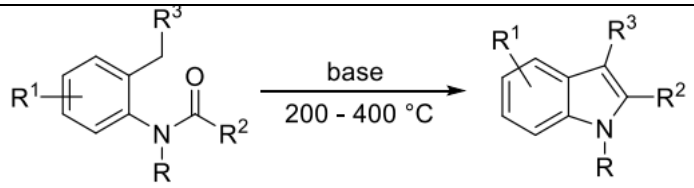
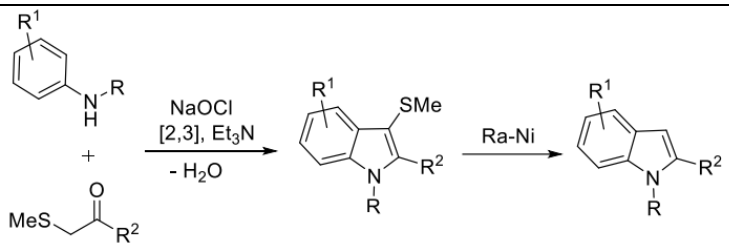
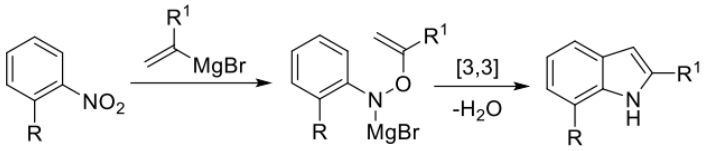
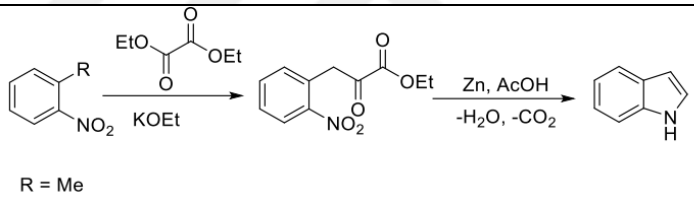
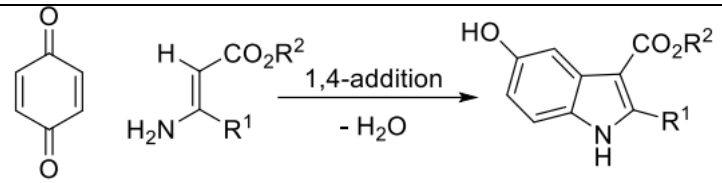
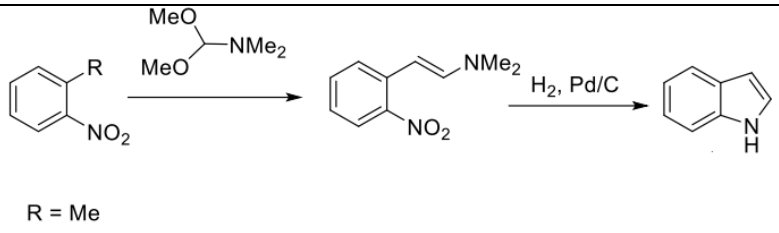
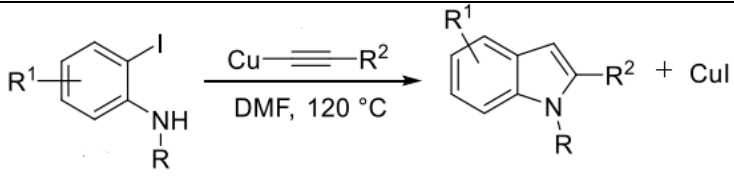
Indole can be oxidized to indoxyl and then to indigo, a blue dye. Indole can also be reduced to indoline which is used as a starting material for the synthesis of other compounds.

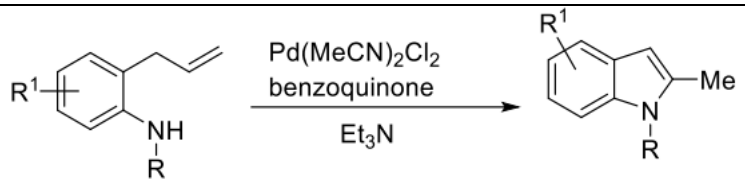
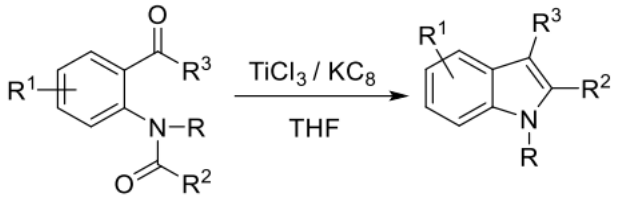
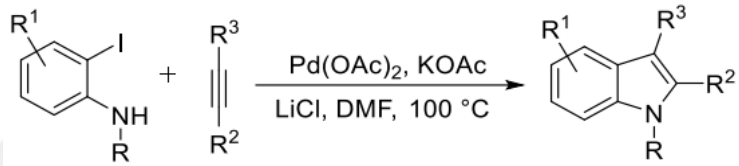
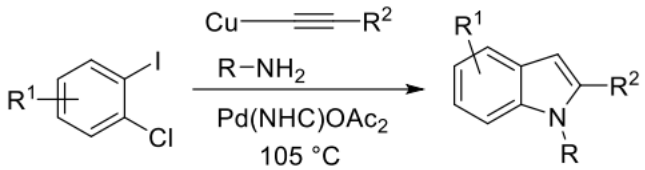
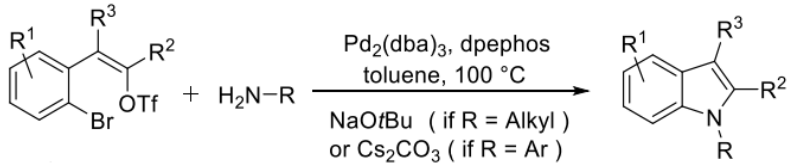
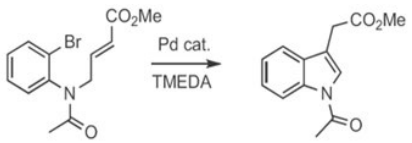
2.3.3 Chemical Synthesis Methods

Several synthesis methods are available for indole and its derivatives. Some of the important indole synthesis methods are summarized in Table 2.6. The Fischer indole synthesis is a classical method for the synthesis of indoles. In this reaction, a substituted phenyl hydrazine reacts with an aldehyde or ketone to form an imine intermediate. The imine then undergoes an acid-catalyzed cyclization to form the indole [56].

Table 2.6 Some synthesis methods for Indole and its derivatives

Name of the Reaction	Schematic Presentation of Reaction
The Fischer Indole Synthesis [57, 58]	
Bischler-Mohlau Synthesis [59]	

Name of the Reaction	Schematic Presentation of Reaction
Madelung Synthesis [60]	
Gassman Synthesis [61]	
Bartoli Synthesis [62]	
Reissert Synthesis [63]	
Nenitzescu Synthesis [64]	
Leimgruber-Batcho Synthesis [65]	
Castro [66]	

Name of the Reaction	Schematic Presentation of Reaction
Hegedus [67]	
Furstner [68]	
Larock [69]	
Ackermann [70]	
Willis [71]	
Mori [72]	

2.3.4 Spectral Analyses

Ultraviolet-Visible (UV-Vis) Spectroscopy

Indole has a characteristic UV-Vis absorption spectrum that is dominated by a strong absorption band in the range of 250-280 nm. This absorption band is attributed to the π-π* transition of the benzene ring in the molecule. In addition, indole also exhibits weaker absorption bands in the range of 280-350 nm, which are attributed to the n-π* transition of the pyrrole ring in the molecule [73,74].

FT-IR (Fourier-transform infrared)

Indole has a characteristic IR absorption spectrum that is dominated by several strong absorption bands.The most prominent absorption band is located at around 1620 cm^{-1}, which is attributed to the C=C stretching vibration of the benzene ring in the molecule. Other prominent absorption bands are located at around 3100-3500 cm^{-1}, which are attributed to the N-H stretching vibration of the pyrrole ring in the molecule for not substituted indole [75].

^{1}H Nuclear Magnetic Resonance (^{1}H NMR) Spectroscopy

Indole has a characteristic ^{1}H NMR spectrum that is dominated by several peaks in the range of 6-9 ppm. The peak at around 7.5 ppm is attributed to the proton on the nitrogen atom in the pyrrole ring, while the peaks in the range of 6-9 ppm are attributed to the protons on the carbon atoms in the benzene ring [76].

^{13}C Nuclear Magnetic Resonance (^{13}C NMR) Spectroscopy

Indole has a characteristic ^{13}C NMR spectrum that is dominated by several peaks in the range of 100-150 ppm. The peak at around 110 ppm is attributed to the carbon atom (C2) in the pyrrole ring, while the peaks in the range of 120-140 ppm are attributed to the other aromatic carbon atoms [76].

2.3.5 Pharmacological Importance of Indole Derivatives

The indole moiety is found in a wide range of substances with diverse biological uses, and it is a component of the pharmacophore structure of many synthetic drug compounds. An essential heterocyclic molecule with broad-spectrum biological activity is created when the indole nucleus is added to pharmaceuticals that have physiologically active pharmacophores.

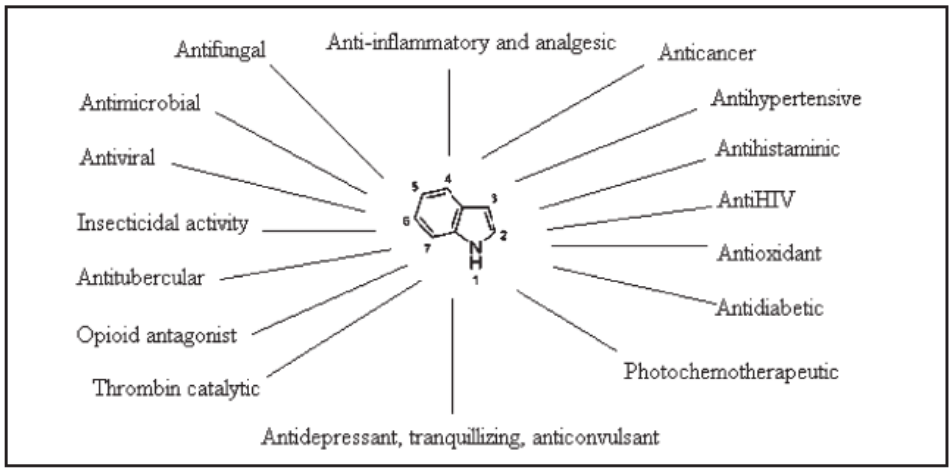
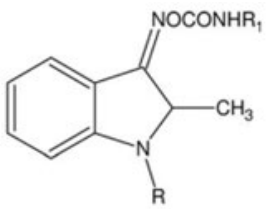


Figure 1.18 Several biological activities of substances with an indole moiety

Because of its numerous biological and therapeutic uses, indole derivatives are of great interest. These diverse biological activities make indole derivatives promising candidates for drug development and therapeutic interventions different biological actions, such as antidiabetic, anticancer, antimicrobial, anti-HIV, antiviral, anti-inflammatory, antioxidant, anticholinesterase, antitubercular, and antimalarial properties. In this section, the significant pharmacological activities of indole derivatives are discussed by presenting the most cited publications of the last twenty years.

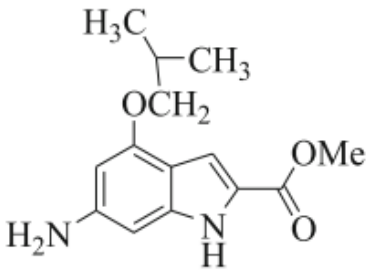
Antiviral Activity

Abele et al. [77] synthesized among the isatin and indole oximes by the indole oxime, which is the carbamoyl derivative of indole-3-oxime, demonstrated the strongest antiviral activity.

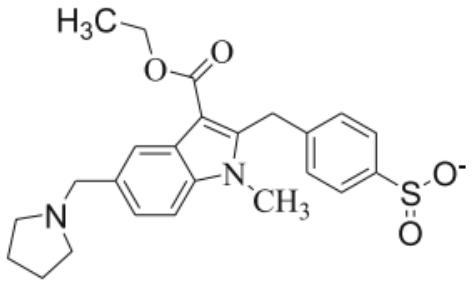


R = H, methyl, ethyl
R_1 = methyl, phenyl

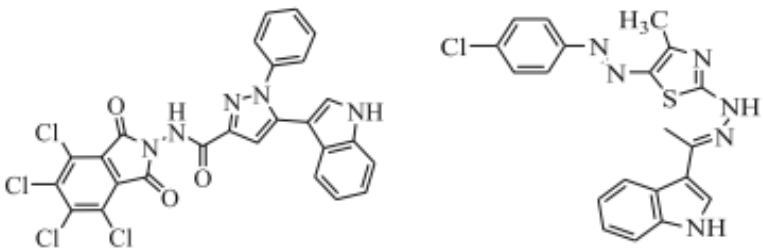
Xue et al. [78] synthesized and evaluated some 6-amino-4-substituted alkyl-1H-indole-2-substituted carboxylate derivatives as antiviral drugs and found that methyl 6-amino-4-isobutoxy-1H-indole-2-carboxylate derivative showed the highest selectivity index (SI) value 17.1 to CoxB3 virus and the highest IC50 value as 7.53 μmol/L against Influenza A.



Sellitto G. et al. [79] synthesized several ethyl 1H-indole-3-carboxylates derivatives and the most effective substance against the hepatitis C virus (HCV) at low concentrations was4-((3-(ethoxycarbonyl)-1-methyl-5-(pyrrolidin-1-ylmethyl)-1H-indole-2-yl) methyl) benzene sulfinate.



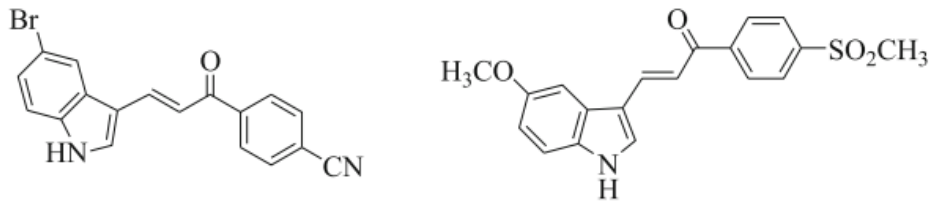
Abdel-gawad et al. [80] studied on 1,3-Thiazole and 1,2,4-triazolo[3,4-b]1,3,4-thiadiazine containing indole nucleus derivatives and test their HSV-1(herpes simplex type1) activity. Among the derivatives 5-(1H-indole-3-yl)-1-phenyl-N-(4,5,6,7-tetrachloro-1,3-dioxoisoindolin-2-yl)-1H-pyrazole-3-carboxamide and 2-(2-((E)-1-(1H-indole-3-yl) ethylidene) hydrazinyl)-5-((E)-(4-chlorophenyl) diazinyl)-4-methylthiazole showed best antiviral activity.



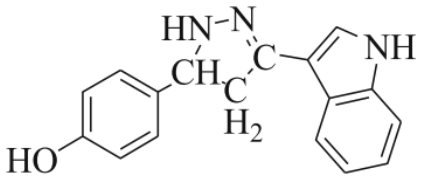
Anti-inflammatory and Analgesic Activity

Indole derivatives have been shown to possess anti-inflammatory and analgesic effects. Some indole derivatives, such as indomethacin and celecoxib, have been used as nonsteroidal anti-inflammatory drugs (NSAIDs) in clinical settings [81,82,83]. Indole derivatives are believed to exert their anti-inflammatory and analgesic effects by inhibiting the activity of cyclooxygenase (COX) enzymes, which are responsible for the biosynthesis of prostaglandins that mediate pain and inflammation [84].

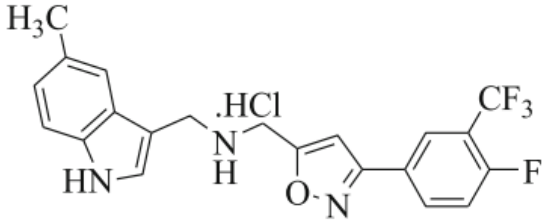
Ozdemir A et al. [85], studied on indole based chalcone derivatives as COX-1 and COX-2 inhibitor and 3-(5- Bromo-1H-indole-3-yl)-1-(4-cyanophenyl) prop-2-en-1-one and 3-(5-methoxy-1H-indole-3-yl)-1-(4-(methyl sulfonyl) phenyl) prop-2-en-1-one derivatives showed significantly active.



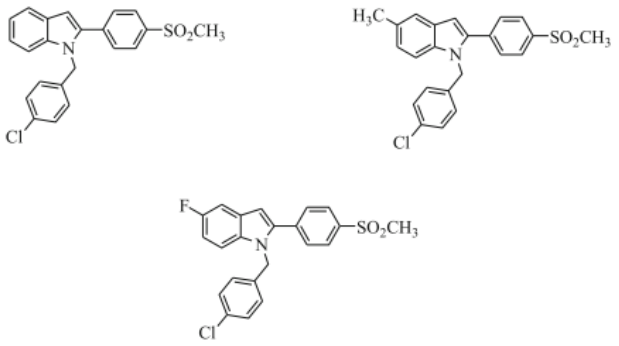
Indole based chalcone derivatives also studied by Rani et al. [86] for their anti-inflammatory activities and the most active synthesized derivative found as 4-(3-(1H-indole-3-yl)-4, 5-dihydro-1H-pyrazol-5-yl) phenol.



Pedada et al. [87] synthesized indole containing isoxazole derivatives and N-((3-(4-fluoro-3-(trifluoromethyl) phenyl) isoxazol-5-yl) methyl) (5-methyl-1H-indole-3-yl) methanamine hydrochloride found as significant sPLA2 inhibition activity.



Shaker AM et al. [88] synthesized 2-(4-(methylsulfonyl) phenyl)-1-substituted-indole derivatives and evaluate their in-vitro COX-2 and in-vivo anti-inflammatory activities. Non-substitue, 5-methyl, and 5-fluoro 1-(4-chlorobenzyl)-2-(4-(methylsulfonyl) phenyl)-1H-indole showed most in-vivo anti-inflammatory activity. Also, they performed a molecular modeling study on targeted derivatives that showed non-substitue, and 5-fluoro 1-(4-chlorobenzyl)-2-(4-(methylsulfonyl) phenyl)-1H-indole derivatives have promising binding interaction to COX-2 enzyme.



Some marketed anti- inflammatory drugs as indole derivatives presented in Figure 2.19 [89].

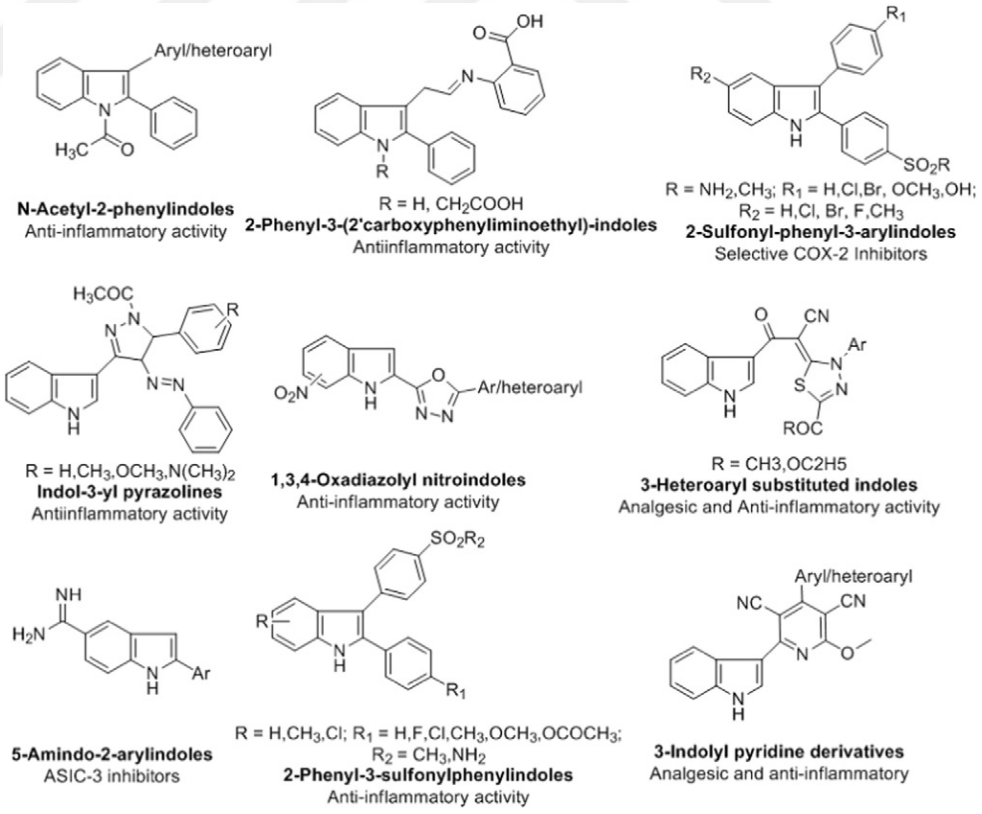
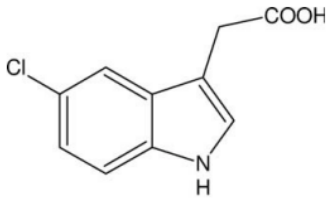


Figure 2.19 Anti- inflammatory indole derivatives drugs available in the market

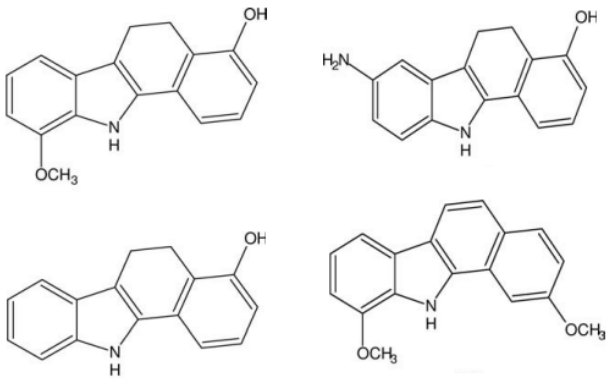
Anticancer Activity

Indole derivatives have been extensively studied for their anticancer activity. Many indole derivatives have been found to possess potent antiproliferative and cytotoxic effects against various cancer cell lines [90]. Some indole derivatives, such as indomethacin known as NSAID, have been used as anticancer drugs in clinical settings. Indole derivatives are believed to spend their anticancer activity by inducing cell cycle arrest, inhibiting angiogenesis, and inducing apoptosis in cancer cells.

