Ministry of Higher Education and Scientific Research
Hassiba Benbouali University of Chlef
Faculty of Technology
Process Engineering Department

THESIS
Presented for graduation from
MASTER 2

Field: Process Engineering
Option: Pharmaceutical Process Engineering

Presented by:
KERROUZI Ichraf

Theme:

QSAR, ADMET, Molecular docking and Dynamic studies of natural products as potential inhibitors for Breast Cancer

Proposed jury members:

DOUANI Mustapha	Professor	University of Chlef	President
OTMANINE Khaled	MCB	University of Chlef	Supervisor
ALLICHE Sid Ahmed	MCB	University of Chlef	Examiner

Scholastic year : 2024 /2025

Gratitude

First and foremost, we thank and gratitude **ALLAH** for giving us

the courage, the will and the patience to complete this work.

Our heartfelt thanks and deep gratitude to our supervisor,

Dr. OTMANINE KHALED, for his valuable advice and guidance throughout the project.

I also extend my sincere appreciation to my

Co-supervisors, Dr. BENDRISS HOUARI & Dr. HENTABLI MOHAMED,

and Dr. HAMMOUDI MOUNIR, for their invaluable guidance and insightful comments, as they shared their scientific expertise and knowledge with us throughout this project.

We address our warm and sincere thanks to the members of the jury for having done us the honor of evaluating this modest work

I would also like to express my sincere gratitude to the laboratory members and the research team at Hassiba Ben Bouali University and Saidal Group for their spirit of cooperation and technical support.

Finally, this research would not have been possible without the collective effort

of all these individuals, and I sincerely appreciate their contributions.

DEDICATION

{ يَرْفَعُ اللَّهُ الَّذِينَ آمَنُوا مِنْكُمْ وَالَّذِينَ أُوتُوا الْعِلْمَ دَرَجَاتٍ}

*Praise be to **ALLAH**, by whose grace I have climbed the ladder of success today, and my proudest dream has come true. I dedicate this humble work...*

To the one who raised me, toiled, and struggled for us, to the one whose name I proudly carry:

***My dear father**.*

*To my first and eternal support after **ALLAH**, to the one whose prayers have been the secret to my success throughout my life's journeys: **My beloved mother.***

*To my sisters, whose presence I cherish and am proud of, and whom God has blessed me with as a support in life: **Marwa**, **Maryam**, and **Fatima Al-Zahra**.*

To my uncle, who has been my motivation and support every step of the way.

I dedicate my graduation to all those I love and who are happy for my success. To my friends, with whom I shared my school days and who were a great support.

*Finally, I thank myself from the bottom of my heart, for persevering through the difficulties and challenges. I ask **ALLAH** to make this a good conclusion to the beginning of a path filled with success.*

ICHRAF -Kr ...

Table of contents

Abstract

Abbreviations list

Figures list

Tables list

General Introduction

CHAPTER I:Bibliographic Studies

I.1 Introduction: 5

I.2 Breast Cancer definition: 5

I.3 Types of Breast Cancer 6

I.3.1 Ductal carcinoma Breast Cancer: 6

I.3.2 Lobular carcinoma Breast Cancer: 7

I.3.3 Inflammatory Breast Cancer (IBC): 7

I.4 Breast Cancer Statistics: 8

I.5 Types of breast cancer treatments: 9

I.6.1 Protein presentation: 11

I.6.1.a HER2 Protein: Human epidermal growth factor receptor 2 11

I.6.1.b MCF-7 Protein: Michigan Cancer Foundation-7 12

I.6.1.c CDK4 Protein: Cyclin-dependent kinase 4 13

I.6.1.d AKT1 Protein: AKT1 - RAC-alpha serine/threonine-protein kinase 13

I.6.2 Ligand presentation: 15

I.6.2.1 *Trigonella foenum-graecum* (الحلبة): 15

I.6.2.2 *Curcuma longa L* (الكركم): 19

I.6.2.3 *Atriplex Halimus L* (القطف): 22

I.7 Conclusion: 25

CHAPTER II: Molecular Docking

II.1 Introduction: 27

II.2 Molecular Docking: 27

II.3 Principle of Molecular Docking: 27

II.4 Optimization of Molecules: 28

II.5 Molecular docking simulation: 29

II.6 Target and Ligand Selection: 29

II.6.1 Target selection: 30
II.6.2 Ligand selection: 33
II.7 Target and Ligand Preparation: 35
II.7.1 Gauss View 6.0.16 and Gaussian 09w: 35
II.7.2 Density functional theory (DFT): 36
II.7.3 AutoDock Vina Tools: 38
II.7.4 BIOVIA Discovery Studio: 39
II.8 Pharmacokinetic and ADMET study: 40
II.8.1 ADMET(Absorption, Distribution, Metabolism, Excretion, and Toxicity) 40
II.8.2 Swiss ADME 41
II.8.3 ProTox 3.0: 42
II.9 Conclusion: 42

CHAPTER III: Quantitative Structure Activity Relationship (QSAR)

III.1 Introduction 44
III.2 Definition 44
III.3 Principal of QSAR 45
III.4 Applications of QSAR Approaches 46
III.4.1 Pharmacological QSAR 47
III.4.2 Toxicological QSAR 47
III.5 Molecular Descriptors 47
III.5.1 Classification of descriptors 47
III.5.2 Feature Selection of Descriptors in QSAR Modeling 50
III.6 Multiple Linear Regression (MLR) 51
III.7 Support Vector Regression (SVR) 51
III.8 Validation Methods for QSAR Models 52
III.8.1 Internal Validation: (Training set) 53
III.8.2 External Validation: (test set) 53
III.9: Key Difference between MLR and SVR in QSAR Modeling 55
III.10 Conclusion 56

CHAPTER IV: Experimental Studies

IV.1 Introduction **58**
IV.2 Experimental Study **58**
IV.2.1 Materials and methods **58**
IV.2.1.a Products **58**
IV.2.1.b Materials **59**
IV.2.2 Extraction method: At Hassiba Ben Bouali University **59**
IV.2.3 Physicochemical study of extract: At Laboratory of Dar El Beida Saidal, Algeria.
a) UV-VIS Spectroscopy **62**
b) Infrared Spectroscopy **63**
IV.2.4 Determination of Antioxidant Potential **64**
a) Principal of the DPPH test **64**
b) Operation mode: Protocol of laboratory Saidal dar el beida. **64**
IV.3 Molecular Docking Study **66**
IV.3.1 Material and methods **66**
IV.3.2 Ligand Preparation **67**
IV.3.2.a Optimization (DFT) **67**
IV.3.3 Target Preparation **71**
IV.3.4 Ligand in AutoDock Tools **71**
IV.3.5 Target in AutoDock tools **72**
IV.3.6 Relationship between target proteins **72**
IV.3.7 Pharmacokinetic Properties **73**
IV.3.8 Docking **73**
IV.3.8.a InSilico interaction between Ligand pdbqt & Target pdbqt **73**
IV.3.8.b Vina AutoDock **74**
IV.3.8.c BIOVIA Discovery Studio **75**
IV.3.8.d CB-Dock **76**
IV.4 Quantitative Structure Activity Relationship **78**
IV.4.1 Databases **78**
IV.4.2 Regression methods **93**
IV.4.2.1 Multiple linear model **94**
IV.4.2.2 Support vector model **95**
IV.5 Conclusion **96**

CHAPTER V: Results and Discussion.

V.1 Introduction 98

V.2 Experimental study results 98

V.2.1 Extraction yield 98

V.2.2 Color/Apparence 98

V.2.3 Visible UV results 99

V.2.4 Infrared spectroscopy results 101

V.2.5 Determination of antioxydant Activity 103

V.2.6 Discussion 105

V.3 Molecular docking results 106

V.3.1 Ligand& Target preparation results 106

V.3.1.a Density Theory Calculation 107

V.3.2 In Silico Predicted ADMET Properties 112

V.3.3 In Silico Prediction Pharmacokinetic and Toxicological 113

V.3.4 Protein-ligands interaction and binding sites 117

V.3.4.1 Discovery study 119

V.3.4.2 CB-Dock study 123

V.3.5 Comparison between the commercial treatment and our molecule 126

V.3.6 Discussion 127

V.4 Quantitative structure–activity relationship results 128

V.4.1 MLR model for pIC50 breast cancer 128

V.4.1.1 Feature selection of descriptors in QSAR modeling 128

V.4.1.2 Selection of representative descriptors 129

V.4.2 SVR model for pIC50 breast cancer 132

V.4.3 Models' comparison 134

V.4.4 Discussion 135

V.5 Conclusion 136

General Conclusion 139

ANNEX

REFERENCE

Abstract

This study was conducted to support the ongoing search for novel molecules effective enough to treat breast cancer derived from *Trigonella foenum-graecum*, *Curcuma longa L.*, and *Atriplex Halimus*. The study combined computational techniques and computer simulations to evaluate their efficacy, as well as Fourier transform infrared (FTIR) and ultraviolet/visible (UV/VIS) spectroscopy of plant extract. Quantitative structure-activity relationship (QSAR) modeling was used to predict bioactivity; with 2D_3D-QSAR models developed using multiple linear regression (MLR) and support vector regression (SVR). Comparative analysis revealed that MLR outperformed SVR, achieving an R^2 value of 0.93, an RMSE value of 0.163, and a high Q^2 value of 0.87, demonstrating superior predictive accuracy. Model predictions were validated using cross-validation. Molecular docking simulations evaluated the binding interactions with key breast cancer targets (HER2, CDK4, AKT1, and MCF-7), revealing strong affinities ranging from -7 to -11.4 kcal/mol. Furthermore, in conjunction with the density functional theory (DFT) method (B3LYP/6-311G(d,p)), molecular docking identified two promising candidates, fenugreekine and diosgenin, derived from fenugreek, which demonstrated potent inhibition of all proteins compared to the FDA-approved drugs Ibrance and Capivasertib, respectively. ADMET (absorption, distribution, metabolism, excretion, and toxicity) analysis of diosgenin, curcumin, and arbutin demonstrated their pharmacokinetic affinity and safety, making them promising candidates for further experimental validation.

Keywords: Breast cancer; 2D_3D-QSAR; Molecular Docking; HER2; CDK4; AKT1; MCF-7.

ملخص

أجريت هذه الدراسة لدعم البحث الجاري عن جزيئات جديدة فعالة بما يكفي لعلاج سرطان الثدي المشتقة من الحلبة و الكركم و القطف المالح. جمعت الدراسة بين التقنيات الحسابية والمحاكاة الحاسوبية لتقييم فعاليتها، بالإضافة إلى مطيافية تحويل فورييه بالأشعة تحت الحمراء (FTIR) والأشعة فوق البنفسجية / المرئية (UV / VIS) لمستخلص النبات. تم استخدام نمذجة العلاقة الكمية بين البنية والنشاط (QSAR) للتنبؤ بالنشاط الحيوي، مع تطوير نماذج QSAR ثنائية وثلاثية الأبعاد باستخدام الانحدار الخطي المتعدد (MLR) وانحدار متجه الدعم (SVR). كشف التحليل المقارن أن MLR تفوق على SVR، محققًا قيمة R^2 تبلغ 0.93 وقيمة RMSE تبلغ 0.163 وقيمة Q^2 عالية تبلغ 0.87، مما يدل على دقة تنبؤية فائقة. تم التحقق من صحة تنبؤات النموذج باستخدام التحقق المتبادل. قيّمت عمليات محاكاة الالتحام الجزيئي تفاعلات الارتباط مع أهداف سرطان الثدي الرئيسية (HER2، CDK4، AKT1، وMCF-7)، كاشفةً عن تقارب قوي يتراوح بين -7 و-11.4 كيلو كالوري/مول. علاوة على ذلك، وبالتزامن مع طريقة نظرية الكثافة الوظيفية (B3LYP/6-311G(d,p)) (DFT) ، حدد الالتحام الجزيئي مرشحين واعدين، وهما فينوجريكين والديوسجينين، المشتقان من الحلبة، واللذان أظهرا تثبيطًا فعالًا لجميع البروتينات مقارنةً بالدواءين المعتمدين من إدارة الغذاء والدواء الأمريكية (FDA) إيبرانس وكابيفاسيرتيب، على التوالي. أظهر تحليل ADMET (الامتصاص، التوزيع، الأيض، الإخراج، والسمية) للديوسجينين والكركمين والأربوتين تقاربهما الحرائكي الدوائي وسلامتهما، مما يجعلهما مرشحين واعدين لمزيد من التحقق.

الكلمات المفتاحية: سرطان الثدي؛ 2D_3D-QSAR ؛ الإلتحام الجزيئي؛ HER2 AKT1, MCF-7 وCDK4.

Abbreviations List

Abbreviations	Meaning
ADMET	Absorption, Distribution, Metabolism, Excretion, and Toxicity.
AKT1	AKT1 - RAC-alpha serine/threonine-protein kinase
BBB	Blood-Brain Barrier
BC	Breast Cancer
CDK4	Cyclin-dependent kinase 4
DCIS	Ductal carcinoma in situ
DFT	Density Functional Theory
DPPH	2,2-diphenyl-1-picrylhydrazyl
ER	Estrogen receptor
HA	Hydrogen Acceptor
HER2	Human epidermal growth factor receptor 2
HD	Hydrogen Donor
HIA	Human Intestinal Absorption.
HOMO	the highest occupied molecular orbitals
IBC	Inflammatory breast cancer
IC_{50}	the concentration which corresponds to 50% inhibition
LCIS	Lobular carcinoma in situ
LUMO	the lowest unoccupied molecular orbitals
Log p	water/ octanol partition coefficient
MCF-7	Michigan Cancer Foundation-7
MD	Molecular docking
MLR	Support vector regression
Mw	Molecular weights
PPI	Protein-Protein Interactions
QSAR	Quantitative structure-Activity Relationship
QSPR	Quantitative Structure-Property Relationships
RMSE	Root Mean Squared Error
SDF	Structure Data File
SVR	Support vector regression
SMILES	Simplified Molecular Input Line Entry System
Vd	Volume of Distribution
WHO	World Health Organization

Figures list

N°	Titel	Page
I-1	Ductal Breast Cancer.	6
I-2	Lobular carcinoma Breast Cancer.	7
I-3	Inflammatory Breast Cancer.	8
I-4	Most common site per country, Absolute numbers, Incidence, age in 2022.	9
I-5	Convergence of HER2 and ER pathways in HR-positive/HER2-positive BC, Green arrows show the stimulated pathways after HER2 and HR stimulation.	11
I-6	HER2 gene Human epidermal growth factor receptor 2.	12
I-7	MCF-7 Protein.	12
I-8	CDK4 gene Cyclin-dependent kinase 4.	13
I-9	AKT signaling pathway in breast cancer development and progression.	14
I-10	AKT1 protein.	15
I-11	Fenugreek (*Trigonella foenum-graecum L*) (الحلبة).	16
I-12	Mechanism of action(s) of diosgenin at the cellular level as a cancer chemopreventive/therapeutic agent.	18
I-13	Curcumin **(***Curcuma longa L*) (الكركم).	20
I-14	Selected curcumin derivatives and their activity against selected cancer diseases.	21
I-15	Curcumin inhibition.	22
I-16	*Atriplex halimus*(القطف) .	23
I-17	A schematic summary of the anticancer potential of arbutin, obtained from different plant families.	24
II-18	Principe of Molecular Docking.	28
II-19	Screenshot of PDB.	30
II-20	Plip databases screenshot.	31
II-21	Molecular interaction of the compound selected from databases CDK4.	31
II-22	Molecular interaction of the compound selected from databases HER2.	32
II-23	Molecular interaction of the compound selected from databases MCF-7.	32
II-24	Molecular interaction of the compound selected from databases AKT1.	32
II-25	**(1)** Screenshot of PubChem**. (2)** Download molecule forma 3D SDF.	33
II-26	**(1)** Gaussian view. **(2)** Gaussian Screenshot.	36

II-27	Screenshot of AutoDock Tools.	39
II-28	Discovery screenshot.	39
II-29	ADMET screenshot.	41
II-30	SwissADME screenshot.	41
II-31	ProTox databases screenshot.	42
IV-32	Drying, grinding and weighing the fenugreek.	59
IV-33	Maceration method.	60
IV-34	Filtration of extract.	60
IV-35	Rotary Evaporator (Rotavaporation).	61
IV-36	Spectrophotometer visible UV.	62
IV-37	Spectrophotometer Infrared.	63
IV-38	Fenugreek extract preparation.	65
IV-39	Preparation of DPPH solution.	65
IV-40	Optimization step with GaussView.	68
IV-41	(1) DFT method (2) Unhook write connectivity. (3) Click the 5 lockers. (4) Submit job. (5) Calculation process for Optimization.	69
IV-42	Calculation period for optimization.	70
IV-43	From GJF converts in to PDB.	70
IV-44	HOMO and LUMO energy orbital.	71
IV-45	**(1)** and **(2)** Protein-protein interaction between the Studies receptors.	72
IV-46	Interface of centr Grid Box.	74
IV-47	**(1)** Structure of command Vina. **(2)** Affinity calculates by Vina.	75
IV-48	Interfaces for ligand-out Pdbqt in program BIOVIA Discovery Studio and add protein Pdbqt, show 2D Diagram and Interaction.	76
IV-49	CB-Dock screenshot, select the protein and ligand and CB-Dock result.	77
IV-50	**(a)** Parthenolide template molecule 1. **(b)** Parthenolide template molecule 2. **(c)** Parthenolide template molecule 3. **(d)** Parthenolide template molecule 4.	79
IV-51	Compounds structures used for QSAR model development. **(a)** Series1 analogues with acetone core. **(b)** Series 2 analogues with cyclopentanone. **(c)** Series 3 analogues with cyclohexanone and 4-oioeridone cores.	79
IV-52	Screenshot of Spyder programe.	95
IV-53	MLR code.	95
IV-54	SVR code.	96

V-55	Fenugreek extracts absorbance.	99
V-56	Determine the absorbance of fenugreek.	100
V-57	Spectre of infrared result.	101
V-58	DPPH incubation of fenugreek at different concentration.	103
V-59	Frontier molecular orbitals and their energy gaps of molecules.	108
V-60	Diagram of boiled egg shows in SwissADME of **(1)** Diosgenin, **(2)** Arbutin, **(3)** Fenugreekine and **(4)** Curcumin.	115
V-61	Radar charts generated by SwissADME with the chemical space of **(1)** Diosgenin, **(2)** Arbutin, **(3)** Fenugreekine and **(4)** Curcumin on the upper and lower limits of crucial features for the drug likeness.	116
V-62	**(1)** 2D diagram of interaction CDK4 with Fenugreekine. **(2)** Interpolated charge. **(3)** H-Bonds. **(4)** Hydrophobicity of interaction.	119
V-63	**(1)** 2D diagram of interaction AKT1 with Diosgenin. **(2)** Interpolated charge. **(3)** H-Bonds. **(4)** Hydrophobicity of interaction.	120
V-64	**(1)** 2D diagram of interaction HER2 with Curcumin. **(2)** Interpolated charge. **(3)** H-Bonds. **(4)** Hydrophobicity of interaction.	121
V-65	2D diagram of interaction MCF-7 with Arbutin. Interpolated charge. H-Bonds. Hydrophobicity of interaction.	122
V-66	MLR performance QSAR results.	130
V-67	MLR feature_importance.	131
V-68	Distribution of residues.	131
V-69	MLR Regression plot.	132
V-70	SVR performance QSAR results.	133
V-71	Distribution of residues.	133

Table list

N° of table	Titel	Page
I-1	Chemical constituents of fenugreek.	16
II-2	3D SDF format of Ligand.	34
III-3	Types of Descriptors.	49
III-4	Common Descriptors in Breast Cancer QSAR.	50
III-5	MLR& SVR model's parameters.	54
III-6	Comparison between MLR& SVR model.	55
IV-7	Extraction and DPPH Products.	58
IV-8	Preparation of different sample concentrations.	66
IV-9	PubChem CID for compounds and experimental PIC50 values, predicted by MLR and SVR.	80
V-10	Extract characterization.	98
V-11	UV visible analyses results.	99
V-12	DPPH absorbance.	103
V-13	Percentage of results of DPPH inhibition for different concentrations.	104
V-14	Molecular plants in Gaussian view.	106
V-15	Energy calculation of molecules.	107
V-16	Inhibition Constant.	109
V-17	Ligand PDB and PDBQT by AutoDock Vina Tools.	110
V-18	Target PDB and PDBQT by AutoDock Vina Tools.	111
V-19	Molecules ADMET prediction.	112
V-20	Applying Swiss ADME and Prediction of Toxicity of chemicals molecules.	114
V-21	Lipinski's application for our molecules.	114
V-22	Vina score of interaction studies.	118
V-23	CB-dock Score results.	123
V-24	The interactions Target& Ligand using CB dock.	124
V-25	Drug In Silico.	126
V-26	Relative importance of the selected descriptors MLR model.	129
V-27	Validation Parameter of MLR model.	130
V-28	Validation Parameter of SVR model.	132
V-29	Statistical criteria for the MLR and SVR model.	134

General Introduction

General Introduction

In current literature, breast cancer is a disease represented by a tumor microenvironment marked by inflammatory cells, involved in the neoplastic process by accelerating proliferation, withstanding and transfer of tumor cells from one part to another using receptors [1]. Breast cancer (BC), a heterogeneous group of tumors, is the most commonly diagnosed cancer in women globally. In spite of improvements in early diagnosis and development of several targeted therapeutic methods, breast cancer-related morbidity still remains high. The existing therapeutic approaches are associated with high toxicity, low efficacy, therapeutic resistance and therapy-related morbidity [2]. Drug discovery and the process of new drug design have been formulated much easier in the past two decades by introducing, proliferation of combined physical, biochemical process from computing capabilities and computational approaches. Since the breast cancer is one of the life-threatening problems globally, no effective prescription is still now invented or not available in the market or medical treatment. Although a few medications are still available on the market, treatment has serious side effects and low efficacy.

In recent years, the global health community has become increasingly interested in harnessing the therapeutic potential of traditional plants. Natural plant products have shown promising results as antitumor and anticancer agents. Their effectiveness has been demonstrated through their low toxicity upon application and reduced resistance to hormone-based anticancer agents. These uses are due to their antioxidant and anti-inflammatory properties, which can be both preventive and therapeutic, and are safe for long-term use [3].

The context of this research includes two main parts. The first theoretical part focuses on understanding breast cancer and its diagnostic processes. This will then identify the various proteins responsible for the disease, such as HER2, AKT1, CDK4, and MCF-7, their functions in the human body, and their association with breast cancer. In addition, we will conduct a theoretical study on three plants—fenugreek, turmeric, and Atriplex halimus—that exhibit the highest levels of specific receptor inhibition. The molecular docking method and tools used will be discussed. Quantitative structure-activity relationships (QSARs) are another method for establishing a mathematical relationship between the chemical structure of compounds and their biological activity.

In the second experimental part of this study, we applied an integrated approach that combines advanced computational modeling techniques and algorithms to study selected plant molecules and evaluate their therapeutic efficacy. We primarily relied on three fundamental methodologies: molecular docking, quantum theoretical calculations, and QSAR modeling. Molecular docking techniques were used to predict the optimal binding sites of drug molecules to the active sites of target proteins. We applied density functional theory (DFT) using the B3LYP/6-31G (2d,p) core package. Accurate quantitative properties such as HOMO and LUMO energies and the energy gap were calculated. These analyses were supported by a comprehensive evaluation of ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties to ensure the efficacy and safety of candidate drug compounds. We also developed advanced QSAR models using machine learning algorithms to predict the biological activity of compounds and optimize their pharmacokinetic profiles. All simulations were performed using high-performance computing platforms, providing high accuracy with significant time and cost savings compared to traditional experimental methods.

Therefore, the question remains: Can CDK4, AKT1, HER2, and MCF-7 receptors be inhibited to completely cure breast cancer, or is it sufficient to inhibit these receptors while mitigating the disease's harmful symptoms…? Is it possible to rely on integrating natural compounds into the targeted drug discovery process and determine their effectiveness…? Are computational methods such as molecular docking and QSAR modeling accurately predict biological activity and binding affinity…?

The main objective of this study is to treat breast cancer using medicinal plants found in Chlef, Algeria with a particular focus on their efficacy ...relying on traditional phytochemistry with modern computational drug design and artificial intelligence

Evaluation of bioactive compounds that can effectively interact with key proteins associated with breast cancer. Contributing to the development of new phytotherapeutic agents, significantly reducing treatment costs, and reducing the serious side effects caused by the use of medications.

Chapter Ⅰ

Bibliographic Studies

I.1 Introduction:

Breast cancer remains a significant global health challenge, representing one of the most prevalent malignancies among women and a leading cause of cancer-related mortality worldwide. Despite advances in early detection and treatment, the complexity of breast cancer at the molecular level continues to hinder efforts toward fully effective and personalized therapies. This study aims to provide a comprehensive overview of breast cancer, including its historical background, classification, epidemiology, and treatment strategies. Particular emphasis is placed on the molecular mechanisms underlying the disease, highlighting key protein targets and ligand interactions involved in tumor progression. By understanding the biological underpinnings and current therapeutic approaches, this study seeks to contribute to the ongoing efforts in improving diagnosis, prognosis, and targeted treatment of breast cancer.

I.2 Breast Cancer definition:

Breast cancer is a well-known malignancy that is recognized as the leading cause of death among cancers in women both nationally and internationally according to the World Health Organization (WHO). Cancer occurs when the immune system fails to function properly, or when there is an uncontrolled increase in cell division, either spontaneously or abnormally. Furthermore, various factors can contribute to the development of cancer within the human body, including the frequency of DNA and RNA mutations, exposure to harmful environmental elements such as radioactive materials, poor nutrition, and an unhealthy cellular environment. In addition, individuals may have a genetic predisposition to such abnormalities. This disease is characterized by its great diversity, and is influenced by the complex interaction of genetic, epigenetic, and environmental factors. In addition to other factors, the development of breast cancer is closely linked to changes in signaling pathways, both direct and indirect, which are regulated by the hormone estrogen and its receptors. This hormonal hypothesis is supported by the fact that breast cancer affects women more frequently than men [4].

Cancer develops when cells become abnormal, expand beyond their normal limits, and then spread across the entire body or specific organs [5,6]. It is caused by abnormal cellular activities that characterize cancer, including division, morphogenesis, movement, and

signaling [7,8]. Chemotherapy is one of the most widely used methods for treating different types of cancer. Recently, inhibitors that stop or slow down the growth of cancer cells have been used as promising options for cancer treatment [9,10][11].

I.3 Types of Breast Cancer

I.3.1 Ductal carcinoma Breast Cancer:

Ductal carcinoma in situ (DCIS) is cancer in your breasts' milk ducts. It isn't aggressive and it typically doesn't spread. This early form of breast cancer is usually curable with appropriate treatment, which often includes lumpectomy and radiation therapy. DCIS is highly treatable, and the outlook is excellent [12]. Ductal carcinoma in situ is the most common type of breast cancer [13].

Ductal carcinoma in situ is a very early form of breast cancer. In ductal carcinoma in situ, the cancer cells are confined inside a milk duct in the breast. The cancer cells haven't spread into the breast tissue. Ductal carcinoma in situ is often shortened to DCIS. It's sometimes called noninvasive, preinvasive or stage 0 breast cancer. However, it does require an evaluation and a consideration of treatment options. Treatment for DCIS often involves surgery. Other treatments may combine surgery with radiation therapy or hormone therapy [14].

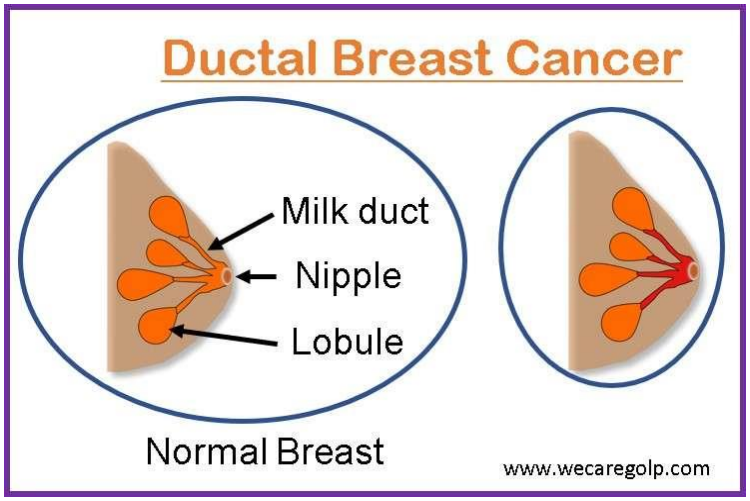


Figure I.1: Ductal Breast Cancer [12].

Ⅰ.3.2 Lobular carcinoma Breast Cancer:

Lobular carcinoma in situ (LCIS) is an uncommon condition in which abnormal cells form in the milk glands (lobules) in the breast. Lobular carcinoma in situ (LCIS) isn't cancer. But being diagnosed with LCIS indicates that you have an increased risk of developing breast cancer [15]. It starts in the lobules of the breast and typically does not spread through the wall of the lobules to the surrounding breast tissue or other parts of the body [13].

Lobular breast cancer (also called invasive lobular carcinoma or ILC) is breast cancer that starts in the milk-producing gland, or lobules, of your breast and has spread into surrounding breast tissue. It accounts for about 10% to 15% of all breast cancers and is the second most common type of breast cancer. Left untreated, lobular breast cancer spreads to nearby lymph nodes, and then to other areas of your body [16].

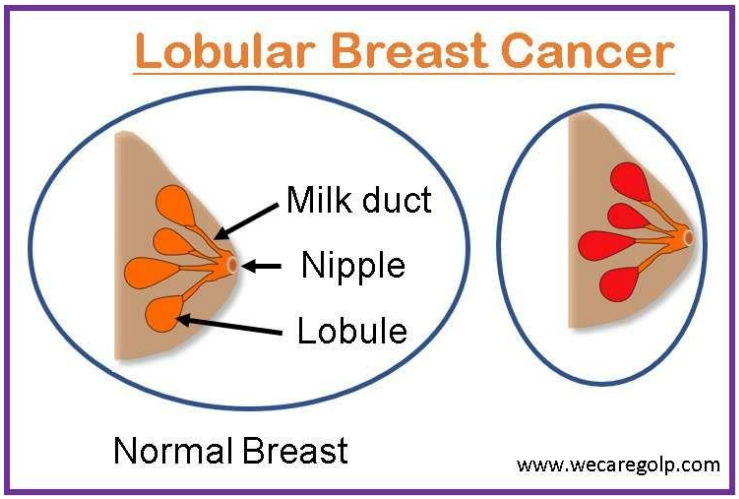


Figure I.2: Lobular carcinoma Breast Cancer [12].

Ⅰ.3.3 Inflammatory Breast Cancer (IBC):

IBC causes pain, redness, swelling and dimpling on the affected breast. IBC results when cancer cells block lymph vessels — the small, hollow tubes that allow lymph fluid to drain out of your breast. The blockage leads to inflammation, causing symptoms that make it easy to mistake IBC for an infection [17]. Inflammatory breast cancer (IBC) is rare. It accounts for only 1% to 5% of all breast cancers. Although it is a type of invasive ductal carcinoma, its symptoms, outlook, and treatment are different. IBC causes symptoms of breast

inflammation like swelling and redness, which is caused by cancer cells blocking lymph vessels in the skin causing the breast to look "inflamed.

Inflammatory breast cancer (IBC) tends to grow quickly, is more likely to have spread at the time it's found, and is more likely to come back after treatment than most other types of breast cancer. Because of this, the survival rates are generally not as high as they are for other types of breast cancer [18].

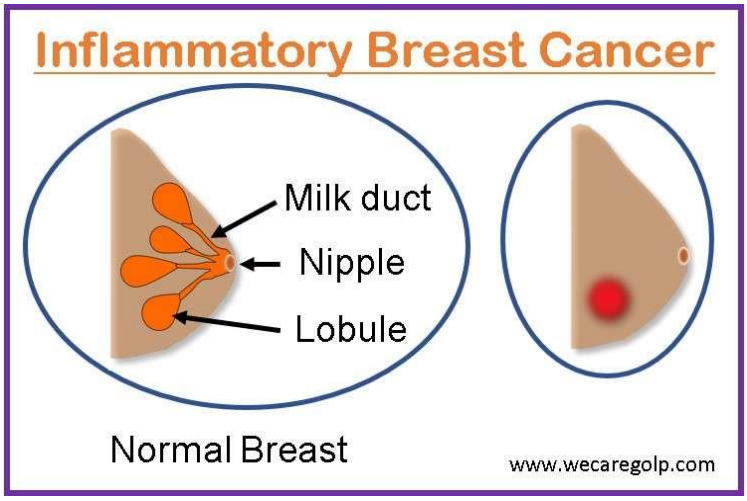


Figure I.3: Inflammatory Breast Cancer [12].

I.4 Breast Cancer Statistics:

Cancer is a major global public health concern, and its incidence has been steadily rising in both developed and developing countries. Even though there are many efficient cancer treatment options available, these treatments can occasionally show resistance and have noticeable adverse effects.

Breast cancer is the most prevalent cancer amongst females in 157 out of 185 countries (Arnold et al. 2022, Bray et al. 2024). It is incidence has increased rapidly and affected more than 2 million women in 2022 worldwide (Arnold et al. 2022, Bray et al. 2024). Due to early detection improvements, most breast cancer cases are identified in the early stages (Centers for Disease Control and Prevention 2024, National Cancer Control Indicators 2018). This study focused on early breast cancer in females. We define breast cancer according to the National Cancer Institute, referring to cancers that have not spread

beyond the breast or the axillary lymph nodes (National Cancer Institute Dictionary of Cancer Terms 2024) [19].

In 2025, an estimated 316,950 women will be diagnosed with breast cancer in the U.S., making it the most common cancer in American women. Every two minutes a woman in the U.S. is diagnosed with the disease [20].

Breast cancer tops the list of cancers prevalent in Algeria, with more than 14,000 new cases recorded annually, a large percentage of which appear before the age of 40, unlike Western countries where breast cancer appears after the age of 60 and above, according to data from the National Cancer Registry [21].

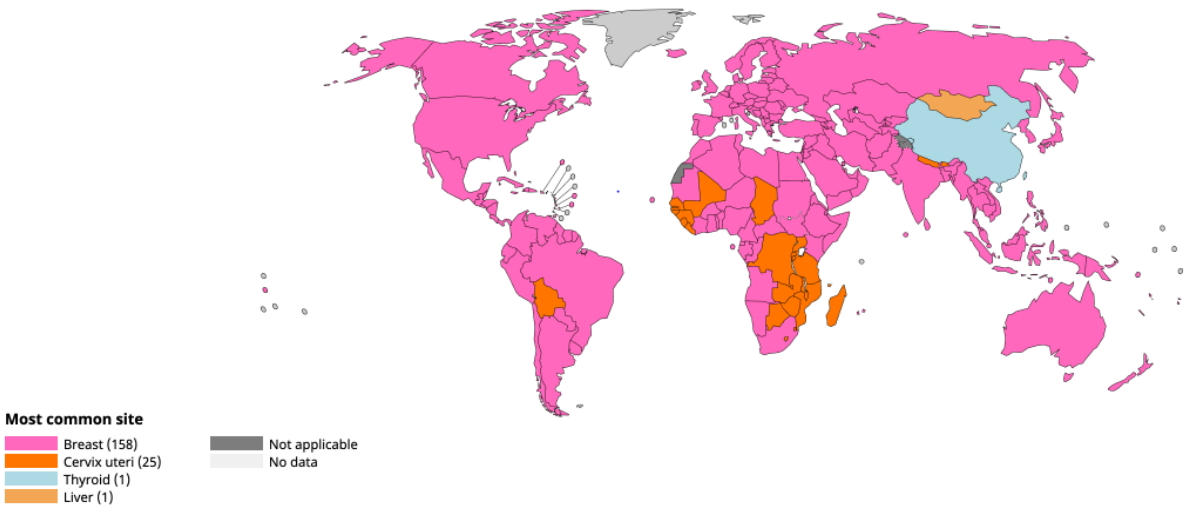


Figure I.4: Most common site per country, Absolute numbers, Incidence, age in 2022 [22]

I.5 Types of breast cancer treatments:

Early breast cancer typically presents with milder symptoms (e.g., a smaller lump and no involvement of lymph nodes) and is associated with prolonged survival compared to advanced stages [23]. Breast cancer treatment often starts with surgery to remove the cancer. Women with breast tumors usually undergo surgery to remove the tumor (lumpectomy or partial mastectomy) [24, 25]. A sample of underarm lymph nodes is also taken (sentinel gland

biopsy). Occasionally, more extensive surgery may be required [26, 27], such as removal of the entire breast (mastectomy or subtotal mastectomy) or a total removal of the underarm lymph nodes. Most people with breast cancer will have other treatments after surgery, such as radiation, chemotherapy and hormone therapy. Some people may have chemotherapy or hormone therapy before surgery. These medicines can help shrink the cancer and make it easier to remove [28].

- **Radiation**: After surgery, additional radiation therapy is usually necessary, especially in the case of partial mastectomy. Radiation can be directed to the entire breast and sometimes to the lymph nodes that drain the breast. In some cases, a single dose of radiation during the surgical removal is sufficient [29,30].
- **Chemotherapy**: Chemotherapy involves drugs that destroy tumor cells. Chemotherapy can cause side effects such as nausea, vomiting, and hair loss. These side effects can be alleviated with medication. It's worth noting that not every patient with breast cancer may need chemotherapy. Chemotherapy can be given before or after surgery [31,32].
- **Biological therapies**: Unlike chemotherapy, biological therapies target tumor cells more specifically and reduce damage to other parts of the body, resulting in fewer side effects. Biologics used against breast cancer include Herceptin and lapatinib [33,34].
- **Antihormonal therapies**: After completing chemotherapy and radiation therapy, some patients are advised to take additional antihormonal therapy in pill form for 5 to 10 years. This treatment is suitable for women whose tumors contain hormone receptors and aims to reduce the risk of tumor recurrence. An example of an antihormonal drug is tamoxifen [35, 36] [37].

I.6 Protein and ligand presentation:

In recent years, the global health community has become increasingly interested in harnessing the therapeutic potential of traditional plants

I.6.1 Protein presentation:

I.6.1.a HER2 Protein: Human epidermal growth factor receptor 2

HER2-positive breast cancer is a breast cancer that tests positive for a protein called human epidermal growth factor receptor 2 (HER2). This protein promotes the growth of cancer cells. In about 1 of every 5 breast cancers, the cancer cells have extra copies of the gene that makes the HER2 protein. HER2-positive breast cancers tend to be more aggressive than other types of breast cancer [38].

Human epidermal growth factor receptor 2 (HER2), a targetable transmembrane glycoprotein receptor of the epidermal growth factor receptor (EGFR) family, plays a crucial role in cell proliferation, survival, and differentiation. Aberrant HER2 signaling is implicated in various cancers, particularly in breast and gastric cancers, where HER2 overexpression or amplification correlates with aggressive tumor behavior and poor prognosis. HER2-activating mutations contribute to accelerated tumorigenesis and metastasis. This review provides an overview of HER2 biology, signaling pathways, mechanisms of dysregulation, and diagnostic approaches, as well as therapeutic strategies targeting HER2 in cancer. Understanding the intricate details of HER2 regulation is essential for developing effective targeted therapies and improving patient outcomes [39].

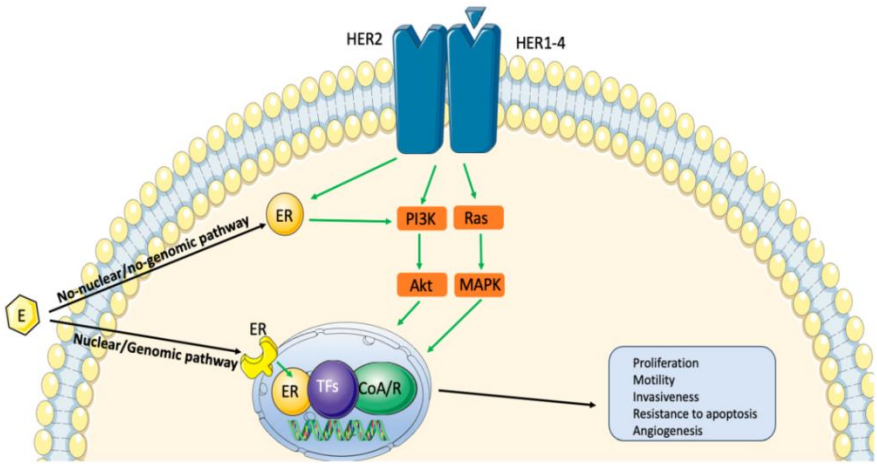


Figure I.5: Convergence of HER2 and ER pathways in HR-positive/HER2-positive BC, Green arrows show the stimulated pathways after HER2 and HR stimulation [40].

Figure I.6: HER2 gene Human epidermal growth factor receptor 2 [41].

I.6.1.b MCF-7 Protein: Michigan Cancer Foundation-7

MCF-7 is a human breast cancer cell line with estrogen, progesterone and glucocorticoid receptors. MCF-7 is a breast cancer cell line isolated in 1970 from a 69-year-old woman. MCF-7 is the acronym of Michigan Cancer Foundation-7, referring to the institute in Detroit where the cell line was established in 1973 by Herbert Soule and co-workers [42].

The MCF-7 cell line is estrogen receptor (ER) positive, progesterone receptor positive, and HER2 negative, which are features commonly seen in luminal subtype breast cancers. Because the MCF-7 line is responsive to estrogen, it is commonly used as a model for estrogen-dependent breast cancers Treatment of MCF7 with oestrogens has been shown to have an anti-apoptotic effect6, 7 whereas treatment with anti-oestrogen chemotherapy drugs (e.g. tamoxifen) can reduce growth of cultures by inhibiting proliferation and inducing apoptosis [43].

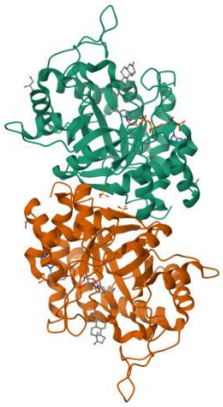


Figure I.7: MCF-7 Protein [41].

I.6.1.c CDK4 Protein: Cyclin-dependent kinase 4

A gene that makes a protein involved in the cell cycle (the process a cell goes through each time it divides). Mutations (changes) in the CDK4 gene may cause cells to divide too quickly or in an uncontrolled way. This may cause abnormal cells, including cancer cells, to grow. Cyclin-dependent kinases (CDKs), in complex with their cyclin partners, modulate the transition through phases of the cell division cycle. Cyclin D–CDK complexes are important in cancer progression, especially for certain types of breast cancer [44].

CDK inhibitors have gained significant interest in the treatment of hormone receptor-positive metastatic breast tumors in recent years. CDKs are a type of protein kinase that specifically phosphorylates serine and threonine amino acid residues to regulate the cell cycle and its progression. The prevailing perspective on the role of CDKs is that CDK4 and CDK6, in conjunction with D-type cyclins, impede the retinoblastoma (Rb) protein pathway to trigger the transition from the G1 to S phases of the cell cycle [45].

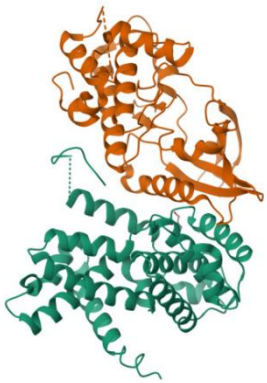


Figure I.8: CDK4 gene Cyclin-dependent kinase 4 [41].

I.6.1.d AKT1 Protein: AKT1 - RAC-alpha serine/threonine-protein kinase

AKT1 - RAC-alpha serine/threonine-protein kinase; AKT1 is one of 3 closely related serine/threonine-protein kinases (AKT1, AKT2 and AKT3) called the AKT kinase, and which regulate many processes including metabolism, proliferation, cell survival, growth and angiogenesis. *AKT1* is a member of the serine-threonine kinase class that plays a key role in cellular processes, including growth, proliferation, survival, and angiogenesis. It is a downstream mediator of phosphatidylinositol 3-kinase which, along with *AKT1*, is a key mediator of proliferation and survival pathways frequently activated in cancer [46].

In general, AKT1 knockdown leads to inhibition of tumor growth via blocking the cell-cycle progression and/or promoting apoptosis in breast cancer model systems [47-49]. Similarly, overexpression and/or constitutive activation of Akt1 in the mammary epithelial cells inhibits the pro-apoptotic signals as well as activates the survival signals to support the process of tumorigenesis [50,51] [52].

