YAHIA FARES UNIVERSITY OF MEDEA

FACULTY OF SCIENCES

DEPARTMENT OF MATERIAL SCIENCES

Thesis

Submitted in fulfillment of the requirements for the **Master's** degree

In Chemistry

Specialty:
Pharmaceutical Chemistry

By:

BOUCHHER ABDELKADER

Topic:

In silico study of the anticancer potential of phenolic compounds from North African medicinal plants

Before the jury composed of:

Dr. TAHRAOUI Hichem	MCB	UYFM	President
Dr. EULDJI Amel	MCB	UYFM	Examiner
Pr. BITAM Said	Prof	UYFM	Supervisor

Année universitaire 2025/2026

ACKNOWLEDGMENTS

First and foremost, we thank Almighty Allah for granting us the courage, patience, and willpower to begin and complete this thesis.

Our thesis has now reached its conclusion after a long journey. We wish to express our sincere gratitude to all those who have helped us in assembling this modest work.

We would like to express our sincere thanks and respect to our supervisor, **Mr. Bitam Said**, for his rigor and scientific honesty, his valuable remarks and advice, and his perseverance in supervising this work.

We also wish to thank the members of the jury who have kindly accepted to evaluate our thesis.

CONTENTS

Acknowledgments i

List of Figures iv

List of Tables vi

Abstract viii

List of Abbreviations viii

1 GENERAL INTRODUCTION 1
- 1.1 General Information on Cancer 1
  - 1.1.1 Characteristics of a Cancer Cell 1
  - 1.1.2 Stages of Carcinogenesis 1
  - 1.1.3 Causes of Cancer 2
  - 1.1.4 Categories of Cancer 2
  - 1.1.5 Cancer Treatments 3
- 1.2 Anticancer Targets 3
  - 1.2.1 Kinases 4
  - 1.2.2 Hormone Receptors 4
  - 1.2.3 Anti-apoptotic Proteins (Bcl-2) 4
  - 1.2.4 DNA Repair Enzymes 4
  - 1.2.5 Cell Cycle Regulators 5
- 1.3 Phenolic Compounds 5
  - 1.3.1 Phenolic Acids 5
  - 1.3.2 Flavonoids 6
  - 1.3.3 Tannins 6
  - 1.3.4 Properties of Phenolic Compounds 7

2 MOLECULAR DOCKING 9
- 2.1 Molecular Docking 9
  - 2.1.1 Tools of Molecular Docking 9
  - 2.1.2 General Docking Protocol 11
  - 2.1.3 Scoring Functions 11
- 2.2 Protein-Ligand Interactions 12
  - 2.2.1 Hydrogen Bonding 12
  - 2.2.2 Ionic Interactions 12
  - 2.2.3 Van der Waals Interactions 14
  - 2.2.4 Hydrophobic Interactions 14
  - 2.2.5 Pi-Pi Interactions 15
- 2.3 ADMET Properties 16
  - 2.3.1 Absorption 16
  - 2.3.2 Distribution 16

2.3.3 Metabolism . . . 16
2.3.4 Excretion . . . 16
2.3.5 Toxicity . . . 17

3 MATERIALS AND METHODS . . . 18
3.1 Materials . . . 18
3.1.1 Computing Resources . . . 18
3.2 Programme . . . 18
3.2.1 MOE (Molecular Operating Environment) 2014 . . . 18
3.2.2 ChemDraw 2016 . . . 19
3.2.3 BIOVIA Discovery Studio 2025 . . . 19
3.2.4 DataWarrior . . . 20
3.3 Databases and Web Tools . . . 21
3.3.1 RCSB Protein Data Bank (PDB) . . . 21
3.3.2 COCONUT Database . . . 21
3.3.3 PubChem . . . 22
3.3.4 The Open Targets Platform . . . 23
3.3.5 SwissADME . . . 23
3.3.6 ProTox 3.0 . . . 24
3.4 Methods . . . 25
3.4.1 Choice of proteins . . . 25
3.4.2 Preparation of the proteins . . . 25
3.4.3 Choice of ligands . . . 27
3.4.4 Preparation of the ligands . . . 30
3.4.5 Docking Protocol: . . . 32

4 RESULTS AND DISCUSSION . . . 33
4.1 ADMET Filtering . . . 33
4.1.1 Interpretation . . . 33
4.2 Molecular Docking . . . 39
4.2.1 5Y2S protein . . . 39
4.2.2 6Q4G protein . . . 40
4.2.3 6UDV protein . . . 41
4.2.4 7NFW protein . . . 42
4.2.5 8A27 protein . . . 43
4.3 Study of protein–ligand interactions . . . 44
4.3.1 5Y2S–ligand interactions . . . 44
4.3.2 6Q4G–ligand interactions . . . 48
4.3.3 6UDV–ligand interactions . . . 51
4.3.4 7NFW–ligand interactions . . . 54
4.3.5 8A27–ligand interactions . . . 57

5 GENERAL CONCLUSION . . . 61
5.1 General conclusion . . . 61

BIBLIOGRAPHY . . . 61

LIST OF FIGURES

Figure 2.1 Receptor-ligand interaction . . . 10
Figure 2.2 hydrogen bond . . . 13
Figure 2.3 Interaction between two molecules with opposite charges . . 14
Figure 2.4 Hydrophobic patch . . . 15
Figure 3.1 MOE interface. . . . 19
Figure 3.2 ChemDraw 2016 structure drawing. . . . 19
Figure 3.3 BIOVIA Discovery Studio molecular modeling interface. . . 20
Figure 3.4 DataWarrior visualization of chemical space. . . . 20
Figure 3.5 RCSB PDB website interface. . . . 21
Figure 3.6 COCONUT database overview. . . . 22
Figure 3.7 PubChem. . . . 22
Figure 3.8 Open Targets Platform scoring interface. . . . 23
Figure 3.9 SwissADME prediction output. . . . 24
Figure 3.10 ProTox 3.0 toxicity prediction interface. . . . 24
Figure 3.11 Three-dimensional representations of the five target proteins. 26
Figure 3.12 Presentation diagram of the computational protocol for molecular docking. . . . 32
Figure 4.1 Molecular docking interactions of OLAPARIB with receptor 5Y2S. . . . 45
Figure 4.2 Molecular docking interactions of compound 3 with receptor 5Y2S. . . . 45
Figure 4.3 Molecular docking interactions of compound 1 with receptor 5Y2S. . . . 46
Figure 4.4 Molecular docking interactions of compound 5 with receptor 5Y2S. . . . 46
Figure 4.5 Molecular docking interactions of compound 6 with receptor 6Q4G. . . . 48
Figure 4.6 Molecular docking interactions of compound 11 with receptor 6Q4G. . . . 49
Figure 4.7 Molecular docking interactions of DINACICLIB with receptor 6Q4G. . . . 49
Figure 4.8 Molecular docking interactions of compound 8 with receptor 6Q4G. . . . 50
Figure 4.9 Molecular docking interactions of compound 8 with receptor 6UDV. . . . 51
Figure 4.10 Molecular docking interactions of compound 11 with receptor 6UDV. . . . 52
Figure 4.11 Molecular docking interactions of compound 1 with receptor 6UDV. . . . 52
Figure 4.12 Molecular docking interactions of compound Obatoclax mesylate with receptor 6UDV. . . . 53

Figure 4.13 Molecular docking interactions of FULVESTRANT with receptor 7NFW. 55
Figure 4.14 Molecular docking interactions of compound 11 with receptor 7NFW. 55
Figure 4.15 Molecular docking interactions of compound 5 with receptor 7NFW. 56
Figure 4.16 Molecular docking interactions of compound 1 with receptor 7NFW. 56
Figure 4.17 Molecular docking interactions of Lapatinib Ditosylate with receptor 8A27. 58
Figure 4.18 Molecular docking interactions of compound 3 with receptor 8A27. 58
Figure 4.19 Molecular docking interactions of compound 5 with receptor 8A27. 59
Figure 4.20 Molecular docking interactions of compound 11 with receptor 8A27. 59

LIST OF TABLES

Table 3.1 Computer specifications used for the study. 18
Table 3.2 Properties of the selected target proteins 25
Table 3.3 The characteristics of Lipinski, Ghose, Veber, Egan, and Muegge drug-likeness rules according to SwissAdme[76]. . 27
Table 3.4 drugs and their 2D structures 31
Table 4.1 ADME/Tox properties of selected phenolic compounds (Compounds 1-15). 34
Table 4.2 List of selected phenolic compounds. 35
Table 4.3 Molecular Docking Results: of receptor 5Y2S 39
Table 4.4 Molecular Docking Results: of receptor 5Y2S (The score of the drug and the highest three scores of the compounds) . . 40
Table 4.5 Molecular Docking Results: of receptor 6Q4G 40
Table 4.6 Molecular Docking Results: of receptor 6Q4G (The score of the drug and the highest three scores of the compounds) . . 41
Table 4.7 Molecular Docking Results: of receptor 6UDV 41
Table 4.8 Molecular Docking Results: of receptor 6UDV (The score of the drug and the highest three scores of the compounds) . . 42
Table 4.9 Molecular Docking Results: of receptor 7NFW 42
Table 4.10 Molecular Docking Results of 7NFW receptor (The score of the drug and the highest three scores of the compounds) . . 43
Table 4.11 Molecular Docking Results: of receptor 8A27 43
Table 4.12 Molecular Docking Results of 8A27 receptor (The score of the drug and the highest three scores of the compounds) . . 44
Table 4.13 MOE interaction report for the top-scoring ligands with the 5Y2S receptor. 44
Table 4.14 MOE interaction report for the top-scoring ligands with the 6Q4G receptor. 48
Table 4.15 MOE interaction report for the top-scoring ligands with the 6UDV receptor. 51
Table 4.16 MOE interaction report for the top-scoring ligands with the 7NFW receptor. 54
Table 4.17 MOE interaction report for the top-scoring ligands with the 8A27 receptor. 57

Abstract

Cancer remains a leading cause of death, necessitating the discovery of novel therapeutic agents. This study employed molecular docking to investigate the binding interactions between 15 phenolic compounds and five anticancer protein targets (5Y2S, 6Q4G, 6UDV, 7NFW, and 8A27). Following ADMET filtering using SwissADME and ProTox 3.0, molecular docking was performed using MOE software. Our results identified compounds 3, 6, 8, and 11 as the most promising inhibitors, with docking scores ranging from -6.84 to -7.44 kcal/mol. Notably, compounds 6 and 11 outperformed the reference drug Dinaciclib against 6Q4G, while compounds 8 and 11 showed superior activity against 6UDV compared to Obatoclax mesylate. Protein-ligand interaction analysis revealed that binding stability is governed by hydrogen bonds and hydrophobic contacts with key active-site residues. All selected compounds exhibited excellent drug-likeness properties with no Lipinski rule violations. This study successfully identifies compounds 3, 6, 8, and 11 as promising lead candidates for further anticancer drug development.

Keywords: Cancer; Phenolic compounds; Molecular docking; ADMET; Drug discovery; Protein-ligand interactions.

résumé

Le cancer demeure l'une des principales causes de mortalité dans le monde, ce qui rend nécessaire la découverte de nouveaux agents thérapeutiques. Cette étude a eu recours au docking moléculaire pour étudier les interactions de liaison entre 15 composés phénoliques et cinq protéines cibles anticancéreuses (5Y2S, 6Q4G, 6UDV, 7NFW et 8A27). Après un filtrage ADMET à l'aide des serveurs SwissADME et ProTox 3.0, le docking moléculaire a été réalisé à l'aide du logiciel MOE. Nos résultats ont identifié les composés 3, 6, 8 et 11 comme étant les inhibiteurs les plus prometteurs, avec des scores de docking allant de -6,84 à -7,44 kcal/mol. Notamment, les composés 6 et 11 ont surpassé le médicament de référence Dinaciclib contre la cible 6Q4G, tandis que les composés 8 et 11 ont montré une activité supérieure contre 6UDV par rapport à l'Obatoclax mésylate. L'analyse des interactions protéine-ligand a révélé que la stabilité de la liaison est régie par des liaisons hydrogène et des contacts hydrophobes avec des résidus clés du site actif. Tous les composés sélectionnés ont présenté d'excellentes propriétés de type médicamenteux (drug-likeness) sans aucune violation de la règle de Lipinski. Cette étude identifie avec succès les composés 3, 6, 8 et 11 comme des candidats principaux prometteurs pour le développement futur de médicaments anticancéreux.

Mots-clés: Cancer ; Composés phénoliques ; Docking moléculaire ; ADMET ; Découverte de médicaments ; Interactions protéine-ligand.

ملخص

يُعد السرطان أحد الأسباب الرئيسية للوفاة في العالم، مما يستلزم اكتشاف عوامل علاجية جديدة. استخدمت هذه الدراسة تقنية الالتحام الجزيئي (Docking) للتحقيق في تفاعلات الارتباط بين 15 مركباً فينولياً وخمسة بروتينات مستهدفة مضادة للسرطان (5Y2S, 6Q4G, 6UDV, 7NFW و 8A27) .بعد إجراء الترشيح المعتمد على خصائص الامتصاص والتوزيع والاستقلاب والإطراح والسمية (ADMET) باستخدام خادمي SwissADME وProTox 3.0 ، تم إجراء الالتحام الجزيئي باستخدام برنامجMOE. كشفت نتائجنا أن المركبات 3، 6، 8، و 11 هي المثبطات الأكثر واعدة، حيث تراوحت درجات الالتحام بين -6.84 و -7.44 كيلو كالوري/مول. ومن الجدير بالذكر أن المركبين 6 و 11 تفوقا على العقار المرجعي ديناسيكليب (Dinaciclib) ضد مستقبل 6Q4G ، بينما أظهر المركبان 8 و 11 نشاطاً متفوقاً ضد مستقبل 6UDV مقارنة بأوباتوكلاكس ميسيلات (Obatoclax mesylate). أظهر تحليل تفاعلات البروتين-ربيطة أن استقرار الارتباط يعتمد على الروابط الهيدروجينية والتماسات الكارهة للماء مع البقايا الأساسية في الموقع النشط. أظهرت جميع المركبات المختارة خصائص دوائية ممتازة (شبيهة بالأدوية) دون أي انتهاكات لقاعدة ليبينسكي (Lipinski). تحدد هذه الدراسة بنجاح المركبات 3، 6، 8، و 11 كمرشحات رئيسية واعدة لمزيد من التطوير في مجال الأدوية المضادة للسرطان.

الكلمات المفتاحية: السرطان؛ المركبات الفينولية؛ الالتحام الجزيئي؛ADMET ؛ اكتشاف الأدوية؛ تفاعلات البروتين-ربيطة.

LIST OF ABBREVIATIONS

ADMET	Absorption, Distribution, Metabolism, Excretion, Toxicity
ATP	Adenosine Triphosphate
BBB	Blood-Brain Barrier
Bcl-2	B-cell Lymphoma 2
BH	Bcl-2 Homology
CADD	Computer-Aided Drug Design
CDK	Cyclin-Dependent Kinase
COCONUT	Collection of Open Natural products
CSM	Computed Structure Model
CYP	Cytochrome P450
DNA	Deoxyribonucleic Acid
EGFR	Epidermal Growth Factor Receptor
ER	Estrogen Receptor
GI	Gastrointestinal
GIT	Gastrointestinal Tract
H-bond	Hydrogen Bond
hERG	human Ether-à-go-go-Related Gene
LD50	Median Lethal Dose
MDR	Multidrug Resistance
MOE	Molecular Operating Environment
MW	Molecular Weight
NCI	National Cancer Institute
NIH	National Institutes of Health
NMR	Nuclear Magnetic Resonance
PDB	Protein Data Bank
P-gp	P-glycoprotein
PSA	Polar Surface Area
QSAR	Quantitative Structure-Activity Relationship
RMSD	Root Mean Square Deviation
RB	Rotatable Bonds
TPSA	Topological Polar Surface Area
WHO	World Health Organization

1 GENERAL INTRODUCTION

1.1 GENERAL INFORMATION ON CANCER

Hippocrates is known to have first referenced cancer describing tumors using the word Karkinos from which the modern term is derived. The reasoning behind this was to relate the precise shape of the tumor to that of a crab, whose shell is often seen to have ray like extensions. Cancer is a disease characterized by uncontrolled cell proliferation and the ability to invade surrounding tissues and metastasize[1].

The World Health Organization (WHO) lists cancer as the second leading cause of death globally. The number of cancer cases has doubled at the beginning of the last two decades and is believed to reach three times that by 2030 [1, 2].

1.1.1 *Characteristics of a Cancer Cell*

Uncontrolled proliferation is manifested by the fact that cancer cells have lower requirements for extracellular growth factors and no longer exhibit contact inhibition. Many of them no longer differentiate normally, which explains their continuous proliferation in vivo. The fact that these cells escape apoptosis also significantly contributes to tumor growth [3].

1.1.2 *Stages of Carcinogenesis*

1.1.2.1 *Initiation*

This stage corresponds to the appearance of transformed cells possessing cancerous characteristics [4].

1.1.2.2 *Promotion*

Cell multiplication increases in parallel with DNA mutations in the nucleus [4].

1.1.2.3 *Progression*

This stage corresponds to the progression phase of carcinogenesis and is typically associated with clinically detectable tumors [4].

1.1.3 *Causes of Cancer*

1.1.3.1 *Extrinsic Factors*

It is difficult to pinpoint the exact cause of cancer. Several well-known environmental factors exist. These include hygiene, which involves diet quality and physical activity [5]. Alcohol consumption and smoking are also important factors [6]. Certain chronic infections, mainly caused by viruses affecting a specific organ, can promote the appearance of a tumor in that organ. Finally, certain industrial chemicals and radiation, particularly ultraviolet rays, can also cause cancer depending on the degree of exposure [7].

1.1.3.2 *Intrinsic Factors*

There are also so-called intrinsic risk factors, linked to genetic factors involving age and sex. The detection of these risk factors also helps to adapt the screening of the pathology [5].

1.1.4 *Categories of Cancer*

Unlike benign tumors, malignant tumors lead to the spontaneous death of the patient. This significant evolutionary difference is strongly correlated with macroscopic and histological criteria [8].

1.1.4.1 *Benign Tumors*

Benign tumors appear locally and remain confined to the tissue of origin. Their growth is slow. However, they can reach a considerable volume and weight. In some cases, due to their location and metabolic disturbances, they can lead to serious complications [8].

1.1.4.2 *Malignant Tumors*

Malignant tumors generally develop rapidly. They lead to distant tumor dissemination, particularly via the lymphatic and blood systems, with the emergence and

development of secondary tumors: metastases. If left untreated, the natural course is inevitably death [8].

1.1.5 *Cancer Treatments*

The choice of treatment depends on the type of cancer, its stage of progression, and the patient. This includes treatments such as radiotherapy, chemotherapy, or immunotherapy, etc. [1].

- **Surgery**: The goal of surgery is to remove the tumor and assess its severity and extent [9].
- **Radiotherapy**: It involves using radiation to destroy cancer cells by blocking their ability to multiply, while preserving healthy tissues and nearby organs as much as possible [10].
- **Chemotherapy**: This is a systemic treatment based on the use of drugs to eliminate cancer cells. Stage 0 and I cancers are not treated with chemotherapy. It is administered after surgery in the case of advanced cancers (stages II, III, and IV) [11].
- **Immunotherapy**: This treatment involves administering substances that stimulate the body's immune defenses to increase the elimination of cancer cells [12].
- **Targeted Therapy**: The term "targeted therapy" refers to a therapy directed against an existing molecular target that is supposed to play a role in the neoplastic transformation of the cancer cell [13].
- **Phytotherapy**: Phytotherapy is the broadest and oldest treatment; it relies on the use of different parts of plants, such as leaves, flowers, roots, and seeds, for their medicinal properties [14].

1.2 ANTICANCER TARGETS

Molecular entities or biological processes that, when modulated pharmacologically, disrupt the survival, proliferation, or dissemination of malignant cells. Unlike conventional cytotoxic agents that non-selectively affect rapidly dividing cells, anticancer targets are typically aberrantly activated, overexpressed, or structurally altered in tumors relative to normal tissues, thereby offering a therapeutic window for targeted interventions.

1.2.1 *Kinases*

Protein kinases are enzymes that regulate cellular signaling by transferring phosphate groups from ATP to target proteins. In cancer, gain-of-function mutations, amplifications, or chromosomal translocations frequently result in constitutive kinase activation, driving uncontrolled proliferation and survival. Consequently, small-molecule ATP-competitive or allosteric inhibitors have emerged as cornerstone targeted therapeutics, exemplified by imatinib in chronic myeloid leukemia [15].

1.2.2 *Hormone Receptors*

Nuclear or membrane-bound proteins that mediate cellular responses to steroid hormones (e.g., estrogen, androgen) or peptide hormones. In hormone-dependent malignancies such as breast and prostate cancers, aberrant expression or ligand-independent activation of these receptors (e.g., estrogen receptor α, androgen receptor) sustains transcriptional programs promoting cell cycle progression and survival. Endocrine therapies including selective receptor modulators, downregulators, and aromatase inhibitors exploit this dependency by antagonizing receptor signaling or depleting ligand availability [16].

1.2.3 *Anti-apoptotic Proteins (Bcl-2)*

A subset of the B-cell lymphoma 2 (Bcl-2) protein family characterized by four Bcl-2 homology (BH) domains that preserve mitochondrial outer membrane integrity by sequestering pro-apoptotic effectors (e.g., BAX, BAK). Overexpression of Bcl-2 and its close homologs (Bcl-xL, Mcl-1) is a common hallmark of numerous hematological and solid tumors, conferring resistance to intrinsic apoptotic stimuli. Therapeutic targeting has been achieved with BH3-mimetics, such as venetoclax, which competitively displace pro-apoptotic proteins from anti-apoptotic guardians, thereby lowering the apoptotic threshold [17].

1.2.4 *DNA Repair Enzymes*

A diverse ensemble of proteins responsible for recognizing and correcting chemically distinct forms of genomic damage, including base mismatches, single-strand breaks, double-strand breaks, and crosslinks. In cancer, certain repair pathways (e.g., base excision repair, homologous recombination) are upregulated in response

to replication stress or therapy-induced genotoxicity, enabling tumor cell survival despite accumulating mutations. Conversely, synthetic lethality strategies exploit pre-existing repair deficiencies for instance, poly(ADP-ribose) polymerase (PARP) inhibitors in *BRCA1/2*-deficient tumors by blocking compensatory repair mechanisms [18].