Rossiter S. et al. [91], synthesized a series of halogenated indole-3-acetic acid derivatives as pro-drug and evaluated their potential targeted cancer therapies against V79 Chinese hamster lung fibroblasts. The study performed in 2002 and showed that one of the derivatives given below has remarkable results.

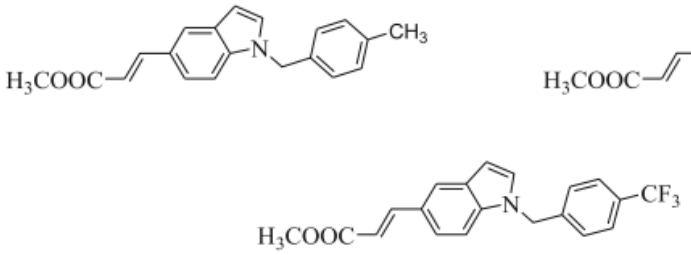


In 2006, Hong et al. [92] synthesized 63 tricyclic and tetracyclic indole derivatives and screened for their anticancer activities against human nasopharyngeal carcinoma (HONE-1) and gastric adenocarcinoma (NUGC-3) cell lines. The study group also performed a SAR (structure activity relationship) study. Four of the derivatives showed significant impact on the in-vitro anticancer activity.

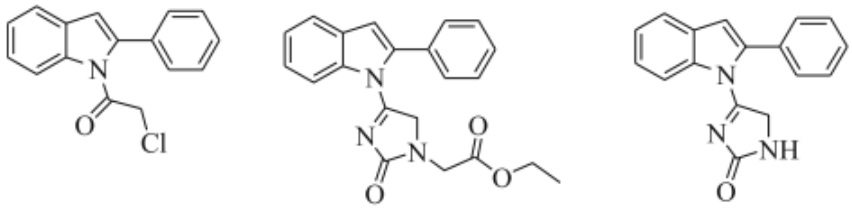


In 2014, Han K. et al. [93], studied on 5- (2-carboxyethenyl) indole derivatives and their anticancer activity on HT-29 and K562 cell lines and derivatives (E)-methyl 3-(1-tolyl-1H-indole-5-yl)acrylate, (E)-methyl 3-(1-benzyl-1H-indole-5-yl)acrylate, and (E)-methyl 3-

(1-(4-(tri-fluoromethyl)benzyl)-1H-indole-5-yl)acrylate found as significant in-vitro results as 4.67, 8.24 and 6.73 µM IC50 values respectively.



In 2019, Yousif M et al. [94], synthesized several 2-phenylindole derivatives ad screened for their colorectal carcinoma, liver carcinoma, prostate cancer, and breast adenocarcinoma anticancer activities. 2-chloro-1-(2-phenyl-1H-indole-1-yl) ethenone, 4-(2-phenyl-1H-indole-1-yl)-1H-imidazol-2(5H)-one and ethyl 2-(2-oxo-4-(2-phenyl-1H-indole-1-yl)-2H-imidazol-1(5H)-yl) acetate founded as significant in-vitro cytotoxic activity.



Some marketed anticancer drugs as indole derivatives presented in Figure 2.20 [28].

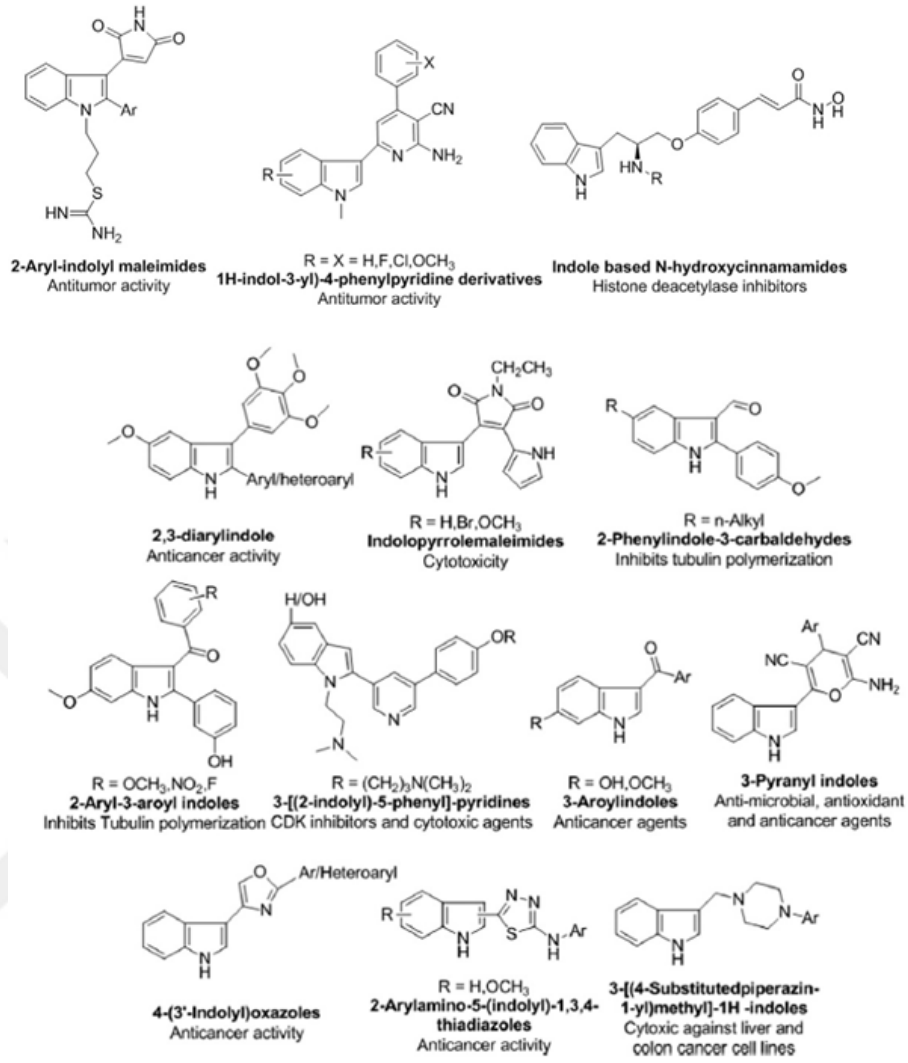


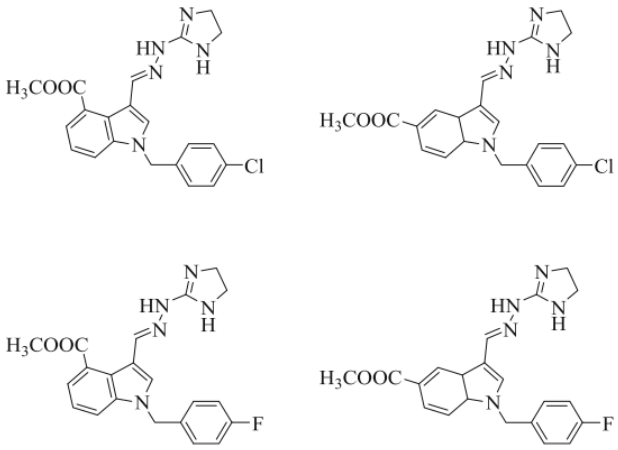
Figure 2.20 Anticancer indole derivatives drugs available in the market

Antimicrobial Activity

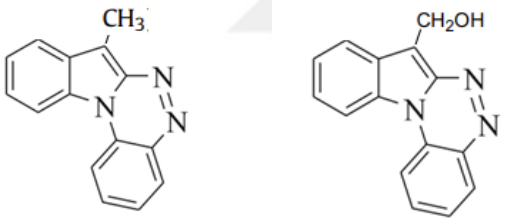
Indole derivatives have also been shown to possess antimicrobial activity against a wide range of pathogens, including bacteria, fungi, and viruses. Some indole derivatives, such as tryptophol and melatonin, have been shown to possess antiviral activity against influenza virus and hepatitis C virus [93,94]. Indole derivatives are believed to exert their antimicrobial activity by disrupting the membrane integrity of microorganisms, inhibiting the activity of enzymes involved in the biosynthesis of essential cellular components, and inducing oxidative stress in microorganisms.

Wei H et. Al [95], synthesized methyl (E)-1-(substituted benzyl)-3-((2-(4, 5-dihydro-1H-imidazol-2-yl) hydrazone) methyl)-1H-indole-5 carboxylate indole derivatives and

evaluated their antibacterial activity for bacterial pathogens in 2017. Some derivatives showed promised results on *Mycobacterium tuberculosis Acinetobacter baumanii*, and *Gram-negative bacteria*.



Hui X. et al., [96] synthesized indole[1,2-c]-1,2,4- butylnitrite benzotriazine derivatives via Sandmeyer reaction analysed their antifungal activities. Two of them given below showed significant activity compared with commercial Hymexazol.



Some marketed antimicrobial drugs as indole derivatives presented in Figure 2.21. [97].

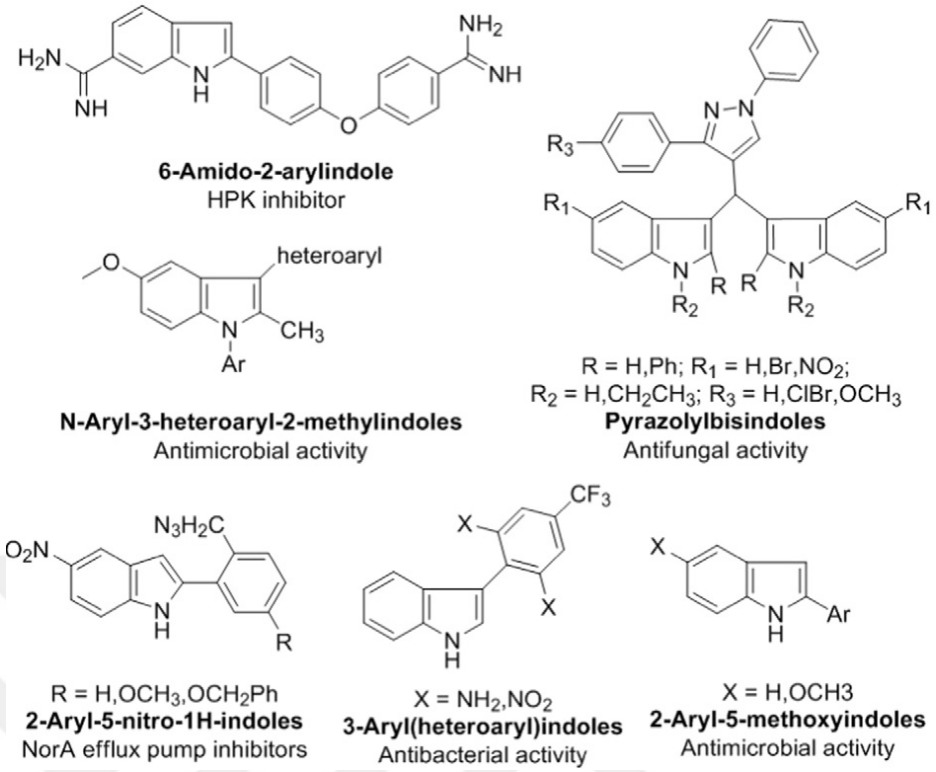


Figure 2.21 Antimicrobial indole derivatives drugs available in the market

In addition to anticancer, antimicrobial, anti-inflammatory, and analgesic effects indole and its derivatives have been shown to possess some CFTR modulators activities. Next section focuses on the CFTR proteins and their regulation.

2.4 Tezacaftor

Tezacaftor is an orally available CFTR potentiator class drug which was developed by Vertex Pharmaceuticals and approved in Europe and the US in 2018. It is used in combination with Ivacaftor. The name of the drug that contains this active pharmaceutical ingredient is Symdeko, which corresponds to a fixed-dose combination of 100 mg Tezacaftor and 150 mg Ivacaftor in a tablet pharmaceutical dosage form.

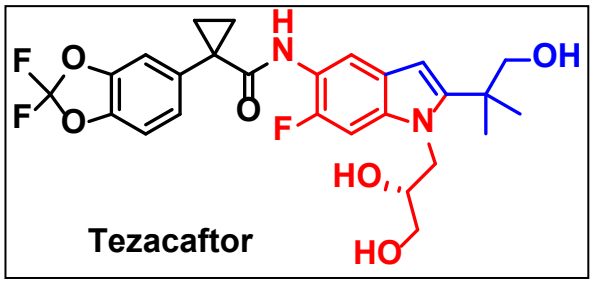


Figure 2.22 Chemical Structure of Tezacaftor ((1- (2,2-difluoro-1,3-benzodioxol-5-yl) -N- [1- [(2R) -2,3-dihydroxypropyl] -6-fluoro-2- (1-hydroxy-2-methylpropan-2- yl) indol-5-yl] cyclopropane-1-carboxamide))

The synthetic pathways of Tezacaftor were first developed by Vertex Pharmaceuticals as first-generation and second-generation (Figure 2.23, Figure 2.24). The first-generation synthesis pathway starts with the Sonogashira reaction of aryl bromide with alkyne in triethylamine as the solvent, resulting in alkyne.

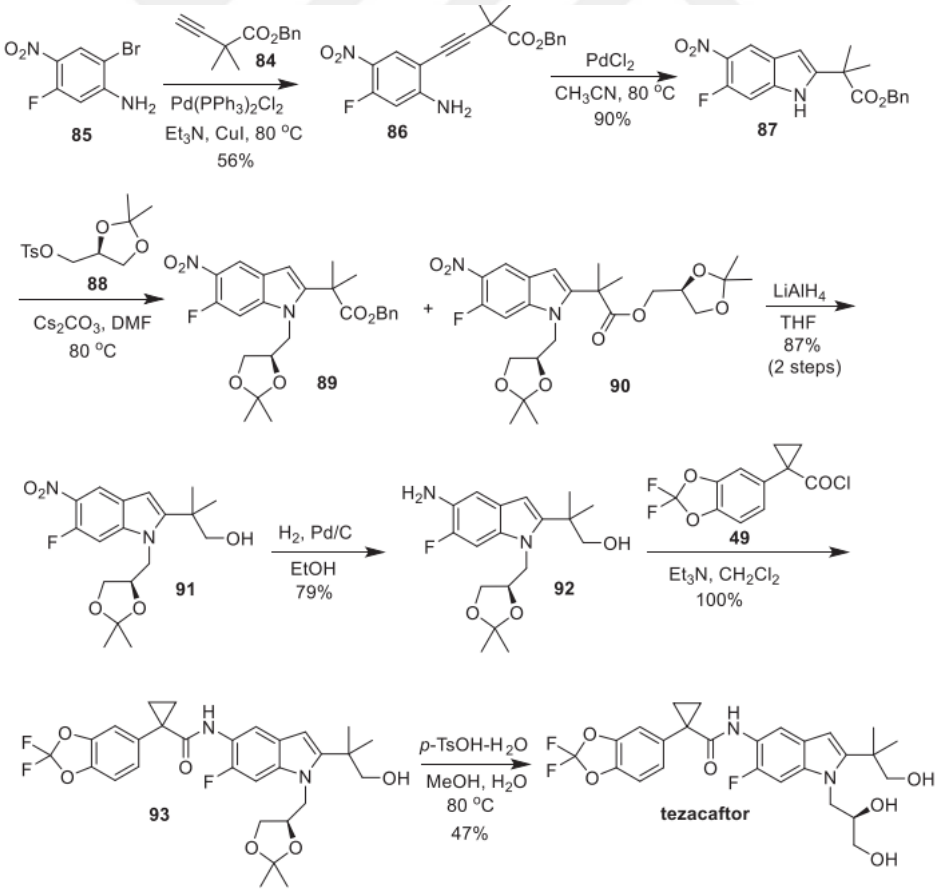


Figure 2.23 First-Generation pathway of Tezacaftor by Vertex Pharmaceuticals [98]

The second-generation pathway follows the similar pathway as the first-generation, with changes in preparation of the key fragments and step order. The synthesis consists of fifteen steps and starts with the bromination of 3-fluoro-4-nitroaniline using N-bromo succinimide in ethyl acetate.

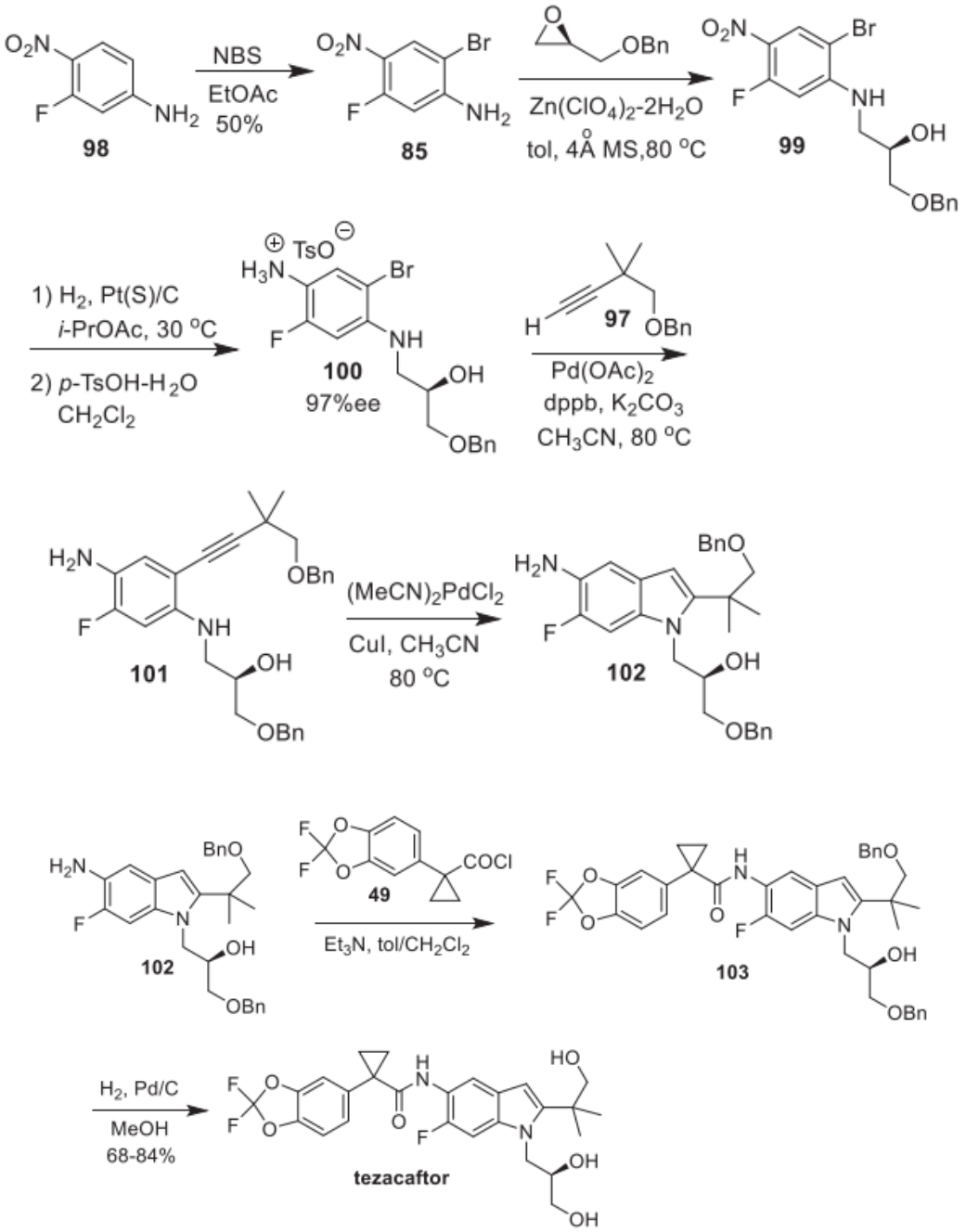


Figure 2.24 Second-generation pathway of Tezacaftor by Vertex Pharmaceuticals [98, 99]

2.5 QSAR and Indole Derivatives

The indole scaffold, defined as a prior structure, is one of the most important heterocyclic systems for pharmaceutics. Thus, numerous computational studies are performed for indole derivatives targeting several pharmacological activities such as antifungal, anti-inflammatory, anticancer etc.

In 2003, Chakraborti A.K, et al. [100], studied with some indole derivatives as phosphodiesterase (PDE) type IV inhibitors analysing by CoMFA (Comparative Molecular Field Analyses) and CoMSIA (Comparative Molecular Similarity Indices Analyses) methods (r^2=0.494 and 0.986 respectively). 7 indole derivatives are tested with 3D-QSAR modelling and good predictivity is founded.

In 2018, Ratchanok P., et al. [101], synthesized thirty four indole containing sulfonamide derivatives and evaluated their antiaromatase activities using QSAR modelling with MLR (Multiple Linear Regression) statistical tool and founded aromatase predicted inhibitory activity as IC50=0.8-15.3μM.

In 2020, Rosa P [102], et al. used already published compounds to address the design of the new molecules carrying indole and isatin pattern with accepting IC50s= 100 μM as ´active compounds´ and IC50s< 100 μM as ´inactive compounds´. The training set was populated by 45 compounds and test set with 14 compounds.