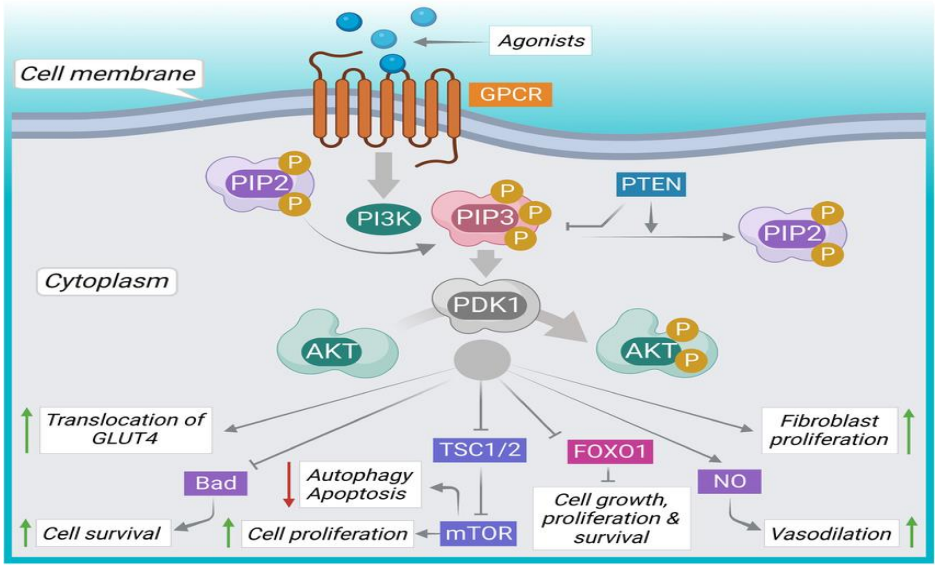


Figure I.9: AKT signaling pathway in breast cancer development and progression [53].

It is a major intracellular pathway leading to cell survival and proliferation. PI3K activation induces the phosphorylation of PIP2 to PIP3, which promotes PDK1 activation. Phosphorylation of FOXO1 by AKT inhibits its transcriptional activities, leading to cell growth, proliferation, and survival. Additionally, AKT inhibits TSC1/2, leading to mTOR activation, which simultaneously inhibits autophagy and apoptosis and stimulates proliferation. Gray arrows represent subsequent events, gray lines represent inhibition, upper green arrowheads represent activation/up-regulation, and lower red arrowheads represent repression/down-regulation [53].

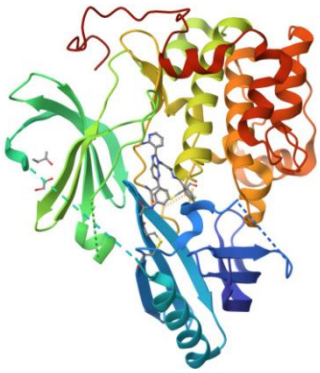


Figure I.10: AKT1 protein [41].

I.6.2 Ligand presentation:

In recent years, the global healthcare community has become increasingly interested in harnessing the therapeutic potential of traditional plants, and new and alternative options are being adopted to combat the initiation and progression of human cancers [54]. Among these approaches and possibilities is the use of molecules isolated from traditional medicinal herbs, which play a pivotal role in cancer prevention and treatment, either alone or in combination with available chemotherapeutic agents. Compounds that modulate these carcinogenic processes are potential candidates for cancer therapy and may eventually reach clinical applications.

I.6.2.1 *Trigonella foenum-graecum* (الحلبة):

Fenugreek is one of the oldest traditional plants in the Mediterranean, The plant has been mainly found on the continents of Asia (India and China), parts of Europe, Africa, Australia, and North and South America and is consumed as a medicinal plant for a variety of ailments [54]. In modern times, with the reemergence of food products as natural medicines, researchers have shown great interest in understanding the therapeutic effects of fenugreek seeds against many diseases, including cancer. Fenugreek is an erect, annual herb with trifoliate leaves, reaching a height of 0.3–0.8 m. It bears white or yellow flowers and produces long, slender pods, ranging in color from yellow to brown (see figure 11) [55]. Several studies indicate that the fenugreek plant contains several bioactive compounds described for their pharmacological effects including steroidal sapogenins. Trigonella foenum-graecum is an

annual plant belonging to the family Fabaceae. The green leaves and seeds of fenugreek are widely used in food and medicinal applications dating back to ancient times [56].

Figure I.11: Fenugreek (*Trigonella foenum-graecum L*) (الحلبة) [57].

Overall, fenugreek's dual role as a culinary herb and a medicinal plant underscores its cultural and medicinal significance in Algeria and beyond. The ongoing research into its therapeutic properties highlights its potential as a valuable natural remedy in modern healthcare practices [58].

Table I.1: Chemical constituents of fenugreek. [59].

Class of Compounds	Phytochemical Constituents
Steroidal sapogenins	Diosgenin, Yamogenin, Smilagenin, Saponaretin, Gitogenin
Flavonoids	Quercetin, Rutin, Vitexin, Isovitexin
Saponins	Graecunins, Fenugrin B, Fenugreekine, Trigofoenosides A-G
Alkaloids	Neurin, Trigonelline, Trimethylamine, Choline, Gentianine

I.6.2.1.a Fenugreekine:

Fenugreek is a leguminous plant used worldwide as a spice to enhance the flavor of foods. Known for its medicinal properties, fenugreek contains numerous chemical constituents, including steroidal sapogenins. Fenugreek also has anticancer properties, and its

chemical constituents are known to induce apoptosis. Furthermore, fenugreek induces a dose-dependent effect on human breast cancer cells [60].

Fenugreekine is a bioactive steroidal saponin primarily isolated from fenugreek *(Trigonella foenum-graecum),* a plant widely used in traditional medicine and culinary practices. This compound has garnered significant attention due to its potent anti-diabetic effects, demonstrated by its ability to enhance insulin sensitivity and inhibit carbohydrate-digesting enzymes, thereby reducing postprandial hyperglycemia (Hamden and al., 2010). Additionally, fenugreekine exhibits hypolipidemic properties by modulating lipid metabolism and decreasing serum LDL cholesterol and triglyceride levels (Sauvaire and al., 1998). Its anti-inflammatory and antioxidant activities have also been reported, attributed to its capacity to scavenge free radicals and suppress pro-inflammatory cytokines (Bin-Hafeez and al., 2003). Recent studies suggest potential anti-cancer effects, where fenugreekine may induce apoptosis and inhibit proliferation in certain cancer cell lines (Raju et al., 2004). Given its multifaceted pharmacological profile, fenugreekine holds promise for developing nutraceuticals and therapeutic agents targeting metabolic and inflammatory disorders.

I.6.2.1.b Diosgenin:

Diosgenin is the most abundant steroidal saponin in fenugreek seeds and is found in the plant's oil. Diosgenin is a steroidal plant saponin and the main bioactive compound found in fenugreek (*T. foenum-graecum*) seeds. Structurally, diosgenin [(25R)-spirost-5-en-3b-ol] is a saponin consisting of a hydrophilic sugar moiety linked to a hydrophobic steroid aglycone. Additionally, diosgenin inhibits cancer cell proliferation and induces apoptosis in a variety of cancer cell lines, including colorectal, liver, breast, osteosarcoma, and leukemia. The primary mechanism of action of diosgenin is to modulate multiple cellular signaling pathways that play prominent roles in cell cycle regulation, differentiation, and apoptosis [61]. Diosgenin is an attractive molecule with versatile properties and is used in pharmaceuticals, functional foods, and cosmetics. The results showed that diosgenin affects the apoptotic pathway in cancer cells. The level of Akt-mediated signaling, which is essential for breast cancer cell growth, was decreased as a result of diosgenin treatment [62].

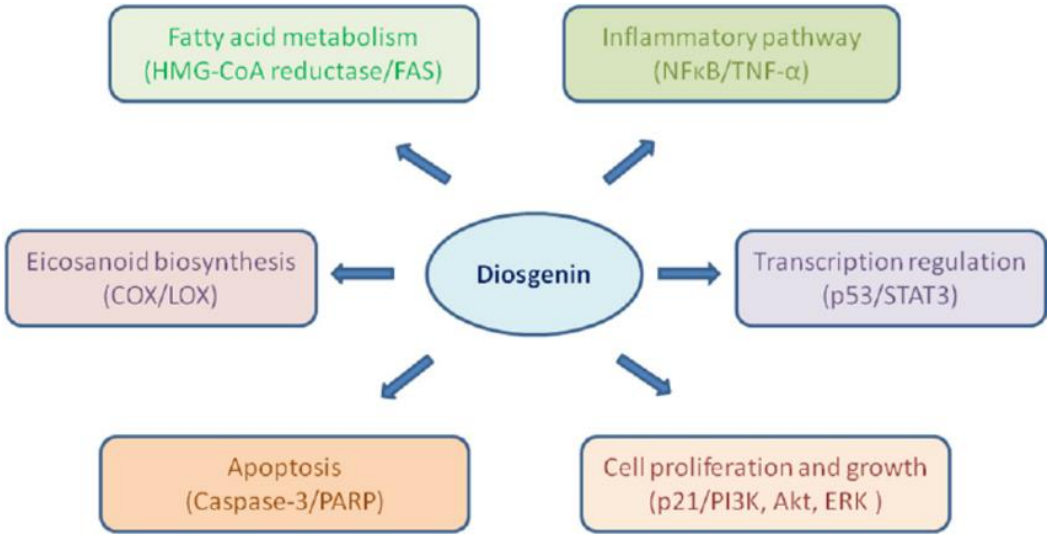


Figure I.12: mechanism of action(s) of diosgenin at the cellular level as a cancer chemopreventive/therapeutic agent [63].

I.6.2.1.c Anti-cancer effect:

Fenugreek is of interest for use in cancer therapy because it has antioxidant and immunostimulatory activity. Both its seeds and leaves contain phytochemicals including flavonoids, alkaloids, and tannins. The putative therapeutic benefits of fenugreek are considered to be due to the phytoestrogens, saponins, and flavonoids present in the plant's seeds and leaves. In this effort, many studies have demonstrated the protective effect of fenugreek seeds in experimental models of cancer using cell lines or experimental animals. Trigonella foenum-graecum, traditionally used to treat disorders such as diabetes, high cholesterol, wounds, inflammation and gastrointestinal ailments, has been recently demonstrated to possess anticarcinogenic potential [64].

In human breast cancer cell line MCF-7 experiment (Li and al., 2019), diosgenin could significantly inhibit the proliferation of breast cancer cell line MCF-7 by regulating the methylation status and activating microRNA-145 (miR-145), and the treatment was safe and effective [65]. Additionally, diosgenin has the potential to inhibit activated pro-inflammatory and pro-survival signaling pathways and promote the death of a variety of cancer cells (Shishodia and Aggarwal, 2006). Additionally, diosgenin inhibits the growth of oestrogen

receptor-positive MCF-7 cells by activating the p53 tumor suppressor gene. Further, BCL2 is downregulated in estrogen receptor-negative MDA-MB-231 breast cancer cells (Srinivasan and al., 2009). Although the potential of diosgenin to treat breast cancer has been demonstrated its interaction with proteins involved in the progenesis of breast cancer has yet to be investigated [66]. Diosgenin could down-regulate FAS expression and induce apoptosis in HER2-overexpressing cancer cells through modulating Akt, mTOR and JNK phosphorylation, and further suggest that diosgenin has therapeutic implications in HER2-overexpressing cancer cells [67].

I.6.2.2 *Curcuma longa L* (الكركم):

Turmeric (*Curcuma longa L.* (CL)), a perennial herb commonly called turmeric or curcumin, belongs to the Zingiberaceae family, native to India, China, Japan and other South Asian countries as traditional medicinal systems for centuries (Jin and al., 2007) [68]. It is a perennial herbaceous plant that grows up to 2 m without stem and rhizomes-stock but with erect leafy shoots bearing up to twelve leaves Curcumin is the yellow pigment that is found in the rhizomes of Curcuma longa (see figure 13). This pigment is the main phytochemical with anticancer properties found in turmeric and belongs to the family of polyphenols (Henrotin and al., 2013; Basnet and Skalko-Basnet, 2011). In the ethanolic extraction of turmeric, there are three main types of curcuminoids; these polyphenols are called curcumin I (diferuloylmethane), curcumin II (desmetoxicurcumin), and curcumin III (bisdemetoxicurcumin). That said, it is important to take into account that the commercial extracts of "curcumin" are actually a mixture of the three different curcuminoids, with curcumin I having the highest concentration (approximately 77%) (Basnet and Skalko-Basnet, 2011). At present, the anti-cancer potential of curcumin has been intensively investigated, compared to other phytochemicals, which has allowed much of its intracellular mechanisms to be defined. The anticancer potential of curcumin is mainly due to its ability to inhibit and/or activate various intracellular transcription factors, thus regulating the expression of various proteins that participate in tumor growth and development. [69].

Figure I.13: Curcumin **(***Curcuma longa L*) (الكركم) [70].

I.6.2.2.a Curcumin:

Curcumin is a natural compound found in the roots of the turmeric plant, especially long turmeric (*Curcuma longa L.*). Curcumin is one of the phytochemicals being intensively studied by pharmacologists and formulation scientists to discover its potential therapeutic effects and address innate biopharmaceutical problems. Its efficacy in inhibiting tumor growth has sparked interest in its potential for treating various types of cancer, such as breast, lung, prostate, and brain cancer. Curcumin has been shown to have anticancer and chemopreventive effects against breast cancer. According to previous studies, curcumin is a useful herb in cancer treatment. It affects cancer growth, progression, and metastasis at the molecular level. In addition, it kills cancer cells and reduces the growth of new blood vessels in tumors (angiogenesis) and the spread of cancer (metastasis). Curcumin exerts its anticancer effects through a complex molecular signaling network, including proliferative pathways, estrogen receptors (ER), and human epidermal growth factor receptor 2 (HER2) [71]. Historically, curcumin's bright color led to its use as a dye, while its flavor made it popular as a kitchen spice.

Curcumin is classified as a diarylheptanoid. Diarylheptanoids are a group of compounds characterized by two aromatic rings connected by a chain of seven carbon atoms [72].

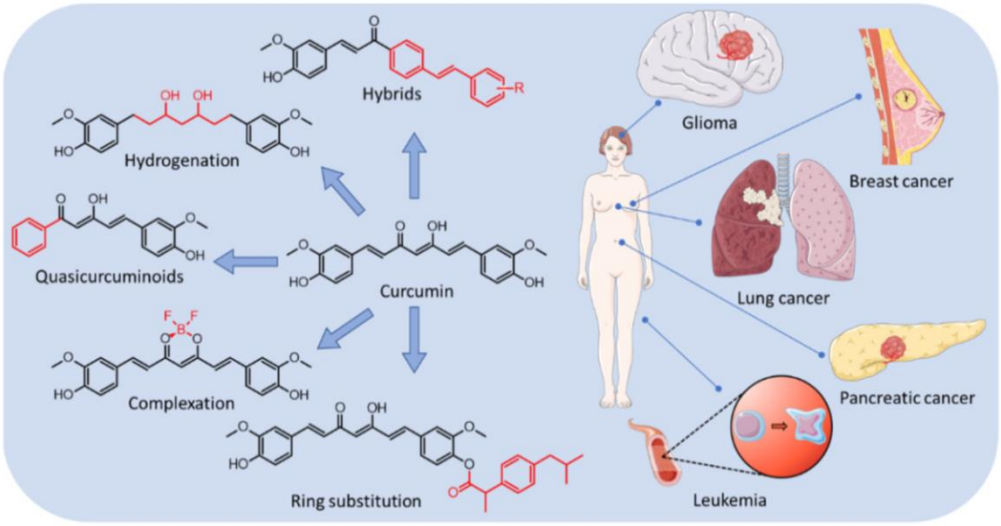


Figure I.14: Selected curcumin derivatives and their activity against selected cancer diseases [72].

I.6.2.2.b Anti-cancer effect:

The antitumoral properties of curcumin relate to the regulation of cellular pathways. Curcumin is recognized for its efficacy in cancer therapy by inducing apoptosis, inhibiting cell proliferation, and preventing the transformation of normal into tumoral cells [73,74]. The induction of apoptosis by inhibiting signaling pathways such as NF-κB, PI3K/Akt, and MAPK prevents tumor cell survival [75] [76].

Curcumin, which is an active compound from turmeric, has been accounted to have an anti-cancer and chemoprevention effect on breast cancer. According to previous studies, curcumin is a useful herb in cancer treatment. It affects cancer growth, development, and spread at the molecular level. Besides, it kills cancer cells and reduces the growth of new blood vessels in tumors (angiogenesis) and spread of cancer (metastasis). Curcumin exerts its anticancer impact through a complicated molecular signaling network, including multiplication, Estrogen Receptor (ER), and human epidermal growth factor receptor 2 (HER2) pathways [71].

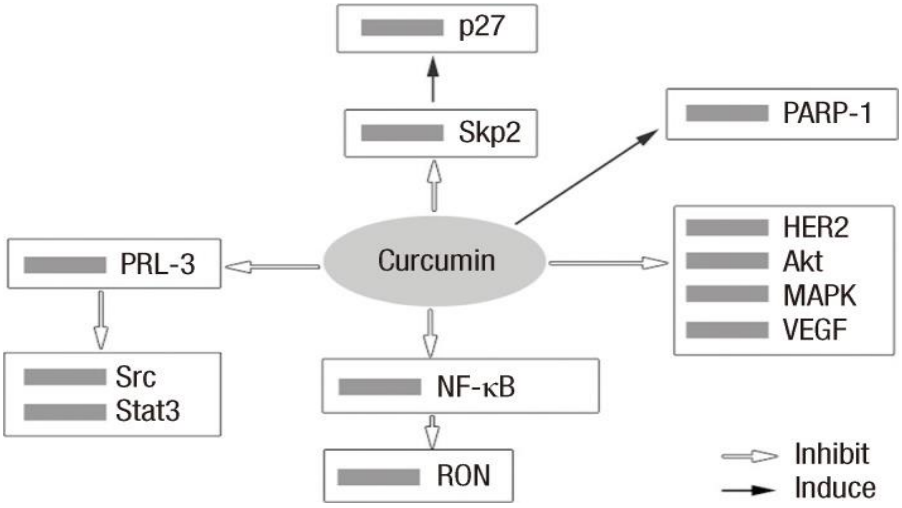


Figure I.15: Curcumin inhibition [77].

Human epidermal growth factor 2 (HER2) is an important oncoprotein, which is overexpressed in about 15% to 25% of breast cancers. Curcumin decreased HER2 oncoprotein, phosphorylation of Akt, MAPK, and the expression of NF-κB in both BT-474 and SK-BR-3-hr cells (Figure 15) [**77**].

Ⅰ.6.2.3 *Atriplex Halimus L* (القطف):

Atriplex Halimus L is an aromatic plant belonging to the Amaranthaceae family. It is commonly known as "Rghel" and "Lgtef" [78]. *A. halimus* is a halophytic shrub widely distributed in arid and semi-arid Mediterranean regions, known for its tolerance to highly saline soils, as well as for exerting allelopathic effects on other plants [79,80]. It grows up to 3 m tall and is branched at the base. The bark is grayish-white, and the leaves are 10–30 mm long and 5–20 mm wide (see figure 16). The leaves are highly variable, ranging from rounded to lanceolate; they are blunt at the base with a short petiole This plant represents a potential source for economical utilization; it can provide forage sources with a good nutritive value during the dry seasons [81;82] [83].

The Mediterranean region is home to Atriplex *halimus*, a halophytic shrub with high protein content [84]. In drought-prone areas, it serves as essential livestock fodder due to its nutritional value [85,86]. Historically, *Atriplex halimus* has been used for its medicinal properties, treating a range of health issues, from inflammation and hormonal imbalance to

cardiovascular disease [84-89]. Recent studies support its potential, highlighting its antioxidant, anticancer, antibacterial, and antidiabetic activities [90-92].

Figure I.16: *Atriplex halimus* (القطف) [93].

In Algeria, A. halimus is mainly used by the local population as a remedy to treat cancer [94], diabetes mellitus and anaemia [95][96].

I.6.2.3.a Arbutin:

Arbutin (C12H16O7), also known as β-arbutin, is a hydroquinone glucoside. This compound was first reported from the leaves of Arbutus unedo L. (family: Ericaceae) [97]. Arbutin structurally differs from its isomer α-arbutin by having a β-glucose unit instead of a α-glucose unit. Since its discovery, arbutin has been found in approximately 50 other plant families. This glycoside is capable of inhibiting melanin production by inhibiting tyrosinase. Chemically, arbutin is a beta-glucoside derived from hydroquinone. Its structure consists of a hydroquinone molecule bound to a glucose molecule [98]. Arbutin has been traditionally used in herbal medicine for its antimicrobial properties [99]. As a competitive inhibitor of the enzyme tyrosinase, which plays a crucial role in melanin production, arbutin is widely used in cosmetics. To date, arbutin has demonstrated potential anticancer properties in several cancers by inducing apoptosis, inhibiting inflammatory markers, and suppressing the phosphoinositide 3-kinase/protein kinase B/mechanistic target of rapamycin (PI3K/Akt/mTOR) signaling pathway [100,101] [102].

I.6.2.3.b Anti-cancer effect:

In traditional medicine, *A. halimus* is used to treat a large number of diseases such as inflammation, cracked hands, and regulates hormones, heart diseases, diabetes, and rheumatism [103,104]. However, there is yet few works regarding the cytotoxic activities of *A. halimus* specifically against breast cancer cell lines and their antioxidant properties.

The crude extract and fractions of *A. halimus* growing in Egypt have cytotoxic activity against MCF-7 and PC3 carcinoma cells and human hepatocellular carcinoma [105]. Moreover, *Atriplex* have a cytotoxicity effect on human cervical cancer cells [106][83].

In addition to its skin-whitening properties, which have been known for at least seven decades, arbutin has been shown to possess numerous other therapeutically relevant biological properties, including antioxidant, antimicrobial, and anti-inflammatory activity. It also has the potential to act as an anticancer agent. Information from the published literature on arbutin shows that this compound possesses cytotoxic properties against several human cancers and tumor cell lines, including bladder, bone, brain, breast, cervical, colon, stomach, liver, prostate, and skin cancers. Most of these activities have been demonstrated in vitro, and in some cases, plausible mechanisms of action have been identified, such as apoptosis [107].

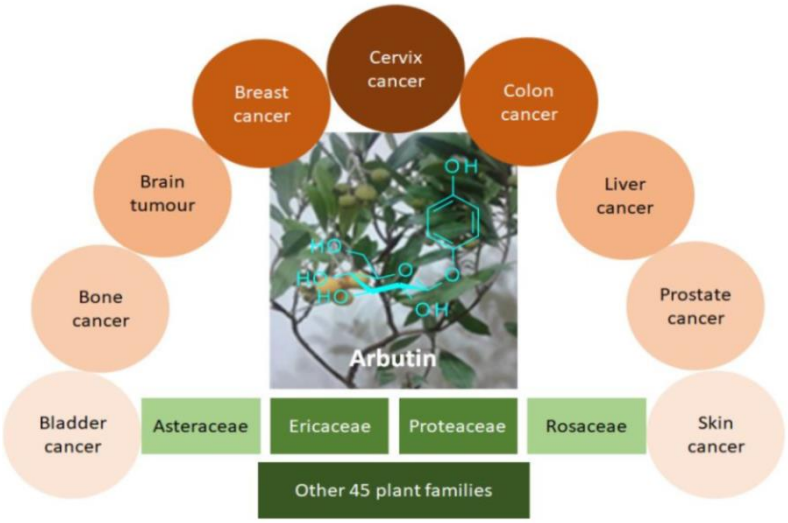


Figure I.17: A schematic summary of the anticancer potential of arbutin, obtained from different plant families [107].

I.7 Conclusion:

Breast cancer remains a major public health challenge due to its high incidence and complex pathophysiology. However, advances in molecular biology, early detection, and targeted therapies have significantly improved patient outcomes in recent decades. This approach supports the study of key anticancer molecular targets such as HER2, CDK4, AKT1, and the MCF-7 cell line, and the analysis of their interactions with natural ligands such as curcumin and diosgenin. It also reinforces the value of integrating traditional medical knowledge with modern computational tools.

Chapter II

Molecular Docking

An advanced study of molecular linkage techniques between biocompounds and protein targets

II.1 Introduction:

Breast cancer is one of the most prevalent diseases and poses a significant challenge to drug therapy. In this study, we explore the potential of natural compounds extracted from fenugreek (*Trigonella foenum-graecum*); turmeric (*Curcuma longa*) and *Atriplex Halimus* (Arbutin) as potential inhibitors of key breast cancer proteins (HER2, CDK4, AKT1, and MCF-7). Through molecular docking and advanced computational analyses.

II.2 Molecular Docking:

Molecular docking is one of the widely adapted methods to predict the binding affinities between the ligand and the target protein and further the lead optimization [108]. Additionally, the molecular docking imparts knowledge on the interactions at the atomic level and predicts the ideal binding mode [109,110]. Molecular docking mechanism generally evaluates the binding conformations, its orientation, and the accommodation of the small molecule at the active site of the proteins binding site and are read as scores [111][60].

Molecular docking is a computational method used to predict the three-dimensional structure of molecular compounds by simulating the interaction between a target molecule, often a protein, and a ligand, such as a potential drug. By identifying the most favorable position for the ligand in the target protein's active site—that is, the most stable position—molecular docking helps predict molecular interactions and improve their efficiency. Molecular docking has become a fundamental aspect of computational drug development in recent years, and its applications extend to many other fields, where optimizing the spatial arrangement of macromolecules can improve their efficacy: in personalized medicine, to predict a patient's response to specific drugs based on their genetic and molecular profile; in the food industry, to improve flavors and food additives [112].

II.3 Principle of Molecular Docking:

In many drug discovery initiatives, molecular docking has become crucial, especially for the virtual screening of phytochemicals or nutraceuticals as possible therapeutic compounds. Molecular Docking has become an essential aspect of in-silico drug development in recent

years. This technique involves predicting the interaction between a small molecule and a protein at the atomic level. This enables researchers to study the behavior of small molecules, such as nutraceuticals, within the binding site of a target protein and understand the fundamental biochemical process underlying this interaction [113]. The technique is structure-based and requires a high-resolution 3D representation of the target protein [114]. One of the most effective methods in structure-based drug design is docking that predicts the preferred orientation of a molecule and when bound to each other to form a stable complex and can indicate binding-conformity of small molecule ligands on suitable target sites. Since protein-ligand interaction plays a significant role in structurally based drug designing, H bonding and hydrophobic bonding are the main reasons for docking scores. It is considered a standard drug if the docking score is above 6.00 kcal/mol [115]. The main goal of molecular docking is to give a prediction of the ligand-receptor complex structure by using computation methods [116].

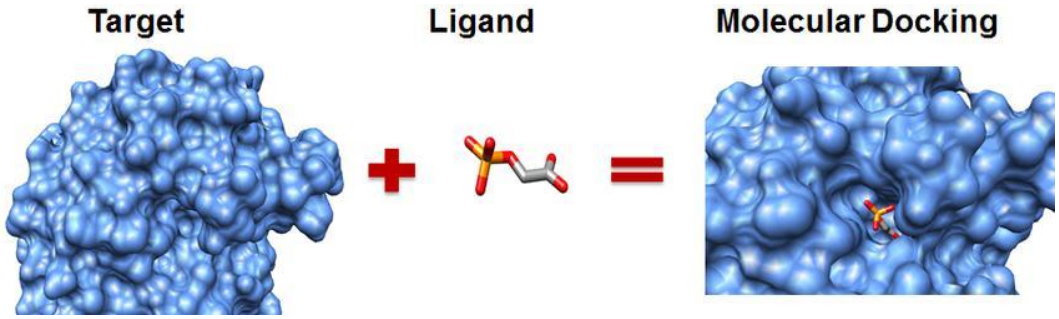


Figure II.18: Principe of Molecular Docking [117].

II.4 Optimization of Molecules:

Molecule optimization is a critical step in drug discovery to improve desired properties of drug candidates through chemical modification. Recent in silico approaches using deep learning have enabled alternative computationally generative processes to accelerate the conventional paradigm. These deep-learning methods learn from string-based molecule representations (SMILES) or molecular graphs and generate new ones accordingly (e.g., via connecting atoms and bonds) with better properties. While computationally attractive, these methods do not conform to the in vitro molecule optimization process in one very important aspect: molecule optimization needs to retain the major scaffold of a molecule, but generating

entire, new molecular structures may not reproduce the scaffold. Therefore, these methods are limited in their potentials to inform and direct in vitro molecule optimization [118]. Structural optimization was performed using DFT/B3LYP with a 6-31G (2d, p) basis set. The molecular geometry and vibrational assignments were analyzed using density functional theory (DFT), using HOMO-LUMO energies and the molecular electrostatic potential energy gap, the reactivity and molecular stability of the molecules were calculated [119]. The tumor inhibitory properties and pharmacokinetics of these molecules were evaluated using a computational approach using molecular docking and ADMET.

II.5 Molecular docking simulation:

Molecular docking techniques simulate the interaction between a molecule and its biological target, typically a protein involved in disease pathways. By predicting how small molecules fit into the binding sites of target proteins, molecular docking helps to ascertain the binding affinity and interaction mechanisms, crucial for assessing the therapeutic potential of new drugs. This approach is instrumental in refining the molecular structures of drug candidates to enhance their specificity and potency against cancer targets [116].

Molecular docking studies involve several key steps: Selection Ligand& protein, preparing the protein and ligand, identifying the binding site, performing the docking, scoring the docked poses, and validating the results. These steps ensure accurate prediction of ligand binding orientations and can be used for various applications

II.6 Target and Ligand Selection:

After a comprehensive bibliographic of the disease, medicinal plants, and major inhibitory compounds, we identified potential therapeutic targets and their corresponding ligands for computational analysis. This preparatory phase paved the way for molecular docking and further computational evaluation, combining evidence from the literature with structural bioinformatics tools.

II.6.1 Target selection:

The three-dimensional structure of the target protein was obtained from the Protein Data Bank (PDB) website at https://www.rcsb.org/ (see figure 19) ensuring selection based on accuracy, biological relevance, and availability of active site information. Four breast cancer-associated proteins were retrieved: HER-2, AKT1, CDK4, and MCF-7 (PDB IDs: 2IOK, 4EJN, 2W96 and 4XO6) , respectively.

- The PDB database was used to obtain detailed information about target protein

Figure II.19: Screenshot of PDB.

- **Plip:** the protein-ligand interaction profiler detects hydrogen bonds, hydrophobic contacts, and halogen bonds between ligands and targets in 3D [120]. PLIP is easy to use, requiring only a Protein Database Identifier (PDB) or PDB file as input (see figure 20).

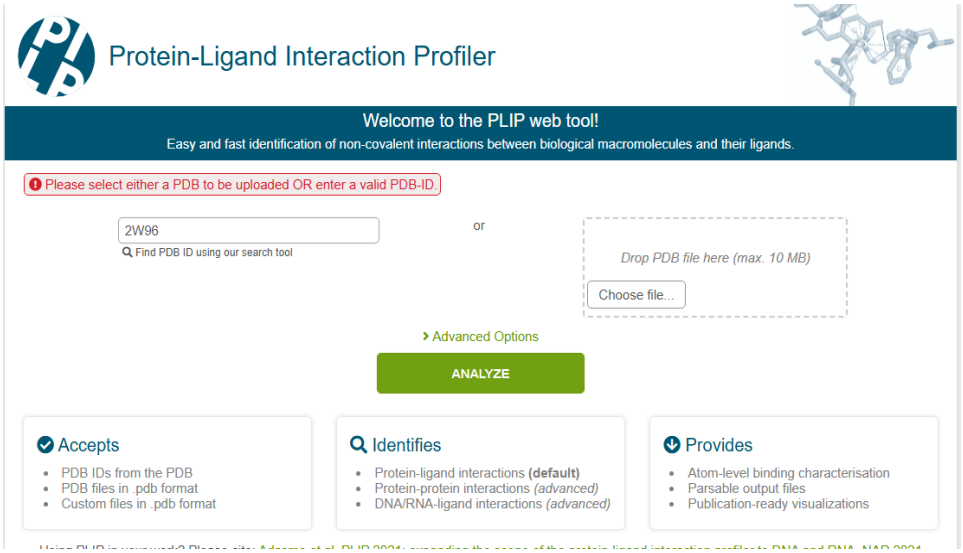


Figure II.20: Plip databases screenshot.

❖ **CDK4:**

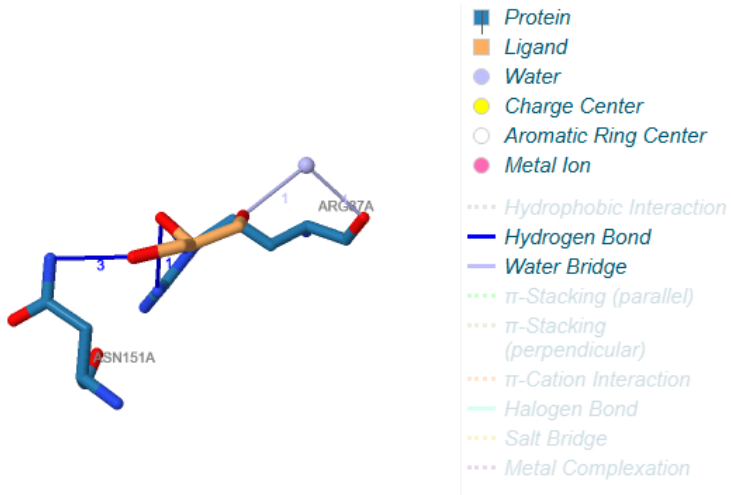


Figure II.21: Molecular interaction of the compound selected from databases CDK4 [120].

❖ **HER2:**

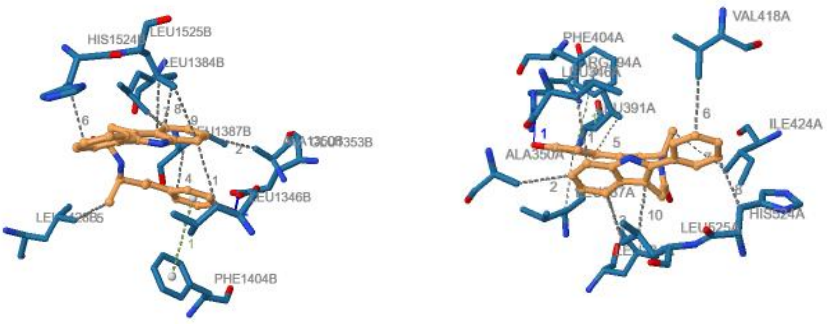


Figure II.22: Molecular interaction of the compound selected from databases HER2 [120].

❖ **MCF-7:**

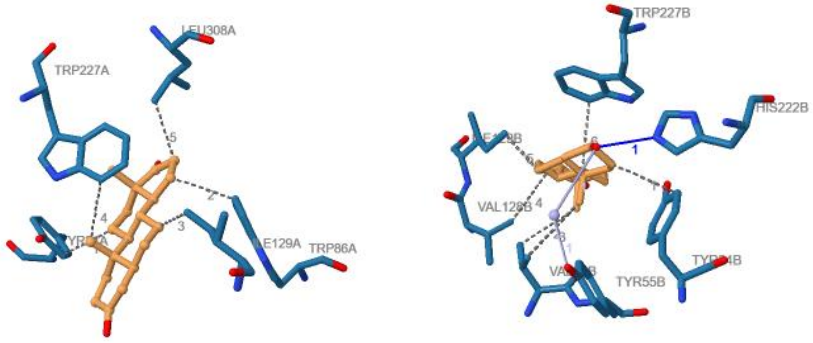


Figure II.23: Molecular interaction of the compound selected from databases MCF-7 [120].

❖ **AKT1:**

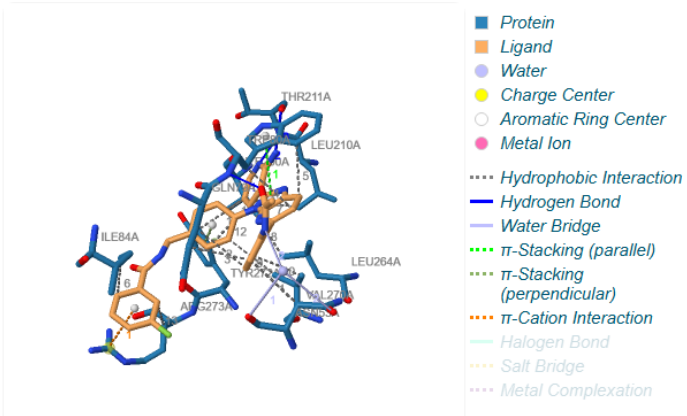
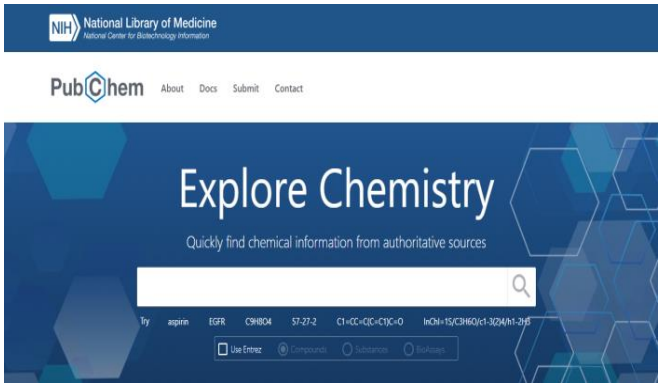


Figure II.24: Molecular interaction of the compound selected from databases AKT1 [120].

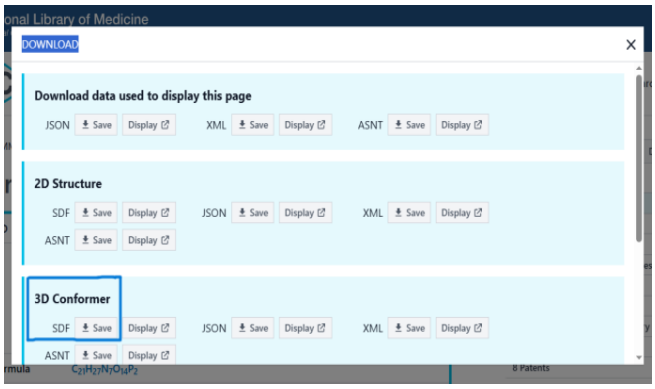
II.6.2 Ligand selection:

The chemical structures of the identified inhibitory molecules, curcumin, fenugreekine, diosgenin and arbutin were obtained from the PubChem database https://pubchem.ncbi.nlm.nih.gov/ in structure data file (SDF) format (see Figure 25). These compounds were selected based on extensive evidence from the scientific literature highlighting their important pharmacological activities. Their importance in the pathophysiology of breast cancer and other malignancies has been justified.

- The PubChem database was used to obtain detailed information about molecule format 3D SDF.



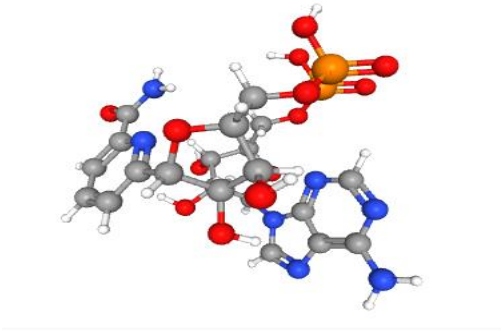
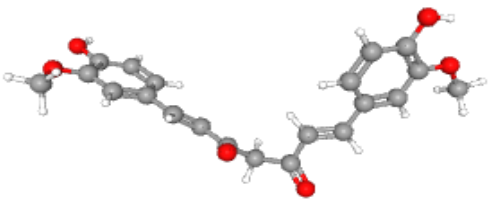
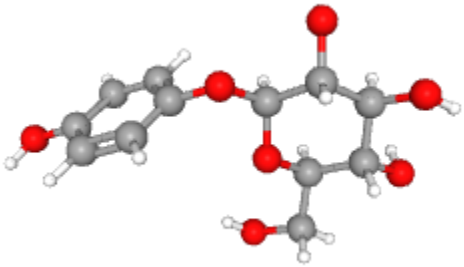
(1)



(2)

Figure II.25: (1) Screenshot of PubChem**, (2)** Download molecule forma 3D SDF.

Table II.2: 3D SDF format of Ligand.

Plants	Ligand	3D SDF Format
Fenugreek	Fenugreekine	
Fenugreek	Diosgenin	
Turmeric	Curcumin	
Atriplex Halimus	Arbutin	

II.7 Target and Ligand Preparation:

Molecular docking is a crucial technique in silico drug design. It predicts the geometries and affinities with which small-molecule ligands (such as drug candidates) might bind to a target (protein). Docking involves placing (docking) a small molecule within a binding pocket of the protein target and calculating (scoring) the interaction geometry to assess the binding probability. Researchers then prioritize compounds with the best docking scores for further analysis or experimental testing [121].

Many tools and computational algorithms are available for molecular docking techniques, both commercial and free. These programs and tools have been developed and are currently used in pharmaceutical research and academic fields, according to Sahoo et al [114]. In this research, we used some of the most popular and widely used molecular docking programs: AutoDock Vina Tools, Discovery and Gauss View 6.0.16 is used to perform DFT calculations for breast cancer drugs and Gaussian 09w for optimization.

II.7.1 Gauss View 6.0.16 and Gaussian 09w:

Gauss View 6.0.16 is a molecular visualization and computational chemistry software developed by Gaussian, Inc. as an integrated graphical interface for Gaussian quantum chemistry (see figure 26). It is an essential tool for computational chemistry researchers, enabling them to construct, visualize, and analyze molecular systems while facilitating quantum mechanical calculations. Its core functionality revolves around three main principles: molecular modeling, computational setup, and result interpretation.

Gaussian 09W is a Windows-compatible version of the Gaussian 09 software suite a leading computational chemistry program used for electronic structure modeling and quantum chemical calculations (Frisch, M. J. et al. (2009) (see figure 26). It enables researchers to predict molecular properties, reaction mechanisms, and spectroscopic data using density functional theory (DFT), and semi-empirical methods. The basic principle of Gaussian 09W is to solve the Schrödinger equation for atoms and molecules using basis sets and electron correlation methods. Users enter molecular structures and specify calculation types via formatted input files or through graphical interfaces such as GaussView. Gaussian 09W then

applies computational algorithms to simulate chemical behavior and predict molecular stability, reactivity, and electronic properties [122].

The theoretical foundation of GaussView 6.0.16 is based on the principles of quantum mechanics, particularly the numerical solution of the Schrödinger equation using the Gaussian computation engine, employing approximations such as the Born–Oppenheimer approximation, as well as various electronic structure methods, including Hartree–Fock theory and density functional theory (Gaussian, Inc. (2021). The program also incorporates molecular mechanics for initial structure refinements. For visualization and analysis, GaussView provides advanced capabilities for displaying and interpreting computational results, including molecular orbitals, electrostatic potential maps, and vertical vibrational modes. These features are supported by robust reference materials, including the Official GaussView 6.0 User Manual (Gaussian, 2021) and the Gaussian 16 Reference Manual (Frisch and al., 2016), which document the program's implementation of these quantum chemical methods. Additional theoretical context can be found in the computational chemistry literature, such as Young's book "Computational Chemistry: A Practical Guide to Applying Techniques to Real-World Problems" (2004), which discusses the broader scientific principles underlying the program's functionality [123] [124].

(1) **(2)**

Figure II.26: (1) Gaussian view and (2) Gaussian Screenshot.

II.7.2 Density functional theory (DFT):

Density functional theory (DFT) is a widely used quantum mechanical modeling method for studying the electronic structure of molecules and condensed matter systems. Calculations are performed using standard functions and basis sets, providing valuable

insights into the structure-activity relationships of the plant-derived molecules under study [125][126].

- **Energy of bonding Orbital and Antibonding Orbital (HOMO and LUMO):**

The most fundamental distinction respectively HOMO and LUMO is that HOMO donates electrons, where LUMO receives electrons. Molecular interaction (HOMO/LUMO) predicts which molecules should have unusually high HOMOs and abnormally low LUMO's ultimately recognize the functional groups and indicate which functional groups are reacted or interacted with each other. [127]

We examined the electronic properties of these molecules, including the electronic band gap (Egap), the highest occupied molecular orbitals, and the lowest unoccupied orbitals (EHOMO and ELUMO, respectively). Based on these relationships, we were able to determine the electronic band gap. [128]

$$\Delta E_{gap} = E_{HOMO} - E_{LUMO} \ldots \text{II (1)}$$

The study also incorporates the computation of theoretical physicochemical characteristics such as energy gap (ΔE), chemical softness (S), chemical hardness (η), electronegativity (χ), electrophilicity index (ω), and chemical potential (μ). Quantum chemical descriptors suggested by Parr and colleagues are widely used by computational chemists for predicting a molecule's chemical behaviour. These descriptors can be calculated using equations: [119]

$$\mu = \frac{1}{2}(EHOMO + ELUMO) \ldots \text{II (2)}$$

$$\eta = \frac{1}{2}(-EHOMO + ELUMO) \ldots \text{II(3)}$$

$$S = \frac{1}{2\eta} \ldots \text{II (4)}$$

$$X = \frac{-(EHOMO + ELUMO)}{2} \ldots \text{II (5)}$$

$$\omega = \frac{\mu^2}{2\eta} \ldots \text{II (6)}$$

- **Bioactivity of Selected Compounds :**

The inhibition constant (Ki) determines the affinity of an inhibitor to its target enzyme or receptor. It is the concentration required to occupy 50% of the binding sites at equilibrium. This constant is a fundamental criterion in drug design, as it reflects the effectiveness of a compound as an inhibitor. Equation 2 illustrates the inverse relationship between binding energy and the inhibition constant; Lower Ki values indicate a stronger association [129].

$$Ki = e^{\left(-\frac{\Delta G}{RT}\right)} \ldots (7)\ \mathbf{II}$$

Ki= Inhibition constant.

