CAPITAL UNIVERSITY OF SCIENCE AND TECHNOLOGY, ISLAMABAD

Study of Anticancerous Compounds of *Nigella sativa* Effective against Breast Cancer

by

Allah Rakha

A thesis submitted in partial fulfillment for the degree of Master of Science

in the

Faculty of Health and Life Sciences
Department of Bioinformatics and Biosciences

2025

Dedicated to my parents, family and my respected teachers who have been a constant source of motivation and encouragement during challenges and supporting me spiritually throughout my life.

CERTIFICATE OF APPROVAL

Study of Anticancerous Compounds of *Nigella sativa* Effective against Breast Cancer

by

Allah Rakha

(MBS231019)

THESIS EXAMINING COMMITTEE

S. No.	Examiner	Name	Organization
(a)	External Examiner	Dr. Shahid Hussain	KU, Murree
(b)	Internal Examiner	Dr. Rizwan ur Rehman	CUST, Islamabad
(c)	Supervisor	Dr. Shaukat Iqbal Malik	CUST, Islamabad

Dr. Shaukat Iqbal Malik
Thesis Supervisor
September, 2025

Dr. Syeda Marriam Bakhtiar
Head
Dept. of Bioinfo. & Biosciences
September, 2025

Dr. Sahar Fazal
Dean
Faculty of Health & Life Sciences
September, 2025

Author's Declaration

I, **Allah Rakha** hereby state that my MS thesis titled "**Study of Anticancerous Compounds of *Nigella sativa* Effective against Breast Cancer**" is my own work and has not been submitted previously by me for taking any degree from Capital University of Science and Technology, Islamabad or anywhere else in the country/abroad.

At any time if my statement is found to be incorrect even after my graduation, the University has the right to withdraw my MS Degree.

(Allah Rakha)
Registration No: MBS231019

Plagiarism Undertaking

I solemnly declare that research work presented in this thesis titled "**Study of Anticancerous Compounds of *Nigella sativa* Effective against Breast Cancer**" is solely my research work with no significant contribution from any other person. Small contribution/help wherever taken has been duly acknowledged and that complete thesis has been written by me.

I understand the zero tolerance policy of the HEC and Capital University of Science and Technology towards plagiarism. Therefore, I as an author of the above titled thesis declare that no portion of my thesis has been plagiarized and any material used as reference is properly referred/cited.

I undertake that if I am found guilty of any formal plagiarism in the above titled thesis even after award of MS Degree, the University reserves the right to withdraw/revoke my MS degree and that HEC and the University have the right to publish my name on the HEC/University website on which names of students are placed who submitted plagiarized work.

(Allah Rakha)
Registration No: MBS231019

Acknowledgement

I humbly thanks to ALLAH ALMIGHTY, the most Merciful and Beneficent who best owed His innumerable blessings upon mankind, one of which is knowledge a distinction for mankind. I offer my gratitude to the Holy Prophet Muhammad (PBUH) who preached us to seek knowledge for the betterment of mankind in particular and other creature in general.

I am deeply indebted to my supervisor Dr. Shaukat Iqbal, Assistant Professor, Department of Bioinformatics and Biosciences, Faculty of Health and Life Sciences, Capital University of Science and Technology Islamabad (CUST), Pakistan. His guidance and continuous encouragement throughout my study have helped me in completion of my thesis. I am grateful to Dr. Erum Dilshad, Assistant Professor, Department of Bioinformatics and Biosciences, Faculty of Health and Life Sciences, Capital University of Science and Technology Islamabad (CUST), Pakistan for her kind cooperation to provide research facilities.

I specially thanks to Muhammad Maaz who helped me in installation and under standing of different softwares.

(Allah Rakha)

Abstract

In today's environment, where stress, toxic waste, and radiation are prevalent, humans have become more vulnerable to various diseases. Researchers are actively searching for therapeutic compounds that can either treat or reduce the onset of these diseases. *Nigella sativa*, a plant from the Ranunculaceae family, has a long history as a medicinal herb. The metabolites found in *Nigella sativa* seeds exhibit promising therapeutic activities for the immune, cardiovascular, endocrine, and respiratory systems, in addition to having antioxidant, anticancer, and anti-ulcer properties. The identification of natural, plant-based, and non-toxic anticancer agents is crucial for the treatment of various types of cancer. In the current study, Topoisomerase I and Topoisomerase II α were chosen as target proteins, while ligands selected for the study included α-pinene, anethol, myristic acid, nigeglaine, nigellaquinomine, nigellicine, palmitic acid, pyrogallol, salfredin B11, and salicylic acid. Molecular docking was employed to evaluate the bond strength between a ligand and its target protein using a specialized scoring function. It also helped determine the optimal orientation of the ligand within the binding site. The docking process utilized the 3 Dimensional structures of target proteins in Protein Data Bank format and the ligand structures in Structure Data format. The best ligand was selected based on the highest docking score, LogP value, hydrogen bond acceptors, hydrogen bond donors, and molecular weight. Salfredin B11 emerged as the lead compound, demonstrating the best docking score, hydrogen bonding, and pharmacokinetic properties compared to the other ligands. Sorafenib, an Food and Drug Administration-approved drug for cancer, was used as a reference for comparison. By evaluating both compounds in terms of ADMET and physicochemical properties, it was concluded that Salfredin B11 exhibited superior activity and safety compared to Sorafenib. This promising active compound, Salfredin B11, identified from *Nigella sativa*, holds potential as a future medicinal agent. To validate its drug potential, it could undergo experimental testing in mice, and, following successful trials, proceed to clinical studies for further confirmation.

Contents

List of Figures

List of Tables

Abbreviations

ADMET	Absorption, Distribution, Metabolism, Excretion, Toxicology
ADMET	Absorption, distribution, metabolism, excretion and toxicity
BBB	Brain Blood Barrier
CL-Tot	Total clearance
DCIS	Ductal Carcinoma In-Situ
FDA	Food and Drug Administration
GC-MS	Gas Chromatography – Mass Spectrometery
GRAVY	Grand average of htdropathycity
HCC	Hepatocellular carcinoma
hERG	Human ether-á-go-go-related gene
MRTD	Maximum recommended tolerated dose
N. sativa	Nigella sativa
OCT2	Organic Cation Transporter 2
PARP	Poly ADP- ribose polymerase
PDB	Protein Databank
SDF	Structure Data Format
WHO	World Health Organization

Chapter 1

Introduction

Nigella sativa, a member of the Ranunculaceae, is a historically renowned medicinal plant. It has been mentioned in various texts. The seeds of this plant are referred to as "Kalonji" in South Asia, "Habbat us Sauda" in the Middle East, and are commonly known in English as "black cumin." [1, 2]. The metabolites found in *N. sativa* seeds exhibit promising therapeutic potential for the immune, endocrine, cardiovascular, and respiratory systems [3]. Because thymoquinone is present in the essential oil of these seeds, several of *N. sativa*'s actions are taken into consideration. [4]. *N. sativa* seeds are mostly used in different food dishes as condiment due to their specific aroma and bitter taste like peppers [5]. The leaves of the green plant *N. sativa* are finely split. Its flowers have five to ten petals and come in a variety of colors, including white and pastel blue. The fruit resembles a capsule that is further separated into three follicles, each of which bears oval, black seeds [6].

Cancer is the term used for the disease which comprises of uncontrolled and abnormal division of cells and they may invade other body cells through blood or lymph. It has become the second major cause of death in the world and over all cancer prevalence has raised [7]. Cancer is the dangerous health problem affecting human beings. At the tissue level it is a variety disease and this variety is huge challenge for the specific diagnosis followed by the treatment [8, 9].

Highest cancer type percentages in men occur in prostate, lung, bronchus, colon, rectum and urinary bladder respectively. While in women highest cancer prevalence is in breast, lung, bronchus, colon, rectum, uterine corpus and thyroid. This information shows that breast and prostate cancer are the major types of cancer found in women and men respectively [10]. In children highest percentage of cancer type is blood cancer, cancer related to brain and lymph nodes cancer respectively [11, 12].

The question "What causes cancer" got the attention of people over generations. There are different causes of cancer which develop different type of cancer.

WHO organized an international symposium in 1950, where the participants were intrigued by the variations in the variety of cancer found in the world [13]. It comes to know that people who travelled to other countries developed the type of cancer common to their country of stay rather than their native country. It showed that most types of cancer are caused by exposures to the environment instead of inherited genes [14].

Cancer is mainly caused by a series of mutations in the genes and these mutations alter the cell function. Chemical compounds are responsible for gene mutations and production of cancer cells. It is proved that Smoking leads to lung cancer because it contain many carcinogens [15]. Environmental chemical substances having carcinogenic properties effect the nucleus and cytoplasm of the cells directly and indirectly leading to gene mutations [16–18]. Bacteria, viruses and radiations contribute for 7% of all type of cancers [19]. Generally cancer breaks the cellular relations leading to dysfunction of vital genes. This disturbance effect the cell cycle which result in abnormal proliferation [20, 21]. Proto-oncogenes which are essential for normal functioning of cell become oncogenes during mutations and this is very dangerous for the cell [22]. If there is no tumor suppressor gene it will trigger uncontrolled and abnormal cell division [23].

Breast cancer incorporates complex molecular mechanisms. Because of the complicated molecular variations, treating breast cancer with radiation treatment or chemotherapy therapy is challenging and often has unintended side effects. For

centuries, plants and their extracts have been utilized to address various ailments, including breast cancer. Herbal remedies are a reliable option for treating cancer due to their minimal toxicity. Additionally, women with breast cancer readily embrace herbal solutions due to their accessibility and affordability. In recent years, numerous plants and their compounds have displayed promising anti-cancer properties against breast cancer cells in both lab and animal studies. However, their efficacy in treating breast cancer remains uncertain due to the absence of randomized clinical trials [24].

Breast cancer exhibits a varied molecular profile. Key molecular features of the disease include the activation of human epidermal growth factor receptor 2 (HER2, encoded by *ERBB2*), activation of estrogen and progesterone receptors, and/or mutations in the BRCA genes. Depending on the molecular subtype, different treatment plans apply. Treatment for breast cancer is interdisciplinary and consists of both systemic therapy and a loco- regional strategy (surgery and radiation therapy). Systemic therapies include hormone therapy for hormone receptor–positive disease, chemotherapy, anti-HER2 treatments for HER2-positive cases, bone-stabilizing agents, poly (ADP-ribose) polymerase (PARP) inhibitors for patients with BRCA mutations, and more recently, immunotherapy. Future strategies for breast cancer treatment aim to minimize complications, enhance effectiveness based on tumor biology and early treatment response, and personalize therapy for each individual patient [25].

Breast cancer development frequently involves the activation of BRACA genes. They are important tumor suppressor genes, which suggest that through controlling cell division and mending broken DNA, they contribute in the prevention of cancer. The cells have capacity to correctly repair DNA may be compromised by mutations in BRACA1 and BRACA2, which increases the likelihood of mutations that could result in cancer. It is possible to inherit these gene mutations from one's parents. Those who carry these mutations are at an increased risk of developing metastases. By identifying these changes through genetic testing, people can assess their risk of developing cancer and decide on preventative and early detection measures like intensified surveillance, risk-reducing operations, or tailored

medicines. Knowing the presence of BRACA mutations in breast cancer patients can guide treatment decisions. Some treatments like PARP inhibitors have been developed specifically for breast cancer with BRCA mutations.

1.1 Problem Statement

Breast cancer is a leading global health concern, with conventional chemotherapy often causing serious side effects. This study explores the potential of *Nigella sativa*-derived bioactive compounds as alternative treatments, using computational modeling to evaluate their efficacy against breast cancer.

1.2 Aim and Objectives

The main aim of this research study was to identify potential compounds using molecular docking of *Nigella sativa* to deal with breast cancer which is a serious threat to life.

- To find the potential compound of *Nigella sativa* having anti-cancerous properties in silico.
- To study the interaction between ligand and protein complex through molecular docking.

Chapter 2

Literature Review

2.1 *Nigella sativa*

N. sativa has used as spices as well as flavoring agent in different food preparations i.e. pickles, sauces, salad etc. It has long been used in Europe, Arabian countries and Africa as traditional remedy [26]. The earlier herbalists consider the *N. sativa* as "The herb from heaven" [27]. The Holy Prophet (PBUH) spoke of the healing properties of black seeds in His Hadith, saying, "'continue to use this black seed, for it is a cure for every disease except death." [28]. Avicenna has recommended the use of black seeds to enhance the body energy in his book "The Canon of Medicine" [29]. The morphology of *Nigella sativa* plant and seeds is shown in Figure 2.1. Taxonomic classification of this plant is given in Table 2.1.

FIGURE 2.1: Morphology of *Nigella sativa* seeds and plant [30]

TABLE 2.1: Taxonomic Classification of *Nigella sativa*

Sr.No.	Taxonomic Rank	Eukarya
1	Kingdom	Plantae
2	Subkingdom	Tracheobionata (Vascular plant)
3	Division	Spermatophyta
4	Class	Magnoliophyta
5	Subclass	Magnolidae
6	Order	Ranunculales
7	Family	Ranunculaceae
8	Genus	Nigella
9	Species	Sativa

2.2 Cultivation and Collection of *Nigella sativa*

N. sativa is cultivated all over the world and it is grown as annual herb in Pakistan and India. It is grown as the same way like wheat during winter season. Land from where crops of green grams, black grams or corn are harvested can be used for its cultivation. Before cultivation 2-3 times plouging is enough to get good yield and to stop the weeds growth in the field. Light soils can be prepared more easily as compared to heavy soils as they required more plouging. Distance between the seeds should be 30cm and to avoid the delayed germination, there should not be deep sowing of seeds. 12-15 kg per hectare seed is required. 3-5 irrigations are enough for the crop at different stages till its harvesting. Crop of *Nigella sativa* matures in April and May. When fruit turns yellowish, crop is harvested and it should be done early in the morning because late harvesting may leads to the scattering of seeds. After harvesting process and drying, crop is threshed by proper thresher or by tractor. After this threshing step seeds should be stored in the proper bags [32].

According to WHO more than 3/4th communities of less resource countries depend on medicinal plants for their basic health care because they are unable to buy or

do not have access to the allopathic medicines [33, 34]. Now a day with the development of optimum nutrition; a new interest is being developed to use plants as food as well as medicine [35, 36]. Recently there is increased use of plant derive medicines not of due to easy access but the belief that phyto-medicines have less side effects as compared to the synthetic medicines [37]. It is reported that about 300,000 herbal species exist in the world but only 15% of these species are examined for their pharmacological activity [38]. *N. sativa* is considered as the much nutritious herb among all herbs around the world and different scientific researchers are going on to validate its traditional claim of uses [39, 40].

2.3 Nutritional Value of *Nigella sativa*

Different researchers reported its nutritional values as 38% of fats, 32% of total carbohydrates, 7-94% of fiber and 20-85% of protein. Different amino acids found in *Nigella sativa* are aspartate, glutamate and arginine whereas methionine and cysteine are the minor and major amino acids respectively. It also contains significant levels of copper, iron, zinc, phosphorus, thiamin, niacin, calcium and folic acid [41, 42].

2.4 Chemical Composition

As this plant is widely used in food as well as medicine, it is extensively studied to analyze the phytochemicals found in it. *N. sativa* seeds composed of different chemicals i.e. fixed oil, alkaloid, saponin, essential oil and proteins. They contain 0.4-2.5% essential oils and 28-36% fixed oil. Main component of fixed oil is unsaturated fatty acids which include arachidonic, linoleic, eicosadienoic and linolenic acid. Stearic, myristic and palmitic acid are the components of saturated fatty acid of fixed oil [43]. Essential oil of *N. sativa* seeds was investigated through gas Chromatography-mass spectrometery (GC-MS). There are many compounds in the oil but pharmacologically most active component found in volatile oil are dithy-

moquinone, thymoquinone, thymol and thymohydroquinone. Chemical structures of these compounds are shown in Figure 2.2 [44, 45].

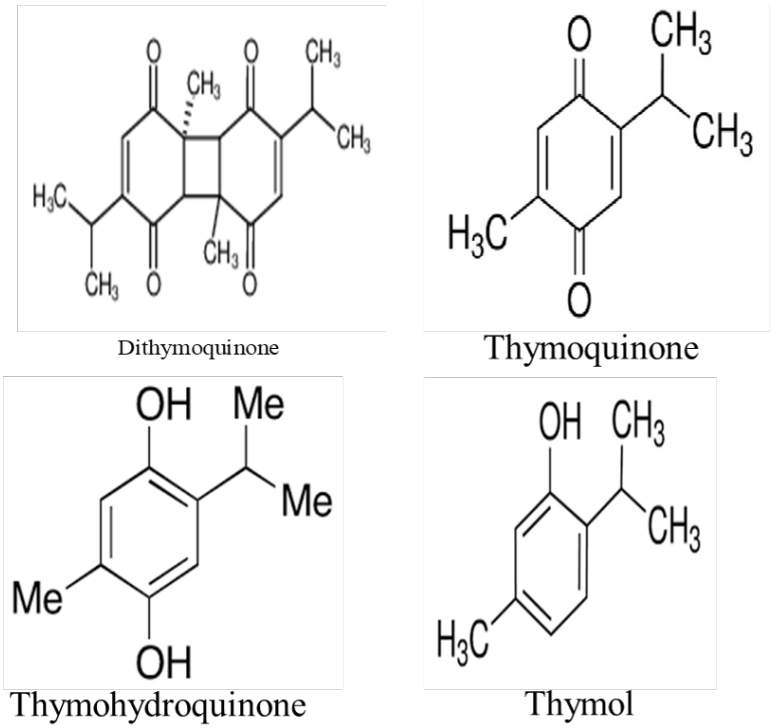


FIGURE 2.2: Chemical structure of active components of *Nigella sativa* [46]

In addition to these active compounds some other chemical compounds are also found in the *Nigella sativa* seeds which have different active components in them performing some specific function in the body when used as a remedy. Summary of these compounds is described in Table 2.2.