1.2.5 *Cell Cycle Regulators*

The cell cycle is regulated by cyclins, CDKs, and their inhibitors, which control progression through different cell cycle phases. Neoplastic transformation frequently involves genetic or epigenetic alterations that dysregulate these controls, such as amplification of *CCND1* (cyclin D1) or loss-of-function mutations in *CDKN2A* (p16). Pharmacological inhibition of CDK4/6 has proven clinically effective in hormone receptor-positive breast cancer, restoring checkpoint control and inducing quiescence in tumor cells [19].

1.3 PHENOLIC COMPOUNDS

Phenolic compounds are secondary metabolites with high molecular weight. They are widely distributed in plants and play a significant role in contributing to the color and taste of fruits and vegetables [20, 21].

Their characteristic basic structure consists of one or more aromatic rings to which one or more hydroxyl groups are directly attached, either free or engaged in another function (ether, ester). Among these metabolites are phenolic acids, flavonoids, and tannins. They are classified according to the presence of different substituents on the rings and their degree of saturation [22].

1.3.1 *Phenolic Acids*

Phenolic acids are widely distributed in fruits, stems, and leaves of vegetables [23]. They are generally soluble in polar organic solvents. Phenolic acids are divided into two classes: those derived from benzoic acid and those derived from cinnamic acid.

- **Phenolic acids derived from benzoic acid** [24]:
  - These are derivatives of benzoic acid.
  - They have a basic general structure of the type (C6-C1).
  - They often exist as esters or glycosides.

- **Phenolic acids derived from cinnamic acid** [24]:
  - These are derived from cinnamic acid.
  - They have a basic general structure of the type (C6-C3).
  - They often exist in combined form with organic molecules.

1.3.2 *Flavonoids*

The term flavonoid derives from the Latin word *flavus* (meaning yellow) [25]. More than 4,000 varieties of flavonoids have been identified in plants [26]. With the exception of chalcones, several are well known as very potent antioxidants in vitro due to their free radical scavenging activity and metal ion chelation capacity [27]. They can be found in all parts of plants: roots, stems, wood, bark, leaves, seeds, flowers, and fruits [28], where they are responsible for the coloration of these organs [29].

The basic structure of these compounds (Figure 4) is diphenylpropane with 15 carbon atoms (C6-C3-C6) [30]. They are characterized by the presence of two benzoic rings (A and B) and an oxygenated heterocycle (ring C) [31, 32].

1.3.3 *Tannins*

Tannins are substances found mainly in barks. They are polymers (polyphenols) present in polymerized form, soluble in water and insoluble in nonpolar organic solvents [33], with molecular weights ranging between 500 and 3000 [34]. After coagulation, they form very stable compounds with proteins. They have the property of transforming fresh animal skin into a non-putrescible material, leather [35]. They also possess other properties such as the ability to precipitate alkaloids, gelatin, and other proteins [36].

According to the nature of their molecular assemblies, tannins are subdivided into two groups: hydrolysable tannins and condensed tannins [37].

- **Hydrolysable Tannins** Hydrolysable tannins are heteropolymers whose chemical hydrolysis releases a sugar, generally glucose, and a phenolic acid. The released acid may be gallic acid in the case of gallotannins, or hexahydroxydiphenic acid and its derivatives such as ellagic acid in the case of ellagitannins [38].

- **Condensed Tannins (Proanthocyanidins)** Condensed tannins are polymers of certain flavonoids, composed of flavan-3-ol units linked together by carbon-carbon bonds, most often between C4 and C8 or C4 and C6. They are generally polymers of catechin or catechol and do not hydrolyze under the action

of dilute mineral acids [39, 40]. They play a role in protection against ultraviolet radiation, drought, and natural predators (insects and herbivores) [41].

1.3.4 *Properties of Phenolic Compounds*

1.3.4.1 *Physicochemical Properties*

Polyphenols are generally aromatic compounds soluble in polar solvents such as ethanol, methanol, butanol, acetone, dimethyl sulfoxide, water, and others [42]. Less polar polyphenols, including isoflavones, flavanones, flavones, and flavonols, exhibit greater solubility in solvents such as diethyl ether and chloroform [43].

1.3.4.2 *Biological Properties*

In plants, phenolic compounds are involved in various processes, including seed germination and plant growth [44]. In recent years, research on phenolic compounds has intensified due to their anti-inflammatory, antibacterial, antifungal, antioxidant, and even anticancer activities [45].

- **Anti-inflammatory Activity** The anti-inflammatory properties of polyphenolic compounds may be attributed to their ability to inhibit enzymes involved in inflammatory processes [46] as well as their antioxidant activities. Matricaria, also known as German chamomile or common chamomile, is a medicinal plant used for its antispasmodic and anti-inflammatory properties [44].

- **Antibacterial Activity** Polyphenols, particularly flavonoids and tannins, are recognized for their toxicity toward microorganisms [47].

- **Anticancer Activity** Flavonoids and other phenols can play a preventive role in cancer development [10]. They intervene during the initiation stage by trapping electrophilic mutagens or by stimulating the repair of mutated DNA. During the promotion and progression stages, they act as tumor-suppressing agents through various mechanisms, including the induction of apoptosis and the inhibition of cell proliferation [48].

- **Antioxidant Activity** The antioxidant activity of phenolic compounds is due to their ability to scavenge free radicals, donate hydrogen atoms and electrons, and chelate metal cations. The structure of phenolic compounds is the determining factor for their activity. For phenolic acids, antioxidant activity increases proportionally with the degree of hydroxylation and the presence

of the C=CH-COOH group [49]. For flavonoids, the structure-activity relationship is generally more complicated than for phenolic acids due to the greater complexity of the flavonoid molecule [50].

2

MOLECULAR DOCKING

2.1 MOLECULAR DOCKING

Molecular docking aims to predict the structure of a molecular complex from individual molecules, combining various approaches to study interaction modes between two molecules. Docking software is therefore highly valuable in biology, pharmacy, and medicine, since most active pharmaceutical ingredients are small molecules (ligands) that interact with a biological target of therapeutic interest. As the macromolecular receptor is most often a protein, the term "docking" alone is commonly used to designate "protein-ligand docking".

The goal of molecular docking is to determine the interaction mode of a complex formed by two or more molecules, by searching for favorable spatial orientations and conformations for the binding of a ligand to a receptor.

A docking simulation essentially comprises two complementary steps: docking and scoring.

- **Docking (first step)**: This is the selection step, which involves placing the ligand into the protein's active site and sampling possible conformations, positions, and orientations (poses), retaining only those that represent the most favorable interaction modes.
- **Scoring (second step)**: This is the ranking step, which evaluates the affinity between the ligand and the protein and assigns a score to the poses obtained during the docking phase. This score allows the best pose to be selected among all those proposed [51].

2.1.1 *Tools of Molecular Docking*

Docking software is highly useful in biology, pharmacy, and medicine because most active ingredients are small molecules (ligands) that interact with a biological target of therapeutic interest (protein).

2.1.1.1 *Ligand*

A ligand (from Latin *ligandum*, binding) is an atom, ion, or molecule bearing chemical functions that allow it to bind to one or more central atoms or ions. The term ligand is most often used in coordination chemistry and organometallic chemistry (branches of inorganic chemistry).

The structure of ligands can be obtained by two means:

- The first involves using ligands from the PDB or from the literature, which can be drawn, optimized, and saved in various formats (pdb, mol, mol2, etc.) using molecular construction software such as Titan.
- The second involves using collections of molecules (chemical libraries), both commercial and virtual, consisting of databases of chemical structures called chemical libraries or chemical spaces. Among the most widely used, we cite PubChem.

2.1.1.2 *Receptor*

The receptor is macromolecular, most often a protein. To obtain receptors, we use 3D structures of proteins that are freely accessible in the PDB database (`http://www.rcsb.org/pdb/`). This is a large archive of structural data of biological macromolecules: proteins and nucleic acids (DNA, RNA). The structural data are obtained by X-ray crystallography and NMR spectroscopy and are freely accessible on the Internet via defined websites of its member organizations including RCSB PDB, PDBj, PDBe, etc.[52].

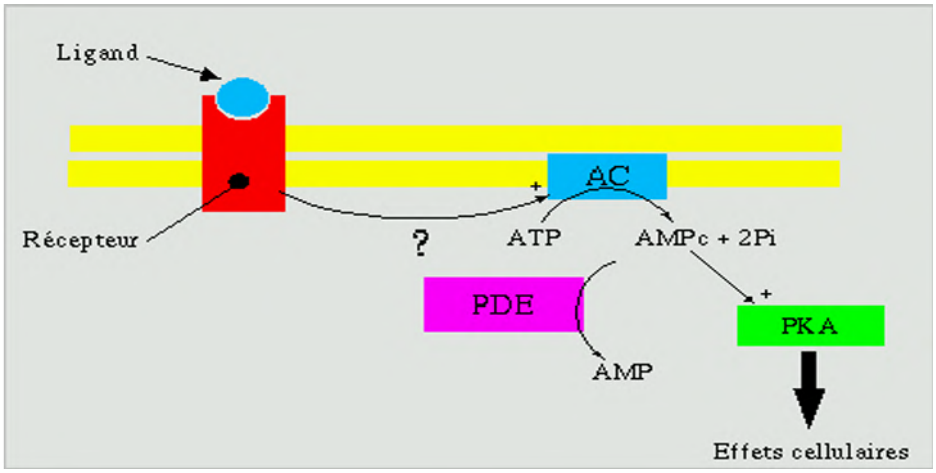


Figure 2.1: Receptor-ligand interaction

2.1.2 *General Docking Protocol*

The approaches currently used are exclusively computational and evaluated by visualization tools. These approaches can be broken down into four to five successive phases:

- Choice of the protein representation mode (all-atom, pseudo-atom, grid, etc.);
- Conformational exploration (rigid-body ligand position/orientation and/or flexible ligand position/orientation/shape);
- Minimization of the evaluation function of interaction energy (or scoring function) for conformations resulting from the exploration;
- Grouping by similarity and classification through finer score evaluation, accompanied by a non-automatic step of visual evaluation of results when the score does not allow discrimination of the native conformation from the various generated conformations;
- An optional step of refining selected complexes by minimization or molecular dynamics;
- A search algorithm to explore possible binding modes, a mechanism to place the ligand in the binding site, and a scoring function to rank the different binding modes [53].

2.1.3 *Scoring Functions*

The scoring function provides a method for ranking the positioning of ligands according to their relative proportion. Ideally, the score should be directly related to the binding affinity of the ligand with the protein, such that ligands with the best score are the best binders [54]. The score is a valuable numerical measure used to evaluate the level of complexation between a ligand and a receptor. This evaluation is based on force fields, which are mathematical functions describing the energy of a system as the sum of various terms derived from molecular mechanics. In other words, the score predicts the affinity, which can be represented by the change in free energy (ΔG) resulting from complex formation:

$$\Delta G = \Delta G_{\text{complex}} - \Delta G_{\text{ligand}} - \Delta G_{\text{protein}}$$

2.2 PROTEIN-LIGAND INTERACTIONS

The functioning of biological systems is based on remarkable assembly and recognition mechanisms involving low-energy non-covalent molecular interactions. The biological messages integrated into their structure can only be expressed through specific, often reversible, interactions established with other selected partners.

The challenge during ligand docking into a protein is predicting the conformation and orientation of the ligand relative to the active site of the target protein, starting from the initial conformation. For each target protein with a known structure, docking can be the key to designing new drugs [55–57]. Numerous types of non-covalent interactions have been identified in protein-ligand complexes. Among the interactions that can exist between the ligand (considered a small organic molecule) and the protein are polar interactions (hydrogen bonds, ionic bonds) and hydrophobic interactions (resulting from contact between hydrophobic groups).

2.2.1 *Hydrogen Bonding*

The hydrogen bond is a non-covalent dipole-dipole chemical bond between two molecules or between two groups within the same molecule. It essentially consists of an interaction between two molecules, one of which possesses an electron-donor atom (O, N, F) and the other possesses an electron-acceptor H atom (OH, NH2). This type of bond results from a partial transfer of an unpaired electron onto the H group. Hydrogen bonds can be intramolecular or intermolecular [58]. This bond is illustrated below with examples. Hydrogen bonds play a very important role in chemistry and biochemistry. They are bonds that give water its unique properties, conferring upon it the ability to associate with large structures. Furthermore, they act as stabilizers of the secondary structure of biological macromolecules. Hydrogen bonds exist in proteins (tertiary and quaternary structure), nucleic acids (double helix stabilized by H-bonds), and molecular interactions (recognition of small molecules by receptors, etc.). Hydrogen bonds are stronger than van der Waals bonds; their energies are estimated between 3 and 9 kcal/mol.

2.2.2 *Ionic Interactions*

Interactions between molecules carrying electric charges are governed by Coulomb's law. The potential energy of interaction between two electric charges Q1 and Q2 separated by a distance d is obtained by calculating the work required to

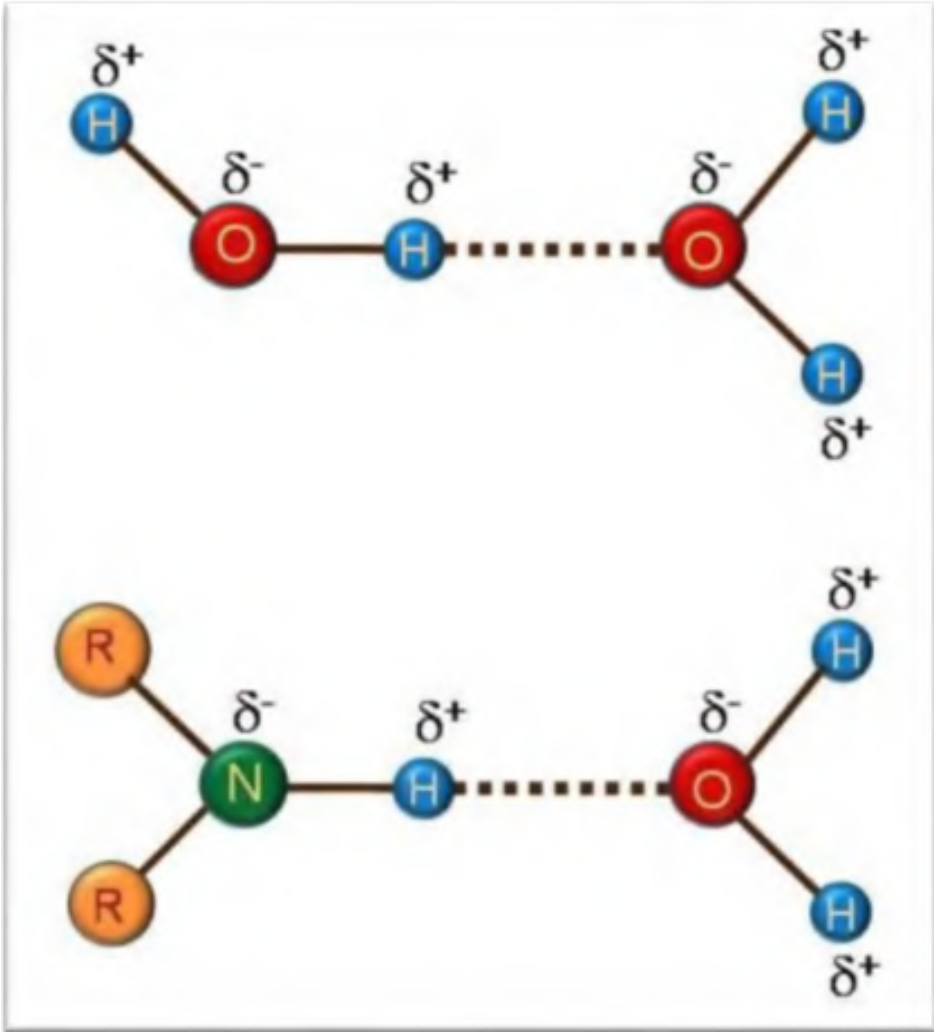


Figure 2.2: hydrogen bond

separate these two charges to an infinite distance. The potential energy (V) of two molecules with opposite charges is expressed by the following equation:

$$V = \frac{Q_1 Q_2}{4\pi\varepsilon\varepsilon_0 d}$$

Where:

- ε_0 is the permittivity of free space,
- ε is the permittivity of the medium in which the charges evolve,
- d is the distance between the two charges.

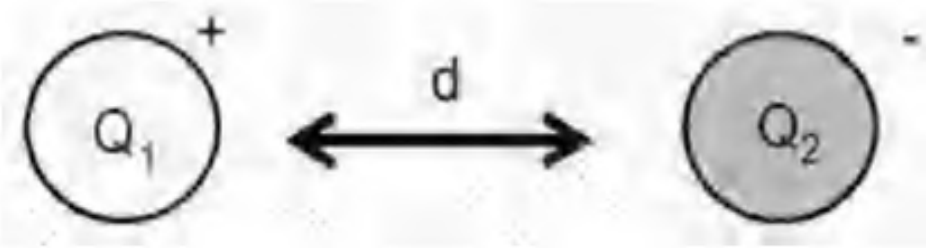


Figure 2.3: Interaction between two molecules with opposite charges

2.2.3 *Van der Waals Interactions*

These interactions were studied by J.D. van der Waals, a Dutch physicist and Nobel Prize in Physics 1910. They are based on interactions between the dipoles formed by molecules. Three types of these interactions are distinguished:

- **Permanent dipole – permanent dipole interaction (Keesom orientation effect)**: This type of interaction develops between two polar molecules.
- **Permanent dipole – induced dipole interaction (Debye induction effect)**: This type of interaction develops between a polar molecule and any molecule (polar or nonpolar).
- **Instantaneous dipole – instantaneous dipole interaction (London dispersion)**: This type of interaction develops between any two molecules (polar or nonpolar).

These interactions are very weak, but in the case of macromolecules, their large number produces a significant total force [59].

2.2.4 *Hydrophobic Interactions*

Nonpolar molecules tend to aggregate in aqueous environments through hydrophobic interactions. These are interactions between molecules or groups that have very little affinity for the solvent in which they are dissolved (water). The groups will bind on the side with the smallest contact surface with water. Thus, the groups attract each other through dispersion forces (London). Several hydrophobic amino acids come into contact at the interface through hydrophobic complementarity. These interactions provide stabilizing force for many complexes. Nonpolar amino acids tend to create "hydrophobic patch" interactions with each other [60].

Various types of hydrophobic interactions exist. Among them are those formed between the centers of two aromatic rings, an amide, or methyl groups. The geo-

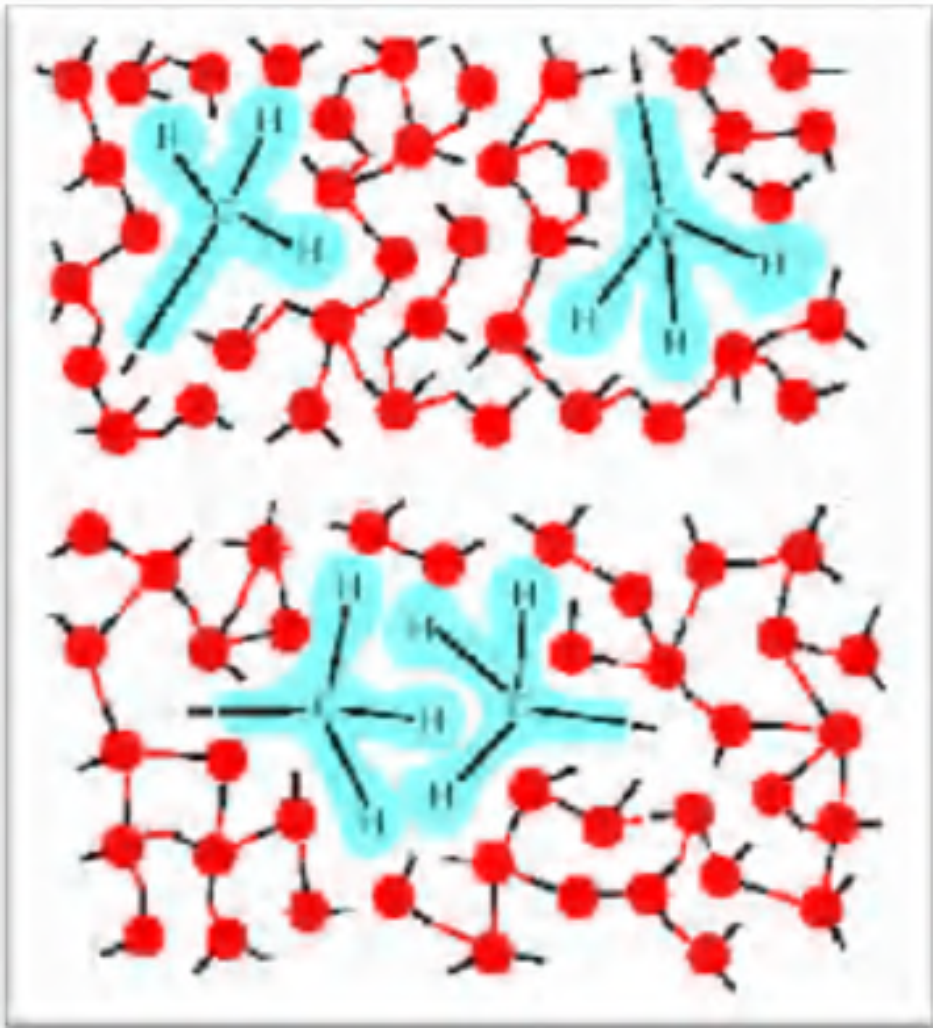


Figure 2.4: Hydrophobic patch

metry of this type of interaction is slightly less restrictive than purely hydrophobic interactions, and they are generally the most common.

Another type of hydrophobic interaction forms during non-specific hydrophobic contacts between aliphatic or aromatic carbon atoms; these are spherical in nature with a radius of 4 Å and cover most of the ligand [61, 62].

For ligand placement in the protein's active site, hydrogen bonds are used first, followed by hydrophobic interactions.