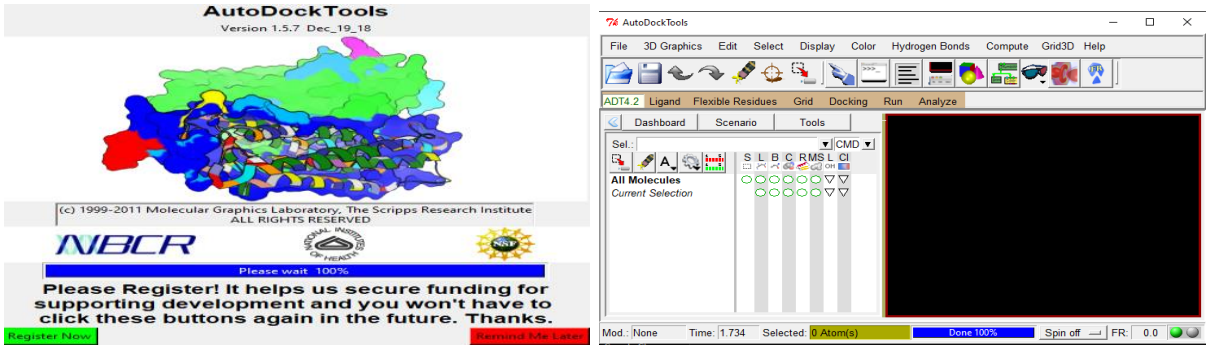
R= Gaz constant (1.987 *10^-3 kcal/k mol).

T= 298.15k (Absolute Temperature).

II.7.3 AutoDock Vina Tools:

AutoDock is a molecular modeling simulation program, particularly effective for protein-ligand binding. AutoDock is one of the most popular protein-binding applications in the research community. AutoDock consists of two programs: AutoGrid and AutoDock. AutoGrid is primarily responsible for calculating the energy required in the grid, while AutoDock is responsible for searching and evaluating configurations. AutoDock includes, but is not limited to, the following applications: X-ray crystallography, structure-based drug design, lead compound optimization, virtual screening, combinatorial library design, protein binding, and chemical mechanism studies. [130].

AutoDock Vina is a turnkey computational docking program based on a simple scoring function and fast gradient-optimized conformational search. It can rapidly and efficiently dock drug-like ligands with protein targets. The molecular docking in silico experiment was conducted by AutoDock Tools 1.5.7 programe (see figure 27). The protein and ligand data were converted into three-dimensional PDBQT files. The AutoGrid tool from the AutoDock software suite was used to create grid maps showing the energy of contact between the target macromolecule and the ligands during the docking experiment [131].

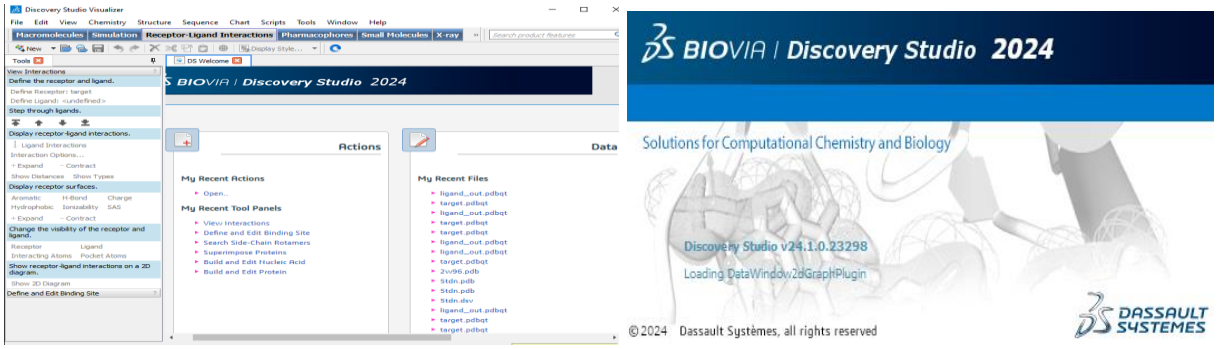


(1) **(2)**

Figure 27: (1) and (2) screenshot of AutoDock Tools.

II.7.4 BIOVIA Discovery Studio:

BIOVIA Discovery Studio is a comprehensive molecular modeling and simulation platform developed by Dassault Systèmes to facilitate drug discovery (see figure 28), analysis of protein-ligand interactions, and structure-based computational research [132]. The software provides an integrated suite of advanced tools for molecular docking, pharmacophores modeling, virtual screening, and molecular dynamics simulations. A key feature is its user-friendly graphical interface, which allows researchers to visualize complex biomolecular structures and interactions in three dimensions, enabling a more intuitive interpretation of docking results and molecular behavior [133].



(1) (2)

Figure 28: (1) and (2) discovery screenshot.

II.8 Pharmacokinetic and ADMET study:

Pharmacokinetic and ADMET (absorption, distribution, metabolism, excretion, and toxicity) studies are essential components of drug discovery and development, providing valuable insights into the biological behavior and safety profiles of candidate compounds. ADMET computational prediction tools offer a rapid and cost-effective alternative to early experimental screening by estimating a molecule's pharmacokinetic properties and potential toxicity based on its chemical structure. These parameters help determine whether a compound has pharmacological properties suitable for further development [134].

We evaluated the physicochemical properties, drug similarity, and pharmacokinetic characteristics of the bioactive compounds using the electronic servers: ADMET, SwissADME and ProTox.

II.8.1 Absorption, Distribution, Metabolism, Excretion, and Toxicity ADMET:

The evaluation of pharmacokinetics and toxicity is crucial for the design of new therapeutic candidates. In silico virtual screens and generative AI output a vast number of molecules that must be filtered to a tractable number for synthesis and experimental validation. An effective primary filter is to evaluate candidate compounds based on their **A**bsorption, **D**istribution, **M**etabolism, **E**xcretion, and **T**oxicity (**ADMET**) properties. https://admet.ai.greenstonebio.com/

1. **Absorption**
    a. **HIA (Human Intestinal Absorption):** Estimates oral bioavailability.

2. **Distribution**
    b. **Blood-Brain Barrier (BBB) Penetration:** Determines CNS drug potential.
    c. **Volume of Distribution (Vd):** Predicts tissue vs. plasma concentration.

3. **Metabolism**
    d. **Cytochrome P450 (CYP) Inhibition/Induction:** Identifies drug-drug interaction risks.

4. **Excretion**
    e. **Renal Clearance:** Estimates elimination pathways.

5. **Toxicity (T)**
    f. **Hepatotoxicity:** Assesses liver damage potential.

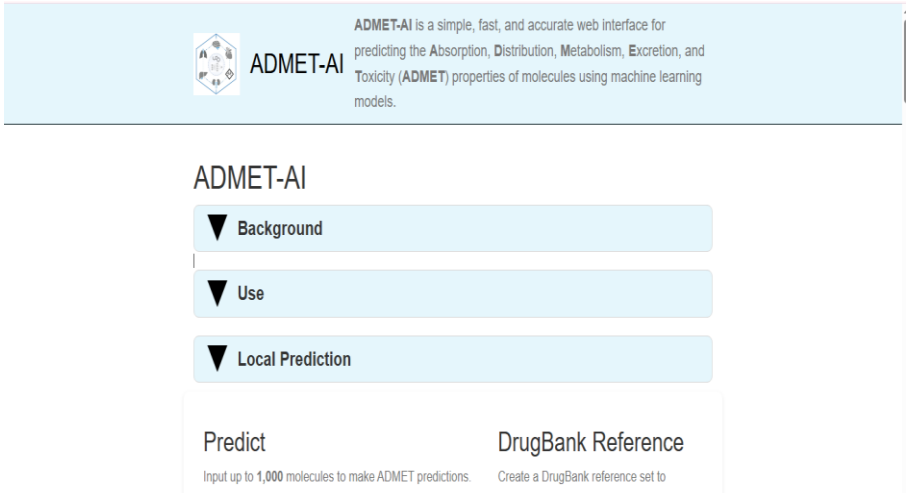


Figure 29: ADMET screenshot.

II.8.2 Swiss ADME:

SwissADME gives free access to a number of parameters and predictive models in order to compute the physicochemistry and estimate the pharmacokinetics, druglikeness and medicinal chemistry friendliness of one or several small molecules (see figure 30) [135]. https://www.expasy.org/resources/swissadme

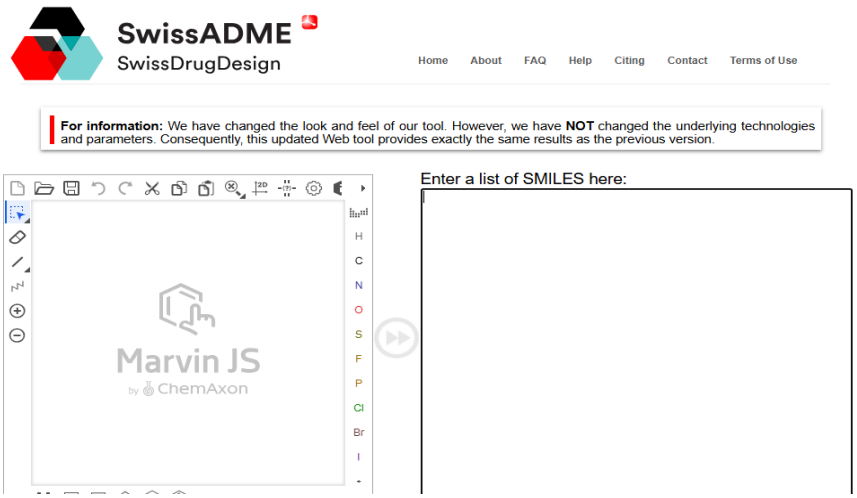


Figure 30: SwissADME screenshot.

II.8.3 ProTox 3.0:

The toxicity model presented relies on comparing the structural characteristics of the compounds under study with known toxic effects to forecast toxic fragments. ProTox 3.0 (see figure 31) integrates molecular similarity and machine learning techniques to predict 45 toxicity endpoints, including Acute Toxicity [136]. https://tox.charite.de/protox3/

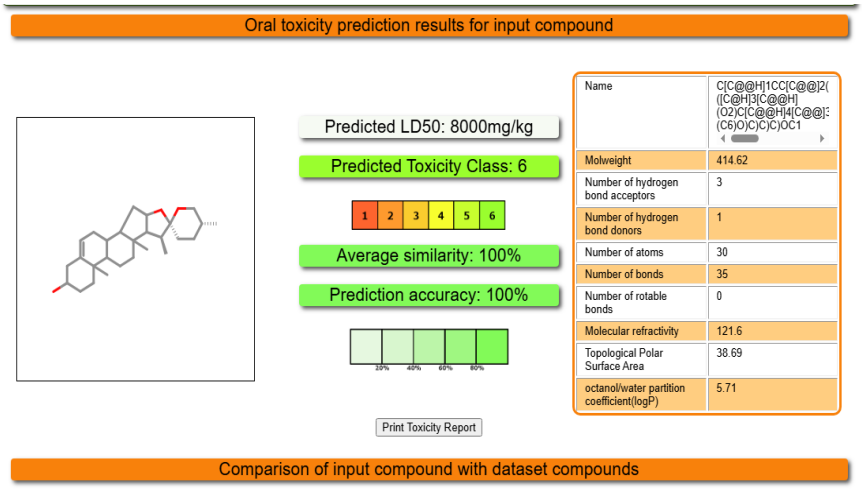


Figure 31: ProTox databases screenshot.

II.9 Conclusion:

This chapter demonstrates that computational drug discovery can efficiently identify and optimize plant-derived anticancer agents, accelerating the development of safer and more effective therapies. Further validation through biological assays is recommended to confirm these predictions and advance potential lead compounds into preclinical development.

Chapter III

Quantitative Structure Activity Relationship (QSAR)

III.1 Introduction:

The process of drug discovery is an extremely costly and time-consuming endeavor aimed at ensuring the safety and quality of new chemical entities entering the market. It is estimated that developing a novel molecule can take up to 14 years from target identification to regulatory approval [137]. Among the main approaches, quantitative structure-activity relationship (QSAR) modeling stands out as a useful tool for predicting the biological activity or properties of a compound by providing a mathematical link to its structural properties [138]. QSAR models have been widely used in the fields of drug design and environmental toxicology and have become fundamental in the molecular interpretation of biological properties. The biological activity of a compound can be described by spatial, hydrophobic, electronic, spatial, and quantum chemical parameters encoded in a set of descriptors. The molecular descriptors required for QSAR can be obtained experimentally (i.e., derived experimentally) or calculated by relevant software packages and platforms, such as the SwissADME web tool [139] based on the Simplified Molecular Input Line Entry System (SMILES) [140-142][143].

A growing number of studies have been conducted to develop anticancer drugs, and some compounds have shown great potential in combating this disease. Our work aims to propose new drugs and design anticancer drugs using computational methods. To show a relationship between a distinct structure and inhibitory activity against breast cancer, we used three-dimensional quantitative structure-activity relationship (2D- and 3D-QSAR). Computational chemistry techniques are used in pharmaceutical research to speed up the drug design and discovery process, reducing the time and cost of discovery research.

III.2 Definition:

Quantitative structure-activity relationship (QSAR) dates back to the 19th century and is a computational tool that seeks to link changes in the structural or molecular properties of compounds to their biological activities. These physicochemical descriptions, which include parameters for hydrophobicity, topology, electronic properties, and steric effects, are determined experimentally or, more recently, computationally. The idea is that the structure of a chemical determines the physical properties and interactions that underlie its biological and

toxicological properties. The ability to predict potential adverse effects not only aids the tailored development of new chemicals but also reduces the need for animal testing. This could ultimately, or potentially, improve health and environmental protection through the strategic application of limited testing resources and available information assets to help screen or identify the most hazardous chemicals. QSAR models are currently applied in many disciplines related to drug design and development. Recently, interest has grown in using this tool in environmental risk assessment to group compounds with similar biological activity.

Biological variability can be demonstrated quite readily by building a QSAR model, which discriminates between chemicals with different toxicological hazard classifications. Application of computer-based QSARs has resulted in developing novel predictive capabilities for representing chemical structures as a distribution of conformations and properties rather than discrete structures. [144]

Quantitative structure-activity relationship (QSAR) approach relies on the basic principle of chemistry that states that the biological activity of any ligand or compound is associated with the arrangement of atoms forming the molecular structure. In other words, structurally related molecules possess similar biological activities. This structural information can be defined in terms of a series of parameters called molecular descriptors. The model thus developed based on the biological activities of known ligands is used to predict the response of new compounds. [145] In QSAR, the biological activity is represented as a function of these molecular descriptors as depicted in Eq.

Biologicalresponseoractivity= f (moleculardescriptors) … III (1)

III.3 Principal of QSAR:

Quantitative structure-activity relationship (QSAR) is a computational method for determining the relationship between chemical structures and biological activity. The foundation of QSAR investigations is the idea that changes in bioactivity are connected to structural and molecular changes in a group of chemicals. To find chemical compounds that potentially have potent inhibitory effects on particular targets while being low toxicity, drug discovery frequently uses QSAR (non-specific activity). In the 1980s, 3D structural information was added into the QSAR approach, known as 3D-QSAR, with the further development of structure-activity connection theory and statistical methodologies. Since the

1990s, as computing power has increased and the precise 3D structures of many biomacromolecules have been determined, structure-based drug design has gradually supplanted quantitative structure-activity relationships (QSAR), despite QSAR's advantages of low computational burden and strong predictive power [146].

By using mathematical models, this analysis enables the prediction of the biological activity of chemical compounds based on their structural, physical, and chemical properties. This method plays a pivotal role in discovering new drug candidates and improving the properties of key compounds, which contributes to the development of potential treatments [137]. To create a trustworthy QSAR model, several specifications are necessary, namely (a) the bioactivity data for at least 20 compounds acquired from a common experimental protocol such that the potency values are comparable; (b) the training and test sets should contain the appropriate selection of compounds; (c) to prevent overfitting, molecular descriptors for the ligands shouldn’t be autocorrelated. Furthermore, to verify the model's applicability and predictability, validation should be conducted using internal or external validation [146].

III.4 Applications of QSAR Approaches:

In drug design, QSAR models are extensively used to screen large libraries of compounds, prioritize lead candidates, and optimize molecular structures to enhance potency, selectivity, and safety (Cherkasov and al., 2014). In toxicology, QSAR models support the prediction of mutagenicity, carcinogenicity, and other toxicity endpoints, reducing reliance on in vivo testing and aligning with ethical principles of the 3Rs (Replacement, Reduction, and Refinement).

The integration of machine learning and AI techniques into QSAR modeling has further expanded its application scope, improving predictive power and model robustness (Roy and al., 2015). As a result, QSAR approaches continue to be indispensable in Pharmacological and Toxicological domain [147][148].

III.4.1 Pharmacological QSAR: Focused on predicting drug activity, receptor binding, or ADMET properties. QSAR models help predict the biological activity of new drug candidates before synthesis. They reduce time and cost by identifying promising compounds early in the manufacturing phase.

III.4.2 Toxicological QSAR: Focused on environmental and human toxicity assessments. QSAR is used to evaluate the environmental fate and impact of chemicals, including bioaccumulation, biodegradation, and aquatic toxicity. Is widely used to predict acute, chronic, and environmental toxicity of chemicals.

III.5 Molecular Descriptors:

Modern QSAR approaches are characterized by the use of multiple chemical structure descriptors, combined with the application of both linear and nonlinear optimization approaches, and a strong emphasis on rigorous model validation to provide robust and predictive QSAR models. The most significant recent developments in this field coincide with the significant increase in the size of experimental datasets available for analysis and the increasing application of QSAR models as virtual screening tools for the discovery of bioactive molecules in chemical databases and/or virtual chemical libraries. In the context of QSAR (quantitative structure-activity relationship) modeling, a descriptor, or more specifically a molecular descriptor, is a numerical value or Boolean attribute that encodes the chemical structure and/or physicochemical properties of a molecule. These descriptors are used as independent variables in QSAR models, which aim to relate them to the biological activity (the dependent variable) of a set of compounds [149].

III.5.1 Classification of descriptors:

The molecular descriptors can be classified in several ways. There are binary, discreet, and continuous values and we can also calculate other new descriptors from previously defined ones. Usually they are classified based on their dimensionality as 0, 1, 2, 3 … nD descriptors.

- **0D descriptors:** are molecular descriptors that do not require information about molecular structure or atomic bonds. This class is characterized by the number of atoms and bonds, as well as the sum or average of the atomic properties. These descriptors can be easily calculated, interpreted naturally, do not require molecular structure optimization, and are independent of any conformational problem. They typically exhibit very high degeneracy, meaning their values are equal for several molecules, such as isomers. Despite their low information content, they can play an important role in modeling many physical and chemical properties, or contribute to more complex models.[150]
- **1D descriptors:** are all molecular descriptors that can be calculated from substructural information of the molecule. The most common one-dimensional descriptors include functional group counts and substructure fragments, as well as atom-centered descriptors. These descriptors are often presented in the form of a fingerprint, a binary vector where 1 indicates the presence of a specific substructure and 0 indicates its absence. An important advantage of describing molecules using fingerprints is the ability to perform rapid calculations for molecular similarity/diversity problems.[150]
- **2D descriptors:** are molecular descriptors based on a graph representation of the molecule and represent graph–theoretical properties that are preserved by isomorphism, that is, properties with identical values for isomorphic graphs. A graph invariant may be a characteristic polynomial, a sequence of numbers, or a single numerical index obtained by the application of algebraic operators to matrices representing molecular graphs and whose values are independent of vertex numbering or labelling.[150]
- **3D descriptors:** are calculated from the 3D representation of the molecule. Thus, 3D chemical structures (conformations) are needed, which in best case can be bioactive conformations of the molecules in the field of drug design. 3D descriptors are as useful and better known as their 2D siblings, but they were typically developed later than 2D descriptors, and their calculation is more time-consuming and computationally demanding. 3D descriptors (like 2D ones) are frequently applied for QSAR and similarity searching.[150]

- **4D descriptors:** are also grid-based descriptors (but they are not derived from molecular interaction fields.) The fourth dimension represents ensemble sampling or conformational flexibility, which is defined by the conformational ensemble profile (CEP). The CEP is calculated with molecular dynamics (MD) simulations. 4D descriptors help identify the active conformations of the flexible molecules, and they are also useful tools in alignment problems. The pioneering work of Hopfinger et al. in 1997 has introduced 4D descriptors to QSAR analysis. Hence, 4D descriptors are among the youngest members of the family of (now thousands of) molecular descriptors. [150]

Table III.3: Types of Descriptors [151].

Type	Description	Examples
0D descriptors	Basic properties independent of structure	Molecular Weight, number of atoms
1D descriptors	Counts of atom types or functional groups	Number of H-bond donors, acceptors
2D descriptors	Topological or graph-based properties	Molecular connectivity index, logP, TPSA
3D descriptors	3D shape and conformation-based properties	Molecular surface area, volume, energy
4D/5D descriptors	Derived from multiple conformers, time and interaction dimensions	Advanced modeling (molecular dynamic)

III.5.2 Feature Selection of Descriptors in QSAR Modeling:

In QSAR modeling, feature selection refers to the process of identifying the most relevant molecular descriptors that contribute significantly to predicting biological activity or physicochemical properties. Because QSAR models often start with a large set of calculated descriptors—many of which may be redundant, irrelevant, or coincident—feature selection is a critical step for improving model accuracy, interpretability, and computational efficiency. Effective feature selection helps prevent overfitting and enhances model generalizability to unseen data [152].

Table III.4: Common Descriptors in Breast Cancer QSAR [153,147].

Descriptors	Description	Relevance in Breast Cancer
LogP	Octanol–water partition coefficient (lipophilicity)	Affects cell membrane permeability & bioavailability
TPSA	Topological Polar Surface Area	Related to H-bonding; influences oral bioavailability
NRotB	Number of rotatable bonds	Reflects molecular flexibility; impacts binding
HBD / HBA	Hydrogen bond donors/acceptors	Crucial for receptor interaction (e.g., ER binding)
Molecular Weight	Total mass of the molecule	Filters drug-likeness (Lipinski's rule)
pKa	Acid dissociation constant	Important for ionization and membrane passage
HOMO / LUMO Energies	Orbital energy levels	Determine electron donation/acceptance ability
Chi1	Connectivity Index, First Order	It measures the contact between atoms.

III.6 Multiple Linear Regression (MLR):

Multiple Linear Regression (MLR) is a fundamental and widely used statistical method in QSAR modeling to relate a set of independent variables (molecular descriptors) to a dependent variable (biological activity, e.g., pIC_{50}). MLR is based on the assumption that the biological activity of a compound can be linearly explained by a combination of molecular descriptors. The aim is to find the best-fitting linear equation that minimizes the difference between predicted and observed activities. Is a statistical technique used to model the linear relationship between a single dependent variable and two or more independent variables. It's an extension of simple linear regression, which only considers one independent variable. In machine learning, MLR is a fundamental algorithm for predictive analysis, allowing us to understand how several factors influence an outcome [146].

The general formula for a multiple linear regression model in QSAR is:

$$\mathbf{Y= a+b1X1+b2X2+\cdots+bnXn+\epsilon \ldots III\ (2)}$$

Where:

- Y: Biological activity (e.g., pIC_{50}, log IC_{50}, Ki)
- a: Intercept (the value of activity when all Xi=0)
- bi: Regression coefficient for the descriptor
- Xi: Molecular descriptor (e.g., LogP, molecular weight, number of H-bond donors)
- ϵ (epsilon): Random error (residual)

III.7 Support Vector Regression (SVR):

Support vector regression (SVR) is a supervised machine learning technique derived from the principles of support vector machines (SVMs) and used to predict continuous outcomes. Unlike traditional regression models that minimize the sum of squared errors, SVR seeks to find a function that approximates target values within a specified error margin, denoted ε, while maintaining model simplicity by minimizing the parameter criterion (i.e.,

flattening the model). Only data points that fall outside this ε-insensitive region—called support vectors—affect the final model. SVR is particularly useful for quantitative structure-activity relationship (QSAR) modeling, where molecular activity often exhibits complex and nonlinear patterns. The ability to incorporate kernel functions enables SVR to effectively capture these nonlinear relationships. The flexibility of the model is controlled by parameters such as ε (tolerance margin), C (regularization coefficient), and the choice of kernel, making SVR a powerful tool in chemistry and bioinformatics for activity prediction and drug discovery [154][154].

- The general formula for a nonlinear regression model in QSAR is:

$$\boldsymbol{f(x) = \sum_{i}^{N}(\alpha i - \alpha i *)K(xi, x) + b \ldots \text{ III (3)}}$$

Where:

- x: Input vector (e.g, molecular descriptors)
- xi: Support vectors from the training data
- αi, αi∗: Lagrange multipliers obtained from solving the SVR optimization problem
- K(xi,x): Kernel function, which maps input data into a higher-dimensional space
- b: Bias term (intercept)
- N: Number of training data

III.8 Validation Methods for QSAR Models:

Validation methods are essential for determining the predictability of a model based on unseen data and for helping to determine the complexity of the equation justified by the amount of data. Using the data that generated the model (an internal method) or a separate dataset (an external method) can help validate a QSAR model. Least squares (R^2), cross-validation (Q^2), and root mean square error (RMSE) are internal methods for model validation. The best method for model validation is external, such as evaluating a QSAR model on a test set of vehicles. These statistical methodologies are used to ensure that the generated models are sound and unbiased (a "good model"). A poor model can cause more harm than good, so confirming a model as a "good model" is crucial [156].

III.8.1 Internal Validation: (Training set)

The most common method for internal model validation is least squares fitting. This validation method is similar to linear regression and uses R^2 (the squared correlation coefficient) to compare predicted and experimental activities. A difference between the R^2 value and the R^2adj value of less than 0.3 indicates that the number of descriptors participating in the QSAR model is acceptable. The fit of QSAR models can be determined using root mean square error (RMSE) methods, which are used to determine whether the model has the predictive quality reflected in R2. The use of RMSE describes the error between the mean of the experimental values and the predicted activities. For a good predictive model, RMSE values should be low (<0.3). Another common method for internal validation of a QSAR model is cross-validation (CV, Q^2 or jack-knifing) [156].

III.8.2 External Validation: (test set)

External validation testing requires only a single split of the entire dataset into structurally similar training and external validation sets. The purpose of this validation is to test the predictive ability of the model. The primary statistical criteria used to evaluate the performance of external validation are the root mean square error of prediction (RMSE test), the external validation correlation coefficient (Q^2ext), and the Pearson correlation coefficient of prediction (R^2ext). When performing external validation, two things must be considered. The first is the number of samples in the external validation set, and the second is the selection procedure. It is recommended to use 30% of the samples for external validation on smaller datasets, and to maintain the same percentage of external samples in the reassignment and external validation phases [157].

The equations for these different internal and external validation parameters are given below [158]:

- **R^2 (Coefficient of Determination):** measures how well the regression model explains the variance in the observed biological activity.

$$R^2 = 1 - \frac{\sum(obs - pred)^2}{\sum(obs - \bar{y})^2} \dots \text{III (4)}$$

- **Adjusted R^2 (R^2 adjusted):** adjusted R^2 corrects R^2 for the number of descriptors used, preventing overfitting.

$$R^2adj = 1 - (\frac{(1-R^2)(n-1)}{n-p-1})\ldots \text{III (5)}$$

- **Q^2 (Cross-validated R^2):** measures the model's predictive power, often computed via Leave-One-Out (LOO) or k-fold cross-validation.

$$Q^2 = 1 - \frac{\sum(yobs-ypred,cv)^2}{\sum(yobs-\bar{y})^2}\ldots \text{III (6)}$$

- **RMSE (Root Mean Squared Error):** measures the average error between predicted and observed values.

$$RMSE = \sqrt{\frac{1}{n}\sum(yobs - ypred)^2}\ldots \text{III (7)}$$

- **MAE (Mean Absolute Error):** measures the average magnitude of the errors between predicted and actual values, without considering their direction.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|yobs,i - ypred,i|\ldots \text{III (8)}$$

Table III.5: MLR& SVR model's parameters [159].

Metric	Meaning	Ideal value	Usage
R^2	Coefcient of determination (Goodness of fit)	$\rightarrow 1$	**Model explains variance**
Q^2	Predictive power (CV)	> 0.5	**Internal validation**
RMSE	Root mean square error (Prediction error)	$\downarrow$ (low)	**Model accuracy (sensitive to error size)**

III.9: Key Difference between MLR and SVR in QSAR Modeling:

In this phase of the research, we expand our analytical framework by applying linear and nonlinear modeling techniques using modern, advanced computational codes. The linear approach, primarily based on multiple linear regressions (MLR), provides interpretable insights into the contribution of each molecular descriptor to biological activity. In parallel, a nonlinear approach, such as support vector regression (SVR), is used to capture more complex and potentially nonlinear interactions that traditional linear models may overlook. By applying this dual modeling strategy, we can systematically compare the performance metrics (such as R^2, RMSE, and MAE) of both methods, identify limitations, and highlight the strengths of each algorithm in modeling molecular activity associated with breast cancer.

Table III.6: Comparison between MLR& SVR model [147,160,161].

Feature	**Multiple Linear Regression (MLR)**	**Support Vector Regression (SVR)**
Model Type	Linear	Non-linear (with kernel trick)
Assumptions	Linear relationships, normal residuals	No distributional assumptions
Feature Selection	Manual/statistical	Automatic via kernel transformations
Overfitting Risk	High with many descriptors	Lower (regularization parameter C controls complexity)
Interpretability	High (clear coefficients)	Low (black-box with kernels)
Computational Cost	Low	High (especially with RBF kernel)
Best for	Small datasets, linear relationships	Complex patterns, high-dimensional data
Hyperparameters	None	C, ε, kernel type (RBF/linear/poly)

III.10 Conclusion:

This QSAR-based modeling approach simplifies early-stage drug development by allowing virtual screening of candidate molecules and predicting their activity prior to synthesis. These computational strategies significantly reduce the cost, time, and reliance on animal experiments and aid in the rational design of targeted anticancer therapies. As breast cancer remains a leading cause of cancer-related deaths globally, the continued development and application of QSAR models provide valuable support in the search for effective and safer treatments.

Chapter IV

Experimental Studies

IV.1 Introduction:

Breast cancer remains a major global health challenge, underscoring the need for innovative therapeutic approaches. This study investigates the potential of bioactive compounds from fenugreek; turmeric and arbutin from *Atriplex Halimus* as natural inhibitors against key breast cancer targets (HER2, CDK4, AKT1, and MCF-7) through a combined experimental and computational approach. Ethanol extraction will be performed, followed by physicochemical characterization and antioxidant activity assessment using a DPPH assay. Furthermore, computational techniques such as molecular docking, DFT optimization, and QSAR modeling will be used to study the interactions of plant-derived compounds with key molecular targets associated with breast cancer. This multidisciplinary approach combines traditional photochemistry with modern computational drug design, providing a rational strategy for identifying natural compounds with anti-breast cancer properties while reducing reliance on costly experimental screening.

IV.2 Experimental Study:

In this study, we will extract the active bioactive compounds from using ethanol as the main solvent in the laboratory of Hassiba Ben Bouali University in Chlef, and then conduct advanced physicochemical analyses in the laboratory of Dar El Beida Saidal in Algiers. The main objective is to isolate and characterize compounds that may have anti-breast cancer effects, based on previous studies indicating that fenugreek contains phenolic compounds, flavonoids, and saponins with antitumor properties.

IV.2.1 Materials and methods:

IV.2.1.a Products:

Table 7: Extraction and DPPH Products.

Products	**Ethanol (C_2H_6O)**	**Distilled Water (H_2O)**	**DPPH ($C_{18}H_{12}N_5O_6$)**
Properties	M= 46g/mol	M= 18g/mol	M= 394.3g/mol

IV.2.1.b Materials:

- **Equipment used:** UV-visible spectrophotometer.
  Infrared spectrophotometer.
  Filter paper.
  Beaker.
  Rota-vapor.
  Funnel.
  Balance.
- **Plant Material:** The plant chose for this study is Dried fenugreek seeds (finely ground) tree in April 2025 from Oued Sly Chlef.

IV.2.2 Extraction method: At Hassiba Ben Bouali University

Extraction is a fundamental process used to isolate bioactive compounds from plant materials like fenugreek by employing selective solvents and techniques. The most common method, maceration, involves soaking powdered plant material in a solvent (typically 60-70% ethanol for fenugreek) for an extended period (24-72 hours) at room temperature, allowing compounds to diffuse into the solvent [162].

a) preparation of the plant:

Fenugreek seeds are dried and ground to enhance the contact surface with the solvent.

Weigh a quantity of 10g to start

Figure 32: Drying, grinding and weighing the fenugreek.

b) Maceration:

Maceration is a cold extraction technique where plant material is soaked in solvent, allowing gradual diffusion of bioactive compounds through cell walls via osmosis and solubilization.

Maceration is used using ethanol (40-60%) for **48–72 hours** with occasional stirring. We will mix 10g of plant with 100ml ethanol solvent (60% ethanol+ 40% Water distilled), stirring for 2 hours and we less at room temperature.

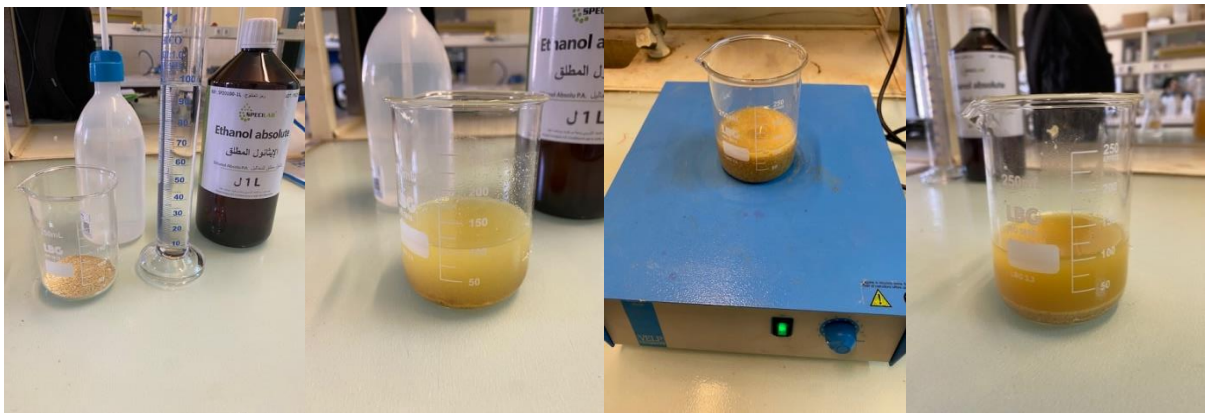


Figure 33: Maceration method

c) Filtration:

To clarify fenugreek extracts by removing particulate matter while maximizing retention of bioactive compounds. After 24h the solution was Separate solids from the liquid extract.

Figure 34: filtration of extract.

d) Solvant evaporation:

In this step we will place the liquid in a rotary evaporator in a water bath over very low heat at 40 °C 100 mbar (no direct flame), let the solvent evaporate slowly and when the liquid becomes more viscous, turn off the heat.

Figure 35: Rotary Evaporator (Rotavaporation)

- **Extract yield:**

$$\mathbf{R = \frac{W_0 - W_1}{W_0} \times 100 \ldots \text{IV (1)}}$$

R: Yield.

W0: Original weight.

W1: Weight after.

e) Stockage: we will Store the extract in a tinted glass bottle. Away from light, in a cool place.

IV.2.3 Physicochemical study of extract: At Laboratory of Dar El Beida Saidal, Algeria.

a) UV-VIS Spectroscopy:

UV-VIS spectroscopy is considered as the most important spectrophotometric technique that is most widely used for the analysis of variety of compounds (see figure 36). This technique works on the basis of the measurement of interaction of electromagnetic radiations (EMR) with matter at particular wavelength [163].

As one of the oldest analytical techniques, ultraviolet (UV) spectroscopy is a versatile tool indispensable in various scientific fields. Its ability to distinguish between different types of substances makes it a valuable tool in analytical chemistry. This method is commonly used to determine the identity, strength, quality, and purity of various samples. The basic principle of UV spectroscopy is based on the absorption of samples at specific wavelengths of light, providing valuable insights into the responses of materials to this absorption. This method is characterized by its accuracy and simplicity, and has a wide range of applications, including drug discovery, structural elucidation of organic molecules, molecular weight determination, and impurity detection. Both quantitative and qualitative analyses can be effectively performed using UV spectroscopy. The instrument operates within a wavelength range of 200–800 nm, allowing the analysis of colored and colorless compounds in both the visible and ultraviolet ranges. In short, UV spectroscopy is a powerful analytical technique with wide-ranging applications, making it a cornerstone of scientific research and analysis [164].

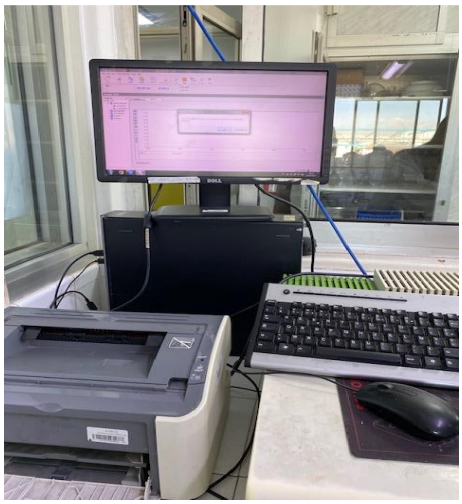


Figure 36: Spectrophotometer visible UV.

b) Infrared Spectroscopy:

Infrared (IR) spectroscopy is a fundamental tool of modern chemical analysis (see figure 37). A technique centered around the absorption of infrared light by molecules, it offers a window into the molecular world that reveals a wide range of details about chemical structures and compositions. IR spectroscopy's ability to identify and analyze many different materials from organic compounds to inorganic substances makes it indispensable across various scientific fields that rely on chemical analysis.

The importance and significance of this technique lie in its ability to detect vibrations of molecular bonds. When infrared light interacts with a sample, specific frequencies are absorbed, corresponding to the vibrational energies of the molecular bonds within the sample. This absorption creates a unique spectral fingerprint for each substance, providing insights into its molecular structure and characteristics [165].

Infrared spectroscopy is the measurement of the interaction of IR radiation with compounds. IR region involves the range between region 400-4000 cm-1 [166].

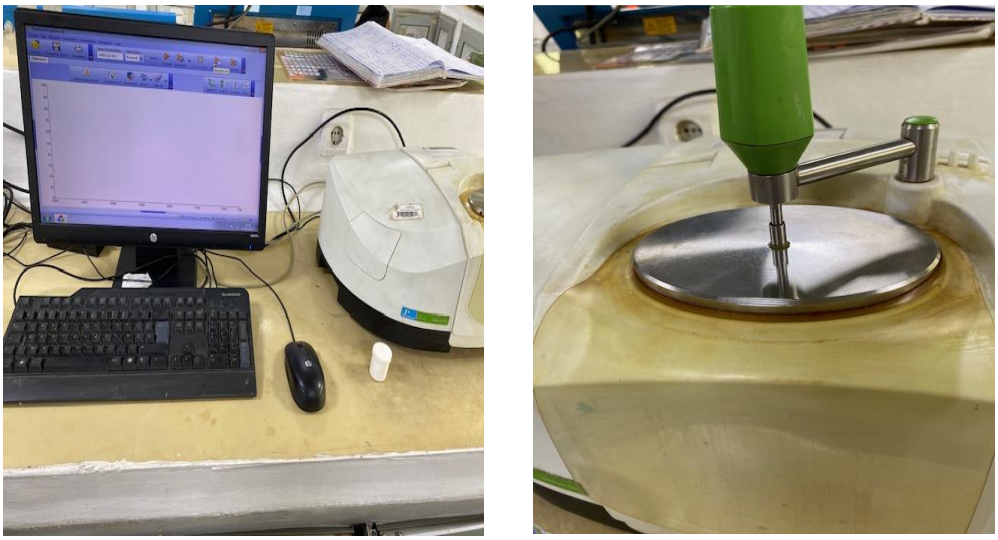


Figure 37: Spectrophotometer Infrared

IV.2.4 Determination of Antioxidant Potential:

Antioxidant potential is a critical parameter for evaluating the ability of plant extracts to neutralize free radicals, which are implicated in various diseases including cancer. This study employs multiple assays to comprehensively assess the antioxidant capacity of fenugreek ethanolic extract.

a) Principal of the DPPH test:

The DPPH (2,2-diphenyl-1-picrylhydrazyl) test evaluates antioxidant capacity through a single electron transfer (SET) and/or hydrogen atom transfer (HAT) reaction:

DPPH• (purple) + AH (antioxidant) →DPPH-H (yellow) +A• (oxidized antioxidant)… IV (2)

- DPPH• (purple-colored stable free radical) accepts an electron or hydrogen atom from an antioxidant compound, becoming reduced to DPPH-H (yellow-colored diamagnetic molecule).
- The degree of color change (purple → yellow) is proportional to the antioxidant capacity of the sample.

The reaction is monitored spectrophotometrically at 517 nm, where the decrease in absorbance indicates radical scavenging activity [167].

b) Operation mode: Protocol of laboratory Saidal dar el beida.

➢ **Preparation of fenugreek extract :**

- In this step we weigh 0.25µg of extract and we mix then with 0.25ml of ethanol (see figure 38).

Figure 38: fenugreek extract preparation.

➢ **Preparation of the DPPH test:**

- In this step, we will obtained by dissolving 4mg of DPPH in 100ml of ethanol (see figure 39).

✓ **Control solution:** 1 ml of ethanol and 1 ml of DPPH.

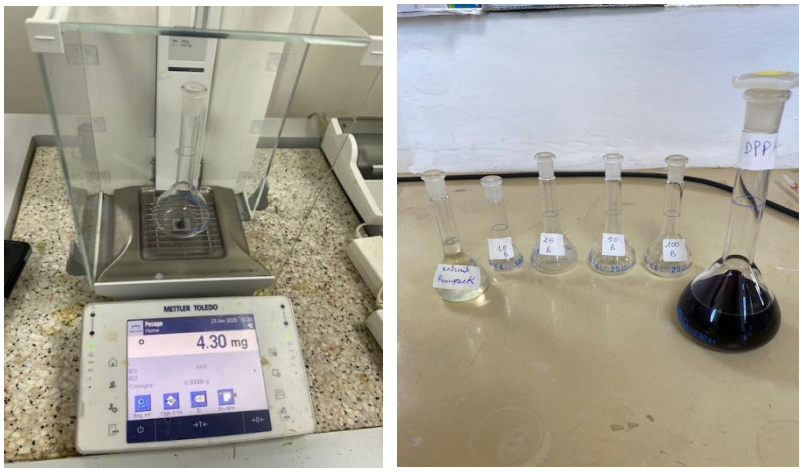


Figure 39: Preparation of DPPH solution

- The DPPH assay was used to assess the antioxidant capacity of the plant extract. 1 ml of different concentrations (10, 25, 50, and 100 µg/ml) of the extract in ethanol was added to 1ml of 0.004% ethanolic DPPH solution. After 30 minutes of incubation at room temperature, the absorbance value was read on a blank sample at a wavelength of 517 nm [168].

Table 8: Preparation of different sample concentrations

Sample	Extract solution (ml)	
	Extract of fenugreek (µg/ml)	**Ethanol (ml)**
10 µg/ml	10 µl	990 µl
25 µg/ml	25 µl	975 µl
50 µg/ml	50 µl	950 µl
100 µg/ml	100 µl	900 µl

IV.3 Molecular Docking Study:

Molecular docking is a key tool in structural molecular biology and computer-assisted drug design. The goal of ligand-protein docking is to predict the predominant binding mode(s) of a ligand with a protein of known three-dimensional structure. Successful docking methods search high-dimensional spaces effectively and use a scoring function that correctly ranks candidate dockings. The best-scoring pose is considered as the most likely binding mode of the ligand to the receptor [169].

IV.3.1 Material and methods:

In this study, we will simulate bioactive compounds from fenugreek (*Trigonella foenum-graecum*), turmeric (*Curcuma longa*) and *Atriplex halimus* as inhibitors of key breast cancer targets (HER2, CDK4, AKT1, and MCF-7) using molecular docking, DFT optimization, and ADMET prediction.

Induced fit docking was carried out using Schrodinger. Molecular docking is an important method for structure-based drug design and screening by studying the interaction of ligand and receptor molecules and predicting their affinity and binding modes. Molecular docking places molecules from a database of known 3D structures one by one at the active site of the target molecule. By continuously optimizing the position of the receptor compound, its conformation, the dihedral angle of the rotatable bonds within the molecule and the side chains and backbone of the amino acid residues of the receptor, the best conformation for the receptor small molecule compound to interact with the target macromolecule is searched and

its binding mode and affinity are predicted. The ligands with the best affinity to the receptor are selected by scoring functions that are close to the natural conformation. We will combine Gaussian 09W, AutoDock Vina, and BIOVIA Discovery Studio to facilitate and provide a powerful computational workflow, from structural optimization to toxicity prediction.

