Université Mohamed Khider de Biskra
Entrez votre faculté
Entrez votre département

MASTER'S GRADUATION THESIS

Domaine:Matter Science
Spinneret:Chemistry
Field: PharmaceuticalChemistry
Réf. : Entrez la référence du document

Présenté et soutenu par :
Zohra SAIDANE

Le :
/06/2023

A computational study of natural inhibitors as potential Anti-cancer Agents

---

Jury :

Mme. Dalal HARKATI	Pr.	University of Biskra	Supervisor
Mme. Afaf ZEKRI	Dr.	University of Biskra	President
Mme. Habiba LARAOUI	Dr.	University of Biskra	Examiner

Année universitaire : 2022/2023

To my dear mother and sister

To my brother

To my dear Dr.Djebbari

To all my colleagues in the medical laboratory of Dr. Djebbari Eps Ben cheaah.

To all those who are dear to me

Acknowledgements

First of all, I pray to Allah, the Almighty for providing me this opportunity and granting me the capability to proceed successfully.

I would like to thank **Dr. Dalal HARKATI** Professor at the University of Biskra, for suggesting this topic to me and for taking the lead in this work. I would like to thank you for stimulating and guiding me throughout my research work, for the trust you have placed in me during this time for which you are responsible. His skills, scientific rigor and availability have always motivated me to do this job. She was also a valuable advisor in answering my different questions and writing this thesis.

I would like to express my deepest gratitude to **Dr. Basil A. SALEH**, Professor at the University of Basrah, Iraq, for his kind help, support, advice, collaboration and availability to evaluate my work.

I wish to extend my sincere thanks to **Dr. Afaf ZEKRI**, Professor at the University of Biskra, for accepting to chair the dissertation of my thesis.

I would like also to thank**Dr. Habiba LARAOUI,** Professor at the University of Biskra, for accepting to examinat this thesis.

I am very grateful to **Dr . Hadjer KHELFAOUI** who helps me during this research.

Finally, I will forever be grateful to my family, especially my mother and sister for supporting me over the years and always being available for when I needed them.

Abstract

Breast cancer stands as one of the most frequently diagnosed cancers among women, with a noteworthy 80% of cases being positive for Estrogen Receptor α (ERα). This receptor subtype holds significant therapeutic importance, prompting our study's focus on it. Selective estrogen receptor modulators (SERMs), such as tamoxifen, and selective estrogen receptor degraders (SERDs), such as fulvestrant, directly target ERα. To explore potential compounds, we turned to *Citrullus colocynthis*, which contains cucurbitacins and flavonoids. Through molecular docking, we accurately assessed their affinity to bind with ERα. Subsequently, we predicted the ADMET properties and chemical reactivity descriptors of these compounds. Our computational approach suggests that Cucurbitacine J and 2S-3',4'-methylene dioxy-5,7-dimethoxy flavan exhibit favorable inhibitory characteristics against the ERα receptor, providing an alternative pathway for understanding the binding mechanism of these compounds.

Keywords: Computer-aided drug design, Estrogen receptor alpha (Erα), Breast cancer, Natural Compounds, Virtual screening, Lead compound, Molecular docking, ADMET, Chemical reactivity descriptor, *Citrullus colocynthis*, Cucurbitacines, flavonoids.

Résumé

Le cancer du sein est l'un des cancers les plus fréquemment diagnostiqués chez les femmes, avec un taux remarquable de 80% des cas présentant une positivité pour le récepteur de l'œstrogène α (ERα). Ce sous-type de récepteur revêt une importance thérapeutique considérable, ce qui motive notre étude à se concentrer sur celui-ci. Les modulateurs sélectifs du récepteur de l'œstrogène (SERM), tels que le tamoxifène, et les dégradeurs sélectifs du récepteur de l'œstrogène (SERD), tels que le fulvestrant, ciblent directement l'ERα. Pour explorer les composés potentiels, nous nous sommes tournés vers *Citrullus colocynthis*, qui contient des cucurbitacines et des flavonoïdes. Par le docking moléculaire, nous avons évalué avec précision leur affinité de liaison avec l'ERα. Ensuite, nous avons prédit les propriétés ADMET et les descripteurs de réactivité chimique de ces composés. Notre approche computationnelle suggère que la Cucurbitacine J et le flavane 2S-3',4'-méthylènedioxy-5,7-diméthoxy présentent des caractéristiques inhibitrices favorables vis-à-vis du récepteur ERα, offrant ainsi une voie alternative pour comprendre le mécanisme de liaison de ces composés.

Mots clés : Conception de médicaments assistée par ordinateur, Récepteur des œstrogènes alpha (Erα), Cancer du sein, Composés naturels, Criblage virtuel, Composé principal, Docking moléculaire, ADMET, Descripteur de réactivité chimique, Citrullus colocynthis, Cucurbitacines, flavonoïdes.

يعتبر مرض سرطان الثدي من أكثر امراض السرطان شهرة التي تصيب النساء. فقد لوحظ بان 80% من المصابات تكون بسبب مستقبلات الهرمون الاستروجين الايجابي (ERα+) مما يجعلها هدفاً لتصميم الدواء المضاد لهذا النوع من السرطانات. وللوصول لهذا الهدف العلاجي يتم اختيار مستقبلات الهرمون الاستروجين الانتقائي المحور (SERM) مثل تاموكسيفين ومستقبلات الهرمون الهدام مثل فالفسترانت. ولهذا الغرض تم اختيار *Citrullus colocynthis* التي تحتوي على مركبات من نوع cucurbitacins و flavonoid. ثم تم تقييم ارتباط هذه المركبات بالمستقبل (ERα+) من خلال دراسة الارساء الجزيئي لها. كما تم دراسة ADMET للمركبات المختارة. من خلال دراستنا هذه يمكن توفير طرق بديلة للحصول على مثبطات مستقبلات ERα+ مشتقة من Cucurbitacine J و S-3',4'-methylene dioxy-2 5,7-dimethoxy flavan

الكلمات المفتاحية: تصميم الأدوية بمساعدة الحاسوب، مستقبل الاستروجين ألفا(Erα) ، سرطان الثدي، مركبات طبيعية، فحص افتراضي، مركب قيادي، ربط جزيئي، ADMET، وصف تفاعلية كيميائية، *Citrullus colocynthis*، الكوكوربيتاسين، الفلافونويدات.

CONTENT :

LIST ABBREVIATIONS

A

ADMET: Absorption, Distribution, Metabolism, Elimination, and Toxicity

ACE: Angiotensin-Converting Enzyme

AP1: Activator protein 1

Ala: Alanine

Asp: Aspartic acid

Arg: Arginine

C

CADD: Computer aided drug design

CASTp: Computed Atlas of Surface Topography of proteins

CA: Carbonic Anhydrase

D

DNA: Deoxyribonucleic acid

DFT: Density Functional Theory

DBD:DNA binding

E

ERα: Estrogen receptor alpha

ERβ: Estrogen receptor beta

EHT: Extended Hückel Theory

G

GF: Growth factor

GFR: Glomerular filtration rate

Glu: Glutamic acid

Gly: Glycine

H

HBD: Hydrogen bond donor

HBA: Hydrogen bond acceptor

HUMO: Highest occupied molecular orbital

HIV: Human immunodeficiency viruses

His: Histidine

I

Ile: Isoleucine

L

LBDD: Ligand-based drug design

LBVS: ligand based virtual screening

LUMO: Lowest unoccupied molecular orbital

LBD: Ligand binding

LHRH: Luteinizing hormone-releasing hormone

Leu: Leucine

M

MAPK: Mitogen-activated protein kinase

MOE: Molecular operating environment

Met: Methionine

N

NMR: Nuclear magnetic resonance

P

PDB: Global protein data bank

PI3K: Phosphoinositide 3-kinase

Pro: Proline

Phe: Phenylalanine

Q

QSAR: Quantitative structure-activity relationship

R

RNA: Ribonucleic acid

RMSD: Root-Mean Squared Deviation

S

SBDD: Structure-Based Drug Design

SBVS: Structure-based virtual screeing

SLN:sentinel lymph node

SP1:specificity protein 1

SERM: Selective estrogen receptor modulator

SERD: Selective estrogen receptor degrader

T

TCE:tetracyanoethylene

Thr: Threonine

Trp: Tryptophan

ΔE:energy gap

ω:electrophilicity

▪: chemical hardness

μ: chemical potential

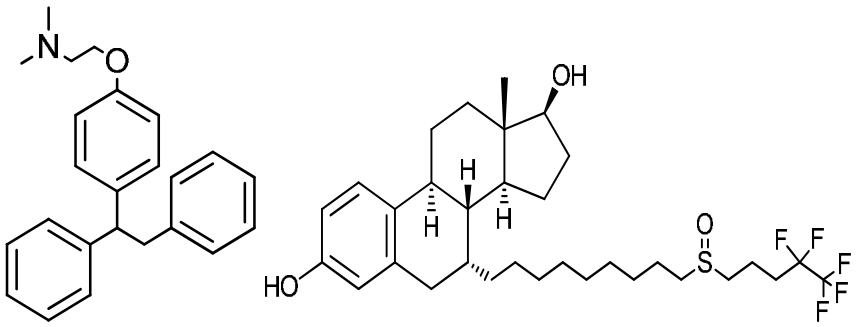
LIST OF FIGURES:

LIST OF TABLES :

GENERAL INTRODUCTION

GENERAL INTRODUCTION

Breast cancer is a prevalent form of cancer among women worldwide. It can be categorized into different subtypes based on the presence of specific tumor receptors, namely estrogen receptor-positive (ER+), human epidermal growth factor receptor 2-positive (HER2+), and triple-negative subtypes. ER+ breast cancer accounts for approximately 80% of newly diagnosed cases. Endocrine therapy, which involves modifying estrogen production, is the primary treatment approach for this subtype. The estrogen receptor (ERα) has been a target for breast cancer treatment since the 1960s, leading to the discovery of selective estrogen receptor modulators (SERMs) and selective estrogen receptor degraders (SERDs). These compounds, such as tamoxifen (a SERM) and fulvestrant (a SERD), directly target the ERα and have been extensively studied in the field. [1,8]



Tamoxifen Fulvestrant

Figure 1:Structure of Tamoxifen and Fulvestrant molecules.

Medicinal plants have been used in healthcare for a long time, and their use to prevent and treat illness is expanding worldwide. The medicinal properties of plants are due to the natural chemicals/compounds they contain. Plants are a source of food and act as raw materials from which a variety of drugs are synthesized. Citrullus colocynthis is a desert plant and a source of several bioactive compounds such as essential oils, glycosides, cucurbitacins, flavonoids, alkaloids, and fatty acids. Citrullus colocynthis also has excellent pharmacological properties, such as being a laxative and purgative; it is anti-diabetic, anti-inflammatory, anthelmintic, and anti-cancerous.[9,10].

Recently, there has been a growing interest in in silico approaches due to their ability to expedite the process of drug discovery in terms of time, labor, and costs. Computational

methods have proven successful in the development of numerous new drug compounds. Conversely, the prediction and comprehension of properties and binding affinity of chemical compounds hold significant importance from both technological and academic perspectives. [11,12].

This study aims to shed light on the potential properties of Citrullus colocynthis using theoretical approaches, contributing to a better understanding of its implications within our body system.Our research is divided onto three chapters:

- ✓ The first chapter is devoted to a literature review covering some hand knowledge of computer aided drug design, an overview of breast cancer and estrogen receptor alpha in humen breast cancer.
- ✓ The second chapter highlights the techniques and methods theoretical assays used or applied during our study.
- ✓ The third chapter illustrates a theoretical approach including optimization and docking investigation of the chosen bioactive compounds.

And finally a general conclusion that summarizes the essence of our research.

References

1. Tong, C. W. S.; Wu, M.; Cho, W. C. S.; To, K. K. W. Recent advances in the treatment of breast cancer. Front. Oncol. 2018, 8, 227.
2. Anderson, W. F.; Katki, H. A.; Rosenberg, P. S. Incidence of breast cancer in the United States: current and future trends. J. Natl. Cancer Inst. 2011, 103, 1397−1402.
3. Ariazi, E. A.; Jordan, V. C. M. *Curr. Top. Med. Chem.* **2006**, 6, 203.
4. Levenson, A. S.; Jordan, V. C. *Eur. J. Cancer* **1999**, 35, 1628.
5. Johnston, S. R. D. *Breast Cancer Res.* **2005**, 7,119.
6. Blizzard, T. A. *Curr. Top. Med. Chem.* **2008**, 8, 792.
7. Jordan, V. C. Tamoxifen: a most unlikely pioneering medicine. Nat. Rev. Drug Discovery 2003, 2, 205−213.
8. Das, S.; Crockett, J. C. Osteoporosis - a current view of pharmacological prevention and treatment. Drug Des., Dev. Ther. 2013, 7, 435−448.
9. Dhama et al., 2018; Bilal et al., 2021; Reda et al., 2021; Saeed et al., 2021.
10. Saeed et al., 2019; Alagawany et al., 2021c; Garg et al., 2021; Zhang et al., 2021.
11. Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, Ruddy K, Tsang J, Cardoso F. Breast cancer. Nat Rev Dis Primers. 2019 Sep 23;5(1):66. doi: 10.1038/s41572-019-0111-2. PMID: 31548545.
12. Surabhi et al Journal of Drug Delivery & Therapeutics. 2018; 8(5):504-509.

Chapter I:

LITERARY REVIEW

Chapter I: LITERARY REVIEW

1 COMPUTER-AIDED DRUG DESIGN

Computational approaches to drug finding and incident advance through brisk finding and achievement of surprises. Bringing a new drug to display is a very complex process that is to say dangerous and high-priced in agreements momentary services and exertion. It is widely recognized that the process of drug finding and happening takes nearly 10-14 age and is as well $1 billion in capital .

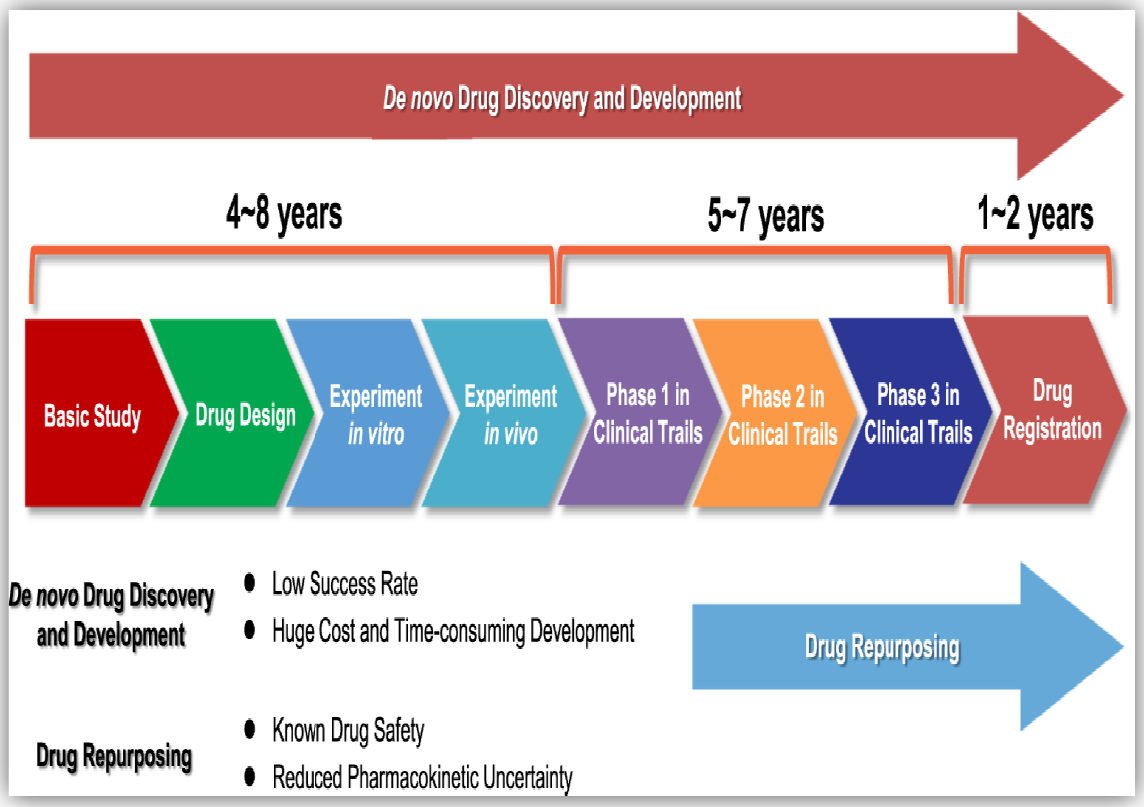


Figure 2: Traditional process of drug discovery and development.

So, for lowering period, cost and risk carried determinants computer aided drug design (CADD) pattern is established as a new drug design approach. It has existed visualized that utilizing CADD approaches we can lower the cost of drug finding and incident until 50%[2]. CADD exist use of some operating system program-located process for establishing a standard to pertain project to construction [1] .

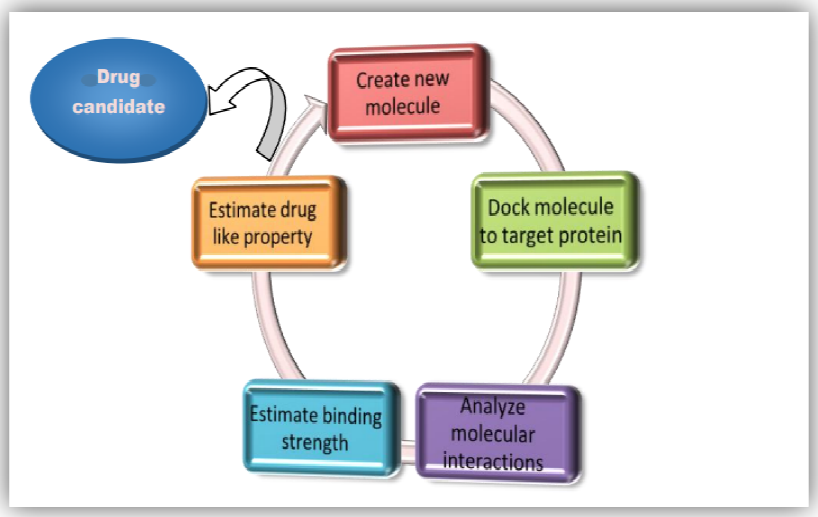


Figure 3: General principal for drug design through CADD

A brief history of CADD

- ✓ In 1900: the concept of receptor and lock-and-key was given by P. Ehrich (1909) and E. Fisher (1894).
- ✓ 1970s: the concept of Quantitative structure-activity relationships (QSAR) was established, it had Limitations: 2- Dimensional, retrospective analysis.
- ✓ 1980s: Beginning of an era of CADD Molecular Biology, X-ray crystallography, multi-dimensional NMR Molecular modeling along with computer graphics.
- ✓ 1990s: modern techniques like Human genome Bioinformatics along with combinatorial chemistry and High-throughput screening were introduced in the world of innovative medical science [2].

1.1 Steps of computer aided drug design

The computer aided drug design process consists of three steps[3].

Step 1: Generate a heterogeneous small molecule library for therapeutic target identification and testing. There is development of virtual screening protocols that are enabled by docking of small molecules.

Step 2: Confirm the specificity of the selected hits by docking at the binding sites of other known drug targets.

Step 3: Selected leads are subjected to computational ADMET profile studies and those who pass these studies are called leads. Target identification is the critical first step in the drug discovery pipeline. The literature indicates that identifying optimal targets from thousands of candidate macromolecules is a challenging process that can be achieved through genetic analysis and pathway analysis.

Goal Validation: Goal setting requires rigorous evaluation to establish that goal modification is desirable. treatment effect. The target validation process determines whether a target modification has the desired therapeutic effect. Lead optimization: Leads can be identified with the help of techniques such as structure-based design. At this point the structure of the target protein in complex with the lead molecule is very useful in suggesting ways to improve the affinity of lead to the target. In this case the paths used may not be optimal and should be optimized to maximize selection at the target site. Optimization can be achieved by changing the structural properties.

1.2 Major types of approaches in CADD

There are types of approaches for drug design through CADD is the following-:

- Structure based drug design / direct approach.
- Legend based drug design / indirect approach.

1.2.1 Structure based drug design / direct approach:

Structure-based drug design is the technique to be used in drug design. Structure-based drug design helps in the discovery process of new drugs [3].

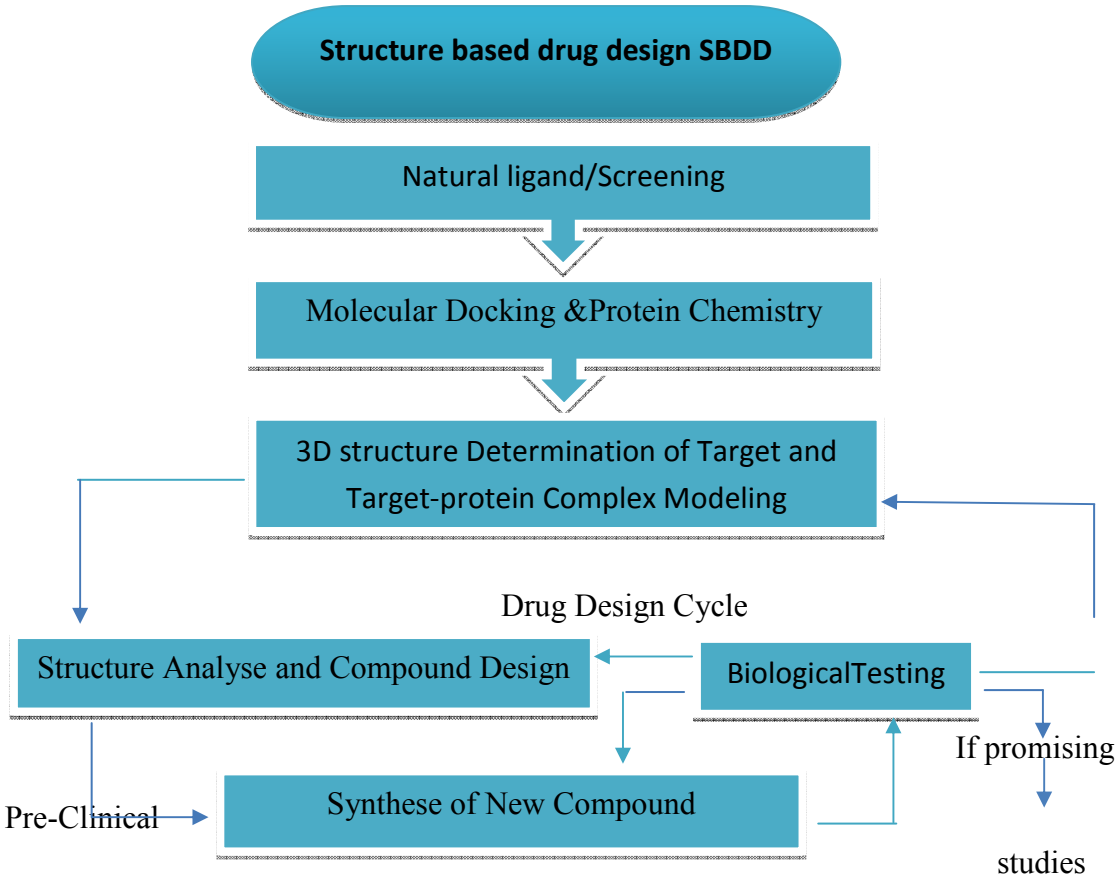


Figure 4: Schematic diagram of Structure-based drug design.

1.2.2 Legend based drug design / indirect approach

Ligand-based drug LBDD finding approaches include the study of ligands that interact accompanying goals. These fundamental orders use a set of connections collected from compound possessions to resolve a target of interest and allure 2D or 3D building for each additional. Alternatively drug finding maybe established a process that uses the target proteins binding characteristics as a beginning if facts about the mark proteins shape is not mainly vacant. This approach is popular as ligand-based drug finding[3].