2.2.5 *Pi-Pi Interactions*

Pi-pi (π-π) interactions are a type of non-covalent interaction that form through lateral overlap between a π orbital of an aromatic ligand (such as a benzene ring) and a π orbital of an aromatic residue of the protein [63].

2.3 ADMET PROPERTIES

ADMET (Absorption, Distribution, Metabolism, Excretion, and Toxicity) represents key pharmacokinetic and safety parameters in drug discovery. [64]. These properties are fundamental in determining both the effectiveness and safety of prospective drug candidates [64].

2.3.1 *Absorption*

Absorption describes the process by which a drug enters the body following administration. It involves assessing how efficiently a drug is taken up into the bloodstream from its administration route (e.g., oral intake, intravenous injection). Key factors affecting drug absorption include solubility, bioavailability, and permeability [65].

2.3.2 *Distribution*

After absorption, a drug must reach its intended site of action within the body. Distribution refers to how the drug disperses throughout various tissues and organs. Factors influencing drug distribution include blood flow, protein binding, and tissue permeability [65].

2.3.3 *Metabolism*

The body's metabolic pathways can modify the chemical structure of a drug. Metabolism occurs mainly in the liver and involves enzymatic reactions that may either activate or deactivate a drug. Understanding drug metabolism is critical for assessing its stability and the potential formation of toxic byproducts [65].

2.3.4 *Excretion*

Excretion pertains to the elimination of a drug and its metabolites from the body. Most drugs are excreted via the kidneys into urine, although other routes such as feces and exhalation also contribute. Knowledge of excretion pathways helps determine the drug's elimination half-life [65].

2.3.5 *Toxicity*

Evaluating the potential harmful effects of a drug is essential for establishing its safety profile. Toxicity studies investigate whether a drug or its metabolites can damage the body's cells or organs. This includes assessing acute and chronic toxicity, genotoxicity, and carcinogenicity [65].

3

MATERIALS AND METHODS

In this chapter, i present the materials and methods used:

3.1 MATERIALS

3.1.1 *Computing Resources*

The following computer specifications were used:

Component	**Specification**
Processor	AMD Ryzen 5 7500F 6-Core Processor (3.70 GHz)
Installed RAM	32.0 GB
Graphics card	AMD Radeon RX 9060 XT
Edition	Windows 11 Pro
System type	64-bit operating system, x64-based processor

Table 3.1: Computer specifications used for the study.

3.2 PROGRAMME

3.2.1 *MOE (Molecular Operating Environment) 2014*

The Molecular Operating Environment (MOE), available from the Canadian-based Chemical Computing Group, is an all-inclusive platform for molecular modeling and drug discovery. It has a large market presence and a large user base. [66].As stated in many journal articles, with MOE, users can take advantage of a wide variety of computational tools in QSAR analysis, molecular docking, and pharmacophore and protein structure modeling. With this range of tools, both ligand-based and structure-based drug design methods can be applied using MOE. [66].

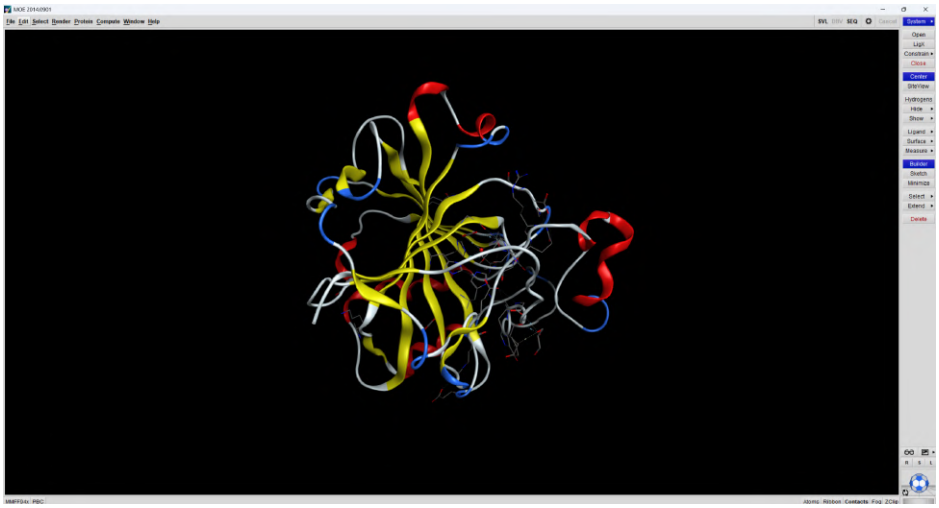


Figure 3.1: MOE interface.

3.2.2 *ChemDraw 2016*

ChemDraw is the most powerful software a chemist or biologist could ask for when working with three-dimensional molecular models. It also provides models visualization through molecular structures, chemical data analysis, and molecular properties calculation [67].Its features allow seamless integration into electronic laboratory notebooks, scientific papers, and chemical databases [67].

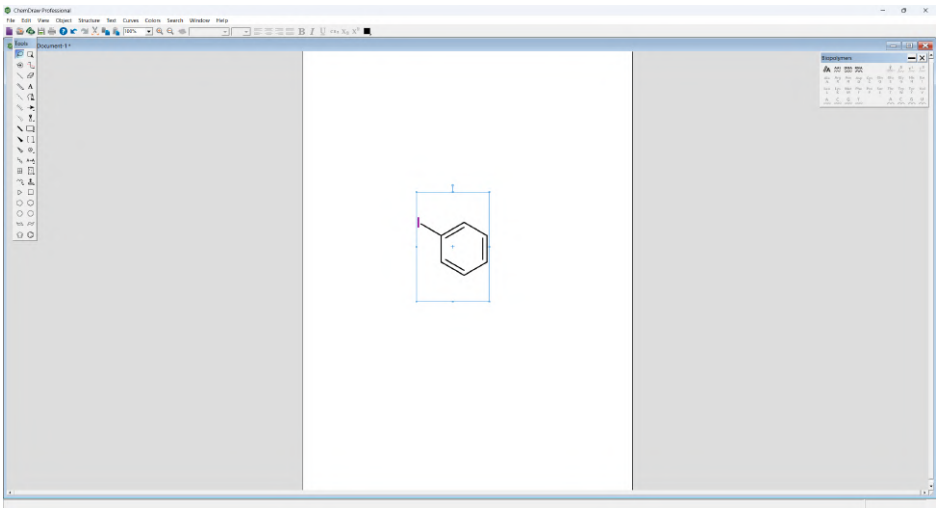


Figure 3.2: ChemDraw 2016 structure drawing.

3.2.3 *BIOVIA Discovery Studio 2025*

In brief, BIOVIA Discovery Studio 2025 provides an integrated environment for molecular modeling and simulation activities supporting all aspects of drug discovery and development, including structure-based modeling of proteins and anti-

bodies, molecular dynamics studies, structure-based virtual screening, and in silico prediction of toxicity and ADME (Absorption, Distribution, Metabolism, and Excretion) properties. [68].

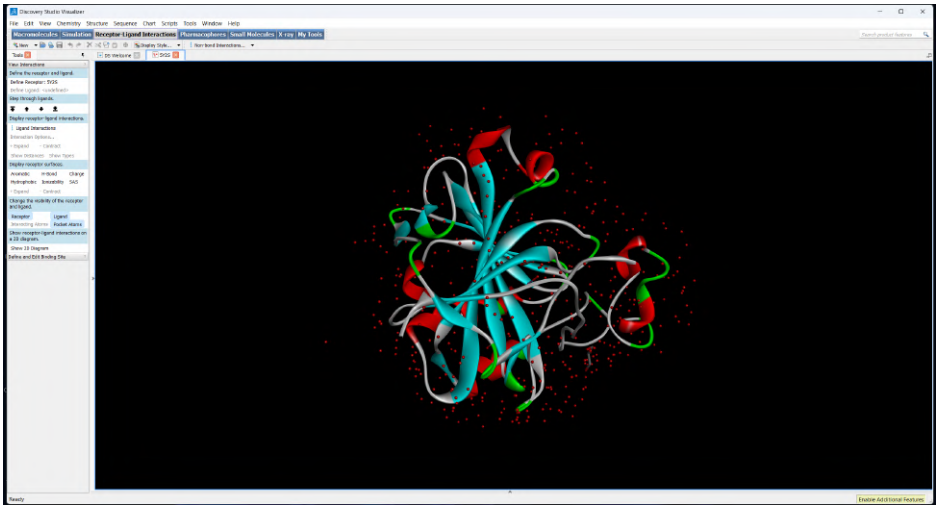


Figure 3.3: BIOVIA Discovery Studio molecular modeling interface.

3.2.4 *DataWarrior*

DataWarrior is a free and open-source software used for chemical data analysis and visualization, facilitating efficient compound filtering in drug discovery. [69].

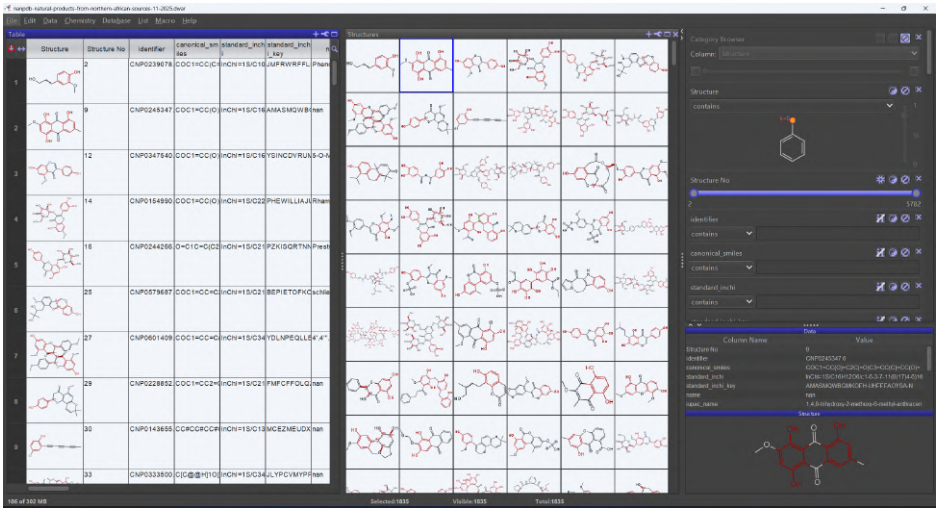


Figure 3.4: DataWarrior visualization of chemical space.

3.3 DATABASES AND WEB TOOLS

3.3.1 *RCSB Protein Data Bank (PDB)*

RCSB PDB is the primary access gateway to the global Protein Data Bank (PDB) archive [70]. The website provides tools for the exploration, visualization, and analysis of experimentally determined three-dimensional structures of proteins and nucleic acids, as well as Computed Structure Models (CSMs) from databases such as AlphaFold DB [70]. It is an essential reference for structure-based drug design. Researchers use it to study enzyme active sites, analyze how existing drugs bind to their biological targets (through available drug-target complex structures), and subsequently design new drug molecules capable of binding more effectively and with greater selectivity [70].

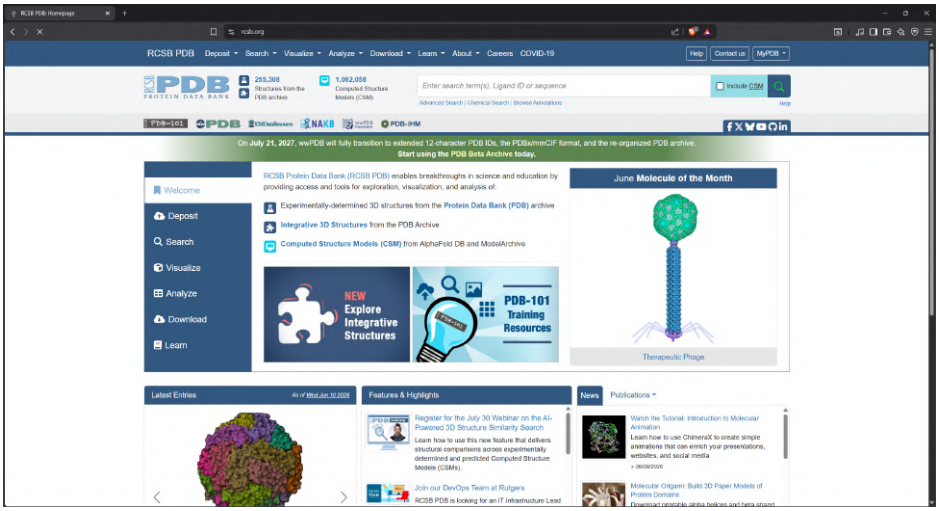


Figure 3.5: RCSB PDB website interface.

3.3.2 *COCONUT Database*

COCONUT (Collection of Open Natural products) is a comprehensive, freely accessible database that collects natural products from various sources [71]. It is designed to facilitate natural product research by providing data, tools, and services for deposition, curation, and reuse [71]. Natural products are a rich source of lead compounds in drug discovery. COCONUT provides researchers with an organized, large-scale library of these compounds to perform virtual screening for new drug discovery, particularly for antibiotics and anticancer agents, via its REST API [71].

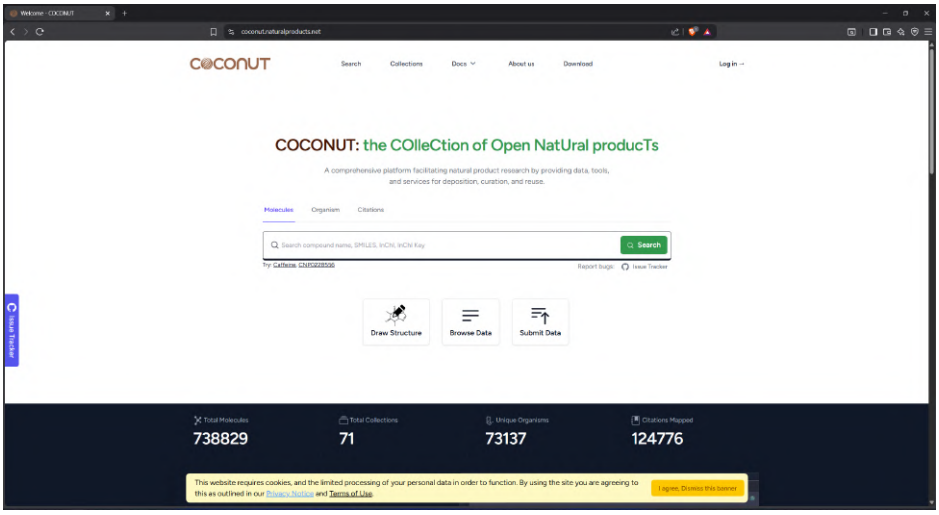


Figure 3.6: COCONUT database overview.

3.3.3 *PubChem*

PubChem is the most extensive repository of chemical information in the world, freely accessible, and maintained by the NIH [72]. It can be searched through name, molecular formula, or structure. Information related to chemical and physical properties, biological activities, safety and toxicity data, patent literature citations, and more is included in its contents [72]. It is a basic starting point in target identification and the processes involved in drug design. It gives full detail on chemical compounds, their biological activities, toxicology records, and molecular properties. It also has applications in virtual screening and in structure-activity relationship analyses [72].

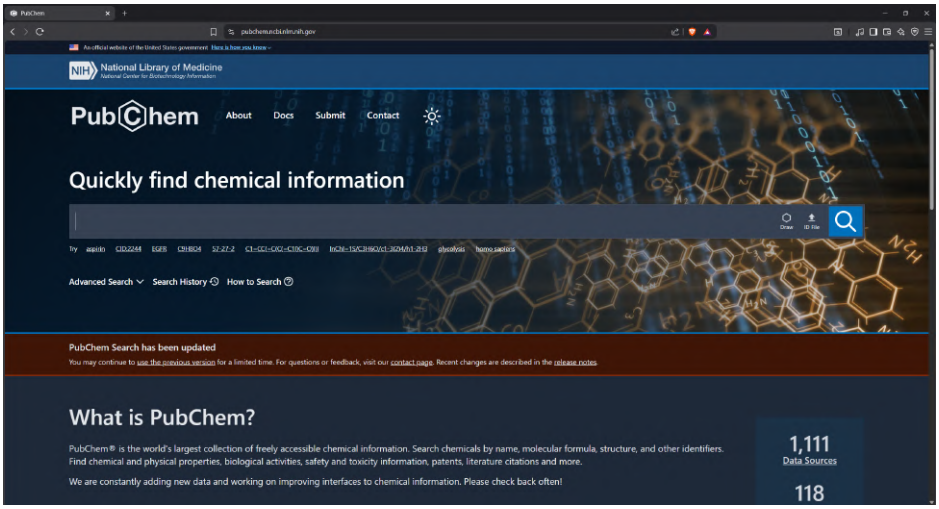


Figure 3.7: PubChem.

3.3.4 *The Open Targets Platform*

The Open Targets Platform is a valuable open-source tool that enables systematic identification and prioritization of putative therapeutic drug targets by integration of publicly available data and scoring of target-disease associations [73].This is used at the target identification stage, a very important and the riskiest early step. Open Targets integrates vast datasets from genomics, proteomics, and drugs; scores the target-disease relationship; and assists researchers in finding the promising and relevant targets to mitigate late-stage failure [73].

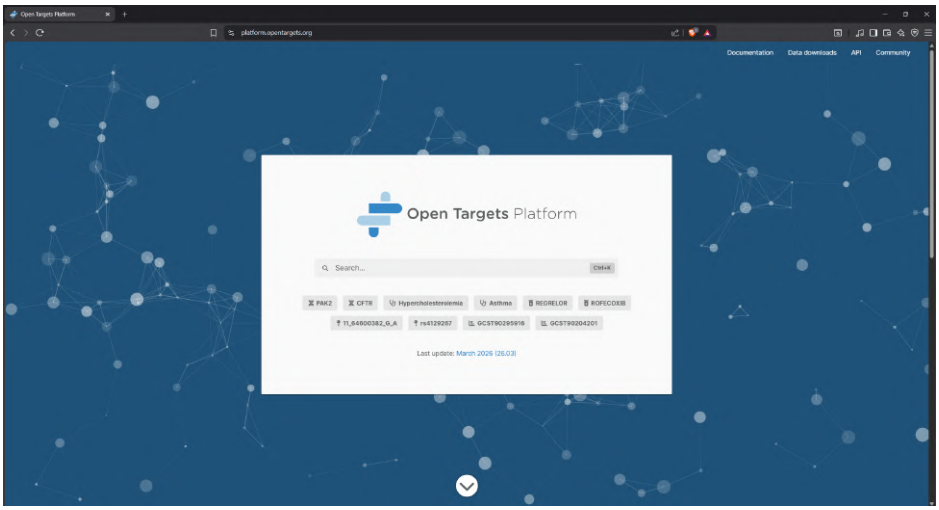


Figure 3.8: Open Targets Platform scoring interface.

3.3.5 *SwissADME*

SwissADME is a free web server that allows the fast and reliable prediction of ADME pharmacokinetic properties, as well as key properties such as solubility and drug likeness, based on the SMILES string of a given compound [74]. It is a very powerful early filtering tool in drug development. Of note is its capability to process thousands of virtually generated compounds, with the program quickly calculating properties related to bioavailability. In this way, early elimination of candidates with poor drug-like properties (for example, poor absorption or synthetic difficulty) allows for a massive gain in time and resources and helps to narrow down efforts on truly promising candidates [74].

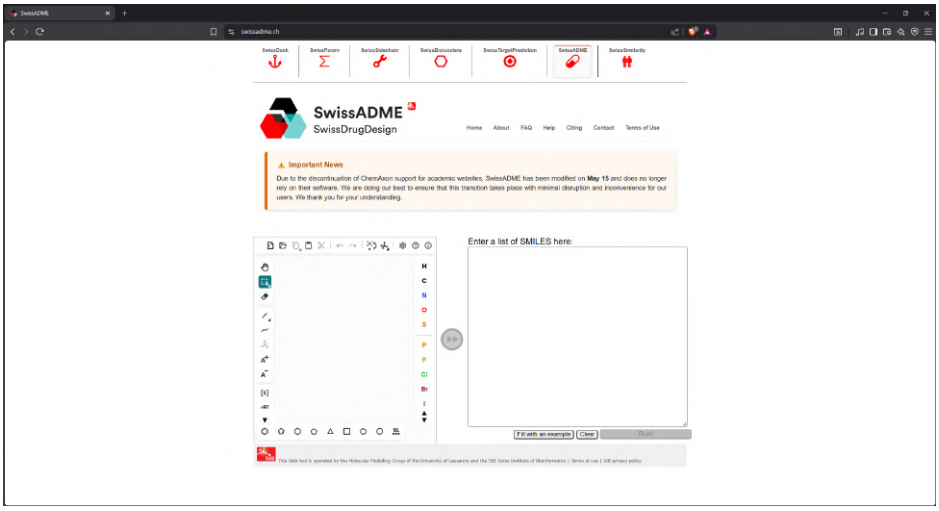


Figure 3.9: SwissADME prediction output.

3.3.6 *ProTox 3.0*

ProTox 3.0 is a virtual lab for predicting small molecules' toxicity with machine learning and molecular similarity [75]. IIt is designed to predict toxicology endpoints such as acute toxicity (LD50), organ toxicity, and toxicological pathways (Tox21) [75].It is used to predict safety because the prediction of toxicity is one of the biggest causes of drug failure in clinical trials. This tool helps researchers to eliminate toxic chemicals (i.e., those that are predicted to be carcinogenic or hepatotoxic) before they are synthesized and tested on animals, aiding in the three Rs (Replacement, Reduction, Refinement) and creating safe drugs faster [75].



Figure 3.10: ProTox 3.0 toxicity prediction interface.