IV.3.2 Ligand Preparation:

In this step, we will enter the molecules into the Gaussian for optimization after to download from PubChem SDF forma (see figure). Then we will prepare the ligand in AutoDock Vina Tools for docking.

IV.3.2.a Optimization (DFT):

Gaussian 16 is used to perform DFT calculations for breast cancer drugs. Geometry optimization of these compounds was performed at the DFT level using the B3LYP hybrid functional.

Density functional theory DFT has become one of the most powerful modern methods for analyzing the properties of many body systems (atoms, molecules, and condensed phases). When using DFT to predict the significance of geochemical systems, the selection of a well-scaled exchange and correlation function and an appropriate basis set are of paramount importance. A basis set is a collection of basis functions that can be used in various calculations within a Gaussian program, such as optimization and energy calculations. The larger the basis set, the more accurate results are produced. Using this hybrid density functional basis set, the optimal scaling factor ensures valid and convergent frequencies for the majority of molecules. Hence, the basis set B3LYP/6-31G (2d, p) show in figure 39 was chosen. Apart from theoretical knowledge, experimental analyses were studied in detail, along with biological assays, to determine the potential of the titled compound as a drug candidate.

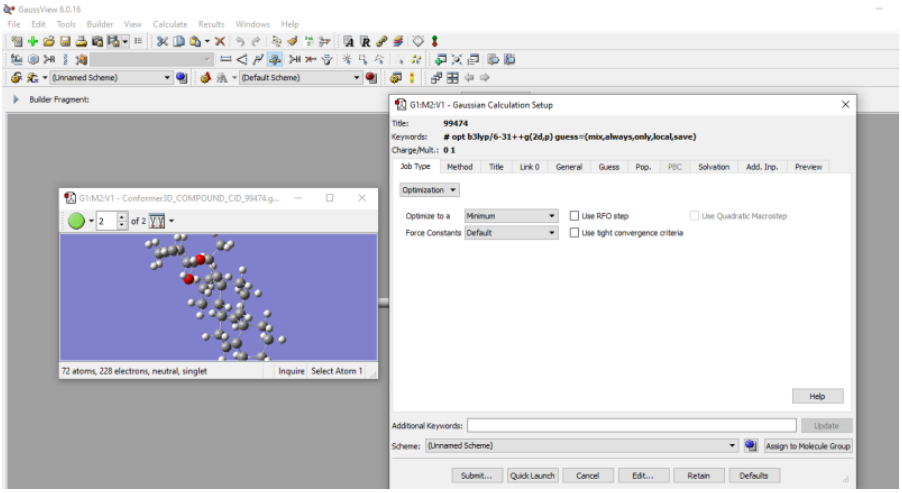
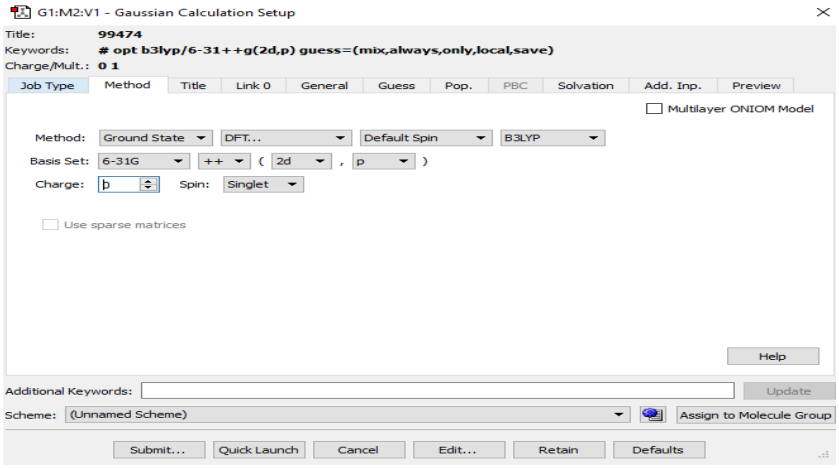
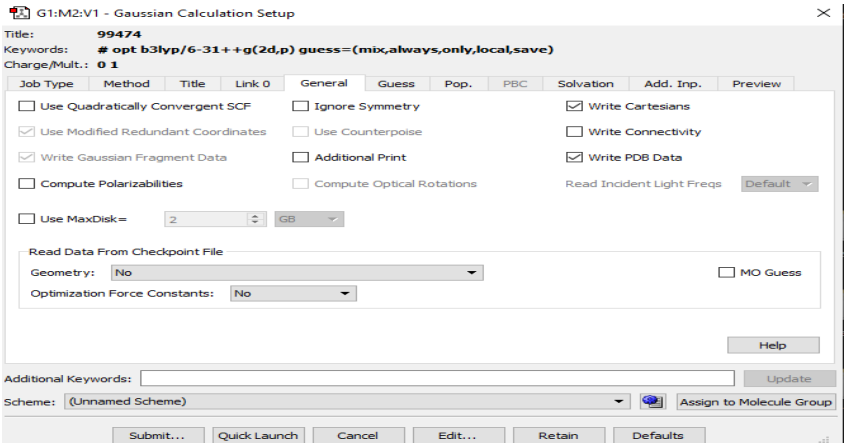


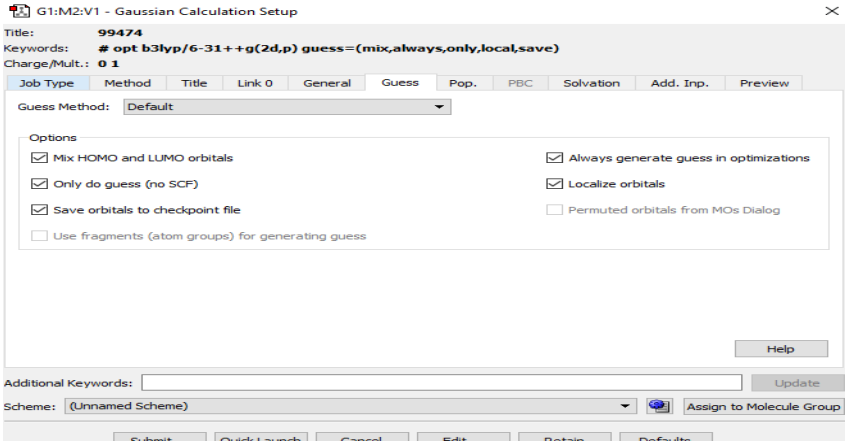
Figure 40: optimization step with GaussView.



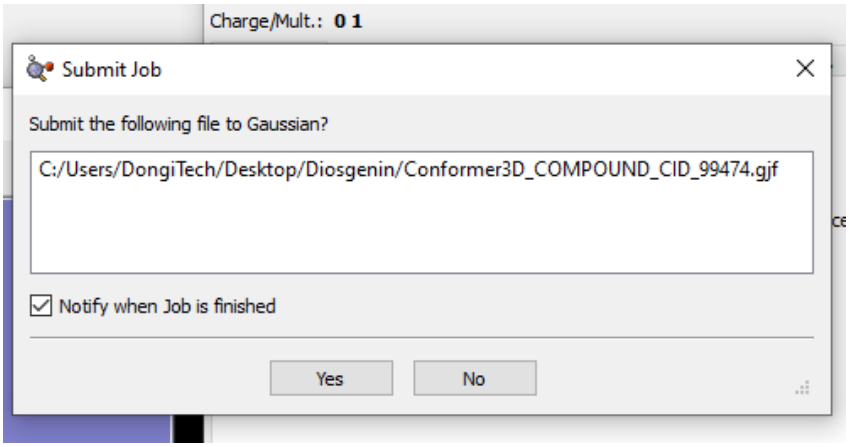
(1)



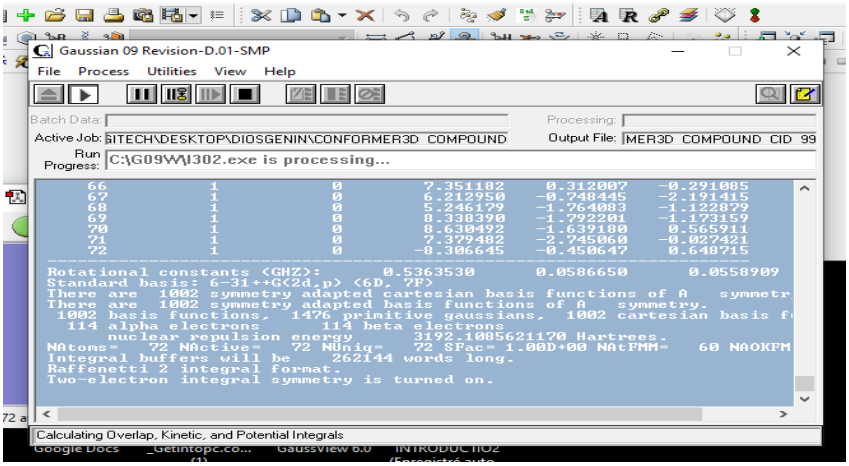
(2)



(3)



(4)



(5)

Figure 41: (1) DFT method. **(2)** Unhook write connectivity. **(3)** Click the 5 lockers.

4) Submit job. **(5)** Calculation process for Optimization.

- After starting the calculation process for optimization, we wait for the calculation period (see figure 42).

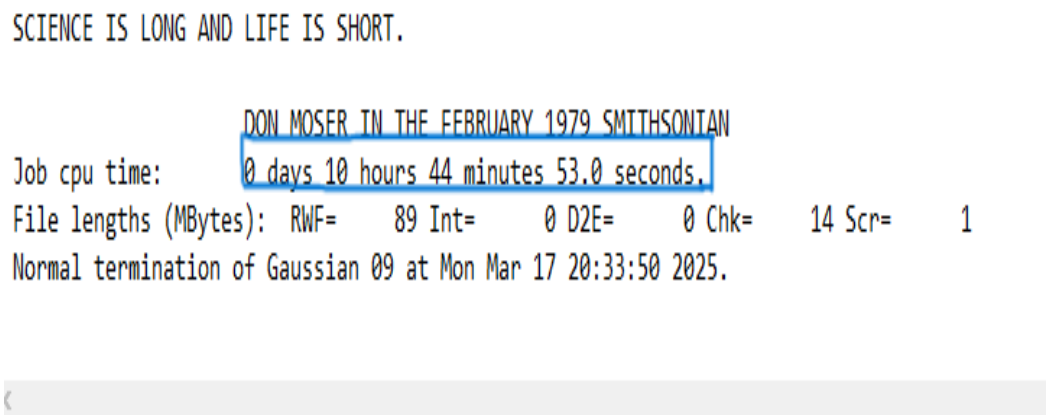


Figure 42: calculation period for optimization.

- After then which account comes out in the form of GJF, after which we convert in to PDB (see figure 43).

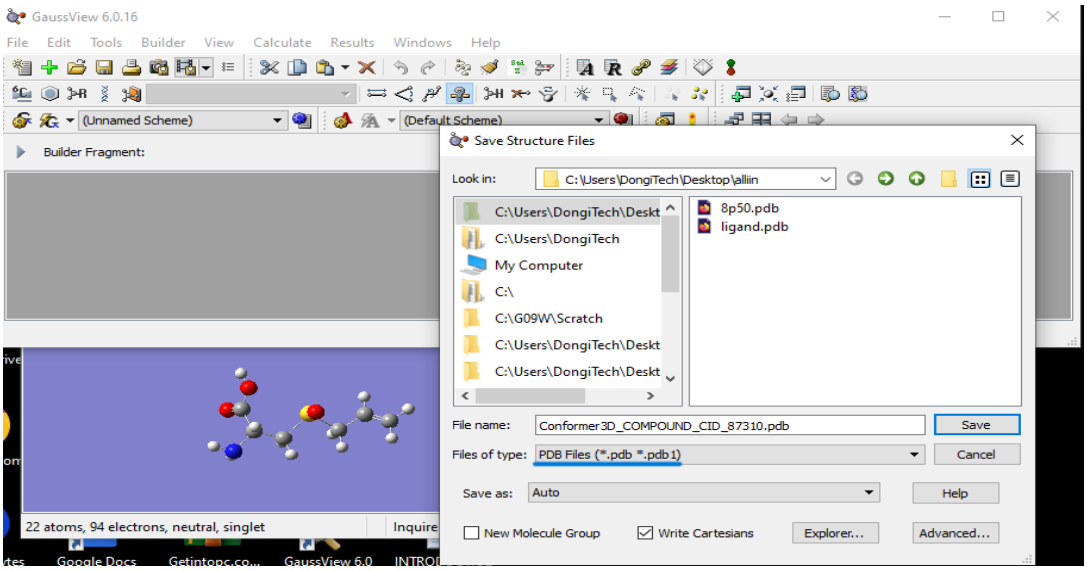


Figure 43: From GJF converts in to PDB.

```
Alpha  occ. eigenvalues --   -0.43085  -0.42728  -0.42327  -0.41778  -0.41302
Alpha  occ. eigenvalues --   -0.41039  -0.40960  -0.40404  -0.40142  -0.39745
Alpha  occ. eigenvalues --   -0.39365  -0.39181  -0.39060  -0.38371  -0.37643
Alpha  occ. eigenvalues --   -0.36966  -0.36303  -0.35955  -0.35706  -0.35555
Alpha  occ. eigenvalues --   -0.35474  -0.35420  -0.35261  -0.34155  -0.33981
Alpha  occ. eigenvalues --   -0.33956  -0.33306  -0.33061  -0.32439  -0.32095
Alpha  occ. eigenvalues --   -0.31045  -0.28120
Alpha virt. eigenvalues --   -0.17978  -0.15497  -0.13255  -0.11351  -0.08004
Alpha virt. eigenvalues --   -0.07636  -0.07331  -0.05558  -0.05100  -0.04487
Alpha virt. eigenvalues --   -0.03714  -0.02091  -0.02012  -0.01203  -0.00682
Alpha virt. eigenvalues --   -0.00243   0.00085   0.00249   0.00357   0.01009
Alpha virt. eigenvalues --    0.01135   0.01303   0.01685   0.02329   0.02727
Alpha virt. eigenvalues --    0.02933   0.03100   0.03586   0.03666   0.03948
Alpha virt. eigenvalues --    0.04261   0.04385   0.04891   0.05412   0.05679
```

Figure 44: HOMO and LUMO energy orbital

IV.3.3 Target Preparation:

In this step, we will download the targets (proteins/receptors) HER2, CDK4, MCF-7, and AKT1 from the rcsb PDB website in the Homo sapiens PDB format. We will then prepare the protein in AutoDock Vina Tools for docking.

IV.3.4 Ligand in AutoDock Tools:

After employing the Gaussian 09 quantic-chemical package, the resulting output files were employed as .pdb input files to molecular docking simulations. The molecular docking in silico experiment was conducted by AutoDock Tools 1.5.7 programe. The protein and ligand data were converted into three-dimensional PDBQT files, setting the follow parameters prior to docking simulations; we go to input and chose the setup ligand Contains Gasteiger charges merged, polar hydrogens, rotatable bonds and aromatic carbons. In the last step we chose output to save as pdbqt forma.

IV.3.5 Target in AutoDock tools:

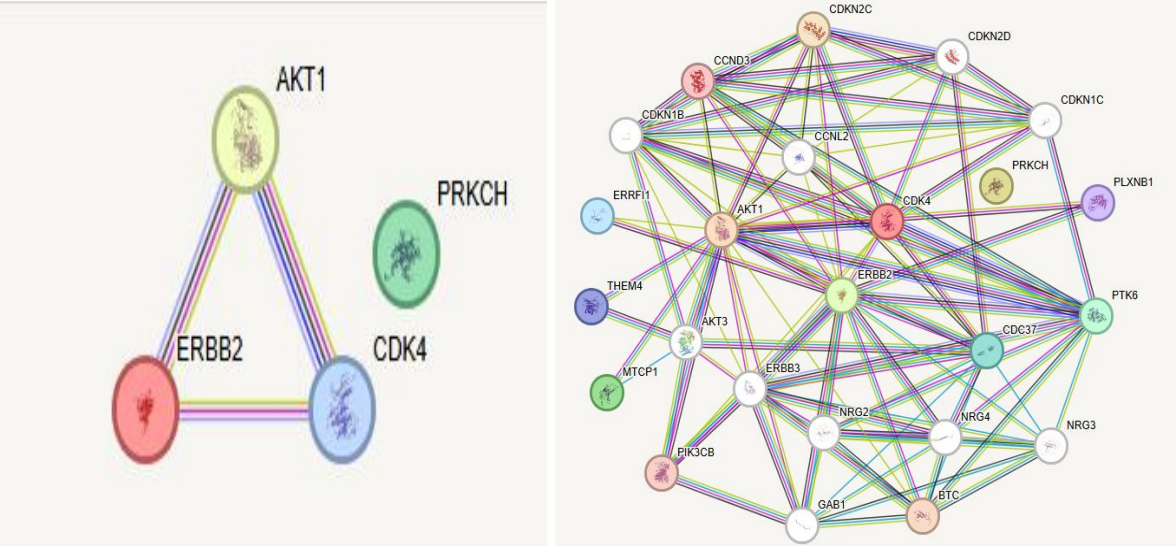
We insert the protein into the base AutoDock Tools and make the following adjustments:

We go to edit and chose in the first Delete Water and Add kollman Charges, then Add Hydrogens Polar Only. In the last step we go to Grid and chose Macromolecule and choose the target and save at pdbqt forma.

IV.3.6 Relationship between target proteins:

The integrated analysis of gene expression and Protein-Protein Interactions (PPI) would help to identify candidates that could serve as therapeutic targets as shown in figure 45. The STRING database was used to obtain data on the potential target genes and their interactions.

STRING is a database of known and predicted Protein-protein interactions. The interactions include direct (physical) and indirect (functional) associations; they stem from computational prediction, from knowledge transfer between organisms, and from interactions aggregated from other (primary) databases. [170]. (https://string-db.org/cgi/about)



(1) **(2)**

Figure 45: (1) and **(2)** Protein-protein interaction between the Studies receptors [170]

IV.3.7 Pharmacokinetic Properties:

Pharmacokinetic properties, including absorption, distribution, metabolism, excretion, and toxicity (ADMET), were evaluated. We also applied Lipinski's rule using established online platforms such as SWISSADME, ADMET-AI and ProTox-II. Initially, the chemical structures of selected phytochemicals were retrieved from the PubChem database, and their corresponding SMILES (Simplified Molecular Input Line Entry System) codes were extracted. These SMILES strings were then fed into relevant computational tools to predict their pharmacokinetic and toxicological properties. This computational approach facilitates a comprehensive assessment of the drug-likeness and safety potential of plant-derived compounds.

IV.3.8 Docking:

In this part, we do the interaction between the target pdbqt and ligand pdbqt. Protein-ligand interaction (PLI) plays an important role in new drug discovery, providing vital information about drug binding to target pathogen proteins. However, it also plays a fundamental role as a therapeutic target. It is currently considered one of the most challenging areas of drug discovery due to the specific structural characteristics of PLI.

IV.3.8.a InSilico interaction between Ligand pdbqt & Target pdbqt:

A grid box is created around the anticipated binding region in AutoDock to identify the active site. This tool aids in choosing the grid's location and size. The protein and ligand files must be properly prepared in PDBQT format. The next step is to set up the grid box to specify the area that the ligand will be docked with utilizing AutoDock tools, such as the active site or possible binding site. You can drag the box to the desired location or manually enter the X, Y, and Z values to change the grid's center [119]. The grid box should have adequate room for the ligand and be big enough to cover the active site (see figure 46). We click on [File, Output grid dimensions file] and save it as grid.txt in the previous file.

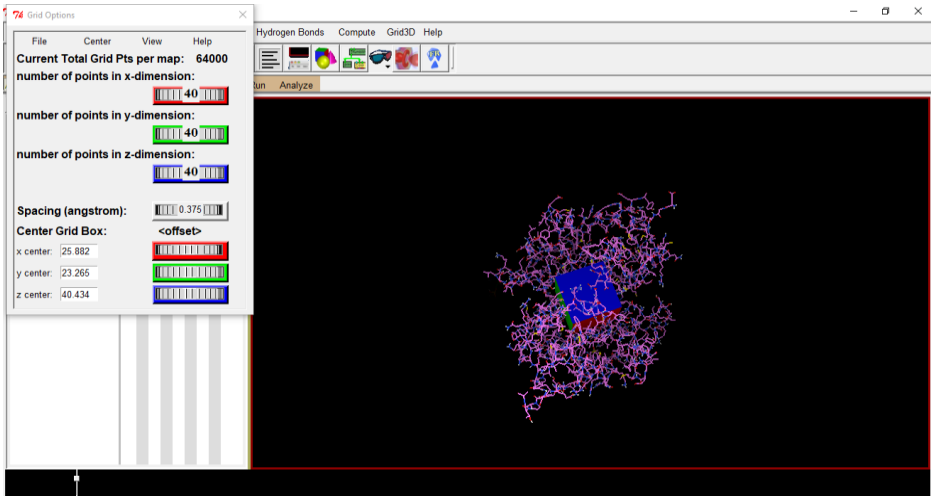
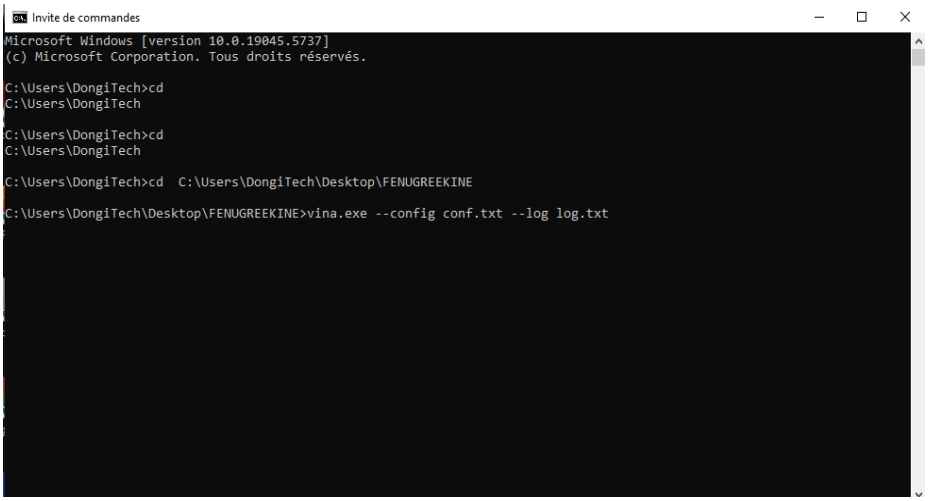


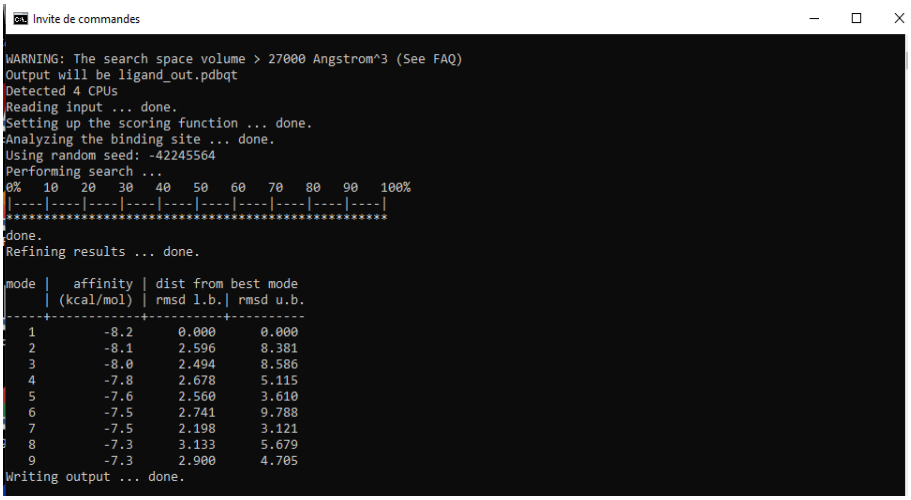
Figure 46: Interface of centr Grid Box

IV.3.8.b Vina AutoDock:

In this step we will created a folder which contains the ligand pdbqt, target pdbqt, conf file and the Vina file (see figure 47), For the purpose of calculation the energy band.



(1)



(2)

Figure 47: (1) structure of command Vina. **(2)** Affinity calculates by Vina.

IV.3.8.c BIOVIA Discovery Studio:

Initially, the target pdbqt protein was prepared by removing all bound ligands and non-essential molecules to ensure a clean docking environment. Subsequently, the desired ligand (ligand_out pdbqt) was introduced into the protein's active site (see figure 48). Optimal binding positions were determined through docking simulations, allowing for the identification of the most favorable ligand conformation within the binding pocket. Detailed interaction analyses were then conducted to examine the nature of the molecular interactions between the ligand and protein, focusing particularly on the amino acid residues involved in the binding. Special attention was given to hotspot regions, where key amino acids formed significant interactions—such as hydrogen bonding, van der Waals forces, and hydrophobic contacts—indicating the stability and specificity of the ligand-protein complex.

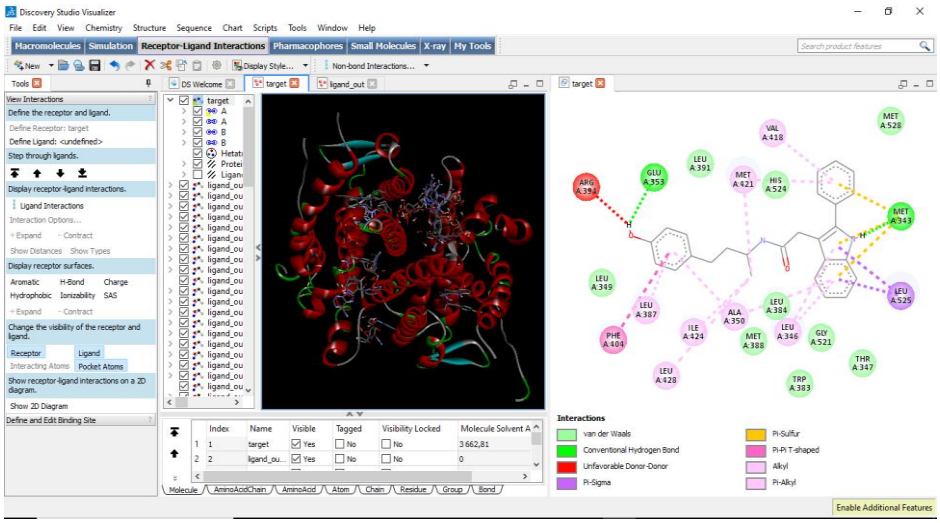
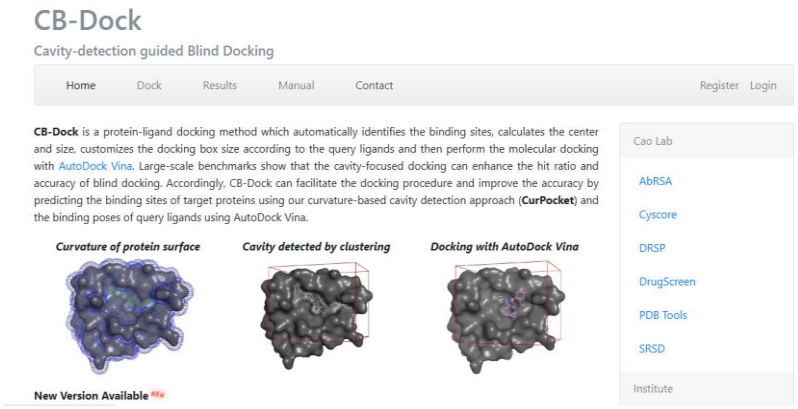


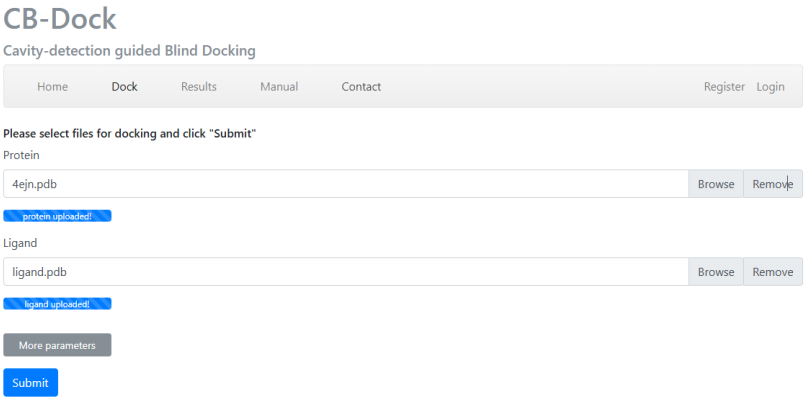
Figure 48: interfaces for ligand-out Pdbqt in program BIOVIA Discovery Studio and add protein Pdbqt, show 2D Diagram and Interaction.

IV.3.8.d CB-Dock:

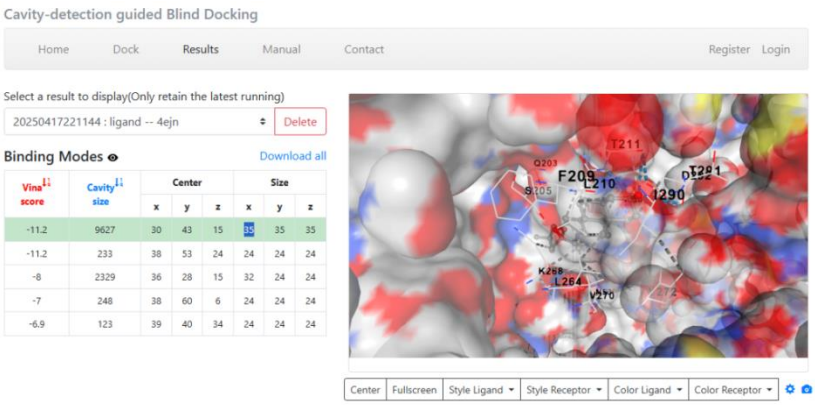
Cavity-detection guided Blind Docking is a protein-ligand docking method that automatically identifies binding sites, calculates the center and volume, and customizes the docking box size according to the query ligands. It predicts protein-binding sites and calculates their centers and sizes using an innovative curvature-based cavity detection approach, and performs docking using the popular docking software AutoDock Vina. This method was carefully optimized and achieved a success rate of approximately 70% for the best-ranked poses, with the root mean square deviation (RMSD) being within 2 angstroms of the X-ray pose, outperforming state-of-the-art blind docking tools in our benchmark tests [171]. CB-Dock requires the input of a protein file to be in the PDB format and a ligand file in the pdb, mol, mol2, sdf format (see figure 49).



(1)



(2)



(3)

Figure 49: (1) CB-Dock screenshot. **(2)** Select the protein and ligand. **(3)** CB-Dock result.

IV.4 Quantitative Structure Activity Relationship:

In the important part of this study, a quantitative structure-activity relationship (QSAR) modeling approach was adopted to build a predictive model linking the chemical structures of the studied compounds to their biological activities against target proteins associated with breast cancer. This model was developed using multiple linear regression (MLR) and support vector regression (SVR) techniques, combining the explanatory capabilities of linear models with the predictive accuracy of nonlinear models. This integrated framework provides a powerful tool for rational drug design, contributing to the identification of the most biologically effective compounds and improving their properties for the development of promising therapeutic agents.

IV.4.1 Databases:

In an effort to validate the molecular docking results regarding the inhibition of key proteins associated with breast cancer. The initial strategy involved identifying structural analogs of the studied phytochemicals. However, the diversity of the targeted proteins posed a significant limitation, preventing the development of a unified dataset aligned with a single protein target. To address this constraint, a curated dataset was selected comprising structurally diverse, bioactive compounds with reported inhibitory activity specifically against the MDA-MB-231 breast cancer cell line, sourced from previously published studies [172][173]. We selected compounds derived from parthenolide (see Figure 50) and the curcumin molecule with its derivatives (see Figure 51).

The QSAR model was constructed based on the biological activity values expressed as pIC50. For each compound, the PubChem Compound Identifier (CID) along with its structures and SMILES notation were retrieved from the PubChem database. The compound identifiers and their associated activity values are presented in Table 9.

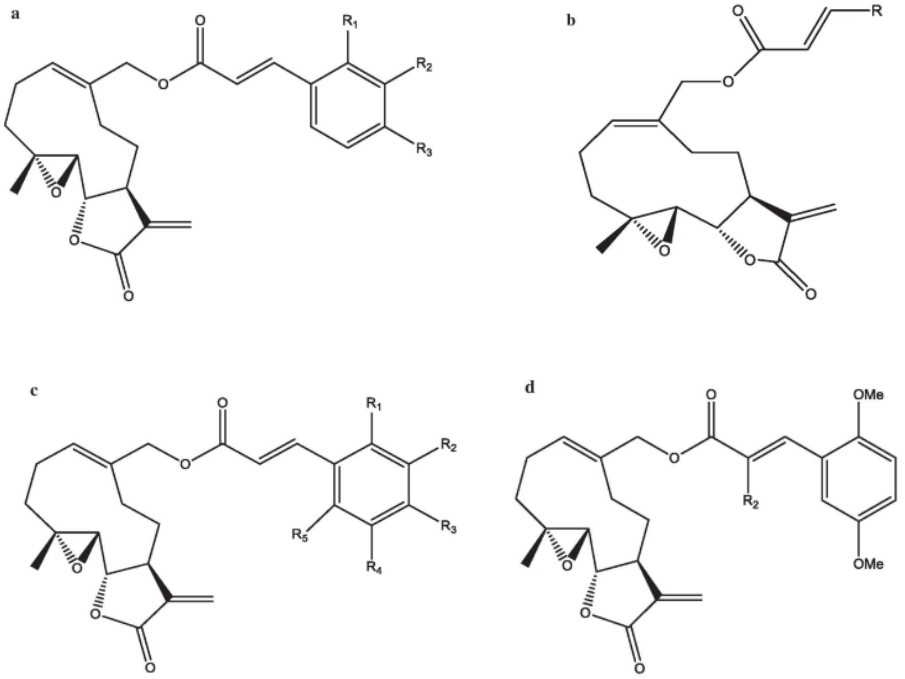


Figure 50: (**a)** Parthenolide template molecule 1. (**b)** Parthenolide template molecule 2.

(c) Parthenolide template molecule 3. **(d)** Parthenolide template molecule 4.

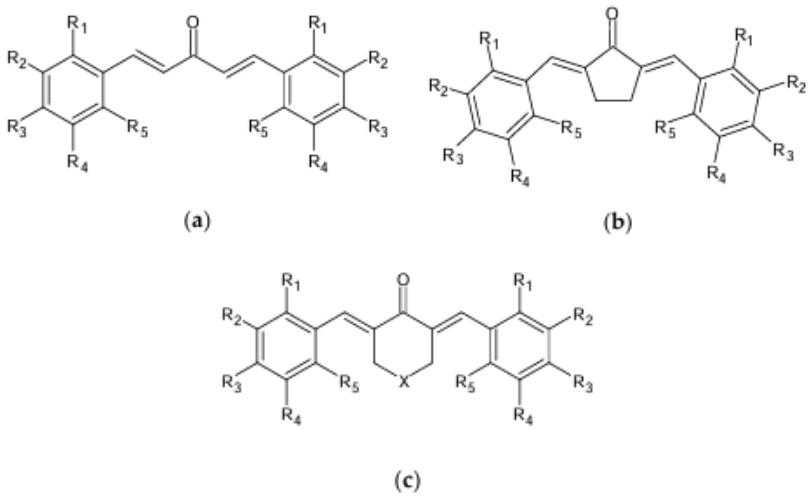
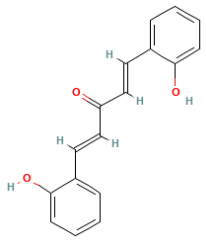
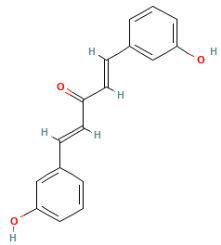
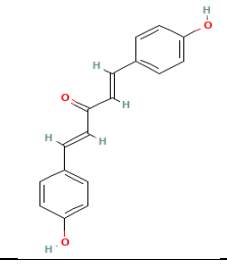
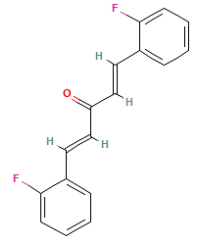
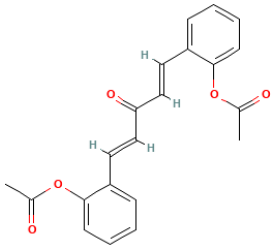
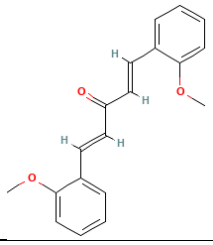
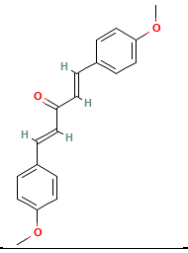
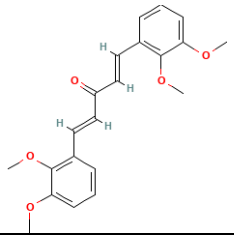
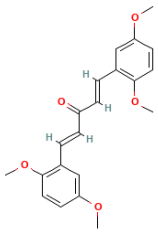
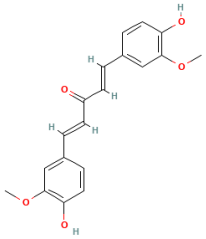
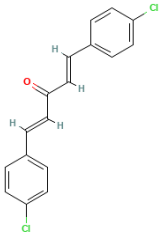
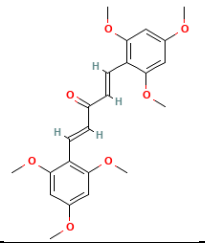
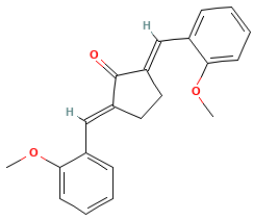
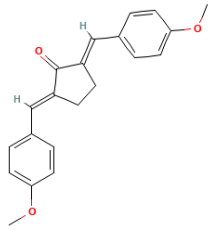
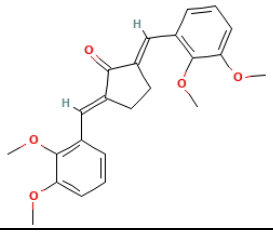
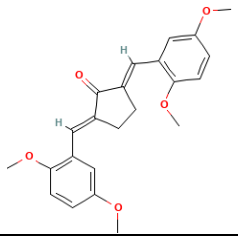
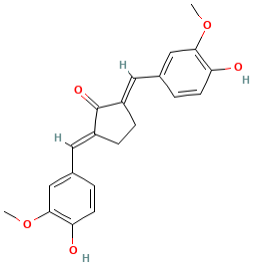


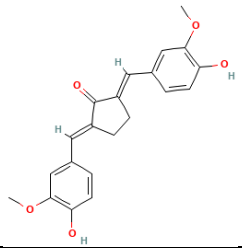
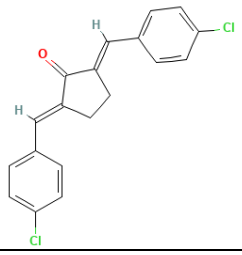
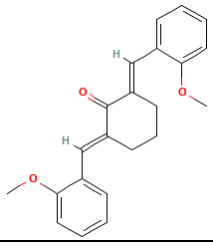
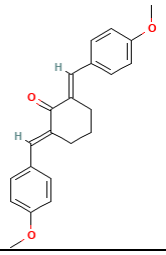
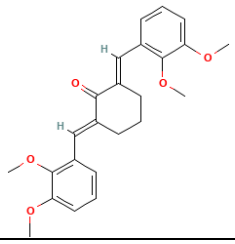
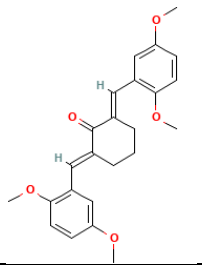
Figure 51: Compounds structures used for QSAR model development. **(a)** Series 1 analogues with acetone core. **(b)** Series 2 analogues with cyclopentanone core. **(c)** Series 3 analogues with cyclohexanone and 4-piperidone cores.

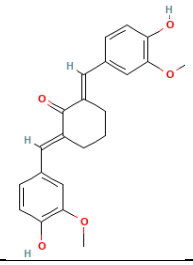
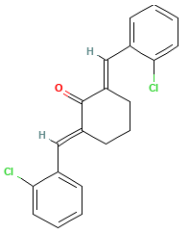
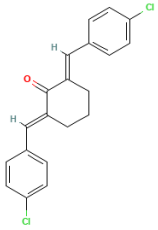
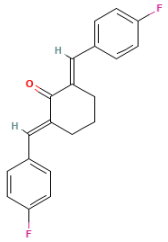
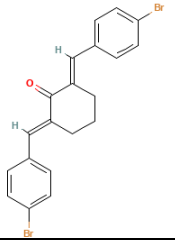
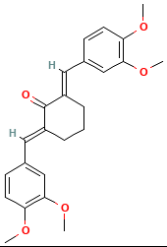
Table 9: PubChem CID for compounds and experimental PIC50 values and predicted by MLR; SVR.