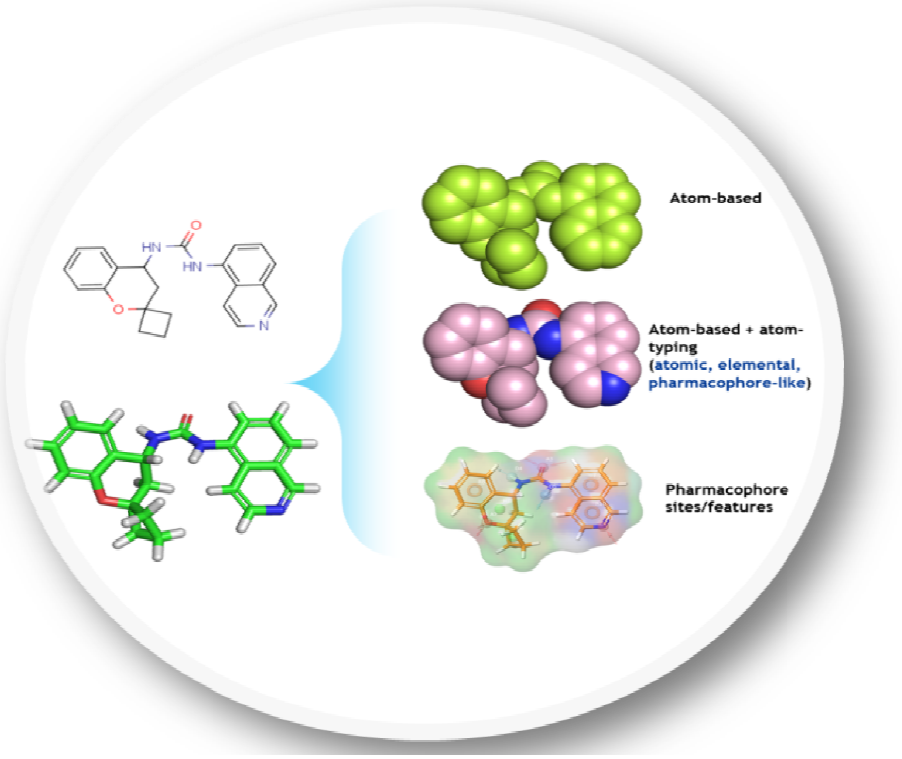


Figure 5: Ligand-based drug design

1.3 Virtual Screening

With the help of news about protein targets or known active ligands virtual screening has existed secondhand as the most appropriate form now to find ultimate appropriate bioactive compounds. Recently virtual scanning has happened thought-out as an alternative to extreme-throughput hide especially in conditions of cost and the feasibility to find the most acceptable novel by protect large complex book repositories.

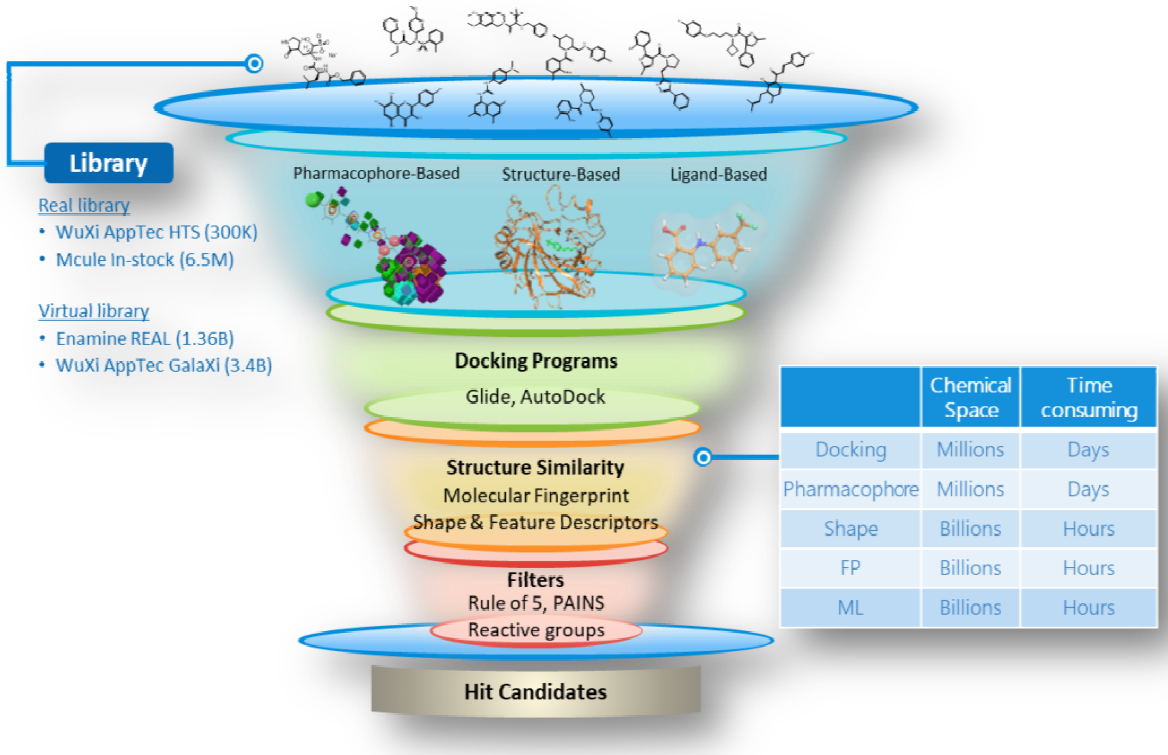


Figure 6: Overview of Virtual screening process

Generally skilled are two types of virtual screening orders as ligand-based virtual hide (SBVS) and ligand-based virtual screening (LBVS) in which the SBVS pattern depends on the active site form of the mark protein during the LBVS plan. Computational likeness estimates are based on popular active ingredients and compounds got from databases [4].

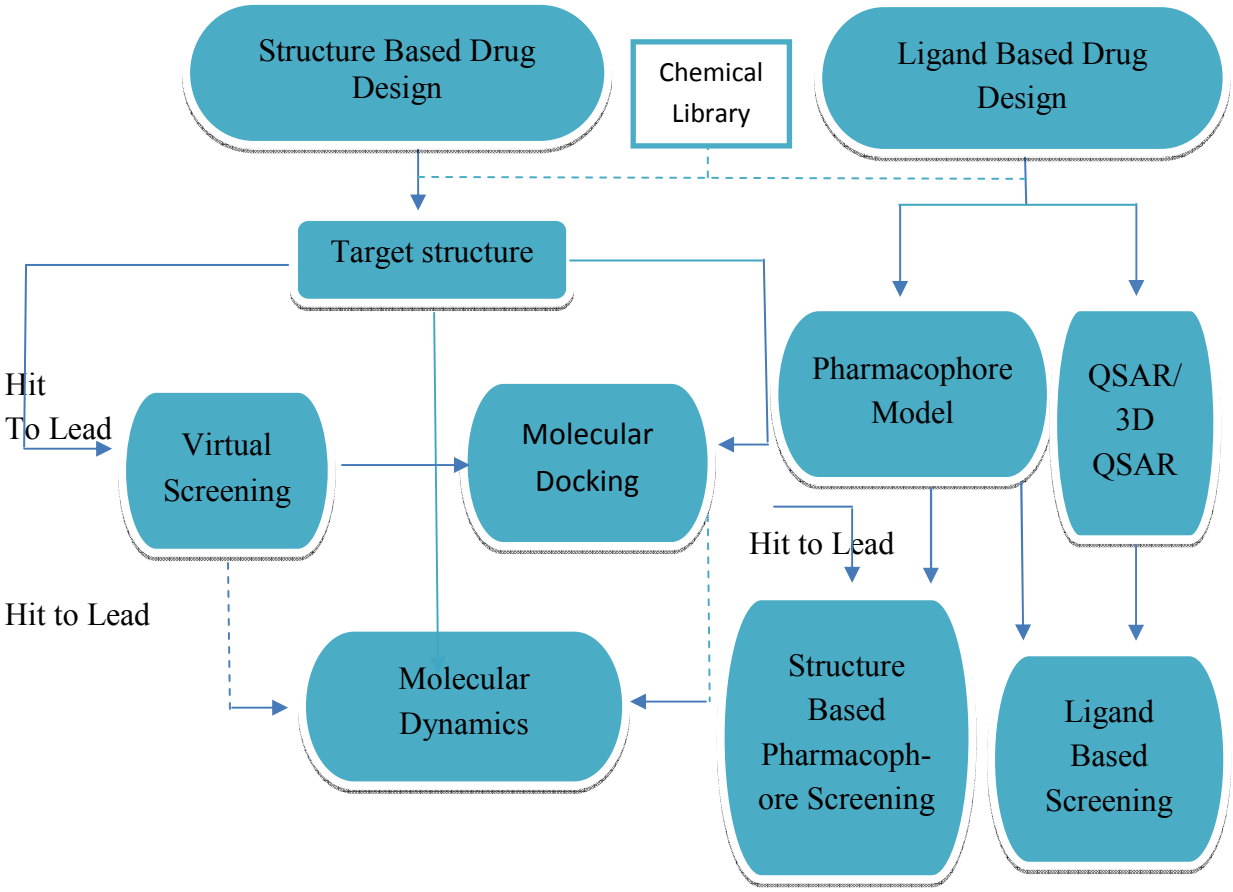


Figure 7: Schematic diagram of VS process for SBDD & LBDD

1.4 Molecular Docking

Docking aims to precisely match the ligand structure to the requirements of the receptor binding site and accurately estimate the binding strength (Adrian-Scotto and Vasilescu 2008). Blind docking is useful when the binding site of the target protein structure is unknown because it considers the entire protein structure as the binding site region and docking site specificity helps predict interactions when the binding site is known. Properties of binding molecules. In general blind binding results in less precision and requires more time and memory than site-specific binding as it targets only selected amino acid residues present in the binding site cavity. However most of the available literature represents case studies in which molecular docking was used to solve specific problems related to ligand design or target recognition (Huang et al. 2010; Yuriev and Ramsland 2013; Pathak et al. 2016; Rana et al. 2019). A summary of widely cited molecular docking programs used in drug discovery is provided in Table 1 [5].

Table 1: summary of the highly cited molecular docking programs used in drug discovery

S. No.	Programs	Description	Availability	References
1	AutoDock	Used for molecular docking. It predicts the binding affinity and poses of a small molecule to a 3D structure target protein	https://autodock.scripps.edu/	Goodsell et al. (1996)
2	AutoDock Vina	Used for virtual screening and molecular docking	https://vina.scripps.edu/index.html	Trott and Olson (2010)
3	Glide Schrodinger	A complete package for molecular modeling and computer-aided drug discovery (CADD)	https://www.schrodinger.com/	Friesner et al. (2004)
4	Hex	Used for docking studies	http://hex.loria.fr/	Ritchie (2003)
5	Molecular operating environment (MOE)	A complete package for molecular modeling and computer-aided drug discovery	https://www.chemcomp.com/	Vilar et al. (2008)

1.4.1 Typsof molecular docking

1.4.1.1 Flexible Docking

In this model both ligand and receptor side chains remain flexible and binding energies are calculated for different poses of the ligand that fit the receptor. The adapter-induced backbone also shifts to incorporate conformational changes in the receptor upon ligand binding (Huang et al. 2010; Yuriev and Ramsland 2013). Although time consuming and computationally intensive this method can estimate many different possible conformations making it more comprehensive more likely to simulate real phenomena and more reliable. Therefore flexible docking is considered a better method as it offers better predictability than conventional docking. The main drawback of other docking methods is that docking scores can be low due to incorrect ligand binding modes [5].

1.4.1.2 Le docking semi-flexible

The semi-flexible connection allows for more accurate results. Semi-flexible molecular bonds with flexible and rigid acceptors. This anchor consists of two main steps. The first called docking allows the small molecule being tested to explore the conformational space around the protein. The molecules adopt different structures and different positions around the protein. The second stage called marking allows you to evaluate the shots you found during the first stage and keep only the best ones. This move generates a result [5].

1.4.1.3 Rigid Docking

This is another molecular docking method where the internal geometry of the ligand and receptor is maintained during the docking simulation.(Huang et al. 2010; Yuriev and Ramsland 2013; de Ruyck et al. 2016). The DOCK program based on rigid docking applied to the aspartic protease of HIV yielded a candidate inhibitor molecule with higher potency, and this molecule can be used as a lead for designing more powerful inhibitors. Besides, with simple bound and unbound target cases, ZDOCK correctly predicted 47% of interface contacts, demonstrating its strength in predicting binding sites. SOFTDOCK, on the other hand, predicted 66 of 83 (Pagadala et al. 2017) [5].

1.4.2 Molecular docking tools

1.4.2.1 Receptor

The first major approach to study and design bioactive molecules through molecular modeling is based on receptor structures. This approach is based on exploiting the three-dimensional molecular structure of the target protein. Three of his experimental methods nuclear magnetic resonance (NMR) electron microscopy and his X-ray crystallography now make it possible to determine the structure of proteins. This last technology is responsible for the bulk of the structures that result in a structural database called the Structure Bank of the Protein Database. (PDB). The PDB is a global repository of information on protein and nucleic acid conformation. These molecules come from all organic states. The PDB is freely accessible via the Internet (http://www.rcsb.org/pdb/). This includes thousands of protein structures obtained by X-ray crystallography or NMR.

If the target is not conserved at the library level and there are subsequent proteins with similar patterning sequences due to homology formation of his 3D structure of the target of interest is prevented [6].

1.4.2.2 ligand

Ligand selection is a very important step in molecular binding. This choice should be important because of the specificity of the targets active site avoiding unnecessary testing of molecules. The ligand must be in a 3D shape for molecular bonding. Currently there are two ways to obtain the chemical structure of a given ligand. the first is often commercial in nature consisting of databases of chemical structures called chemical libraries or chemical blanks. The second way is to use PDB or literature ligands that can be drawn in different formats (pdb mol mol2...etc) thanks to molecular structure software such as chemDraw Arguslab Titan or Sybyl...etc [6].

1.4.2.3 The Molecular Bonding Process

Molecular bonding is accomplished in two complementary steps. The first is to find the conformation of the ligand that makes the desired interaction with the receptor. The second is an evaluation function that allows these conformations to be evaluated by rapidly calculating the energy of their interaction with that receptor [7].

1.4.3 The docking process

Involves the docking process of allowing a small organic molecule with a receptor usually a protein in nature. Studies have shown that some algorithms are more reliable than others in producing experimental binding methods (GLIDE GOLD). Basically compensating for these techniques increases the computation time. It is therefore necessary to concentrate on the purpose. These slow and accurate methods are recommended if the goal is to screen a chemical library of ~10 compounds. On the other hand projects of virtual screening of millions of products cannot be used with such algorithms but can be carried out using a simpler code (LigandFit FlexX) with minimal savings in computer time. However the work of Warren et al. the predicted docking algorithm can generate conformations similar to those experimentally determined by crystallography [8].

1.4.3.1 Active site identification

A fundamental step in virtual screening design is the ligand recognition process. Crystallization of receptor binding can provide information about the active site. However this information should be interpreted with caution as the active site for the same receptor may vary depending on the pharmacological profile of the molecule (agonist antagonist inverse agonist). A method that is capable of detecting cavities is a good option if the ligand has not crystallized into the receptor. CASTp is said to measure the number of active sites. An algorithm available online measures easily accessible but inaccessible voids. Methods used include the estimation of solvation surfaces (Richard surfaces 17) and molecular surfaces (Connolly surfaces 18). Finally with CASTp you can measure the size and circumference of the opening. In general algorithms of this type are known to be possible in several domains. All of these potential and available binding sites will then be tested by the user against the known pharmaceutical shape of the molecule [9].

1.4.3.2 Intermolecular interactions

1.4.3.2.1 Ionic interactions (charge-charge interactions)

- At physiological pH the positive environment is provided by protons in the side chains of essential amino acids in proteins such as arginine.
- The anionic environment is usually provided by amino acids such as aspartic acid with acidic side chains.
- The drug particles may contain acidic and/or basic groups.
- The strength of ionic bonds is inversely proportional to the distance between the charges.

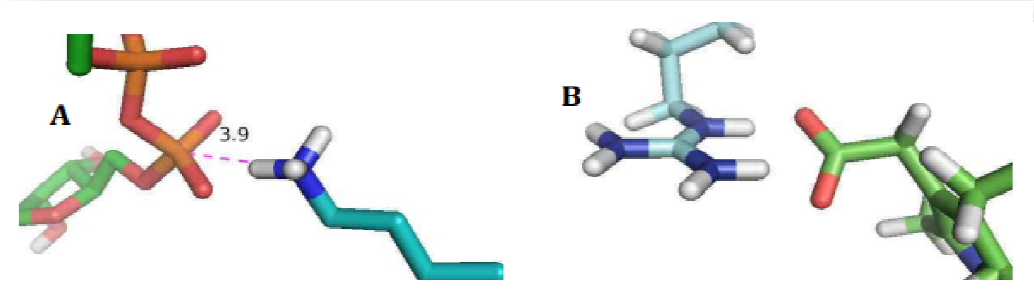


Figure 8: Example of ionic interaction in a protein complex (cyan) ligand (green) phosphate group and a quaternary amine. (PDB code: 2y27) B) Guanidinium case and carboxylic acid (PDB code: 3EQ1)

1.4.3.2.2 Ionic Dipole Interactions

- Molecular groups such as carbonyl (C=O) have permanent dipole moments due to the difference in electronegativity of the atoms in the group.
- Ionic dipole interactions are electrostatic interactions between ions and neutral groups and dipole.
- The dipole is represented by crossed arrows.
- The side ends represent the positive ends and the arrows represent the negative ends of the dipole.

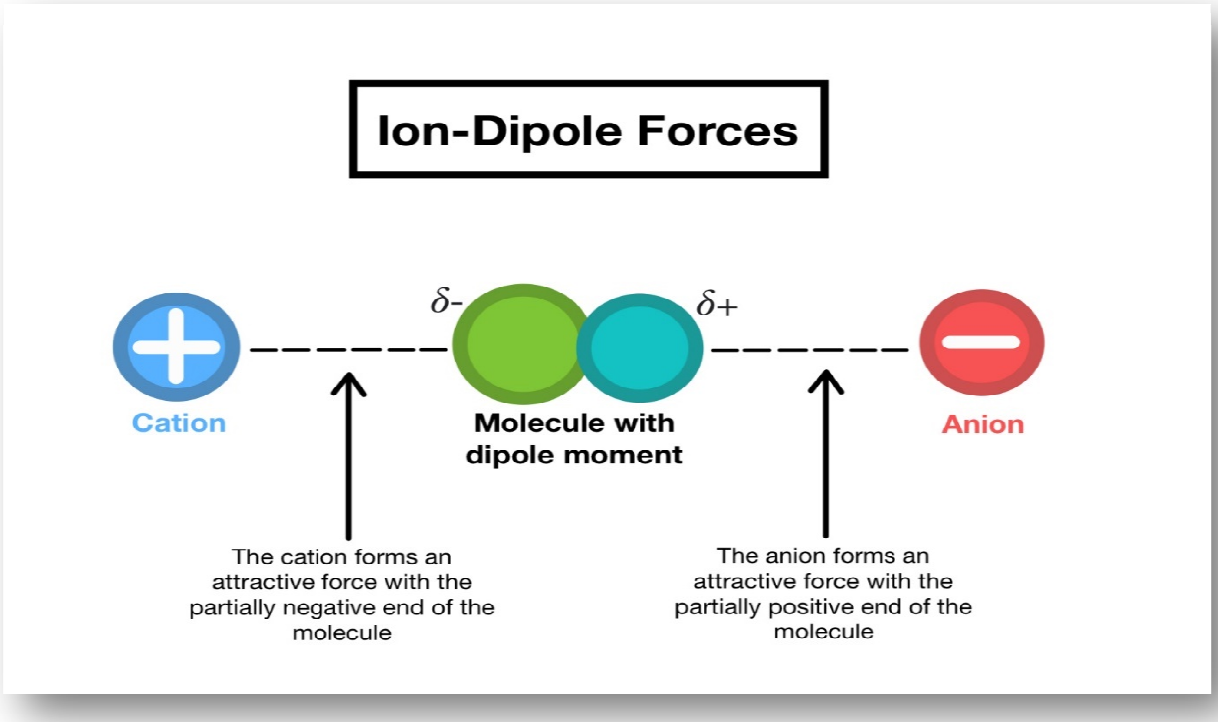


Figure 9:Interactions ion-dipole

1.4.3.2.3 Dipole-dipole interactions

Dipole-dipole interactions are electrostatic interactions between permanent dipoles.

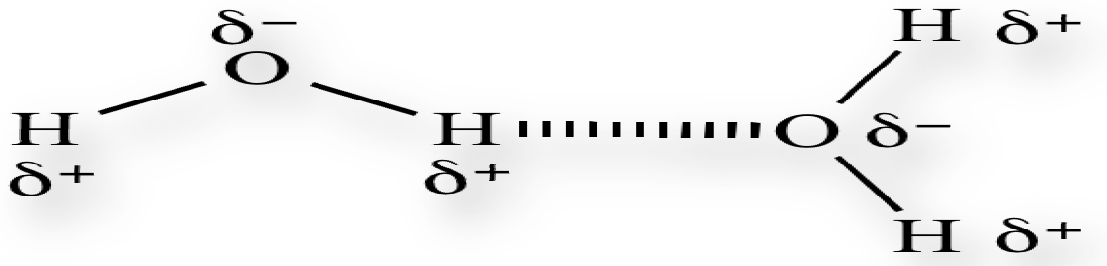


Figure 10: Interactions dipole-dipole

1.4.3.2.4 Ion-induced dipole interactions

Ion-induced dipole interactions occur when: The electric field of ions induces dipoles in non-polar molecules.

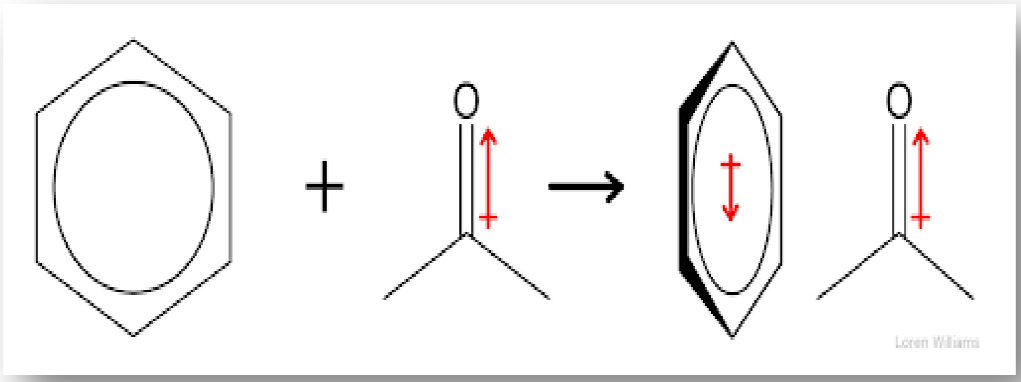


Figure 11: Ion-induced dipolar interactions

1.4.3.2.5 Hydrogen Bonding

Hydrogen bonding interactions are attractive interactions that include two groups:

Hydrogen bond donors (HBD): functional groups in which an electron-deficient hydrogen electron is covalently bonded to an atom.

Hydrogen bond acceptor (HBA): A functional group that contains an electron-rich heteroatom and a hydrogen bond acceptor.

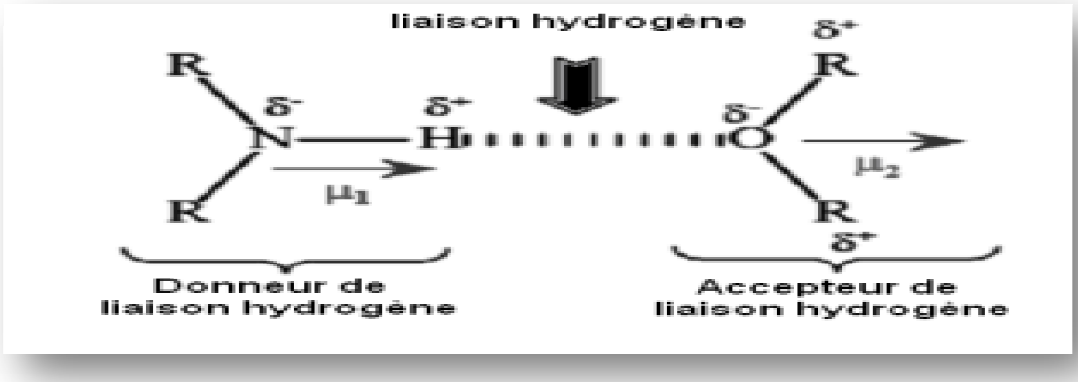


Figure 12: Liaison à l'hydrogène

1.4.3.2.6 π-π interactions Attractive non-covalent

interactions between aromatic rings are often supported by dispersion forces π-π interactions and hydrogen bonds are abundant in nucleic acids.