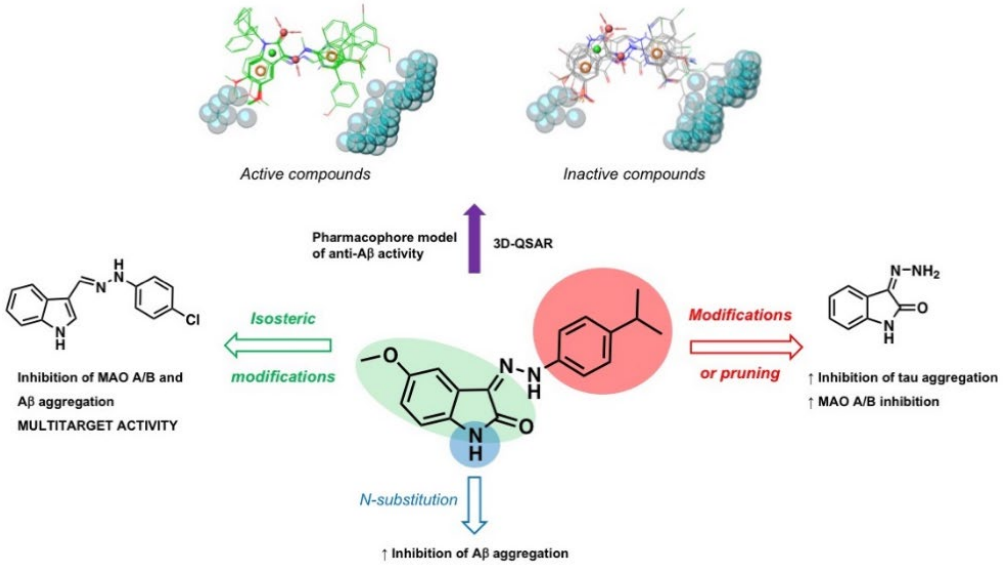


Figure 2.25 Graphical summary of the Indole and Isatin derivatives 3D-QSAR study

The gained 3D-QSAR modelling created with PLS regression (r^2=0.887, RMSE= 0.270) used to predict 17 newly synthesized compounds, most of them bearing the arylhydrazone moiety, and showed good correlation.

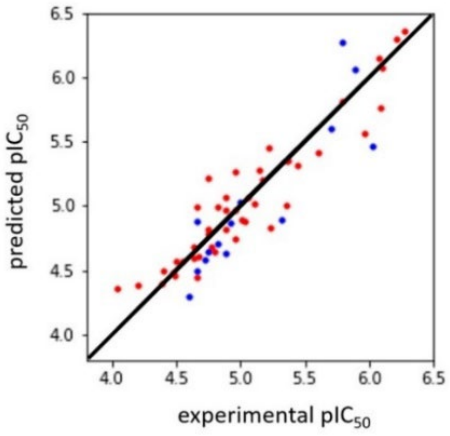


Figure 2.26 Experimental vs Predicted pEC50 values (blue circles show test set and red circles show training set)

One of the recent studies in 2023, Niousha S, et al. [103], performed a study to predict the inhibitory activity of 81 isatin and indole-based compounds against SARS Cov 3CLpro inhibitors using Monte Carlo tool and ChEMBL database, and developed four proposed models (r^2=0.81-0.92), pIC50 values are calculated. Two targeted compounds containing indole scaffold found highest drug-score as selected lead compounds.

The aim of this thesis consists of developing a QSAR modelling to predict the molecular activities of some Tezacaftor analogues and more precisely a series of 5-substituted 6-fluoro indole derivatives as CFTR modulators that could potentially show some anticancer activity. The synthesizes and evaluates their cytotoxic activities are also aimed. The derivatization of the 5-amino group of the indole heterocycle being the key modification for the generation of the molecules (Figure 2.27).

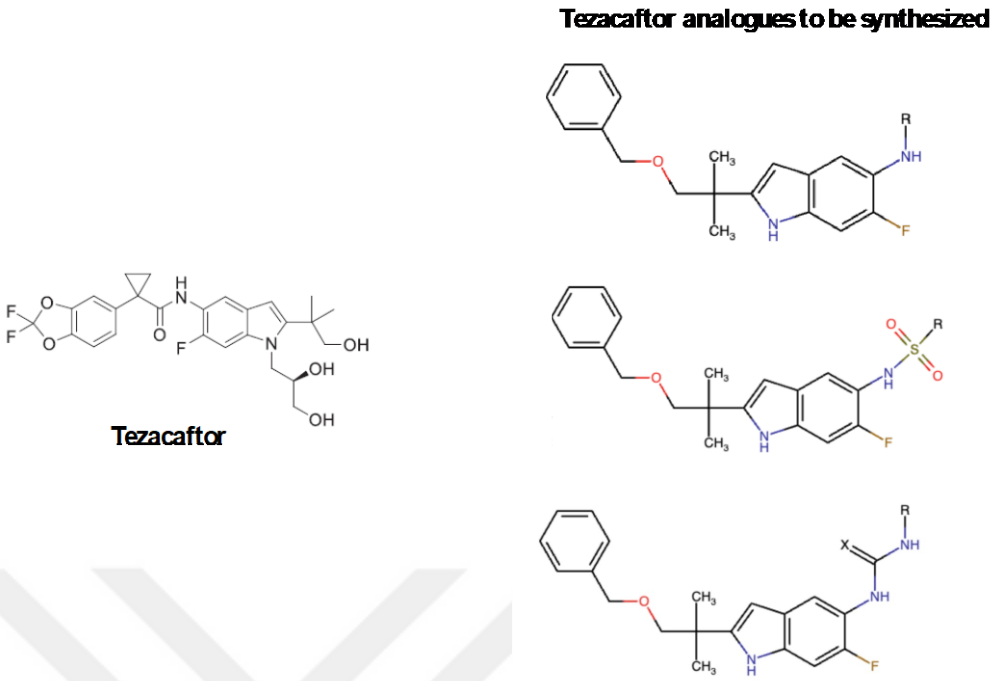


Figure 2.27 Structures of Tezacaftor and the molecules to be developed in this study

3. MATERIALS and METHODS

3.1 The QSAR Study

The details of the QSAR study are described stepwise and a flow chart that summarizes these steps is given in Figure 3.1.

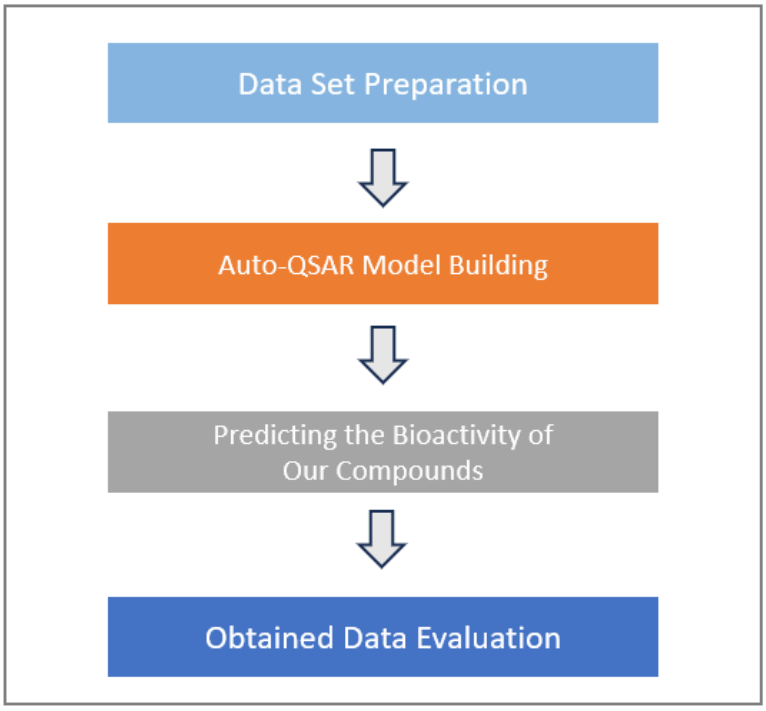


Figure 3.1 QSAR study flow chart following in this study

3.1.1 Data Set Preparation

UniProtID Determination

The F508del-CFTR protein number was searched *via* one of the biological databases [104] to find the unique identification code (UniProtID) of the CFTR protein: P13569.

Data Set Creation

The data set of active and inactive compounds for the P13569 protein was obtained *via* the Reaxys database [105]. The obtained results were filtered using single protein, wild type protein filters, as well as by eliminating molecules without pEC50 values to create the final data set.

3.1.2 Auto-QSAR Model Building

Auto-QSAR calculations were performed with the Schrödinger Suite 2017-1 release.

Model Building Process

Schrodinger database [106] was used to build an Auto-QSAR Model. Randomly, 50% of the model was trained, 40% tested and 10% validation set was created.

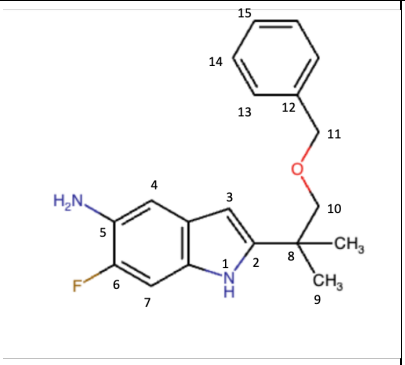
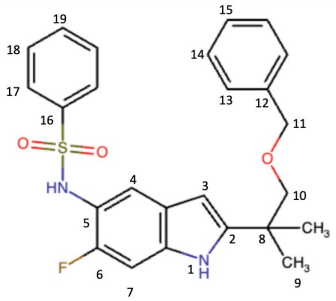
Model Evaluation

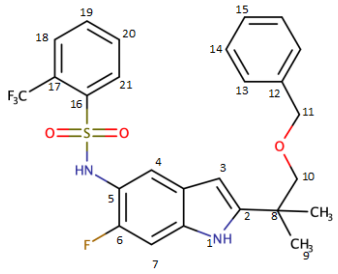
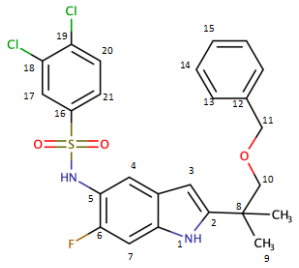
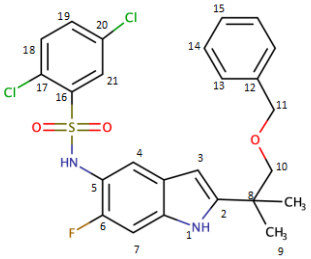
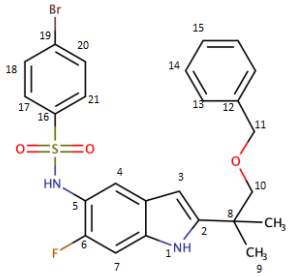
The aim of this step was to select the optimum model for our QSAR study. Different models using MLR, PLS, PCR, KPLS, Bayes, and RP machine learning methods were created via the Schrodinger database.

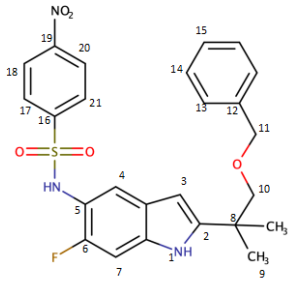
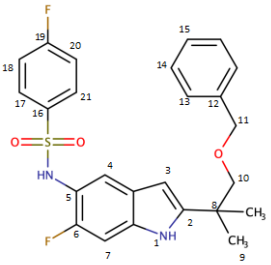
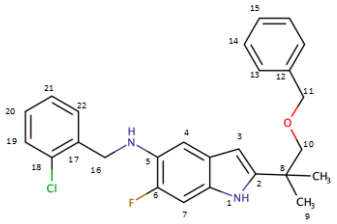
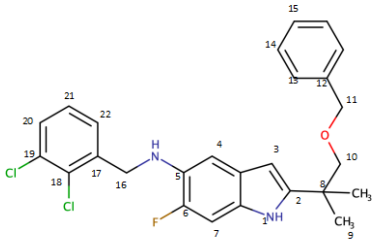
3.1.3 Predicting the Bioactivity of 5-substituted 6-fluoro indole derivatives

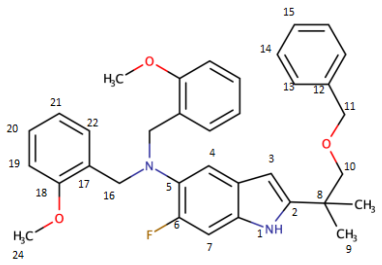
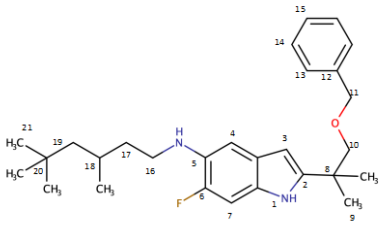
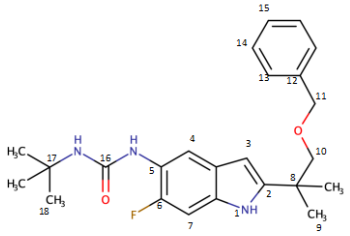
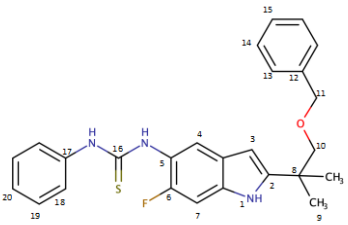
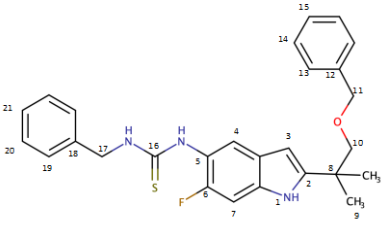
The SMILES (Simplified Molecular Input Line Entry Specification) format for the fifteen compounds (**1-15**), were generated using Marvin Pro, 2022 [107] software is given in Table 3.1.

Table 3.1 The chemical names, structures, and SMILES strings of **1-15**.

COMPOUND #	IUPAC NAME	STRUCTURE	SMILES
1	2-[1-(benzyloxy) -2-methylpropan-2-yl] -6-fluoro-1H-indol-5-amine		CC(C)(COCC1=CC=CC=C1)C1=CC2=C(N1)C=C(F)C(N)=C2
2	N-{2-[1-(benzyloxy) -2-methylpropan-2-yl] -6-fluoro-1H-indol-5yl}benzene sulfonamide		[H]N(C1=CC2=C(NC(=C2)C(C)(C)COCC2=CC=CC=C2)C=C1F)S(=O)(=O)C1=CC=CC=C1

COMPOUND #	IUPAC NAME	STRUCTURE	SMILES
3	N-{2-[1-(benzyloxy) -2-methylpropan-2-yl]-6-fluoro-1H-indol-5-yl}-2-(trifluoromethyl)benzene-1-sulfonamide		[H]N(C1=CC2=C(NC(=C2)C(C)(C)COCC2=CC=CC=C2)C=C1F)S(=O)(=O)C1=C(C=CC=C1)C(F)(F)F
4	N-{2-[1-(benzyloxy) -2-methylpropan-2-yl] -6-fluoro-1H-indol-5-yl} -3,4-dichlorobenzene-1-sulfonamide		[H]N(C1=CC2=C(NC(=C2)C(C)(C)COCC2=CC=CC=C2)C=C1F)S(=O)(=O)C1=C(C=CC=C1)C(F)(F)F
5	N-{2-[1-(benzyloxy) -2-methylpropan-2-yl] -6-fluoro-1H-indol-5-yl}-2,5-dichlorobenzene-1-sulfonamide		[H]N(C1=CC2=C(NC(=C2)C(C)(C)COCC2=CC=CC=C2)C=C1F)S(=O)(=O)C1=CC(Cl)=CC=C1Cl
6	N-{2-[1-(benzyloxy) -2-methylpropan-2-yl]-6-fluoro-1H-indol-5-yl}-4-bromobenzene-1-sulfonamide		[H]N(C1=CC2=C(NC(=C2)C(C)(C)COCC2=CC=CC=C2)C=C1F)S(=O)(=O)C1=CC=C(Br)C=C1

COMPOUND #	IUPAC NAME	STRUCTURE	SMILES
7	N-{2-[1-(benzyloxy) -2-methylpropan-2-yl]-6-fluoro-1H-indol-5-yl}-4-nitrobenzene-1-sulfonamide		[H]N(C1=CC2=C(NC(=C2)C(C)(C)COCC2=CC=CC=C2)C=C1F)S(=O)(=O)C1=CC=C(C=C1)[N+]([O-])=O
8	N-{2-[1-(benzyloxy) -2-methylpropan-2-yl]-6-fluoro-1H-indol-5-yl}-4-florobenzene-1-sulfonamide		[H]N(C1=CC2=C(NC(=C2)C(C)(C)COCC2=CC=CC=C2)C=C1F)S(=O)(=O)C1=CC=C(F)C=C1
9	2-[1-(benzyloxy) -2-methylpropan-2-yl] -N-[(2-chlorophenyl) methyl] -6-fluoro-1H-indol-5-amine		CC(C)(COCC1=CC=CC=C1)C1=CC2=C(N1)C=C(F)C(NCC1=C(Cl)C=CC=C1)=C2
10	2-[1-(benzyloxy) -2-methylpropan-2-yl] -N-[(2,3-dichlorophenyl) methyl] -6-fluoro-1H-indol-5-amine		CC(C)(COCC1=CC=CC=C1)C1=CC2=C(N1)C=C(F)C(NCC1=C(Cl)C(Cl)=CC=C1)=C2

COMPOUND #	IUPAC NAME	STRUCTURE	SMILES
11	2-[1-(benzyloxy) -2-methylpropan-2-yl] -6-fluoro-N, N-bis[(2-methoxyphenyl) methyl] -1H-indol-5-amine		COC1=C(CN(CC2=C(OC)C=CC=C2)C2=CC3=C(NC(=C3)C(C)(C)COCC3=CC=CC=C3)C=C2F)C=CC=C1
12	2-[1-(benzyloxy) -2-methylpropan-2-yl] -6-fluoro-N-(3,5,5-trimethylhexyl) -1H-indol-5-amine		CC(CCNC1=CC2=C(NC(=C2)C(C)(C)COCC2=CC=CC=C2)C=C1F)CC(C)(C)C
13	1-{2-[1-(benzyloxy) -2-methylpropan-2-yl] -6-fluoro-1H-indol-5-yl} -3-tert-butylurea		CC(C)(C)NC(=O)NC1=CC2=C(NC(=C2)C(C)(C)COCC2=CC=CC=C2)C=C1F
14	3-{2-[1-(benzyloxy) -2-methylpropan-2-yl] -6-fluoro-1H-indol-5-yl} -1-phenylthiourea		CC(C)(COCC1=CC=CC=C1)C1=CC2=C(N1)C=C(F)C(NC(=S)NC1=CC=CC=C1)=C2
15	1-benzyl-3-{2-[1-(benzyloxy) -2-methylpropan-2-yl] -6-fluoro-1H-indol-5-yl} thiourea		CC(C)(COCC1=CC=CC=C1)C1=CC2=C(N1)C=C(F)C(NC(=S)NCC1=CC=CC=C1)=C2

3.2 Synthesis of 5-Substituted 6-Fluoro Indole Derivatives

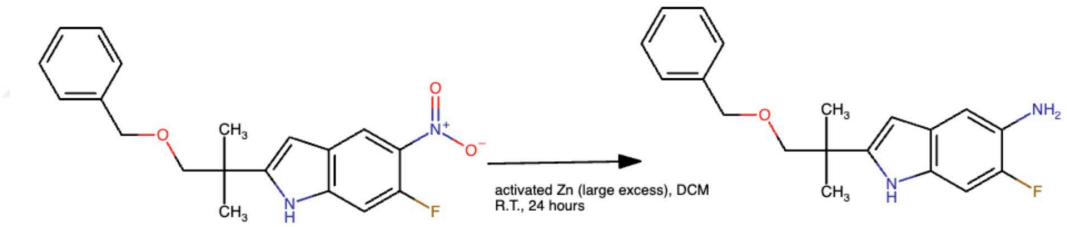
3.2.1 Chemicals

Sulfonyl chlorides, aldehydes, isocyanates and isothiocyanates used in this study were purchased from Sigma Aldrich along with ethyl acetate, n-hexane, DMSO, triethylamine, sodium sulfate and sodium chloride. Zinc was purchased from Fisons. Ethanol and dichloromethane were purchased from Isolab. TLC plates (Aluminum Sheet 20×20 cm Silica gel 60 F254) and silica were purchased from Merck. 2-(1-(benzyloxy)-2-methylpropan-2-yl)-6-fluoro-5-nitro-1H-indoIe (Teza A00) and pic-BH_3 were provided from Chemo Iberica.

3.2.2 Methods of Synthesis

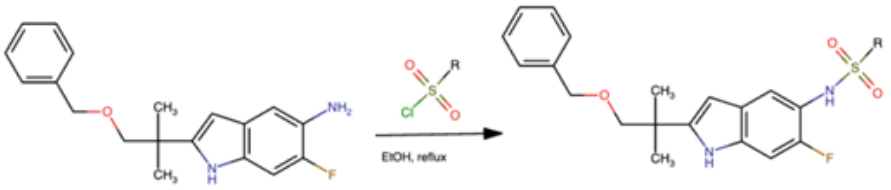
3.2.2.1 Nitro Reduction: synthesis of compound 1

Activation of Zinc Dust: 200 g Zn is taken in a flask with 75 mL of 10% HCI (aq), stirred for a while, filtered washed with 150 mL water followed by 50 mL acetone, dried in a vacuum oven.



1 equivalent Teza A00 and 50 equivalent activated zinc dust dissolved in DCM (10 mL). The reaction stirred at room temperature for 1 day. Acetic acid is added dropwise and stirred for 1 hour. Completion of reaction checked by TLC analyses (n-hexane: ethyl acetate 1:1 (v:v)). DCM is evaporated. The celite washed with ethyl acetate, the filtrate neutralized with $NaHCO_3$ (aq) solution (0.2 g/ mL), organic phase dried on $MgSO_4$ (s), filtered as brown oil, and purified by silica column [109].

3.2.2.2 General Procedure for the Synthesis of Sulfonamide Derivatives (2-8)



One equivalent of compound **1** and 2 equivalents of the corresponding sulfonyl chloride were refluxed in ethanol (5 mL) for 2 hours before being stirred for one week at room temperature. Completion of reaction checked by TLC analyses (n-hexane: Ethyl acetate 1:1 (v:v)). After completion of the reaction, the solution evaporated, and the crude was dissolved in DMSO, and product was precipitated by water and isolated with filtration. Obtained products were further purified on silica gel using n-hexane/ethyl acetate (1/1, v: v) mobile phase solvent system [110].

3.2.3 Analytical Methods

3.2.3.1 Controls By Thin Layer Chromatography

Material

Plates: Aluminum Sheet 20×20 cm Silica gel plaques 60 F254

Solvents: Three different solvents systems were prepared to be used as mobile phases for chromatographic controls of compounds.

S1: n-hexane: Ethyl acetate 1:1 (v:v), T:25°C

S2: Dichloromethane: Methanol 99:1 (v:v), T:25°C

S3: n-hexane: Ethyl acetate 1:1 (v:v), T:25°C

Method: Solvent systems were poured into chambers and kept for approximately 1 hour for saturation. Reactions were monitored with Thin Layer Chromatograph (TLC) after dissolving the synthesized compounds and starting materials with suitable solvents and application of them with Pasteur pipettes onto silica gel plates. The plates were dragged for 10 cm at room temperature. Retention Factor (Rf) values of compounds were calculated. Stains of synthesized compounds and their starting materials were determined by UV light (254/365 nm).