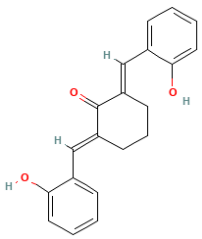
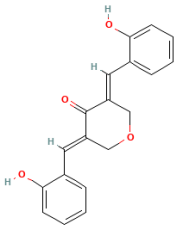
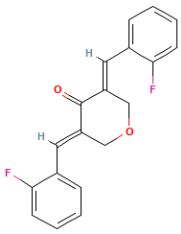
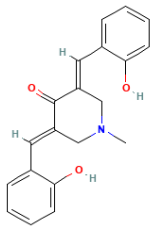
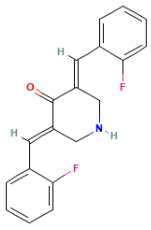
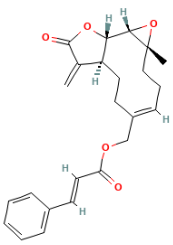
N°	CID compound	pIC50	MLR predict	SVR predict
1	5472867 	5,64	5,718	5,6930
2	5893800 	5,41	5,4294	5,3915
3	6437306 	5,12	5,2183	5,2288
4	2308751 	4,90	4,7704	4,7870
5	5468558 	5,57	5,5975	5,5624

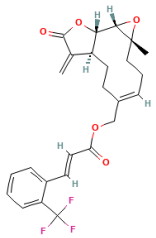
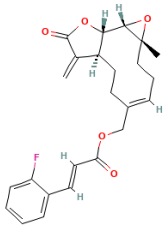
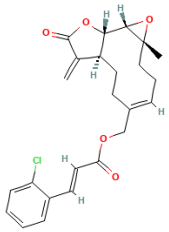
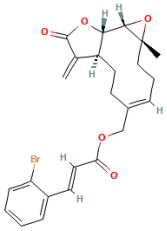
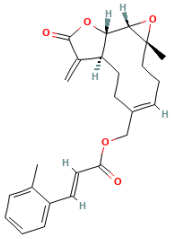
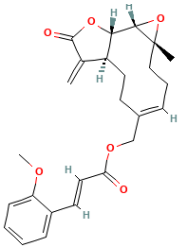
6	830608 	5,68	5,7736	5,7729
7	715840 	4,96	4,9537	4,9541
8	1125827 	5,02	4,8463	4,9027
9	10831969 	5,82	5,8259	5,8182
10	6474893 	5,04	5,1745	5,1038
11	5378584 	4,85	4,7384	4,7925

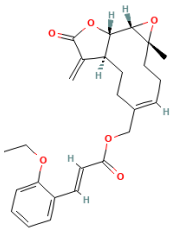
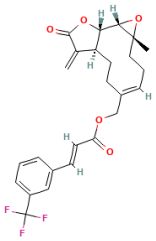
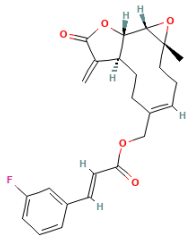
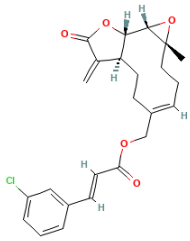
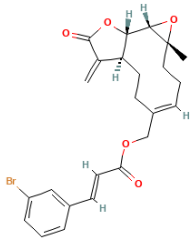
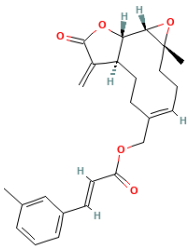
12	10501916 	5,51	5,5098	5,5204
13	1719442 	4,74	4,7022	4,6782
14	1551688 	4,52	4,5398	4,6204
15	1877197 	4,52	4,8898	4,9950
16	2055274 	4,56	4,6853	4,6841
17	5470829 	5,19	5,1183	5,1306

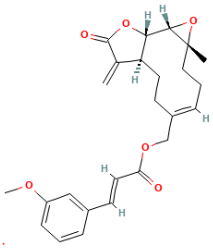
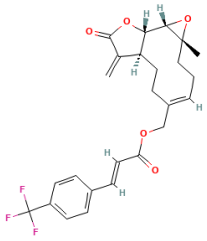
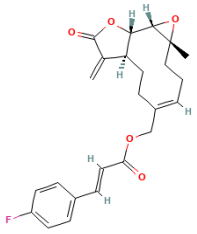
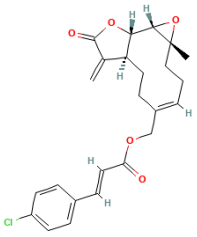
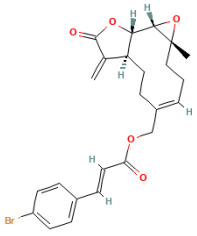
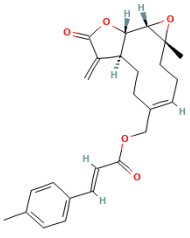
18	5714141 	5,27	5,3975	5,3968
19	1551005 	4,52	4,6192	4,6127
20	1714168 	5,52	5,4014	5,3902
21	1550699 	4,19	4,6266	4,6220
22	5781779 	5,02	5,1656	5,1107
23	1811454 	5,64	5,6392	5,6227

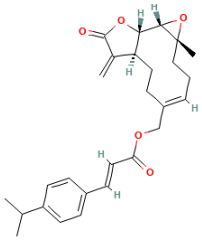
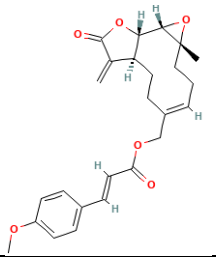
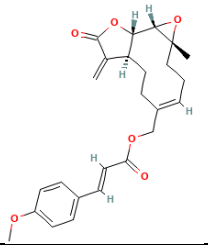
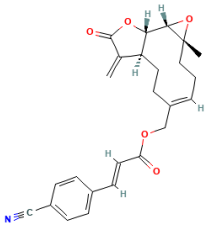
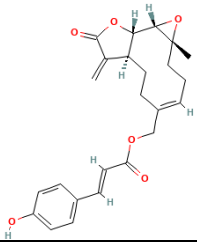
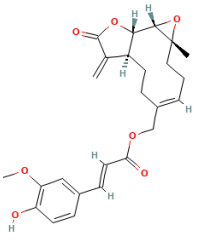
24	1550234 	5,52	5,2440	5,1967
25	1736139 	4,52	4,5937	4,6191
26	1551008 	5,13	5,1540	5,1474
27	1551384 	5,52	5,3562	5,4195
28	5716584 	5,20	5,3537	5,2757
29	1714173 	4,92	4,9192	4,9241

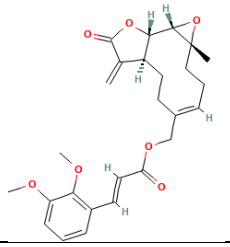
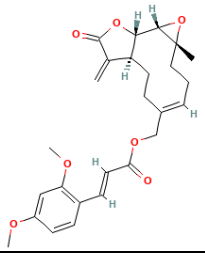
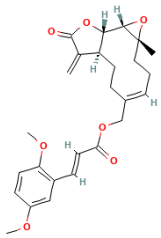
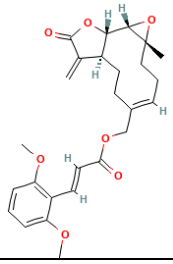
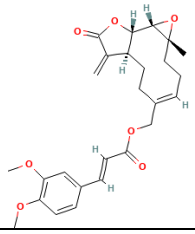
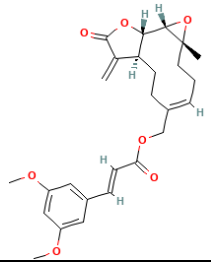
30	10267144 	5,44	5,3325	5,3184
31	9904691 	6,10	6,1249	6,0892
32	6123787 	5,19	5,2222	5,2219
33	171343759 	5,72	5,9127	5,8823
34	9885748 	6,10	6,2366	6,1520
35	137642970 	5,8697	5,8761	5,8639

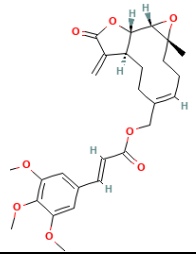
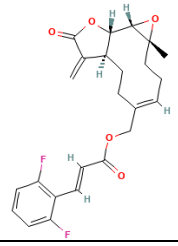
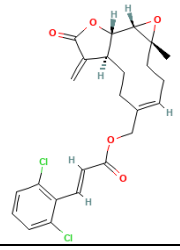
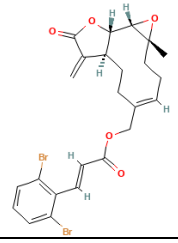
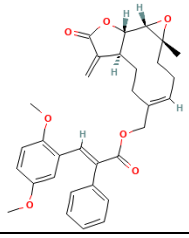
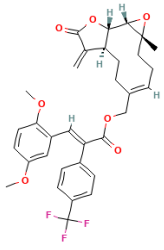
36	155538480 	6,1308	6,2769	6,1743
37	155566636 	5,7447	5,7099	5,7281
38	155519783 	5,8477	5,8099	5,8176
39	155557656 	5,8761	5,9020	5,8945
40	155557594 	5,7471	5,7547	5,7703
41	155554857 	6,3372	6,2571	6,3571

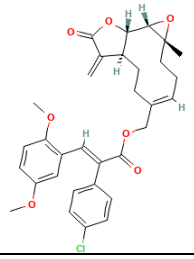
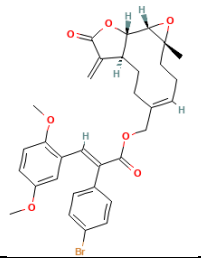
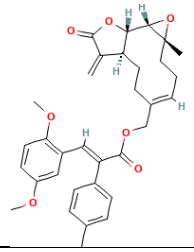
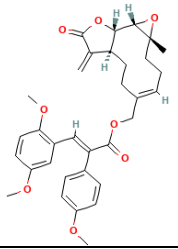
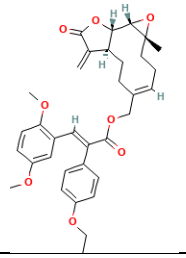
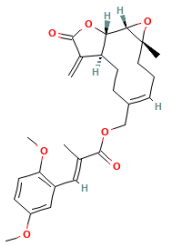
42	155557229 	6,4685	6,2895	6,9632
43	155512509 	6,0809	6,0296	6,0724
44	155537931 	5,7610	6,1262	6,1307
45	155522898 	5,7496	5,7137	5,7063
46	155510682 	5,7852	5,8248	5,7154
47	155547105 	5,9666	5,7281	5,7044

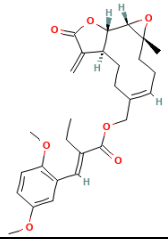
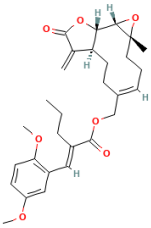
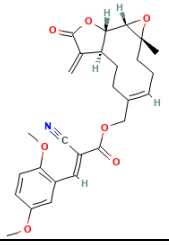
48	155554277 	5,7352	6,0096	5,9789
49	155532446 	5,5986	5,7521	5,7645
50	155548781 	5,8794	5,6309	5,6027
51	155521805 	5,9136	5,8578	5,8634
52	155562971 	5,9914	5,9272	5,8621
53	155557766 	5,6430	5,6519	5,6610

54	155562847 	5,8761	5,8228	5,7919
55	155535337 	5,7610	5,8039	5,8201
56	155553179 	5,8013	5,7131	5,7033
57	155566672 	5,7447	5,8091	5,8216
58	155524389 	5,7235	5,8416	5,8526
59	155535995 	5,6968	6,0905	5,8707

60	155534508 	5,7595	5,7114	5,7885
61	155556346 	6,2596	6,2767	6,8037
62	155561714 	6,5850	6,3046	6,5697
63	155530276 	6,6021	6,3960	6,1662
64	155558666 	5,7090	5,8431	5,8542
65	155512879 	5,6946	5,7629	5,7709

66	155529923 	6,0223	5,7629	6,6768
67	155520457 	5,7986	5,7046	6,0199
68	155553834 	5,7986	6,0968	5,8193
69	155541645 	5,8239	5,7868	5,8789
70	155515824 	5,6968	5,8088	5,8185
71	155540990 	5,7033	5,8198	5,8594

72	155541926 	5,6946	5,7076	5,6869
73	155515336 	5,6968	5,8481	5,7358
74	155525692 	5,5302	5,7089	5,7331
75	155535553 	5,6003	5,7067	5,7181
76	155563987 	5,7282	5,5771	5,8322
77	155537222 	5,7545	5,8417	5,7689

78	155515655 	5,6478	5,7109	5,8502
79	155529681 	5,1232	5,2501	5,2227
80	155525864 	5,6289	5,6112	5,6208

IV.4.2 Regression methods:

In this phase of the research, we expand our analytical framework by applying linear and nonlinear modeling techniques using modern, advanced computational codes. We relied on specialized artificial intelligence platforms to extract appropriate algorithms for data analysis and the creation of predictive models. However, most of these algorithms contained programming errors that prevented accurate and reliable results. Therefore, we made great efforts to review and correct these codes with the help of algorithm programming and data analysis specialists. After several attempts and successive modifications, we were able to successfully address these technical issues, ultimately enabling us to obtain the desired results and achieve the study's objectives. In this phase of the research, we expand our analytical framework by applying linear and nonlinear modeling techniques using modern, advanced computational codes. We relied on specialized artificial intelligence platforms to extract appropriate algorithms for data analysis and the creation of predictive models. However, most of these algorithms contained programming errors that prevented accurate and reliable results. Therefore, we made great efforts to review and correct these codes with the help of algorithm

programming and data analysis specialists. After several attempts and successive modifications, we were able to successfully address these technical issues, ultimately enabling us to obtain the desired results and achieve the study's objectives.

The linear approach, primarily based on multiple linear regressions (MLR), provides interpretable insights into the contribution of each molecular descriptor to biological activity. In parallel, a nonlinear approach, such as support vector regression (SVR), is used to capture more complex and potentially nonlinear interactions that traditional linear models may overlook.

By applying this dual modeling strategy, we can systematically compare the performance metrics (such as R^2, RMSE, and MAE) of both methods, identify limitations, and highlight the strengths of each algorithm in modeling molecular activity associated with breast cancer.

IV.4.2.1 Multiple linear model:

At this stage of the QSAR modeling workflow, multiple linear regression (MLR) algorithms for molecular dataset is used are implemented using the Spyder integrated development environment (IDE) (see figure 52,53), part of the Anaconda distribution, which is widely used in scientific computing using Python and RDKit (a Python library): a powerful open-source cheminformatics toolkit with extensive descriptor computation capabilities. It is essential to carefully review and modify any file paths or file names within the code. This includes specifying the correct input files for molecular descriptors and biological activity data. After successful code execution, the program processes the descriptor dataset. Finally, the descriptors are computed. Typically, the dataset is split into training and test sets, after which a model is trained. After the model is trained, MLR is evaluated using performance metrics, including the coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE).

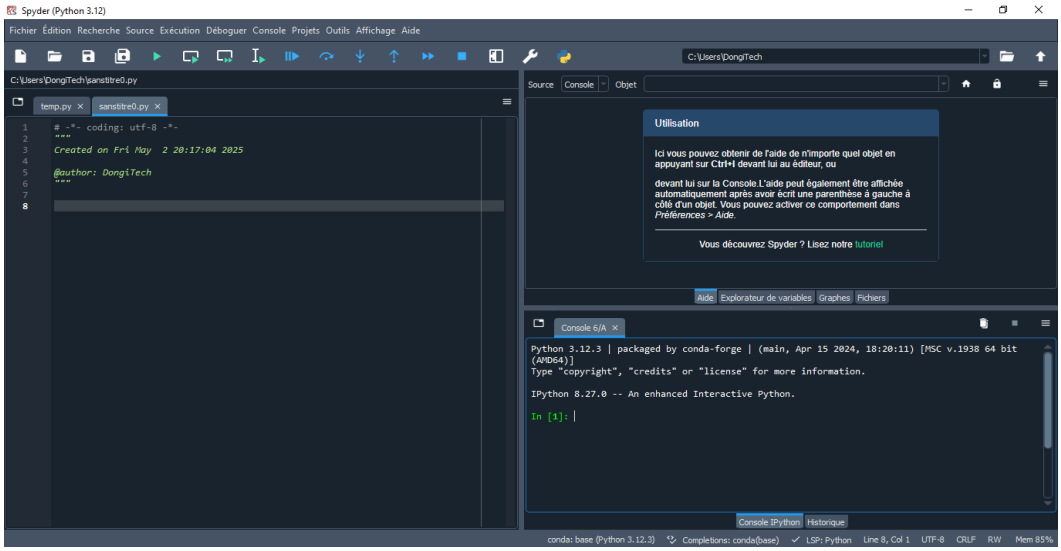


Figure 52: screenshot of Spyder programe.

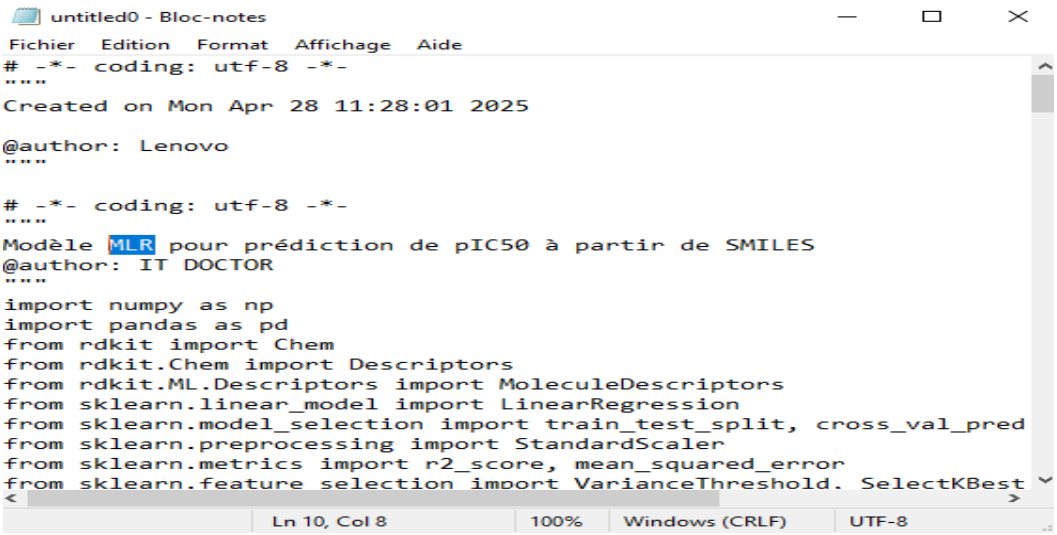


Figure 53: MLR code

IV.4.2.2 Support vector model:

In the case of support vector regression (SVR), a methodological framework similar to that of multiple linear regressions (MLR) is followed. Initially, the same molecular dataset is used. The SVR algorithm is implemented using appropriate computing libraries, such as Spyder in Python (see figure 54). Next, molecular descriptors are calculated. These descriptors serve as independent variables representing the chemical structure and

physicochemical properties of each compound. Typically, the dataset is split into training and test sets, after which an SVR model is trained to learn the underlying nonlinear relationships between the molecular descriptors and the biological activity under study. After model training, SVR is evaluated using the same performance metrics used for MLR, including the coefficient of determination (R^2), root mean square error (RMSE).

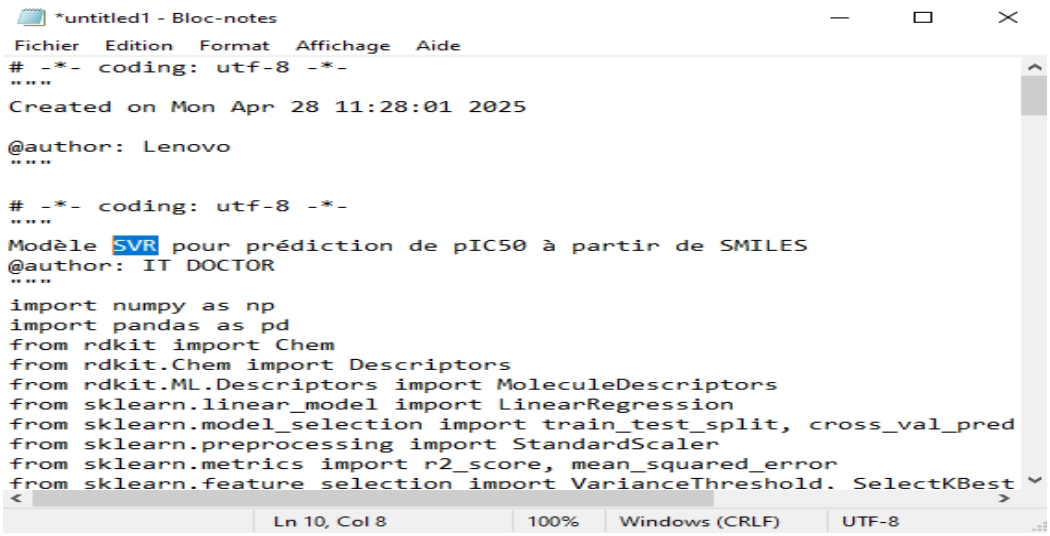


Figure 54: SVR code.

IV.5 Conclusion:

This study presents an integrated approach to assessing the potential of fenugreek and turmeric extracts in the treatment of breast cancer. Computational analyses further supported their potential by identifying specific compounds with favorable interactions with proteins and targets associated with breast cancer. The combination of in vitro analysis and computational simulations not only demonstrated the bioactivity of these natural products but also highlighted their promising potential as complementary or alternative agents in breast cancer treatment. These findings pave the way for future preclinical and clinical studies aimed at developing safer, more affordable, and natural therapeutic options derived from traditional medicinal plants.

Chapter V

Results and Discussion

V.1 Introduction:

This chapter presents the results collected from the laboratory studies, molecular docking simulations and QSAR study we conducted in previous chapters and their interpretations.

V.2 Experimental study results:

In this study, we conducted a comprehensive investigation into the bioactive potential of fenugreek (*Trigonella foenum-graecum*) through ethanolic extraction and advanced spectroscopic analyses, revealing significant results with therapeutic implications.

V.2.1 Extraction yield:

The extraction yield represents the efficiency of the extraction process in obtaining bioactive compounds from plant material, calculated as the percentage ratio of the dry extract mass to the initial dry plant mass. For fenugreek (*Trigonella foenum-graecum*), typical extraction yields range between 12-25% depending on the method and solvent system used.

For fenugreek, 60% ethanol provides a balanced yield (12–20%), effectively extracting phenolics, saponins (e.g., diosgenin).

V.2.2 Color/Apparence:

Extract visual inspection. These properties—purity, color, and aroma—provide valuable initial information about the quality and composition of the extract before proceeding with further analytical evaluations, as shown in Table 10.

Table 10: Extract characterization.

Extract	Appearance	Color	Smells
Fenugreek Extract	Clear liquid	Yellow	Full of ethanol

V.2.3 Visible UV results:

The UV-Visible spectroscopy analysis of the fenugreek ethanolic extracts revealed characteristic absorption patterns indicative of its phytochemical composition.

- In this step we dilute the extract and read the samples in the UV-Visible spectroscopy.

Table 11: UV visible analyses results.

Concentration	**0.50**	**0.45**	**0.42**	**0.38**	**0.34**	**0.30**	**0.28**	**0.26**	**0.24**	**0.20**	**0.10**
Absorbance	0.965	0.936	0.878	0.728	0.699	0.510	0.419	0.381	0.333	0.202	0.151

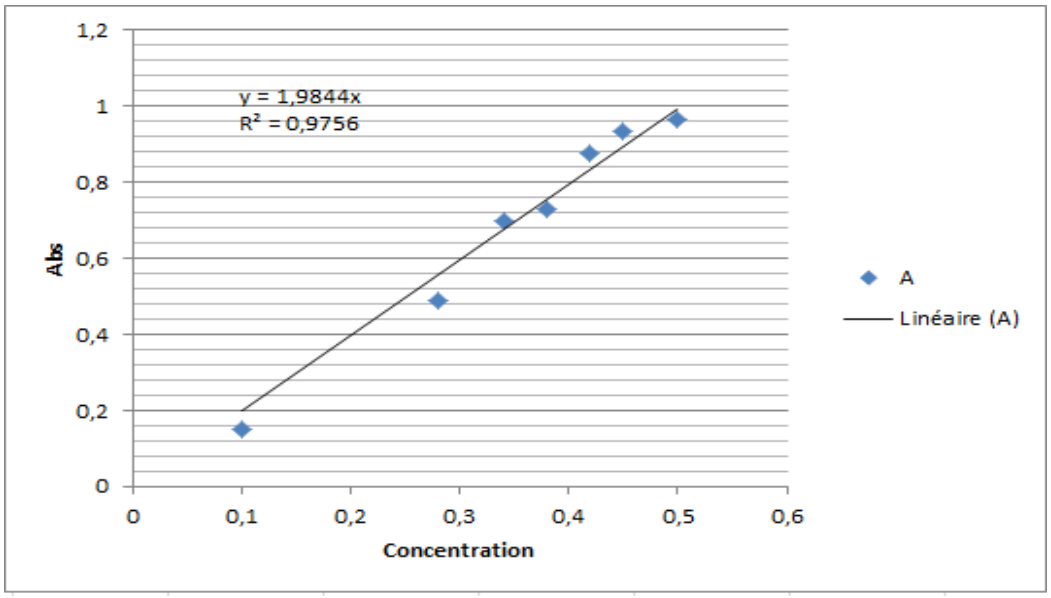


Figure 55: Fenugreek extracts absorbance.

- In this step we will read in to visible UV spectrophotometry the extract of fenugreek

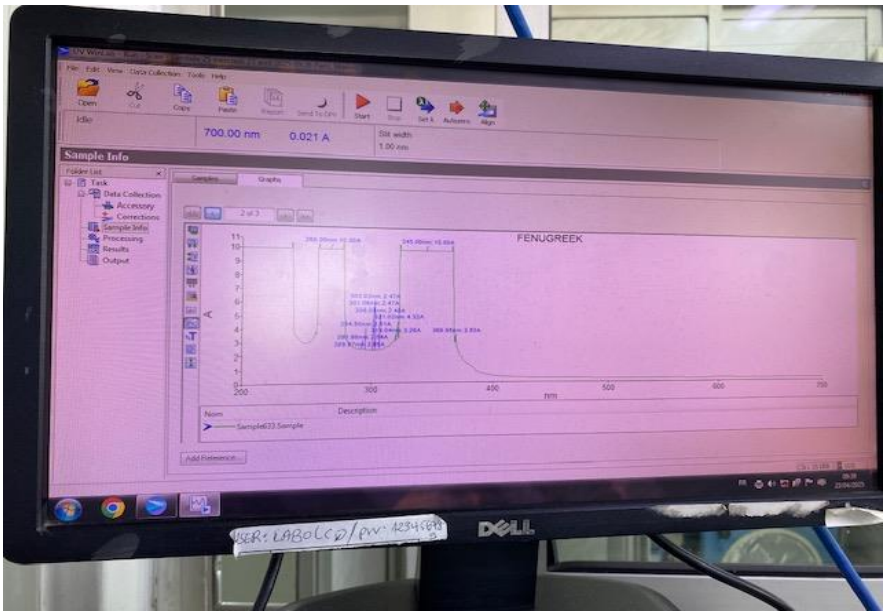


Figure 56: Determine the absorbance of fenugreek.

Interpretation:

UV-Vis spectroscopic analysis of the fenugreek extract revealed distinct absorption patterns indicative of its saponin content, albeit with limited resolution. The spectrum showed absorption in the 200-nm range, consistent with the aliphatic and steroidal backbones of triterpenoid saponins such as diosgenin and yamugenin. The peak observed around 280-300 nm suggests the possible presence of saponin systems or associated phenolic compounds. The broad, poorly defined peaks indicate a complex mixture of saponin glycosides with varying sugar groups. While these UV characteristics are suggestive of the presence of saponins, they are not conclusive due to the potential for overlapping of extracted compounds and the poor UV absorption of most saponins. This underscores the need for complementary techniques such as HPLC-MS.

V.2.4 Infrared spectroscopy results:

The functional groups and bond structures of the fenugreek seed extract were determined using a PerkinElmer Spectrum IR version 10.6.2 at Saidal Laboratory. Infrared spectra were measured in the range of 400 to 4000 cm-1 with a selected resolution of 4 cm-1 at room temperature (25°C). Approximately one drop of the extract was placed in the instrument to form a thin layer, and then directly measured using the infrared spectrometer.

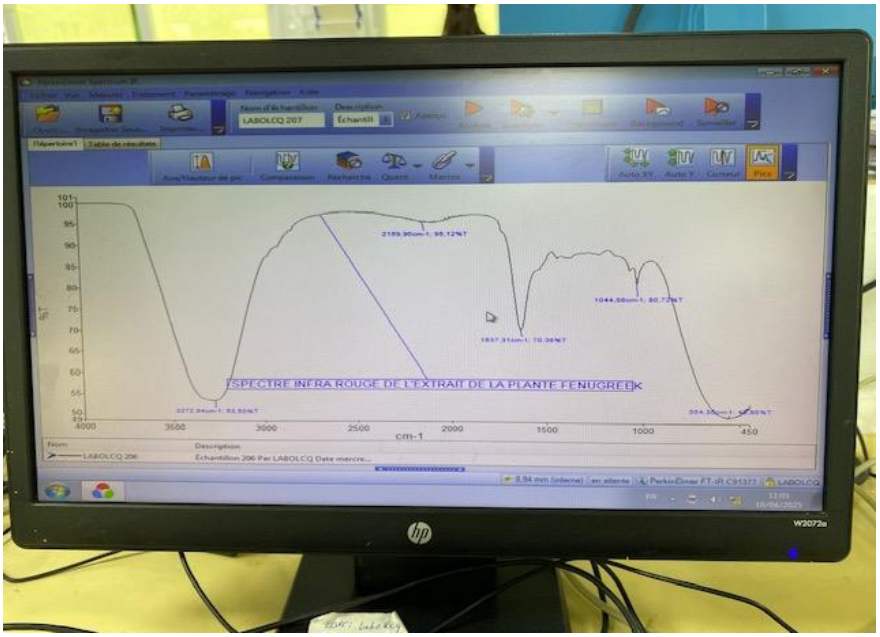


Figure 57: Spectre of infrared result.

Interpretation:

The functional groups of the fenugreek seed extract were identified using infrared spectra, as shown in Figure 49. Functional groups were identified at an intermediate peak, groups at a strong peak, and one functional group at a weak peak between 500 and 3500 cm^{-1}. The infrared (IR) spectrum of fenugreek extract reveals several characteristic peaks, indicating the presence of specific functional groups associated with the bioactive compounds of this plant. Here is a detailed analysis of the observed absorption bands:

- Peak at 1637.31 cm^{-1}

(C=O) group: Carbonyl bonds of flavonoids or carboxylic acids. Probable origin: Carbonyl compounds (e.g., diosgenin, a steroidal saponoside of fenugreek). It confirms the presence of saponins (e.g., diosgenin) and flavonoids, which are responsible for antioxidant activity.

Diosgenin : Pic C=O à 1637 cm^{-1} confirmé par Singh and al. (2021) dans *J. Ethnopharmacol.*

- Peak at 3272.94 cm^{-1}

Hydroxyl (OH) bonds of phenolic compounds, alcohols, or carboxylic acids. Hydroxyl groups (3272 cm^{-1}): Confirm the presence of polyphenols and/or polysaccharides, consistent with the literature on fenugreek (Patel and al., 2020).

Silverstein and al. (2014) report that polyphenol-rich plant extracts exhibit a broad OH band around 3200–3400 cm^{-1}.

- Peak at 1044 cm^{-1}:

Trigonelline: Vibrations cycles 500–600 cm^{-1} (Kumar and al., 2022, Phytochem. Anal).

The IR spectrum is consistent with the known composition of fenugreek: Polyphenols (broad OH band). Saponins (C=O at 1637 cm^{-1}). Polysaccharides (C-O-C at 1044 cm^{-1}).

- Peak at 554, 36 cm^{-1}:

The out-of-plane bend represents the presence of a C-H aromatic range stretch.

V.2.5 Determination of antioxydant Activity:

- After 30min of incubation, we read the samples in Spectrophotometer visible UV.

Figure 58: DPPH incubation of fenugreek at different concentration.

Table 12: DPPH Absorbance.

Concentration (%)	Absorbance (517 nm)
10	0.454
25	0.422
50	0.381
100	0.346
Control	0.4931

❖ **Control:** ethanol + DPPH.

The evaluation of antioxidant activity using the DPPH method is expressed as percentage according to the following formula:

$$\% \textbf{ Inhibition} = \frac{\textbf{Abs Controle} - \textbf{Abs Extrai}}{\textbf{Abs Controle}} \times \textbf{100... V (1)}$$

Table 13: Percentage of results of DPPH inhibition for different concentrations

Concentration (%)	Inhibition (%)
10	7.93%
25	14.42%
50	22.73%
100	29.83%

Interpretation:

The results of the free radical scavenging assay using the DPPH extract demonstrated remarkable antioxidant activity, with a clear concentration-dependent response pattern. Absorbance values, after subtracting the reference solution sample (0.4931), showed negligible free radical scavenging capacity at all tested concentrations. At 10µg/mL, the extract showed 7.93 % inhibition, increasing to 29.83% at 100µg/mL, indicating a lack of efficacy. The persistence of the purple color of the DPPH solution suggests several possible explanations that deserve careful consideration. First, the apparent variability may indicate delayed reaction kinetics, as some antioxidant compounds—particularly high-molecular-weight polyphenols or complex saponins—require long incubation periods (up to two to four hours) to completely reduce the DPPH radical or to a higher concentration.

V.2.6 Discussion:

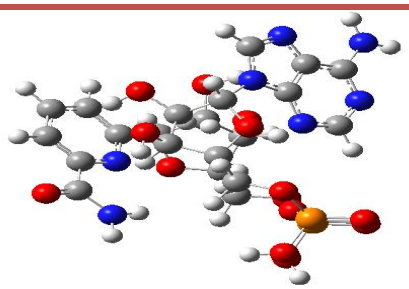
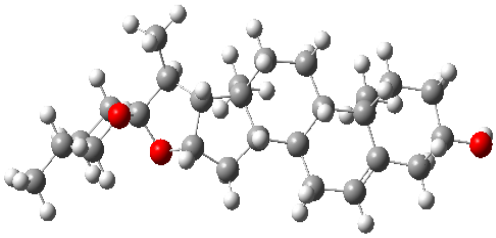
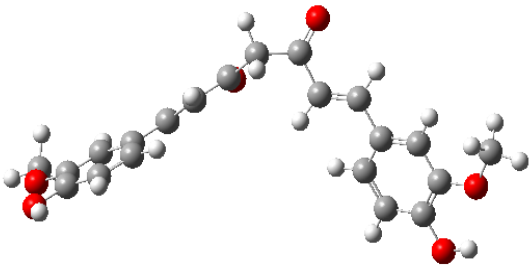
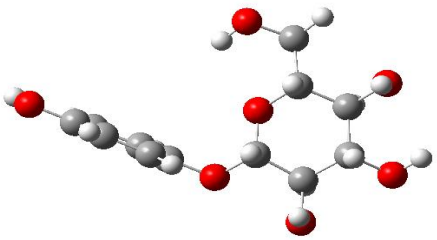
- The results of this study provide compelling evidence of the therapeutic potential of fenugreek (Trigonella foenum-graecum) as a rich source of bioactive compounds with antioxidant properties. The ethanolic extraction rate of 12–20% is consistent with previous reports (Patel et al., 2020). UV–Vis spectroscopy revealed distinct absorption patterns indicating the presence peaks at **200–300 nm** suggest polyphenols/flavonoids.

- The Fourier transforms infrared (FTIR) analysis results were invaluable, identifying the key functional groups associated with the bioactive components of fenugreek. The broad O–H stretch at 3272 cm^{-1} and the C=O peak at 1637 cm^{-1} strongly indicate the presence of phenolic compounds and saponins, respectively, confirming the findings of Singh and al. (2021). These findings gain additional significance when considering the reported anticancer properties of fenugreek. Previous studies have shown that identified compounds, particularly diosgenin and flavonoids, target several cancer hallmarks, including proliferation, angiogenesis, and metastasis (Srinivasan, 2020). The current findings provide a phytochemical basis for these observed biological activities.
- Furthermore, although the DPPH assay is widely used, it represents only one mechanism of antioxidant activity. The discrepancy between the strong DPPH-scavenging activity and the minimal color change warrants further investigation. This phenomenon could be explained by several factors: (1) the presence of colored compounds that mask the color change, (2) unconventional antioxidant mechanisms that do not directly involve free radical neutralization, or (3) the formation of intermediates that maintain the purple color. This finding suggests that visual assessment alone may not reliably indicate the antioxidant activity of synthesized plant extracts. This study contributes significantly to the growing body of evidence supporting the potential of fenugreek in preventive and therapeutic applications, particularly in the context of an integrative approach to the management of cancer and oxidative stress-related disorders.

V.3 Molecular docking results:

V.3.1 Ligand& Target preparation results:

Table 14: Molecular plants in Gaussian view.

Plants	Molecules	Gaussian output molecules
Fenugreek	Fenugreekine	
	Diosgenin	
Curcuma	Curcumin	
Saltbush	Arbutin	

V.3.1.a Density Theory Calculation:

Four potential breast cancer therapeutics: fenugreek, diosgenin, curcumin, and arbutin, were optimized using dynamic differential finite-dimensional (DFT) analysis at the B3LYP/6-31G (2d,p) level. Frontier molecular orbital (FMO) analysis was used to evaluate how a molecule interacts with other species. The frontier orbital consists of the HOMO, the highest occupied molecular orbital, and the LUMO, the lowest unoccupied molecular orbital. The HOMO denotes an electrophilic nucleus, while the LUMO represents an electron-loving acceptor.

The various electronic properties of each of these molecules were calculated using this method. Table 15 summarizes the calculated dipole moment (α) and energies of the frontier molecular orbitals. Figure 59 shows the energies of the frontier molecular orbitals (HOMO and LUMO) and their corresponding 3D diagrams for each anticancer molecule, and the inhibition constant for our molecules is shown in Table 16.

Table 15: Energy calculation of molecules

N°	E_{LUMO}	E_{HOMO}	μ	H	X	ω	S	ΔE_{gap}
M1	-0.18245	-0.28484	-0.2336	0.0511	0.2336	0.5339	9.7847	-0.10059
M2	-0.13521	-0.33242	-0.2338	0.0986	0.2338	0.2772	5.0713	-0.19721
M3	-0.08433	-0.21895	-0.1516	0.0673	0.1516	0.1707	7.4291	-0.13462
M4	-0.03331	-0.22849	-0.1309	0.0976	0.1309	0.0877	5.1229	-0.19518

M1: Fenugreekine.

M2: Diosgenin.

M3: Curcumin.

M4: Arbutin.

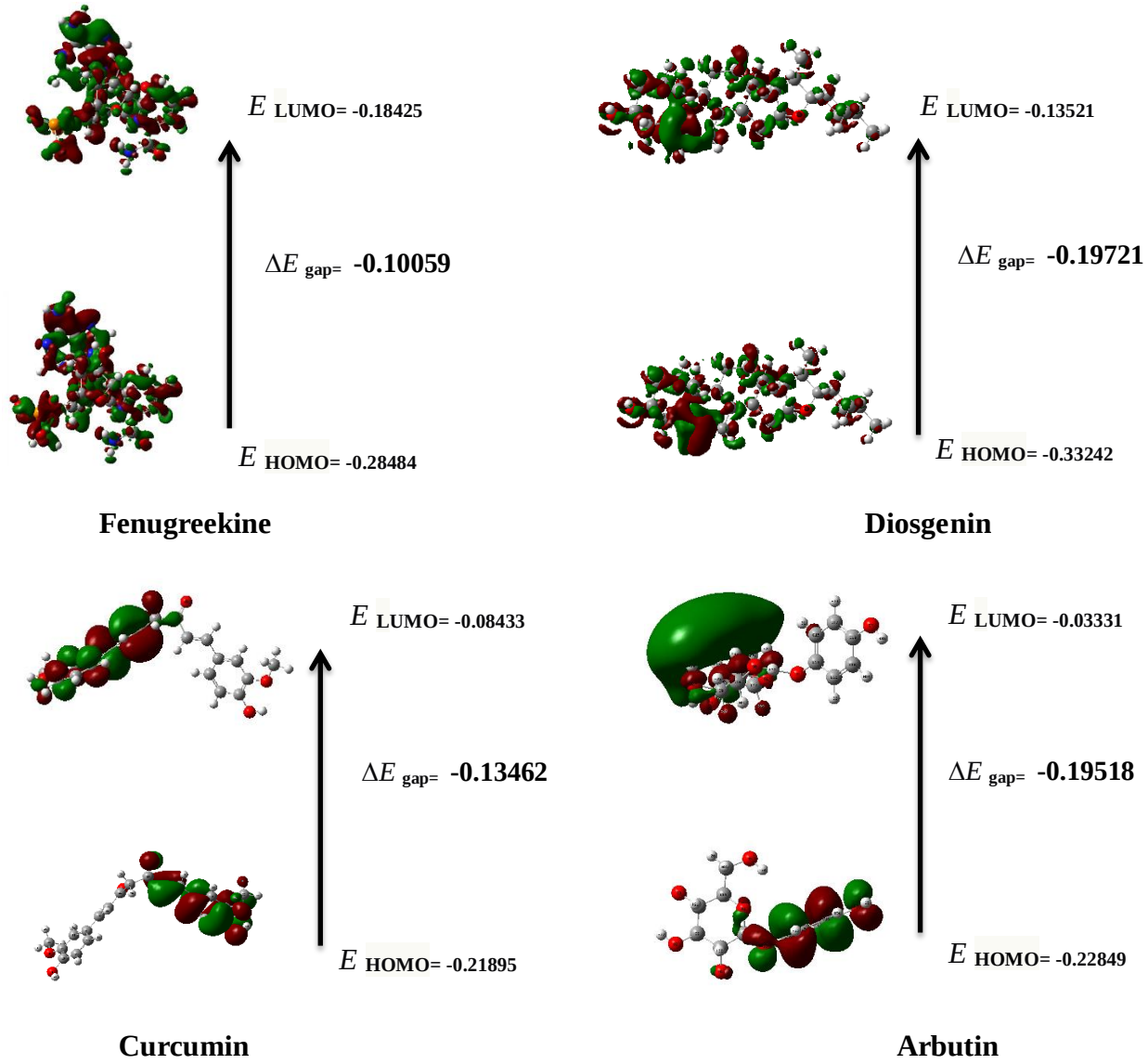


Figure 59: Frontier molecular orbitals and their energy gaps of molecules.

Interpretation:

The time required for complete optimization of each compound varied depending on its molecular complexity and structural flexibility. Fenugreekine required the longest computational time (10 hours, 44 minutes, and 53 seconds), reflecting its relatively complex structure. Diosgenin also showed a longer optimization time (3 hours, 38 minutes, and 27.0 seconds). Curcumin took (1 hour, 53 minutes, and 10.0 seconds). In contrast, arbutin showed a much shorter optimization time (54 minutes and 14 seconds), likely due to its more compact structure.

The quantum chemical analysis of the four molecules (M1–M4) using DFT reveals distinct electronic and reactivity profiles. M1 stands out as the most electrophilic and electron-accepting compound, with the lowest LUMO energy, highest chemical hardness, and lowest softness, indicating high stability and low reactivity. In contrast, M3 shows greater chemical reactivity due to its narrow HOMO-LUMO gap and higher softness, making it the most reactive among the four. M2 has the deepest HOMO, suggesting strong resistance to oxidation, while M4 exhibits the least electrophilicity and lowest chemical potential, implying weaker electron affinity and lower electron-donating ability. Overall, M1 is the most stable and electron-seeking, whereas M3 is the most chemically active.

Table 16: Inhibition Constant.

Composants	**KI (µM)**
Fenugreek	1.1850
Diosgenin	1.3949
Curcumin	1.2551
Arbutin	1.3902

Interpretation:

The inhibition constant (KI) further supports these findings, with fenugreek exhibiting the lowest KI value (1.1850 µM), meaning it required the lowest concentration to inhibit the target compared to the other compounds. Interestingly, curcumin, despite its weaker binding affinity compared to diosgenin, exhibited a slightly better KI (1.2551 µM vs. 1.3949 µM), suggesting differences in binding kinetics.

Table 17: Ligand PDB and PDBQT by AutoDock Vina Tools.