TABLE 2.2: Phytochemical Composition

Component	**Type**	**Constituents**
Fixed oil (32–40%)	Unsaturated fatty acids [47]	Arachidonic, eicosadienoic, linoleic, linolenic, oleic, palmitoleic, palmitic, stearic, and myristic acids
	Sterols	Beta-sitosterol, cycloeucalenol, cycloartenol, sterol esters, and sterol glucosides
Volatile oil (0.4–0.45%)	Saturated fatty acids [49, 50]	Thymoquinone, nigellone, thymol, carvacrol, dithymoquinone, thymohydroquinone, *p*-cymene, *d*-limonene, α- and β-pinene, *d*-citronellol, and 2-(2-methoxypropyl)-5-methyl-1,4-benzenediol
Alkaloids [51]	-	Nigellicine, nigellidine, nigellimine-N-oxide

continued on next page

Table 2.2 continued from previous page

Component	Type	Constituents
Coumarins [51]	-	6-methoxy-coumarin, 7-oxy-coumarin, 7-hydroxy-coumarin
Saponins [52]	Triterpenes and Steroidal	Alpha-Hedrin, steryl-glucosides, acetyl-steryl-glucoside
Minerals (1.79–3.74%) [53]	-	Calcium, phosphorous, potassium, sodium, and iron
Carbohydrates (33.9%) Fiber (5.5%), Water (6%) [54]	-	-

2.5 Pharmacological Activities of *N. sativa*

There are so many pharmacological activities of *N. sativa* of which some are being discussed here.

2.5.1 Antioxidant Activity

Free radicals may be produced in many human diseases. *N. sativa* has antioxidant property and to find the antioxidant compounds in the essential oil of *N. sativa*, it was tested. Essential oil along with other components of oil performs respectable radical scavenging property [67].

2.5.2 Anti-cancer Activity

Research shows that methanolic *Nigella sativa* seeds extract has strong cytotoxic effects on Elrich ascites carcinoma, sarcoma 180 and Daltons ascites lymphoma with least cytotoxic effects on normal lymphocytes [68]. Other research shows that alcoholic and aqueous extracts of *N.sativa* in pure form or along with hydrogen

per oxide, as an oxidative stressor were found useful for stoping breast cancer cells [69].

The cytotoxic action of *Nigella sativa* seeds was performed on human hepatoma HepG2 cell lines and then incubated for 24 hours with distinct quantities of the *N.sativa* extract [70]. Oral administration of thymoquinone was found useful to increase the actions of quinine reductase and glutathione transferase and acts as prophylactic source for the toxicity produced in chemical carcinogenesis and hepatic cancer [71].

2.5.3 Anti-ulcer Activity

Ulcer caused due to Aspirin can be reduced by the extract of *N. sativa* seeds by 36% [49]. Recent clinical studies support the use *Nigella sativa* for the eradication of *Helicobactor pylori* in non-ulcer dyspepsia patients [72].

2.5.4 Effects on Circulatory System

Nigella sativa is recommended alone or along with honey or garlic for treating hypertension. El-Tahir et al. [73] investigated the activity of volatile oil of *Nigella sativa* along its active compound; the thymoquinone on the heart of anesthetizes rats and arterial blood pressure.

2.5.5 Effects on Reproductive System

2 months administration of *N .sativa* seeds enhanced the sperm motility, sperm count in the testicular ducts and caudal epidydimus as well as increase in the weight of reproductive organs. Spermatogenesis increased at primary and secondary spermatocytes hence curing the male infertility [74, 75].

2.6 Breast Cancer

Breast cancer is a prevalent malignancy that affects both men and women globally. This occurs when aberrant cells in the breast tissue proliferate out of control, resulting in a tumor that has the ability to infiltrate neighboring tissue and maybe spread to other body areas. The milk ducts, also known as lobules, are the organs of the breast that produce and distribute milk. This kind of cancer frequently starts in these structures. Although the precise causes of breast cancer are still unknown, a number of risk factors have been found. Gender is significant because women have higher levels of estrogen and progesterone than men, which puts them at greater risk. Certain types of breast cancer cells may develop more quickly as a result of these hormones.

Age is another important factor as the risk of developing breast cancer increases with age. Family history and genetics also play a role; people whose close relatives have breast cancer are at increased risk of developing the disease, and some genetic mutations are associated with a significantly higher risk [76].

Early diagnosis is the key to improving the prognosis of breast cancer. A regular breast self-exam and a clinical breast exam by a healthcare professional can help detect any unusual changes in your breasts. Mammograms, or x-rays of breast tissue, are useful tools for early diagonosis of breast cancer, even before physical symptoms appear [77].

2.6.1 Types of Breast Cancer

Breast cancer is not a single disease but encompasses several different types, each with distinct characteristics.

The two primary categories of breast cancer are invasive and noninvasive. Within these categories, there are various subtypes of breast cancer. Some of the most common types include:

- Ductal Carcinoma in Situ (DCIS): This is a non-invasive type of cancer in which abnormal cells are located within the lining of a breast duct. It is regarded as the earliest stage of breast cancer.
- Invasive Ductal Carcinoma (IDC): This is the most common type of invasive breast cancer. It begins in the milk ducts and then invades surrounding tissues in the breast.
- Invasive Lobular Carcinoma (ILC): ILC starts in the milk-producing glands (lobules) and can also spread to nearby tissues.
- Triple-Negative Breast Cancer: This is a subtype of breast cancer that lacks estrogen receptors (ER), progesterone receptors (PR), and HER2/Neu receptors.
- HER2Positive Breast Cancer: Some breast cancers over express the HER2/Neu protein, making them more aggressive. Targeted therapies like Herceptin are used to treat this subtype.
- Luminal A and Luminal B: These are subtypes based on presence of hormone receptors (ER and PR). Luminal A cancers are often low-grade, while Luminal B cancers tend to be more aggressive.
- Inflammatory Breast Cancer: This uncommon and aggressive form of breast cancer is marked by breast redness and swelling, and is often misdiagnosed as an infection.
- Phyllodes Tumours: These are rare tumors that develop in the stroma (connective tissue) of the breast.

Male Breast Cancer: Breast cancer is more common in women; men can also develop breast cancer. It can be of various types similar to those in women.

- Metastatic Breast Cancer: This refers to breast cancer that has spread to other parts of the body, typically the bones, liver, lungs, or brain. It can be any subtype of breast cancer [76][78].

2.7 Available Treatment

To diagnose breast cancer, a biopsy—in which a small tissue sample is taken from the suspicious area and examined under a microscope—is essential, along with imaging tests such as mammography and ultrasound. Once diagnosed, breast cancer is classified into different stages based on the tumor's size, how far it has spread, and involvement of nearby lymph nodes or organs. This staging system helps guide appropriate treatment decisions.

Treatment choices depend on several factors, including the type and stage of the cancer, as well as the patient's overall health. One common approach is surgery, which may involve:

- Lumpectomy, the removal of the tumor along with a small margin of surrounding tissue, or
- Mastectomy, the complete removal of the breast.

Following surgery, radiation therapy is often used to destroy any remaining cancer cells through high-energy radiation. In addition, systemic treatments such as chemotherapy and targeted therapy are employed to address cancer cells that may have spread throughout the body [79].

Hormone therapy, particularly for hormone receptor-positive breast cancers, blocks hormones that fuel cancer growth. In recent years, targeted therapies have shown significant promise in treating certain types of breast cancer. HER2-positive breast cancers, for instance, over express the HER2 protein, which promotes cancer growth. Targeted therapies like Herceptin specifically target HER2-positive cells, improving treatment outcomes and reducing side effects [76].

Treatment for breast cancer is very individualized and is based on the patient's choices, general health, and the type and stage of her cancer. Variety of therapies are typically used in combination for treatment, including hormone therapy, radiation therapy, chemotherapy, targeted therapy, and surgery. Chemotherapy

is a popular treatment for breast cancer that is frequently combined with radiation, surgery, or other forms of treatment. To target and eradicate cancer cells, potent medications are used. Depending on following conditions, overall health, the stage and type of your breast cancer, and your personalized treatment plan, your particular chemotherapy regimen may change. Chemotherapy medications for breast cancer often consist of cyclophosphamide, Texans (like paclitaxel), and anthracyclines (like doxorubicin).

Treatment cycles may vary in duration and frequency and may cause side effects such as hair loss, nausea, fatigue and reduced blood cell counts [80][77]. Precision medicine has gained importance in recent years, tailoring treatments based on the individual genetic and molecular profiles of tumors. This approach is particularly effective in certain types of cancer, such as HER2-positive breast cancer, where targeted therapies specifically target the overactive HER2 protein.

Immunotherapy, another innovative treatment, uses the immune system to recognize and attack cancer cells. As our knowledge of the genetic and molecular mechanisms of cancer increases, the development of personalized and targeted therapies offers hope for more effective and less invasive treatment strategies for various cancer types [76].

2.8 Medicinal Plants

Plant extracts, such as Taxol derived from the Pacific yew tree, have shown promise in treating breast cancer by disrupting cancer cell division [77]. However, rigorous clinical trials are needed to establish the safety and efficacy of botanical treatments. *Nigella sativa* is a medicinal plant with potential bioactive compounds.

2.9 Proteins Found in Breast Cancer Cells

Breast cancer cells can contain a wide range of proteins, and the specific proteins found in individual people can vary depending on the type of breast cancer.

According to immune-histochemical studies and molecular tests, some proteins commonly associated with breast cancer are human epidermal growth receptor 2 (HER2), estrogen receptor (ER) and progesterone receptor (PR), BRACA1 and BRACA2, Ki-67, p53 and cyclin D1 [81]. Table 2.3 indicates the mutated genes in breast cancer and their reported inhibitors.

TABLE 2.3: Mutants and Their Inhibitors

MUTANT	**INHIBITOR**
HER2	Trastuzumab (Herceptin)
	Pertuzumab (Perjeta, monoclonal antibodies)
ER	Tamoxifen, Fulvestrant
PR	No possible inhibitors BRCA1
BRCA2	PARP (Poly ADP-ribose Polymerase) inhibitors: Olaparib,
	Rucaparib
Ki-67	No specific inhibitors known
P53	Nutlin-3
Cyclin D1	CDK4/6 Inhibitors:
	Palbociclib, Ribociclib, Abemaciclib

Chapter 3

Procedure and Methodology

3.1 Selection of Disease

Breast cancer is a complex and heterogeneous disease defined by a range of molecular features, including the overexpression of human epidermal growth factor receptor 2 (HER2, encoded by *ERBB2*), activation of hormone receptors (estrogen and progesterone), and/or mutations in the BRCA genes. These molecular subtypes influence distinct therapeutic strategies. This study focuses on breast cancer due to its growing incidence within South Asian populations [76]. Deep knowledge of molecular intricacies of breast cancer is paramount for developing targeted therapies tailored to individual patients' needs.

3.2 Selection of Proteins

In the context of breast cancer, numerous proteins with reported mutations were meticulously chosen for further investigation and analysis. Notable examples include BRACA1, ER, PR, and Ki67 [16][17]. The selection criteria were based on existing literature reports documenting their involvement in breast cancer pathology. These proteins play crucial roles in cell growth regulation, hormone signaling,

and proliferation, making them key targets for therapeutic interventions. Comprehensively analyzing the molecular alterations in these proteins, we aim to uncover potential biomarkers and therapeutic targets for breast cancer management. By comprehensively analyzing the molecular alterations in these proteins, we aim to uncover potential biomarkers and therapeutic targets for breast cancer management.

3.3 Cleaning of the Downloaded Protein

Extra constituents attached to the protein have to be removed after downloading protein structure. This was performed by using an open source system Pymol. Linear chain consisting of 1-306 amino acids was kept and referred as A-chain. All other constituents were removed so further process can be done effectively [81].

3.4 Determination of Functional Domains of Target Proteins

To determine the functional domains of the target proteins, InterPro database was used. It can analyze the protein and provides protein information regarding functional sites, families and domains of the protein of interest [83].

3.5 Selection of Active Metabolic Ligands

Those active ligands were selected which have shown some anticancerous properties in the past. These active ligands were α-pinene, anethol, myristic acid, nigeglaine, nigellaquinomine, nigellicine, palmitic acid, pyrogallol, salfredin-B11 and salicylic acid [84].

3.6 Ligand Preparation

3-dimensional structures of the selected ligands were downloaded using PubChem database which is running under National Centre of Biotechnology Information (NCBI). It contains the information of chemical molecules which is stored with reference to their name, simple or 3-dimensional structures, their isomers, molecular formulas, canonical smiles and activities of molecules against the biological assays [84]. Structures of the ligands were searched from PubChem, downloaded and MM2 energy minimization was done using Chem3D ultra. When energy was minimized we selected the sdf format at the end to save the energy minimized structure of the ligand.

3.7 Molecular Docking

CB-dock (Cavity detection guided blind docking) is used to perform the molecular docking between ligands and the target protein. CB-dock automatically finds the sites of docking. It is a method of protein and ligand docking which indicates about the sites of bonding, the size and the center is calculated. Docking was performed by adjusting the box size according to the ligand. As it is cavity binding focused docking so ratio of accuracy is higher [85]. To perform docking we uploaded 3D structure of ligand in sdf format and of target protein in pdb format [86].

3.8 Visualization of Result via PyMol

Pymol has emerged as an efficient visualization molecular tool over the past few years. Its graphics and ability to visualize 3D structures is extraordinary [87]. The result of docking can be captured and somehow Pymol provides a plugin which have access to the result as well as make their visualization more clear so that docking result can be studied easily [88]. For whole process docking results were saved in pdb format and after visualization in the Pymol were also being saved in the pdb file format.

3.9 Analysis of Docked Complex via Ligplot

Once we got the docked complex with the minimum vina score, the next step was the analysis of the docked complex. This analysis was done by the software LigPlot. For the given pdb file format the schematic diagrams of the protein and ligand interactions were generated automatically. These interactions were modified by hydrogen bonds and through hydrophobic contacts. LigPlot provided the analysis of the hydrophobic and hydrogen bonding interactions. LigPlot plus generated the 2D representation of the protein-ligand complex [89].

3.10 Ligand ADMET Properties

As analysis had been done the next step was the study of pharmacokinetic and toxicity properties. The weak ligands of the drug were eliminated during preclinical ADMET. The remaining ligands were selected as potential drugs against the disease. Optimization of the ADMET which is absorption, distribution, metabolism, excretion and toxicity related to human body were done using the PkCSM online server [90].

3.11 Lead Compound Identification

Out of these potential ligands, lead compound was identified on the basis of Lipinski's rule of 5's and pharmacokinetic properties.

1. The molecular weight should also be lesser than 500.
2. Hydrogen bond acceptors maximum number should be 10.
3. Hydrogen bond donors' maximum number should be 5.
4. The log value of the drug-like compound must be limited to 5 [91].

3.12 Comparison with the Standard Drug

Sorafenib drug which has shown anti-tumor properties against Breast cancer caused due to ER, PR mutations, has been selected as a standard drug for comparison against the lead compound. Olaparib drug could be selected in case of breast cancer developed due to mutations in BRACA1 genes [65].

Sorafenib is a selective estrogen receptor (recurrence) modulator used in early recurrence of breast cancer in women or men after surgery to reduce the risk of hormone receptor positivity. Selective Estrogen Receptor Modulator (SERM) for the treatment of advanced stage hormone receptor positive breast cancer in women or men. To reduce the risk of invasive breast cancer in women diagnosed with hormone receptor-positive DCIS (ductal carcinoma in situ) and to reduce the risk of breast cancer in women with above-average risk of undiagnosed disease after surgery [34]. Physicochemical properties and molecular binding results determine drug candidate decisions.

Sorafenib undergoes oxidative metabolism by CYP3A4 in the liver, as well as glucuronidation by UGT1A9 in the liver and kidneys. At steady-state, Sorafenib accounts for 70-85% of the circulating analytes in plasma. About eight metabolites of Sorafenib have been identified, of which five were detected in plasma. The main circulating metabolite was the pyridine N-oxide form, which comprises approximately 9–16% of the total circulating dose at steady-state. [66].

Chapter 4

Results and Discussion

4.1 Structure Modeling

4.1.1 Acquisition of Primary Sequence

The primary sequences of the target proteins, Topoisomerase I and Topoisomerase IIα, were retrieved in FASTA format from the UniProt database (www.uniprot.org) using accession numbers P11387 and P11388, respectively. The sequences contain 764 and 1531 amino acid residues

>sp P11387 TOP1 HUMAN DNA Topoisomerase 1 OS=Homo sapiens OX=9606 GN=TOP1 PE=1 SV=2

```
MSGDHLHNDSQIEADFRLNDSHKHKDKHKDREHRHKEHKKEKDREKSKHSNSEHKDSEKK
HKEKEKTKHKDGSSEKHKDKHKDRDKEKRKEEKVRASGDAKIKKEKENGFSSPPQIKDEP
EDDGYFVPPKEDIKPLKRPRDEDDADYKPKKIKTEDTKKEKKRKLEEEEDGKLKKPKNKD
KDKKVPEPDNKKKKPKKEEEQKWKWWEEERYPEGIKWKFLEHKGPVFAPPYEPLPENVKF
YYDGKVMKLSPKAEEVATFFAKMLDHEYTTKEIFRKNFFKDWRKEMTNEEKNIITNLSKC
DFTQMSQYFKAQTEARKQMSKEEKLKIKEENEKLLKEYGFCIMDNHKERIANFKIEPPGL
FRGRGNHPKMGMLKRRIMPEDIIINCSKDAKVPSPPPGHKWKEVRHDNKVTWLVSWTENI
QGSIKYIMLNPSSRIKGEKDWQKYETARRLKKCVDKIRNQYREDWKSKEMKVRQRAVALY
FIDKLALRAGNEKEEGETADTVGCCSLRVEHINLHPELDGQEYVVEFDFLGKDSIRYYNK
VPVEKRVFKNLQLFMENKQPEDDLFDRLNTGILNKHLQDLMEGLTAKVFRTYNASITLQQ
QLKELTAPDENIPAKILSYNRANRAVAILCNHQRAPPKTFEKSMMNLQTKIDAKKEQLAD
ARRDLKSAKADAKVMKDAKTKKVVESKKKAVQRLEEQLMKLEVQATDREENKQIALGTSK
LNYLDPRITVAWCKKWGVPIEKIYNKTQREKFAWAIDMADEDYEF
```

FIGURE 4.1: Protein sequence Topoisomerase I

>sp|P11388|TOP2A HUMAN DNA topoisomerase 2-alpha OS=Homo sapiens OX=9606 GN=TOP2A PE=1 SV=3