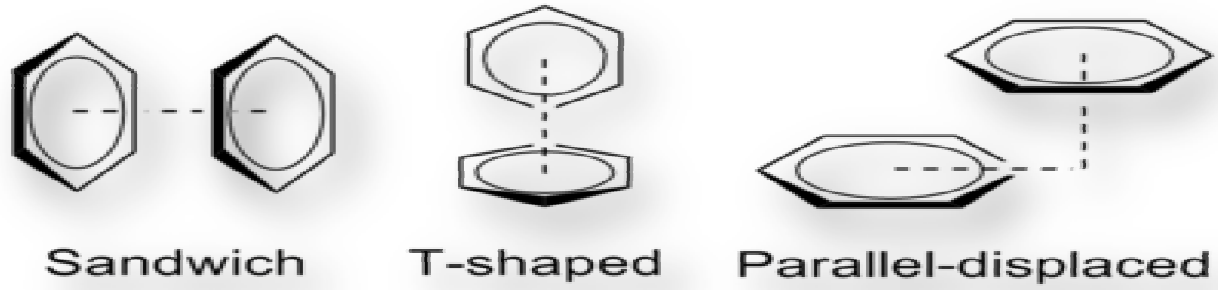


Figure 13: Interactions π-π

1.4.3.3 Constructive research

A major concern of naval operations is to consider the flexibility of both entities. Each learning algorithm requires a more complete organic molecule search flow structure. Creating and selecting convention devices can be done in several ways. When two vectors A and B are considered in space they can undergo translation and rotation in three dimensions. As the number of single bonds (especially acyclic bonds) theoretically increases the number of organic molecules formed increases. Therefore constructive problems cannot be high. The greater the degrees of freedom of the chemical entity the larger the conformational space and the greater the approximation. So it makes sense to work with molecules whose degrees of freedom (number of atoms and bonds subject to rotation) are limited. This method will be based on the available design strength. Therefore molecules that are too flexible were excluded from our study. A number of methods have been developed to determine the lowest energy configuration [10].

1.4.3.4 Scoring fonctions

Concept and Application of Scoring Functions Scoring functions are the most important element of structure-based drug design for evaluating the binding ability of a ligand to a target protein. Molecular docking experiments require rapid and accurate characterization of protein-ligand complexes. Molecular docking experiments generate thousands of ligand binding orientations/conformations so we use a score function to rank these complexes and distinguish between correct and incorrect binding mode predictions. The purpose of the ideal unit operation is to classify empirically determined complexes.

And the scoring function should predict the absolute affinity of the complex to facilitate the identification of potential hit/lead candidates for any therapeutic target from large combinatorial libraries used in virtual screening. Score functions are particularly useful for screening libraries of compounds or single compounds based on binding mode and affinity [11].

1.5 Applications and importance of molecular docking

The use of docking programs to reveal the properties of atoms and functional groups in a 3D (three-dimensional) structure allows the study of drug binding to their target site.

1.6 Calculation of HOMO and LUMO energies

These are molecular orbitals that play an important role in drug discovery and design. HOMO (highest occupied molecular orbital) energy provides a region for small molecules to donate electrons in complex formation while LUMO (lowest unoccupied molecular orbital) capacity refers to the molecules ability to accept electrons from bound proteins. The energy difference between HOMO and LUMO known as the HOMO-LUMO energy gap represents the electronic excitation energy to measure the chemical stability and reactivity of a mixture (Banavath et al. 2014) [5].

1.5 ADMET Prediction and Analysis

ADMET is an important step in any drug discovery program because it provides information on the pharmacokinetics (ADME ie absorption distribution metabolism and excretion) and pharmacodynamics ie toxicity of the lead molecule prior to wet lab experiments (T). Thus it reduces the time of experimental costs and the risk of drug failure. Literature studies have shown that poor pharmacokinetics and pharmacodynamics are major causes of high drug development costs and late-stage failures and it is generally accepted that computational ADMET should be considered in drug discovery programs in in vitro and in vivo studies. Prediction. Currently it is limited to some other theoretical descriptions such as Lipinskis law and van der Waals surface refractive index of polar surface (2D) polarization etc. Advances in combinatorial chemistry and high-throughput screening have greatly increased the number of small molecules. Preliminary ADMET data can be used as a reference. Additionally the accuracy of ADMET tools can be improved by

incorporating new theoretical descriptors and improved algorithms using sophisticated computational tools that can model the most relevant pharmacokinetic and pharmacodynamic information for any given molecule. Drug-like compounds. for cost-effective development (Pathak et al. 2018; Singh and Pathak 2020)[5].

1.6 Advantages of CADD

- Through it we can reduce the synthetic and biological testing efforts.
- It gives the most promising drug candidate by eliminate the compounds with undesirable properties (poor efficacy, poor ADMET etc.) through in silico filters.
- It is a Cost-effective, time saving, Rapid and automatic process.
- Through it we can know about the drug-receptor interaction pattern.

1.7 Examples of Drugs Synthesized Using CADD

Molecular modeling has also been used in the development of drugs that have passed clinical trials and in the treatment of a number of diseases have become modern therapies. In 2003, the quest for novel transforming growth factor-β1 receptor kinase inhibitors was one of the most compelling examples of the possibilities presented by molecular modeling methods in drug discovery. One group at Eli Lilly used a conventional high-throughput screening method to investigate a lead compound, which was subsequently improved by analyzing structure–activity relationship via in vitro assays.

Table 2: List of drugs identified through computational approaches.

S.No.	Drug	Drug Target	Disease	Approved year	references
1	Captopril	Angiotensin-converting enzyme (ACE)	Hypertension	1981	Talele et al (2010)
2	Dorzolamide	Carbonic Anhydrase (CA)	Glaucoma and ocular hypertension	1995	Vijayacrishnan (2009)
3	Saquinavir	HIV-1 and HIV-2 Proteases	AIDS	1996	Van Drie (2007)
4	Indinavir	HIV Protease	AIDS	1996	Van Drie (2007)
5	Ritonavir	HIV Protease	AIDS	1996	Van Drie (2007)
6	Tirofiban	Glycoprotein □b/□a receptor	Blocked coronary artery,	1998	Hartman et al (1992)

			antiplatelet drug		
7	Zanamivir	Neuraminidase	Influenza	1999	Kim et al (1997)
8	Oseltamivir	Neuraminidase	Influenza	1999	An et al (1997)
9	Raltegravir	Integrase	AIDS	2007	Schames et al (2004)
10	Aliskiren	Renin	Hypertension high blood pressure	2007	Cohen (2007)

While the Biogen Idec group used a molecular modeling approach involving virtual high-throughput screening based on the structural interactions among the weak inhibitor and altering growth factor-β1 receptor kinase. The group at Biogen Idec found 87 hits after the virtual screening of compounds, the best hit being similar in structure with the lead compound discovered by Eli Lilly's conventional high-throughput screening approach. In this case, molecular modeling, a process with reduced costs and tons of effort, was able to investigate the same lead for drug development (Sawyer et al. 2003; Singh et al. 2003; Sliwoski et al. 2014). Some of the earliest examples of drugs synthesized after their discovery through the molecular modeling methods are listed in Table 2 [5].

1.8 Success and Limitations

Discovery of new drug molecules using computational approaches has a focused research area due to advances in integrated omics, i.e. Genomics, proteomics, metabolomics, and bioinformatics, it has many successful stories. Recently, the concept of pharmacogenomics is introduced to focus on personalized medicine. The key advantages of pharmacogenomics are to produce drugs based on the Table □.2 List of drugs identified through computational approaches organizational structures of individual genomes. It is mainly used to address difficult tasks. It should not surprise that success is sometimes limited. Furthermore, several key problems related to computational complexities that have been on the agenda for decades remain to be resolved (Bajorath 2015; Hassan Baig et al. 2016; Singh 2014) [12].

- In drug discovery practice, the potential of in silico methods should not be overestimated because it affects the credibility of serious computer work in the academic and pharmaceutical industries.
- In many cases, computational approaches can advance drug discovery projects significantly only if they are carefully selected and employed to problems, such as the selection of small molecules with a probability of displaying a specific activity, identification of novel compounds for optimization, or design of analogs that positively interact with a specified binding site in the cavity of a receptor.
- Biological systems are very complex and directed by numerous significant parameters. So, there are certain restrictions, and it is not possible to copy and simulate the whole biological system on a PC using cutting edge techniques.
- One of the major challenges in drug discovery is target flexibility because most of the software provides only ligand flexibility.
- It is exceptionally hard to give total molecular flexibility to the protein as these augments the time and space complexity of the computation.
- Besides, the major limitation of pharmacophore modeling is dependent on pre-computed databases that hold a less number of low-energy conformations per compound.

2 AN OVERVIEW OF BREST CANCER

Breast cancer is a disease in which cells in the breast grow out of control. There are different kinds of breast cancer. The kind of breast cancer depends on which cells in the breast turn into cancer. Breast cancer can begin in different parts of the breast. A breast is made up of three main parts: lobules, ducts, and connective tissue. The lobules are the glands that produce milk. The ducts are tubes that carry milk to the nipple. The connective tissue (which consists of fibrous and fatty tissue) surrounds and holds everything together. Most breast cancers begin in the ducts or lobules. Breast cancer can spread outside the breast through blood vessels and lymph vessels. When breast cancer spreads to other parts of the body, it is said to have metastasized.

Who is mainly affected by breast cancer?

Breast cancer is one of the most common cancers among women, second only to skin cancer. It is most likely to affect women over the age of 50.

Though rare, men can also develop breast cancer. Approximately 2,600 men develop male breast cancer every year in the United States, making up less than 1% of all cases.

Transgender women are more likely to develop breast cancer compared to cisgender men. Additionally, transgender men are less likely to develop breast cancer compared to cisgender women [13].

2.1 Management of Breast Cancer

Following approaches are to be made for the management of breast cancer. They are as follows [14]:

2.1.1 Surgery

Depending on the stage and type of the tumor, lumpectomy (removal of the lump only), or surgical removal of the entire breast (mastectomy) is performed. Standard practice requires the surgeon to establish that the tissue removed in the operation has margins clear of cancer, indicating that the cancer has been completely excised. If the removed tissue does not have clear margins, further operations to remove more tissue may be necessary. This may sometimes require removal of part of the pectoralis major muscle, which is the main muscle of the anterior chest wall. More recently, the technique of sentinel lymph node (SLN) dissection has become popular, as it requires the removal of far fewer lymph nodes, resulting in fewer side effects. Advances in sentinel lymph node mapping over the past decade have increased the accuracy of detecting sentinel lymph node from 80% using blue dye alone to between 92% and 98% using combined modalities.

2.1.2 Radiation Therapy Radiation therapy

Involves using high-energy X-rays or gamma rays that target a tumor or post surgery tumor site. These radiations are very effective in killing cancer cells that may remain after surgery or recur where the tumor was removed. In addition to this treatment

implanted radioactive catheters (brachytherapy), similar to those used in prostate cancer treatment, can be used. However, this treatment option has been superseded by electron beam radiotherapy to the breast scar. Radiation therapy for breast cancer is usually performed after surgery and is an integral component of breast-conserving therapy. The dose of radiation must be strong enough to ensure the elimination of cancer cells. Treatments are typically given over a period of five to seven weeks, performed five days a week. Each treatment takes about 15 minutes.

2.1.3 Chemotherapy

Chemotherapy is the use of anti-cancer drugs to treat cancerous cells. Specific treatment for the breast cancer will be based on; overall health, medical history, age (whether menstruation is there or not), type and stage of the cancer, tolerance for specific medications and procedures etc. Chemotherapy treatments are often given in cycles; a treatment for a period of time, followed by a recovery period, then another treatment. Chemotherapy can be given before surgery to shrink the tumor and sometimes make breast conserving surgery possible rather than a mastectomy. Many times, it is given after surgery and may be given every three weeks or every two weeks in a “dose dense” fashion.

2.1.4 Hormone therapy for breast cancer

Some types of breast cancer use hormones — such as estrogen and progesterone — to grow. In these causes, hormone therapy can either lower estrogen levels or stop estrogen from attaching to breast cancer cells. Most often, healthcare providers use hormone therapy after surgery to reduce the risk of breast cancer recurrence. However, they may also use it before surgery to shrink the tumor or to treat cancer that has spread to other parts of your body.

2.1.5 Cell-Target Suicide

A conversion of a pro drug to a toxic metabolite by genetically engineering tumor cells is an attractive way to create an artificial difference between normal and neoplastic tissue. This can be achieved by the expression of a gene that confers a dominant, negatively selectable phenotype to the cancer cells, such as cell death imparted by expression of a prodrug – metabolism enzyme. Greater selectively in killing malignant cells will be obtained by transferring a gene that is not normally found inhuman beings (e.g. HSVthymidine kinase), rather then by overexpression an endogenous gene.

3 ESTROGEN RECEPTOR ALPHA IN HUMAN BREAST CANCER

3.1 Occurrence and Significance

Estrogens have long been recognized as being important for stimulating the growth of alarge proportion of breast cancers. Now it is recognized that two receptors, and the presence of estrogen receptor α (ERα) 3 correlates with better prognosis and the likelihood of response to hormonal therapy mediate estrogen action. Over half of all breast cancers overexpress ERα and around 70% of these respond to anti-estrogen (for exampletamoxifen) therapy. In addition, the presence of elevated levels of ERα in benign breastepithelium appears to indicate an increased risk of breast cancer, suggesting a role forERα in breast cancer initiation, as well as progression. However, a proportion of ERα-positive tumors does not respond to endocrine therapy and the majority of those that dorespond eventually become resistant. Most resistant tumors remain ERα-positive andfrequently respond to alternative endocrine treatment, indicative of a continued role forERα in breast cancer cell proliferation. The problem of resistance has resulted in the search for and the development of diverse hormonal therapies designed to inhibit ERαaction, while research on the mechanisms, which underlie resistance, has shed light on thecellular mechanisms, other than ligand binding, which control ERα function [15].

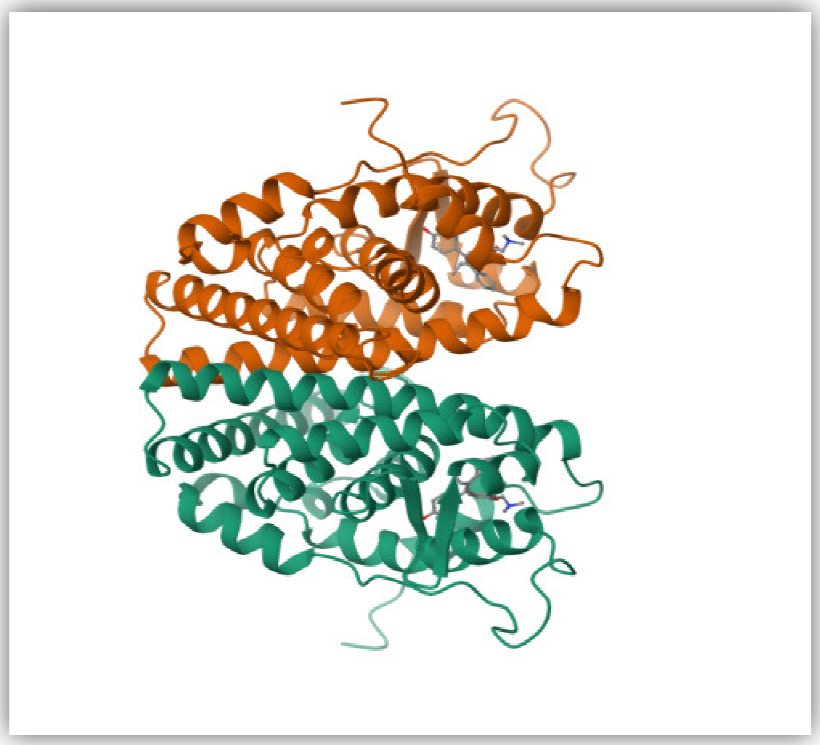


Figure14:Structure 3D estrogen receptor (ERα) (3ERT)

3.2 Structure

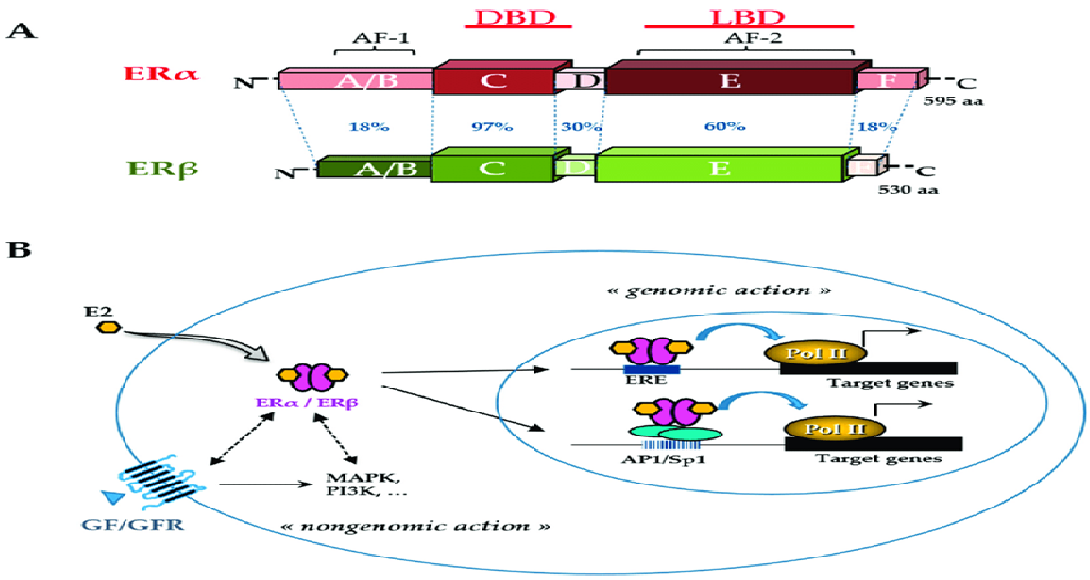


Figure 15: Estrogens receptor alpha (ERα) structure and action

Estrogen receptor (ERα) structure and action. The schematic structures of the two human ERαand ERβ and the percentage of homology between the different domains (annotated by theletters A to F) are indicated (A). Domains involved in DNA binding (DBD), ligand binding(LBD), ligandindependent transactivation function 1 (AF1), and ligand-dependenttransactivation function 2 (AF2) are shown. The number of amino acids for each receptor isalso indicated on the right side. Estradiol (E2) mediates numerous phenotypic effects in cellsby binding to and activating ERs (B). E2 enters the cell through the lipid membranes andbinds ER, which can be present in the cytoplasm and the nucleus. The activated ER forms dimer to tightly fix chromatin directly at the estrogen-responsive element (ERE) sites orindirectly at activator protein 1 (AP1) or specificity protein 1 (Sp1) sites. ER is then able toremodel chromatin by recruiting cofactors and activating RNA polymerase II (Pol II), at targetgenes (genomic action). Besides, ERs can use rapid non-genomic action through theinteraction with intracellular kinases (mitogen-activated protein kinase (MAPK)). Phosphatidylinositol 3-kinase (PI3K)and the growth factor (GF) receptor (GFR)pathways [16].

3.3 Roles

Estrogen receptors influence many physiological processes in mammals such as thoseinvolved in reproduction, cardiovascular disease, bone development, cognition, and

behavior.Given their roles, it is not surprising that estrogen receptors are also involved in thedevelopment of many cancers such as breast cancer, cancer ovaries, the uterus or prostate, orin other diseases such as osteoporosis, cardiovascular or neurodegenerative diseases, orobesity. Receptors in Estrogens (REs), especially REα, are heavily involved in developmentmammary tumor [17].

3.4 Involved mechanisms

The mechanism underlying the mitogenic action of estrogens has been widely studied in celllines and probably results from the complex modulation by the ERs of different transcriptionalpathways—thus involving the regulation of a multitude of genes. The initial and currenthypothesis is that estrogens control the growth of primary breast cancers by inducing estrogen-regulated proteins that function as autocrine, paracrine or intracrine growth factors. Estrogensactivate (and antiestrogens block) genes controlled by estrogen-responsive elements (EREs).In addition to these classical transcriptional effects, these ligands can also modulate othergenes, not containing ERE, via direct protein-protein interaction of ER with other transcriptionfactors. For example, ER interferes with AP1-directed gene activity through a protein-proteininteraction with c-Jun. In cell cultures, genes such as cyclin D1 are stimulated by estrogensthrough an AP1 pathway and repressed by tamoxifen. By contrast, the non-genomic effects ofestrogens on signal transduction do not appear implicated in their mitogenic action, since allkey events in cell cycle stimulation can occur in the presence of a MAP kinase-activatinginhibitor. The genes responsible for the mitogenic effect of estrogens have not beendefinitively determined but they probably include secreted growth factors, growth factorreceptors, proteases such as cathepsin D and cyclin/cdk factors. The implication of moleculesinterfering with the cytoskeleton, such as E-cadherin, a mediator of cell-cell interactions, asbeen also suggested. Ecadherin is down-regulated by estrogens in normal and tumorigenicbreast epithelial cells. Moreover, most of these estrogen-regulated proteins are differentiallyexpressed in ER-positive and ER-negative tumors and this probably contributes to theirdifferent metastatic potentials. Actual molecular profiling of breast tumors based on newscreening technologies would complete the set of genes associated with different phenotypes [18].

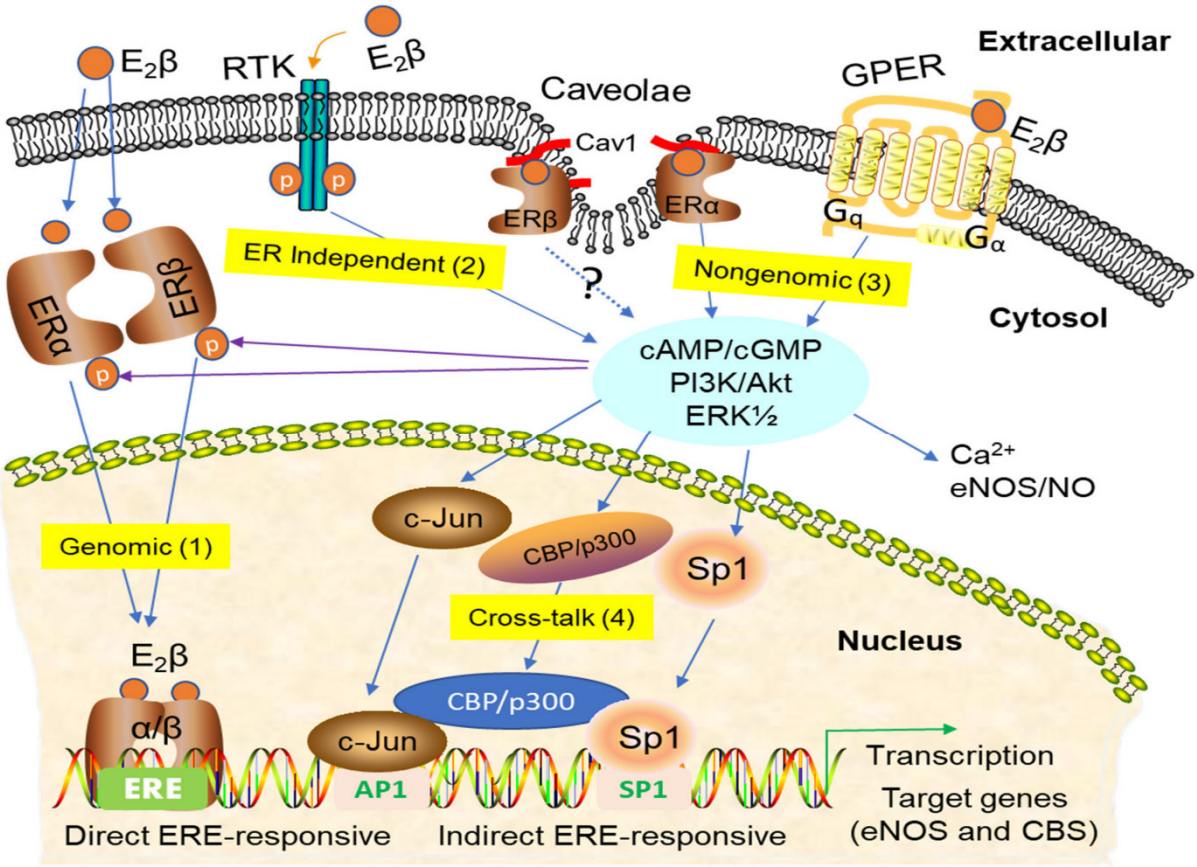


Figure 16: Mechanism of estrogen receptor action

3.5 Drugs that block estrogen receptors

These drugs work by stopping estrogen from fueling breast cancer cells to grow [19].