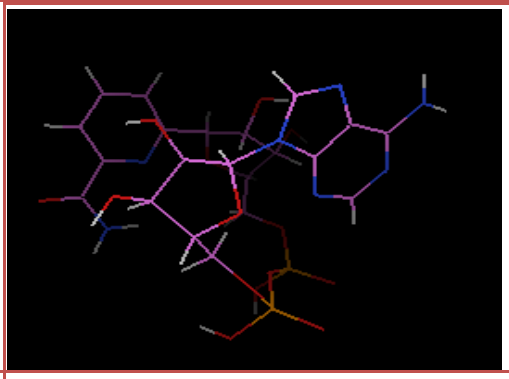
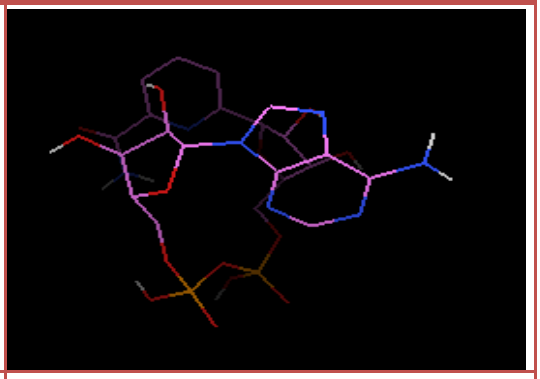
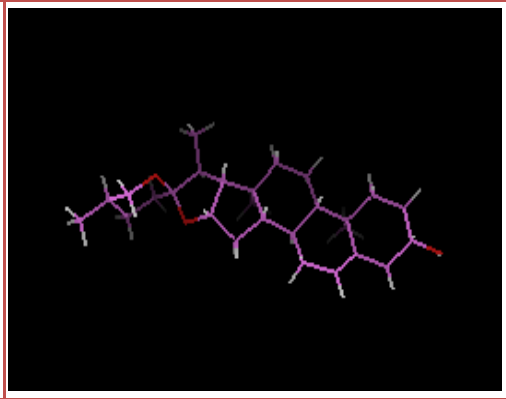
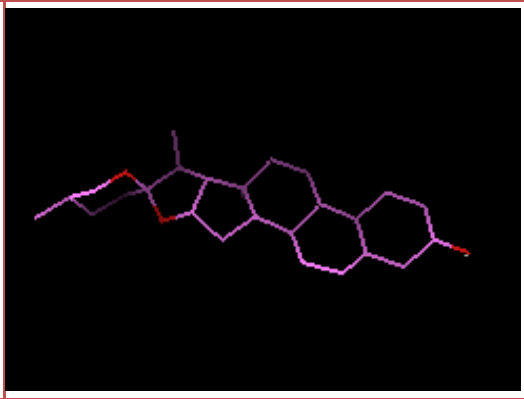
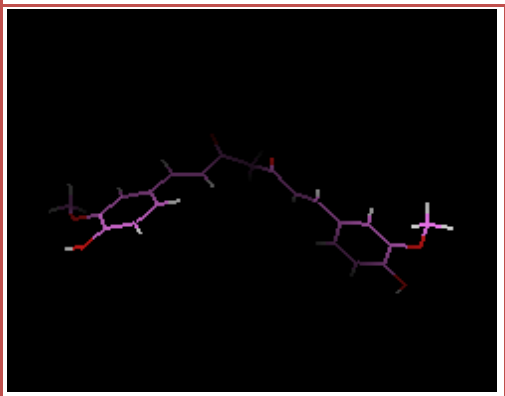
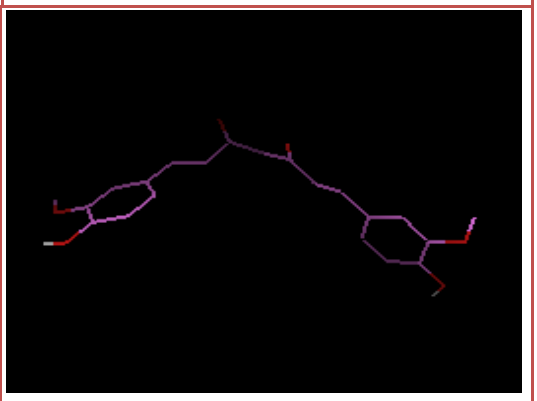
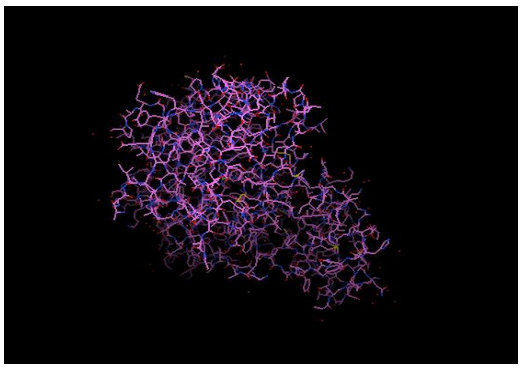
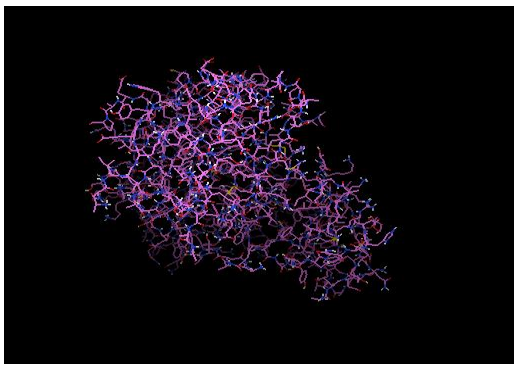
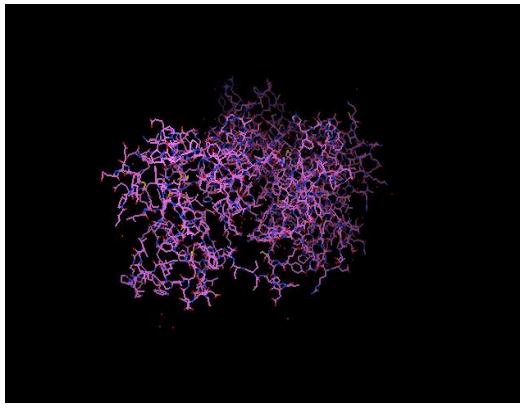
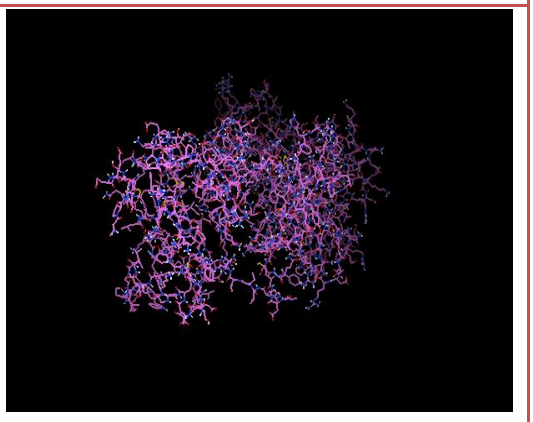
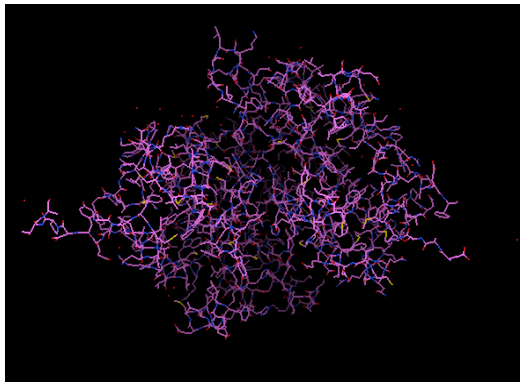
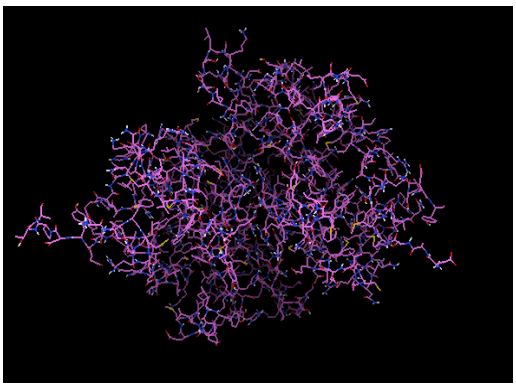
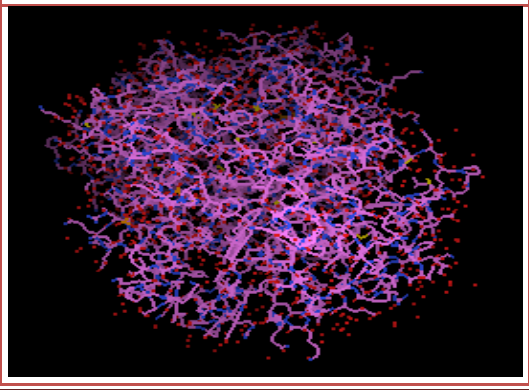
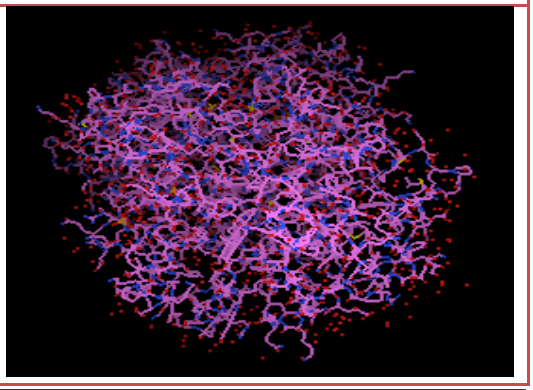
	Ligand PDB	Ligand PDBQT
Fenugreekine		
Diosgenin		
Curcumin		
Arbutin		

Table 18: Target PDB and PDBQT by AutoDock Vina Tools.

	Target pdb	Target pdbqt
AKT1		
CDK4		
HER2		
MCF-7		

V.3.2 In Silico Predicted ADMET Properties:

Computational pharmacology (ADMET) is a rapidly growing field involving the development of software to integrate and capture medical and biological data from around the world. Studying the ADMET properties of compounds, which can be performed using computer technology, is important for finding and creating new drugs. Table 19 shows the ADMET predictions. The results show that all compounds have an absorption value of approximately 90%, meaning that they are well absorbed by the human gut. The blood-brain barrier (BBB) protects the brain and maintains good health. Furthermore, it blocks the flow of hazardous chemicals, preventing them from easily entering the brain.. Importantly, the high toxicity rating ensures the safety of the molecules when used.

Table 19: Molecules ADMET prediction.

	Fenugreekine	**Diosgenin**	**Curcumin**	**Arbutin**
A				
Intestinal absorption	0.09	1	0.98	0.44
Oral Bioavailability	0.30	0.87	0.46	0.44
D				
BBB	0.50	0.84	0.34	0.37
VD	0.01 L/Kg	0.00 L/Kg	0.01 L/Kg	0.02 L/Kg
M				
Inhibition				
CYP1A2	No	No	0.44	No
CYP2C19	No	0.12	0.58	No
CYP2C9	No	No	0.57	No
CYP2D6	No	No	0.12	No
CYP3A4	No	0.16	0.75	No
Substrate				
CYP2C9	No	No	0.18	No
CYP2D6	No	0.15	No	No
CYP3A4	0.2	0.69	0.57	0.24
E				
Clearance (Hepatocyte)	8.26 uL/min/10 cells	77.60	75.97	68.94
Half Life	51.15 hr	16.88 hr	0.01 hr	0.00 hr
T				
Clinical Toxicity	0.47	0.08	0.06	0.05

V.3.3 In Silico Prediction Pharmacokinetic and Toxicological:

The effects of selected compounds on nuclear signaling pathways were predicted using ProTox-II. The SwissADME Web site identifies physicochemical, pharmacokinetic, drug-like, and related criteria for multiple molecules; therefore, SwissADME studies were conducted on molecules proven to have an effect against breast cancer.

Lipinski's Rule: Drug-like and non-drug-like substances can be categorized using the Lipinski rule of five. Drug-likeness criteria have been used to more rapidly determine a molecule's drug-like qualities when it comes to structural properties of compounds. Determine (Hydrogen bond acceptor, Hydrogen bond donor, TPSA, Bioavailability Score, etc.) is the main goal of the Lipinski Five Rule [174].

The rule specifies that, typically, an orally active drug should not have more than one violation of these criteria: a MW over 500 g/mol, a calculated logP (logP) over 5, more than five hydrogen bond donors, or more than ten hydrogen bond acceptors.

- ✓ Hydrogen donors: No more than 5. Although an increasing number of hydrogen donors may indicate the potential for hydrogen bonding, Lipinski's rule does not set a maximum limit for hydrogen donors, such as fenugreekine.
- ✓ Hydrogen acceptors: no more than 10 Similarly, Lipinski's rule does not set a maximum limit for hydrogen acceptors.

Among the studied compounds, diosgenin demonstrates the most favorable pharmacokinetic and safety profile, while fenugreekine shows poor absorption but low toxicity. Curcumin and arbutin balance moderate toxicity with good absorption, highlighting their potential for therapeutic application in cancer-related treatments. The results are shown in table 20 and 21.

Table 20: Applying Swiss ADME and Toxicity prediction of chemicals molecules.

Composants	Compounds CID	Swiss ADME	Protox	Lipenski Rule
Fenugreekine (C21H27N7O14P2)	444170	BBB HIA	Classe : 6 LD50%= 7000 Mg/kg	NO
Diosgenin ($C_{27}H_{42}O_3$)	99474	BBB	Classe : 6 LD50%=8000 Mg /kg	YES
Curcumin (C21H20O6)	969516	HIA	Classe : 4 LD50%= 2000 Mg/kg	YES
Arbutin (C12H16O7)	440936	HIA	Classe : 5 LD50%= 2500 Mg/kg	YES

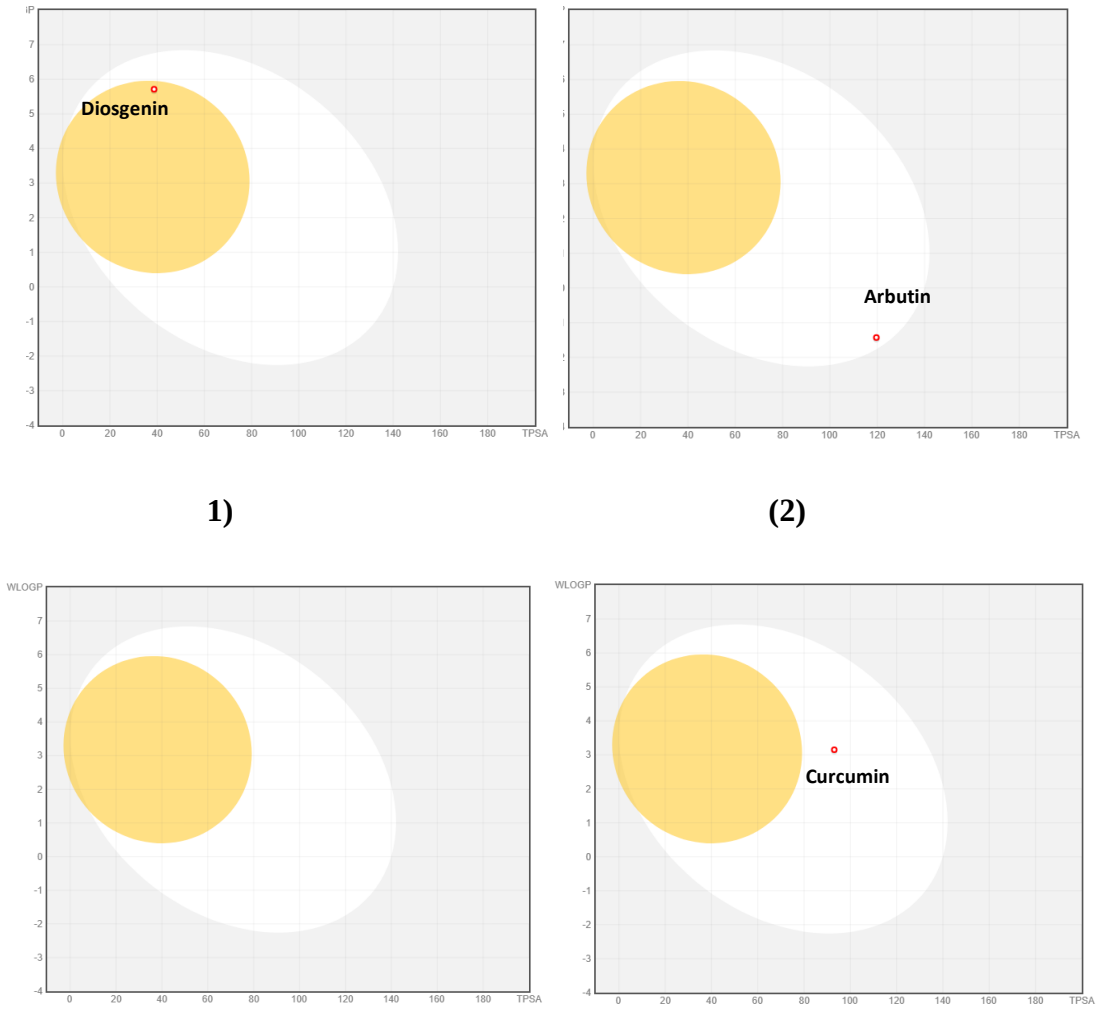
Table V.21: Lipinski's application for our molecules.

Composants	Mw	LogP	HA	HD	TPSA
Fenugreekine	663.43 g/mol	-3.77	18	8	346.89 Å²
Diosgenin	414.62 g/mol	5.02	3	1	38.69 Å²
Curcumin	368.38 g/mol	3.07	6	2	93.06 Å²
Arbutin	272.25 g/mol	1.64	7	5	119.61 Å²

HA: Hydrogen Acceptor (≤ 10).

HD: Hydrogen Donor (≤ 5).

Log p: water/ octanol partition coefficient (-2 ≤ LogP ≤ 5).



1) **(2)**

(3) **(4)**

Figure 60: Diagram of boiled egg shows in SwissADME of (1) Diosgenin (2) Arbutin (3) Fenugreekine (4) Curcumin.

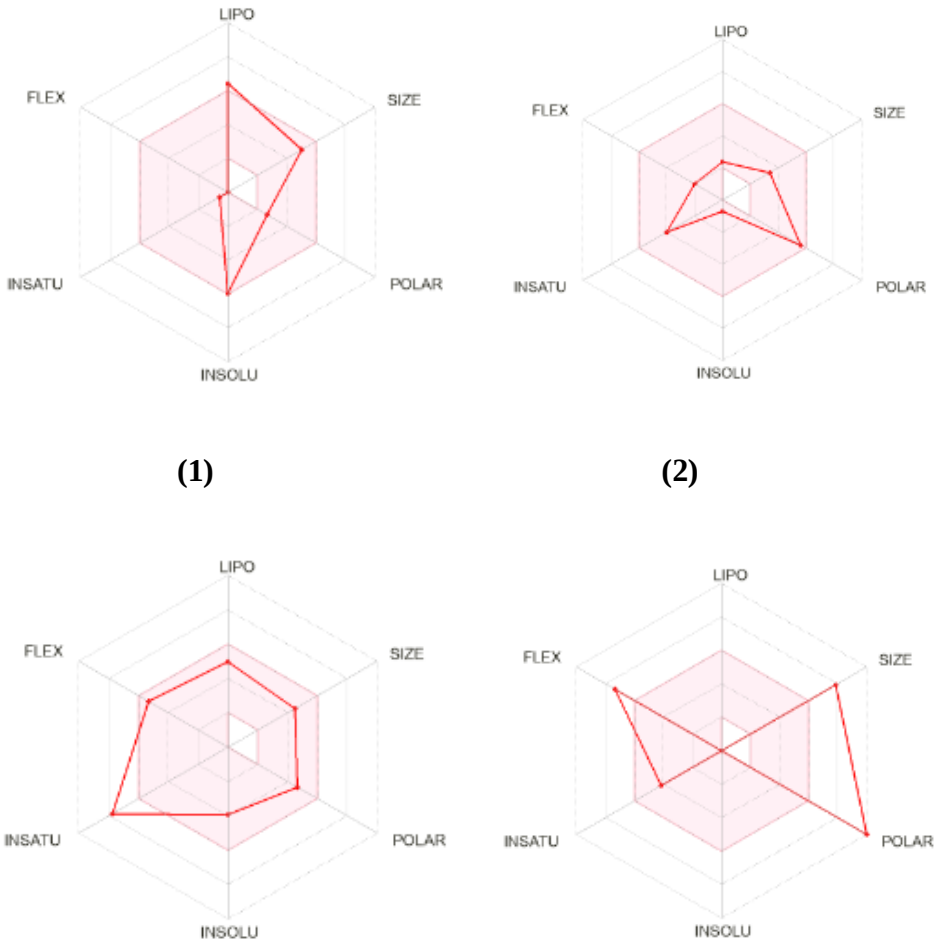


Figure 61: Radar charts generated by SwissADME with the chemical space of **(1)** diosgenin**, (2)** arbutin**, (3)** fenugreekine and curcumin **(4)** on the upper and lower limits of crucial features for the drug likeness.

Interpretation:

Based on ADMET analysis (table 19), diosgenin appears to be the most effective compound in terms of absorption, metabolic stability, and toxicity. Despite its good absorption, curcumin may require optimization in its formulation due to its rapid clearance. Fenugreekine, despite its long half-life, suffers from poor absorption. Arbutin, with its moderate pharmacokinetic properties and low toxicity, is an effective secondary candidate. In terms of oral bioavailability, diosgenin again outperformed the others (0.87), followed by

fenugreekine (0.30), while curcumin and arbutin had relatively low values (0.46 and 0.44, respectively).

Among the compounds, diosgenin shows the greatest membrane permeability due to its high lipophilicity and low polarity, whereas fenugreekine may face absorption challenges due to its high polarity and molecular weight. Curcumin and arbutin offer a more balanced physicochemical profile, favorable for drug development, with moderate size, polarity, and lipophilicity. Figure 61 is bioavailability radar, used to visually assess the drug-likeness of a molecule. It evaluates how well a compound fits within the optimal range for oral bioavailability by plotting six physicochemical properties such as the red line represents the actual properties of the compound. The pink area represents the optimal range for good oral bioavailability.

- **LIPO (Lipophilicity -0.7 to +5.0):** Reflects how well the compound can cross cell membranes. Too high or too low values may reduce absorption.
- **SIZE (150 to 500 g/mol):** Molecular weight influences absorption and distribution. Larger molecules often show poorer oral bioavailability.
- **POLAR (Polarity TPSA 20–130 Å²):** Polarity affects solubility and membrane permeability. Excessive polarity can hinder passive diffusion across membranes.
- **INSOLU (Insolubility Should not be too high):** Relates to aqueous solubility. Poor solubility limits bioavailability.
- **INSATU (Insaturation):** Reflects molecular complexity and 3D character. Low values suggest high aromaticity or flat molecules, which may affect bioavailability.
- **FLEX (Flexibility Number of rotatable bonds):** Highly flexible molecules may suffer from reduced membrane permeability and binding specificity.

V.3.4 Protein-ligands interaction and binding sites:

A molecular docking study was conducted to investigate the potential interactions between the protein complex and the ligand of the designed compounds. The AutoGrid tool from the AutoDock software suite was used to create grid maps showing the energy of contact between the target macromolecule and the ligands during the docking experiment.

To examine the binding efficacy of all the designed compounds, molecular docking studies were performed for the proteins AKT1, CDK4, MCF-7 and HER2 [defined in the Protein Database: 4EJN, 2W96, 4XO6 and 2IOK] with ligand molecules. All the designed compounds showed optimal binding scores ranging from -7.7 to -9.1 kcal/mol (see table 22).

Table 22: Vina score of interaction studies

Target pdbqt	**Ligand pdbqt**	**Grid box**	**Vina score (kcal/mol)**
CDK4	Fenugreekine	X= 5.956 Y=3.663 Z=59.279	-8.7 kcal/mol
AKT1	Diosgenin	X= 32.073 Y= 41.35 Z= 13.658	-9.1 kcal/mol
HER2	Curcumin	X= 25.882 Y= 23.256 Z= 40.434	-8.3 kcal/mol
MCF-7	Arbutin	X= 134.943 Y= 138.656 Z= 135.403	-7.7 kcal/mol

V.3.4.1 Discovery study:

Molecular docking output file was used to generate visual representations of protein–ligand interactions. BIOVIA Discovery Studio 2021 was employed to display these interactions, enabling analysis of bond types, bond distances, and the amino acid residues involved. By loading both the protein and ligand into the software's graphical interface, the interactions were visualized, with ligands positioned within the protein's receptor region. Discovery studies show the results of interactions between target and ligand-out-pdbqt (see figure 62, 63, 64 and 65).

❖ **Fenugreekine with CDK4 interaction:**

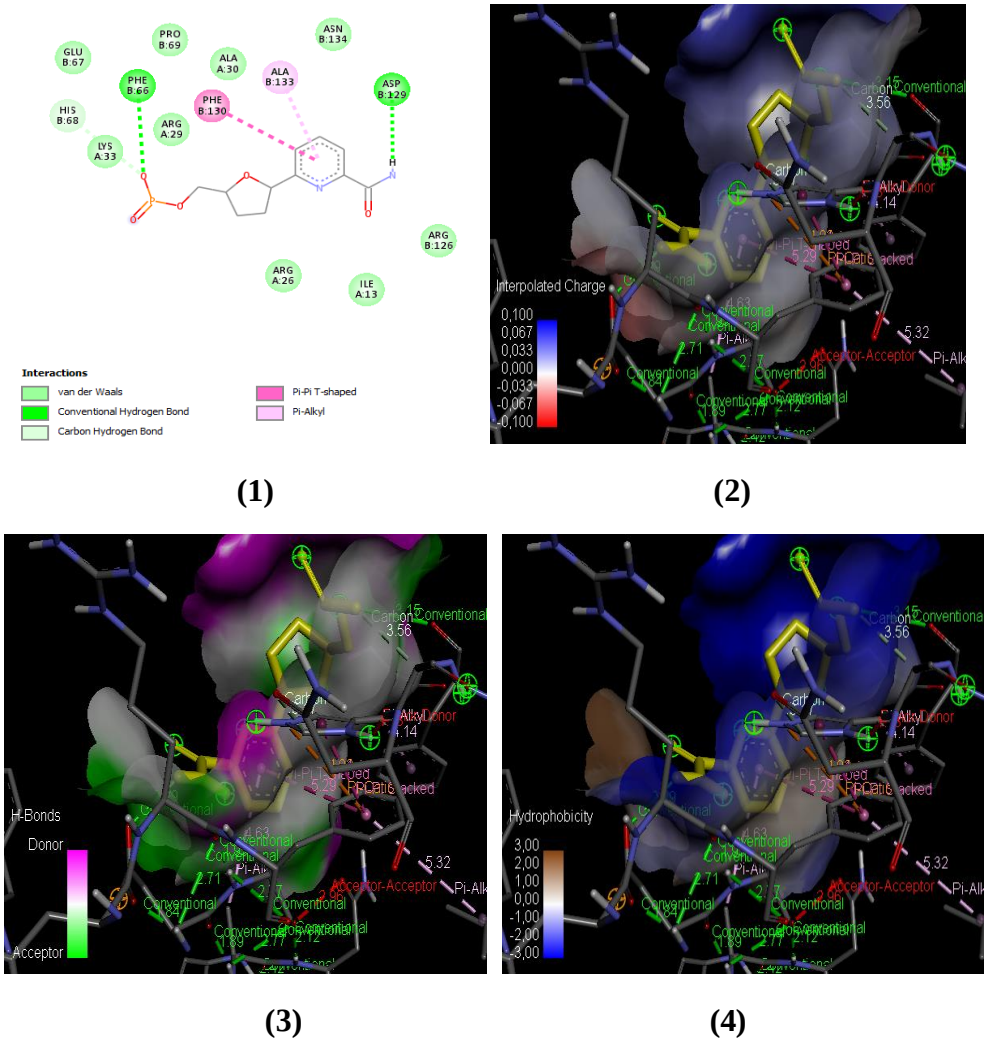
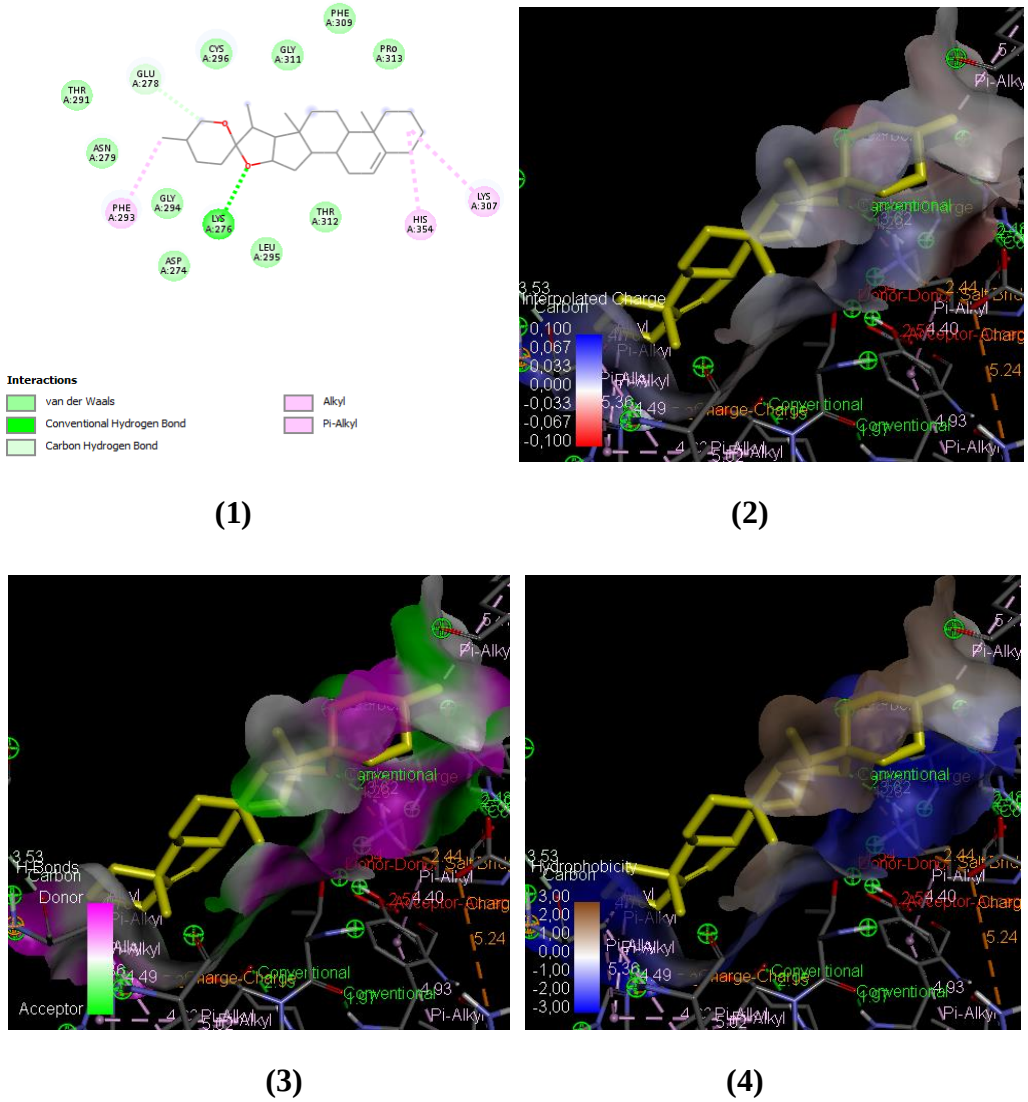


Figure 62: (1) 2D diagram of interaction CDK4 with Fenugreekine. **(2)** Interpolated charge **(3)** H-Bonds. **(4)** Hydrophobicity of interaction.

Interpretation:

Molecular docking analysis revealed that the fenugreekine forms stable interactions with CDK4 through multiple binding mechanisms. Key residues involved in the binding pocket include PHE66, HIS68, ARG126, and ASP229, which participate in hydrogen bonding (conventional and carbon-hydrogen bonding). Additionally, hydrophobic contacts with the alkyl chains enhance the stability of the complex.

❖ **Diosgenin with AKT1 interaction:**



(1) **(2)**

(3) **(4)**

Figure 63: (1) 2D diagram of interaction AKT1 with Diosgenin. **(2)** Interpolated charge **(3)** H-Bonds. **(4)** Hydrophobicity of interaction.

Interpretation:

Diosgenin forms a stable complex with AKT1 protein through several key interactions. It engages in a conventional hydrogen bond with LYS A:276, supporting strong binding. Additionally, carbon-carbon hydrogen bonds were observed with ASP A:274 and THR A:291, enhancing specificity. The molecule also forms bi-alkyl and alkyl interactions with hydrophobic residues such as LYS A:307, HIS A:354, and PHE A:293, contributing to the hydrophobic stabilization of the complex.

❖ **Curcumin with HER2 interaction:**

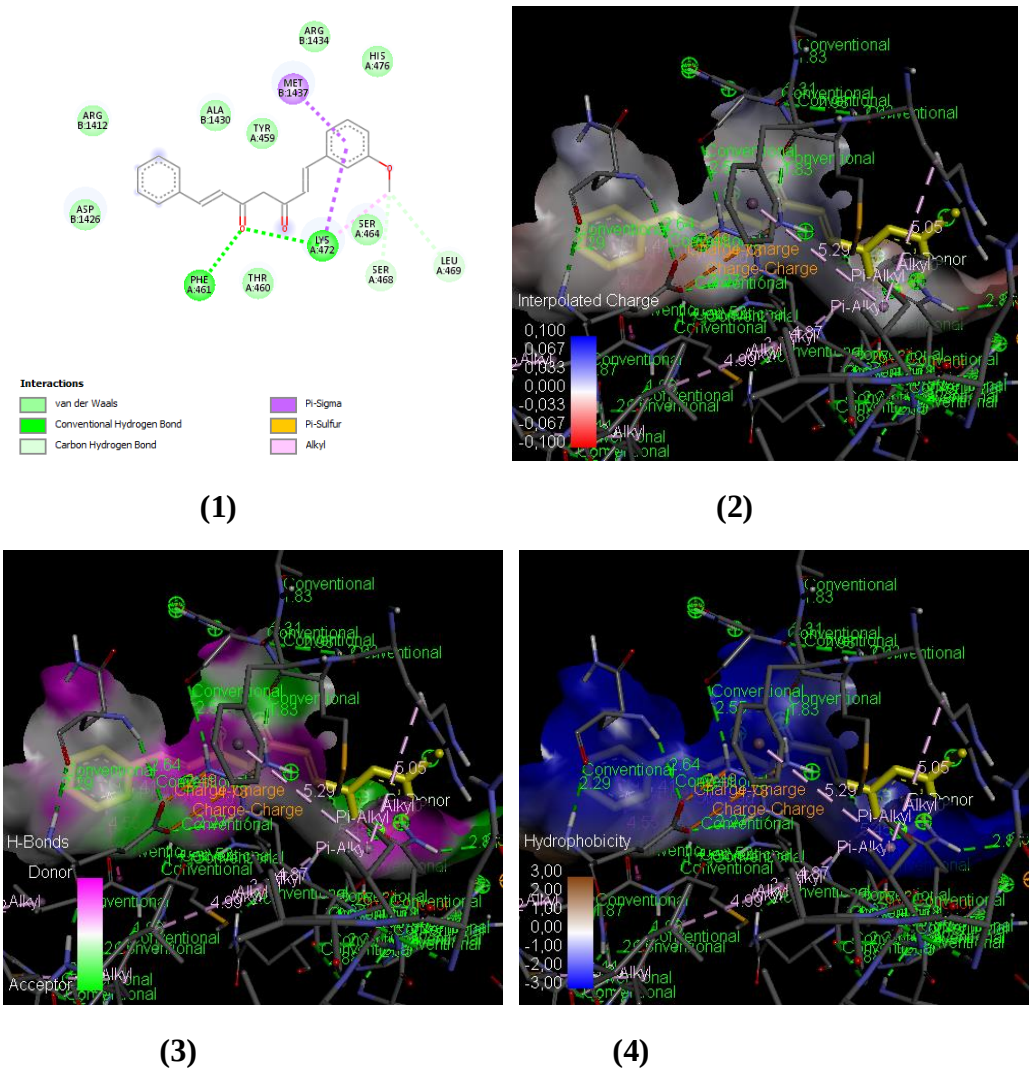
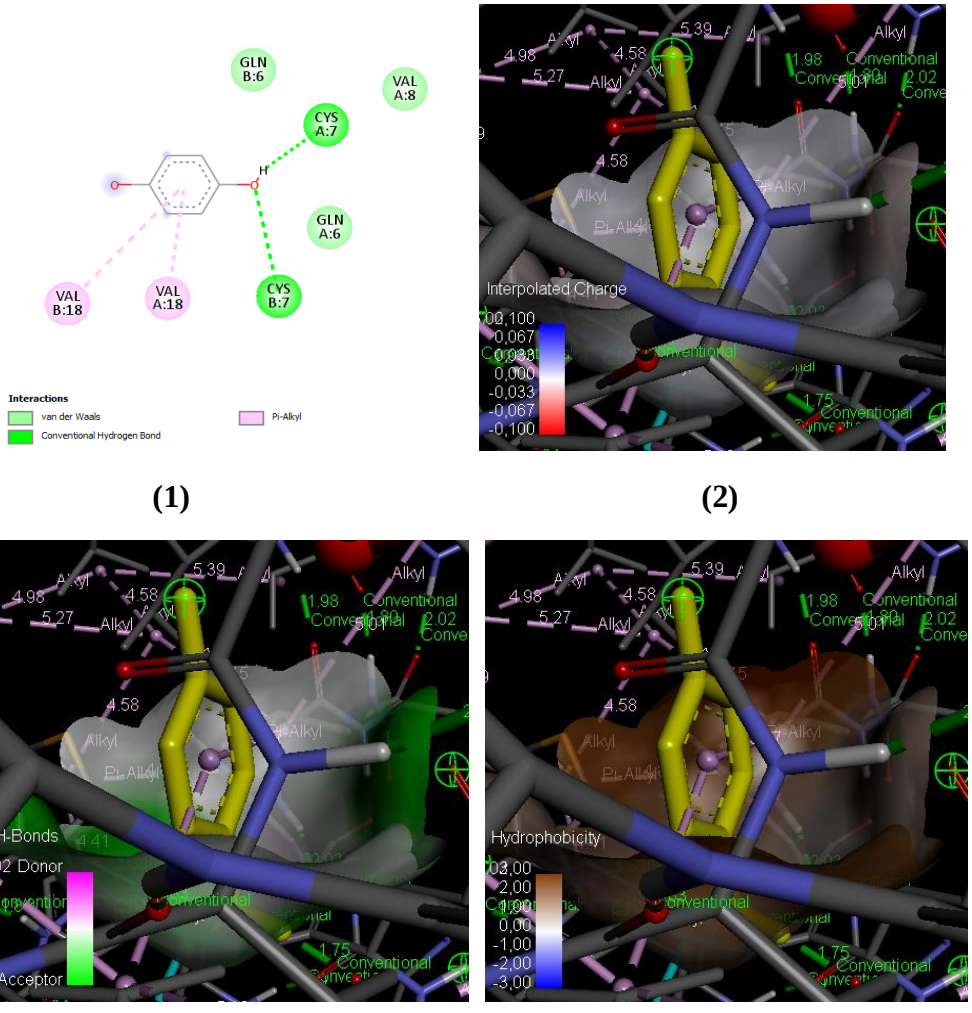


Figure 64: (1) 2D diagram of interaction HER2 with Curcumin. **(2)** Interpolated charge **(3)** H-Bonds. **(4)** Hydrophobicity of interaction.

Interpretation:

The molecular docking analysis of curcumin with the HER2 (2IOK) protein demonstrates a stable and well-fitted interaction within the active binding pocket. Curcumin forms several key non-covalent interactions that contribute to its binding affinity. Notably, conventional hydrogen bonds are observed with residues such as Lys472 and Ser464, supporting strong polar interactions. Hydrophobic interactions, such as Pi-Sigma and Alkyl contacts with Met1437 and Leu469, further enhance the ligand's binding within the protein's hydrophobic core.

❖ **Arbutin with MCF-7 interaction:**



(1) **(2)**

(3) **(4)**

Figure 65: (1) 2D diagram of interaction MCF-7 with Arbutin. **(2)** Interpolated charge **(3)** H-Bonds. **(4)** Hydrophobicity of interaction.

Interpretation:

Molecular docking results of arbutin with the MCF-7 breast cancer protein (4XO6) reveal a modest but statistically significant interaction pattern. Arbutin interacts primarily through conventional hydrogen bonding with Cys A:7 and Cys B:7, which are essential for stabilizing the ligand within the active site. In addition, bi-alkyl interactions with Val A:18 and Val B:18 contribute to hydrophobic stabilization. Several van der Waals interactions were observed with residues such as Val A:8, Gln A:6, Gln B:6, and Val B:18, suggesting favorable but nonspecific contacts within the binding pocket.

V.3.4.2 CB-Dock study:

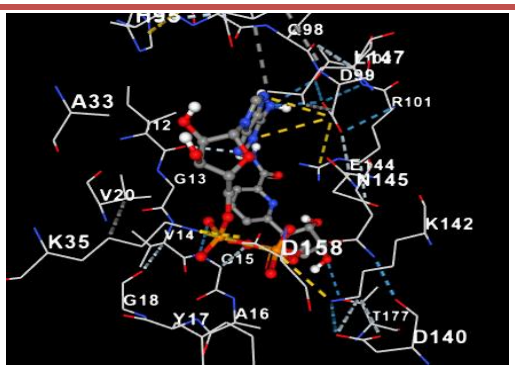
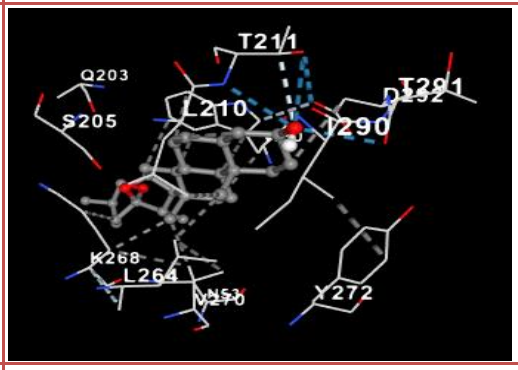
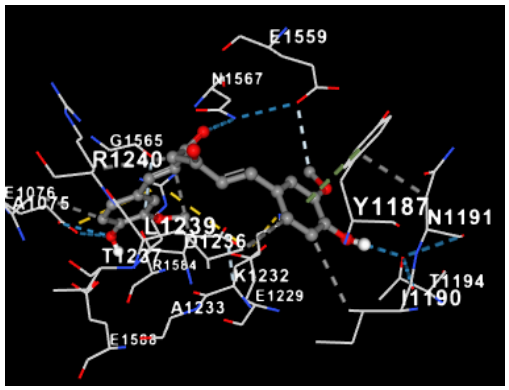
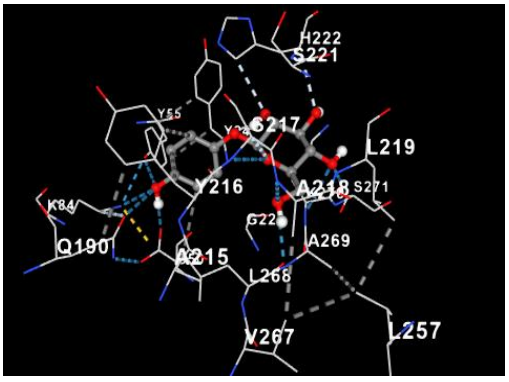
The results showed that all compounds reacted favourably; with overall scores ranging from -7 to -11.4 kcal/mol as shown in Table 23 below and the results for amino acids interaction in Table 24.

Table 23: CB-dock results.

		Fenugreekine	**Curcumin**	**Diosgenin**	**Arbutin**
CB-dock Score kcal/mol	CDK4 (2W96)	-9.3	-8.4	-9.7	-7.3
	HER2 (2IOK)	-9.2	-8.6	-9	-7
	AKT1 (4EJN)	-11.4	-9.5	-10.3	-7.8
	MCF-7 (4XO6)	-10.9	-9.7	-11	-8.4

• Protein-ligand interaction (PLI) plays an important role in the discovery of new drugs, providing vital information about drug binding to target pathogen proteins. Key breast cancer proteins were studied, and molecular docking analysis revealed critical interactions between bioactive compounds and their target proteins, providing insights into their potential inhibitory mechanisms against breast cancer pathways. The observed amino acid interactions explain the binding affinity and indicate the structural basis for the compounds' biological activities.

Table 24: The interactions Target& Ligand using CB dock.

Target& Ligand	CB-Dock Interaction	Amino Acid
Fenugreekine with CDK4 (2W96)		**D140 ; K142 ; R101 ; A33, D158 ; V96 ; K35 ; N145, V20 ; T102 ; E144 ; H95, G13 ; L147 ; Y17 ; F93 ; I12 ; G18.**
Diosgenin with AKT1 (4EJN)		**T211 ; Q203 ; S205 ; Y272 ; K268 ; L264 ; I290 ; L210 ; D292 ; N151 ; V270 ; T291.**
Curcumin with HER2 (2IOK)		**E1559 ; G1565 ; Y1187 N1191; I1190 ; E1229; A1233 F1556 ; K1232 ; R1240 ; L1239 ; E1076.**
Arbutin with MCF-7 (4XO6)		**S217 ; Q190 ; V267 ; D50 K84 ; L257 ; H222 ; G22 Y216 ; S221; A269; K270 V216; Y55; L219;**

Interpretation:

- ✓ For fenugreekine compounds binding to CDK4 (2W96), the extensive interactions with catalytic residues (D140, K142, R101) and hydrophobic pocket residues (V96, F93, and I12) are particularly noteworthy. These interactions likely interfere with ATP binding and kinase activity, as D140 and K142 are known to participate in phosphate transfer. The presence of multiple hydrogen bonds with polar residues (N145, T102) and hydrophobic contacts with conserved residues (Y17, L147) suggests stable binding within the active site.

- ✓ Diosgenin interaction with AKT1 (4EJN) showed strategic positioning near the PH domain (T211, Q203) and catalytic cleft (K268, D292). The hydrogen bond with T291 and hydrophobic contacts with L264/I290 may explain its predicted inhibition of AKT1 activation, as these residues are involved in the conformational changes required for kinase activation.

- ✓ The curcumin-HER2 (2IOK) complex demonstrated ionic interactions with critical HER2 residues (E1559, K1232) that participate in dimerization. The aromatic stacking with F1556 and hydrogen bonds with Y1187/N1191 suggest stabilization of HER2 in an inactive conformation, potentially inhibiting downstream signaling.

- ✓ Arbutin with the MCF-7 (4XO6) protein revealed significant interactions with key amino acid residues, suggesting potential anti-breast cancer activity. Arbutin formed strong hydrogen bonds with polar residues S217, Q190, and D50, which likely contribute to its stable binding within the protein's active site. Additionally, hydrophobic interactions with V267, L257, and L219 residues indicate favorable van der Waals contacts that enhance binding affinity. The presence of potential π-stacking between arbutin phenolic ring and Y55 further stabilizes the complex. These interactions are particularly noteworthy as they involve residues located in functionally important regions of the protein, potentially interfering with cancer cell proliferation pathways.

V.3.5 Comparison between the commercial treatment and our molecule:

We compared the molecular docking results of natural compounds, four standard anticancer drugs Lapatinib, Ibrance, Capivasertib, and Tamoxifen were analyzed for their binding affinities, molecular weights (Mw), and predicted toxicological properties using CB-Dock and ProTox-II.

Table 25: Drug In Silico.

Composants	**Mw**	**Protox**	**CB-Dock Score**	**Amino Acid**
Lapatinib (**$C_{29}H_{26}ClFN_4O_4S$**) With **HER2**	581.1 g/mol	Classe : 4 LD50%=1500 Mg/kg	-10.3 kcal/mol	W1383 ;L1354 ; D1351 ; G1521 ; F1404 ; L1349 ; L1536 ; M1388 ; H1524 ; V1418; A1350; G1420; M1343; N1348
Ibrance (C24H29N702) With **CDK4**	447.53g/mol	Classe : 5 LD50%=2500 Mg/kg	-9.1 kcal/mol	D140 ; T177 ; V72 ; K142 ; R101 ; A33, D158 ; V96 ; K35 ; N145, V20 ; T102 ; E144 ; H95, G13 ; L147 ; A157 ; F93 ; I12
Capivasertib (C21H25ClN6O2) With **AKT1**	428.9 g/mol	Classe : 4 LD50%=400 Mg/kg	-10.1 kcal/mol	T211 ; Q203 ; S205 ; N54 ; L210 ; F293 ; D292 ; R273 ; T211 ; V270 ; K268 ; Q79 ; G294 ; N204 ; I84 ; Y272 ; W80 ; K179
Tamoxifen (**$C_{26}H_{29}NO$**) With **MCF-7**	371.5 g/mol	Classe : 5 LD50%=2500 Mg/kg	-8.9 kcal/mol	H222 ; Y24 ; Y55 ; F311 ; Q190 ; N167 ; I129 ; S166 ; V126 ; W86; E224; H117; L308; A27; V54; Y216

V.3.6 Discussion:

- ✓ The docking affinity of fenugreekine, curcumin, diosgenin, and arbutin against key breast cancer-related protein targets was evaluated using CB-Dock. The docking scores, expressed in kcal/mol, indicate the binding strength of each ligand to its respective receptor, with more negative values representing stronger binding affinities. Fenugreekine showed the highest affinity for AKT1 (PDB ID: 4EJN) with a score of –11.4 kcal/mol, followed by strong interactions with MCF-7 (–10.9 kcal/mol), CDK4 (–9.3 kcal/mol), and HER2 (–9.2 kcal/mol). Diosgenin also demonstrated strong binding, particularly with MCF-7 (–11.0 kcal/mol) and AKT1 (–10.3 kcal/mol), indicating its potential anticancer properties. Curcumin exhibited consistent, though slightly lower, affinities across all targets, with the best score against MCF-7 (–9.7 kcal/mol). Arbutin displayed the weakest interactions overall, with scores ranging from –7.0 to –8.4 kcal/mol across all targets. These results suggest that fenugreekine and diosgenin are the most promising candidates for further investigation due to their superior binding affinities with crucial breast cancer targets.