MEVSPLQPVNENMQVNKIKKNEDAKKRLSVERIYQKKTQLEHILLRPDTYIGSVELVTQQ
MWVYDEDVGINYREVTFVPGLYKIFDEILVNAADNKQRDPKMSCIRVTIDPENNLISIWN
NGKGIPVVEHKVEKMYVPALIFGQLLTSSNYDDDEKKVTGGRNGYGAKLCNIFSTKFTVE
TASREYKKMFKQTWMDNMGRAGEMELKPFNGEDYTCITFQPDLSKFKMQSLDKDIVALMV
RRAYDIAGSTKDVKVFLNGNKLPVKGFRSYVDMYLKDKLDETGNSLKVIHEQVNHRWEVC
LTMSEKGFQQISFVNSIATSKGGRHVDYVADQIVTKLVDVVKKKNKGGVAVKAHQVKNHM
WIFVNALIENPTFDSQTKENMTLQPKSFGSTCQLSEKFIKAAIGCGIVESILNWVKFKAQ
VQLNKKCSAVKHNRIKGIPKLDDANDAGGRNSTECTLILTEGDSAKTLAVSGLGVVGRDK
YGVFPLRGKILNVREASHKQIMENAEINNIIKIVGLQYKKNYEDEDSLKTLRYGKIMIMT
DQDQDGSHIKGLLINFIHHNWPSLLRHRFLEEFITPIVKVSKNKQEMAFYSLPEFEEWKS
STPNHKKWKVKYYKGLGTSTSKEAKEYFADMKRHRIQFKYSGPEDDAAISLAFSKKQIDD
RKEWLTNFMEDRRQRKLLGLPEDYLYGQTTTYLTYNDFINKELILFSNSDNERSIPSMVD
GLKPGQRKVLFTCFKRNDKREVKVAQLAGSVAEMSSYHHGEMSLMMTIINLAQNFVGSNN
LNLLQPIGQFGTRLHGGKDSASPRYIFTMLSSLARLLFPPKDDHTLKFLYDDNQRVEPEW
YIPIIPMVLINGAEGIGTGWSCKIPNFDVREIVNNIRRLMDGEEPLPMLPSYKNFKGTIE
ELAPNQYVISGEVAILNSTTIEISELPVRTWTQTYKEQVLEPMLNGTEKTPPLITDYREY
HTDTTVKFVVKMTEEKLAEAERVGLHKVFKLQTSLTCNSMVLFDHVGCLKKYDTVLDILR
DFFELRLKYYGLRKEWLLGMLGAESAKLNNQARFILEKIDGKIIIENKPKKELIKVLIQR
GYDSDPVKAWKEAQQKVPDEEENEESDNEKETEKSDSVTDSGPTFNYLLDMPLWYLTKEK
KDELCRLRNEKEQELDTLKRKSPSDLWKEDLATFIEELEAVEAKEKQDEQVGLPGKGGKA
KGKKTQMAEVLPSPRGQRVIPRITIEMKAEAEKKNKKKIKNENTEGSPQEDGVELEGLKQ
RLEKKQKREPGTKTKKQTTLAFKPIKKGKKRNPWSDSESDRSSDESNFDVPPRETEPRRA
ATKTKFTMDLDSDEDFSDFDEKTDDEDFVPSDASPPKTKTSPKLSNKELKPQKSVVSDLE
ADDVKGSVPLSSSPPATHFPDETEITNPVPKKNVTVKKTAAKSQSSTSTTGAKKRAAPKG
TKRDPALNSGVSQKPDPAKTKNRRKRKPSTSDDSDSNFEKIVSKAVTSKKSKGESDDFHM
DFDSAVAPRAKSVRAKKPIKYLEESDEDDLF

FIGURE 4.2: Protein sequence Topoisomerase IIα

Topoisomerase I and Topoisomerase II α were selected as the target proteins and α-pinene, anethol, myristic acid, nigeglaine, nigellaquinomine, nigellicine, palmitic acid, pyrogallol, salfredin B11 and salicylic acid were selected as ligands for the current study.

4.2 Physiochemical Characterization of Topioisomerase I and Topioisomerase II α

ProtParam is an online tool which let on the calculation of different physical & chemical properties of any given protein stored in any protein database (Swiss-Prot, TrEMBL) or for protein sequence entered by the user. ProtParam calculates several physicochemical parameters, including theoretical pH, molecular weight, composition of amino acid (both positively and negatively charged residues), estimated half-life, extinction coefficient, aliphatic index, instability index, , and the grand average of hydropathicity (GRAVY). A theoretical pH value above 7 indicates that the protein is basic in nature, while a value below 7 suggests it is

acidic. Light absorption is represented by extinction coefficient. Instability index if less than 40 indicates the stability of the protein while greater than 40 indicates the instability of protein. The aliphatic index reflects the aliphatic side chain content of a protein, with higher values indicating greater thermostability. Molecular weight accounts for the presence of both positively and negatively charged residues in the protein. A lower GRAVY (Grand Average of Hydropathicity) value suggests stronger interaction with water molecules. The physicochemical properties of Topoisomerase I and Topoisomerase IIα are presented in Table 4.1 and 4.2.

TABLE 4.1: Physiochemical Properties of Topoisomerase I

Sr.No.	**Parameters**	**Computed values**
1	Molecular weight	90594.53
2	Theoretical pI	9.33
3	Positively charged Residues	176
4	Negatively charged Residues	142
5	Extinction coefficient 1	103290
6	Extinction coefficient 2	102790
7	Estimated half-life	1.9 hrs (Mammals, in vitro) >20 hours (Yeast, in vivo) >10 hours (E. coli, in vivo)
8	Instability index	45.21
9	Aliphatic index	58.99
10	Grand average of hydropathicity (GRAVY)	-1.295

Topoisomerase I has theoretical pH greater than 7 which represents that it is basic in nature, Instability index is >40 represents protein instability. It has low GRAVY values showing better interactions with water molecules.

TABLE 4.2: Physiochemical Properties of Topoisomerase II α

Sr.	**Parameters**	**Computed values**
1	Molecular weight	174385.1
2	Theoretical pI	8.82
3	Positively charged Residues	246
4	Negatively charged Residues	226

continued on next page

Table 4.2 continued from previous page

Sr.	Parameters	Computed values
5	Extinction coefficient 1	163820
6	Extinction coefficient 2	163070
7	Estimated half-life	30hrs (Mammals, *in vitro*) 20 hrs (Yeast, *in viv*o) 10hrs (*E. coli*, *in vivo*)
8	Instability index	40.29
9	Aliphatic index	75.62
10	Grand average of hydropathicity (GRAVY)	-0.695

As theoretical pI is greater than 7 which represents that Topoisomerase II α is basic in nature, Instability index is greater than 40 which shows that protein is unstable. It has low GRAVY values showing better interactions with water molecules.

4.3 Functional Domains Identification of Proteins

Active part of a protein is termed as the functional domain which is involved in the interaction of proteins with other compounds. Proteins can have more than one functional domain that shows different functions.

Functional domains are identified using Interpro; an online server which uses FASTA format of protein as input [89].

Functional domains of Topoisomerase I and Topoisomerase II α are shown in Figure 4.1 and 4.2 respectively. Topoisomerase I contains four functional domains i.e. First domain TopoI DNA bd euk starting from residue number 215 and ends at residue number 429, second domain starting from residue number 432 and ends at 663 with name TopoI cat euk, third domain named TopoI euk starts from residue number 360 and ends at 737.

Fourth and last domain of Topoisomerase I is TopoI C dom starting from residue number 695 and ends at 765. These domains are enlisted in Table 4.3.

TABLE 4.3: Functional Domains of Topoisomerase I

Sr. No.	Name of protein	Functional Domains	Residues Length
1	Topoisomerase I	TopoI DNA bd euk	215 To 429
2		TopoI cat euk	432 To 663
3		TopoI euk	360 To 737
4		TopoI C dom	695 To 765

Topoisomerase II α has seven functional domains. First domain HATPase C starting from residue number 79 and ends at 224, second domain Topo IIA bsu dom starts from residue number 266 and ends at 426, third domain starting from residue number 455 and ends at 572 named TOPRIM domain, fourth domain of this protein starts from residue number 455 and ends at 575 with name title TOPRIM TopoII.

Fifth domain starting from residue number 573 and ends at 711 named TOPRIM C, sixth domain Topo IIA dom A starts from residue number 693 and ends at 1174 and seventh domain DTHCT starts from residue number 1435 and ends at 1522. Functional domains of Topoisomerase II α are represented in Table 4.4.

TABLE 4.4: Functional Domains of Topoisomerase II α

Sr.No.	Name of protein	Domain	Residues Length
1	Topoisomerase II α	HATPase C	79 To 224
2		Topo IIA bsu dom	432 To 663
3		TOPRIM Domain	455 To 572
4		TOPRIM TopoII	455 To 572
5		TOPRIM C	573 To 711
6		Topo IIA dom A	693 To 1174
7		DTHCT	1435 To 1522

Figure 4.1 and 4.2 are showing functional domains of Topoisomerase I and Topoisomerase II α

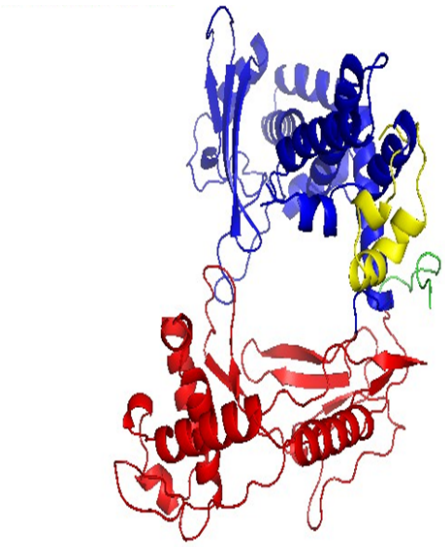


FIGURE 4.3: 3D structure of Topoisomerase I showing functional domains.

Topoisomerase I contains four functional domains i.e. First domain TopoI DNA bd euk starting from residue number 215 and ends at residue number 429, second domain starting from residue number 432 and ends at 663 with name TopoI cat euk, third domain named TopoI euk starts from residue number 360 and ends at 737. Fourth and last domain of TopoisomeraseI is TopoI C dom starting from residue number 695 and ends at 765.

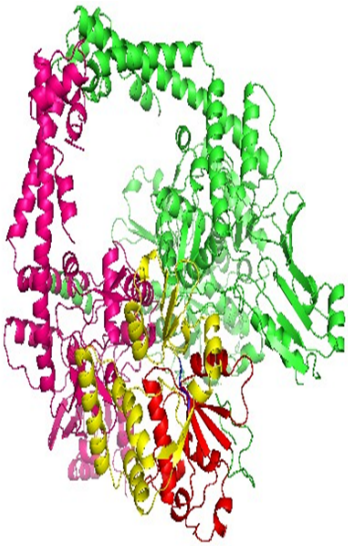


FIGURE 4.4: 3D structure of Topoisomerase II α showing functional domains.

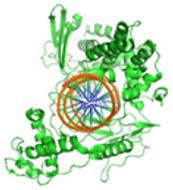
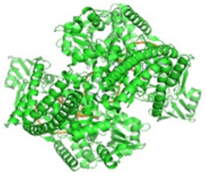
Topoisomerase II α has seven functional domains. First domain HATPase C starting from residue number 79 and ends at 224, second domain Topo IIA bsu dom starts from residue number 266 and ends at 426, third domain starting from residue number 455 and ends at 572 named TOPRIM domain.

Fourth domain of this protein starts from residue number 455 and ends at 575 with name title TOPRIM TopoII, fifth domain starting from residue number 573 and ends at 711 named TOPRIM C, sixth domain Topo IIA dom A starts from residue number 693 and ends at 1174 and seventh domain DTHCT starts from residue number 1435 and ends at 1522.

4.4 Template Selection

After generating a list of potential templates through the search method, it was essential to select one or more that were most suitable for molecular docking. Several factors must be considered during template selection. The most basic criterion is to choose a structure that closely aligns with the target sequence [90]. The structures of the selected templates were obtained from the Protein Data Bank (PDB) and are presented in 4.5.

TABLE 4.5: PDB Template Structures

Sr.	Templates	Resolution	PDB ID	Structure
1	Crystal Structure of Human Topoisomerase I DNA Complex	2.60 Å	1EJ9	
2	Human Topoisomerase II α in complex with DNA and etoposide	3.15 Å	5GWK	

These structures are representing the templates choosen for target proteins. Templates have ligands and water molecules attached to them and we had removed

to refine the structure so that, this refined structure of protein can be used for molecular docking.

4.5 Refined Protein Structures for Docking

3D structures of proteins were refined by the use of PyMol software. Refining was done by removing water molecules as well as other ligands attached to the proteins. Finally, refined structure of protein was obtained, which was used for Molecular Docking. Refined structures of Topoisomerase I and Topoisomerase II α are shown in Fig. 4.3 and Fig. 4.4 respectively.

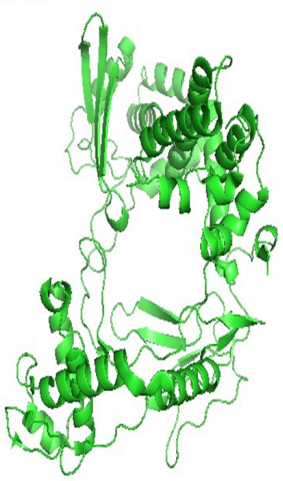


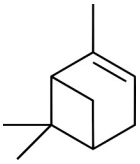
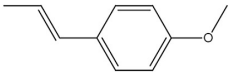
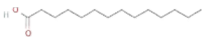
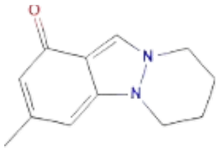
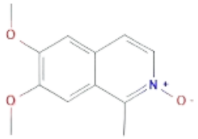
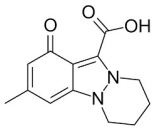
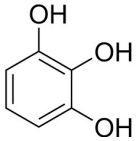
FIGURE 4.5: Refined structure of Topoisomerase I without any water molecule and extra ligands.

FIGURE 4.6: Refined structure of Topoisomerase II α without any water molecule and extra ligands.

4.6 Ligand Selection

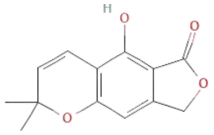
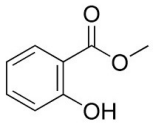
If we need to select a single protein-ligand complex for structure based drug designing, we need to look at the ligand present in the active site. There must be a good affinity between ligand and the receptor. Ligand should be able to adjust the function of proteins (interact with potential residues) and should be able to be used as a drug molecule or lead compound. On the basis of these characteristics we have selected best ligands. Bioactive compounds of Nigella sativa were selected as ligands for the present research and they are enlisted in Table 4.6.

TABLE 4.6: Detailed information about Selected Ligands

Sr.	Name	Molecular Formula	Molecular Weight g/mol	Structure
1	α Pinene	$C_{10}H_{16}$	136.23	
2	Anethol	$C_{10}H_{120}$	148.20	
3	Myristic acid	$C_{14}H_{2802}$	228.37	
4	Nigeglaine	$C_{12}H_{14}N_0$	202.25	
5	Nigellaquinomine	$C_{19}H_{25}N_0$	283.4	
6	Nigellicine	$C_{13}H_{14}N_2O_3$	246.26	
7	Palmitic acid	$C_{16}H_{32}O_2$	256.42	
8	Pyrogallol	$C_6H_6O_3$	126.11	

continued on next page

Table 4.6 continued from previous page

Sr.	Name	Formula	Weight g/mol	Structure
9	Salfredin B11	$C_{13}H_{12}O_4$	232.23	
10	Salicylic acid	$C_7H_6O_3$	138.12	

The selected ligands were α-pinene, anethol, myristic acid, nigeglaine, nigellaquino mine, nigellicine, palmitic acid, pyrogallol, salfredin B11, and salicylic acid. PubChem is a public repository for experimental data that identifies the biological activity of small molecules. The structure of ligands and other information related to ligands were obtained from PubChem database.

4.7 Virtual Screening and Toxicity Prediction

PkCSM is an online tool used for toxicity prediction of ligands. It was used to find out the (ADMET) properties of the ligands; absorption, distribution, metabolism, excretion and toxicity. In this study, the Lipinskis rule has employed for screening of ligands. Application of Lipinskis rule on ligands is mentioned in Table 4.7. All ligands followed Lipinskis rule of 5's.

TABLE 4.7: Application of Lipinskis Rule on Ligands

Sr.	Ligands	LogP Value	Molecular Weight	H-Bond Acceptor	H-Bond Donor
1	α Pinene	2.9987	136.23	0	0
2	Anethol	2.7283	148.2	1	0
3	Myristic acid	4.7721	228.37	1	1
4	Nigeglaine	1.85682	202.25	3	0
5	Nigellaquinomine	4.4277	283.4	2	0
6	Nigellicine	1.55502	246.26	4	1

continued on next page

Table 4.7 continued from previous page

Sr.	Ligands	LogP	Weight	Acceptor	Donor
7	Palmitic acid	5.5523	256.42	1	1
8	Pyrogallol	0.8034	126.11	3	3
9	Salfredin B11	2.2468	232.23	4	1
10	Salicylic acid	1.0904	138.12	2	2

According to Lipinski's rule, the number of Hydrogen bond donor must be less than 5, the maximum number of Hydrogen bond acceptors must be 10, the logP value must be limited to 5 and the molecular weight must be less than 500 g/mol.

4.7.1 Toxicity Prediction

PkCSM is an online tool which provides an integrated platform for rapid evaluation of pharmacokinetics and toxicity properties of drugs. So this tool was used for the toxicity measurements of ligands against Topoisomerase I and Topoisomerase II α which were the target proteins in the present study. Mutagenic potential of a compound is being checked through AMES toxicity test using bacteria. If it shows positive result, then ligand is mutagenic and can act as a carcinogenic compound and vice verse [91]. The hERG I and II inhibitor models are used to assess a compound's potential to block potassium channels associated with the human ether-à-go-go-related gene (hERG). Inhibition of these channels can lead to prolonged QT syndrome and, in severe cases, sudden cardiac death. Several pharmaceutical products have been withdrawn from the market due to their hERG channel-inhibiting properties [93].

In silico toxicology involves assessing chemical toxicity using computational methods to predict or analyze potential harmful effects. Toxicity tests estimate the impact of a substance on humans, animals, or plants after single or repeated exposure. Several factors influence chemical toxicity, including dose, duration, frequency of exposure, and ADME (absorption, distribution, metabolism, and excretion) characteristics [94]. The toxicity profiles of the ligands are summarized in Table 4.8.