3.5.1 Selective estrogen receptor modulators (SERMs)

These drugs block estrogen from connecting to the cancer cells and telling them to grow anddivide. While they have anti-estrogen effects in breast cells, they act like an estrogen in othertissues, like the uterus and the bones.

3.5.2 Selective estrogen receptor degraders (SERDs)

Like SERMs, these drugs attach to estrogen receptors. But SERDs bind to the receptors moretightly and cause them to be broken down. These drugs have anti-estrogen effects throughoutthe body.SERDs are used most often in women who are past menopause. When given to pre-menopausal women, they need to be combined with a luteinizing-hormone releasing hormone(LHRH) agonist to turn off the ovaries (see ovarian suppression below).

References

1. Surabhi et al Journal of Drug Delivery & Therapeutics. 2018; 8(5):504-509.
2. Begum, Sameera, and S. M. Shahidulla. "Role of Computer Aided Drug Design in Drug Development and Discovery: An Overview." Int J Res Eng Sci Manag 2.2 (2019): 445-450.Hoque, Injamul, et al. "An approach of computer-aided drug design (CADD) tools for in silico pharmaceutical drug design and development." Int. J. Adv. Res. Biol. Sci 4.2, 2017 : 60-71.
3. Hoque I, Chatterjee A, Bhattacharya S, Biswas R. An approach of computer-aided drug design (CADD) tools for in silico pharmaceutical drug design and development. Int. J. Adv. Res. Biol. Sci. 2017;4(2):60-71.
4. Lill M. Virtual screening in drug design. In Silico Models for Drug Discovery. 2013 Vol. 993, pp. 1-12. Humana Press, Totowa, NJ.
5. BOOK Singh, Dev Bukhsh, ed. Computer-aided drug design. Singapore: Springer, 2020.
6. Krid A. Modélisation de nouvelles molécules biologiques actives liées au stress oxydant. Thèse de magister d'université : chimie theorique. Constantine: Université Mentouri de Constantine. Algérie. 2008. 75p
7. Bensegueni A. Etude théorique des métabolites secondaires des végétaux et des composés de synthèse sur le plan de l'activité biologique : simulation par docking (arrimage) moléculaire sur la lipoxygénase et la cyclooxygénase. Thèse de doctorat d'université :Biochimie appliquée. Constantine : Université Mentouri de Constantine. Algérie. 2007. 76p.
8. Arrault, A. ; Monge, A. ; Violas, S. ; Genest, M. ; Genest, D. ; Garnier, N. ; Morin-Allory, L. ; Marot, C. Molecular dynamic assisted docking. Application to the screening on PPAR gamma
9. 10th Anniversary of MipTec – The Leading European Event for Drug Discovery Technologies 05-2007 - Basel (Switzerland).
10. Kollman, P.A. Massova, I. Reyes, C. Kuhn, B. Huo, S. Chong, L. Lee, M.Lee, T.Duan,Y. Wang, W. Domini, O. Cieplak, P. Srinivasan, J. Case, D.A. Cheatham, T.E.«Calculating structures and free energies of complex molecules : combining molecularmechanics and continuum models Acc ». Chem. Res. 33.12, 2000 : 889-897.
11. Bajorath J (2015) Computer-aided drug discovery. F1000Res 4:630.

12. Łukasiewicz S, Czeczelewski M, Forma A, Baj J, Sitarz R, Stanisławek A. Breast Cancer-Epidemiology, Risk Factors, Classification, Prognostic Markers, and Current Treatment Strategies-An Updated Review. Cancers (Basel). 2021 Aug 25;13(17):4287. doi: 10.3390/cancers13174287. PMID: 34503097; PMCID: PMC8428369.
13. Rahman GA. Breast Conserving Therapy: A surgical Technique where Little can Mean More. J Surg Tech Case Rep. 2011 Jan;3(1):1-4.
14. 10. Begg AC, Stewart FA, Vens C. Strategies to improve radiotherapy with targeted drugs. Nat Rev Cancer. 2011;11:239–253Sak K. Chemotherapy and dietary phytochemical agents. Chemother Res Pract. 2012;2012:282570. doi: 10.1155/2012/282570. Epub 2012 Dec 20. PMID: 23320169; PMCID: PMC3539428.
15. Sharma GN, Dave R, Sanadya J, Sharma P, Sharma KK. Various types and management of breast cancer: an overview. J Adv Pharm Technol Res. 2010 Apr;1(2):109-26. PMID: 22247839; PMCID: PMC3255438.
16. Huang B, Omoto Y, Iwase H, Yamashita H, Toyama T, Coombes RC, et al. Differential expression of estrogen receptor α, β1, and β2 in lobular and ductal breast cancer. Proc Natl Acad Sci U S A (2014) 111(5):1933–8. doi: 10.1073/pnas.1323719111
17. Pagano MT, Ortona E, Dupuis ML. A role for estrogen receptor alpha36 in cancer progression. *Front Endocrinol* (2020) 11. doi: 10.3389/fendo.2020.00506
18. Szostakowska M, Trębińska-Stryjewska A, Grzybowska EA, Fabisiewicz A. Resistance to endocrine therapy in breast cancer: molecular mechanisms and future goals. Breast Cancer Res Treat (2019) 173(3):489–97. doi: 10.1007/s10549-018-5023-4
19. Wijayaratne, Ashini L., and Donald P. McDonnell. "The human estrogen receptor-α is a ubiquitinated protein whose stability is affected differentially by agonists, antagonists, and selective estrogen receptor modulators." *Journal of Biological Chemistry* 276.38 (2001): 35684-35692.

Chaptre □

MATERIALS AND METHODS

CHAPTER □: MATERIALS AND METHODS

Molecular docking was the first step, which applied to the cucurbitacins and flavonoids compounds to study their affinity to the ERα **(PDB ID: 3ERT).** Meanwhile, the second step computational approach for some selected compounds were applied to characterize their chemical properties Absorption, Distribution, Metabolism, Excretion and Toxicity analysis. The third step was global reactivity descriptors of the selected compounds were calculated to understand their structures, stability and reactivity. The methodology of this work was illustrated in the **figure 17**.

1 General protocol

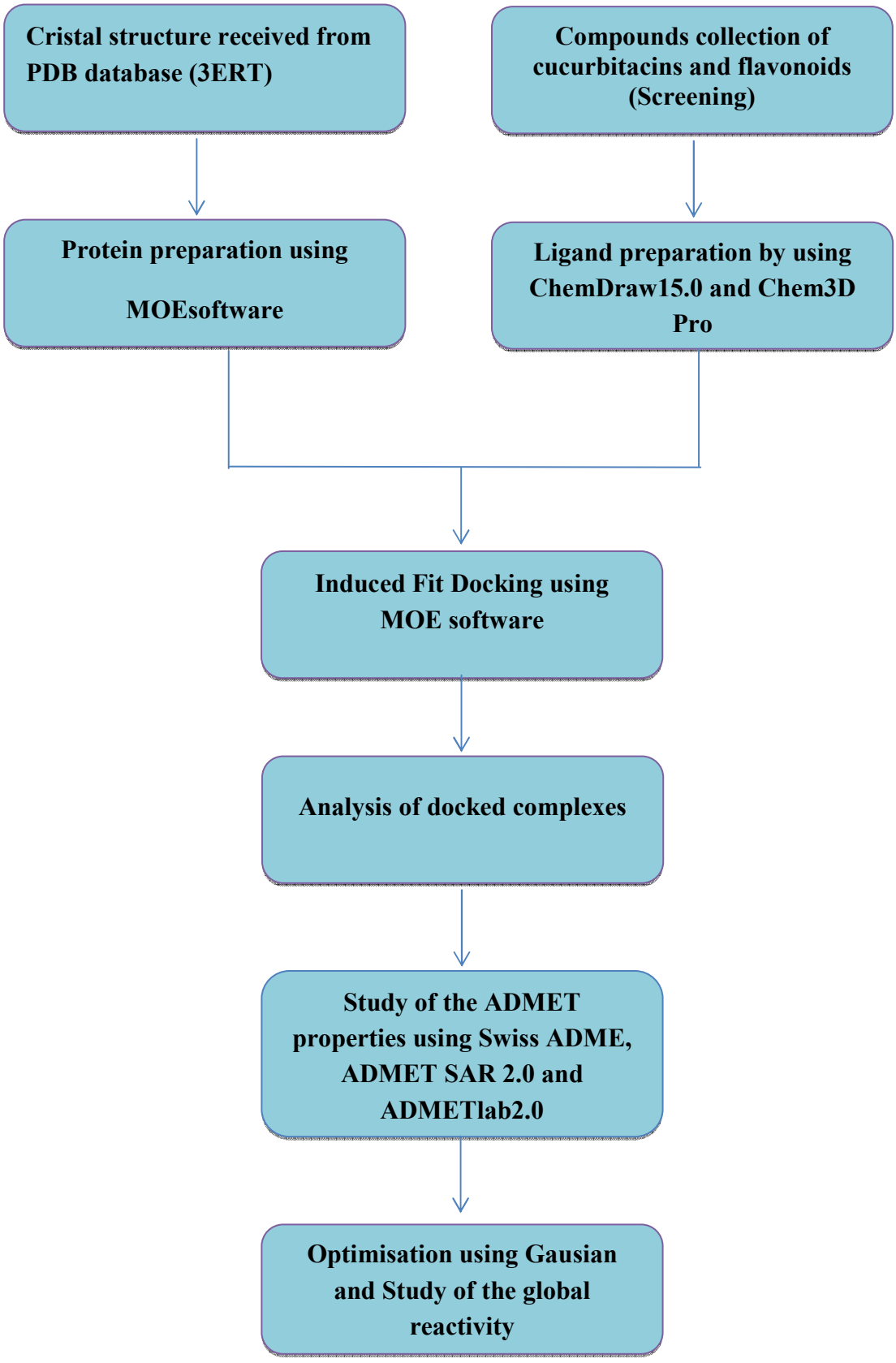


Figure 17: Schematic representation of the methodology used in the study.

2 Compound collections

An online research of published articles related to cucurbitacins have been isolated from C. colocynthis **[2]**. Several compounds have been isolated from the fruits of C. colocynthis. These different products belong to the groups of cucurbitacins, flavonoids, alkaloids and acids phenolic. Among these natural chemical groups, cucurbitacins are the main constituents of C. colocynthis. In this study, wefocused cucurbitacins and flavonoids structures that collected from the library **[2]**. The selected structures were designed by ChemDraw 15.0 **[3]**.

3 Citrullus colocynthis Colocynth the source of cucurbitacins and flavonoids

The colocynth, owned by dry soils, is very common in wet or fairly dry sweltering domains, it is not very present in temperate zones. It resides a very colossal domain that extends from North Africa, to the Sahara, Egypt, and Saudi Arabia to India in addition to the Mediterranean border**figure 18**. Traditional curative requests of C. colocynthis have inspired many abundant pharmacological searches. Several extracts and compounds unique from crops were evaluated for their organic actions, that is to say antidiabetic project, antimicrobial, anti-instigative, antioxidant, antiallergic, hypolipidemic poison and anticancer exercise (cytotoxicity) **[6]**.

Figure 18: Citrullus colocynthis whole plant

3.1 Structure of cucurbitacins

Cucurbitacins are particularly oxygenated tetracyclic triterpenes. Their basic structure is composed of a tetracyclic cucurbitane nucleus, the 19, 2-(10→9β))-abeo-10α-lanost-5-ene also known as 9β-methyl-19-norlanosta-5-ene **Figure 19**. Cucurbitacins are arbitrarily divided into twelve categories according to their structural characteristics. All cucurbitacins identified with a date have a double bond between C-5 and C-6. They differ from each other by the presence of a double bond at C-1 and/or C-23, a hydroxyl group at C-2, C-3,C-9, C-16, C-20, C-24 and/or C-25, a ketone function (C=O) at C-2, C-3, C-11 and/or C22 and by acetylation of the hydroxyl functional group at C-26. In addition, the derivatives heterosides exhibit glycosylations at the C-2 or C-3 carbons C-2 or C-3 **[4]**.

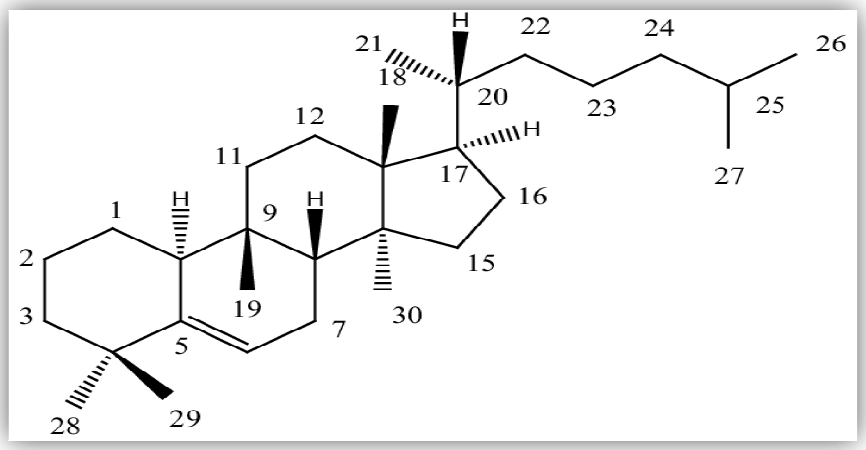


Figure 19: Basic structure of cucurbitans [19-(10→9β)-abeo-10α-lanost-5-ene].

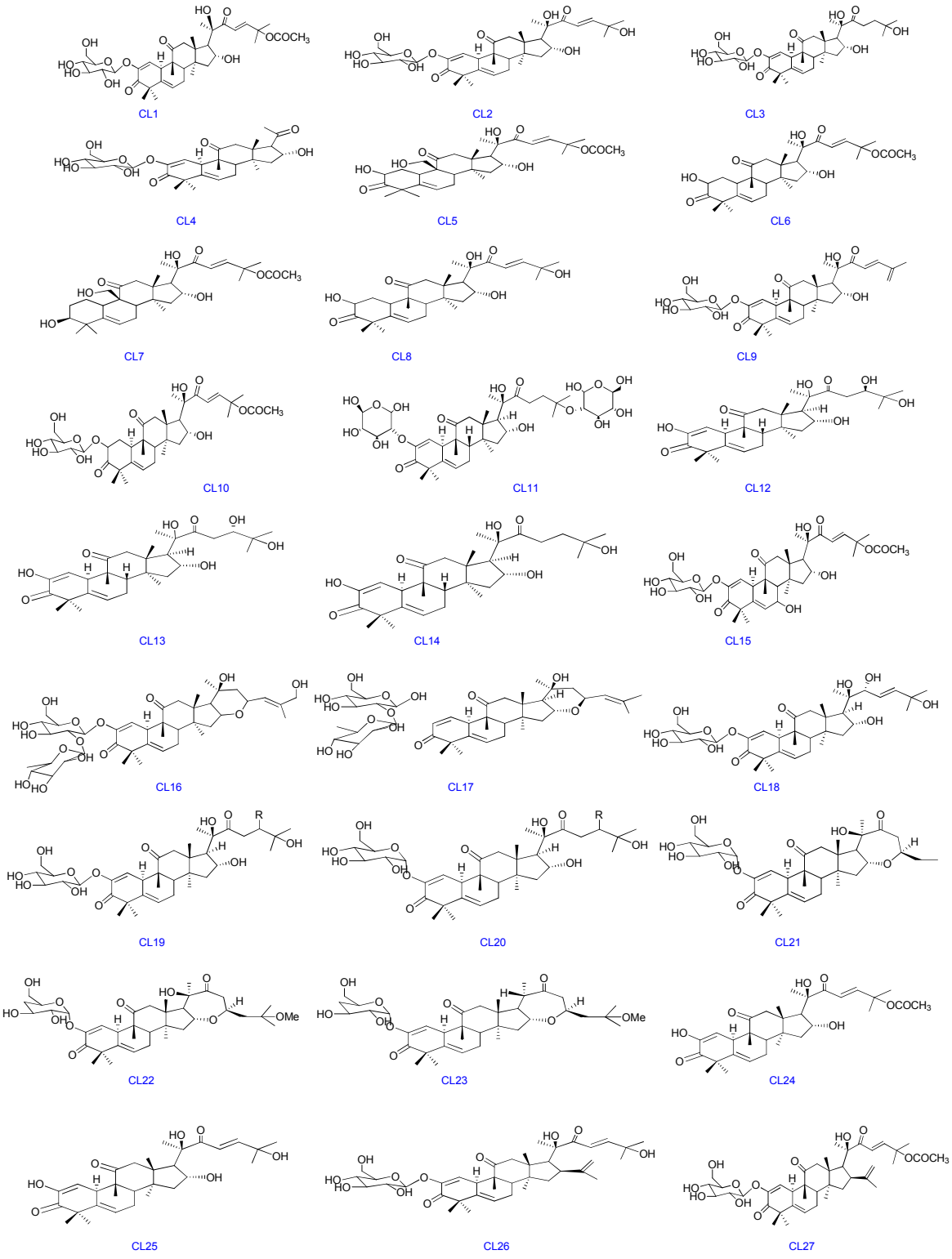


Figure 20: Chemical 2D structures of cucurbitacins the major constituents of C. colocynthis.

3.2 Structurs of flavonoids

Chemically flavonoids are based upon a fifteen-carbon skeleton involving two benzene rings (A and B as proved in **Figure 21**) connected by way of a heterocyclic pyrane ring (C). They maybe detached into a type of classes to a degree flavones (such as, flavone, apigenin, and luteolin), flavonols (for instance, quercetin, kaempferol, myricetin, and fisetin), flavanones (for instance, flavanone, hesperetin, and naringenin), and others. General structure of flavonoids was presented in the Figure 4. The various classes of flavonoids disagreein the level of oxidation and pattern of substitution of the C ring, while individual compounds inside a class clash in the pattern of replacement of the A and B rings **[5]**.

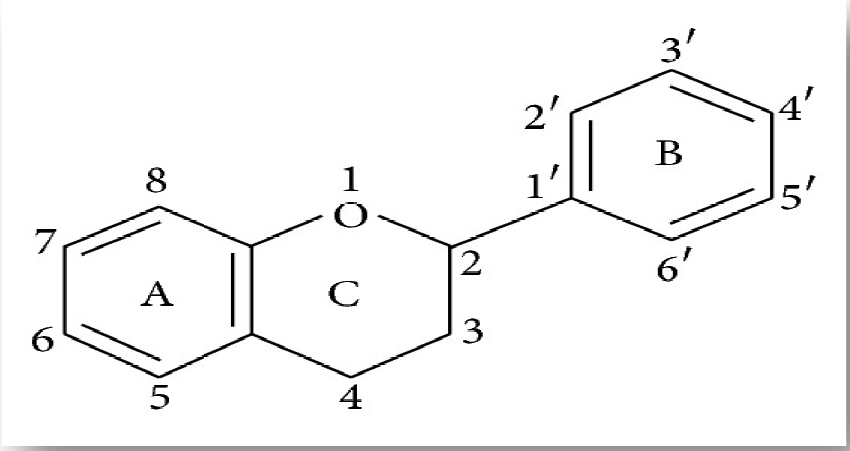


Figure 21: Basic structure of flavonoids.

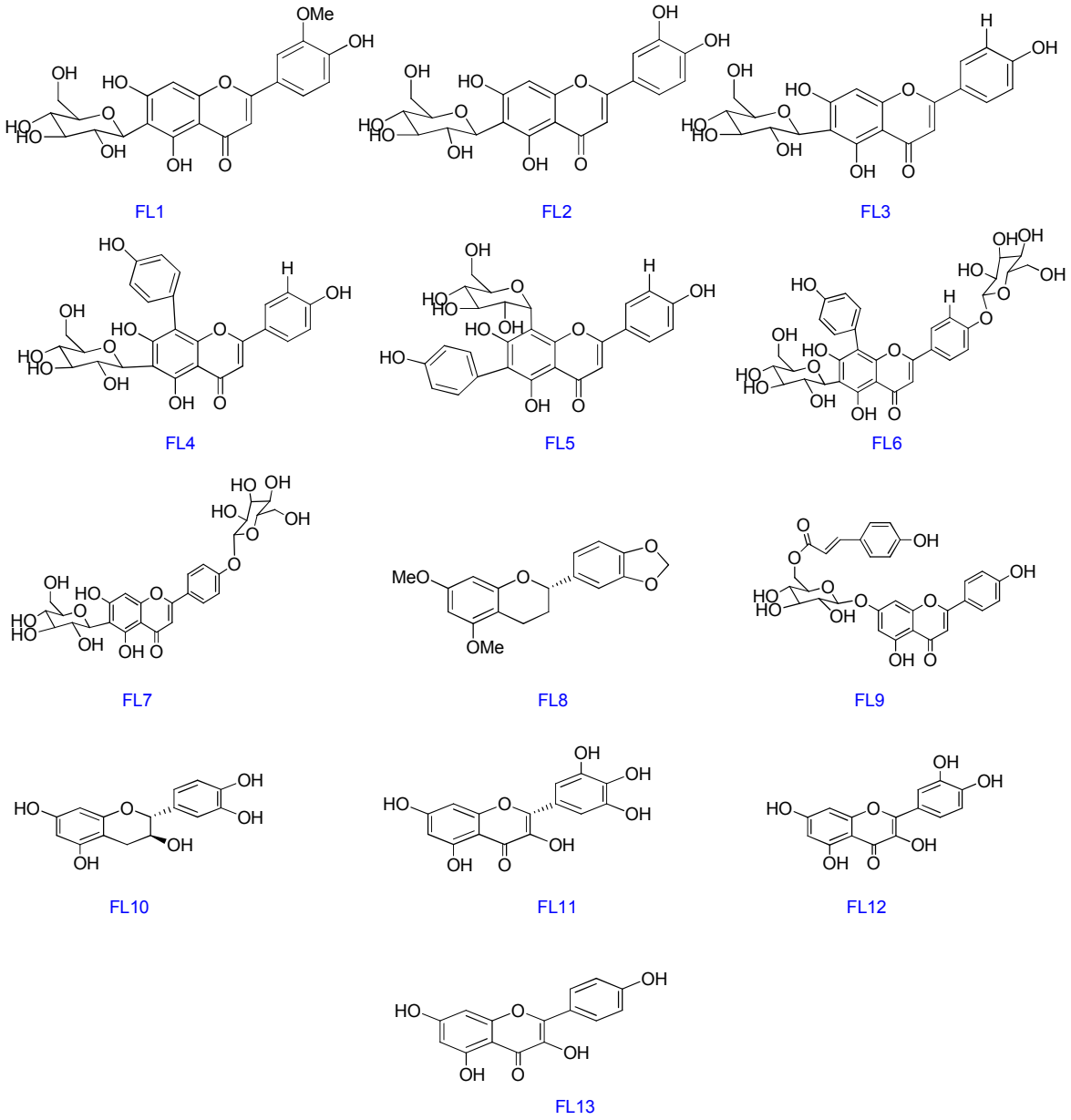


Figure 22:Chemical 2D structures of flavonoids.

4 Compututional details

The theoretical quantum chemical calculations were performed on a computer of a Processor: Intel (R) 2Core (R) N3350, up to 2.4 GHz , Installed memory (RAM): 4GB and System type: Dual-Core , with windows 10 platform, by the mean Gaussian 09 **[7]**, ChemDraw 15.0 **[3]**, and MOE software program **[8]**.