3.4 METHODS

3.4.1 *Choice of proteins*

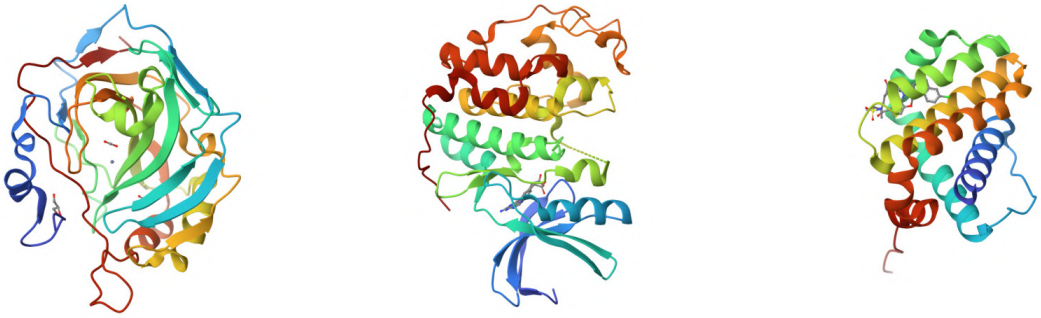
The three-dimensional structures of the proteins investigated in this study were retrieved from the Protein Data Bank (PDB). These structures encompass the following target classes: kinases, hormone receptors, anti-apoptotic proteins, DNA repair enzymes, and cell cycle regulators. All selected proteins are of *Homo sapiens* origin, and possess the highest available resolution and other characteristics which are presented in the table below:

Table 3.2: Properties of the selected target proteins

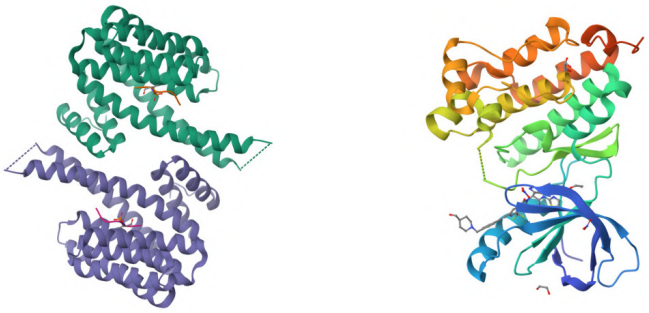
PDB ID	Target Type	Target	Method	Organism	Resolution (Å)	Mutation(s)
8A27	Kinases	EGFR	X-RAY DIFFRACTION	Homo sapiens	1.07	no
7NFW	Hormone Receptors	Estrogen receptor (ER)	X-RAY DIFFRACTION	Homo sapiens	1.19	no
6UDV	Anti-Apoptotic Proteins	Mcl-1	X-RAY DIFFRACTION	Homo sapiens	1.35	no
5Y2S	DNA Repair Enzymes	ATM	X-RAY DIFFRACTION	Homo sapiens	0.90	no
6Q4G	Cell Cycle Regulators	CDK2	X-RAY DIFFRACTION	Homo sapiens	0.98	no

3.4.2 *Preparation of the proteins*

- Proteins were downloaded from the Brookhaven Protein Data Bank (PDB).
- The proteins were simplified by keeping one chain alone.
- All water molecules were stripped off.
- All co-crystallized inhibitors were deleted.
- Active site definition: Every protein molecule may have several sites that could serve as an active site. In order to define this site, residues of the active site were highlighted and shown in MOE's Site_Finder module.



(a) 5Y2S (b) 6Q4G (c) 6UDV



(d) 7NFW (e) 8A27

Figure 3.11: Three-dimensional representations of the five target proteins.

3.4.3 *Choice of ligands*

The ligands were selected by downloading the dataset named "natural-products-from-northern-african-sources" from the COCONUT database, which contains 5,787 compounds. These were filtered to extract only the phenolic compounds, resulting in 1,835 compounds. The ADMET rules will be applied to these remaining compounds, and they will be filtered using the SwissADME and ProTox web servers. The filtration was carried out according to the following rules:

Table 3.3: The characteristics of Lipinski, Ghose, Veber, Egan, and Muegge drug-likeness rules according to SwissAdme[76].

Drug-Likeness Rules				
Lipinski	**Ghose**	**Veber**	**Egan**	**Muegge**
MW $\leq$ 500 Da	160 $\leq$ MW $\leq$ 480 Da			200 $\leq$ MW $\leq$ 600 Da
MlogP $\leq$ 4.15	$-0.4 \leq$ WlogP ≤ 5.6			$-2 \leq$ XlogP ≤ 5
N or O $\leq$ 10	40 $\leq$ MR $\leq$ 130	RB $\leq$ 10	WlogP $\leq$ 5.88	TPSA $\leq$ 150
NH or OH $\leq$ 5	20 $\leq$ atoms $\leq$ 70	TPSA $\leq$ 140	TPSA $\leq$ 131.6	No. of rings ≤ 7
				No. of carbon atoms > 4
				No. of heteroatoms > 1
				No. of RB $\leq$ 15
				H-bond acceptors $\leq$ 10
				H-bond donors ≤ 5

- **MW (Molecular Weight):** Sum of all atoms' masses in a molecule in terms of Daltons or g/mol. MW is one of the crucial physicochemical properties in the drug design field since most orally administered medications should have MW under 500 Da [77].
- **WLOG (Likely logP – Partition Coefficient):** Logarithm of partitioning coefficient (P) of a substance in 1-octanol and water solution system. LogP estimates lipophilicity; higher logP means a higher degree of lipophilicity or fat solubility while lower value of logP implies higher hydrophilicity (water solubility). LogP can be used for estimating absorption, permeability through cell membranes and distribution [78].
- **TPSA (Topological Polar Surface Area):** Surface area provided by polar groups in a molecule that include oxygen, nitrogen atoms, and their bound hydrogen atoms based on 2D molecule geometry. TPSA below 90 Å^2 contributes to membrane permeability and blood-brain barrier penetration; high TPSA (over 140 Å^2) correlates with poor absorption [76–79].
- **P-glycoprotein:** Permeability glycoprotein also known as P-glycoprotein (P-gp), which belongs to the ATP-binding cassette (ABC) superfamily of trans-

porters, can actively transport a wide range of structurally and mechanistically diverse endogenous and xenobiotic chemical agents across the cell membrane at the energy expense of ATP hydrolysis. P-gp, a 170-kDa plasma membrane protein encoded by the multidrug resistance gene (MDR1/ABCB1), is expressed at high levels in various tissues such as blood–brain-barriers (BBB), gastrointestinal tract (GIT), liver, kidney, and placenta. In addition, P-gp plays significant roles in cell and tissue detoxification and elimination of harmful substances per se [80].

- **A Bioavailability Score** is formulated as the probability that a compound will have >10% bioavailability in rat or measurable Caco-2 permeability. ABS is 0.11 for anions for which PSA is >150 $Å^2$, 0.56 if PSA is between 75 and 150 $Å^2$, and 0.85 if PSA is <75 $Å^2$. For the remaining compounds ABS is 0.55 if it passes the rule-of-five and 0.17 if it fails. ABS also identifies poorly- and well-absorbed compounds tested in humans [81].

- **GI absorption (Gastrointestinal Absorption):** The ability of an orally administered drug to be taken up by the gastrointestinal tract into the bloodstream. High GI absorption means the compound is well absorbed upon oral administration [82].

- **Cardiotoxicity:** Drug-induced cardiotoxicity can be caused by inhibiting the human ether-à-go-go-related gene (hERG) channel—a potassium ion channel that plays a crucial role in regulating the cardiac action potential. The hERG channel is essential for the repolarization of the cardiac cells, and its inhibition by certain small molecules can lead to the prolongation of the QT interval on an electrocardiogram (ECG) [83].

- **Blood–brain barrier (BBB)** Some toxins, by size, lipophilicity, or specific transport mechanisms, can bypass or compromise the BBB and enter the brain. These toxicants may include environmental pollutants, drugs, and certain chemicals. Though clinical experiments are the most accurate methods to predict the BBB permeability of a molecule, they are, on the other hand, time-consuming and labour-intensive [83].

- **Carcinogenicity** Substances classified as carcinogens can either induce or increase the incidence of cancers. The Carcinogenic Potency Database (CPDB) and the CEBS database provide information for predicting carcinogenicity [84].

- **Immunotoxicity:** The detrimental effects of xenobiotics on the immune system are referred to as immunotoxicity. The National Cancer Institute (NCI) in

the United States provided the immune cell cytotoxicity data that were used in the immunotoxicity model [84].

- **Mutagenicity** The capacity to result in genetic changes. Mutagens are chemicals that cause abnormal genetic mutations, like changes to a cell's DNA. These changes can harm cells and cause certain illnesses, such as cancer [84].
- **Cytotoxicity** Predicting a substance's potential to cause either desired or undesirable cell damage—the latter being the case with tumor cells—is essential for screening compounds [84].
- **Hepatotoxicity** or liver damage, is caused by hepatotoxins, which may source from chemicals, dietary supplements, pharmaceutical drugs, and medicinal plants. Notably, numerous medicinal plants are used to alleviate illness [85].
- **CYP3A4** is the most widely expressed and clinically significant cytochrome responsible for the biotransformation of 30.2% of xenobiotics that undergo hepatic clearance. CYP3A4 is capable of oxidizing a wide array of bulky substrates with many different structural features. The versatile functionality of CYP3A4 is due to its large substrate-binding cavity. The active site cavity of CYP3A4 is open and flexible, capable of accommodating volumes ranging from 1173 to 2862 cubic angstroms ($Å^3$). The CYP3A4 ligand complexes formed upon substrate binding are incredibly flexible, explaining the observed ligand promiscuity and diverse range of substrates capable of high-affinity binding interactions [86].
- **CYP2C19** metabolizes 6.8% of drugs that undergo hepatic biotransformation. Despite being considered a minor hepatic cytochrome based on its relative contribution, CYP2C19 is responsible for metabolizing several clinically important drugs, including omeprazole and phenytoin. As such, many pharmacologists view CYP2C19 as a major cytochrome. Inhibitors of CYP2C19 share frequently observed structural features; structures contain numerous aromatic moieties, heterocycles, carbonyl groups, and aromatic nitrogen atoms [86].
- **CYP2D6** comprises 2% of hepatocyte CYP content and is responsible for the biotransformation of 20% of drugs that undergo hepatic metabolism. The active site cavity of CYP2D6 has a more restricted volume of 510 cubic angstroms ($Å^3$). The active site of CYP2D6 is acidic in nature and contains several phenylalanine residues capable of pi-pi stacking with substrates [86].
- **CYP2C9** is responsible for the biotransformation of 12.8% of drugs that undergo hepatic metabolism. In particular, CYP2C9 plays a key role in

the metabolism of commonly used medications, such as non-steroidal anti-inflammatory drugs (NSAIDs) and warfarin. The active site of CYP2C9 is rather large and has a cavity volume ranging from 978 to 1271 cubic angstroms ($Å^3$). Inhibitors of CYP2C9 are typically aromatic compounds with heterocycles, aromatic nitrogens, primary amines, and halogens [86].

- **CYP1A2** is responsible for the biotransformation of 8.9% of drugs that undergo hepatic metabolism, including antipsychotics and antibiotics. Additional CYP1A2 substrates include caffeine, paracetamol, and aflatoxin B1. Interestingly, CYP1A2 activity is commonly induced by dietary constituents, including various vegetables and polyaromatic hydrocarbons, whereas ingredients such as cumin and turmeric can inhibit enzyme function. The active site of CYP1A2 is small and planar with a volume of 375 to 390 cubic angstroms ($Å^3$), and thus CYP1A2 substrates tend to be small, planar molecules [86].

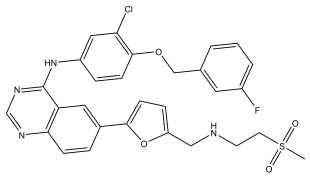
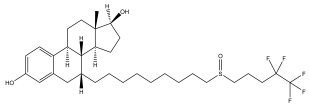
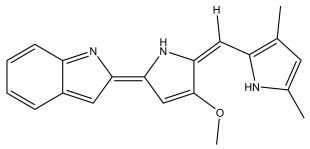
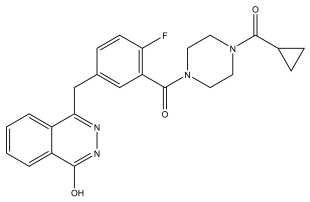
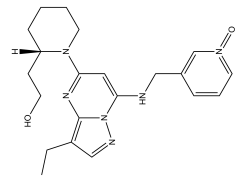
Furthermore, 5 drugs were selected via the Open Targets platform for each protein to serve as a comparison with the compounds obtained after filtration, as shown in the table 3.4.

3.4.4 *Preparation of the ligands*

After completing the ADMET filtration process of the compounds, the following is performed:

- Construction of ligands was done using ChemDraw.
- Optimization process was performed through MOE (Molecular Operating Environment).

Table 3.4: drugs and their 2D structures

Target	PDB ID	Drug	PubChem CID	2D Structure
EGFR	8A27	LAPATINIB DITOSYLATE	11557040	
Estrogen receptor (ER)	7NFW	FULVESTRANT	104741	
Mcl-1	6UDV	Obatoclax mesylate	16681698	
ATM	5Y2S	OLAPARIB	23725625	
CDK2	6Q4G	DINACICLIB	46926350	

3.4.5 *Docking Protocol:*

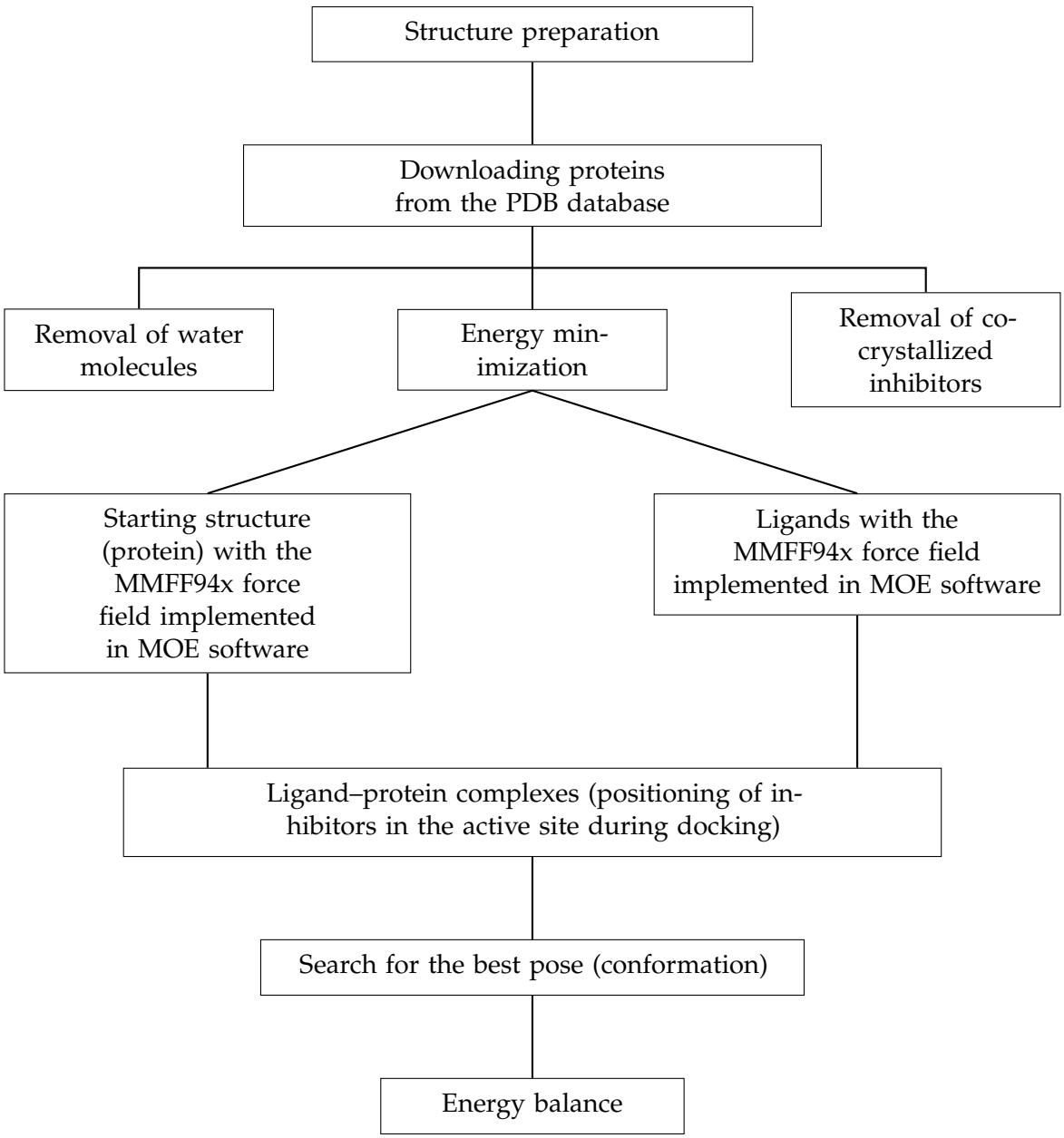


Figure 3.12: Presentation diagram of the computational protocol for molecular docking.

4

RESULTS AND DISCUSSION

4.1 ADMET FILTERING

A pharmacokinetic analysis of the inhibitors predicted via molecular docking is performed to confirm the selection of the most promising compound for inhibiting cancerous proteins. The *Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADME/T)* criteria help evaluate the drug's efficacy and safety by verifying that it properly reaches its therapeutic targets and is free from unacceptable risks. Absorption is a fundamental factor, as the drug's failure to be absorbed toward its target compromises its effectiveness. Moreover, water solubility is crucial for enabling absorption and bioavailability, and it can be adjusted through galenic (pharmaceutical) formulations. Distribution assesses the drug's dispersion throughout tissues and organs, with the apparent volume of distribution serving as a key indicator in this regard. Metabolism, which primarily occurs in the liver, is a cornerstone of drug activity. Furthermore, excretion via the excretory system is necessary to prevent adverse effects that may arise after use. Toxicity is a critical aspect, assessed through several tests, such as hepatotoxicity and acute oral toxicity in rats, in order to identify the potentially harmful effects of the drug on the body. Overall, these analyses aim to ensure that the drug efficiently reaches its target, is properly metabolized, is excreted without harmful effects, and avoids any unacceptable toxicity risks.

After applying ADMET filtering to the database of phenolic compounds (1835 compounds) and extracting the best compounds that met all the acceptance criteria on the SwissADME and ProTox webservers, And the obtained results are presented in Table 4.1.

4.1.1 *Interpretation*

An *in silico* analysis of ADME/Tox characteristics for the chosen phenolic derivatives (Compounds 1–15) demonstrates a positive trend for pharmacokinetics and safety for all compounds tested. Although the basic physicochemical properties,

Table 4.1: ADME/Tox properties of selected phenolic compounds (Compounds 1-15).

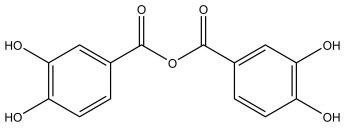
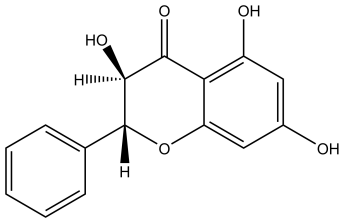
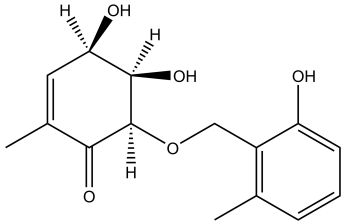
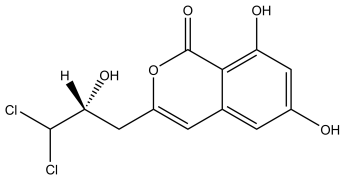
Property	Cmpd 1	Cmpd 2	Cmpd 3	Cmpd 4	Cmpd 5	Cmpd 6	Cmpd 7	Cmpd 8	Cmpd 9	Cmpd 10	Cmpd 11	Cmpd 12	Cmpd 13	Cmpd 14	Cmpd 15
MW	290.23	272.25	278.30	305.11	280.27	300.26	278.26	298.29	288.25	260.08	344.32	260.29	210.23	212.20	246.26
TPSA	124.29	86.99	86.99	90.90	111.13	96.97	107.97	86.99	107.22	72.55	124.29	69.64	77.76	86.99	85.35
WLOG	1.51	1.45	0.68	1.91	0.88	2.03	1.09	2.54	1.16	1.19	2.04	-0.34	2.03	0.92	0.83
GI absorption	High	High	High	High	High	High	High	High	High	High	High	High	High	High	High
BBB permeant	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Pgp substrate	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
CYP1A2 Inh.	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
CYP2C19 Inh.	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
CYP2C9 Inh.	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
CYP2D6 Inh.	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
CYP3A4 Inh.	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
Lipinski Viol.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ghose Viol.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Veber Viol.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Egan Viol.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Muegge Viol.	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Bioavailability	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55	0.55
Hepatotoxicity	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Cardiotoxicity	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Carcinogenicity	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Immunotoxicity	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Mutagenicity	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Cytotoxicity	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive

Abbreviations: Cmpd = Compound; Inh. = Inhibitor; Viol. = Violations.

such as molecular weight, topological polar surface area (TPSA), and lipophilicity (WLOGP), vary, all fifteen derivatives exhibit high intestinal absorption properties and are unable to permeate through the blood–brain barrier (BBB). Furthermore, they display good drug-likeness, as reflected by the absence of any violations of pharmacological rules, namely Lipinski, Ghose, Veber, Egan, and Muegge. Consequently, the standardized bioavailability score for each compound equals 0.55. From a toxicity perspective, all derivatives do not interact with P-glycoprotein (Pgp) and are inactive as inhibitors of essential Cytochrome P450 (CYP) enzymes. They are also classified as "Inactive" with respect to all stringent toxicity criteria, including hepatotoxicity, cardiotoxicity, carcinogenicity, mutagenicity, and cytotoxicity. Collectively, the obtained data imply that these phenolic compounds represent promising candidates for further research, owing to their excellent bioavailability and favorable safety profile.