- ✓ Comparison Between the Reference Drugs and our Compounds Based on Binding Energy and Amino Acid Interactions. Curcumin, with a binding energy of -8.6 kcal/mol, showed relatively strong interactions with HER2 target residues such as W1383, L1354, D1351, and H1524, mimicking the binding pattern of lapatinib (-10.3 kcal/mol), a clinically approved HER2 inhibitor. Fenugreekine showed a stronger binding affinity (-9.3 kcal/mol) compared to Ibrance (-9.1 kcal/mol), targeting key residues of the CDK4 binding pocket, including D140, K142, R101, and D158. These results suggest that fenugreek may be a potential CDK4/6 inhibitor. Meanwhile, diosgenin matched the binding energy of capevercetib (-10.3 kcal/mol), an AKT pathway inhibitor, with specific interactions with residues T211, K268, D292, and Y272 within the AKT1 active site, supporting its high affinity and potential as an AKT1-targeting agent. Finally, arbutin exhibited slightly weaker binding (-8.4 kcal/mol) compared to tamoxifen (-8.9 kcal/mol), yet maintained favorable interactions with estrogen receptor residues such as H222, Y24, F311, and E224, suggesting its potential role as a selective modulator of estrogenic activity.

V.4 Quantitative structure–activity relationship results:

QSAR modeling is a fundamental tool in computational medicinal chemistry, aiming to establish mathematical relationships between the structural properties of chemical compounds and their biological activities. In this context, QSAR models have been developed to predict the efficacy of selected compounds against protein targets associated with breast cancer. These models aim to establish accurate predictive relationships between the structural properties of tested chemical compounds (represented by a set of molecular descriptors) and their biological activity, expressed as pIC_{50} values, against breast cancer cells. Model performance was evaluated using multiple statistical indices, such as the coefficient of determination (R^2), the predictive coefficient of determination (Q^2), and the root mean square error (RMSE), to ensure the reliability and predictive quality of the models. The results demonstrated that some models exhibit high predictive power, reflecting a logical link between the chemical structure of the compounds and their target biological activity. This supports their use as an effective tool in the design and development of potential new drugs for the treatment of breast cancer.

V.4.1 MLR model for pIC50 breast cancer:

- **MLR equation:**

PIC50 = 25,9661(+/-2,7573) +0,5103(+/-0,1325) Chi2v +0,0016(+/-0,0005) Chi3v

- 2,2567(+/-0,2635) Chi4n -0,0912(+/-0,0303) SlogP_VSA5

-1,0689(+/-0,5084) Chi4v.

V.4.1.1 Feature selection of descriptors in QSAR modeling:

Feature selection is a critical step in quantitative structure-activity relationship (QSAR) modeling to improve model performance, reduce dimensionality, and eliminate irrelevant or redundant information. In this study, molecular descriptors were initially calculated using chemistry informatics tools, capturing various structural, electronic, and physicochemical properties of compounds. To identify the most relevant descriptors for biological activity, a combination of statistical and machine learning-based methods, such as

correlation analysis, variance thresholding, and iterative feature elimination, was used. Only descriptors that significantly contributed to the model's predictive ability were retained.

V.4.1.2 Selection of representative descriptors:

In this research, we relied on the QSAR study, noting that chemical descriptors fall specifically within the QSPR field because the QSAR database depends on the biological activity, unlike QSPR, which depends on the physicochemical property of the molecule.

The selection was performed using MLR model, which resulted in the identification of the 5 most representative descriptors: Chi2v, Chi3v, Chi4n, SlogP_VSA5 and Chi4v (see table 26).

Table 26: Relative Importance of Selected Descriptors for the MLR Model.

Descriptors	Definition	Reference
Chi2v	Valence molecular connectivity index of order 2; reflects molecular branching considering valence state of atoms.	Kier & Hall, 1986
Chi3v	Valence molecular connectivity index of order 3; measures molecular complexity and branching at a longer atomic path length (3 bonds).	Kier & Hall, 1986
Chi4n	Non-valence molecular connectivity index of order 4; topological descriptor related to the molecule's shape and branching without considering valence.	Kier & Hall, 1986
SlogP_VSA5	Surface area (VSA) descriptor in a specific logP (octanol–water partition coefficient) range; quantifies hydrophobic surface contributions.	Wildman & Crippen, 1999
Chi4v	Valence molecular connectivity index of order 4; measures atomic valence contributions at a four-bond distance.	Kier & Hall, 1986

Table 27: Validation Parameters of the MLR Model

Validation Criterion	MLR model
R^2_ train	0,890
R^2_test	0,930
Q^2	0,873
RMSE_train	0,158
RMSE_test	0,163

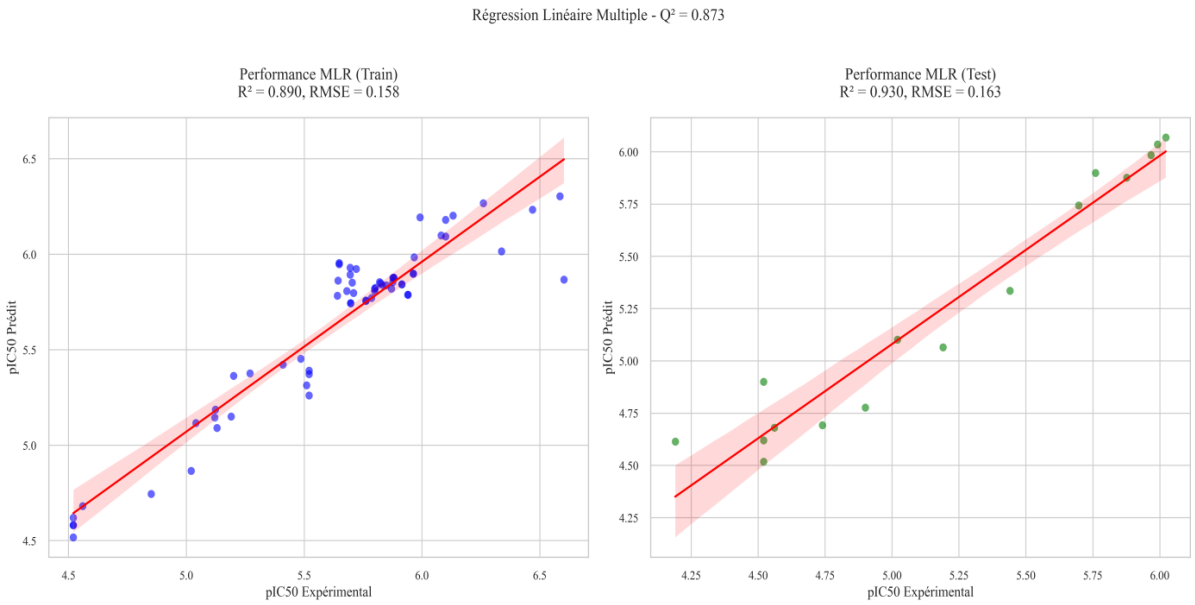


Figure 66: MLR performance QSAR results.

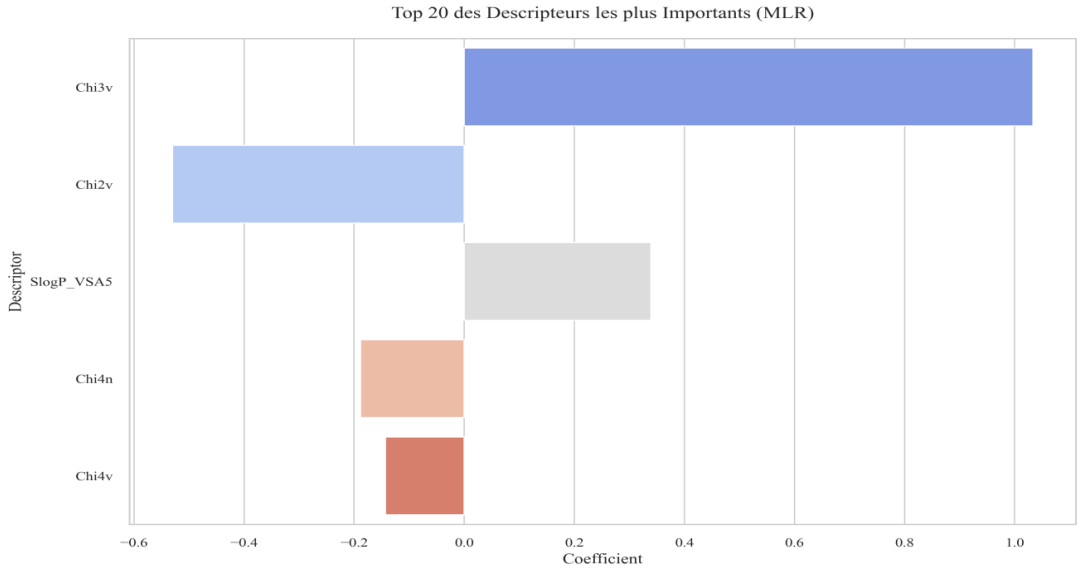


Figure 67: MLR feature_importance.

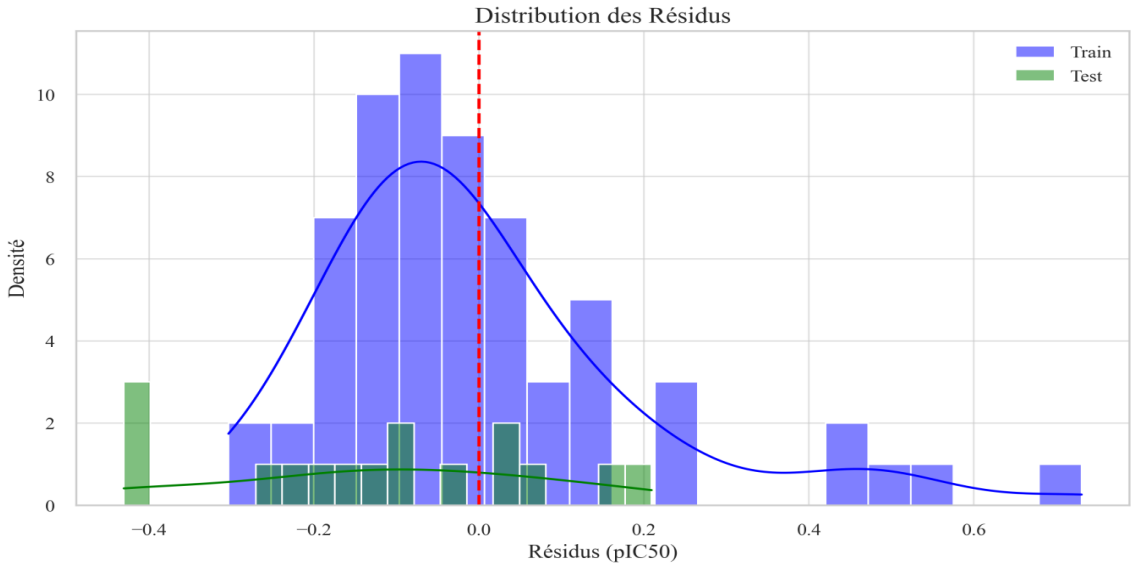


Figure 68: Distribution of residues.

- After collecting the equation, we retrieve the pIC50 predict (see table 9) and plotted the curve shown in the figure.

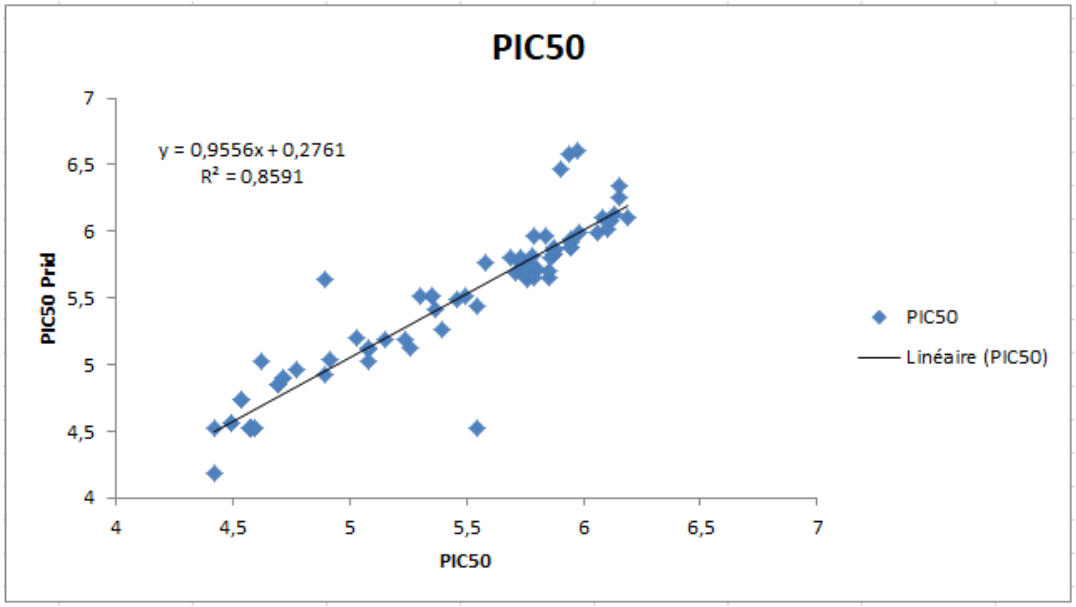


Figure 69: MLR Regression plot.

V.4.2 SVR model for pIC50 breast cancer:

Table 28: Validation Parameter of SVR model

Validation Criterion	SVR model
R^2_ train	0,875
R^2_test	0,912
Q^2	0,859
RMSE_train	0,169
RMSE_test	0,183

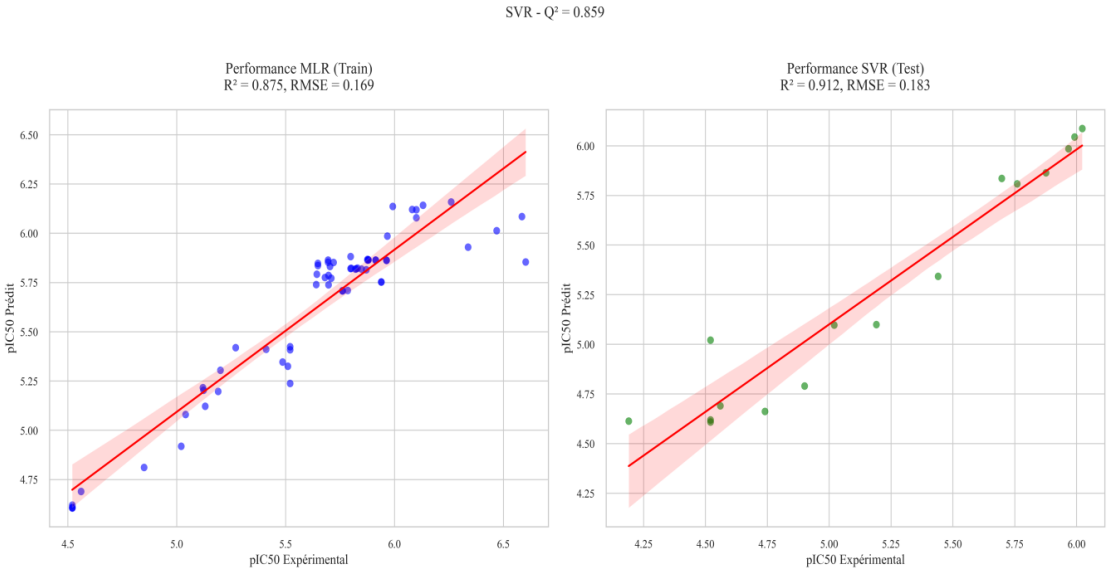


Figure 70: SVR performance QSAR results.

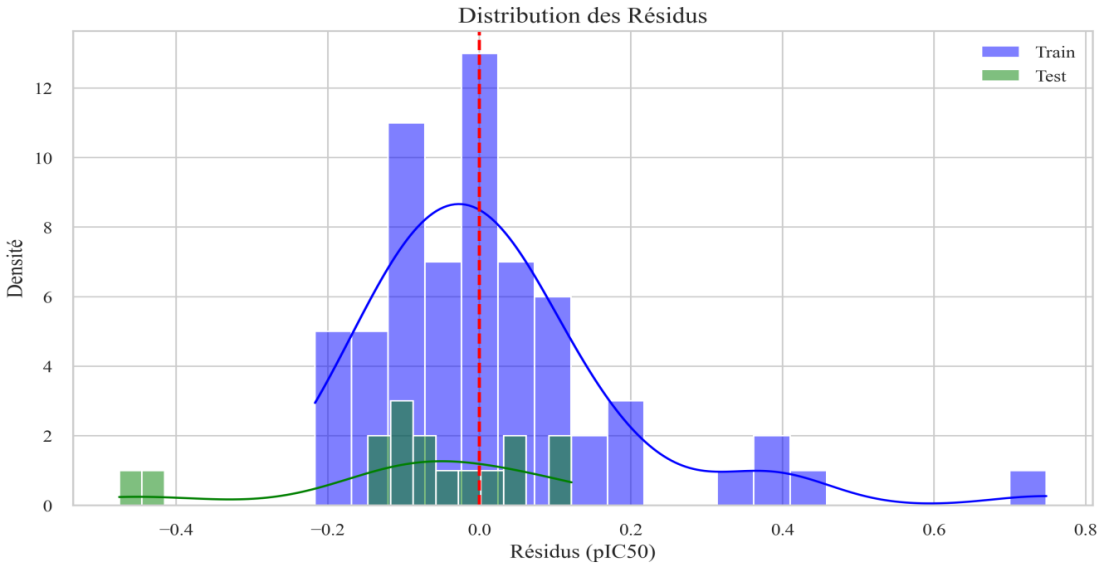


Figure 71: Distribution of residues.

V.4.3 Models' comparison:

The performance of the multiple linear regression (MLR) and support vector regression (SVR) models was systematically compared to assess their effectiveness in predicting the biological activity of the studied compounds. Both models used the same dataset and we chose five molecular descriptors to ensure consistency.

Table 29: Statistical criteria for the MLR and SVR models

Validation criterion	**Model**	
	MLR	**SVR**
R^2	0,93	0,912
Q^2	0,873	0,859
RMSE	0.163	0,183
Number of features	5	5

Interpretation:

A cross-validation comparison showed that both multiple linear regression (MLR) and support vector regression (SVR) models performed exceptionally well in predicting bioactivity, with MLR showing slightly superior performance across all evaluated criteria. MLR achieved a slightly higher coefficient of determination (R^2 = 0.93 versus 0.912 for SVR), indicating that it explained approximately 93% of the variance in the bioactivity data, compared to 91.2% for SVR. Both models demonstrated strong predictive ability in cross-validation, with MLR's Q^2 of 0.873 slightly outperforming SVR's Q^2 of 0.859, suggesting that MLR may generalize better to new data. Root mean square error (RMSE) values were comparable but favored MLR (0.163 versus 0.183), indicating slightly more accurate predictions. Notably, both models used the same five molecular descriptors, eliminating feature selection bias in the comparison. The results suggest that, for this particular dataset, the simpler MLR model may outperform the more complex SVR model, achieving similar or

better performance with greater interpretability. However, the slight differences in performance suggest that both modeling approaches are valid for this QSAR application.

V.4.4 Discussion:

- ✓ The Multiple Linear Regression (MLR) model demonstrated robust predictive performance based on standard validation metrics. The coefficient of determination for the training set (R^2_train) was 0.890, indicating that approximately 89% of the variance in the biological activity could be explained by the selected descriptors. The model also exhibited excellent external predictive power, with an R^2_test value of 0.930, suggesting high reliability in predicting unseen data. The cross-validated coefficient (Q^2 = 0.873) further supports the model's strong internal consistency and predictive ability. Additionally, the root mean square error (RMSE) was 0.158 for the training set and 0.163 for the test set, reflecting low prediction errors and good generalization performance. These results collectively confirm the reliability and robustness of the MLR model for structure–activity relationship (SAR) predictions.

- ✓ The predictive performance of the QSAR model developed using a Support Vector Regression (SVR) approach was evaluated based on standard statistical validation metrics. The coefficient of determination for the training set (R^2_train) was 0.875, indicating a strong correlation between the predicted and experimental activities during model training. The model exhibited even higher predictive accuracy on the test set with an R^2_test value of 0.912, suggesting good generalizability. The cross-validated coefficient (Q^2) was 0.859, further confirming the model's robustness and predictive reliability. The root mean square error (RMSE) values were 0.169 and 0.183 for the training and test sets, respectively, reflecting low prediction errors and consistent performance across datasets. These validation criteria demonstrate that the SVR-based QSAR model is both statistically reliable and predictive for the designed molecular dataset.

- ✓ The performance of the developed QSAR models was evaluated using Multiple Linear Regression (MLR) and Support Vector Regression (SVR) approaches. The MLR model achieved a coefficient of determination (R^2) of 0.93, slightly outperforming the SVR model with an R^2 of 0.912. Both models demonstrated strong predictive power, as indicated by high cross-validated coefficients (Q^2), with values of 0.873 for MLR and 0.859 for SVR. The root mean square error (RMSE) was 0.163 for MLR and 0.183 for SVR, indicating low prediction errors in both cases. Each model was constructed using five molecular descriptors, reflecting a balanced complexity with reliable predictive capability. These results confirm that both MLR and SVR models are statistically robust and suitable for predicting the biological activity of the selected compounds.
- ✓ Distribution of residuals for the QSAR model on both the training and test data. Residuals are represented as a measure of the prediction difference (ΔpIC_{50}) between the experimental and predicted values. The blue bars represent the distribution of residuals in the training set, while the green bars represent the distribution of residuals in the test set. The distribution of the data in the training set appears close to a normal distribution, with a dense center around zero, indicating good performance of the model during training. This result indicates good generalization ability of the model, as it does not show any obvious bias or highly skewed residual pattern.

Ⅴ.5 Conclusion:

This comprehensive computational study demonstrates the promising potential of fenugreek (fenugreekine, diosgenin), turmeric (curcumin), and *Atriplex halimus* (arbutin) as natural inhibitors against key breast cancer targets (HER2, CDK4, AKT1, and MCF-7). The molecular docking results revealed strong binding affinities, with fenugreekine and diosgenin exhibiting the highest interactions (docking scores: −9.1 to −11.4 kcal/mol), comparable to or even surpassing some commercial drugs like tamoxifen (−8.9 kcal/mol) and lapatinib (−10.3 kcal/mol). While these results highlight the therapeutic potential of plant-derived compounds, experimental validation (e.g., cell-based assays, pharmacokinetic studies) is essential to confirm efficacy. Future work should focus on: In vivo studies to evaluate tumor inhibition and safety and Synergistic combinations with existing therapies to combat resistance.

The evaluation of the QSAR models developed using both Multiple Linear Regression (MLR) and Support Vector Regression (SVR) methods demonstrated strong statistical performance and predictive reliability. The MLR model slightly outperformed SVR in terms of R^2 and RMSE, indicating a marginally better fit to the data and lower prediction error. However, both models exhibited high Q^2 values, confirming their robustness and generalization capability. The use of five molecular descriptors in each model ensured a good balance between model simplicity and predictive power. Collectively, these findings validate the suitability of both MLR and SVR approaches for accurately predicting the biological activity of the studied compounds, supporting their application in rational drug design.

General Conclusion

General Conclusion

In conclusion, this study combines traditional phytochemistry with modern computational drug design, providing a cost-effective strategy for identifying natural anticancer agents. The strong binding interactions and favorable ADMET properties of fenugreek, diosgenin, and curcumin make them promising candidates for further preclinical development in breast cancer treatment. The methodology yielded accurate and reliable results, enabling us to identify the most therapeutically potent compounds, with a comprehensive evaluation of their pharmacokinetic and toxicological properties. This approach opens new avenues in computational modeling-based drug discovery, which will contribute to accelerating the development of new drugs and improving their therapeutic properties.

The experimental study of the ethanolic extract of fenugreek (*Trigonella foenum-graecum*) demonstrated its high phytochemical richness and antioxidant capacity. The extraction rate (12–20%) was consistent with published values in the scientific literature, confirming the efficient recovery of bioactive compounds. UV–Vis spectroscopy revealed absorption patterns (200–300 nm) characteristic of polyphenols, flavonoids, and saponins, while Fourier transform infrared (FTIR) analysis identified key functional groups, including hydroxyl (3272 cm^{1}) and carbonyl (1637 cm^{1}) bonds, confirming the presence of functional groups characteristic of diosgenin and other steroidal saponins. These results highlight the therapeutic potential of fenugreek, particularly in cancer prevention.

This comprehensive computational study evaluated the potential of fenugreekine, diosgenin, curcumin, and arbutin as natural inhibitors of key breast cancer targets: HER2, CDK4, AKT1, and MCF-7. Fenugreekine showed the strongest binding affinity (-11.4 kcal/mol for AKT1), closely followed by diosgenin, demonstrating superior inhibitory capacity compared to the standard drug Capivasertib (-10.1 kcal/mol). Curcumin also showed good and consistent affinity with all targets, while arbutin showed intermediate results. DFT-based quantitative chemical analysis revealed that fenugreekine has the highest electrophilicity (ω =0.5339) and chemical rigidity (η = 0.0511), making it a stable electron acceptor, while curcumin exhibited higher reactivity due to its narrow HOMO-LUMO gap.

ADMET predictions confirmed favorable pharmacokinetic properties, with diosgenin showing perfect conformity to Lipinski's criteria (LogP = 5.02, MW = 414.62 g/mol) and low toxicity (LD_{50} = 8000 mg/kg), similar to curcumin and arbutin, making it suitable for oral administration. while fenugreekine encountered absorption difficulties due to its high polarity. These results confirm the therapeutic potential of plant-derived compounds and warrant further in vitro and in vivo studies to validate their efficacy. Future research could explore structural optimization to enhance bioavailability and target specificity.

The quantitative structure-activity relationship (QSAR) study successfully developed and validated two predictive models, multiple linear regression (MLR) and support vector regression (SVR), to evaluate the inhibitory potential of phytochemicals against breast cancer-associated proteins. The MLR model demonstrated superior performance with an R^2_test of 0.930, a Q^2 of 0.873, and a low RMSE (0.163), confirming its strong linear predictive ability. Meanwhile, the SVR model demonstrated robust performance (R^2_test = 0.912, Q^2 = 0.859, and RMSE = 0.183), enhancing its effectiveness in detecting potential nonlinear relationships. The feature selection process identified five key molecular descriptors (Chi2v, Chi3v, Chi4n, SlogP_VSA5, and Chi4v) that significantly influence biological activity, providing mechanistic insights into structure-activity relationships. This study highlights the reliability of QSAR modeling in drug discovery, with MLR yielding interpretable linear trends, while SVR allows for complex nonlinear patterns. These results underscore the therapeutic potential of turmeric derivatives (curcumin) and provide a computational framework for optimizing novel breast cancer inhibitors. Future research could expand the dataset and explore additional machine learning techniques to enhance predictive power.

Perspective

Develop safer, more affordable, and natural treatment options derived from traditional medicinal plants… In vitro (cell culture) and in vivo (animal models) studies are needed to confirm computational predictions.

Early diagnosis of breast cancer improves survival rates and reduces the risk of it spreading to other parts in the body.

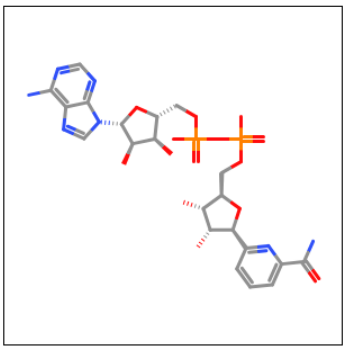
In summary, these computational predictions align with existing literature on similar phytochemicals, though experimental validation is needed to confirm the exact binding modes. The identified interacting residues serve as valuable guides for future structure-activity relationship studies and rational design of more potent analogs.

ANNEX

ANNEX

ANNEX ProTox 1

Oral toxicity prediction results for input compound



Predicted LD50: 7000mg/kg

Predicted Toxicity Class: 6

1 2 3 4 5 6

Average similarity: 61.34%

Prediction accuracy: 68.07%

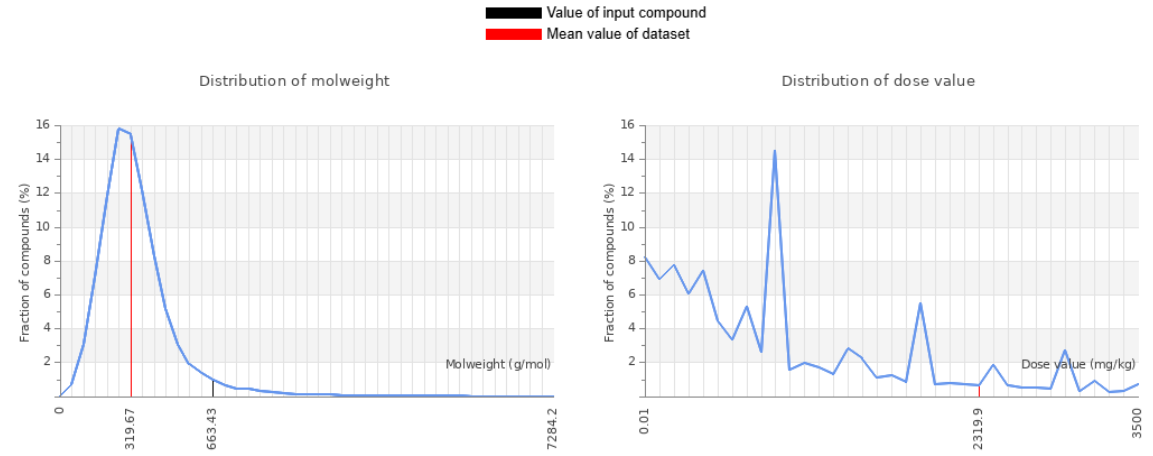
20% 40% 60% 80%

Print Toxicity Report

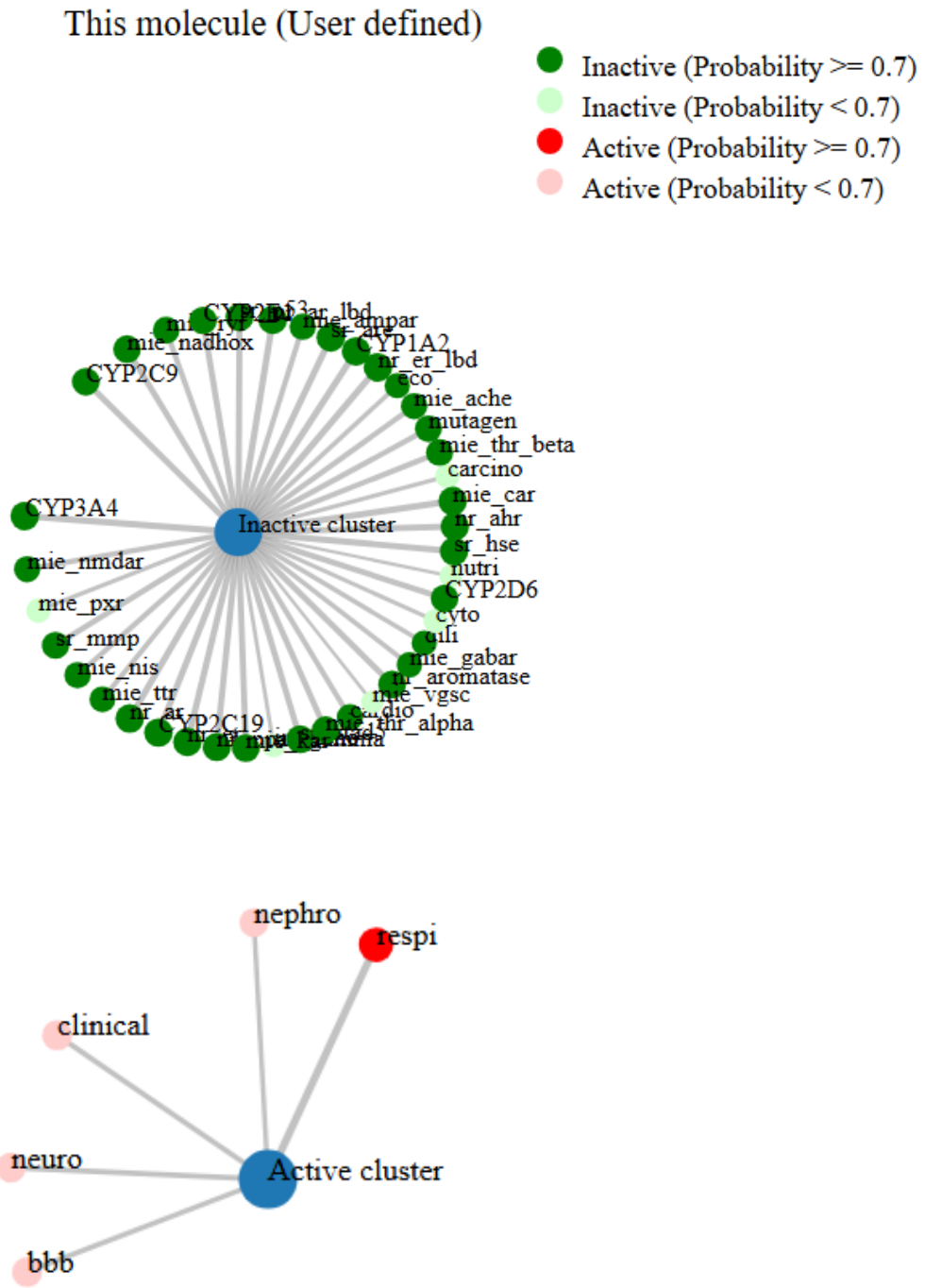
Name	C1=CC(=NC(=C1)C(=O)[C@H]2[C@@H]([C@@H]([C@H](O2)COP(=O)(O)OP(=O)(O)OC[C@@H]3[C@H]([C@H]([C@@H](O3)N4C=NC5=C(N=CN
Molweight	663.43
Number of hydrogen bond acceptors	20
Number of hydrogen bond donors	8
Number of atoms	44
Number of bonds	48
Number of rotable bonds	11
Molecular refractivity	140.12
Topological Polar Surface Area	346.89
octanol/water partition coefficient(logP)	-1.08

(1)

Comparison of input compound with dataset compounds



(2)



(3)

Figure: (1) class Toxicity of fenugreekine. **(2)** Comparison with dataset compounds. **(3)** The connection between the selected compound and predicted activities

ANNEX MLR 2

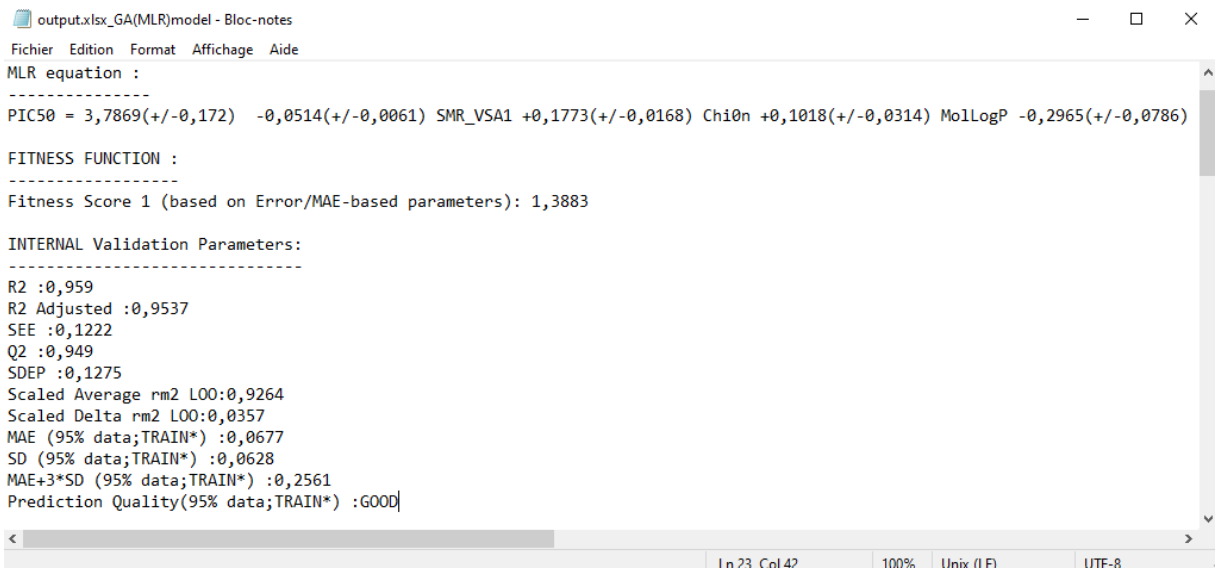
MLR algorithm:

```
untitled0 - Bloc-notes
Fichier  Edition  Format  Affichage  Aide
# -*- coding: utf-8 -*-
"""
Created on Mon Apr 28 11:28:01 2025

@author: Lenovo
"""

# -*- coding: utf-8 -*-
"""
Modèle MLR pour prédiction de pIC50 à partir de SMILES
@author: IT DOCTOR
"""
import numpy as np
import pandas as pd
from rdkit import Chem
from rdkit.Chem import Descriptors
from rdkit.ML.Descriptors import MoleculeDescriptors
from sklearn.linear_model import LinearRegression
from sklearn.model_selection import train_test_split, cross_val_predict
from sklearn.preprocessing import StandardScaler
from sklearn.metrics import r2_score, mean_squared_error
from sklearn.feature_selection import VarianceThreshold, SelectKBest, f_regression
import matplotlib.pyplot as plt
import seaborn as sns
from scipy import stats
import os
#%%
# Configuration des graphiques pour publication scientifique
plt.style.use('seaborn-v0_8')
sns.set(style="whitegrid", font_scale=1.2)
plt.rcParams.update({
    'font.family': 'Times New Roman',
    'figure.dpi': 300,
    'savefig.dpi': 300,
    'axes.titlesize': 16,
    'axes.labelsize': 14,
Ln 13, Col 19  100%  Windows (CRLF)  UTF-8
```

Figure: MLR Code.



```
output.xlsx_GA(MLR)model - Bloc-notes
Fichier  Edition  Format  Affichage  Aide
MLR equation :
---------------
PIC50 = 3,7869(+/-0,172)  -0,0514(+/-0,0061) SMR_VSA1 +0,1773(+/-0,0168) Chi0n +0,1018(+/-0,0314) MolLogP -0,2965(+/-0,0786)

FITNESS FUNCTION :
------------------
Fitness Score 1 (based on Error/MAE-based parameters): 1,3883

INTERNAL Validation Parameters:
-------------------------------
R2 :0,959
R2 Adjusted :0,9537
SEE :0,1222
Q2 :0,949
SDEP :0,1275
Scaled Average rm2 LOO:0,9264
Scaled Delta rm2 LOO:0,0357
MAE (95% data;TRAIN*) :0,0677
SD (95% data;TRAIN*) :0,0628
MAE+3*SD (95% data;TRAIN*) :0,2561
Prediction Quality(95% data;TRAIN*) :GOOD
Ln 23, Col 42  100%  Unix (LF)  UTF-8
```

Figure: MLR Genetic calculation results.

ANNEX SVR 3

SVR algorithm:

Figure: SVR code.

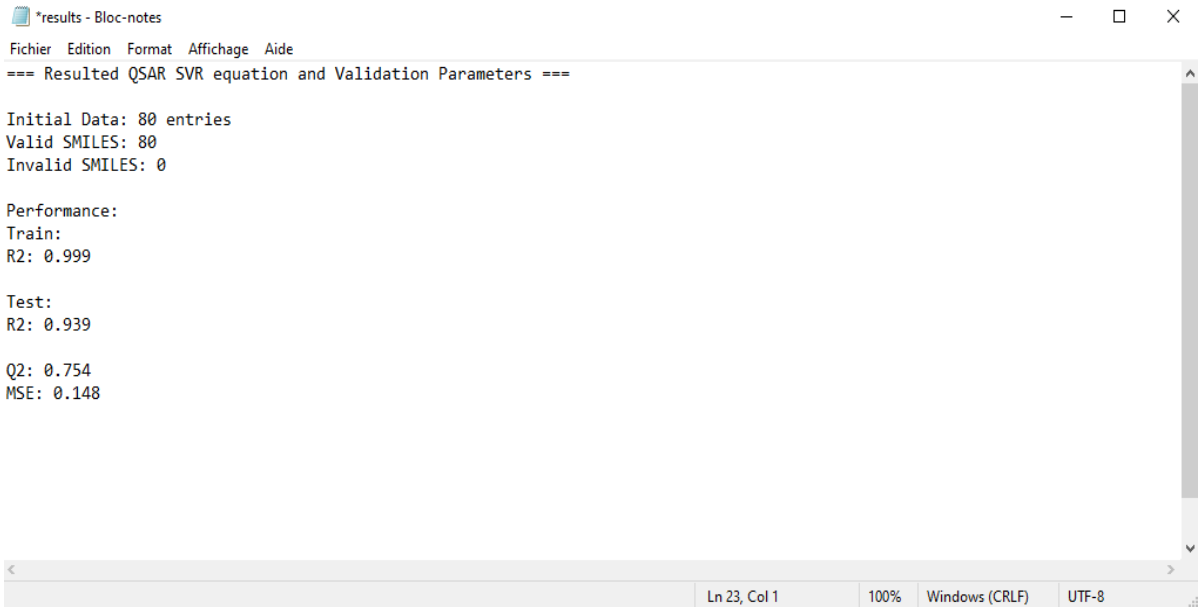


Figure: SVR calculation results.

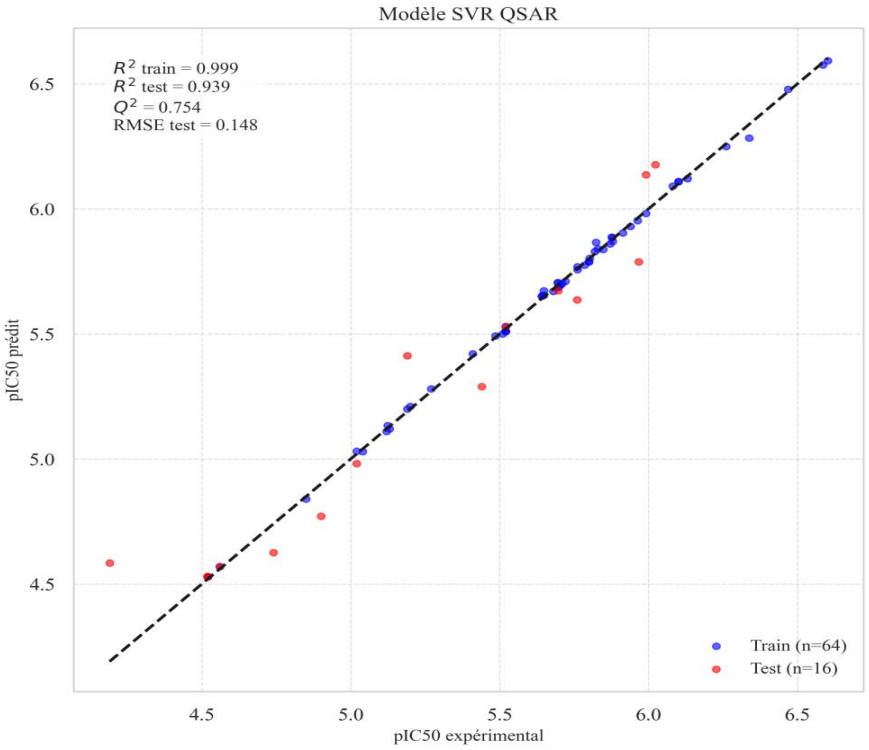
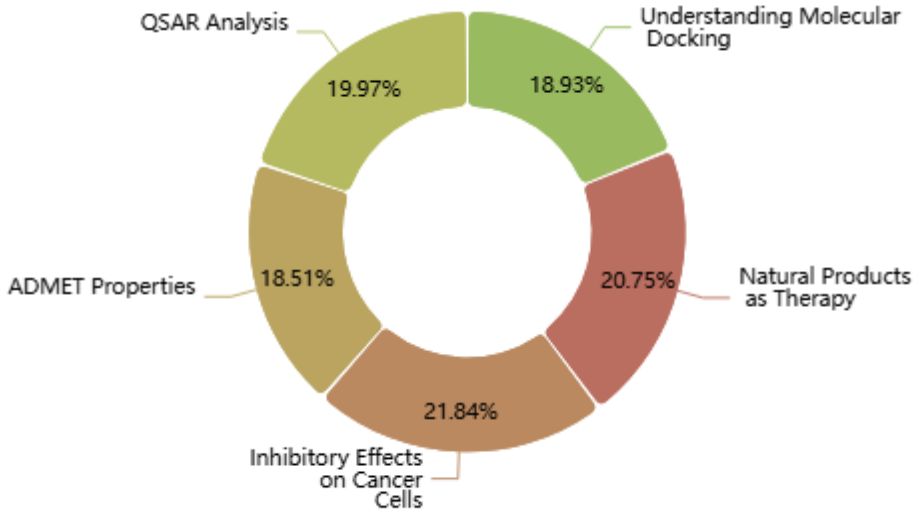


Figure: SVR QSAR results.