5 Molecular library preparation

5.1 The 2D and 3D structures

The 2D structures of 27 Cucurbitacins and 13 of flavonoids molecules shown in (figure2 and 3)in orderwere drawn by using ChemDraw professional 15.0 **[3]** then converted to 3D and pre-optimized by ChemDraw 3D pro15.0 software **[3]**. The 3D ligand structures were saved in MDL format and then imported to the Gaussian09 **[7]** program to do the optimization and global reactivity calculation.

5.2 Receptor preparation

The three-dimensional (3D) crystal structure of ERα mutant protein (PDB ID: 3ERT) were downloaded from the protein data bank (PDB)**[10]**.The Protein Data Bank archive contains thousand protein structures obtained either by crystallography X-ray or by NMR. The protein was imported into MOE software for visualizing the binding domain of this complex and identifying the amino acids in the binding pocket (1:(GLU323 PRO324 PRO325 ILE326 LEU327 MET343 LEU346 THR347 LEU349 ALA350 ASP351 GLU353 HIS356 MET357 TRP360 TRP383 LEU384 ILE386 LEU387 MET388 GLY390 LEU391 TRP393 ARG394 PHE404 GLY420 MET421 ILE424 LEU428 PHE445 LYS449 GLY521 HIS524 LEU525).

The enzyme were prepared by removing the repeated chains, leaving water molecules within the active site to ensure the formation of a hydrogen bond between the ligand and the target using MOE software **[8]**, the missing bonds in the protein structure, which were broken through X-ray diffraction were corrected, and hydrogen atoms were added. After that, the residue of protein active site were found using a site finder, they are presented in Table 1. The energy minimization of protein was done by applying the assisted model building and energy refinement (Amber 10): Extended Hückel Theory (EHT) force field.

6 Molecular docking

All the molecular docking and scoring calculations were performed using the molecular operation environment software (MOE.2015)[**9**]. After 30 poses, the ligands will attacked the protein internal grooves, resulting in the most stable docking ligand-receptor complexes. The scoring energies were increased by two unrelated adjustments by the triangular Matcher techniques, which were the mean values of trials utilizing the London dG scoring function. In addition to important interaction characteristics, the interacting complexes were retrieved. The level of inhibition was determined using extracted characteristics such ligand locations, receptor backbones (amino acids), interaction type, bond lengths, and internal and scoring energies. It is well known that the optimal RMSD score is near 2 with an energy score of less than or equal to -7 Kcal/mol**[10,11]**. These two numbers are frequently used as a criterion for evaluating the molecular docking results. Also the bond length must be not exceeding than 3.5 Å to be effective.

The molecular docking process inserted in software (MOE) was implemented for the selected drugs from Drug Bank database. The tested inhibitors were chosen based on their structural similarities tofulvestrant to provide a broad overview of their interactions with selected receptor.. The crystal structure of estrogen receptor alpha (PDB entry:3ERT) at a resolution of 2.20 Å. A resolution between 1.5 and 2.5 Å is considered as a good quality for docking studies.

7 Computationel pharmacokinetics

Due to weak ADMET (absorption, distrubition, metabolisme, excretion, toxcicty) traits, the majority of healing drugs abandoned in dispassionate tests. QSAR models are now usually applyied to call ADMET studies for medications inthe CADD stage. ADMETlab **[12]**, SwissADME **[13]** and admetSAR **[14]** are now appropriate databases for anticipating ADMET traits. The ADMET characteristics werecalled utilizing ADMETlab, that determined a more precise forecasting than the SwissADME and the admet SAR. For cucurbitacins and flavonoids derivative inhibitors of the ERα family, absorption, distribution, metabolism, excretion, and toxicity analyses were determined. To realize spoken bioavailability, ADMET was a critical step. ADMETwas an important step to attain spoken bioavailability. Those limits are main for absorption (caco-2 permeability > -5.15: caco 2 cells are a human colon epithilial cancer cell line used as a model of human intestinal absorption of drug, Pgp-inhibitor, Pgp-substrat “P-glycoprotein (P-gp), the

permeability glycoprotein or plasma glycoprotein is an active, efflux, membrane bound transport protein pump " and human intestinal absorption), Distribution (PPB: plasma protein binding, VD: volume distribution and BBB: blood-brain barrier), metabolism CYP450 enzyme (1A2-inhibitor and substrate, 3A4-inhibitor and substrate, 2C9-inhibitor and substrate, 2C19-inhibitor and substrate and 2D6-inhibitor and substrate), Excretion (T1/2: half-life and CL: clearance) and for toxicity (hERG:human ether-à-go-go related gene, AMES:The Ames test for mutagenicity and LD50: Lethal Dose 50). ADMET characteristics were predicted utilizing ADMETlab 2.0 connected to the internet spreadsheet. The results of ADMET characteristics are classified by (1: inhibitor, substrate or blocker; 0: non inhibitor, no substrate or no blocker).

8 Optimization and global reactivity calculation

The makeups of the compounds complicated in the current study were enhanced by using the commonness and addition (opt+frq) parameters in Gaussian 09 **[6]**. The density functional theory (DFT) method was second-hand in this place study to envision the miscellaneous physicochemical properties. As a result, the compounds were re-optimized at the DFT/B3LYP/6- 311G level of theory by using the Gaussian 09 program **[7]**. The highest occupied molecular orbital energy (E_{HOMO}), the lowest unoccupied molecular orbital energy (E_{LUMO}), the energy gap **($\Delta E = E_{LUMO} - E_{HOMO}$)** the global electrophilicity index **($\omega = \mu^2/2\eta$)**, the chemical potential **($\mu = [E_{LUMO} + E_{HOMO}]/2$)**, the chemical hardness **($\eta = [E_{LUMO} - E_{HOMO}]/2$)**, the chemical softness **($s = 1/2\eta$)**, and the nucleophilicity **($N = E_{HOMO}$ (Nucleophile) $- E_{HOMO}$ (tetracyanoethylene, TCE))**.

References

1 Harbeck N, Penault-Llorca F, Cortes J, Gnant M, Houssami N, Poortmans P, Ruddy K, Tsang J, Cardoso F. Breast cancer. Nat Rev Dis Primers. 2019 Sep 23;5(1):66. doi: 10.1038/s41572-019-0111-2. PMID: 31548545.

2 Watson L., Dallwitz M.J. 1992. The families of flowering plants: descriptions, illustrations, identification, and information retrieval. delta-intkey.com.

3 HyperChem. 1115 NW 4th, Street, Gainesville, FL 32601, USA,: molecular modelling system.Hypercube Inc; 2009.

4 Chen J.C., Chui M.H., Nie R.L., Cordell G.A., Qui S.X. 2005. Cucurbitacins and cucurbitane

5 glycosides: structures and biological activities. Nat. Prod. Rep., 2005, 22: 386-399.

6 Duke J.A. 1983. Citrullus colocynthis (L.) Schrad. Handbook of Energy Crops.

7 Becke AD. Density-functional thermochemistry. V. Systematic optimization of exchange-correlation functionals. J Chem Phys. 1997;107(20):8554–60.

8 Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, et al.gaussian. Wallingford, CT; 2009.

9 Molecular Operating Environment. (MOE). Montreal, QC, Canada, H3A 2R7,: Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #910; 2019.

10 Ramalho TC, Caetano MS, da Cunha EFF, Souza TCS, Rocha MVJ. Construction and assessment of reaction models of class i epsp synthase: Molecular docking and density functional theoretical calculations. J Biomol Struct Dyn. 2009;27(2):195–207.

11 Kellenberger E, Rodrigo J, Muller P, Rognan D. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. Proteins Struct Funct Genet. 2004;57(2):225–42.

12 DrugBank [Internet]. 2020. Available from: https://www.drugbank.ca/

13 Dong J, Wang N-N, Yao Z-J, Zhang L, Cheng Y, Ouyang D, et al. ADMETlab: a platform for systematic ADMET evaluation based on a comprehensively collected ADMET database. J Cheminform [Internet]. 2018; Available from: http://admet.scbdd.com/

14 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 [Internet]. 2017;7(March):1–13. Available from:.

15 Cheng F, Li W, Zhou Y, Shen J, Wu Z, Liu G, et al. AdmetSAR: A comprehensive source and free tool for assessment of chemical ADMET properties. Vol. 52, Journal of Chemical Information and Modeling. ACS Publications; 2012. p. 3099–105.

Chapter III
RESULTS AND DISCUSSION

CHAPTRE III: RESULTS AND DISCUSSION

1 RESULTS AND DISCUSSION OF CUCURBITACINS COMPOUNDS

1.1 Molecular docking results

Molecular docking simulation of cucurbitacins compounds, docked into ERα active sites **PDB ID: 3ERT [1,2]** performed to assess their abilities to inhibit the disease exactly breast cancer disease. **Table 3** represent the result obtained, the selection of the best-docked ligands based on both the binding scores and RMSDs value. The results show that the three cucurbitacins **CL18, CL17** and **CL12** had the bust interactions with the site1, with a score ranging between **(-6.9656kcal/mol)** and **(-6.8937 kcal/mol).**

Due to that the score must be more than **-6** and equal or less than **-7** , we could say that all these ligand **CL18, CL17, CL12, CL8, CL6, CL13, CL4, CL14, CL9, CL5, CL25, CL2, CL22, CL11, CL7, CL21 and CL3** had a good score fonction they ranging between **(-6.8912 kcal/mol)** and **(-6.0743 kcal/mol)** as the table 1 shows.

Because of the value of the RMSDs must be near to 2.50Å we could say that all these ligands had a good resolution they ranging between **(1.4918Å)** and **(2.4663 Å).**

From the **table 3**the compound **CL18** interact with the amino acid **Met522** with H-donor interaction bond between **O-26** and **SD** with a length **(3.66Å)** which was more than the accepted limit **3.5Å** and with the amino acid **Leu536** with H-acceptor bond between **O-28** and **N(2.71Å)**, while the compound **CL17** interact with the amino acid **Asp351** with tow bond type H-donor the first between **C-12** and **OD2** with a length **(3.28Å)** and the second between **O-63** and **OD2(3.34Å)**. For the compound**CL12**, it was interact with the amino acid **Asp351** with tow bond type H-donor the first between **O-39** and **OD1 (2.94Å)** which was effective**(<3.5Å)** and near to the length value of reference molecule **2.96Å** and the second between **O-39** and **OD2 (3.10Å)** .

When compared these compounds with the reference ligand (HOT) we noticed that the **CL27** have a score value **(-6.3859 kcal/mol)** approximately equal to the score of the reference ligand **(-6.3009 kcal/mol)**, also we observed that the ligands **CL5**,**CL7**,**CL2**,**CL19** and **CL26** interact with the solvent **HOH 31** with a interaction type H-donorsame to the reference molecule.

Table 3: The results obtained from docking of compound with ERα.

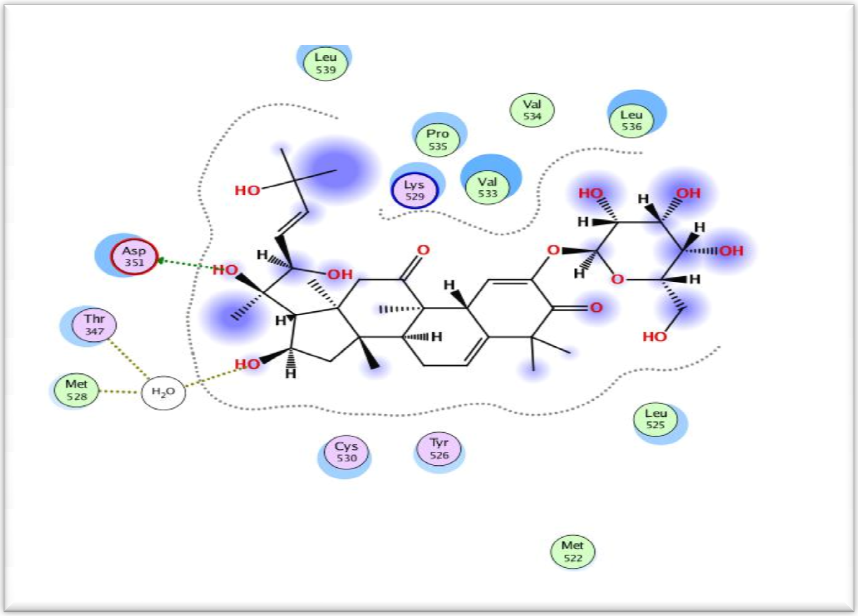
Bonds between atoms of compounds and residues of the active site								
Ligands	**S score (kcal/mol)**	**RMSD (Å)**	**Atom of compounds**	**Atom of receptor**	**Evolved receptor residues**	**Type of interaction bond**	**Distance (Å)**	**E (kcal/mol)**
L_{Ref}	-6.3009	1.3903	O-15	O	HOH31	H-donor	2.96	-1.7
CL1	-5.4114	2.2326	O-81 O-81	SG N	Cys530 Cys531	H-donor H-acceptor	3.07 2.91	-1.6 6.7
CL2	-6.2446	2.3360	O-58 O-81	O OD2	HOH31 Asp351	H-donor	2.73 3.74	-0.5 -0.9
CL3	-6.0743	2.4603	O-24 O-24 O-101 O-8	OD1 OD2 SD N	Asp351 Asp351 Met543 Cys530	H-donor H-donor H-donor H-acceptor	3.23 3.12 4.16 3.02	-0.9 -0.9 -0.8 -2.2
CL4	-6.6261	2.2913	O-81	OD1	Asp351	H-donor	2.50	0.9
CL5	-6.4687	2.0471	O45	O	HOH31	H-donor	2.67	-1.6
CL6	-6.8363	1.4918	O-82 O-79	SG N	Cys530 Cys530	H-donor H-acceptor	3.54 3.08	-0.8 -1.2
CL7	-6.0865	1.8146	O-44 O-43	O N	HOH31 Cys530	H-donor H-acceptor	3.03 2.86	-0.7 -2.6
CL8	-6.8912	2.4221	O-13	N	Leu536	H-acceptor	3.10	-1.7
CL9	-6.5703	2.3479	O-81	OD1	Asp351	H-donor	3.11	-0.8
CL10	-6.4897	2.5787	O-26 O-74	SG SG	Cys530 Cys530	H-donor H-acceptor	3.03 3.73	-0.9 -0.9

CL11	-6.1878	2.3515	O-78 O-78	SG N	Cys530 Cys530	H-donor H-acceptor	3.24 3.22	-1.0 -0.4
CL12	-6.8937	2.4663	O-39 O-39	OD1 OD2	Asp351 Asp351	H-donor H-donor	2.94 3.10	-1.0 -1.7
CL13	-6.7090	1.7366	O-1 O-39	SG OD2	Cys 530 Asp 351	H-donor H-donor	3.12 3.22	-1.7 -1.1
CL14	-6.5992	1.7328	O-40 O-78	OD1	Asp 351	H-donor	2.63	-2.5
CL15	-6.7843	2.7108	C-8 O-57	SG ND2	Cys 530 Asn 348	H-donor H-acceptor	3.26 3.09	-0,8 -1,5
CL16	16.1091	2.5610	O-74	SD	Met343	H-donor	3.25	-0.3
CL17	-6.9345	2.0484	C-12 O-63	OD2 OD2	Asp351 Asp351	H-donor H-donor	3.28 3.34	-0.8 -0.6
CL18	-6.9656	2.3535	O-26 O-28	SD N	Met522 Leu536	H-donor H-accepteur	3.66 2.71	-0.7 -0.7
CL19	17.5486	1.9715	O-81 O-101	OD2 O	Asp 351 HOH 31	H-donor H-donor	2.47 2.67	-2.6 -0.1
CL20	-4.8893	3.7204	O-70 O-104	OD1 O	Asp351 HOH20	H-donor H-acceptor	2.79 2.84	-2.7 -0.6
CL21	-6.2230	2.3476	O-86	OD2	Asp351	H-donor	2.72	-2.5
CL22	-6.4722	2.3477	O-86	OD2	Asp351	H-donor	2.99	-1.8
CL23	-5.5754	3.6167	O-8	N	Cys530	H-acceptor	2.98	-2.6
CL24	-3.9284	1.8731	O-38 O-67	O	HOH31 HOH58	H-donor H-donor	2.52 2.92	-0.1 -0.9
CL25	-6.4468	1.6705	O-61	OD1	Asp351	H-donor	2.82	-3.2

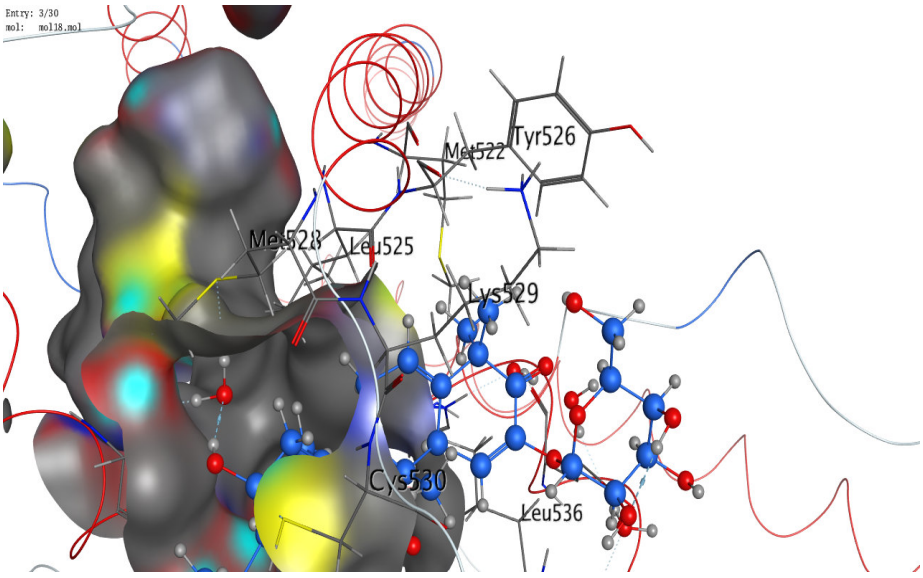
CL26	-6.9969	2.7844	C-81 O-97 O-97 O-28	O SD OG1 CA	HOH20 Met 343 Thr 347 Thr 347 HOH59	H-donor H-donor H-donor H-acceptor	2.77 2.97 2.88 3.20	1.8 -0.8 -0.7 -0.7
CL27	-6.3859	2.7844	O-95	SD	Met522	H-donor	3.17	-2.3

(Notes: Ref: the reference molecule is the original molecule of the ERα.)

The bond length and hydrogen bonding interactions in the active site were further investigated and characterized in **figure 23**whichshowed that C**L18** interact with an amino acid **Asp351** residue and a water molecule (solvent) with H-donor interaction with two. For the **Asp351** the interaction between the functions (O-H….O-C) with a length 2.74 Å. For the water molecule (O-H….O-H) with a length 2.67 Å.



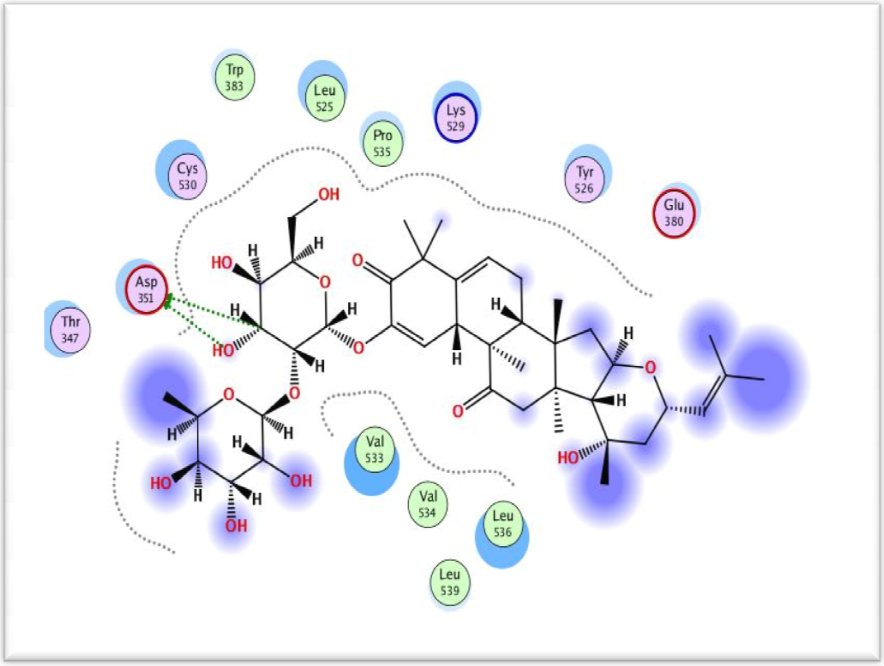
(a)



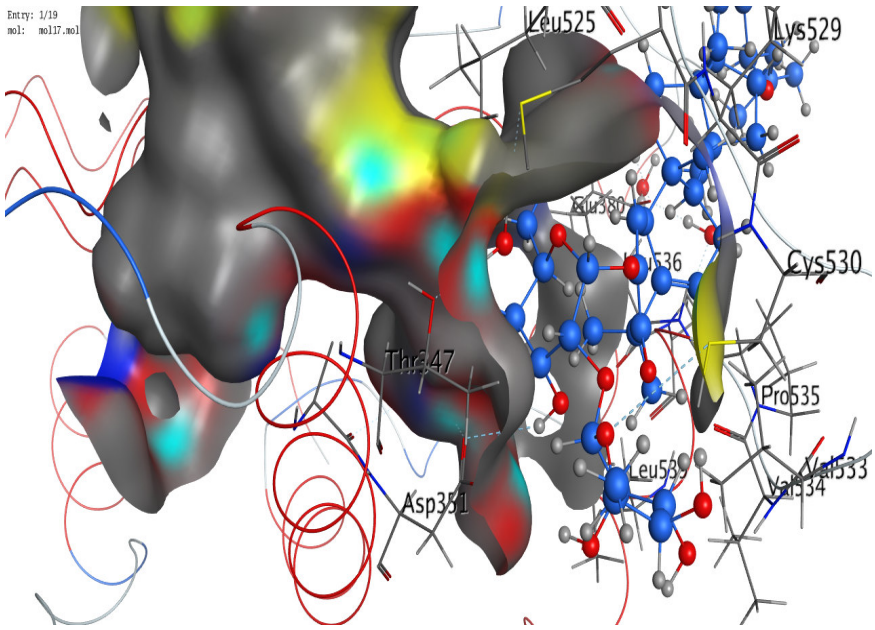
(b)

Figure 23: Interaction of the CL18 with the ERα (3ERT)

In the **figure24** witch showed that the C**L17** interact with an amino acid **Asp351**residue with tow bonds types H-donor: (C-H….O-C) with a length 3.28 Å and (O-H….O-C) with a length 3.34 Å.