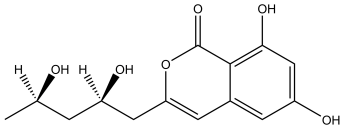
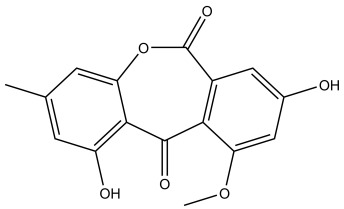
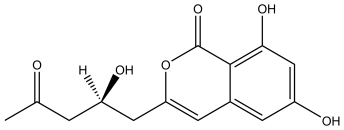
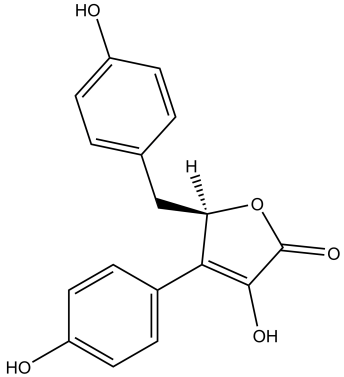
The table below shows the two-dimensional structures of the 15 selected compounds for the docking process:

Table 4.2: List of selected phenolic compounds.

Compound	2D Structure	Identifier (Coconut)	Canonical SMILES
1		CNP0123642.0	O=C(c1ccc(cc1)O)OOC(=O)c1ccc(cc1)O
2		CNP0226937.1	Oc1cc2O[C@H](c3cccc3)[C@H](C=O)c2c(1)O
3		CNP0075219.1	O[C@H]1C=C(C)C(O)=[C@H](C=O)[C@H](C=O)Oc1c(CC)ccc1O
4		CNP0104458.1	Oc1cc(O)c2c(1)cc(Oc2=O)C[C@H](CCl)CO

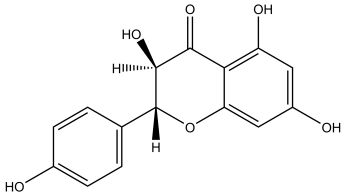
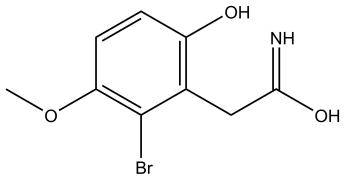
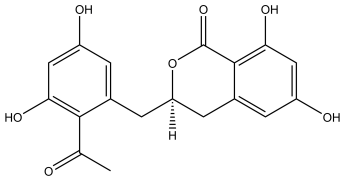
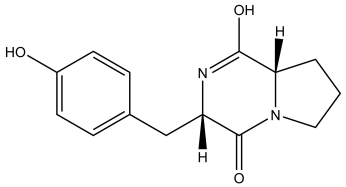
Continued on next page

Table 4.2 – continued from previous page

Compound	2D Structure	Identifier (Coconut)	Canonical SMILES
5		CNP0272195.4	Cl[C@H](C[C@H](Cc1oc(=O)c2c(1)cc(Oc2O)O)O
6		CNP0154427.0	COc1cc(O)c2c1c(=O)c1c(O)cc(Oc1c2=O)C
7		CNP0259856.1	CC(=O)C[C@H](Cc1oc(=O)c2c(1)cc(Oc2O)O
8		CNP0155543.1	Oc1ccc(cc1)[C@H]1OC(=O)C=C1c1ccc(cc1)O

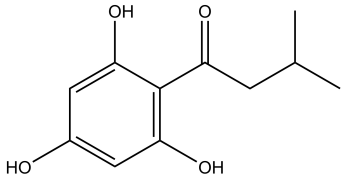
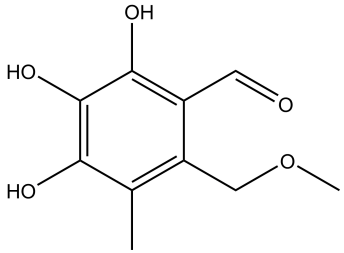
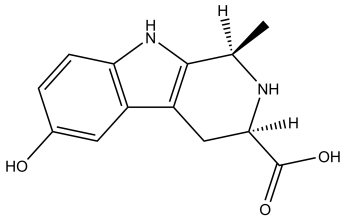
Continued on next page

Table 4.2 – continued from previous page

Compound	2D Structure	Identifier (Coconut)	Canonical SMILES
9		CNP0230258.2	Oc1ccc(cc1)[C@H]1Oc2cc(O)c2c(O)[C@H]1O
10		CNP0132890.0	COc1ccc(cc1Br)CC(=O)N
11		CNP0232713.1	CC(=O)C1=C(O)C=C(O)C=C1C[C@@H]1CC2=CC(O)=CC(O)=C2C(=O)O1
12		CNP0320469.1	O=C1N[C@@H](CC2=CC=C(O)C=C2)C(=O)N2CCC[C@@H]12

Continued on next page

Table 4.2 – continued from previous page

Compound	2D Structure	Identifier (Coconut)	Canonical SMILES
13		CNP0210081.0	CC(C)CC(=O)C1=C(O)C=C(O)C=C1O
14		CNP0356682.0	COCC1=C(C)C(O)=C(O)C(O)=C1C=O
15		CNP0284879.2	C[C@H]1N[C@@H](C(=O)O)CC2=C1NC1=CC=C(O)C=C21

4.2 MOLECULAR DOCKING

The molecular docking results yielded twenty possible poses per ligand, where the first pose typically corresponds to the lowest interaction energy between the ligand and the target protein, suggesting that this conformation may represent the most favorable binding mode within the protein's active site. The molecular docking calculations also assessed the interactions between the active sites of the proteins and each corresponding reference drug, as well as the fifteen selected compounds. The ligand binds optimally to a given receptor when the root-mean-square deviation (RMSD) is close to 2 Å and the docking score is ≤ 7 kcal/mol.These two values were used as criteria to validate the molecular docking results. [87].

Following the docking process, the following results were obtained for each (protein–ligand) interaction:

4.2.1 *5Y2S protein*

Table 4.3: Molecular Docking Results: of receptor 5Y2S

Compound N°	**S-score (kcal/mol)**	**RMSD**
OLAPARIB	-5.62	1.46
1	-5.02	2.39
2	-4.42	2.07
3	-5.27	1.46
4	-4.76	2.36
5	-4.97	0.91
6	-4.70	1.33
7	-4.72	3.46
8	-4.69	3.24
9	-4.59	1.63
10	-4.42	0.76
11	-4.84	1.44
12	-4.58	1.47
13	-4.25	1.95
14	-4.44	1.04
15	-4.28	0.85

- **Docking studies on the receptor 5Y2S reveal that the reference drug OLAPARIB was bound to the receptor with the best binding affinity,**

Table 4.4: Molecular Docking Results: of receptor 5Y2S (The score of the drug and the highest three scores of the compounds)

Compound N°	S-score (kcal/mol)	RMSD
OLAPARIB	-5.62	1.46
3	-5.27	1.46
1	-5.02	2.39
5	-4.97	0.91

judging from its S-score value of 5.62 kcal/mol and RMSD value of 1.46 Å. The best binding affinity among the analyzed compounds was obtained for compound 3 with an S-score value of 5.27 kcal/mol, while compound 1 and 5 had S-score values of 5.02 kcal/mol and 4.97 kcal/mol, respectively.

4.2.2 *6Q4G protein*

Table 4.5: Molecular Docking Results: of receptor 6Q4G

Compound N°	S-score (kcal/mol)	RMSD
DINACICLIB	-6.54	2.16
1	-6.10	1.87
2	-6.14	1.97
3	-5.90	0.97
4	-5.40	2.64
5	-5.94	1.46
6	-6.84	2.63
7	-5.73	1.92
8	-6.53	1.84
9	-5.90	2.13
10	-5.20	1.68
11	-6.74	1.72
12	-5.73	1.20
13	-5.35	0.83
14	-5.09	1.64
15	-5.36	1.49

- **Receptor 6Q4G: The results of the molecular docking analysis have indicated that compound 6 possesses the greatest binding affinity to receptor**

Table 4.6: Molecular Docking Results: of receptor 6Q4G (The score of the drug and the highest three scores of the compounds)

Compound N°	S-score (kcal/mol)	RMSD
6	-6.84	2.63
11	-6.74	1.72
DINACICLIB	-6.54	2.16
8	-6.53	1.84

6q4g. Its S-score was equal to 6.84 kcal/mol, which is better compared to the affinity of compound 11 (6.74 kcal/mol), and compound 8 (6.53 kcal/-mol). It should be noted that all these scores were rather close to those of reference drug DINACICLIB: 6.54 kcal/mol. As for the docking results, compound 6 and compound 11 have shown greater affinity than the reference drug, while the former demonstrated a lower RMSD score.

4.2.3 *6UDV protein*

Table 4.7: Molecular Docking Results: of receptor 6UDV

Compound N°	S-score (kcal/mol)	RMSD
Obatoclax mesylate	-4.97	5.22
1	-6.63	2.17
2	-6.31	0.94
3	-6.52	1.55
4	-5.73	1.70
5	-6.13	1.13
6	-6.40	1.39
7	-6.04	1.19
8	-6.99	1.92
9	-6.28	1.65
10	-5.37	1.36
11	-6.98	2.15
12	-6.00	2.48
13	-5.46	1.66
14	-5.17	0.84
15	-5.19	1.23

Table 4.8: Molecular Docking Results: of receptor 6UDV (The score of the drug and the highest three scores of the compounds)

Compound N°	S-score (kcal/mol)	RMSD
8	-6.99	1.92
11	-6.98	2.15
1	-6.63	2.17
Obatoclax mesylate	-4.97	5.22

- **Receptor 6UDV: According to the docking analysis, the binding affinities of compound 8 and 11 were the greatest, with their scores being equal to 6.99 kcal/mol, and 6.98 kcal/mol, respectively, while compound 1 had the greatest score 6.63 kcal/mol. At the same time, their results significantly outperform the values observed for reference drug Obatoclax mesylate. The latter has the binding score equal to 4.97 kcal/mol, and RMSD of 5.22 Å.**

4.2.4 *7NFW protein*

Table 4.9: Molecular Docking Results: of receptor 7NFW

Compound N°	S-score (kcal/mol)	RMSD
FULVESTRANT	-6.31	2.05
1	-5.16	1.23
2	-4.55	4.06
3	-5.15	2.37
4	-5.04	2.23
5	-5.37	1.45
6	-5.09	4.19
7	-4.85	1.81
8	-4.84	2.27
9	-4.43	1.42
10	-4.46	2.46
11	-5.46	1.63
12	-4.18	3.14
13	-4.76	1.15
14	-4.48	0.94
15	-4.43	4.37

Table 4.10: Molecular Docking Results of 7NFW receptor (The score of the drug and the highest three scores of the compounds)

Compound N°	S-score (kcal/mol)	RMSD
FULVESTRANT	-6.31	2.05
11	-5.46	1.63
5	-5.37	1.45
1	-5.16	1.23

- **Receptor 7NFW: The docking analysis showed that the reference drug FULVESTRANT possesses the greatest binding affinity to receptor 7NFW, having the score of 6.31 kcal/mol, which is significantly higher than that of other compounds (e.g., 5.46 kcal/mol, 5.37 kcal/mol, and 5.16 kcal/mol, respectively, for compounds 11, 5, and 1).**

4.2.5 *8A27 protein*

Table 4.11: Molecular Docking Results: of receptor 8A27

Compound	S-score (kcal/mol)	RMSD
Lapatinib Ditosylate	-9.10	1.68
1	-6.35	1.66
2	-5.99	0.84
3	-7.44	1.15
4	-6.01	1.15
5	-6.79	1.04
6	-6.21	1.24
7	-6.27	1.70
9	-6.59	0.94
10	-5.66	1.02
11	-6.75	1.25
12	-6.42	1.98
13	-5.48	1.37
14	-5.55	1.23
15	-5.31	1.64

- **Receptor 8A27: According to the docking analysis, the drug Lapatinib Ditosylate had the greatest binding affinity score, which was 9.10 kcal/mol,**

Table 4.12: Molecular Docking Results of 8A27 receptor (The score of the drug and the highest three scores of the compounds)

Compound Nº	**S-score (kcal/mol)**	**RMSD**
Lapatinib Ditosylate	-9.10	1.68
3	-7.44	1.15
5	-6.79	1.04
11	-6.75	1.25

and it also demonstrated a relatively low RMSD, which was 1.68 Å. As for other drugs under consideration, compound 3 provided the greatest S-score equal to 7.44 kcal/mol, and compound 5 and 11 scored 6.79 and 6.75 kcal/mol, respectively.

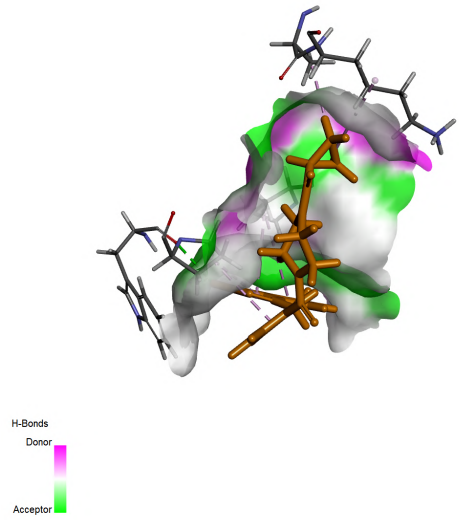
4.3 STUDY OF PROTEIN–LIGAND INTERACTIONS

In order to suggest new inhibitors for the selected proteins, a study was conducted between the top three ligands and the reference compound. Binding interactions between these molecules were determined by performing two-dimensional and three-dimensional interactions through Biovia Discovery Studio. This provided information about the residue involved in binding and the chains that contain them. These active phytochemicals bind with active site residues through hydrogen bonds and hydrophobic interactions.

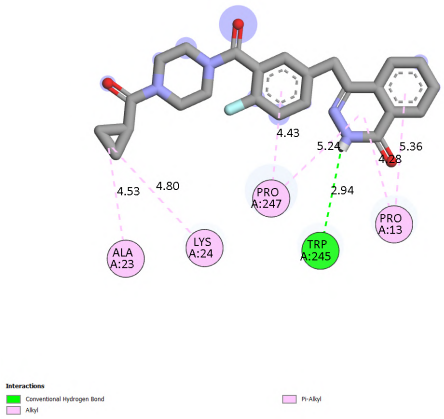
4.3.1 *5Y2S–ligand interactions*

Complexes	**Ligand**	**Receptor**	**Amino acid types**	**Interactions**	**Distance (Å)**	**Energy (kcal/mol)**
5Y2S-L1	O 30	OE2	GLU 239 (A)	H-donor	2.9	-4.9
	O 1	NE2	GLN 103 (A)	H-acceptor	3.2	-1.0
5Y2S-L5	O 31	OE1	GLU 14 (A)	H-donor	2.92	-1.5
	O 35	OE2	GLU 239 (A)	H-donor	2.76	-2.5
	O 35	NZ	LYS 9 (A)	H-acceptor	2.89	-6.0
	6-ring	CD	PRO 13 (A)	pi-H	4.26	-0.6
5Y2S-OLAPARIB	6-ring	CD	PRO 13 (A)	pi-H	3.9	-0.6
5Y2S-L3	-	-	-	-	-	-

Table 4.13: MOE interaction report for the top-scoring ligands with the 5Y2S receptor.

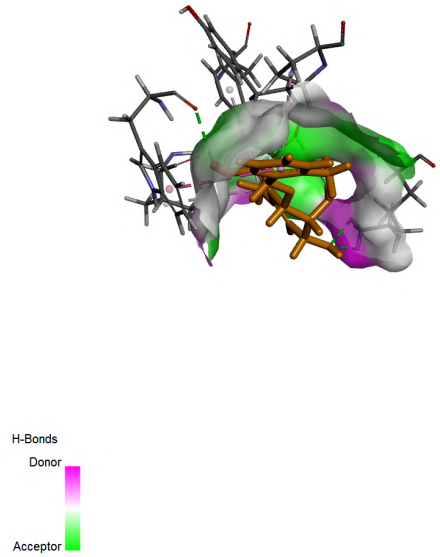


(a) 3D interaction of OLAPARIB with receptor 5Y2S

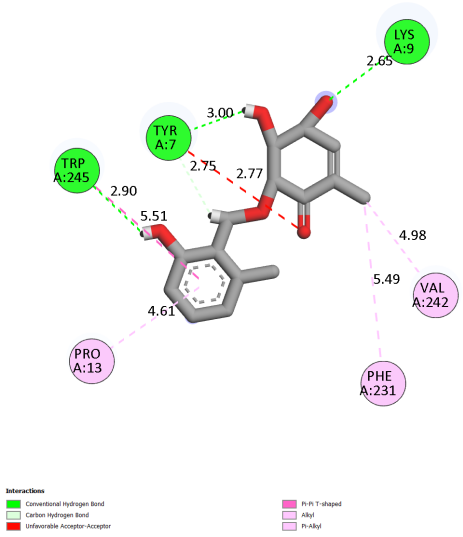


(b) 2D interaction of OLAPARIB with receptor 5Y2S

Figure 4.1: Molecular docking interactions of OLAPARIB with receptor 5Y2S.

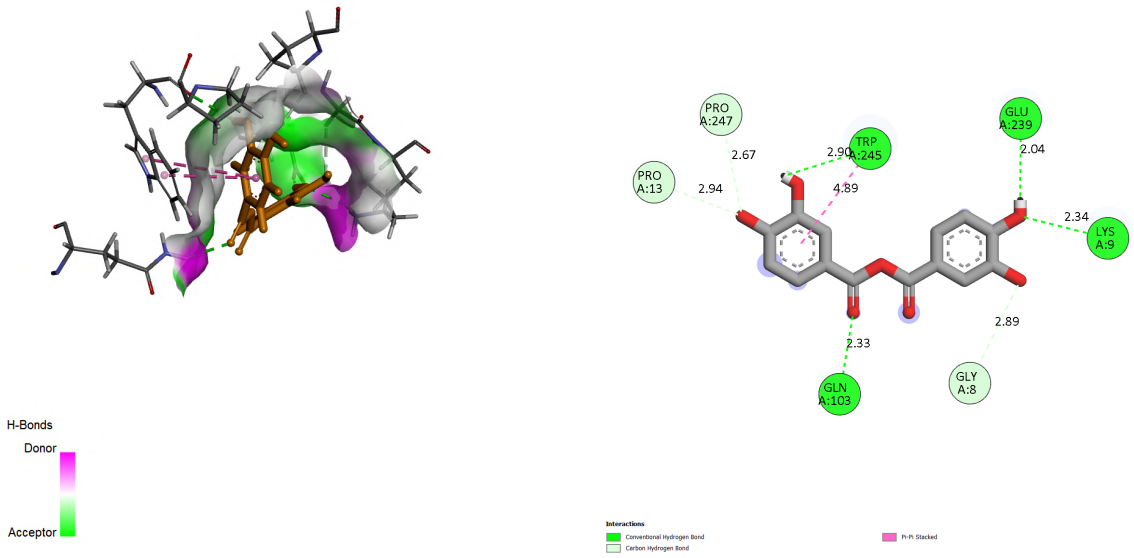


(a) 3D interaction of compound 3 with receptor 5Y2S



(b) 2D interaction of compound 3 with receptor 5Y2S

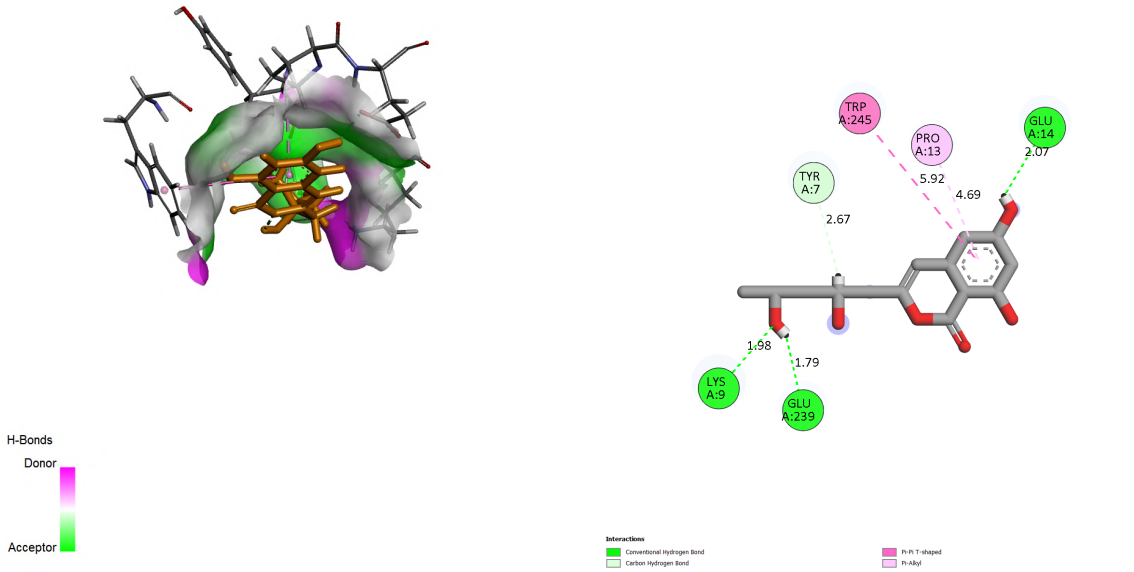
Figure 4.2: Molecular docking interactions of compound 3 with receptor 5Y2S.



(a) 3D interaction of compound 1 with receptor 5Y2S

(b) 2D interaction of compound 1 with receptor 5Y2S

Figure 4.3: Molecular docking interactions of compound 1 with receptor 5Y2S.



(a) 3D interaction of compound 5 with receptor 5Y2S

(b) 2D interaction of compound 5 with receptor 5Y2S

Figure 4.4: Molecular docking interactions of compound 5 with receptor 5Y2S.

- 5Y2S-L5:The interaction of compound 5 shows a stable binding interaction with strong conventional hydrogen bonds formed with LYS 9 (1.98 Å), GLU 239 (1.79 Å), and GLU 14 (2.07 Å), a carbon-hydrogen bond with TYR 7 (2.67 Å), a Pi-Pi T-shaped interaction with TRP 245 (5.92 Å), and a hydrophobic Pi-Alkyl interaction with PRO 13 (4.69 Å).

- 5Y2S-L1: The interaction of compound 1 is represented by numerous conventional hydrogen bonds with TRP 245 (2.90 Å), GLN 103 (2.33 Å), GLU 239 (2.04 Å), and LYS 9 (2.34 Å), carbon-hydrogen bonds with PRO 247 (2.67 Å), PRO 13 (2.94 Å), and GLY 8 (2.89 Å), along with a Pi-Pi Stacked interaction with TRP 245 (4.89 Å).