TABLE 4.8: Toxicity models for selected compounds

Sr.	Model Name	α-Pinene	Anethol	Myristic acid	Niger- inone	Nigella alol	Nigellicine	Palmitic	Progr. B11	Salfredin acid	Salicylic	No. acid
1	AMES toxicity	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
2	Max tolerated dose (Human)	0.48	0.824	-0.559	0.19	-0.138	0.283	-0.708	-0.269	-0.051		
3	hERG I inhibitor	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
4	hERG II inhibitor	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
5	Oral Rat Acute Toxicity (LD_{50})	1.77	1.798	1.477	2.094	1.853	2.265	1.44	2.049	1.701		
6	Oral Rat Chronic Toxicity (LOAEL)	2.262	2.217	3.034	1.037	0.872	1.17	3.181	2.374	2.419	2.483	
7	Hepatotoxicity	NO	NO	YES	NO	NO	YES	NO	NO	NO	NO	NO
8	Skin sensitisation	NO	YES	YES	NO	NO	NO	YES	NO	NO	NO	NO
9	T. Pyriformis toxicity	0.45	0.807	0.978	1.016	1.456	0.237	0.84	1.027	0.494		
10	Minnow toxicity	1.159	0.869	-0.601	1.432	0.488	1.776	-1.083	2.734	1.492	1.812	

All ligands are non carcinogenic. α-pinene, Anethol and Salicylic acid show high maximum tolerated dose. All ligands are supporter of potassium channels. Except Nigellicine, all ligands are non hepatotoxic. Anethol, Myristic acid and Palmitic acid are skin sensitive and other shows no skin sensitivity. Minnow toxicity values of myristic acid, nigellaquinomine and palmitic acid predicted them toxic.

4.8 Molecular Docking

Molecular Docking is technique used to estimate the bond strength between a ligand and a target protein through a special scoring function and used to determine the correct structure of the ligand within the target binding site. It also helps in the recognition of new small molecular compounds, revealing the essential properties, such as high interaction between binding with target protein having reasonable absorption, distribution, metabolism and excretion which help in the selection of lead compound for the target [94].

The 3D structure of the target proteins in Pdb format and the ligands in Sdf format was used as input for molecular docking. It represents a frequently used approach in structure-based drug designing. The docking was performed among Topoisomerase I, Topoisomerase IIα and ligands which are α-pinene, anethol, myristic acid, nigeglaine, nigellaquinomine, nigellicine, palmitic acid, pyrogallol, salfredin B11 and salicylic acid. Ligands with best binding score values with Topoisomerase I and Topoisomerase IIα are represented in Table 4.9 and 4.10 respectively.

To automatically predict binding modes without information about binding sites, used a user friendly blind docking web server called CB Dock, which predicted and estimated a binding site for a given protein and calculated centers and sizes with a novel rotation cavity detection method and performed docking with the popular docking program named Auto dock Vina [95]. CB Dock gave 5 best interacting confirmations for each ligand molecule. These confirmations were arranged based on binding affinity and then finest confirmation selection was made on the basis of highest affinity score of protein-ligand interaction.

Table 4.9: Ligands exhibiting the strongest binding affinity to Topoisomerase I

Sr. No.	**Parameters**	**α-Pinene**	**Anethol**	**Myristic acid**	**Nigeglaine**	**Nigellaquinomine**	**Nigellicine**	**Palmitic acid**	**Pyrog-Allol**	**Salfredin B11**	**Salicylic acid**
1	Binding score	-5.8	-5.9	-5.8	-6.1	-7.1	-7	-5.6	-6	-7.1	-6.6
2	Cavity size	963	1940	1940	256	1940	1940	1940	1940	1940	1940
3	HBD	0	0	1	0	0	1	1	3	1	2
4	HBA	0	1	1	3	2	4	1	3	4	2
5	logP	2.9987	2.7283	4.7721	1.85682	4.4277	1.55502	5.5523	0.8034	2.2468	1.0904
6	Molecular weight (g/mol)	136.23	148.2	228.37	202.25	283.4	246.26	256.42	126.11	232.23	138.12
7	Rotatable bonds	0	2	12	0	0	1	14	0	0	1
8	Grid map	24	33	33	39	33	33	33	33	33	33
9	Min.Energy Kcal/mol	0.1965	0.0005	0.39	5.1561	-0.409	11.581	0.3901	-2.2322	2.9755	2.9688
10	Max.Energy Kcal/mol	38.7381	5.6305	5.9051	33.2906	25.5747	50.7663	7.1828	-15.3188	19.2704	5.6109

TABLE 4.10: Ligands exhibiting the strongest binding affinity to Topoisomerase II α

Sr. No.	**Parameters**	**α-Pinene**	**Anethol**	**Myristic acid**	**Nigeglaine**	**Nigellaquinomine**	**Nigellicine**	**Palmitic acid**	**Pyrog-Allol**	**Salfredin B11**	**Salicylic acid**
1	Binding score	-6.1	-6.1	-5.4	-6.8	-8.2	-7.2	-5	-5.6	-7.4	-6.5
2	Cavity size	1448	1448	1448	1448	45026	1448	45026	1448	1448	1448
3	HBD	0	0	1	0	0	1	1	3	1	2
4	HBA	0	1	1	3	2	4	1	3	4	2
5	logP	2.9987	2.7283	4.7721	1.85682	4.4277	1.55502	5.5523	0.8034	2.2468	1.0904
6	Molecular weight (g/mol)	136.23	148.2	228.37	202.25	283.4	246.26	256.42	126.11	232.23	138.12
7	Rotatable bonds	0	2	12	0	0	1	14	0	0	1
8	Grid map	25	25	24	25	35	25	35	25	25	25
9	Min. Energy Kcal/mol	0.1965	0.0005	0.39	5.1561	-0.409	11.581	0.3901	-2.2322	2.9755	2.9688
10	Max. Energy Kcal/mol	38.7381	5.6305	5.9051	33.2906	25.5747	50.7663	7.1828	-15.3188	19.2704	5.6109

The Docked structures after docking process were selected for further analysis; on the basis of docking score, binding energy, cavity size and Grid map, we selected best docked structure.

Molecular docking was performed using Topoisomerase I, Topoisomerase II α as receptors and 15 selected compounds as ligands. Out of which only 10 ligands showed interactions with target proteins i.e. Topoisomerase I and Topoisomerase II α and these compounds were taken for further proceedings. Data of all these ligands was arranged after performing molecular docking. Docking result gave five confirmations based and arranged on the basis of binding score and we had selected the confirmation with minimum binding score. All these ligands are showing minimum binding score in their corresponding confirmations.

4.9 Ligands Interaction with Target Proteins

For interpretation of docked results, interaction of the active pockets of the ligands and proteins were calculated. Two types of interactions were studied to analyze the docking results; hydrogen bonding and hydrophobic interaction using Ligplot+. Hydrogen bonding & hydrophobic interactions of active ligands are shown in Table 4.11 and 4.12.

TABLE 4.11: Hydrogen and Hydrophobic Interactions of Topoisomerase I

Sr.	**Ligand Name**	**Binding Energy**	**No. of HBs**	**H-Bonds (Amino:Bond length)**	**Hydrophobic Amino Acids**
1	α-Pinene	-5.8	0	—	Trp203, Glu209, Glu208, Glu438, Tyr211, Arg434, Asp344
2	Anethol	-5.9	0	—	Gly503, Cys504, Leu487, Gly531, Ile535, Cys630, Asn631, Arg590, Arg488, Phe529

continued on next page

Table 4.11 continued from previous page

Sr.	Ligand Name	Binding Energy	No. of HBs	H-Bonds (Amino:Bond length)	Hydrophobic Amino Acids
3	Myristic acid	-5.8	1	Arg590:2.81	Ala486, Leu629, Cys630, Tyr537, Arg488, Lys532, Gln633, Asp533, His632, Ile535, Gly531, Asn631
4	Nigeglaine	-6.1	1	Phe240:2.95	Pro235, Glu254, Leu234, Val238, Lys239, Met247, Glu236, Lys248, Val246
5	Nigellaquinomine	-7.1	0	—	Val502, Ser506, Arg508, Asp500, Gly496, Thr498, Ala499, Glu494, Lys493
6	Nigellicine	-7.0	6	Arg364:2.88, Lys493:2.94, Asp533:3.10, His367:2.95, Ala499:3.33, Ser534:2.99, :3.11	Thr498, Gly363, Lys532, Thr501
7	Palmitic acid	-5.6	0	—	Arg590, Cys630, Gly531, Tyr537, Lys532, Lys493, Asn491, Gly490, Gly503, Arg488, Phe529, Asn631, Val502
8	Pyrogallol	-6.0	2	Arg590:2.94, Arg488:3.04	Cys630, Asn631, Tyr537, Phe529, Gly531, Ala486, Leu487

continued on next page

Table 4.11 continued from previous page

Sr.	Ligand Name	Binding Energy	No. of HBs	H-Bonds (Amino:Bond length)	Hydrophobic Amino Acids
9	Salfredin B11	-7.1	4	His367:3.20, Thr501:3.20, Ser534:2.73, Arg364:2.85, :3.08	Phe361, Gln421, Asp533, Lys493, Ala499, Thr498
10	Salicylic acid	-6.6	3	Ile535:2.70, Arg488:2.91, :3.05, Cys630:2.71	Gly531, Asn631, Ala486, Arg590, Leu487, Phe529, Tyr537

α-Pinene made 6 hydrophobic interactions with Topoisomerase I and no hydrogen bonding as illustrated in figure 4.5. Anethol made 10 hydrophobic interactions with Topoisomerase I and no hydrogen bonding as illustrated in figure 4.6.

Myristic acid made 12 hydrophobic interactions with Topoisomerase I and 1 hydrogen bond with Arg590 residue as illustrated in figure 4.7. Nigeglaine made 09 hydrophobic interactions with Topoisomerase I and 1 hydrogen bond with Phe240 residue as illustrated in figure 4.8. Nigellaquinomine made 09 hydrophobic interactions with Topoisomerase I and no hydrogen bonding as shown in figure 4.9.

Nigellicine made 04 hydrophobic interactions with Topoisomerase I and 06 hydrogen bonds with Arg364, His367, Lys493, Ala499, Asp533 and Ser534 residues as shown in figure 4.10. Palmitic acid made 13 hydrophobic interactions with Topoisomerase I and no hydrogen bonding as shown in figure 4.11. Pyrogallol made 07 hydrophobic interactions with Topoisomerase I and 02 hydrogen bonds with Arg590 and Arg488 residues as illustrated in figure 4.12.

Salfredin B11 made 06 hydrophobic interactions with Topoisomerase I and 04 hydrogen bonds with His367, Arg364, Ser534 and Thr501 residues as illustrated in figure 4.13. Salicylic acid made 07 hydrophobic interactions with Topoisomerase I and 03 hydrogen bonds with Arg488, Cys630 and Ile535 residues as illustrated in figure 4.14.

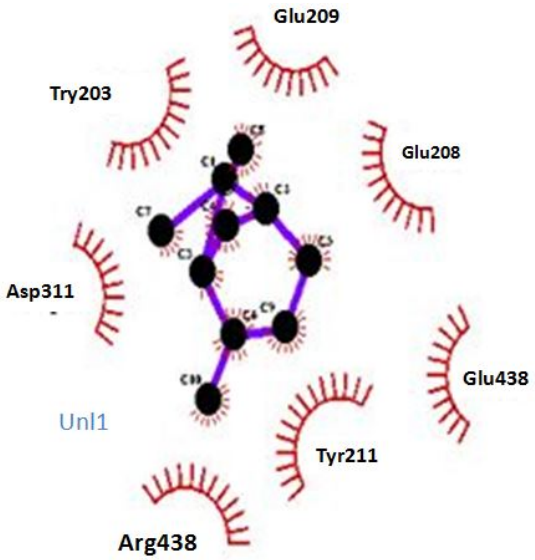


FIGURE 4.7: Interaction of α-Pinene with Topoisomerase I

α-Pinene made 07 hydrophobic interactions with Trp203, Glu209, Glu208, Glu438, Tyr211, Arg434 and Asp344 residues and no Hydrogen bond.

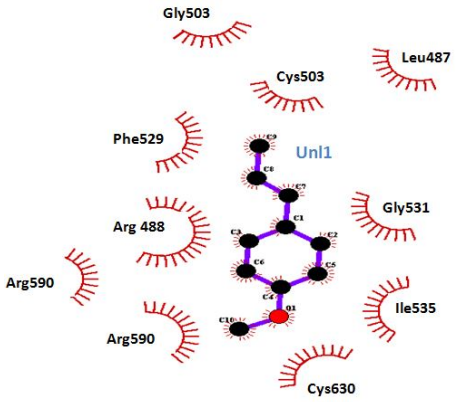


FIGURE 4.8: Interaction of Anethol with Topoisomerase I

Anethol made 10 hydrophobic interactions with Gly503, Cys504, Leu487, Gly531, Ile535, Cys630, Asn631, Arg590, Arg488 and Phe529 residues and no Hydrogen bond.

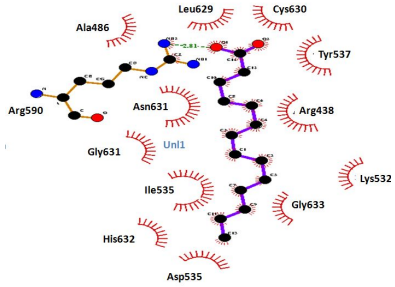


FIGURE 4.9: Interaction of Myristic acid with Topoisomerase I

Myristic acid made 12 hydrophobic interactions with Ala486, Leu629, Cys630, Tyr537, Arg488, Lys532, Gln633, Asp533, His632,Ile535, Gly531 and Asn631 residues and 1 Hydrogen bond with Arg590 residue having bond length of 2.81.

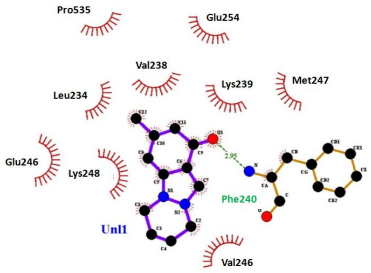


FIGURE 4.10: Interaction of Nigeglaine with Topoisomerase I

Nigeglaine made 09 hydrophobic interactions with Pro235, Glu254, Leu234, Val238, Lys239, Met247, Glu236,Lys248 and Val246 residues and 01 Hydrogen bond with Phe240 residue having bond length of 2.95.

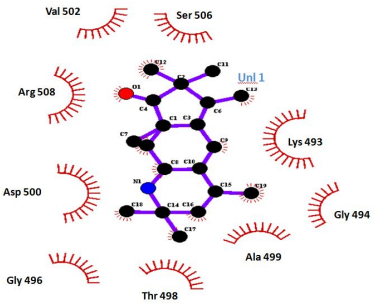


FIGURE 4.11: Interaction of Nigellaquinomine with Topoisomerase I

Nigellaquinomine made 09 hydrophobic interactions with Val502, Ser506, Arg508, Asp500, Gly496, Thr498, Ala499, Glu494 and Lys493 residues and no Hydrogen bond.

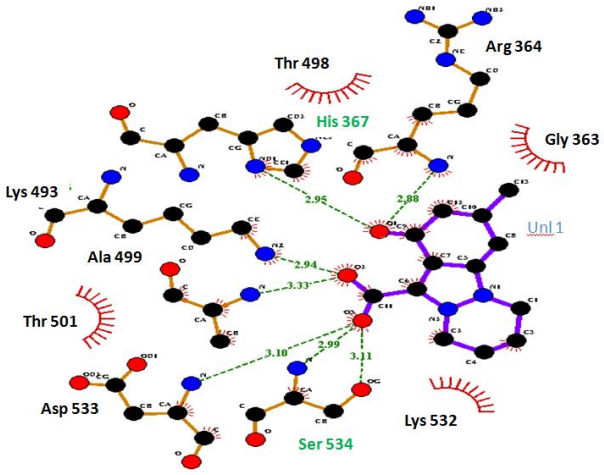


FIGURE 4.12: Interaction of Nigellicine with Topoisomerase I

Nigellicine made 04 hydrophobic interactions with Thr498, Gly363, Lys532 and Thr501 residues and 06 Hydrogen bonds with residues Arg364 having bond length of 2.88, His367 having bond length of 2.95, Lys493 having bond length of 2.94, Ala499 having bond length of 3.33, Asp533 having bond length of 3.10 and Ser534 having bond lengths of 2.99 and 3.11.

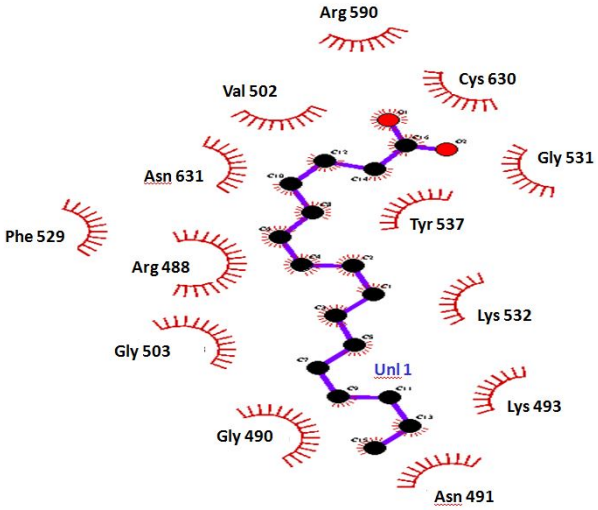


FIGURE 4.13: Interaction of Palmitic acid with Topoisomerase I

Palmitic acid made 13 hydrophobic interactions with Arg590, Cys630, Gly531, Tyr537, Lys532, Lys493, Asn491, Gly490, Gly503, Arg488, Phe529, Asn631 and Val502 residues and no Hydrogen bond.

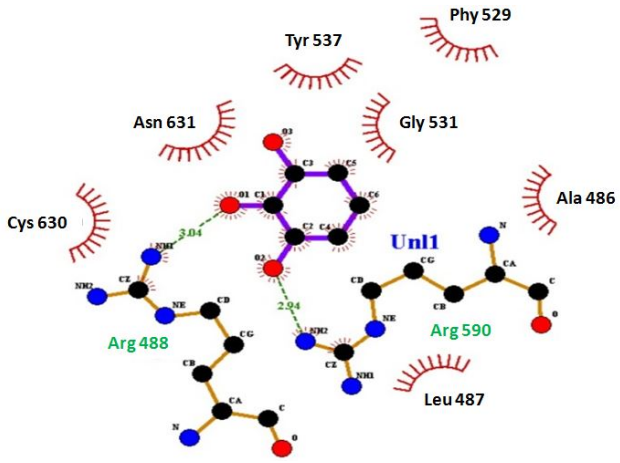


FIGURE 4.14: Interaction of Pyrogallol with Topoisomerase I

Pyrogallol made 07 hydrophobic interactions with Cys630, Asn631, Tyr537, Phe529, Gly531, Ala486 and Leu487 residues and 02 Hydrogen bonds with Arg590 having bond length of 2.94 and Arg488 having bond length of 3.04.