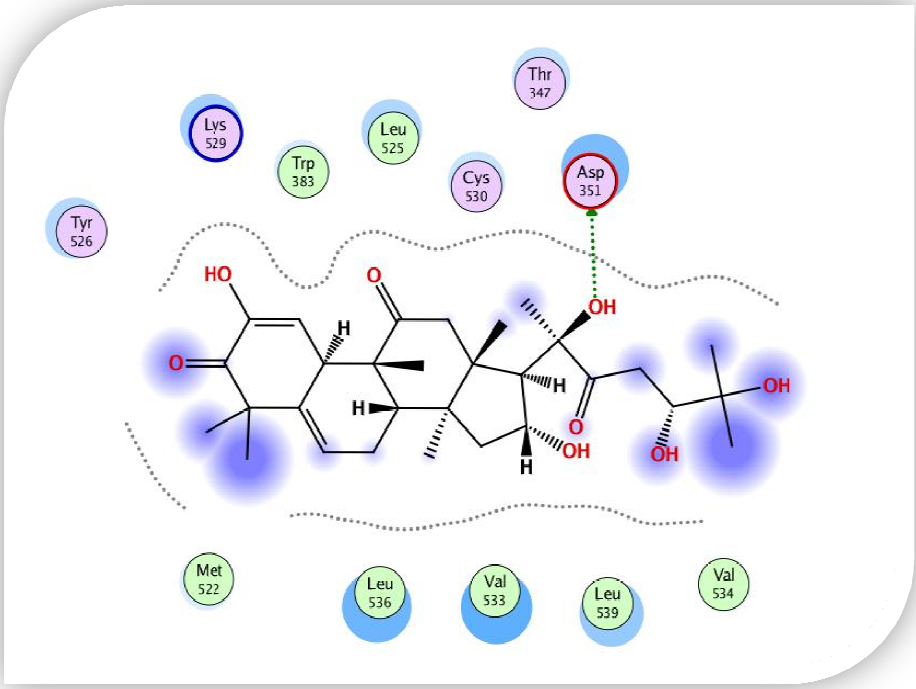
(a)



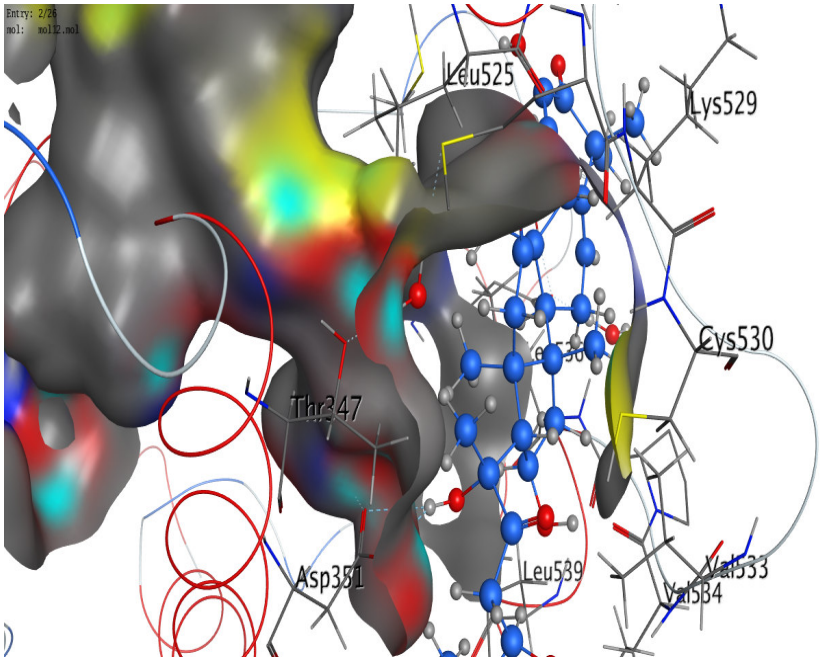
(b)

Figure 24:Interaction of the CL17 with the receptor ERα (3ERT).

The **figure 25** showed that, the ligand **CL12** interacted with the amino acid **Asp351** residue with tow bond types H-donor (O-H ….C-H) with a length 2.94 and 3.10 Å.



(a)



(b)

Figure 25: Interaction of the ligand CL12 with the receptor ERα (3ERT).

1.2 ADMET properties of cucurbitacins compounds

Absorption, distribution, metabolism, elimination, toxicity, properties for the ligands chosen from docking were cited in **table 4**. **CL18 and CL17** had value of caco2 <-5.15 exepte the **CL12**had a caco2 value a little higher. All ligand classified as p-gp inhibitor and substrate which means that all ligand have ability to be removed from intestines back into the gut lumen and from kindneys and liver into urine and bile respectively also by this fonction we can maintaining the integrity of the blood brain barrier**[3]**.

On the other hand, the ligand **CL12**could be administered by the oral route which take a good bioavailability (F>50%).However, **CL18 and CL17** had a low bioavailability (F<50%).

All the protein plasma binding values of ligands was less than 90% (low PPB bond) which means that plasma proteins play their role by virtue of their high concentration, control the free drug concentration in plasma and in compartments in equilibrium with plasma, thereby, effectively attenuating drug potency in vivo [**4**]. All ligand passed BBB and from the VD part <0.07L/kg highly hydrophilic, 0.07-0.7 evenly distributed and >0.7 highly lipophilic, all ligand are highly lipophilic (the VD optimal between 0.04-20 L/Kg).

All ligand considered as inhibitors of CYP-1A2, 2C9, 2D6, 2C19, 3A4, and inhibitory promiscuity. In addition all ligands were considered also as substrat of CYP2C9, 2D6, 3D4, all of these CYPs responsible for phase I reactions are concentrated in the liver.

So all ligands eligible for submit to biotransformation reaction (metabolism reactions).

From excretion side, all ligand had a clairance value less than 5 ml/min/kg and had a half time value higher than 0.5h for the ligand **CL18, CL12 and CL17.**

For toxicity, all ligand are hERG blockers and had Ames mutagenicity. The LD50 acute of toxicity value should be > 500 mg/kg to considered as low toxic, 50-500 mg/kg toxic and <50 high toxic, so all ligand considered as a toxic compound because they have LD50 value low than 50 also all ligand have a IGC50 value less than 50.

By comparing the compouns ligands with the references molcule we noticed that the ligandshad a diffrents values than the values of referencs molecule.

Table 4:The results obtained from ADMET predict of cucurbitacins compounds

Comp	Absorption				Distribution				Metabolism CYP450									Excretion		Toxicity			
	Caco-2 permeability	HIA	P-gp inhibitor	P-gp substrate	F%	PPB%	BBB	VD (L /Kg)	2C9 Substrate	2D6 Substrate	3A4 Substrate	1A2 Inhibitor	2C9 Inhibitor	2D6 Inhibitor	2C19 Inhibitor	3A4 Inhibitor	Inhibitory Promiscuity	CL (ml/min/kg)	T (h)	LD50(mol/kg)	AMES	hERG	pIC50(ug/L)
L_{ref}	-4.565	0.99	0.99	0.0	0.01	95.95	0.69	1.71	0.13	0.46	0.92	0.53	0.13	0.98	0.33	0.92	0.91	9.55	0.07	5.59	0.07	0.88	4.76
CL18	-5.47	0.82	0.69	0.87	0.17	73.46	0.63	0.43	0.85	0.90	0.72	0.86	0.84	0.94	0.94	0.77	0.91	1.38	0.60	3.52	0.79	0.98	0.99
CL17	-5.46	0.80	0.69	0.87	0.17	74.75	0.62	0.50	0.87	0.90	0.71	0.92	0.87	0.95	0.94	0.79	0.94	1.06	0.23	3.63	0.87	0.98	0.99
CL12	-5.07	0.98	0.89	0.83	0.55	76.81	0.82	0.43	0.85	0.90	0.75	0.93	0.89	0.95	0.81	0.73	0.90	4.00	0.59	3.71	0.75	0.98	1.09

1.3 Reactivity

Analysis of DFT descriptors allows knowing more about the characteristics stability **(ΔE)**, electrophilic **(μ)** and nucleophilic **(N)** compound. Negative values for $\mathbf{E_{HOMO}}$ and $\mathbf{E_{LUMO}}$**;** refer to the stability of the examined compounds [**4]**.

Thechemical reactivity descriptor calculatedto the L_{Ref} and **CL12**which selected from the docking and ADMET prrdiction steps, by comparing the obtained values of compound with the values of reference molecule we noticed that the compound had a values of **(ΔE=4.9364 eV andη = 2.2673 eV)** little more than the reference molecule value**(ΔE=4.3073 eV and η = 2.1536 eV)**.**CL12** had an **(S=0.2026 eV)** value approximately equal to the $\mathbf{L_{Ref}}$**(S=0.2026 eV)** which means that this compound are reactive. **CL12**had an electrophilicity index more than 2eV that was mean that this compound could participate easily in polar reaction **(ω=2.36874eV)**value higher than the value of references molecule**(ω=2.02021eV**). **CL12** had a value of electronic chemical potential**(μ=-3.4195 eV)**less than the $\mathbf{L_{Ref}}$**(μ=-2.9499eV)**it could exchange electron density with the environment efficiently. A further classification of organic molecules as strong **(N > 3 eV)**, moderate **(2.0 eV ≤ N ≤3.0 eV)** and marginal nucleophile **(N < 2.0 eV)** were obtained by analysis of a series of common nucleophilic species participating in polar organic reaction, the value of **CL12(N= 2.76822 eV)** less than the values of reference molecule**(N= 3.55245eV),** so the compound classified as organic molecule moderate**[5]**.

2 RESULTS AND DISCUSSION OF FLAVONOIDS COMPOUNDS

2.1 Molecular docking results

In this time the affinity of flavonoids ligands to the ERα **(PDB: 3ERT)** were studied to identify receptor-ligand interactions in the binding pocket. The results of docking showed in the **table 5**. As we said, the bust result of docking based on both the binding scores and RMSDs value. We observed that the bust docking results for ligands **FL2, FL5, FL7, FL8** with a score ranging between **(-6.5655 kcal/mol) and (-6.1770kcal/mol).** For the pervious ligands, the RMSDs values were between (1.1506Å) and (2.6513Å).

On the other hand, the others ligands were excluded because they have invalid score, which were out of the validation range.

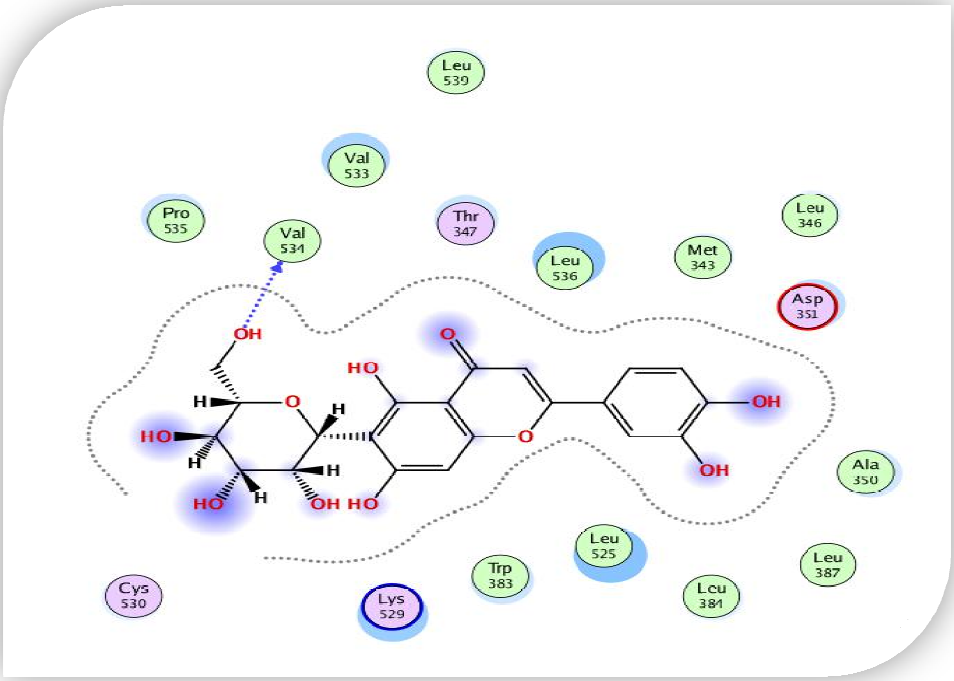
As we noticed in the **table 4**. The **FL2** interact with an amino acid **Val534** by an interaction type **H-donor** between **O-43** and **O** with a length 2.99 Å that is near to the reference value. The **FL5** interact with tow amino acids and a solvent by an interaction type **H-donor**. The first amino acid residue **Val534** between **O-22** and **O** with a length 2.70 Å, the second amino acid residue **Met522** between **O-25** and **SD** with a length 4.12 Å and the third with the solvent **HOH31** between **O-41** and **O** with a length 3.15 Å that had the same solvent interaction like the reference ligand. **FL7** interact with two amino acids residues by interaction type **H-donor**, the first amino acid **Leu525** that the bond was between **O-49** and **O** with a length 2.91 Å that is near to the reference value and the second amino acid **Glu353** between **O-65** and **OE2** with a length 2.72 Å. The **FL8** interact with an amino acid **Met343** by interaction type **Pi-H** between **6-ring** and **CE**.

Table 5: Docking results of flavonoids compounds with ERα.

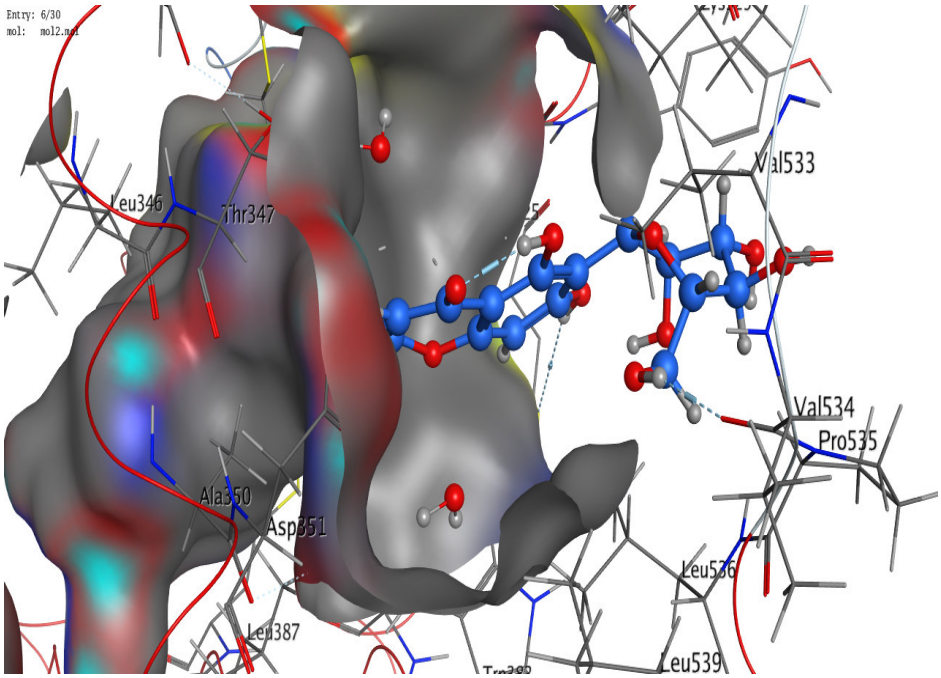
Bonds between atoms of compounds and residues of the active site of 3ERT								
Ligands	S score (kcal/mol)	RMSD (Å)	Atom of compound	Atom of receptor	Evolved receptor residues	Type of interaction bond	Distance (Å)	E (kcal/mol)
L_{Ref}	-6.3009	1.3903	O-15	O	HOH31	H-donor	2.96	-1.7
FL1	-4.4371	1.1241	O-23 O-45 O-43	OD2 O N	Asp351 Val534 Cys530	H-donor H-acceptor H-acceptor	3.12 3.37 4.52	-2.2 -0.7 -2.3
FL2	-6.1770	3.7526	O-43	O	Val 534	H-donor	2.99	-1.9
FL3	-3.6709	-3.6709	O-23 O-26 O-46 O-46 O-48	SD O OD1 OD2 O	Met543 Leu387 Asp351 Asp351 HOH31	H-donor H-donor H-donor H-donor H-donor	2.90 3.16 3.03 3.23 2.67	-2.0 -0.9 -1.3 -0.9 -1.2
FL4	-5.4088	1.9577	O-41	SD	Met522	H-donor	3.15	-0.19
FL5	-6.1903	2.0471	O-22 O-25 O-41	O SD O	Val534 Met522 HOH31	H-donor H-donor H-donor	2.70 4.12 3.15	-3.3 -0.8 -0.6
FL6	-5.1743	2.0690	O-40 6-ring	ND2 N	Asn348 Cys530	H-acceptor Pi-H	3.24 4.35	-1.5 -0.9
FL7	-6.7189	2.2022	O-49 O-65	O OE2	Leu525 Glu353	H-donor H-donor	2.91 2.72	-1.7 -2.8
FL8	-6.5655	1.9951	6-Ring	CE	Met343	Pi-H	3.96	-0.6

FL9	-4.2260	3.4482	O-67	SD	Met522	H-donor	3.49	-2.5
FL10	-5.6543	2.9668	O-32 O-34	O OE2	Leu387 Glu353	H-donor H-donor	3.21 2.76	-0.9 -7.1
FL11	-5.3842	-5.3842	O-25 O-27	O OE2	Glu353 Leu387	H-donor H-donor	2.62 2.71	-4.3 -0.8
FL12	-5.7448	1.7534	O-26	OE2	Glu353	H-donor	3.12	-1.2
FL13	-5.5063	1.9738	O-27	OE2	Glu353	H-donor	2.64	-5.3

The figure 26 showed that the**FL2** interact with an amino acid residue **Val534** with an interaction type H-donor between (O-H….O-C) with a length 2.99Å.



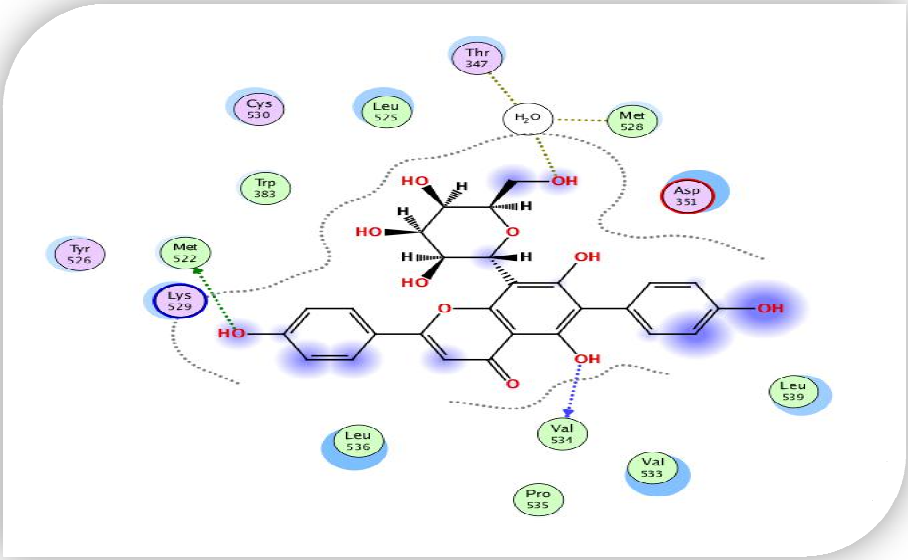
(a)



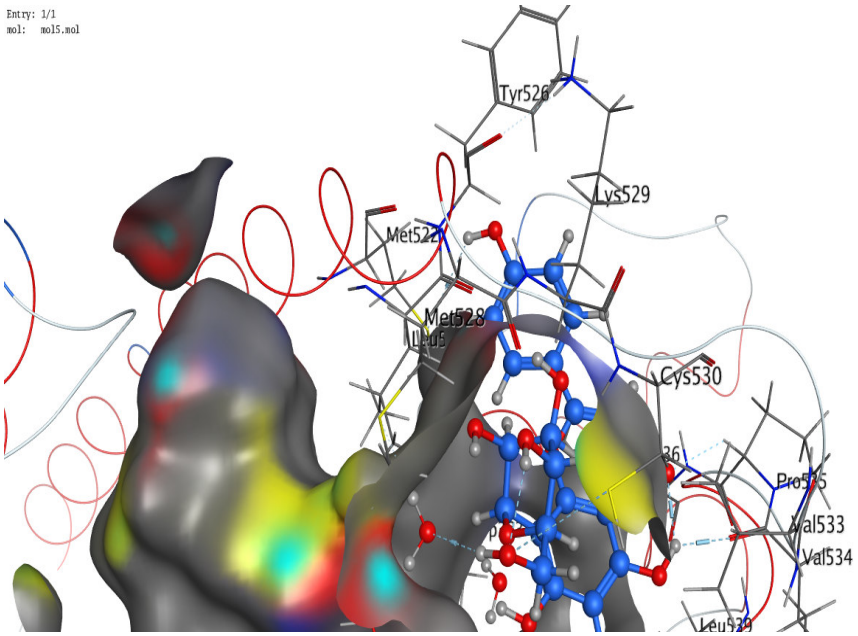
(b)

Figure 26:Interaction of FL2 with the receptor ERα (3ERT).

We noticed in the **figure 27** that **FL5** interact with tow amino acids and solvent. The first amino acid **Val534** with type of interaction H-donor between (O-H….O-C) with a length2.70 Å. The second amino acid was **Met522** with type H-donor between (O-H…SD) with a length 4.12Å. The third interaction was with the solvent HOH by interaction type H-donor with a length 3.15Å.



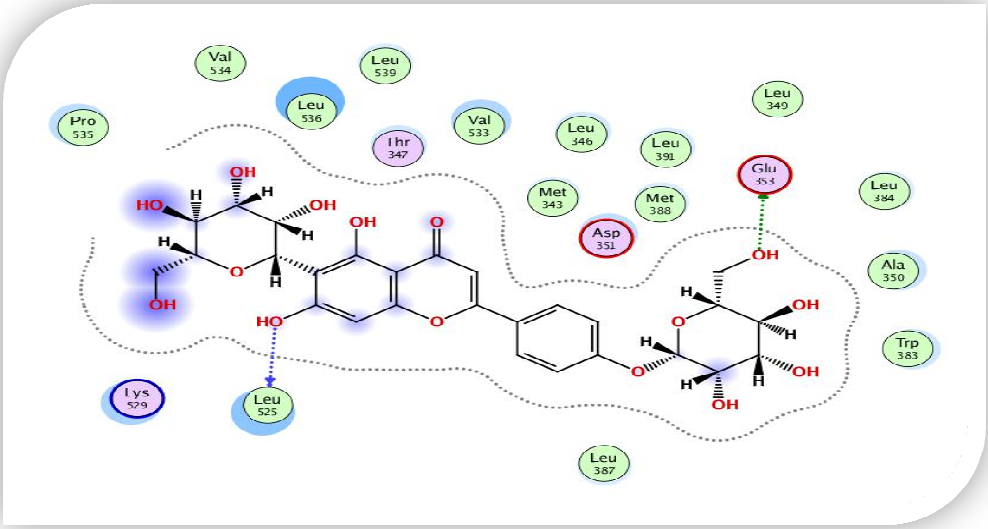
(a)



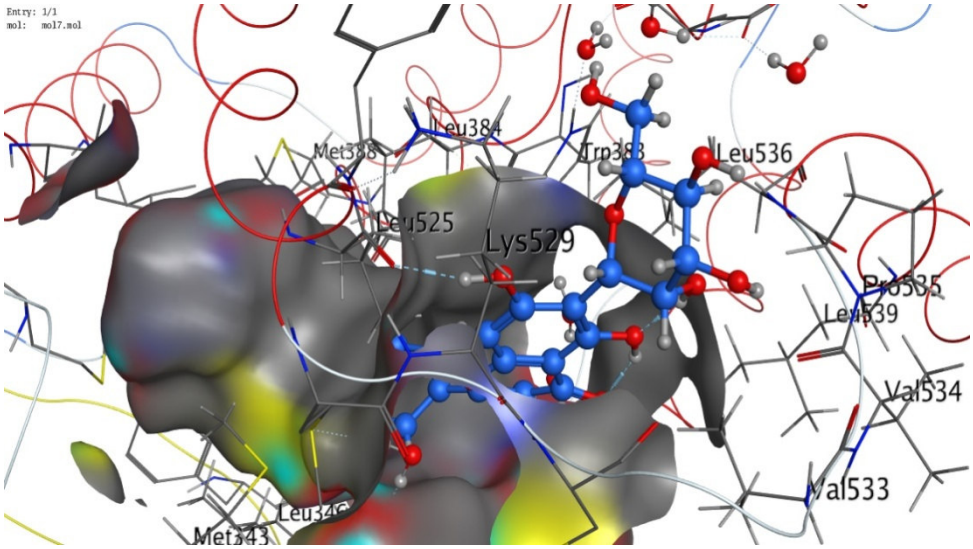
(b)

Figure 27:Interaction of FL5 with the receptor ERα (3ERT).

The **figure 28** indicated that **FL7** interact with tow amino acids residues. The first amino acid was **Leu525** by interaction type H-donor between (O-H….O-C) with a length 2.91Å. The second amino acid was **Glu353** with interaction type H-donor between (O-H….O-C) with a length 2.72 Å.