3.2.3.2 Purification By Column Chromatography

Material

Stationary phase: Silica gel 60 mesh

Mobile phase: Two different solvent systems were prepared to be used in chromatographic controls of compounds.

Mobile phase system 1: Dichloromethane (100%)

Mobile phase system 2: n-hexane: Ethyl acetate (1:1) (v:v)

Method: The column was filled by a wet loading method. Elution was controlled with TLC using selected suitable solvents as mobile phase systems.

3.2.3.3 Spectral Analysis

^{1}H-NMR Spectra

^{1}H-NMR spectra were recorded with a Bruker BioSpin GmbH instrument, at 500 MHz resonance frequency, using DMSO-d_6 as solvent, 298 K temperature.

^{13}C APT (Attached Proton Test)-NMR Spectra

APT-NMR spectra were recorded with a Bruker BioSpin GmbH instrument, at 125.76 MHz resonance frequency, using DMSO-d_6 as solvent, 298 K temperature.

COSY (Correlated Spectroscopy)-NMR Spectra

COSY-NMR spectra were recorded with a Bruker BioSpin GmbH instrument, at 500 MHz resonance frequency, using DMSO-d_6 as solvent, 298 K temperature, (1H, 1H) nucleus, (2048, 512) spectral size.

HSQC (Heteronuclear Single Quantum Coherence)-NMR Spectra

HSQC-NMR spectra were recorded with a Bruker BioSpin GmbH instrument, at 500 MHz resonance frequency, using DMSO-d_6 as solvent, 298 K temperature, (^{1}H, ^{13}C) nucleus, (1024, 1024) spectral size.

HPLC-UV-MS Analyses (High Performance Liquid Chromatography-Ultraviolet - Mass Spectrum Analyses)

The analyses of compound were recorded with LC/MS Bruker Compact ESI with acquisition mode: Full scan-MS (ESI+). The samples were prepared in CD3CN dilute 1:100 with ACN (injection volume 5µL).

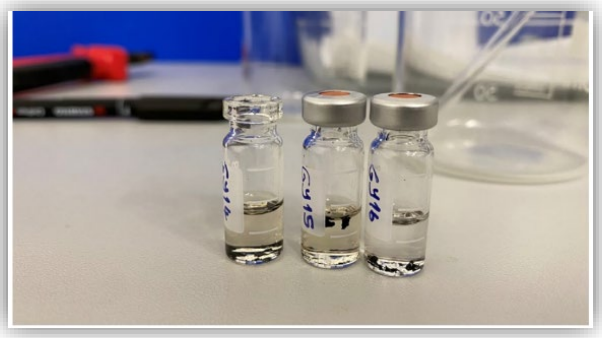


Figure 3.2 A visual of analytical samples of some synthesized compounds

3.3 Biological Activity Test

Cytotoxicity Screening With NCI-Sulforhodamine B Assay Method

Sulforhodamine B (SRB) assay is a commonly used method for in vitro cytotoxicity screening developed by Skehan et al. in 1990. SRB assay can be utilized for cytotoxicity testing, and it is a bright pink aminoxanthene dye bearing two sulphonic groups that bind stoichiometrically to basic amino acid residues under mild acidic conditions and dissociate under basic conditions. Colorimetric measurement provides an estimate of total protein mass, which is related to cell number. The greater the number of cells, the greater the amount of dye is taken up, then after fixing with trichloroacetic acid, as the cells are lysed, the released dye will give a more intense color and greater absorbance.

OVSAHA (Human High-Grade Serous Ovarian Cancer Cell Line) and OVCAR-3 (Human Ovarian Cancer Cell Line) as colon cancer cell lines were inoculated into 96-well plates as 3000 cells/well. After 24 hours incubation, cells were treated with the compounds in increasing concentrations (2.5 µM to 40 µM). Each drug treatment was performed in triplicate. Dimethyl sulfoxide was used as negative control. After 72 hours of incubation time, medium was discarded, and plates were washed twice with 1xPBS. Then the cells were fixed with 10% (w/v) trichloroacetic acid solution for 1 hour in dark condition at + 4ºC. To remove trichloroacetic acid solution, cells were then washed with ddH_2O about four to five times and left air-dry at room temperature. The plates were then stained using 0.4% sulforhodamine B (SRB) solution in acetic acid solution (1 %) and incubated in dark at room temperature for 10 minutes. Finally, excess dye was discarded by washing off by acetic acid solution (1%) four to five times until no dye comes out and left air dry at room temperature.

4. EXPERIMENTAL DATA

4.1 The Auto-QSAR Study

4.1.1 Data Set Preparation Step

UniProtID Determination

The search for the F508del CFTR protein resulted in a list where proteins of a few species commonly encountered, such as HUMAN, BOVIN, ECOLI, MOUSE, etc. appeared. The P13569 protein was the one that was selected, as this code was the one dedicated to human organisms.

Data Set Creation

UniProtID P13569 was searched *via* the Reaxys database [105] to build a structure activity dataset. The obtained results were filtered with single protein and wild type proteins. First an initial data set comprising 5404 compounds was obtained, containing both active and inactive structures. Then, the dataset was filtered to eliminate compounds for which the pEC50 data was missing to result in 3801 compounds. Finally, the selected data set, which contains 1735 compounds, was created.

4.1.2 Auto-QSAR Model Building Step

Model Building Process

The pEC50 values of dataset gained between 3.5 to 10.2. The median value of pEC50 was founded as 5.13, therefore, below the median was accepted as inactive compounds, and above the median was accepted as active compounds (Table 4.1).

Table 4.1 The distribution of datasets into training and test sets as Inactive and Active in the Auto-QSAR Model

	Active	**Inactive**	**Total**	**Active / Inactive**
Training set	793	768	1561	1.03
Test set	86	88	174	0.98
Total	879	856	1735	

Model Evaluation

The results of different models presented in Table 4.2 were obtained via the Schrodinger database.

Table 4.2 The models properties obtained in this Auto-QSAR study

Model Code	Score	SD	R^2	RMSE	Q^2
kpls_dendritic_10	0.5914	0.2528	0.8865	0.3985	0.7142
kpls_radial_10	0.5808	0.2890	0.8526	0.4139	0.6917
kpls_linear_10	0.5691	0.2376	0.8998	0.4044	0.7057
kpls_dendritic_8	0.5640	0.2412	0.8989	0.4070	0.6956
kpls_dendritic_31	0.5606	0.2630	0.8799	0.4145	0.6836
kpls_linear_4	0.5566	0.2475	0.8912	0.4128	0.6938
kpls_dendritic_4	0.5535	0.2594	0.8805	0.4175	0.6867
kpls_linear_8	0.5518	0.2472	0.8937	0.4142	0.6848
kpls_dendritic_44	0.5518	0.2692	0.8740	0.4202	0.6747
kpls_linear_31	0.5455	0.2505	0.8910	0.4178	0.6785

The dendritic model (kpls_dendritic_10) was found to give the best results, as for this method, the R^2 and Q^2 values were found to be greater than 0.6, the RMSE value was between 0.4 and 0.5 as ideally, the SD value was the lowest, and the Score value was the highest when compared to the other models.

4.1.3 Predicting the Bioactivity of Our Compounds

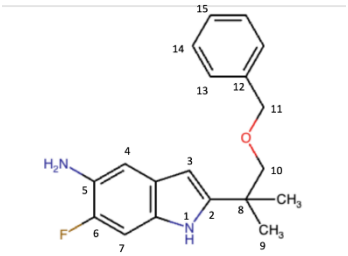
After running the ′kpls_dendritic_10′ Auto-QSAR model, predicted pEC50 values were obtained for the fifteen 5-substituted 6-fluoro indole derivatives given in Table 4.3.

Table 4.3 pEC50 values of fifteen 5-substituted 6-fluoro indole derivatives

Compound #	**pEC50 Value**
1	4.92
2	4.83
3	4.84
4	5.06
5	5.00
6	4.96
7	4.90
8	4.85
9	4.98
10	5.11
11	5.10
12	4.79
13	4.80
14	4.87
15	4.89

4.2 Experimental Data for Synthesized Molecules

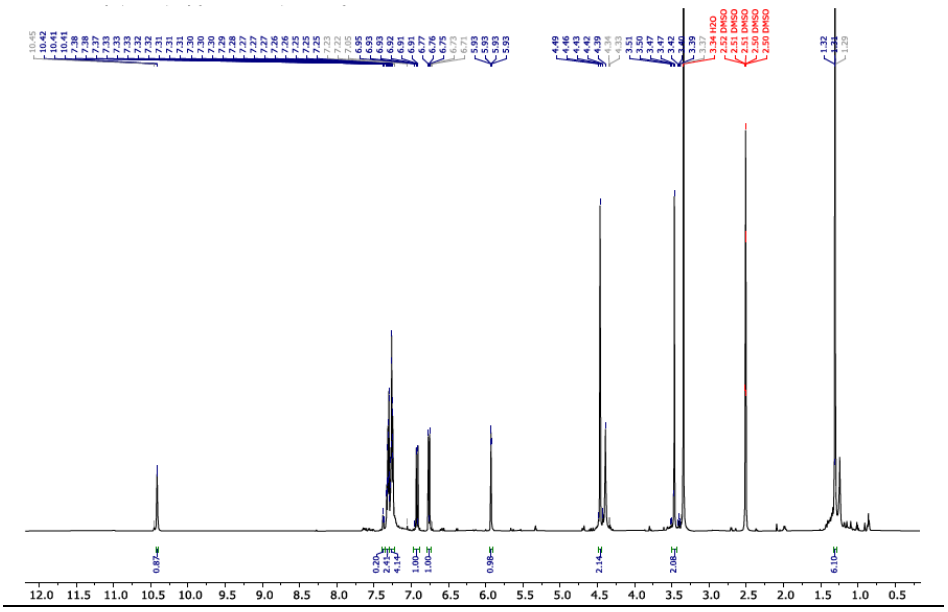
2-[1-(benzyloxy) -2-methylpropan-2-yl] -6-fluoro-1H-indol-5-amine (1)



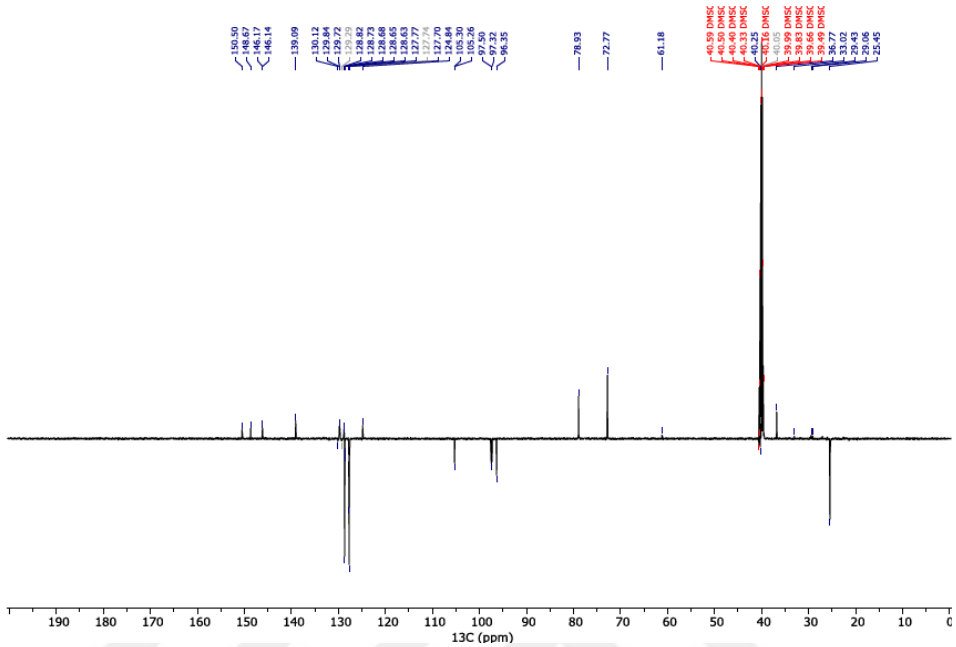
Teza A00 (0.45 mmol) and activated zinc (22.5 mmol) were reacted in 10 mL DCM as described in the general method at 3.2.2.1. The form of compound is brown oil.

Yield (%)	:	98
Physical appearance	:	Brown oily solid
Solubility	:	Acetone, DMSO, slightly soluble in n-hexane
Molecular formula	:	$C_{19}H_{21}FN_2O$
Molecular weight (g/mol)	:	312.1638

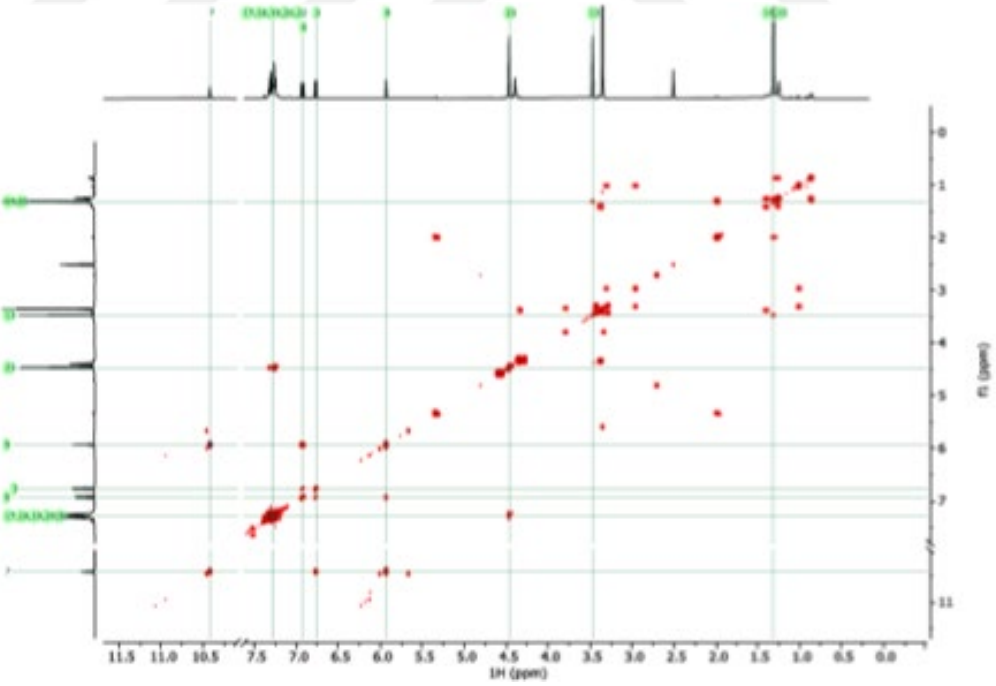
^{1}H NMR (500 MHz, DMSO-d_6) δ (ppm) : 10.42 (s, 1H, H-1), 7.38-7.25 (m, 5H, H-13+H-14+H-15), 6.93 (dd, J_1 = 11.7Hz, J_2=0.8 Hz, 1H, H-7), 6.77 (d, J =8.9 Hz, 1H, H-4), 5.93 (dd, J_1 = 2.3, J_2=0.8 Hz, 1H, H-3), 4.46 (s, 2H, H-11), 3.47 (s, 2H, H-10), 1.31 (s, 6H,H-9).



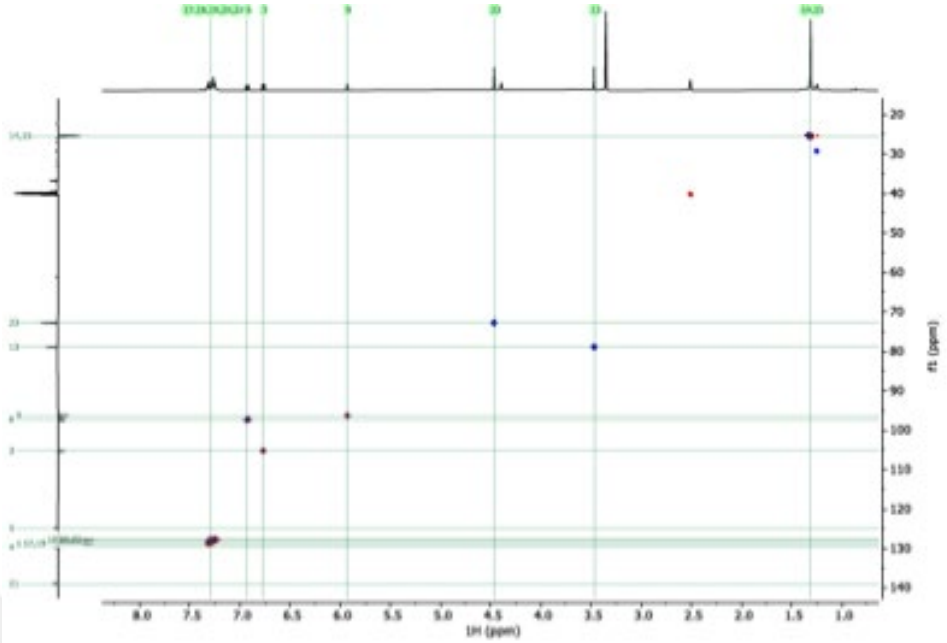
^{13}C APT NMR (126 MHz, DMSO-d_6) δ (ppm) : 150.51 (d, ^{1}J=230 Hz, C-6), 148.67 (C-2), 139.09 (C-12), 129.84, 129.72, 128.74, 128.71, 128.68, 127.76, 127.70, 124.85, 105.27 (d, ^{3}J=5.0 Hz, C-4), 97.4 (d, ^{2}J=22.7 Hz, C-7), 96.35 (C-3), 78.93 (C-11), 72.77 (C-10), 36.7 (C-8), 25.45 (C-9).



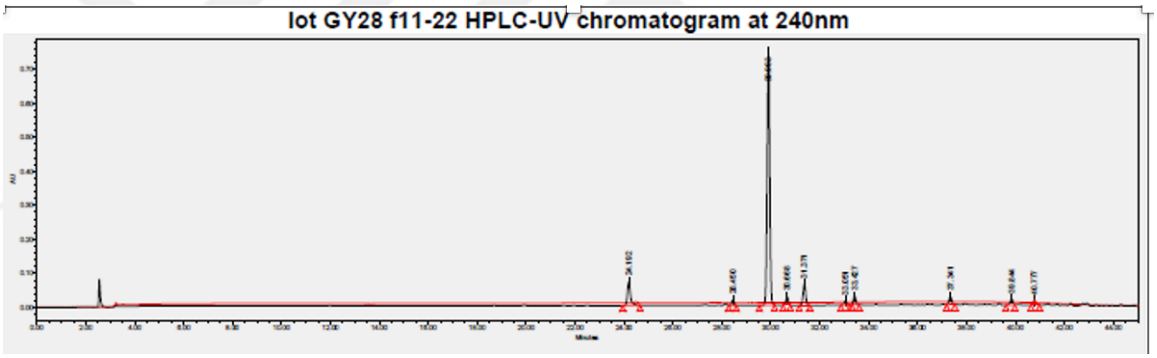
COSY :

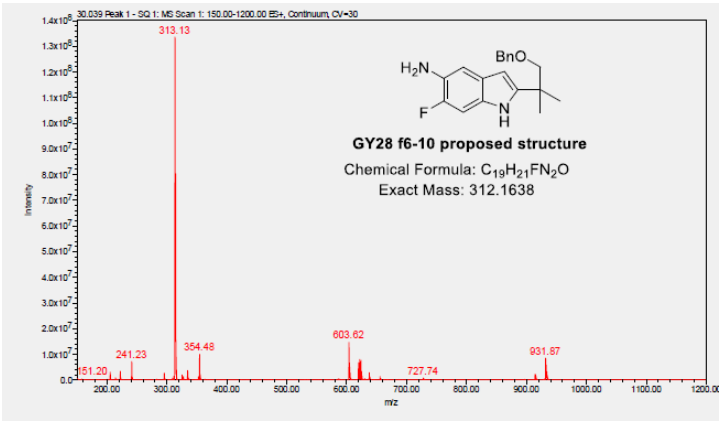


HSQC :

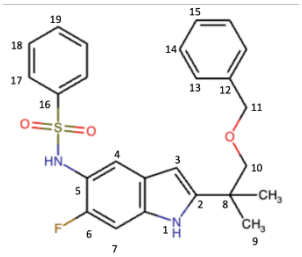


LC-MS : RT=29.9 min., m/z, 313 $[M+H]^+$





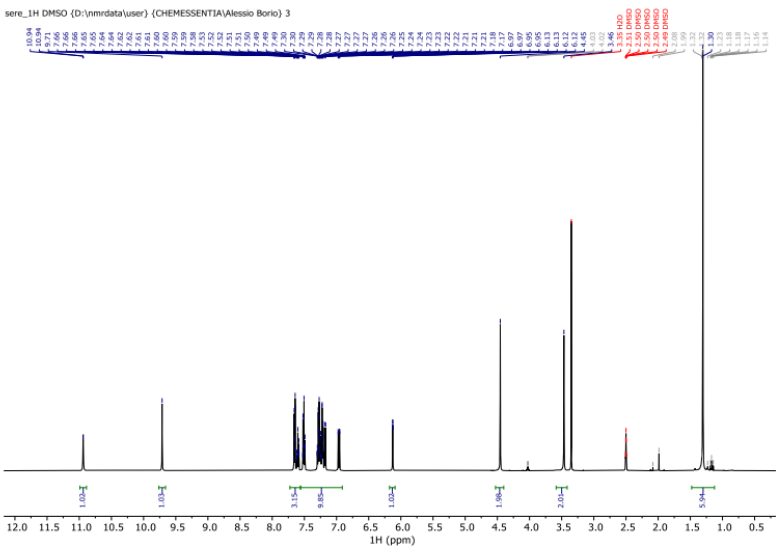
N-{2-[1-(benzyloxy)-2-methylpropan-2-yl]-6-fluoro-1H-indol-5yl}benzenesulfonamide (2)



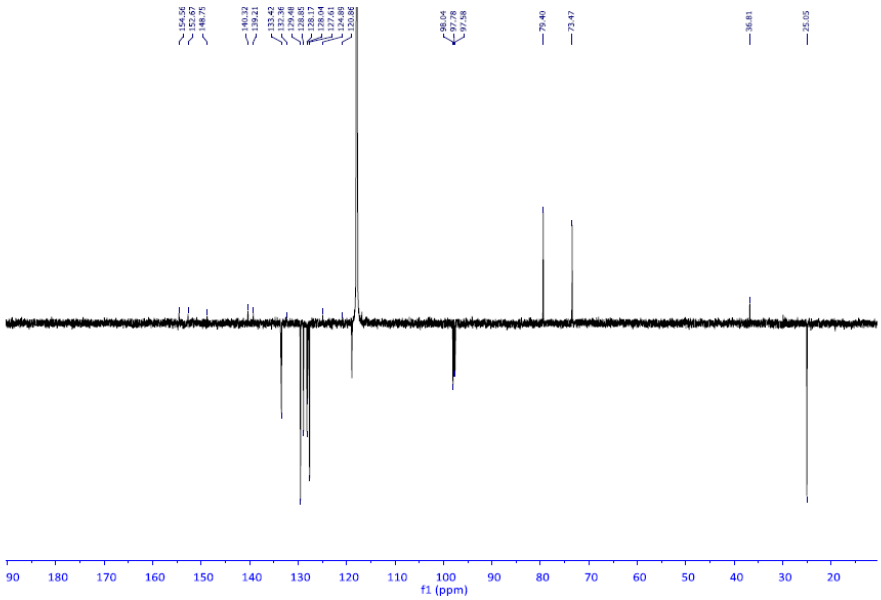
Compound 1 (0.50 mmol) and benzene sulfonyl chloride (1.0 mmol) were reacted in 5 mL ethanol as described in the general method at 3.2.2.2. The form of the compound is brown solid.