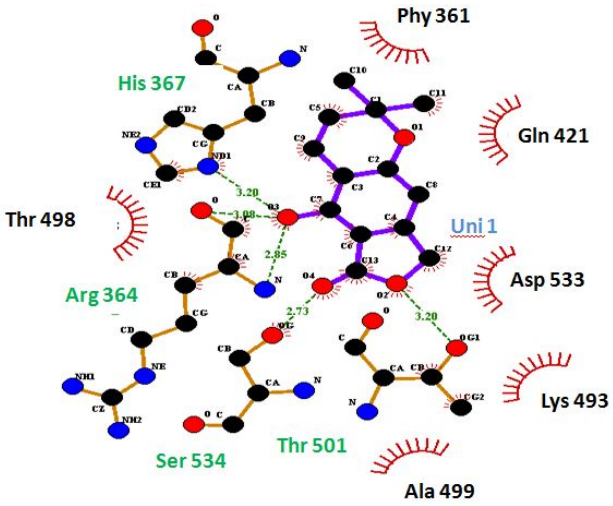


FIGURE 4.15: Interaction of Salfredin B11 with Topoisomerase I

Salfredin B11 made 06 hydrophobic interactions with Phe361, Gln421, Asp533, Lys493, Ala499 and Thr498 residues and 04 Hydrogen bonds with residues His367 having bond length of 3.20, Thr501 having bond length of 3.20, Ser534 having bond length of 2.73 and with Arg364 having bond lengths of 2.85 and 3.08.

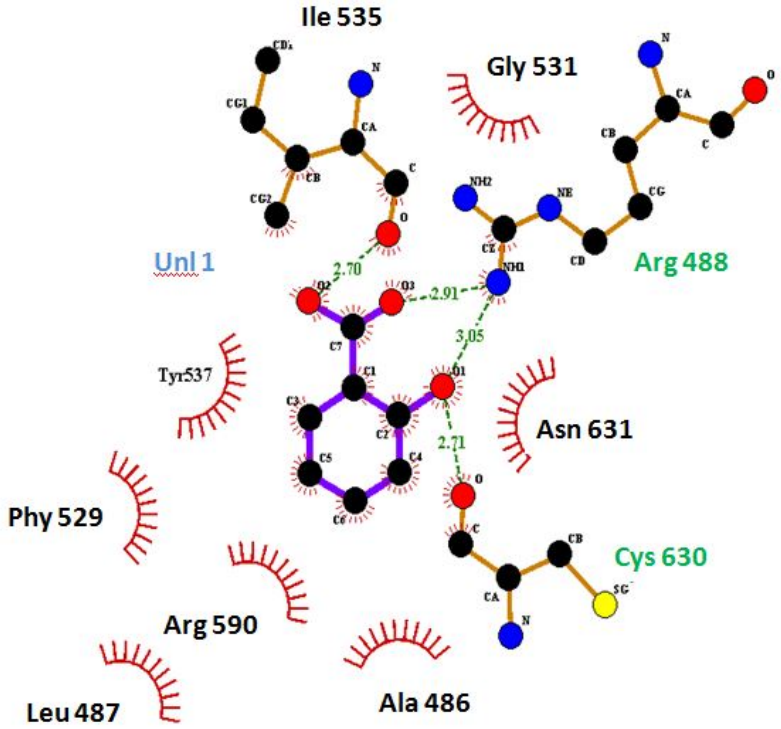


FIGURE 4.16: Interaction of Salicylic acid with Topoisomerase I

Salicylic acid made 07 hydrophobic interactions with Gly531, Asn631, Ala486, Arg590, Leu487,Phe529 and Tyr537 residues and 03 Hydrogen bonds with Ile535 having bond length of 2.70, Arg488 having bond lengths of 2.91 and 3.05 and with residue Cys630 having bond length of 2.71.

TABLE 4.12: Hydrogen and Hydrophobic Interactions of Topoisomerase II α

Sr.	**Ligand Name**	**Binding Energy**	**No. of HBs**	**H-Bonds (Amino:Bond Length)**	**Hydrophobic Interactions (Amino Acids)**
1	α-Pinene	-6.1	0	NIL:NIL	Phe1003, Glu712, Pro724, Pro716, Ser717, Ile715
2	Anethol	-6.1	0	NIL:NIL	Arg673, Gly1007, Pro724, Ile715, Arg727, Leu722, Ser717, Pro716, Phe1003, Glu712

continued on next page

Table 4.12 continued from previous page

Sr.	Ligand Name	Binding Energy	No. of HBs	H-Bonds (Amino:Bond Length)	Hydrophobic Interactions (Amino Acids)
3	Myristic acid	-5.4	2	Arg727:3.07, :3.15; Leu722:2.77	Val1006, Glu839, Gly1007, Leu829, Pro724, Glu712, Ser717, Lys723, Pro716, Phe1003, Ile715, Asp1004, Trp840
4	Nigeglaine	-6.8	1	Trp840:3.20	His1005, Glu839, Val1006, Gly1007, Glu712, Phe1003
5	Nigellaquinomine	-8.2	0	NIL:NIL	Thr618, Phe807, Tyr805, Gly617, Gly615, Lys614
6	Nigellicine	-7.2	3	Trp840:3.10, Asp1004:2.68, Ser717:3.17, :3.32	His1005, Phe1003, Glu712, Gly1007, Val1006, Lys676, Glu839
7	Palmitic acid	-5.0	2	Thr618:2.70, Gly617:3.16	Lys614, Gly615, Tyr805, Ala801, Leu616, Gln789, Leu468, Thr467, Ser619, Ser464
8	Pyrogallol	-5.6	1	Arg727:3.30, :2.82, :3.09	Pro724, Ile715, Phe1003, Glu712, Glu839, Ser717
9	Salfredin B11	-7.4	2	Glu839:2.85, Arg727:3.22	Trp840, Val1006, Gly1007, Glu712, Phe1003, Pro724
10	Salicylic acid	-6.5	1	Arg727:3.06, :2.89	Pro724, Ser717, Pro716, Ile715, Glu712, Phe1003

α-Pinene made 06 hydrophobic interactions with Topoisomerase II α and no hydrogen bonding as illustrated in figure 4.15. Anethol made 10 hydrophobic interactions with Topoisomerase II α and no hydrogen bonding as illustrated in figure 4.16. Myristic acid made 13 hydrophobic interactions with Topoisomerase II α and 02 hydrogen bonds with Arg727 and Leu722 residues as illustrated in figure 4.17. Nigeglaine made 07 hydrophobic interactions with Topoisomerase II α and 01 hydrogen bonds with Trp840 residue as illustrated in figure 4.18. Nigellaquinomine made 06 hydrophobic interactions with Topoisomerase II α and no hydrogen bonding as illustrated in figure 4.19. Nigellicine made 07 hydrophobic interactions with Topoisomerase II α and 03 hydrogen bonds with Trp840, Asp1004 and Ser717 residues as shown in figure 4.20. Palmitic acid made 10 hydrophobic interactions with Topoisomerase II α and 02 hydrogen bonds with Gly617 and Thr618 residues as shown in figure 4.21. Pyrogallol made 06 hydrophobic interactions with Topoisomerase II α and 01 hydrogen bond with Arg727 residue as shown in figure 4.22. Salfredin B11 made 07 hydrophobic interactions with Topoisomerase II α and 02 hydrogen bonds with Arg727 and Glu839 residues as shown in figure 4.23. Salicylic acid made 07 hydrophobic interactions with Topoisomerase II α and 01 hydrogen bond with Arg727 residue as illustrated in figure 4.24.

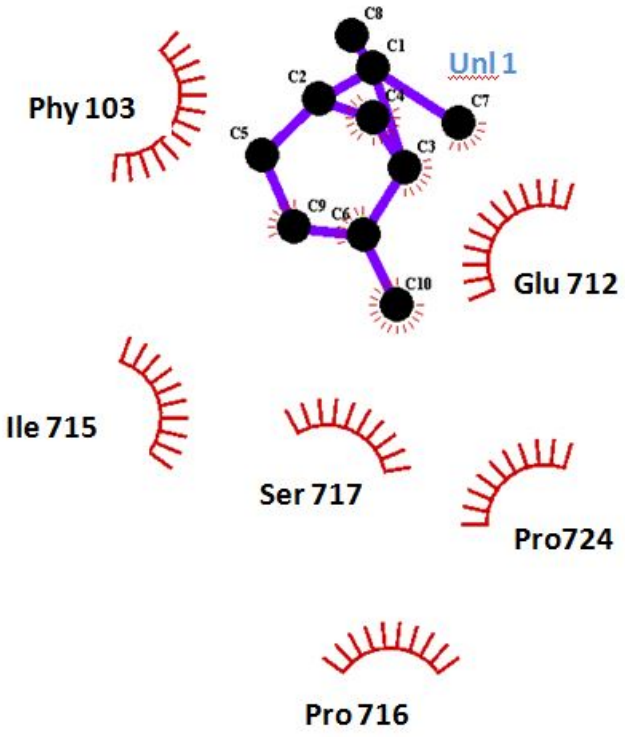


FIGURE 4.17: Interaction of α-Pinene with Topoisomerase II α

α-Pinene made 06 hydrophobic interactions with Phe1003, Glu712, Pro724, Pro716, Ser717 and Ile715 residues and no Hydrogen bond.

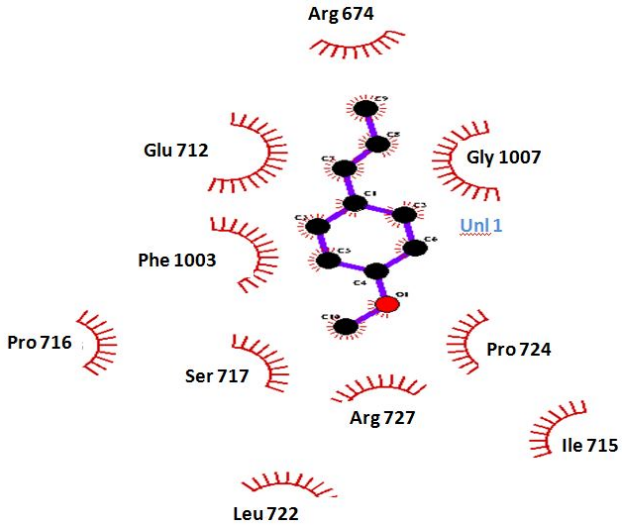


FIGURE 4.18: Interaction of Anethol with Topoisomerase II α

Anethol made 10 hydrophobic interactions with Arg673, Gly1007, Pro724, Ile715, Arg727, Leu722, Ser717, Pro716, Phe1003 and Glu712 residues and no Hydrogen bond.

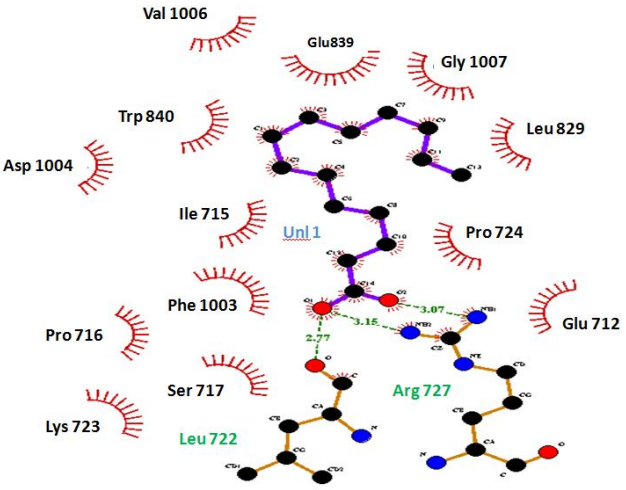


FIGURE 4.19: Interaction of Myristic acid with Topoisomerase II α

Myristic acid made 13 hydrophobic interactions with Val1006, Glu839, Gly1007, Leu829, Pro724, Glu712, Ser717, Lys723, Pro716, Phe1003, Ile715, Asp1004 and Trp840 residues and 02 Hydrogen bonds with Arg727 having bond lenghts of 3.07 and 3.15 and with residue Leu722 having bond length of 2.77.

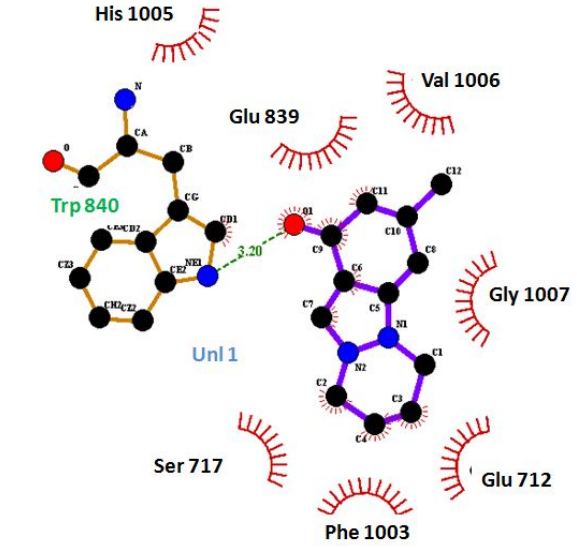


FIGURE 4.20: Interaction of Nigeglaine with Topoisomerase II α

Nigeglaine made 07 hydrophobic interactions with His1005, Glu839, Val1006, Gly1007, Glu712, Phe1003 and Ser717 residues and 01 Hydrogen bond with Trp840 residue having bond length of 3.20.

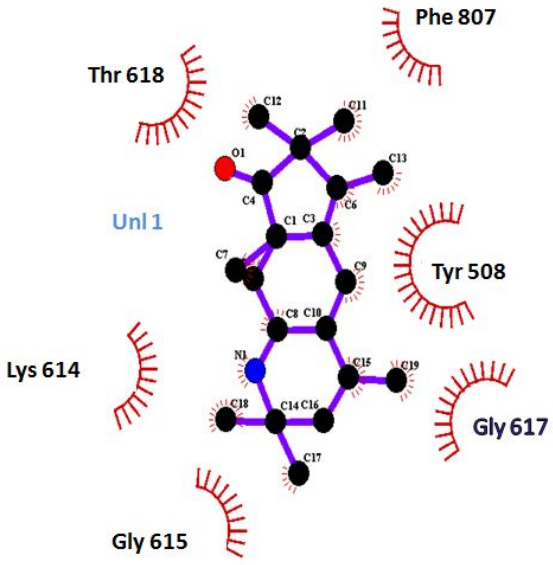


FIGURE 4.21: Interaction of Nigellaquinomine with Topoisomerase II α

Nigellaquinomine made 06 hydrophobic interactions with Thr618, Phe807, Tyr805, Gly617,Gly615 and Lys614 residues and no Hydrogen bond.

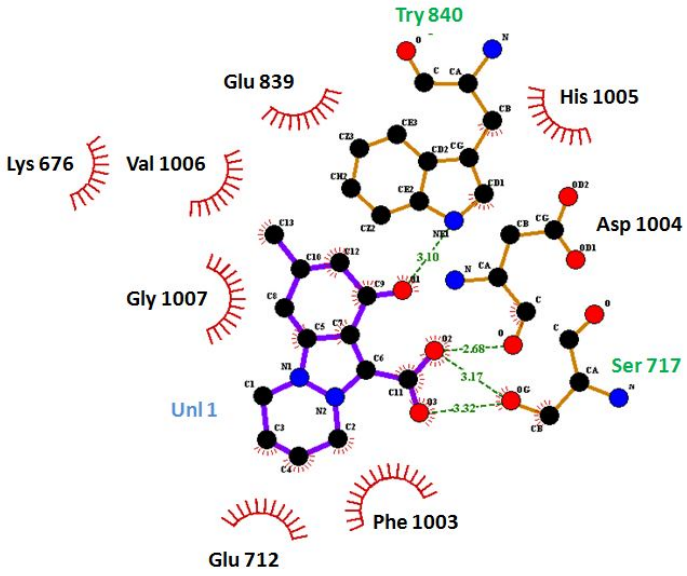


FIGURE 4.22: Interaction of Nigellicine with Topoisomerase II α

Nigellicine made 07 hydrophobic interactions with His1005, Phe1003, Glu712, Gly1007, Val1006, Lys676 and Glu839 residues and 03 Hydrogen bonds with Trp840 having bond length of 3.10, Asp1004 having bond length of 2.68 and with residue Ser717 having bond length of 3.17 and 3.32.

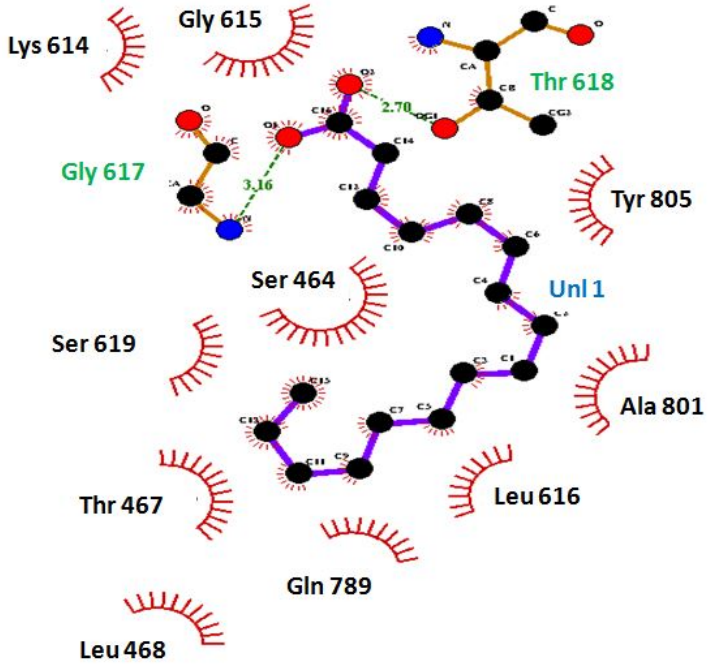


FIGURE 4.23: Interaction of Palmitic acid with Topoisomerase II α

Palmitic acid made 10 hydrophobic interactions with Lys614, Gly615, Tyr805, Ala801, Leu616, Gln789, Leu468, Thr467, Ser619 and Ser464 residues and 02

Hydrogen bonds with Thr618 having bond length of 2.70 and with Gly617 having bond length of 3.16.