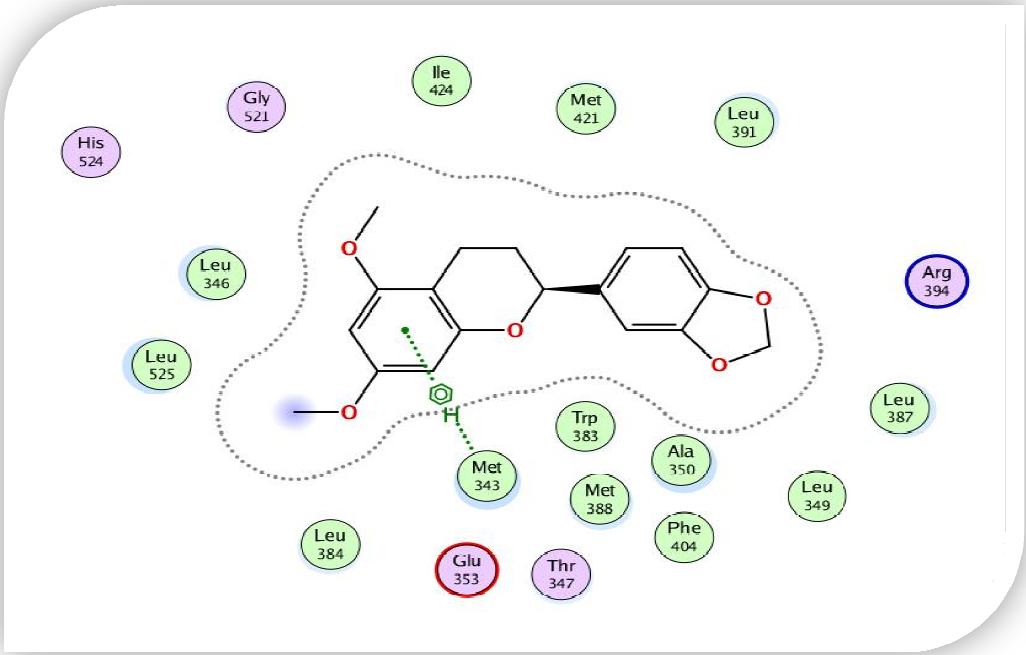
(a)



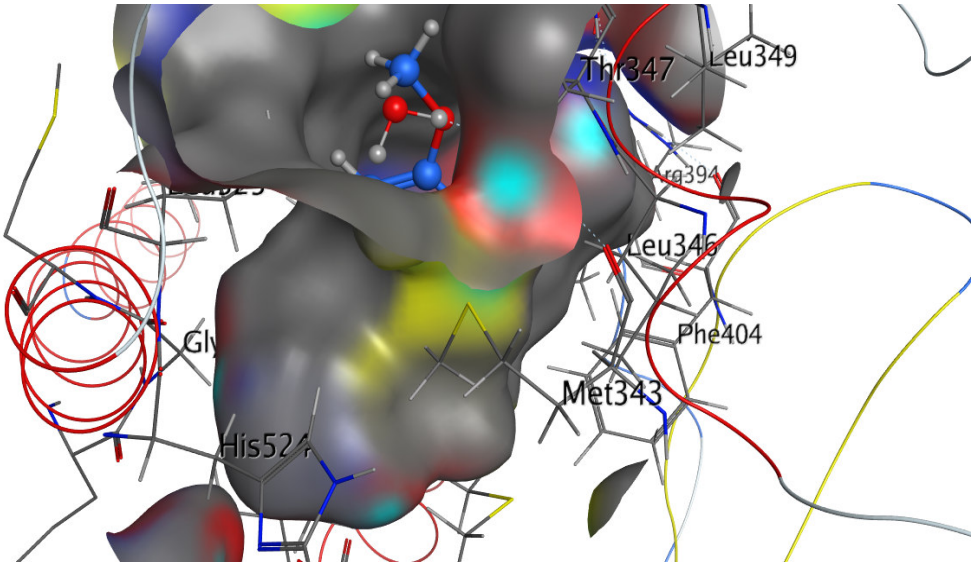
(b)

Figure 28: Interaction of FL7 with the receptor ERα (3ERT).

FL8 in the **figure 29** interact with **Met343** between (C=C….C-H) with a length 3.96Å.



(a)



(b)

Figure 29:Interaction of FL8 with the receptor ERα (3ERT).

2.2 ADMET proprieties results

ADMET properties for the flavonoids ligands chosen from docking cited in **table 6**. **FL8** had value of caco2 >-5.15, all ligand classified as p-gp inhibitor and substrate which means that all ligand had ability to be removed from intestines back into the gut lumen and from kindneys and liver into urine and bile respectively also by this fonction we can maintaining the integrity of the blood brain barrier**[5].**

From other way, all ligands can be administered by the oral route **FL7 and FL8** take a good bioavailability (F>50%).However, **FL2** and **FL5**with a low bioavailability (F<50%).

The PPB values of ligands **FL2** and **FL7**is less than 90% (low PPB bond) which means that plasma proteins play their role by virtue of their high concentration, control the free drug concentration in plasma and in compartments in equilibrium with plasma, there by effectively attenuating drug potency in vivo. All ligand passed BBB and from the VD part <0.07L/kg highly hydrophilic, 0.07-0.7 evenly distributed and >0.7 highly lipophilic, the ligands **FL2**, **FL5**and **FL8** are highly lipophilic (the VD optimal between 0.04-20 L/Kg) however, **FL7** was highly hydrophilic.

All ligand considered as inhibitors of CYP-1A2, 2C9, 2D6, 2C19, 3A4, and inhibitory promiscuity. In addition all ligands were considered also as substrat of CYP2C9, 2D6, 3D4, all of these CYPs responsible for phase I reactions are concentrated in the liver.

So all ligands eligible for submit to biotransformation reaction (metabolism reactions).

From excretion side, all ligand have a clairance value less than 5 ml/min/kg except **FL8** anda half time value less than 0.5 h for the **FL2**, **FL5**, and **FL7.**

For toxicity, all ligand are hERG blockers and have Ames mutagenicity. The LD50 acute of toxicity value should be > 500 mg/kg to considered as low toxic, 50-500 mg/kg toxic and <50 high toxic, so all ligand considered as a high toxic compound because they have LD50 value low than 50 also all ligand have a IGC50 value less than 50.

By comparing the compouns ligands with the references molcule we noticed that the ligands had a diffrents values than the values of referencs molecule.

Table 6: ADMET properties of flavonoids compounds.

Comp	Absorption					Distribution			Metabolism CYP450									Excretion		Toxicity			
	Caco-2 permeability	HIA	P-gp inhibitor	P-gp substrate	F%	PPB%	BBB	VD (L /Kg)	2C9 Substrate	2D6 Substrate	3A4 Substrate	1A2 Inhibitor	2C9 Inhibitor	2D6 Inhibitor	2C19 Inhibitor	3A4 Inhibitor	Inhibitory Promiscuity	CL (ml/min/kg)	T (h)	LD50(mol/kg)	AMES	hERG	pIGC50(ug/L)
L_{ref}	-4.56	0.99	0.99	0.0	0.01	95.95	0.69	1.71	0.13	0.46	0.92	0.53	0.13	0.98	0.33	0.92	0.91	9.55	0.07	5.59	0.07	0.88	4.76
FL2	-6.14	0.94	0.92	0.58	0.17	89.39	0.64	0.91	0.80	0.87	0.60	0.83	0.90	0.94	0.92	0.83	0.78	3.64	0.71	2.36	0.72	0.97	0.22
FL5	-6.20	0.94	0.92	0.58	0.17	97.56	0.64	0.70	0.80	0.87	0.60	0.83	0.90	0.94	0.92	0.83	0.78	2.07	0.40	2.36	0.72	0.97	0.22
FL7	-6.27	0.70	0.86	0.63	0.63	71.35	0.72	0.62	0.81	0.88	0.61	0.91	0.93	0.94	0.90	0.93	0.75	1.08	0.19	2.11	0.53	0.97	0.28
FL8	-4.79	0.98	0.58	0.58	0.55	93.58	0.86	1.70	0.86	0.82	0.60	0.75	0.69	0.63	0.74	0.81	0.90	14.22	0.18	2.78	0.81	0.94	0.54

2.3 Reactivity of flavonoids compounds

As we said, analysis of density functional theory descriptors allows knowing more about the characteristics stability **(ΔE)**, electrophilic **(μ)** and nucleophilic **(N)** compound. Negative values for $\mathbf{E_{HOMO}}$ and $\mathbf{E_{LUMO}}$**;** refer to the stability of the examined compounds.

The chemical reactivity descriptor calculated to **FL8** and $\mathbf{L_{Ref}}$ligand that selected from the docking and ADMET properties steps, by comparing the obtained values of compound with the values of reference molecule we noticed that the compound had a values of **(ΔE=5.5457eV and η = 2.7728eV)**more than the reference molecule value**(ΔE=4.3073 eV and η = 2.1536 eV)**.**FL8** had an **(S=0.1803eV)** value less than the $\mathbf{L_{Ref}}$value **(S=0.2322eV)** which means that this compound are reactive. **FL8**had an electrophilicity index less than 2eV that was mean that this compound couldn't participate in polar reaction **(ω=1.26226eV)**value less then the value of references molecule**(ω=2.02021eV**). **FL8** had a value of electronic chemical potential**(μ=-2.6458eV)**) was higher then the $\mathbf{L_{Ref}}$**(μ=-2.9499eV)**it could exchange electron density with the environment efficiently. A further classification of organic molecules as strong **(N > 3 eV)**, moderate **(2.0 eV ≤ N ≤3.0 eV)** and marginal nucleophile **(N < 2.0 eV)** were obtained by analysis of a series of common nucleophilic species participating in polar organic reaction, the value of **FL8(N= 3.23734eV)** less than the values of reference molecule**(N= 3.55245eV),** so the compound classified as organic molecule strong.

Referencs:

1. Towler P, Staker B, Prasad SG, Menon S, Tang J, Parsons T, et al. ACE2 X-Ray Structures Reveal a Large Hinge-bending Motion Important for Inhibitor Binding and Catalysis. J Biol Chem. 2004;279(17):17996–8007.
2. DS W, YD F, AC G, EJ L, A M, JR G, et al. DrugBank [Internet]. 2020. Available from: https://www.drugbank.ca/.
3. Bohnert, Tonika, and Liang-Shang Gan. "Plasma protein binding: from discovery to development." *Journal of pharmaceutical sciences* 102.9.2013: 2953-2994.
4. Srivastava R. Chemical reactivity theory (CRT) study of small drug-like biologically active molecules. J Biomol Struct Dyn. 2020 : 1–10.
5. Matheny, Christopher J., et al. "Pharmacokinetic and pharmacodynamic implications of P□glycoprotein modulation." *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy* 21.7. 2001: 778-796.
6. Wang, Hengliang, et al. "DFT study of new bipyrazole derivatives and their potential activity as corrosion inhibitors." *Journal of Molecular Modeling* 13. 2007 : 147-153.

Conclusion:

The objective of this study was to investigate the interaction between cucurbitacins and flavonoid compounds with the ERα receptor using docking analysis. The ranking of the docking results suggested that certain ligands from these two compounds might possess the capability to inhibit the ERα receptor.

The docking results were analyzed based on the various interactions observed between the ligands of the compounds and the ERα receptor. The findings indicated that all the compounds studied exhibited interactions with the ERα receptor.

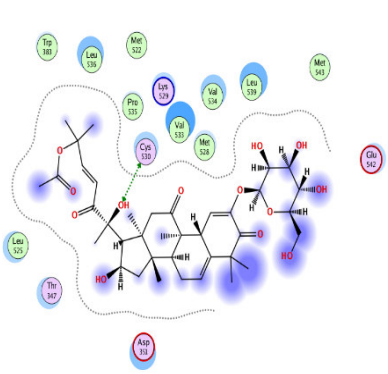
Furthermore, the docking analysis revealed that among the cucurbitacin compounds, **CL18** exhibited the lowest score value and a good RMSD value, indicating an H-donor and H-acceptor interaction. This was followed by **CL17** with an H-donor interaction and **CL12** with another H-donor interaction.
ADMET prediction for the three selected cucurbitacin compounds based on the docking results indicated that **CL12** possessed the most favorable pharmacokinetic properties.
To evaluate the chemical reactivity of **CL12** from the cucurbitacin compounds, the DFT/B3LYP/6-31G method was employed, and the obtained results indicated that this ligand displayed favorable descriptor values.

Regarding the flavonoid compounds, **FL7** exhibited the lowest score value and an acceptable RMSD value, with an H-donor interaction. This was followed by **FL8** with a Pi-H interaction, **FL5** with an H-donor interaction, and **FL2** with an H-donor interaction.
ADMET prediction for the four selected flavonoid ligands based on the docking results indicated that **FL8** possessed the most favorable pharmacokinetic properties.
Similar to the cucurbitacins, the chemical reactivity of **FL8** from the flavonoid compounds was assessed using the DFT/B3LYP/6-31G method, and the results indicated favorable descriptor values for this ligand.

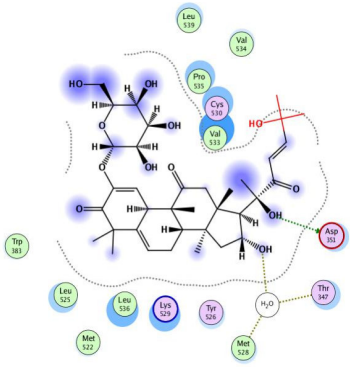
In conclusion, all the findings suggest that a combined computational approach involving molecular docking, ADMET properties, and global reactivity descriptors can offer an alternative method to understand the binding mechanism of **Cucurbitacine J** and **2S-3',4'-methylene dioxy-5,7-dimethoxy flavan** as effective inhibitors of the ERα receptor.

APPENDIX

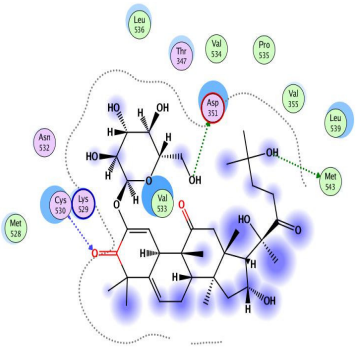
Appendix



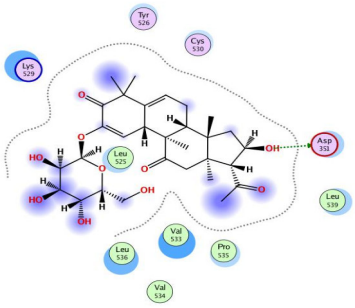
CL1



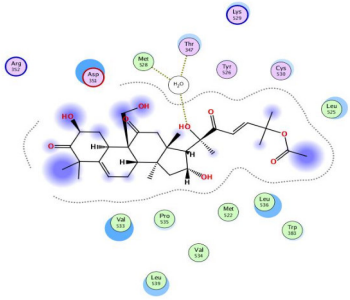
CL2



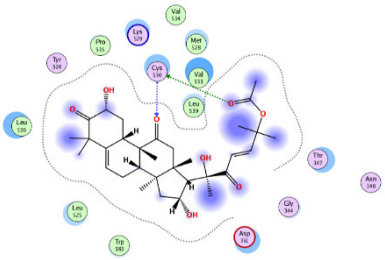
CL3



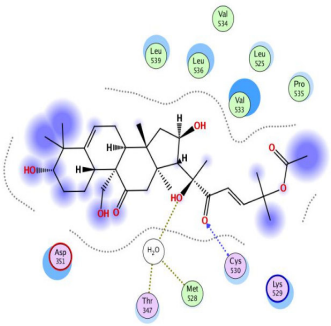
CL4



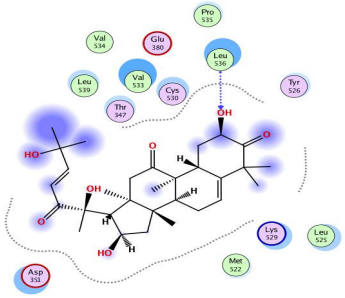
CL5



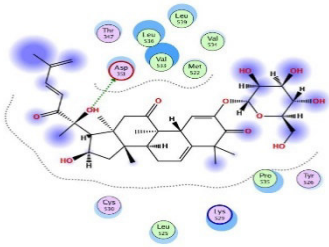
CL6



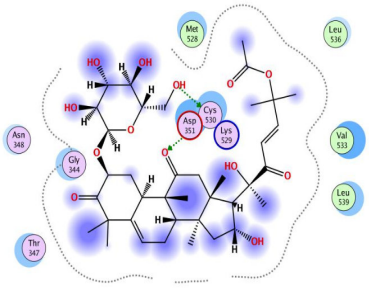
CL7



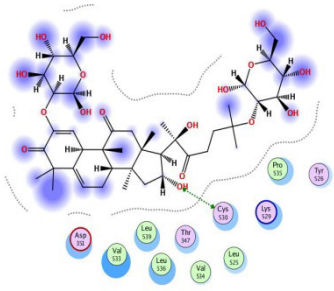
CL8



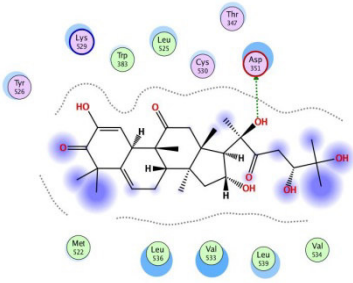
CL9



CL10

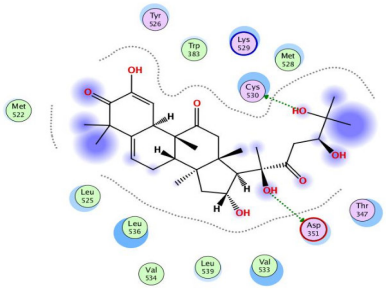


CL11

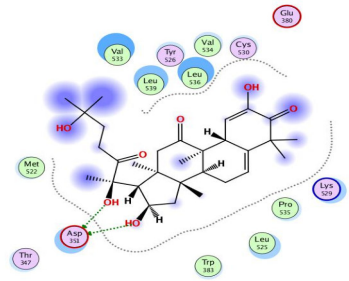


CL12

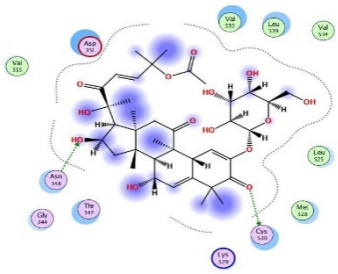
Appendix



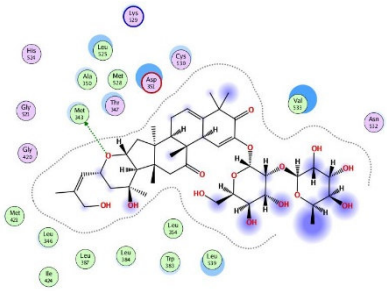
CL13



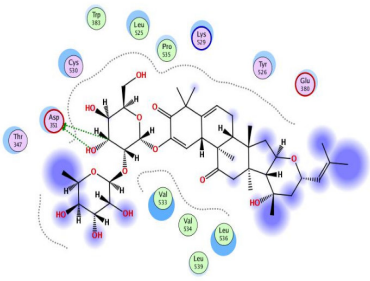
CL14



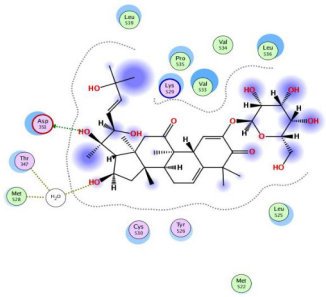
CL15



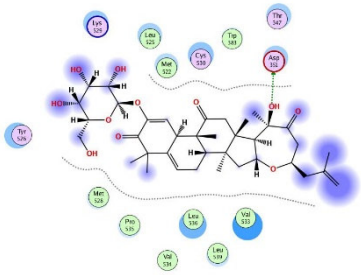
CL16



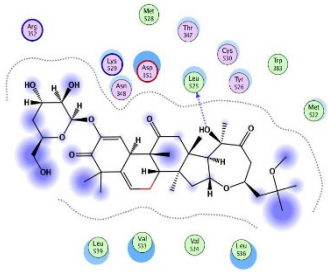
CL17



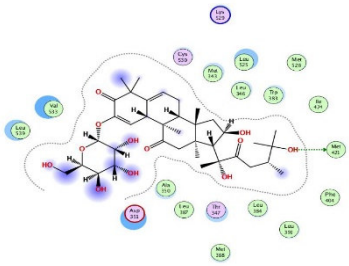
CL18



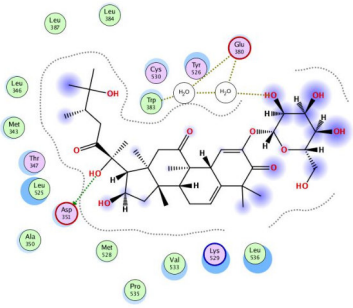
CL19



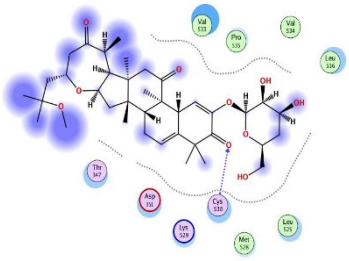
CL20



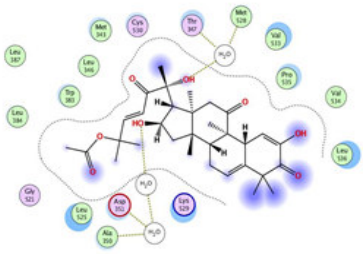
CL21



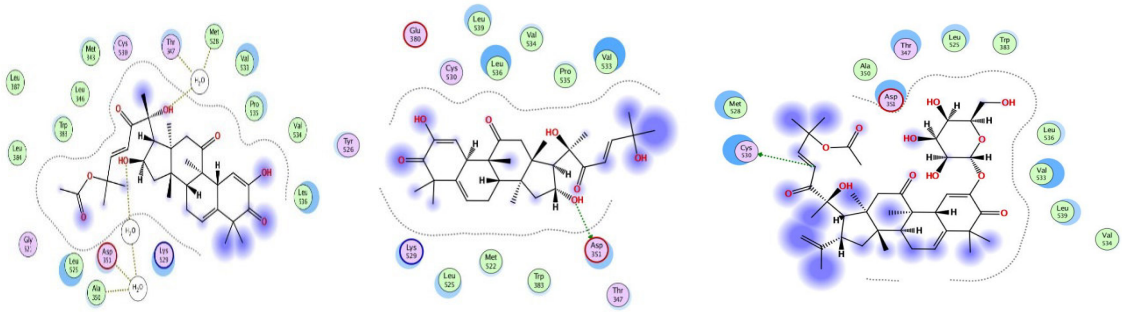
CL22



CL23



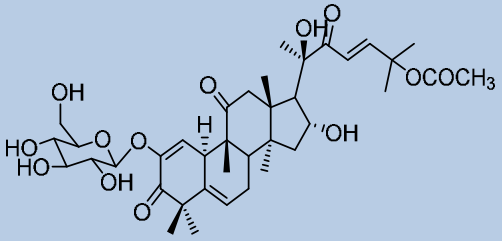
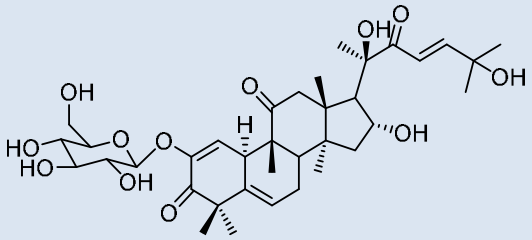
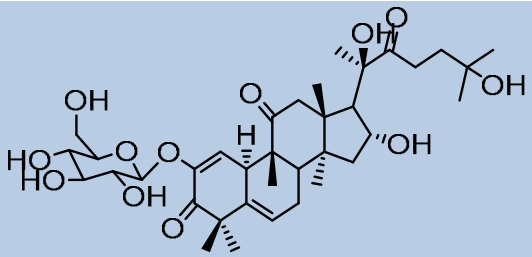
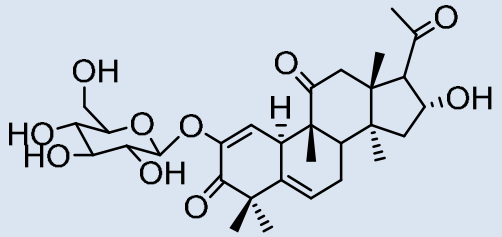
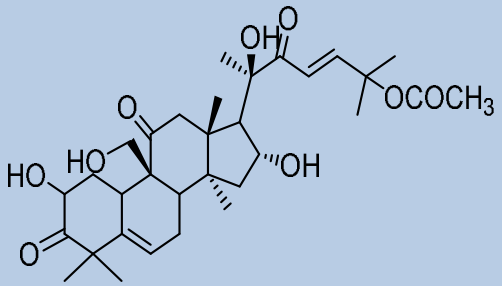
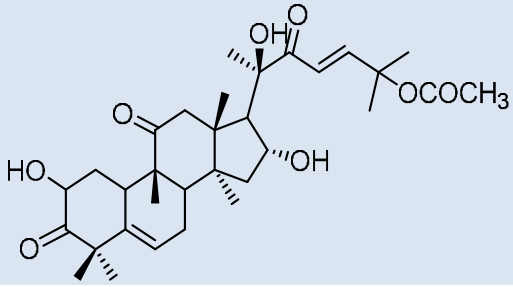
CL24