Yield (%)	:	98
Physical appearance	:	Brown oily solid
Solubility	:	Slightly soluble in DMSO, practically insoluble in ethyl acetate, n-hexane, and acetone.
Molecular formula	:	C25H25FN2O3S
Molecular weight (g/mol)	:	452.16

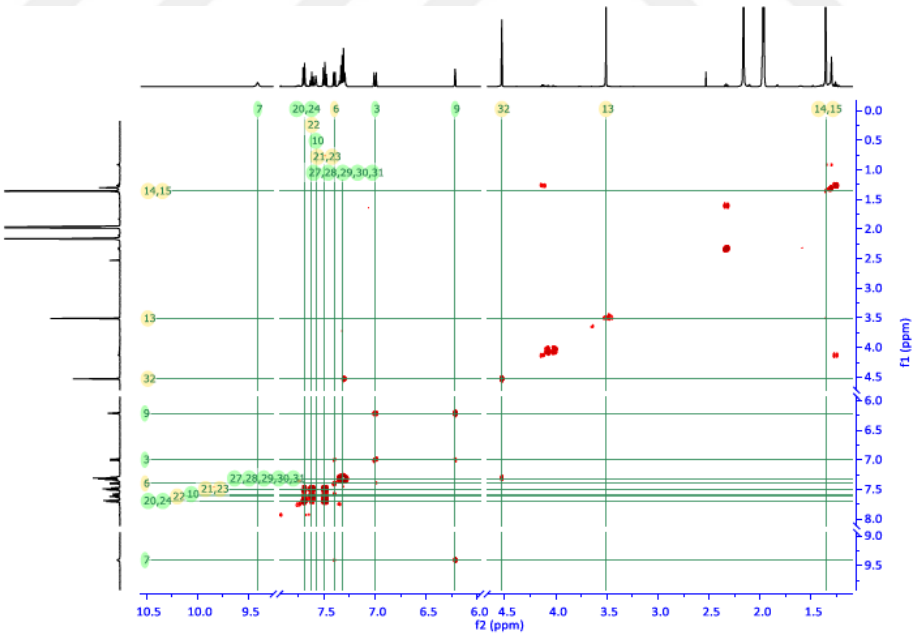
^{1}H NMR (500 MHz, DMSO-d_6) δ (ppm) : 10.94 (s, 1H, NH), 9.41 (s, 1H, H-1), 7.66-7.58 (m, 5H, H-13+H-14+H-15), 7.40 (d, J = 7.6 Hz, 1H, H-4), 7.32-7.17 (m, 5H, H-17+H-18+H-19), 7.00 (dd, J_1 = 10.8Hz, J_2=0.8 Hz, 1H, H-7), 6.12 (dd, J_1 = 2.3, J_2=0.9 Hz, 1H, H-3), 4.53 (s, 2H, H-11), 3.51 (s, 2H, H-10), 1.36 (s, 6H,H-9).



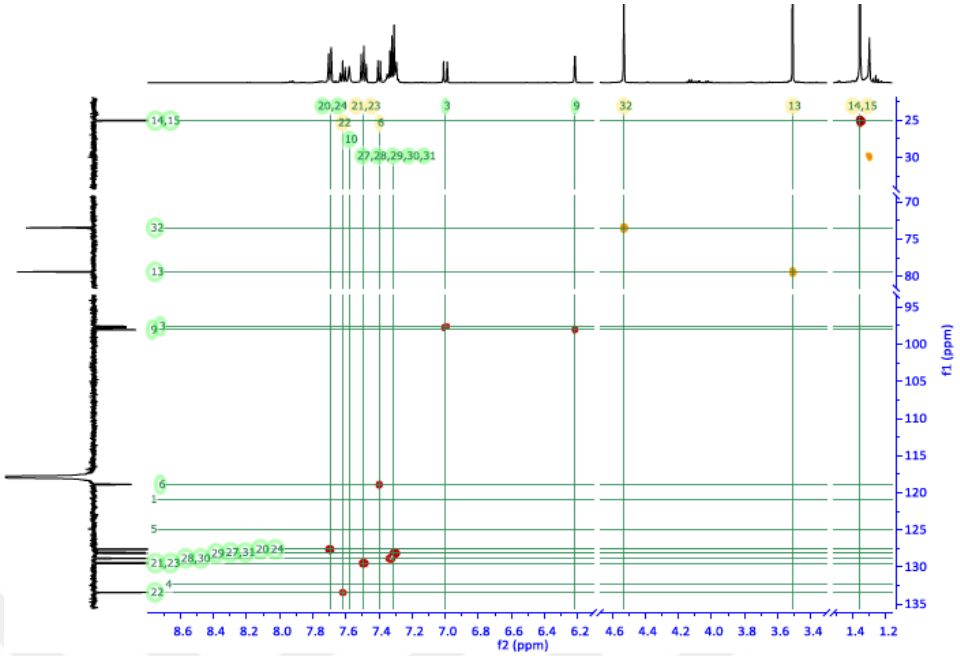
^{13}C APT NMR (126 MHz, DMSO-d_6) δ (ppm) : 153.61 (d, ^{1}J=230 Hz, C-6), 148.75 (C-2), 140.32 (C-12), 139.21 (C-X), 133.42, 132.36, 129.48, 128.85, 128.17, 128.04, 127.61, 124.89, 120.86 (C-4), 98.04 (C-3), 97.68 (d, ^{2}J=24.9 Hz, C-7), 79.40 (C-11), 73.47 (C-10), 36.81 (C-8), 25.05 (C-9).



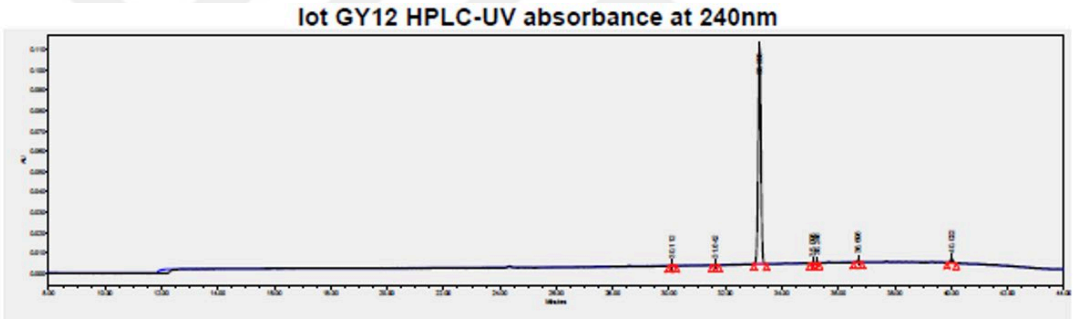
COSY :

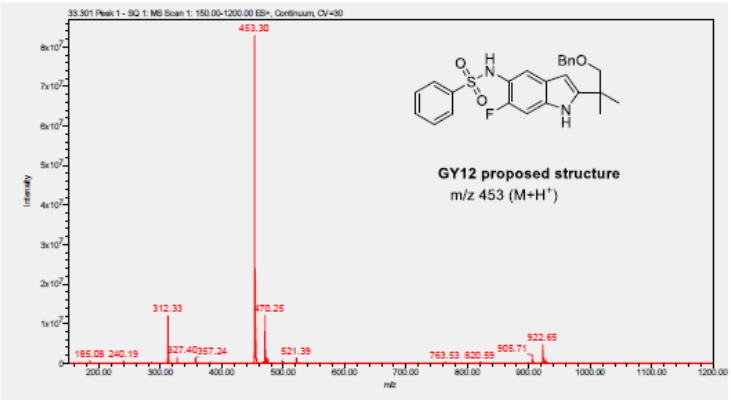


HSQC :

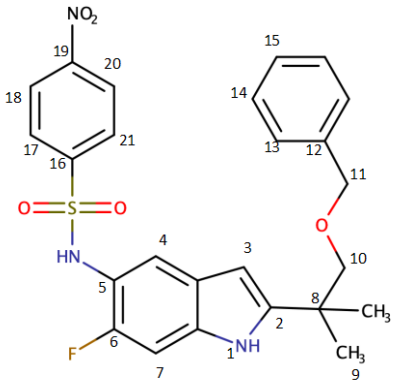


LC-MS : RT=33.2 min., m/z, 453 $[M+H]^+$





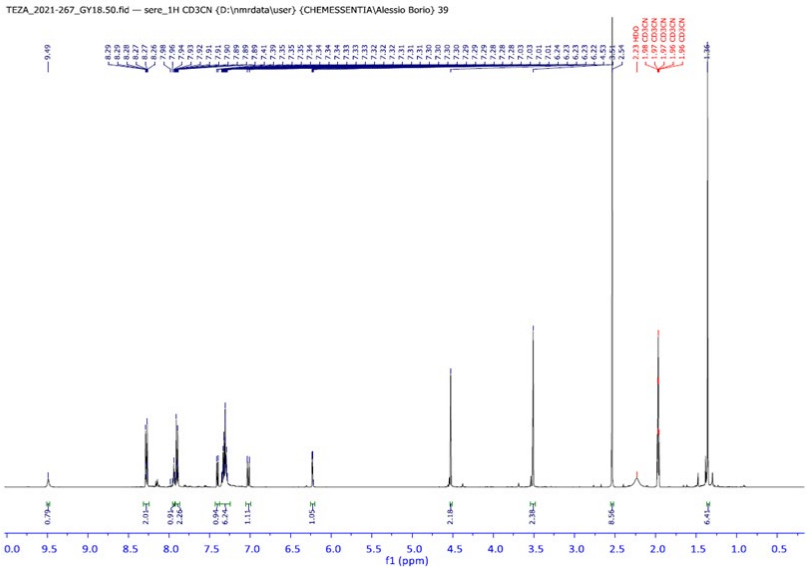
N-{2-[1-(benzyloxy)-2-methylpropan-2-yl]-6-fluoro-1H-indol-5-yl}-4-nitrobenzene-1-sulfonamide (7)



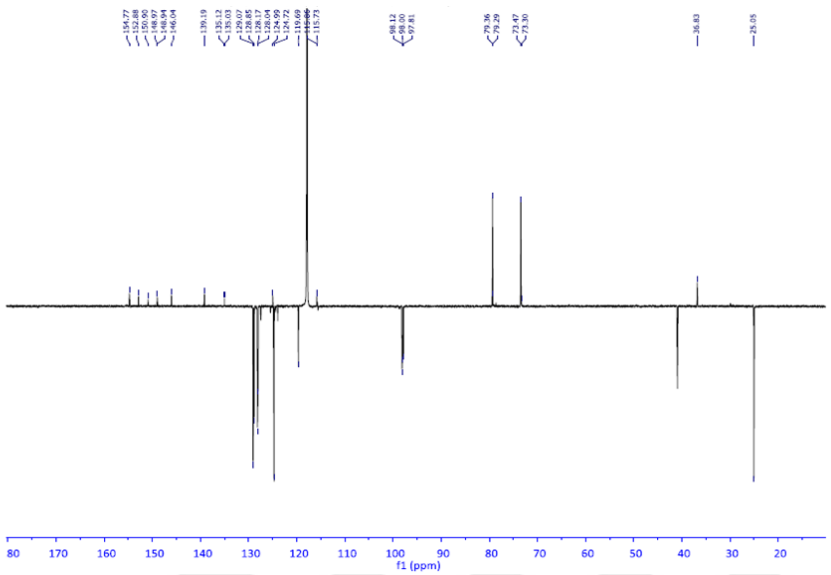
Compound 1 (0.50 mmol) and 4-nitrobenzene sulfonyl chloride (1.0 mmol) were reacted in 5 mL ethanol as described in the general method at 3.2.2.2. The form of the compound is brown solid.

Yield (%)	:	70
Physical appearance	:	Brown oily solid
Solubility	:	DMSO
Molecular formula	:	C25H24FN3O5S
Molecular weight (g/mol)	:	497.54

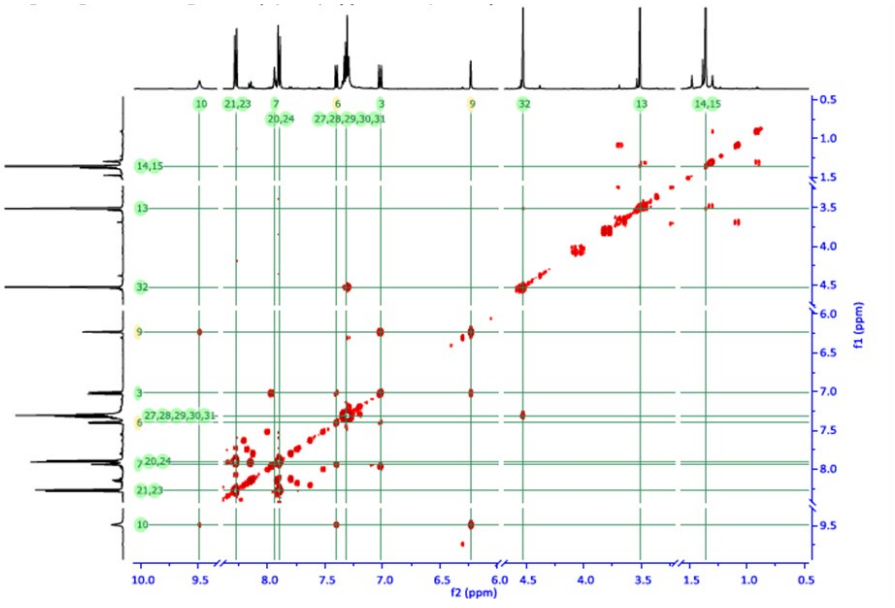
^{1}H NMR (500 MHz, DMSO-d_6) δ (ppm) : 9.49 (s, 1H, H-1), 8.29-8.26 (m, 2H, H-17+H-21), 7.98-7.89 (m, 3H, H-arom), 7.40 (d, J= 7.6Hz, 1H, H-4), 7.35-7.28 (m, 5H, H-arom), 7.02 (dd, J_1= 10.8Hz, J_2= 0.9 Hz, 1H, H-7), 6.23 (dd, J_1= 2.3Hz, J_2= 0.8 Hz, 1H, H-3), 4.53 (s, 2H, H-11), 3.51 (s, 2H, H-10), 1.36 (s, 6H,H-9).



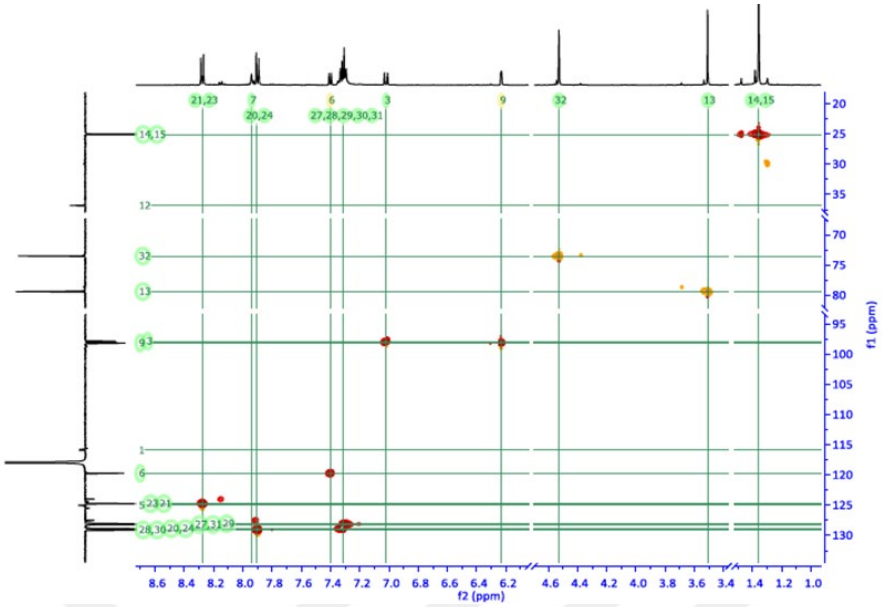
^{13}C APT NMR (126 MHz, DMSO-d_6) δ (ppm) : 153.83 (d, ^{1}J=237.8 Hz, C-6), 150.90 (C-19), 148.96 (C-2), 146.04 (C-16), 139.19, 135.07, 129.07, 128.85, 128.17, 128.04, 124.99, 124.72 (C-17+C-21), 119.69 (C-4), 115.80 (d, J=15.8Hz, C5), 98.12 (C-3), 97.86 (d, ^{2}J=24.9 Hz, C-7), 79.36 (C-11), 73.47 (C-10), 36.83 (C-8), 25.05 (C-9).



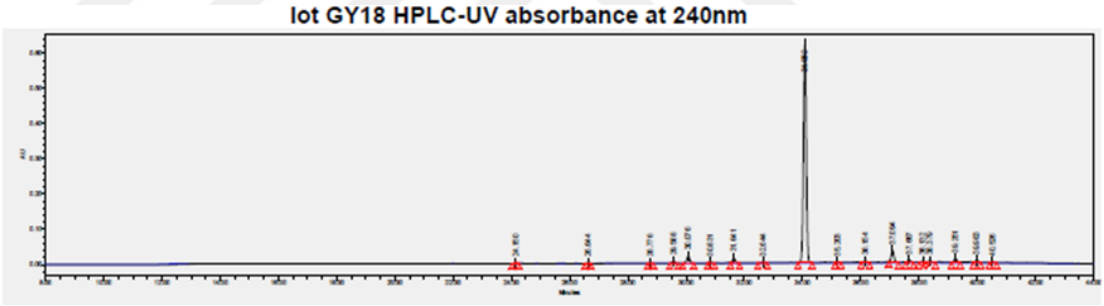
COSY :

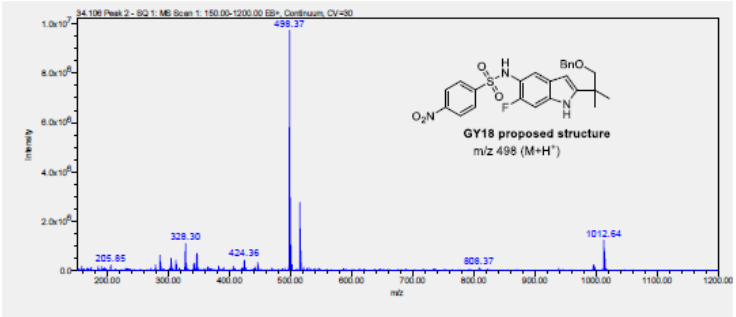


HSQC :

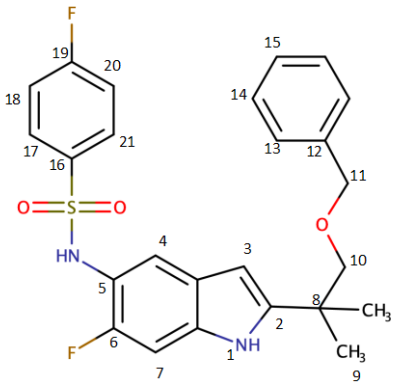


LC-MS : RT=34.1 min., m/z, 498 $[M+H]^+$





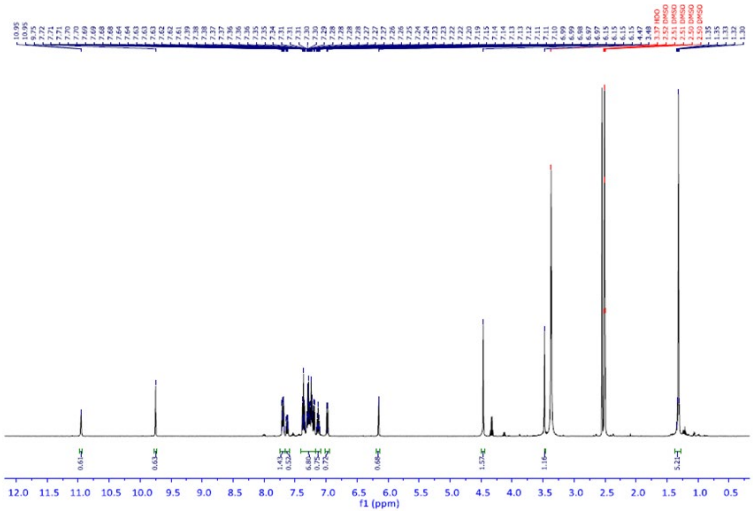
N-{2-[1-(benzyloxy)-2-methylpropan-2-yl]-6-fluoro-1H-indol-5-yl}-4-florobenzene-1-sulfonamide (8)



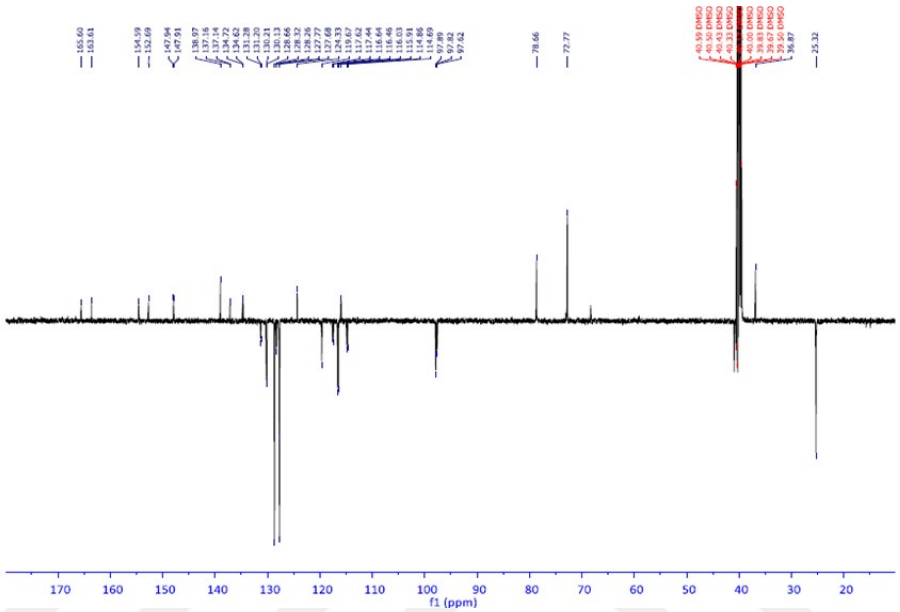
Compound 1 (0.50 mmol) and 4-florobenzene sulfonyl chloride (1.0 mmol) were reacted in 5 mL ethanol as described in the general method at 3.2.2.2. The form of the compound is brown solid.