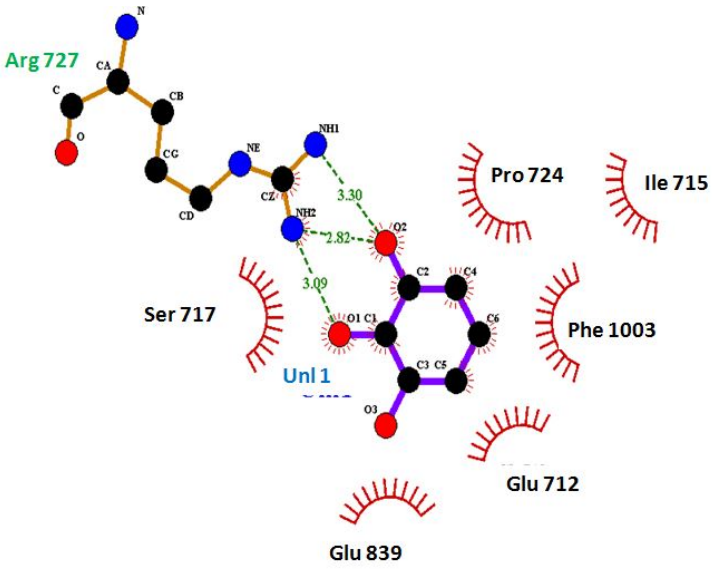


FIGURE 4.24: Interaction of Pyrogallol with Topoisomerase II α

Pyrogallol made 06 hydrophobic interactions with Pro724, Ile715, Phe1003, Glu712, Glu839 and Ser717 residues and 01 Hydrogen bond with residue Arg727 having bond lengths of 3.30,2.82 and 3.09.

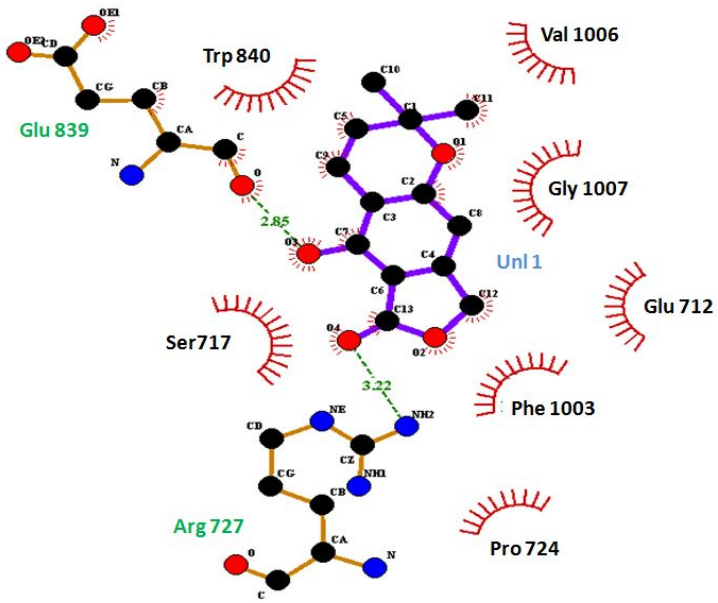


FIGURE 4.25: Interaction of Salfredin B11 with Topoisomerase II α

Salfredin B11 made 07 hydrophobic interactions with Trp840, Val1006, Gly1007, Glu712, Phe1003, Pro724 and Ser717 residues and 02 Hydrogen bonds with Glu839 having bond length of 2.85 and with residue Arg727 having bond length of 3.22.

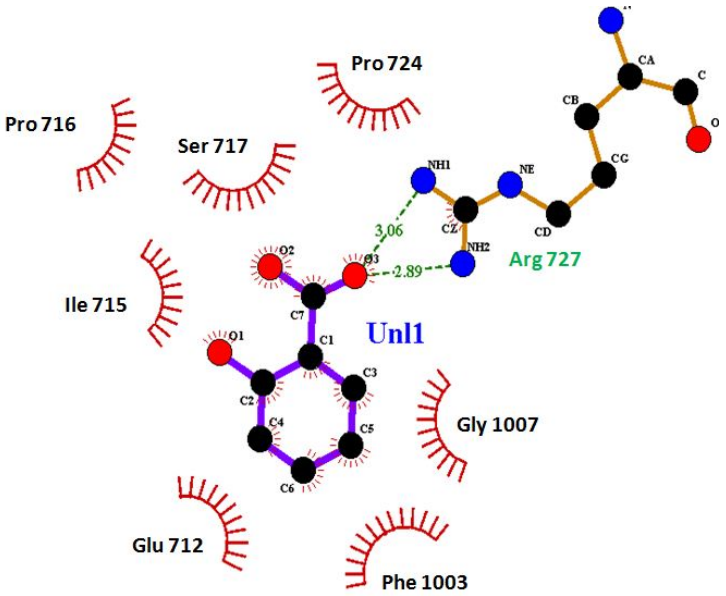


FIGURE 4.26: Interaction of Salicylic acid with Topoisomerase II α

Salicylic acid made 07 hydrophobic interactions with Pro724, Ser717, Pro716, Ile715, Glu712, Phe1003 and Gly1007 residues and 01 Hydrogen bond with Arg727 having bond lengths of 3.06 and 2.89.

4.10 ADMET Properties of Ligands

Ligands toxicity and ADMET properties are obtained from pkCSM online tool. Canonical SMILES of the ligands obtained from PubChem and are used as input for pkCSM server. Toxicity measurements provide information regarding the nature of ligands, which must be considered before drug designing. Toxicity of a compound must be tested to use it as a therapeutic agent. ADMET properties of the selected ligands extracted from this server are as follows.

4.10.1 Absorption

In pharmacology, absorption refers to the movement of a drug from the bloodstream into the tissues. Both the chemical nature of the drug and the surrounding environment influence the rate and extent of absorption. For a drug to be absorbed into tissues, it must cross cellular barriers such as epithelial or endothelial cells. Most drugs cross these barriers via passive transport, moving from areas of higher concentration to lower concentration by diffusing through cell membranes. Only a few drugs use active transport, a process that requires energy in the form of ATP to move the drug against its concentration gradient—from lower to higher concentration.

Absorption is one of the key ADMET properties and is evaluated using seven models: water solubility, $CaCO_2$ permeability, intestinal absorption, skin permeability, P-glycoprotein substrate, and P-glycoprotein inhibitors types I and II. Water solubility refers to the compound's ability to dissolve in water at 25°C and is expressed as the logarithm of its molar concentration (log mol/L). Water solubility refers to the compound's ability to dissolve in water at 25°C and is expressed as the logarithm of its molar concentration (log mol/L). Drugs that are water-soluble tend to dissolve better in aqueous environments compared to lipid-soluble drugs [97].

The $CaCO_2$ permeability model estimates the logarithm of the apparent permeability coefficient (log Papp; log cm/s). A compound is considered to have high $CaCO_2$ permeability if its predicted Papp value exceeds 0.9, according to pkCSM predictions. Intestinal absorption refers to the percentage of a compound expected to be absorbed in the small intestine; compounds with absorption values below 30% are considered poorly absorbed. Skin permeability is represented by the log Kp value and is important in the design of transdermal drugs. A log Kp value greater than -2.5 indicates low skin penetration.

The $CaCO_2$ permeability model estimates the logarithm of the apparent permeability coefficient (log Papp; log cm/s). A compound is considered to have high $CaCO_2$ permeability if its predicted Papp value exceeds 0.9, according to pkCSM

predictions. Intestinal absorption refers to the percentage of a compound expected to be absorbed in the small intestine; compounds with absorption values below 30% are considered poorly absorbed. Skin permeability is represented by the log Kp value and is important in the design of transdermal drugs. A log Kp value greater than -2.5 indicates low skin penetration. Skin permeability is represented by the log Kp value and is important in the design of transdermal drugs. A log Kp value greater than -2.5 indicates low skin penetration.

P-glycoprotein (P-gp) substrates function as natural barriers that help eliminate toxins from cells. This model predicts whether a compound acts as a P-gp substrate, which may reduce its oral absorption. The P-glycoprotein I and II inhibitor models determine whether a compound inhibits P-gp I or II. P-gp inhibitors reduce the efflux activity of P-gp, potentially enhancing drug absorption.

The absorption properties of the ligands are detailed in 4.13. All listed ligands exhibit low water solubility. $CaCO_2$ permeability values for all ligands fall within the normal range, except for nigellicine, which falls below the recommended threshold. Intestinal absorption values are high for all ligands, with each exceeding 90%.

Among them, nigellicine has the highest absorption at 100%. Myristic acid is the only ligand with a skin permeability value within the effective range, indicating strong transdermal potential. Nigellaquinomine is predicted to act as a P-glycoprotein I inhibitor

4.10.2 Distribution

In pharmacology, distribution is a branch of pharmacokinetics that focuses on the movement of a drug within the body from one location to another. In pharmacology, distribution is a branch of pharmacokinetics that focuses on the movement of a drug within the body from one location to another.

After a drug enters the systemic circulation either through direct administration or absorption, it is distributed into both intracellular and interstitial fluids [98].

Table 4.13: Absorption Properties of Ligands

Sr.	Model Name	α-Pinene	Anethol	Myristic acid	Nigeglaine	Nigellaquinomine	Nigellicine	Palmitic acid	Pyrogallol	Salfredin B11	Salicylic acid
1	Water solubility	-3.733	-2.936	-4.952	-3.052	-5.129	-2.148	-5.562	-1.408	-3.081	-1.808
2	$CaCO_2$ permeability	1.38	1.669	1.56	1.217	1.42	0.453	1.558	1.122	1.201	1.151
3	Intestinal absorption (human)	96.041	95.592	92.691	99.353	97.773	100	92.004	83.549	94.058	83.887
4	Skin permeability	-1.827	-1.139	-2.705	-2.497	-2.472	-2.73	-2.717	-2.751	-3.236	-2.723
5	P-glycoprotein substrate	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO
6	P-glycoprotein I inhibitor	NO	NO	NO	NO	YES	NO	NO	NO	NO	NO
7	P-glycoprotein II inhibitor	NO	NO	NO	NO	NO	NO	NO	NO	NO	NO

The distribution properties of the ligands are summarized in Table 4.14. Distribution, a key aspect of ADMET, is characterized using four models: volume of distribution in humans (VDss, expressed as log L/kg), fraction unbound in humans (Fu), and blood-brain barrier (BBB) permeability, expressed as log BBB [100].

Model-I represents the theoretical volume in which the total amount of a drug would need to be evenly distributed to achieve the same concentration as in the blood plasma.

A VDss (volume of distribution at steady state) value below 0.71 L/kg is considered low, while a value above 2.81 L/kg is considered high. A high VDss indicates that a greater proportion of the drug is distributed into tissues rather than remaining in the plasma.

A higher Fu (fraction unbound) value suggests greater effectiveness of the compound. The blood-brain barrier (BBB) protects the brain from external substances, making BBB permeability a critical parameter. A predicted log BB value greater than 0.3 indicates that the substance can cross the BBB, whereas a value less than -1 suggests it is unlikely to affect the brain.

Log P, which combines blood-brain permeability and surface area, is used to predict central nervous system (CNS) penetration. A log P value above -2 suggests the compound can penetrate the CNS, while a value below -3 indicates it is likely safe from CNS effects.

All ligands exhibited low VDss (human) values, suggesting that a larger proportion of the drug remains in the plasma rather than distributing into tissues. The fraction unbound (human) values for these ligands fall within the recommended range.

BBB permeability values for α-pinene, anethol, nigeglaine, nigellaquinomine, and salfredin B11 exceed 0.3, indicating these compounds have the potential to cross the blood-brain barrier and may affect the brain. Additionally, CNS permeability values for anethol, myristic acid, and palmitic acid are greater than -2, suggesting these compounds can penetrate the central nervous system [99].

4.10.3 Metabolism

Metabolism is a process of converting one form of compound into another form; that is conversion of complex to simpler compounds and vice versa. Most of the drug metabolism occurs in liver, blood plasma, lungs and intestine. Metabolism generally converts the drug into more water soluble compound by increasing its polarity. Cytochrome P450 is an important cleansing enzyme found in the liver has different isoforms whose models are included in metabolism of ADMET properties which are CYP1A2, CYP2C19, CYP2C9, CYP2D6 and CYP3A4.

TABLE 4.14: Distribution Properties of Ligands

Sr.	**Model Name**	**VDss (human)**	**Fraction unbound (human)**	**BBB permeability**	**CNS permeability**
1	α-Pinene	0.667	0.425	0.791	-2.201
2	Anethol	0.343	0.266	0.529	-1.659
3	Myristic acid	-0.578	0.171	-0.027	-1.925
4	Nigeglaine	0.312	0.414	0.673	-2.362
5	Nigellaquinomine	0.527	0.225	0.498	-2.645
6	Nigellicine	-0.801	0.489	-0.144	-2.944
7	Palmitic acid	-0.543	0.101	-0.111	-1.816
8	Pyrogallol	0.13	0.712	-0.441	-3.252
9	Salfredin B11	0.363	0.465	0.747	-2.827
10	Salicylic acid	-1.57	0.563	-0.334	-3.21

TABLE 4.15: Metabolic Properties of Ligands

Model Name	**CYP2 D6**	**CYP3 A4**	**CYP1 A2**	**CYP2 C19**	**CYP2 C9**	**CYP2 D6**	**CYP3 A4**
α-Pinene	NO	NO	NO	NO	NO	NO	NO
Anethol	NO	NO	YES	NO	NO	NO	NO
Myristic acid	NO	NO	NO	NO	NO	NO	NO
Nigeglaine	NO	NO	NO	NO	NO	NO	NO
Nigellaquinomine	NO	YES	NO	YES	NO	NO	NO
Nigellicine	NO	NO	NO	NO	NO	NO	NO
Palmitic acid	NO	YES	NO	NO	NO	NO	NO
Pyrogallol	NO	NO	NO	NO	NO	NO	NO

continued on next page

Table 4.15 continued from previous page

Model Name	CYP2 D6	CYP3 A4	CYP1 A2	CYP2 C19	CYP2 C9	CYP2 D6	CYP3 A4
Salfredin B11	NO	NO	NO	NO	NO	NO	NO
Salicylic acid	NO	NO	NO	NO	NO	NO	NO

This enzyme interacts with xenobiotics to aid in their elimination. While it activates certain drugs, it neutralizes most of them [94]. Tables 4.15 show the metabolic properties of ligands. Out of the ligands mentioned in Table 4.15, Nigellaquinomine and palmitic acid act as the substrate of isoform CYP3A4, anethol act as inhibitor of CYP1A2 isoform and Nigellaquinomine act as inhibitor of CYP2C19 isoform. No other ligand act as inhibitor or substrate of any other isoform.

4.10.4 Excretion

The liver and kidneys are the primary organs responsible for drug excretion, although other organs can also contribute. For instance, the lungs eliminate gaseous drug byproducts, and drugs may also be excreted through tears, sweat, and saliva. Excretory properties are often evaluated using models that include total clearance (CL_tot), expressed as log CL_tot in ml/min/kg, and renal OCT2 substrate prediction, indicated as either Yes or No. OCT2 (organic cation transporter 2) is a renal uptake transporter involved in the kidney's role in drug elimination. A negative value for total clearance suggests poor drug elimination, whereas a positive value indicates effective clearance. The excretory properties of the ligands are summarized in Table 4.16. According to the table, all ligands show negative results for the renal OCT2 substrate, but they have positive total clearance values, suggesting good overall drug excretion.

TABLE 4.16: Excretory Properties of Ligands

Sr.No.	Model Name	TC	Renal OCT2 substrate
1	α-Pinene	0.043	NO
2	Anethol	0.268	NO
3	Myristic acid	1.693	NO

continued on next page

Table 4.16 continued from previous page

Sr.No.	Model Name	TC	Renal OCT2 substrate
4	Nigeglaine	0.526	NO
5	Nigellaquinomine	0.903	NO
6	Nigellicine	0.55	NO
7	Palmitic acid	1.763	NO
8	Pyrogallol	0.104	NO
9	Salfredin B11	0.481	NO
10	Salicylic acid	0.607	NO

4.11 Lead Compound Identification

Physiochemical properties or Lipinskis rule of five acts as a primary filter and then pharmacokinetics properties sorts further potential compounds as drug or non drug. After analysing all ligands carefully, Salfredin B11 has selected as lead compound because it is active of all the ligands.

4.12 Drug Identification against Breast Cancer

4.12.1 Sorafenib

Sorafenib is an FDA-approved medication for the treatment of advanced renal cell carcinoma and is marketed by Bayer under the brand name Nexavar. It also received a 'Fast Track' designation from the FDA for the treatment of advanced hepatocellular carcinoma (HCC) and has shown improved outcomes in Phase III clinical trials.

Sorafenib is a small-molecule inhibitor that targets multiple kinases, including Raf kinase, platelet-derived growth factor (PDGF) receptor, VEGF receptors 2 and 3, and c-Kit—the receptor for stem cell factor. While many emerging drugs focus on these signaling pathways, Sorafenib is distinct in its simultaneous inhibition of the Raf/Mek/Erk signaling cascade [69].

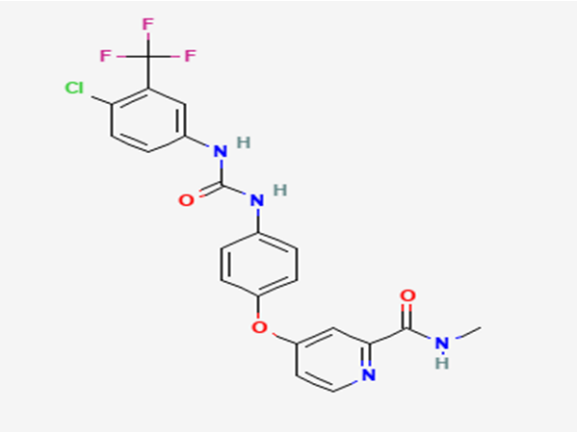


FIGURE 4.27: 2D structure of Sorafenib Drug (Pubchem Database)

Physiochemical properties of Sorafenib drug are shown in Table 4.17.

TABLE 4.17: Physiochemical properties of Sorafenib Drug

log P value	**Rotatable Bonds**	**H-Bond acceptor**	**H-Bond donor**	**Molecular formula**	**Molecular weight**
5.5497	5	4	3	$C_{21}H_{16}CIF_3N_4O_3$	464.8

4.13 ADMET Properties of Drug

PkCSM is used to predict the ADMET properties of Sorafenib drug. Predicted values of sorafenib are given in Table 4.18.