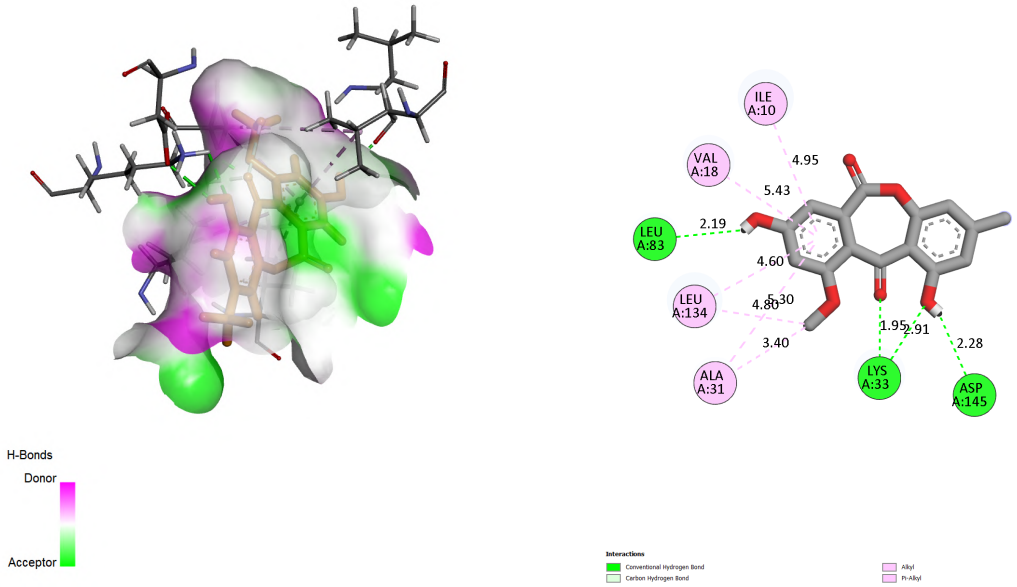
- 5Y2S-L3:The interaction of compound 3 demonstrates binding interactions assisted by conventional hydrogen bonds with TRP 245 (2.90 Å), TYR 7 (3.00 Å), and LYS 9 (2.65 Å), supplemented with carbon-hydrogen bonds with TYR 7 (2.75 Å) and TRP 245 (5.51 Å). Despite a slight local strain due to an unfavorable Acceptor-Acceptor interaction with TYR 7 (2.77 Å), the whole binding structure is maintained by hydrophobic Pi-Alkyl interactions with PRO 13 (4.61 Å), VAL 242 (4.98 Å), and PHE 231 (5.49 Å).

- 5Y2S-OLAPARIB: The interaction of the reference drug presents specific protein-ligand interactions based on non-polar interactions, forming a single crucial conventional hydrogen bond with TRP 245 (2.94 Å). The complex is mainly stabilized by a hydrophobic binding network consisting of Alkyl and Pi-Alkyl interactions, with multiple contacts occurring at PRO 247 (4.43 Å and 5.24 Å), PRO 13 (4.28 Å and 5.36 Å), LYS 24 (4.80 Å), and ALA 23 (4.53 Å).

4.3.2 *6Q4G–ligand interactions*

Complexes	Ligand	Receptor	Amino acid types	Interactions	Distance (Å)	Energy (kcal/mol)
6Q4G-DINACICLIB	O 43	NZ	LYS 129 (A)	H-acceptor	2.96	-3.9
	O 43	N	THR 165 (A)	H-acceptor	3.08	-1.9
	6-ring	CA	THR 14 (A)	pi-H	4.38	-0.6
	5-ring	CB	GLN 131 (A)	pi-H	4.16	-0.7
	6-ring	CB	GLN 131 (A)	pi-H	3.64	-1.3
	6-ring	CA	VAL 164 (A)	pi-H	4.90	-0.8
6Q4G-L6	O 10	O	LEU 83 (A)	H-donor	2.90	-1.4
	O 17	NZ	LYS 33 (A)	H-acceptor	2.82	-5.3
6Q4G-L11	O 1	OD2	ASP 86 (A)	H-donor	2.98	-1.8
	O 14	NZ	LYS 33 (A)	H-acceptor	3.11	-5.9
6Q4G-L8	-	-	-	-	-	-

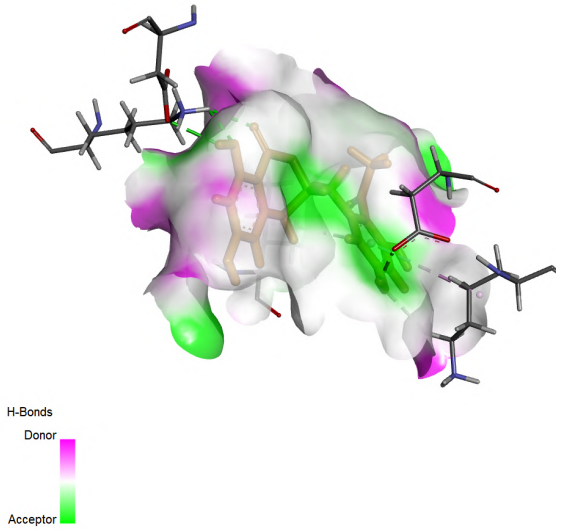
Table 4.14: MOE interaction report for the top-scoring ligands with the 6Q4G receptor.



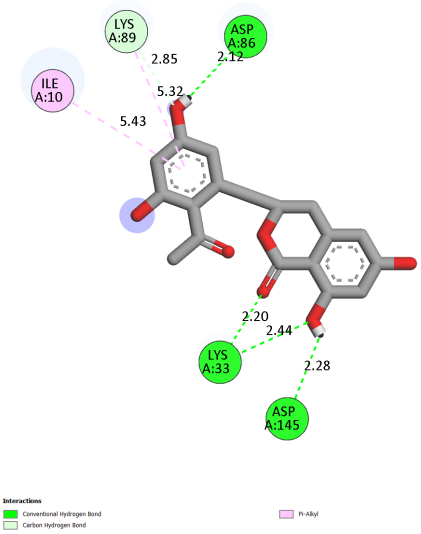
(a) 3D interaction of compound 6 with receptor 6Q4G

(b) 2D interaction of compound 6 with receptor 6Q4G

Figure 4.5: Molecular docking interactions of compound 6 with receptor 6Q4G.

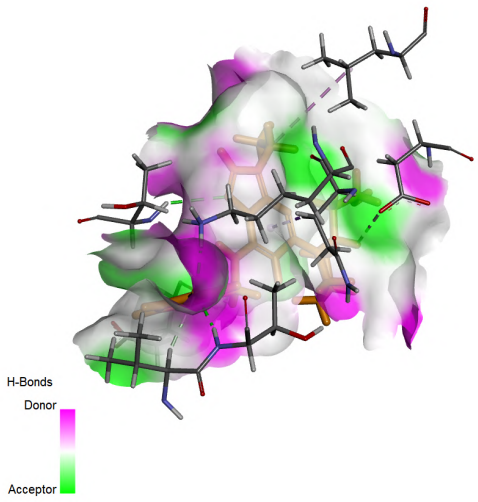


(a) 3D interaction of compound 11 with receptor 6Q4G

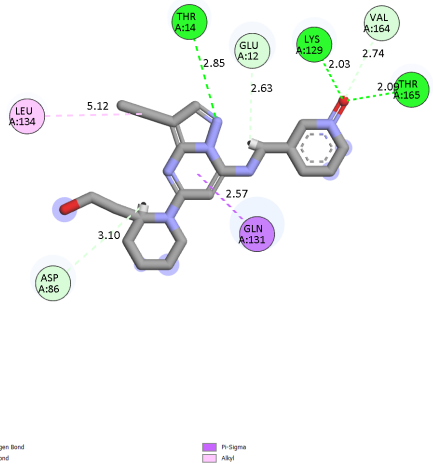


(b) 2D interaction of compound 11 with receptor 6Q4G

Figure 4.6: Molecular docking interactions of compound 11 with receptor 6Q4G.

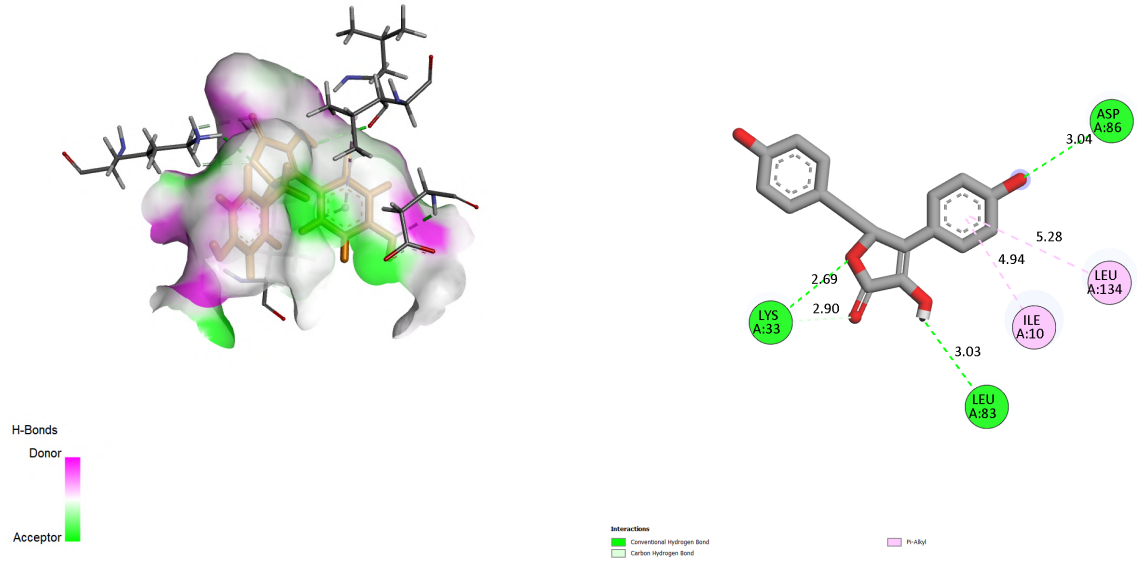


(a) 3D interaction of DINACICLIB with receptor 6Q4G



(b) 2D interaction of DINACICLIB with receptor 6Q4G

Figure 4.7: Molecular docking interactions of DINACICLIB with receptor 6Q4G.



(a) 3D interaction of compound 8 with receptor 6Q4G

(b) 2D interaction of compound 8 with receptor 6Q4G

Figure 4.8: Molecular docking interactions of compound 8 with receptor 6Q4G.

- 6Q4G-L6: The interaction of compound 6 with the active site shows a stable orientation inside the active pocket, mainly due to conventional hydrogen bonds with LEU 83 (2.19 Å), LYS 33 (1.95 Å and 2.91 Å), and ASP 145 (2.28 Å), along with hydrophobic Alkyl and Pi-Alkyl interactions with ILE 10 (4.95 Å), VAL 18 (5.43 Å), LEU 134 (4.60 Å, 4.80 Å, and 5.30 Å), and ALA 31 (3.40 Å).

- 6Q4G-L8: The interaction of compound 8 shows high binding affinity due to polar and hydrophobic interactions, forming conventional hydrogen bonds with LYS 33 (2.69 Å), LEU 83 (3.03 Å), and ASP 86 (3.04 Å), a carbon-hydrogen bond with LYS 33 (2.90 Å), and Pi-Alkyl interactions with LEU 134 (4.94 Å and 5.28 Å) and ILE 10.

- 6Q4G-L11: The interaction of compound 11 forms a stable binding interaction based on conventional hydrogen bonds with ASP 86 (2.12 Å), LYS 33 (2.20 Å and 2.44 Å), and ASP 145 (2.28 Å), further stabilized by a carbon-hydrogen bond with LYS 89 (2.85 Å), along with hydrophobic Pi-Alkyl interactions with LYS 89 (5.32 Å) and ILE 10.

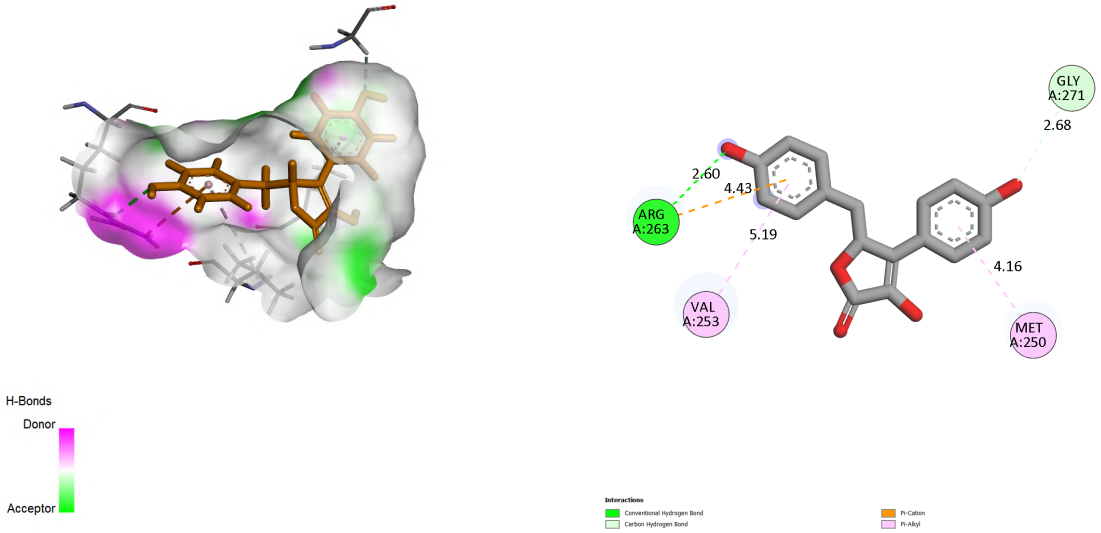
- 6Q4G-DINACICLIB: The interaction of the reference drug shows stable and complex interactions inside the active pocket, forming conventional hydro-

gen bonds with THR 14 (2.85 Å), LYS 129 (2.03 Å), and THR 165 (2.09 Å), carbon-hydrogen bonds with GLU 12 (2.63 Å), VAL 164 (2.74 Å), and ASP 86 (3.10 Å), a Pi-Sigma contact with GLN 131 (2.57 Å), and an Alkyl interaction with LEU 134 (5.12 Å).

4.3.3 *6UDV–ligand interactions*

Complexes	Ligand	Receptor	Amino acid types	Interactions	Distance (Å)	Energy (kcal/mol)
6UDV-Obatoclax mesylate	5-ring	CB	ALA 227 (A)	pi-H	4.57	-0.7
6UDV-L1	O 12	O	LEU 267 (A)	H-donor	2.74	-2.8
6UDV-L11	O 1	O	LEU 267 (A)	H-donor	3.11	-2.2
6UDV-L8	-	-	-	-	-	-

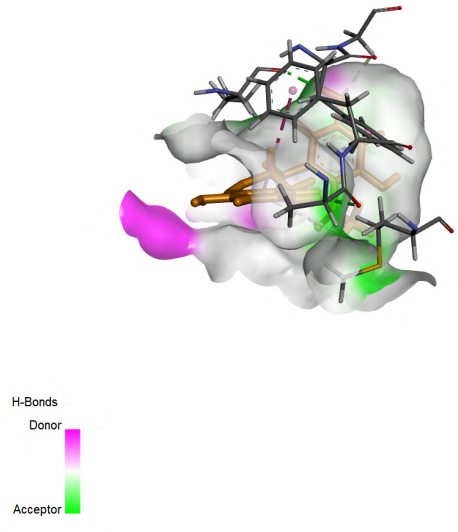
Table 4.15: MOE interaction report for the top-scoring ligands with the 6UDV receptor.



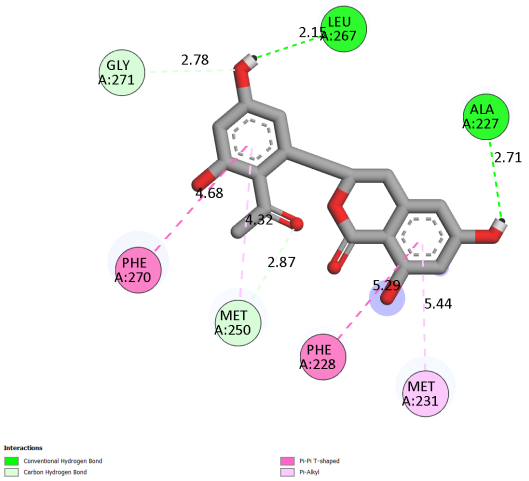
(a) 3D interaction of compound 8 with receptor 6UDV

(b) 2D interaction of compound 8 with receptor 6UDV

Figure 4.9: Molecular docking interactions of compound 8 with receptor 6UDV.

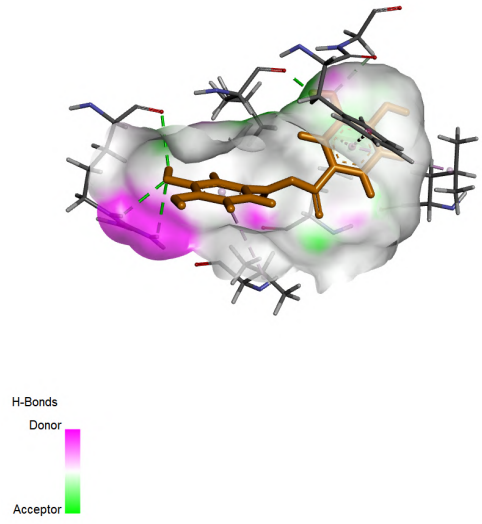


(a) 3D interaction of compound 11 with receptor 6UDV

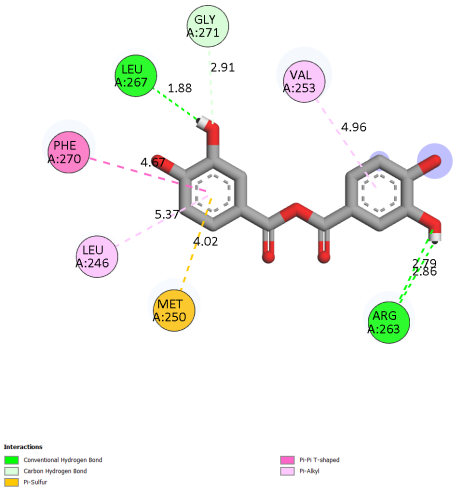


(b) 2D interaction of compound 11 with receptor 6UDV

Figure 4.10: Molecular docking interactions of compound 11 with receptor 6UDV.

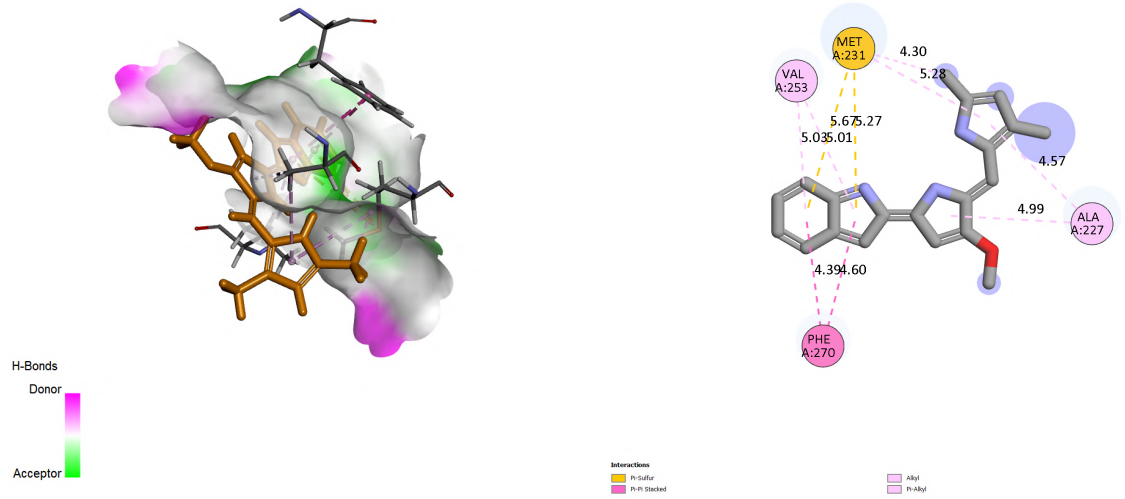


(a) 3D interaction of compound 1 with receptor 6UDV



(b) 2D interaction of compound 1 with receptor 6UDV

Figure 4.11: Molecular docking interactions of compound 1 with receptor 6UDV.



(a) 3D interaction of compound Obatoclax mesylate with receptor 6UDV

(b) 2D interaction of compound Obatoclax mesylate with receptor 6UDV

Figure 4.12: Molecular docking interactions of compound Obatoclax mesylate with receptor 6UDV.

- 6UDV-L1: The interaction of compound 1 exhibits a stable conformation formed by conventional hydrogen bonds with LEU 267 (1.88 Å) and ARG 263 (2.79 Å and 2.86 Å), a carbon-hydrogen bond with GLY 271 (2.91 Å), a Pi-Sulfur interaction with MET 250 (4.02 Å), a Pi-Pi T-shaped interaction with PHE 270 (4.67 Å), and hydrophobic Pi-Alkyl interactions with LEU 246 (5.37 Å) and VAL 253 (4.96 Å).

- 6UDV-L8: The interaction of compound 8 is anchored by a conventional hydrogen bond with ARG 263 (2.60 Å), supplemented by a carbon-hydrogen bond with GLY 271 (2.68 Å), a Pi-Cation interaction with ARG 263 (4.43 Å), and hydrophobic Pi-Alkyl interactions with VAL 253 (5.19 Å) and MET 250 (4.16 Å).