Yield (%)	:	76
Physical appearance	:	Brown oily solid
Solubility	:	DMSO
Molecular formula	:	C25H24F2N2O3S
Molecular weight (g/mol)	:	470.15

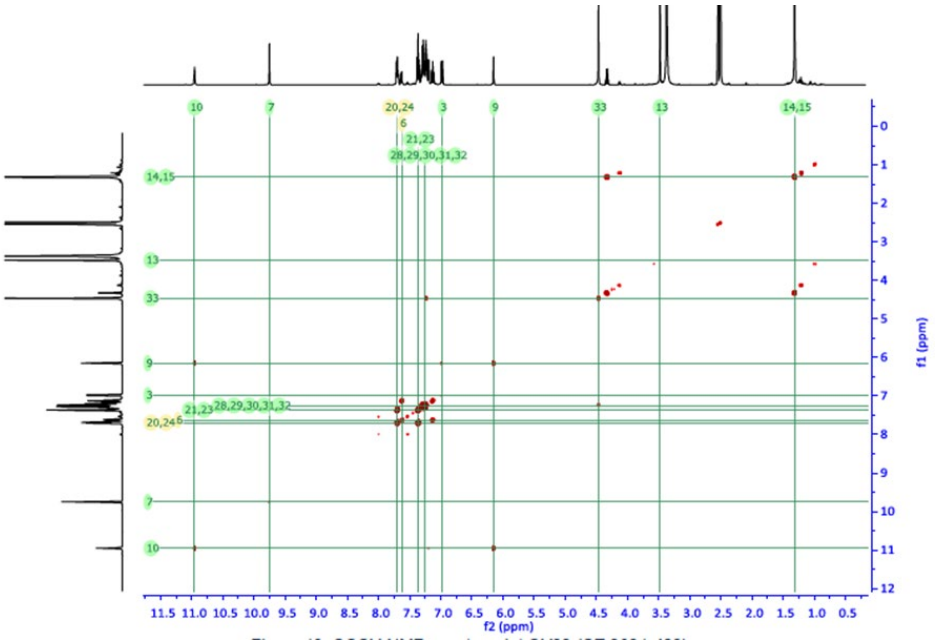
^{1}H NMR (500 MHz, DMSO-d_6) δ (ppm) : 10.95 (d, J = 2.3 Hz, 1H,NH), 9.75 (s, 1H, H-1), 7.72-7.10 (m, 10H, H-arom), 6.98 (dd, J_1 = 11.7Hz, J_2=0.8 Hz, 1H, H-7), 6.15 (dd, J_1 = 2.3 Hz, J_2=0.9 Hz, 1H, H-3), 4.47 (s, 2H, H-11), 3.48 (s, 2H, H-10), 1.32 (s, 6H,H-9).



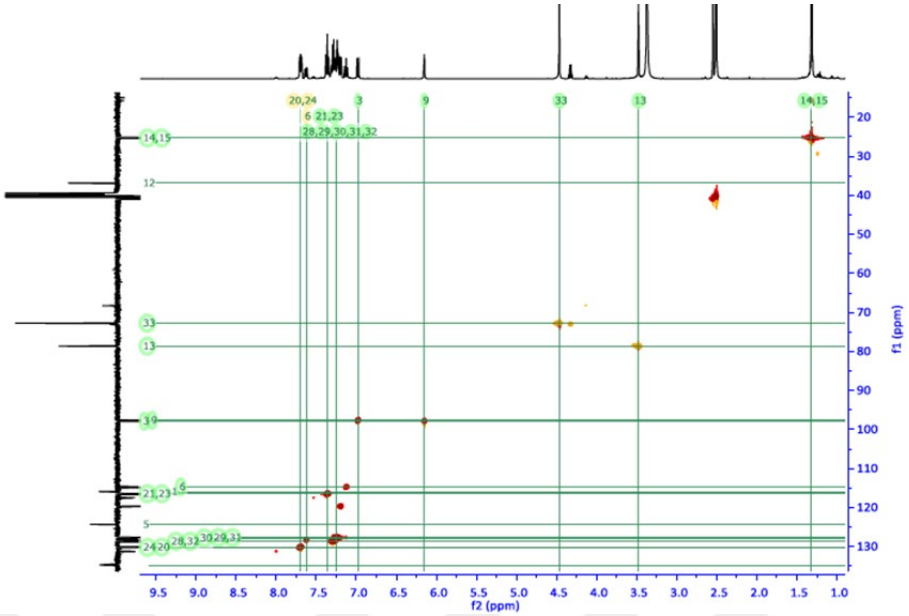
^{13}C APT NMR (126 MHz, DMSO-d_6) δ (ppm) : 163.61 (d, ^{1}J=250.7 Hz,C-19), 152.69 (d, ^{1}J=239.4 Hz, C-6), 147.9, 138.97, 130.21, 130.13, 127.77, 127.68, 124.33, 119.31 (C-4), 116.46, 115.91, 114.69, 97, .89 (C-3), 97.79 (d, ^{2}J=25.2 Hz, C-7), 78.66 (C-11), 72.77 (C-10), 36.97 (C-8), 25.32 (C-9).



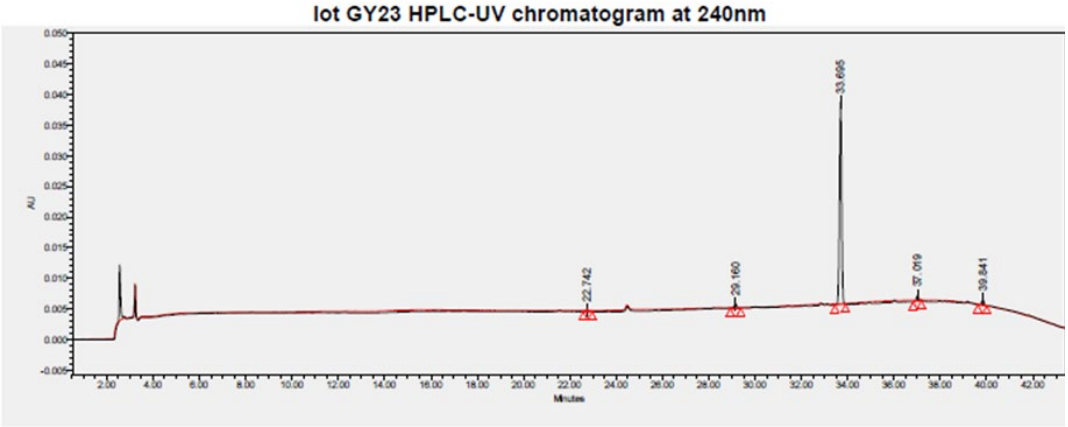
COSY :

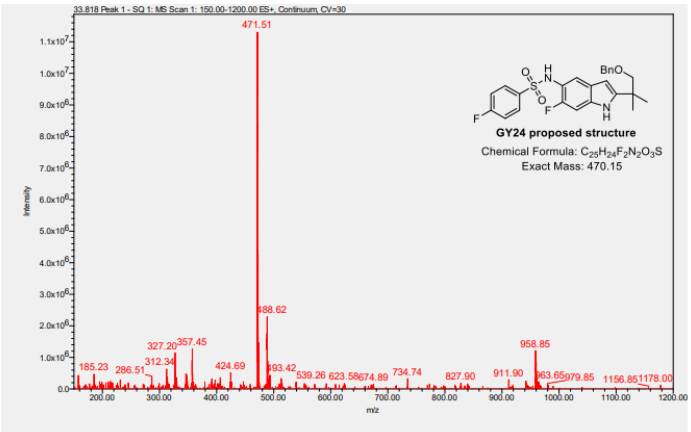


HSQC :



LC-MS : RT=33.7 min., m/z, 471 $[M+H]^+$





5. RESULTS AND DISCUSSION

5.1 Auto-QSAR

5.1.1 Data Set Preparation

In this study, we aimed to develop a QSAR model to predict the molecular CFTR activities of some 5 substituted 6-fluoro indole derivatives that were derived from the Tezacaftor CFTR modulator drug. A predictive QSAR model has been developed using regression-based learning techniques as KPLS. All obtained results will be shared and discussed in this section.

UniProtID Determination

The phenylalanine 508 (F508del) CFTR proteins number was searched *via* UniProt biological database [50] to find unique identification code of the CFTR protein. The results were listed for a few species commonly encountered such as HUMAN, BOVIN, ECOLI, MOUSE, etc. and ´P13569´ identification code was selected which dedicated for human organism.

Dataset Creation

UniProtID= P13569 was searched *via* Reaxy database [51] to build structure activity dataset. Obtained results were filtered with single protein and wild type (in a natural population) protein. First initial data set contained 5404 compounds. The obtained data was filtered to eliminate compounds without pEC50 value. Then, molecules were separated into two groups as active and inactive. This segregation was obtained by calculating the median of the pEC50 values of all structures: compounds below the median were accepted as inactive, and compounds above the median as active. This filtration reduced the number of structures of our dataset to 3801 compounds. A final tuning step consisted of calculating the average structural similarity of the molecules in order to gather structures with similar skeletons. As a result, 1735 compounds constituted our final dataset.

5.1.2 Auto-QSAR Model Building

Auto-QSAR calculations were performed with Schrödinger Suite 2017-1 release.

Model Building Process

To build the Auto-QSAR model the Schrodinger database [52] was used.

Schrodinger database is an automated machine-learning application to help to select ideal combinations of descriptors and learning methods. Figure 2.10 provides a full workflow of the Auto-QSAR implementation. Randomly, 50% of the model was trained, 40% tested and 10% validation set was created. Calculated EC50 values were between 3.5 to 10.2. The median value of pEC50 was founded as 5.13, therefore, below of median is accepted as inactive compounds and upper of median is accepted as active compounds.

Model Evaluation: The models presented in Table 5.1 is obtained via Schrodinger database.

Table 5.1 The models properties obtained this QSAR study

Model Code	Score	SD	R^2	RMSE	Q^2
kpls_dendritic_10	0.5914	0.2528	0.8865	0.3985	0.7142
kpls_radial_10	0.5808	0.2890	0.8526	0.4139	0.6917
kpls_linear_10	0.5691	0.2376	0.8998	0.4044	0.7057
kpls_dendritic_8	0.5640	0.2412	0.8989	0.4070	0.6956
kpls_dendritic_31	0.5606	0.2630	0.8799	0.4145	0.6836
kpls_linear_4	0.5566	0.2475	0.8912	0.4128	0.6938
kpls_dendritic_4	0.5535	0.2594	0.8805	0.4175	0.6867
kpls_linear_8	0.5518	0.2472	0.8937	0.4142	0.6848
kpls_dendritic_44	0.5518	0.2692	0.8740	0.4202	0.6747
kpls_linear_31	0.5455	0.2505	0.8910	0.4178	0.6785

The dendritic model (kpls_dendritic_10) was found best option as R^2 and Q^2 values >0.6, RMSE value was between 0.4-0.5 as ideally, SD value was low comparing other models. This means that Q^2 and RMSE statistics for the predictions were comparable for the model.

Table 5.2 The distribution of dataset into training and test set as Inactive and Active in Auto-QSAR Model

	Active	Inactive	Total	Active / Inactive
Training set	793	768	1561	1.03
Test set	86	88	174	0.98
Total	879	856	1735	

5.1.3 Predicting the Bioactivity of 5-substituted 6-fluoro indole derivatives

After creating the SMILES files for the fifteen 5-substituted 6-fluoro indole compounds using the Marvin Sketch software [53], ´kpls_dendritic_10´ Auto-QSAR model was run and predicted EC50 values were obtained for the 5-substituted 6-fluoro indole derivatives **1**-**15**. Obtained results are provided in Table 5.3.

Table 5.3 pEC50 values of fifteen 5-substituted 6-fluoro indole derivatives

Compound #	pEC50 Value
1	4.92
2	4.83
3	4.84
4	5.06
5	5.00
6	4.96
7	4.90
8	4.85
9	4.98
10	5.11
11	5.10
12	4.79
13	4.80
14	4.87
15	4.89

First of all, all of the fifteen 5-substitued indole derivatives were found to show similar pEC50 values varying within the 4.83 and 5.11 range, a promising range close to the pEC50 values of the approved CFTR modulators Tezacaftor and Ivacaftor that are of >5, 5.2 respectively [45].

Compound **10** showed the highest pEC50 value (5.11) being followed by compound **11**, compound **4** and compound **5** which are respectively secondary amines and sulfonamide derivatives respectively.

To evaluate the effects of different substitutions on the indole structure on pEC50 values, results have also been gathered on a pEC50 scale as follows (Figure 5.1).

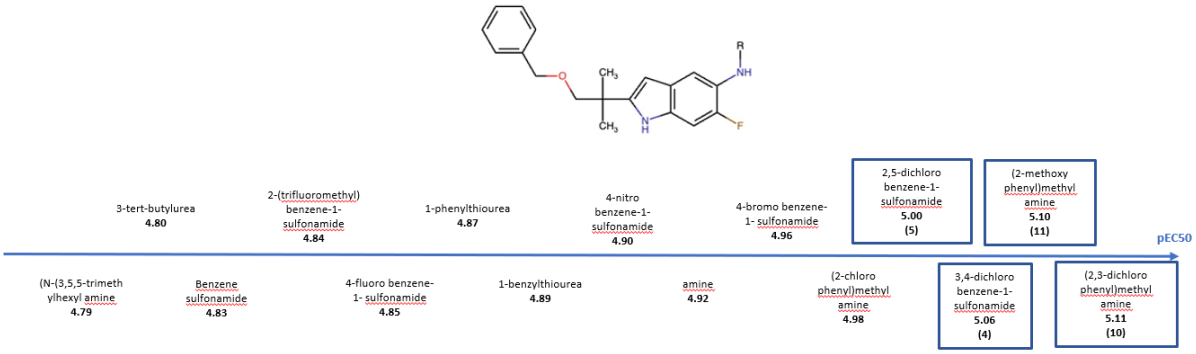


Figure 5.1 The effects of the different substitutions on the pEC50 value

The analysis of the results suggested that electron-withdrawing groups might increase the bioactivity of sulfonamide derivatives since higher pEC50 values were obtained for all the halide substituted phenyl sulfonamide derivatives (**3**-**5**) when compared to the phenyl sulfonamide **2**. Also, the higher values obtained for di-halosubstituted molecules confirmed this hypothesis for this series. Concerning the secondary amine structures (**6**-**12**), electron-donating groups were found, this time, to increase the bioactivity as the highest pEC50 value was predicted for the methoxy substituted derivative. Besides, the di-substituted molecules of this series provided similar results when compared to those obtained for the sulfonamide derivatives. Finally, for molecules with an urea or thiurea function as a linker (**13**-**15**) no significant correlation was established between the bioactivity compared and the structure of the molecules of this series.

5.2. Chemistry

Promising pEC50 values were obtained for all of the fifteen 5-substituted 6-fluoro indole derivatives that were tested from developed QSAR model. The obtained pEC50 values are between 4.79 and 5.11. Thus, these Tezacaftor analogues were also aimed to be synthesized.

5.2.1. Synthesis of compound 1, the primary amine further derivatized

In order to synthesis analogues of the Tezacaftor molecule, first the core 5-amino-6-fluoroindole ring was synthesized starting from the commercial 2-(1-(benzyloxy)-2-methylpropan-2-yl)-6-fluoro-5-nitro-1H-indole compound. The starting material was reduced into the corresponding 5-amino-6-fluoroindole structure using activated zinc in acetic acid and the expected compound was successfully isolated with a yield of 98% (Figure 5.2).

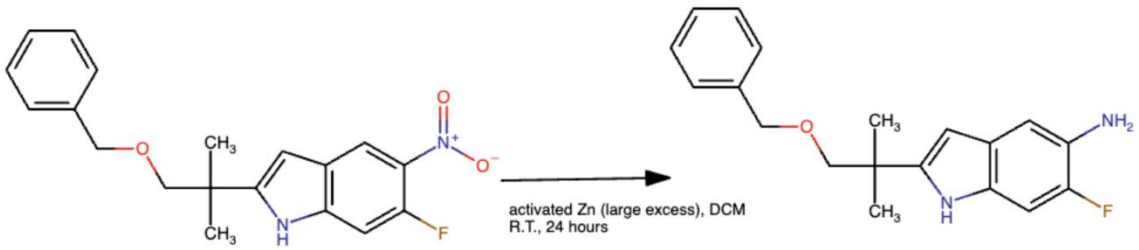
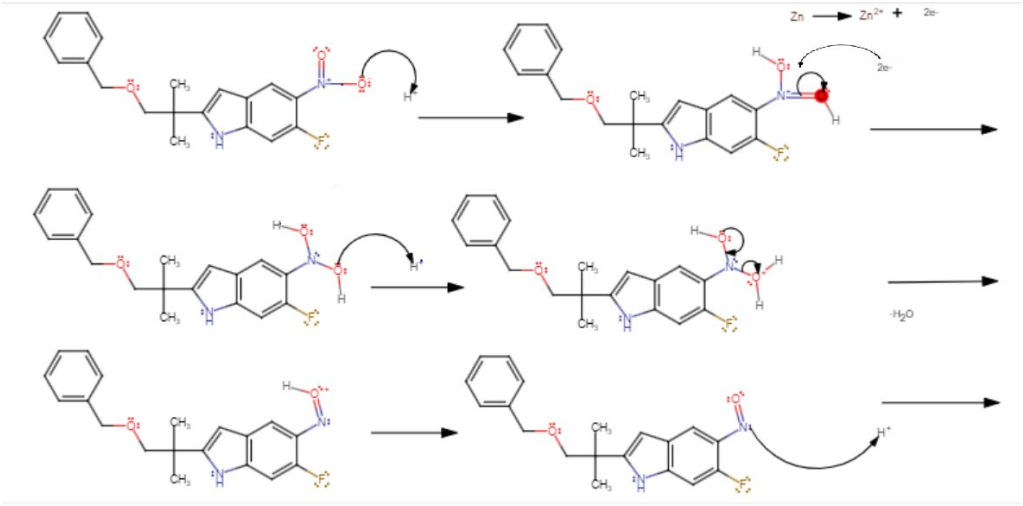


Figure 5.2 Reduction of 2-(1-(benzyloxy)-2-methylpropan-2-yl)-6-fluoro-5-nitro-1H-indole

The reduction is thought to be undertaken according to the mechanism depicted in Figure 5.3, the source of the electrons responsible for the reduction of the nitro group being provided by 3 equivalents of Zn that itself is oxidized into Zn^{2+}.



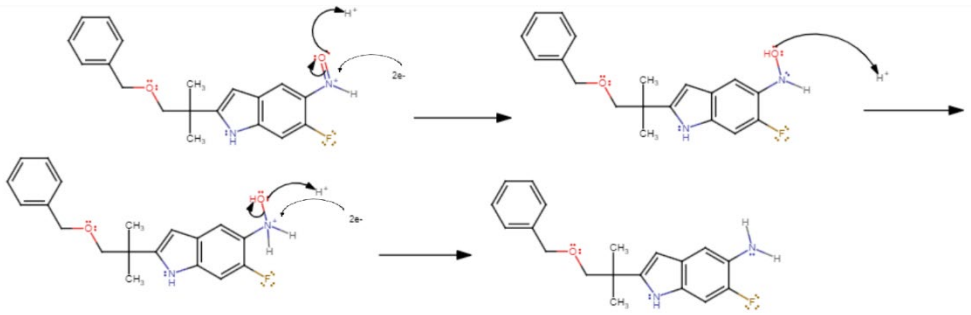


Figure 5.3 Reduction mechanism of the nitro group into the amine

The conversion of the nitro group into the primary amine was monitored by TLC and confirmed by spectral analyses. Indeed, the conversion was confirmed with the analysis of the NMR spectra as the peak at 8.33 ppm that corresponded to the aromatic proton of the indole ring neighboring the nitro group disappeared and a new peak at 6.77 ppm appeared. This signal found to appear up field confirmed the transformation of the electron-attracting nitro group into the electron donating amino (Figure 5.4).