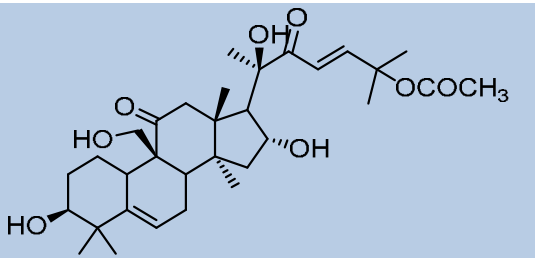
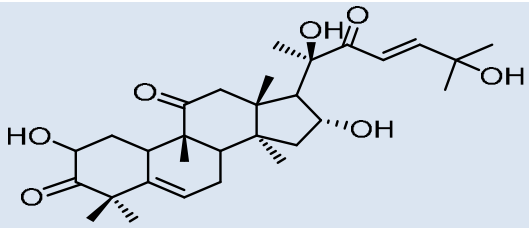
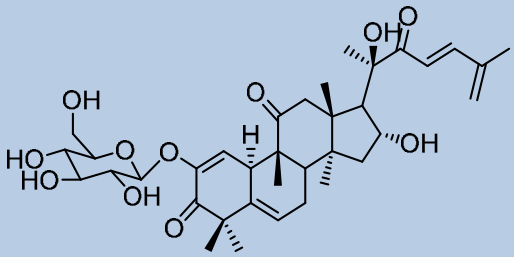
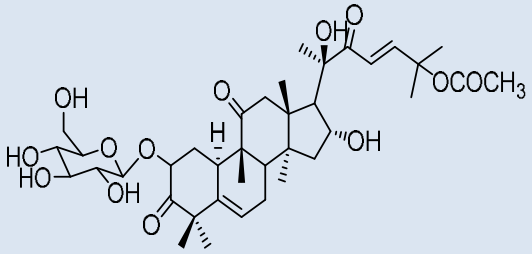
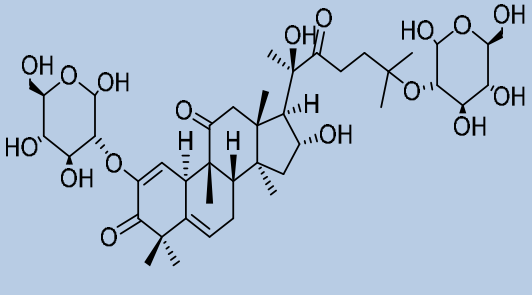


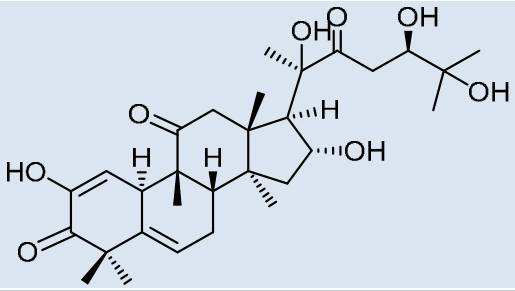
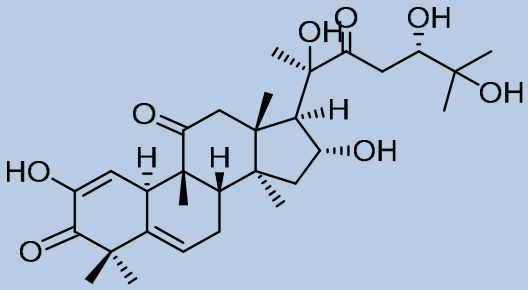
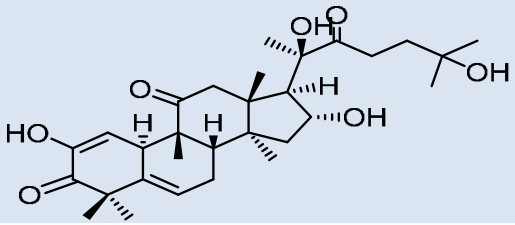
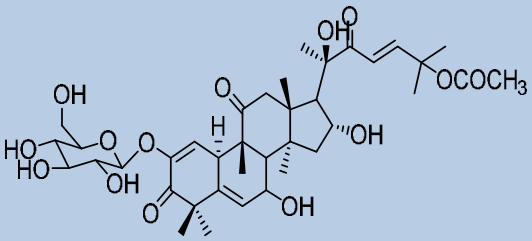
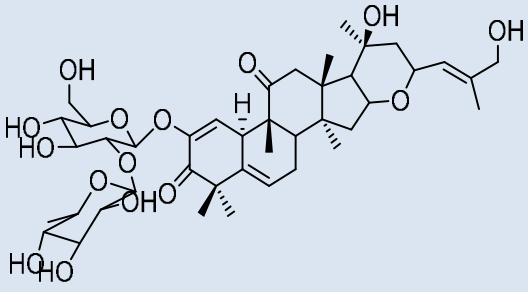
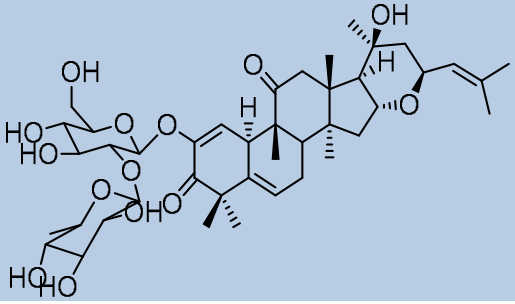
CL25 CL26 CL27

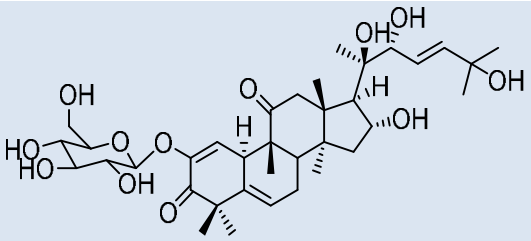
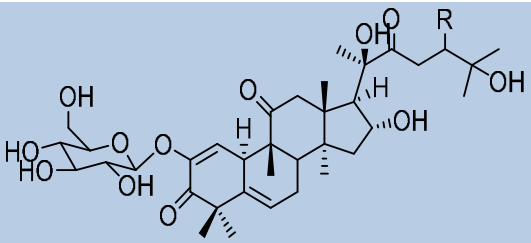
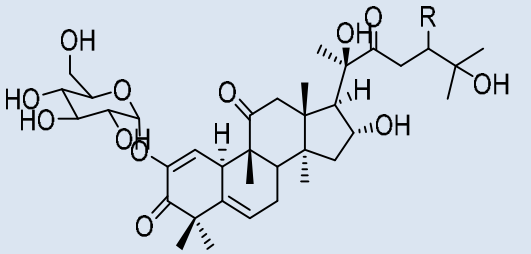
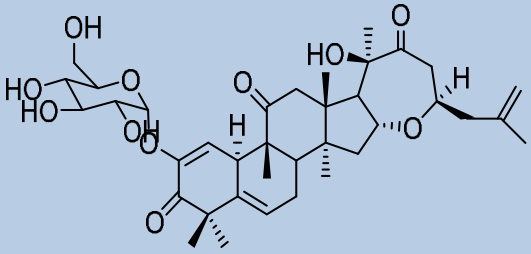
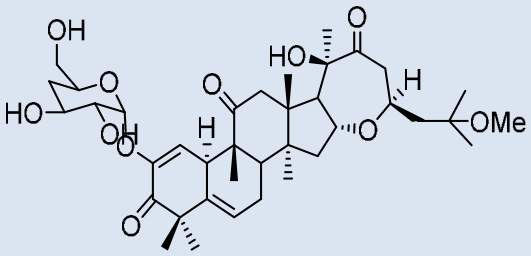
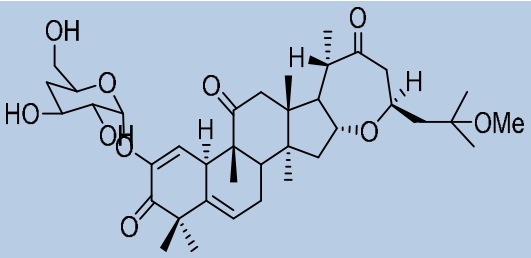
Figure 30: Interaction of cucurbitacins compound with the ERα receptor.

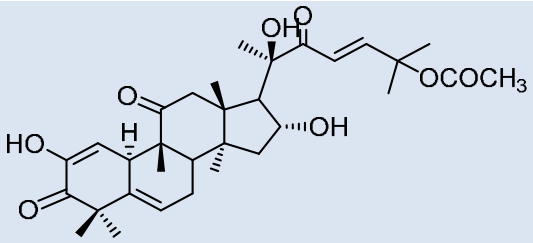
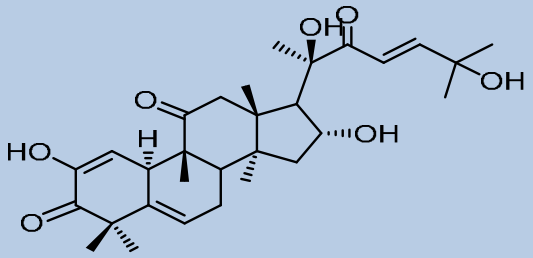
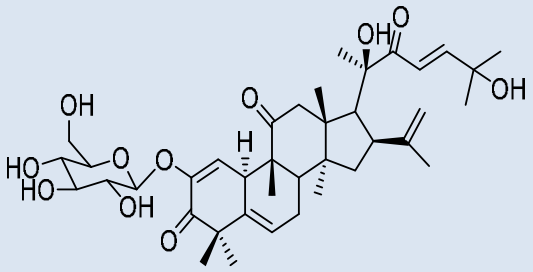
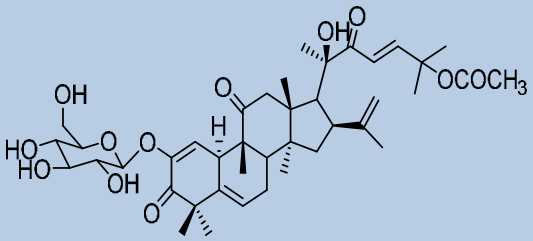
Table 7: The structures and names of cucurbitacins compound.

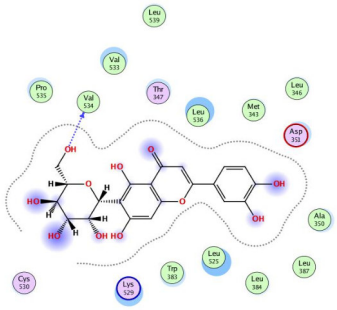
Compounds N°	Name	structure
CL1	2-O-β-D-Glucopyranosyl-cucurbitacine E	
CL2	2-Oβ-D-Glucopyranosyl-cucurbitacine L	
CL3	2-Oβ-D-Glucopyranosyl-cucurbitacine L	
CL4	2-O-β-D-Glucopyranosyl-(22,27)-hexanocucurbitacine	
CL5	CucurbitacinesA	
CL6	Cucurbitacine B	

CL7	Cucurbitacine C	
CL8	CucurbitacineD	
CL9	2-Oβ-D-Glucopyranosyl-16α-20R-dihydroxy-cucurbita-1,5,23E,25(26)-tetraen-3,11,22-trione	
CL10	2-O-β-D-Glucopyranosyl-cucurbitacine B (avernine I)	
CL11	2,25-di-O-β-DGlucopyranosyl-cucurbitacine L	

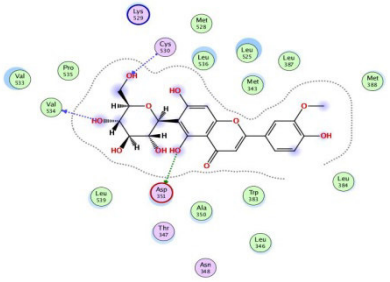
CL12	Cucurbitacine J	
CL13	Cucurbitacine K	
CL14	Cucurbitacine L	
CL15	Colocynthoside A	
CL16	ColocynthosideB	
CL17	Déoxocucurbitacine B	

CL18	khekadaengoside E	
CL19	2-O-β-D-Glucopyranosylcucurbitacine J	
CL20	2-O-β-D-Glucopyranosylcucurbitacine K	
CL21	Colocynthines A	
CL22	Colocynthines B	
CL23	Colocynthines C	

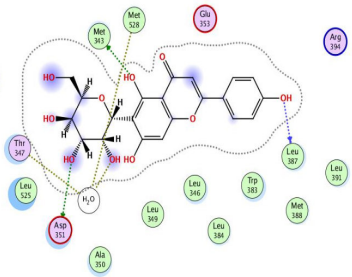
CL24	Cucurbitacine E	
CL25	Cucurbitacine I	
CL26	16-(2-Prop-1-en-yl)-2-O-βD-glucopyranosylcucurbitacine I	
CL27	16-(2-Prop-1-en-yl)-2- O-β-D-glucopyranosylcucurbitacine E	



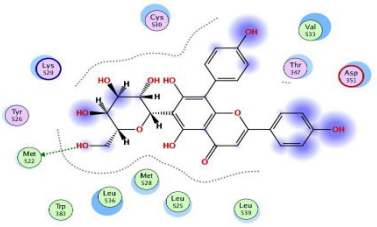
FL1



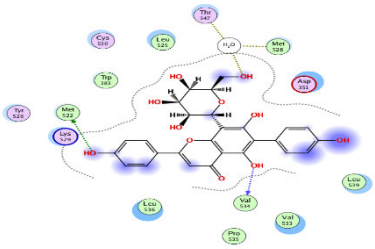
FL2



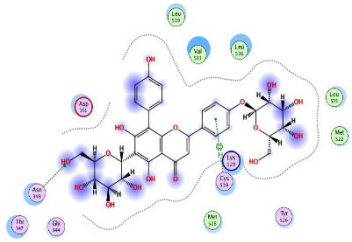
FL3



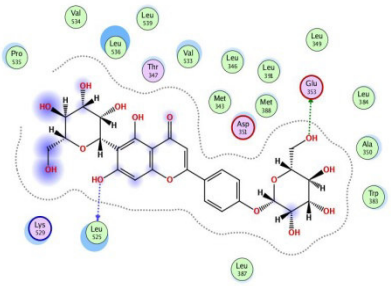
FL4

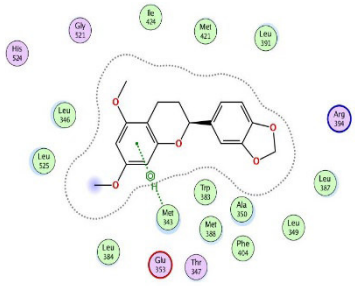


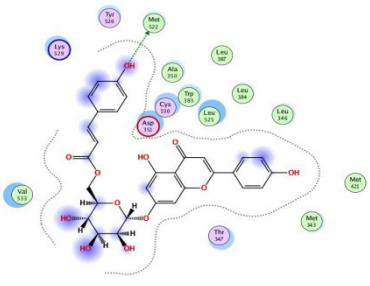
FL5



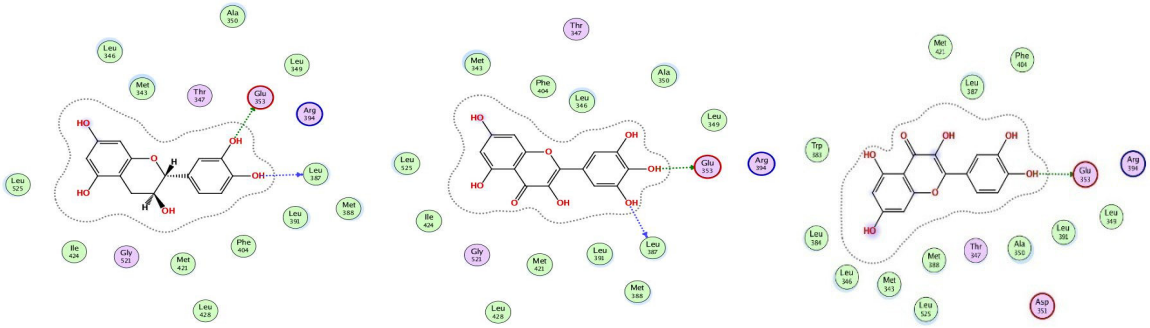
FL6



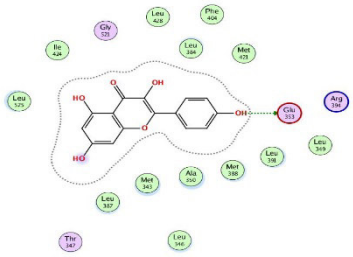




FL7 FL8 FL9



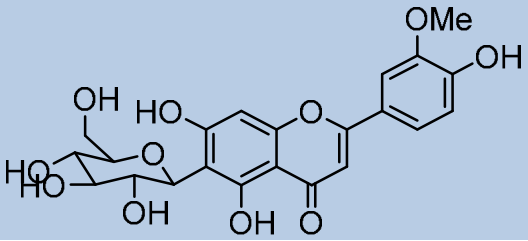
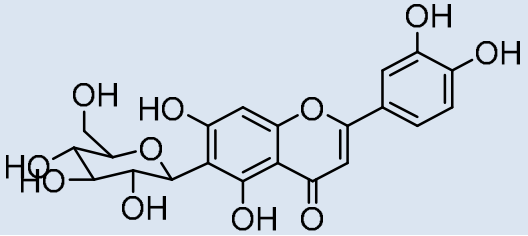
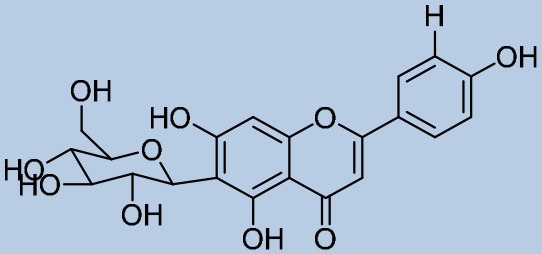
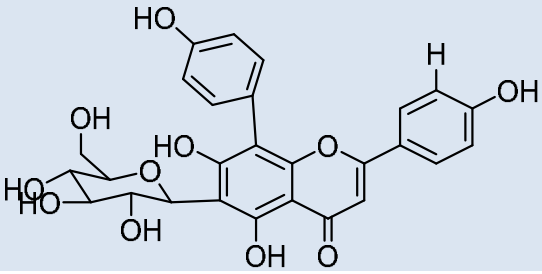
FL10 FL11 FL12

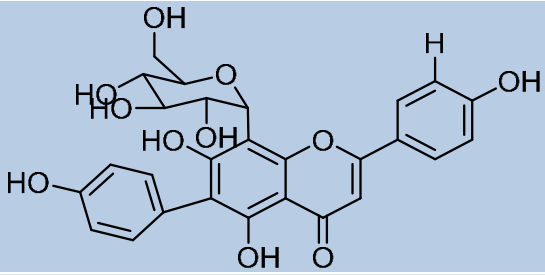
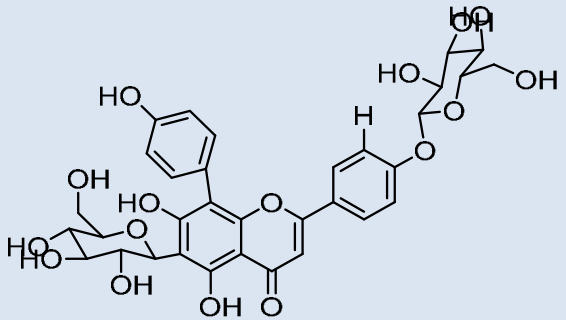
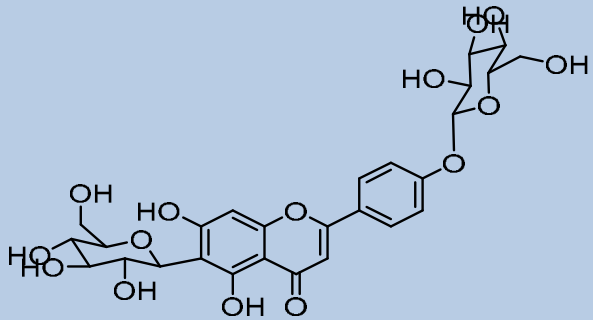
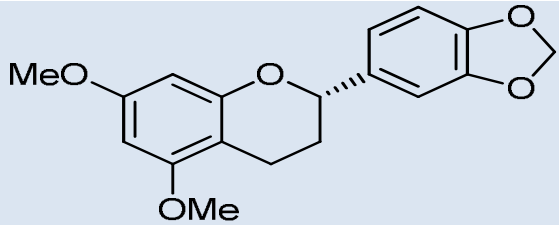
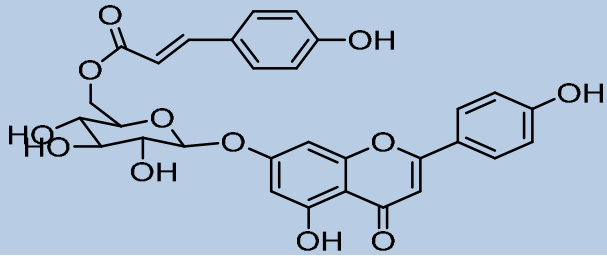
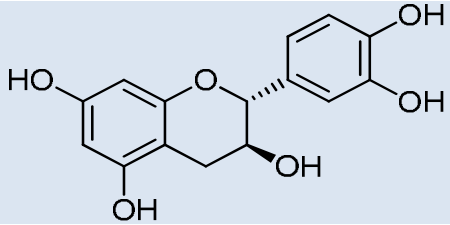


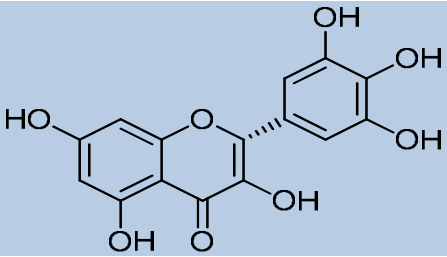
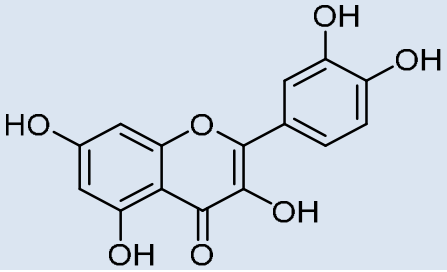
FL13

FIGURE 31:THE STRUCTURES AND NAMES OF CUCURBITACINS COMPOUND.

Table 8: Structures and names of flavonoids compound

Compound N° :		Structure
FL1	**Isoorientin-3'-methyl ether**	
FL2	**Isoorientin**	
FL3	**Isovitexine**	
FL4	**8-C-p-hydroxybenzoyl-isovitexin**	

FL5	**6-C-p-hydroxybenzoyl-vitexin**	
FL6	**8-C-p-hydroxybenzoyl-4'-O-β-D-glycosyl-isovitexin**	
FL7	**Isosaponarin**	
FL8	**2S-3',4'-methylene dioxy-5,7-dimethoxy flavan**	
FL9	**Hispidulin-7-(6-E-p-coumaroyl-β-D-glucopyranosid)**	
FL10	**Catechol**	

FL11	Myricetol	
FL12	Quercetol	
FL13	Kamferol