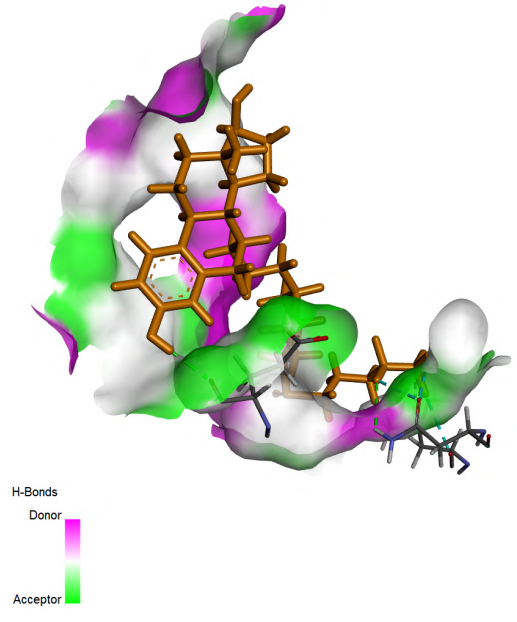
- 6UDV-L11: The interaction of compound 11 exhibits a robust binding mode through strong conventional hydrogen bonds with LEU 267 (2.15 Å) and ALA 227 (2.71 Å), further strengthened by carbon-hydrogen bonds with GLY 271 (2.78 Å) and MET 250 (2.87 Å), a Pi-Pi T-shaped interaction with PHE 228 (5.29 Å), and hydrophobic Pi-Alkyl interactions with PHE 270 (4.68 Å), MET 250 (4.32 Å), and MET 231 (5.44 Å).

- 6UDV-Obatoclax Mesylate: The interaction of the reference drug indicates a hydrophobic mechanism of binding with the absence of conventional hydrogen bonds, forming Pi-Sulfur interactions with MET 231 (5.01 Å, 5.27 Å, and 5.67 Å), Pi-Pi Stacked interactions with PHE 270 (4.39 Å and 4.60 Å), and numerous Alkyl and Pi-Alkyl interactions with VAL 253 (5.03 Å), ALA 227 (4.57 Å and 4.99 Å), and MET 231 (4.30 Å and 5.28 Å).

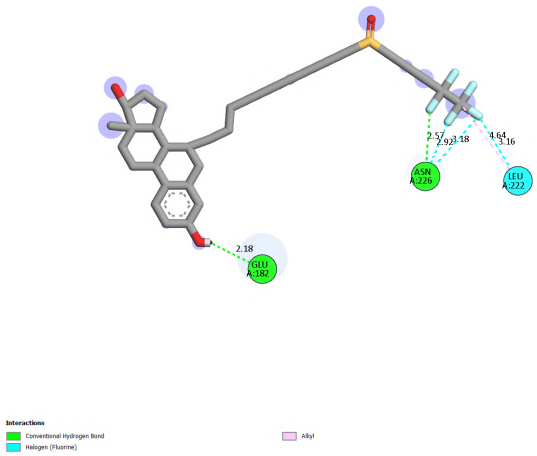
4.3.4 *7NFW–ligand interactions*

Complexes	Ligand	Receptor	Amino acid types	Interactions	Distance (Å)	Energy (kcal/mol)
7NFW-FULVESTRANT	O 55	O	GLU 182 (A)	H-donor	3.16	-1.6
7NFW-L1	-	-	-	-	-	-
7NFW-L5	O 18	NH2	ARG 129 (A)	H-acceptor	3.03	-2.9
	O 18	OH	TYR 130 (A)	H-acceptor	3.07	-2.4
	O 33	NZ	LYS 122 (A)	H-acceptor	3.10	-5.9
	O 33	ND2	ASN 175 (A)	H-acceptor	3.00	-2.4
	6-ring	CG2	VAL 178 (A)	pi-H	4.43	-0.6
7NFW-L11	O 26	OE1	GLU 182 (A)	H-donor	3.29	-1.7
	O 28	OE2	GLU 133 (A)	H-donor	3.31	-0.7
	O 34	NH2	ARG 129 (A)	H-acceptor	3.38	-0.5
	O 37	NH2	ARG 56 (A)	H-acceptor	3.63	-1.4
	O 37	NH1	ARG 129 (A)	H-acceptor	3.02	-4.0

Table 4.16: MOE interaction report for the top-scoring ligands with the 7NFW receptor.

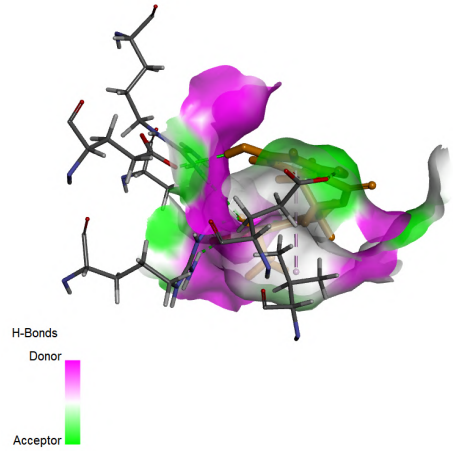


(a) 3D interaction of FULVESTRANT with receptor 7NFW

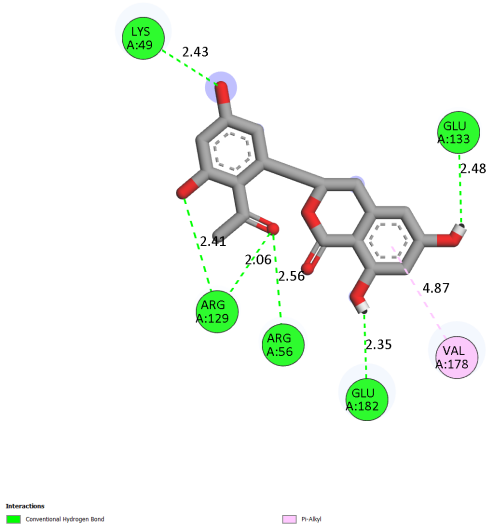


(b) 2D interaction of FULVESTRANT with receptor 7NFW

Figure 4.13: Molecular docking interactions of FULVESTRANT with receptor 7NFW.

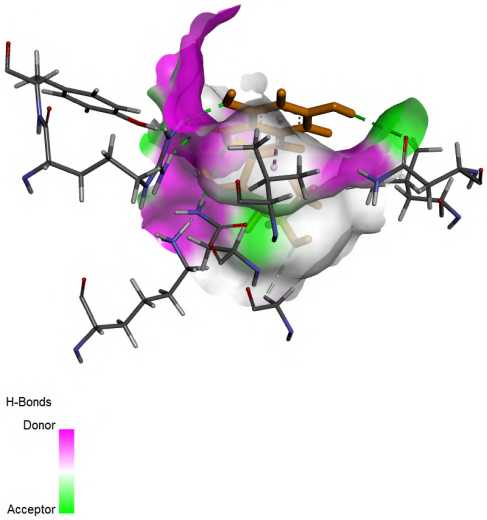


(a) 3D interaction of compound 11 with receptor 7NFW

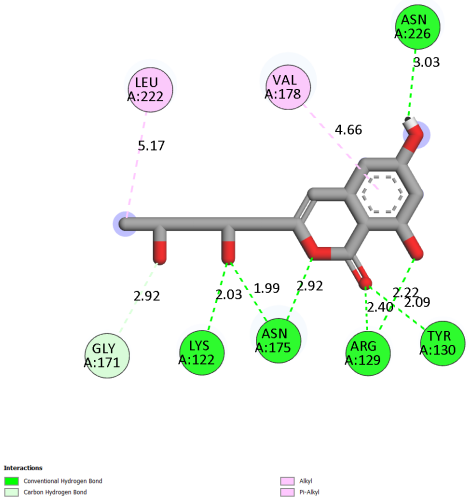


(b) 2D interaction of compound 11 with receptor 7NFW

Figure 4.14: Molecular docking interactions of compound 11 with receptor 7NFW.

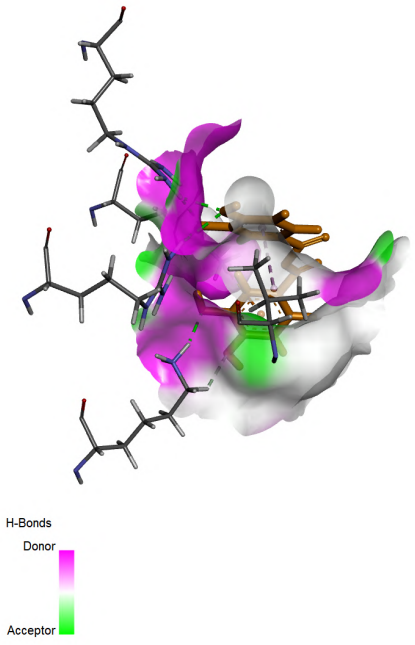


(a) 3D interaction of compound 5 with receptor 7NFW

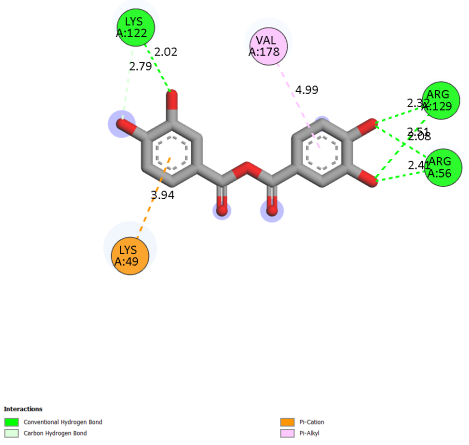


(b) 2D interaction of compound 5 with receptor 7NFW

Figure 4.15: Molecular docking interactions of compound 5 with receptor 7NFW.



(a) 3D interaction of compound 1 with receptor 7NFW



(b) 2D interaction of compound 1 with receptor 7NFW

Figure 4.16: Molecular docking interactions of compound 1 with receptor 7NFW.

- 7NFW-L1: The interaction of compound 1 with the active site yields a complex stabilized by a dense hydrogen-bond network. It forms conventional hydrogen bonds with LYS 122 (2.02 Å), ARG 129 (2.32 Å and 2.51 Å), and ARG 56 (2.08 Å and 2.41 Å), a carbon-hydrogen bond with LYS 122 (2.79 Å), a Pi-Cation interaction with LYS 49 (3.94 Å), and a Pi-Alkyl interaction with VAL 178 (4.99 Å).

- 7NFW-L5: The interaction of compound 5 displays a high binding affinity through extensive hydrogen bonds with ASN 226 (3.03 Å), LYS 122 (2.03 Å), ASN 175 (1.99 Å and 2.92 Å), ARG 129 (2.40 Å), and TYR 130 (2.09 Å and 2.22 Å), a carbon-hydrogen bond with GLY 171 (2.92 Å), and Pi-Alkyl interactions with LEU 222 (5.17 Å) and VAL 178 (4.66 Å).

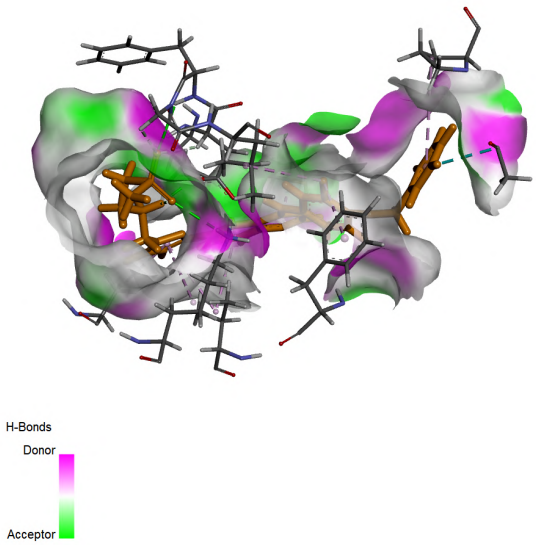
- 7NFW-L11: The interaction of compound 11 is mainly stabilized by a dense hydrogen-bond network with LYS 49 (2.43 Å), ARG 129 (2.06 Å and 2.41 Å), ARG 56 (2.56 Å), GLU 182 (2.35 Å), and GLU 133 (2.48 Å), along with a Pi-Alkyl interaction with VAL 178 (4.87 Å).

- 7NFW-FULVESTRANT: The interaction of the reference drug forms complexes in the terminal regions of its steroidal structure, with conventional hydrogen bonds to GLU 182 (2.18 Å) and ASN 226 (2.57 Å), and halogen (fluorine) interactions with ASN 226 (2.92 Å and 3.18 Å) and LEU 222 (3.16 Å and 4.64 Å).

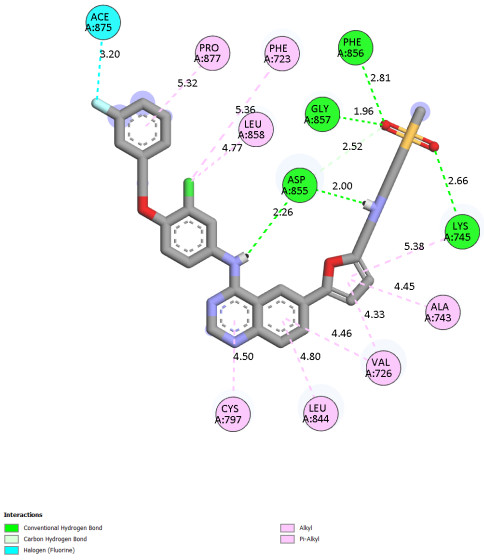
4.3.5 *8A27–ligand interactions*

Complexes	Ligand	Receptor	Amino acid types	Interactions	Distance (Å)	Energy (kcal/mol)
8A27-LAPATINIB DITOSYLATE	N 26	OD2	ASP 855 (A)	H-donor	3.19	-2.0
	N 52	OD1	ASP 855 (A)	H-donor	2.93	-12.6
	O 61	N	GLY 857 (A)	H-acceptor	2.90	-2.9
	N 52	OD1	ASP 855 (A)	ionic	2.93	-4.9
	6-ring	CG2	VAL 726 (A)	pi-H	4.31	-0.6
8A27-L3	O 18	O	GLY 857 (A)	H-donor	2.85	-2.0
	O 37	6-ring	PHE 856 (A)	H-pi	3.66	-1.6
8A27-L5	O 33	OD1	ASP 855 (A)	H-donor	3.39	-0.7
8A27-L11	O 26	O	ILE 759 (A)	H-donor	2.92	-2.8
	O 37	N	GLY 857 (A)	H-acceptor	3.15	-2.9

Table 4.17: MOE interaction report for the top-scoring ligands with the 8A27 receptor.

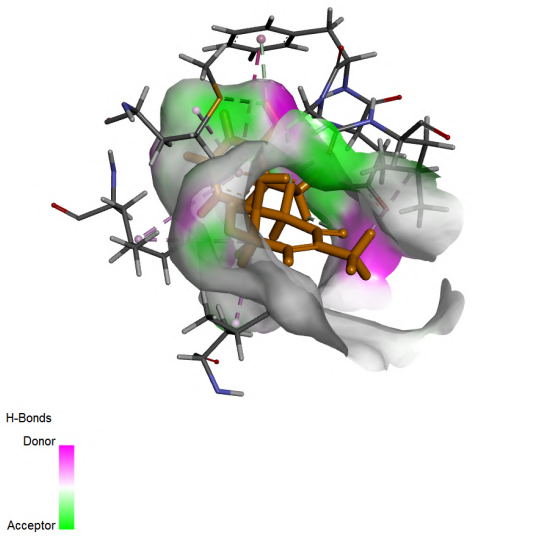


(a) 3D interaction of Lapatinib Ditosylate with receptor 8A27

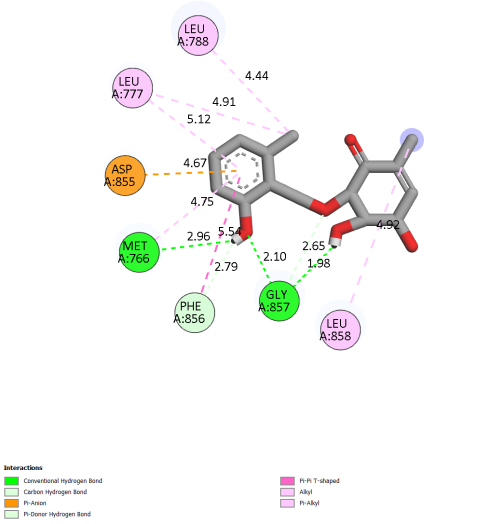


(b) 2D interaction of Lapatinib Ditosylate with receptor 8A27

Figure 4.17: Molecular docking interactions of Lapatinib Ditosylate with receptor 8A27.

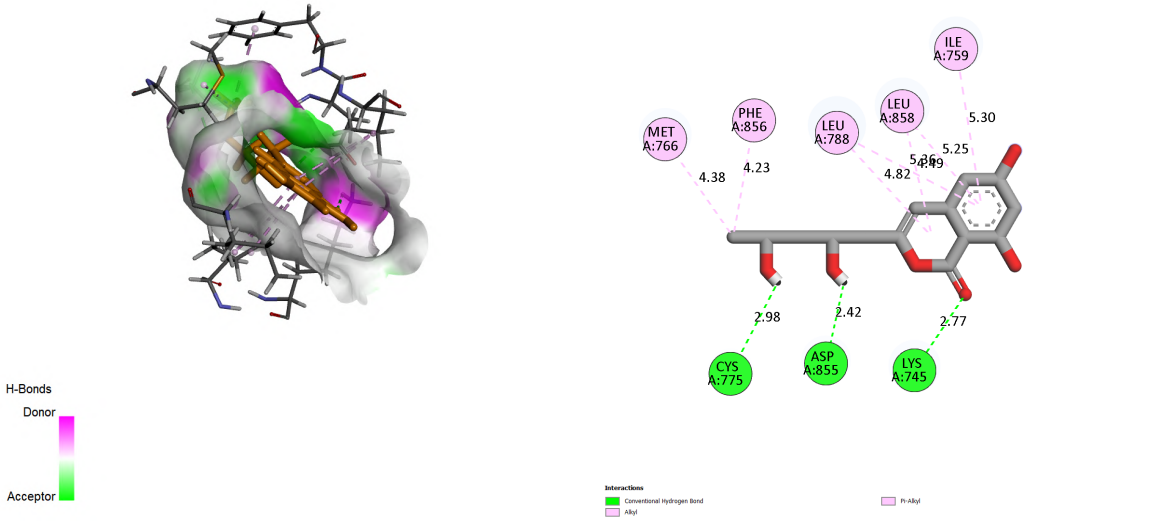


(a) 3D interaction of compound 3 with receptor 8A27



(b) 2D interaction of compound 3 with receptor 8A27

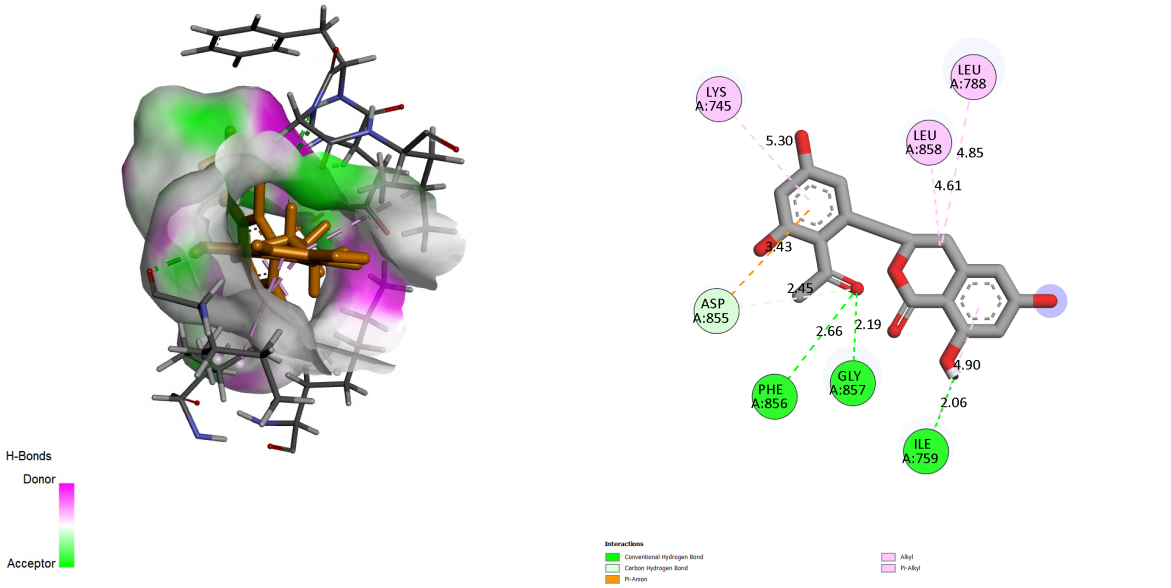
Figure 4.18: Molecular docking interactions of compound 3 with receptor 8A27.



(a) 3D interaction of compound 5 with receptor 8A27

(b) 2D interaction of compound 5 with receptor 8A27

Figure 4.19: Molecular docking interactions of compound 5 with receptor 8A27.



(a) 3D interaction of compound 11 with receptor 8A27

(b) 2D interaction of compound 11 with receptor 8A27

Figure 4.20: Molecular docking interactions of compound 11 with receptor 8A27.

- 8A27-L3: The interaction of compound 3 with the active site reveals multiple conventional hydrogen bonds with MET 766 and GLY 857 (1.98–2.96 Å), an electrostatic interaction with ASP 855 (4.67 Å), and hydrophobic Pi-Alkyl interactions with LEU 777, LEU 788, and LEU 858.

- 8A27-L5: The interaction of compound 5 shows specific conventional hydrogen bonds with CYS 775 (2.98 Å), ASP 855 (2.42 Å), and LYS 745 (2.77 Å), along with Pi-Alkyl interactions between its aromatic ring and MET 766, PHE 856, LEU 788, LEU 858, and ILE 759.

- 8A27-L11: The interaction of compound 11 demonstrates considerable stabilization through hydrogen bonds with ILE 759 (2.06 Å), PHE 856 (2.66 Å), and GLY 857 (2.19 Å), a carbon-hydrogen bond with ASP 855 (2.45 Å), a Pi-Anion interaction with ASP 855 (3.43 Å), and hydrophobic Pi-Alkyl interactions with LYS 745, LEU 788, and LEU 858.

- 8A27-LAPATINIB DITOSYLATE: The interaction of the reference drug is quite complex, forming important hydrogen bonds with ASP 855, PHE 856, GLY 857, and LYS 745, further enhanced by a fluorine-halogen interaction with ACE 875 (3.20 Å), carbon-hydrogen bonds, and hydrophobic Pi-Alkyl interactions with PRO 877, LEU 858, VAL 726, and LEU 844.