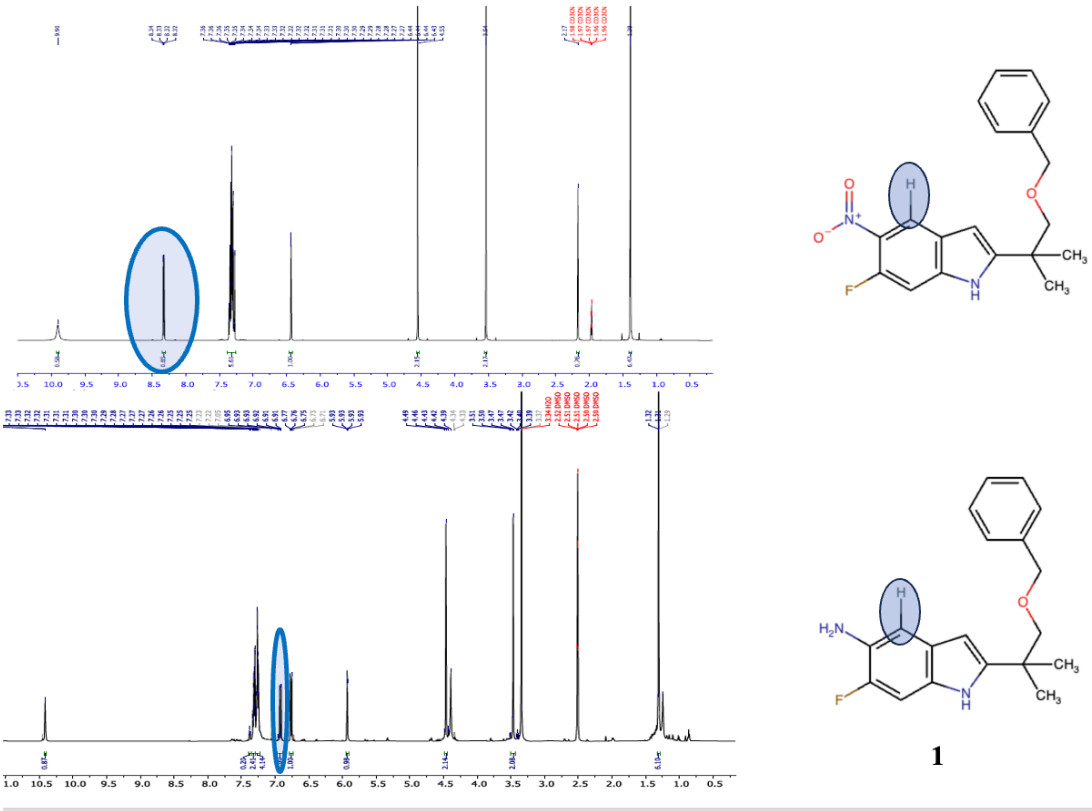


Figure 5.4 ^{1}H NMR spectra of the starting material and compound **1**

The reduction was also confirmed by ^{13}C APT NMR and LC-MS analyses with the appearance of a peak at 105.3 ppm (signal of the carbon next to the one bearing the amino group) and the disappearance at 118.8 ppm (signal of the carbon next to the one bearing the nitro group) (Figure 5.5) and the expected m/z= 313 value on the mass spectrum corresponding to the M+H peak.

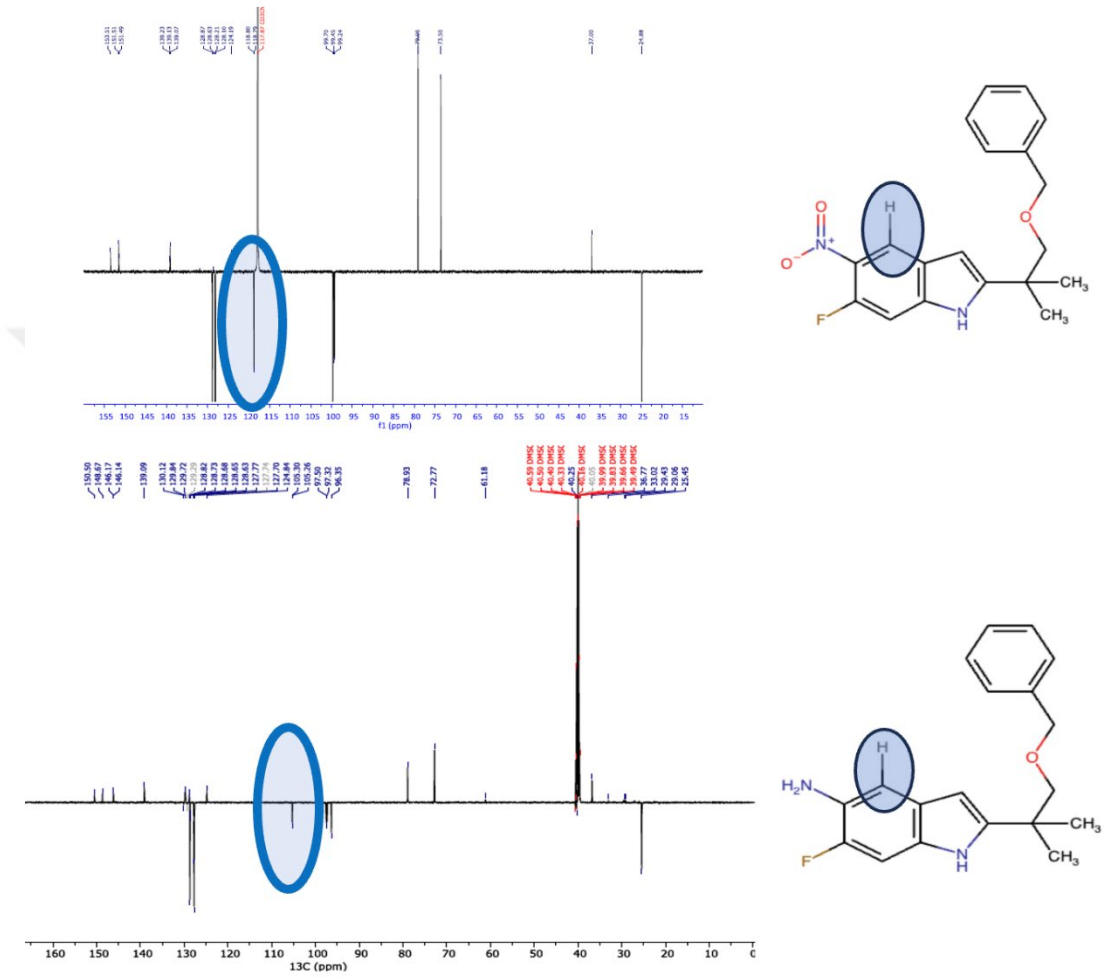


Figure 5.5 ^{13}C APT NMR spectra of the starting material and compound **1**

The successfully obtained compound **1** was then further derivatized in order to generate the targeted series of 5-substituted 6-fluoroindole derivatives. The derivatization has been carried out through the amino functional group located on the 5th position of the indole ring. The aim consisted of generating three different classes of molecules from this starting material. For the sulfonyl chlorides were used to convert the primary amine of the starting material into sulfonamides (**2-8**). For the second group a series of aldehydes were used to synthesize secondary amines via a reductive amination (compounds **9-12**). To prepare the last class of compounds the amine was reacted with a series of isocyanates and

isothiocyanates to obtain the corresponding ureas and thioureas (**13)** and (**14-15**). The following parts detail the procedure and results obtained for each of the generated molecule classes.

5.2.2 Synthesis of sulfonamides (2-8) with a series of sulfonyl chlorides

Compound **1** was reacted with a series of sulfonyl chlorides in order to synthesis the corresponding sulfonamides (Figure 5.6).

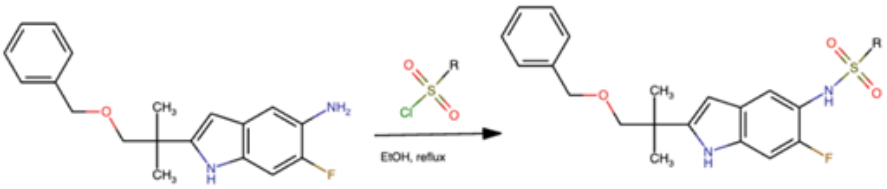


Figure 5.6 Reaction scheme to synthesis compounds **2-8**

The reaction was carried out in ethanol under reflux overnight and compounds **2-8** were obtained with satisfactory yields (Table 5.4).

Table 5.4 Series of sulfonyl chlorides to synthesis compounds **2-8**

Sulfonyl Chloride	Yield (%)	Obtained Compounds
Benzene sulfonyl chloride	71	2
4-nitro benzene sulfonyl chloride	70	7
4-fluoro benzene sulfonyl chloride	61	8

The structures resulted from an addition-elimination reaction where first the amine attacked the electrophilic sulfur of the sulfonyl chloride followed by the elimination of the chloride ion to synthesis the sulfonamide moiety (Figure 5.7).

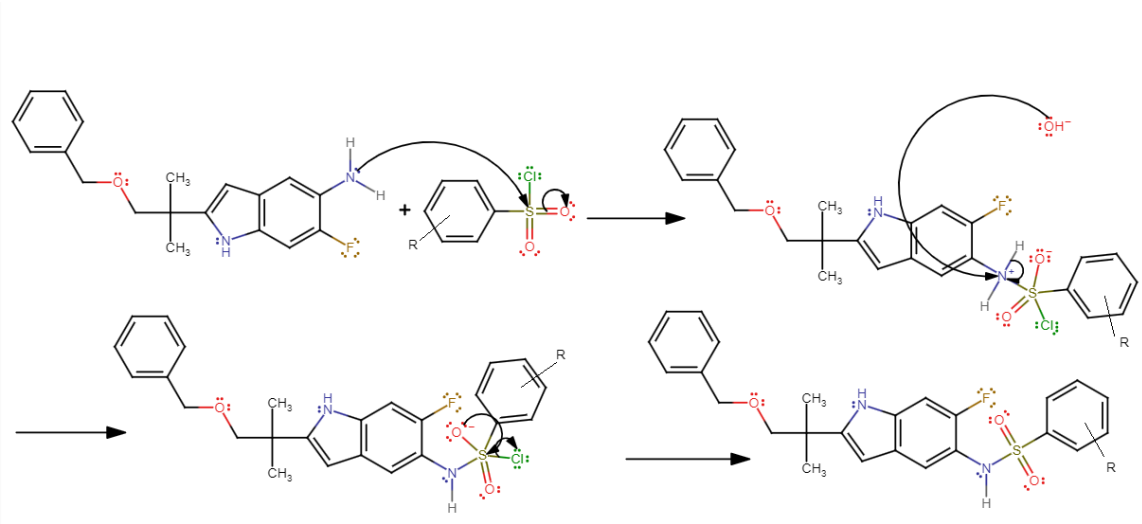


Figure 5.7 Mechanism of the addition of the amine **1** on sulfonyl chloride derivatives

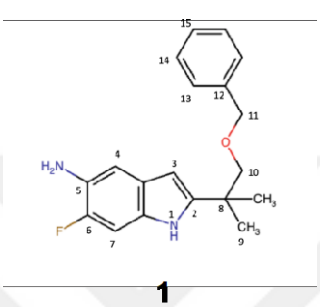
The procedure was applied to a total of nine sulfonamides, products were purified on silica column and sent to spectral analyses. Unfortunately, only three products were found to be obtained.

The conversion of compound 1 into sulfonamide derivatives was monitored by spectral analyses.

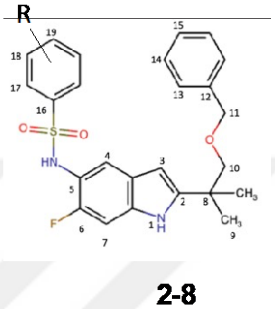
The shift of the signal of the aromatic indole proton (H-4) on carbon C-4 next to the one bearing the nitrogen atom (C-5) played a crucial role in confirming the conversion of the amine group into a sulfonamide. Indeed, the protons numbered as 3, 4 and 7 and the corresponding carbons to which these hydrogens are the bound were easy to detect and follow. The signal of C-7 was the most distinct as this carbon neighbors a carbon bearing a fluorine atom, thus appears as a doublet (^2J=25Hz). In compound **1**, its chemical shift was found to be at 97.7 ppm. The signal of the corresponding H-7 proton was then determined via the HSQC spectrum and was identified at 6.93 ppm. Since the only carbons with similar shifts were C-3 and C-7, and given that the H-3 proton was expected to yield a signal up field compared to the signal of the H-4 proton, the HSQC spectrum facilitated the identification of these signals: C-3 at around 98 ppm and the corresponding H-3 proton at 6.15 ppm, and C-4 at 106 ppm with the corresponding H-4 at 6.77 ppm. The spectra obtained

for the sulfonamide derivatives also exhibited the characteristic peaks for C-3 and C-7 and their corresponding protons at similar shifts whereas the signals for C-4 and H-4 were found to be in downfield. Thus, shifts identified for these three positions, provided in the following table (Table 5.5), demonstrated the successful conversion of the amino group into the expected sulfonamide.

Table 5.5 Chemical shifts obtained for C3, C6 and C9 and their corresponding protons.



1



2-8

Compound #	δ C 7 (ppm)	δ H7 (ppm)	δ C 3 (ppm)	δ H 3 (ppm)	δ C 4 (ppm)	δ H 4 (ppm)
1	97,4	5,93	96,3	6,93	105,3	6,77
2	97,7	6,22	98,04	7	120,86	7,4
7	97,86	6,23	98,12	7,02	119,69	7,4
8	97,62	6,15	97,89	6,98	114,69	7,63

Also, the presence of additional aromatic signals in both ^{1}H and ^{13}C ATP NMR and the presence of the expected m/z values in LC-MS data allowed to validate the obtention of the sulfonamide derivatives.

5.2.3. Synthesis of the secondary amines (9-12) *via* reductive amination of compound 1 with a series of aldehydes

A series of aldehydes were selected to derivatize the primary amino function of compound **1** (Table 5.3) to obtain the corresponding secondary amines (Figure 5.8).

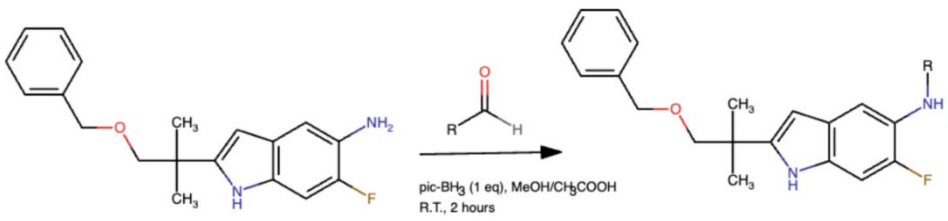


Figure 5.8 Synthesis of compounds **9-12** from a series of aldehydes

The compounds were expected to be obtained through a reductive amination for which the mechanism is provided in Figure 5.9.

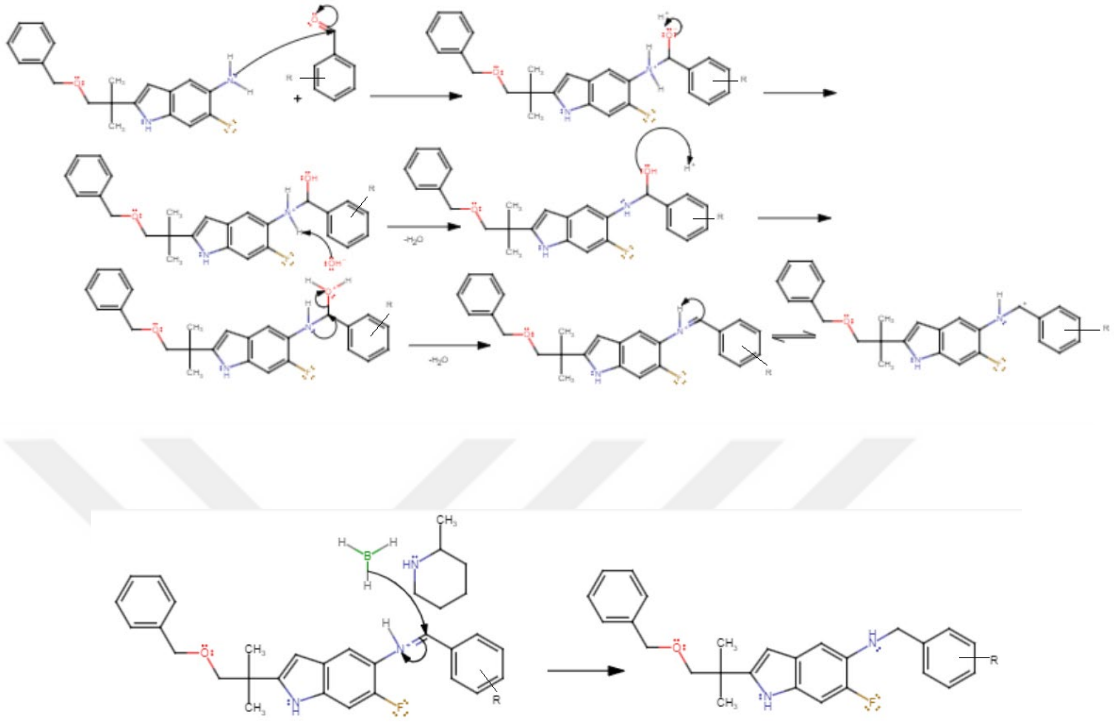


Figure 5.9 Mechanism of the reductive amination

Initially sodium cyanoborohydride was chosen as the reductive agent to convert the imines into the corresponding amines. However, as the expected products were not obtained with satisfying yields (only 3.5% for 4-fluorobenzaldehyde), further investigations were carried out and the procedure was optimized using pic-BH_3 (2-methylpyridine borane) as the reducing agent [50]. Unfortunately, the expected amines were not isolated purely with this agent neither.

Indeed, even if some of these compounds were isolated and appeared pure before being sent to analyses (single spots isolated from silica column), subsequent data revealed that they were actually in mixtures (undetermined structures).

5.2.4. Synthesis of the ureas and thioureas with a series of isocyanate and isothiocyanates

In this study a series of ureas and thioureas were also aimed to be obtained by

deriving the primary amine of the 6th position of the indole ring with a series of isocyanate and thioisocyanates (Figure 5.10).

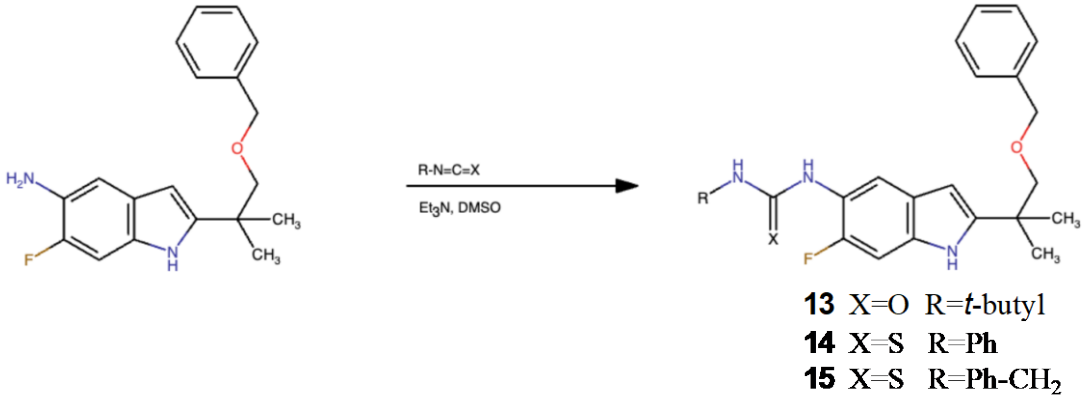


Figure 5.10 Derivatization of the amino group into urea and thioureas

The coupling reactions were conducted in alkali medium to synthesis the expected urea and thiourea derivatives after tautomerization of the iminoalcohol/thiols according to the following mechanistic scheme (Figure 5.11).

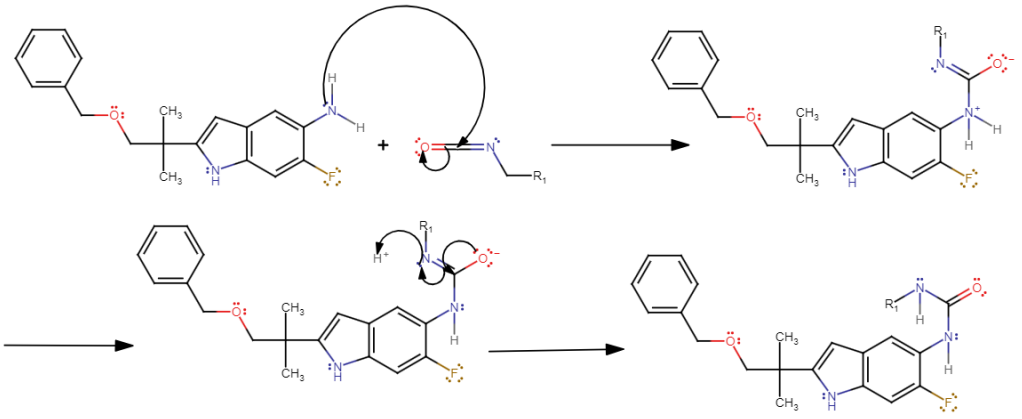


Figure 5.11 Mechanism for the synthesis of urea derivatives

The procedure was applied to a total of nine iso/ thioisocyanates, products were purified on silica column and sent to spectral analyses. Unfortunately, data revealed that complex mixtures were obtained indicating a poor stability of the urea and thiourea structures synthesized from the starting 6-fluoro-5-amino indole.

5.3. Biology

CFTR plays a significant role in the initiation and progression of various cancers, and it is reasonable to consider the administration of CFTR modulators in cancer treatment of CF patients. Successfully synthesized 5-substituted 6-fluoro indole derivatives´ cytotoxic activities were evaluated for this purpose on some cancer lines.

The cytotoxicity of the synthesized compounds was examined on OVSAHO and OVCAR-3, two distinct ovarian cancer cell lines to assess their potential efficacy. Data indicated that the synthesized derivatives (**1, 2, 7, 8**) did not exhibit any detectable activity. These results underline the necessity to further explore, aiming to refine the activity of these 5-substituted 6-fluoro indole derivatives and enhance their cytotoxic potential (Table 5.6).

Table 5.6 Evaluation of the cytotoxic activity of synthesized compounds on OVSAHO and OVCAR-3 ovarian cancer lines

Compound #	**OVSAHO (IC50 µM)**	**OVCAR-3 (IC50 µM)**
1	N.I.	N.I.
2	N.I.	N.I.
7	N.I.	N.I.
8	N.I.	N.I.

N.I.: No Inhibition

6. CONCLUSION

In conclusion, in this study we developed a QSAR model to predict the EC50 values of some Tezacaftor analogues, more precisely a series of 5-substituted 6-fluoro indole derivatives as CFTR modulators. Promising pEC50 values were obtained for all of the fifteen 5-substituted 6-fluoro indole derivatives that were tested. The obtained pEC50 values are between 4.79 and 5.11.