TABLE 4.18: ADMET Properties of Drug

Sr. No.	**Property**	**Model Name**	**Predicted Values**
1		Water solubility	-4.255
2		$CaCO_2$ permeability	0.762
3		Intestinal absorption (human)	85.494
4	Absorption	Skin permeability	-2.74
5		P-glycoprotein substrate	YES
6		P-glycoprotein I inhibitor	YES
7		P-glycoprotein II inhibitor	YES
1		VDss (human)	-0.009

continued on next page

Table 4.18 continued from previous page

Sr. No.	Property	Model Name	Predicted Values
2		Fraction unbound (human)	0
3	Distribution	BBB permeability	-1.473
4		CNS permeability	-2.025
1		CYP2D6 substrate	NO
2		CYP3A4 substrate	YES
3		CYP1A2 inhibitor	NO
4	Metabolism	CYP2C19 inhibitor	YES
5		CYP2C9 inhibitor	YES
6		CYP2D6 inhibitor	NO
7		CYP3A4 inhibitor	YES
1	Excretion	Total Clearance	-0.213
2		Renal OCT2 substrate	NO
1		AMES toxicity	NO
2		Max. tolerated dose (Human)	0.253
3		hERG I inhibitor	NO
4		hERG II inhibitor	YES
5	Toxicity	Oral Rat Acute Toxicity	2.14
6		Oral Rat Chronic Toxicity	1.068
7		Hepatotoxicity	YES
8		Skin sensitisation	NO
9		T.pyriformis toxicity	0.307
10		Minnow toxicity	-0.515

4.14 Mechanism of Action of Sorafenib

Sorafenib targets a range of intracellular kinases (including CRAF, BRAF, and mutant BRAF) as well as cell surface kinases such as KIT, FLT-3, VEGFR-2, VEGFR-3, and PDGFR- . Many of these kinases play key roles in angiogenesis, the process of new blood vessel formation. By inhibiting them, Sorafenib effectively reduces blood supply to the tumor. Notably, Sorafenib specifically inhibits the Raf/MEK/ERK signaling pathway, disrupting genetic transcription processes involved in cell proliferation and thereby hindering tumor growth [102].

4.15 Sorafenib Effects on Body

Sorafenib is known to cause occasional, temporary increases in serum aminotransferase levels during treatment, which are typically mild and without symptoms. However, in rare cases, Sorafenib has been associated with clinically significant liver injury, which can be severe and potentially fatal. The maximum clinically studied dose of Sorafenib is 800 mg twice daily.

However, in rare cases, Sorafenib has been associated with clinically significant liver injury, which can be severe and potentially fatal. The maximum clinically studied dose of Sorafenib is 800 mg twice daily.

At this dosage, the most commonly reported side effects include diarrhea and skin-related reactions [102].

4.16 Comparison Between Sorafenib and Lead Compound Salfredin B11

Comparison between Sorafenib and Salfredin B11 is done to identify the better treatment for Breast cancer. Comparison is done on the basis of physiochemical properties and ADMET properties of Sorafenib and Salfredin B11. Application of Lipinski rule of five on Sorafenib and Salfredin B11 is shown in 4.19.

TABLE 4.19: Application of Lipinski rule of five on Sorafenib and Salfredin B11

Sr.No.	**Compound**	**LogP value**	**Molecular weight**	**H-Bond Acceptor**	**H-Bond Donor**
1	Sorafenib	5.5497	464.8	4	3
2	Salfredin B11	2.2468	232.23	4	1

LogP value, molecular weight and hydrogen bond donor of Salfredin B11 is less than Sorafenib.

4.17 ADMET Properties Comparison

The ADMET properties of a drug—Absorption, Distribution, Metabolism, Excretion, and Toxicity—are essential for evaluating its activity and overall effectiveness. A comparative analysis of Sorafenib and Salfredin B11 based on these properties is presented in Table 4.20.

TABLE 4.20: Comparison of ADMET Properties

Sr. No.	Property	Model Name	Sorafenib	Salfredin B11
1		Water solubility	-4.255	-3.081
2		$CaCO_2$ permeability	0.762	1.201
3		Intestinal absorption (human)	85.494	94.058
4	Absorption	Skin permeability	-2.74	-3.236
5		P-glycoprotein substrate	YES	NO
6		P-glycoprotein I inhibitor	YES	NO
7		P-glycoprotein II inhibitor	YES	NO
1		VDss (human)	-0.009	0.363
2		Fraction unbound (human)	0	0.465
3	Distribution	BBB permeability	-1.473	0.747
4		CNS permeability	-2.025	-2.827
1		CYP2D6 substrate	NO	NO
2		CYP3A4 substrate	YES	NO
3		CYP1A2 inhibitor	NO	NO
4	Metabolism	CYP2C19 inhibitor	YES	NO
5		CYP2C9 inhibitor	YES	NO
6		CYP2D6 inhibitor	NO	NO
7		CYP3A4 inhibitor	YES	NO
1	Excretion	Total Clearance	-0.213	0.481
2		Renal OCT2 substrate	NO	NO
1		AMES toxicity	NO	NO
2		Max. tolerated dose (Human)	0.253	-0.051
3		hERG I inhibitor	NO	NO
4		hERG II inhibitor	YES	NO
5	Toxicity	Oral Rat Acute Toxicity	2.14	1.701
6		Oral Rat Chronic Toxicity	1.068	2.419
7		Hepatotoxicity	YES	NO
8		Skin sensitization	NO	NO

continued on next page

Table 4.20 continued from previous page

Sr. No.	Property	Model Name	Sorafenib	Salfredin B11
9		T. pyriformis toxicity	0.307	0.494
10		Minnow toxicity	-0.515	1.492

Both are showing low water solubility, $CaCO_2$ permeability value of salfredin B11 is greater than 0.9 indicating that it has high absorbance, intestinal absorption and skin permeability of salfredin B11 is higher than sorafenib. As sorafenib is P-glycoprotein substrate depicting its low oral absorption.

Both of these are showing low VDss (human) but value of salfredin B11 is higher, fraction unbound (human) of salfredin is also greater than sorafenib making salfredin B11 more effective. BBB permeability value of sorafenib is good as compared to salfredin B11. CNS permeability values of both compounds are in recommended range. Total clearance value of salfredin B11 is positive which shows good drug clearance while of sorafenib, it is negative indicating poor drug clearance and none is OCT2 substrate.

According to toxicity values both of the compounds are non carcinogenic, MRTD of sorafenib is high as compared to salfredin B11. Sorafenib act as hERG II inhibitor which is more dangerous as it is known to us that many drugs have had withdrawn from the market just because of hERG inhibitor which cause the dysfunction of potassium channels.

Sorafenib is hepatotoxic which means it damage the liver and it is the main concern of safety for drug designing because mostly drug metabolism is done in the liver. Sorafenib also indicate minnow toxicity while salfredin B11 does not.

4.18 Comparison of Physiochemical Properties

The comparison between physiochemical properties of Sorafenib and Salfredin B11 is important step that help us to find out the drug activity manner and biochemical reactivity. Comparison between physiochemical properties of Sorafenib and Salfredin B11 is shown in Table 4.21.

TABLE 4.21: Comparison between physiochemical properties of Sorafenib and Salfredin B11

Sr.	**Compound**	**LogP**	**Rotatable Bonds**	**H-Bond Acceptor**	**H-Bond Donor**	**Molecular Weight**
1	Sorafenib	5.5497	5	4	3	464.8
2	Salfredin B11	2.2468	0	4	1	232.23

LogP value, rotatable bonds, hydrogen bond donor and molecular weight of salfredin B11 is less than sorafenib.

Chapter 5

Conclusions and Future Prospects

The objective of this study was to identify a potential compound for breast cancer treatment using a computational approach, with the goal of developing an effective drug for future use. Following an extensive literature-based data mining process, ten ligands were selected for investigation. The target proteins used for virtual screening were Topoisomerase I and Topoisomerase IIα.

Molecular docking was conducted using CB-Dock, an automated implementation of AutoDock Vina. Protein-ligand interactions were further examined with LigPlot (version 1.4.5). Based on detailed analyses of binding scores, physicochemical characteristics, and ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) properties, Salfredin B11 emerged as a promising inhibitor against breast cancer.

When compared to Sorafenib, Salfredin B11 demonstrated superior activity and safety profiles, as evidenced by favorable physicochemical and ADMET parameters. All software and tools utilized in this study are validated and widely recognized for their reliability.

5.1 Future Prospects

This identified active compound of *Nigella sativa*; Salfredin B11 can be used as medicine in near future. To prove its rank as a drug it can be used on mice for experiments and after these successful experiments, it can be introduced in the clinical trials for its validation.

Bibliography

[1] A.-u.-H. Gilani, Q. Jabeen, and M. A. U. Khan, "A review of medicinal uses and pharmacological activities of Nigella sativa," *Pak. J. Biol. Sci.*, vol. 7, no. 4, pp. 441–451, Apr. 2004.

[2] A. Tavakkoli, A. Ahmadi, B. M. Razavi, and H. Hosseinzadeh, "Black seed and its constituent thymoquinone as an antidote or a protective agent against natural or chemical toxicities," *Iran. J. Pharm. Res.*, vol. 16, no. Suppl, pp. 2–23, 2017.

[3] L. Ait Mbarek, H. A. Mouse, N. Elabbadi, M. Bensalah, A. Gamouh, R. Aboufatima, A. Benharref, A. Chait, M. Kamal, and A. Dalal, "Anti-tumor properties of black seed (Nigella sativa) extracts," *Braz. J. Med. Biol. Res.*, vol. 40, no. 6, pp. 839–847, Jun. 2007.

[4] B. H. Ali and G. Blunden, "Pharmacological and toxicological properties of Nigella sativa," *Phytother. Res.*, vol. 17, no. 4, pp. 299–305, Apr. 2003.

[5] U. P. Hedrick, *Sturtevant's Edible Plants of the World*. New York, NY, USA: Dover Publications, 1919.

[6] M. A. Randhawa, "Thymoquinone, an active principle of Nigella sativa, inhibited Fusarium solani," *Pak. J. Med. Res.*, vol. 44, no. 1, pp. 1–3, 2005.

[7] R. A. Fisher, L. Pusztai, and C. Swanton, "Cancer heterogeneity: Implications for targeted therapeutics," *Br. J. Cancer*, vol. 108, no. 3, pp. 479–485, Feb. 2013.

[8] A. Jemal, K. D. Miller, J. Ma, R. L. Siegel, S. A. Fedewa, F. Islami, S. S. Devesa, and M. J. Thun, "Higher lung cancer incidence in young women

than young men in the United States," *N. Engl. J. Med.*, vol. 378, no. 21, pp. 1999–2009, May 2018.

[9] D. Schottenfeld and J. F. Fraumeni, Jr., *Cancer Epidemiology and Prevention*. New York, NY, USA: Oxford Univ. Press, 2006.

[10] J. V. Tapia-Vieyra and J. Mas-Oliva, "Discovery of the Arp2 protein as a determining molecule in tumor cell death," *Gac. Med. Mex.*, vol. 155, no. 5, pp. 504–510, 2019.

[11] C. B. Blackadar, "Historical review of the causes of cancer," *World J. Clin. Oncol.*, vol. 7, no. 1, pp. 54–86, Feb. 2016.

[12] M. B. Shimkin, *Contrary to Nature: Being an Illustrated Commentary on Some Persons and Events of Historical Importance in the Development of Knowledge Concerning Cancer*, vol. 76, no. 720. Washington, DC, USA: U.S. Dept. Health, Educ., Welfare, Public Health Service, 1977.

[13] K. Aizawa, C. Liu, S. Tang, S. Veeramachaneni, K.-Q. Hu, D. E. Smith, and X.-D. Wang, "Tobacco carcinogen induces both lung cancer and non-alcoholic steatohepatitis and hepatocellular carcinomas in ferrets which can be attenuated by lycopene supplementation," *Int. J. Cancer*, vol. 139, no. 5, pp. 1171–1181, Sep. 2016.

[14] S. L. Poon, J. R. McPherson, P. Tan, B. T. Teh, and S. G. Rozen, "Mutation signatures of carcinogen exposure: Genome-wide detection and new opportunities for cancer prevention," *Genome Med.*, vol. 6, no. 3, pp. 1–14, Mar. 2014.

[15] J. Trafialek and W. Kolanowski, "Dietary exposure to meat-related carcinogenic substances: Is there a way to estimate the risk?" *Int. J. Food Sci. Nutr.*, vol. 65, no. 6, pp. 774–780, Sep. 2014.

[16] M. G. Cumberbatch, A. Cox, D. Teare, and J. W. F. Catto, "Contemporary occupational carcinogen exposure and bladder cancer: A systematic review and meta-analysis," *JAMA Oncol.*, vol. 1, no. 9, pp. 1282–1290, Dec. 2015.

[17] S. O. Antwi, E. C. Eckert, C. V. Sabaque, E. R. Leof, K. M. Hawthorne, W. R. Bamlet, K. G. Chaffee, A. L. Oberg, and G. M. Petersen, "Exposure to environmental chemicals and heavy metals, and risk of pancreatic cancer," *Cancer Causes Control*, vol. 26, no. 11, pp. 1583–1591, Nov. 2015.

[18] D. M. Parkin, "The global health burden of infection-associated cancers in the year 2002," *Int. J. Cancer*, vol. 118, no. 12, pp. 3030–3044, Jun. 2006.

[19] M. Seto, K. Honma, and M. Nakagawa, "Diversity of genome profiles in malignant lymphoma," *Cancer Sci.*, vol. 101, no. 3, pp. 573–578, Mar. 2010.

[20] J. C. Cigudosa, N. Z. Parsa, D. C. Louie, D. A. Filippa, S. C. Jhanwar, B. Johansson, F. Mitelman, and R. S. K. Chaganti, "Cytogenetic analysis of 363 consecutively ascertained diffuse large B-cell lymphomas," *Genes Chromosomes Cancer*, vol. 25, no. 2, pp. 123–133, Jun. 1999.

[21] E. Shtivelman, B. Lifshitz, R. P. Gale, and E. Canaani, "Fused transcript of abl and bcr genes in chronic myelogenous leukaemia," *Nature*, vol. 315, no. 6020, pp. 550–554, Jun. 1985.

[22] G. Matlashewski, P. Lamb, D. Pim, J. Peacock, L. Crawford, and S. Benchimol, "Isolation and characterization of a human p53 cDNA clone: Expression of the human p53 gene," *EMBO J.*, vol. 3, no. 13, pp. 3257–3262, Dec. 1984.

[23] World Health Organization, "WHO global meeting to accelerate progress on SDG target 3.4 on non-communicable diseases and mental health, 9–12 December 2019, Muscat, Oman," World Health Organization, Geneva, Switzerland, Tech. Rep., 2020.

[24] R. X. Zhu, W.-K. Seto, C.-L. Lai, and M.-F. Yuen, "Epidemiology of hepatocellular carcinoma in the Asia-Pacific region," *Gut Liver*, vol. 10, no. 3, pp. 332–339, May 2016.

[25] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA Cancer J. Clin.*, vol. 68, no. 6, pp. 394–424, Nov. 2018.

[26] D. R. Rice and T. M. Pawlik, "ASO author reflections: Optimizing end-of-life care for patients dying from hepatocellular carcinoma," *Ann. Surg. Oncol.*, vol. 28, no. 9, pp. 5423–5424, Sep. 2021.

[27] J. Wands, "Hepatocellular carcinoma and sex," *N. Engl. J. Med.*, vol. 357, no. 19, pp. 1974–1976, Nov. 2007.

[28] S. Mittal, J. R. Kramer, R. Omino, M. Chayanupatkul, P. A. Richardson, H. B. El-Serag, and F. Kanwal, "Role of age and race in the risk of hepatocellular carcinoma in veterans with hepatitis B virus infection," *Clin. Gastroenterol. Hepatol.*, vol. 16, no. 2, pp. 252–259, Feb. 2018.

[29] J. D. Yang, P. Hainaut, G. J. Gores, A. Amadou, A. Plymoth, and L. R. Roberts, "A global view of hepatocellular carcinoma: Trends, risk, prevention and management," *Nat. Rev. Gastroenterol. Hepatol.*, vol. 16, no. 10, pp. 589–604, Oct. 2019.

[30] C. A. Hudis and L. Gianni, "Triple-negative breast cancer: An unmet medical need," *Oncologist*, vol. 16, no. S1, pp. 1–11, 2011.

[31] A. Jindal, A. Thadi, and K. Shailubhai, "Hepatocellular carcinoma: Etiology and current and future drugs," *J. Clin. Exp. Hepatol.*, vol. 9, no. 2, pp. 221–232, Mar. 2019.

[32] Z. A. Lone, Y. Lone, S. S. Khan, A. A. Wani, and M. I. Reshi, "Hepatoprotective medicinal plants used by the Gond and Bhill tribals of district Raisen Madhya Pradesh, India," *J. Med. Plants Res.*, vol. 9, no. 12, pp. 400–406, Apr. 2015.

[33] B. Adamczyk, M. Antczak, and M. Szachniuk, "RNAsolo: A repository of cleaned PDB-derived RNA 3D structures," *Bioinformatics*, vol. 38, no. 14, pp. 3668–3670, Jul. 2022.

[34] J. Shabnam, A. S. Ahmad, S. H. Muhammad, U. Aysha, A. Rauf, and M. Sobia, "Nutritional, phytochemical potential and pharmacological evaluation of Nigella sativa (Kalonji) and Trachyspermum ammi (Ajwain)," *J. Med. Plants Res.*, vol. 6, no. 5, pp. 768–775, Feb. 2012.

[35] I. Ahmad, J. Tripathi, S. Manik, L. Umar, and J. Rabia, "Preliminary phytochemical studies of the miracle herb of the century, Nigella sativa (black seed)," *Indo Amer. J. Pharm. Res.*, vol. 3, no. 4, pp. 3000–3007, 2013.

[36] E. M. Yimer, K. B. Tuem, A. Karim, N. Ur-Rehman, and F. Anwar, "Nigella sativa L. (black cumin): A promising natural remedy for wide range of illnesses," *Evid.-Based Complement. Altern. Med.*, vol. 2019, pp. 1–16, 2019.

[37] K. M. Nadkarni, *Indian Materia Medica*. Bombay, India: Popular Prakashan, 1976, pp. 386–411.