5

GENERAL CONCLUSION

5.1 GENERAL CONCLUSION

The work presented in this thesis concerns the application of molecular docking in the discovery of new anticancer agents.

The main objective is to elucidate the binding interactions between a series of phenolic compounds and five target proteins (5Y2S, 6Q4G, 6UDV, 7NFW, and 8A27) implicated in cancer progression.

Firstly, we presented cancer in general, including its characteristics, causes, and current treatments. In the second chapter, we addressed the basic principles of molecular docking, protein–ligand interactions, as well as different ADMET filtering concepts that help us to deepen the study of the proposed compounds, such as drug-likeness properties. In the third chapter, we presented the most important part of this work, which is the development of the molecular docking protocol using MOE software (version 2014) for the proposed inhibitors. Molecular docking was performed on 15 selected phenolic compounds extracted from the COCONUT database to predict the affinity of each compound with the five target proteins.

Following the docking process, four compounds were identified as the best inhibitors: **compound 3** against 5Y2S (-5.27 kcal/mol) and 8A27 (-7.44 kcal/mol), **compound 6** against 6Q4G (-6.84 kcal/mol), **compound 8** against 6UDV (-6.99 kcal/mol), and **compound 11** against 6UDV (-6.98 kcal/mol) and 7NFW (-5.46 kcal/mol). Notably, compounds 6 and 11 against 6Q4G, and compounds 8 and 11 against 6UDV, outperformed their respective reference drugs (Dinaciclib and Obatoclax mesylate). These results were confirmed by the analysis of ligand–protein interactions, which revealed that the binding stability is predominantly governed by conventional hydrogen bonds, carbon–hydrogen bonds, and hydrophobic contacts (Alkyl, Pi-Alkyl, Pi-Pi stacked, and Pi-Sulfur interactions) with key active-site residues.

At the end of this work, we can conclude that the identified compounds (**3, 6, 8, and 11**) are promising inhibitors for the selected targets and can be proposed as potential new drug candidates for cancer therapy.

BIBLIOGRAPHY

[1] A. Chamandy, 'Study of bioactive molecules and their activities in two species of astragalus used in traditional medicine: Astragalus emarginatus labill and astragalus coluteoides will', Ph.D. dissertation, Faculté des Sciences Chimiques, Université de Strasbourg, 2022.

[2] Y. Asses, 'Design by modeling and in silico screening of c-met receptor inhibitors', Ph.D. dissertation, Faculté de Chimie Informatique et Théorique, Université Henri Poincaré, 2011.

[3] G. Cooper, *The cell: a molecular approach*, 3rd ed. Paris, 1997.

[4] H. Adjailia, 'Nutritional risk factors for breast cancer', M.S. thesis, Université Abdelhamid Ibn Badis – Mostaganem, 2018, p. 68.

[5] A. Aïm-Eusébi, E. Prothon, C. Majerholc, D. Barger, Y. Yazdanpanah and J. Aubert, 'The acceptability and effectiveness of a questionnaire for the identification of risk factors for hiv and hepatitis b and c: An observational study in general practice', *European Journal of General Practice*, vol. 24, pp. 60–67, 2018.

[6] J. Jégu, 'Subsequent cancer in survivors of a first cancer: Incidence and impact on survival', Ph.D. dissertation, École Doctorale des Sciences de la Vie et de la Santé, Université de Strasbourg, 2015.

[7] L. Mathias, *Larousse medical: A medical encyclopedia on cd-rom*, 2006. [Online]. Available: https://www.larousse.fr/encyclopedie/medical/cancer/11745.

[8] N. Bouchemella, 'Contribution to the study and mapping of some types of cancer (breast cancer, prostate cancer and colorectal cancer) recorded in the wilaya of jijel', M.S. thesis, Université Med-Seddik Benyahia – Jijel, 2018.

[9] M. Sehil, 'In silico design of new egfr tyrosine kinase inhibitors', M.S. thesis, Université Mohamed Khider de Biskra, 2021.

[10] E. Genoux, 'Flavonoid and verapamil derivatives as ligands of mrp1 and abcg2 transporters: From design to anticancer activity', Ph.D. dissertation, Université de Grenoble, 2011.

[11] R. J. Van Vuuren, H. V. Michelle, E. T. Anne and M. J. Annie, 'Antimitotic drugs in the treatment of cancer', *Cancer Chemotherapy and Pharmacology*, vol. 76, pp. 1101–1112, 2015.

[12] K. Esfahani, L. Roudaia, N. Buhlaiga, S. V. Del Rincon, N. Papneja and W. H. Miller, 'A review of cancer immunotherapy: From the past, to the present, to the future', *Oncology*, pp. 87–97, 2020.

[13] J. Blay, 'Targeted cancer therapies: Not lost in translation', *John Libbey*, vol. 93, pp. 799–804, 2006.

[14] L. Bouhaous, 'Ethnobotanical, phytochemical and biological activities study of some plants used against cancer in traditional herbal medicine in algeria', Ph.D. dissertation, Faculté des Sciences de la Nature et de la Vie, Université Ibn Khaldoun de Tiaret (Algeria), 2022.

[15] B. J. Druker et al., 'Activity of a specific inhibitor of the bcr-abl tyrosine kinase in the blast crisis of chronic myeloid leukemia and acute lymphoblastic leukemia with the philadelphia chromosome', *New England Journal of Medicine*, vol. 344, no. 14, pp. 1038–1042, 2001.

[16] E. V. Jensen and V. C. Jordan, 'The estrogen receptor: A model for molecular medicine', *Clinical Cancer Research*, vol. 9, no. 6, pp. 1980–1989, 2003.

[17] A. W. Roberts et al., 'Targeting bcl2 with venetoclax in relapsed chronic lymphocytic leukemia', *New England Journal of Medicine*, vol. 374, no. 4, pp. 311–322, 2016.

[18] H. E. Bryant et al., 'Specific killing of brca2-deficient tumours with inhibitors of poly(adp-ribose) polymerase', *Nature*, vol. 434, no. 7035, pp. 913–917, 2005.

[19] R. S. Finn et al., 'Palbociclib and letrozole in advanced breast cancer', *New England Journal of Medicine*, vol. 375, no. 20, pp. 1925–1936, 2016.

[20] F. Manolaraki, 'Anthelmintic properties of sainfoin (onobrychis viciifoliae): Analysis of variation factors and the role of involved phenolic compounds', Ph.D. dissertation, Institut National Polytechnique de Toulouse, 2011.

[21] V. Riou, A. Vernhet, T. Doco and M. Moutounet, 'Aggregation of grape seed tannins in model wine—effect of wine polysaccharides', *Food Hydrocolloids*, vol. 16, no. 1, pp. 17–23, 2002.

[22] M. H. Moussa, 'Analytical and biological study of natural flavonoids', Ph.D. dissertation, Université Badji Mokhtar de Annaba, 2011.

[23] P. Andrade, F. Ferreres, M. I. Gil and F. A. Tomás-Barberán, 'Determination of phenolic compounds in honeys with different floral origin by capillary zone electrophoresis', *Food Chemistry*, vol. 60, no. 1, pp. 79–84, 1997.

[24] S. Brianceau, 'Towards a quantitative and qualitative improvement of phenolic compound extraction from red grape pomace using electrotechnologies', Ph.D. dissertation, Compiègne, 2015.

[25] D. Malešev and V. Kuntić, 'Investigation of metal-flavonoid chelates and the determination of flavonoids via metal-flavonoid complexing reactions', *Journal of the Serbian Chemical Society*, vol. 72, no. 10, pp. 921–939, 2007.

[26] M. Medić-Šarić, I. Jasprica, A. Mornar, A. Smolčić-Bubalo and P. Golja, 'Quantitative analysis of flavonoids and phenolic acids in propolis by two-dimensional thin layer chromatography', *JPC–Journal of Planar Chromatography–Modern TLC*, vol. 17, no. 100, pp. 459–463, 2004.

[27] C. Kaur and H. C. Kapoor, 'Anti-oxidant activity and total phenolic content of some asian vegetables', *International Journal of Food Science & Technology*, vol. 37, no. 2, pp. 153–161, 2002.

[28] C. Rice-Evans, N. Miller and G. Paganga, 'Antioxidant properties of phenolic compounds', *Trends in Plant Science*, vol. 2, no. 4, pp. 152–159, 1997.

[29] F. Ferreres, C. García-Viguera, F. Tomás-Lorente and F. A. Tomás-Barberán, 'Hesperetin: A marker of the floral origin of citrus honey', *Journal of the Science of Food and Agriculture*, vol. 61, no. 1, pp. 121–123, 1993.

[30] S. Bessaoud, *Evaluation of some oxidant/antioxidant balance parameters in diabetic rats receiving quercetin*, Thesis, university not specified, 2015.

[31] A. El Kabouss, Z. Charrouf, H. Oumzil, M. Faid, D. Lamnaouer, Y. Miyata and K. Miyahara, 'Characterization of flavonoids from argan tree leaves (argania spinosa (l.) skeels, sapotaceae) and study of their antimicrobial activity', *Revue Marocaine des Sciences Agronomiques et Vétérinaires*, vol. 21, no. 3, pp. 157–162, 2011.

[32] O. Carvalho, C. Botelho, C. Ferreira, H. Ferreira, M. Santos, M. Diaz, T. Oliveira, J. Soares-Martins, M. Almeida and A. S. Júnior, 'In vitro inhibition of canine distemper virus by flavonoids and phenolic acids: Implications of structural differences for antiviral design', *Research in Veterinary Science*, vol. 95, no. 2, pp. 717–724, 2013.

[33] A. Aguilera-Carbo, C. Augur, L. A. Prado-Barragan, E. Favela-Torres and C. N. Aguilar, 'Microbial production of ellagic acid and biodegradation of ellagitannins', *Applied Microbiology and Biotechnology*, vol. 78, no. 2, pp. 189–199, 2008.

[34] J. Doat, 'Tannins in tropical woods', *Bois et Forêts des Tropiques*, no. 182, pp. 37–54, 1978.

[35] D. Vandi, E. N. Nga, J. L. Betti, G. M. E. Loe, P. B. M. Ottou, R. J. Priso, T. N. Foze, P. C. N. Boumsong, S. D. Dibong and E. M. Mpondo, 'Contribution of populations from the cities of yaoundé and douala to the knowledge of tannin- and anthocyanin-containing plants', *Journal of Animal & Plant Sciences*, vol. 30, no. 3, pp. 4797–4814, 2016.

[36] S. D. Dibong, P. B. M. Ottou, D. Vandi, R. C. Ndjib, F. M. Tchamaha and E. M. Mpondo, 'Ethnobotany of anti-hemorrhoidal medicinal plants from markets and villages of the center and littoral regions of cameroon', *Journal of Applied Biosciences*, vol. 96, pp. 9072–9093, 2015.

[37] R. V. Barbehenn, C. P. Jones, A. E. Hagerman, M. Karonen and J.-P. Salminen, 'Ellagitannins have greater oxidative activities than condensed tannins and galloyl glucoses at high ph: Potential impact on caterpillars', *Journal of Chemical Ecology*, vol. 32, no. 10, pp. 2253–2267, 2006.

[38] S. Mole, J. C. Rogler and L. G. Butler, 'Growth reduction by dietary tannins: Different effects due to different tannins', *Biochemical Systematics and Ecology*, vol. 21, no. 6-7, pp. 667–677, 1993.

[39] H. Mehansho, L. G. Butler and D. M. Carlson, 'Dietary tannins and salivary proline-rich proteins: Interactions, induction, and defense mechanisms', *Annual Review of Nutrition*, vol. 7, no. 1, pp. 423–440, 1987.

[40] V. De Freitas, E. Carvalho and N. Mateus, 'Study of carbohydrate influence on protein–tannin aggregation by nephelometry', *Food Chemistry*, vol. 81, no. 4, pp. 503–509, 2003.

[41] J. Aufrere, K. Theodoridou and R. Baumont, 'Nutritional value for ruminants of legumes containing condensed tannins in temperate environments', *Productions Animales*, vol. 25, no. 1, p. 29, 2012.

[42] R. Benkrief, 'Ethnobotanical inventory of medicinal plants of eastern algeria: Chemical study of hammada articulata (moquin) iljin ssp. scoparia pomel. chemical study of 3 new caledonian plants with monoterpenoids', Ph.D. dissertation, Paris 5, 1990.

[43] E. Verykokidou-Vitsaropoulou and C. Vajias, 'Methylated flavones from teucrium polium', *Planta Medica*, vol. 52, no. 05, pp. 401–402, 1986.

[44] J. Marcheix, A. Fleuriet and C. Jay-Allemand, *Phenolic compounds of plants*, No journal given, 2005.

[45] P. Montoro, A. Braca, C. Pizza and N. De Tommasi, 'Structure–antioxidant activity relationships of flavonoids isolated from different plant species', *Food Chemistry*, vol. 92, no. 2, pp. 349–355, 2005.

[46] M. Škerget, P. Kotnik, M. Hadolin, A. R. Hraš, M. Simonič and Knez, 'Phenols, proanthocyanidins, flavones and flavonols in some plant materials and their antioxidant activities', *Food Chemistry*, vol. 89, no. 2, pp. 191–198, 2005.

[47] A. Basli, M. Chibane, K. Madani and N. Oukil, 'Antibacterial activity of polyphenols extracted from a medicinal plant of the algerian flora: Origanum glandulosum desf.', *Phytothérapie*, vol. 10, no. 1, pp. 2–9, 2012.

[48] M. Gerber, M.-C. Boutron-Ruault, S. Hercberg, E. Riboli, A. Scalbert and M.-H. Siess, 'Current news in cancerology: Fruits, vegetables and cancers. a synthesis of the nacre network', *Bulletin du Cancer*, vol. 89, no. 3, pp. 293–312, 2002.

[49] N. Balasundram, K. Sundram and S. Samman, 'Phenolic compounds in plants and agri-industrial by-products: Antioxidant activity, occurrence, and potential uses', *Food Chemistry*, vol. 99, no. 1, pp. 191–203, 2006.

[50] W. Bors, C. Michel and K. Stettmaier, 'Antioxidant effects of flavonoids', *Biofactors*, vol. 6, no. 4, pp. 399–402, 1997.

[51] K. Said, 'Contribution to the study of enzyme inhibition by pyrazolic tripods using molecular modeling', Original French title: Contribution à l'étude de l'inhibition d'enzyme par des Tripodes pyrazoliques par modélisation moléculaire, M.S. thesis, University of Tlemcen, 2016.

[52] Z. Dekkiche and S. Benaiche, 'Molecular docking modeling of enzyme-inhibitor interaction: Case of tyrosine kinase jak2 as a therapeutic target involved in blood pathologies', Original French title: Modélisation par Docking moléculaire de l'interaction Enzyme-Inhibiteur : Cas de la Tyrosine kinase-JAK2 en tant-que cible thérapeutique impliquée dans les pathologies du sang, M.S. thesis, University of Constantine, 2018.

[53] W. Lalout Soufi, 'Contribution to the modeling of interactions in biomolecules', Original French title: Contribution à la modélisation des interactions dans les biomolécules, Ph.D. dissertation, University of Tlemcen, 2015.

[54] K. Chaudhary and N. Mishra, 'A review on molecular docking: Novel tool for drug discovery', *JSM Chem*, vol. 4, no. 3, p. 1029, 2016.

[55] A. Leach and I. Kuntz, 'Unknown title (please check original)', *Journal of Computational Chemistry*, vol. 13, pp. 730–748, 1992, Title missing in provided citation.

[56] H. Böhm, 'Unknown title (please check original)', *Journal of Computer-Aided Molecular Design*, vol. 6, no. 6, pp. 593–606, 1992, Title missing.

[57] H. Böhm, 'Unknown title (please check original)', *Journal of Computer-Aided Molecular Design*, vol. 8, no. 3, pp. 243–256, 1994, Title missing.

[58] M. Atkins, *Physical Chemistry*, 6th ed. 1997, Original French title: Chimie Physique 6ème édition.

[59] A. Gerschel, *Intermolecular interactions: forces in condensed matter*, 2nd ed. EDP Sciences, Paris, 2012, p. 288, Original French title: Liaisons intermoléculaires Les forces en jeu dans la matière condensée.

[60] J. Israelachvili, *Intermolecular and surface forces*, 2nd ed. London: Academic Press, 1997.

[61] M. Rarey, B. Kramer, T. Lengauer and G. Klebe, 'Unknown title (please check original)', *Journal of Molecular Biology*, vol. 26, pp. 470–489, 1996, Title missing.

[62] M. Rarey, B. Kramer and T. Lengauer, 'Unknown title (please check original)', *Bioinformatics*, vol. 15, no. 3, pp. 243–250, 1999, Title missing.

[63] M. Kumar and P. Balaji, 'C-h... interactions in proteins: Prevalence, pattern of occurrence, residue propensities, location, and contribution to protein stability', *Journal of Molecular Modeling*, vol. 20, p. 2136, 2014.

[64] A. Khaldan et al., 'Prediction of potential inhibitors of sars-cov-2 using 3d-qsar, molecular docking modeling and admet properties', *Heliyon*, vol. 7, 2021. DOI: 10.1016/j.heliyon.2021.e06603.

[65] E. Edache, A. Uzairu, P. Mamza and G. Shallangwa, 'Qsar, homology modeling, and docking simulation on sars-cov-2 and pseudomonas aeruginosa inhibitors, admet, and molecular dynamic simulations to find a possible oral lead candidate', *Journal of Genetic Engineering and Biotechnology*, vol. 20, p. 88, 2022. DOI: 10.1186/s43141-022-00362-z.

[66] S. Vilar, G. Cozza and S. Moro, 'Medicinal chemistry and the molecular operating environment (moe): Application of QSAR and molecular docking to drug discovery', *Current Topics in Medicinal Chemistry*, vol. 8, no. 18, pp. 1555–1572, 2008.

[67] PerkinElmer, Inc., *Chemdraw professional 16.0*, Computer software, Waltham, MA, USA, 2016.

[68] Dassault Systèmes BIOVIA, *BIOVIA discovery studio® overview*, PDF document, 3DEXPERIENCE.

[69] T. Sander, J. Freyss, M. von Korff and C. Rufener, 'Datawarrior: An open-source program for chemistry aware data visualization and analysis', *Journal of Chemical Information and Modeling*, vol. 55, no. 2, pp. 460–473, 2015.

[70] RCSB Protein Data Bank, *Homepage*, https://www.rcsb.org, Accessed: 2026, 2026.

[71] COCONUT, *Homepage*, https://coconut.naturalproducts.net, Accessed: 2026.

[72] PubChem, *Quickly find chemical information*, https://pubchem.ncbi.nlm.nih.gov, Accessed: 2026, 2026.

[73] Open Targets Platform, *About the open targets platform*, https://platform.opentargets.org, Accessed: 2026.

[74] SwissADME, *Homepage*, https://www.swissadme.ch, Accessed: 2026.

[75] P. Banerjee, E. Kemmler, M. Dunkel and R. Preissner, 'ProTox 3.0: A webserver for the prediction of toxicity of chemicals', *Nucleic Acids Research*, vol. 52, no. W1, W504–W511, 2024.

[76] O. L. Moldovan, A. Sandulea, I. A. Lungu, Ș. A. Gâz and A. Rusu, 'Identification of some glutamic acid derivatives with biological potential by computational methods', *Molecules*, vol. 28, no. 10, p. 4123, 2023.

[77] C. A. Lipinski, F. Lombardo, B. W. Dominy and P. J. Feeney, 'Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings', *Advanced Drug Delivery Reviews*, vol. 46, no. 1-3, pp. 3–26, 2001.

[78] A. K. Ghose, V. N. Viswanadhan and J. J. Wendoloski, 'A knowledge-based approach in designing combinatorial or medicinal chemistry libraries for drug discovery. 1. A qualitative and quantitative characterization of known drug databases', *Journal of Combinatorial Chemistry*, vol. 1, no. 1, pp. 55–68, 1999.

[79] P. Ertl, B. Rohde and P. Selzer, 'Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties', *Journal of Medicinal Chemistry*, vol. 43, no. 20, pp. 3714–3717, 2000.

[80] C. Chen, M. H. Lee, C. F. Weng and M. K. Leong, 'Theoretical prediction of the complex P-glycoprotein substrate efflux based on the novel hierarchical support vector regression scheme', *Molecules*, vol. 23, no. 7, p. 1820, 2018.

[81] Y. C. Martin, 'A bioavailability score', *Journal of Medicinal Chemistry*, vol. 48, no. 9, pp. 3164–3170, 2005.

[82] A. Dahan, H. J. Lemmer and G. L. Amidon, 'Intestinal permeability and drug absorption: Predictive experimental, computational and In Vivo approaches', *Pharmaceutics*, vol. 11, no. 8, p. 411, 2019.

[83] P. Banerjee, E. Kemmler, M. Dunkel and R. Preissner, 'ProTox 3.0: A webserver for the prediction of toxicity of chemicals', *Nucleic Acids Research*, vol. 52, no. W1, W513–W520, 2024.

[84] S. D. Kumar, A. Rajasekaran and K. S. Kumar, 'In silico toxicity prediction by using ProTox-II computational tools', *Asia-Pacific Journal of Pharmaceutical Toxicology*, 2024.

[85] M. Thompson, Y. Jaiswal, I. Wang and L. Williams, 'Hepatotoxicity: Treatment, causes and applications of medicinal plants as therapeutic agents', *Journal of Phytopharmacology*, vol. 6, no. 3, pp. 186–193, 2017.

[86] T. C. Beck, K. R. Beck, J. Morningstar, M. M. Benjamin and R. A. Norris, 'Descriptors of cytochrome inhibitors and useful machine learning based methods for the design of safer drugs', *Pharmaceuticals*, vol. 14, no. 5, p. 472, 2021.

[87] H. A. Abdulhassan, B. A. Saleh, D. Harkati, H. Khelfaoui, N. L. Hewitt and G. A. El-Hiti, 'In silico pesticide discovery for new anti-tobacco mosaic virus agents: Reactivity, molecular docking, and molecular dynamics simulations', *Applied Sciences*, vol. 12, no. 6, p. 2818, 2022. DOI: 10.3390/app12062818.