These Tezacaftor analogues were also aimed to be synthesized afterwards. Unfortunately, only four compounds were successfully obtained, as the rest of them were not isolated purely.

The synthesized compounds were also evaluated for their cytotoxicity. The results indicated that the obtained sulfonamide derivatives did not present any cytotoxicity.

In future studies, new 5-substituted 6-fluoro indole derivatives may be tested for their ability to bind CFTR *via* the developed QSAR model and these structures could be synthesized to generate structures with enhanced cytotoxic potential. Furthermore, to compare *in-vivo* CFTR function to the predicted EC50 values, a more appropriate biological activity test may be used in upcoming research.

7. REFERENCES

[1] Simmonds et al., Is it cystic fibrosis? The challenges of diagnosing cystic fibrosis, Pediatric Respiratory Reviews, 2019; 31, 6-81.

[2] Agency, European Medicines. Assessment report. London, United Kingdom: Committee for Medicinal Products for Human Use (CHMP), 2018.

[3] Balfour Lynn et al., CFTR modulator therapies-Effect on life expectancy in people with cystic fibrosis, Pediatric Respiratory Reviews, 2020; 3-8.

[4] Ong T. et al., Cystic Fibrosis: A Review, The Journal of the American Medical Association, 2023; 329(21):1859-1871.

[5] J. Ferlay, I. Soerjomataram, R. Dikshit, et al., Cancer incidence and mortality worldwide: sources, methods, and major patterns, International Journal of Cancer, 2015.

[6] Timothy A. et al., Cancer in 2022, American Association for Cancer Research, 2022; 2(12): 2733-2738.

[7] Margarida D. Amaral, Margarida C. Quaresma and Ines Pankonien, What Role Does CFTR Play in Development, Differentiation, Regeneration and Cancer, International Journal of Molecular Sciences, 2020.

[8] Qi Wang, Shubing Jia, Jie Zheng, Rongwu Xiang, Yong Cui, Jinghai Zhang, Yijia Xu, Mingyi Zhao, Analysis of CFTR gene expression as an immunological and prognostic biomarker in pan-cancers, Computers in Biology and Medicine, 2022.

[9] Giada R. et al., New Insights into the Binding Features of F508del CFTR Potentiators: A Molecular Docking, Pharmacophore Mapping and QSAR Analysis Approach, Pharmaceuticals, 2020.

[10] Giada R. et al., Molecular Docking and QSAR Studies as Computational Tools Exploring the Rescue Ability of F508del CFTR Correctors, International Journal of Molecular Sciences, 2020.

[11] Hughes, David L., Patent Review of Synthetic Routs and Crystalline Forms of the CFTR-Modulator Drugs Ivacaftor, Lumacaftor, Tezacaftor, and Elexacaftor, OPR&D, 2019.

[12] https://www.trikaftahcp.com/mechanism-of-action (accessed 2023).

[13] Mermis, J. et al., Impact of CFTR modulators (lumacaftor/ivacaftor and tezacaftor / ivacaftor) on long-term patient outcomes, Pediatric Pulmonology, 2019; 350-351.

[14] Enrico C, et al., The impact of FDA and EMA regulatory decision-making process on the access to CFTR modulators for the treatment of cystic fibrosis, Journal of Rare Diseases, 2022; 17(188).

[15] David L. et al., “real world” impact of cftr modulators (lumacaftor/ivacaftor and tezacaftor/ ivacaftor) on long-term patient outcomes review of synthetic routes and crystalline forms of the CFTR-modulator drugs Ivacaftor, Lumacaftor, Tezacaftor, and Elexacaftor, Organic Process Research and Development, 2019; 2302-2322.
[16] Krysa J et al., Pancreas and Cystic Fibrosis: The Implications of Increased Survival in Cystic Fibrosis, Pancreatology, 2007; 7(5-6):470-478.
[17] Pharma, Rx Media. [Online] 2022.
[18] https://www.medchemexpress.com/Targets/CFTR.html (accessed 2023).
[20] Huang W. et al., Epigenetic regulation of CFTR during cancer development and implications in cancer treatment, FASEB journal, April 2014; 28(1).
[21] Basso, D., Zamboni, G., & Padoan, Cancer and cystic fibrosis: biological interactions and clinical implications, Future Oncology, 2020; 16(11):591–601.
[22] Hanrahan J. W, Tacharanite, J. A et al., CFTR and cancer, American Journal of Physiology-Lung Cellular and Molecular Physiology, 2017; 312(6), L711–L718.
[23] Sala, A., Bettini, S., Fabbri et al., Cystic fibrosis transmembrane conductance regulator and cancer: signaling, metabolism, and new therapeutic targets, International Journal of Molecular Sciences, 2020; 21(21):8025.
[24] N Malats, T Casals, M Porta, L Guarner et al., Cystic fibrosis transmembrane regulator (CFTR) F508 mutation and 5T allele in patients with chronic pancreatitis and exocrine pancreatic cancer, 2001; 48:70-74.
[25] Robert R. McWilliams, M.D et al., Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) Gene Mutations and Risk for Pancreatic Adenocarcinoma, NIH Public Access, 2010; 116(1): 203-209.
[26] Tümmler B. et al., Post-approval studies with the CFTR modulators Elexacaftor, Tezacaftor, and Ivacaftor, Frontiers in Pharmacology, 2020;14.
[27] Stephan U. et al., Mucoviscidosis (cystic fibrosis, cystic pancreatic fibrosis), Monatsschrift Kinderheilkunde, 1969; 117(9):574-577.
[28] JF Linnekamp, TK Starr, DA Largaespada et al., CFTR is a tumor suppressor gene in murine and human intestinal cancer, Oncogene advance online publication, 2015;1-9.
[29] Liu C. et al., CFTR Functions as a Tumor Suppressor and Is Regulated by DNA Methylation in Colorectal Cancer, Cancer Management and Research, 2020; 12:4261-4270.
[28] Patricia Scott, Kyle Anderson, Mekhla Singhania and Robert Cormier, Cystic Fibrosis, CFTR, and Colorectal Cancer, International Journal of Molecular Sciences, 2020; 21.
[30] Palma, Alejandra Graciela et al., Functional relationship between CFTR and RAC3

expression for maintaining cancer cell stemness in human colorectal cancer, Cellular Oncology, 2021; 44(3):627-641.
[31] Bhattacharya B, Zachary S. et al., CFTR and Gastrointestinal Cancers: An Update, Journal of Personalized Medicine, 2022;12(6).
[32] Xue Peng, Zhao Wu et. al., Overexpression of cystic fibrosis transmembrane conductance regulator (CFTR) is associated with human cervical cancer malignancy, progression, and prognosis, Gynecologic Oncology, 2012; 125:470–476.
[33] Xu Jiao et al., High level of CFTR expression is associated with tumor aggression and knockdown of CFTR suppresses proliferation of ovarian cancer in vitro and in vivo, Oncology Reports, 2015; 33(5):2227-34.
[34] Madelung, W. et al., Berichte der deutschen chemischen Gesellschaft, 1912; 45(1128).
[35] Meyer, H. et al., Arch. Exp. Path. Pharm, 1899;42(109-118).
[36] Bruice, E. et al., R. J., Arch. Biochem. Biophys., 1956; 62(305 -317).
[37] Free J. et al., Journal of Medicinal Chemistry, 1964; 7(395-399).
[36] Hansch, C. et al., Ace. Chem. Res., 1969; 2(232-239).
[38] Hugo Kubinyi, QSAR: Hansch Analysis and Related Approaches. VCH; 1993
[39] Khan Au. et al., Descriptors and their selection methods in QSAR analysis: paradigm for drug design, Drug Discovery Today, June 2016
[40] Cumming JG, Davis AM, Muresan S, Haeberlein M, Chen H, Chemical predictive modelling to improve compound quality, Nat. Rev. Drug Discovery, 2013; 12(12), 948–962.
[41] Todeschini R, Molecular Descriptors for Chemoinformatics, Wiley-VCH; 2009.
[42] Sastry M, et al., Large scale systematic analysis of 2D fingerprint methods and parameters to improve virtual screening enrichments, J. Chem. Inf. Model,2010; 50(5), 771–784.
[43] Duan J, Dixon SL, Lowrie JF, Sherman W, Analysis and comparison of 2D fingerprints: insights into database screening performance using eight fingerprint methods. J. Mol. Graph. Model, 2010; 29(2), 157–170.
[44] Dixon SL et al., A new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results, J. Comp. Aided Mol. Des, 2006; 20(10–11), 647–671.
[45] An Y et al., Kernel-based partial least squares: application to fingerprint-based QSAR with model visualization, J. Chem. Inf. Model, 2013; 53(9), 2312–2321.
[46] Hideyuki S. et al., Kernel Analysis of Partial Least Squares (PLS) Regression Models, Applied Spectroscopy, 2011; 65(5).

[47] Dixon SL, Villar Ho, Investigation of classification methods for the prediction of activity in diverse chemical libraries, J. Comp. Aided Mol. Des., 1999; 13(5), 533–545.
[48] Cumming JG, Davis AM, Muresan S, Haeberlein M, Chen H, Chemical predictive modelling to improve compound quality, Nat. Rev. Drug Disc, 2013; 12(12), 948–962.
[49] Steven L Dixon et al., AutoQSAR: an automated machine learning tool for best-practice QSAR modelling, Future Medicinal Chemistry, 2016.
[50] EMA Assessment Report of Symkevi (Tezacaftor / Ivacaftor), 26 July 2018
[51] A. Berichte et al., der deutschen chemischen Gesellschaft von Baeyer, 1866;2: 555-558.
[52] A. Nature. Lapworth., A. Nature, 1896;53: 347-348.
[53] Lednicer, D., Strategies for Organic Drug Synthesis and Design, 2nd ed.: Hoboken; 2009.
[54] Katritzky, A. R. et al., Indoles, Part One: The Chemistry of Heterocyclic Compounds, 1996.
[55] Sharma V et al, Biological importance of the indole nucleus in recent years, Journal of Heterocyclic Chemistry, 2010; 47(3): 491-502.
[56] Nakazaki M. et al., American Chemical Society Direct synthesis of indole by the Fischer indole synthesis, Journal of Organic Chemistry, 1976;41(10): 1877-1877.
[57] Robinson, B. et al., Chemical Reviews, 1969; 69(227).
[58] Shelke G, Development of Synthetic Methods for Selected Indole Based Heterocyclic Compounds from Indole, Isatin and Indoline-2-thione [Doctoral thesis dissertation], Birla institute of technology and science Pilani; 2016.
[59] Möhlau, R. et al., Berichte der deutschen chemischen Gesellschaft, 1881; 14(171).
[60] Crum-Brown A. et al., Trans. Roy. Soc. Edinburgh, 1868; 25(151-203).
[61] Gassman, P. et al., W. Journal of the American Chemical Society, 1974; 96(5495).
[62] Bartoli, G. et al., Tetrahedron Letters, 1989; 30(2129).
[63] Velezheva, V. et al., Journal of Heterocyclic Chemistry, 2006; 43(873).
[64] Gribble, G. et al., W. Journal of the Chemical Society, 2000; 1045.
[65] Castro, C. et al., Journal of the American Chemical Society, 1969; 91(6464).
[66] Hegedus, L. et al., Journal of the American Chemical Society, 1976; 98(2674).
[67] Fürstner, A. et al., Angewandte Chemie International, 1996; 35(2442).
[68] Larock, R. et al., Journal of the American Chemical Society, 1991; 113(6689).
[69] Kaspar, L. et al., Tetrahedron, 2005; 61(11311).
[70] Willis, M. et al., Angewandte Chemie International Edition, 2005; 44(403).
[71] Li L. et al., Mori-Ban indole synthesis, Springer Berlin Heidelberg, 2006;180: 401-402.

[72] Li, J. et al., Name Reactions: A Collection of Detailed Mechanisms and Synthetic Applications, Springer, 2003.
[73] Sanyal N et al., Spectrochimica Acta Part A: Molecular Spectroscopy, Ultraviolet and infrared spectra of substituted indole, 1982; 38(8):933-935.
[74] Pavia D. L. et al., Introduction to Spectroscopy, 2015; 231-233.
[75] Lee, H. K. Encyclopedia of Analytical Chemistry. UK: John Wiley & Sons Ltd: Chichester, 2006; 1733-1749.
[76] Takeuchi H. et al., Normal coordinate analysis of the indole ring, Spectrochemical Acta Part A: Molecular Spectroscopy, 1986;42(9):1069-1078.
[77] Abele, E. et al., E. Chem. Heterocyclic Compounds, 2003; 39(3).
[78] Xue S, Ma L, Gao R, Lin Y, Linn Z, Synthesis and antiviral activity of some novel indole-2-carboxylate derivatives, Acta Pharmaceutica Sinica B, 2014; 4(4):313–321.
[79] Sellitto G, Faruolo A, Caprariis PD, Altamura S, Paonessa G, Ciliberto G, Synthesis and anti-hepatitis C virus activity of novel ethyl-1H-indole-3- carboxylates in vitro, Bioorg Med Chem, 2010; 18:6143–6148.
[80] Abdel-gawad H, Mohamed HA, Dawood KM, Badria FAR, Synthesis and antiviral activity of new indole-based heterocycles, Chem Pharm Bull, 2010; 58(11):1529–1531.
[81] Rudrapal M. et al., Analgesic and Anti-Inflammatory Potential of Indole Derivatives, Polycyclic Aromatic Compounds, 2022;1-22.
[82] Siddique S. et al., Evaluation of the anti-inflammatory, analgesic, anti-pyretic and anti-ulcerogenic potentials of synthetic indole derivatives, Scientific Reports, 2023;13(1).
[83] Palmer, R. M. J., Ashton, D. S. and Moncada, S., Vascular endothelial cells synthesize nitric oxide from L-arginine, Nature, 1988;333(6174):664-666.
[84] Da Silva G. et al., Anti-inflammatory and antinociceptive activities of indole-imidazolidine derivatives, International Immunopharmacology, 2011; 11(11):1816-1822
[85] Ozdemir A, Altıntop MD, Zitouni GT, Çiftçi GA, Ertorun I, Alatas O, Kaplancıklı ZA, Synthesis and evaluation of new indole-based chalcones as potential antiinflammatory agents, Eur J Med Chem, 2015; 89:304–309.
[86] Rani P, Srivastava VK, Kumar A. Synthesis and anti-inflammatory activity of heterocyclic indole derivatives, Eur J Med Chem, 2004; 39:449–452.
[87] Pedada SR, Yarla NS, Tambade PJ, Dhananjaya BL, Bishayee A, Arunasree KM, Philip GH, Dharmapuri G, Aliev G, Putta S, Rangaiah G, Synthesis of new secretory phospholipase A2-inhibitory indole containing isoxazole derivatives as anti-inflammatory and anticancer agents, Eur J Med Chem, 2004; 112:289–297.

[88] Shaker AM, Abdelall EK, Abdellatif KR, Abdel-Rahman HM, Design, synthesis, and biological evaluation of 2-(4-(methylsulfonyl) phenyl) indole derivatives with promising COX-2 inhibitory activity, J Appl Pharm Sci, 2018; 8(11): 001–008.
[89] Sravanth T. et al., Indoles -A promising scaffold for drug development, European Journal of Pharmaceutical Sciences, 2016; 91(1-10).
[90] Panda S. S. et al., new frontiers in the therapeutic potential of indole derivatives, Mini Rev. Med. Chem, 2017;17(6):500-533.
[91] Rossiter S. et al., Halogenated Indole-3-acetic Acids as Oxidatively Activated Prodrugs with Potential for Targeted Cancer Therapy, Bioorganic & Medicinal Chemistry Letters, 2002; 2523–2526.
[92] Hong B. et al., Synthesis and Cytotoxicity Sudies of Cyclohepta[b]indoles, Benzo [6,7] Cyclohepta[1,2-b] Indoles, Indeno[1,2-b] Indoles, and Benzo[a]Carbazoles, Journal of the Chinese Chemical Society, 2006; 53, 647-662.
[93] Kailin H. et al., Design, synthesis, and biological activity evaluation of novel anticancer agent 5-(2-carboxyethenyl) indole derivatives, Journal of Chemical and Pharmaceutical Research, 2014; 6(9):376-380.
[94] Yousif M et al., Synthesis and anticancer activity of novel 2-phenylindole linked imidazolothiazole, thiazolo-s-triazine and imidazolyl-sugar systems, Journal of applied pharmaceutical science, 2019; 9(01):6–14.
[95] Bhat M. et al., Novel indole derivatives of dihydropyrimidinone: Synthesis, characterization, molecular docking and antimicrobial activity, Journal of Molecular Structure, 2023;1291.
[96] Reiter, R. J., et al., Biochemical reactivity of melatonin with reactive oxygen and nitrogen species: a review of the evidence, Cell Biochem. Biophys, 2001;34(2):237-256.
[97] Hong W. et al., Synthesis and biological evaluation of indole core-based derivatives with potent antibacterial activity against resistant bacterial pathogens, The Journal of Antibiotics, 2017; 70:832–844.
[98] Hui X. et al., Antifungal agents. Part 4: Synthesis and antifungal activities of novel indole[1,2-c]-1,2,4-benzotriazine derivatives against phytopathogenic fungi in vitro, European Journal of Medicinal Chemistry, 2011; 46:364-369.
[99] Sravanth T. et al., Indoles -A promising scaffold for drug development, European Journal of Pharmaceutical Sciences, 2016; 91(1-10).
[100] David L. et al., Hughes, Patent Review of Synthetic Routes and Crystalline Forms of the CFTR-Modulator Drugs Ivacaftor, Lumacaftor, Tezacaftor, and Elexacaftor, Organic

Process Reserach&Development, 2019; 23(2302-2322).
[101] Vertex Pharmaceutcals, Process of producing cycloalkylcarboxamido-indole compounds, WO 2011/133751, 2011
[102] Chakraborti, A.K, et al., 3D-QSAR studies of indole derivatives as phosphodiesterase IV inhibitors, European Journal of Medicinal Chemistry, 2003; 38(11-12):975 – 982.
[103] Ratchanok P, et al., Synthesis, molecular docking, and QSAR study of sulfonamide-based indoles as aromatase inhibitors, European Journal of Medicinal Chemistry, 2018; 143:1604-1615.
[104] Rosa P, et al., Pharmacophore Modeling and 3D-QSAR Study of Indole and Isatin Derivatives as Antiamyloidogenic Agents Targeting Alzheimer's Disease, MDPI molecules, 2020; 25(23): 5773.
[105] Niousha S, et al., QSAR and molecular docking studies of isatin and indole derivatives as SARS 3CLpro inhibitors, BMC Chemistry, 2023; 32.
[106] UniProt Database; https://www.uniprot.org/uniprotkb?query=F508 (accessed 2023).
[106] Reaxys Database; https://www.reaxys.com/#/login (accessed 2023).
[107] Schrodinger Software; https://www.schrodinger.com/ (accessed 2023).
[108] Marvin Database; https://chemicalize.com/app/drawing (accessed 2023).
[109] Altenbach Robert J. et al., Preparation of substituted tricyclic compounds as CFTR modulators, PCT Int. Appl., 2017009804, 2016.
[110] Kavala Güneş, Synthesis of a series of new sulfonamide substituted isoindolo[2,1-a] quinazoline-6,12-dione derivatives and evaluation of their biological activity, Istanbul, Yeditepe University, 2021.

8. CURRICULUM VITAE

Personal Informations

Name	Gülay	**Surname**	Yelken Demirel

Education

Degree	Department	The name of the Institution Graduated From	Graduation year
Doctorate	Pharmaceutical Chemistry	Yeditepe University	2023
Master	Pharmaceutical Chemistry	Gazi University	2008
University	Department of Chemistry	Gazi University	2005
High school	-	Ankara High School	2001

Languages	**Grades (#))**
English	C2
German	A2

All the grades must be listed if there is more than one (KPDS, ÜDS, TOEFL; EELTS vs),

Work Experience (Sort from present to past)

Position	Institute	Duration (Year - Year)
CMC Manager, Formulation and Drug Delivery	BioNTech SE	April 2023-Present
R&D and Patent Manager	Exeltis Pharmaceuticals	August 2018-March 2023
R&D Project Group Executive	Sanovel Pharmaceuticals	March 2014-August 2018
R&D Senior Specialist	Sanovel Pharmaceuticals	April 2009-March 2014
R&D Formulation Scientist	Nobel Pharmaceuticals	2006-2009

Computer Skills

Program	Level
MS Office	Excellent

***Excellent , good, average or basic**

Scientific works

The articles published in the journals indexed by SCI, SSCI, AHCI

Gulay Yelken Demirel, et al., Quality by Design (QbD) Approach for An Orally Disintegrating Tablet Analytical Method Validation, International Journal of Pharmaceutical Sciences and Drug Research, 2018; 10(1):12-19
Gulay Yelken, et al., Synthesis and Analgesic and Anti-inflammatory Activity 6-Substituted-3(2H)-pyridazinone-2-acetyl-2-(p-substituted/nonsubstituted benzal) hydrazone Derivatives, European J. Medicinal Chem., 2009, 44(9), 3760-3764

Articles published in other journals

-

Proceedings presented in international scientific meetings and published in proceedings book.

Gulay Yelken Demirel, Research and Development Activities for Food Supplement Products and Important Points to Consider, Journal of Food and Nutrition, April 2018, Issue 2, p. 24-26
Gulay Yelken Demirel, Fast Dissolving Drug Delivery Systems, Drug Formulation & Bioavailability Congress, September 05-07, 2016, Beijing China, Volume 5, Issue 6, p. 58
Gulay Yelken Demirel, Challenges in Parenteral Formulation Development Studies and an Evaluation from QbD Point of View, International Conference and Expo on Parenterals & Injectables, August 17-19, 2015, Chicago, USA, Volume 3, Issue 3, p. 30
Gulay Yelken Demirel, Fast Dissolving Drug Delivery Systems, 4th International Conference and Exhibition on Pharmaceutics & Novel Drug Delivery Systems, March 24-26, 2014, San Antonio USA, Volume 5, Issue 3, p. 77

Gulay Yelken, et al., Synthesis and Antimicrobial Evaluation of 6-Substituted-3(2h)-Pyridazinone-2-Yl Acetohydrazide Derivatives. International Symposium on Drug Research & Development from Chemistry to Medicine May 17-20, 2007, Antalya, Abstract Book P.76

Journals in the proceedings book of the refereed conference / symposium

-

Others (Projects / Certificates / Rewards)

-