Insights into the Role of Natural Products in Cancer Therapy(Effectiveness Score (%))

REFERENCE

REFERENCE

Reference

[1] Al-Hadi, Ahmed, et al. "In Silico Evaluation of the Inhibitory Potential of Natural Compounds against the SARS-CoV-2 Main Protease (Mpro) and Angiotensin-Converting Enzyme 2 (ACE2)." BioMed Research International, vol. 2022, 2022, p. 3338549. Wiley Online Library, doi:10.1155/2022/3338549.

[2] Yadav, Dharmendra K., et al. "Molecular docking, QSAR and ADMET studies of withanolide analogs against breast cancer." Drug Design, Development and Therapy, vol. 11, 2017, pp. 1859-70. PubMed, doi:10.2147/DDDT.S130601.

[3] McGrowder, Donovan A., et al. "Medicinal Herbs Used in Traditional Management of Breast Cancer: Mechanisms of Action." Medicines (Basel), vol. 7, no. 8, 2020, p. 47. PubMed Central, doi:10.3390/medicines7080047.

[4] El Rhabori, Said, et al. "Exploring innovative strategies for identifying anti-breast cancer compounds by integrating 2D/3D-QSAR, molecular docking analyses, ADMET predictions, molecular dynamics simulations, and MM-PBSA approaches." Journal of Molecular Structure, vol. 1297, 2024, p. 139500. doi:10.1016/j.molstruc.2024.139500.

[5] Rashmi, S., et al. "Predicting the molecular subtype of breast cancer based on mammography and ultrasound findings." Indian Journal of Radiology and Imaging, vol. 28, no. 3, 2018, pp. 278-83. PubMed Central, doi:10.4103/ijri.IJRI_78_18.

[6] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA A Cancer J Clin. 2019 Jan; 69(1):7–34.

[7] Sana S, Tokala R, Bajaj DM, Nagesh N, Bokara KK, Kiranmai G, et al. Design and synthesis of substituted dihydropyrimidinone derivatives as cytotoxic and tubulin polymerization inhibitors. Bioorg Chem.

[8] Sun M, Xu Q, Xu J, Wu Y, Wang Y, Zuo D, et al. Synthesis and bioevaluation of N,4-diaryl-1,3-thiazole-2-amines as tubulin inhibitors with potent antiproliferative activity.

[9] Fan L, Goss PE, Strasser-Weippl K. Current status and future projections of breast cancer in Asia. Breast Care (Basel). 2015 Dec;10(6):372–8.

[10] Ceramella J, Caruso A, Occhiuzzi MA, Iacopetta D, Barbarossa A, Rizzuti B, et al. Benzothienoquinazolinones as new multi-target scaffolds: Dual inhibition of human Topoisomerase I and tubulin polymerization. Eur J Med Chem. 2019 Nov;181:111583.

REFERENCE

[11] EL-Mernissi, Reda, et al. "3D-QSAR, molecular docking, ADMET, simulation dynamic, and retrosynthesis studies on new styrylquinolines derivatives against breast cancer." Open Chemistry, vol. 22, no. 1, 2024. De Gruyter, doi:10.1515/chem-2024-0041.

[12] Cleveland Clinic. "Ductal Carcinoma In Situ (DCIS)." Cleveland Clinic, my. https://my.clevelandclinic.org/health/diseases/17869-ductal-carcinoma-in-situ-dcis.

[13] Cleveland Clinic. "Breast Cancer." Cleveland Clinic, my. https://my.clevelandclinic.org/health/diseases/3986-breast-cancer.

[14] Mayo Clinic Staff. "Ductal Carcinoma In Situ (DCIS): Symptoms & Causes." Mayo Clinic, www.mayoclinic.org/diseases-conditions/dcis/symptoms-causes/syc-20371889

[15] Cleveland Clinic. "Lobular Breast Cancer." Cleveland Clinic, my. https://my.clevelandclinic.org/health/diseases/21180-lobular-breast-cancer.

[16] **Breast Cancer Now.** "Invasive Lobular Breast Cancer." Breast Cancer Now, https://breastcancernow.org/about-breast-cancer/diagnosis/types-of-breast-cancer/invasive-lobular-breast-cancer.

[17] University Hospitals. "Breast Cancer Surgery Appointment." University Hospitals, www.uhhospitals.org/services/cancer-services/breast-cancer/surgery-appointment.

[18] American Cancer Society. "Inflammatory Breast Cancer | Details, Diagnosis, and Signs." American Cancer Society, www.cancer.org/cancer/types/breast-cancer/about/types-of-breast-cancer/inflammatory-breast-cancer.html.

[19] Liu, Jing, et al. "Decision Regret About Treatment Amongst Women With Early Breast Cancer: An Integrative Review." Journal of Advanced Nursing, 2025. PubMed, doi:10.1111/jan.16767.

[20] Breast Cancer Statistics and Resources Breastcancer.org, www.breastcancer.org/.

[21] International Agency for Research on Cancer. "Algeria: Statistics at a Glance, 2022. Global Cancer Observatory Fact Sheet." 2022. International Agency for Research on Cancer, gco.iarc.who.int/media/globocan/factsheets/populations/12-algeria-fact-sheet.pdf

[22] Union for International Cancer Control. "Top cancer per country." UICC, 2 Aug. 2023, www.uicc.org/news-and-updates/resources/top-cancer-country.

[23] Liu, Jing, et al. "Decision Regret About Treatment Amongst Women With Early Breast Cancer: An Integrative Review." Journal of Advanced Nursing, 22 Jan. 2025. PubMed, doi:10.1111/jan.16767.

REFERENCE

[24] Wen X, Guo X, Wang S, Lu Z, Zhang Y. Breast cancer diagnosis: A systematic review. In: Biocybernetics and Biomedical Engineering, vol. 44. Elsevier BV; (2024). p. 119–48.

[25] He Z, Chen Z, Tan M, Elingarami S, Liu Y, Li T, et al. A review on methods for diagnosis of breast cancer cells and tissues. In: Cell Proliferation, vol. 53. Wiley; (2020).

[26] Magnoni F, Alessandrini S, Alberti L, Polizzi A, Rotili A, Veronesi P, et al. Breast cancer surgery: new issues. Curr Oncol. (2021) 28:4053–66.

[27] Riis M. Modern surgical treatment of breast cancer. In: Annals of Medicine and Surgery, vol. 56. Ovid Technologies (Wolters Kluwer Health; (2020). p. 95–107.

[28] Mayo Clinic Staff. "Breast Cancer: Symptoms & Causes." Mayo Clinic, www.mayoclinic.org/diseases-conditions/breast-cancer/symptoms-causes/syc-20352470.

[29] Purswani JM, Hardy-Abeloos C, Perez CA, Kwa MJ, Chadha M, Gerber NK. Radiation in early-stage breast cancer: moving beyond an all or nothing approach. Curr Oncol. (2022) 30:184–95.

[30] Poland S, Ebina W, Muggia F, Guth A. Breast radiation-associated secondary Malignancies: A review. Clin Surg Oncol. (2023) 2:100010. doi: 10.1016/j.cson.2023.100010.

[31] Wang J, Wu S-G. Breast Cancer: An Overview of Current Therapeutic Strategies, Challenge, and Perspectives. In: Breast Cancer: Targets and Therapy, vol. 15. Informa UK Limited; (2023). p. 721–30.

[32] Mutebi M, Anderson BO, Duggan C, Adebamowo C, Agarwal G, Ali Z, et al. Breast cancer treatment: A phased approach to implementation. Cancer. (2020) 126:2365–78.

[33] Lawson M, Cureton N, Ros S, Cheraghchi-Bashi A, Urosevic J, Darcy S, et al. The next-generation oral selective estrogen receptor degrader camizestrant (AZD9833) suppresses ER+ Breast cancer growth and overcomes endocrine and CDK4/6 inhibitor resistance. Cancer Res. (2023) 83:3989–4004.

[34] Neupane N, Bawek S, Gurusinghe S, Ghaffary EM, Mirmosayyeb O, Thapa S, et al. Oral SERD, a novel endocrine therapy for estrogen receptor-positive breast cancer. Cancers. (2024) 16:619.

[35] Debela DT, Muzazu SG, Heraro KD, Ndalama MT, Mesele BW, Haile DC, et al. New approaches and procedures for cancer treatment: Current perspectives. In: SAGE Open Medicine, vol. 9. SAGE Publications; (2021). p. 205031212110343.

REFERENCE

[36] Naser R, Dilabazian H, Bahr H, Barakat A, El-Sibai M. A guide through conventional and modern cancer treatment modalities: A specific focus on glioblastoma cancer therapy (Review). Oncol Rep. (2022) 48.

[37] Nave, Ophir. "Asymptotic analysis of mathematical model describing a new treatment of breast cancer using AZD9496 and palbociclib." Frontiers in Oncology, vol. 14, 2024, p. 1482223. PubMed Central, doi:10.3389/fonc.2024.1482223.

[38] "HER2-Positive Breast Cancer: Hormone Therapy and Anti-HER2 Treatment: An Update on Treatment Strategies." Mayo Clinic, www.mayoclinic.org/diseases-conditions/breast-cancer/expert-answers/breast-cancer/faq-20058066.

[39] Cheng, Xiaoqing. "A Comprehensive Review of HER2 in Cancer Biology and Therapeutics." Genes, vol. 15, no. 7, 2024, p. 903. MDPI, doi:10.3390/genes15070903.

[40] Tommasi, Chiara, et al. "Hormone Receptor-Positive/HER2-Positive Breast Cancer: Hormone Therapy and Anti-HER2 Treatment: An Update on Treatment Strategies." Journal of Clinical Medicine, vol. 13, no. 7, 2024, p. 1873. MDPI, doi:10.3390/jcm13071873.

[41] RCSB PDB. RCSB PDB, www.rcsb.org/.

[42] Goffin, Vincent, et al. "Identification of Gain-of-Function Variants of the Human Prolactin Receptor." Methods in Enzymology, vol. 484, Academic Press, 2012. ScienceDirect, doi:10.1016/B978-0-12-381298-8.00017-4.

[43] Public Health England. "MCF-7 (ECACC 86012803) - Cell line profile." Culture Collections, 16 May 2024, www.culturecollections.org.uk/products/cell-cultures/ecacc-cell-line-profiles/mcf7/.

[44] Fassl, Anne, et al. "CDK4 and CDK6 kinases: From basic science to cancer therapy." Science, vol. 375, no. 6577, 2022, p. eabc1495. PubMed, doi:10.1126/science.abc1495.

[45] Abd-Elghany, A. A., et al. "Chemosensitization and Molecular Docking Assessment of Dio-NPs on Resistant Breast Cancer Cells to Tamoxifen." Pharmaceuticals, vol. 18, no. 4, 2025, p. 452. MDPI, doi:10.3390/ph18040452.

[46] Rudolph, M., et al. "AKT1 (E17K) mutation profiling in breast cancer: prevalence, concurrent oncogenic alterations, and blood-based detection." BMC Cancer, vol. 16, 2016, p. 622. BioMed Central, doi:10.1186/s12885-016-2626-1.

[47] Hinz, N.; Jücker, M. Distinct Functions of AKT Isoforms in Breast Cancer: A Comprehensive Review. Cell Commun. Signal. 2019, 17, 154.

REFERENCE

[48] Watson, K.L.; Moorehead, R.A. Loss of Akt1 or Akt2 Delays Mammary Tumor Onset and Suppresses Tumor Growth Rate in MTB-IGFIR Transgenic Mice. BMC Cancer 2013, 13, 375.

[49] Gargini, R.; Cerliani, J.P.; Escoll, M.; Antón, I.M.; Wandosell, F. Cancer Stem Cell-like Phenotype and Survival Are Coordinately Regulated by Akt/FoxO/Bim Pathway. Stem Cells 2015, 33, 646–660.

[50] Hutchinson, J.; Jin, J.; Cardiff, R.D.; Woodgett, J.R.; Muller, W.J. Activation of Akt (Protein Kinase B) in Mammary Epithelium Provides a Critical Cell Survival Signal Required for Tumor Progression. Mol. Cell. Biol. 2001, 21, 2203–2212.

[51] Dillon, R.L.; Marcotte, R.; Hennessy, B.T.; Woodgett, J.R.; Mills, G.B.; Muller, W.J. Akt1 and Akt2 Play Distinct Roles in the Initiation and Metastatic Phases of Mammary Tumor Progression. Cancer Res. 2009, 69, 5057–5064.

[52] AHe, Yiming, et al. "The Landscape of RNA m6A Modification in Cancer." Molecular Cancer, vol. 22, no. 1, 2023, p. 95. BioMed Central, doi:10.1186/s12943-023-01805-y..

[53] Ye, F., Dewanjee, S., Li, Y. et al. Advancements in clinical aspects of targeted therapy and immunotherapy in breast cancer. Mol Cancer **22**, 105 (2023). https://doi.org/10.1186/s12943-023-01805-y

[54] Khan, Amjad, et al. "Herbal Spices as Food and Medicine: Microscopic Authentication of Commercial Herbal Spices." Plants, vol. 13, no. 8, 2024, p. 1067. doi:10.3390/plants13081067.

[55] Yadav, U.C.S., and N.Z. Baquer. "Pharmacological effects of Trigonella foenum-graecum L. in health and disease." Pharmaceutical Biology, vol. 52, no. 2, 2014, pp. 243-54. PubMed, doi:10.3109/13880209.2013.826247.

[56] Abdel Salam, Shaimaa G., et al. "Phytochemical screening and in-vitro biological properties of unprocessed and household processed fenugreek (Trigonella foenum-graecum Linn.) seeds and leaves." Scientific Reports, vol. 13, no. 1, 2023, p. 5747. PubMed Central, doi:10.1038/s41598-023-31888-y.

[57] Ewalia. "Fenugreek." Ewalia Magic Tipps, 4 May 2022, www.ewalia.com/en/ewalia-magic-tips/fenugreek.

[58] Altuntas E., Ozgoz E., Taser O.F. Some physical properties of fenugreek (Trigonella foenum-graceum L.) seeds. J. Food Eng. 2005;71:37–43.

REFERENCE

[59] Wani, Sajad Ahmad, and Pradyuman Kumar. "Fenugreek: A review on its nutraceutical properties and utilization in various food products." Journal of the Saudi Society of Agricultural Sciences, vol. 17, no. 2, 2018, pp. 97-106. ScienceDirect, doi:10.1016/j.jssas.2016.01.007.

[60] Rampogu, Shailima, et al. "Exploring the Therapeutic Ability of Fenugreek against Type 2 Diabetes and Breast Cancer Employing Molecular Docking and Molecular Dynamics Simulations." Evidence-Based Complementary and Alternative Medicine, vol. 2018, 2018, p. 1943203. PubMed Central, doi:10.1155/2018/1943203.

[61] Yatoo, Gulam Nabi, and Javid Ahmad Banday. "Synthesis, antioxidant, antiproliferative activity, molecular docking and DFT studies of novel isoxazole derivatives of diosgenin, a steroidal sapogenin from Dioscorea deltoidea." Fitoterapia, vol. 172, 2024, p. 105621. doi:10.1016/j.fitote.2023.105621.

[62] Chiang, Chun-Te, et al. "Diosgenin, a naturally occurring steroid, suppresses fatty acid synthase expression in HER2-overexpressing breast cancer cells through modulating Akt, mTOR and JNK phosphorylation." FEBS Letters, vol. 581, no. 28, 2007, pp. 5700-706. doi:10.1016/j.febslet.2007.11.021.

[63] Khan, Muhammad Rashid, Muhammad Riaz, and Nauman Aziz. "Diosgenin, a Steroid Saponin Constituent of Yams and Fenugreek: Emerging Evidence for Applications in Medicine." Fenugreek: A Wonder Herb, edited by Wadud Maisha, IntechOpen, 2012. doi:10.5772/26700.

[64] Magallanes, Sara, et al. "Cytotoxicity of Fenugreek Sprout and Seed Extracts and Their Bioactive Constituents on MCF-7 Breast Cancer Cells." Molecules, vol. 27, no. 4, 2022, p. 1319. PubMed Central, doi:10.3390/molecules27041319.

[65] Gouveia, Bruno A., et al. "Diosgenin: an important natural pharmaceutical active ingredient." Food Science and Technology, vol. 42, 2022, p. e94521. doi:10.1590/fst.94521.

[66] Khanal, Pukar, et al. "Systems and in vitro pharmacology profiling of diosgenin against breast cancer." Frontiers in Pharmacology, vol. 13, 3 Jan. 2023, p. 1052849. Frontiers Media S.A., doi:10.3389/fphar.2022.1052849.

[67] Chiang, Chun-Te, et al. "Diosgenin, a naturally occurring steroid, suppresses fatty acid synthase expression in HER2-overexpressing breast cancer cells through modulating Akt, mTOR and JNK phosphorylation." FEBS Letters, vol. 581, no. 28, 2007, pp. 5700-06. PubMed, doi:10.1016/j.febslet.2007.11.021.

REFERENCE

[68] Mukherjee, P. K., et al. "Immunomodulatory leads from medicinal plants." Bioactive Food as Dietary Interventions for Diabetes, edited by Ronald Ross Watson and Victor R Preedy, Academic Press, 2012, pp. 235-56. doi:10.1016/B978-0-12-397153-1.00022-6.

[69] Martín Ortega, A. M., and M. R. Segura Campos. "Medicinal Plants and Their Bioactive Metabolites in Cancer Prevention and Treatment." Bioactive Compounds: Health Benefits and Potential Applications, Elsevier Inc., 2019, pp. 85-109. doi:10.1016/B978-0-12-814774-0.00005-0.

[70] "Turmeric Curcumin Powder." IndiaMART, www.indiamart.com/proddetail/turmeric-curcumin-powder-2849912806373.html.

[71] Hosseini, Zohreh, and Parvin Darvish. "Protective Effect of Curcumin on Breast Cancer, Type 2 Diabetes, and Obesity." Progress in Biological Research, vol. 6, no. 1, 2022, pp. 27-36. doi:10.18502/pbr.v6i1.3425.

[72] Kuzminska, Joanna, et al. "Curcumin Derivatives in Medicinal Chemistry: Potential Applications in Cancer Treatment." Molecules, vol. 29, no. 22, 2024, p. 5321. MDPI, doi:10.3390/molecules29225321.

[73] Kłosowska, Aleksandra, et al. "Fenugreek (Trigonella foenum-graecum L.) as a Source of Bioactive Compounds for Modern Therapy." Nutrients, vol. 11, no. 10, 2019, p. 2376. MDPI, doi:10.3390/nu11102376.

[74] Pricci, M.; Girardi, B.; Giorgio, F.; Losurdo, G.; Ierardi, E.; Di Leo, A. Curcumin and Colorectal Cancer: From Basic to Clinical Evidences. Int. J. Mol. Sci. 2020, 21, 2364.

[75] Zoi, V.; Galani, V.; Lianos, G.D.; Voulgaris, S.; Kyritsis, A.P.; Alexiou, G.A. The Role of Curcumin in Cancer Treatment. Biomedicines 2021, 9, 1086.

[76] Prakash, Anubhav, et al. "Modulating the Progression of Breast Cancer by Regulating the Microtubule-Associated Protein Tau." International Journal of Molecular Sciences, vol. 25, no. 21, 2024, p. 11492. MDPI, doi:10.3390/ijms252111492.

[77] Liu, Dongwu, and Zhiwei Chen. "The Effect of Curcumin on Breast Cancer Cells." Journal of Breast Cancer, vol. 16, no. 2, 2013, pp. 133-37. PubMed Central, doi:10.4048/jbc.2013.16.2.133. Accessed 20 June 2025.

[78] Bencheikh, N.; Elbouzidi, A.; Kharchoufa, L.; Ouassou, H.; Alami Merrouni, I.; Mechchate, H.; Es-Safi, I.; Hano, C.; Addi, M.; Bouhrim, M. Inventory of Medicinal Plants

REFERENCE

Used Traditionally to Manage Kidney Diseases in North-Eastern Morocco: Ethnobotanical Fieldwork and Pharmacological Evidence. Plants 2021, 10, 1966.

[79] Elbouzidi, A.; Bencheikh, N.; Seddoqi, S.; Bouhrim, M.; Bouramdane, Y.; Addi, M. Investigation of the Allelopathic Effect of Matricaria Chamomilla L. Parts' Aqueous Extracts on Germination and Seedling Growth of Two Moroccan Varieties of Durum Wheat. Int. J. Agron. 2021, 2021, 4451181.

[80] Bouchikh-Boucif, Y.; Labani, A.; Benabdeli, K.; Boidielouane, S. Allelopathic Effects of Shoot and Root Extracts From Three Alien and Native Chenopodiaceae Species on Lettuce Seed Germination. Ecol. Balk. 2014, 6, 51–55.

[81] Slama, K.; Boumendjel, M.; Taibi, F.; Boumendjel, A.; Messarah, M. Atriplex halimus Aqueous Extract Abrogates Carbon Tetrachloride-Induced Hepatotoxicity by Modulating Biochemical and Histological Changes in Rats. Arch. Physiol. Biochem. 2020, 126, 49–60.

[82] Benhammou, N.; Bekkara, F.A.; Panovska, T.K. Antioxidant Activity of Methanolic Extracts and Some Bioactive Compounds of Atriplex halimus. Comptes Rendus Chim. 2009, 12, 1259–1266.

[83] LC–MS/MS Phytochemical Profiling, Antioxidant Activity, and Cytotoxicity of the Ethanolic Extract of Atriplex Halimus L. against Breast Cancer Cell Lines: Computational Studies and Experimental Validation', accessed 15 June 2025, https://www.mdpi.com/1424-8247/15/9/1156.

[84] Walker, D. J., et al. "Atriplex halimus L.: Its biology and uses." Journal of Arid Environments, vol. 98, 2013, pp. 1-13. doi:10.1016/j.jaridenv.2013.09.004.

[85] Sghaier, M., et al. "Chromosome numbers in Tunisian populations of Atriplex halimus L. (Chenopodiaceae)." African Journal of Biotechnology, vol. 5, no. 12, 2006, pp. 1195-98.

[86] Gharaibeh, M. A., et al. "Reclamation of highly calcareous saline sodic soil using Atriplex halimus and by-product gypsum." Journal of Plant Nutrition, vol. 34, no. 11, 2011, pp. 1709-21. PubMed, doi:10.1080/15226514.2011.573821.

[87] Slama, Kheira, et al. "Atriplex halimus aqueous extract abrogates carbon tetrachloride-induced hepatotoxicity by modulating biochemical and histological changes in rats." Drug and Chemical Toxicology, vol. 42, no. 6, 2019, pp. 642-49. PubMed, doi:10.1080/13813455.2018.1489852.

REFERENCE

[88] Zouari, Nizar, et al. "Antioxidant activity of methanolic extracts and some bioactive compounds of Atriplex halimus." Comptes Rendus Chimie, vol. 12, no. 12, 2009, pp. 1259-66. doi:10.1016/j.crci.2009.02.004.

[89] Dahmani, Yacine, et al. "Traditional knowledge and diversity of wild medicinal plants in El Kantara's area (Algerian Sahara gate): An ethnobotany survey." Acta Ecologica Sinica, vol. 42, no. 1, 2022, pp. 24-34. doi:10.1016/j.chnaes.2021.01.007.

[90] Bouaziz, Samia, et al. "Protoscolicidal activity of Atriplex halimus leaves extract against Echinococcus granulosus protoscoleces." Experimental Parasitology, vol. 228, 2021, p. 108155. PubMed, doi:10.1016/j.exppara.2021.108155.

[91] Bentaha, B., et al. "Phytochemical screening and evaluation of the antioxidant and antibacterial activity of Atriplex halimus from two regions Algeria (El Oued and Tlemcen)." Genetics and Biodiversity Journal, vol. 5, no. Special Issue (Characterization and valorisation of plants: Virtues and development prospects), 2021, pp. 59-67. doi:10.46325/gabj.v5i2.72.

[92] Bekkouche, Mohamed Amine, et al. "Antidiabetic effect of Atriplex halimus long and short term treatment against streptozotocin induced diabetes in rats." Anales de Biología, vol. 44, 2022, pp. 15-22. doi:10.6018/analesbio.44.03.

[93] "Atriplex halimus." Les Secrets du Sahara, lessecretsdusahara.wordpress.com/tag/atriplex-halimus/.

[94] Mohammedi Z. Resistance, pharmacology properties and nutritional value of a shrub from arid environment s Atriplex halimus. Res J Med Plant, 2016; 10:10-8.

[95] Alves-Silva JM, Romane A, Efferth T, Salgueiro L. North African medicinal plants traditionally used in cancer therapy. Frontiers in pharmacology, 2017; 8:383.

[96] Khaoula, Zekrouk, and Djahra Ali Boutlelis. "Preventive and Curative Effects of Atriplex Halimus L. Aqueous Extract on Benzene Provoked Hepatic Injury in Rats." Journal of Drug Delivery and Therapeutics, vol. 10, no. 3, 2020, pp. 217-22. doi:10.22270/jddt.v10i3.4053.

[97] Thies, H.; Sulc, D. Arbutus unedo L. I. Determination of arbutin in the leaves of the strawberry tree. Pharmazie 1950, 5, 553–555.

[98] Nahar L, Al-Groshi A, Kumar A, Sarker SD. Arbutin: Occurrence in plants, and its potential as an anticancer agent. Molecules. 2022;27(24):8786.

REFERENCE

[99] Migas P, Krauze-Baranowska M. The significance of arbutin and its derivatives in therapy and cosmetics. Phytochem Lett. 2015;13:35- 40.

[100] Zhang S, Yao K, Pi Y, Yang S, Huang Z, Pan X, et al. Role of arbutin in the inhibition of FBXO5 in hepatocellular carcinoma. Discov Oncol. 2024;15(1):827.

[101] Hazman Ö, Sarıova A, Bozkurt MF, Ciğerci İH. The anticarcinogen activity of β-arbutin on MCF-7 cells: Stimulation of apoptosis through estrogen receptor-α signal pathway, inflammation and genotoxicity. Mol Cell Biochem. 2021;476(1):349-60.

[102] Kılınç, Mehmet Halil, et al. "Investigation of the Cytotoxicity of Arbutin Combined with Doxorubicin In Vitro." Yeditepe Journal of Health Sciences, vol. 12, no. 4, 2024, pp. 1047-52. doi:10.37989/yeditepe.v12i4.1047.

[103] Mechaala, S.; Bouatrous, Y.; Adouane, S. Traditional Knowledge and Diversity of Wild Medicinal Plants in El Kantara's Area (Algerian Sahara Gate): An Ethnobotany Survey. Acta Ecol. Sin. 2021, 42, 33–45.

[104] Walker, D.J.; Lutts, S.; Sánchez-García, M.; Correal, E. Atriplex halimus L.: Its Biology and Uses. J. Arid Environ. 2014, 100, 111–121.

[105] Al-Senosy, N.K.; Abou-Eisha, A.; Ahmad, E.S. In Vitro Antiproliferation Effect of Atriplex halimus L. Crude Extract on Human Cell Lines by Induction of Apoptosis and G2/M Phase Arrest. Egypt. Acad. J. Biol. Sci. C Physiol. Mol. Biol. 2018, 10, 115–126.

[106] Capua, C.J.; Hopson, N.P.; Stewart, C.M.M.; Johnston, G.R.; O'Neill, K.L.; Schaalje, G.B.; Lee, C.M.; Booth, G.M. Cytotoxicity of Atriplex confertifolia. J. Toxicol. 2010, 2010, 976548.

[107] Nahar, Lutfun, et al. "Arbutin: Occurrence in Plants, and Its Potential as an Anticancer Agent." Molecules, vol. 27, no. 24, 2022, p. 8786. MDPI, doi:10.3390/molecules27248786.

[108] Wang G., Zhu W. Molecular docking for drug discovery and development: a widely used approach but far from perfect. Future Medicinal Chemistry. 2016;8(14):1707–1710. doi: 10.4155/fmc-2016-0143.

|109] Rampogu S., Son M., Baek A., et al. Targeting natural compounds against {HER}2 kinase domain as potential anticancer drugs applying pharmacophore based molecular modelling approaches. Computational Biology and Chemistry. 2018;74:327–338.

REFERENCE

[110] Rampogu S., Baek A., Zeb A., Lee K. W. Exploration for novel inhibitors showing back-to-front approach against VEGFR-2 kinase domain (4AG8) employing molecular docking mechanism and molecular dynamics simulations. BMC Cancer. 2018;18(1).

[111] Meng X.-Y., Zhang H.-X., Mezei M., Cui M. Molecular docking: a powerful approach for structure-based drug discovery. Current Computer-Aided Drug Design. 2011;7(2):146–157

[112] Agu, P. C., et al. "Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management." Scientific Reports, vol. 13, no. 1, 2023, p. 13398. Nature, doi:10.1038/s41598-023-40160-2.

[113] Meng, X. Y., Zhang, H. X., Mezei, M. & Cui, M. Molecular docking: A powerful approach for structure-based drug discovery. Curr. Comput. Aided Drug Des. 7, 146–157 (2011).

[114] Agu, PC, et al. "Molecular docking as a tool for the discovery of molecular targets of nutraceuticals in diseases management." Scientific Reports, vol. 13, 2023, p. 13398. Nature, doi:10.1038/s41598-023-40160-2.

[115] Kumer, Ajoy, et al. "Modified d-glucofuranose computationally screening for inhibitor of breast cancer and triple breast cancer: chemical descriptor, molecular docking, molecular dynamics and qsar." Journal of the Chilean Chemical Society, vol. 63, no. 3, 2022, pp. 5623-35. SciELO Chile, doi:10.4067/S0717-97072022000305623.

[116] Moussaoui, Mohamed, et al. "QSAR, ADMET, molecular docking, and dynamics studies of 1,2,4-triazine-3(2H)-one derivatives as tubulin inhibitors for breast cancer therapy." Scientific Reports, vol. 14, 2024, p. 16418. Nature, doi:10.1038/s41598-024-66877-2.

[117] IAAnalysis. (n.d.). Principles, processes and types of molecular docking. Retrieved June 21, 2025, from https://www.iaanalysis.com/principles-processes-and-types-of-molecular-docking.html

[118] Chen, Z., Min, M.R., Parthasarathy, S. et al. A deep generative model for molecule optimization via one fragment modification. Nat Mach Intell **3**, 1040–1049 (2021). https://doi.org/10.1038/s42256-021-00410-2

[119] Quantum DFT Analysis and Molecular Docking Investigation of Various Potential Breast Cancer Drugs - Journal of Materials Chemistry B (RSC Publishing) DOI:10.1039/D4TB01803F.

REFERENCE

[120] Adasme, Melissa F., et al. "PLIP 2021: Expanding the Scope of the Protein-Ligand Interaction Profiler to DNA and RNA." Nucleic Acids Research, vol. 49, no. W1, 2021, pp. W530-34. PubMed Central, doi:10.1093/nar/gkab294.

[121] Kochnev, Yuri, et al. "MolModa: accessible and secure molecular docking in a web browser." Nucleic Acids Research, vol. 52, no. W1, 2024, pp. W498-W506. Oxford Academic, doi:10.1093/nar/gkae406.

[122] Frisch, M. J., et al. (2009). Gaussian 09, Revision D.01. Gaussian, Inc., Wallingford CT. Open Journal of Inorganic Chemistry, vol. 6, no. 2, 2016, pp. 39–50. https://doi.org/10.4236/ojic.2016.62004.

[123] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., et al. (2016). Gaussian 16, Revision C.01. Gaussian, Inc., Wallingford CT. Gaussian.com, gaussian.com/citation/.

[124] Young, David. Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems. John Wiley & Sons, Inc., 2001.

[125] van Mourik, Tanja, Michael Bühl, and Marie-Pierre Gaigeot. "Density Functional Theory across Chemistry, Physics and Biology." Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, vol. 372, no. 2011, 2014, p. 20120488. Royal Society Publishing, doi:10.1098/rsta.2012.0488.

[126] What Is Density Functional Theory and How Does It Work?" Synopsys, www.synopsys.com/glossary/what-is-density-functional-theory.html. Accessed 21 June 2025.

[127] Pradhan, Tathagata, et al. "Identification of novel thiazolidine-4-one based hits as potential PPARγ modulators through in silico workflow and validation through in vitro studies." Journal of Molecular Structure, vol. 1339, 2025, p. 142391. doi:10.1016/j.molstruc.2025.142391.

[128] EL-Mernissi, Reda, et al. "3D-QSAR, molecular docking, ADMET, simulation dynamic, and retrosynthesis studies on new styrylquinolines derivatives against breast cancer." Open Chemistry, vol. 22, no. 1, 2024, p. 20240041. De Gruyter, doi:10.1515/chem-2024-0041..

[129] Cheng, Y., & Prusoff, W. H. (1973). Relationship between the inhibition constant (Ki) and the concentration of inhibitor which causes 50 percent inhibition (IC50) of an enzymatic reaction. Biochemical Pharmacology, 22(23), 3099–3108.

REFERENCE

[130] Subramani, Arun Kumar, et al. "QSAR and Molecular Docking Studies of Pyrimidine-Coumarin-Triazole Conjugates as Prospective Anti-Breast Cancer Agents." Molecules, vol. 27, no. 6, 2022, p. 1845. PubMed, doi:10.3390/molecules27061845.

[131] Abbas, Hassan Ahmed. "Phytochemical, antioxidant, and anticancer assessments of Atriplex halimus extracts: In silico and in vitro studies." Heliyon, vol. 9, no. 10, 2023, article e20078. ScienceDirect, doi:10.1016/j.heliyon.2023.e20078.

[132] Dassault Systèmes. "BIOVIA Discovery Studio." Dassault Systèmes, www.3ds.com/products/biovia/discovery-studio.

[133] Shao, Xiaochun, et al. "Applications of BIOVIA Materials Studio, LAMMPS, and GROMACS in Various Fields of Science and Engineering." Materials Modeling: From First-Principles to Calphad and Beyond, Elsevier, 2019, pp. 209-40. doi:10.1016/B978-0-12-816954-4.00007-3.

[134] Roubi, Mohammed, et al. "Phytochemical, Antioxidant, and Anticancer Assessments of Atriplex halimus Extracts: In Silico and In Vitro Studies." Scientific African, 2023. ScienceDirect, doi:10.1016/j.sciaf.2023.e01959. Accessed 21 June 2025.

[135] Daina, Antoine, et al. "SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules." Scientific Reports, vol. 7, 2017, p. 42717. Expasy, doi:10.1038/srep42717.

[136] Alomar, Hatun A., et al. "Computational antidiabetic assessment of Salvia splendens L. polyphenols: SMOTE, ADME, ProTox, docking, and molecular dynamic studies." Results in Chemistry, vol. 7, 2025, p. 102081. doi:10.1016/j.rechem.2025.102081.

[137] Vasilev, Boris, and Mariyana Atanasova. "A (Comprehensive) Review of the Application of Quantitative Structure–Activity Relationship (QSAR) in the Prediction of New Compounds with Anti-Breast Cancer Activity." Applied Sciences, vol. 15, no. 3, 2025, p. 1206. MDPI, doi:10.3390/app15031206.

[138] Gramatica, Paola, et al. "Predicting the Bioconcentration Factor (BCF) of Organic Chemicals: Integration of a Hierarchical QSAR Approach into a Web-Based Read-Across System." Molecular Informatics, vol. 30, no. 1, 2011, pp. 27-37. doi:10.1002/minf.201000061.

[139] Daina, A., Michielin, O. & Zoete, V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep **7**, 42717 (2017). https://doi.org/10.1038/srep42717

REFERENCE

[140] E. Anderson, G.D. Veith, D. Weininger, SMILES: A line notation and computerized interpreter for chemical structures. Report No. EPA/600/M-87/021. U.S. Environmental Protection Agency, Environmental Research Laboratory-Duluth, Duluth, MN 55804, 1987.

[141] Weininger, David. "SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules." Journal of Chemical Information & Computer Sciences, vol. 28, no. 1, 1988, pp. 31-36. doi:10.1021/ci00057a005.

[142] Weininger, David, Arthur Weininger, and Joseph L. Weininger. "SMILES. 2. Algorithm for generation of unique SMILES notation." Journal of Chemical Information and Computer Sciences, vol. 29, no. 2, 1989, pp. 97-101. doi:10.1021/ci00062a008.

[143] Quantitative Structure-Activity Relationship (QSAR) Modeling to Predict the Transfer of Environmental Chemicals across the Placenta - ScienceDirect', accessed 17 June 2025, https://www.sciencedirect.com/science/article/pii/S2468111321000578.

[144] 'Computational Toxicology - ScienceDirect', accessed 17 June 2025, https://www.sciencedirect.com/science/article/abs/pii/B012369400000260X.

[145] Hansch, Corwin, and Rajeshwar P. Verma. "A QSAR Review on Melanoma Toxicity." Bioorganic & Medicinal Chemistry, vol. 13, no. 19, 2005, pp. 5508-26. doi:10.1016/j.bmc.2005.06.044..

[146] Roy, Kunal, editor. Cheminformatics, QSAR and Machine Learning Applications for Novel Drug Development. Academic Press, an imprint of Elsevier, 2023. doi:10.1016/C2022-0-00080-5.

[147] Cherkasov, Artem, et al. "QSAR Modeling: Where Have You Been? Where Are You Going To?" Journal of Medicinal Chemistry, vol. 57, no. 12, 2014, pp. 4977–5010. ACS Publications, doi:10.1021/jm4004285.

[148] Organisation for Economic Co-operation and Development. Guidance Document on the Validation of (Quantitative) Structure-Activity Relationship [(Q)SAR] Models. OECD Series on Testing and Assessment, no. 69, OECD Publishing, 2014. doi:10.1787/9789264085442-en.

[149] Tropsha, A., and A. Golbraikh. "Predictive QSAR Modeling Workflow, Model Applicability Domains, and Virtual Screening." Current Pharmaceutical Design, vol. 13, no. 34, 2007, pp. 3494-504. PubMed, doi:10.2174/138161207782794257.

[150] Bajusz, Dávid, Anita Rácz, and Károly Héberger. "Chemical Data Formats, Fingerprints, and Other Molecular Descriptions for Database Analysis and Searching."

REFERENCE

Comprehensive Medicinal Chemistry III, edited by Samuel Chackalamannil, et al., Elsevier, 2017, pp. 329-78. doi:10.1016/B978-0-12-409547-2.12345-5.

[151] Gasteiger, Johann. "Molecular Descriptors." Chemoinformatics: From Concept to Practice, Springer, 2017, pp. 119-27. doi:10.1007/978-3-319-27282-5_51.

[152] Cai, Zhengguo, et al. "Quantitative Structure–Activity Relationship (QSAR) Study Predicts Small-Molecule Binding to RNA Structure." Journal of Medicinal Chemistry, vol. 65, no. 10, 2022, pp. 7262-77. doi:10.1021/acs.jmedchem.2c00254.

[153] Todeschini, Roberto, and Viviana Consonni. Molecular Descriptors for Chemoinformatics. Wiley-VCH, 2009. Wiley Online Library, doi:10.1002/9783527628766.

[154] Smola, Alex, and Bernhard Schölkopf. "A tutorial on support vector regression." Statistics and Computing, vol. 14, no. 3, 2004, pp. 199-222. Springer Nature, doi:10.1023/B:STCO.0000035301.49549.88.

[155] Support Vector Regression (SVR) Model: A Regression-Based Machine Learning Approach." Analytics Vidhya, Medium, medium.com/analytics-vidhya/support-vector-regression-svr-model-a-regression-based-machine-learning-approach-f4641670c5bb.

[156] Veerasamy, Ravichandran, et al. "Validation of QSAR Models - Strategies and Importance." International Journal of Drug Design and Discovery, vol. 2, no. 3, 2011, pp. 511-19. ResearchGate, doi:10.37285/ijddd.2.3.1.

[157] Araújo, Gustavo Henrique de Almeida, et al. "A Critical Review on the Use of the QSPR/ QSAR Approach in the Prediction of the Physicochemical and Biological Properties of Organometallic Compounds." Journal of the Brazilian Chemical Society, vol. 20, no. 4, 2009, pp. 683–695. https://doi.org/10.1590/S0103-50532009000400021

[158] Madani, Achouak, Othmane Benkortbi, and Maamar Laidi. "In silico prediction of the inhibition of new molecules on SARS-CoV-2 3CL protease by using QSAR: PSOSVR approach." Brazilian Journal of Chemical Engineering, vol. 41, no. 1, 2023, pp. 427-42. Springer Nature, doi:10.1007/s43153-023-00332-z.

[159] **Tropsha, A.** (2010). Best practices for QSAR model development, validation, and exploitation. **Molecular Informatics**, 29(6-7), 476–488.

[160] Gramatica, P. (2007). Principles of QSAR model validation: Internal and external. **QSAR & Combinatorial Science**, 26(5), 694–701.

REFERENCE

[161] Roy, Kunal, Supratik Kar, and Rudra Narayan Das. Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment. Academic Press, 2015.

[162] Türker, İzzet, and Hilal İşleroğlu. "EXTRACTION OF BIOACTIVE COMPOUNDS FROM FENUGREEK LEAVES BY MACERATION WITH D-OPTIMAL DESIGN." GIDA / The Journal of Food, vol. 48, no. 2, 2023. DergiPark, dergipark.org.tr/en/pub/gida/issue/75830/1199417.

[163] "UV-Vis Spectroscopy: Principle, Strengths and Limitations and Applications." Technology Networks, 30 June 2021, www.technologynetworks.com/analysis/articles/uv-vis-spectroscopy-principle-strengths-and-limitations-and-applications-349865.

[164] Jee, Krishna, et al. "A SHORT REVIEW ON ULTRAVIOLET AND VISIBLE SPECTROSCOPY." International Journal of Scientific Development and Research, vol. 7, no. 12, 2022, pp. 683-87. IJSDR, doi:10.1729/Journal.32451

[165] Stuart, B. H. Infrared Spectroscopy: Fundamentals and Applications. Analytical Techniques in the Sciences, Wiley, 2004. doi:10.1002/0470011149.

[166] "Infrared Spectroscopy- Definition, Principle, Parts, Uses." Science Info, 13 July 2022, scienceinfo.com/infrared-spectroscopy/.

[167] Brand-Williams, W., et al. "Use of a Free Radical Method to Evaluate Antioxidant Activity." LWT - Food Science and Technology, vol. 28, no. 1, 1995, pp. 25-30. doi:10.1016/S0023-6438(95)80008-5.

[168] Reza Ghafarzadegan and others, 'Process Optimization for Green Synthesis of Iron Nanoparticles by Extract of Fenugreek (Trigonella Foenum-Graecum L.) Seeds', Journal of Medicinal Plants, 21.81 (2022), pp. 22–32, doi:10.52547/jmp.21.81.22.

[169] ComputaBio, "Molecular Docking Service," accessed June 20, 2025, https://www.computabio.com/molecular-docking-service.html.

[170] STRING Consortium. STRING: Functional Protein Association Networks. SIB Swiss Institute of Bioinformatics, Novo Nordisk Foundation Center for Protein Research, European Molecular Biology Laboratory, 2023. https://string-db.org/cgi/about

[171] Liu, Y., et al. "CB-Dock: a web server for cavity detection-guided protein–ligand blind docking." Acta Pharmacologica Sinica, vol. 41, no. 1, 2020, pp. 138-44. doi:10.1038/s41401-019-0228-6. http://clab.labshare.cn/cb-dock/ .

REFERENCE

[172] Lawal, Hadiza Abdulrahman, et al. "QSAR, molecular docking studies, ligand-based design and pharmacokinetic analysis on Maternal Embryonic Leucine Zipper Kinase (MELK) inhibitors as potential anti-triple-negative breast cancer (MDA-MB-231 cell line) drug compounds." Bulletin of the National Research Centre, vol. 45, no. 90, 2021. SpringerLink, doi:10.1186/s42269-021-00541-x.

[173] A Combined 2D- and 3D-QSAR Study, Design and Synthesis of Some Monocarbonyl Curcumin Analogs as Potential Inhibitors of MDA-MB-231 Breast Cancer Cells', accessed 16 May 2025, https://www.mdpi.com/2673-4583/12/1/5.

[174] Nhlapho, Samukelisiwe, et al. "Druggability of Pharmaceutical Compounds Using Lipinski Rules with Machine Learning." Sciences of Pharmacy, vol. 3, no. 4, 2024, pp. 177-92. doi:10.58920/sciphar0304264.