[38] M. F. Ahmad, F. A. Ahmad, S. A. Ashraf, H. H. Saad, S. Wahab, M. I. Khan, M. Ali, S. Mohan, K. R. Hakeem, and M. T. Athar, "An updated knowledge of black seed (Nigella sativa Linn.): Review of phytochemical constituents and pharmacological properties," *J. Herb. Med.*, vol. 25, pp. 100404:1–100404:14, Feb. 2021.

[39] S. V. Tembhurne, S. Feroz, B. H. More, and D. M. Sakarkar, "A review on therapeutic potential of Nigella sativa (Kalonji) seeds," *J. Med. Plants Res.*, vol. 8, no. 3, pp. 167–177, Jan. 2014.

[40] A. Zahoor and G. Abdul, "Nigella sativa—A potential commodity in crop diversification traditionally used in healthcare," in *Breeding of Neglected and Under-Utilized Crops, Spices and Herbs*, S. N. Saxena, Ed. Enfield, NH, USA: Science Publishers, 2007, pp. 215–230.

[41] World Health Organization, "Traditional medicine," World Health Organization, Geneva, Switzerland, Fact Sheet 134, 2008.

[42] F. Jamshidi-Kia, Z. Lorigooini, and H. Amini-Khoei, "Medicinal plants: Past history and future perspective," *J. Herbmed Pharmacol.*, vol. 7, no. 1, pp. 1–7, Jan. 2018.

[43] F. Anwar and G. Muhammad, "Capparis spinosa L.: A plant with high potential for development of functional foods and nutraceutical," *Int. J. Pharmacol.*, vol. 12, no. 3, pp. 201–207, 2016.

[44] M. F. R. Hassanien, A. M. Assiri, A. M. Alzohairy, and H. F. Oraby, "Health-promoting value and food applications of black cumin essential oil: An overview," *J. Food Sci. Technol.*, vol. 52, no. 10, pp. 6136–6142, Oct. 2015.

[45] M. Adib-Hajbaghery and S. Rafiee, "Medicinal plants use by elderly people in Kashan, Iran," *Nurs. Midwifery Stud.*, vol. 7, no. 2, pp. 67–73, Apr. 2018.

[46] V. De Luca, V. Salim, S. M. Atsumi, and F. Yu, "Mining the biodiversity of plants: A revolution in the making," *Science*, vol. 336, no. 6089, pp. 1658–1661, Jun. 2012.

[47] O. A. Ghosheh, A. A. Houdi, and P. A. Crooks, "High performance liquid chromatographic analysis of the pharmacologically active quinones and related compounds in the oil of the black seed (Nigella sativa)," *J. Pharm. Biomed. Anal.*, vol. 19, no. 5, pp. 757–762, Apr. 1999.

[48] M. El-Dakhakhny, "Studies on the chemical constitution of Egyptian Nigella sativa L. seeds. II: The essential oil," *Planta Med.*, vol. 11, no. 4, pp. 465–470, Dec. 1963.

[49] V. K. Babayan, D. Koottungal, and G. A. Halaby, "Proximate analysis, fatty acid and amino acid composition of Nigella sativa L. seeds," *J. Food Sci.*, vol. 43, no. 4, pp. 1314–1315, Jul. 1978.

[50] S. Enomoto, R. Asano, Y. Iwahori, T. Narui, Y. Okada, A. N. B. Singab, and T. Okuyama, "Hematological studies on black cumin oil from the seeds of Nigella sativa L.," *Biol. Pharm. Bull.*, vol. 24, no. 3, pp. 307–310, Mar. 2001.

[51] A.-u.-R. Atta-ur-Rahman, S. Malik, S. S. Hasan, M. I. Choudhary, C.-Z. Ni, and J. Clardy, "Nigellidine—A new indazole alkaloid from the seeds of Nigella sativa," *ChemInform*, vol. 26, no. 30, pp. 1–2, Jul. 1995.

[52] A. A. Ansari, S. Hassan, L. Kenne, and T. Wehler, "Structural studies on a saponin isolated from Nigella sativa," *Phytochemistry*, vol. 27, no. 12, pp. 3977–3979, 1988.

[53] A. Haq, P. I. Lobo, M. Al-Tufail, N. R. Rama, and S. T. Al-Sedairy, "Immunomodulatory effect of Nigella sativa proteins fractionated by ion exchange chromatography," *Int. J. Immunopharmacol.*, vol. 21, no. 4, pp. 283–295, Apr. 1999.

[54] I. Kruk, T. Michalska, K. Lichszteld, A. Kladna, and H. Y. Aboul-Enein, "The effect of thymol and its derivatives on reactions generating reactive oxygen species," *Chemosphere*, vol. 41, no. 7, pp. 1059–1064, Oct. 2000.

[55] N. J. Salomi, S. C. Nair, K. K. Jayawardhanan, C. D. Varghese, and K. R. Panikkar, "Anti-tumour principles from Nigella sativa seeds," *Cancer Lett.*, vol. 63, no. 1, pp. 41–46, Feb. 1992.

[56] I. O. Farah and R. A. Begum, "Effect of Nigella sativa (N. sativa L.) and oxidative stress on the survival pattern of MCF-7 breast cancer cells," *Biomed. Sci. Instrum.*, vol. 39, pp. 359–364, 2003.

[57] M. I. Thabrew, R. R. Mitry, M. A. Morsy, and R. D. Hughes, "Cytotoxic effects of a decoction of Nigella sativa, Hemidesmus indicus and Smilax glabra on human hepatoma HepG2 cells," *Life Sci.*, vol. 77, no. 12, pp. 1319–1330, Aug. 2005.

[58] M. N. Nagi and H. A. Almakki, "Thymoquinone supplementation induces quinone reductase and glutathione transferase in mice liver: Possible role in protection against chemical carcinogenesis and toxicity," *Phytother. Res.*, vol. 23, no. 9, pp. 1295–1298, Sep. 2009.

[59] E. M. Salem, T. Yar, A. O. Bamosa, A. Al-Quorain, M. I. Yasawy, R. M. Alsulaiman, and M. A. Randhawa, "Comparative study of Nigella sativa and triple therapy in eradication of Helicobacter pylori in patients with non-ulcer dyspepsia," *Saudi J. Gastroenterol.*, vol. 16, no. 3, pp. 207–214, Jul. 2010.

[60] K. E. H. El Tahir, M. M. S. Ashour, and M. M. Al-Harbi, "The cardiovascular actions of the volatile oil of the black seed (Nigella sativa) in rats: Elucidation of the mechanism of action," *Gen. Pharmacol. Vasc. Syst.*, vol. 24, no. 5, pp. 1123–1131, Sep. 1993.

[61] J. A. Al-Sa'aidi, A. L. D. Al-Khuzai, and N. F. H. Al-Zobaydi, "Effect of alcoholic extract of Nigella sativa on fertility in male rats," *Iraqi J. Vet. Sci.*, vol. 23, no. 2, pp. 123–128, 2009.

[62] M. A. Mohammad, M. I. Mohamad, and H. Dradka, "Effects of black seeds (Nigella sativa) on spermatogenesis and fertility of male albino rats," *Res. J. Med. Med. Sci.*, vol. 4, no. 2, pp. 386–390, 2009.

[63] S. E. O'Connor, J. M. Ward, M. Watson, B. Momin, and L. C. Richardson, "Hepatocellular carcinoma—United States, 2001–2006," *Morb. Mortal. Wkly. Rep.*, vol. 59, no. 17, pp. 517–520, May 2010.

[64] J. Ferlay, H.-R. Shin, F. Bray, D. Forman, C. Mathers, and D. M. Parkin, "Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008," *Int. J. Cancer*, vol. 127, no. 12, pp. 2893–2917, Dec. 2010.

[65] S. F. Altekruse, K. A. McGlynn, and M. E. Reichman, "Hepatocellular carcinoma incidence, mortality, and survival trends in the United States from 1975 to 2005," *J. Clin. Oncol.*, vol. 27, no. 9, pp. 1485–1491, Mar. 2009.

[66] R. T.-P. Poon, S. T. Fan, C. M. Lo, C. L. Liu, and J. Wong, "Long-term survival and pattern of recurrence after resection of small hepatocellular carcinoma in patients with preserved liver function: Implications for a strategy of salvage transplantation," *Ann. Surg.*, vol. 235, no. 3, pp. 373–382, Mar. 2002.

[67] C. Verslype, E. Van Cutsem, M. Dicato, N. Arber, J. D. Berlin, D. Cunningham, A. De Gramont, E. Diaz-Rubio, M. Ducreux, T. Gruenberger, and others, "The management of hepatocellular carcinoma, current expert opinion and recommendations derived from the 10th World Congress on Gastrointestinal Cancer, Barcelona, 2008," *Ann. Oncol.*, vol. 20, no. S7, pp. vii1–vii6, Jul. 2009.

[68] J. M. Llovet, S. Ricci, V. Mazzaferro, P. Hilgard, E. Gane, J.-F. Blanc, C. A. de Oliveira, A. Santoro, J.-L. Raoul, A. Forner, and others, "Sorafenib in advanced hepatocellular carcinoma," *N. Engl. J. Med.*, vol. 359, no. 4, pp. 378–390, Jul. 2008.

[69] D. Poon, B. O. Anderson, L.-T. Chen, K. Tanaka, W. Y. Lau, E. Van Cutsem, H. Singh, W. C. Chow, L. L. Ooi, P. Chow, and others, "Management of hepatocellular carcinoma in Asia: Consensus statement from the Asian Oncology Summit 2009," *Lancet Oncol.*, vol. 10, no. 11, pp. 1111–1118, Nov. 2009.

[70] L. H. Reddy and P. Couvreur, "Nanotechnology for therapy and imaging of liver diseases," *J. Hepatol.*, vol. 55, no. 6, pp. 1461–1466, Dec. 2011.

[71] S. K. Singh, S. Singh, J. W. Lillard, Jr., and R. Singh, "Drug delivery approaches for breast cancer," *Int. J. Nanomed.*, vol. 12, pp. 6205–6218, Aug. 2017.

[72] M. E. Davis, Z. G. Chen, and D. M. Shin, "Nanoparticle therapeutics: An emerging treatment modality for cancer," in *Nanoscience and Technology: A Collection of Reviews from Nature Journals*, P. Rodgers, Ed. Singapore: World Scientific, 2010, pp. 239–250.

[73] S. K. Singh, J. W. Lillard, Jr., and R. Singh, "Reversal of drug resistance by planetary ball milled (PBM) nanoparticle loaded with resveratrol and docetaxel in prostate cancer," *Cancer Lett.*, vol. 427, pp. 49–62, Jul. 2018.

[74] G. Agarwal, P. V. Pradeep, V. Aggarwal, C.-H. Yip, and P. S. Y. Cheung, "Spectrum of breast cancer in Asian women," *World J. Surg.*, vol. 31, no. 5, pp. 1031–1040, May 2007.

[75] W. F. Anderson, K. C. Chu, S. Chang, and M. E. Sherman, "Comparison of age-specific incidence rate patterns for different histopathologic types of breast carcinoma," *Cancer Epidemiol. Biomarkers Prev.*, vol. 13, no. 7, pp. 1128–1135, Jul. 2004.

[76] C. M. Perou, T. Sørlie, M. B. Eisen, M. Van De Rijn, S. S. Jeffrey, C. A. Rees, J. R. Pollack, D. T. Ross, H. Johnsen, L. A. Akslen, and others, "Molecular portraits of human breast tumours," *Nature*, vol. 406, no. 6797, pp. 747–752, Aug. 2000.

[77] F. Cardoso, L. Cataliotti, A. Costa, S. Knox, L. Marotti, E. Rutgers, and M. Beishon, "European Breast Cancer Conference manifesto on breast centres/units," *Eur. J. Cancer*, vol. 72, pp. 244–250, Feb. 2017.

[78] H. Nagase, R. Visse, and G. Murphy, "Structure and function of matrix metalloproteinases and TIMPs," *Cardiovasc. Res.*, vol. 69, no. 3, pp. 562–573, Feb. 2006.

[79] D. Anwanwan, S. K. Singh, S. Singh, V. Saikam, and R. Singh, "Challenges in cancer and possible treatment approaches," *Biochim. Biophys. Acta Rev. Cancer*, vol. 1873, no. 1, pp. 188314:1–188314:11, Jan. 2020.

[80] Y.-J. Zhu, B. Zheng, H.-Y. Wang, and L. Chen, "New knowledge of the mechanisms of sorafenib resistance in liver cancer," *Acta Pharmacol. Sin.*, vol. 38, no. 5, pp. 614–622, May 2017.

[81] S. Hunter, P. Jones, A. Mitchell, R. Apweiler, T. K. Attwood, A. Bateman, T. Bernard, D. Binns, P. Bork, S. Burge, and others, "InterPro in 2011: New developments in the family and domain prediction database," *Nucleic Acids Res.*, vol. 40, no. D1, pp. D306–D312, Jan. 2012.

[82] E. Yuriev, J. Holien, and P. A. Ramsland, "Improvements, trends, and new ideas in molecular docking: 2012–2013 in review," *J. Mol. Recognit.*, vol. 28, no. 10, pp. 581–604, Oct. 2015.

[83] L. Santana Azevedo, F. Pretto Moraes, M. Morrone Xavier, E. Ozorio Pantoja, B. Villavicencio, J. Aline Finck, A. Menegaz Proenca, K. Beiestorf Rocha, and W. Filgueira de Azevedo, "Recent progress of molecular docking simulations applied to development of drugs," *Curr. Bioinf.*, vol. 7, no. 4, pp. 352–365, Dec. 2012.

[84] W. L. DeLano, "PyMOL: An open-source molecular graphics tool," *CCP4 Newsl. Protein Crystallogr.*, vol. 40, pp. 82–92, 2002.

[85] S. Yuan, H. S. Chan, and Z. Hu, "Using PyMOL as a platform for computational drug design," *WIREs Comput. Mol. Sci.*, vol. 7, no. 2, pp. e1298:1–e1298:9, Mar. 2017.

[86] S. I. Mostafa, "Mixed ligand complexes with 2-piperidine-carboxylic acid as primary ligand and ethylene diamine, 2,2'-bipyridyl, 1,10-phenanthroline and 2-(2-pyridyl) quinoxaline as secondary ligands: Preparation, characterization and biological activity," *Transit. Met. Chem.*, vol. 32, no. 6, pp. 769–775, Sep. 2007.

[87] S. Farabi, N. R. Saha, N. A. Khan, and M. Hasanuzzaman, "Prediction of SARS-CoV-2 main protease inhibitors from several medicinal plant compounds by drug repurposing and molecular docking approach," 2020, arXiv:2006.12349. [Online]. Available: https://arxiv.org/abs/2006.12349

[88] E. Quevillon, V. Silventoinen, S. Pillai, N. Harte, N. Mulder, R. Apweiler, and R. Lopez, "InterProScan: Protein domains identifier," *Nucleic Acids Res.*, vol. 33, no. S2, pp. W116–W120, Jul. 2005.

[89] A. Fiser, "Template-based protein structure modeling," in *Computational Biology*, vol. 22, D. Fenyö, Ed. New York, NY, USA: Humana Press, 2010, pp. 73–94.

[90] P. Petkov, H. Ivanova, T. W. Schultz, and O. Mekenyan, "Criteria for assessing the reliability of toxicity predictions: I. TIMES Ames mutagenicity model," *Comput. Toxicol.*, vol. 17, pp. 100143:1–100143:11, Feb. 2021.

[91] H. G. Stampfer, G. M. Gabb, and S. B. Dimmitt, "Why maximum tolerated dose?" *Br. J. Clin. Pharmacol.*, vol. 85, no. 10, pp. 2213–2217, Oct. 2019.

[92] A. Kolarič and N. Minovski, "Novel bacterial topoisomerase inhibitors: Challenges and perspectives in reducing hERG toxicity," *Future Med. Chem.*, vol. 10, no. 19, pp. 2241–2244, Oct. 2018.

[93] D. E. Pires, T. L. Blundell, and D. B. Ascher, "pkCSM: Predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures," *J. Med. Chem.*, vol. 58, no. 9, pp. 4066–4072, May 2015.

[94] O. Deeb and M. Goodarzi, "In silico quantitative structure toxicity relationship of chemical compounds: Some case studies," *Curr. Drug Saf.*, vol. 7, no. 4, pp. 289–297, Oct. 2012.

[95] J. Meiler and D. Baker, "RosettaLigand: Protein–small molecule docking with full side-chain flexibility," *Proteins Struct. Funct. Bioinf.*, vol. 65, no. 3, pp. 538–548, Nov. 2006.

[96] A. Umar, A. Uzairu, G. A. Shallangwa, and S. Uba, "Docking-based strategy to design novel flavone-based arylamides as potent V600E-BRAF inhibitors with prediction of their drug-likeness and ADMET properties," *Bull. Natl. Res. Cent.*, vol. 44, no. 1, pp. 179:1–179:16, Oct. 2020.

[97] A. J. Onetto and S. Sharif, "Drug distribution," in *StatPearls*. Treasure Island, FL, USA: StatPearls Publishing, 2021.

[98] J. Fan, J. Yang, and Z. Jiang, "Prediction of central nervous system side effects through drug permeability to blood–brain barrier and recommendation algorithm," *J. Comput. Biol.*, vol. 25, no. 4, pp. 435–443, Apr. 2018.

[99] D. E. Rollins and C. D. Klaassen, "Biliary excretion of drugs in man," *Clin. Pharmacokinet.*, vol. 4, no. 5, pp. 368–379, Sep. 1979.

[100] Y.-J. Zhu, B. Zheng, H.-Y. Wang, and L. Chen, "New knowledge of the mechanisms of sorafenib resistance in breast cancer," *Acta Pharmacol. Sin.*, vol. 38, no. 5, pp. 614–622, May 2017.

[101] D. S. Wishart, Y. D. Feunang, A. C. Guo, E. J. Lo, A. Marcu, J. R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, and others, "DrugBank 5.0: A major update to the DrugBank database for 2018," *Nucleic Acids Res.*, vol. 46, no. D1, pp. D1074–D1082, Jan. 2018.

[102] V. Law, C. Knox, Y. Djoumbou, T. Jewison, A. C. Guo, Y. Liu, A. Maciejewski, D. Arndt, M. Wilson, V. Neveu, and others, "DrugBank 4.0: Shedding new light on drug metabolism," *Nucleic Acids Res.*, vol. 42, no. D1, pp. D1091–D1097, Jan. 2014.